



Universidade do Minho
Escola de Engenharia

Marisa da Mota Pinheiro

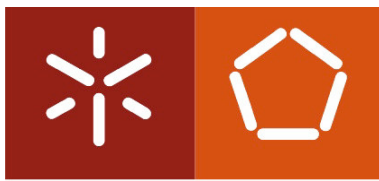
**Numerical Methodology to Characterise
Heterogeneous Rock Masses**

Marisa da Mota Pinheiro **Numerical Methodology to Characterise Heterogeneous Rock Masses**

UMinho | 2017



May 2017



Universidade do Minho

Escola de Engenharia

Marisa da Mota Pinheiro

**Numerical Methodology to Characterise
Heterogeneous Rock Masses**

Tese de Doutoramento

Engenharia Civil

Trabalho efetuado sob a orientação do

Professor Doutor Tiago F. S. Miranda

Doutor Engenheiro Luís N. Lamas

maio de 2017

DECLARAÇÃO

Nome Marisa da Mota Pinheiro

Endereço eletrónico: marisamotapinheiro@gmail.com

Número do Cartão de Cidadão: 13541641

Título da tese:

Numerical Methodology to Characterise Heterogeneous Rock Masses

Orientador(es):

Professor Doutor Tiago Filipe da Silva Miranda

Doutor Engenheiro Luís Lamas

Ano de conclusão: 2017

Designação do Doutoramento: Doutoramento em Engenharia Civil

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA TESE APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE.

Universidade do Minho, ___/___/_____

Assinatura: _____

STATEMENT OF INTEGRITY

I hereby declare having conducted my thesis with integrity. I confirm that I have not used plagiarism or any form of falsification of results in the process of the thesis elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

University of Minho, _____

Full name: Marisa da Mota Pinheiro

Signature: _____

*You don't write because you want to say something; you
write because you've got something to say.*

F. Scott Fitzgerald

ABSTRACT

The uncertainties related with spatial variability and heterogeneities, naturally present in rock masses, play an important role in geotechnical engineering practice. A more accurate assessment and characterisation is increasingly in need due to their impact on the mechanical behaviour of heterogeneous rock masses. The aim of this work was to contribute with innovative methodologies to be used in the geomechanical characterisation of this type of rocks, as well as to reduce the uncertainty associated with the definition of prospection plans.

The main methodology hereby proposed combines geostatistical techniques, which have proved to be efficient in identifying rock mass heterogeneities and in considering the spatial variability, with clustering techniques, so that a more realistic geomechanical model of the rock mass is obtained. Since the geotechnical information used to perform the numerical analysis of underground works has different stages, the proposed methodology was developed to take that feature into account. As a first step, the geotechnical parameters are simulated using geostatistical techniques, followed by their conversion into geomechanical parameters in a second step. A third step concerns the selection of representative realisations, with a minimum of three, to be used in the numerical analysis. Finally, the rock mass characterisation models are imported to a finite differences software to perform a mechanical behaviour analysis of the geotechnical structure. The numerical results are compared with the ones obtained from a deterministic approach that assumes the rock mass as a homogeneous medium. For the methodology validation, real data from two case studies were used: a Chilean gold deposit and the Salamonde II hydroelectric complex recently built in the North of Portugal.

Concerning the methodology for optimizing prospection plans, the same geostatistical techniques were used to simulate the rock mass and to define the uncertainty metrics to use in the optimisation process, which relied on the Simulated Annealing algorithm. The randomness that exists in the definition of boreholes location was the stepping stone of this innovative methodology that intends to be a helping tool for professionals. As an output, one obtains the optimal spatial location for new boreholes, as well as their lengths, in order to decrease the uncertainty of the numerical model and to increase the geomechanical detail. It is worthy of notice that the two above mentioned case studies were also used in this methodology validation.

ACKNOWLEDGEMENTS

Five years ago, I received the news that would mark my professional and personal life and would form me as a better person, that was the acceptance to do my PhD in the University of Minho. That was my institution since 2007 and I could not be prouder to belong to their family. All the accomplishments made during this intense and long journey could not be achieved without the help and support of my supervisors, family and friends.

First of all, I would like to express my gratitude for all the time, suggestions and guidance given to me by my supervisor Professor Tiago Miranda. The regular meetings where we discussed not only the technical parts of this thesis but also the mental force needed to accomplish them are one of the important things that make this work possible. He taught me how to embrace new challenges and to concentrate most of my mental energy in the positive things that happens during my PhD. The rigorousness and innovation were always mentioned as the characteristics to aimed for my work. During these five years working together, I can only hope to work with him in the future.

I would like to acknowledge my co-supervisor, Eng. Luís Lamas for all the sincere discussions and constructive criticism that I channelled into this work, as well as Eng. Margarida Espada for all the help and discussions made, in person and by skype, that were crucial to my work.

I would like to make a special thanks to prof. Xavier Emery, crucial to this work once he taught me one of the most interesting techniques to use in the future geomechanics, the geostatistics. During my time in Chile, he helped me to understand the complexity of the techniques and how to face the scientific community in a fearless way. All the crazy and inspiring ideas for new developments will always be remembered. Another special thank you is intended to Prof. Ana Rocha for teaching and supporting me in the optimisations techniques used in Chapters 6 and 7.

I would also like to acknowledge the University of Minho for welcoming me during this past 10 years and to the Portuguese company, EDP for providing me the information to use in this work, without which I could not reached such detail and rigour.

I also need to thank my friends, without whom I could not ended this long journey, for all their constant support and help: Hélder, Alice, Vânia, Filipa, Vítor, Zeca, Liliana, Pedro. With a special thank you to my dearest friend Diana for all the “look forward” references and confidence given to me and to my work. Also, I could not forget my Chilean friends, Nasser, Mohammad, Amir, Amin

and Daisy for your welcome and support showed during my stay in Chile. Last but, not least, to my Parisian friend Maria, with who I learned and shared so much about this research world.

Finally, I would like to acknowledge and dedicate my PhD thesis to my family and boyfriend for all the patience, support, love and understanding showed throughout this difficult but grateful journey. Diogo thank you for believing in me even when I do not, and for remembering me that we can always do better. Your love and support will always be remembered. *Avó, Mãe, Pai, Cris e Bete obrigada por tudo e, acima de tudo, por acreditarem em mim.*

This research work has been developed in the scope of the Doctoral Programme in Civil Engineering, at the University of Minho, Portugal, and fully supported by the Portuguese Science and Technology Foundation (FCT), under grant no. SFRH/BD/89627/2012. This is gratefully acknowledged.

RESUMO

A incerteza associada à variabilidade espacial e às heterogeneidades, naturalmente presentes aquando a caracterização geomecânica dos maciços rochosos, assumem um papel cada vez mais relevante na prática da engenharia geotécnica. Como tal, existe uma necessidade cada vez maior em obter caracterizações deste tipo de rochas heterogéneas de forma mais precisa e real. Assim, o principal objetivo deste trabalho foi contribuir com o desenvolvimento de metodologias inovadoras para serem usadas, tanto na caracterização geomecânica deste tipo de rochas, como na redução da incerteza associada à elaboração dos planos de prospeção geotécnica.

Assumindo que as técnicas geoestatísticas são eficientes na identificação das heterogeneidades e na consideração da variabilidade espacial dos maciços rochosos, a principal metodologia deste trabalho combina estas técnicas com técnicas de *clustering*, utilizadas para a seleção de modelos geomecânicos mais realistas. Assim, numa primeira fase, é simulada a informação geotécnica utilizando ferramenta geoestatísticas que, numa segunda fase, é convertida em informação relativa aos parâmetros geomecânicos do maciço rochoso. Uma terceira fase da metodologia considera a seleção das realizações representativas para utilizar na análise numérica. Por fim, os modelos numéricos do maciço rochoso são importados para o software de diferenças finitas onde é efetuada uma análise do comportamento mecânico da estrutura geotécnica. Estes resultados são comparados com os obtidos quando considerada uma abordagem determinística que assume o maciço rochoso como um meio homogéneo. Na validação da metodologia foram utilizados dois casos de estudo, um depósito de ouro chileno e o complexo hidroelétrico de Salamonde II, contruído recentemente no norte de Portugal.

No que concerne à metodologia para otimizar os planos de prospeção, foram utilizadas as mesmas técnicas geoestatísticas para a simulação do maciço rochoso e assim definir métricas de avaliação da incerteza a utilizar no processo de otimização que utiliza o algoritmo *Simulated Annealing*. A aleatoriedade inerente à definição dos locais de sondagens foi o ponto de partida para o desenvolvimento de uma segunda metodologia que pretende ser uma ferramenta de apoio aos profissionais da área. Como *output* são obtidos os locais ótimos para a execução de novas sondagens, assim como os seus comprimentos com o objetivo de reduzir a incerteza associada ao modelo numérico do maciço rochosos e aumentar o detalhe geotécnico. É importante referir que os casos de estudo já mencionados também são utilizados na validação desta metodologia.

CONTENTS

Chapter 1 - Introduction.....	1
1.1. Motivation	3
1.2. Scope of the work/Objectives of the thesis	5
1.3. Methodologies and outline of the thesis.....	6
1.3.1. First segment.....	6
1.3.2. Second segment	8
Chapter 2 - Overview on Rock Mass Characterisation Methodologies.....	11
2.1. Introduction.....	13
2.2. Deterministic approaches	15
2.2.1. Rock Quality Designation	16
2.2.2. Rock Mass Rating system.....	17
2.2.3. Q system	18
2.2.4. Geological Strength Index	19
2.3. Probabilistic methods	20
2.3.1. General.....	20
2.3.2. Monte Carlo	21
2.3.3. First Order Second Moment Method	23
2.3.4. Point Estimation Method	24
2.4. Random Field Theory.....	26
2.5. Geostatistical techniques	30
2.5.1. General.....	30
2.5.2. Exploratory and covariance analyses of regionalised variables	33
2.5.3. Variogram	34
2.5.4. Estimation.....	40
2.5.5. Simulation.....	43
2.5.6. Model validation	50

2.5.7.	Geostatistics applied to Geotechnics	51
2.6.	Scenario reduction.....	53
2.6.1.	Introduction	53
2.6.2.	Kernel clustering technique	56
2.7.	Conclusions.....	58
Chapter 3 - Geostatistical Simulation of Geomechanical Parameters		61
3.1.	Introduction.....	63
3.2.	Geostatistical simulation models	63
3.2.1.	General.....	63
3.2.2.	First and second approaches: continuous variables.....	64
3.2.3.	Third approach: discrete variable	66
3.3.	Case study – Chilean deposit.....	67
3.3.1.	Introduction	67
3.3.2.	Exploratory analysis.....	69
3.3.3.	Modelling univariate distributions.....	72
3.3.4.	Modelling spatial continuity.....	75
3.3.5.	Conditional simulation results.....	79
3.3.6.	Post-processing simulations.....	83
3.3.7.	Split-sample validation.....	88
3.3.8.	Discussion of the results.....	91
3.4.	Conclusions.....	92
Chapter 4 - Numerical Methodology to Characterise Heterogeneous Rock Masses....		95
4.1.	Introduction.....	97
4.2.	Methodology outline.....	98
4.2.1.	General.....	98
4.2.2.	Geostatistical simulation	100
4.2.3.	Scenario reduction methodology	101
4.3.	Case study - Chilean deposit.....	105
4.3.1.	Geotechnical data	105

4.3.2.	Three-dimensional numerical model	105
4.3.3.	Geostatistical simulation results.....	106
4.3.4.	Reduction scenario methodology applied to RMR simulations	108
4.3.5.	From geotechnical data to geomechanical parameters	116
4.3.6.	Models in analysis.....	117
4.3.7.	Numerical results.....	122
4.3.8.	Discussion	128
4.4.	Conclusions.....	132
Chapter 5 - Salamonde II Powerhouse Complex: Geomechanical Characterisation and Modelling		135
5.1.	Introduction.....	137
5.2.	Geotechnical prospection and tests.....	139
5.3.	The powerhouse cavern.....	144
5.4.	Numerical model.....	149
5.4.1.	Numerical Model presentation.....	149
5.4.2.	Numerical modelling results	152
5.5.	Application of the new characterisation methodology.....	153
5.5.1.	Introduction	153
5.5.2.	Geostatistical simulation.....	155
5.5.3.	From geotechnical data to geomechanical parameters.....	162
5.5.4.	Scenario reduction methodology.....	164
5.6.	Numerical modelling results.....	174
5.6.1.	Introduction	174
5.6.2.	Data set 1 vs. deterministic model.....	176
5.6.3.	Data set 2 vs. deterministic model.....	185
5.6.4.	Data set 3 vs. deterministic model.....	195
5.6.5.	Discussion of the results.....	202
5.7.	Conclusions.....	205
Chapter 6 - Methodology for Optimisation of Boreholes Plans		209
6.1.	Introduction.....	211

6.2.	Optimisation concepts	214
6.2.1.	Simulated Annealing.....	214
6.2.2.	Multi-criteria optimisation	216
6.3.	Proposed methodology	218
6.3.1.	Overview	218
6.3.2.	Optimisation process.....	219
6.3.3.	Uni-objective approach	222
6.3.4.	Multi-criteria approach.....	223
6.4.	Theoretical application - Chilean deposit.....	225
6.4.1.	Data presentation.....	225
6.4.2.	Geostatistical simulation	227
6.4.3.	Simulated Annealing internal parameters study.....	229
6.5.	Optimisation results.....	231
6.5.1.	Uni-objective results	231
6.5.2.	Multi-criteria results.....	239
6.5.3.	Discussion of the results.....	244
6.6.	Conclusions.....	246
Chapter 7 - Salamonde II Powerhouse Complex: Borehole Optimisation		249
7.1.	Introduction.....	251
7.2.	Data presentation	252
7.3.	Optimisation process	255
7.4.	Geostatistical simulation	258
7.4.1.	Scenario 1	258
7.4.2.	Initial data.....	261
7.5.	Simulated Annealing input parameters study.....	263
7.5.1.	Introduction	263
7.5.2.	Step size value (Scenario 2).....	264
7.5.3.	Points generating function study (Scenarios 2 and 3)	265
7.6.	Optimisation results.....	271

7.6.1.	Scenario 1	271
7.6.2.	Analysis of the results.....	276
7.7.	Conclusions.....	278
Chapter 8 - Conclusions and Future Works.....		281
8.1.	Summary and Main contributions	283
8.2.	Future works	286
Chapter 9 - References		289
Appendix 1		303
Appendix 2		313

LIST OF FIGURES

Figure 1.1 Outline and organisation of the thesis.....	9
Figure 2.1 Scheme with the uncertainties related to geotechnical engineering (adapted from Christian and Baecher (2003)).	15
Figure 2.2 Monte Carlo method principle for an example of three variables used as input.	22
Figure 2.3 Principle of Point Estimate method assuming two random variables (adapted from Rocscience (2016)).	25
Figure 2.4 Example of two random fields Xt with distinctive correlation lengths (θ) (Fenton, 1997).	28
Figure 2.5 Difference between classical statistics and geostatistics: a) independent realisations (classical); and b) dependent realisations (geostatistics).....	30
Figure 2.6 General scheme showing the estimation process in geostatistics (Zhang, 2011).....	32
Figure 2.7 General scheme showing the simulation process in geostatistics (Zhang, 2011).....	32
Figure 2.8 Flow predictions based on geostatistical simulations (Zhang, 2011).....	33
Figure 2.9 Regionalised variables example in $1D$	34
Figure 2.10 Variogram general representation of a stationary random variable (Imanzadeh, 2013).	36
Figure 2.11 Variogram map examples for a random field with: a) an isotropic behaviour; and b) an anisotropic behaviour.....	36
Figure 2.12 Variogram and realisation examples for an anisotropic field (different ranges for N55E and N350).	37
Figure 2.13 Representation image of the variogram parameters (adapted from Deutsch (2015))....	37
Figure 2.14 Variogram models' examples set for a nugget effect equal to zero (adapted from Jaksa et al. (1997)).	38
Figure 2.15 <i>Nested structures</i> theoretical example considering a small <i>nugget</i> effect.....	39
Figure 2.16 Variograms and realisations examples for three fields with distinct spatial behaviour. ...	39
Figure 2.17 Cokriging cases: a) Isotopic; b) Partially heterotopic; and c) Heterotopic (adapted from Wackernagel (2006)).	43

Figure 2.18 Representative images to compare preliminary data configuration, the non-conditional simulation, kriging estimation and conditional simulation (Wackernagel, 2006).	44
Figure 2.19 General steps required in a geostatistical simulation process using the Gaussian random field model.....	46
Figure 2.20 Examples of: a) <i>1D</i> truncation flag; and b) a <i>3D</i> truncation flag.....	47
Figure 2.21 Turning Bands Method geometrical projection (adapted from Wackernagel (2006)).....	48
Figure 2.22 TBM simulation using 1, 10, 100 and 1000 lines.....	49
Figure 2.23 Accuracy plot example for true Hydraulic conductivity values fall within the probability interval. The black points represent the variable prediction using a Gaussian algorithm and the white points an indicator algorithm (Goovaerts, 2001).	51
Figure 2.24 Proposed methodology for uncertainty quantification- (A) distance between two models, (B) distance matrix D, (C) models mapped in Euclidean space, (D) feature space, (E) pre-image construction, (F) P10, P50 and P90 quantile estimations (Scheidt and Caers, 2009b).....	55
Figure 3.1 Flow charts for the variable simulation under: a) the Gaussian model (Approaches 1 and 2); and b) the truncated Gaussian model (Approach 3). 65	
Figure 3.2 Scheme showing the three approaches considered in the RMR simulation.	67
Figure 3.3 <i>2D</i> maps of spatial distribution at elevation 3439 m for: a) RMR original values; b) Parameter P2; and c) Parameter P3.....	69
Figure 3.4 Data histograms for the RMR individual parameters in a rating scale for: a) P1; b) P2; c) P3; and d) P5.	71
Figure 3.5 Data histograms for the RMR individual parameters in their original scale for: a) P1 (in MPa); b) P2; c) P3 (in mm); and d) P5 (wet or damp).....	72
Figure 3.6 Anamorphosis function used for Approach 1. The ordinate indicates the RMR value and the abscissa the associated Gaussian value.....	73
Figure 3.7 Experimental (crosses) and theoretical (solid lines) variograms for RMR (Approach 1) along the main anisotropy directions, horizontal plane (black) and vertical direction (blue).....	77
Figure 3.8 Experimental (crosses) and theoretical (solid lines) variograms for Approach 2 along the main anisotropy directions, horizontal plane (black) and vertical direction (blue) for parameters: a) P1; b) P2; c) P3; and d) P5.....	77

Figure 3.9 Experimental (crosses) and theoretical (solid lines) variograms for Approach 3 along the main anisotropy directions, horizontal plane (black) and vertical direction (blue) for parameters: a) P1; b) P2; c) P3; and d) P5.....	78
Figure 3.10 Maps of RMR at elevation 3560 m of the average of 100 realisations obtained with: a) Approach 1; b) Approach 2; and c) Approach 3.	80
Figure 3.11 Maps of RMR at elevation 3560 m for realisation #1 obtained with: a) Approach 1; b) Approach 2; and c) Approach 3.....	81
Figure 3.12 Maps of RMR at elevation 3560 m of the standard deviation of 100 realisations obtained with: a) Approach 1; b) Approach 2; and c) Approach 3.	82
Figure 3.13 Maps of probability (between 0 and 1) that the RMR is less than a threshold of 65 at elevation 3560 m, obtained with: a) Approach 1; b) Approach 2; and c) Approach 3.	84
Figure 3.14 Maps of discontinuity parameters at elevation 3560 m for realisation #1 of a) P2; and b) P3.	85
Figure 3.15 Maps of discontinuity parameters at elevation 3560 m for the average of 100 realisations of a) P2; and b) P3.	85
Figure 3.16 Maps of discontinuity parameters at elevation 3560 m for the standard deviation of 100 realisations of: a) P2; and b) P3.	86
Figure 3.17 Maps of deformation modulus (in GPa) at elevation 3560 m obtained with Approach 1, for: a) realisation #1; and b) average of 100 realisations.....	87
Figure 3.18 Standard deviation of 100 realisations map of deformation modulus (in GPa) at elevation 3560 m obtained with Approach 1.....	87
Figure 3.19 Scatter plots between true and expected RMR values for: a) Approach 1; b) Approach 2; and c) Approach 3.	89
Figure 3.20 Accuracy plots showing the uncertainty modelling of the simulations for: a) Approach 1; b) Approach 2; and c) Approach 3.....	91
Figure 4.1 General workflow containing the general steps of the proposed characterisation methodology.....	99
Figure 4.2 Scheme to follow in order to apply the scenario reduction methodology proposed in this section.	104
Figure 4.3 Finite difference grid of tunnel model: a) xyz perspective; b) xz plan view of the tunnel zone.	106

Figure 4.4 Experimental (crosses) and theoretical (solid lines) variograms along the main anisotropy directions: horizontal plane (black) and vertical (blue): a) RMR; and b) UCS.....	107
Figure 4.5 3D maps of RMR simulated on Flac ^{3D} mesh, for: a) realisation #1; and b) average of 100 realisations.	108
Figure 4.6 2D spatial representation of the RMR 100 realisations (black points) using the Euclidean distance.....	110
Figure 4.7 Average silhouette width values for the performed clusters evaluation.	111
Figure 4.8 Clusters final configuration (points) with the matching medoids (squares) for: a) 2 clusters; and b) single cluster.....	112
Figure 4.9 Validation of the two clusters configurations compared with the 100 realisations set using the percentiles 10, 50 and 90, for: a) average percentiles values for the 2 clusters configuration; b) point by point percentile sum for the 2 clusters configuration; and c) average percentiles values for the single cluster configuration.....	114
Figure 4.10 Workflow applied to build the different models to represent the rock mass characterisation of the Chilean rock mass using as an input the RMR system.....	119
Figure 4.11 Flac ^{3D} xyz perspective with Em values (colour scale in Pa) used in: a) model 1; b) model 2 – cluster 1; c) model 2 – cluster 2; and d) model 3.....	121
Figure 4.12 XZ plane at y=0 with Flac ^{3D} contour of displacement (colour scale in m) of the rock mass after the excavation for: a) model 1; b) model 2 – cluster 1; c) model 2 – cluster 2; d) model 3; and e) model 4.....	125
Figure 4.13 XZ plane at y=0 with Flac ^{3D} contour of the principal maximum stresses (compression-colour scale in Pa) after the excavation process for: a) model 1; b) model 2 – clusters 1; c) model 2 – cluster 2; d) model 3; and e) model 4.....	126
Figure 4.14 Representation of the displacements (in mm) obtained for all the analysed models (104), the single cluster of model 1 (Single), cluster 1 (C1) and 2 (C2) of model 2, the deterministic model 4 (Homogeneous) and the 100 realisations, in: a) centre of the tunnel arch; b) centre of the tunnel invert.....	127
Figure 4.15 Representation of the displacements (in mm) obtained for all the analysed models (104), the single cluster of model 1 (Single), cluster 1 (C1) and 2 (C2) of model 2, the deterministic model 4 (Homogeneous) and the 100 realisations, in: a) centre of left sidewall; b) centre of the right sidewall.	128

Figure 4.16 Representation of the principal stresses (in MPa) obtained for all the analysed models (104), the single cluster of model 1 (Single), cluster 1 (C1) and 2 (C2) of model 2, the deterministic model 4 (Homogeneous) and the 100 realisations: a) maximum principal stress; b) minimum principal stress.....	128
Figure 4.17 100 realisations histograms and distribution fitting curve and models 1 to 5 values (lines) of displacements in: a) centre of the tunnel arch; b) centre of the tunnel invert.	130
Figure 4.18 100 realisations histograms and distribution fitting curve and models 1 to 5 values (lines) of displacements in: a) centre of left sidewall; b) centre of the right sidewall.....	131
Figure 4.19 100 realisations histograms and distribution fitting curve and models 1 to 5 values (lines) of the principal stresses (in MPa) for: a) maximum principal stress; b) minimum principal stress.	131
Figure 5.1 General overview of the Salomonde II hydroelectric circuit implementation (adapted from EDP (2009)).	137
Figure 5.2 Salomonde II underground structures detailed scheme (adapted from EDP (2009)).	138
Figure 5.3 Salomonde II powerhouse general plant with the <i>in situ</i> tests identification (LNEC, 2012).	140
Figure 5.4 Histograms obtained using the phase 1 information of Salomonde II, for: a) E_i ; b) RQD; and c) W	144
Figure 5.5 Powerhouse complex geometry and elements (EDP, 2009).	145
Figure 5.6 Histograms obtained using data set 1 of Salomonde II phase 2 information, for: a) RMR; b) Log Q; and c) GSI.	147
Figure 5.7 Cross section of the powerhouse cavern showing the monitoring plan adopted (adapted from EDP (2009)).	148
Figure 5.8 Histograms obtained using data set 2 of Salomonde II phase 2 information, for: a) E_i (in GPa); b) RQD (%); and c) W	149
Figure 5.9 3D mesh developed for the Salomonde II powerhouse complex (adapted from Espada and Lamas (2014)).....	150
Figure 5.10 Detail of Salomonde II numerical mesh with the identification of the considered damage zone near the cavern (adapted from LNEC (2013)).....	152
Figure 5.11 General workflow containing the general steps of the characterisation methodology using in Salomonde II.	155

Figure 5.12 Experimental (crosses) and theoretical (solid lines) variograms along the horizontal plane xy plane for E_m using information from data set 1, for: a) E_i ; b) W ; and c) RQD.....	157
Figure 5.13 Experimental (crosses) and theoretical (solid lines) variograms along the horizontal plane xy plane for E_m using information from data set 2, for: a) RMR; b) Log Q ; and c) GSI.	157
Figure 5.14 Experimental (crosses) and theoretical (solid lines) variograms along the horizontal plane xy plane for E_m using information from data set 2, for: a) E_i ; b) W ; and c) RQD.....	158
Figure 5.15 Accuracy plots of data set 1 showing the uncertainty modelling for: a) E_i (in GPa); b) RQD; and c) W	160
Figure 5.16 Accuracy plots of data set 2 showing the uncertainty modelling for: a) E_i (in GPa); b) RQD; and c) W	161
Figure 5.17 Accuracy plots of data set 2 showing the uncertainty modelling for: a) RMR; b) Log Q ; and c) GSI.	162
Figure 5.18 Summary of the geotechnical information that compose each considered data set...	163
Figure 5.19 Histograms of the E_m values (in GPa) of Salamonde II obtaining for all the 100 realisation (each colour represents one realisation), for: a) data set 1; b) data set 2; and c) data set 3.	164
Figure 5.20 2D spatial representation of the E_m 100 realisations (black points) using the Euclidean distance computed with data from: a) data set 1; b) data set 2; and c) data set 3.	165
Figure 5.21 Average silhouette width values for the performed clusters evaluation using the data from: a) data set 1; b) data set 2; and c) data set 3.	166
Figure 5.22 Clusters final configuration (points) with the matching medoids (squares) for: a) data set 1 (3 clusters); b) data set 2 (9clusters); and c) data set 3 (4 clusters).....	167
Figure 5.23 Validation of clusters configurations in comparison with the 100 realisations set using the average values of E_m for percentiles 10, 50 and 90, for: a) data set 1 (3 clusters); b) data set 2 (9clusters); and c) data set 3 (4 clusters).....	168
Figure 5.24 Validation of clusters configurations in comparison with the 100 realisations set after summing point by point the E_m values of percentiles 10, 50 and 90, for: a) data set 1 (3 clusters); b) data set 2 (9 clusters); and c) data set 3 (4 clusters).	169

Figure 5.25 Validation of clusters configurations using a kernel smoothing function to represent the selected and full realisations sets along with the average of the 100 realisations, for: a) data set 1 (3 clusters); b) data set 2 (9 clusters); and c) data set 3 (4 clusters).	171
Figure 5.26 Workflow applied to build the data sets to represent the rock mass characterisation of Salamonde II using as input the geotechnical information from two distinct phases.	176
Figure 5.27 2D contours of the displacements magnitude (in m) and vectors for section A in the last excavation stage, for: a) the Homogeneous model; and b) cluster 3 of data set 1.	180
Figure 5.28 2D contours of the displacements magnitude (in m) and vectors for section B in the last excavation stage, for: a) the Homogeneous model; and b) cluster 1 of data set 1.	180
Figure 5.29 3D contours of the displacements magnitude (in m) with a deformed factor equal to 1000 in the last excavation stage, for: a) the Homogeneous model; and b) cluster 2 of data set 1.	180
Figure 5.30 3D maximum principal stresses (in Pa) contours at the last excavation stage, for: a) the Homogeneous model; and b) cluster 3 of data set 1.	181
Figure 5.31 Displacements evolution measured by all the extensometers installed in: a) section A; and b) section B.	182
Figure 5.32 Total displacements (in mm) in the longer rod measured (dashed line) and computed (dotted and solid lines) for data set 1, for: a) E4A; and b) E4B.	183
Figure 5.33 Total displacements (in mm) in the longer rod measured (dashed line) and computed (dotted and solid lines) for data set 1, for: a) E5A; and b) E5B.	184
Figure 5.34 Total displacements (in mm) in the longer rod measured (dashed line) and computed (dotted and solid lines) for data set 1 in section A, for: a) E6; and b) E7.	184
Figure 5.35 2D contours of the displacements magnitude (in m) and vectors for section A in the last excavation stage, for: a) the LNEC model (displacements with a scale of 400); b) cluster 5 of data set 2; and c) the Homogeneous model	189
Figure 5.36 2D contours of the displacements magnitude (in m) and vectors for section B in the last excavation stage, for: a) the LNEC model (displacements with a scale of 400); b) cluster 6 of data set 2; and c) the Homogeneous model.	190
Figure 5.37 3D contours of the displacements magnitude (in m) with a deformed factor equal to 500 in the last excavation stage, for: a) the LNEC model; b) cluster 1 of data set 2; and c) the Homogeneous.	191

Figure 5.38 2D maximum principal stresses (in Pa) contours in section A at the last excavation stage, for: a) the LNEC model; b) cluster 8 of data set 2; and c) the Homogeneous model.	192
Figure 5.39 Total displacements (in mm) in the longer rod measured (dashed line) and computed (dotted and solid lines) for data set 2, for: a) E4A; and b) E4B.	193
Figure 5.40 Total displacements (in mm) in the longer rod measured (dashed line) and computed (dotted and solid lines) for data set 2, for: a) E5A; and b) E5B.	194
Figure 5.41 Total displacements (in mm) in the longer rod measured (dashed line) and computed (dotted and solid lines) for data set 2 in section A, for: a) E6; and b) E7.	194
Figure 5.42 2D contours of the displacements magnitude (in m) and vectors for section A in the last excavation stage, for: a) the LNEC model (displacements with a scale of 400); b) cluster 2 of data set 3; and c) the Homogeneous model.	199
Figure 5.43 2D contours of the displacements magnitude (in m) and vectors for section B in the last excavation stage, for: a) the LNEC model (displacements with a scale of 400); b) cluster 3 of data set 3; and c) the Homogeneous model.	200
Figure 5.44 3D contours of the displacements magnitude (in m) with a deformed factor equal to 500 in the last excavation stage, for: a) the LNEC model; b) cluster 2 of data set 3; and c) the Homogeneous.	201
Figure 5.45 2D maximum principal stresses (in Pa) contours in section A at the last excavation stage, for: a) the LNEC model; b) cluster 1 of data set 3; and c) the Homogeneous model.	202
Figure 5.46 Graphical representation of the minimum, mean and maximum values of the maximum displacements registered in each analysed numerical model (data set 1, 2 and 3 and corresponding Homogeneous models). The red line represents the maximum displacement obtained by LNEC(2013) model for terms of compression.	205
Figure 6.1 Identification of all the optimisation methodologies and their limitations along the years.	213
Figure 6.2 Simulated Annealing functioning method - hill climbing (adapated from Zhigljavsky, 1991).	215
Figure 6.3 Diagram of the proposed methodology combining geostatistical simulation and simulated annealing.	219
Figure 6.4 Simulated annealing workflow adapted to the proposed methodology.	221

Figure 6.5 Mapping of the initial boreholes data used in the isolated points scenario in: a) XY plane; and b) XYZ perspective (X, Y and Z in meters).....	226
Figure 6.6 Mapping of the initial boreholes data used in the vertical alignments scenario in: a) XY plane; and b) XYZ perspective (X, Y and Z in meters).....	226
Figure 6.7 Experimental (crosses) and theoretical (solid line) variograms along the horizontal plane using: a) 22 isolated points; and b) 6 point alignments.	228
Figure 6.8 RMR preliminary data geostatistical simulation results: a) first realisation; and b) variance of 100 realisations.....	229
Figure 6.9 Convergence study for the number of additional points considering the average width of 95% probability intervals (black line) and the average variance (red line).....	233
Figure 6.10 3D location of 12 additional points (colour points) with preliminary data represented by black points for: a) average width of 95% probability intervals; and b) average variance.	234
Figure 6.11 3D representation of the block used in the search when 12 points were added, for: a) average width of 95% probability intervals; and b) average variance.....	235
Figure 6.12 Convergence study for the number of additional alignments considering the average width of 95% probability intervals (black line) and the average variance (red line).....	237
Figure 6.13 3D location of 3 additional point alignments (colour points) with preliminary data represent by black points for: a) average width of 95% probability intervals; and b) average variance.....	237
Figure 6.14 Convergence study relating the maximum number of iterations with the objective function value using: a) and b) variance as the objective function; c and d) 95% probability interval as the objective function obtained for 12 isolated points (a and c) and 3 point alignments (b and d).....	239
Figure 6.15 Isolated points objective function values for each weight combination test.....	241
Figure 6.16 Representation of the objective functions values (95% probability interval, variance and both functions) obtained in the multi-criteria approach.	241
Figure 6.17 Point alignments objective function values for each weight combination test.	243
Figure 6.18 Representation of the objective functions values (95% probability interval, variance and both functions) obtained in the multi-criteria approach of the point alignments information.	243
Figure 6.19 Spatial representation (XY plane) of the optimal points alignments (colour points) obtained for the variance (black plus) and 95% probability interval (red circles) as the objective functions.	245

Figure 6.20 Spatial representation (XY plane) of the optimal points alignments (colour points) obtained in all the tests performed in a multi criteria optimisation.	245
Figure 7.1 Section of the geological and prospection location plant with the identification of the selected boreholes (colour circles) (adapted from EDP (2009)).	253
Figure 7.2 Spatial distribution of the boreholes with codes S1, S2, S5, S8, S9, S15 and S16 with a colour scale representing: a) RQD; b) F; and c) W.	255
Figure 7.3 Spatial distribution of the remaining boreholes with a colour representing the RQD in: a) scenario 1; b) scenario 2; and c) scenario 3.	258
Figure 7.4 Experimental (crosses) and theoretical (solid line) variograms along the horizontal plane using the information of 6 boreholes, for: a) RQD; b) F; and c) W.	260
Figure 7.5 Experimental (crosses) and theoretical (solid line) variograms using the initial data set of 7 boreholes, for: a) RQD; b) F; and c) W.	262
Figure 7.6 Convergence study for a total of seven tests using as information the scenario 2.	265
Figure 7.7 Convergence analysis comparing the number of iterations with the OF value for tests number 2a and 4a in each run for: a) scenario 2; and b) scenario 3.	268
Figure 7.8 Spatial representation (XY plane) of the optimal points obtained for test number 2a and 4a in all the runs along with the original locations of the two removed boreholes (blue circles), for: a) scenario 2; and b) scenario 3.	270
Figure 7.9 Scenario 2 spatial representation (XY plane) of the optimal points movements (coloured arrows) obtained from run to run, in: a) test 2a; and b) test 4a.	270
Figure 7.10 Scenario 3 spatial representation (XY plane) of the optimal points movements (coloured arrows) obtained from run to run using data from scenario 2, in: a) test 2a; and b) test 4a.	270
Figure 7.11 Spatial representation (XY plane) of the optimal points (colour points) obtained in all the tests, for: a) trio 1; b) trio 2; c) trio 3; d) trio 4; and e) trio 5.	274
Figure 7.12 Convergence study for all the tests performed assuming the weights values from: a) trio 1; b) trio 2; c) trio 3; d) trio 4; and e) trio 5.	275
Figure 7.13 Spatial representation (XY plane) of the optimal points (colour points) obtained in all the tests for all the five trios.	277
Figure A1.1 Histograms obtained using all the phase 1 information of Salamonde II, for: a) F; and b) σ_c	306

Figure A1.2 Representative scheme identifying all the zones from where the geotechnical mappings were recovered (adapted from EDP (2009)).....	307
Figure A1.3 The cell declustering mean versus 25 cell sizes (each value represents a $3D$ square in m) for: a) data set 1; and b) data set 2.....	309
Figure A1.4 Anamorphosis function with the E_m values in GPa (ordinate) and the associated Gaussian value (abscissa) for: a) data set 1; and b) data set 2.....	310
Figure A2.1 Spatial plot for test number 1 using the 'v4' type of interpolation to obtain the colour surface for; a) PI95%; and b) Var.....	315
Figure A2.2 Spatial plot for test number 2 using the 'v4' type of interpolation to obtain the colour surface for; a) PI95%; and b) Var.....	315
Figure A2.3 Spatial plot for test number 3 using the 'nearest' and 'v4' type of interpolation to obtain the colour surface, respectively for; a) PI95%; and b) Var.....	316

LIST OF TABLES

Table 2.1 Empirical formulas to obtain E_m using the RQD index value.....	16
Table 2.2 RMR rock quality scale and description.	17
Table 2.3 Empirical formulas to obtain E_m using the RMR value.	18
Table 2.4 Empirical formulas to obtain E_m using the Q system value.	19
Table 2.5 Empirical formulas to obtain E_m using the GSI system value.	20
Table 2.6 Mathematical functions and respective models to use in variogram fitting (Jaksa et al., 1997).....	38
Table 2.7 General procedure to adapt for k-medoid algorithm implementation.	58
Table 3.1 Information about average UCS and average density by lithological unit.	68
Table 3.2 Basic statistics on RMR ratings and original data (3969 samples).	69
Table 3.3 Correlation matrix between parameters P1, P2, P3 and P5.....	70
Table 3.4 Calculated proportions for P2 data with the corresponding Gaussian thresholds.	74
Table 3.5 Calculated proportions for P3 with the corresponding Gaussian thresholds.....	75
Table 3.6 Statistics on the average of 100 conditional realisations of RMR obtained with Approaches 1, 2 and 3.	79
Table 4.1 Silhouette reference values used to evaluate the optimum number of clusters (Struyf et al., 1997).....	103
Table 4.2 RMR and UCS (in MPa) statistics for the initial values and the average of 100 conditional realisations.	108
Table 4.3 Extract of the RMR 100 realisations dissimilarity matrix (each column and line represent the realisation number).	109
Table 4.4 Basic statistics of the RMR variable for the initial values, 100 realisations and selected realisation data sets.	115
Table 4.5 Basic statistics of the RMR for the individual realisations of the two clusters, single cluster configurations and 100 realisations average values.....	115
Table 4.6 Empirical expressions used to obtain E_m and the corresponding authors.	117

Table 4.7 Rock mass deformation modulus (E_m in GPa) basic statistics of model 1 (single cluster), model 2 (two clusters) and model 3 (100 realisations average).....	120
Table 4.8 Displacements and stresses values obtained for the 100 individual realisations.....	122
Table 4.9 Summary of the displacements and stresses results obtained in each one of the first four models.	123
Table 4.10 Percentage differences between the displacements values obtained for model 1 and models 2 in relation to model 4 (homogeneous model).....	129
Table 4.11 Distribution curve fitting details (first three moments) regarding the 100 realisations obtained value for displacements and principal stresses.	132
Table 5.1 STT tests results and rock characteristics (adapted from LNEC (2009)).	141
Table 5.2 SFJ tests results (adapted from LNEC (2012)).....	141
Table 5.3 List of the geotechnical variables that compose the data set 1 and data set 2 of phase 1 information.	143
Table 5.4 Statistical analysis of the phase 1 geotechnical information of Salamonde II (E_i , W and RQD).	143
Table 5.5 Geotechnical zoning characteristics (EDP, 2009).....	146
Table 5.6 Statistics analysis of data set 1 of the phase 2 geotechnical information of Salamonde II (RMR , $\log Q$ and GSI).	146
Table 5.7 Statistical analysis of data set 2 of the phase 2 geotechnical information of Salamonde II (F , W , RQD and E_i).	148
Table 5.8 Adopted construction stages for Salamonde II 3D numerical model.	151
Table 5.9 Summary of the displacements and stresses for the last excavation stage (adapted from LNEC (2013)).	153
Table 5.10 Correlation matrix between parameters P_1 , P_2 , P_3 and P_5	156
Table 5.11 Statistical analysis of the variable from phase 1 and 2 after the geostatistical simulation.	159
Table 5.12 Statistical analysis of E_m (in GPa) obtained values for data set 1, 2 and 3 of the geotechnical information.	164
Table 5.13 Basic statistics of the E_m values (in GPa) for the initial values, average of the 100 realisations and selected realisations for data set 1.	171

Table 5.14 Basic statistics of the E_m values (in GPa) for the initial values, average of the 100 realisations and selected realisations for data set 2.	172
Table 5.15 Basic statistics of the E_m values (in GPa) for the initial values, average of the 100 realisations and selected realisations for data set 3.	172
Table 5.16 Basic statistics of the E_m values (in GPa) for the individual realisation that compose the three-clusters configuration of data set 1.	173
Table 5.17 Basic statistics of the E_m values (in GPa) for the individual realisation that compose the nine-clusters configuration of data set 2.	173
Table 5.18 Basic statistics of the E_m values (in GPa) for the individual realisation that compose the four-clusters configuration of data set 3.	174
Table 5.19 Summary of the maximum displacements (in mm) obtained for data set 1 - 3 clusters, Mean and Homogeneous models and the differences (in percentage) in relation to the Homogeneous.	178
Table 5.20 Summary of the maximum principal stresses (in MPa) obtained for data set 1 – 3 clusters, Mean and Homogeneous models (all the values should be negative corresponding to compressive stresses) and the differences (in percentage) in relation to the Homogeneous.	179
Table 5.21 Summary statistics of the displacements (in mm) regarding the 3 clusters of data set 1 in all the analysed zones.	179
Table 5.22 MAD values (in mm) obtain between the computed and the measured displacements for all the extensometers for data set 1.	185
Table 5.23 Summary of the maximum displacements (in mm) obtained for data set 2 - 9 clusters, LNEC (2013), Mean and Homogeneous models and the differences (in percentage) in relation to the Homogeneous.	187
Table 5.24 Summary of the maximum principal stresses (in MPa) obtained for data set 2 – 9 clusters, LNEC (2013), Mean and Homogeneous models (all the values should be negative corresponding to compressive stresses) and the differences (in percentage) in relation to the Homogeneous.	188
Table 5.25 Summary statistics of the displacements (in mm) regarding the 9 clusters of data set 2 in all the analysed zones.	189
Table 5.26 MAD values (in mm) obtain between the computed and the measured displacements for all the extensometers for data set 2.	195

Table 5.27 Summary of the maximum displacements (in mm) obtained for data set 3 - 4 clusters, LNEC (2013), Mean and Homogeneous models and the differences (in percentage) in relation to the Homogeneous.....	196
Table 5.28 Summary of the maximum principal stresses (in MPa) obtained for data set 3 – 4 clusters, LNEC (2013), Mean and Homogeneous models (all the values should be negative corresponding to compressive stresses) and the differences (in percentage) in relation to the Homogeneous.	198
Table 5.29 Summary statistics of the displacements (in mm) regarding the 4 clusters of data set 3 in all the analysed zones.	198
Table 5.30 Average values of the MAD mean values (in mm) obtain between the computed and the measured displacements for all the extensometers and the standard average values.....	204
Table 6.1 First step to calculate the variance value.	223
Table 6.2 Basic statistics on RMR preliminary data and their Gaussian transforms.	228
Table 6.3 SA internal parameters values assumed for the sensitivity test performed.	230
Table 6.4 SA internal parameters.	230
Table 6.5 Isolated points optimisation results when the 95% probability interval (PI95%) was the objective function.....	232
Table 6.6 Isolated points optimisation results when the variance (Var) average value was the objective function.....	232
Table 6.7 Summary of the SA results for each additional point using two different objective functions (PI95% and Var).....	233
Table 6.8 Point alignments optimisation results when the 95% probability interval (PI95%) was the objective function.....	236
Table 6.9 Point alignments optimisation results when the variance (Var) average value was the objective function.....	236
Table 6.10 Summary of the SA results for each additional alignment using two different objective functions (PI95% and Var).	236
Table 6.11 Isolated points optimisation results for each combination of weights for both objective functions along with the combination of the two objective functions (Both).	240
Table 6.12 Point alignments optimisation results for each combination of weights for both objective functions along with the combination of the two objective functions (Both).	242

Table 7.1 Basic statistics of the initial data set composed by boreholes with codes S1, S2, S5, S8, S9, S15 and S16. 254	
Table 7.2 Basic statistics of boreholes with codes S1, S2, S8, S9, S15 and S16 (6 boreholes). ... 259	
Table 7.3 Correlation matrix between RQD, F and W. 259	
Table 7.4 Basic statistics using the 6 boreholes data set to perform the geostatistical simulation of variables RQD, F and W. 261	
Table 7.5 Basic statistics using the 7 boreholes data set to perform the geostatistical simulation of variables RQD, F and W. 263	
Table 7.6 Input parameters values for all the seven tests performed using information from scenario 2 to study the step size value. 264	
Table 7.7 Input parameters values for tests number 2 and 4 using information from scenarios 2 and 3 to study the points generating function. 266	
Table 7.8 Scenario 2 optimisation results of test number 2a. 266	
Table 7.9 Scenario 3 optimisation results of test number 2a. 267	
Table 7.10 Scenario 2 optimisation results of test number 4a. 267	
Table 7.11 Scenario 3 optimisation results of test number 4a. 267	
Table 7.12 Input parameters values adopted to perform the optimisation using the data from scenario 1. 271	
Table 7.13 Optimisation results using the weight values from trio number 1. 272	
Table 7.14 Optimisation results using the weight values from trio number 2. 272	
Table 7.15 Optimisation results using the weight values from trio number 3. 273	
Table 7.16 Optimisation results using the weight values from trio number 4. 273	
Table 7.17 Optimisation results using the weight values from trio number 5. 273	
Table 7.18 Results comparison between the optimised objective function values and the initial ones considering the weight values from all the five trios. 277	
Table 7.19 Comparative analysis using the optimisation results of scenarios 2 and 3 with the OF value obtained from the initial set simulation. 278	
Table A1.1 Summary of the boreholes main characteristics (EDP, 2009). 304	
Table A1.2 Summary of the uniaxial compressive strength tests performed on the borehole 305	
Table A1.3 Statistical analysis of the geotechnical information that composes data set 1 of phase 1 (RQD, F and W). 306	

Table A1.4 Statistical analysis of the geotechnical information that composes data set 2 of phase 1 (E_i , W , σ_c and ν).	306
Table A1.5 List of all the structures considered in this phase 2 information recovery.....	307
Table A1.6 Information regarding the coordinates of the powerhouse cavern edge points for an elevation equal to 126 m.	308
Table A1.7 Details about the mapping registered for the left side wall of the cavern.	308
Table A1.8 Final layout of the recovered information for the left side Wall of the cavern.	308
Table A1.9 Calculated proportions for W of data set 1 with the corresponding Gaussian thresholds.	310
Table A1.10 Calculated proportions for W of data set 2 with the corresponding Gaussian thresholds.	310
Table A1.11 Empirical expressions used to obtain E_m and the corresponding authors for both phases of geotechnical information.	311
Table A1.12 E_m (in GPa) basic statistics after applying the empirical formulas selected for data set 3.	311
Table A1.13 Statistical analysis of E_m (in GPa) obtained values for the deterministic approaches to use for comparison with data sets 1, 2 and 3.	312
Table A2.1 Input parameters for all the tests.....	314

SYMBOLS

ABREVIATIONS

<i>1D</i>	One dimension
<i>2D</i>	Two dimensions
<i>3D</i>	Three dimensions
SAW	Silhouette Average Width
D	Disturbance factor
E_i	Deformability modulus of the intact rock mass
E_m	Deformability modulus of the rock mass
H	Depth
k_0	Ratio between the horizontal and vertical effective stresses
LFJ	Large Flat Jack test
LNEC	National Laboratory of Civil Engineering
EDP	Electricity of Portugal
ISRM	International Society of Rock Mechanics
MAD	Mean Absolute Deviation
Q	Q- system index
R^2	Determination coefficient
RQD (P2)	Rock Quality Designation
RMSE	Root Mean Squared Error
RMR	Rock Mass Rating
SFJ	Small Flat Jack test
UCS (P1)	Uniaxial Compressive Strength
STT	Strain Tensor Tube
PLT	Point Load Test
GSI	Geological Strength Index

JS (P3)	Joint Spacing parameter of the RMR system
JC (P4)	Joint Condition parameter of the RMR system
GW (P5)	Groundwater condition parameter of the RMR system
W/WD	Weathering degree of the rock mass
F	Fracturing degree of the rock mass
J_r	Joint roughness
J_w	Degree of water inflow on the joint
J_a	Degree of alteration or filling of the joint
J_n	Number of joint sets
SRF	Rock mass stress condition
FOSM	First Order Second Moment method
PEM	Point Estimate Method
MCM	Monte Carlo Method
SF	Safety Factor
TBM	Turning Bands Method
DZ	Damage Zone
SGS	Sequential Gaussian Simulation
PCA	Principal Component Analysis
MDS	Multi-Dimensional Scaling
<i>FISH</i>	Scripting language used in Flac ^{3D}
R	Euclidean space
F	Featured space
AI	Admissible Interval methodology
PSO	Particle Swarm Optimisation algorithm
DIRECT	Dividing Rectangles algorithm
T	Temperature parameter of the SA algorithm
S_n	New sampling design for the boreholes
SA	Simulated Annealing algorithm

OF	Objective Function
Var	Variance as an objective function
PI	Probability interval
$\overline{PI}_{95\%}$	95% probability interval as an objective function
P_{10}	10 th percentile
P_{50}	50 th percentile
P_{90}	90 th percentile
$Z(x)$	Random variable in location x
C	Covariance between two points
E	Expectation of a random variable
Y	Gaussian field representation
G	Standard Gaussian cumulative distribution function
Flac ^{3D}	Fast Lagrangian Analysis of Continua in three dimensions
LW	Left sidewall of the cavern
RW	Right sidewall of the cavern
A	Arch of the cavern
F	Floor of the cavern
NT	North top of the cavern
ST	South top of the cavern
SA	Section A of the cavern
SB	Section B of the cavern

GREEK LETTERS

μ	Mean
σ^2	Variance
σ	Standard deviation
ν	Poisson coefficient
σ_c	Unconfined compressive strength of the intact rock
γ	Variogram

ρ	Correlation coefficient
s	Skewness
ψ	Anamorphosis function

Chapter 1

INTRODUCTION

1.1. MOTIVATION

The characterisation of rock masses in a more accurate and realistic way, mainly in what refers to geomechanical parameters, is an important task to accomplish in geotechnical engineering. The role that a good cost management plays in the engineering world is increasingly important; however, this management should be balanced with a more detailed and complete characterisation of the rock masses under analysis. Therefore, managing these inversely proportional factors is, nowadays, a difficult task for engineers, especially in underground works where the uncertainties must be reduced and the safety of the workers guaranteed at the smallest cost possible. So, the urgent need of new characterisation methodologies to use in geotechnical works, namely in rock masses that show some heterogeneities (like the granites that could be found in the north of Portugal), is undeniable. These heterogeneities imply the transitions from zones with higher resistance and rigidity to others with lower resistance and rigidity, and all that can happen at small distances making the mechanical behaviour of these type of rock masses more difficult to predict. For this reason, the rock mass characterisation has been beheld as a very complex task and subject of study for many years. According to Karim et al. (2007), most of the accidents that happens during the construction phase are due to the adverse circumstances not detected during the design phase, mainly because the rock mass characterisation model is incomplete and not able to consider in a joint way, the uncertainties associated with the characterisation, the spatial variability of the geomechanical parameters and the heterogeneities (this last point is more difficult to assess).

In a 2006 report from the National Research Council the difficulty in dealing with the uncertainty was referred, mentioning that "*it is poorly understood and practiced*", as well as the need for "*new methods to assess the potential impacts of these uncertainties on engineering decisions*". Moreover, and according to the prognoses of the European Commission, the growth in traffic is expected to double in 2012. Consequently, the exploration of underground space can be an answer to solve the problems arising from this traffic growth; however, the complexity associated with the geotechnical area requires the use of new methodologies that are able to reduce, in a considerable way, the uncertainties and spatial variability in order to minimise the problems that could emerge from the underground excavation (Lei, 2016).

In addition, another challenge existing in underground works lies in the difficulty in dealing with deterministic values, be they displacements in the rock mass, stresses or even geomechanical

parameters. The approval of using these deterministic values in one of the most uncertain areas of the engineering is difficult to accept. In this context, the use of probabilistic techniques has been increasing in the last decade (Christian and Baecher, 2003; Le et al., 2014; Popescu et al., 2005; Sejnoba et al., 2007 and Fenton, 1997), aiming to obtain a range of values and safety factors instead of the deterministic values. Yet, the proposed methodologies still show some practical and implementation limitations justifying their underuse.

Although, the most used probabilistic techniques, namely the First Order and Second Moment method or even the Monte Carlo method, consider the variable as spatially independent, which does not correspond to the reality of ground formations, since they show some spatial dependency that should be taken into account. This creates one of the points to investigate throughout this thesis.

Apart from that, and still in the geotechnical characterisation subject, the definition of the borehole plans is a critical point in what concerns cost optimisation. Currently, and in order to obtain a detailed characterisation of the rock masses, the executed plans are mainly composed by mechanical boreholes executed from the surface. The boreholes location is somewhat random, namely because they are based on the professional experience and therefore could be considered as user-dependent. In addition, the costs associated with these prospection plans are considerably high and most of the times a relevant slice of the underground work budget. Due to that, a more optimised and systematic methodology to define the boreholes location is of need. This methodology should allow obtaining the optimal position of new boreholes that lead to a more detailed characterisation model of the rock mass, but also their total length. Note that any type of optimisation methodology, mostly the ones dealing with random variables, should be assumed as helping tool for the professionals that always should be involved in the decision process.

Due to these mentioned factors, the key question of how to achieve a more detailed characterisation of the rock masses without needing to perform an extensive prospection plan was raised. For that reason, this was identified as the second main question to which this thesis intends to answer.

In this regard, the use of geostatistics as the main toolbox was the way found to solve the raised problems, once it considers the spatial behaviour of the parameters in contrast to the traditional probability techniques. Geostatistics is able to offer a more accurate prediction of the variables always with a certain associated level of uncertainty, suited to reduce the impact of a characterisation of the rock masses in underground works in a deterministic way. Concurrently, this uncertainty assessment is the perfect feature to use in boreholes optimisation since it can be adopted as the minimisation objective function. Therefore, it is intended that this thesis could be

seen as a step forward in implementing geostatistical techniques in geotechnical engineering, mainly at the rock mass characterisation level.

1.2. SCOPE OF THE WORK / OBJECTIVES OF THE THESIS

The lack of methodologies to use in heterogeneous rock masses characterisation is large. Therefore, and due to the complexity of underground works, this characterisation must be as accurate as possible. The existing probabilistic methodologies are only able to reduce the uncertainty associated with the rock mass characterisation and spatial variability of the geomechanical parameters; however, the most critical point, and to which these methodologies do not give answer, are the existence of heterogeneities. These are classified as zones where the rock mass shows lower resistance parameters and that can, most of times, lead to accidents.

Therefore, this thesis aims to develop an integrated methodology that could be used in heterogeneous rock masses characterisation combining the uncertainty reduction, spatial variability and heterogeneities normally present in the rock masses. At the same time, it is important to develop an easy and practical methodology to allow its use by a geotechnical engineer, not only to increase the safety of the workers but also to bring some economic advantages resulting from these new detailed models. For the methodology validation, it is necessary to use geotechnical data from different case studies increasing the reliability of the methodology. The mentioned methodology is outlined as the leading objective of this thesis.

Furthermore, in an early stage of this research work, the potential of the stochastic techniques, used throughout the thesis, is identified. As a consequence, a secondary objective is thought and formulated to fill the major gap of optimisation methodologies for geotechnical engineering. This methodology requires some preliminary information gathered from mechanical boreholes and whose objective is to find out the optimal locations on the rock mass to perform additional boreholes. With this, not only it is possible to reduce the cost of this type of prospection but, at the same time, one can guarantee that these new information is collected in strategic points in order to increase the geotechnical detail of upcoming characterisation models. The purpose of the methodology is to help the professional in an objective and scientific way, diminishing the arbitrariness existing in this geotechnical process.

1.3. METHODOLOGY AND OUTLINE OF THE THESIS

The present thesis proposes two methodologies to be applied in geotechnical works. First, a numerical one aiming to characterise the rock mass heterogeneities in order to take into account an uncertainty reduction, a quantification of the geomechanical parameters variability and, if it is the case, to identify the existence of heterogeneities – first segment of the thesis workflow (Chapter 3 to Chapter 5). The second methodology addresses a boreholes optimisation problem, whose goal goes through identifying the optimal locations where to execute new boreholes in order to reduce the uncertainty of an upcoming numerical model of the rock mass - second segment of the thesis (Chapter 6 and Chapter 7). Each one of the six main Chapters follows the same organisational structure, starting with an introduction section followed by the body of text and, finally some drawn conclusions (see Figure 1.1). All the chapters are designed to answer some scientific and engineering questions developed at the start and during the thesis development.

1.3.1. First segment

- *How many methodologies are able to integrate the uncertainty, spatial variability of the characterisation and the heterogeneities existing in rock masses? and which techniques could be used to overcome the limitations of the mentioned methodologies?*

Following this Chapter 1, where an introduction to the thesis is presented, a comprehensive literature review on the existing characterisation methodologies to use in rock mechanics is presented in Chapter 2.

The also called state-of-art intends to list all the deterministic techniques and the most recent ones, the probabilistic techniques to use in the rock mass characterisation. A special emphasis is given to geostatistics, found to be the most adequate and robust approach to enforce the identification of the heterogeneities in the rock masses, as well as the uncertainty reduction that always exists in geotechnical engineering. All the work developed after this Chapter shows novelty and interesting approaches to be applied in real underground works.

- *Would the geostatistical techniques be able to consider the spatial variability and identify the heterogeneities of the rock masses? Can the geomechanical parameters and geotechnical systems be simulated conditionally using geostatistics? and which type of information can provide the simulation output?*

Chapter 3 assesses the potential of the mentioned geostatistical techniques to conditionally simulate the geotechnical data. Thereby, the simulation results of the Rock Mass Rating (RMR) empirical system obtained in a Chilean deposit are used in this validation. Three different approaches are carried out to understand the existing differences in the simulation algorithms. Provided with the geostatistical simulation of the geotechnical parameters, some probability and uncertainty maps are created as output, as well as the rock deformation modulus map. In this Chapter, the potential of the geostatistical techniques to pursue with the development of the new numerical methodology to use in the characterisation of heterogeneous rock masses is proven.

- *Are the realisations average able to consider the spatial variability of the variables? How can the extreme values resulting from the simulations be maintained for further analysis? and how will it be made the migration of information obtained from the geostatistical simulation to the numerical modelling software?*

Once proven the value of the geostatistical simulation, in Chapter 4 a new numerical methodology to use in heterogeneous rock masses characterisation is proposed. This methodology works as a new combination of techniques that uses the outputs of the conditional simulation to build more realistic models of the rock mass for numerical analysis. Also, an additional technique like the scenario reduction methodology is adapted and applied as a middle step of the main methodology (all the details are presented in Chapter 2), since the realisations average tends to smooth the amplitude of the simulated values. The migration of the realisations chosen to represent the variable will then be imported to a numerical modelling software at a zone centroid level (each zone of the mesh shows a different parameter value). Thus, this Chapter 4 can be considered as the backbone of this thesis, concurrently with its validation using the geotechnical information of Salamonde II powerhouse complex described in Chapter 5.

- *In comparison with deterministic methods, how can the new methodology bring more advantages to the geotechnical community in terms of rock masses characterization? How easy and practical it is to apply the methodology? and how reliable and valid is the proposed methodology?*

In this fifth Chapter, all details regarding the powerhouse complex and the 3D numerical model used are presented. The structural behaviour of the powerhouse cavern is analysed using this numerical characterisation methodology and then compared with observational data, as well as with the numerical results obtained at the time of its construction. The aim of the chapter is to identify the strength of the proposed methodology when compared with a deterministic one, highlighting all the differences between them in terms of displacements and principal stresses values. All the steps of the methodology are applied in a detailed way.

1.3.2. Second segment

The second segment of this thesis, which corresponds to the second proposed methodology, is presented and validated in Chapters 6 and 7 (see Figure 1.1) and represents a considerable different work from the one presented until then. In common, both proposed techniques have the use of geostatistical techniques and the application in the same cases studies; however, they have very different goals and applications in geotechnical engineering.

- *What are the existing optimisation methodologies? Is it possible to have a numerical methodology to help the professionals to define the positions of new boreholes? How will this methodology work?*

In this regard, Chapter 6 starts with a literature review of existing optimisation methodologies and the most used algorithms in engineering optimisation problems. Then, the main steps of the proposed methodology are presented with all the needed details. Since the number of existing methodologies is almost null, and both case studies consider different types of geotechnical information, more than one type of optimisation method are carried out, namely an uni-objective and a multicriteria, this last more adequate when exist information regarding the Rock Quality designation index (RQD), fracturing degree (F) and weathering degree (W) at the same exact points.

- *Which geotechnical gain is it possible to obtain with the new optimisation methodology and how reliable it is?*

Chapter 7 concerns the application of the optimisation methodology in Salomonde II powerhouse complex. In there some tuning of the Simulated Annealing optimisation algorithm are presented, as well different weights assigned to each one of the three objectives to minimise (RQD, F and W). Some data from mechanical boreholes are intentionally removed and their location optimised by the proposed methodology and then compared in terms of location and geotechnical gain in a future characterisation model of the rock mass to build.

Finally, some conclusions of the research conducted in the previous chapters are drawn in Chapter 8. In addition, and since the work presented in here can be considered as innovative and relevant to use in the reality of the geotechnical engineering, some topics that need further developments are likewise outlined.

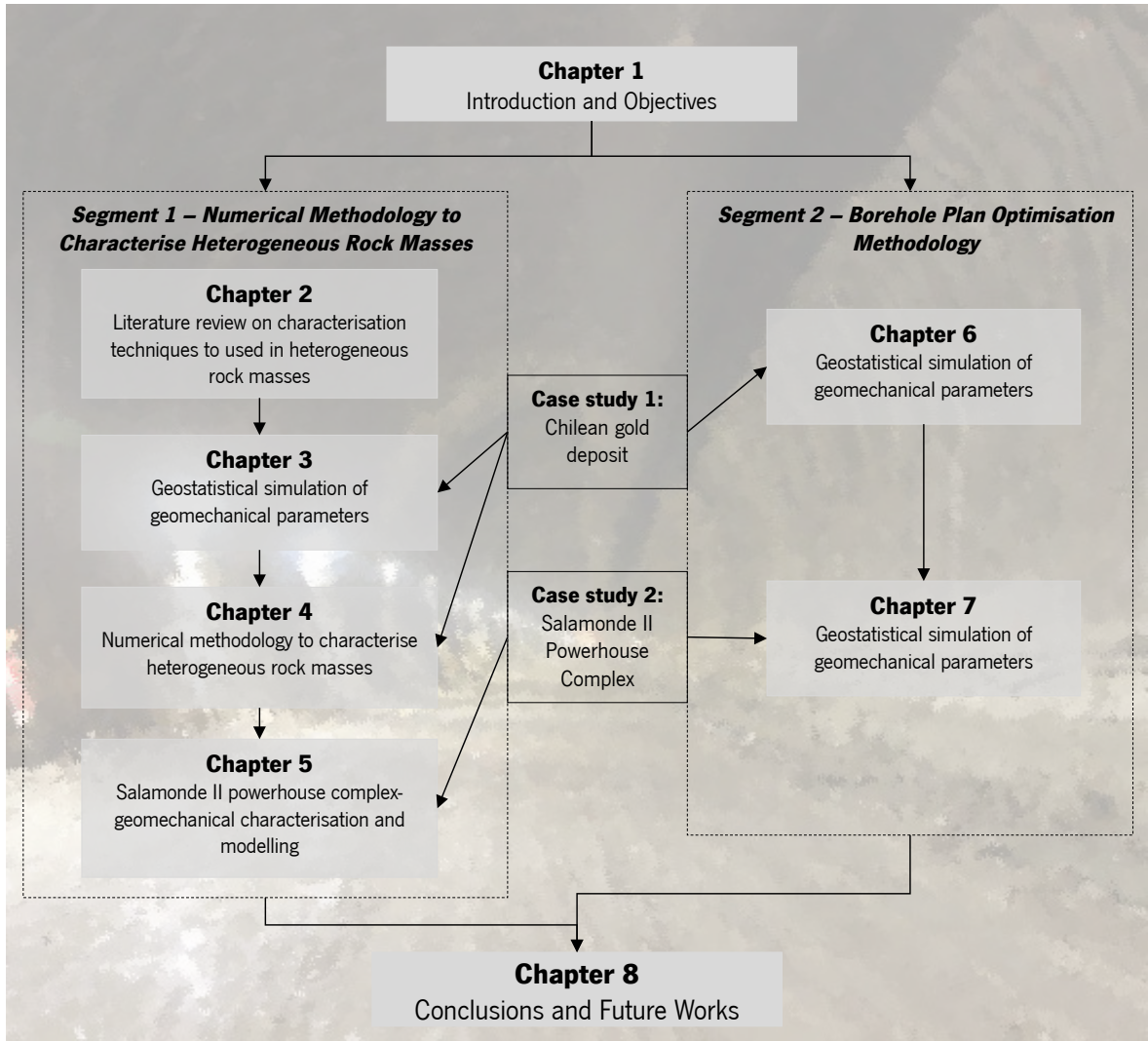


Figure 1.1 Outline and organisation of the thesis.

Chapter 2

OVERVIEW ON ROCK MASS CHARACTERISATION METHODOLOGIES

2.1. INTRODUCTION

The risk associated with underground works, not only for people but also for assets, should be minimised, and more accurate and reliable methodologies that result in more complete and realistic analyses of the structures should be applied. The biggest part of underground works accidents happens during the construction phases and are commonly related to unfavourable conditions of the ground formations that can only be identified *in situ*. Therefore, the margin in what concerns design and construction errors associated with extra costs is small and difficult to deal by the engineers. For these reasons, in recent years, new techniques that are able to characterise the rock mass in a more accurate way than in the past emerged; although, they are not duly adapted to be used in the characterisation of the rock masses present in all underground works and a greater effort should be done towards that.

According to Christian and Baecher (2003) there are three types of uncertainty that should be considered in underground works, the one related with the natural variations of rock mass properties also called as inherent variability, the one resulting from measurement errors and the third one related to the numerical model construction (see Figure 2.1). Regarding the natural variations uncertainty type, linked to the random effect of the natural processes, it can be divided into spatial variability and temporal variability. The first, and more difficult to assess, reflects the natural processes (*i.e.* structural changes) in different locations, while the second comprises the natural variations that happens with time in the same locations.

The uncertainties that can result from measurement errors have three main causes, the ones resulting from bad equipment calibration that leads to incorrect information of the rock mass properties, the human errors, *i.e.* the poor knowledge in using the equipment or in the results analysis, and finally, the residual dispersion of the tests results that can be of difficult identification (Christian and Baecher, 2003). Finally, the third type of uncertainty, and easier to minimise, is related to the rock mass characterisation model, the so-called uncertainty model, that appears in the conceptual and in the numerical part of the model, which should be as close as possible to reality (see Figure 2.1). In fact, the model uncertainty can be, most of the times, associated to the limited knowledge about the rock masses that results from poor geotechnical and geological prospections and from its pointwise nature, *i.e.* only some zones of the rock mass are indeed characterised.

Regarding the first source of uncertainty, natural variability, it is important to distinguish *variability* from *heterogeneity*. From a rock mechanics point of view, variability comprises small variations from point to point of the rock mass characteristics (parameters), while heterogeneities can be classified as important and significant variations in the rock mass characteristics for small distances, *i.e.* abrupt transitions from zones with better mechanical characteristics to zones with worst characteristics and *vice-versa* (Popescu et al., 2005).

To overcome these incomplete and poor characterisation of the rock mass, and to take into account the associated uncertainty (model and measurements errors), probabilistic techniques are being applied; however, the deterministic models that characterises the rock mass using overall values, are still in use. The complexity and novelty of the probabilistic techniques are some of the reasons associated to their lack of use in geotechnical engineer and, as consequence, an effort should be done towards their simplification and validation.

To address this issue, in the very first sections of this chapter, traditional methods used to classify the rock masses, namely, the main *in situ* and laboratory tests along with the empirical classification systems, are briefly described. Also, some of the most used empirical formulas to derive the values of some geomechanical parameters, such as the rock mass deformation modulus (E_m) are outlined. In the following sections, an exhaustive list of the most commonly used probabilistic techniques is presented; however, more emphasis is given to the geostatistics. Note that this latter has been proving capable to integrate the uncertainty related with the geomechanical parameters, their spatial variability and the existence of the heterogeneities in a single rock mass model. In contrast to the traditional probabilistic techniques, geostatistics treats the variable under study as a realisation of a spatial random field, through the evaluation of the data spatial dependence and behaviour. This technique allows the prediction of the unknown values of a variable at the target points in a regular or irregular grid presented in *two* or *three dimensions*, always with an uncertainty measurement associated with the prediction. As output, numerous realisations can be obtained from the technique; however, it becomes inviable to analyse all of them or even perform the realisation average, which tends to smooth the extreme value and consequently reduce the spatial variability. Thereby, the use of scenario reduction techniques is an essential step to follow in the geostatistical simulation. For this reason, in the last sections, scenario reduction methodologies are mentioned and pointed out as a workable option to reduce the smoothness problem of using the geostatistical realisations average and as a statistical based solution to select a small number of realisations.

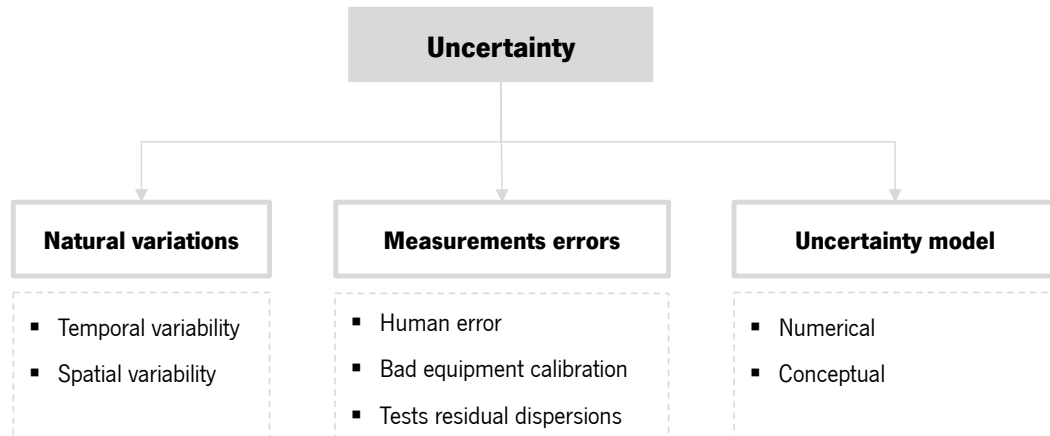


Figure 2.1 Scheme with the uncertainties related to geotechnical engineering (adapted from Christian and Baecher (2003)).

2.2. DETERMINISTIC APPROACHES

Regardless of the fact that geotechnical engineering had, in the last years, gone through some vital transformations, notably due to the emergence of more powerful techniques and methods used to obtain the geomechanical parameters, the deterministic are still in use. In every geomechanical characterisation it is essential to carry out *in situ* and/or laboratory tests, coupled with the use of some indirect methods like empirical systems and formulas, mainly in rock formations. Examples of these classification systems are the Rock Quality Designation (RQD) presented by Deere, 1964, the Rock Mass Rating (RMR) proposed by Bieniawski (1976), the Geological Strength Index (GSI) proposed by Hoek and Brown (1997) and Q-system presented in Barton et al. (1974). The use of these systems for rock mass classification facilitates not only the knowledge of the rock mass quality but also allows obtaining other relevant information such as the choice of the support systems to apply in underground works. The process of applying these systems involves the collection of rock mass information from different sources, as boreholes logs, scanline surveys, geological structures mapping and rock testing (Cai et al., 2007). The possibility to obtain the geomechanical parameters of the rock mass using these systems is one of the main advantages in applying them, being the most commonly used the RQD, the RMR, the GSI and the Q systems. A brief explanation of these systems will be given next, including a more detailed information about the RMR system, which will be used in thoroughly in this work. Besides, for each of them, empirical formulas for deriving values of the rock mass deformation modulus are presented, because they will be relevant for the remaining work.

2.2.1. Rock Quality Designation

The Rock Quality Designation (RQD) provides a quantitative index of the rock mass based on the drill core logs and it is a leading indicator for low-quality rock zones (Deere, 1964). It is determined by measuring the core recovery percentage of intact core pieces that are greater than 100 mm in length. To employ this index, the core must have a minimum diameter of 54.7 mm and should be drilled with a double-tube core barrel. This index (equation (2.1)) is here presented in detail since it is a standard parameter used in the RMR and Q classification systems.

$$RQD = \frac{\sum \text{Core pieces with length} > 100 \text{ mm}}{\text{Total length of core}} \times 100 \quad (2.1)$$

It is worth mentioning that the RQD is directional dependent, *i.e.* it can change considerably with the borehole orientation. Also, it is important to be aware that the drilling process can cause some additional fractures in the rock core. That is why the index is normally associated with other classification systems. However, some authors proposed empirical formulas relating the RQD index value and the rock mass deformation modulus as presented in Table 2.1. All the presented formulas relate the deformation modulus of the intact rock (E_i) with other rock parameters, such as the uniaxial compressive strength (σ_c) and the weathering degree of the rock mass (W) defined according to the International Society of Rock Mechanics (ISRM).

Table 2.1 Empirical formulas to obtain E_m using the RQD index value.

Author	Limitations	Equation (E_m in GPa)
Zhang and Einstein (2004)	-	$\frac{E_m}{E_i} = 0.2 \times 10^{0.018RQD-1.91}$ Lower Bound
		$\frac{E_m}{E_i} = 1.8 \times 10^{0.018RQD-1.91}$ Upper Bound
		$\frac{E_m}{E_i} = 10^{0.018RQD-1.91}$ Mean
Kayabasi et al. (2003)	-	$E_m = 0.135 \left[\frac{E_i \left(1 + \frac{RQD}{100}\right)}{W} \right]^{1.1811}$
Gokceoglu et al. (2003)	-	$E_m = 0.001 \left[\frac{(E_i/\sigma_c^1)(1 + RQD/100)}{W} \right]^{1.5528}$

¹ σ_c is measured in MPa.

2.2.2. Rock Mass Rating system

The Rock Mass Rating (RMR) empirical system was initially proposed by Bieniawski (1976) and a new version was proposed by the same author in 1989. The system is composed by the individual evaluation of six rock related parameters and the final value varies between 0 and 100. For each underlying parameter, different ratings (with a maximum of 20) were defined according to parameters scale of variation. Subsequently, five parameter ratings are summed resulting in a so-called basic RMR (equation (2.2)).

$$RMR_{basic} = P1 + P2 + P3 + P4 + P5 \quad (2.2)$$

Those parameters are:

- P1: Uniaxial compressive strength of rock material (UCS);
- P2: Rock quality designation (RQD);
- P3: Spacing of discontinuities (JS);
- P4: Condition of discontinuities (JC);
- P5: Groundwater conditions (GW).

In addition, a sixth parameter should be subtracted from the RMR_{basic} . This parameter refers to the correction related with the orientation of the discontinuities. In underground structures if the orientation of the discontinuities is unfavourable to the excavation, a value of 10 should be subtracted from RMR_{basic} .

Based on the value of RMR the rock mass can be classified into five rock quality classes (see Table 2.2).

Table 2.2 RMR rock quality scale and description.

RMR value	Rock quality description
< 21	Very poor
[21; 40]	Poor
[41; 60]	Fair
[61; 80]	Good
[81; 100]	Very good

As mentioned, the deformability parameters of the rock mass can be assessed using the RMR system value. Therefore, in the past years, a wide range of formulas have been proposed by

numerous authors. However, it is important to be aware of the formulas origin since, in many cases, they were developed for specific types of rocks. Table 2.3 presents a list of empirical formulas to obtain the rock mass deformation modulus (E_m) using the RMR value and other rock parameters presented before. In Verman (1993) empirical formula an additional parameter is required, the overburden (H), which enters in meters.

Table 2.3 Empirical formulas to obtain E_m using the RMR value.

Author	Limitations	Equation (E_m in GPa)
Bieniawski (1989)	RMR > 50	$E_m = 2 \times RMR - 100$
Serafim and Pereira (1983)	$RMR \leq 50$	$E_m = 10^{(RMR-10)/40}$
Mohammadi and Rahmancejad (2010)	-	$E_m = 0.0003 \times RMR^3 - 0.0193 \times RMR^2 + 0.315 \times RMR + 3.4064$
Read et al. (1999)	-	$E_m = 0.1 \times \left(\frac{RMR}{10}\right)^3$
Hoek and Brown (1997)	$\sigma_c \leq 100 \text{ MPa}$	$E_m = \frac{\sqrt{\sigma_c}}{10} 10^{(RMR-10)/40}$
Nicholson and Bieniawski (1990)	-	$E_m(\text{MPa}) = \frac{E_i}{100} \times (0.0028RMR^2 + 0.9e^{(RMR/22.28)})$
Mitri et al. (1994)	-	$\frac{E_m}{E_i}(\text{MPa}) = 0.5 \times (1 - \cos(\pi \times RMR/100))$
Verman (1993)	$\sigma_c > 100 \text{ MPa}$ and $H > 50 \text{ m}$	$E_m = 0.3 \times H^\alpha \times 10^{(RMR-20)/38}$
Galera et al. (2007)	-	$E_m = E_i \times e^{(RMR-100)/36}$
Sonmez et al. (2004)	-	$E_m = E_i 10^{((RMR-100)(100-RMR))/4000e^{(-\frac{RMR}{100})}}$

2.2.3. Q system

The Q system was proposed by Barton et al. (1974) aiming to help the engineers in designing the underground works support systems. Indeed, the system provides a quality index for the rock mass classification combining six different parameters, such as the RQD system value, the number of joint sets (J_n), the roughness of the most unfavourable joint or discontinuity (J_r), the degree of alteration or filling of the most unfavourable joint or discontinuity (J_a), the degree of water inflow (J_w), and the stress condition (SRF) resulting in the following expression:

$$Q = \left(\frac{RQD}{J_n}\right) \times \left(\frac{J_r}{J_n}\right) \times \left(\frac{J_w}{SRF}\right) \quad (2.3)$$

In what concerns the Q system scale, it varies from 0.001 to 1000 corresponding to a rock mass with poor quality and exceptional good quality, respectively. A list of formulas that allows the use of Q system value to calculate the rock mass deformation modulus is presented in Table 2.4.

Table 2.4 Empirical formulas to obtain E_m using the Q system value.

Author	Limitations	Equation (E_m in GPa)
Barton et al. (1983)	$Q > 1$	$E_m = 25 \times \log(Q)$
Barton and Quadros (2002)	$Q \leq 1$	$E_m = 10 \times Q_c^{1/3}$, $Q_c = Q \times \frac{\sigma_c}{100}$
Singh (1997)	$H > 50 \text{ m}$	$E_m = H^{0.2} \times Q^{0.36}$
Singh (1997)	$Q \leq 500$	$E_m = 1.5 \times Q^{0.6} \times E_i^{0.14}$

2.2.4. Geological Strength Index

In order to apply the Mohr-Coulomb and Hoek-Brown failure criteria the rock mass strength parameters are required. Thus, Hoek and Brown (1997) proposed the Geological Strength Index (GSI) that can also be used to obtain the value of E_m . Later, in 2002, Hoek et al. proposed a new version of the system, which consists in a qualitative description considering the rock mass structure (from an intact to laminated or sheared) and its discontinuities conditions (from very good to very poor). Regarding the system scale, alike the RMR system, it varies from 0 to 100, corresponding to a very poor and very good quality rock mass, respectively. As a matter of fact, both systems show a high correlation using the updated version of the RMR system (Bieniawski, 1989) that could be translated by the following relation:

$$GSI = RMR_{basic} - 5 \quad (RMR \geq 23) \quad (2.4)$$

For RMR values under 23 the correlations between both systems cannot be built, since this was the minimum value obtained for the RMR by Bieniawski (1976, 1989). In the same way as presented for RMR and Q systems, Table 2.5 lists empirical formulas that require the GSI value in order to obtain the value of E_m . Some of the empirical formulas use the disturbance factor (D), introduced by Hoek et al. (2002) as an additional parameter.

Table 2.5 Empirical formulas to obtain E_m using the GSI system value.

Author	Limitations	Equation (E_m in GPa)
Hoek and Brown (1997)	$\sigma_c \leq 100 \text{ MPa}$	$E_m(\text{MPa}) = \left(1 - \frac{D}{2}\right) \times \sqrt{\sigma_c/100} \times 10^{(GSI-10)/40}$
Hoek and Brown (1997)	$\sigma_c > 100 \text{ MPa}$	$E_m(\text{MPa}) = \left(1 - \frac{D}{2}\right) \times 10^{(GSI-10)/40}$
Hoek and Diederichs (2006)	-	$E_m = 100000 \times \left[\left(1 - \frac{D}{2}\right) / \left(1 + e^{\frac{75+25D-GSI}{11}}\right) \right]$
Hoek and Diederichs (2006)	-	$E_m = E_i \left(0.02 + \frac{1 - \frac{D}{2}}{1 + e^{\left(\frac{60+15D-GSI}{11}\right)}} \right)$
Sonmez et al. (2004)	-	$E_m = E_i \times (s^a)^{0.4}$ with $s = e^{\left(\frac{GSI-100}{9}\right)}$ $a = \frac{1}{2} + \frac{1}{6} \left(e^{-GSI/15} - e^{-\frac{20}{3}} \right)$
Carvalho (2004)	-	$E_m = E_i \times s^{1/4}$

2.3. PROBABILISTIC METHODS

2.3.1. General

The emergence of new methodologies and computational tools to use in geotechnical engineering has emphasised the limitations associated with deterministic models. These models, most of the times, do not result in satisfactory results as the characterisation model of the ground, since their outcome is rather simplistic when compared with complex realities; however, the easiness associated with these models and lack of better alternatives led to the widespread use of this approach. According to El Ramly (2001), the estimated error, meaning the difference between the estimated values and the real values, are assumed as zero, which corresponds to a distant assumption for the rock masses reality, where variability and heterogeneity are present and significant.

Therefore, probabilistic methods emerged recently as a good alternative to spatially model the geotechnical variables in the sense that the input parameters are no longer single values and jump to assume any value inside a given range. As insinuated by Christian and Baecher (2003), the uncertainty and risk associated with the geotechnical models can only be expressed according to a probabilistic approach and the risk assessed by stochastic deductions. Therefore, the use of probabilistic techniques is particularly useful to compare, analyse and combine all types of uncertainty providing a relationship between the safety and the failure probability. Among the significant number of techniques, the most commonly used are the methods that derived from the *Taylor* series, the First Order Second Moment (FOSM), the Point Estimate Method (PEM) and the

Monte Carlo simulation. The first two mentioned techniques only require as inputs the mean and standard deviation of the variable(s), while the Monte Carlo simulation allows manipulating a higher number of data resulting in a more detailed analysis. At this point it is important to define the term *simulation*, which in a general dictionary is stated as "...pretend to be like, imitate a phenomenon or generic process..." (Oxford Dictionary, 1976). From a mathematical point of view the term defines the design of a mathematical model of an actual or theoretical system to reproduce outcomes that are close to the real ones. Some concepts and details of probabilistic techniques, relevant in geotechnical engineering, are presented next.

2.3.2. Monte Carlo

The Monte Carlo method (MCM) most known version was proposed by Metropolis and Ulam (1949) and allows, in a general way, the random generation of values resulting in a probability distribution for a set of independent random variables and can be considered as a statistical method. The MCM appeared in 1940's in the context of the development of the atomic bomb to use in the World War II (Ulam, 1947). The method has, currently, a vast number of algorithms that could be applied to several areas (physics, chemistry, engineering, finance, etc.). Examples of these applications that are worth noticing and distinguished are: 1) the simulation of processes from a stochastic point of view with the goal to mimic the stochasticity of the input system; 2) the optimisation of complex problems. In some cases, the objective functions are deterministic and some randomness should be introduced to efficiently search the domain of the objective function (Griffiths et al., 2002).

In most cases, the method can be applied without the formulation of mathematical equations, normally used to describe their behaviour. Therefore, the only requirement is the use of probability functions of the independent variables. Also, these variables can have different types of probabilistic distribution: normal, exponential, etc. Regarding the MCM the obtained results can be controlled by the variance value that can result in more simulations in order to achieve a preferred error value. However, in order to understand the system results other metrics can be computed, such as mean, variance and coefficient of skewness, as well as probabilistic values like probability density functions and cumulative density functions (see Figure 2.2).

One of the main steps required to apply the MCM is the number of realisations or simulations. In order to achieve a realistic distribution of the results the MCM needs to be repeated many times, which sometimes translates in a considerable computation time. The convergence of the model can be verified through the analysis of the simulations fluctuations that should be minor.

In what concerns geotechnical engineering, the MCM is applied, mainly, to generate simulations combining some geomechanical parameters that can be expressed in a deterministic or probabilistic way. Among the many types of post analysis outputs, the most relevant in geotechnics are rock mass displacements, stresses distributions and safety factors (Viana, 2014).

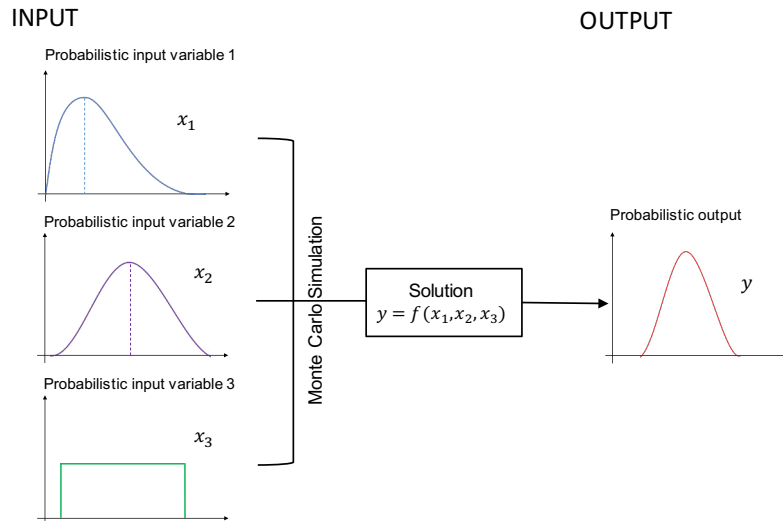


Figure 2.2 Monte Carlo method principle for an example of three variables used as input.

Moreover, the MCM can be applied, in a general geomechanical problem (*i.e.* a tunnel), through the following steps:

1. Build the numerical model of the structure;
2. Identify the variables that contribute to the model uncertainty;
3. Calculate the probability distributions of each variable (cumulative distribution functions);
4. Define the range of values for each variable;
5. Generate random numbers inside each predefined range (sample simulation);
6. Compute the model for several combinations of numbers obtained in the previous step until the convergence is reached;
7. Perform a statistical analysis of the results (e.g. histograms, confidence intervals, etc.).

Like all probabilistic methods, the MCM presents some disadvantages, like the fact of being directly related to the number of performed simulations. The accuracy of the model increases as the number simulations also increases resulting in significant processing time. However, the mathematical simplicity, the small number of required inputs, the versatility in considering different numerous variables are the most noteworthy advantages of the MCM that results in an increasing application of the method (El-Ramly, 2001).

2.3.3. First Order Second Moment Method

The First Order Second Moment method (FOSM) derived from *Taylor* series (characterisation function) provides an analytic approximation for the mean and standard deviation of the interest variable Y . This method is normally used to estimate the variables uncertainty.

Jointly with the characterisation function, the definition of the first and second moments (corresponding to the mean and variance, respectively) are required. Therefore, Y is assumed as a function of random variables x_1, x_2, \dots, x_n , given by:

$$Y = f(x_1, x_2, \dots, x_n) \quad (2.5)$$

As input, the series considers the probability density function of the random variables; however, and because the variable information can, many times, be limited to their mean and variance, the *Taylor* series is truncated after the linear term (first order) to obtain the second moment of the probability function, the variance value (Ang and Tang, 1984). Hence, the mean and variance are given by Equations (2.6) and (2.7), respectively.

$$\mu_Y = f(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n}) \quad (2.6)$$

$$\sigma_Y^2 = \sum_{i=1}^n \sum_{j=1}^n \rho_{x_i x_j} \sigma_{x_i} \sigma_{x_j} \left(\frac{\partial Y}{\partial x_i} \frac{\partial Y}{\partial x_j} \right) \quad (2.7)$$

where μ_{x_i} represents the mean value of x_i , $\rho_{x_i x_j}$ symbolises the correlation coefficient between x_i and x_j , σ_{x_i} is the standard deviation of x_i and σ_{x_j} the standard deviation of x_j . In the case that the random variables do not show correlation, the variance is given by:

$$\sigma_Y^2 = \sum_{i=1}^n \sigma_{x_i}^2 \left(\frac{\partial Y}{\partial x_i} \right)^2 \quad (2.8)$$

In what concerns the number of calculations, the minimum should be set to $2n + 1$, where n represents the number of random variables that composes Y (equation (2.5)).

As referred by Christian (1996) a standard procedure (list of steps) should be applied in order to use the FOSM method in an accurate and efficient way:

1. Identify the variables that contribute to the model uncertainty;
2. Define the mean, standard deviation and correlation coefficient values of the variables;
3. Estimate the uncertainty of the variables and compute its variance;
4. Perform a sensitivity analysis in order to understand the variation in the dependent variable compared to the remaining ones;
5. Calculate the variance (using equation (2.8)) and, consequently, the standard deviation value of the model.

As well as the Monte Carlo method, the FOSM method shows as a drawback the fact that it is not able to identify the heterogeneities typically presented in rock masses since they use as input the variable global statistics and do not consider their spatial behaviour. On the other hand, the precision in the results and the use of linear functions are one the main advantages of the method.

Moreover, regarding the geotechnical applications, the FOSM method is often used to calculate the Safety Factor (SF) of a slope, which normally depends on various parameters, dependent or independent of each other, related to the geomechanical properties of the ground.

2.3.4. Point Estimation Method

In 1975, Rosenblueth proposed the Point Estimation Method (PEM), which in contrast with the FOSM method, does not require the probability density function (mean and variance values are sufficient as input). In accordance with Christian and Baecher (2003), the method numerically approximates the moments value of the variables functions and by using these moments of an independent variable x , the method is able to obtain an approximation for the same order moments for the dependent variable y . The referred moments, considered as output, are the mean, variance and skewness.

PEM is a weighted average method that generates several solutions for multiple points after combining them in different ways until the desired distribution is reached (solution). As a reference, the PEM method uses two points to perform the estimation, at plus ($x +$) and minus ($x -$) one standard deviation from the variable mean value corresponding to $P+$ and $P-$ probability masses, as exposed in Figure 2.3. Regarding the number of calculations to carry out, it is related to the number of random variables used in the process, *i.e.* if a total of n random variables is used the total number of combinations is 2^n .

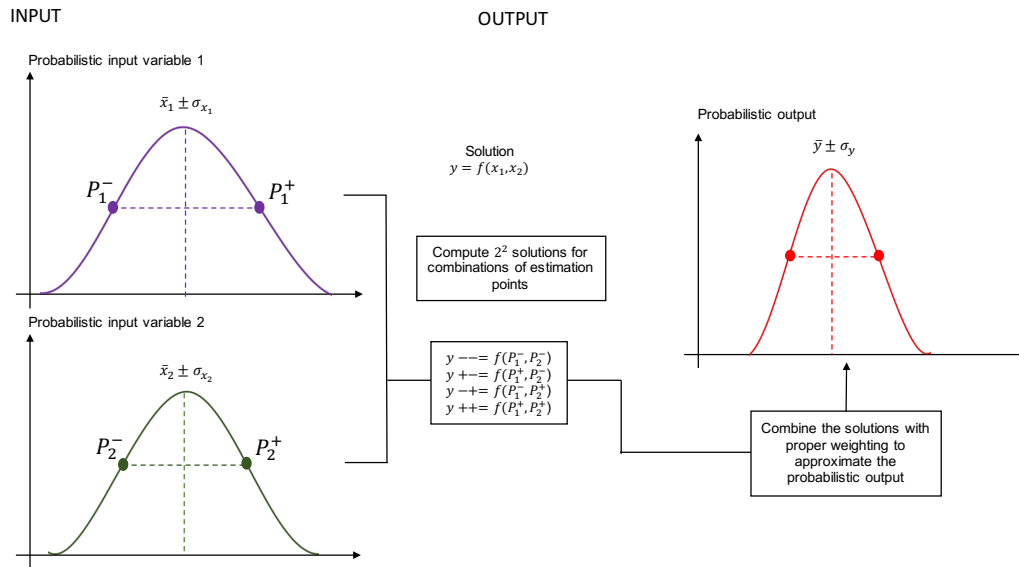


Figure 2.3 Principle of Point Estimate method assuming two random variables (adapted from Rocscience (2016)).

Once executed the 2^n combinations, the probabilistic distribution of the dependent variable, represented as $y = (x_1, x_2, \dots, x_n)$, is given by:

$$\mu_y = \sum_{i=1}^{2^n} W_i f(x_i) \quad (2.9)$$

$$\sigma_y^2 = \sum_{i=1}^{2^n} W_i f(x_i)^2 - \left(\sum_{i=1}^{2^n} W_i f(x_i) \right)^2 \quad (2.10)$$

where W_i values represent the point estimation weights for each random variable x_i given by $\frac{1}{2^n}$ and $f(x_i)$ represents the number of successive evaluations of a given function f in every possible combination for the random variables at point estimate locations, *i.e.* at $\bar{x}_n - \sigma_{x_n}$ and $\bar{x}_n + \sigma_{x_n}$ (Rocscience, 2016).

In the same way that was listed for the previously mentioned methods, the recommended steps are:

1. Identify the variables that contribute to the model uncertainty;
2. Compute the mean and standard deviation for each variable;
3. Perform 2^n iterations combining the maximum and minimum values of each random variable;

4. Calculate the dependent variable value (y) for each 2^n combinations (in this step a deterministic method is used);
5. Compute the mean and variance of the dependent variable (using equations (2.9) and (2.10)).

In terms of computational time, the PEM method results in very efficient times when compared with the two previous methods due to the fact that it is only required an approximation to two estimate points. However, as a rule, the method uses random variables with normal distributions making the use of other distributions result in less accurate results.

2.4. RANDOM FIELD THEORY

The probabilistic techniques previously mentioned have as a main limitation the fact that they are not able to identify the heterogeneities or spatial variability of the rock masses and soils, mainly due to the simplifications and global statistics (mean and variance) that twist and hide information about the variable spatial distribution. Even though, to assess the model uncertainty and quantify the parameters' variability the use of these techniques has been proving enough. To overcome the increasing difficulty in characterising the randomness in physical and mechanical properties, in recent years, the geotechnical community has been developing applications based in stochastic methods. In essence, this method involves the generation of random numbers aiming to explore the possibilities field of a given variable that can be formulated mathematically but may not be predicted precisely.

Furthermore, these methods can be divided into two vast domains, the stochastic processes and the stochastic fields or random fields. The first is applied when the variable to model can be expressed using a temporal function, while the second domain exists for variables with a spatial behaviour. This second domain is the most commonly used in geotechnics, since it requires the use of probabilistic distributions. This distribution can be divided into two types, the Gaussian distribution where the variable(s) present(s) well defined thresholds (nonnegative values) and where the first two moments (mean and variance) can be extracted, and the second group that combines all the remaining distributions (from lognormal to beta or gamma), here called non-Gaussian. This type of distributions is used for more spatial complex variables where it is required all the variable moments (mean, variance, skewness and kurtosis). The powerfulness of the random field theory lies in providing statistical results used to draw inference from *in situ* and laboratory observations and, the most

relevant in geotechnics, incorporating the natural variability in the geomechanical models (Christian and Baecher, 2003). The latter, in accordance with Fenton (1997) can be assessed assuming that points that are close to each other have a high probability of showing similar material properties and the reverse happens for distant points. As a consequence of the existing limitations in geotechnical engineering, in performing a large number of *in situ* or laboratory tests, the use of techniques powerful enough to minimise the unknown effect of the natural variability is crucial. An accurate knowledge of some rock mass or soil zones could be sufficient to model all the volume in study; however, some assumptions have to be made, such as: 1) to assume that the rock mass or soil properties can vary from point to point and the mathematical function used to characterise this variable requires the mean and the variance; 2) the soil and rock mass properties show spatial dependence.

To elucidate this problem let us assume, as an example, Z as a rock variable, such as the deformation modulus (E_m) within a rock mass volume. Z can exhibit natural variability and the variable can take different values from point to point within the rock mass volume. In a random field theory approach, Z will be generated in order to reproduce a random field over the rock mass volume. The natural variability concept, as already explained in section 2.1, is here contemplated by the definition of the autocorrelation function of Z , that allows understanding how a defined variable Z behaves in a selected volume (ρ). The function can be obtained after calculating the coefficient of autocorrelation (ρ) that relates the values of Z in two points located at a distance r ; this coefficient should be calculated from all possible combinations of r considering all the volume points through Equation (2.11).

$$\rho(r) = \frac{Cov[Z(x), Z(x')]}{Var[Z]} \quad (2.11)$$

where r represents the distance between the two points obtained from $|x - x'|$, x and x' the two points coordinates, $Cov[Z(x), Z(x')]$ the covariance value between the two points and $Var[Z]$ the variance of variable Z .

From the analysis of the autocorrelation function behaviour, as the value of r increases, the correlation between the points often decreases until they do not have correlation ($\rho = 0$). Several mathematical models can be used to fit the autocorrelation function, although one important feature to point out is the possibility of the random field to be anisotropic, that means, it can have different correlation lengths along the vertical and horizontal directions. Note that a correlation length

corresponds to a maximum value beyond the random field is effectively uncorrelated. According to Fenton (1997), the soil properties are usually anisotropic, having usually, the vertical correlation length, a smaller value than the horizontal correlation length, as consequence of the soil formation process in the case of sedimentary soil deposits (layers deposition). The differences of two random fields ($X(t)$) for each point in the field (t) with distinctive correlation lengths ($\theta = 0.04$ and $\theta = 2.0$) can be observed in Figure 2.4. From it is possible to detect that for higher correlation lengths, the random field varies in a continuous way. On the contrary, a low correlation length results in a rough random field where a large amount of points in the field become uncorrelated and the field variability more accentuated (Phoon and Kulhawy, 1999).

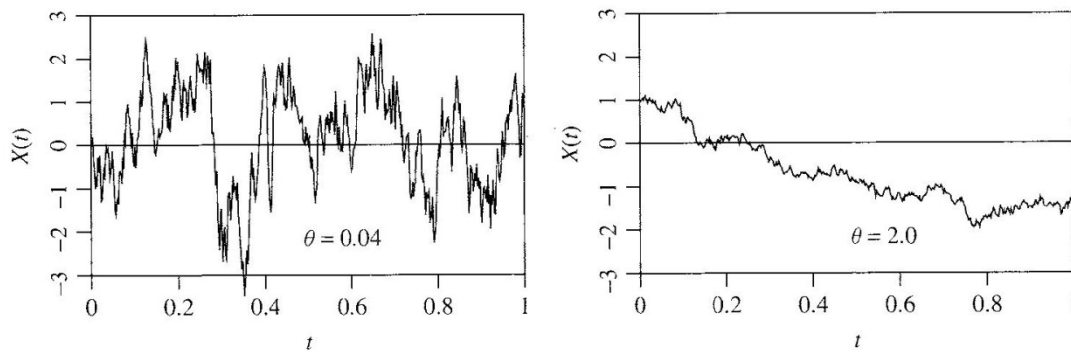


Figure 2.4 Example of two random fields ($X(t)$) with distinctive correlation lengths (θ) (Fenton, 1997).

Furthermore, the random fields technique has been, in the late years, used alongside the probabilistic techniques mentioned before in very different areas. Hence, from a geotechnical point of view the numerous works that have been published use random fields, mainly to analyse the material physical properties, to reproduce the spatial variability of the geomechanical parameters, to study instability process (slopes in rock or soil), to model rock mass fractures networks, etc. Some details about the mentioned works and their scopes are presented next.

In terms of random fields works applied to slopes, Gravanis et al. (2014) studied the failure probability in a slope combining the random fields technique with a finite elements analysis. In this analysis, the geomechanical parameters, such as the friction angle and cohesion were assumed as random variables characterised by their mean, standard deviation and correlation lengths. As a result, the author perceived the significant influence of considering different correlation lengths on the slope performance, namely in the slope probability of failure values. This study had allowed the author to conclude that the spatial variability should not be ignored. In the same domain, Le Goc et al. (2014), Popescu et al. (2005), Sejnoba et al. (2007) used the random fields combined with Monte Carlo method to obtain the failure probabilities and safety factors values related with the

spatial variability degree of the soil (more variability, more probability of failure), with the correlation lengths and the geomechanical parameters influence.

Borghi et al. (2012) uses random field techniques to construct the geological reservoirs model using different inputs, such as maps of fractures, hydrogeological information, geological mappings, etc., and from that identify reservoir zones with the presence of ore.

In underground works, these random fields techniques had not been fully exploited justified by the reduced number of works in this area. However, in the early 2000s, Thacker and Painter (2000) studied the effect of different correlation lengths of the geologic medium on a tunnel response. Therefore, different correlations were considered in different layers of the rock mass and simulated using the Monte Carlo simulation method. As conclusions, the authors confirmed the influence on the tunnel performance when different correlations lengths are considered, particularly in large and heterogeneous rock masses.

More recently, Hu and Huang (2007) applied the Monte Carlo Random field method in three types of soils in order to define the soil types transition zones along a tunnel alignment. The goal was identifying the relevant role that the uncertainty arising from soil transition can play in tunnelling. In the same line of work, Chong et al. (2013) used a lower and upper bound limit analysis in a pressurised tunnel in order to investigate the most critical tunnel pressure in the presence of spatial variability.

Furthermore, Huber and Hicks (2010) presented a work studying the influence of soil spatial variability in surface settlements affected by tunnelling. A Sequential Gaussian Simulation (SGS) algorithm was used to generate the random field of the elasticity modulus, and then, combining FOSM and Monte Carlo, the probability of damage due to differential settlements on the surface was evaluated.

On similar grounds, Luo et al. (2011) proposed a simplified approach to consider the effect of the spatial variability in a clay excavation using a *two-dimensional* random field. Combining the Monte Carlo simulation with the Random fields method, the authors could prove the over-estimated values of the based-heavy failure probability obtained using a deterministic method that does not take into account the spatial variability.

2.5. GEOSTATISTICAL TECHNIQUES

2.5.1. General

Considering all the mentioned methods it is worth noticing that they have in common the same limitation in identifying the rock masses natural heterogeneities. This happens due to the assumption that a variable is normally distributed and is spatially independent from sample to sample, which is not the case in earth sciences where the data do not often satisfy these assumptions. Therefore, geostatistical techniques that incorporate the evaluation of the random variable statistical distribution and model the “interaction” between all the variable samples, i.e. considers the spatial dependence between them (resulting in a prediction of unknown points) can be used to overcome this limitation of the classical statistics (see Figure 2.5).

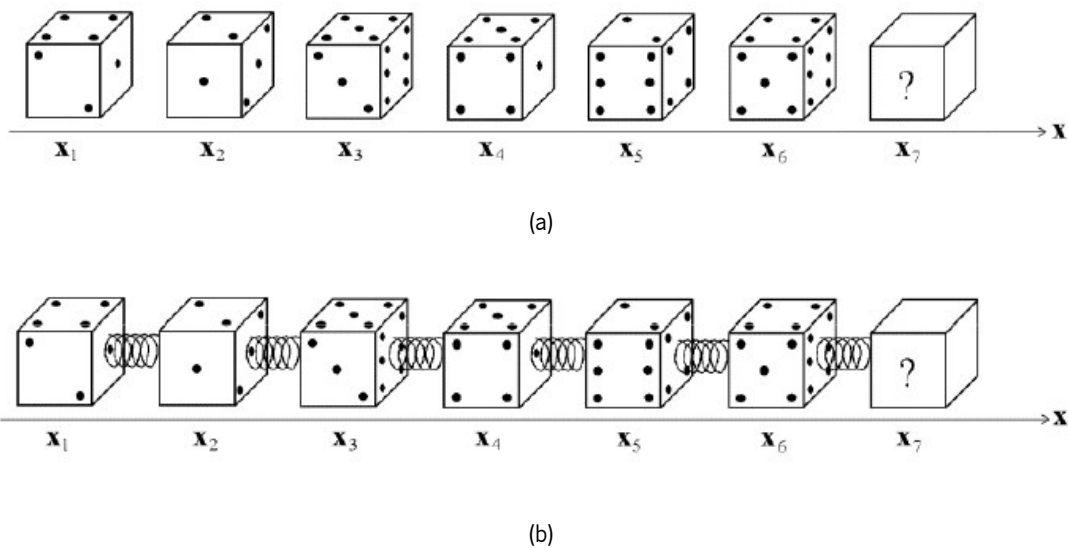


Figure 2.5 Difference between classical statistics and geostatistics: a) independent realisations (classical); and b) dependent realisations (geostatistics).

Hence, since this thesis addresses the heterogeneity identification problem, geostatistical techniques appear as the most robust and adequate to reduce the uncertainty, assess the spatial variability and identify the heterogeneities of the rock masses in the characterisation model to build. However, the application of this technique in engineering, namely in geotechnics has been rare.

The geostatistical technique arose firstly in 1951 in South Africa through a mining engineer named Krige (1951). The technique was initially applied in gold mines deposits after the engineer noted a trend to undervalue the estimation of gold for lower grade categories and the reverse for higher grade categories, *i.e.* a phenomenon now classified as conditional bias. This effect happens when

the expected value (average) of the true grade conditional on the estimated grade is not equal to the estimated grade, in this case, the true value is normally lower than the estimated value. After the problem identification, was recognised the necessity of analysing the ore grade not only by considering the mine block but a larger area (higher variances). As a consequence, the neighbouring data was set as part of the estimation process and the spatial concept was introduced. To minimise the effect of mentioned conditional bias, Krige implemented some corrections in ore block estimation (*i.e.* the estimation is performed not a point level but for a block) through the consideration of a weighted average of the neighbouring data and the global mean of the samples over all the mine. The weighting is defined according to the spatial correlation of the data. From this correction emerged the technique nowadays called kriging.

In 1963, the French engineer named Matheron using Krige's previous works, presented a new theory about regionalised variables that are characterised by a numerical function that varies in space and which variance can be defined through a variogram and not by a simple numerical function (Matheron 1963, 1971). The variogram analysis allows the classification of a random variable as spatially continuous or discontinuous and as stationary (invariance of the spatial distribution) or non-stationary (variance of the spatial distribution); more details about the variogram definition will followed in the next section. In what concerns the regionalised variables forms, according to Chilès and Delfiner (2012) there are three different types, such as continuous, categorical and an object. In the first group are placed the variables that show a continuous scale resulting in a continuous histogram (physical properties), while in the second group are presents the variables whose scale is a rating, indicator or disorganised rankings. In the last variables group, can be included the ones modelled using objects, as is the case of the fractures that are displayed as discs.

Subsequently, when the goal is the analysis, characterisation and spatial prediction of one or more regionalised variables, the geostatistical modelling can be adopted in two different forms, the estimation (using kriging technique or its multivariate extension, cokriging) and the simulation (simulation or cosimulation). It is worth of noticing that more details about the mentioned prediction techniques will be given later in section. Figure 2.6 and Figure 2.7 present the general process to adopt in each process and it is possible to observe that the main differences between both techniques occur in the outputs, while in the estimation process the results are composed by two maps, one with the “best” estimation of the regionalised variable and other with the error variance, in the case of the simulation processes some randomness is added and the result is constituted by numerous equivalent maps, also called realisations. Each one of these realisations is a possible

representation of the reality were the regionalised variable structure is honoured. Moreover, if analysed together, these realisations can be used to measure the uncertainty associated with the prediction process (Chilès and Delfiner, 2012). In addition, and giving a closer look in Figure 2.7, it is possible to observe the addition of a data histogram when compared with the estimation process (see Figure 2.6). The use of the histogram in the simulation process exists to control and compare the final distribution of the data in each computed realisation since more randomness is introduced in the process. In Figure 2.8 one example of uncertainty modelling for a hydrogeological oil reservoir simulation is presented.

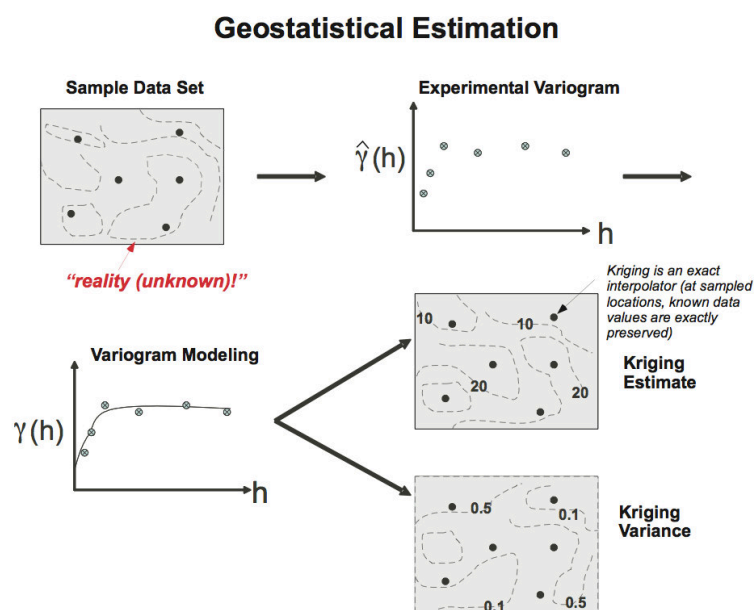


Figure 2.6 General scheme showing the estimation process in geostatistics (Zhang, 2011).

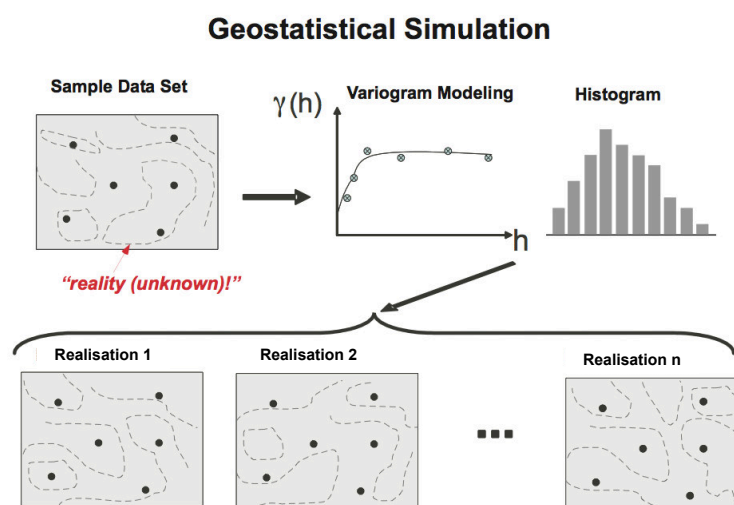


Figure 2.7 General scheme showing the simulation process in geostatistics (Zhang, 2011).

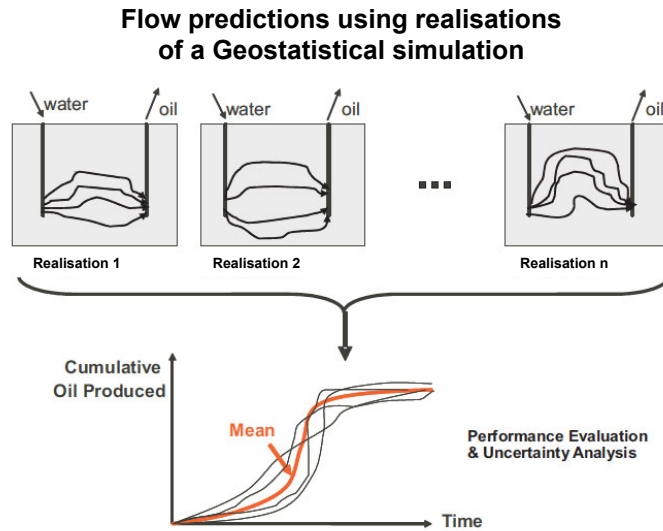


Figure 2.8 Flow predictions based on geostatistical simulations (Zhang, 2011).

A RocScience report untitled “Geomechanics software solutions used worldwide by geotechnical engineers” dated from 2003, states the following “...*Given the potential improvements to design and the successes of geostatistics in resource estimation, we believe that geotechnical engineering should seriously consider the discipline. Geostatistics will facilitate accurate interpretation of ground conditions based on the sparse input information characteristics of geotechnical engineering*”. In the same line, Rouhani et al. (1996) exposed the pertinent role of geostatistical tools in moving forward to build more realistic models in geotechnics and to solve environmental problems. Petroleum, hydrogeology and mining are examples of areas where this technique has been widely used, in the last case mainly to analyse ore grades (Krige, 1951, 1966; Krige et al., 1989), while in the petroleum area, the technique allows modelling reservoir using geophysical data.

2.5.2. Exploratory and covariance analyses of regionalised variables

As mentioned, a regionalised variable can be defined through a deterministic function that can vary in space and that is normally associated with a natural phenomenon. Examples of regionalised variables are: ore grades, rock density, rock type, etc. Figure 2.9 displays a regionalised variable represented in $1D$ and shows the irregularity in its spatial variation.

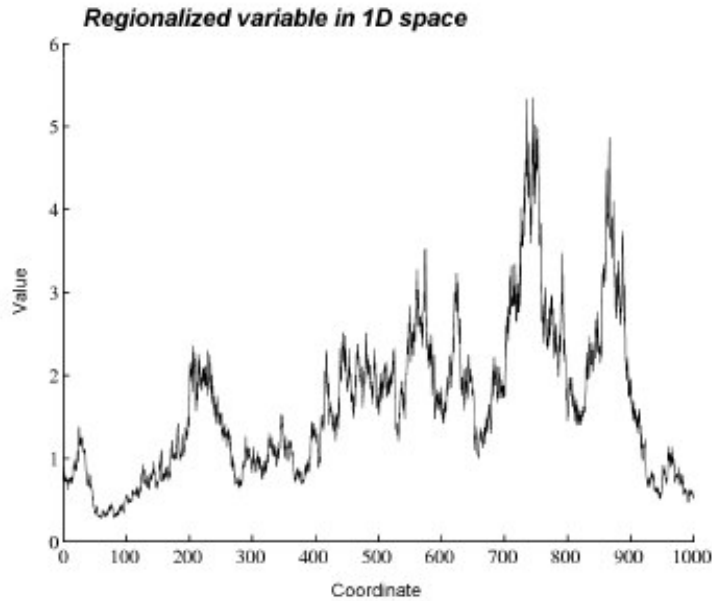


Figure 2.9 Regionalised variables example in 1D.

Let us assume $z(x)$ the regionalised variable value in a location x and $Z(x)$ the parent random field. Once again, assume stationarity in the case of invariance of the spatial distribution under translation, *i.e.* the average value of the regionalised variable. In such a case, the covariance function between two points, can be obtained through Equation (2.12).

$$Cov(h) = cov[Z(x), Z(x + h)] = E[Z(x)Z(x + h)] - m^2 \quad (2.12)$$

where m is the mean value (constant in space) of the random field Z .

This covariance function allows measuring the similarity degree between the values at different locations (distance equal to h) in a defined direction. In the case of a change of direction, the covariance can have a different value and can represent an anisotropic phenomenon. If the covariance value is the same for all directions the phenomenon can be classified as isotropic. To verify all of this, a directional study should be performed calculating the covariance for different directions.

2.5.3. Variogram

In contrast to the covariance function, the variogram is a mathematical function used to evaluate the dissimilarity existing between values at different locations and, as such, to characterise the spatial behaviour of a random field $Z(x)$. This function acts as an alternative to the covariance function,

and is one of the most used tools to classify the random field as stationary or non-stationary. In geostatistics the most common case is the second-order stationary model, where the random field is defined by a zero expectation or a constant value m and a stationary variance, covariance function or variogram (independent of location x). On the contrary, in the case of non-stationary random fields the opposite happens, and it is impossible to define a unique deterministic value to characterise the random field moments.

Therefore, the experimental variogram of a variable $Z(x)$, also acknowledged as semi-variogram (Deutsch, 2002), can be translated by the squared difference of the values in two points separated by a distance vector h , as presented in Equation (2.13).

$$\gamma(h) = \frac{1}{2N(h)} \sum (Z(x) - Z(x + h))^2 \quad (2.13)$$

where $N(h)$ represents the number of pairs of points at a distance h , h denotes the distance between the pairs of points and $Z(x)$ represents the random variable value at location x .

According to Figure 2.10, a variogram model can be defined according to two main characteristics, the sill and the range. The first one represents the difference between the most correlated points (variogram origin) with the points with the highest variance (most distant point). It can also be translated into the value or level at which the variance between the points starts to be constant. The sill ($C + C_0$) is reached when the distance between the points tends to infinity, *i.e.* where $Z(x)$ and $Z(x + h)$ does not display any spatial correlation. The second characteristic considers the variogram range (a) that is representative of a distance value at which the sill value is reached. Presumably, beyond the range the autocorrelation between points is essentially zero. Furthermore, an additional feature that could exist in the variogram is the *nugget effect* (C_0). This effect counteracts the theory of the variogram origin being equal to zero (which should be the expected configuration) due to measurement and positions errors or some natural spatial variability of the random variable.

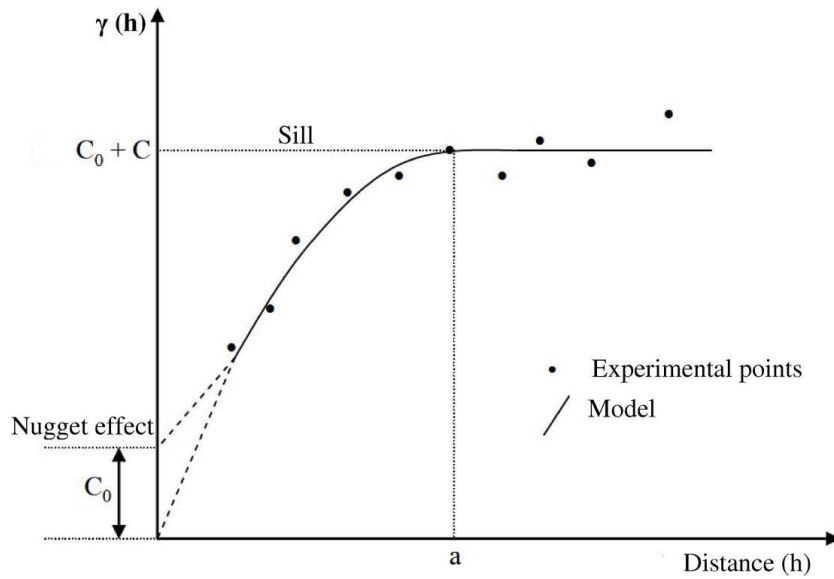


Figure 2.10 Variogram general representation of a stationary random variable (Imanzadeh, 2013).

The number of pairs of points adequate to build the experimental variogram can vary considerably, being 30 the minimum number used in the mining industry (Journel, 1977); however, in geotechnical engineering the available data do not always satisfy this condition. Note that in geostatistics, the samples configuration can vary from regular sampling to purely random (Imanzadeh, 2013).

As mentioned, the variogram can also be used to test the anisotropy of the random field by computing an experimental variogram in several directions, resulting in the called variogram map. The output is a colour map like the ones presented in Figure 2.12, where it is possible to distinguish an isotropic (see Figure 2.11a) of an anisotropic variogram (see Figure 2.11b).

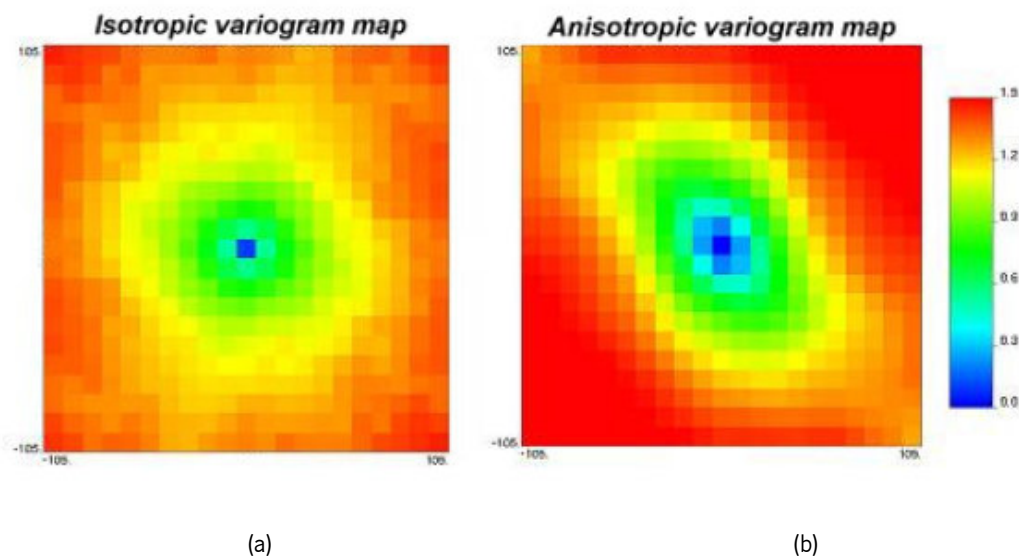


Figure 2.11 Variogram map examples for a random field with: a) an isotropic behaviour; and b) an anisotropic behaviour.

A similar analysis can be performed by computing the variogram in different directions and performing a comparison between the sill and range values. In the case of similarity, the field presents an isotropic behaviour in the computed directions, otherwise some anisotropy should be considered for further prediction (as the example presented in Figure 2.12).

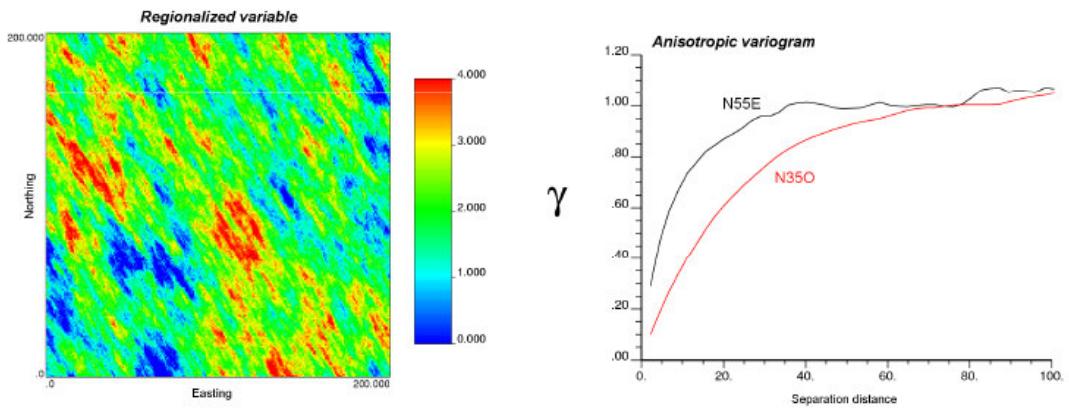


Figure 2.12 Variogram and realisation examples for an anisotropic field (different ranges for N55E and N35O).

At this stage, it is worth mentioning the parameters required in the experimental variogram computation, all represented in Figure 2.13. The azimuth and dip are needed to define the direction of interest and the respective tolerances, as well as the bandwidth and band height used to limit the number of associations of data pairs (A). Regarding distances, the lag parameter (elementary variogram distance), the spacing and tolerance should be defined (B). The first is obtained based on the data spacing in the direction of interest and, consequently, also allows defining the number of lags such that the variogram considers less than half the data domain size; the lag tolerance is usually half of the lag spacing (Deutsch, 2015).

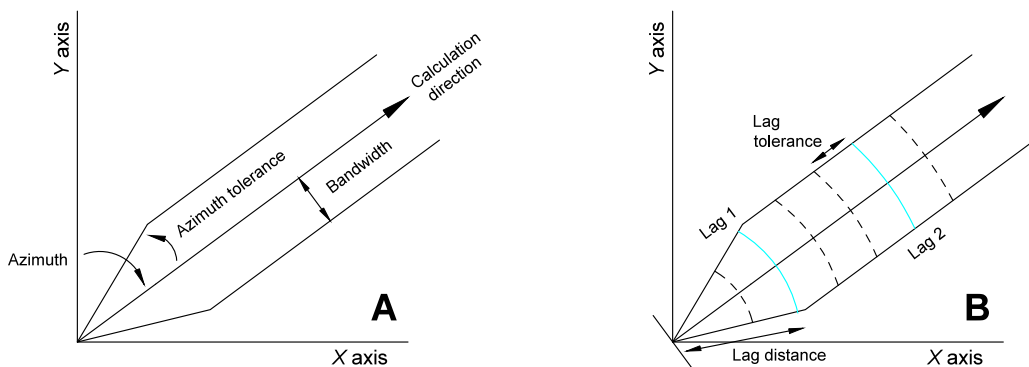


Figure 2.13 Representation image of the variogram parameters (adapted from Deutsch (2015)).

Moreover, after the experimental variogram computation and for the sake of prediction (estimation or simulation), it should be modelled using mathematical functions. The chosen mathematical model should be as fitted as possible, more importantly in the initial part of the variogram (high correlation values), in order to avoid errors or translate erroneous behaviours of the random field prediction. The most commonly used models represented in Figure 2.14 are also the mathematical functions displayed in Table 2.6.

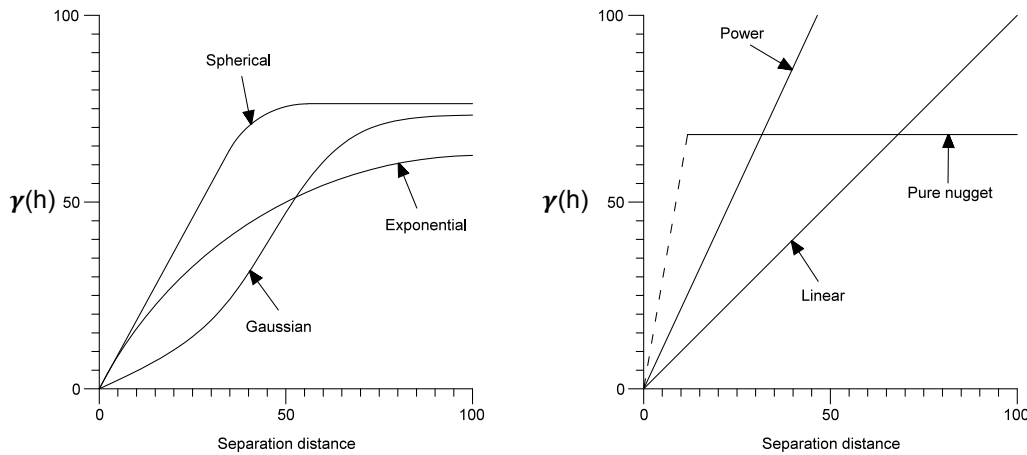


Figure 2.14 Variogram models' examples set for a nugget effect equal to zero (adapted from Jaksa et al. (1997)).

Table 2.6 Mathematical functions and respective models to use in variogram fitting (Jaksa et al., 1997).

Model	Mathematical functions	
Pure nugget	$\gamma(h) = C_0$	$C_0 =$ nugget effect
Spherical	$\gamma(h) = C \left(\frac{3h}{2a} - \frac{h^3}{2a^3} \right) + C_0$	for $h \leq a$, $a =$ range
	$\gamma(h) = C + C_0$	for $h \geq a$, $C + C_0 =$ sill
Exponential	$\gamma(h) = C \left(1 - e^{-\frac{3h}{a}} \right) + C_0$	
Gaussian	$\gamma(h) = C \left(1 - e^{-\left(\frac{\sqrt{3}h}{a} \right)^2} \right) + C_0$	
Linear	$\gamma(h) = ph + C_0$	$p =$ slope
Power	$\gamma(h) = ph^a + C_0$	$0 < a < 2$
Cubic	$\gamma(h) = C \left[7 \left(\frac{h}{a} \right)^2 - \frac{35}{4} \left(\frac{h}{a} \right)^3 + \frac{7}{2} \left(\frac{h}{a} \right)^5 - \frac{3}{4} \left(\frac{h}{a} \right)^7 \right]$	for $h \leq a$
	$\gamma(h) = C + C_0$	for $h > a$

In the case of more irregular variograms the use of a single mathematical function could not be enough for the model fitting; therefore, two or more basic variograms functions can be combined, resulting in the so-called *nested structures*, as presented in Figure 2.15.

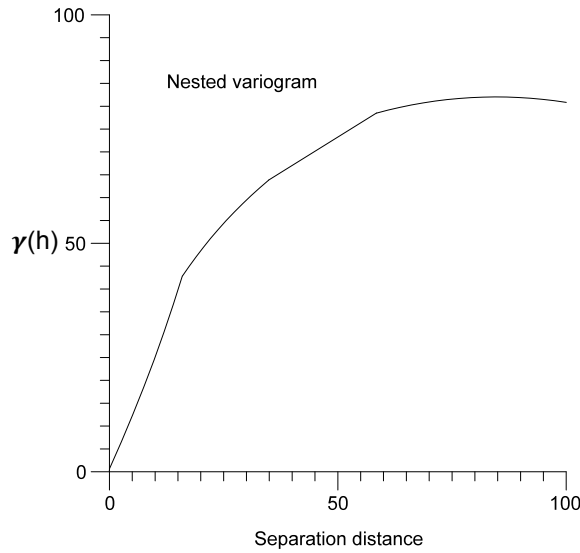


Figure 2.15 *Nested structures* theoretical example considering a small *nugget* effect.

As a reference, Figure 2.16 shows the difference in behaviour for three random fields with distinct variograms near the origin. In a spatially regular or continuous field the transitions from value to value are made in a smoother way, while in an erratic field the variation is more random and with no mathematical order.

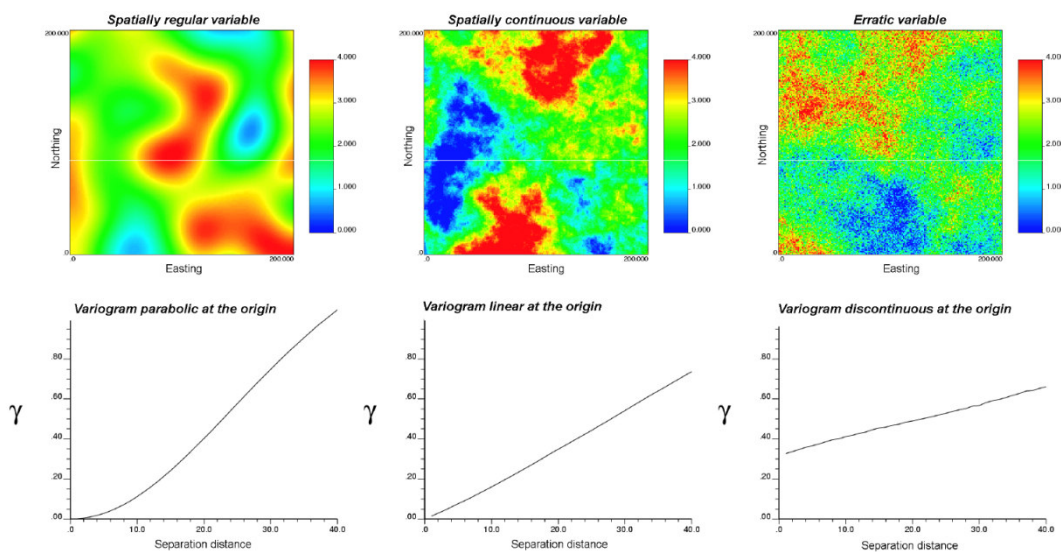


Figure 2.16 Variograms and realisations examples for three fields with distinct spatial behaviour.

2.5.4. Estimation

The geostatistical estimation of one or more random fields can be performed through the Kriging interpolator (Krige, 1951, 1966; Krige et al., 1989). This technique allows the estimation of the variable and the associated errors at locations without any type of information, using only the spatial structure of preliminary data, through the covariance function or the variogram. In contrast to other linear interpolators, like moving average or inverse distance weighting, developed in a deterministic context, Kriging relies on a probabilistic model (covariance or variogram).

Based on the random field model, the kriging weights must be calculated depending on the importance or influence that each sample point takes in the target point estimation value. The weight values are defined according to: 1) the distances between the data and the target locations, *i.e.* for close points the weights should be higher than for distant points; 2) the redundancies between the data, *i.e.* data points that are spatially close should receive smaller individual weights; and 3) the spatial continuity of the regionalised variable, so that more weight is given to closest data in the case of a regular variogram, more balanced weighting is given when a nugget effect exists, and in case of anisotropy more importance is given to the data located along the direction with larger correlation range with respect to the target location (Emery and Cornejo, 2010).

In geostatistics the kriging technique, also designated as the best interpolator or estimator, is obtained by imposing three restrictions, a linearity restriction, an unbiasedness restriction and an optimality restriction.

The linearity restriction imposes writing the random field estimation ($Z^*(x)$) at a target location x_0 as a linear combination of the random field data at locations x_1, \dots, x_n like,

$$Z^*(x_0) = a + \sum_{\alpha=1}^n \lambda_{\alpha} Z(x_{\alpha}) \quad (2.14)$$

Concerning the second restriction, it is guaranteed by assuming the estimation error as a random variable with an expectation equal to zero, as presented in Equation (2.15).

$$E[Z^*(x_0) - Z(x_0)] = 0 \quad (2.15)$$

being $Z^*(x_0)$ the estimation value at the target location and $Z(x_0)$ the true value at the target location.

For the last restriction, the variance of the error should be as small as possible in order to increase the estimation precision, so the dispersion between the estimated value and the true value in the target location should be given by:

$$\sigma_k^2(x_0) = \text{var}[Z^*(x_0) - Z(x_0)] \quad (2.16)$$

In order to perform the Kriging estimation, it is important to be certain about the amount of data to use in each estimation, therefore, two types of implementations can be applied. The one that considers a *unique neighbourhood*, where all the available data are used in each target point estimation, and the *moving neighbourhood*, which considers a limited number of data inside an ellipse to be defined by the practitioner.

In addition, there are several types of kriging like Simple kriging, where the mean value of the random field is known under a stationary assumption, the Ordinary kriging with unknown mean, the Universal kriging more fitted for non-stationary random field and the Indicator kriging headed for categorical variables (Krige, 1966). Following the equations presented before, in Simple kriging the linearity restriction is given by:

$$Z^*(x_0) = a + \sum_{\alpha=1}^n \lambda_{\alpha} Z(x_{\alpha}) \quad (2.17)$$

while the unbiasedness restriction results in:

$$0 = E[Z^*(x_0) - Z(x_0)] = a + \sum_{\alpha=1}^n \lambda_{\alpha} E[Z(x_{\alpha})] - E[Z(x_0)] = a + \sum_{\alpha=1}^n (\lambda_{\alpha} - 1) m \quad (2.18)$$

implying that the additive coefficient a is equal to the random field mean value, m weighted by the complement of the weights given to the data. Moreover, the variance minimisation results in:

$$\text{var}[Z^*(x_0) - Z(x_0)] = \sigma^2 + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} C(x_{\alpha} - x_{\beta}) - 2 \sum_{\alpha=1}^n \lambda_{\alpha} C(x_{\alpha} - x_0) \quad (2.19)$$

being $C(x_{\alpha} - x_{\beta})$ the covariance between the point located at x_{α} and the point located at x_{β} and λ_{α} and λ_{β} the corresponding weights.

Equations (2.17), (2.18) and (2.19) can be combined in an $AX = B$ form that could be solved using a matrix inversion or Gauss pivot.

Another estimation type is Ordinary Kriging, which is settled considering an unknown mean m but a known variogram that can or not present a sill. This fact makes Ordinary Kriging most robust, and therefore, the most applied estimator when compared with Simple Kriging. The linear combination of the neighbourhood data is made in the same way as was presented for Simple Kriging in Equation (2.17); however, since the mean is unknown the only way to solve the unbiasedness restriction is to assume that the weights λ_α add to 1. Regarding the optimality restriction, an additional unknown called Lagrange multiplier must be added to solve the Equation (2.20).

$$\text{var}[Z^*(x_0) - Z(x_0)] = - \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_\alpha \lambda_\beta C(x_\alpha - x_\beta) + 2 \sum_{\alpha=1}^n \lambda_\alpha C(x_\alpha - x_0) \quad (2.20)$$

and the final system of equation is obtained:

$$\left\{ \begin{array}{l} \sum_{\alpha=1}^n \lambda_\alpha = 1 \\ \forall \alpha = 1..n, \sum_{\beta=1}^n \lambda_\beta C(x_\alpha - x_\beta) - \mu = C(x_\alpha - x_0) \end{array} \right. \quad (2.21)$$

A multivariate version of kriging can be constructed, the designated cokriging (Krige, 1966). In this type of kriging the aim is to estimate the random field at a target location by considering the data about itself and other correlated covariates. To do that, cokriging requires the variogram (covariance) and cross variogram (covariance) computation for each pair of variables in order to measure the cross-correlation between them. The data locations of each variable can be scattered differently, all the variables in the exact same locations forming the so-called isotopic case. Secondly, in the partially heterotopic case, the variables share some of the data locations and finally, in the totally heterotopic case, the variables are sampled in disjoint sets of locations (see Figure 2.17). A more striking case can emerge from the heterotopic case, the collocated cokriging that considers the variable of interest in a few points and the covariates exhaustively known in the considered domain. In this case, the moving neighbourhood choice is a vital step.

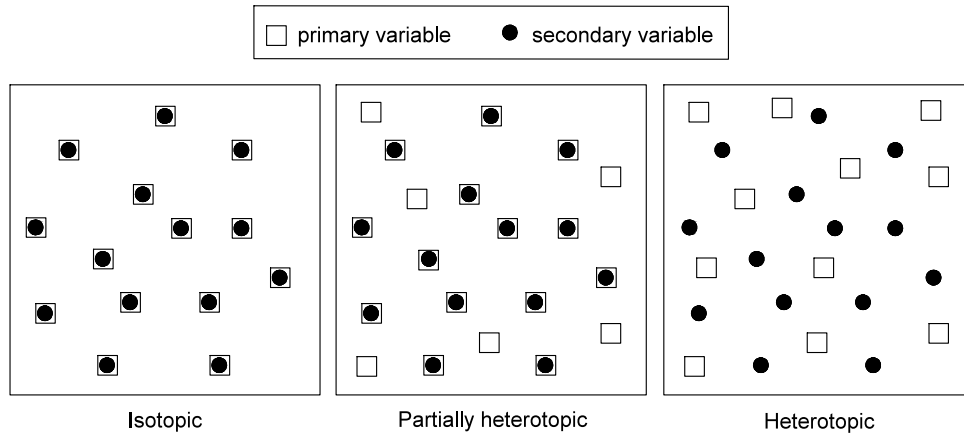


Figure 2.17 Cokriging cases: a) Isotopic; b) Partially heterotopic; and c) Heterotopic (adapted from Wackernagel (2006)).

It is worth mentioning that one of the main advantages of applying cokriging is to obtain more precise estimations and reduced values for the associated errors (cokriging variance and standard deviation) although the calculation time increases (Wackernagel, 2006).

To summarise, the kriging and cokriging estimators can be defined by the following properties:

- Exact interpolators: the estimation at the data location coincides with the data values, delivering a zero-error variance;
- Smoothing effect: the estimated values are less dispersed than the true values. Accordingly, the estimation tends to over-estimate the low values and under-estimate the high values;
- Additivity: the estimation of an average value over a region is the same as the average of the point-support estimation;
- Unbiasedness: the average of the estimation errors over a region tends to zero as the region becomes infinitively large.

2.5.5. Simulation

As mentioned before, the estimation presents a smoothing effect where the high values are under-estimated and the low values are over-estimated, which affects the identification and assessment of the heterogeneities and spatial variability of the random field. In order to achieve that, it is important to consider these extreme values and the geostatistical simulation is a viable alternative to the estimation. The aim of simulation is to generate as output multiples realisations (*i.e.* different interpretations of the reality) of a random field for one or more regionalised variables (Chilès and Delfiner, 2012). In detail, the simulation can be divided into two types, the conditional simulation

where the preliminary data used for simulation are honoured in value and location in all the output realisations, and the non-conditional simulation where the preliminary data are only used to define the random variable structural behaviour (variogram) for the simulation process. Consequently, this last kind of simulation can lead to higher variability between the realisations (Chilès and Delfiner, 2012; Imanzadeh, 2013). The main differences between the two types of simulation and a kriging estimation are presented in Figure 2.18.

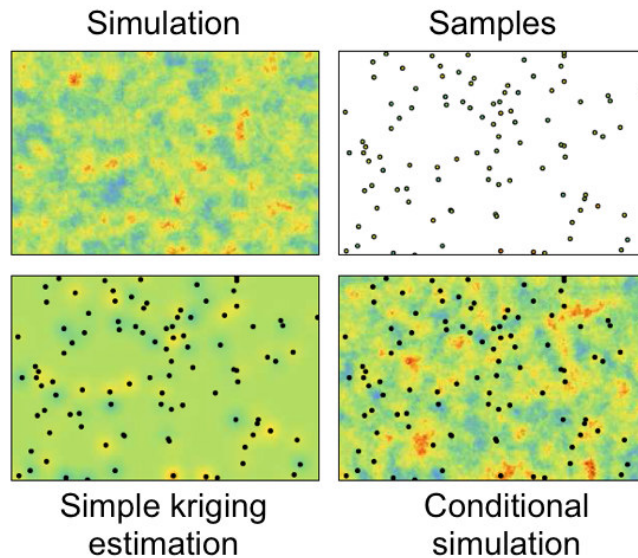


Figure 2.18 Representative images to compare preliminary data configuration, the non-conditional simulation, kriging estimation and conditional simulation (Wackernagel, 2006).

Assuming that a simulation algorithm is applied n times, n independent realisations are generated and their average result in a smaller range between the extreme values (low and high values) until a close solution to the geostatistical estimation (Kriging) is reached (Dubost, 2009). Even though, and as already mentioned, geostatistical simulation emerges as the most adequate alternative to the estimation when the goal is to identify the extreme values of a random field for cases where the information is not abundant.

The number of realisations to calculate can be variable; however, according to a wide range of the published works, the most used and adequate number is of the order of 100. As stated by Chilès and Delfiner (2012) this is the number that allows a trustful post-analysis process. However, if desired, a convergence study can be performed with the intention to reduce the number of realisations. The fact that a large number of realisations is obtained, in the post process analysis an additional map can be computed, the probability map. This map provides information about the probability of finding zones where the field of interest shows a higher or lower value when compared with a predefined threshold.

The variety of simulation algorithms is quite vast according to the type of random field to consider. As such, they can be grouped into two large types, the Gaussian random fields and the non-Gaussian ones.

The Gaussian random field model

The Gaussian random field model is convenient, since there are many algorithms for constructing non-conditional realisations, *e.g.* sequential, autoregressive, spectral and turning bands (Chilès and Delfiner, 2012).

Furthermore, in this model, a conditional simulation is reached by using the estimator kriging to condition the random field to the preliminary data samples, *i.e.* firstly a non-conditional simulation is performed, only using the variogram information and then a kriging estimation is performed in the conditional points (preliminary data locations). The conditional simulation of each point can be achieved after applying Equation (2.22).

$$Y_{CS}(x) = Y^{SK}(x) + [Y_s(x) - Y_s^{SK}(x)] \quad (2.22)$$

where $Y^{SK}(x)$ represents the random field obtained from Simple kriging at location x , $Y_s(x)$ is the result of the random field non-conditional simulation at location x , and $Y_s^{SK}(x)$ denotes the random field value obtained after performing a kriging using the information of the non-conditional simulation.

In practice, as a prior step the Gaussian methods require the normal transformation of the field to be simulated, so that it has a standard normal distribution with zero mean and unit variance, $Y(x) \sim N(0,1)$. To this effect, an anamorphosis function ψ is used to transform the original data (random field $Z(x)$) into Gaussian data (random field $Y(x)$) (equation (2.23)).

$$Z(x) = \psi(Y(x)) \quad (2.23)$$

wherein ψ represents a bijective function mandated to draw a relation between the field $Y(x)$ with a normal distribution $N(0,1)$ and the field $Z(x)$. In practice, this function can be modelled using a piecewise linear interpolation between empirical points and exponential functions for tail extrapolation, as explained in Emery and Lantuéjoul (2006).

In summary, in order to perform a geostatistical simulation using the Gaussian methods, the following general steps are required (see Figure 2.19).

1. Exploratory analysis of the input data (basic statistics of the random variable(s));
2. Transform the data into a normal Gaussian distribution using the anamorphosis function;
3. Perform an anisotropy study by computing the experimental variogram or covariance in several directions of interest;
4. Chose a variogram (covariance) model to use in the remaining simulation process;
5. Define a simulation grid with the target points coordinates (can be a $2D$ or a $3D$ grid);
6. Chose the simulation algorithm adequate for the variable(s) of interest; condition to the preliminary data by kriging;
7. Back-transform the normal Gaussian data into their normal scale;
8. Validate the simulation results and post-process the results.

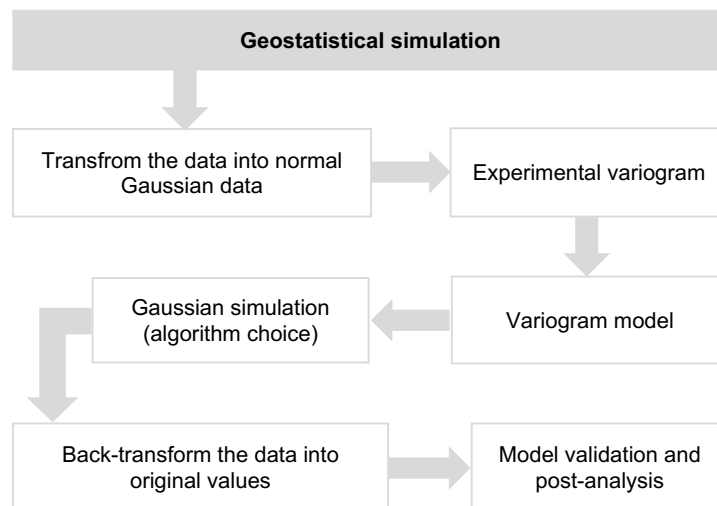


Figure 2.19 General steps required in a geostatistical simulation process using the Gaussian random field model.

Truncated Gaussian and Plurigaussian Models

Unlike the Gaussian model, the truncated Gaussian and the Plurigaussian models are fitted for categorical variable like rock types (Armstrong et al., 2011; Emery and Cornejo, 2010). These models intend to reproduce domains partitioning the space by truncating one or more Gaussian random fields. To this end, one has to establish a truncation rule that defines the contacts between domains and the proportion of space covered by each domain. In a simple case, the number of Gaussian fields used is 1 (Y_1) but to reproduce complex rock type relationships it is often necessary to construct a truncation rule using two (Y_1 e Y_2) or three (Y_1 , Y_2 and Y_3) Gaussian fields (see Figure 2.20).

Therefore, the essential steps of truncated Gaussian or Plurigaussian simulation are as follows:

1. Establish the domain proportions and their contact relationships by constructing a truncation rule. As an example, consider two Gaussian random fields Y_1 and Y_2 characterised by the following truncation rule:

$$\begin{cases} 1 \text{ if } Y_1(x) < t_1 \\ 2 \text{ if } Y_1(x) \geq t_1 \text{ and } Y_2(x) < t_2 \\ 3 \text{ if } Y_1(x) \geq t_1 \text{ and } Y_2(x) \geq t_2 \end{cases} \quad (2.24)$$

where t_1 and t_2 represent thresholds;

2. Using the truncation rule, a variogram analysis should be performed by codifying the domains as indicators, meaning the use of categorical variables;
3. Simulate each Gaussian random field $\{Y_1(x), Y_2(x)\}$ at the data locations conditionally to the domains. This part is guaranteed by using the Gibbs sampler algorithm (Armstrong et al., 2011);
4. Simulate each Gaussian random field at the target locations, the predefined simulation grid, conditioned to the output realisations from the previous step. These realisations are then truncated according to the truncation rule in order to be converted into indicators values.

In closer detail, the Gibbs sampler algorithm is used to transform the indicators values into Gaussian data by using an iterative process of acceptance or rejection. So, the Gaussian values are randomly generated and then verified if they are inside the Gaussian thresholds and approximated to the variable distribution.

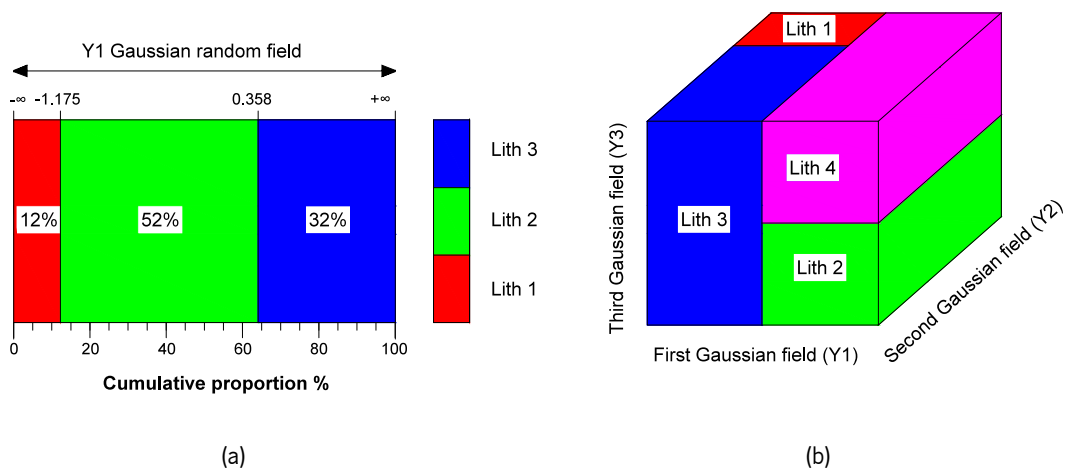


Figure 2.20 Examples of: a) 1D truncation flag; and b) a 3D truncation flag.

Simulation algorithm

Both previously presented models require a simulation algorithm, therefore the most common is the turning bands method (TBM). This algorithm, initially proposed by Chentov (1957) and later improved and used for stationary Gaussian fields by Matheron (1973), aims to simplify the simulation problem in multidimensional spaces by performing simulation in only one direction. To do that, the algorithm rely in lines that span the $2D$ or $3D$ space. After that, the $1D$ simulation is projected into a $2D$ or $3D$ space, in detail the $1D$ simulations in a location x are summed after projecting x into several directions as given in Equation (2.25).

$$Y(x) = \frac{1}{\sqrt{N}} \sum_{k=1}^N Y_k^{(1)}(\langle x | \theta_k \rangle) \quad \text{for } x \in \mathcal{D} \quad (2.25)$$

where $Y_k^{(1)}$ is the $1D$ random function with covariance C_1 for realisations $k = 1 \dots N$, θ_k a vector in \mathcal{R}^d and $\langle x | \theta_k \rangle$ represents the projection of x onto the line spanned by θ_k , as shown in Figure 2.21.

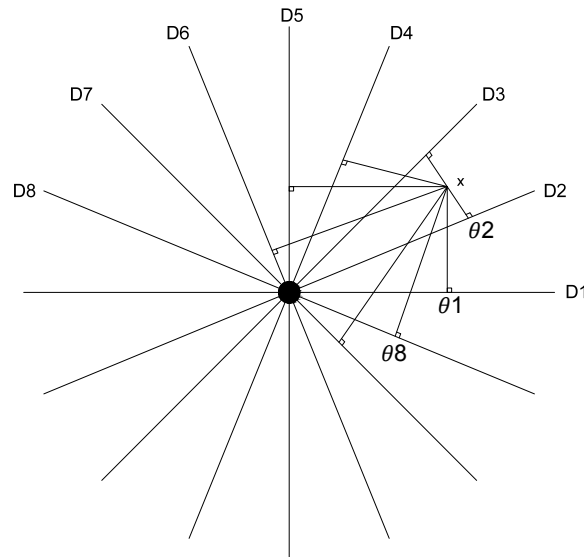


Figure 2.21 Turning Bands Method geometrical projection (adapted from Wackernagel (2006)).

The relation between the covariance function in 1 and d dimensions for an isotropic case can be defined in the $3D$ space as follows:

$$C_{1D}(h) = \frac{d}{dh} (hC_D(h)) \quad \text{and} \quad C_{3D}(h) = \int_0^1 C_{1D}(th) dt \quad (2.26)$$

While for a $2D$ case it is given by a more complex relation like:

$$C_{1D}(h) = 1 + h \int_0^{\pi/2} \frac{dC_{2D}}{dh}(h \sin \theta) d\theta \text{ and } C_{2D}(h) = \frac{1}{\pi} \int_0^{\pi} C_{1D}(h \sin \theta) d\theta \quad (2.27)$$

Moreover, the number of lines to use in this type of method can vary; however, Emery and Lantuéjoul (2006) refer an ideal value between 1000 and 1500 bands. Figure 2.22 shows four simulations of a random field with an exponential covariance function where it is possible to evaluate the differences when distinctive number of lines are used. In a case of using 1000 lines, the $2D$ random field simulation does not present any artificial features, contrary to the previous images where the number of lines is considerably smaller and some artificial features can be detected.

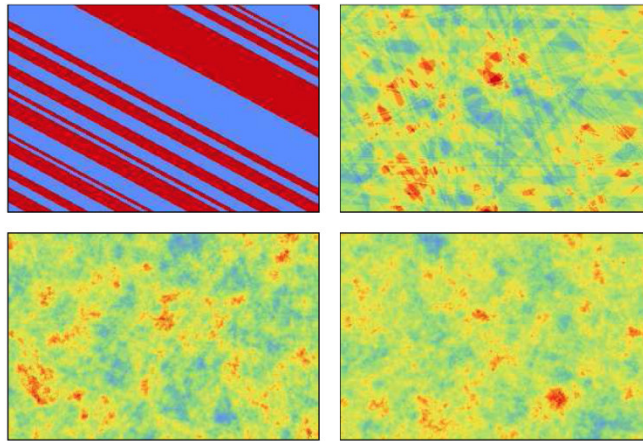


Figure 2.22 TBM simulation using 1, 10, 100 and 1000 lines.

In what concerns the TBM advantages, they are essentially the method speed that requires low computational time, the fact that it can simulate as many locations as desired and the exactly reproduction of the desired covariance function (or variogram). On the other hand, because a finite number of lines is used, some artificial features (the errors in the image construction) can be seen in the output maps if this number is insufficient (see Figure 2.22).

2.5.6. Model validation

Despite that several realisations of the variable of interest result from the geostatistical simulation, this does not mean that the model is geologically realistic or good for further investigations. Therefore, and before going into complex model validations, some easy checks also designated as *minimum criteria* should be completed. They include confirming that the preliminary data values are in the correct locations, analysing the distribution (histograms and basic statistics) of each realisation, confirming if it is similar to the one obtained from the preliminary data and, finally, confirming that the variogram of the realisations is, in average, the same as the one used for the simulation. Also, a visual check should be made for each realisation to confirm if the higher and lower values are placed in accordance with low and high variance values, respectively (Deutsch and Srinivasan, 1996). Similar criteria can be considered for estimation, although the reproduction of the data distribution and variogram is expected to be poor due to the smoothing effect of Kriging estimators.

In what concerns the validation of more complex models, in geostatistics two kinds of validations are usually tested, one designated as cross validation and another using the Jack-knife technique (Chilès and Delfiner, 2012). This model validation was the way found to answer to the following complex questions, “*How good is the estimated or simulated model?*” and “*How can the uncertainty of the model be assessed?*”.

Regarding the cross validation this is a technique that temporarily removes the observations, *i.e.* in each iteration each one of the preliminary data is removed and an estimation or simulation at the data location is performed. All this process is repeated for all preliminary data points, culminating in a comparison between the true values and the estimated or simulated values. Moreover, the validation using the Jack-knife technique consist in dividing the preliminary data in two data sets. The first data set, called training group, is used to perform the geostatistical estimation or simulation on the second data set locations, while the second group designated test group, is used to compare the estimated or simulated group with the true values. The preliminary data set division can vary, although the most common ratios are $2/3$ for the training group and $1/3$ for the test group.

In order to assess the model uncertainty using the previously mentioned techniques some performance criteria must be used. The most used ones are the linear regression between the true values and the simulated or estimated values, as well as the determination coefficient (R^2) calculation, and a probability interval (*PI*) accuracy plot. The latter relies on the calculation, at each test location x , of a series of symmetric p-probabilistic intervals (Goovaerts, 2001). For example, in the case of a *PI* equal to 0.9 it is bounded by quantile 0.05 and 0.95 of the set of simulated values;

ideally, the proportion of times that the true value is inside the probability interval should be close or equal to 90%.

Therefore, the final accuracy plot is constructed representing in the X -axis and Y -axis the probability interval values and the proportion of true values in the interval, respectively. The desirable accuracy plot would be the one where the points are coincident with the 45° line. An example is shown in Figure 2.23 for a hydraulic conductivity simulation, where the points are located slightly below the 45° line or identity line and, therefore slightly over estimated or simulated.

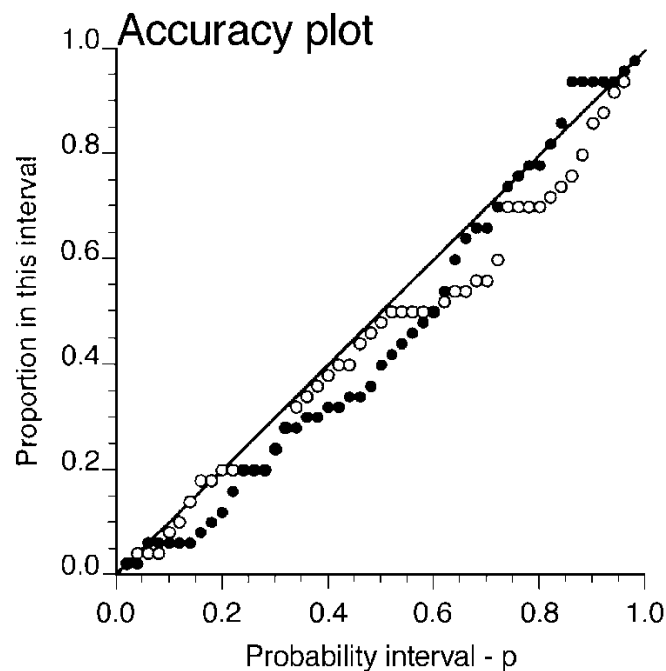


Figure 2.23 Accuracy plot example for true Hydraulic conductivity values fall within the probability interval. The black points represent the variable prediction using a Gaussian algorithm and the white points an indicator algorithm (Goovaerts, 2001).

2.5.7. Geostatistics applied to Geotechnics

The use of geostatistical technique in geotechnical engineering is a recent subject and therefore, the volume of published works in this fields is not as large as anticipated bearing in mind the power of these techniques. However, in recent years, some authors used these techniques to estimate or simulate geotechnical properties or features, as is the case of lithology facies whose works had been developed by Rosenbaum et al. (1997), Mao and Journel (1999) and Emery et al. (2008). In what regards rock mass empirical systems, the number of published works increases and it is important to highlight the following:

- Rock Quality Designation (RQD): the most relevant publication of the RQD estimation using a kriging technique was made by Ozturk and Nasuf (2002). In this work, a cokriging type was performed using correlated variables. Following, a more recent work of Ozturk and Simdi (2014), used kriging techniques to estimate not only the RQD values but also the rock mass and intact rock elasticity modulus. Similar work, of Madani and Asghari (2013), applied the sequential Gaussian simulation (SGS) to predict RQD values in order to detect rock mass faults in 3D blocks (RQD<20 were interpreted as fault zones);
- Rock Mass Rating (RMR): the number of works regarding this system is meaningful. Since 2002, Choi et al. (2002), Ryu et al. (2003), You (2003), Oh et al. (2004), Stavropoulou et al. (2007) and Exadaktylos and Stavropoulou (2008) applied kriging techniques to map the RMR values. Some of the mentioned works (Stavropoulou et al., 2007 and Exadaktylos and Stavropoulou, 2008), used the obtained estimation in numerical analysis of underground works (twin tunnels in China and metro tunnel sections in Barcelona). Some progress on RMR prediction has been achieved by Jeon et al. (2009), Egaña and Ortiz (2013), Ferrari et al. (2014) and, more recently, Pinheiro et al. (2016a, 2016b), whose works applied geostatistical simulation algorithms like SGS or TBM assuming the RMR final value as a continuous field, or simulating the RMR parameters individually as categorical fields;
- Geological Strength Index (GSI): once the GSI is correlated through the RMR system all the previously mentioned works can also be mentioned here; however, works like Deisman et al. (2013) or Ozturk and Simdi (2014) are worth mentioning regarding the GSI simulation;
- Joint frequency: the work of Ellefmo and Eidsvik (2009) employed kriging techniques to estimate the spatial frequency of joints and the associated variability in an iron ore located in Norway. Furthermore, Etminan and Seifi (2008) combined sequential Gaussian simulation (SGS) and sequential indicator simulation (SIS) to build a fracture model of the rock mass considering their azimuth, depth and density integrated with the porosity and permeability of the rock mass.

Though the mentioned works confirm the advantages in using geostatistical techniques to predict geotechnical parameters, most of them only apply estimation techniques like kriging; however, as previously mentioned in this chapter, simulation is the most effective and robust technique to identify, at the same time, the spatial variability of a random field and the existence of heterogeneities. Due to this fact, more work and practical applications in geostatistical simulation should be developed, starting by this PhD thesis.

2.6. SCENARIO REDUCTION

2.6.1. Introduction

One of the referred advantages regarding geostatistical simulation is the numerous realisations obtained as output, which can be, in a post process point of view, a serious problem due to the high computational time required to process all of them. As such, some works have been published proposing scenario reduction methodologies whose goal is the minimisation of the realisations set into a smaller set. These methodologies are considerably different from area to area, while in the petroleum side the highlighted works are Scheidt and Caers (2009a, 2009b), in the mining industry Armstrong et al. (2013) is worth mentioning.

Starting with the petroleum area, it is important to refer that the geostatistically generated realisations are then submitted to flow simulation to assess the reservoir flow performance, so in this case the scenario reduction is an essential step. As a consequence, Deutsch (1998, 1999) and more recently, McLennan and Deutsch (2005) proposed a simpler methodology based on the realisations rankings. In this case, the realisations are sorted in ascending order according with (Deutsch and Srinivasan, 1996; McLennan and Deutsch, 2005): i) statistical static measures (statistical average of geological parameters, e.g. the average net permeability); ii) fractional static measure (define the active fraction of the reservoir); or iii) volumetric static measures (volume of a reservoir capacity of oil transport, e.g. net oil in place also known as net hydrocarbon volume) and a subset of 10 realisations are evenly selected, more precisely realisations 1, 12, 23, 34, 45, 56, 67, 78, 89 and 100. In this subset, the realisations with worst and best performance are present (1 and 100). A more recent work from Rahim (2015) proposes a derived methodology from Armstrong et al. (2013), introducing the calculation of a Kantorovich distance between each realisation to quantify the probability distance between the superset and the subset of realisations. The main goal is to minimise this distance (for each removed realisation) using a linear optimisation algorithm until an optimal subset of realisations is found showing a similar reservoir performance (statistical distribution) in comparison with the superset of realisations. This methodology follows the line of previously developed methodologies in a stochastic optimisation field proposed by Armstrong et al. (2013), Heitsh and Romish (2003, 2009) and Dupacova et al. (2003).

Scheidt and Caers (2009a) proposed a methodology based on distance based kernel clustering, aiming to find a set of realisations able to statistically represent the whole set (distribution, quartiles, mean, variance, percentiles, etc.). In detail the methodology starts with multiple realisations (N_R)

generated previously using a geostatistical algorithm, and a dissimilarity matrix is computed between all the realisations. For the dissimilarity matrix, the Euclidean distance is used resulting in a $N_R \times N_R$ matrix. This matrix is then represented in a $2D$ space, called Euclidean space R using a multidimensional scaling. To consider a more linear behaviour for the points (realisations), a Kernel method is used resulting in a representation in the designated Featured Space F . The referred step is imperative to apply Principal Component Analysis (PCA) techniques to detect some patterns in the $2D$ representation. As a result, some points are selected among all the set and the percentiles P10, P50 and P90 are compared with the whole set (see Figure 2.24). A subsequent work of Scheidt and Caers (2009b) introduced the realisations clustering using a kernel k-means clustering algorithm to select the subset of realisations. In the same work a comparison is made between the traditional ranking method, the kernel k-means method and the random selection method, concluding that the distance based kernel methods provides more robust, accurate and less time-consuming results. Likewise, Singh et al. (2014) also uses the kernel k-means clustering method to quantify the uncertainty associated with various history matched geological models and to forecast production information.

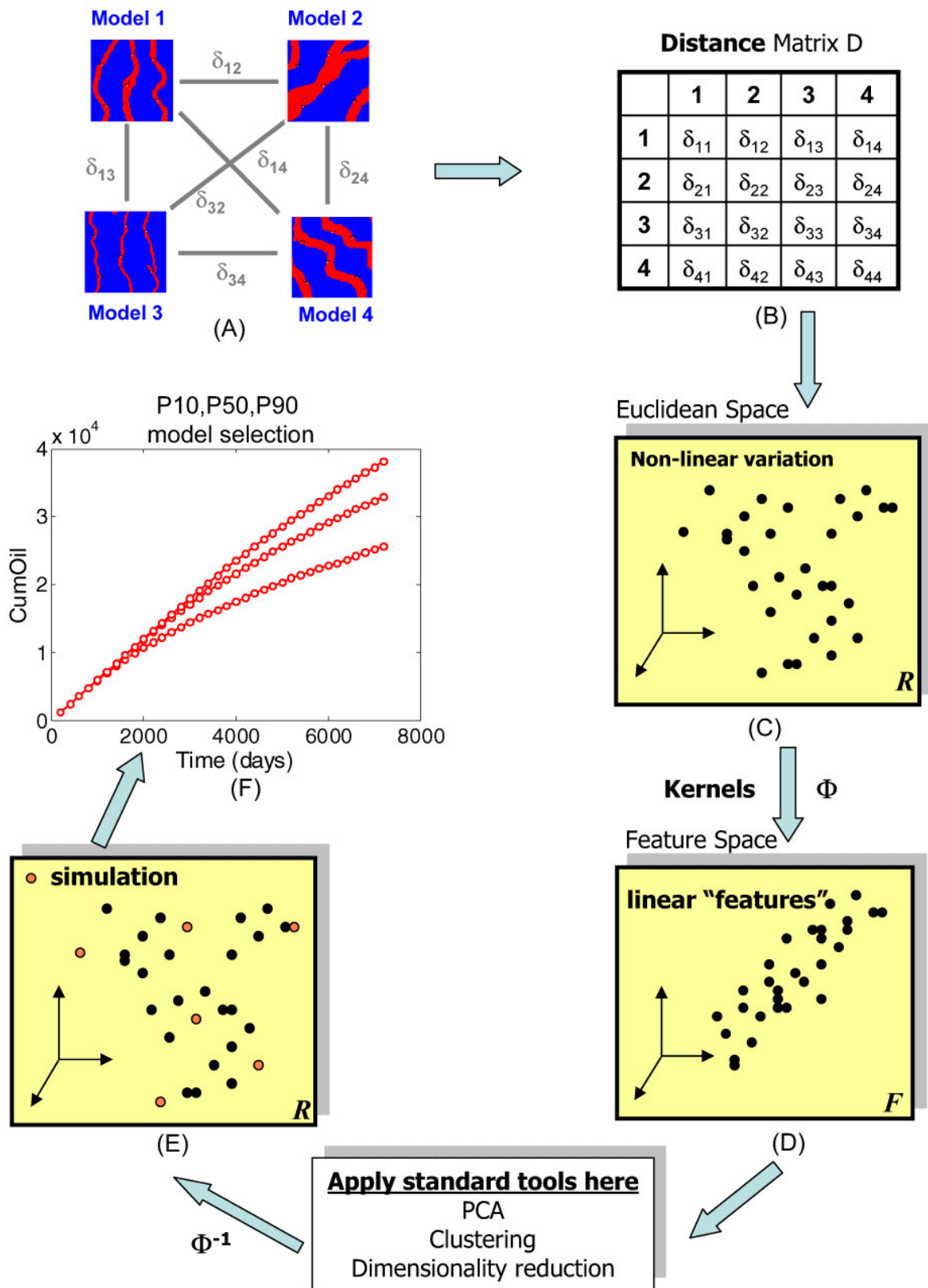


Figure 2.24 Proposed methodology for uncertainty quantification- (A) distance between two models, (B) distance matrix D, (C) models mapped in Euclidean space, (D) feature space, (E) pre-image construction, (F) P10, P50 and P90 quantile estimations (Scheidt and Caers, 2009b).

2.6.2. Kernel clustering technique

Since the clustering method shows the best results in reducing the super set of geostatistical realisations, more details will be given in this section. A cluster analysis divides the data into groups based on their relationships, so similar data should be placed in the same cluster while unrelated data are placed in different clusters. However, the definition of these clusters is imprecise and relies, mainly on the nature of the data and on the goal defined for the results. Due to that fact, the clusters are sometimes classified as unsupervised classification because the division is made with no *a priori* knowledge about the object classes.

There are two types of clustering, the hierarchical and the partitional. In the first type the clusters are organised as a tree, where each cluster (node) is the union of its sub clusters (children), while in the second type, the data are divided into non-overlapping clusters (subsets). The most relevant type is partitioning cluster, more precisely the squared-error based that include k-means and k-medoid techniques. Even though, in all clusters types the aim is to define some properties such as: 1) the number of clusters (that can also be an automatic step); 2) the absolute and relative positions of the clusters; 3) the clusters size, shape and density.

In what concerns the similarity measures used to compare clusters a wide range of formulas can be used; however, the most common is the Euclidean distance as given in Equation (2.28).

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.28)$$

where $d(x, y)$ represents the distance between realisation x and realisation y while x_i and y_i are the variable values in each i point of the realisation x and y , respectively.

Unlike the cluster based distance, the use of Euclidean distance for nonlinear data set can lead to clustering mistakes, therefore, applying the kernel transformation, all the non-linearity of the data set can be identified. This is an essential step before starting the clustering division using the k-medoid algorithm. Different types of kernel functions can be used to that effect, such as polynomial, Gaussian, chi-square, etc., being the Gaussian the most usual type since it has proved to be more robust. As an example, Equation (2.29) shows the Gaussian kernel function that basically consists in dividing the previously computed distance between the realisations (x and y) by a scale factor h carrying more linearity for the data.

$$K(x, y) = \exp\left(-\frac{\|x_i - y_i\|^2}{h^2}\right) \text{ with } h > 0 \quad (2.29)$$

In its turn, the scale factor h , also known as kernel bandwidth or smoothing parameter, controls the smoothness of the resulting points spatial representation that are less smooth for lower h values (the kernel becomes close to the identity matrix) and smoother for the contrary (the kernel is reduced to a constant matrix). To help this parameter definition, several authors have proposed reference values, most of the times based on the estimation errors minimisation; however, no reference values were found. For the purpose of clustering, Shi and Malik (2000) recommended the execution of several attempts using as reference 10% to 20% of the range of distance between points, namely the Euclidean distance.

In what concerns the generic formulation of k-medoid, it starts with a random partition of the data set in the Featured Space F using the kernel matrix values. Then, in an iterative way, the method improves the partition until the optimal arrangement for the clusters is found (equation (2.30)).

$$E = \sum_{i=1}^k \sum_{x \in C_i} d(x, med_i) \quad (2.30)$$

where E represents the sum of squared errors between the objects in the clusters in relation to their centre, also known as medoid med_i that should be minimised, C_i the number of generated clusters k and x the data objects.

In detail, the phases involved in the kernel k-medoid application are the following (see Table 2.7):

1. The algorithm initiates generating a total of k clusters randomly organised and the corresponding centres, called medoids, are defined (med_i);
2. The generated k clusters are evaluated through a distance calculation given by:

$$d(x, med_i) = k(x, x) - 2 \times k(x, k(med_i)) + k(k(med_i), k(med_i)) \quad (2.31)$$

where $k(x, x)$ corresponds to the kernel distance between the object x and itself, $k(x, k(med_i))$ the kernel distance between the object x and their medoids;

3. A Hill Climbing technique is applied where the bad medoids are replaced by new ones, which is chosen accordingly to the minimal average distance between itself and the clusters objects;
4. A new distance evaluation is made between the new medoids and the other objects that compose each cluster, resulting in an object re-assigning to new clusters accordingly to the closest medoid.

Table 2.7 General procedure to adapt for k-medoid algorithm implementation.

K-medoid algorithm**Phase 1- Initialisation**Set $x = x_1, \dots, x_n$ Randomly generated $C = \{C_1, \dots, C_k\}$ Objective function = ∞ **Phase 2 – iterative process**for all clusters $C_i \in C$ Calculate cluster medoid $med_i \subseteq \mathcal{R}^n$ for all object $x_i \in x$ for all clusters $C_i \in C$ Calculate $d(x, C_i) = d(x, med_i)$ Let $best(x, C_x) = \forall C_j \in C: [d(x, med_{C_x}) \leq d(x, med_{C_j})]$ end_for end_forReturn $best(x, C_x)$

End procedure

2.7. CONCLUSIONS

The present Chapter started by presenting the commonly used methods to assess the geomechanical parameters to use in a future characterisation model (empirical systems like RMR, GSI, RQD, etc.) As it has been presented in this Chapter, deterministic methods and tools are very limited to realistically characterise rock masses, mainly regarding their heterogeneity and spatial variability. This is due to the mistaken starting assumption that there is no estimation error in the geomechanical parameters of the rock mass.

In the last years, probabilistic techniques have been applied to overcome this limitation and provide range of values for the geomechanical parameters instead of single ones. However, they undertake the variable(s) global values (mean and variance) and only contemplate the pointwise variability of the parameters. By doing that, the extreme values are hardly identified and, therefore, the natural variability of the rock mass is not considered. Examples of these techniques are the MC simulation, FOSM and PEM, that apart from their simplicity are associated with significant computational times.

As the main goal of this work is to study the effect of the heterogeneity and spatial variability on the rock mass mechanical behaviour, the use of stochastic techniques, namely random fields, that are able to characterise the variable (s) spatial behaviour, is essential. An example of these techniques is geostatistics that, contrarily to the traditional probabilistic techniques, considers the variable as spatially dependent, which result in a more accurate and realistic prediction of the variable(s). As previously stated, the random field simulation is the most adequate approach when the goal is to quantify the spatial variability and heterogeneity since the extreme values are not smoothed, unlike the kriging estimation. As a result of the geostatistical simulation, n equivalent realities can be obtained for the random field(s); however, the use of the full realisations set could not be practical and a reduction scenario methodology based on a clustering algorithm was presented. Since the amount of techniques that can be used to reduce the full realisations set is limited, the ones commonly used in mining and petroleum engineer were presented; nonetheless, some changes need to be made to adapted them into the geomechanical practise. In what concerns the geostatistical simulation, there are a wide range of simulation algorithms fitted to different type of variables (continuous, discrete and objects), hence two of them were detailed exposed, along with their advantages and limitations (turning bands method and truncated Gaussian methods). Finally, aiming to assess the accuracy of the geostatistical simulation some validation techniques were presented, namely the cross validation and the jack knife technique.

Even though geostatistics can be used to predict the geomechanical parameters of the rock masses there are still insufficient works and research made in this field, mainly in the application of the technique in geotechnical engineering.

Chapter 3

GEOSTATISTICAL SIMULATION OF GEOMECHANICAL PARAMETERS

3.1. INTRODUCTION

The characterisation of rock masses in a more accurate and realistic way, mainly regarding the geomechanical parameters is, nowadays, an important task to accomplish in geotechnical engineering. Therefore, geostatistical techniques emerged as an alternative to traditional probabilistic approaches once they consider the spatial behaviour of the parameters.

In this Chapter, it is intended to highlight the main steps required to simulate a random field representing a geomechanical parameter along with an application example, as it will be of interest in the remaining chapters. In section 3.2 the used random field, in this case for the RMR empirical system, is presented alongside with the simulation approaches. Following, in section 3.3, all the details concerning the random field simulation methodology are presented, with modelling approaches based on a continuous scale and on a discrete one.

In Section 3.4, the approaches are applied to a case study aiming the Rock Mass Rating (RMR), which is used to geomechanically characterise the rock mass in geotechnical works. The results are compared in terms of computation time, practical implementation, level of details and post-processing outputs. Besides the RMR mapping and associated uncertainty, the deformation modulus of the rock mass is subsequently obtained based on these maps together with empirical expressions. Finally, all the simulations are validated through a split-sample jack-knife technique.

3.2. GEOSTATISTICAL SIMULATION MODELS

3.2.1. General

Geostatistical models can be used to develop an approach that is able to reduce the uncertainty associated with the rock mass characterisation, with the existing variability in this characterisation and with the natural heterogeneities of the rock masses. Indeed, in these models, the geomechanical parameters are viewed as outcomes (realisations) of spatial random fields, the properties of which can be inferred from the available *in situ* measurements and laboratory tests. Kriging techniques (Matheron, 1971) can be used to predict the values of the parameters of interest at any specific location, based on the information available at neighbouring locations and on the spatial correlation structure of the underlying random fields. These techniques aim to minimise the

expected squared error between predicted and true values, but, in return, they provide over-smoothed maps that do not reflect the actual variability of the true parameters. To avoid this drawback, conditional simulation techniques have been developed to construct numerical models that reproduce the spatial variability at all scales and allow a better understanding of the rock mass heterogeneities (Journel, 1974; Chilès and Delfiner, 2012). Unlike kriging, that provides a single prediction for each parameter of interest, simulation yields as many case scenarios as desired, which are helpful to assess the uncertainty in the actual (unknown) parameter values at any specific location or jointly over several locations. The use of geostatistical techniques in geotechnical engineering was already dealt with in Chapter 2.

The next section presents three geostatistical approaches to use in the geomechanical parameters simulation, depending on whether one considers that the properties are measured on a continuous quantitative scale or on a discrete scale. In the first and second approaches, the most straightforward and usual ones, the chosen variables are viewed as variables measured on a continuous scale and are directly simulated with a Gaussian algorithm; however, in brief, since more details will be given below, while the first approach simulates only the final variable, the second approach considers the underlying parameters and simulates them in their original scale. In contrast, the third approach is more complete, as the variables are considered on a discrete scale, and a specific geostatistical model (truncated Gaussian model) is used for the purpose of simulation.

3.2.2. First and second approaches: continuous variables

In these approaches, the variables are viewed in a continual scale (see Figure 0.1a). To this end, the Gaussian random field model is used, through the following steps (Chilès and Delfiner, 2012):

1. First, a representative distribution of the variables values is calculated, by weighting each value depending on the geometrical configuration of the data. This procedure aims at down-weighting the data that are spatially clustered, which contain redundant information (Deutsch and Journel, 1998). In case of a regular sampling design, the data value can be assigned the same weights;
2. The variables data are then transformed into data with a standard Gaussian distribution, accounting for the previously calculated declustering weights. These transformed data are associated with a parent second-order stationary Gaussian random field, which is fully characterised by its covariance function or, equivalently, by its variogram (Lantuéjoul, 2002);

3. The experimental variogram of the Gaussian data is computed and subsequently fitted with a theoretical model. At this stage, the study can be performed in one or more directions of space, in order to identify a possible anisotropy and to better understand the spatial behaviour of the data;
4. A Gaussian random field is then simulated at the target locations, conditionally to the available data (*i.e.* such that the values simulated at the data locations match the data values). In the present case, the turning bands algorithm (Emery and Lantuéjoul, 2006) is used for simulation;
5. The simulated Gaussian values are back-transformed to the original scale.

Similar approaches, which differ in the specific simulation algorithm used at step (4), have been proposed by Ryu et al. (2003), Jeon et al. (2009), Egaña and Ortiz (2013) and Ferrari et al. (2014), among others, for the spatial prediction of RMR and for uncertainty quantification.

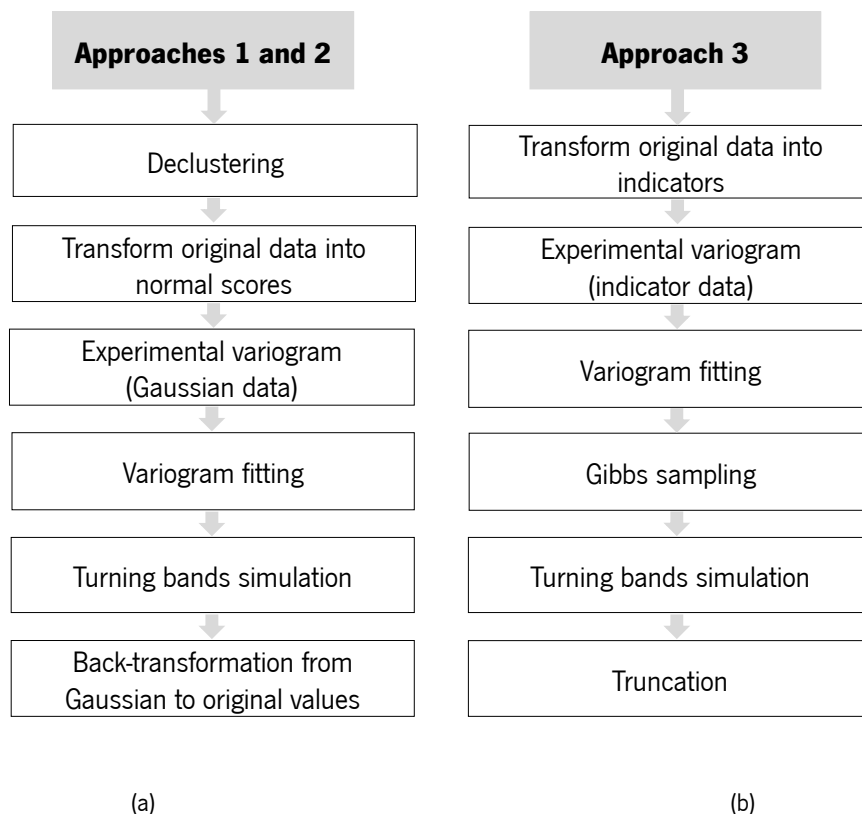


Figure 0.1 Flow charts for the variable simulation under: a) the Gaussian model (Approaches 1 and 2); and b) the truncated Gaussian model (Approach 3).

3.2.3. Third approach: discrete variable

The third approach is more innovative and consists in simulating the variable assigned by ratings, viewed as discrete variables (*i.e.* variables that only take integer values). For the simulation, the truncated Gaussian model (Armstrong, et al., 2011) is used, which relies on the truncation of second-order stationary Gaussian random fields. The application of this model is carried out through the following steps (see Figure 0.1b):

1. First, the original data are transformed into class-indicator data, *i.e.* data that take the value 0 or 1 depending on each parameter score;
2. A set of truncation thresholds is defined for the variable, which allows the proportion of each class to be reproduced by the simulation (Armstrong, et al., 2011);
3. Given the experimental variograms of the class-indicator data, a variogram can be calculated for the underlying Gaussian random field associated with the variable, based on the existing relationships between the indicator and Gaussian variograms (Emery and Cornejo, 2010). The obtained Gaussian variogram can be fitted with theoretical models;
4. The class-indicator data are then transformed into simulated Gaussian data, using an iterative algorithm known as the Gibbs sampler (Armstrong et al., 2011; Lantuéjoul, 2002);
5. The Gaussian random fields are simulated at the target locations, conditionally to the Gaussian values obtained at the previous step, and are truncated in order to get back to simulated class-indicator values. As for the first approach, the turning bands algorithm is used for the simulation of Gaussian fields;
6. The indicators are converted into the variable.

3.3. CASE STUDY – CHILEAN DEPOSIT

3.3.1. Introduction

The Rock Mass Rating (RMR) proposed by Bieniawski (1989) and already introduced in Chapter 2, was the system used to perform the geostatistical simulation using the previously mentioned approaches. This system allows classifying the rock mass into five classes (very good, good, fair, poor and very poor) using a continuous scale that varies from 0 to 100 obtained after weighting six individual parameters regarding the rock mass and its discontinuities. In this case study, the sixth parameter (P6) will not be used because it does not depend only on the characteristics of the rock discontinuities but also on their relationship with the structure and this is unknown. The RMR under consideration is therefore the so-called basic RMR, which is obtained considering only the contribution of parameters P1 to P5.

Therefore, the RMR system will be handled for the geostatistical simulation in three different ways (see Figure 0.2). The first two approaches consider the rating value of each one of the five parameters and the basic RMR final value as continuous variables. On the other hand, the third approach considers the parameters rating as discrete variables.

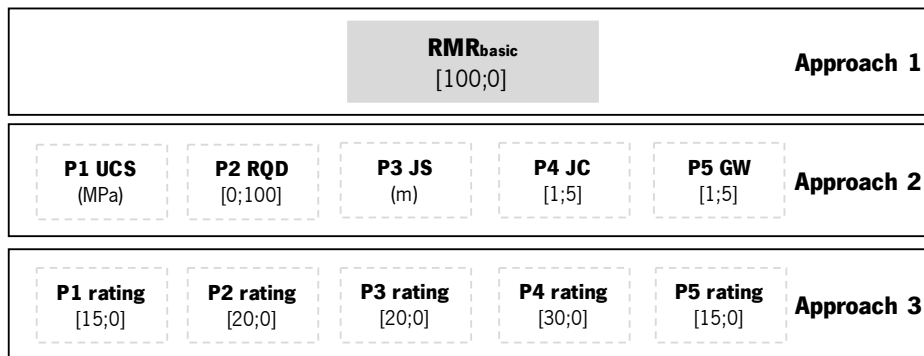


Figure 0.2 Scheme showing the three approaches considered in the RMR simulation.

The data set used in this case study was obtained from exploration boreholes performed in an epithermal gold deposit located in the “Cordillera de Los Andes”, region of Atacama, northern Chile, and surveyed through a set of exploration boreholes. The regional geology of the area is characterised by a group of intrusive, volcanic and sedimentary rocks, affected by fault zones that control the mineralisation, allowing the identification of four main lithological units of sedimentary rocks.

The available data comprise 3969 samples obtained from boreholes with a horizontal spacing of 40 m × 40 m and depths ranging from 96 m to 390 m. Along these boreholes, the samples were taken with a spacing of 20 m, yielding a regular sampling design of 40 m × 40 m × 20 m. The uniaxial compressive strength (P1) was measured from laboratory tests. Cylindrical rock samples were prepared according to the standard ASTM D4543–08, and then tested under uniaxial compressive conditions using the standards included in ASTM D7012–04. The average prepared sample has an aspect ratio (H/D) of 2. Table 0.1 presents the average density and the uniaxial compressive strength normalised to a diameter of 50 mm (UCS_{50mm}) for each lithological unit. The RQD (P2) was estimated directly from borehole logging. To estimate the average discontinuity spacing (P3), the average frequency of fractures (FF/m) was estimated. Bias correction was then applied by considering the average angle of each measured discontinuity. The condition of discontinuities (P4) was not quantitatively measured at the field. A regular condition was assumed for all lithological units. Finally, the water condition (P5) was assigned in agreement with the level of water determined at different depths at each borehole; two classes were mainly identified: *wet* and *damp*.

Table 0.1 Information about average UCS and average density by lithological unit.

Lithological unit	Description	Average density (t/m³)	Average UCS (MPa)
Silty and clayey limestone	Medium to fine calcareous sandstone and limestone	2.68±0.05	215±56
Calcareous sandstone	Fine bioclastic calcareous sandstones, limestones and bioclastic clams	2.68±0.09	154±47
Sandstone	Calcareous sandstone with rounded fragments of quartz	2.64±0.02	143±20
Calcareous sandstone	Sandstones with interblended limestones; levels of fine multicolored calcareous sandstones	2.63±0.06	152±40

According to the results of rock mechanics laboratory tests and the interpreted RMR values from the borehole samples, the rock mass is classified with a quality of fair to good (mostly in the range of 50 to 60). This range of RMR values is used in the mine design process. Plan views of the data are shown in Figure 0.3.

As mentioned above, in the first approach, the RMR was viewed as a variable measured on a continuous scale (0 to 100) and its value was directly simulated with a Gaussian algorithm, still the second approach uses the same algorithm to simulate each parameter individually (P1 to P5) also in a continuous scale. In contrast, the third approach uses a truncated Gaussian model, insofar as each one of the five parameters (P1 to P5) was simulated individually in a discrete scale. At the end of approach 2 and 3 simulations, the results are then summed to obtain the final mapping of RMR.

The RMR simulation should result in a better and improved understanding of the spatial distribution and an easier identification of heterogeneities and uncertainty levels. In addition to RMR, the underlying geomechanical parameters (P1 to P5) could also be mapped and consequently used in numerical models and mine design process in order to obtain a more accurate zoning of the rock mass.

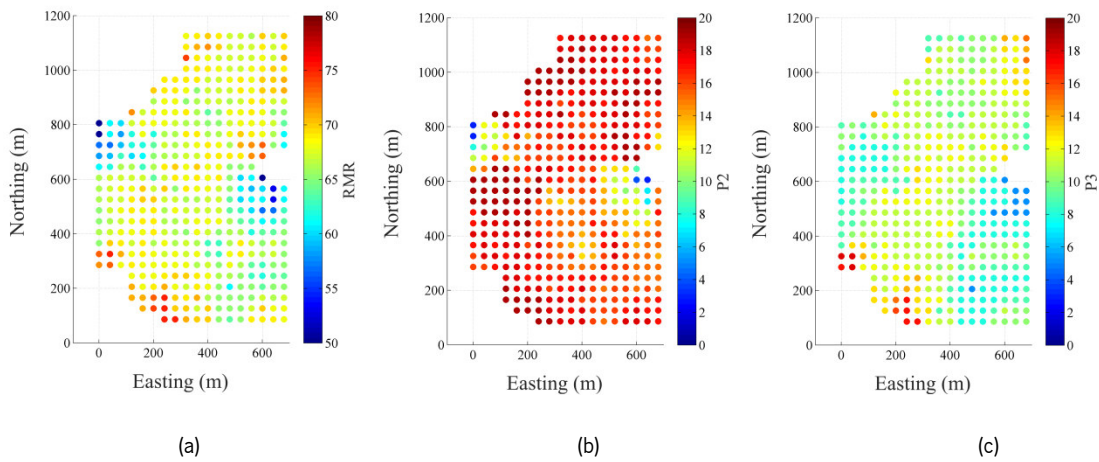


Figure 0.3 2D maps of spatial distribution at elevation 3439 m for: a) RMR original values; b) Parameter P2; and c) Parameter P3.

3.3.2. Exploratory analysis

Table 0.2 displays the basic statistics of the RMR data resulting from the 3969 samples. For the RMR variable a variation of 0 to 100 was considered, while the information of the parameters regarding their rating was obtained from the application of the RMR system.

Table 0.2 Basic statistics on RMR ratings and original data (3969 samples).

	RMR	P1 (UCS)		P2 (RQD)		P3 (JS)		P4 (JC)		P5 (GW)	
		Rating	MPa	Rating	%	Rating	mm	Rating	JC	Rating	GW
Minimum	48.0	12.0	138.0	3.0	0.0	5.0	23.0	20.0	1.0 ¹	7.0	1.0 ²
Maximum	78.0	14.0	208.0	20.0	100.0	19.0	1923.0	20.0	1.0 ¹	10.0	2.0 ²
Mean	66.7	13.1	177.8	16.4	87.3	10.2	220.5	20.0	1.0 ¹	7.0	1.0
Standard deviation	3.8	1.0	32.5	2.7	12.3	2.0	170.2	0.0	0.0	0.3	0.1 ¹²

¹ 1.0 was the value used to represent the intermediate rating of the joint condition

² 1.0 and 2.0 represent a groundwater condition of wet and damp, respectively.

According to the data statistics, with a minimum RMR value of 48 and a maximum value of 78, the geomechanical quality of the rock mass varies from fair to good. Concerning the individual parameters ratings (see Figure 0.4), P1 varies within a short range, meaning that the UCS of the intact rock was almost constant, unlike P2 and P3 that vary in a much wider range showing very different levels of rock mass fracturing. In contrast, for all the samples, the fourth parameter (P4) was constant and equal to 20. This parameter is related to the condition of the discontinuities, so they were all classified as having slightly rough surfaces with a separation smaller than 1 mm and a highly-weathered wall rock. Accordingly, the same score (20) was assumed for all the points of the target simulation grid. Lastly, like P1, parameter P5 varies within a short range, with only two different scores, representing a groundwater condition that was mostly *wet* (7) and punctually *damp* (10). In addition, Figure 0.5 displays the histograms of Approach 2, computed by assuming the RMR individual parameters in their original scale. While P1 and P5 show a concentration of data samples in a few classes (four different UCS values for P1 and 2 for P5), parameters 2 and 3 show some more distributed histograms with a wide range of values.

Before performing an individual simulation of the four RMR parameters, it was necessary to observe the existing correlations between all of them, in order to make sure that they were not (or weakly) cross-correlated. Otherwise, the separate parameter simulation in the second approach should be replaced by a joint simulation (cosimulation), which would make the model quite more complex (Emery and Cornejo, 2010). The correlation matrix, which contains the Pearson product-moment correlation coefficients between all the parameters, is presented in Table 0.3.

Table 0.3 Correlation matrix between parameters P1, P2, P3 and P5.

	P1	P2	P3	P5
P1	1.000	-0.096	-0.164	-0.113
P2	-0.096	1.000	0.292	0.045
P3	-0.164	0.292	1.000	-0.032
P5	-0.113	0.045	-0.032	1.000

Thus, it was possible to notice that the correlation between all four parameters was rather weak. Analysing these coefficients, the only parameters with a positive correlation, although weak, were P3 and P5 comparatively to P2. Whereas, the other parameters show a slightly negative correlation between them. However, these correlations are rather weak from a statistical point of view (less than 0.3 in absolute value), so that the information on a parameter actually brings little information on the other parameters. The low correlation between P2 (RQD) and P3 (discontinuity spacing) can be explained by the good quality of the rock mass, which translates into a wide spacing of the

discontinuities and high RQD values, making the latter parameter less sensitive to closer discontinuities. As such, in Approaches 2 and 3, the simulation of the four RMR parameters can be performed separately as individual variables; cosimulation, which enhances the simulation of a set of variables in order to reproduce their cross-correlation, was not necessary here.

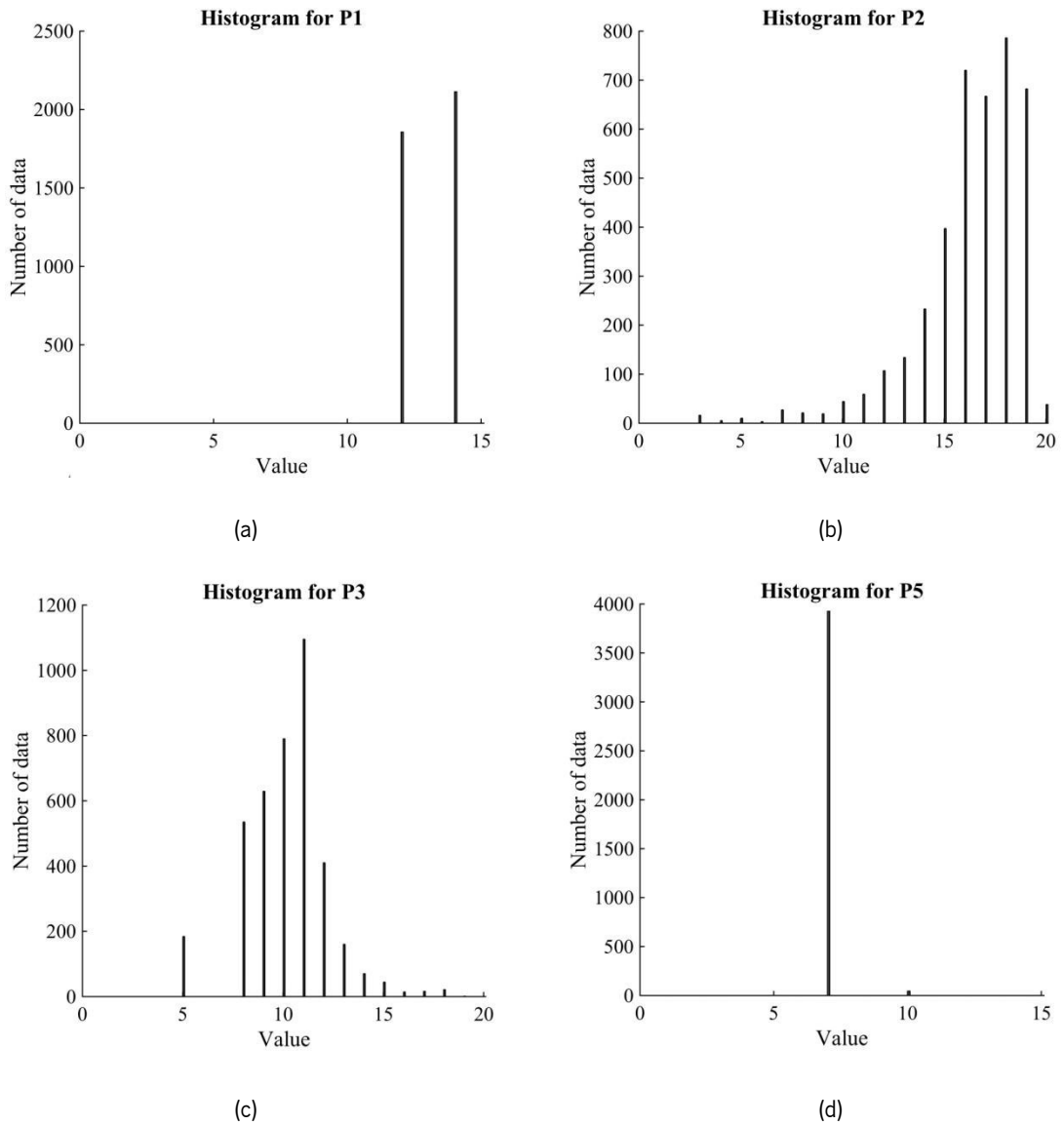


Figure 0.4 Data histograms for the RMR individual parameters in a rating scale for: a) P1; b) P2; c) P3; and d) P5.

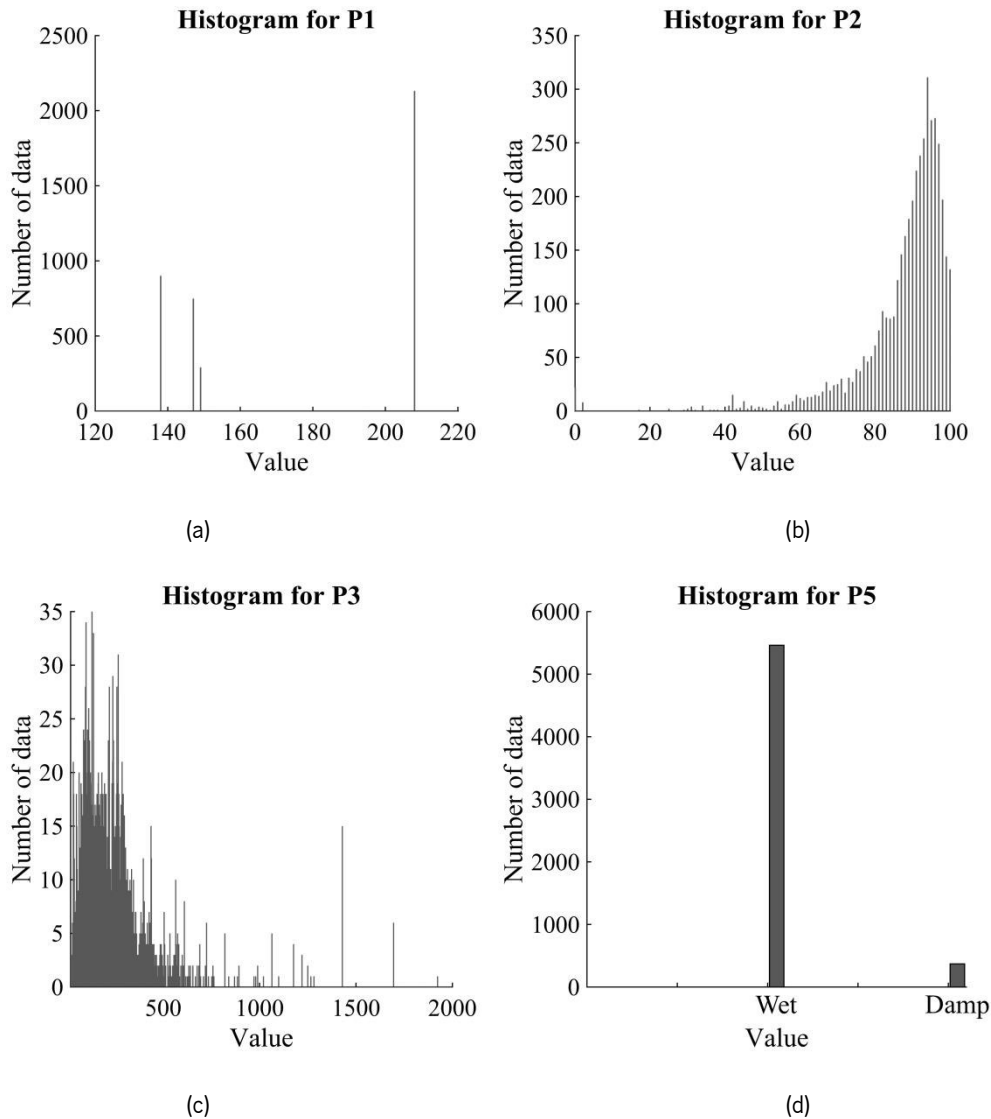


Figure 0.5 Data histograms for the RMR individual parameters in their original scale for: a) P1 (in MPa); b) P2; c) P3 (in mm); and d) P5 (wet or damp).

3.3.3. Modelling univariate distributions

As previously mentioned (3.2.2), for the first approach, the data should be transformed into normal scores. Since the sampling design is regular, there is no need for declustering, *i.e.* all the data are assigned the same weighting. The function that relates the original RMR values and the associated Gaussian values (anamorphosis function) can be calculated empirically with the available data. This function was modelled by using a piecewise linear interpolation between the empirical points and exponential functions for tail extrapolation, as explained in Emery and Lantuéjoul (2006). The parameters of these exponential functions are chosen in order to fit a model as continuous as possible. The fit of the anamorphosis should be done as good as possible in order to improve the simulation result of the variables and, in this case, as showed in Figure 0.6 the continuity and shape

of the function obtained from the original data was deemed satisfactory. The minimum and maximum values for RMR are set to 45 and 80, respectively.

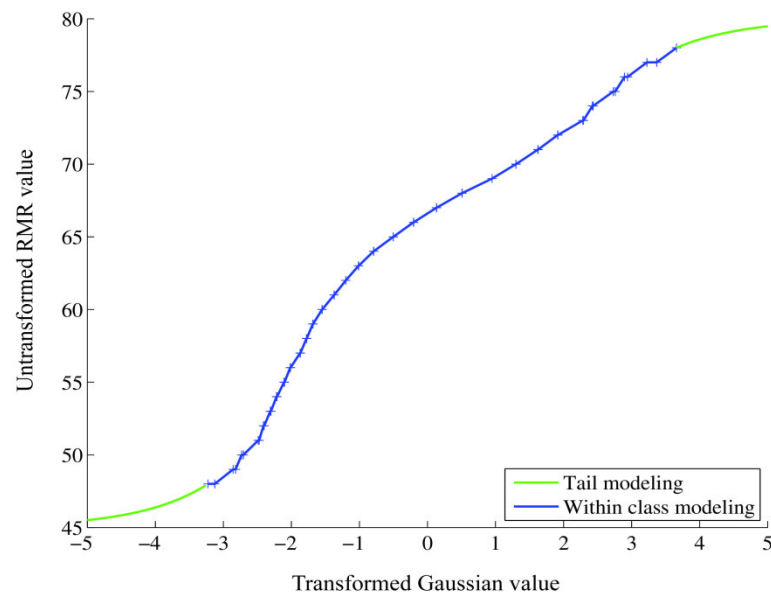


Figure 0.6 Anamorphosis function used for Approach 1. The ordinate indicates the RMR value and the abscissa the associated Gaussian value.

Similarly, for approach two all the four parameters values had to be transformed into normal scores and, consequently, an anamorphosis function was computed to all. The maximum and minimal values of each parameter are presented in Table 0.2.

Regarding Approach 3, the data related to parameter P1 only present two different scores, 12 and 14, with relative proportions of 0.468 and 0.532, respectively. This distribution can therefore be modelled by truncating a standard normal distribution using a single truncation threshold set to $G^{-1}(0.468) = -0.08$, where G is the standard normal cumulative distribution function. In other words, the probability for a standard Gaussian random variable to be less than -0.0803 is 0.468, and the probability to be more than -0.0803 is 0.532, coinciding with the proportions of the two scores of parameter P1. Likewise, the data of Parameter P5 only assume two different scores, 7 and 10, with relative proportions of 0.989 and 0.011, respectively. Again, the model uses a single truncation threshold, here equal to $G^{-1}(0.989) = 2.2904$.

In a different way, the data of Parameters P2 and P3 assume almost every score. As result, a larger number of truncation thresholds had to be defined, as shown in Table 0.4 and

Table 0.5, respectively. These truncation thresholds are such that the proportion of data with a given score coincides with the probability for a standard Gaussian random variable to be between the lower and upper thresholds associated with this score.

Table 0.4 Calculated proportions for P2 data with the corresponding Gaussian thresholds.

Category	Cumulative proportion	Lower threshold	Upper threshold
1	0.0030	$-\infty$	-2.7478
2	0.0032	-2.7478	-2.7266
3	0.0033	-2.7266	-2.7164
4	0.0043	-2.7164	-2.6276
5	0.0073	-2.6276	-2.4422
6	0.0080	-2.4422	-2.4089
7	0.0150	-2.4089	-2.1701
8	0.0200	-2.1701	-2.0537
9	0.0250	-2.0537	-1.9600
10	0.0360	-1.9600	-1.7991
11	0.0510	-1.7991	-1.6352
12	0.0780	-1.6352	-1.4187
13	0.1120	-1.4187	-1.2160
14	0.1710	-1.2160	-0.9502
15	0.2710	-0.9502	-0.6098
16	0.4520	-0.6098	-0.1206
17	0.6200	-0.1206	0.3055
18	0.8180	0.3055	0.9078
19	0.9900	0.9078	2.3263
20	1.0000	2.3263	$+\infty$

Table 0.5 Calculated proportions for P3 with the corresponding Gaussian thresholds.

Category	Cumulative proportion	Lower threshold	Upper threshold
1	0	$-\infty$	$-\infty$
2	0	$-\infty$	$-\infty$
3	0	$-\infty$	$-\infty$
4	0	$-\infty$	$-\infty$
5	0.0500	$-\infty$	-1.6449
6	0.0501	-1.6449	-1.6439
7	0.0502	-1.6439	-1.6429
8	0.1842	-1.6429	-0.8995
9	0.3432	-0.8995	-0.4037
10	0.5422	-0.4037	0.1060
11	0.8152	0.1060	0.8972
12	0.9182	0.8972	1.3931
13	0.9582	1.3931	1.7302
14	0.9762	1.7302	1.9809
15	0.9872	1.9809	2.2322
16	0.9912	2.2322	2.3739
17	0.9952	2.3739	2.5899
18	0.9998	2.5899	3.5401
19	1.0000	3.5401	$+\infty$
20	1.0000	$+\infty$	$+\infty$

3.3.4. Modelling spatial continuity

For all approaches, following the methodology explained in 3.2.2, the variograms of the Gaussian random fields to simulate had to be calculated along the main directions of anisotropy. Because of the sampling design (vertical boreholes), it was not possible to experimentally calculate variograms in inclined directions, thus calculations are restricted to the vertical direction and to the horizontal plane. Furthermore, isotropic variograms were calculated on this plane, insofar as no clear anisotropy was detected in the experimental variograms associated with different horizontal directions. For calculations, the lag distances were multiple of 20 m along the vertical, which corresponds to the data spacing along the boreholes, and of 40 m along the horizontal (borehole spacing), with a tolerance of 20 m.

The experimental variograms (hereafter denoted with the Greek letter γ) calculated were then fitted using combinations of basic nested structures (exponential, spherical, cubic and Gaussian, see Chapter 2 for details on these basic models), as follows.

- Variogram model for Approach 1 (normal score transform of RMR):

$$\gamma = 0.7 \text{ Exponential}(250 \text{ m}; 250 \text{ m}) + 0.1 \text{ Gaussian}(2000 \text{ m}; 300 \text{ m})$$

- Variogram models for Approach 2:

$$\text{P1: } \gamma_1 = 0.63 \text{ Exponential}(500 \text{ m}; 500 \text{ m}) + 0.63 \text{ Exponential}(\infty; 150 \text{ m}) \\ + 0.42 \text{ Gaussian}(700 \text{ m}; 700 \text{ m})$$

$$\text{P2: } \gamma_2 = 0.28 \text{ Exponential}(180 \text{ m}; 100 \text{ m}) + 0.39 \text{ Exponential}(180 \text{ m}; 180 \text{ m}) \\ + 0.46 \text{ Gaussian}(400 \text{ m}; 400 \text{ m})$$

$$\text{P3: } \gamma_3 = 0.19 \text{ Exponential}(180 \text{ m}; 100 \text{ m}) + 0.37 \text{ Exponential}(180 \text{ m}; 180 \text{ m}) \\ + 0.55 \text{ Gaussian}(400 \text{ m}; 400 \text{ m})$$

$$\text{P5: } \gamma_5 = 0.41 \text{ Gaussian}(500 \text{ m}; 200 \text{ m}) + 0.44 \text{ Cubic}(\infty; 250 \text{ m}) \\ + 0.14 \text{ Exponential}(300 \text{ m}; 300 \text{ m})$$

- Variogram models for Approach 3:

$$\text{P1: } \gamma_1 = 1.0 \text{ Cubic}(700 \text{ m}; 100 \text{ m})$$

$$\text{P2: } \gamma_2 = 1.0 \text{ Exponential}(400 \text{ m}; 350 \text{ m})$$

$$\text{P3: } \gamma_3 = 0.8 \text{ Exponential}(300 \text{ m}; 250 \text{ m}) + 0.2 \text{ Gaussian}(300 \text{ m}; 350 \text{ m})$$

$$\text{P5: } \gamma_5 = 0.3 \text{ Exponential}(500 \text{ m}; 200 \text{ m}) + 0.7 \text{ Gaussian}(700 \text{ m}; 150 \text{ m})$$

In the above equations, the coefficient preceding a basic nested structure indicates the sill of this structure (contribution to the total variance), while the distances written between brackets represent the correlation ranges of the structure along the horizontal plane and the vertical direction, respectively. The experimental and theoretical variograms for Approaches 1, 2 and 3 are shown in Figure 0.7, 3.8 and 3.9, respectively.

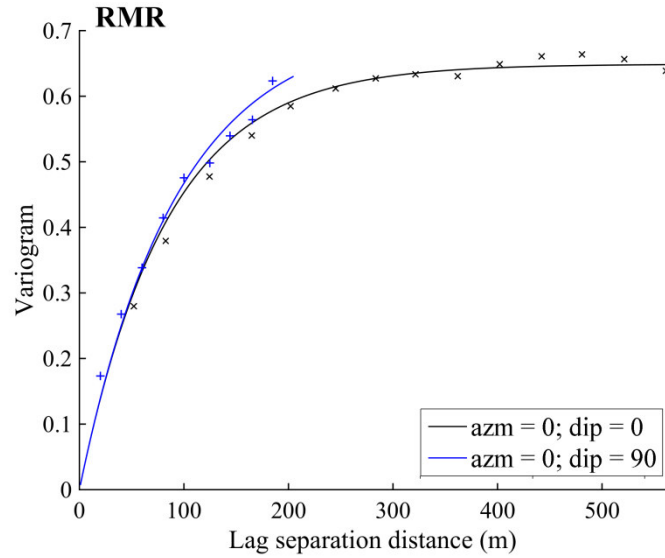


Figure 0.7 Experimental (crosses) and theoretical (solid lines) variograms for RMR (Approach 1) along the main anisotropy directions, horizontal plane (black) and vertical direction (blue).

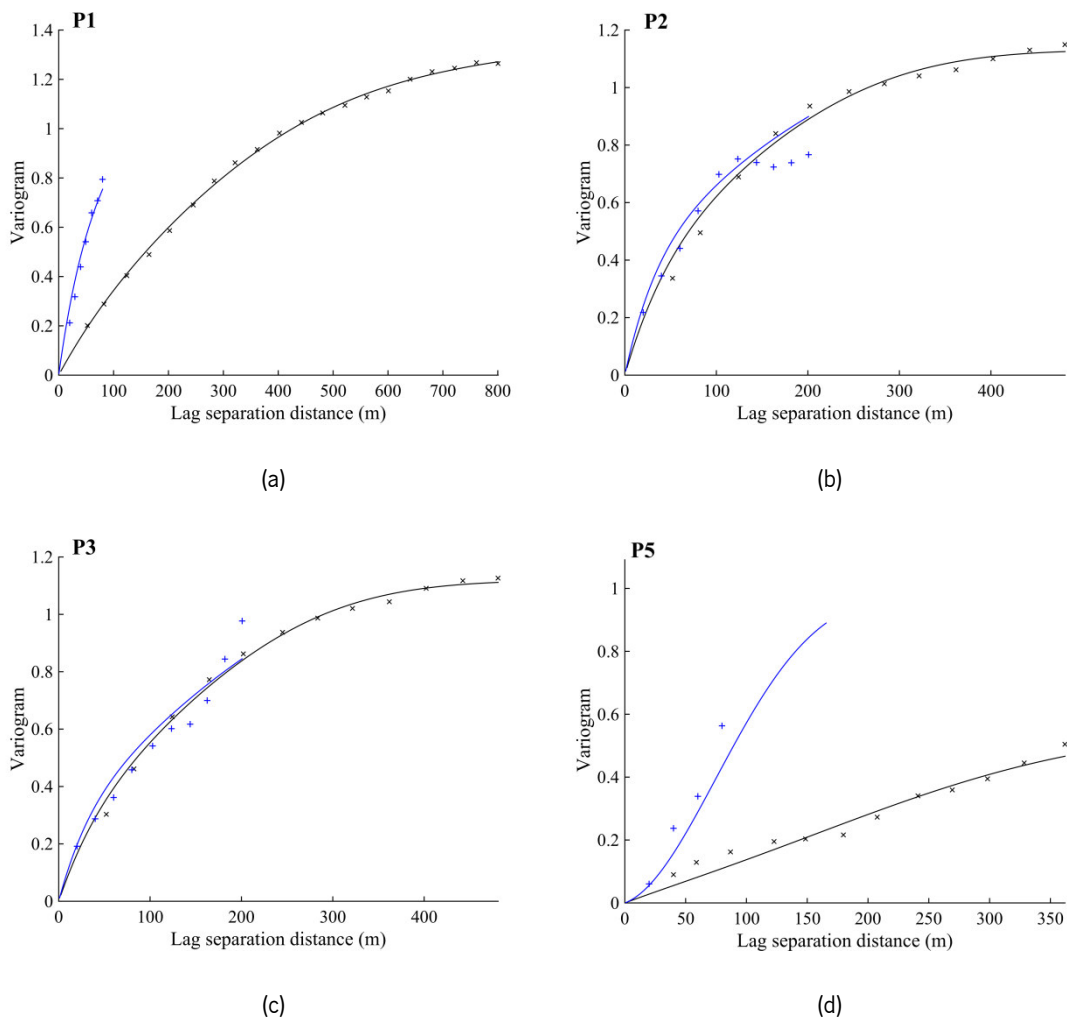


Figure 0.8 Experimental (crosses) and theoretical (solid lines) variograms for Approach 2 along the main anisotropy directions, horizontal plane (black) and vertical direction (blue) for parameters: a) P1; b) P2; c) P3; and d) P5.

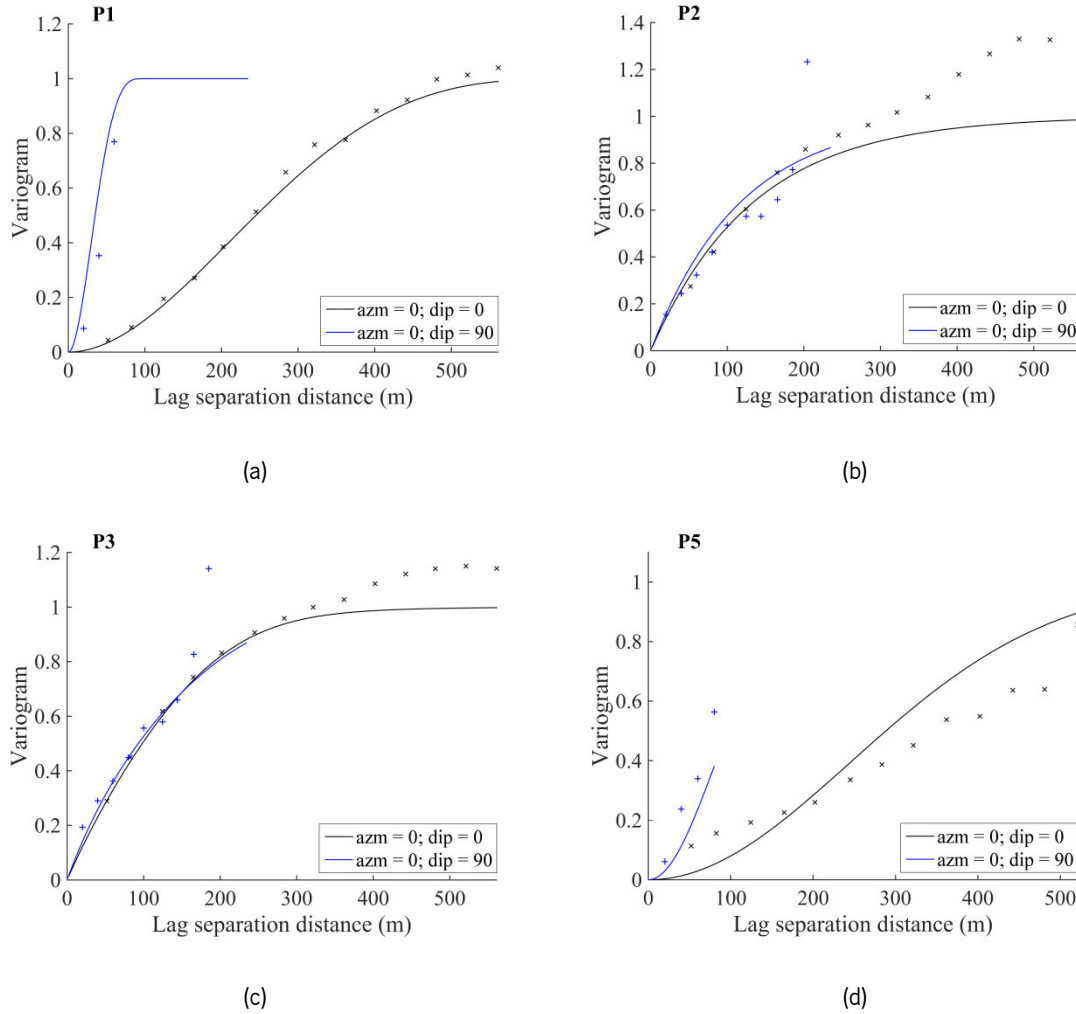


Figure 0.9 Experimental (crosses) and theoretical (solid lines) variograms for Approach 3 along the main anisotropy directions, horizontal plane (black) and vertical direction (blue) for parameters: a) P1; b) P2; c) P3; and d) P5.

Through the variogram analysis, it can be noticed that none of the variograms exhibit a nugget effect (discontinuity near the origin), indicating that the RMR and its underlying parameters were continuous in space. Even more, the variogram models for P1 and P5 have a smooth behaviour near the origin and a large correlation range along the horizontal direction (500 m), which indicates that the regions where these parameters were constant had smooth boundaries and a large spatial extent or spatial connectivity (recall that both P1 and P5 only assume two different scores). The same effect happens either in their original scale or in their rating scale. Moreover, it is important to highlight the existence of anisotropy for parameters P1, P3 and P5 where the range of both curves to reach the same sill (variance equal to 1) is different. As expected, the variograms of Approach 2 showed a similar behaviour as for Approach 3. Contrarily, for Approach 1, the RMR variogram shows small differences in the range for the horizontal plan and the vertical direction, similarly to parameter P2 of Approach 3.

3.3.5. Conditional simulation results

Once the model parameters were specified (anamorphosis function in Approach 1 and 2, truncation thresholds in Approach 3, and variograms of the underlying Gaussian random fields in all approaches), conditional realisations of the RMR and of its parameters can be constructed. For these simulations, adaptations of previously published computer programs were used for simulating Gaussian and truncated Gaussian random fields (Emery and Lantuéjoul, 2006; Emery, 2007). The number of realisations was set to one hundred, so that the post-processing outputs (average and conditional probabilities) could be calculated with a reasonable approximation. In both cases, the turning bands algorithm was applied with 1500 turning lines to generate the Gaussian random fields, and simple kriging was used to condition the realisations to the borehole data. For Approach 3, the Gibbs sampler was stopped after one hundred iterations and the simulated Gaussian random fields were truncated, based on the thresholds indicated in Tables 3.5 and 3.6, resulting in realisations of parameters P1 to P5 that were subsequently summed to obtain realisations of the RMR values (more details on algorithms are given in Chapter 2).

In order to perform simulation, it was necessary to define the target locations that could be a regular two-dimension or three-dimension grid. For ease of display, the locations targeted for simulation correspond to a regular two-dimensional grid placed at elevation 3560 m, with a mesh of 5 m × 5 m and a total of 160 nodes along the east direction (*X*-axis) and 240 nodes along the north direction (*Y*-axis).

As result of the geostatistical simulation, one obtained 22500 target locations with RMR simulated values for the three approaches (see Table 0.6) and it was possible to compare the minimum and maximum values resulting from the average of the realisations with the original data. The difference between them was quite important namely regarding the minimum values that are higher in the simulation (realisation average) when comparing with the original values; this can be explained by the smoothing effect provoked by the averaging of the realisations.

Table 0.6 Statistics on the average of 100 conditional realisations of RMR obtained with Approaches 1, 2 and 3.

	Approach 1	Approach 2	Approach 3
Number of target locations	22500	22500	22500
Minimum RMR value	54.8	54.0	57.5
Maximum RMR value	76.2	73.0	74.1
Mean RMR value	66.6	65.8	66.8
Variance of RMR values	4.7	4.8	3.5

The average of the RMR realisations was used to map the expected RMR over the region of interest resulting from Approaches 1, 2 and 3 (see Figure 0.10). Furthermore, in order to demonstrate the spatial variability existing in this rock mass, the first realisation was mapped as an example (see Figure 0.11). In these figures, the neighbouring data values were superimposed on the maps to highlight the effect of conditioning the RMR realisations to the borehole data: when a target grid node coincides with a data location, the simulated RMR value exactly matches the conditioning data value.

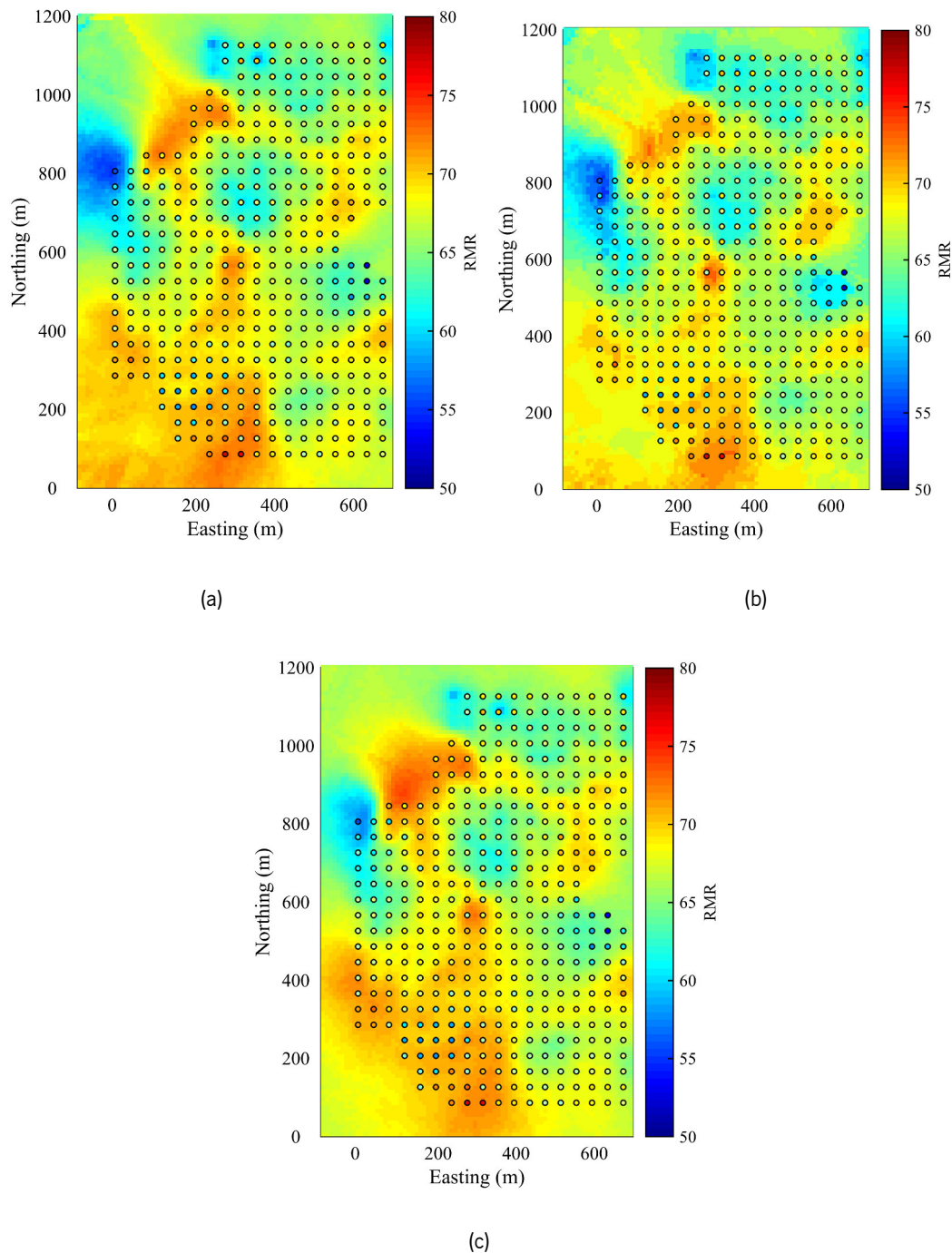


Figure 0.10 Maps of RMR at elevation 3560 m of the average of 100 realisations obtained with: a) Approach 1; b) Approach 2; and c) Approach 3.

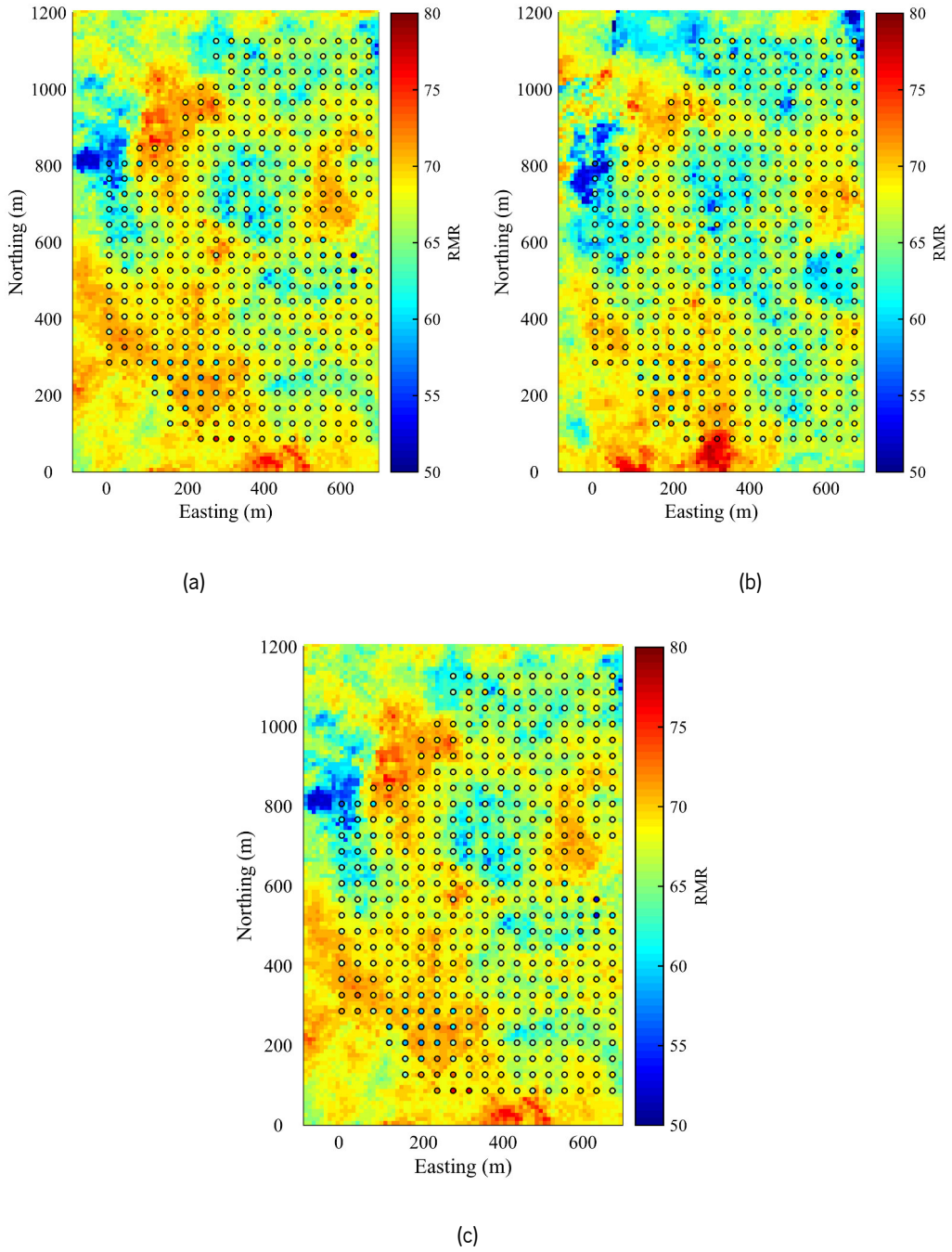


Figure 0.11 Maps of RMR at elevation 3560 m for realisation #1 obtained with: a) Approach 1; b) Approach 2; and c) Approach 3.

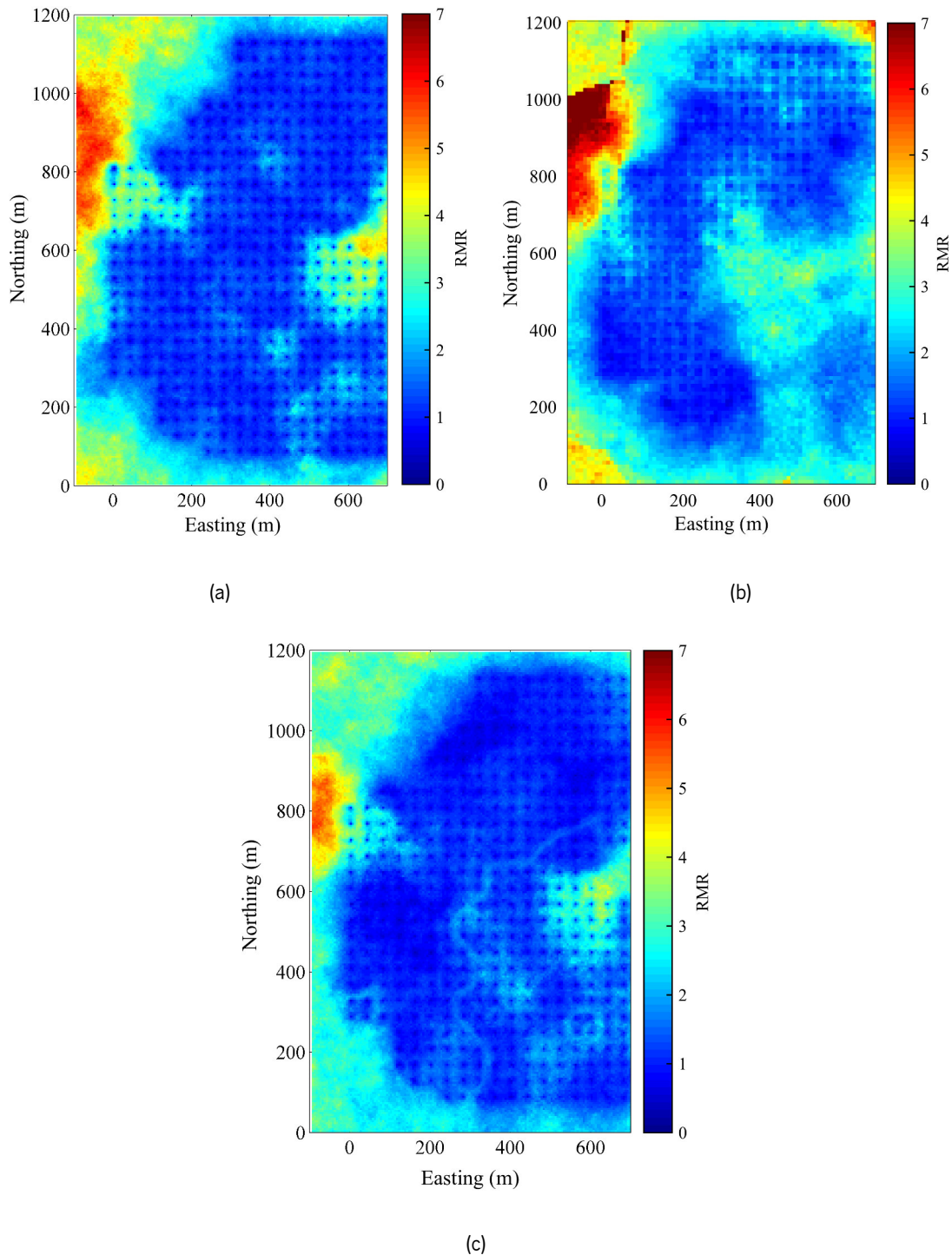


Figure 0.12 Maps of RMR at elevation 3560 m of the standard deviation of 100 realisations obtained with: a) Approach 1; b) Approach 2; and c) Approach 3.

By the figures analysis, it is possible to state, at this stage, that Approaches 1 and 2 do not show a significant difference, not only in the RMR maps but also in the RMR resulting statistics.

The visual comparison suggests that the two approaches (Approach 1 and 3) produce similar results for RMR unlike Approach 2. It is worth noticing that the map of the realisation average tends to smudge the contrasts and that spatial heterogeneities appear as much more faded. The analysis of

the individual realisations is therefore important to visualize the real heterogeneity that it is expected to be observed in the field, whereas the average of the realisations shows an overall trend, which was much smoother.

3.3.6. Post-processing simulations

More outputs can be represented, such as the probability that the RMR exceeds or falls short of a predefined value, which can be estimated by the frequency of threshold exceedance or non-exceedance observed over the realisations. This representation is of great interest if one wants to identify regions on the rock mass where very high or low geomechanical properties could be present, and with which probability (0 to 1). As an example, for the three approaches, Figure 0.13 shows the map with the probability that the actual (unknown) RMR exceeds a threshold of 65. Comparing all three probability maps, it was possible to perceive minor differences between Approach 1 and 3, however, comparing the last two maps with the one obtained from Approach 2 the same cannot be stated. In detail, in the upper central zone, Approach 2 showed a higher probability to find an RMR value above 65.

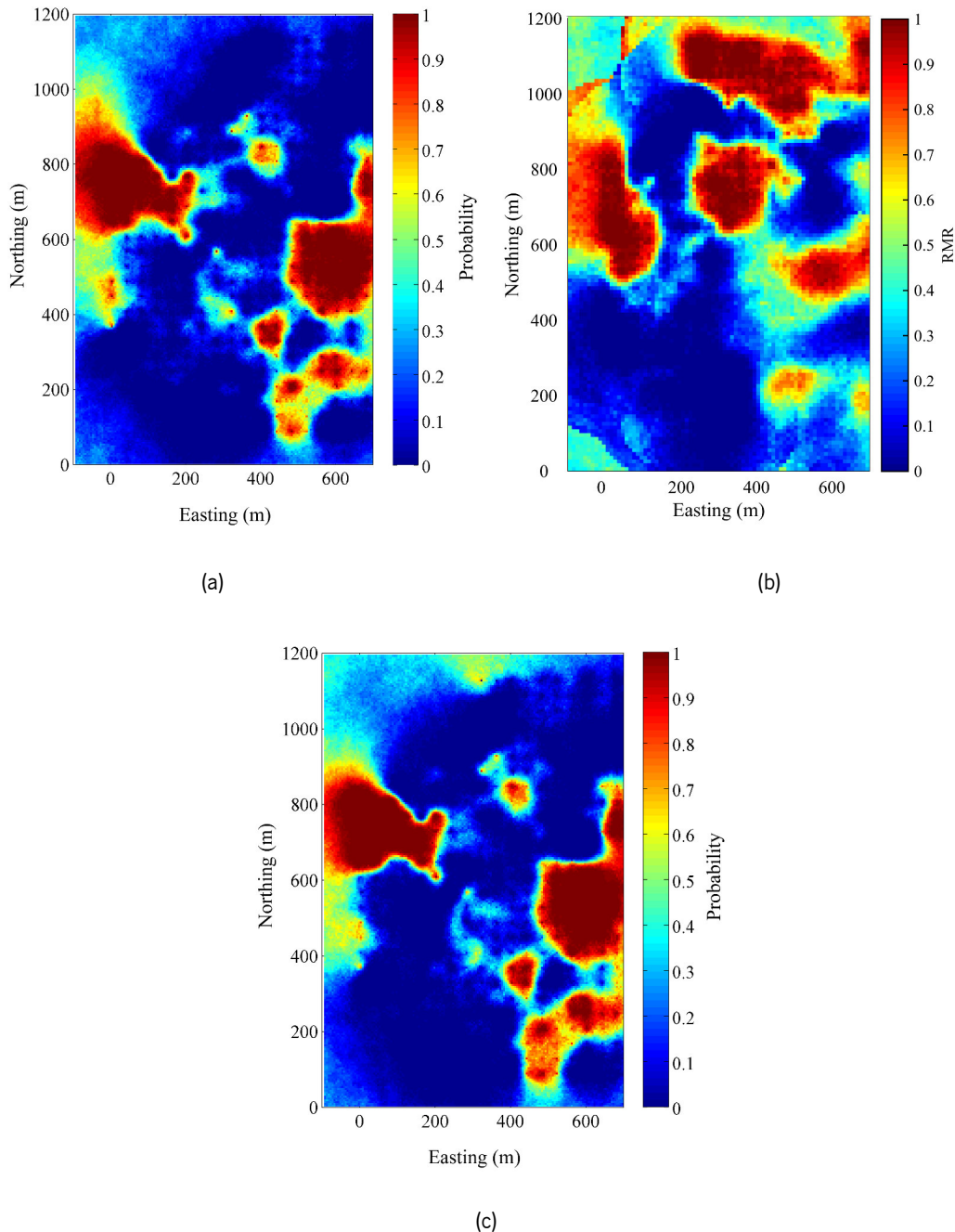


Figure 0.13 Maps of probability (between 0 and 1) that the RMR is less than a threshold of 65 at elevation 3560 m, obtained with: a) Approach 1; b) Approach 2; and c) Approach 3.

Figure 0.14 and Figure 0.15 shows the maps of parameters P2 and P3 for the first realisation and for the average of 100 realisations obtained with Approach 3, which helps to visualise the spatial distribution and variability of these discontinuity parameters. Comparing realisation #1 with the average of 100 realisations for both parameters, the pattern in lower or higher score were similar, however, as already referred, the average mapping exhibits smoother values. The individual analysis of the discontinuity parameters can, by itself, result in a powerful tool in geotechnical works to understand the regions where the rock mass can be more or less fractured. Attentively, comparing the P2 and P3 maps (average and first realisation), it was possible to identify a small region in the

western part of the grid, with low values of RQD (P2) and intermediate values of the discontinuity spacing (P3). These incoherent values were truly present in the borehole data used for conditioning the realisations, therefore they were not a problem of the proposed simulation approach, but rather a problem of the input data. This could be explained by human errors in the measurement of parameter P2.

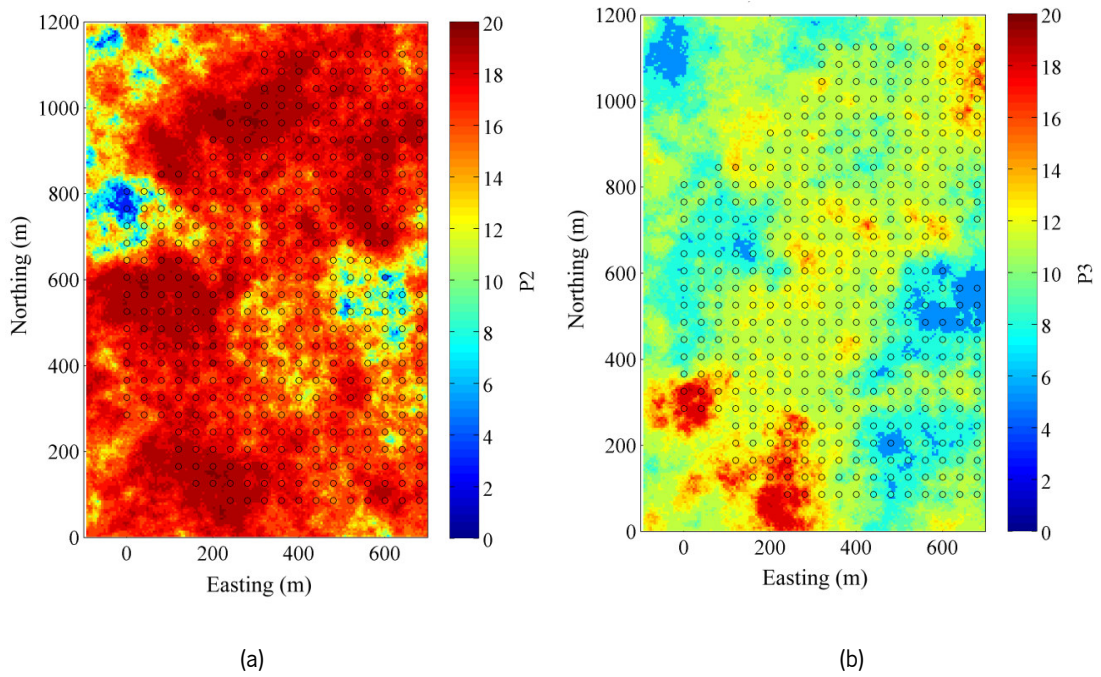


Figure 0.14 Maps of discontinuity parameters at elevation 3560 m for realisation #1 of a) P2; and b) P3.

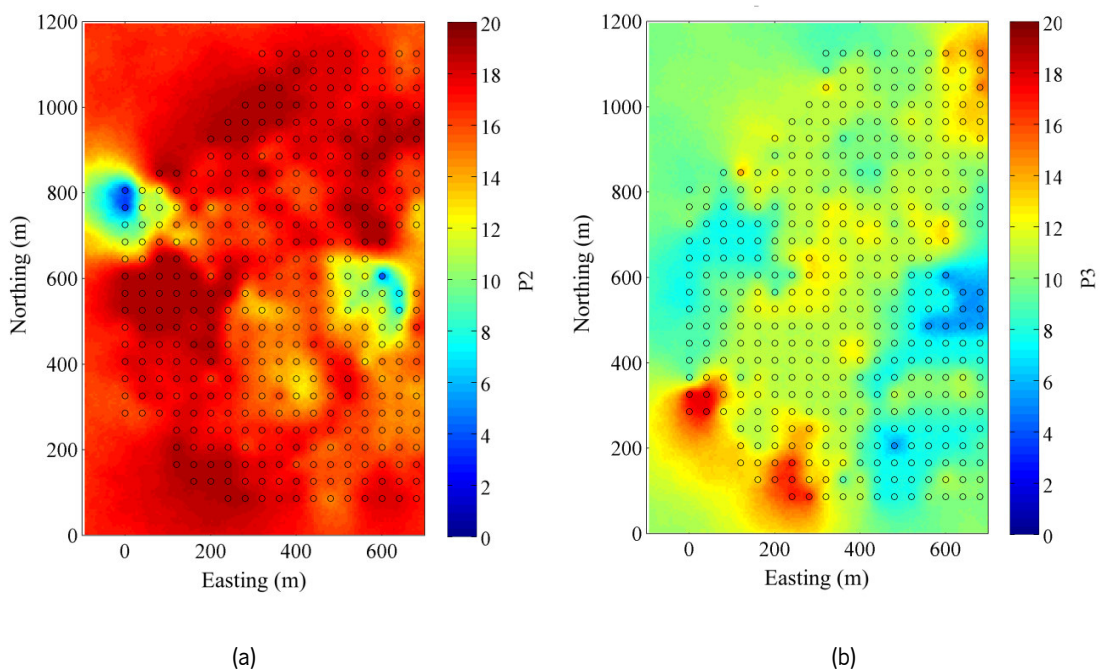


Figure 0.15 Maps of discontinuity parameters at elevation 3560 m for the average of 100 realisations of a) P2; and b) P3.

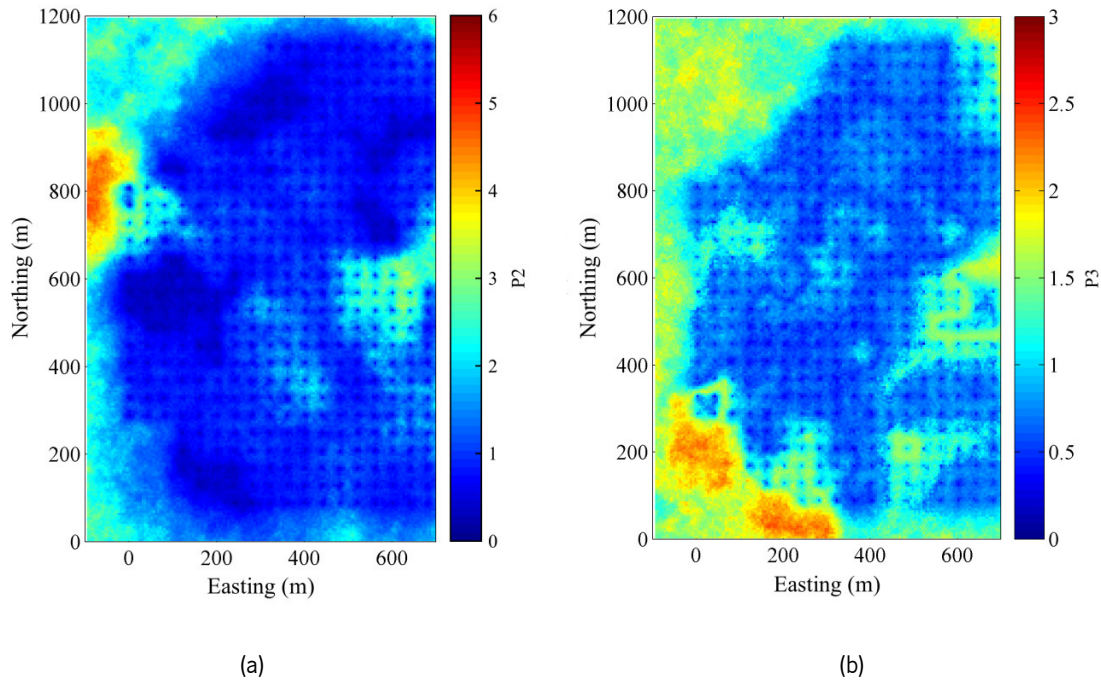


Figure 0.16 Maps of discontinuity parameters at elevation 3560 m for the standard deviation of 100 realisations of: a) P2; and b) P3.

Maps of other geomechanical parameters can be obtained based on the realisations of RMR. In this work, considering that the RMR values are higher than 50, the distribution maps of the deformation modulus (E_m) were developed with the Bieniaskwi (1976) empirical formula presented in Equation (0.1).

$$E_m = 2 \times RMR - 100 \quad (0.1)$$

This formula uses the simulated values of RMR to obtain the E_m values, in GPa at the same locations. Alike the maps computed for the RMR values, Figure 0.17 shows the average of E_m obtained from the 100 conditional realisations of Approach 1, as well as the first realisation, as an example. With these maps, it is possible to distinguish zones where the rock mass is significantly stiffer and zones with lower rigidity.

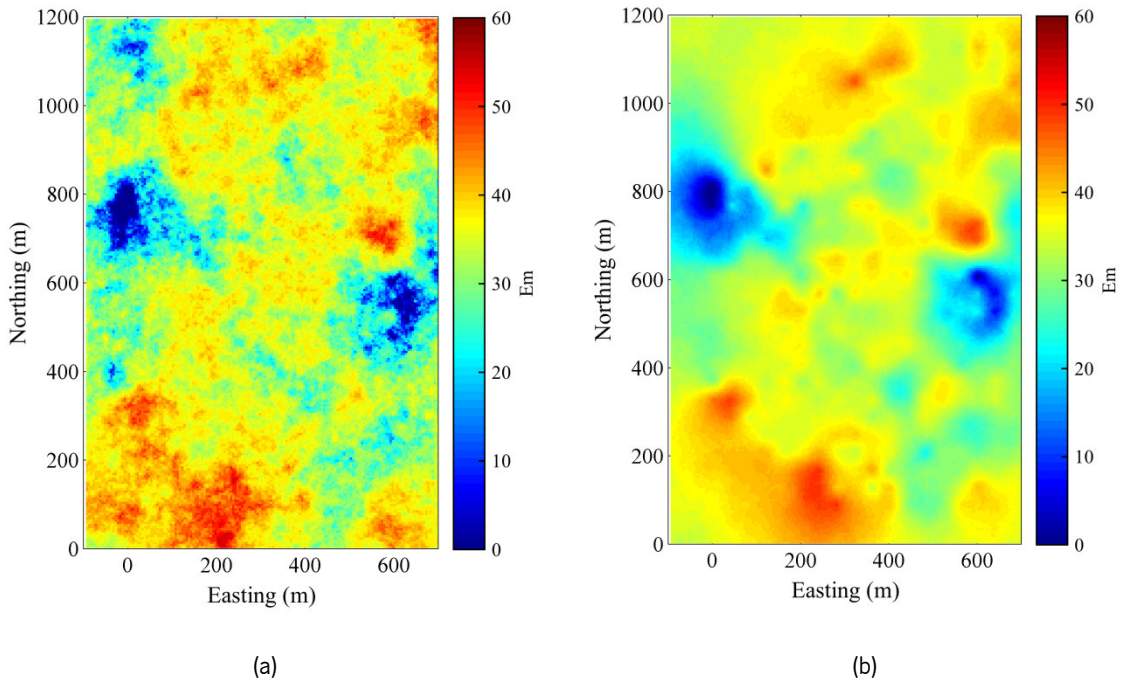


Figure 0.17 Maps of deformation modulus (in GPa) at elevation 3560 m obtained with Approach 1, for: a) realisation #1; and b) average of 100 realisations.

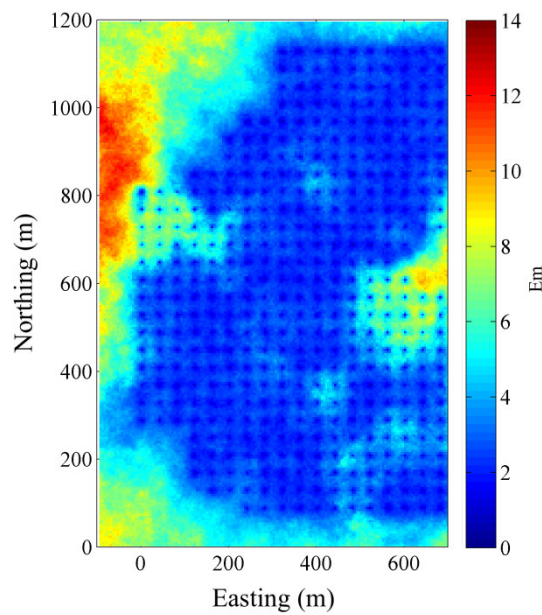


Figure 0.18 Standard deviation of 100 realisations map of deformation modulus (in GPa) at elevation 3560 m obtained with Approach 1.

Finally, to visualise the uncertainty in the simulated values, the standard deviation (or any other uncertainty measure, such as the coefficient of variation or the limits for a given level of confidence) can be calculated at each target node over the 100 realisations and mapped throughout the grid of interest. As an example, the standard deviations of RMR obtained with both approaches, of

parameters P2 and P3 obtained in the truncated Gaussian model (Approach 3) and of the deformation modulus (E_m) obtained in the Gaussian model (Approach 1) were mapped in Figure 0.12, Figure 0.16 and Figure 0.18, respectively. Such maps indicate how much the true unknown values are likely to deviate from their expected values (averages of the realisations) at each target grid node, therefore quantify the local uncertainty concerning the true values. The mapped standard deviations depend on the number and location of the surrounding borehole data (they increase in under-sampled areas and peripheral areas without data) and, to a lesser extent, on the parameter values: for P3, whose distribution was positively skewed, the standard deviation tends to increase in high-valued areas, while the reverse happens for RMR, P2 and E_m whose distributions were negatively skewed, a phenomenon known as proportional effect (Manchuk et al., 2009) or regressive effect (David, 1988). Furthermore, the maximum standard deviation for RMR occurs in the north-western side of the grid of interest and was about 7 (see Figure 0.12), which represents a small deviation for a variable that varies from 0 to 100. This suggests a relatively low uncertainty in the true RMR values at unsampled locations.

The standard deviations mapped in Figure 0.12, Figure 0.16 and Figure 0.18 should not be confused with the ones presented in Table 2: the former measures the variability across the realisations at a given location, therefore depend on the location under consideration, whereas the latter measures the variability of a data set across the region of interest, without distinguishing any specific location.

3.3.7. Split-sample validation

To validate the three proposed approaches, the original data set was randomly divided into two subsets, each containing one half of the data. Thereby, the first subset (training subset) was then used to simulate the RMR at the locations of the data belonging to the second subset (validation subset). Two types of validation were performed aiming to assess the prediction capability and the modelling of uncertainty via the simulation approaches.

First, in order to validate the prediction capability of all the approaches, the expected RMR, calculated as the average of the simulated RMR values, was compared with the real values at the locations of the validation subset (see Figure 0.19). To analyse the results, a linear regression and coefficient of determination (R^2) between expected and real RMR values were calculated. For all the approaches, the resulting points of the scatter plot were distributed close to the diagonal line and the coefficient of determination shows a high value (0.72, 0.73 and 0.74, respectively, for approaches 1, 2 and 3). This indicates that the simulations allow an accurate prediction of RMR,

with small error fluctuation and no conditional bias (Chilès and Delfiner, 2012). Regarding the accuracy, it can be confirmed by calculating the Root Mean Squared Error (RMSE) presented in Equation (0.2).

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}}, \tag{0.2}$$

where N denotes the number of data in the validation subset, y_i the true value and \hat{y}_i the expected value (average of the realisations). The RMSE values were 1.74 for Approach 1, 1.98 for Approach 2 and 1.81 for Approach 3. Since RMR varies from 0 to 100, an error less than 2 was almost residual.

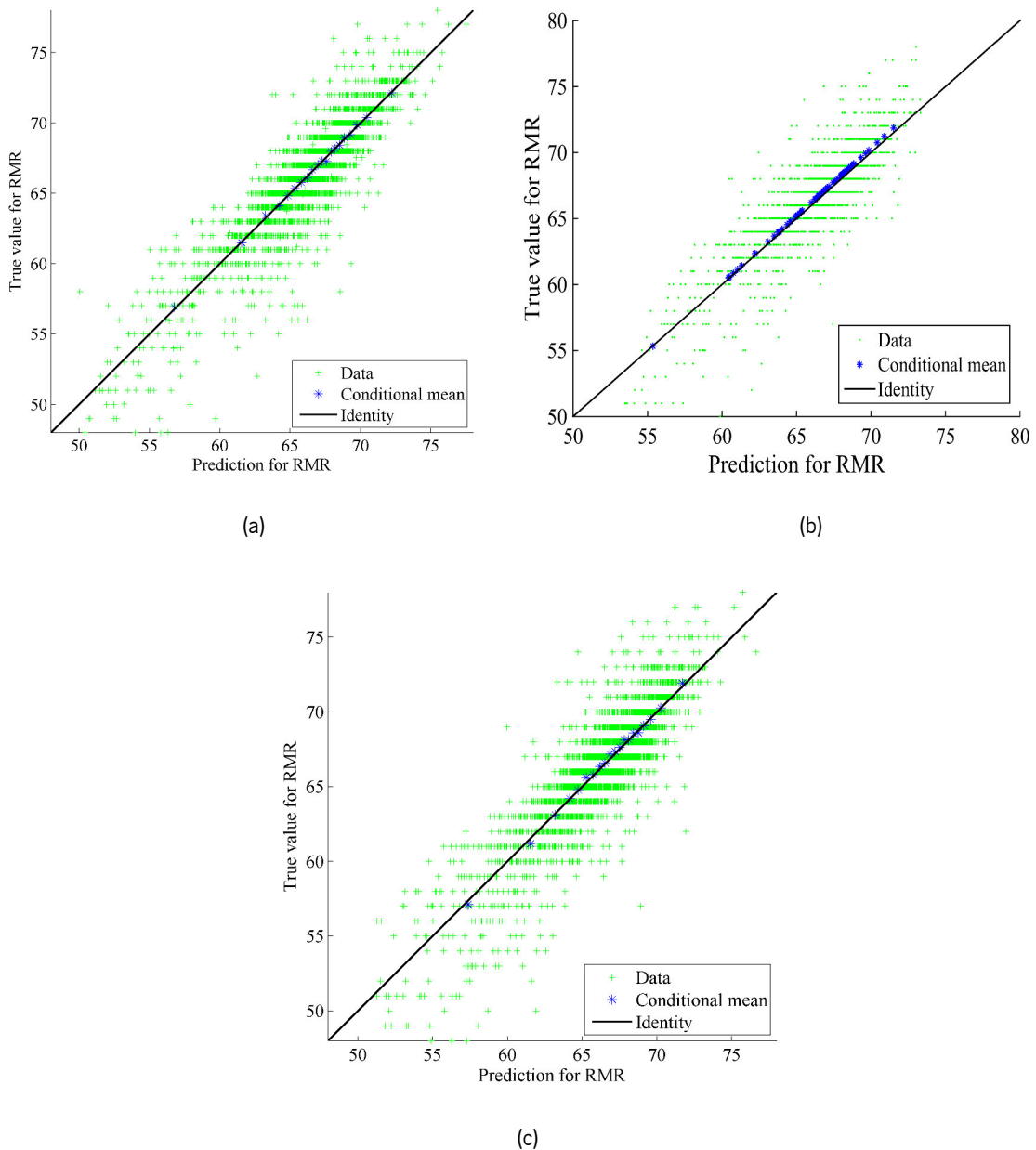


Figure 0.19 Scatter plots between true and expected RMR values for: a) Approach 1; b) Approach 2; and c) Approach 3.

In order to validate the capability of modelling uncertainty, accuracy plots (Goovaerts, 2001) were constructed. In these plots, one considers a given probability p and, based on the obtained realisations, one can define at each target location an interval with such a probability (the interval bounds are the quantiles $1-p/2$ and $1+p/2$ of the set of simulated values). Subsequently, the location was assigned a value of 1 if the true RMR belongs to the interval and 0 otherwise. It was expected that, on average over all the locations of the validation subset, the proportion of 1 should be close to the probability p under consideration. This procedure has been applied with p varying from 0 to 1 (see Figure 0.20). In Approaches 1 and 3, the observed proportion was close to the theoretical probability (points close to the diagonal line, with a slightly better coincidence for Approach 3), indicating that the realisations accurately assess the uncertainty in the actual RMR values. In contrast, the same cannot be stated regarding Approach 2, since this model showed a lower accuracy justified by the existing separation between the theoretical probability and the observed proportion (see Figure 0.20b). In this case, the model tends to simulate higher values in comparison with the expected ones.

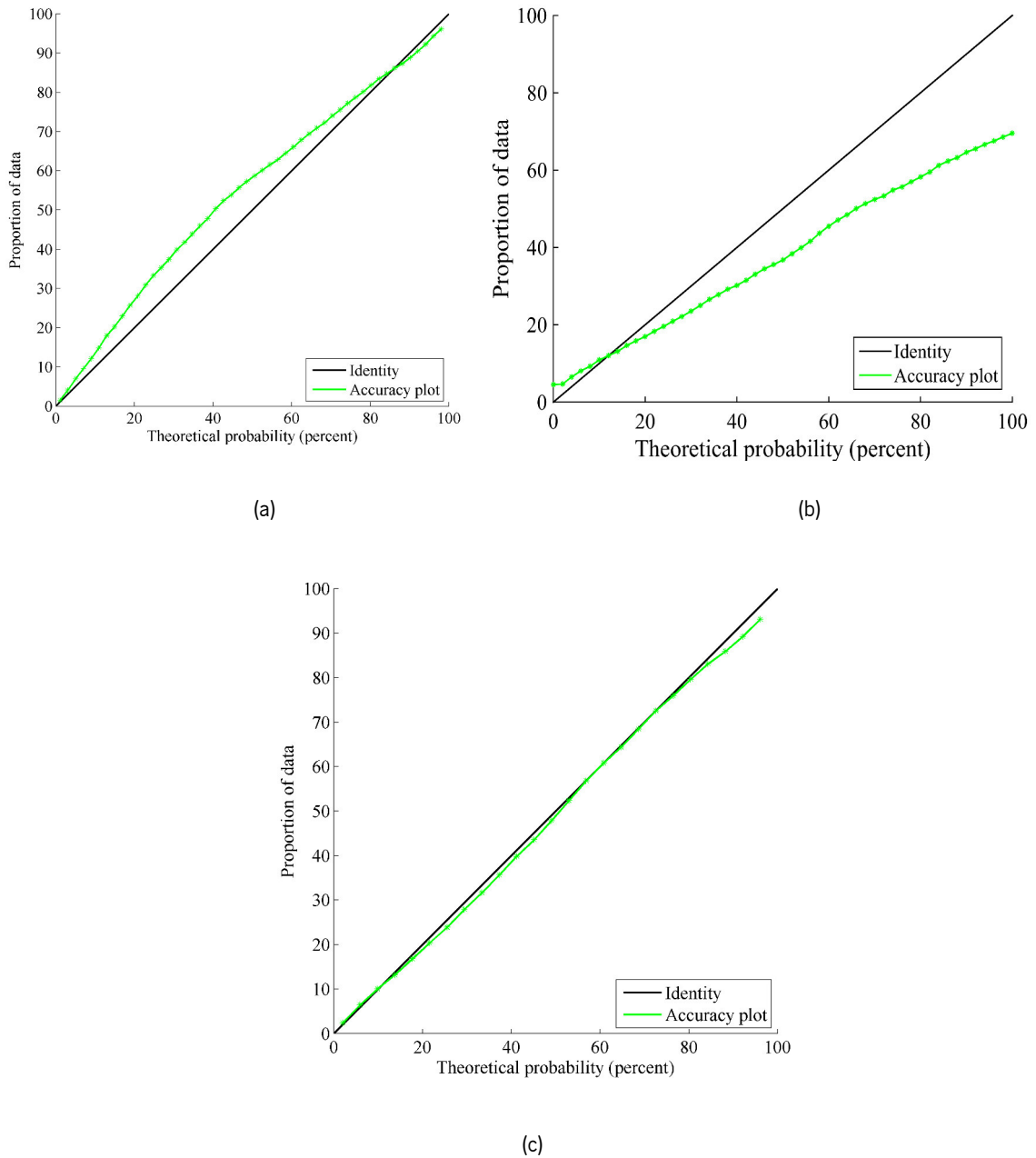


Figure 0.20 Accuracy plots showing the uncertainty modelling of the simulations for: a) Approach 1; b) Approach 2; and c) Approach 3.

3.3.8. Discussion of the results

All the simulation approaches give an insight into two characteristics of the geomechanical parameters of interest: (1) their heterogeneity at all spatial scales, especially at short scale, which can be assessed on each individual realisation; and (2) the uncertainty in the true values at unsampled locations. The latter can be assessed by comparing a set of realisations at the same location or jointly over several locations, e.g., by calculating the standard deviation of the realisations at each location, which measures how much the true unknown values may deviate from their expected values (average of the realisations).

There are several differences between these three proposed approaches, mainly regarding pre and post processing. While Approach 1 allows directly simulating the RMR as a continuous variable, the quantity and detail of the geomechanical information is more limited when compared with Approaches 2 and 3 that allow the simulation of the parameters underlying the definition of RMR. Also, the difference in modelling and computational efforts for these approaches is substantial, Approach 1 being a faster and simpler alternative than Approaches 2 and 3. Regarding the results, the observed differences are generally not significant when mapping one realisation or the average of a set of realisations, but differences are perceptible when mapping the standard deviations of the realisations: Approach 3 yields a lower standard deviation (reflecting less uncertainty) than Approach 1 in the peripheral zones. All things considered, the choice of the best approach should be made based on the resources and needs of the practitioner (degree of required geotechnical detail, understanding of the models, software and time availability). Also, as mentioned previously, despite the computational time, quality and detail of the geomechanical information, Approaches 1 and 2 are almost similar, which emphasises the importance of comparing Approaches 1 and 3 since the simulation process are distinct.

3.4. CONCLUSIONS

The use of geostatistical techniques is very common in the mining and petroleum industry, being its use in geotechnical problems still scarce; however recent works have been able to perform the geomechanical parameters estimation using Kriging techniques. Although, the use of geostatistical simulation in this field is still rare. As such, in this Chapter the Gaussian and truncated Gaussian models were the ones chosen to simulate the empirical system, RMR. This is a commonly used system in rock mechanics since it allows obtaining a rock mass quality classification according to five classes (from very poor to very good). The RMR is presented in a continuous scale varying from 0 to 100 and is obtained by summing five basic parameters, most of them acquired throughout a visual inspection (P2, P3, P4 and P5) and other using the results of the laboratory tests (P1).

In detail, to perform the geostatistical simulation of the RMR, three different approaches were idealised. The first considers the direct simulation of the RMR final values, viewed as a variable measured on a continuous scale. In the second approach the same assumption is made but in turn the underlying RMR parameters were simulated individually, while the third approach simulates them as variables measured in a discrete scale, *i.e.* the simulation of the RMR final ratings was

performed. At the end of Approaches 2 and 3, the simulated values were transformed and summed until the RMR final value was obtained.

As mentioned, the specific algorithm used in all three approaches was the Turning bands method, however, while Approaches 1 and 2 model the variables as Gaussian with a normal distribution, Approach 3 requires the application of the truncated Gaussian model, which simulates class-indicators.

All the developed approaches were applied to the same data set, in this case a real data set obtained from an epithermal gold deposit located in Chile. The data comprise the RMR information and the underlying parameters for a total of 3969 points.

In all the three approaches, it was required for the simulation the computation of the variogram models as well as the statistical treatment of the data to use as conditioning information in the simulation. As a result, several types of maps were obtained, the average values of 100 realisations, the individual realisations, as well as standard deviations and probabilities maps. The latter allow identifying the zones where the rock mass has a higher or lower probability of having a RMR value under a predefined threshold.

To validate the results a split-sample technique was adopted and the RMSE metric was calculated between the simulated values and the true RMR values.

Regarding the obtained outcomes, all the approaches lead to similar results in terms of RMR values and processed outputs. In fact, according to the split-sample validation technique and the calculated RMSE, the three approaches show small differences in terms of prediction accuracy and measurement of uncertainty, which makes them viable for RMR modelling. Although it is worth mentioning that Approach 2 resulted in the highest RMSE value.

The first approach presents the advantage of needing a lower computation and pre-processing time, maintaining a good predictive accuracy, which is an interesting feature from a practical implementation point of view. On the other hand, even though with higher computational costs needed for its implementation, the third approach presents a slightly higher accuracy and, jointly with the second approach, provides information on the RMR individual parameters, which can be useful for geotechnical analyses. For example, the simulation of parameters P2 and P3, which are related with fracturing of the rock mass, can provide an overview about the regions where higher permeability is expected. Furthermore, since it simulates variables measured on discrete scales, this third approach is consistent with the nature of the geomechanical parameters to be modelled, which are ratings rather than variables defined on a continuous scale. If these parameters were cross-correlated, they should be jointly simulated in order to reproduce such cross-correlations.

The three proposed approaches prove that individual realisations were much more accurate in defining the heterogeneities at short scale, while the average of the realisations tends to smooth these heterogeneities. Besides the mapping of the RMR, the realisations allow mapping the probability that the actual RMR is above or below a defined threshold, thus are helpful to quantify uncertainties and risks. In further Chapters, new applications of the conditional simulation validated steps will exist, this time taking into account the concerns and limitations referred throughout this Chapter.

Chapter 4

NUMERICAL METHODOLOGY TO CHARACTERISE HETEROGENEOUS ROCK MASSES

4.1. INTRODUCTION

This chapter addresses a new methodology developed to model the geomechanical parameters of heterogeneous rock masses. This methodology is the core of this PhD thesis, since it combines different techniques, starting with a probabilistic technique, more specifically, the geostatistical simulation of geomechanical parameters, already explored in Chapter 3. Thereafter, the adaptation of a scenario reduction methodology that uses a clustering technique to apply in geostatistical simulations in order to obtain reduced number of representative realisations. Subsequently, the outputs of the proposed methodology are imported to a finite difference software and a mechanical behaviour analysis of a case study is performed.

As mentioned in Chapter 2, geostatistics is a powerful tool that allows the simulation of geomechanical parameters using some preliminary data. In these conditional simulations, the interaction between the sample data is considered, *i.e.* the sample structure (variogram) is honoured and their correlation maintained. As result, several realisations of the random fields representing the variables of interest are obtained, in this case the geomechanical parameters or the empirical classification system used to classify the quality of the rock mass.

In order to avoid averaging the realisations, it is of extreme importance to consider a scenario reduction methodology. Some authors have been applying clustering techniques, mainly in the mining industry and the petroleum area (Scheidt and Caers, 2009a, 2009b; Armstrong et al., 2013; Deutsch and Srinivasan, 1996; McLennan and Deutsch, 2005). Therefore, the contents presented in section 4.2.2 are addressed to an adapted scenario reduction methodology to apply to the geostatistical realisations of the rock mass geomechanical parameters.

Then, to validate the previously mentioned methodologies, an analysis of the rock mass mechanical behaviour during the excavation of a tunnel was examined. To this end, the Flac^{3D} (Itasca, 2012) finite difference software was used. Hence, the main purpose of applying the referred methodologies consists in mapping, simultaneously, the uncertainty associated with the geomechanical parameters quantification, their spatial variability and the heterogeneities existing in the rock masses at different spatial scales.

For this purpose, real data from a Chilean ore deposit, located in the Atacama Region in Chile, was used. The data include the Rock Mass Rating (RMR) information extracted from mechanical boreholes performed in a spatial regular grid. Using this information, a theoretical tunnel was

modelled, using the developed methodology to simulate the rock mass properties, and the results were then compared with a similar analysis performed using a more traditional approach, that means, assuming the rock as a homogeneous mass.

In the final section, some conclusions will be drawn regarding the potential of this innovative characterisation methodology for heterogeneous rock masses, highlighting the developed work to give a step ahead in implementing geostatistical techniques, which, until now, were assumed complex and unpractical, mainly in what concerns the geomechanical characterisation of rock masses.

4.2. METHODOLOGY OUTLINE

4.2.1. General

The present methodology starts with a data collection, *i.e.* all the available geotechnical information originating from different sources, such as boreholes, *in situ* and laboratory tests, should be assembled together (see Figure 4.1:step 1). Hence, a statistical analysis of the available data should be performed in order to outline the behaviour of the variable(s) of interest through the calculation of the mean, variance, minimum and maximum, among other geostatistical parameters required for simulation.

Alongside the aforementioned step, another primary step concerns the development of the grid model selected to represent the rock mass block and the underground structures using Flac^{3D} software. Once all the shapes and elements of the model in study are defined, the corresponding 2D or 3D spatial information from all zones centroids is exported so that it could be used in the following rock mass characterisation methodology (see Figure 4.1:step 2).

Thus, a conditional simulation of the chosen variable(s) on the target grid points (zones centroids of the Flac^{3D} grid) is performed. The chosen variable(s) can represent either a geomechanical parameter or an empirical system commonly used to characterise the rock mass. If the variable(s) is an empirical system, the use of empirical formulas to obtain the geomechanical parameter values is an additional required step (see Figure 4.1:step3).

Due to the smoothing effect arising from averaging the geostatistical realisations, a scenario reduction methodology based on clustering was developed and used. This methodology allows

obtaining a reduced number of realisations from the full set of realisations for the subsequent numerical analysis of the mechanical behaviour.

Furthermore, to connect the information resulting from the geostatistical realisations to the Flac^{3D} grid, a command routine was programmed using *FISH*, an embedded language in Flac^{3D}. This routine allows assigning a set of different values to each zone centroid in the whole model, *i.e.* each centroid shows different values for the geomechanical parameters (see Figure 4.1:step 4).

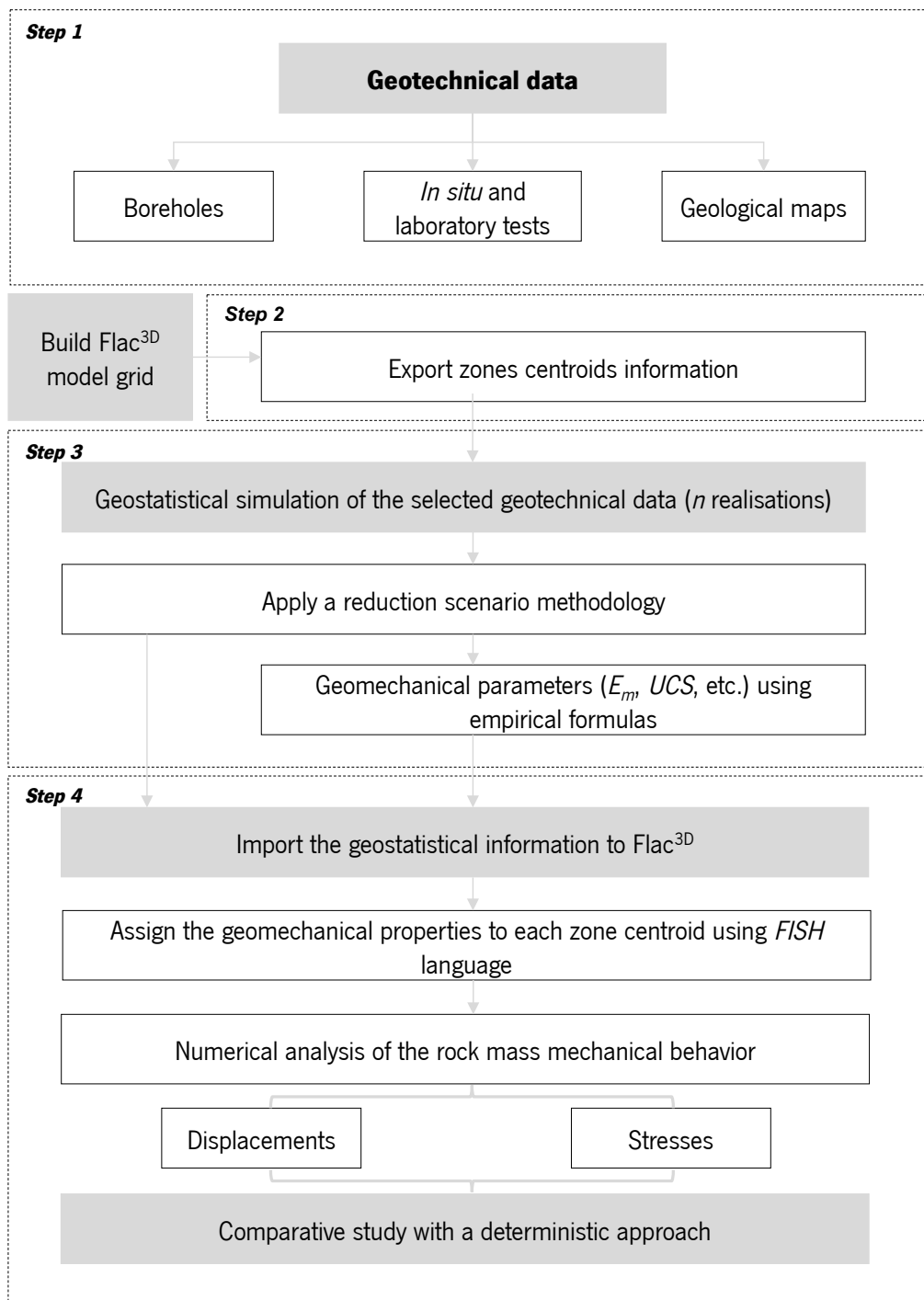


Figure 4.1 General workflow containing the general steps of the proposed characterisation methodology.

4.2.2. Geostatistical simulation

As mentioned in Chapter 2, where details are given regarding the geostatistical techniques and algorithms, simulation is an increasingly used technique that allows the prediction of rock mass properties and, at the same time, identifies different scales of heterogeneities and quantifies the uncertainty in the predicted values. Geostatistical simulation results in multiples representations of the random field(s) or variable(s), called realisations, which can be obtained in a conditional or non-conditional way, depending on whether or not the initial data of the variable are used and honoured in the simulation process. Throughout this thesis, the conditional simulation is the only applied type of simulation since it results in more accurate and realistic outputs always in line with the initial data and their true values.

To perform conditional simulation, it is important to analyse the structural behaviour of the data by computing their variogram (γ), which is a tool that relates the values observed at two points as a function of their distance in space (h) (see details in Chapter 2). Several steps should be followed from the time when the initial data is available until obtaining the final number of realisations of the chosen variable(s). Therefore, in this work the general steps proposed by Chilès and Delfiner (2012) and adapted by Pinheiro et al. (2015), are used and listed below:

1. Primarily, the data are analysed and basic statistics are calculated;
2. The data are then transformed into data with a standard Gaussian distribution (normal scores transformation);
3. The experimental variograms are then computed and accordingly fitted using mathematical basic structures (Gringarten and Deutsch, 2001). At this point, a spatial behaviour study of the data can be performed in order to identify the main anisotropy directions in the parameters spatial behaviour. It is important to notice that this anisotropy is related with the directional variation of the data in space and is not equivalent to the anisotropy of the geomechanical parameters;
4. Using the Flac^{3D} zone centroids as a simulation grid, a Gaussian random field is simulated (non-conditional simulation) at the target points using the Turning Bands Method (Matheron 1973; Emery and Lantuéjoul, 2006). In addition, a post-processing stage based on ordinary kriging is used to condition the simulation to the initial data;
5. Back-transform the Gaussian values of each realisation into their original scale;
6. Perform a model validation by calculating some error metrics and linear regressions.

In addition, a scenario reduction can be applied to the obtained number of realisations, for which further details are given in the following section.

4.2.3. Scenario reduction methodology

It is worth noticing, once again, that the scenario reduction methodology was required due to the smoothing effect existing when the realisations average is calculated, which precludes reproducing the spatial variability and heterogeneity in the rock mass. Therefore, and considering that the main goal of this thesis is to assess the uncertainty, to quantify the spatial variability and to identify the rock mass heterogeneities, it was of utmost importance not to eliminate the extreme values of the simulated variable. Since a high number of realisations is used to calculate an average, the values tend to the initial data mean value and the most unfavourable, the most favourable or even the most realistic scenarios would be eliminated. However, the use of all the obtained realisations would be unpractical, especially in geotechnics, where numerical analysis of the mechanical behaviour is currently carried out. As alternative, individual realisations can be used to characterise the variable of interest, although this selection cannot be made randomly and without any type of criteria. Consequently, two relevant questions arose: 1) How can the selection of the individual realisations be carried out?; 2) How many realisations are needed to achieve a full set of possibilities for the variable (geomechanical parameter) of interest?.

For this reason, it was important to apply a methodology able to reduce the number of realisations to a reasonable number for practical applications and, at the same time, to realistically characterise the random field. A list of this kind of methodologies was proposed in Chapter 2; however, none of them could be directly applied in this problem without making some adjustments.

Using as a reference the work of Scheidt and Caers (2009a), the methodology proposed here also applies clustering techniques to reduce the number of geostatistical realisations, more precisely a partitioning type of clustering, the K-medoid algorithm (Scheidt and Caers, 2009a). It is worth noticing that the main goal is to find an optimum number of realisations that is capable to statistically represent the whole set of realisations without needing a high computational time and enabling a field application in underground works. The general steps outlined for the methodology are as follows:

1. Calculate the Euclidean distance of all the grid points between each one of the n realisations;

2. Define the dissimilarity matrix between the n realisations using the Euclidean distance resulting from them;
3. Apply a Multi-Dimensional Scaling technique (MDS) to represent the dissimilarity matrix in a $2D$ space, and herewith a spatial representation of the n realisations is accomplished, in which each point represents one realisation;
4. Apply the kernel function to obtain the Featured Space (*i.e.* a more linear representation of the points, meaning the realisations) (see Chapter 2 for details);
5. Perform the kernel k-medoid clustering for a previously defined maximum value of clusters;
6. Calculate the silhouette value for each cluster (from 1 until the maximum of clusters) and evaluate the optimum number of clusters to use;
7. Plot the optimum number of clusters and their respective medoids (centres);
8. Perform the post-processing analysis to understand the validity of each cluster.

As a note, it is important to underline that each point represented in the $2D$ space corresponds to one of the n geostatistical realisations. Also, in what concerns numbers 6 and 7 above, as regards the optimum number of clusters, it is noteworthy that in the developed methodology the maximum number of clusters is not assumed as an input variable but as an automatic process, *i.e.* using optimisation methods. In this regard, a Matlab toolbox called *evalclusters* (Tibshirani et al., 2001). was used, resulting in the computation and plotting of each cluster configuration adopting as evaluation criterion the silhouette value.

The silhouette method was firstly introduced by Rousseeuw (1987) aiming to be a measure of similarity between the points within the same cluster compared with the points of the other clusters. For this purpose, it is required to have the partition obtained after applying the clustering technique and using the dissimilarity between each realisation. From here, the silhouette value for each i^{th} object can be obtained through

$$S_i = (b_i - a_i) / \max(a_i, b_i) \quad (4.1)$$

where a_i represents the average distance from the i^{th} point to the other points in the same cluster i , and b_i is the minimum average distance from the i^{th} point to all the other points placed in a different cluster, minimised over all the clusters.

By definition, the value of S_i can vary between -1 to +1. A positive value closer to the unit is desirable because it means that the i^{th} point is well-matched to its own cluster and poorly-matched with the neighbouring clusters, indicating that the point is classified in the right cluster.

On other hand, a negative value for the silhouette (-1) corresponds to a case in which a_i is greater than b_i reflecting that the clustering solution has too many or too few clusters.

In an intermediate case, if S_i has a value near 0, a_i and b_i are approximately equal and it is not clear whether i should have been assigned to cluster A or B, *i.e.* the i^{th} point lies in-between both clusters.

For each cluster, a unique silhouette value can be defined by performing the average of the S_i values for all objects i belonging to the cluster. For example, if a cluster is composed by a total of 100 points, an average of 100 silhouette values is performed, resulting in the so-called *average silhouette width*. An *overall average silhouette width* can be obtained for the entire plot considering all the points that compose the whole data set (Rousseeuw, 1987).

After plotting the *average silhouette width* for all the clusters, the optimum number is chosen for the configuration that results in the higher value, taking as a reference the ranges presented in Table 4.1.

Table 4.1 Silhouette reference values used to evaluate the optimum number of clusters (Struyf et al., 1997).

Silhouette value	Designation
[0.71 - 1]	A strong structure has been found
[0.51 - 0.7]	A reasonable structure has been found
[0.26 - 0.50]	A weak structure has been found and could be artificial
< 0.25	No substantial structure

It is important to stress that all the mentioned techniques and algorithms were programmed in Matlab (MathWorks, 2016) and were entirely adapted and rewritten to allow the integration of the geostatistical information and to consider the geotechnical type of variables used during this thesis development. Also, some internal functions, like the Multi-Dimensional Scaling (MDS) and *evalclusters* (Tibshirani et al., 2001) were integrated into the new developed routines (see Figure 4.2).

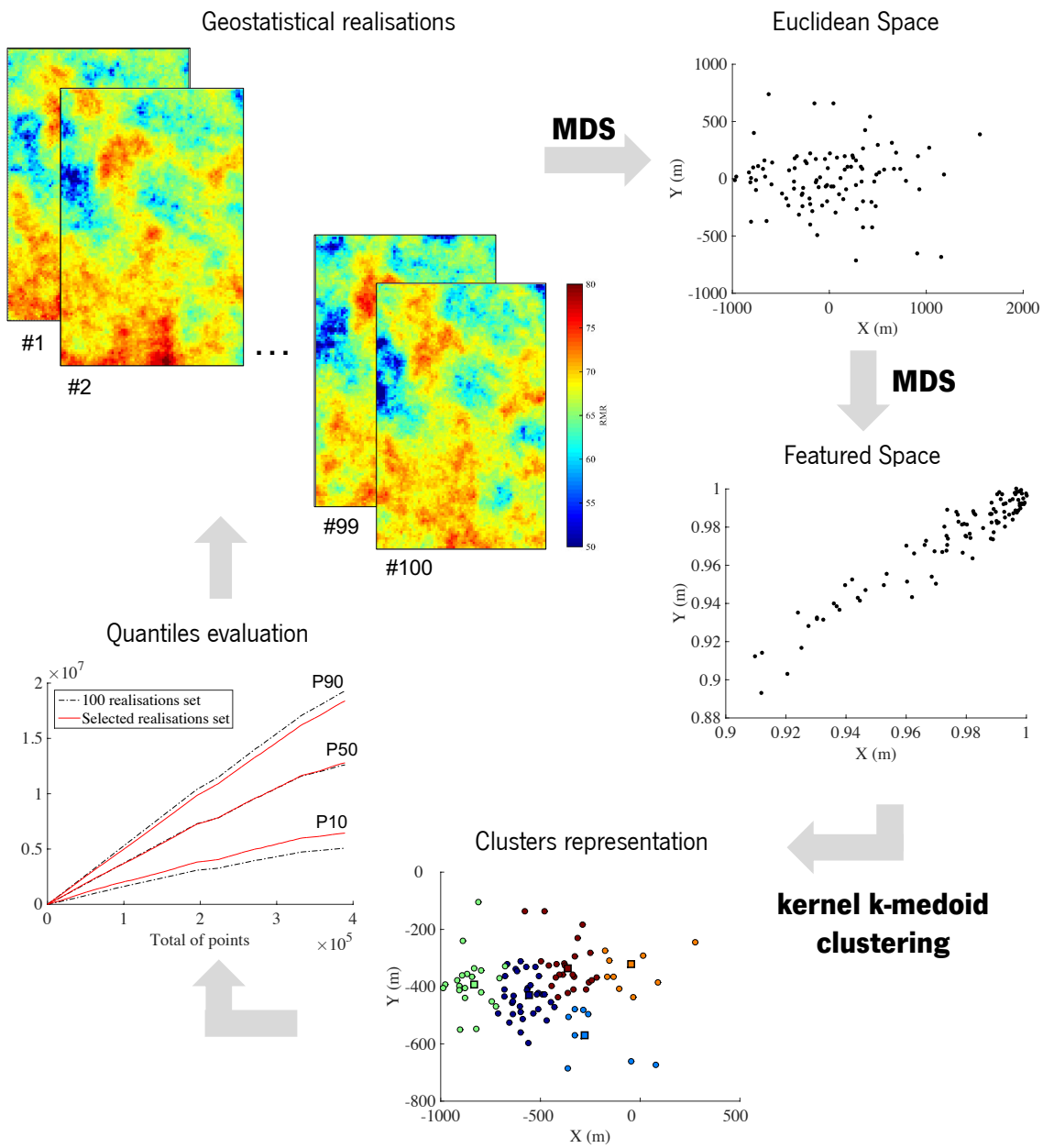


Figure 4.2 Scheme to follow in order to apply the scenario reduction methodology proposed in this section.

4.3. CASE STUDY - CHILEAN DEPOSIT

4.3.1. Geotechnical data

Once again, the data resulting from an epithermal gold deposit located in the Atacama Region in Chile will be used here to validate the proposed methodology. The available data set comprises a total of 3669 samples obtained from a regular grid, spaced 40 m in directions x and y , in the horizontal plane and 20 m in z direction. These data were used to obtain the information regarding the RMR five parameters (all excepting parameter P6 related to the orientation of discontinuities) and, subsequently, used to measure the UCS through the execution of laboratory tests using cylindrical rock samples showing an isotropic behaviour. According to the data statistics, the RMR varies from 48 to 78, *i.e.* from a fair to a good geomechanical quality of the rock mass. More details about the data set were given in Chapter 3.

4.3.2. Three-dimensional numerical model

A 3D numerical model was developed using Flac^{3D} software in order to investigate the differences in the tunnel behaviour when a heterogeneous approach is used to obtain the geomechanical parameters of the rock mass. The modelled tunnel presents a length of 90 m and is composed by a 6 m radius arch with vertical sidewalls with the same height. The Flac^{3D} model grid is 108 m wide along the direction x , 90 m along direction y and 200 m along the direction z and is composed, mostly, of brick elements, which size increases as one moves far away from the tunnel, resulting in a finer mesh near the excavation. A total of 100,800 zones and 105,742 grid points compose the model. The tunnel central axis is located 96 m below the ground surface (see Figure 4.3).

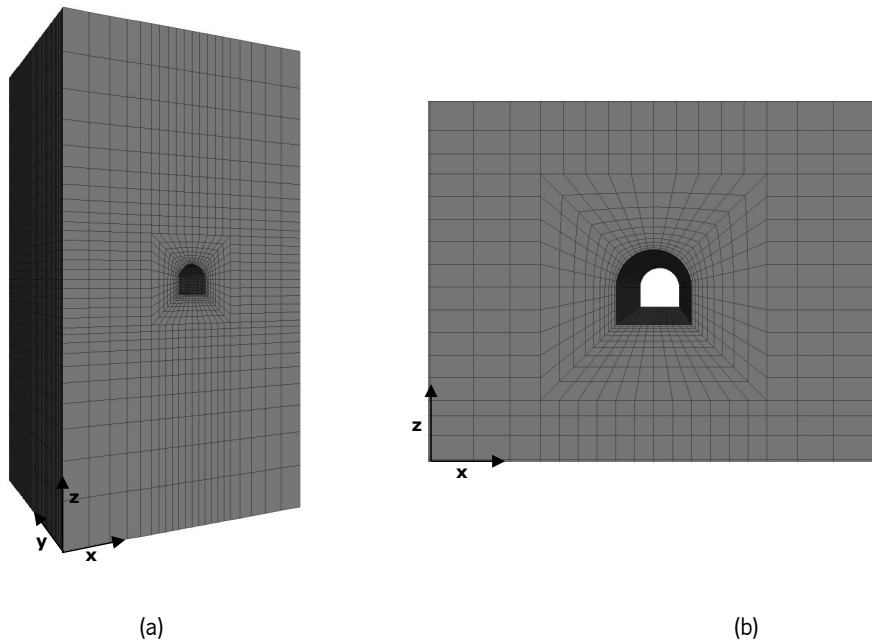


Figure 4.3 Finite difference grid of tunnel model: a) xyz perspective; b) xz plan view of the tunnel zone.

In what concerns the model boundary conditions, the horizontal displacements in the model vertical boundaries were blocked, along with all displacements in the inferior boundary.

The adopted tunnel support system is composed of 0.20 m thick shotcrete simulated with shell elements with a linear elastic isotropic behaviour, with a deformation modulus of 20 GPa and a Poisson's ratio of 0.25. The construction process begins with the excavation of the tunnel arch in a 3 m length followed by the application of shotcrete in the arch. Then, the remaining part of the tunnel is excavated and finally the shotcrete is applied in the walls of the tunnel. A gravitational initial stress field with a horizontal to vertical stress ratio (K_0) of 0.5 was adopted.

4.3.3. Geostatistical simulation results

Taking into account the available data of the case study, the geostatistical simulation was performed for RMR and UCS variables on the previously constructed Flac^{3D} mesh (*i.e.* considering the zone centroids spatial information). As mentioned in section 1.2, the required variogram of both variables (previously transformed into normally-distributed variables) was built along two main directions of anisotropy (regarding the spatial variation of the data), namely, the horizontal plane (xy plane) and the vertical direction (z), considering lag distances multiples of 40 m for the horizontal plane and 20 m for the vertical direction (see Figure 4.4).

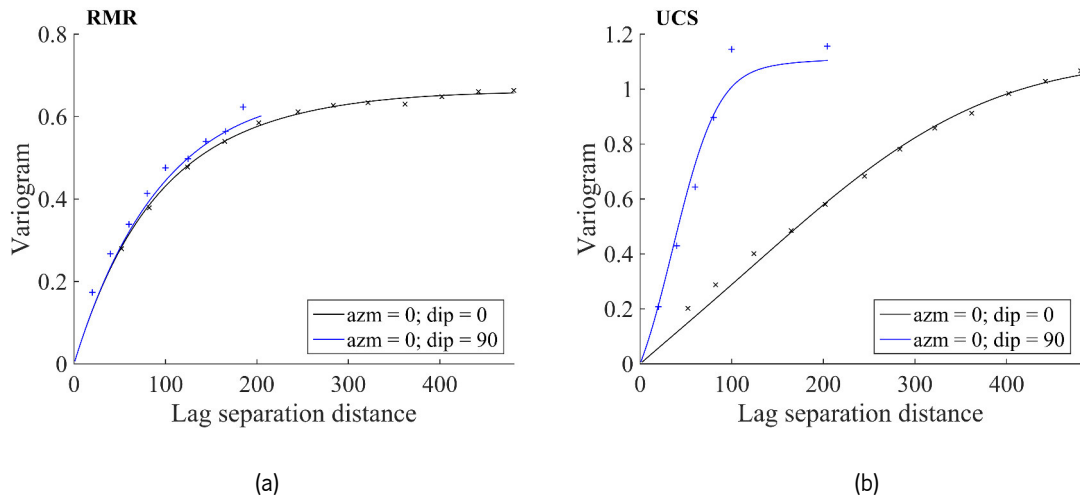


Figure 4.4 Experimental (crosses) and theoretical (solid lines) variograms along the main anisotropy directions: horizontal plane (black) and vertical (blue): a) RMR; and b) UCS.

Analysing Figure 4.4a, it is possible to observe that the RMR shows an almost isotropic behaviour since the variogram range is similar for both analysed directions. However, the same does not happen for UCS, with the vertical direction showing a smaller range than the horizontal plane, *i.e.* confirming the distinctive spatial behaviour in both directions (high variability for smaller distances in the vertical direction). Nonetheless, in Figure 4.4 it is also possible to observe in solid lines the variogram models. The variograms are modelled with exponential and Gaussian basic structures, as presented below (the distances between brackets represent the correlation range for each anisotropy direction and the number preceding the basic structure indicates the sill of the structure):

- RMR variogram model:

$$\gamma = 0.61 \text{ Exponential } (250 \text{ m}; 250 \text{ m}) + 0.05 \text{ Gaussian } (450 \text{ m}; 250 \text{ m})$$

- UCS variogram model:

$$\gamma = 0.44 \text{ Exponential } (450 \text{ m}; 150 \text{ m}) + 0.66 \text{ Gaussian } (500 \text{ m}; 105 \text{ m})$$

Once obtained the variogram models, the realisations of both variables can be constructed using the TBM mentioned before (with a total of 1500 turning lines) and conditioned to the sampling data. The number of realisations was set to 100, resulting in a total of 100 RMR and UCS realisations.

For the simulation process the average of the 100 realisations for each of the 100,800 zone centroids of the Flac^{3D} mesh (x , y and z coordinates) was calculated. In Table 4.2 the RMR basic statistics for the initial values and for the 100 realisations average are provided, as well as for the UCS variable. Analysing the simulations closely, it is possible to detect a good fit between the initial statistics and the simulated ones for the UCs variable; however, the same cannot be stated for the

RMR variable that shows a variation scale considerably narrow, probably justified by the high RMR values in zones with more mesh elements (a more refined mesh was used in the elements near the tunnel). Furthermore, in order to demonstrate the spatial variability existing in this rock mass, the RMR first realisation and average of 100 realisations were mapped in three-dimensions as shown in Figure 4.5. The scale of variation for the RMR is considerably low and the rock mass can be assumed as nearly homogeneous.

Table 4.2 RMR and UCS (in MPa) statistics for the initial values and the average of 100 conditional realisations.

	RMR		UCS	
	Initial values	Average of simulated RMR	Initial values	Average of simulated UCS
Number of grid points	3969.0	100,800	3969.0	100,800
Minimum	48.0	68.0 ¹	137.9	139.3 ¹
Maximum	78.0	73.0 ¹	207.7	189.8 ¹
Mean	66.7	70.8	176.9	163.5
Variance	14.2	5.6	1054.7	908.7

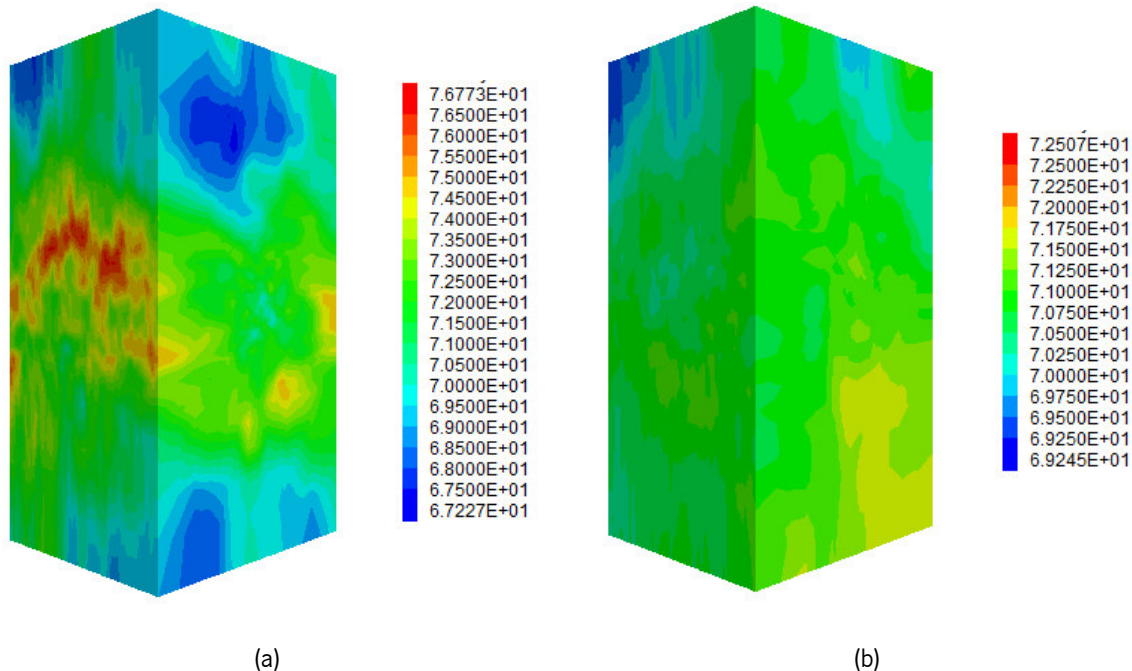


Figure 4.5 3D maps of RMR simulated on Flac^{3D} mesh, for: a) realisation #1; and b) average of 100 realisations.

4.3.4. Reduction scenario methodology applied to RMR simulations

¹ The minimum and maximum values were evaluated after averaging the 100 realisation values.

As perceived from the content of Table 4.2 by performing the 100 realisations average the extreme values of the RMR were, when compared with the initial data set values, much closer to the mean. In the previous Chapters, it was already stated the smoothing effect existing after averaging the geostatistical realisations, which cannot translate, in a precise and accurate way, the variable spatial model. Consequently, it is of interest to use a scenario reduction methodology, as previously exposed, in order to keep with a few realisations instead of their average.

Firstly, it was necessary to spatially represent the 100 RMR realisations, achieved after computing the Euclidean distance between the 100 realisations using all the realisations grid points. To make it clear, the Euclidean distance $d(x, y)$ is computed between realisation x and realisation y using the sum of the square difference of both realisations values in all the grid points (x_i and y_i) as,

$$d(x, y) = \sqrt{\sum_{i=1}^N (x_i - y_i)^2} \tag{4.2}$$

where N represents the number of grid points, and x_i and y_i take, in this case, values of RMR in each i point of the realisation x and y , respectively. From here, and using as reference the Euclidean distances between all the realisations, a dissimilarity matrix was calculated (see Table 4.3).

Table 4.3 Extract of the RMR 100 realisations dissimilarity matrix (each column and line represent the realisation number).

	1	2	3	4	5	6	7	8	9	...	95	96	97	98	99	100
1	0	695	956	706	979	708	1229	768	662		1322	963	1237	683	792	687
2	695	0	955	834	744	658	1234	761	659		1431	1190	1218	711	926	740
3	956	955	0	1183	1254	682	658	756	1009		733	871	1843	867	686	1084
4	706	834	1183	0	1103	917	1497	947	821		1569	1066	1185	776	1064	715
5	979	744	1254	1103	0	925	1411	1120	863	...	1674	1567	1128	1126	1178	1045
6	708	658	682	917	925	0	981	714	769		1093	993	1451	734	731	808
7	1229	1234	658	1497	1411	981	0	991	1257		633	1129	2152	1138	883	1380
8	768	761	756	947	1120	714	991	0	746		1167	847	1621	624	777	762
9	662	659	1009	821	863	769	1257	746	0		1445	1090	1186	677	874	673
...						
95	1322	1431	733	1569	1674	1093	633	1167	1445		0	1036	2289	1260	944	1506
96	963	1190	871	1066	1567	993	1129	847	1090		1036	0	1867	885	782	1036
97	1237	1218	1843	1185	1128	1451	2152	1621	1186	...	2289	1867	0	1472	1612	1256
98	683	711	867	776	1126	734	1138	624	677		1260	885	1472	0	829	632

99	792	926	686	1064	1178	731	883	777	874		944	782	1612	829	0	964
100	687	740	1084	715	1045	808	1380	762	673		1506	1036	1256	632	964	0

Through the application of a Multi-Dimensional Scaling technique (MDS), a 2D spatial representation of the dissimilarity values obtained between realisations was obtained (see Figure 4.6). This representation is accomplished through a random generation of coordinates (X and Y in meters) and consecutive spatial arrangement in a way that the distance between the realisations (in the figure represented by black points) have the same spatial distance as the dissimilarity distance computed before.

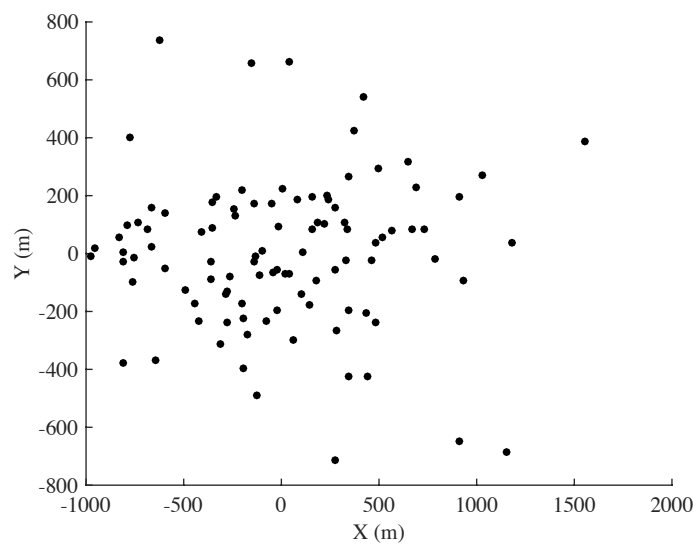


Figure 4.6 2D spatial representation of the RMR 100 realisations (black points) using the Euclidean distance.

In order to apply the kernel k-medoid algorithm an essential step lies in the transformation of the Euclidean space representation into a more linear one, the Featured Space. In this regard, a Kernel width of 2000 was adopted to guarantee the linearity of the points that represent the RMR 100 realisations (see more details in Chapter 2). Before setting the final number of clusters, a cluster evaluation was performed. For this purpose, a Matlab toolbox called *evalclusters* (Tibshirani et al., 2001) was used and the cluster evaluation was made considering the calculation and plot of the clusters Silhouette values (Rousseeuw, 1987). As result, Figure 4.7 shows the graphic obtained – relating in the x axis the number of evaluated clusters and in y axis the results for the *average silhouette width*.

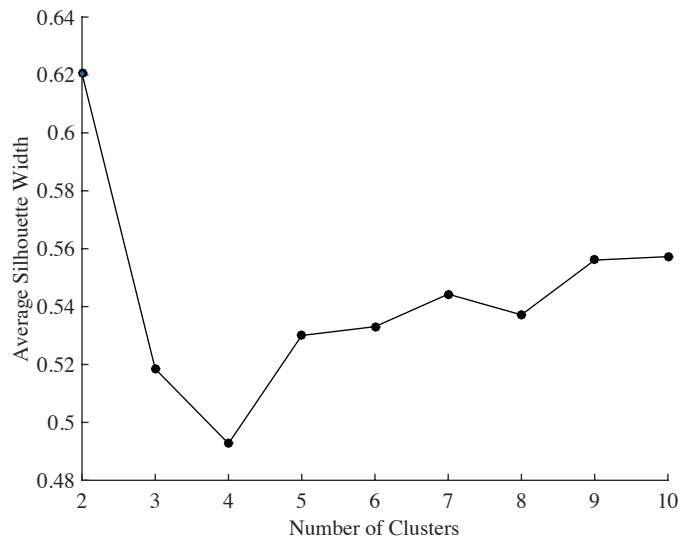


Figure 4.7 Average silhouette width values for the performed clusters evaluation.

According to Figure 4.7, the optimum number of clusters obtained for this case study was 2, which corresponds to an *average silhouette width* value of 0.62. However, the differences in the *average silhouette width* values for other number of clusters are not pronounced, meaning that almost every number of clusters (with an *average silhouette width* value higher than 0.50) could be chosen in this case.

Hence, and considering a total of 2 clusters to perform the reduction scenario of the RMR full realisation set, Figure 4.8 shows the final configuration for both clusters with the identification of the associated medoids (centres). This configuration was set by applying the kernel k-medoid for a maximum of 500 iterations. By way of comparison, a single cluster with no data set division was also considered. It is noted that, in Figure 4.8, the *X* and *Y* axes show the scale resulting from the Euclidean distance representation (see Figure 4.6).

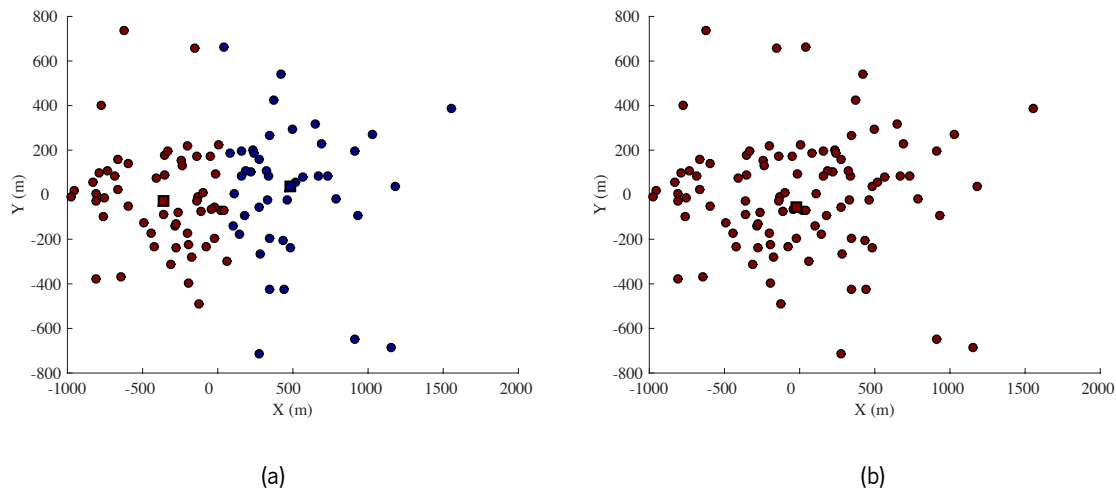


Figure 4.8 Clusters final configuration (points) with the matching medoids (squares) for: a) 2 clusters; and b) single cluster.

Along with the clusters graphical representation, the use of this methodology allows the identification and listing of the medoids corresponding realisations to be used in the following steps of this characterisation methodology. As mentioned, the main goal of this scenario reduction methodology is to select individual realisations of the random field with a mathematical foundation, avoiding loose statistical and meaningless information.

Hence, a validation of the selected clusters configurations was carried out to analyse the differences with the full realisation set. Therefore, the percentiles 10, 50 and 90 of each realisation that composes both data sets (full realisation set and the selected clusters) were computed. This metric indicates the value below which a given percentage of realisations falls. Since in the present problem, the chosen variable (RMR) cannot be related with a time step, as happens in the petroleum and mining studied fields (e.g. net present value of oil production for each time interval), the ways defined to graphically represent the percentiles values obtained for each grid point and computed between the realisations (full realisation set and selected clusters) were: 1) the percentiles 10, 50 and 90 values were summed point by point in order to represent the percentile evolution for all the grid points. The goal was to understand if with a smaller set of realisations (two and one) the percentiles value show the same trend. As an example, for percentile 10 of grid point number 1 was represented in the graph, then for point number 2 the corresponding percentile 10 value was summed with percentile 10 value of grid point 1 and so on until the total of grid points is achieved; and 2) the percentiles 10, 50 and 90 values of each grid point were summed until the total of grid points are reached. Then the total is divided by the number of grid points so that an average value of percentile 10, 50 and 90 is obtained. This last type of graphical representation allows obtaining a more linear overview of the sets differences in all the simulated block (grid points). Figure 4.9

displays the referred two graphics for a two-clusters and a single cluster structure in comparison with the full realisation set. From figure analysis, it is possible to observe that, in both cases, the RMR average values for all the percentiles are never higher than 1.5 (see Figure 4.9a and Figure 4.9c). Indeed, analysing in detail Figure 4.9a and Figure 4.9c, it is possible to notice that the differences between the selected realisations set and the 100 realisations set are smaller in a case of a single cluster.

With regard to the percentile point by point representation, Figure 4.9b allows perceiving that the percentiles plot of the selected realisation set, when compared with the full realisation set, are following the same evolution and are almost coincident, proving that by assuming a two clusters configuration the full set of the 100 realisations can be statistically well represented. In the following table (see Table 4.4) the basic statistics (mean, variance, minimum, maximum and percentiles) of all the sets, initial values, 100 realisations and selected realisations, are presented. As presented in Table 4.4, when the average of the 100 realisations is performed the minimum and maximum value are closer to the RMR mean value, so the extreme values are not considered into the rock mass model, as anticipated. Henceforth, Table 4.5 displays the basic statistics of the RMR for the central realisations (corresponding medoids) of the two clusters and single cluster configurations along with the 100 realisations average values for comparison.

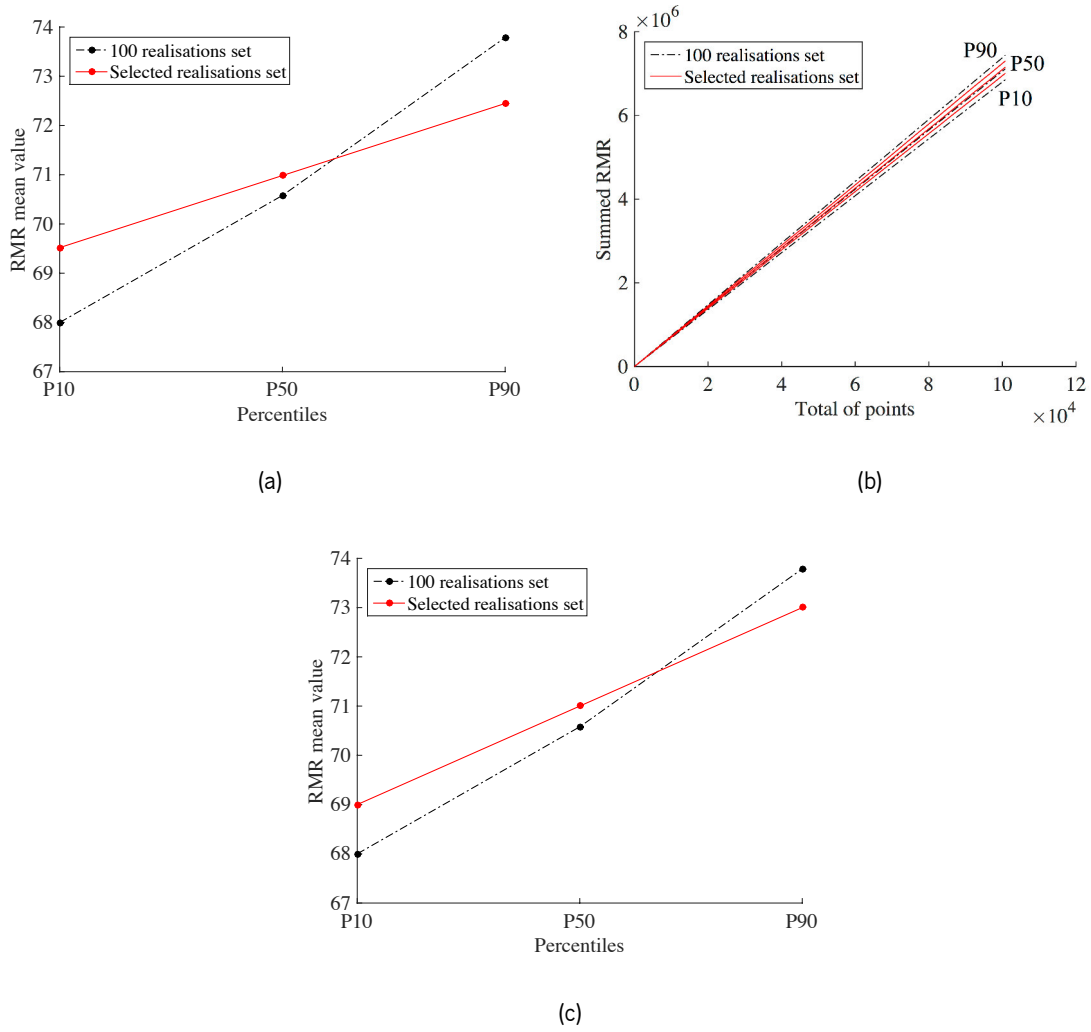


Figure 4.9 Validation of the two clusters configurations compared with the 100 realisations set using the percentiles 10, 50 and 90, for: a) average percentiles values for the 2 clusters configuration; b) point by point percentile sum for the 2 clusters configuration; and c) average percentiles values for the single cluster configuration.

Table 4.4 Basic statistics of the RMR variable for the initial values, 100 realisations and selected realisation data sets.

	Initial values	100 realisations (average values)	2 Clusters (average values)	Single cluster
Number of grid points	3969	100,800	100,800	100,800
Mean	66.69	70.78	70.64	70.71
Variance between points	14.21	0.12 ²	1.02 ²	1.93
Variance between realisations	-	5.63 ³	4.13 ⁴	-
Minimum	48.00	68.00 ⁴	67.50 ⁴	64.00
Maximum	78.00	73.00 ⁴	75.50 ³	77.00
Percentile 10	-	67.99	69.44	69.00
Percentile 50	-	70.58	70.64	71.00
Percentile 90	-	73.79	71.84	73.00

Table 4.5 Basic statistics of the RMR for the individual realisations of the two clusters, single cluster configurations and 100 realisations average values.

	Cluster 1	Cluster 2	Single Cluster	100 realisations (average values)
Number of grid points	100,800	100,800	100,800	100,800
Realisation number	15	83	41	-
Number of realisations per cluster	53.00	47.00	100.00	-
Mean	71.71	69.56	70.71	70.03
Variance between points	2.23	1.63	1.93	0.12
Standard deviation	1.50	1.28	1.39	0.35
Minimum	66.00	65.00	64.00	68.00
Maximum	81.12	76.00	77.00	73.00

As displayed in Table 4.5 by considering individual realisations, the maximum values stay closer to the initial values (see first column of Table 4.4); however, the same statement cannot be made regarding the minimum values since the initial set shows a minimum RMR equal to 48, while the two and single clusters resulted in minimum values of 65 and 64, respectively. When a different metric is compared, namely the variance between the points, it is possible to observe that along with the 100 realisations average, the individual realisation set are still far from the variance that

² Variance obtained after performing the average of the realisations for each grid point and then compute the variance between the grid points.

³ Variance obtained after computing the variance between realisations for each grid point and then calculate the average variance values.

⁴ the minimum and maximum values were obtained after performing the realisations average and, therefore, the values are higher and lower than the ones obtained in each individual realisation.

composed the initial set. Even though, seems reasonable to assume that the best approach to construct the rock mass model is to assume geostatistical realisations as possible characterisation scenarios for the rock mass, namely the individual realisations instead of the average values.

4.3.5. From geotechnical data to geomechanical parameters

Once having finished the Flac^{3D} mesh construction, the simulation of the geotechnical information and the realisations selection through the application of a reduction scenario methodology, a subsequent and crucial step consists in obtaining the geomechanical parameters to use in Flac^{3D}, namely the rock mass deformation modulus, since a linear elastic behaviour was adopted to represent the rock mass. A unit weight of 25 kN/m³ and a Poisson's ratio of 0.20 were assumed for the rock mass.

With the purpose of obtaining the geomechanical parameters using the RMR information, several empirical expressions were considered. Then, by applying a statistical methodology proposed by Miranda (2003) all the empirical expressions were used to obtain the rock mass deformation modulus (E_m). The aim of the methodology is to reduce the uncertainty of applying only one empirical formula to obtain E_m and combines different expressions found in the literature.

The empirical expressions that required the RMR as an input parameter were already listed in Chapter 2, thus four formulas proposed by different authors were used to calculate the E_m . It is important to notice that all the used formulas were developed using data from similar rock types as the one in this case. Since the information regarding the Uniaxial Compressive Strength (UCS) exists and the geostatistical simulation was performed, some of the selected formulas also consider its value as an input. Likewise, the same happens for the Geological Strength Index (GSI) that could be obtained from the RMR using a correlation formula (Hoek et al., 1995). Table 4.6 shows the selected E_m expressions along with their limitations of application and their authors.

Table 4.6 Empirical expressions used to obtain E_m and the corresponding authors.

Number	Author	Required Parameters	Limitations	Equation (E_m in GPa)
1	Bieniawski (1989)	RMR	$RMR > 50$	$E_m = 2 \times RMR - 100$
2	Mohammadi and Rahmancejad (2011)	RMR	-	$E_m = 0.0003 \times RMR^3 - 0.0193 \times RMR^2 + 0.315 \times RMR + 3.4064$
3	Hoek et al. (2002)	UCS, GSI	$UCS > 100$ MPa	$E_m = \left(1 - \frac{D}{2}\right) \times 10^{(GSI-10)/40}$
4	Read et al. (1999)	RMR	-	$E_m = 0.1 \times \left(\frac{RMR}{10}\right)^3$

In the third formula, an additional parameter represented by the D letter is required. This coefficient is dependent on the disturbance degree of the rock mass due to the excavation process and, in this case, an intermediate value was adopted ($D=0.5$).

The admissible interval (AI) methodology (Miranda, 2003) uses as an input the data obtained when different formulas are applied and combined them in order to get an admissible interval, which can be calculated using Equation (4.3).

$$AI = mean \pm \sigma \quad (4.3)$$

where the mean term corresponds to the average values of E_m using all the chosen formulas and σ represents the standard deviation of the average values. The range of the interval is given by the E_m mean plus and minus one standard deviation.

Subsequently, a filter on the E_m values for each one of the four formulas was applied, aiming to identify the values located outside the range of the AI. These outside values were then eliminated and a new average of E_m was computed using the remaining and accepted values. As consequence, a single E_m value was calculated for each grid point of the Flac^{3D} mesh. In the case of all the values being rejected, the E_m value was obtained with the average of the four formulas. Note that the methodology was applied to each point of the Flac^{3D} mesh and then to each one of the 100 realisations.

4.3.6. Models in analysis

In order to understand the advantages of these heterogeneous approaches, the results of the numerical models obtained from individual realisations needed to be compared with the traditional

approach of assuming a homogeneous medium (deterministic values for the geomechanical parameters). Indeed, this comparison allows assessing the main differences in terms of accuracy and uncertainty reduction in the rock mass characterisation process. For this model, called as model 4, the deformation modulus was obtained from the average value of the RMR in the 100,800 zone centroids, using the four empirical formulas above and averaging the obtained E_m results. Therefore, apart from the single cluster (model 1) and two clusters (model 2) configurations, a third model was considered using the realisations average values calculated in each grid point of the Flac^{3D} mesh, resulting in the called *Mean* model (model 3). The inclusion of this model aims to show the limitation in using the realisations average values instead of individual realisations since the extreme values are smoothed.

In addition, to confirm the advantages of using the scenario reduction methodology instead of the realisation averages or another methodology, the selected clusters results (two clusters and single cluster configurations) were compared with the numerical results obtained considering each one of the 100 realisations. It is important to point out that these calculations were carried out because the theoretical tunnel excavated in the Chilean rock mass was a simple case and therefore the computational time was modest.

For the sake of clarification of all the analysed models, Figure 4.10 presents a scheme illustrating the path followed on the mentioned process, starting with the variable selection and its geostatistical simulation, then the application of the scenario reduction methodology and, finally, the E_m output values used in the Flac^{3D} analysis. Moreover, Table 4.7 summarises the adopted values for models 1, 2 and 3 (see Figure 4.10). Regarding model 4, the deterministic value adopted to represent E_m was 32.73 GPa.

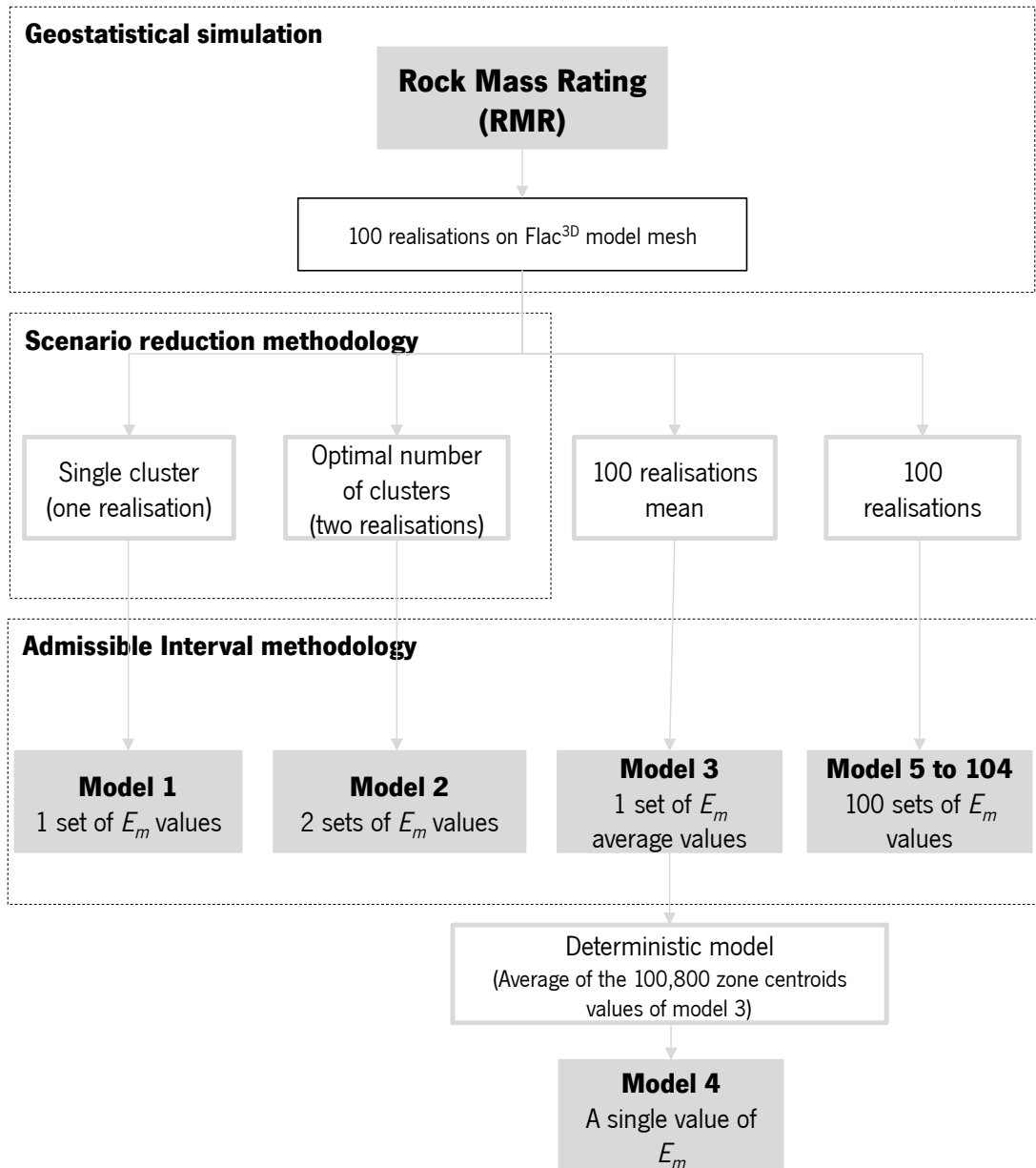


Figure 4.10 Workflow applied to build the different models to represent the rock mass characterisation of the Chilean rock mass using as an input the RMR system.

Table 4.7 Rock mass deformation modulus (E_m in GPa) basic statistics of model 1 (single cluster), model 2 (two clusters) and model 3 (100 realisations average).

	Model 1	Model 2		Model 3
	Single Cluster (real. 41)	Cluster 1 (real. 15)	Cluster 2 (real. 83)	Mean
Number of grid points	100,800	100,800	100,800	100,800
Number of real. per cluster	100.00	53.00	47.00	-
Mean	36.03	39.39	35.27	32.73
Variance between points	5.21	8.31	5.53	0.32
Minimum	25.79	29.00	27.39	29.98
Maximum	49.95	59.25	47.82	38.02

Figure 4.11 presents a xyz perspective with the spatial distribution of the E_m values for the previously defined models where it is possible to observe, in the case of models 1 and 2, zones with lower and higher deformation modulus marking the variability of the E_m values.

In particular, looking at model 2, it is possible to observe that both clusters are almost complementary, *i.e.* in cluster number 1 zones with lower E_m value correspond to zones with higher E_m value in the case of cluster number 2, and *vice-versa*. In the case of model 1, the E_m values pattern are quite different from the one presented in the 2 clusters of model 2; however, the latter were the ones that presented the highest and lowest E_m value.

Concerning model 3, it is important to stress out that the resulting heterogeneity level is very small, since the E_m minimum value is 33 GPa and the maximum value is 36 GPa.

For the remaining 100 models (model 5 to 104) the spatial distribution was not represented due to their large number; although the numerical analysis was performed and the results will be commented further.

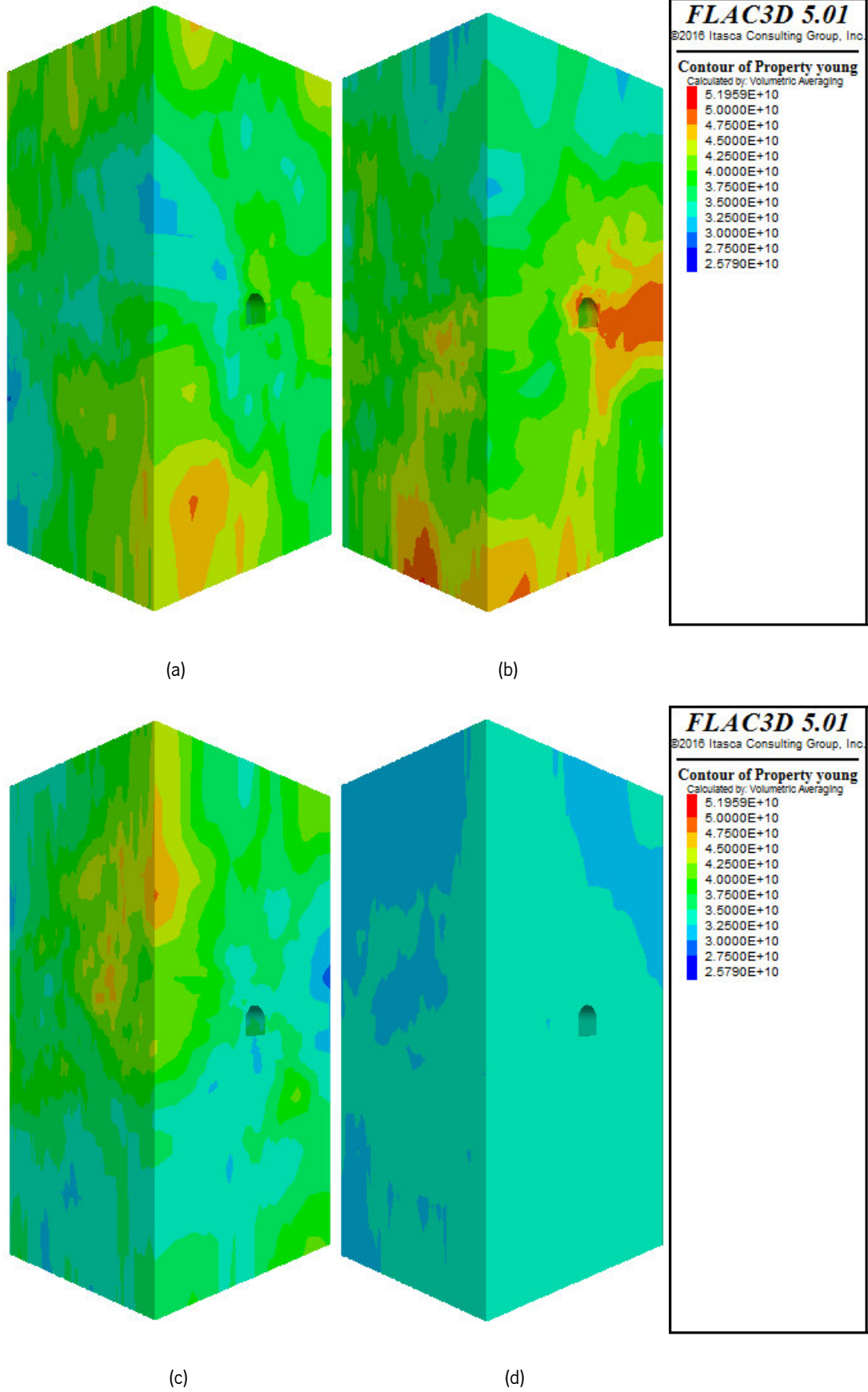


Figure 4.11 Flac^{3D} xyz perspective with E_m values (colour scale in Pa) used in: a) model 1; b) model 2 – cluster 1; c) model 2 – cluster 2; and d) model 3.

4.3.7. Numerical results

In this section, the numerical results obtained from the calculations carried out using the finite difference software Flac^{3D} are presented. Emphasis will be given to the results in terms of displacements since the deformation modulus affects them in a direct way; however, some reference will be made to the maximum and minimum principal stress values calculated in the rock mass. This analysis was performed for all the mentioned models to allow a further comparison between the values. In detail, regarding the displacements on the rock mass, the maximum values obtained on the top of the tunnel arch, at the mid-point of the invert, and at mid-height on the left and right side walls were registered.

Primarily, Table 4.8 presents a summary of the maximum, minimum and percentile values obtained from the calculations for all the 100 individual realisations. Since models 1 and 2 concern to individual realisations, they are also included in the 100 realisations values. Note that these values should serve as a reference for further comparisons.

Table 4.8 Displacements and stresses values obtained for the 100 individual realisations.

	Vertical displacement (mm)		Horizontal displacement (mm)		Max. stress (MPa)	Min. stress (MPa)
	Arch	Invert	Left wall	Right wall		
Minimum	5.20	6.82	1.50	1.42	6.67	1.77
Maximum	8.34	11.31	2.75	2.63	7.91	1.84
Mean	6.60	8.62	2.00	2.00	6.97	1.80
P10	5.92	7.76	1.73	1.71	6.81	1.77
P50	6.60	8.62	2.00	2.00	6.97	1.79
P90	7.49	9.80	2.28	2.33	7.21	1.82

Comparing the 100 realisations results, some differences are worthy of reference, namely the differences between the maximum and minimum values for the vertical and horizontal displacements. In terms of principal stresses, those differences are subtler with variations smaller than 1 MPa. If analysed the mean values, the displacement values differences between the left and right sidewalls are zero, confirming that the heterogeneity in this rock mass is truly smaller.

Looking in detail models 1, 2, and 4, Table 4.9 shows a summary of the displacements and stresses obtained from the numerical analysis.

Table 4.9 Summary of the displacements and stresses results obtained in each one of the first four models.

Model	Corresponding realisation	Maximum vertical displacement (mm)		Maximum horizontal displacement (mm)		Principal stress (MPa)	
		Centre of the arch	Centre of the invert	Centre of the arch	Centre of the invert	Maximum	Minimum
Model 1	Realisation 41	6.741	8.027	1.887	2.091	6.972	1.791
Model 2 - Cluster 1	Realisation 15	6.059	7.988	1.820	1.932	6.890	1.807
Model 2 - Cluster 2	Realisation 83	6.977	8.934	2.188	2.110	6.941	1.782
Model 3	Average	7.159	9.478	2.211	2.176	6.914	1.781
Model 4	-	5.860	9.360	1.790	1.790	6.910	1.770

Regarding the vertical displacements of model 1 (single cluster), the maximum value was observed on the centre of the tunnel invert with a magnitude of 8.03 mm, while the central point of the tunnel arch shows a smaller value of 6.74 mm (see Figure 4.12). Moreover, the tunnel walls show, as expected, maximum horizontal displacements of different values, recording a 1.89 mm displacement on the middle of the left wall and 2.10 mm on the right wall. The observed disparity between the horizontal displacements can be explained by the assumed variability in the E_m values (see Table 4.9).

Concerning the results of model 2 (two clusters) the same qualitative displacement pattern as model 1 can be observed. Indeed, cluster number 1 of model 2 shows lower values for all the displacements controlled points, while cluster number 2 registered higher values in the same points. The maximum vertical displacement occurs in the centre of the tunnel invert with, approximately, 9.00 mm (cluster 2), whereas the maximum displacement on the tunnel arch admits a magnitude of 6.06 mm (cluster 1). For the sidewalls, the range of values vary between 1.82 mm (left) and 1.93 mm (right) for cluster number 1 and 2.19 mm (left) and 2.11 mm (right) for cluster number 2. Indeed, an opposite effect happens between both clusters, since the maximum horizontal displacement is in contrary sidewalls. As anticipated, the displacements value for cluster 1 are smaller when compared with cluster 2, corroborating the pattern of the rock mass deformation modulus contour that are higher in cluster 1 (see Table 4.9).

As expected, for model 3 (100 realisations average) the differences registered between the horizontal displacements were residual. The same tendency as the previous models was followed for the maximum vertical displacements happening in the centre of the tunnel invert with a value of 9.48 mm. In similar grounds, model 4 shows similar results presenting horizontal displacements on the tunnels walls with the same value, 1.79 mm. For the vertical displacement, the maximum value

was of 9.36 mm on the tunnel invert while in the tunnel arch the registered value was 5.86 mm (see Table 4.9).

Note that all the previously presented values were obtained after the tunnel excavation and in the same section of the tunnel, with an y equal to 0. Figure 4.12 shows the displacements contour in all the three models for comparison. In the figure, it is possible to observe that the overall displacements are considerably small due to the fact that the rock mass shows a good geomechanical quality, resulting in high E_m values.

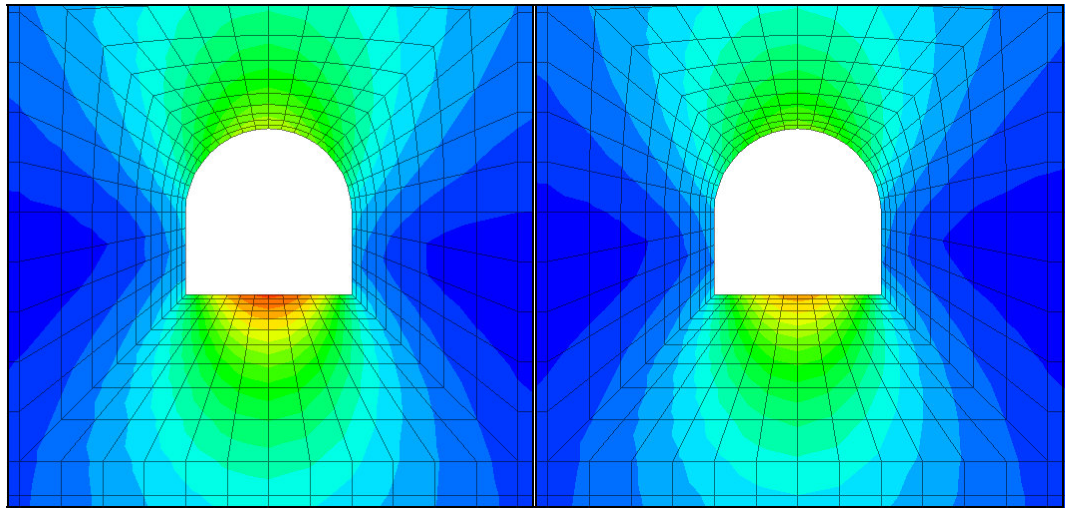
In terms of stress distribution, and based on the maximum and minimum principal stresses, it is possible to conclude that the rock mass surrounding the tunnel is predominantly in compression for all the three models. In model 1 the tunnel walls are the most compressed zones, with values of 6.0 MPa and 1.8 MPa for the maximum and minimum principal stresses, respectively. Moreover, the tunnel arch and invert show compression stresses ranging between 4.0 MPa and 0.4 MPa. Also, a value of 7.0 MPa is obtained in the connection between the arch and the tunnel walls, a small zone where the compression stresses are concentrated.

In what concerns model 2, cluster 1, the most compressed zones are the same as referred before and with the same magnitude of values for the maximum principal stress and minimum principal stress values for the tunnel walls, while the arch and invert show values ranging between 2.0 MPa and 0.5 MPa. Likewise, for cluster 2 of model 2, the range of the stress values show slight difference from cluster 1, presenting smaller values, mostly in the tunnel arch and invert (see Table 4.9).

Similarly, the stresses of model 3 show the same range of values as model 1. Indeed, this similarity in the stress values can be justified by the lower influence of the E_m variation on the rock mass principal stresses. As expected, the stresses are smaller on the tunnel invert and arch where the maximum displacements were observed in all the 104 models.

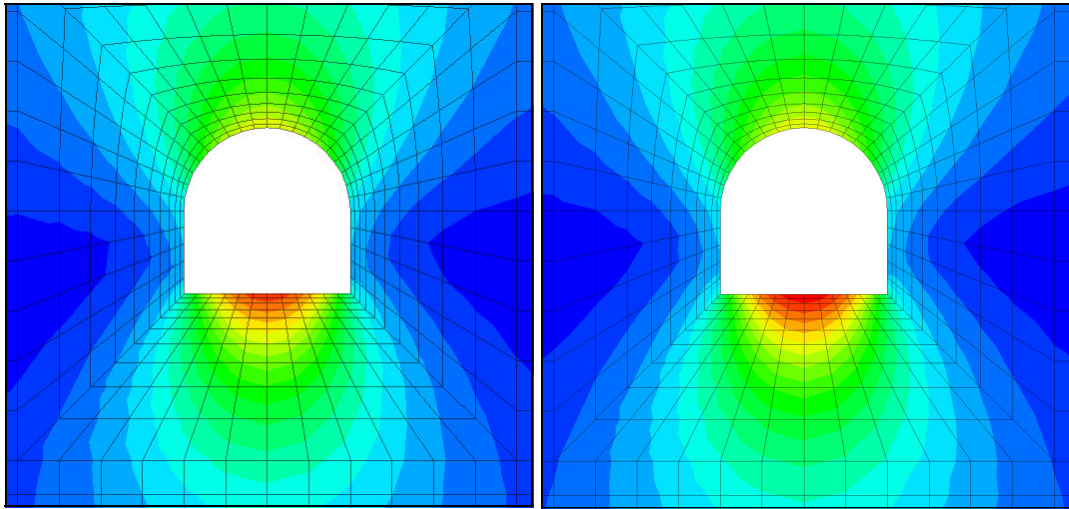
In addition, in all the models a small zone located between the arch and the walls presents compression of almost 7.0 MPa.

Regarding the tensile stresses, they only appear in the centre of the walls and invert with values under 0.2 MPa. Figure 4.13 presents the contour for the maximum principal stress of the rock mass after the excavation process for the same section as previously presented for the displacements contour of model 1, 2, 3 and 4.



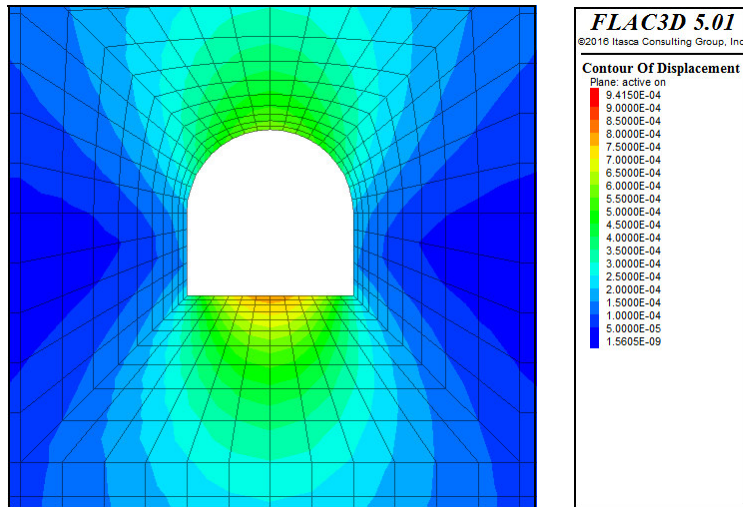
(a)

(b)



(c)

(d)



(e)

Figure 4.12 XZ plane at $y=0$ with $Flac^{3D}$ contour of displacement (colour scale in m) of the rock mass after the excavation for: a) model 1; b) model 2 – cluster 1; c) model 2 – cluster 2; d) model 3; and e) model 4.

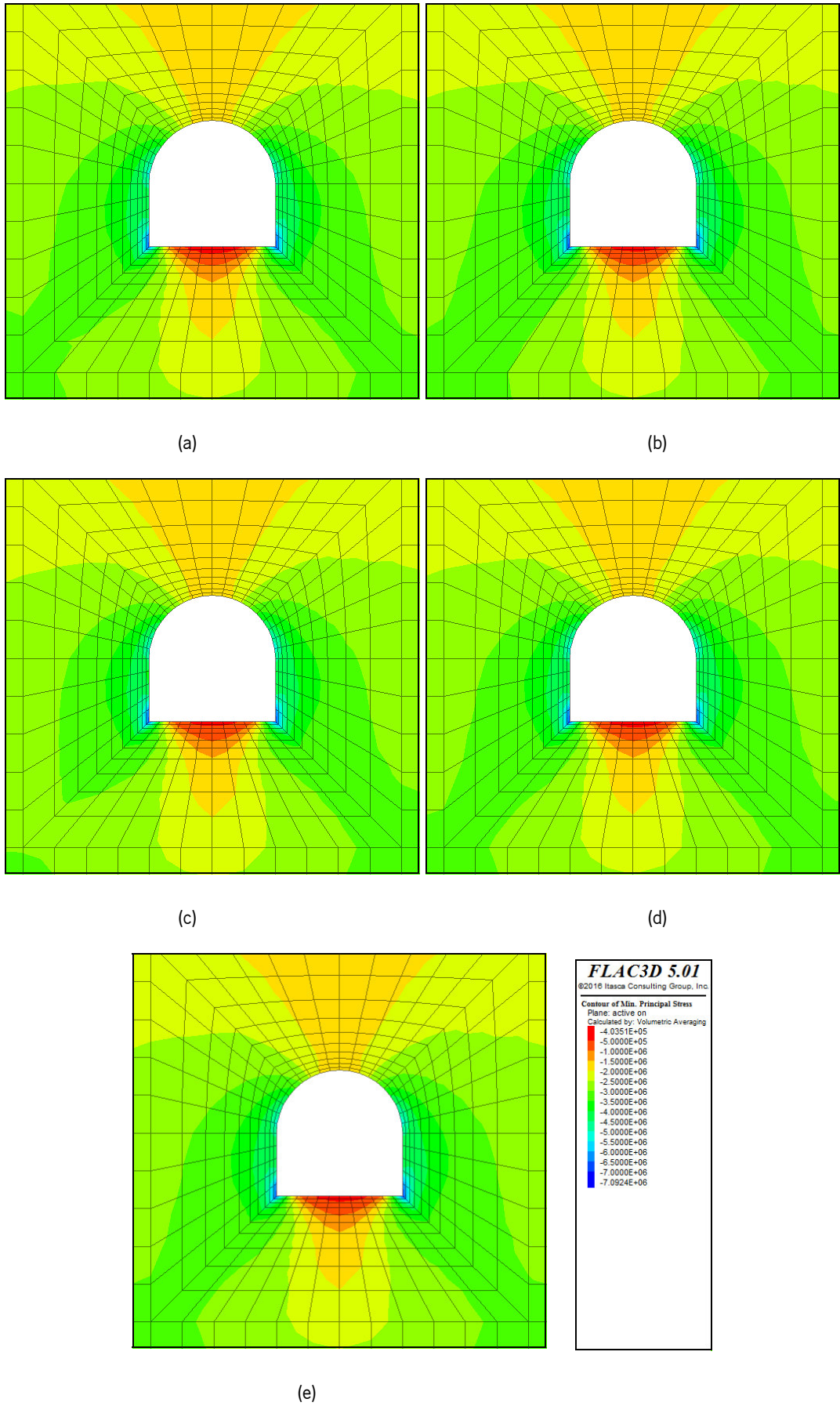


Figure 4.13 XZ plane at $y=0$ with Flac^{3D} contour of the principal maximum stresses (compression- colour scale in Pa) after the excavation process for: a) model 1; b) model 2 – clusters 1; c) model 2 – cluster 2; d) model 3; and e) model 4.

In order to understand the full data set representation using the clusters, an individual analysis was completed for each displacement and stress value. The aim was to represent and identify the percentiles, maximum and minimum values of all the analysed models including the mean and homogeneous models.

Therefore, in Figure 4.14 to Figure 4.16 all the models' values for the displacements and principal stresses obtained after the tunnel excavation are represented. Observing the mentioned figures, it is possible to draw some conclusions regarding the clusters' models and compare them. As such, in the tunnel displacements graphics it is possible to conclude that cluster 1 and cluster 2 from model 2 always result in values limited by percentile 10 and 90, respectively, with exception of the right sidewall displacements, where the clusters show values closer to percentile 50 (see Figure 4.15b). Concerning model 1, the chosen cluster is most of the times near percentile 50 but the same does not happen in the tunnel invert displacements, with model 1 getting closer to percentile 10. Comparing models 3 and 4, except for the tunnel invert displacements where both models show similar values, in the remaining control points both models are quite different from each other.

Regarding the principal stress results, since the values for the maximum principal stress are less disperse, model 2 clusters show similar values, closer to percentile 10. In the case of the minimum principal stress this does not happen and both clusters are quite distant from each other. In this last case, cluster 1 that has been showing the lowest values for the maximum principal stress and displacements, show here the highest value, switching place with cluster 2.

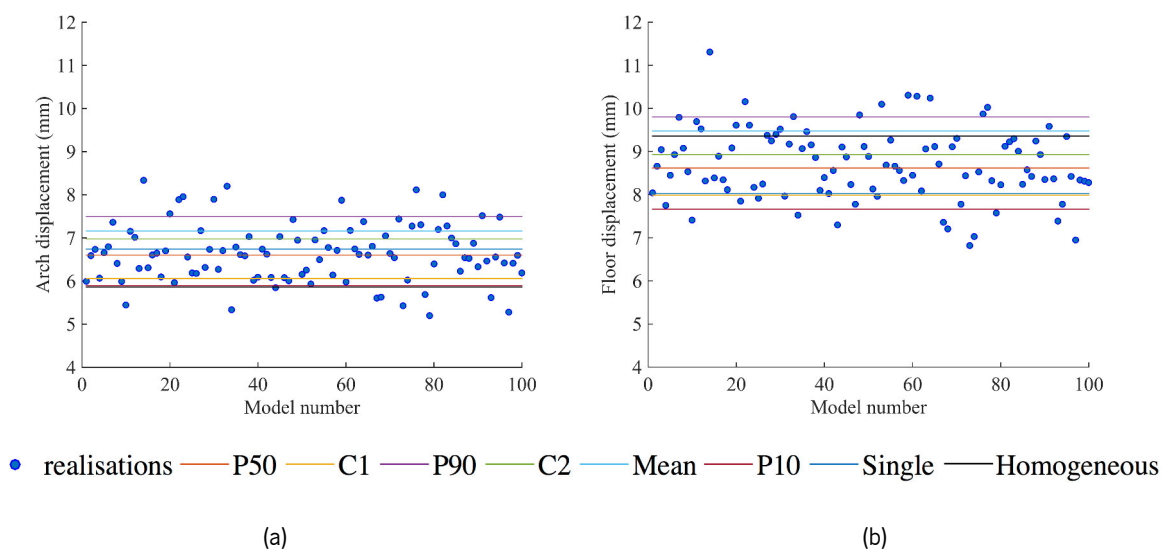


Figure 4.14 Representation of the displacements (in mm) obtained for all the analysed models (104), the single cluster of model 1 (Single), cluster 1 (C1) and 2 (C2) of model 2, the deterministic model 4 (Homogeneous) and the 100 realisations, in: a) centre of the tunnel arch; b) centre of the tunnel invert.

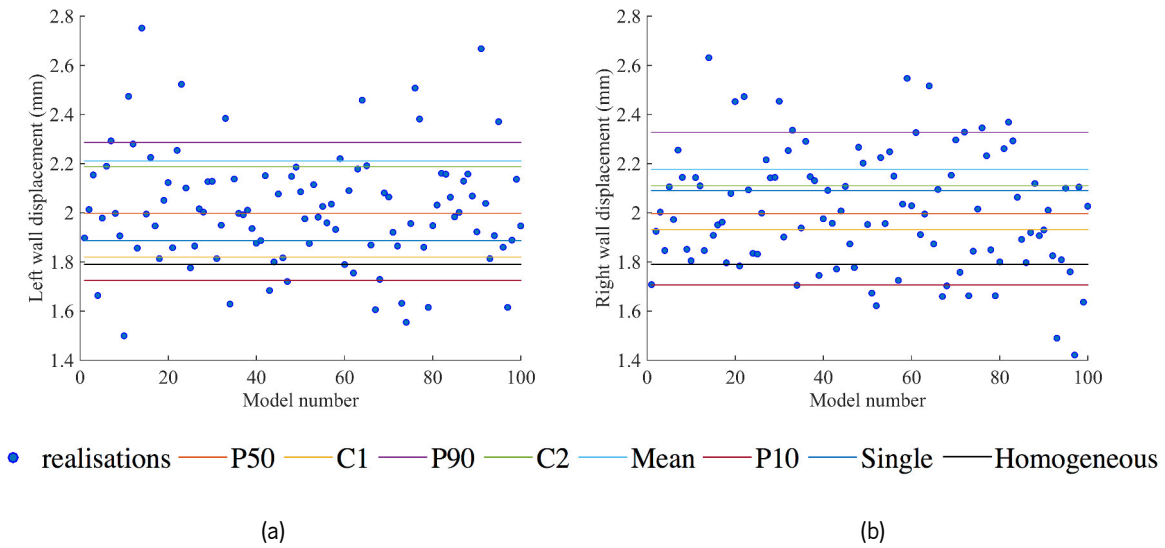


Figure 4.15 Representation of the displacements (in mm) obtained for all the analysed models (104), the single cluster of model 1 (Single), cluster 1 (C1) and 2 (C2) of model 2, the deterministic model 4 (Homogeneous) and the 100 realisations, in: a) centre of left sidewall; b) centre of the right sidewall.

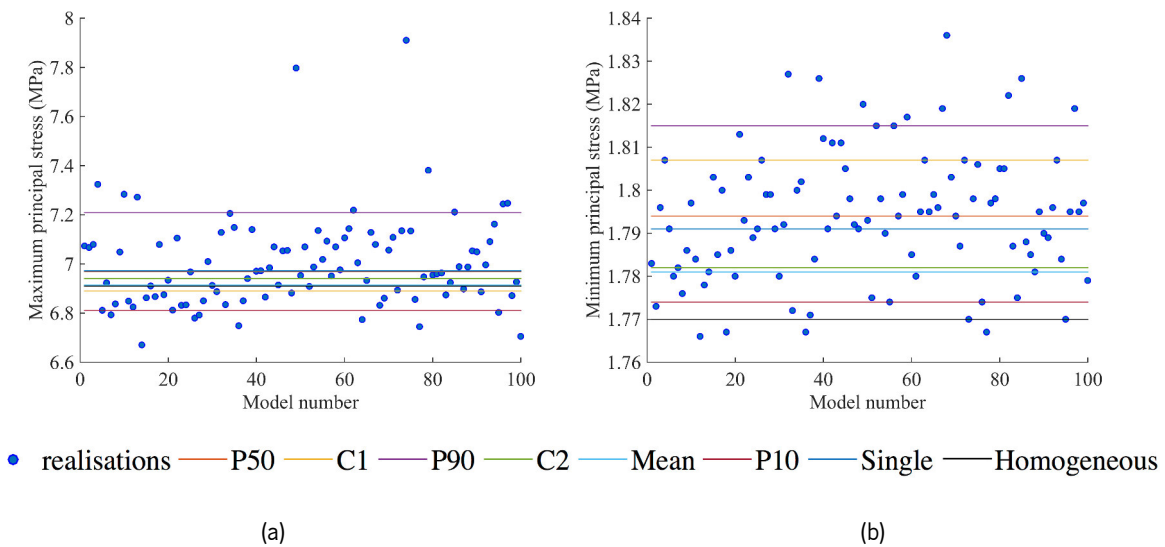


Figure 4.16 Representation of the principal stresses (in MPa) obtained for all the analysed models (104), the single cluster of model 1 (Single), cluster 1 (C1) and 2 (C2) of model 2, the deterministic model 4 (Homogeneous) and the 100 realisations: a) maximum principal stress; b) minimum principal stress.

4.3.8. Discussion

In this section an exhaustive analysis will be made comparing the obtained values for the displacements and principal stresses of this tunnel case, aiming to address three main points: 1) the heterogeneity representation of the individual realisations when compared with a traditional approach that assumes the rock mass as an homogeneous medium; 2) the advantage in using the scenario reduction methodology in comparison with the realisations average or even the 100 realisations; and 3) comparing the results, in terms of values and computational time, of using the

scenario reduction optimal cluster configuration with respect to all the 100 realisations, *i.e.* giving an answer to the question if the optimal clusters configuration is able to integrate the most conservative and optimistic values and embody a good solution to integrate in the numerical analysis.

Regarding the first point, it was possible to observe from Figure 4.12 that the heterogeneity of the rock mass is manifested by the asymmetrical pattern of the contour of displacements in models 1 and 2. Indeed, the disparity of the values of the horizontal displacements came to confirm this statement. In relation to model 4, models 1 and 2 show percentage differences in displacements values ranging from 18% to 2%, being either positive or negative. In detail, these differences can be consulted in Table 4.10 where the presented percentages were obtained using model 4 (homogeneous model) as a reference. In terms of principal stresses the maximum difference between the heterogeneities models and the homogeneous model was of only 2%.

It seems important to mention that, even if small, the differences between a heterogeneous and a homogeneous model can be relevant in geotechnical analysis of underground works; in fact, these differences should be higher as the rock mass spatial variability increases.

Table 4.10 Percentage differences between the displacements values obtained for model 1 and models 2 in relation to model 4 (homogeneous model).

	Model 1	Model 2 – C1	Model 2 – C2
Maximum horizontal displacement (mm) – left sidewall	2%	18%	6%
Maximum horizontal displacement (mm) – right side wall	7%	15%	-5%
Maximum vertical displacement - centre of the arch (mm)	3%	16%	2%
Maximum vertical displacement- centre of the invert (mm)	-17%	-5%	-16%

In what concerns point number 2, as already mentioned by calculating the 100 realisations average, the representation of the heterogeneities is smoothed, which in this case results in equal or higher values for the displacements in comparison with the individual realisations. This proves that the heterogeneities models, namely model 2 by considering two clusters, provides a range of values for the displacements and stress resulting in the reduction of uncertainty and, consequently, in a more optimised design phase for geotechnical works.

To perform the analysis regarding point number 3, histograms containing all the models results had to be computed (see Figure 4.17 to Figure 4.19). Like the graphical representation of the displacements and principal stress value, using the histograms it is possible to confirm the same tendency observed in model 2, where cluster 1 shows values closer to percentile 10 and cluster 2

closer to percentile 90, which allows assuming a more conservative and optimist scenario, respectively. Also, it is possible to state that a big part of the 100 realisations show values that are located between the referred clusters, confirming the capability of the scenario reduction methodology to represent, statistically, the full set of realisations. To complete this analysis in Table 4.11 can be consulted the first three moments (mean, standard deviation and skewness) obtained from the 100 realisations distribution fitting. The skewness value should be close to zero to translate a symmetrical distribution and, for all the zones analysis the displacements show a symmetrical distribution with some tendency to the right side, that is higher displacements.

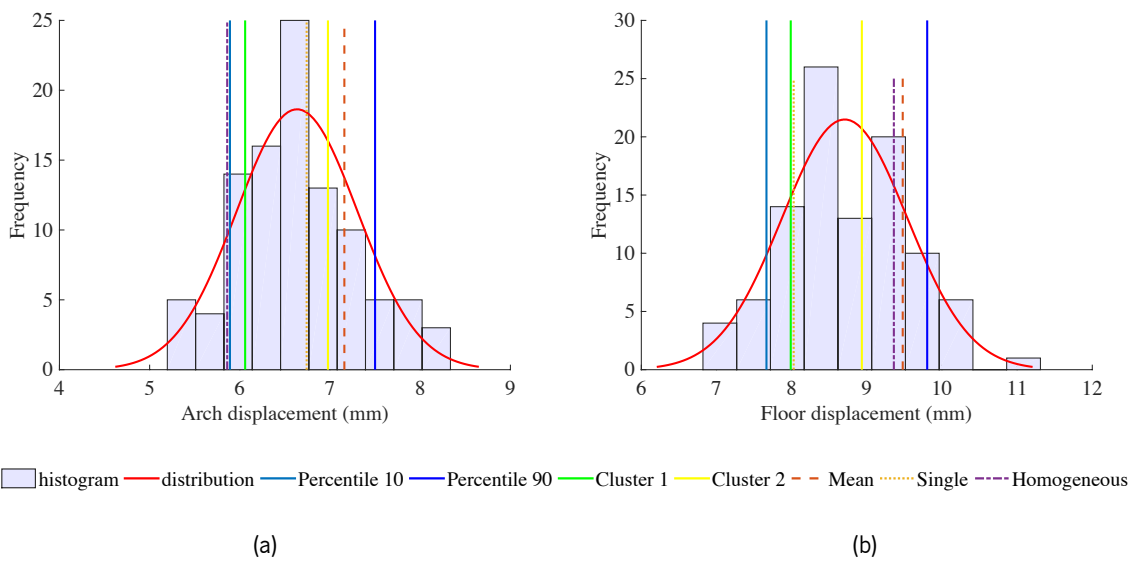


Figure 4.17 100 realisations histograms and distribution fitting curve and models 1 to 5 values (lines) of displacements in: a) centre of the tunnel arch; b) centre of the tunnel invert.

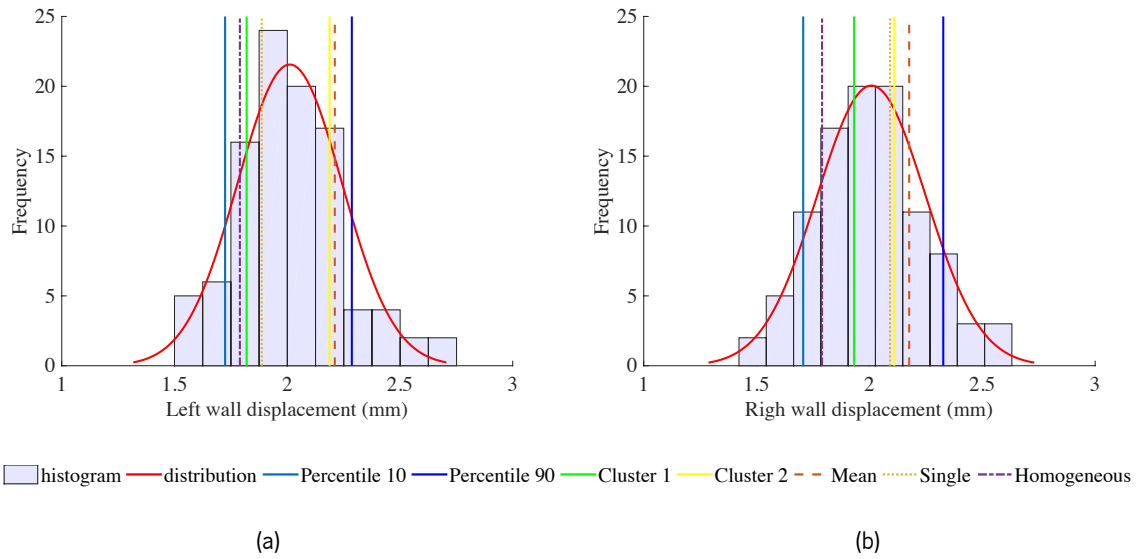


Figure 4.18 100 realisations histograms and distribution fitting curve and models 1 to 5 values (lines) of displacements in: a) centre of left sidewall; b) centre of the right sidewall.

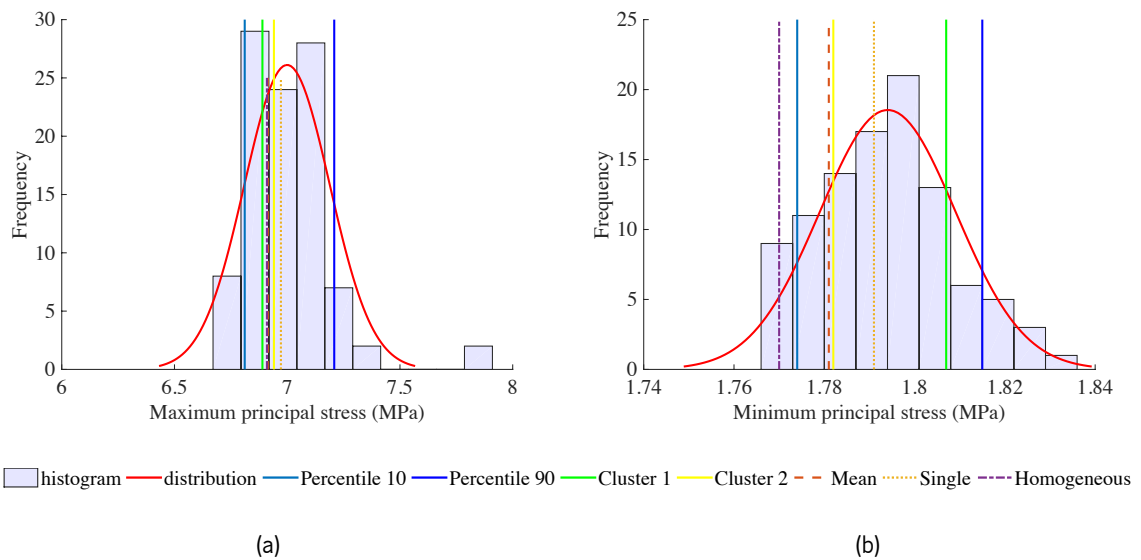


Figure 4.19 100 realisations histograms and distribution fitting curve and models 1 to 5 values (lines) of the principal stresses (in MPa) for: a) maximum principal stress; b) minimum principal stress.

Table 4.11 Distribution curve fitting details (first three moments) regarding the 100 realisations obtained value for displacements and principal stresses.

	Mean	Standard deviation	Skewness
Maximum horizontal displacement (mm) – left sidewall	2.01	0.67	0.091
Maximum horizontal displacement (mm) – right side wall	2.00	0.24	0.100
Maximum vertical displacement - centre of the arch (mm)	6.63	0.67	0.006
Maximum vertical displacement - centre of the invert (mm)	8.71	0.83	0.003
Maximum principal stress (MPa)	6.99	0.19	0.013
Minimum principal stress (MPa)	1.79	0.02	1.07e-6

4.4. CONCLUSIONS

In the context of underground works, the use of more complete and detailed rock mass models is more and more fundamental, especially in the design and construction phases, where a balance between costs and safety must be found. The deterministic approaches used to model the rock masses are somehow limited since they assume the rock mass as a homogeneous medium, instead of considering the spatial variability and heterogeneities that exist. The use of probabilistic techniques has been a step forward in reducing the characterisation uncertainty and spatial variability of the geomechanical parameters, but leaving apart the identification of the heterogeneities intrinsic to rock masses.

Therefore, geostatistics was the key technique identified to be used in the characterisation methodology developed within the scope of this thesis. The methodology aims to reduce the uncertainty associated with geomechanical parameters, quantify their spatial variability and, more importantly, identify the heterogeneities present in the rock masses.

In this Chapter, the main steps, inputs and techniques required to apply the mentioned methodology were detailed. The methodology was outlined to combine the geostatistical simulation and the finite difference software Flac^{3D}. This type of simulation allows the use of geotechnical information commonly obtained from field survey and honour its location and value. A less relevant step consists in using a statistical method to convert the geotechnical data into geomechanical parameters, namely the deformation modulus of the rock mass (E_m), decreasing the uncertainty linked to the use of empirical formulas.

The interface between both techniques is made in a zone centroid level, meaning that it is necessary to firstly create the mesh in Flac^{3D} and then perform the geostatistical simulation of the geotechnical information to obtain the parameters values at the zones centroids. Then, and because there is a smoothness problem in using the 100 realisations average values, a scenario reduction methodology had to be developed to overcome the problem. In this step, the 100 realisations set is statistically represented by a more reduced set of realisations obtained using a clustering technique. Finally, and after a clusters validation, the clusters medoids (individual realisations) are imported to Flac^{3D} and a numerical analysis is performed.

To validate all the numerical methodology mentioned steps a case study composed of real data obtained from an epithermal gold deposit, located in Chile, was used. The data set contains information about the empirical classification systems RMR and values of the UCS. In order to validate the numerical part of the methodology using these real data, a theoretical tunnel had to be idealised and constructed. The analysis was carried out using: i) the optimal and second best clusters configurations obtained from the scenario reduction methodology, ii) the 100 realisations average, iii) a homogeneous approach using deterministic values for the geomechanical parameters, and iv) each individual realisation obtained from the geostatistical simulation. The goal was to compare the results in terms of displacements and principal stresses between all the assumed models and to conclude on the validity of applying the scenario reduction methodology and difference between the heterogeneity and homogeneous models (models 1 and 2 *versus* model 4). For both points the results had been proving encouraging, initially having the two clusters from model 2 able to provide a range of values, which can represent at the same time, a conservative and an optimistic reality of the rock mass condition, and then confirming the limitations of the traditional approach that, for this case, had most of the times resulted in higher displacements and, consequently higher costs for the work.

Therefore, the findings of this Chapter show that, for rock masses with more pronounced heterogeneities/variability, the proposed methodology tends to show better results and their difference should increase significantly and that when the variability of the rock mass parameters is small their influence in the rock mass mechanical behaviour is also small and the rock mass can be modelled using average values. However, since the influence of the geomechanical parameters in the mechanical behaviour of rock masses in underground works is unquestionable, this methodology should provide distinctive results for rock mass with higher levels of variability/heterogeneity. This fact can bring numerous advantages to the geotechnical community once the characterisation models of the rock masses will become more accurate and closer to reality. In this sense, it is

important to highlight the short time required to apply this methodology, which only increases in the geostatistical simulation compared to the traditional approach.

Chapter 5

SALAMONDE II POWERHOUSE COMPLEX GEOMECHANICAL CHARACTERISATION AND MODELLING

5.1. INTRODUCTION

The construction of the Salamonde Dam located in the North of Portugal, namely in the Cávado river was finished in 1953. This complex structure resulted in an arch dam with a height of 75.0 m totalling a reservoir area of 2.4 km². In the downstream of the Salamonde dam, which is situated at an elevation of approximately 270.4 m, another dam was constructed (at an elevation equal to 152.5 m), named Caniçada. Thereby, due to an existing gap in height between the two dams of approximately 120.0 m, the Portuguese electrical company, Electricity of Portugal (EDP) took advantage of this gap by promoting the reinforcement of the first electrical power of Salamonde. As a result, an almost parallel power station was constructed giving rise to a new hydroelectric complex named afterward as Salamonde II. (see Figure 5.1).

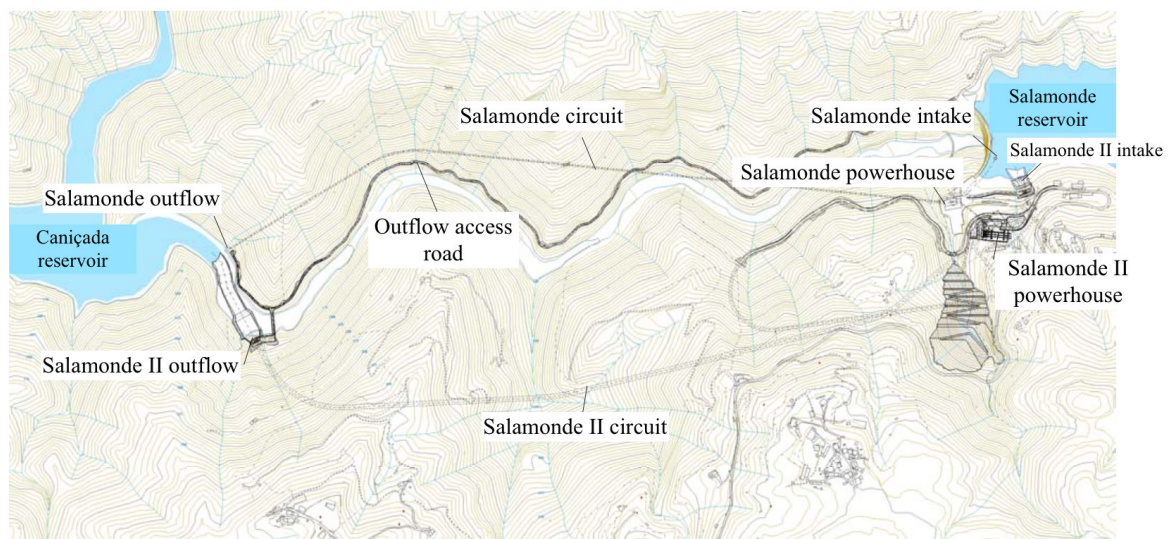


Figure 5.1 General overview of the Salamonde II hydroelectric circuit implementation (adapted from EDP (2009)).

Considering that the Salamonde dam is located closely to the National Park of Peneda-Geres (situated in one side of the river), the construction for this new electrical power station, Salamonde II, was limited to the left side of the existing dam. At this location, the rock mass under consideration was a typical granite from the region.

The new powerhouse of Salamonde II entails the construction of complexes underground elements, like caverns and access tunnels. More precisely, the geotechnical underground structures comprise (see Figure 5.2):

- The central containing the power complex, built as a cavern at a depth of 200.0 m and a high that varies from 27.5 m to 67.5 m extended by 65.7 m;
- The access tunnel to the powerhouse with an extension of 985.0 m and a diameter of 8.2 m;
- The tailrace tunnel with an extension of 1909.0 m with a nominal diameter of 11.8 m;
- The headrace tunnel with a total extension of approx. 200.0 m composed by a vertical segment of 130.0 m and nominal diameter of 8.3 m followed by a 70.0 m horizontal segment and with 7.8 m as nominal diameter that finishes in the powerhouse cavern;
- The hydraulic circuit with 8.3 m of diameter with an extension of 200.0 m.

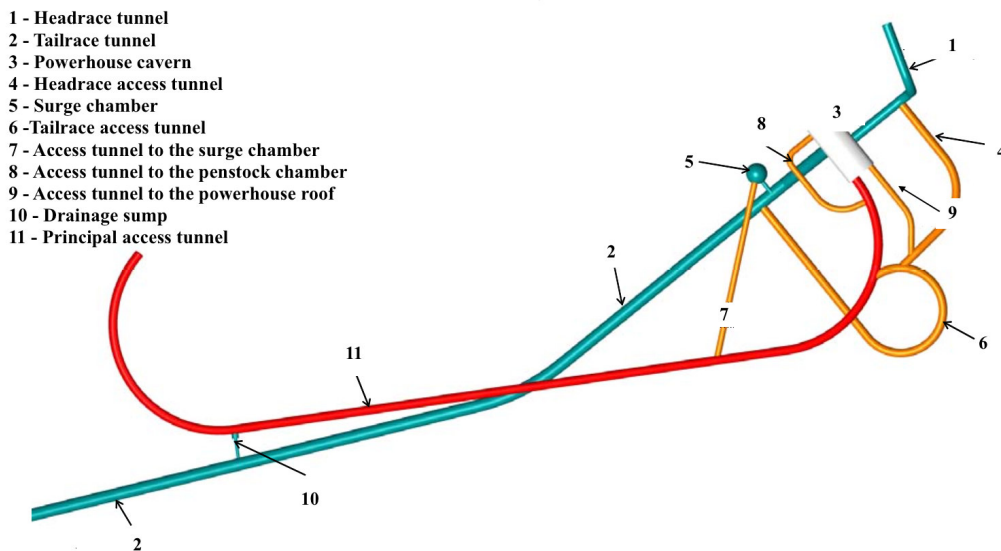


Figure 5.2 Salamonde II underground structures detailed scheme (adapted from EDP (2009)).

Among all the underground structures previously described, the study will focus only on the powerhouse of Salamonde II. As such, the geotechnical data surrounding the powerhouse was taken into analysis and further used to perform a geostatistical conditional simulation as a new method of rock mass characterisation and modelling. To do that, all the steps described in Chapter 3 were considered, simply making some small adjustments, mainly in what concerns the type of variable chosen for the simulation. The collected data comprises the information about the empirical systems used to classify the rock masses encountered on the project, as well as all the *in situ* and laboratory tests performed to complement the rock mass characterisation. Even with different information sources, from there, the geomechanical parameter, namely the deformation modulus (E_m) of the granite formation was calculated using empirical formulas found in the literature. Though, to reduce the uncertainty associated with the use of only one empirical formulas, different data sets of E_m

were created. Subsequently, a 3D numerical model was built, using the finite differences software, Flac^{3D} (Itasca, 2012) and the geomechanical information was imported for further analysis. In essence, this model aims to describe and analyse the deformational behaviour of the rock mass during the excavation of the powerhouse cavern. All the excavation phases were simulated in the numerical model to allow a more detailed analysis during the excavation process. To show the advantages in using the numerical methodology proposed in Chapter 4, the obtained results were compared with the ones obtained by assuming a deterministic, *i.e.* the rock mass as a homogeneous mean. Also, and because Salamonde II is, at this stage, almost active, the data obtained from monitoring campaigns was also used in this validation process.

In detail, in this Chapter the following sections will contain, firstly the information about the geotechnical data available for this study, as well as all the statistical treatment adopted for the data. Then, some relevant aspects regarding the powerhouse cavern geometry and auxiliary structures will be given. In a third section, and using the geotechnical information previously gathered, a conditional simulation of the geomechanical parameters are performed and the details adequately exposed. The following sections are then reserved for the final application of the numerical methodology proposed within the scope of this thesis. Ultimately, an emphasis is given to the numerical results and a comparison is made with the deterministic approach results and real data obtained from monitoring. In the last section, some conclusions are drawn and statements are made regarding the utility and importance of this methodology to help in the planning and construction of underground works.

5.2. GEOTECHNICAL PROSPECTION AND TESTS

Bearing in mind Salamonde II tender phase, EDP had performed a total of 15 mechanical boreholes (S1, S2, S4 to S16) to characterise the area. Boreholes S3 was not executed due to adverse weather conditions. Regarding the area lithology, as already mentioned, a granite from the region was the most abundant rock type. It was formed after the third hercynian orogeny, which corresponds to a recent granite, namely post tectonic. In brief, this rock mass is composed by porphyrioids of medium to coarse grains.

Regarding the 15 boreholes, they show a diameter of 98 mm in the zones with residual soils and 78 mm in zones of rock mass. To characterise the rock mass, the weathering and fracturing degrees as defined by the International Society of Rock Mechanics (ISRM) were considered. Furthermore, to

classify the quality of the rock mass the Rock Quality Designation system (Deere, 1968) was used. Those three geotechnical variables were obtained for all the boreholes whose depth varies from 30 to 200 m (see Appendix 1 for more details about the geotechnical information).

In engineering projects with this dimension different detail levels of information are normally required. Therefore, in order to correctly characterise the rock formation interesting in Salamonde II, laboratory and *in situ* tests were performed in some boreholes to assess the rock mass deformation and strength properties. These tests were carried out by the National Laboratory of Civil Engineering (LNEC) in two distinct phases of the hydroelectric reinforcement project. Firstly, in 2009 corresponding to the beginning of the design phase, six STT tests were performed in two boreholes with the goal to perform a preliminary analysis of the rock mass stress field (LNEC, 2009). Later in 2012, and already in a construction phase where the areas of interest could already be accessed, ten small flat jacks (SFJ) were performed in three different areas of the rock mass (LNEC, 2012). The tests information was used to estimate the initial stress field of the rock mass near the powerhouse cavern. Figure 5.3 shows the places where the mentioned tests were performed for both campaigns. In this section, the tests results are summarily presented, firstly for the ones performed in 2009 and secondly in 2012.

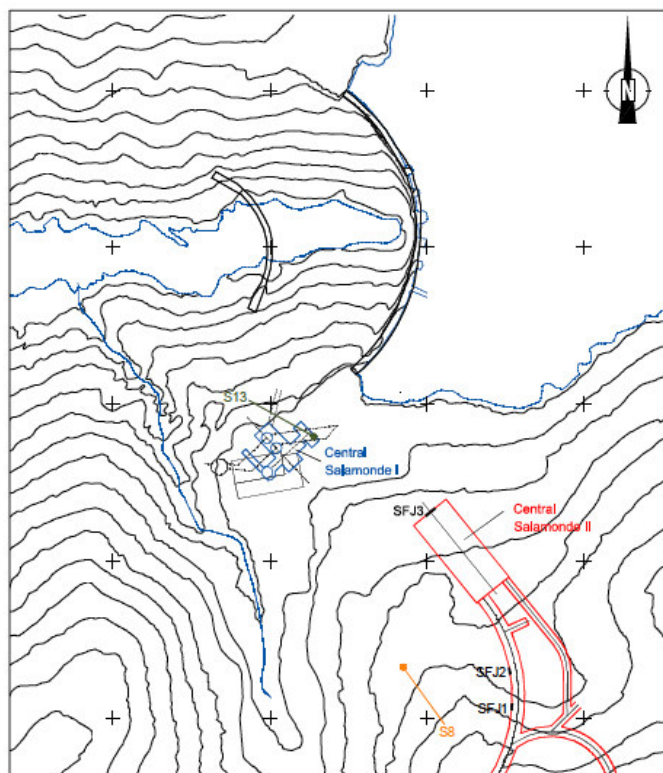


Figure 5.3 Salamonde II powerhouse general plant with the *in situ* tests identification (LNEC, 2012).

In 2009 six STT tests were performed in two boreholes, S8 and S13 located near the powerhouse cavern as shown in Figure 5.3. The first borehole was vertical and the second presented an angle of 70° with the vertical and was executed from the inside of the valve chamber. Table 5.1 presents the obtained results along with the rock mass formation characteristics.

Table 5.1 STT tests results and rock characteristics (adapted from LNEC (2009)).

Boreholes	Depth (m)	Lithological description	Fracturing/Weathering degree	Recovery (%)	RQD
S8	79.33	Porphyry granite (grey colour) of coarse grain, biotitic.	F2/W2	100	100
	93.48		F3-4/W2-3	100	100
	95.98		F3/W2	100	100
S13	28.29	Porphyry granite (grey colour) of coarse grain, biotitic.	F1/W1	100	100
	37.97		F1/W1	100	100
	48.06		Bedrock without fractures	F1/W1	100

To complement the presented tests, additional uniaxial strength tests were performed on boreholes S7 and S8. As a result, an average value of 50 GPa and 0.21 were obtained for the elastic parameters, the elasticity modulus and the Poisson ratio, respectively.

In March of 2012, three SFJ were performed, namely near the left wall of the powerhouse access tunnel (SFJ1 and SFJ2) and in the north part of the powerhouse cavern (SFJ3). In Table 5.2 are displayed the obtained values for all three SFJ tests.

Table 5.2 SFJ tests results (adapted from LNEC (2012)).

Code	Number of slots	Slot	Termination stress (MPa)	Azimuth/Stress slant (°)	E_m (GPa)
SFJ1	2	Horizontal	17.0	0/90	39.7
		135°	13.8	175/45	39.0
SFJ2	3	Vertical	6.5	350/0	47.9
		45°	7.8	350/45	43.5
		Horizontal	12.1	0/90	31.9
SFJ3	5	Horizontal	6.2	0/90	28.8
		135°	9.4	225/45	26.4
		Vertical	6.3	45/0	29.6
		45°	6.8	45/45	31.6
		Vertical	5.3	45/0	33.2

This geotechnical information was used by LNEC to assess the initial stress of the rock mass through the construction of a 2D Flac^{3D} model. In this model, a vertical section perpendicular to the

cavern axis with 210 m width, 350 m large and a range of value between 193 m and 350 m on the vertical direction was adopted. For the rock mass an elastic behaviour was assumed and, contrarily to the value obtained from the uniaxial strength tests, for this stress field test the average value obtained from the SFJ tests was assumed, that is 35 GPa, a Poisson ratio of 0.20 and a unit weight equal to 27 kN/m³.

At this stage, it is important to mention that since all the geotechnical prospection, laboratory and *in situ* tests mentioned above were performed during the design phase, from now, they will be addressed as phase 1 information.

As a common practice to analyse the initial stresses in the normal plane, on the lateral boundaries the perpendicular displacements were blocked, as well as all the displacements in the inferior boundary. On the other hand, to analyse the shear stresses in the horizontal plane, the horizontal displacements and all the displacements were blocked in the lateral boundaries and in the inferior boundary, respectively. To perform this analysis a methodology described in LNEC (2009) had to be developed, from which it was possible to conclude the following:

- Near the surface the minimum and maximum principal stresses were mainly parallel and perpendicular to the Cávado river direction, respectively. Also, a maximum principal stress two times higher than the minimum principal stress was obtained;
- Between the maximum principal stress and longitudinal axis of the powerhouse cavern an angle of 45° is made;
- Regarding the intermediate principal stresses, they were almost vertical and correspond to the weight of the overlaying terrains;
- At the powerhouse cavern depth, a value of approximately 1.30 times the weight of the overlaying terrains was obtained for the maximum principal stress, while the minimum principal stress shows a ratio of 0.60 in relation to the overlaying terrains.

As presented above, the phase information can also be divided into two different data sets. First, the information that directly results from the boreholes analysis was called data set 1 and is composed by 1036 information points resulting from the 9 boreholes that are located near the powerhouse cavern (S2, S5, S6, S7, S8, S12, S14 and S16). This amount of points was obtained after reading the logs of the 15 boreholes from where the geotechnical information was recovered at every 1 m depth. Then, the second data set comprises the results of all the laboratory and *in situ* tests executed in the same 15 boreholes. As a result, was achieved a total of 32 points with common geotechnical information as displayed in Table 5.3.

Table 5.3 List of the geotechnical variables that compose the data set 1 and data set 2 of phase 1 information.

Set 1	Set 2
<ul style="list-style-type: none"> ▪ Rock Quality Designation (RQD) ▪ Fracture degree (F) ▪ Weathering degree (W) 	<ul style="list-style-type: none"> ▪ Weathering degree (W) ▪ Deformation modulus of the intact rock (E_i) ▪ Uniaxial Compressive Strength (σ_c) ▪ Poisson's ratio (ν)

After combining both data sets from phase 1, it was possible to obtain the basic statistics of the geotechnical variable to use in further analysis. These are the deformation modulus of the intact rock (E_i) measured in laboratory, the weathering degree (W) and the Rock Quality Designation system (RQD). Therefore, Table 5.4 shows the variables statistics and Figure 5.4 presents the obtained histograms. Also, and since the other geotechnical variables were mentioned, in Appendix 1 the matching statistics and histograms can be found.

Table 5.4 Statistical analysis of the phase 1 geotechnical information of Salamonde II (E_i , W and RQD).

Variable	Number of samples	Minimum	Maximum	Mean	Variance
E_i (GPa)	81.00	18.10	77.20	56.40	101.07
W	1570.00	1.00	5.00	2.33 ¹	0.87
RQD (%)	1568.00	0.00	100.00	69.64	951.26

¹ The mean value of the weathering degree was computed to give a notion of the averaging value, however, the most correct way to represent this type of variables is computing the histograms.

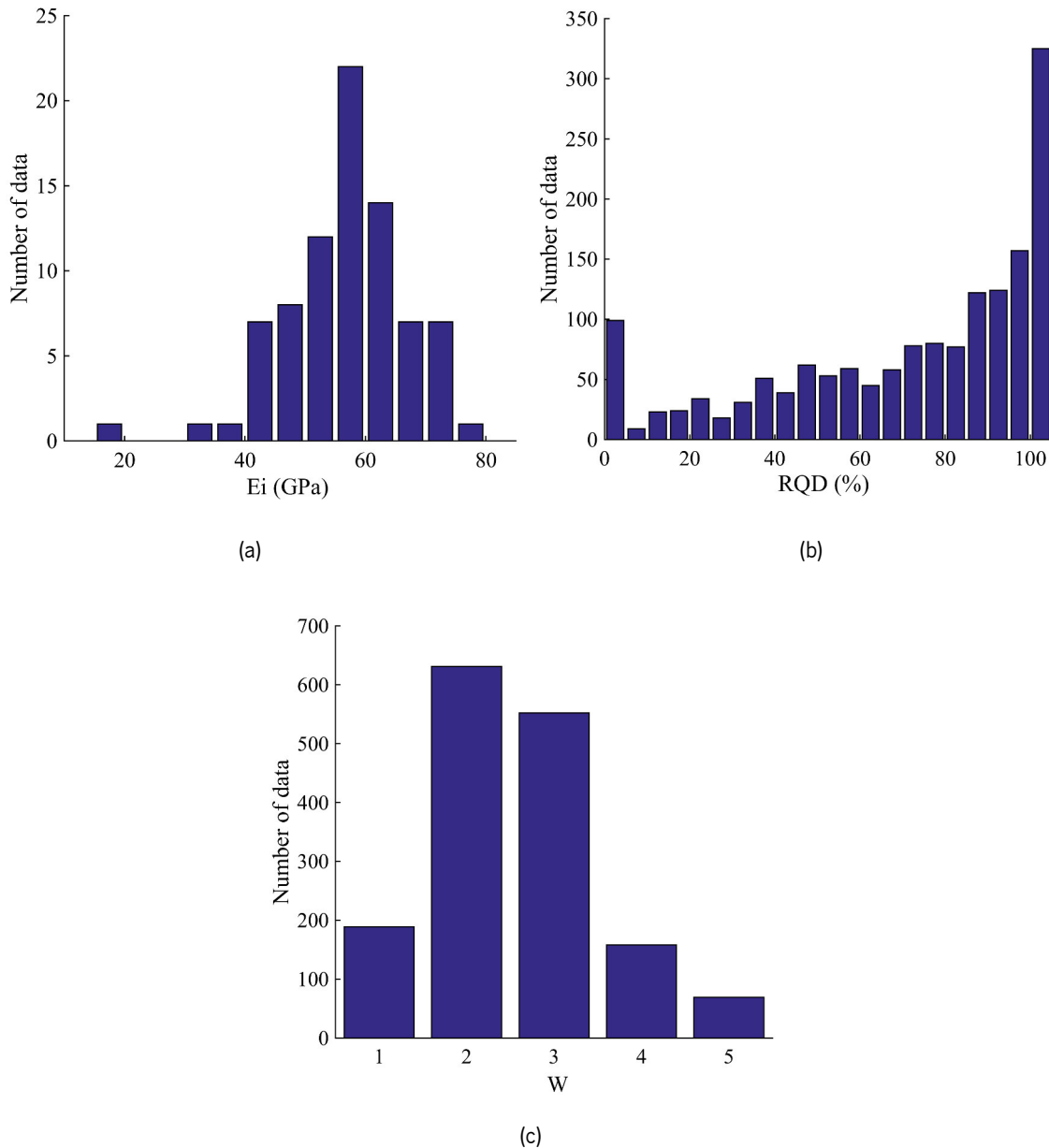


Figure 5.4 Histograms obtained using the phase 1 information of Salamonde II, for: a) E_i ; b) RQD; and c) W.

5.3. THE POWERHOUSE CAVERN

The Salamonde II powerhouse is located in a cavern built in a depth of 200 m in the left side of the Salamonde dam. As already referred, the cavern was built with a vault section with a total high of 27.5 m in the south area and 44.7 m in the north area, presenting a plant area of 65.7 x 26.5 m². The power group was placed in the centre of the cavern, more precisely at an elevation between 115.0 m and 97.0 m. At the north side of the power group there is numerous equipment distributed by six different floors. The access to the cavern is done through a tunnel with a diameter of 8.2 m

and a high of 8.4 m extended by 985.0 m with a slope of 10% (this was contemplated as an auxiliary underground work).

A second and smaller cavern can be found 40 m downstream the main cavern with an area of 7.0 x 15.0 m² and design to allocate the motor responsible for the engaging the downstream group gate. The connection between both caverns is made through a tunnel with length of 130 m and a horseshoe section with 5.5 x 6.0 m² (see Figure 5.5).

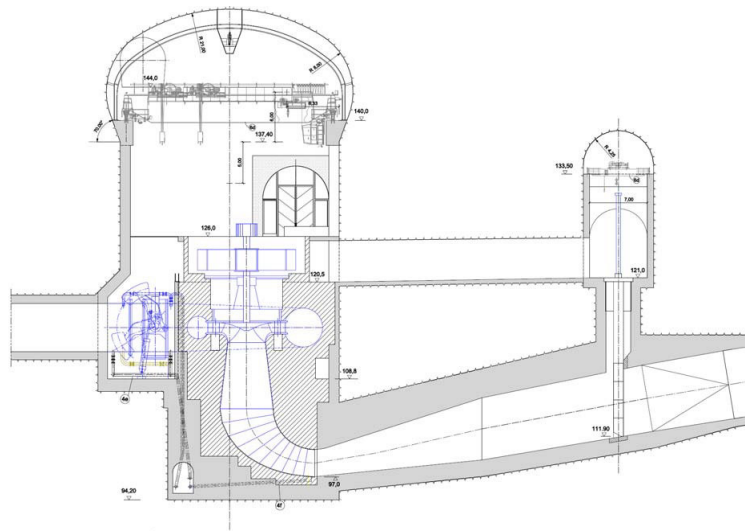


Figure 5.5 Powerhouse complex geometry and elements (EDP, 2009).

In what concerns the geotechnical conditions near the powerhouse complex, four geotechnical zones were defined based on information from phase 1. Closer to the cavern, until 30 m depth the rock mass was classified as varying between a high to medium weathering while the fracturing degree gives a classification from very close to medium fractures. This zone was named as geotechnical zone number 3 and the geotechnical characteristics can be found in Table 5.5. From that depth value, the rock mass starts to show better geotechnical conditions, with a low weathering degree and very distant fractures (geotechnical zone number 1 and 2).

Taking as a reference the overall good conditions of the rock mass, the powerhouse cavern excavation was made using explosives, mostly techniques that result in a low disturbance for the rock mass like smooth blasting and pre-cut in bench dismantling.

Table 5.5 Geotechnical zoning characteristics (EDP, 2009).

Zone	Weathering degree (W)	Fracturing degree (F)	RQD (%)	σ_c (MPa)	γ (kN/m ³)	GSI
ZG1	$\leq W_2$	$\leq F_2$	> 70	100	27.0	70-85
ZG2	W_2 to W_3	F_2 to $F_3^2(**)$		60	26.0	50-70
ZG3	W_3 to W_4	F_3 to $F_4^3(***)$	> 20	45	25.0	30-50
ZG4	W_4 to $W_5^4(*)$	F_4 to F_5	< 20	0.5	24.0	<30

During the powerhouse construction, new geotechnical information was recovered. As so, the entire cavern (excavation front and walls) has been detailed and characterised using three different empirical systems for rocks characterisation, the Rock Mass Rating (RMR), the Q-Barton and the Geological Strength Index (GSI). These mappings were one of the two types of information that composes Salomonde II second phase of geotechnical information (data set 1). In detail, considering that this information was registered in individual documents as the advance step increases, it was necessary to identify the corresponding coordinates in order to reproduce this information in space. To this effect, the kilometre points was used and transformed into X , Y and Z coordinates. In order to characterise the all area and not only the powerhouse cavern, additional mappings were also included, mainly from underground works located near the cavern. More details regarding the method adopted to recover this information from phase 2 can be found in Appendix 1. As a result, a total of 205 points with information about the RMR, the Q-Barton value in a logarithmic scale and GSI were obtained and the systems basic statistics are provided in Table 5.6, along with the empirical systems histograms showed in Figure 5.6.

Table 5.6 Statistics analysis of data set 1 of the phase 2 geotechnical information of Salomonde II (RMR, Log Q and GSI).

	RMR	Log Q	GSI
Total of points with information	205	205	205
Mean	68.93	1.03	67.50
Variance	12.50	0.05	12.50
Standard deviation	3.54	0.23	3.54
Minimum	62.00	0.15	60.00
Maximum	67.00	1.22	65.00

² Sometimes F4

³ Sometimes F4-5

⁴ Sometimes W3

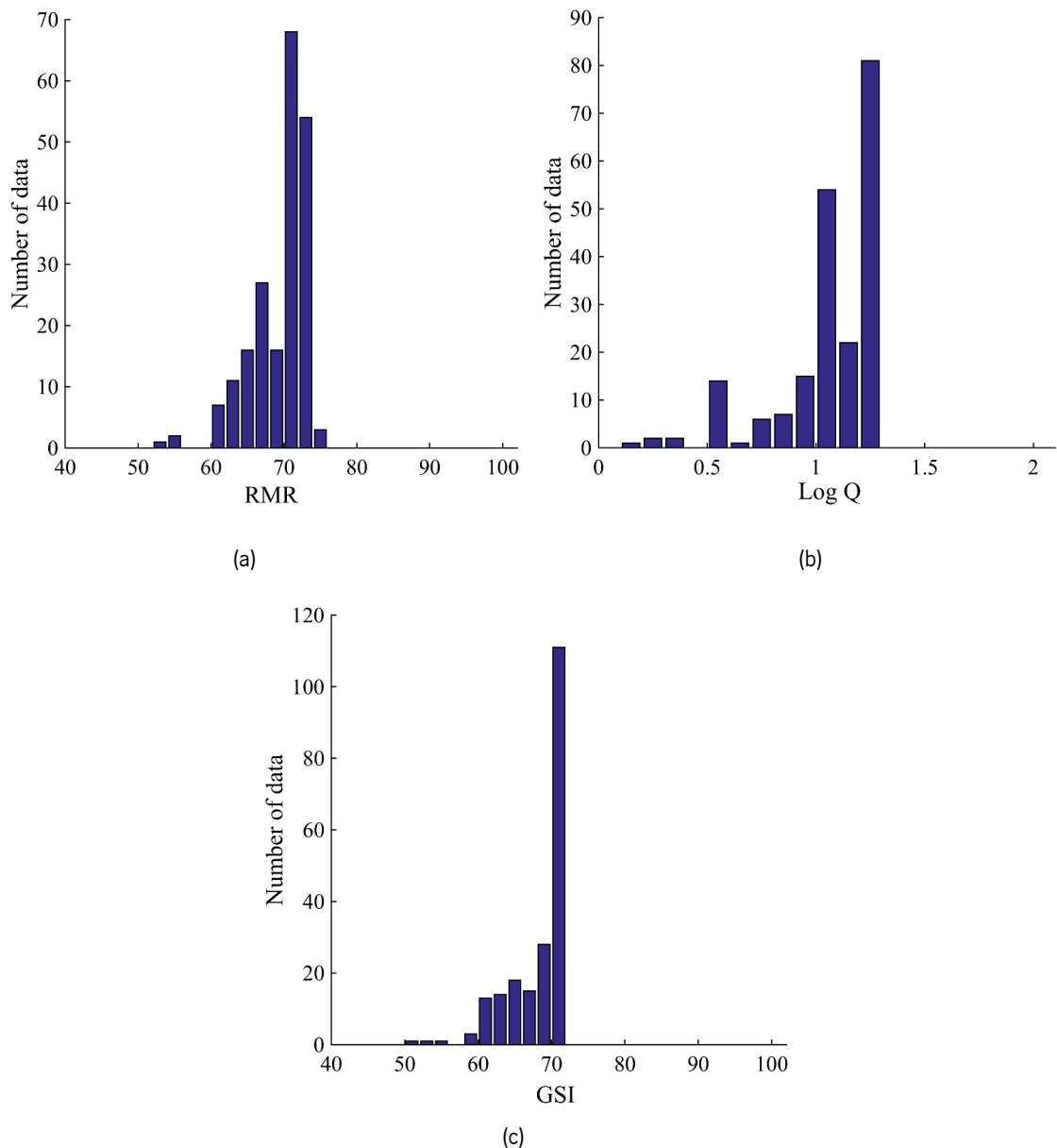


Figure 5.6 Histograms obtained using data set 1 of Salamonde II phase 2 information, for: a) RMR; b) Log Q; and c) GSI.

In addition, in order to evaluate the behaviour of the rock mass and support system during the excavation, a monitoring system using extensometers and piezometers was installed in the cavern. Therefore, some supplementary information was obtained from laboratory tests executed in some piezometers and extensometers holes recovered at the time. These elements were installed in two different sections of the cavern, namely the section A (SA) located at 38.4 m and section B (SB) at 19.2m, both distances were measured from the cavern south top wall. A total of 7 extensometers were installed in section A, three of them in the cavern arch (E1A, E2A and E3A), while the remaining four were placed in the side walls (E4A, E5A, E6 and E7). In section B, the number of extensometers was reduced to five, three placed in the cavern arch (E1B, E2B and E3B) and two installed in the side walls. Also, the number of installed piezometers was set to eleven, all located in

the cavern sidewalls in random sections over the cavern. The mentioned extensometers and piezometers adopted for the monitoring plan of Salomonde II can be seen in Figure 5.7.

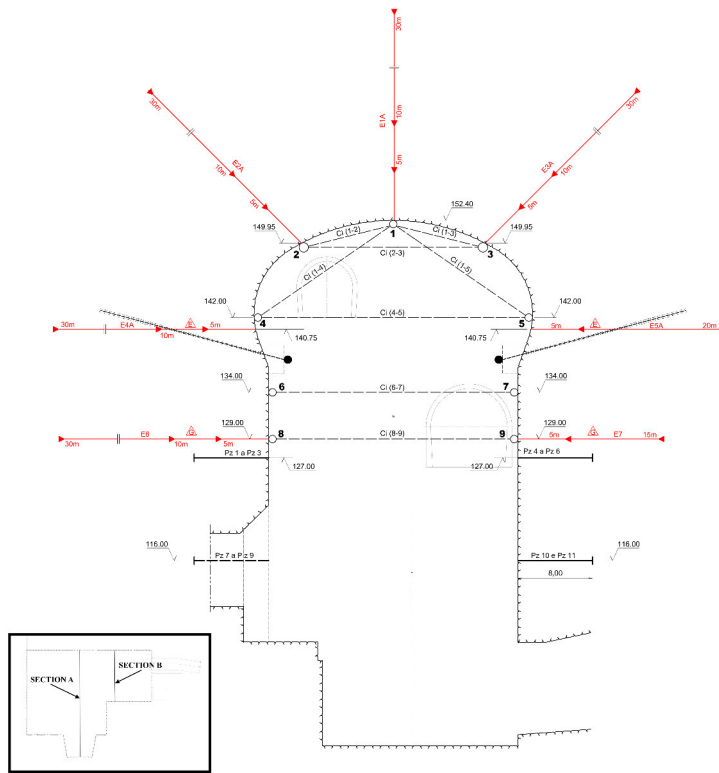


Figure 5.7 Cross section of the powerhouse cavern showing the monitoring plan adopted (adapted from EDP (2009)).

The laboratory tests carried out in some piezometers and extensometers allowed the measurement of the intact rock deformation modulus (E_i) through uniaxial compressive tests. In addition to this information, the weathering and fracturing degree, as well as the RQD system was also recovered and the geotechnical variables statistics are provided in Table 5.7 and the corresponding histograms can be found in Figure 5.19.

Table 5.7 Statistical analysis of data set 2 of the phase 2 geotechnical information of Salomonde II (F, W, RQD and E_i).

	F	W	RQD (%)	E_i (GPa)
Total of points with information	21	21	21	21
Mean	2.10	1.60	90.50	53.20
Variance	0.47	0.14	394.79	369.09
Standard deviation	0.68	0.38	19.87	19.21
Minimum	1.00	1.00	45.00	9.70
Maximum	3.00	3.00	100.00	71.40

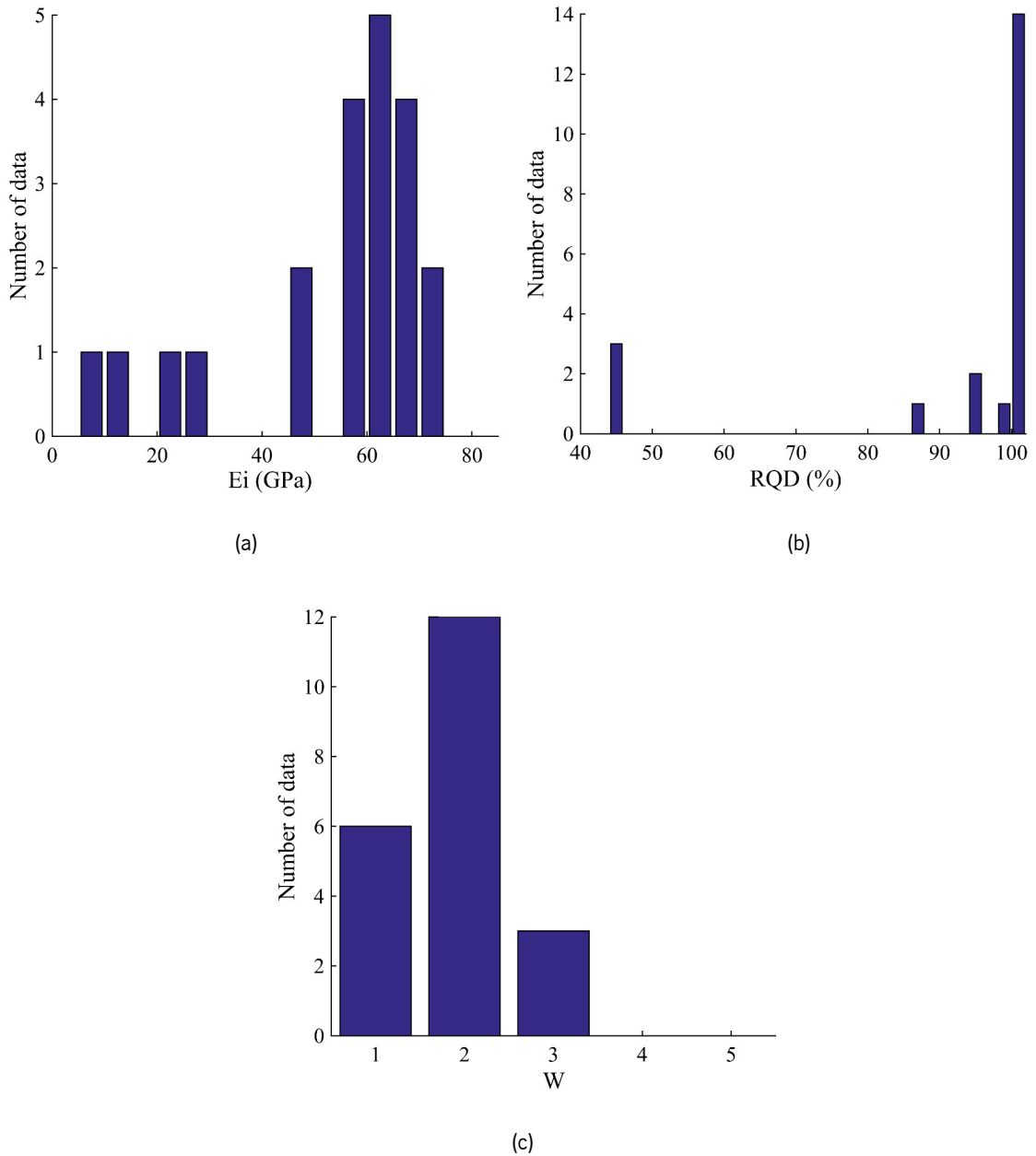


Figure 5.8 Histograms obtained using data set 2 of Salamonde II phase 2 information, for: a) E_i (in GPa); b) RQD (%); and c) W .

5.4. NUMERICAL MODEL

5.4.1. Numerical Model presentation

The numerical model used in this Chapter was developed during the study of the rock mass initial stress field by LNEC in 2009. However, some changes were made, namely the remove of some auxiliary tunnels and the addition of the construction sequence adopted for the powerhouse

complex. The numerical model was carried out using the finite difference software Flac^{3D} (Itasca, 2012) and is composed by a total of 664,022 grid points and 642,086 elements. In addition to the main cavern, also some nearby underground works were added into the model, namely the floodgate chamber (FC), the access tunnel (AFC) between the main cavern (C) and the floodgate chamber and a smaller tunnel design only for pedestrian access (PAT) (see Figure 5.9).

In detail, the Flac^{3D} mesh is 210 m wide along direction x , 350 m along direction y , a range between 303 and 460 m along direction z and is composed by elements, whose size increase as one moves far away from the powerhouse cavern (Espada and Lamas, 2014).

In what concerns the model boundary conditions, the horizontal displacements in the model vertical boundaries were blocked, along with all displacements in the inferior boundary.

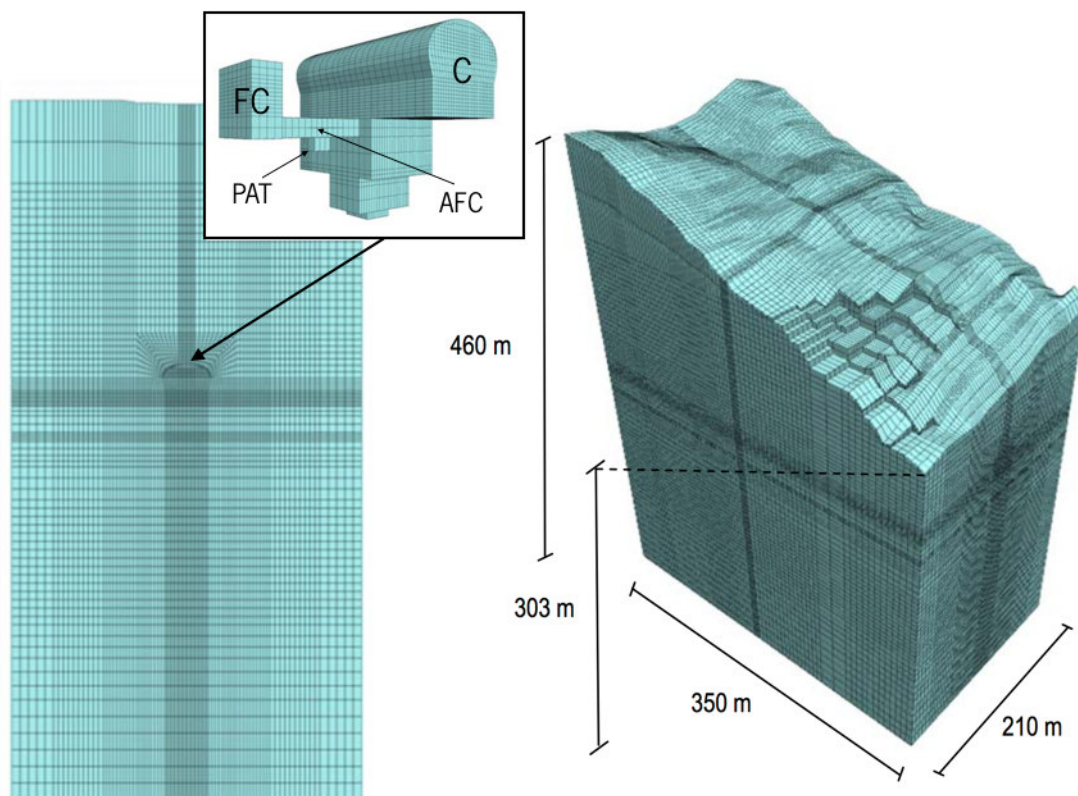


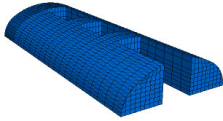
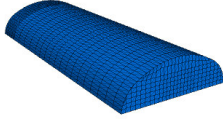
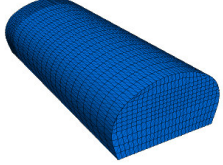
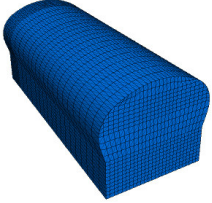
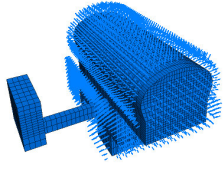
Figure 5.9 3D mesh developed for the Salomonde II powerhouse complex (adapted from Espada and Lamas (2014)).

Regarding the powerhouse cavern excavation, a total of fifteen stages were considered and are described in Table 5.8. The cavern excavation was modelled in a way simpler than the reality, which had a minor impact in the computed displacements since the rock mass had shown good geomechanical quality during the excavation phase. In brief, the excavation sequence starts with the cavern arch, where two longitudinal galleries were excavated followed by three transversal galleries, remaining three rock piles to excavate in further stages (stage 3, 4 and 5). Stage 6 ends the

excavation of the cavern arch, while stages 7 to 10 reach the cavern base level. The auxiliary tunnels and structures were excavated from stages 11 to 15.

Concerning the adopted support system, it was mainly composed by shotcrete with a variable thickness (250 mm in the cavern arch and 100 mm in the walls and in the tops walls) and reinforced with metallic fibres, which results in a material with Young modulus equal to 20 GPa, a Poisson ratio of 0.20 and a unit weight of 25 kN/m³. In the numerical model, this material was modelled using a *shell* type of element with a linear elastic and isotropic behaviour. Beside the shotcrete, a 2 to 3 m spaced mesh of rock bolts (their length varies from 4 to 12 m) with a diameter of 25 mm was also considered. To simulate the rock bolts into the Flac^{3D} model, a cable type of elements was considered, which can yield tensile strength with two nodes and one axial degree of freedom.

Table 5.8 Adopted construction stages for Salamonde II 3D numerical model.

Stage	Model	Description
1 and 2		Excavation of two longitudinal galleries in the cavern arch (stage 1) and 3 transversal galleries (stage 2). Application of 25 cm fibre reinforced shotcrete and 25 mm rockbolts.
3, 4 and 5		Excavation of the remaining rock piles on the cavern arch. Application of 25 cm fibre reinforced shotcrete and 25 mm rock bolts.
6		Excavation of the cavern arch from elevation 143 m to 136 m Application of 25 cm fibre reinforced shotcrete and 25 mm rockbolts.
7, 8, 9 and 10		Excavation of the main cavern until the base level. Application of 10 cm fibre reinforced shotcrete and 25 mm rockbolts.
11, 12, 13, 14 and 15		Excavation between elevation 126 m and 89,8 m including all the auxiliary galleries and access tunnels Application of 10 cm fibre reinforced shotcrete and 25 mm rockbolts.

5.4.2. Numerical modelling results

In LNEC (2013) numerical analysis, a homogeneous and continuous mean for the rock mass was assumed along with an isotropic and linear elastic behaviour. Due to the use of explosives as an excavation method, a 2 m band was created to take into account some damage that could exist in the rock mass as a consequence of the explosions. In this way, for this 2 m zone a rock mass deformation modulus half of the one adopted for all the rock mass was assumed, that was 12 GPa (see Figure 5.10).

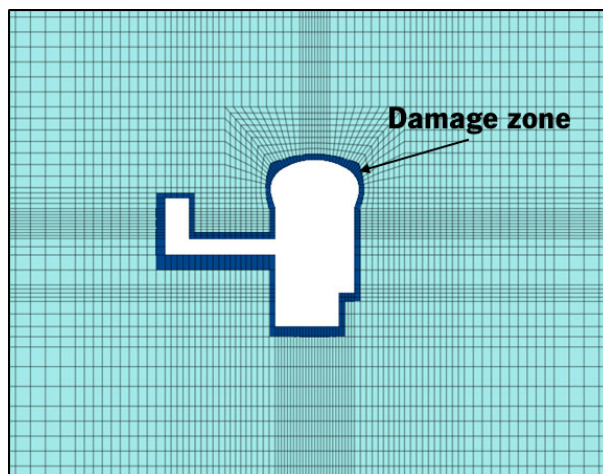


Figure 5.10 Detail of Salomonde II numerical mesh with the identification of the considered damage zone near the cavern (adapted from LNEC (2013))

The method used to achieve the rock mass deformation modulus was an inverse analysis technique performed using the displacements results in all the excavations stages of the six extensometers installed in two sections of the cavern, as explained in section 5.3. However, in the same report (LNEC, 2013) it is stated that the SFJ tests had registered higher values for E_m , such as 30 GPa near the cavern and 40 GPa near the access tunnel to the cavern, all these values as harmonic averages. Likewise, all the remaining tests (Uniaxial compressive strength test) and boreholes analysis corroborate the E_m high values (45 GPa). The gap between the E_m values obtained from the inverse analysis and the ones obtained from *in situ* and laboratory tests, may exist in the fact that the tests samples are not adequate to fully characterise the rock mass since normally the best samples are picked for analysis. Also, a meaningful number of zones showing an elevated weathering degree were found during the excavation coming to confirm this gap (LNEC, 2013).

In what concerns the numerical results, and assuming Section A and Section B adopted in the monitoring process, the maximum displacements after the cavern excavation were of 12.9 mm and

11.8 mm in the left and right side walls respectively for section A. In section B, the maximum displacement was on the cavern invert with a magnitude of 11.2 mm.

In terms of stress distribution, it was possible to conclude that the rock mass surrounding the tunnel is, predominantly, under compression with the maximum value recorded in the last excavation stage. In detail, stresses of 7 MPa can be detected in the cavern arch, while in the sidewalls the values ranging from 3 to 5 MPa. In what concerns the minimum principal stresses the maximum value was of 0.3 MPa located in sidewalls (shear stresses).

Apart from the two analysed sections, the maximum horizontal displacements and principal stresses were analysed in the cavern side walls and both tops for the last excavation stage (stage 15) and the results are summarised in Table 5.13.

Table 5.9 Summary of the displacements and stresses for the last excavation stage (adapted from LNEC (2013)).

Zone	Maximum displacement (mm)
Left sidewall (horizontal towards the cavern)	14
Right sidewall (horizontal towards the cavern)	13
North top (horizontal towards the cavern)	10
South top (horizontal towards the cavern)	8

Zone	Maximum principal stresses (MPa)
Arch	7
Left sidewall	5
Right sidewall	4

5.5. APPLICATION OF THE NEW CHARACTERISATION METHODOLOGY

5.5.1. Introduction

In order to characterise the rock mass of Salamonde II a new proposed methodology will be used. All the details regarding this characterisation methodology were already provided in Chapter 4 along with its validation using a theoretical case study. This methodology combines the geostatistical simulation and used the difference finite software Flac^{3D} to perform the following numerical analysis of the rock mass when the powerhouse cavern is excavated. As first step, the methodology required the existence of geotechnical data, which in this case study has been obtained in two different

phases (Figure 5.11: step 1). Then, a 3D numerical mesh created by Espada and Lamas (2014) was used and the zone centroids information exported for the following steps of the methodology (see Figure 5.11: step 2).

A third step entails the conditional simulate of each geotechnical parameter in the target grid points (zone centroids of the Flac^{3D} mesh). The geotechnical parameters considered for simulation were, for phase 1 the E_i , the W and the RQD value, while phase 2 considers the conditional simulation of the empirical systems, RMR, Log Q and GSI, as well as the previously mentioned parameters of phase 1 (see Figure 5.11: step 2). At this stage, is possible to obtain for each individual variable, a total of 100 realisations for the all rock mass. Then, and in order to perform the numerical analysis of the Salamonde II model, the previously simulated geotechnical data had to be converted into the geomechanical parameter E_m through the application of empirical formulas.

At this stage, it is important to note, that from now on the E_m values obtained using the geotechnical data from phase 1 will be addressed as data set 1, while data set 2 are composed by the E_m values obtained after combining the geotechnical data from phase 1 and phase 2. Subsequently, and still in step 4, a scenario reduction methodology is applied in order to reduce the number of realisations to import for Flac^{3D} for both data sets. This is a required step, since performing the average of the 100 realisations leads to a smoothing effect and the extreme values disappeared (see Figure 5.11: step 4). As a fifth and final step, the selected data sets are imported to Flac^{3D} and the E_m values assigned to each zone centroid of the mesh throughout a command routine programed suing *FISH* language (see Figure 5.11: step 5).

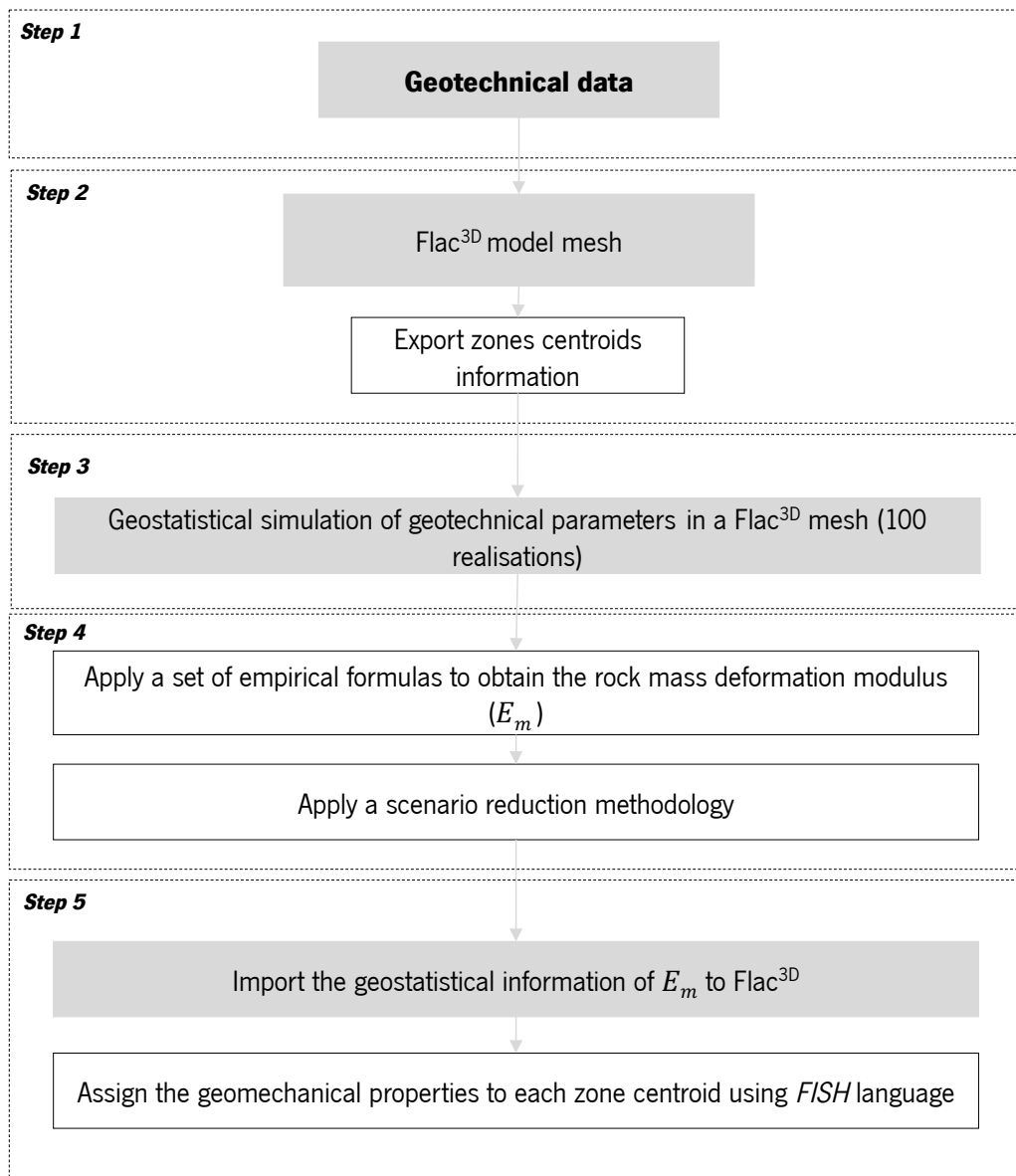


Figure 5.11 General workflow containing the general steps of the characterisation methodology using in Salamonde II.

5.5.2. Geostatistical simulation

Considering both phases of the geotechnical data, the geostatistical simulation of all the parameters was performed individually. Firstly, the data from phase 1 composed by 81 values of E_i , 1570 values of W and 1568 values of RQD.

Before performing an individual simulation of the three parameters, it was necessary to observe the existing correlations between all of them, specially between the E_i and the W parameters, in order to make sure that they were not (or weakly) cross-correlated. Otherwise, a joint simulation

(cosimulation) should be performed. The correlation matrix, which contains the Pearson product-moment correlation coefficients between all the parameters, is presented in Table 5.10.

Table 5.10 Correlation matrix between parameters P1, P2, P3 and P5.

	E_i	W	RQD
E_i	1.000	-0.402	-0.431
W	-0.402	1.000	0.009
RQD	-0.431	0.009	1.000

As a following step, and before computing the experimental variograms of these variables, a declustering cell size study had to be performed in order to weight the initial data to use in this type of simulation (see details in Appendix 1). From this study, a 7 m x 7 m x 7 m was identified as the optimal cell size for the presented variables. Then, as an imperative step, the variables had to be transformed into a normally-distributed variable (more details could be found in Appendix 1, as well as in the previous Chapters). Finally, the experimental variogram was built along the horizontal plane (xy plane), considering a maximum lag distance of 100 m. Here an isotropic behaviour for the variable was assumed for all variables (see Figure 5.12).

Secondly, the empirical systems gathered during phase 2 totalling 205 values of RMR, Log Q and GSI, were simulated. Then, and after combining the in common information from phase 1 and phase 2 (data set 2), a total of 102 points with information of E_i , 1795 points with information of W and 1793 points with information of RQD, was obtained. In the same way as the data from phase1, in this case a declustering cell size study was also performed (see details in Appendix 1); however, for this data the cell size increased to 8 m x 8 m x 8 m (seeFigure 5.13 and Figure 5.14).

Except for the W variable, all the others were modelled in a continuous scale. In the case of W, and since it varies in a discrete scale a different type of simulation had to be used, a truncated Gaussian model (Armstrong et al., 2011). All the steps required to the application of this model were detailed in section 3 of Chapter 3. In brief, a set of truncation threshold needed to be defined using the preliminary information and, with the help of the iterative algorithm known as Gibbs sampler, the class-indicator data are transformed into Gaussian data and the application of the TBM possible.

In what concern the experimental variogram computation, the lag distances were multiple of 10 m along the horizontal with a tolerance of 5 m for all the variables. All the experimental variograms were fitted using exponential, spherical and Gaussian basic structures, as presented below (the distances between brackets represent the correlation range of the structure along the horizontal plane and the number preceding the basic structure indicates the sill of the structure):

- Variogram models of the data from phase 1:

RQD: $\gamma = 0.27 + 0.10 \text{ Spherical } (12 \text{ m}; 12 \text{ m}) + 0.40 \text{ Gaussian } (80 \text{ m}; 80 \text{ m})$

W: $\gamma = 0.06 + 0.94 \text{ Exponential } (100 \text{ m}; 100 \text{ m})$

E_i : $\gamma = 0.16 + 0.97 \text{ Exponential } (25 \text{ m}; 25 \text{ m})$

- Variogram models of the data from phase 1 plus phase 2:

RQD: $\gamma = 0.23 + 0.18 \text{ Gaussian } (8 \text{ m}; 8 \text{ m}) + 0.54 \text{ Gaussian } (150 \text{ m}; 150 \text{ m})$

W: $\gamma = 0.17 + 0.83 \text{ Spherical } (150 \text{ m}; 150 \text{ m})$

E_i : $\gamma = 0.72 \text{ Exponential } (8 \text{ m}; 8 \text{ m}) + 0.40 \text{ Exponential } (50 \text{ m}; 50 \text{ m})$

RMR: $\gamma = 0.09 + 0.38 \text{ Spherical } (20 \text{ m}; 20 \text{ m}) + 0.18 \text{ Gaussian } (80 \text{ m}; 80 \text{ m})$

LogQ: $\gamma = 0.16 + 0.50 \text{ Spherical } (18 \text{ m}; 18 \text{ m})$

GSI: $\gamma = 0.23 + 0.90 \text{ Gaussian } (120 \text{ m}; 120 \text{ m})$

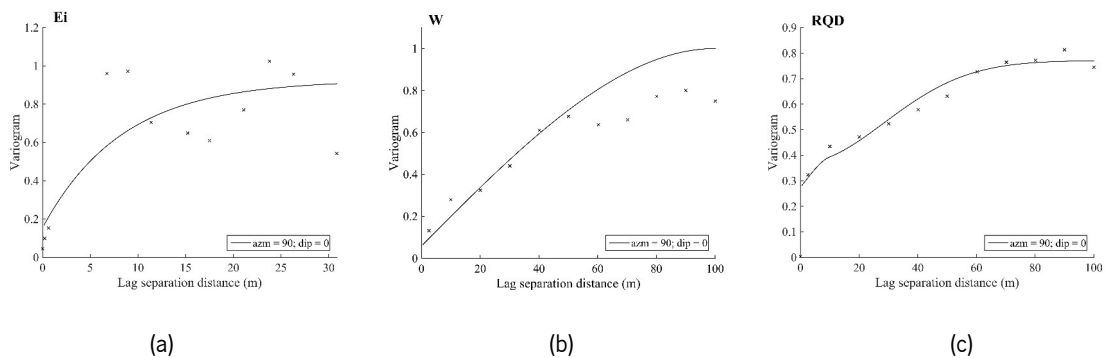


Figure 5.12 Experimental (crosses) and theoretical (solid lines) variograms along the horizontal plane xy plane for E_m using information from data set 1, for: a) E_i ; b) W; and c) RQD.

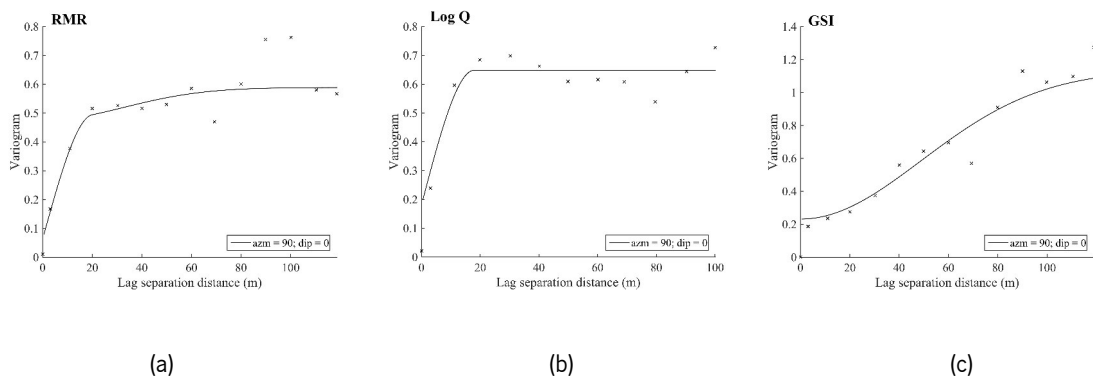


Figure 5.13 Experimental (crosses) and theoretical (solid lines) variograms along the horizontal plane xy plane for E_m using information from data set 2, for: a) RMR; b) Log Q; and c) GSI.

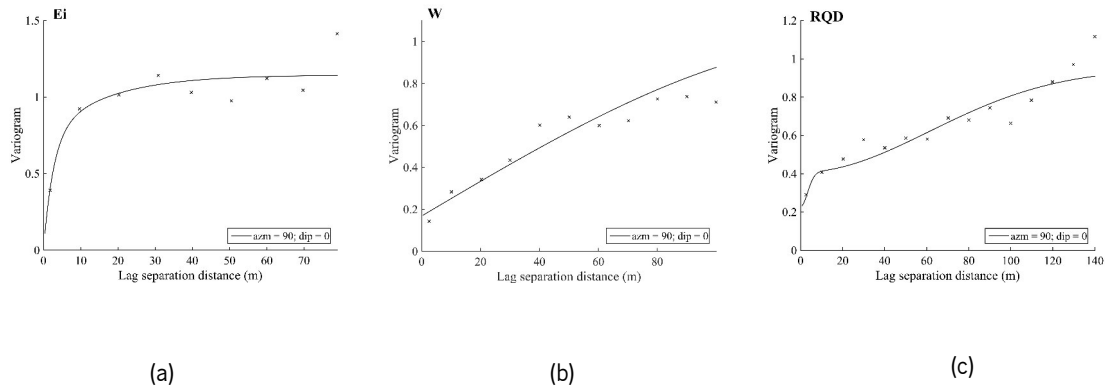


Figure 5.14 Experimental (crosses) and theoretical (solid lines) variograms along the horizontal plane xy plane for E_m using information from data set 2, for: a) E_i ; b) W_i ; and c) RQD.

All the variograms shows a nugget effect, proving that in all the geotechnical variables exist some significant errors of measurement and, at the same time, questions the spatial continuity of the variables. Even more, analysing the variograms of variable E_i , W and RQD was detected some considerable differences on the variables behaviour, which was more linear, if more data are added. like showed by the variograms in. Also, this statement can be confirmed by the nugget effect values that are lower in the variograms showed in Figure 5.14. Regarding the empirical systems experimental variograms, the GSI system was the one showing a more linear behaviour with a sill near one. The remaining two systems show lower sills, and consequently lower ranges, proving the higher spatial variability of the systems for lower distances.

As already mentioned, in order to apply the TBM the information of each variable should be transformed into normal scores. Therefore, the computation of the anamorphosis function is utmost important to improve the results of the simulated variables. All the details regarding the anamorphosis of both data sets variables can be found in Appendix 1. Regarding the W variable, the Gaussian thresholds adopted for the variable proportions can be consulted in Appendix 1, for both data sets.

Once the model parameters were specified (anamorphosis function for all the parameters with exception of W , truncation thresholds for variable W and the variograms of the underlying Gaussian random fields), conditional realisations of the variables in both phases can be constructed using the TBM mentioned before. For the algorithm, a total of 1500 turning lines was used and the number of realisations was set to one hundred. Also, it is important to mention that all the realisations were conditioned to the initial data and simulated on the mesh obtained from Flac^{3D} mesh (642,086 zone centroids).

Table 5.11 Statistical analysis of the variable from phase 1 and 2 after the geostatistical simulation.

Information phase	Variable	Number of samples	Minimum	Maximum	Mean	Variance
Phase 1	E_i (GPa)	642,086	10.00	79.99	55.77	174.46
	W	642,086	1.00	5.00	2.12 ⁵	0.94
	RQD (%)	642,086	0.00	100.00	69.19	314.64
Phase 1 + Phase 2	RMR	642,086	50.23	74.92	63.35	13.45
	Log Q	642,086	0.02	1.30	1.05	0.04
	GSI	642,086	50.01	70.72	65.25	20.86
	E_i (GPa)	642,086	5.02	84.99	54.52	264.24
	W	642,086	1.0	5.0	2.01 ⁶	0.86
	RQD (%)	642,086	0.0	100.0	85.94	456.46

Before going further with both data sets simulation results, they should be validated throughout the application of the cross-validation technique. Contrarily to the technique used in Chapter 3 (jack knife), in here, the limited number of samples had led to opt for the cross-validation technique. This technique temporarily removes one by one all the observations and simulate the variable in the removed locations, from which it is possible to relate the simulated values and compare them with the true ones. Using this technique as a basis, the quantification of uncertainty was evaluated through accuracy plots computed using a probability p varying from 0 to 1, from where based on the obtained realisations, an interval bounded by the quantiles $1-p/2$ and $1+p/2$ at each data location is computed (more details can be found in Chapter 2). A value of 1 is assigned to the data location if the true variable value belongs to the previously defined interval and 0 otherwise. A desirable accuracy plot would be the one where the points are coincident with the 45° line (diagonal line). As shown in Figure 5.15, for the variable of data set 1, the one that results in a poor assessment of the uncertainty was E_i . Since the number of input data for the simulation was smaller (81 points with information) the simulation resulted in an optimistic model, giving higher slightly higher E_i values than the reality. Regarding the RQD and W variables, the uncertainty assessment was good, being the W variable resulted in a more conservative model.

In what concerns data set 2, the same variables were simulated, this time with an increase in the input information used for the conditional simulation. Therefore, and as expected, the uncertainty modelling of the three variables (E_i , W and RQD) are better than the one obtained for data set 1.

⁵ The mean value of the weathering degree was computed to give a notion of the averaging value, however, the most correct way to represent this type of variables is computing the histograms.

⁶ The mean value of the weathering degree was computed to give a notion of the averaging value, however, the most correct way to represent this type of variables is computing the histograms.

Indeed, the E_i variable now resulted in a more balance accuracy plot, giving lower values than reality for lower E_i and higher value for higher E_i . The same presented effect happens with variable W, where for higher W values the model are able to give slightly higher W than the true ones (see Figure 5.16).

Regarding the empirical systems validation, the accuracy plots are very good in all the three variables, being the best uncertainty assessment achieved by the RMR, while GSI give slightly conservative values (see Figure 5.17). Although the difference between the three variables are residual.

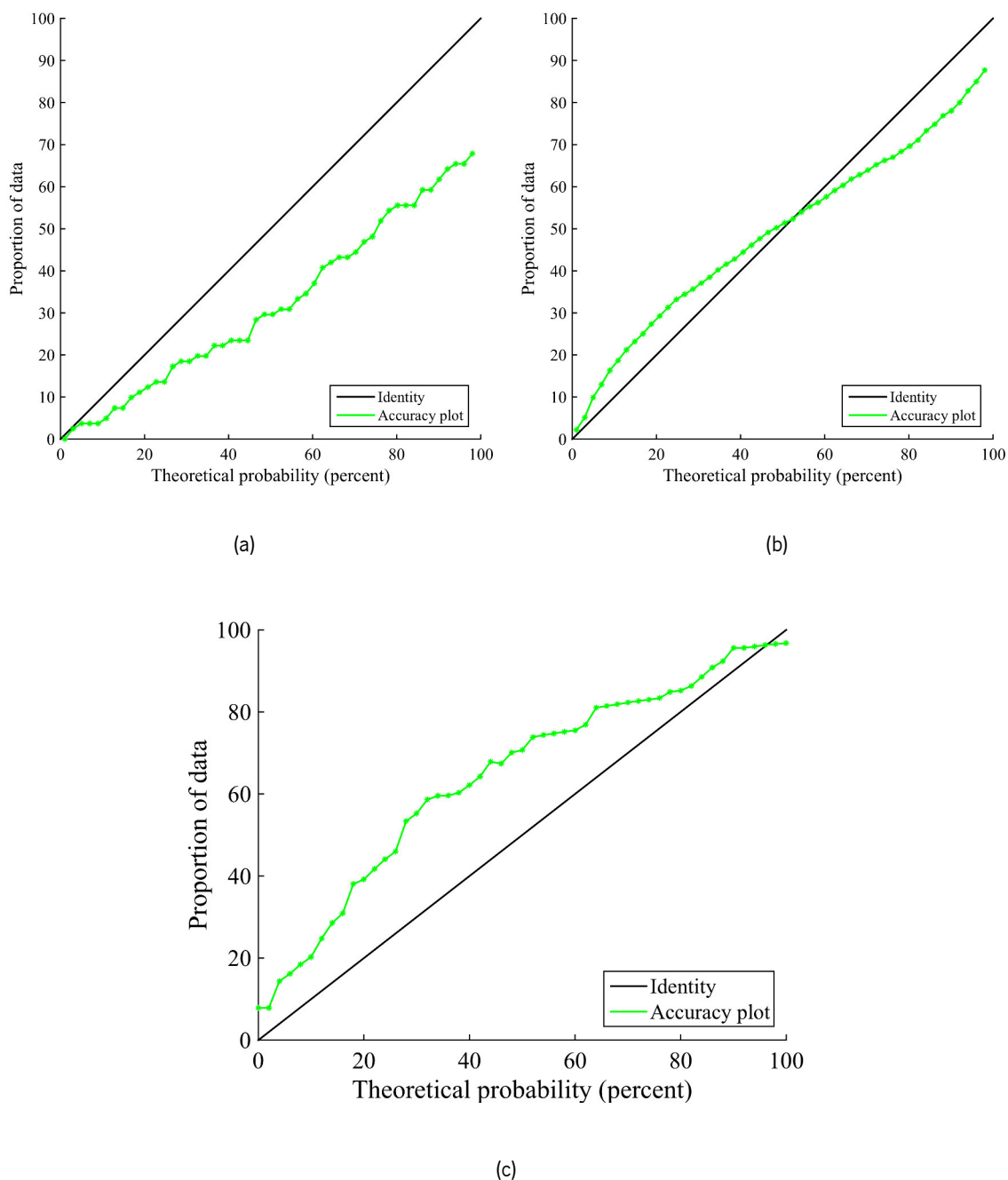
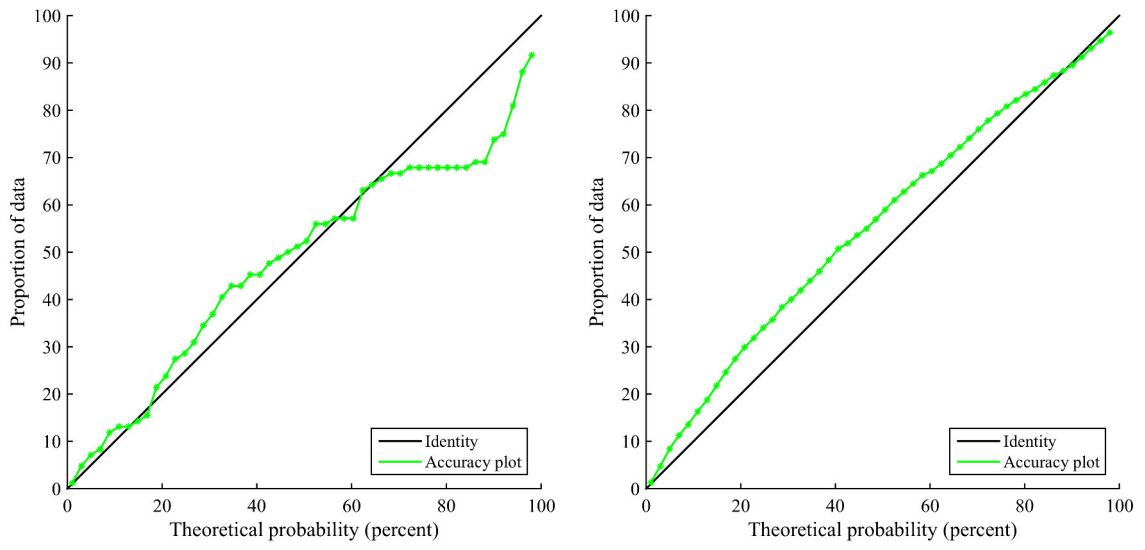
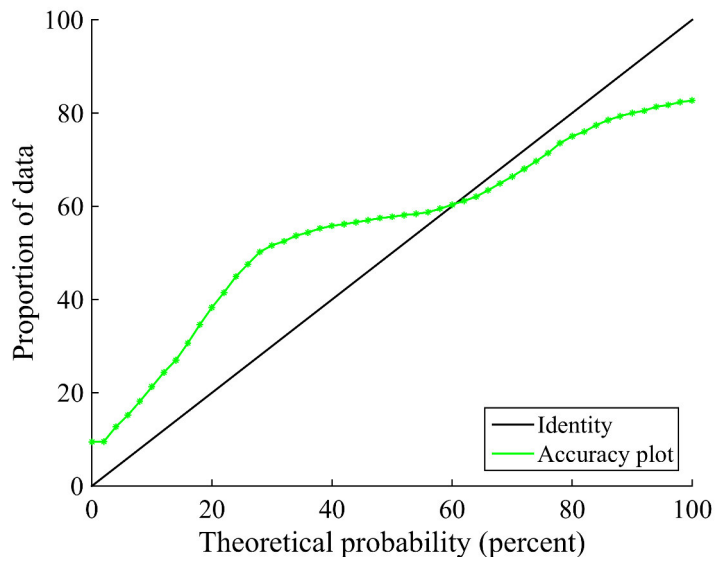


Figure 5.15 Accuracy plots of data set 1 showing the uncertainty modelling for: a) E_i (in GPa); b) RQD; and c) W.



(a)

(b)



(c)

Figure 5.16 Accuracy plots of data set 2 showing the uncertainty modelling for: a) E_i (in GPa); b) RQD; and c) W .

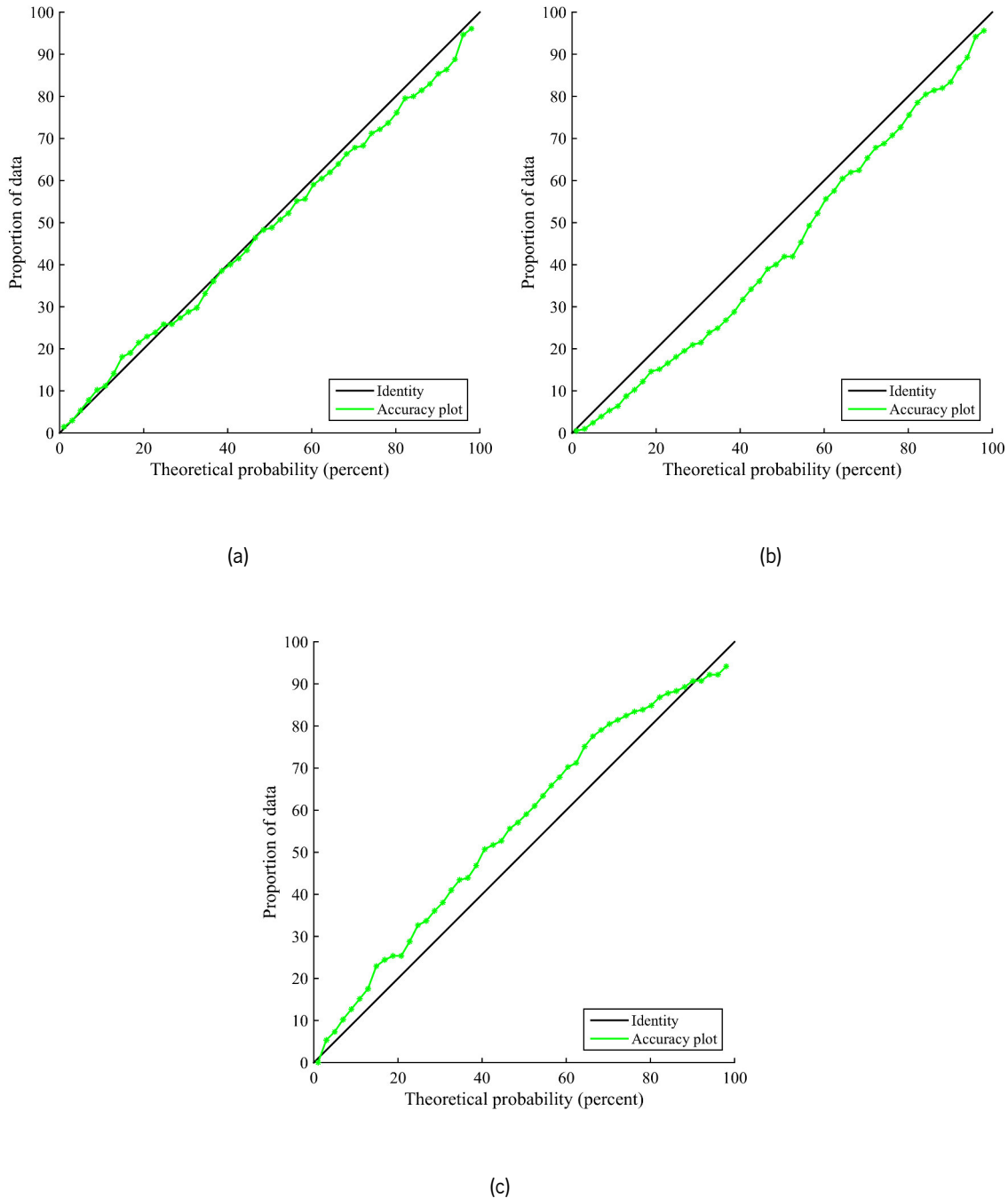


Figure 5.17 Accuracy plots of data set 2 showing the uncertainty modelling for: a) RMR; b) Log Q; and c) GSI.

5.5.3. From geotechnical data to geomechanical parameters

For the numerical models to use in this work, the geomechanical parameters of the rock mass are required, namely the rock mass deformation modulus (E_m). Although, the collected information in both phases are geomechanically different, while in phase 1 information about the E_i , W and RQD are available, in phase 2 an additional geomechanical classification was performed using the empirical systems, RMR, Q and GSI.

Since the geotechnical information obtained is quite distinct in both phases of the underground works, a division was made. For the first group of geotechnical information (phase 1) that contemplates information of the E_i , W and RQD was used a single empirical formula to obtain the E_m for further calculations. This formula was proposed by Kayabasi et al. (2003) and related all the three geotechnical parameters available in phase 1. On the other hand, using the geotechnical information from phase 2, and since there are two data sets that are geotechnically distinct, a separate analysis was performed. In the first data set 1 and, similarly to the data from phase 1, the Kayabasi et al. (2003) formulas was applied, while in data set 2, where all the available information is used, an average of some empirical formulas was performed (see Figure 5.18). The empirical formulas selected to use in this process can be consulted in Appendix 1. The basic statistics of rock mass deformation modulus can be seen in Table 5.12 accordingly divided by their geotechnical information phase. Likewise, Figure 5.19 shows the histograms for the E_m values in all the 100 realisations, and it is possible to observe that for data set 1 and 2 (that uses the geotechnical information obtained from the boreholes) the E_m values are considerably lower in comparison with data set 3, where there is an offset higher than 10 GPa (more data with a deformation modulus of 20 and 30 GPa value). Also, in this last data set, the range of E_m value is considerable lower, since the maximum is equal to 42.15 GPa, while data set 1 and 2 presents maxima close to 60 GPa. It is important to stress out that these E_m values were assumed in all the remaining process until the numerical results are achieved.

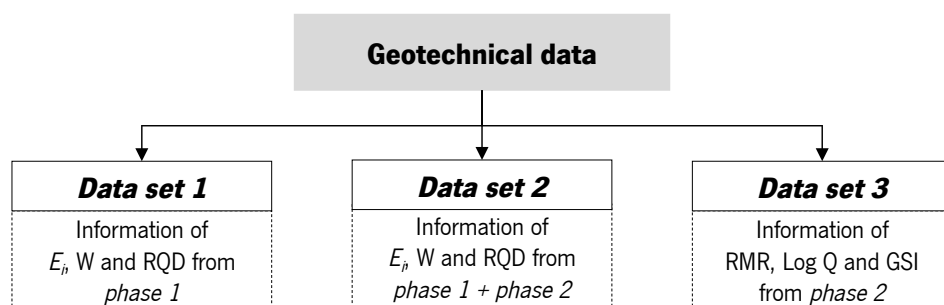
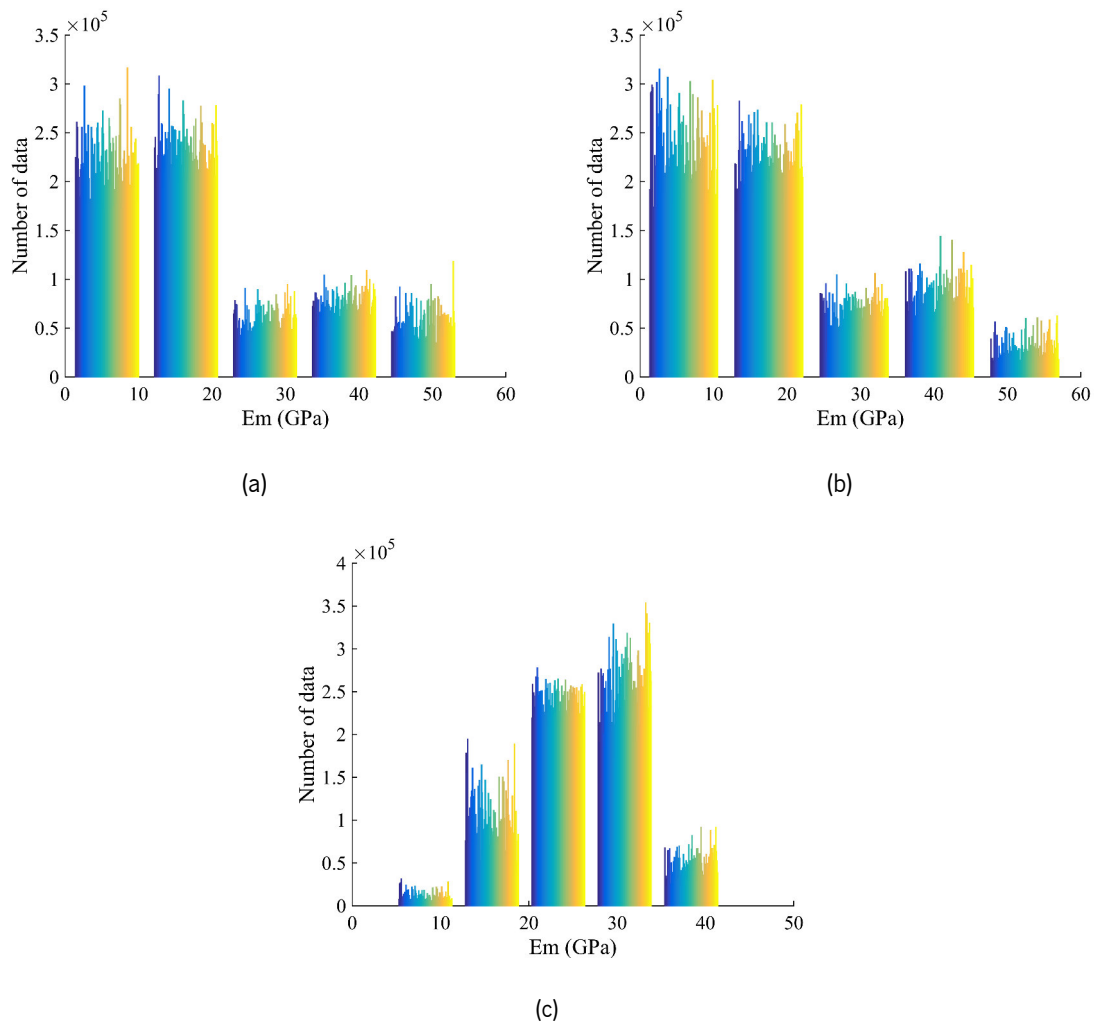


Figure 5.18 Summary of the geotechnical information that compose each considered data set.

Table 5.12 Statistical analysis of E_m (in GPa) obtained values for data set 1, 2 and 3 of the geotechnical information.

	Data set 1	Data set 2	Data set 3
Total of points with information	642,086	642,086	642,086
Mean	19.27	19.23	25.83
Variance	149.56	165.38	32.30
Standard deviation	12.23	12.86	5.68
Minimum	0.31	0.14	4.53
Maximum	54.14	58.16	42.15

Figure 5.19 Histograms of the E_m values (in GPa) of Salamonde II obtaining for all the 100 realisation (each colour represents one realisation), for: a) data set 1; b) data set 2; and c) data set 3.

5.5.4. Scenario reduction methodology

A first step required to apply the scenario reduction methodology presented in Chapter 4, is the representation of all the 100 realisations of the E_m in a $2D$ space using the Euclidean distance as metric. Therefore, in Figure 5.20 the 100 realisations obtained considering the E_m values from data

set 1, data set 2 and data set 3, are represented. In the following graphics, the x and y axes represent the distances of the dissimilarity matrix computed after the Euclidean distance calculation.

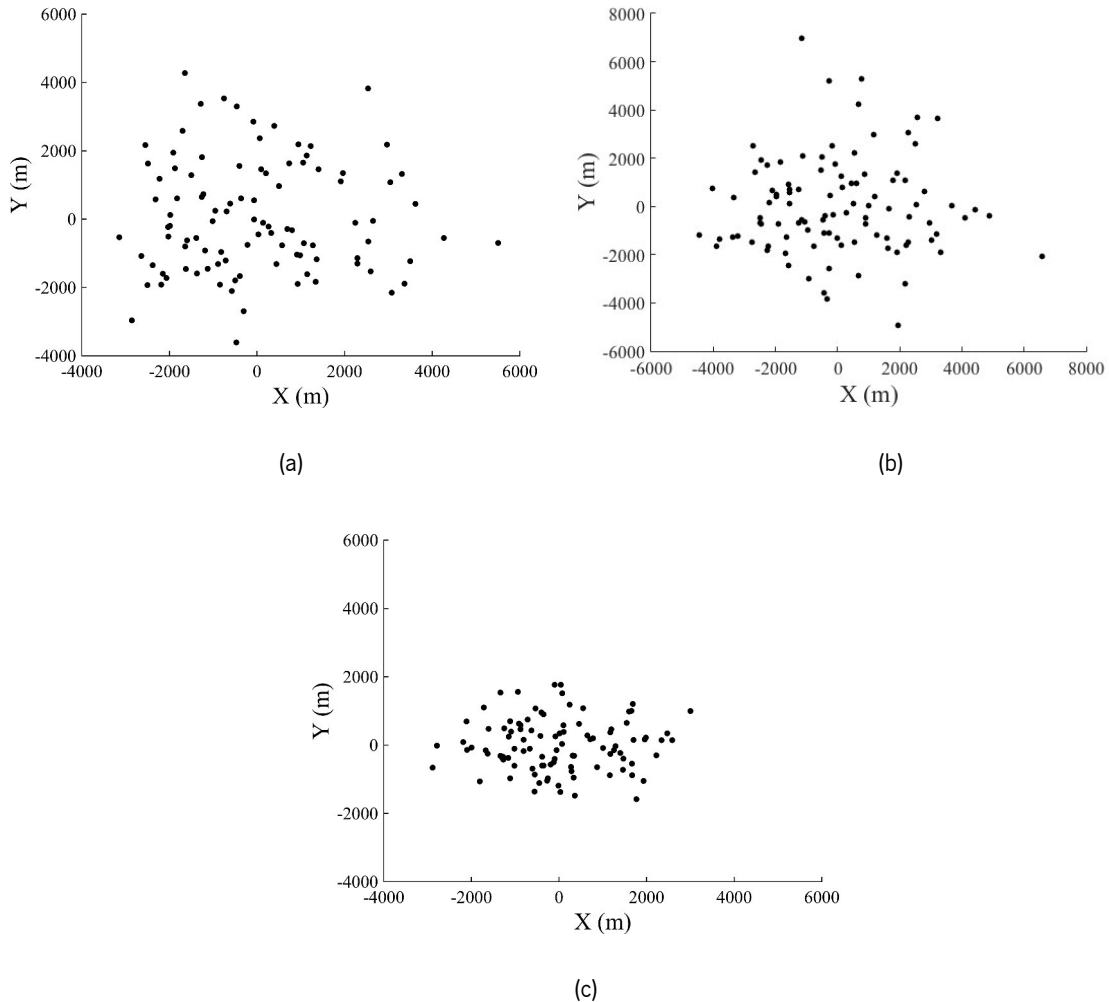


Figure 5.20 2D spatial representation of the E_m 100 realisations (black points) using the Euclidean distance computed with data from: a) data set 1; b) data set 2; and c) data set 3.

The following step of the methodology goes through the transformation of the Euclidean space into a more linear space, the Featured space and that is done using the Kernel algorithm with a kernel bandwidth of approximately 20% of the maximum Euclidean distance (more details can be found in Chapter 2 and Chapter 4). Subsequently, to choose the optimal number of clusters, a *silhouette average width* test was performed (Rousseeuw, 1987). This test allows assessing the number of clusters that best represents the all group of realisations and was performed for 2 until 10 clusters (*i.e.* the maximum number allowed was predefined as 10). Normally a *silhouette average width* (SAW) value higher than 0.5 gives a satisfactory division for the 100 realisations set. As a result, the optimal number of clusters obtained using data set 1 was 3 clusters with a SAW value of 0.55, while for data set 2 the number increased to 9 clusters (SAW equal to 0.56) and, for the last data set 3,

the number decreases again giving an optimal number of clusters of two (see Figure 5.21). Even though, the subsequent number of clusters for data set 3 was 4, with a final SAW value of 0.58.

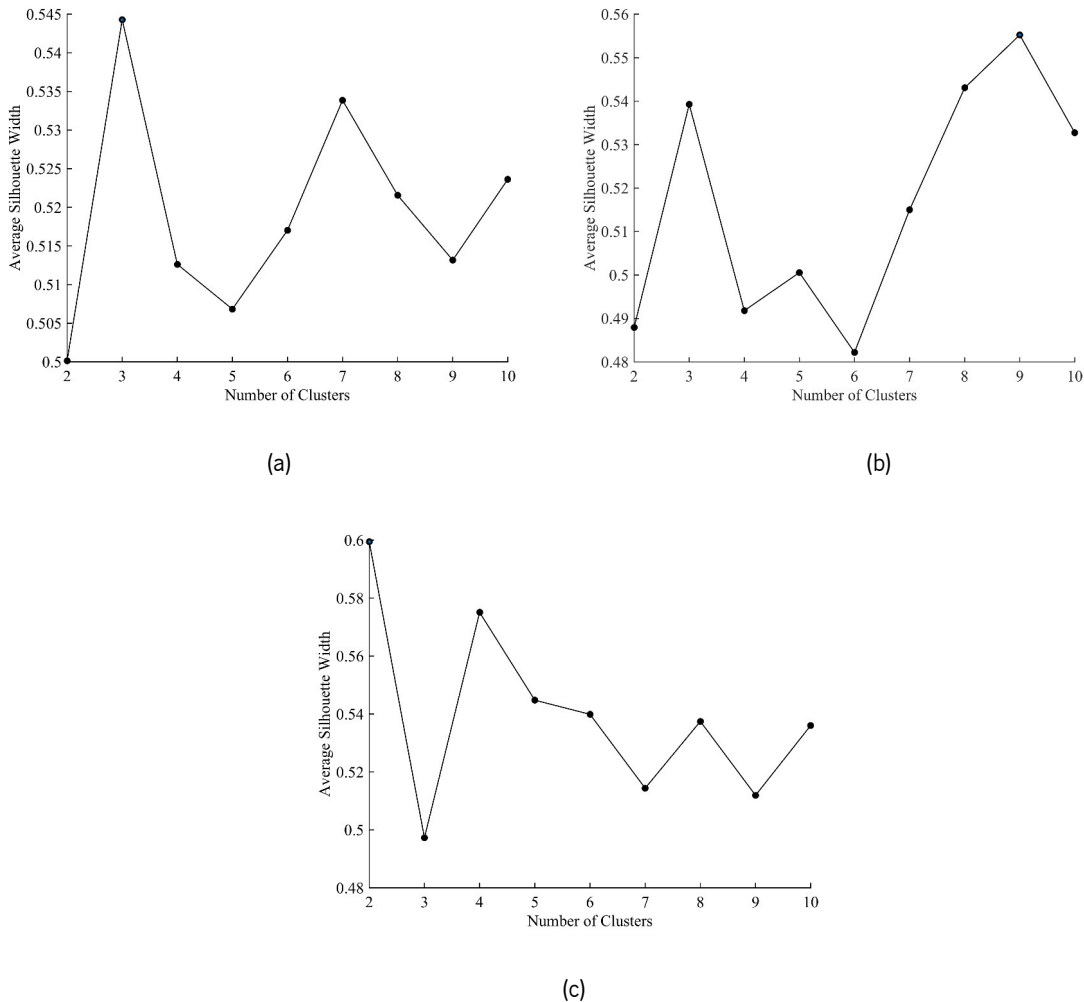


Figure 5.21 Average silhouette width values for the performed clusters evaluation using the data from: a) data set 1; b) data set 2; and c) data set 3.

Since the main goal is to identify the rock mass heterogeneities by assuming only two individual realisations this goal may be limited. Therefore, in the case of data set 3, other configuration of clusters was chosen, where a higher number of clusters (at least 3) are here assumed as a valid configuration. Certainly, the scenario reduction methodology could be defined to start with a minimal number of clusters set to three, however, the methodology would be at start eliminating one of the possibilities. Consequently, for the case study under consideration, the number of clusters chosen to represent the full realisations set of the E_m values were 3 clusters in the case of data set 1, 9 clusters for data set 2 and 4 clusters for data set 3. Accordingly, Figure 5.22 shows in colours the division made by the kernel k-medoid algorithm after performing 500 iterations of all three data sets, where some difference between the data sets can be noticed, namely the dispersions of the points that is higher for data set 1, which difficult the clusters division. In Figure 5.22 each cluster medoid

is identified with a square form with the corresponding cluster colour. It is important to stress out that each cluster centre corresponds to individual realisations chosen after a mathematical foundation to represent the full realisation sets. Furthermore, note that the x and y axes of Figure 5.22 show the same scale resulting from the Euclidean distance representation (see Figure 5.20).

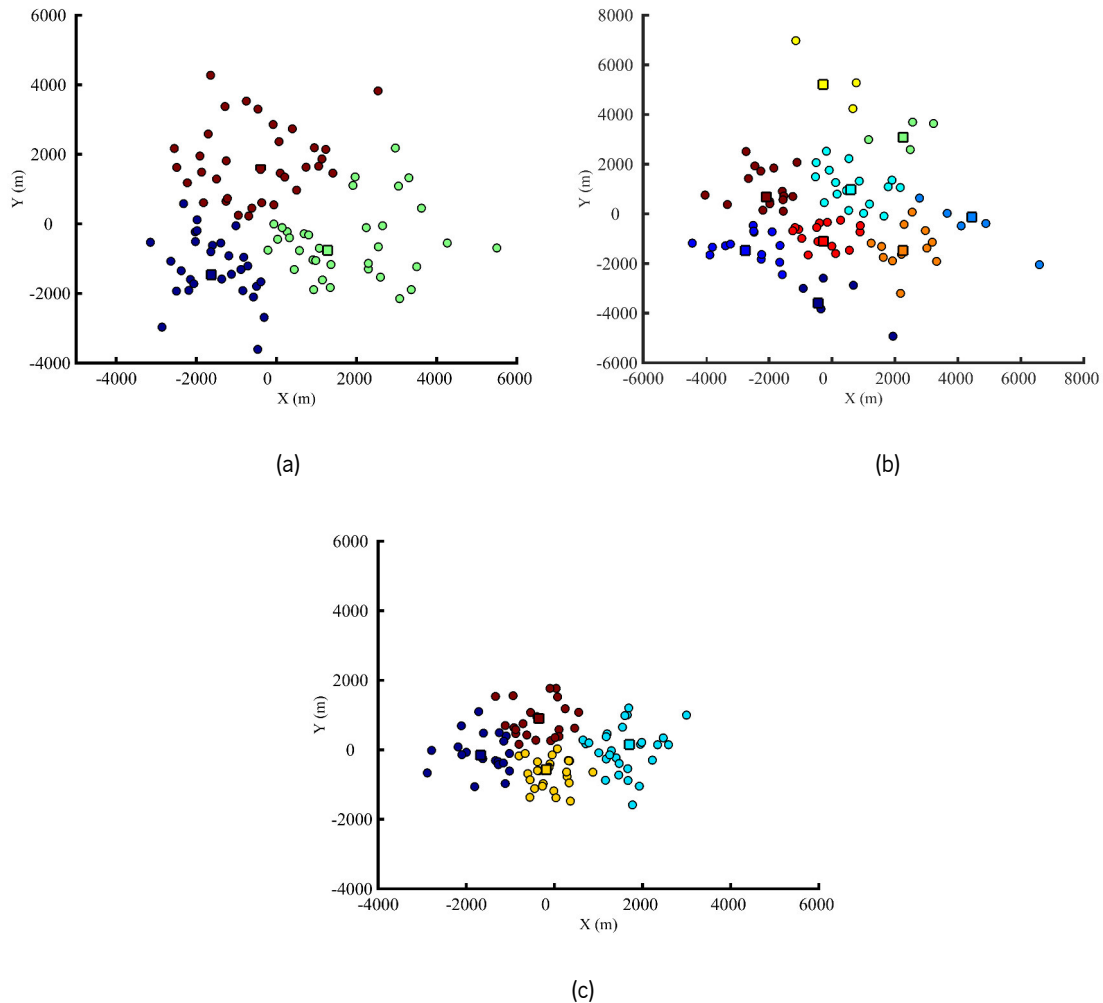


Figure 5.22 Clusters final configuration (points) with the matching medoids (squares) for: a) data set 1 (3 clusters); b) data set 2 (9 clusters); and c) data set 3 (4 clusters).

Once the clusters configuration is defined its validation is the next step to execute in this scenario reduction methodology. As a matter of fact, the main goal of this validation goes through a comparison for percentiles 10, 50 and 90 between the full realisations sets (data set 1, 2 and 3) and the selected realisations sets (data set 1, 2 and 3). These percentiles are computed in two ways: first, the percentiles are computed for each point of the grid and then their values are divided by total number of grid points. In this way, it is possible to obtain a single (average) value to represent each percentile (10, 50 and 90); secondly the percentile value is summed point by point resulting in a relation between the summed E_m value and the total of grid points. This last type of

graphical representation allows obtaining a more linear overview of the sets differences. Figure 5.23 and Figure 5.24 provides the mentioned graphics for all three data sets.

In the case of a three clusters configuration and since three is a relatively small number to represent the full set of realisations, the adjustment shows, as expected, some mismatches, namely in percentile 10 and 90 with difference of almost 5 GPa. On the other hand, for a nine clusters configuration, the correspondence between the full realisation set and the selected set is almost coincident.

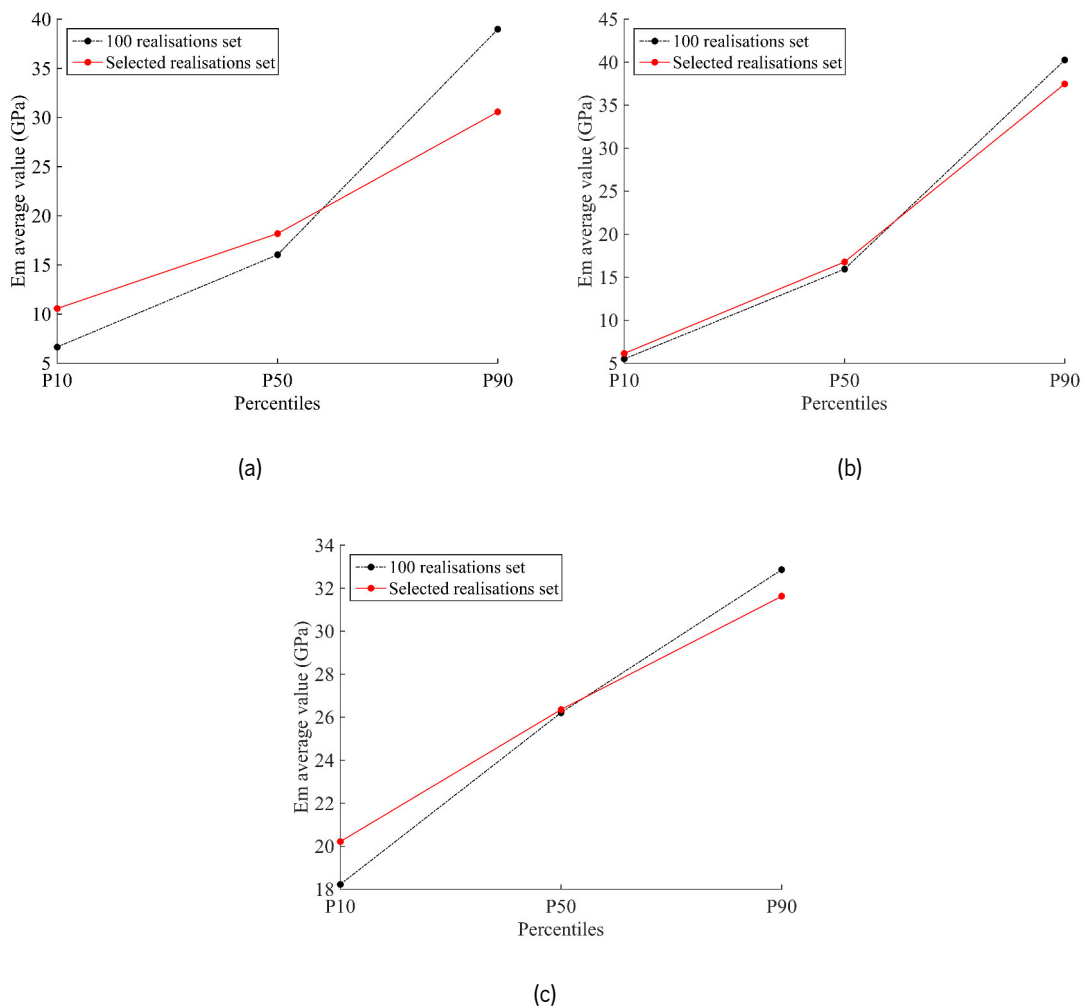


Figure 5.23 Validation of clusters configurations in comparison with the 100 realisations set using the average values of E_m for percentiles 10, 50 and 90, for: a) data set 1 (3 clusters); b) data set 2 (9 clusters); and c) data set 3 (4 clusters).

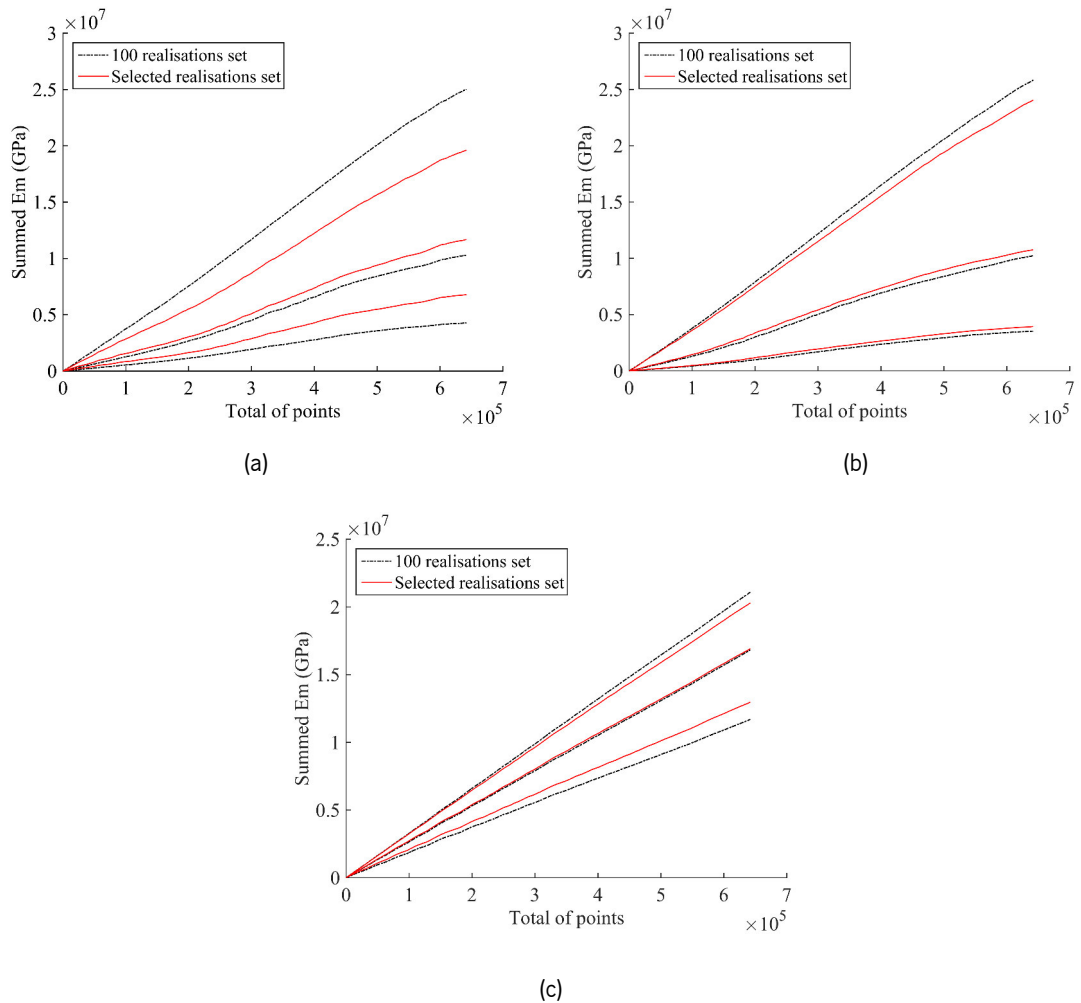


Figure 5.24 Validation of clusters configurations in comparison with the 100 realisations set after summing point by point the E_m values of percentiles 10, 50 and 90, for: a) data set 1 (3 clusters); b) data set 2 (9 clusters); and c) data set 3 (4 clusters).

From Figure 5.23 analysis, it is possible to observe that the difference in the E_m average value between the full set of realisations and the selected realisations of both data sets is always lower than 2 GPa, which can be considered as a residual value in a universe of E_m between 5 GPa and 45 GPa, except for data set 1 that shows a difference of 5 GPa. Indeed, analysing in detail Figure 5.23, it is possible to notice that the major difference between the selected realisations sets and the 100 realisations set happens in percentile 10 and 90, while percentile 50 shows correspondence. Regarding data set 3 (see Figure 5.23c), contrarily to data set 1, the major difference happens for percentile 10, however, in percentile 50 and 90 both sets are close to the full realisations set. An almost perfect correspondence is achieved by data set 2 (see Figure 5.23).

In regard to the representation of the percentiles point by point, Figure 5.24 allows perceiving that for data set 2 and data set 3 a better tuning is achieved between the selected realisations and full realisations sets when compared with data set 1; however, the differences are very small.

Particularly, it is important to stress out that in three data sets, the E_m values of percentile 50 are the quasi coincident between the sets. Furthermore, and to complete the clusters validation process, the E_m values were computed using a kernel smoothing function (Bowman and Azzalini, 1997) that returns the probability distribution of the E_m values in each set using a kernel function instead of a normal Gaussian one. This analysis was made for the three data sets (see Figure 5.25a, b and c) to compare the density function of the E_m values in each set of realisations, the selected, full and average of the 100 realisations. Through Figure 5.25 analysis, it can be stated that the selected realisations for all data sets of information follow the full realisations set density, while in the 100 realisations average there is a concentration of points with E_m values between 30 and 40 GPa. All the presented facts come to reinforce the option of using cluster configurations to statistically represent the full set of realisations. Due to this reason, the option of using the 3, 9 and 4 clusters to represent data set 1, 2 and 3, respectively, is strengthened.

In the following tables (see Table 5.13, Table 5.14 and Table 5.15) the basic statistics (mean, variance, minimum, maximum and percentiles values) of all data sets are displayed. The columns contain information concerning the initial data, the average values of the 100 realisations and of the selected realisations (see Table 5.13, Table 5.14 and Table 5.15). In addition, the basic statistics of each realisation that composed the scenario reduction chosen to represent each data set are provided in Table 5.16, Table 5.17 and Table 5.18, respectively for data set 1, data set 2 and data set 3.

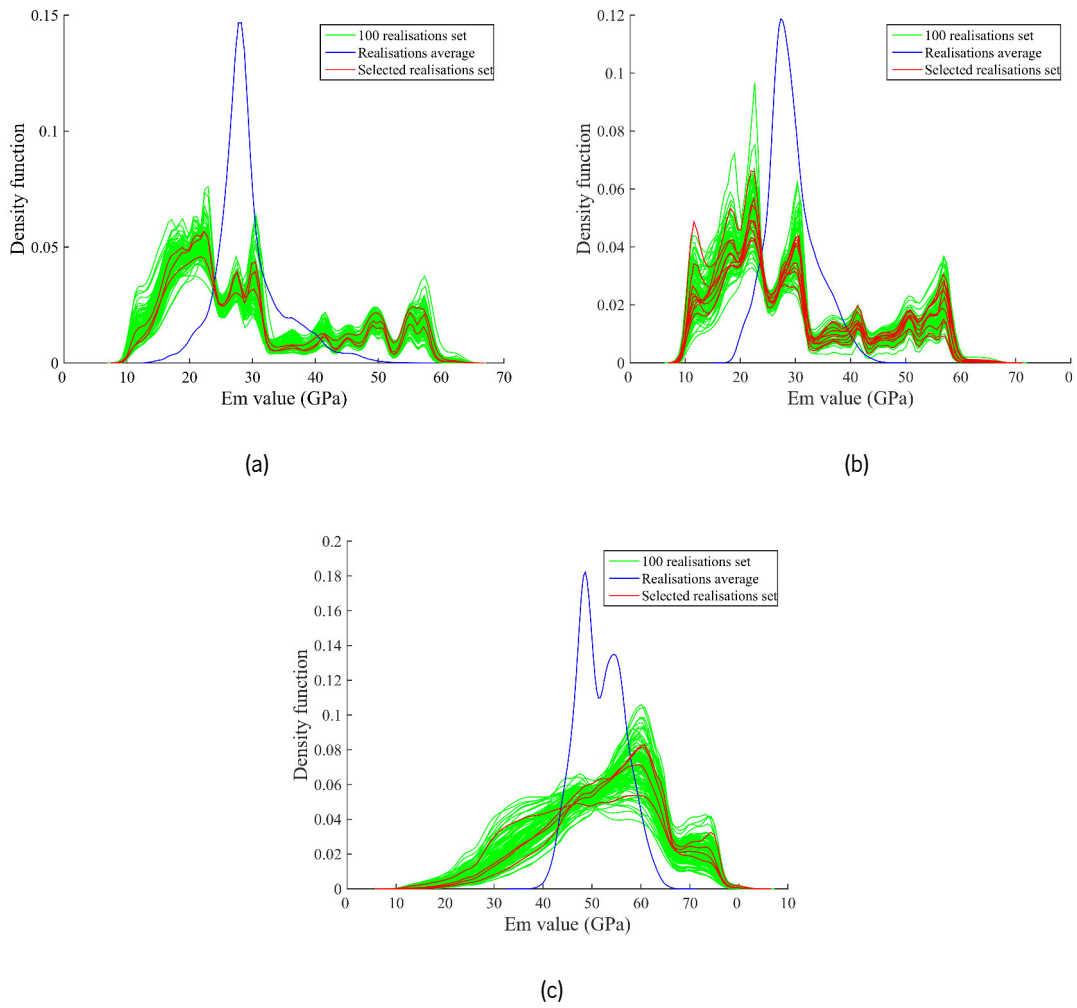


Figure 5.25 Validation of clusters configurations using a kernel smoothing function to represent the selected and full realisations sets along with the average of the 100 realisations, for: a) data set 1 (3 clusters); b) data set 2 (9 clusters); and c) data set 3 (4 clusters).

Table 5.13 Basic statistics of the E_m values (in GPa) for the initial values, average of the 100 realisations and selected realisations for data set 1.

	100 realisations (average values)	3 clusters (average values)
Number of grid points	642,086	642,086
Mean	19.27	19.78
Variance between realisations	149.56	153.81
Variance between points	27.51*	75.44*
Standard deviation	12.23	12.28
Minimum	0.31	0.40
Maximum	54.14	50.55
Percentile 10	6.66	10.57
Percentile 50	16.03	18.19
Percentile 90	38.98	30.57

* Variance obtained after performing the average of the realisations and compute the variance between the grid points

Table 5.14 Basic statistics of the E_m values (in GPa) for the initial values, average of the 100 realisations and selected realisations for data set 2.

	100 realisations (average values)	9 clusters (average values)
Number of grid points	642,086	642,086
Mean	19.24	19.35
Variance between realisations	165.38	167.98
Variance between points	20.00*	37.28*
Standard deviation	12.86	12.96
Minimum	0.14	3.37
Maximum	58.16	48.27
Percentile 10	5.50	6.13
Percentile 50	15.94	16.78
Percentile 90	40.25	37.27

* Variance obtained after performing the average of the realisations and compute the variance between the grid points

Table 5.15 Basic statistics of the E_m values (in GPa) for the initial values, average of the 100 realisations and selected realisations for data set 3.

	100 realisations (average values)	4 clusters (average values)
Number of grid points	642,086	642,086
Mean	25.83	26.14
Variance between realisations	32.30	31.21
Variance between points	5.86*	13.57*
Standard deviation	5.68	5.59
Minimum	4.53	4.65
Maximum	42.15	41.96
Percentile 10	18.23	20.22
Percentile 50	26.21	26.36
Percentile 90	32.86	31.62

* Variance obtained after performing the average of the realisations and compute the variance between the grid points

Table 5.16 Basic statistics of the E_m values (in GPa) for the individual realisation that compose the three-clusters configuration of data set 1.

	C1	C2	C3
Number of points	642,086	642,086	642,086
Realisation n.	69	76	93
Number of real. per cluster	27	31	42
Mean	21.59	19.68	18.07
Variance between points	198.08	171.60	158.05
Standard deviation	14.07	13.10	12.57
Minimum	0.31	0.31	0.31
Maximum	53.92	54.10	53.94

Table 5.17 Basic statistics of the E_m values (in GPa) for the individual realisation that compose the nine-clusters configuration of data set 2.

	C1	C2	C3	C4	C5	C6	C7	C8	C9
Number of points	642,086	642,086	642,086	642,086	642,086	642,086	642,086	642,086	642,086
Realisation n.	5	43	46	59	61	63	68	80	83
Number of real. per cluster	6	15	6	18	5	4	13	16	17
Mean	16.09	16.27	21.68	20.02	21.27	21.15	20.31	18.27	19.12
Variance between points	157.94	139.23	201.93	179.82	205.17	213.50	185.58	170.58	190.59
Standard deviation	12.57	11.80	14.21	13.41	14.32	14.61	13.62	13.06	13.80
Minimum	0.16	0.15	0.16	0.21	0.17	0.15	0.24	0.16	0.14
Maximum	58.03	57.93	58.04	57.83	57.88	57.96	57.81	57.79	58.13

Table 5.18 Basic statistics of the E_m values (in GPa) for the individual realisation that compose the four-clusters configuration of data set 3.

	C1	C2	C3	C4
Number of points	642,086	642,086	642,086	642,086
Realisation n.	28	38	40	62
Number of real. per cluster	22	29	26	23
Mean	27.51	24.57	26.16	26.31
Variance between points	31.15	44.82	33.95	33.63
Standard deviation	5.58	6.69	5.82	5.80
Minimum	5.20	4.65	4.96	5.19
Maximum	41.83	41.57	41.96	41.47

5.6. NUMERICAL MODELLING RESULTS

5.6.1. Introduction

Once defined the number of complementary and optimal clusters, a Flac^{3D} numerical analysis can take place. For this analysis, the three different data sets will be considered and their analysis will be performed separately, although they will be compared with the numerical results obtained when the rock mass is modelled as a homogeneous mean, *i.e.* deterministic values are assumed for the geomechanical parameters. Indeed, this analysis will allow understanding the advantages and powerfulness of using this new characterisation methodology instead of a deterministic one. For the homogeneous model, and since there are two different phases of information, a similar process to the one adopted for the heterogeneous models was considered, *i.e.* the same empirical formulas used for the heterogeneous approach were also applied to obtain the average value of the E_m to represent the Homogeneous model. In detail, the E_m values used in this last model, was 19 GPa for data set 1 and 2, and 29 GPa for data set 3 (consult Appendix 1 for more details).

Therefore, apart from the clusters and Homogeneous models, an additional one considering the 100 realisations average was also taken into account. This model was defined to understand the differences in considering the individual realisations rather than their average and, consequently, the limitations of using the latter. For this model was adopted the term *Mean*, following the same line of terms used in Chapter 4.

In order to perform an accurate comparison with LNEC (2013) numerical results, only data set 2 and data set 3 were used since they shared the same type of geotechnical information used in LNEC

(2013); however, an individual analysis of data set 1 will be performed in order to understand if, with less information, the new characterisation methodology is able to give similar results as if more geotechnical information is contemplated.

Moreover, it is worth mentioning, once again, the adoption of a damage zone in the LNEC numerical model to take into account the micro fractures that could result from the explosive use. Therefore, and since the information from phase 2 considers the excavation front mapping gathered after the powerhouse cavern excavation, the damage zone was not integrated in the new models.

For the sake of clarification, Figure 5.26 presents a scheme illustrating the path followed and the main steps required to build each data set model for analysis, being the first step, the geostatistical simulation of the geotechnical parameters, followed by the application of empirical formulas to obtain the rock mass deformation modulus. Subsequently, and due to the smoothing effect of averaging the 100 realisations, a scenario reduction methodology had to be applied. As output of the process were generated: 1) the data set 1 composed by three clusters; 2) the data set 2, where 9 clusters were assumed to represent the 100 realisations of E_m ; and 3) the data set 3 represented by four clusters.

The main goal of this section is to compare the results in terms of displacements and principal stresses maximum values between all the built models. In detail, the displacements maxima were analysed to last excavation stage in the cavern left and right walls, arch and invert, as well as in section A and B previously presented in section 5.3. Likewise, the principal maximum and minimum stresses were also evaluated in the same sections and zones of the powerhouse cavern. Moreover, aiming to prove, once again, the limitations of using the 100 realisations average or a deterministic model, a control point located in the middle of the rock mass was defined, specifically in the upper left of the powerhouse cavern.

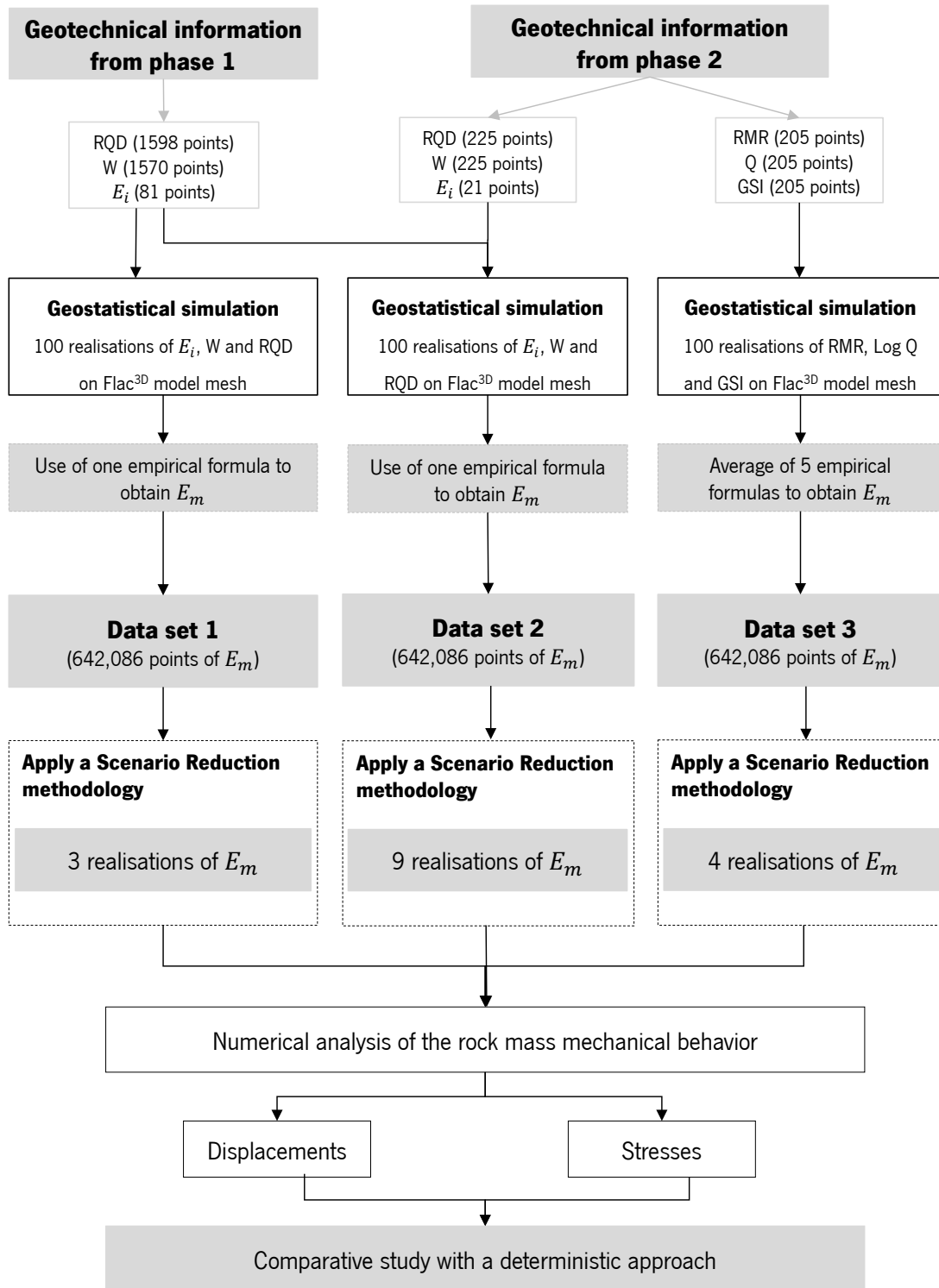


Figure 5.26 Workflow applied to build the data sets to represent the rock mass characterisation of Salomonde II using as input the geotechnical information from two distinct phases.

5.6.2. Data set 1 vs. deterministic model

In this subsection, the numerical results of data set 1 when 3 realisations are chosen to represent the full set of realisations, will be presented. Emphasis will be given to the results in terms of

displacements and principal stresses in the cavern walls, arch and invert, as well as in sections A and B. In addition, and as mentioned before, the same type of evaluation was made in the selected control point. Apart from the 3 individual realisations, the Mean and Homogeneous models will also be taken into analysis. Primarily, Table 5.19 presents a summary of the maximum displacements obtained for data set 1, as well as in the Mean and Homogeneous models. In the table second column, a section of the numerical model (was selected section A) is displayed from where it is possible to observe the spatial distribution of the E_m . The parameter variation is represented using a colour scale in which the blue colour embodies a E_m equal to 0 GPa and the red colour the upper limit with a deformation modulus of 60 GPa. From the table analysis, some differences should be highlighted, such as: 1) the ones obtained between the three clusters, where it is possible to notice that the maximum displacement value is always registered on the right sidewall for section A and jumps between the left sidewall and the cavern invert for section B, always with different magnitudes. This proves the ability of the scenario reduction methodology, that is part of the proposed characterisation methodology, in covering a wide range of possibilities (higher and lower values of E_m); 2) the differences between the clusters and the Mean model, where for the latter the registered displacements values in all the analysed sections were, significantly, inferior. On the other hand, the maximum displacement for all three clusters are always higher than the ones obtained with the Homogeneous model, showing values 110% higher in the vertical displacement of the tunnel invert (cluster 1) and 109% in the horizontal displacements of the right sidewall (cluster 3); and 3) the Homogeneous and Mean models resulted in similar values, proving, once again, the inability of the realisations average in consider the extreme parameter values. Furthermore, the differences between the displacements observed in all data set 1 clusters and the Homogeneous one can also be found in Table 5.19. These percentages were obtained considering the homogeneous displacements as reference. The disparity between models that are worthy of reference are: 1) the maximum absolute difference of 119% registered by cluster 2 in the south top of the cavern with a magnitude of 6.8 mm, while the homogeneous model showed a magnitude of 3.1 mm; and 2) in the south top and cavern invert the displacements exhibited magnitudes that are, sometimes, twice higher than the values of the Homogeneous model. These can be considered as significant differences, mainly in geotechnical engineering since they can translate in a more accurate mechanical analysis of the underground works.

With regard to the principal stresses, the difference between all the analysed models follows the same line as previously inferred for the displacements (see Table 5.20), with a maximum being registered for cluster 1 on the tunnel invert with a magnitude of 18 MPa (compression stress). For

the minimum stresses, that is shear stresses, the values vary from 0.5 MPa to 2.7 MPa in all clusters.

To have a visual representation of these differences, Figure 5.27 and Figure 5.28 display, for section A and section B, the magnitude of the displacements along with their directional vector (black arrow) for the homogeneous model and the clusters that present, for each section, the maximum value in the displacement (see Table 5.19). Similarly, Figure 5.29 shows the magnitude of the displacements along the cavern surface for the last excavation stage, where a deformational command is used to give an idea of how the cavern deforms. Also, and because some references were made to the maximum principal stresses, Figure 5.30 exhibits a 3D perspective of the cavern for the Homogeneous model and cluster number 1 of data set 1, where can be perceived the heterogeneous distribution of the stresses around the cavern.

With the purpose to summarise the main results obtained for data set 1, Table 5.21 indicates the average, standard deviation, minimum and maximum values obtained between all three clusters. These summary values should be used as a reference by the engineers in the underground design since they give an idea of what to encounter *in situ*, in terms of displacements.

Table 5.19 Summary of the maximum displacements (in mm) obtained for data set 1 - 3 clusters, Mean and Homogeneous models and the differences (in percentage) in relation to the Homogeneous.


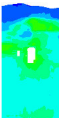
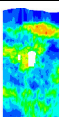
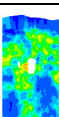
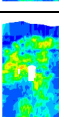
Scenario	Model	SA	SB	Control point	Left sidewall (LW)	Right sidewall (RW)	Arch (A)	Invert (I)	North top (NT)	South top (ST)
Homo.		5.40 RW	4.65 LW	0.32	5.21	5.66	4.10	5.60	4.27	3.10
Mean		5.31 (-1%) LW	5.51 (+18%) LW	0.35 (+9%)	5.80 (+11%)	5.20 (-8%)	3.78 (-10%)	5.80 (+4%)	3.80 (-11%)	4.23 (+36%)
C1		6.90 (+28%) RW	11.50 (+58%) I	0.22 (-31%)	8.21 (+58%)	7.22 (+28%)	3.11 (-26%)	11.77 (+110%)	3.20 (-26%)	5.88 (-25%)
C2		6.49 (+20%) RW	6.62 (+42%) LW	0.39 (+22%)	6.90 (+32%)	7.09 (+25%)	6.30 (+50%)	5.40 (-4%)	3.90 (-9%)	6.80 (+119%)
C3		7.21 (+34%) RW	5.98 (+29%) LW	0.44 (+38%)	6.00 (+15%)	11.84 (+109%)	7.60 (+81%)	6.40 (+14%)	6.20 (+45%)	5.50 (+77%)

Table 5.20 Summary of the maximum principal stresses (in MPa) obtained for data set 1 – 3 clusters, Mean and Homogeneous models (all the values should be negative corresponding to compressive stresses) and the differences (in percentage) in relation to the Homogeneous.

Scenario	SA	SB	Control point	Left sidewall (LW)	Right sidewall (RW)	Arch (A)	Invert (I)	North top (NT)	South top (ST)
Homo.	10.0 I	11.0 I	2.9	8.0	11.0	9.0	10.0	11.0	11.0
Mean	11.0 (+9%) I	12.0 (+9%) I	2.8 (-1%)	12.0 (+50%)	10.0 (-9%)	11.0 (+22%)	13.0 (+30%)	13.0 (+18%)	11.0 (0%)
C1	11.0 (+9%) I	19.0 (+73%) A	2.9 (0%)	14.0 (+75%)	11.0 (0%)	10.0 (+11%)	18.0 (+78%)	13.0 (+18%)	12.0 (+9%)
C2	10.0 (0%) A	11.0 (0%) RW	2.9 (0%)	14.0 (+75%)	16.0 (+45%)	11.0 (+22%)	17.0 (+70%)	15.0 (+36%)	10.0 (-9%)
C3	12.0 (+18%) I	13.0 (+18%) RW	2.9 (0%)	11.0 (+38%)	15.0 (+36%)	10.0 (+11%)	16.0 (+60%)	15.0 (+36%)	15.0 (+36%)

Table 5.21 Summary statistics of the displacements (in mm) regarding the 3 clusters of data set 1 in all the analysed zones.

	SA	SB	Control point	Left sidewall (LW)	Right sidewall (RW)	Arch (A)	Invert (I)	North top (NT)	South top (ST)
Clusters μ	6.87	6.64	0.35	7.04	8.71	5.67	7.86	4.43	6.06
Clusters σ	0.29	0.55	0.09	0.91	2.21	1.88	2.80	1.28	0.55
Minimum	6.49 ⁷	5.98 ⁷	0.22	6.00	7.09	3.11	5.40	3.20	5.50
Maximum	7.21 ⁸	7.33 ⁸	0.44	8.21	11.84	7.60	11.77	6.20	7.10

⁷ These minima values represent the minimum between the maximum displacements registered for section A and B in all the selected clusters.

⁸ These maxima values represent the minimum between the maximum displacements registered for section A and B in all the selected clusters.

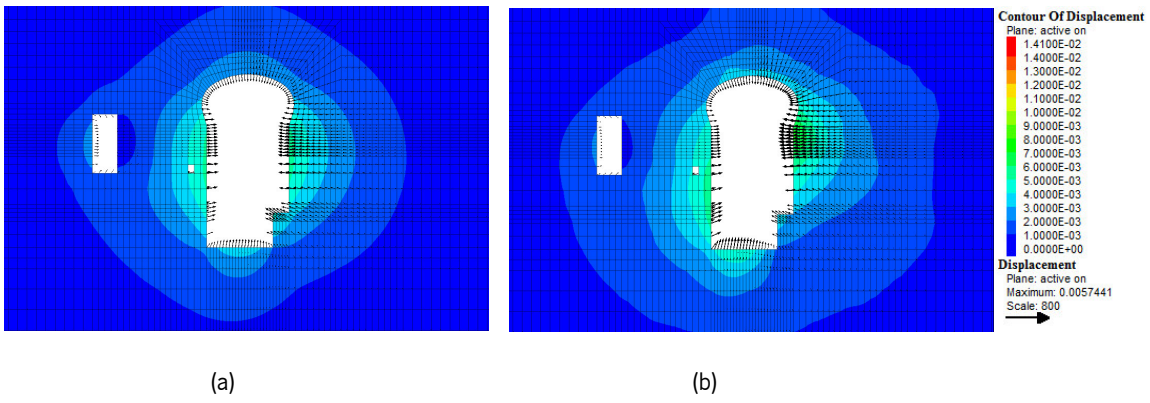


Figure 5.27 2D contours of the displacements magnitude (in m) and vectors for section A in the last excavation stage, for: a) the Homogeneous model; and b) cluster 3 of data set 1.

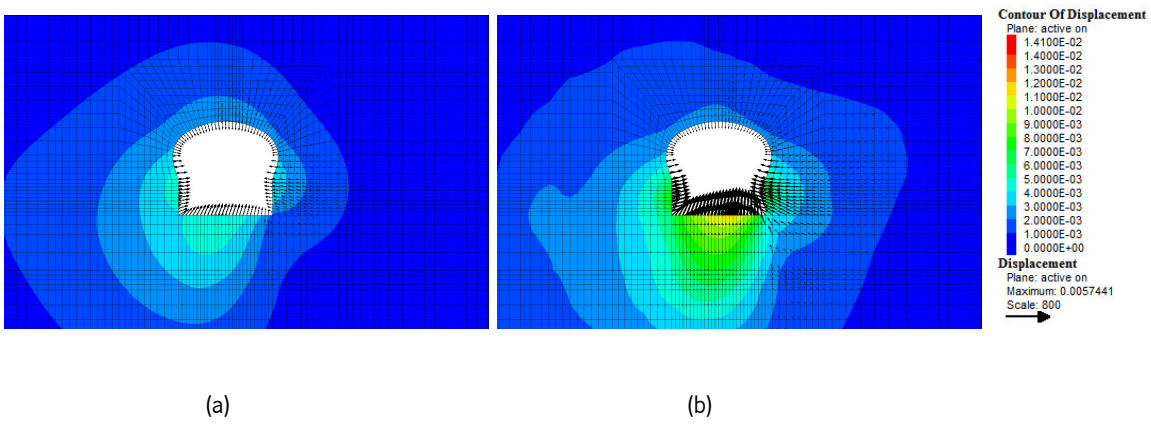


Figure 5.28 2D contours of the displacements magnitude (in m) and vectors for section B in the last excavation stage, for: a) the Homogeneous model; and b) cluster 1 of data set 1.

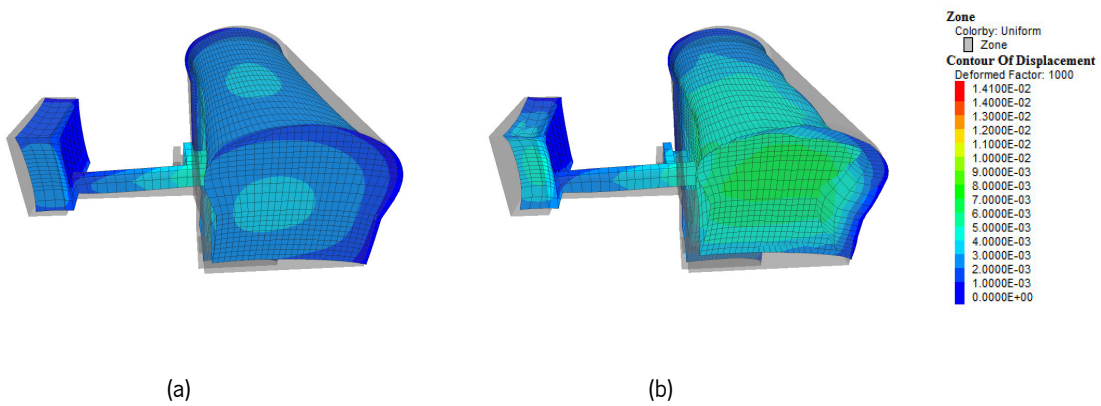


Figure 5.29 3D contours of the displacements magnitude (in m) with a deformed factor equal to 1000 in the last excavation stage, for: a) the Homogeneous model; and b) cluster 2 of data set 1.

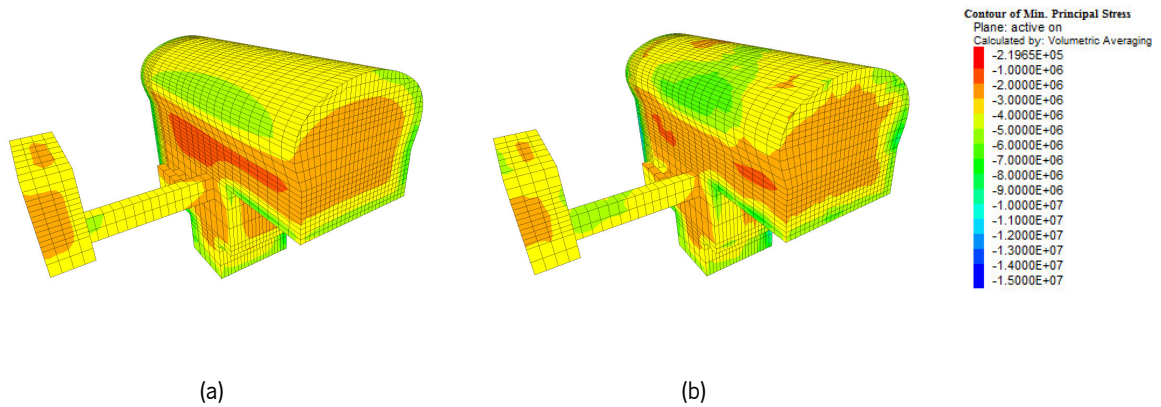


Figure 5.30 3D maximum principal stresses (in Pa) contours at the last excavation stage, for: a) the Homogeneous model; and b) cluster 3 of data set 1.

5.1.1.1 Comparison with the observed results

In order to assess the methodology accuracy, the data obtained from the monitoring campaign was used. This campaign was composed by 12 extensometers, located in section A (extensometers E1A, E2A, E3A, E4A, E5A, E6 and E7) and in section B (extensometers E1B, E2B, E3B, E4B and E5B) of the cavern. However, extensometers 1, 2 and 3 presented very small displacements readings, confirming the inadequacy of using these extensometers in this analysis; therefore, they were removed in both sections. Since the extensometers were installed in different timelines, some additional calculations had to be performed on the numerical results extracted from the software Flac^{3D}. Each extensometer is composed by 3 rods, with 5 m, 10 m and 30 m long, with exception of extensometers E5A and E5B that are composed by two rods of 5 m and 20 m, and extensometer E7 with two rods of 5 m and 15 m. All the displacements readings were made in an automatic way, except for the first readings that were registered manually. To facilitate the analysis and further comparisons, the measured displacements are represented assuming at the end of the rod (30 m and 20 m) a null displacement. This type of representation was executed since it was the one adopted in LNEC (2013). Thereby, the measured displacements are shown in Figure 5.31, from where it is possible to state that they are coherent with the cavern geometry. In brief, in section A higher values were measured in comparison with section B, and the same happens for the extensometers located in the lower elevations (6 and 7). To complement it is worth mentioning the decompression effect that can happen in the walls with higher height that can justify the highest horizontal displacements (section A).

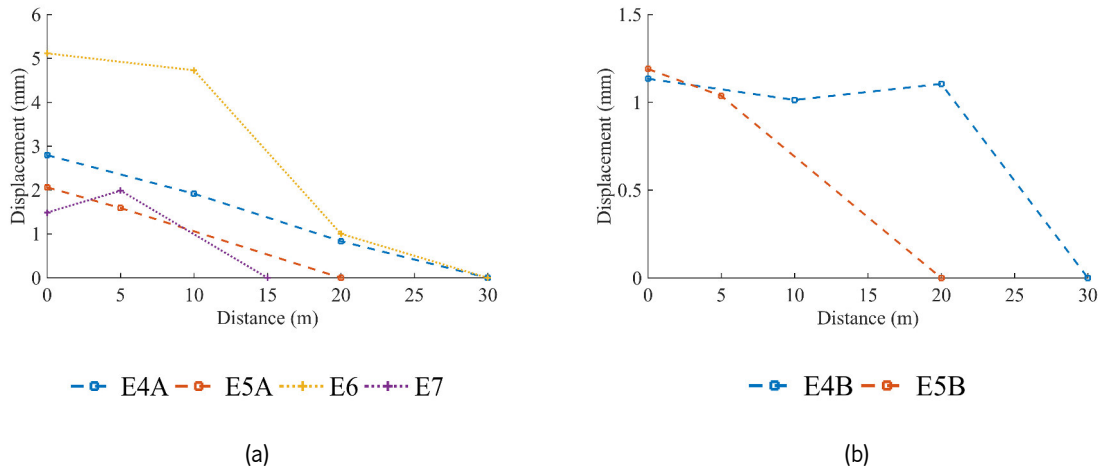


Figure 5.31 Displacements evolution measured by all the extensometers installed in: a) section A; and b) section B.

All the previously represented extensometers will be used for data set 1 comparison analysis. Therefore, the results computed for extensometer 4 installed after the sixth excavation stage on both sections, are presented in Figure 5.32, where the dashed line represents the displacement measured at 0 m, 10 m, 20 m and 30 m, the dotted line represents the displacements computed using the Homogeneous model and the remaining lines represent the calculated displacements for each one of the three clusters. As can be observed, the maximum displacements achieved by the three clusters and by the Homogeneous model, is half the value measured by E4A; however, it seems relevant to point out that all the three clusters are able to follow the same qualitative trend as the measured displacements and the Homogeneous line is, most of the times, located between clusters or under them (with lower displacements). In what concerns E4B, the case is different, since cluster 3 is able to track the measured displacements line. The main difference exists at a 20 m distance, where there is an increase in the measured displacement. Similarly, Figure 5.33 and Figure 5.34 show the results obtained for extensometers 5A and 5B and 6 and 7, the firsts with a total length of 20 m and the lasts with a total length of 15 m. From the figures analysis, it is possible to observe that for a length equal to 0 m, meaning near the cavern surface in section A, all the clusters lines are far from the measured displacements confirming the ability to underestimate the displacements by the considered models - although in the middle and end points the adjustments are almost achieved. The dispersion in the first meters is more accentuated for the extensometers installed in section A, showing section B extensometers (E4B and E5B) displacements value almost equal to the one obtained by the clusters.

Analysing closely, extensometers 5A and 6 are the ones with a higher difference between the measured and calculated values showing Mean Absolute Deviation (MAD) values higher than 2 mm for E5A and 5 mm for E6; however, and as can be proved by the MAD average values between all

the extensometers (see Table 5.22), the clusters provide more adjusted displacements when compared with the Homogeneous model (with all clusters showing an average MAD value lower than 2.327 mm obtained for the Homogeneous). In a practical point of view, and by computing the average MAD values of all the 3 clusters for each extensometer, it is possible to obtain values, most of the times, lower than the ones obtained for the Homogeneous model (with exception of E5A). Also, the maximum standard deviation of the three clusters absolute deviation in each individual extensometer is of 0.635 mm, which can be considered residual since the displacements are in a millimetre's scale. Even though, these relevant differences in the measured *versus* computed displacements, can be justified by some disturbance of the rock mass at the time of the excavation provoked by the use of explosives, as well as some considerable errors associated with the used equipment.

In sum, considering a wider range of E_m values, it is possible to state from a qualitative point view, that the contemplated clusters are able to reproduce the spatial variability existing in the rock mass, namely in more interior zones of the rock mass.

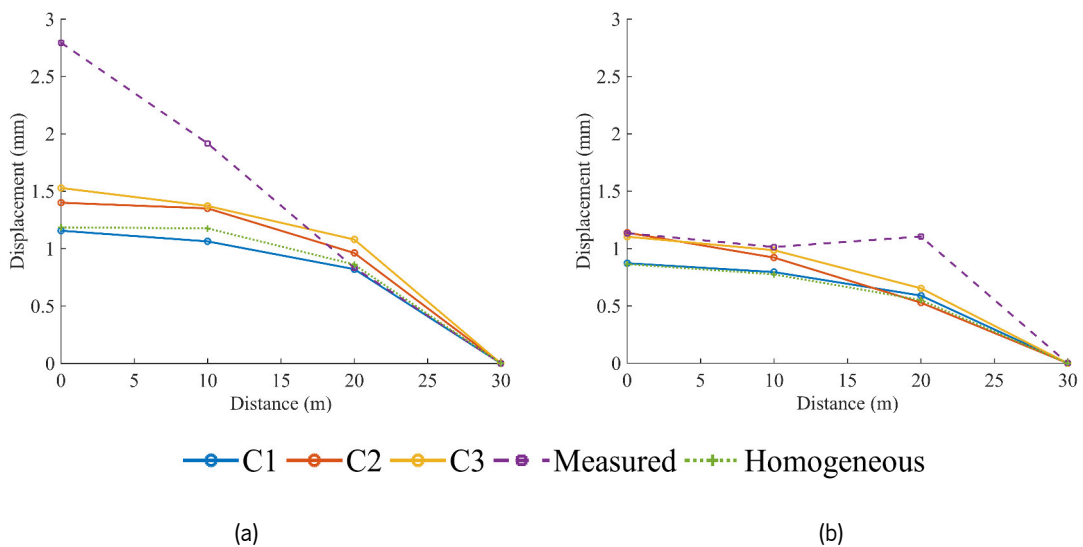


Figure 5.32 Total displacements (in mm) in the longer rod measured (dashed line) and computed (dotted and solid lines) for data set 1, for: a) E4A; and b) E4B.

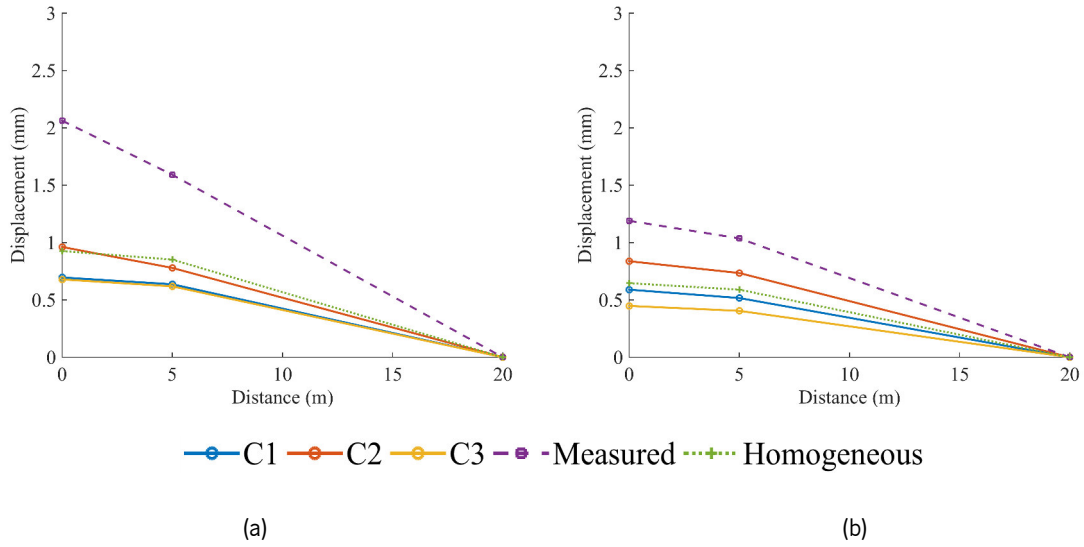


Figure 5.33 Total displacements (in mm) in the longer rod measured (dashed line) and computed (dotted and solid lines) for data set 1, for: a) E5A; and b) E5B.

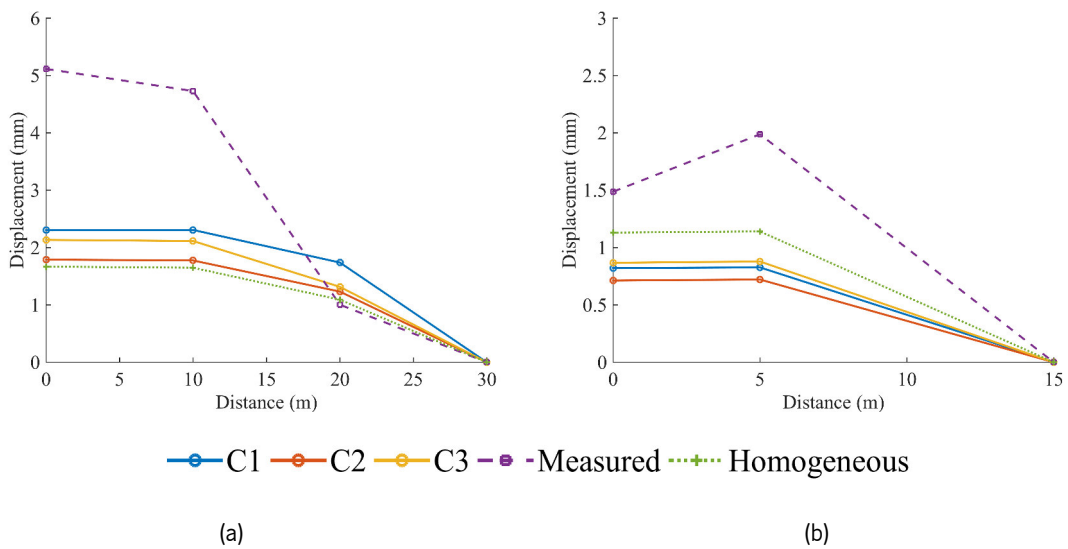


Figure 5.34 Total displacements (in mm) in the longer rod measured (dashed line) and computed (dotted and solid lines) for data set 1 in section A, for: a) E6; and b) E7.

Table 5.22 MAD values (in mm) obtain between the computed and the measured displacements for all the extensometers for data set 1.

	Homogeneous	C1	C2	C3	Clusters μ (mm)	Clusters σ (mm)
E4A	-2.389	-2.505	-1.833	-1.565	-1.968	0.395
E4B	-1.062	-0.996	-0.664	-0.510	-0.723	0.203
E5A	-1.877	-2.322	-1.913	-2.357	-2.197	0.201
E5B	-0.991	-1.120	-0.654	-1.374	-1.050	0.298
E6	-6.441	-4.497	-6.051	-5.286	-5.278	0.635
E7	-1.204	-1.827	-2.041	-1.729	-1.878	0.415
μ (mm)	-2.327	-2.211	-2.193	-2.137	-	-
σ (mm)	1.905	1.268	1.983	1.604	-	-

5.6.3. Data set 2 vs. deterministic model

In this subsection, and similarly to the subsection focused on data set 1, the numerical results of data set 2 are presented. They contain the results from all the nine clusters previously chosen in the scenario reduction methodology, as well as for the Homogeneous and Mean models. Once the model used in LNEC (2013) was developed using the same information as data set 2, the deterministic numerical results will also be compared with the ones obtained here; however, it is important to mention, that the rock mass deformation modulus assumed by LNEC is significantly lower (12 GPa and 6 GPa for the damage zone) than the one assumed for this data set 2 numerical analysis (which is in average 19 GPa). Thereby, any comparison made should take this E_m disparity into account. The main numerical results of all the mentioned models are provided in Table 5.23 and Table 5.24, where the first column shows the models cross section with the E_m distribution in the same scale used before, varying from 0 GPa (blue) to 60 GPa (red).

In general, the displacements are higher than the ones obtained for data set 1, with special emphasis for the difference between the left and right sidewalls that, in some cases is very relevant, proving, once again, the advantages of considering the spatial variability of the rock mass in the mechanical analysis.

In what concerns maximum displacements, in-depth look, the differences between the models to be highlighted are: 1) the Homogeneous model shows lower displacements when compared with each individual cluster; 2) by cross-checking the Mean and Homogeneous models, the differences are not as relevant as expected, corroborating the inadequacy of using the 100 realisations mean for this numerical analysis (with differences average of approximately 10%); 3) comparing the clusters and

the LNEC (2013) results, as expected, the displacements are lower than the values obtained by LNEC (see Table 5.23). In detail, in LNEC model, the maximum displacement reached a magnitude of 14.0 mm in the left sidewall, while cluster number 3 shows a maximum magnitude of 8.41 mm in the cavern arch (see Table 5.23). Also, in the cavern sidewalls, LNEC model presented as maximum displacements values of 14 mm and 13 mm, while cluster 1 showed nor very detached value of 12 mm and 15.30 mm for the left and right sidewalls. These displacements magnitudes prove, once again, the ability of the methodology in reaching the same displacements that were obtained assuming lower E_m values obtained from a back analysis of the measured displacements. Yet, in the cavern north top, the displacement magnitude was 10 mm in LNEC model and 16.2 mm in cluster number 1, being this last higher than expected according to the rock mass deformation modulus.

In terms of displacements contours, it is possible to confirm in Figure 5.35 and Figure 5.36 that due to the spatial variability of the deformation modulus, this contour is not as linear as the one obtained in LNEC. Indeed, by analysing the cavern deformation (see Figure 5.37) it is possible to confirm some irregularity of the cavern surface, justified by the lower values of the rock mass deformation modulus.

In what concerns the principal stresses, since the displacements are higher, one would expect to have lower magnitude of values in comparison with the results from data set 1. Indeed, Table 5.24 displays the maximum values of the principal stresses, from where it is possible to confirm that they are, in overall, lower; however, a maximum of 16 MPa was registered for cluster 3. Furthermore, it is important to highlight the relevant stresses value that appeared in the cavern walls in comparison with the lower value obtained by LNEC. Regarding the minimum principal stresses the magnitudes vary from 0.8 to 3.2 MPa.

As with the table shown for data set 1, Table 5.25 shows a summary of the values that a geotechnical engineer should use for the underground design to be aware of the displacements possibilities that can exist in the field.

Table 5.23 Summary of the maximum displacements (in mm) obtained for data set 2 - 9 clusters, LNEC (2013), Mean and Homogeneous models and the differences (in percentage) in relation to the Homogeneous.



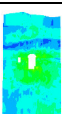
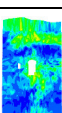
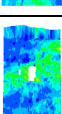
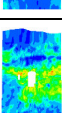
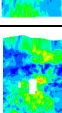
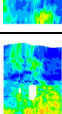
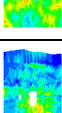
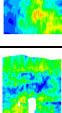
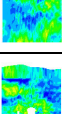
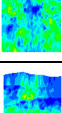
Scenario	Model	SA	SB	Control point	Left sidewall (LW)	Right sidewall (RW)	Arch (A)	Invert (I)	North top (NT)	South top (ST)
Homo.		5.40 RW	4.65 LW	0.32	5.21	5.66	4.10	5.60	4.27	3.10
LNEC (2013)		12.90 (+139%) LW	11.20 (+124%) I	-	14.00 (+169%)	13.00 (+130%)	-	-	10.00 (+134%)	8.00 (+158%)
Mean		4.95 (-8%) RW	4.10 (-18%) I	0.31 (-19%)	4.60 (-12%)	5.30 (+6%)	3.70 (-12%)	5.60 (-18%)	4.12 (-4%)	3.10 (0%)
C1		7.70 (+43%) RW	6.62 (+32%) LW	0.40 (+54%)	12.00 (+130%)	15.30 (+170%)	6.3 (+50%)	5.80 (+4%)	16.4 (+284%)	4.20 (+35%)
C2		6.73 (+25%) RW	5.6 (+12%) I	0.28 (+8%)	6.20 (+19%)	7.50 (+33%)	5.60 (+34%)	8.70 (+55%)	4.10 (-4%)	4.10 (+32%)
C3		4.79 (-11%) LW	3.98 (-20%) I	0.33 (+27%)	5.51 (+6%)	4.90 (-13%)	5.20 (-7%)	3.90 (-7%)	4.90 (+15%)	2.70 (-13%)
C4		5.40 (0%) RW	4.36 (-13%) LW	0.33 (+27%)	4.90 (-6%)	6.20 (+10%)	5.40 (+29%)	5.70 (+2%)	5.80 (+36%)	3.70 (+19%)
C5		8.41 (+56%) A	5.00 (0%) LW	0.48 (+85%)	6.50 (+25%)	5.50 (-3%)	9.05 (+116%)	5.20 (-7%)	6.60 (+55%)	3.50 (+16%)
C6		5.54 (+8%) RW	7.70 (+54%) I	0.59 (+127%)	9.70 (+86%)	8.00 (+41%)	5.0 (+22%)	12.00 (+114%)	3.00 (-30%)	5.70 (+84%)
C7		6.46 (+20%) RW	5.52 (+10%) I	0.23 (-12%)	7.10 (+36%)	8.90 (+57%)	4.40 (+5%)	7.50 (+34%)	5.40 (+26%)	3.70 (+52%)
C8		7.72 (+43%) RW	5.39 (+8%) RW	0.35 (+35%)	7.30 (+40%)	9.12 (+61%)	6.30 (+77%)	7.40 (+13%)	14.68 (+244%)	3.10 (0%)
C9		5.07 (-6%) RW	3.97 (-21%) I	0.25 (-4%)	3.50 (-33%)	5.10 (-10%)	4.12 (-2%)	4.50 (-20%)	2.70 (-35%)	3.10 (0%)

Table 5.24 Summary of the maximum principal stresses (in MPa) obtained for data set 2 – 9 clusters, LNEC (2013), Mean and Homogeneous models (all the values should be negative corresponding to compressive stresses) and the differences (in percentage) in relation to the Homogeneous.

Scenario	SA	SB	Control point	Left sidewall (LW)	Right sidewall (RW)	Arch (A)	Invert (I)	North top (NT)	South top (ST)
Homo.	10.0 I	11.0 I	2.9	8.0	11.0	9.0	10.0	11.0	11.0
LNEC (2013)	7.0 (-30%) A	-	-	5.0 (-38%)	4.0 (-64%)	7.0 (-22%)	-	-	-
Mean	9.0 (0%) A/I	10.0 (+11%) A/I	2.9 (0%)	11.0 (-7%)	12.0 (+18%)	11.0 (-10%)	13.0 (+15%)	12.0 (0%)	10.0 (0%)
C1	10.0 (0%) I	10.0 (-9%) I	2.9 (0%)	12.0 (+50%)	12.0 (+9%)	11.0 (+22%)	15.0 (+50%)	12.0 (+9%)	11.0 (0%)
C2	10.0 (0%) A/I	10.0 (-9%) A	2.9 (0%)	12.0 (+50%)	10.0 (-9%)	10.0 (+11%)	13.0 (+30%)	10.0 (-9%)	12.0 (+9%)
C3	10.0 (0%) A/I	11.0 (0%) I	2.9 (0%)	16.0 (+100%)	13.0 (+18%)	11.0 (+22%)	12.0 (+20%)	13.0 (+18%)	12.0 (+9%)
C4	10.0 (0%) I	12.0 (+9%) I	2.9 (0%)	11.0 (+38%)	14.0 (+27%)	12.0 (+33%)	13.0 (+30%)	13.0 (+18%)	10.0 (-9%)
C5	11.0 (+10%) A	11.0 (0%) A/I	2.8 (-3%)	14.0 (+75%)	13.0 (+18%)	11.0 (+22%)	12.0 (+30%)	12.0 (+18%)	10.0 (-9%)
C6	10.0 (0%) I	10.0 (-9%) A/I	2.9 (0%)	15.0 (+88%)	13.0 (+18%)	11.0 (+22%)	14.0 (+40%)	14.0 (+27%)	11.0 (0%)
C7	10.0 (0%) A/I	10.0 (-9%) A/I	2.9 (0%)	15.0 (+88%)	14.0 (+27%)	12.0 (+33%)	17.0 (+70%)	11.0 (0%)	12.0 (+9%)
C8	11.0 (+10%) A/I	10.0 (-9%) A/I	2.8 (-3%)	12.0 (+50%)	13.0 (+18%)	12.0 (+33%)	12.0 (+20%)	11.0 (0%)	12.0 (+9%)
C9	10.0 (0%) A/I	12.0 (+9%) I	2.9 (0%)	15.0 (+88%)	12.0 (+9%)	12.0 (+33%)	14.0 (+40%)	12.0 (+9%)	10.0 (-9%)

Table 5.25 Summary statistics of the displacements (in mm) regarding the 9 clusters of data set 2 in all the analysed zones.

	SA	SB	Control point	Left sidewall (LW)	Right sidewall (RW)	Arch (A)	Invert (I)	North top (NT)	South top (ST)
Clusters μ	6.45	5.35	0.36	6.97	7.84	5.70	6.77	7.06	3.87
Clusters σ	1.21	1.16	0.11	2.41	3.03	1.58	2.21	4.70	0.88
Minimum	4.79 ⁷	3.97 ⁷	0.23	3.50	4.90	3.90	4.50	2.70	2.70
Maximum	8.41	7.70	0.59	12.00	15.30	9.05	12.00	16.40	5.70

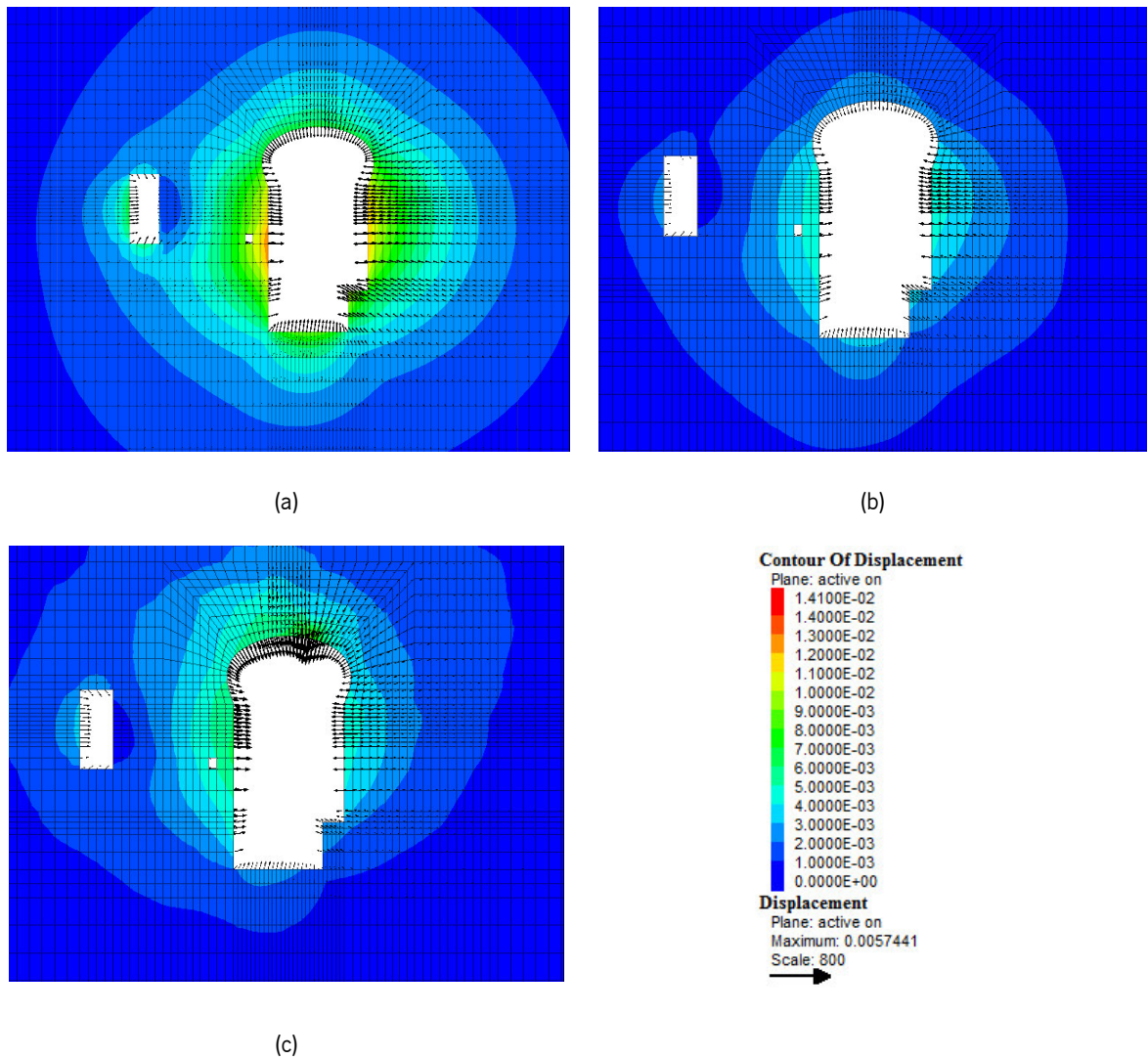


Figure 5.35 2D contours of the displacements magnitude (in m) and vectors for section A in the last excavation stage, for: a) the LNEC model (displacements with a scale of 400); b) cluster 5 of data set 2; and c) the Homogeneous model

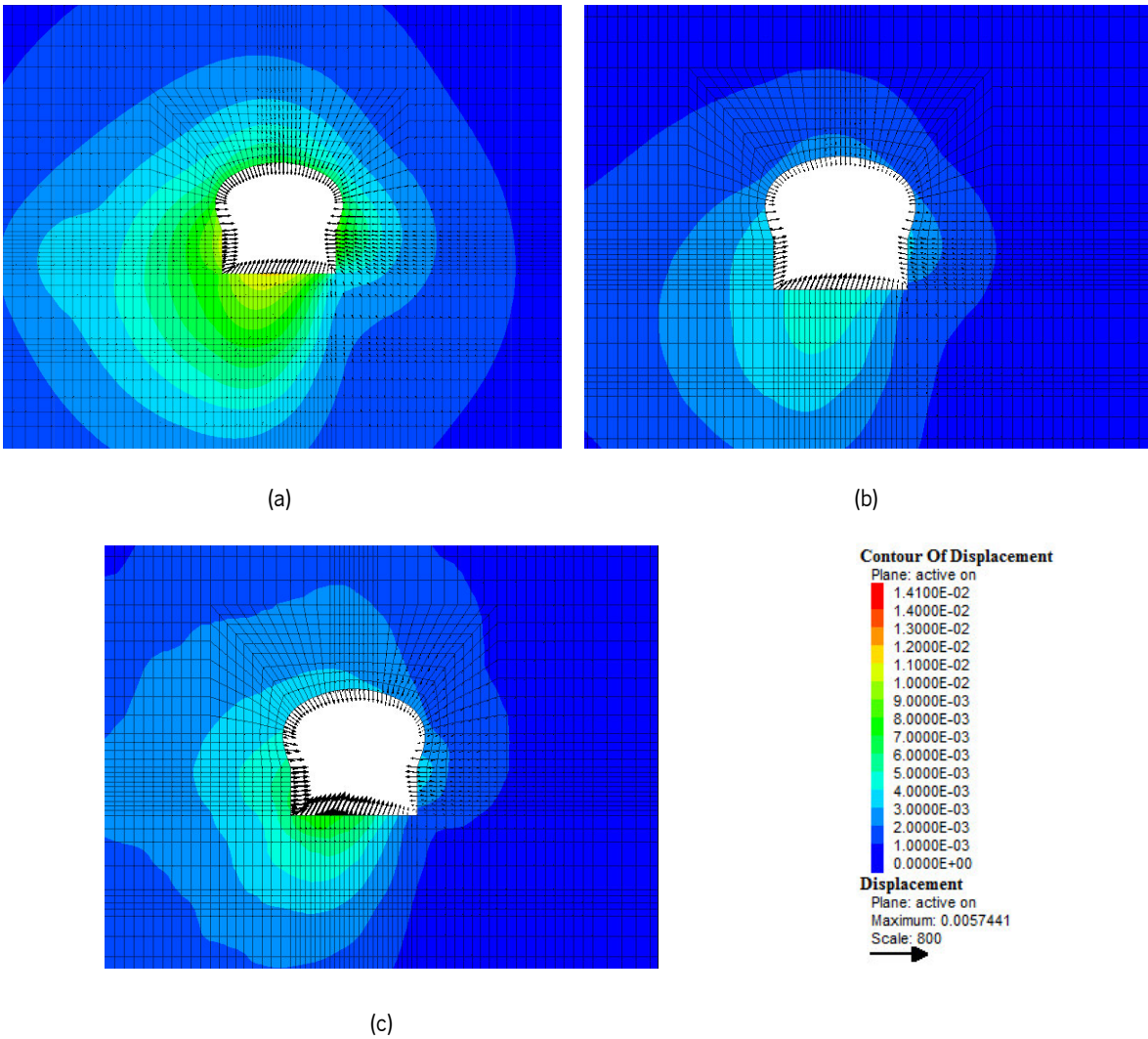


Figure 5.36 2D contours of the displacements magnitude (in m) and vectors for section B in the last excavation stage, for: a) the LNEC model (displacements with a scale of 400); b) cluster 6 of data set 2; and c) the Homogeneous model.

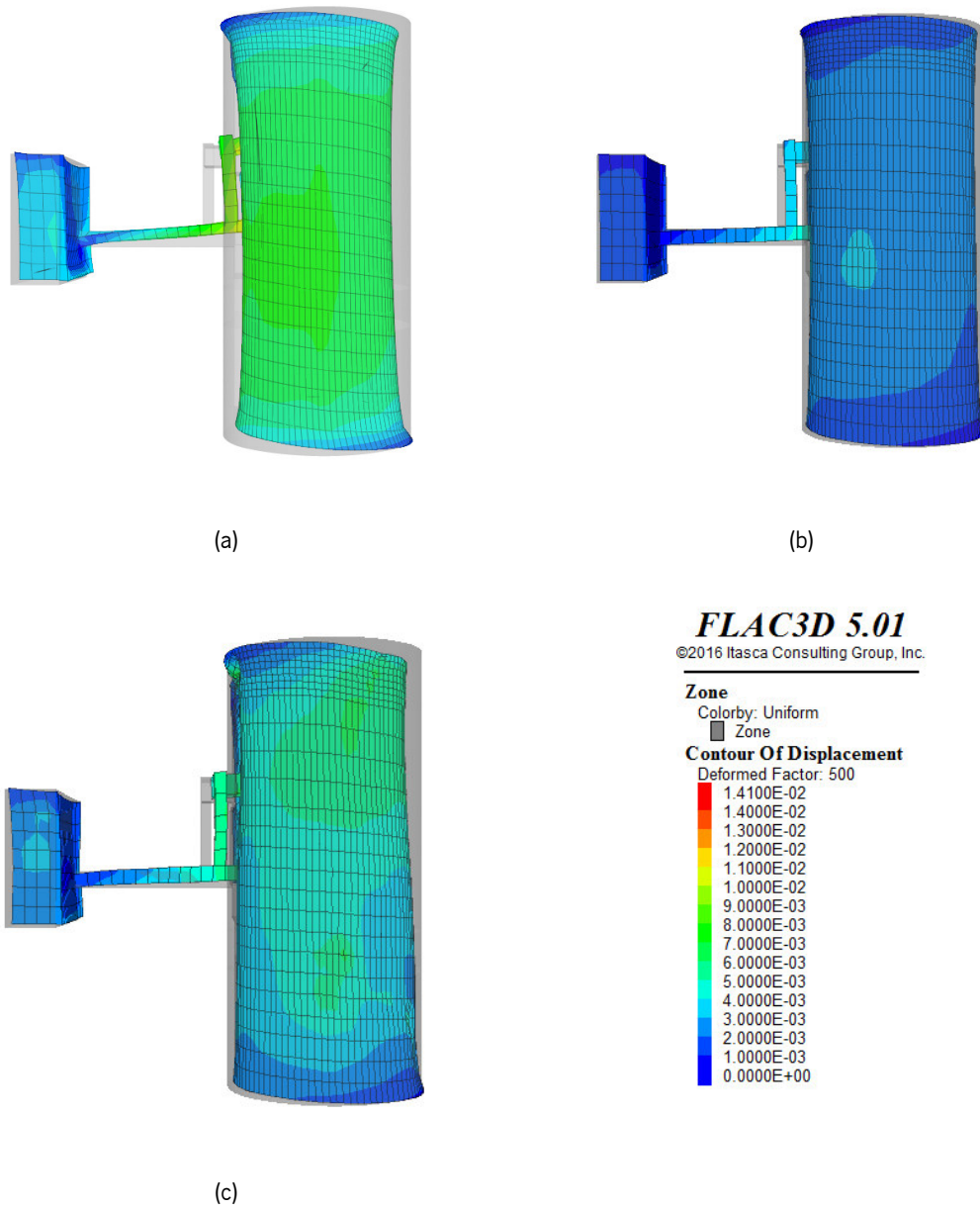


Figure 5.37 3D contours of the displacements magnitude (in m) with a deformed factor equal to 500 in the last excavation stage, for: a) the LNEC model; b) cluster 1 of data set 2; and c) the Homogeneous.

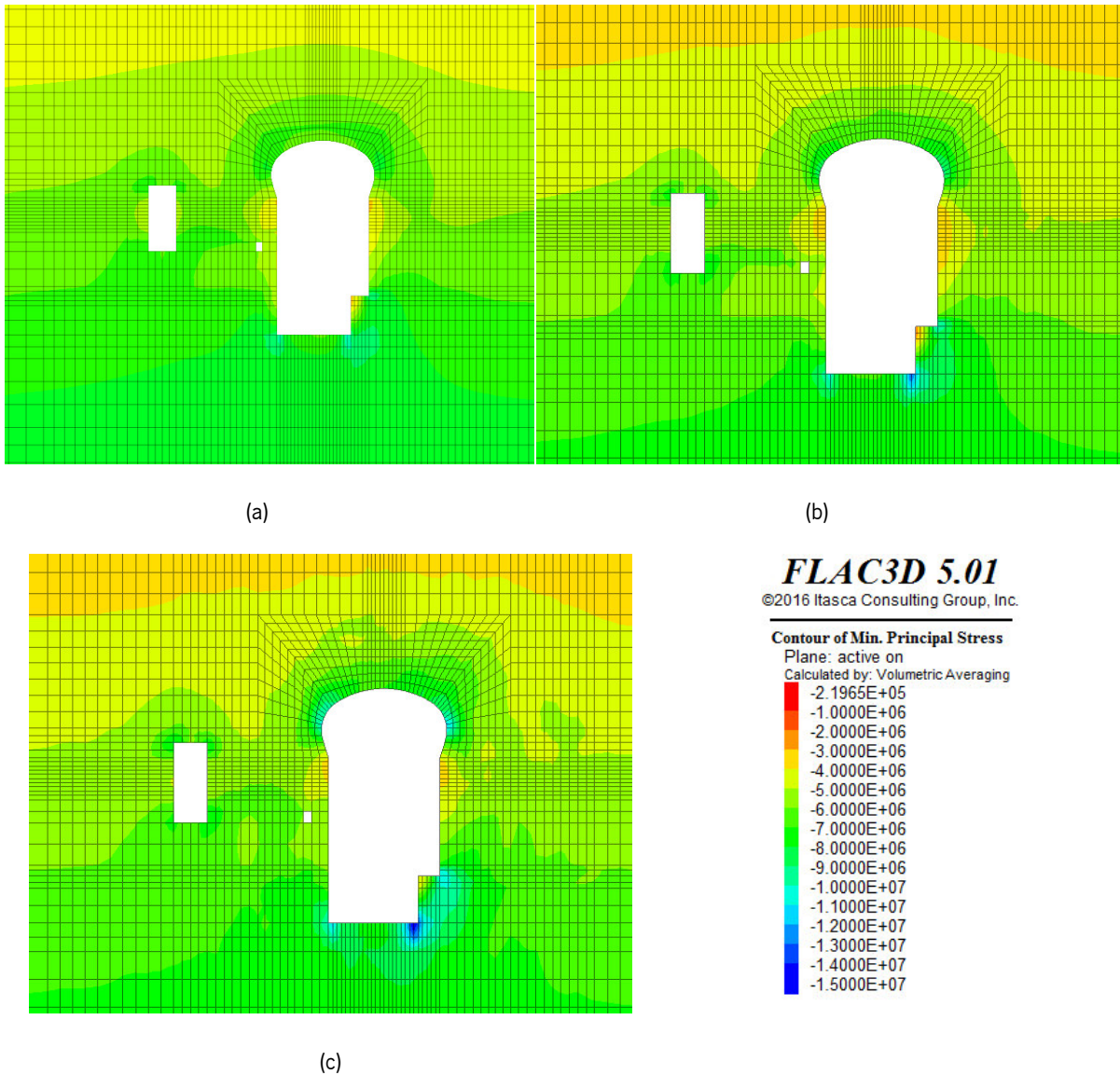


Figure 5.38 2D maximum principal stresses (in Pa) contours in section A at the last excavation stage, for: a) the LNEC model; b) cluster 8 of data set 2; and c) the Homogeneous model.

5.1.1.2 Comparison with the observed results

Similarly to subsection 5.1.1.1, in this subsection only the extensometers installed in section A and B will be presented for analysis. In this second data set, the results from LNEC (2013) will also be considered for the displacements comparison. Even though, note that the rock mass deformation modulus assumed in LNEC (2013) was lower than the one assumed for data set 1 and 2. The E_m value of 12 GPa was the result of a back-analysis performed using the extensometers displacements, therefore, one expects an almost perfect coincidence between the measured displacements and the ones displayed in LNEC (2013).

From Figure 5.39 to Figure 5.41 the results of all models are compared with the measured ones for extensometers E4A, E4B, E5A, E5B, E6 and E7 installed in sections A and B of the powerhouse cavern for the last excavation stage. From a qualitative point of view, all the clusters are able to

reproduce perfectly the trend of the measured displacements in all the extensometers unlike the LNEC resulting lines (dotted blue line). Looking closely, in extensometers E4A, E4B, E5A and E5B the clusters show better adjustments when compared with the Homogeneous model. In comparison with the LNEC results, for E6, cluster number 8 achieved a better adjustment than LNEC, and the same happens for E7 with cluster number 1 showing a MAD value lower than the one obtained by LNEC. It is important to stress out that the E_m values assumed for each cluster are at almost twice the ones assumed by LNEC.

The biggest difference happens for E6, with the clusters showing an average MAD value around 7 mm, while LNEC presents a value around 2 mm (see Table 5.26). In almost all the extensometers, the results from LNEC (2013) show adjustments above the measured displacements for extensometers E4B, E5A and E7 and the opposite for the remaining ones. In the middle sections of the extensometers, LNEC results are, most of the times, worse than the one obtained by the clusters, achieving a good fit near the cavern surface but missing the value as going deeper in the rock mass.

In an overall analysis, the average of the MAD values indicates LNEC (2013) as the best fitted model, with average values under 0.5 mm, followed by clusters 1 and 8 with average value of 2.012 mm and 2.154 mm, respectively. All the made statements can be confirmed in Table 5.26, where the clusters can be analysed individually and in average.

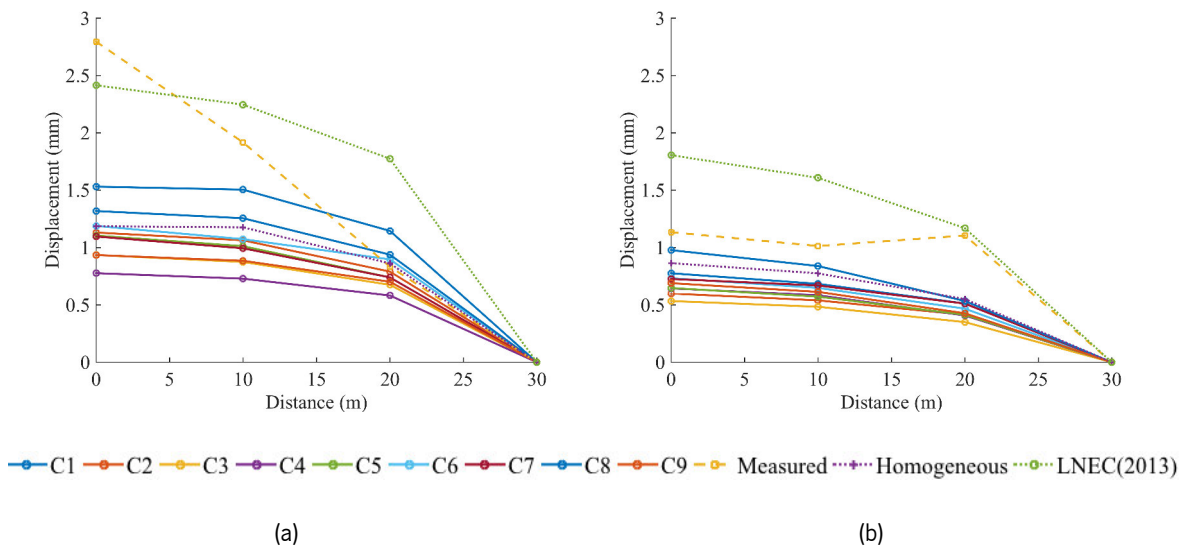


Figure 5.39 Total displacements (in mm) in the longer rod measured (dashed line) and computed (dotted and solid lines) for data set 2, for: a) E4A; and b) E4B.

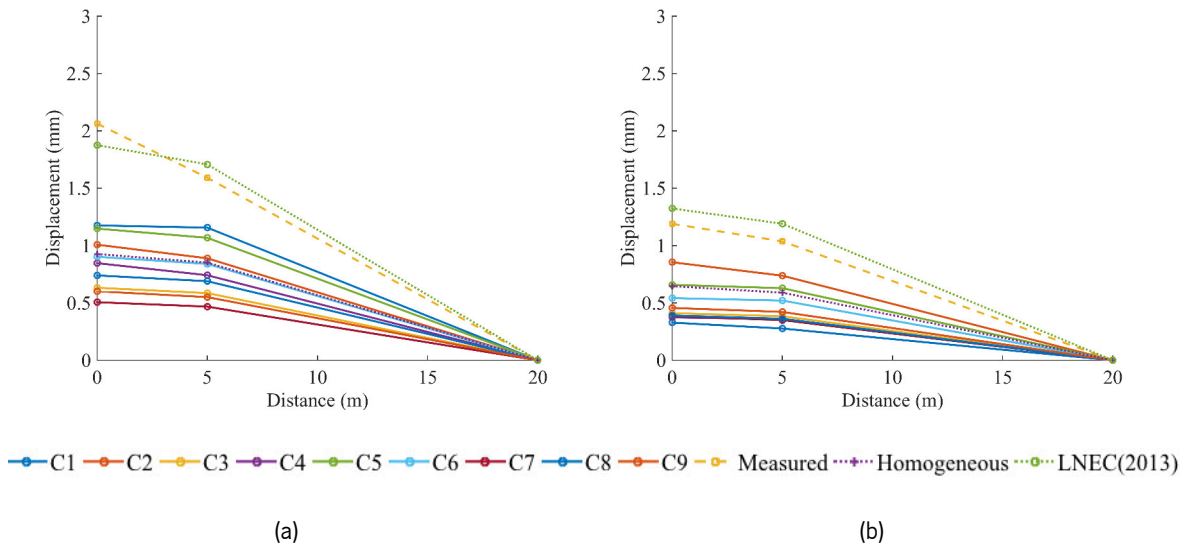


Figure 5.40 Total displacements (in mm) in the longer rod measured (dashed line) and computed (dotted and solid lines) for data set 2, for: a) E5A; and b) E5B.

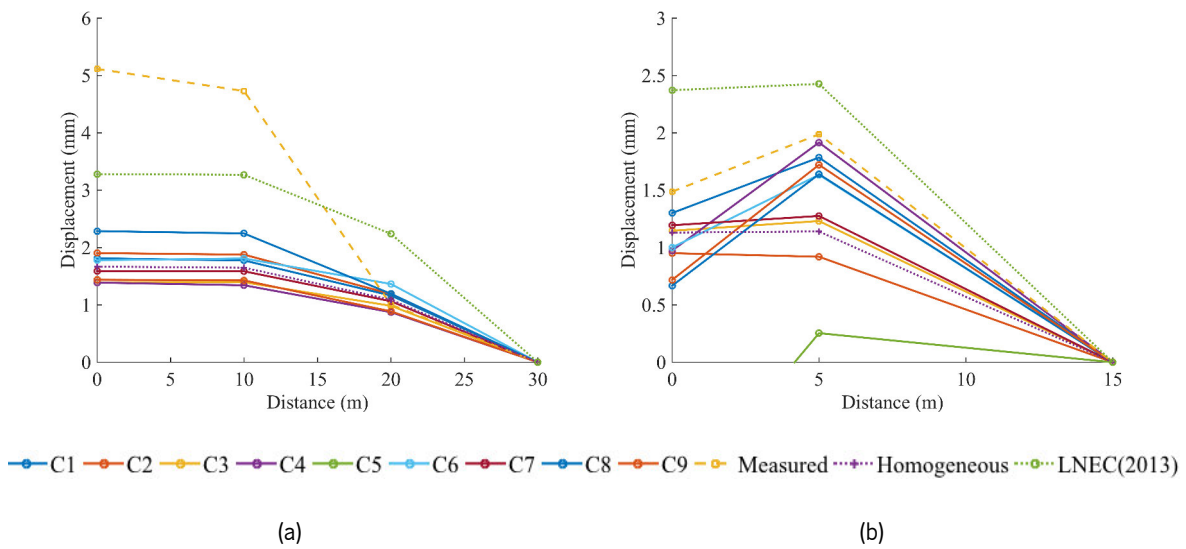


Figure 5.41 Total displacements (in mm) in the longer rod measured (dashed line) and computed (dotted and solid lines) for data set 2 in section A, for: a) E6; and b) E7.

Table 5.26 MAD values (in mm) obtain between the computed and the measured displacements for all the extensometers for data set 2.

	Hom o.	LNEC (201 3)	C1	C2	C3	C4	C5	C6	C7	C8	C9	Clust ers μ (mm)	Clust ers σ (mm)
E4A	- 3.015	0.889	- 1.365	- 2.561	- 3.061	- 3.458	- 2.690	- 2.387	- 2.716	- 2.034	- 3.025	- 2.589	0.582
E4B	- 2.802	1.330	- 1.281	- 1.700	- 1.885	- 1.617	- 1.619	- 1.404	- 1.346	- 0.902	- 1.522	- 1.457	0.269
E5A	- 1.840	0.070	- 1.322	- 1.757	- 2.436	- 2.066	- 1.439	- 1.912	- 2.680	- 2.225	- 2.502	- 2.038	0.446
E5B	- 1.926	0.288	- 1.621	- 0.634	- 1.432	- 1.495	- 0.940	- 1.164	- 1.492	- 1.474	- 1.348	- 1.289	0.301
E6	- 6.592	2.057	- 6.096	- 5.876	- 7.078	- 7.243	- 13.28 8	- 5.893	- 6.608	- 5.120	- 7.095	- 7.144	2.271
E7	- 1.336	1.324	- 0.388	- 1.604	- 1.097	- 0.585	- 4.495	- 0.840	- 1.004	- 1.168	- 1.035	- 1.357	1.157
μ (mm)	- 2.919	0.348	- 2.012	- 2.355	- 2.832	- 2.744	- 4.079	- 2.267	- 2.641	- 2.154	- 2.755	-	-
σ (mm)	1.740	1.165	1.866	1.671	2.004	2.187	4.278	1.697	1.899	1.403	2.059	-	-

5.6.4. Data set 3 vs. deterministic model

Finally, in this subsection, and similarly to the subsections focused in data set 1 and 2, the numerical results of data set 3 are presented. They contain the results from all the five clusters previously chosen in the scenario reduction methodology, as well as for the Homogeneous and Mean models. Similarly to data set 2 analysis, the model used in LNEC (2013) was used for comparison.

The main numerical results of all the mentioned models are provided in Table 5.27 and Table 5.28, where the first column shows the models cross section with the E_m distribution in a scale that varies from 0 GPa (blue) to 60 GPa (red). The main differences between data set 2 and data set 3 are the inclusion of the information obtained from the powerhouse front mapping, that had resulted in an increase of the E_m value. Therefore, and as expected, the displacements are somewhat higher than the ones obtained for data set 1 and data set 2.

The main differences to be highlighted for this last data set are that: 1) the Homogeneous and Mean models show similar displacements values in all the analysed points; 2) all four clusters give displacements magnitude that are twice higher than the one obtained by LNEC (2013) model, which would be expected since the E_m value is two times higher; 3) the differences between the four




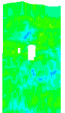
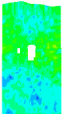
clusters are not as higher as expected, proving that , at the end, the simulated realisations had a lower dispersion of the E_m values. Looking closely, and in comparison with the Homogeneous model, the displacements are 50% higher for section A (cluster 4) and 21% higher for section B (cluster 1), while the absolute maximum difference happens between cluster 4 and Mean model, namely in the north wall, with the clusters giving a displacement 56% higher.

In what concerns to the stresses, the values are higher than the ones obtained by LNEC (2013); although, there are no important differences between the cluster, Mean and Homogeneous models (see Table 5.28 and Figure 5.45). The compression stresses reached 14 MPa in the left sidewall, while the shear stresses vary from 0.5 MPa to 1.2 MPa.

In terms of displacements contours, it is possible to confirm by Figure 5.42 and Figure 5.43 the one obtained for section A and B of the powerhouse cavern, as well as the cavern deformation in Figure 5.44. From the figure analysis, it is possible to observe that lower deformations happen for the homogeneous and data set 3 clusters, contrarily to LNEC (2013).

As with the table shown for data set 1 and 2, Table 5.29 shows a summary of the displacements reference values, namely the average, standard deviation, maximum and minimum values to use by the engineers *in situ*.

Table 5.27 Summary of the maximum displacements (in mm) obtained for data set 3 - 4 clusters, LNEC (2013), Mean and Homogeneous models and the differences (in percentage) in relation to the Homogeneous.

Scenario	Model	SA	SB	Control point	Left sidewall (LW)	Right sidewall (RW)	Arch (A)	Invert (I)	North top (NT)	South top (ST)
Homo.		4.46 RW	4.12 I	0.26	4.32	4.67	3.19	4.60	3.51	2.65
LNEC (2013)		12.90 (+189%) LW	11.20 (+172%) I	-	14.00 (+224%)	13.00 (+178%)	-	-	10.00 (+185%)	8.00 (+202%)
Mean		4.47 (-13%) RW	4.24 (-6%) I	0.24 (-8%)	4.60 (+6%)	4.70 (+1%)	3.20 (0%)	4.73 (+3%)	3.73 (+6%)	2.70 (+2%)
C1		4.75 (20%) RW	4.43 (-21%) I	0.21 (-19%)	4.40 (+2%)	4.90 (+5%)	2.95 (-8%)	4.80 (+4%)	3.82 (+9%)	2.60 (-2%)
C2		4.87 (49%) RW	4.55 (-18%) I	0.26 (0%)	4.91 (+14%)	5.14 (+10%)	3.70 (+16%)	5.00 (+9%)	4.80 (+37%)	2.90 (+9%)

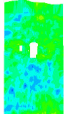
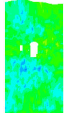
C3		4.73 (33%) RW	4.92 (20%) I	0.24 (-8%)	4.73 (+9%)	4.72 (+1%)	3.23 (+1%)	5.45 (+18%)	3.85 (+10%)	2.81 (+6%)
C4		4.62 (50%) LW	4.19 (20%) I	0.25 (-4%)	4.65 (+8%)	4.60 (-1%)	3.20 (0%)	4.60 (0%)	5.49 (+56%)	2.80 (+6%)

Table 5.28 Summary of the maximum principal stresses (in MPa) obtained for data set 3 – 4 clusters, LNEC (2013), Mean and Homogeneous models (all the values should be negative corresponding to compressive stresses) and the differences (in percentage) in relation to the Homogeneous.

Scenario	SA	SB	Control point	Left sidewall (LW)	Right sidewall (RW)	Arch (A)	Invert (I)	North top (NT)	South top (ST)
Homo.	10.0 	9.0 A/I	2.9	14.0	11.0	10.0	13.0	12.0	11.0
LNEC (2013)	7.0 (-30%) A		-	5.0 (-64%)	4.0 (-64%)	7.0 (-30%)	-	-	-
Mean	10.0 (0%) 	10.0 (+11%) 	2.9 (0%)	13.0 (-7%)	13.0 (+18%)	9.0 (-10%)	15.0 (+15%)	12.0 (0%)	11.0 (0%)
C1	10.0 (0%) 	10.0 (+11%) 	2.9 (0%)	10.0 (-29%)	10.0 (-9%)	9.0 (-10%)	14.0 (+8%)	11.0 (-8%)	10.0 (-9%)
C2	10.0 (0%) 	10.0 (+11%) 	2.9 (0%)	12.0 (-14%)	12.0 (+9%)	9.0 (-10%)	15.0 (+15%)	12.0 (0%)	11.0 (0%)
C3	10.0 (0%) 	10.0 (+11%) 	2.8 (-3%)	12.0 (-14%)	11.0 (0%)	11.0 (+10%)	14.0 (+8%)	12.0 (0%)	12.0 (+9%)
C4	10.0 (0%) 	12.0 (+33%) 	2.9 (0%)	14.0 (0%)	12.0 (+9%)	10.0 (0%)	13.0 (0%)	10.0 (-17%)	11.0 (0%)

Table 5.29 Summary statistics of the displacements (in mm) regarding the 4 clusters of data set 3 in all the analysed zones.

	SA	SB	Control point	Left sidewall (LW)	Right sidewall (RW)	Arch (A)	Invert (I)	North top (NT)	South top (ST)
Clusters μ	4.74	4.52	0.24	4.67	4.84	3.27	4.96	4.49	2.78
Clusters σ	0.09	0.26	0.02	0.18	0.20	0.27	0.31	0.70	0.11
Minimum	4.62 ⁷	4.19 ⁷	0.21	4.40	4.60	2.95	4.60	3.82	2.60
Maximum	4.87	4.91	0.26	4.73	5.14	3.70	5.45	5.49	2.90

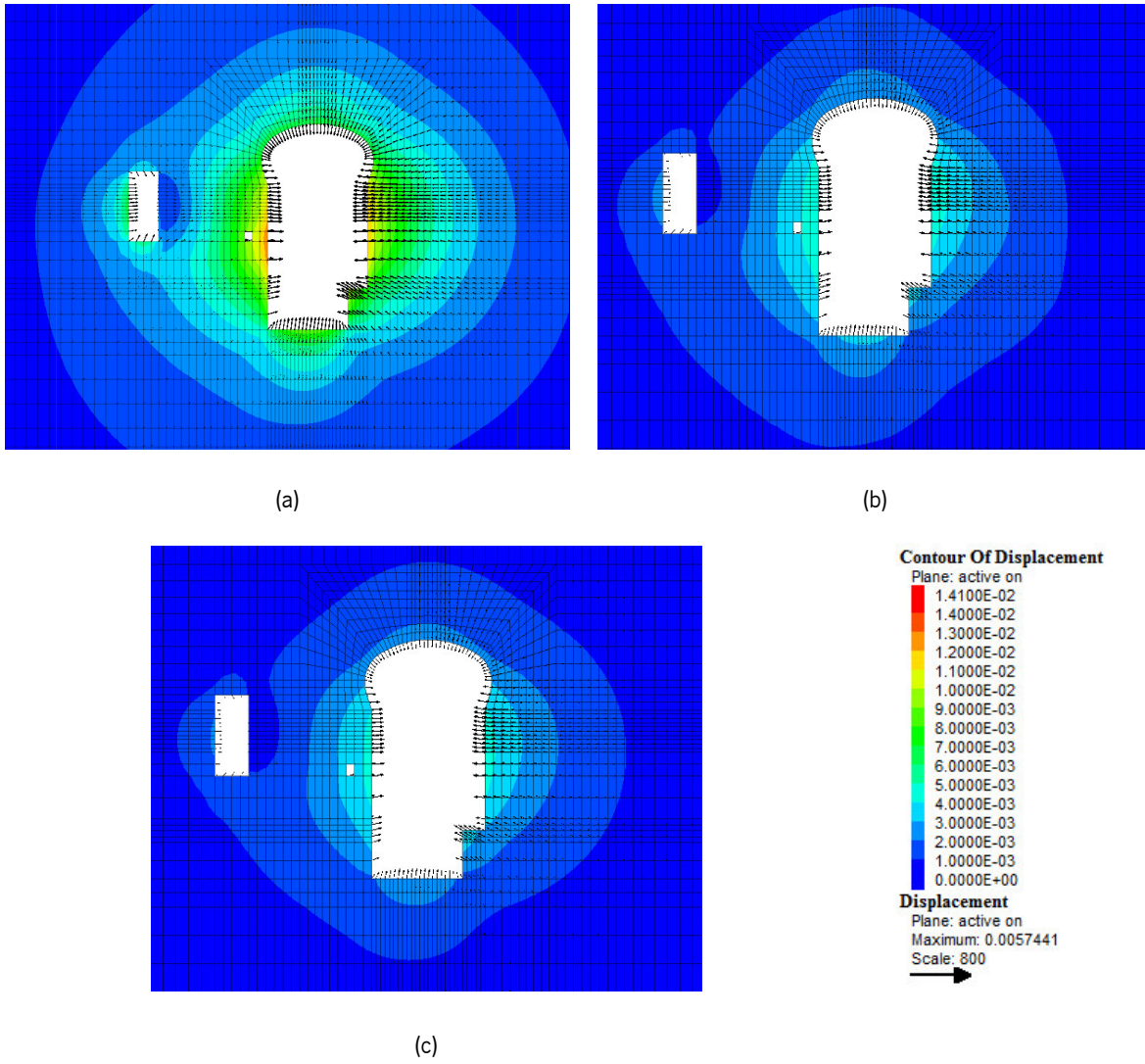


Figure 5.42 2D contours of the displacements magnitude (in m) and vectors for section A in the last excavation stage, for: a) the LNEC model (displacements with a scale of 400); b) cluster 2 of data set 3; and c) the Homogeneous model.

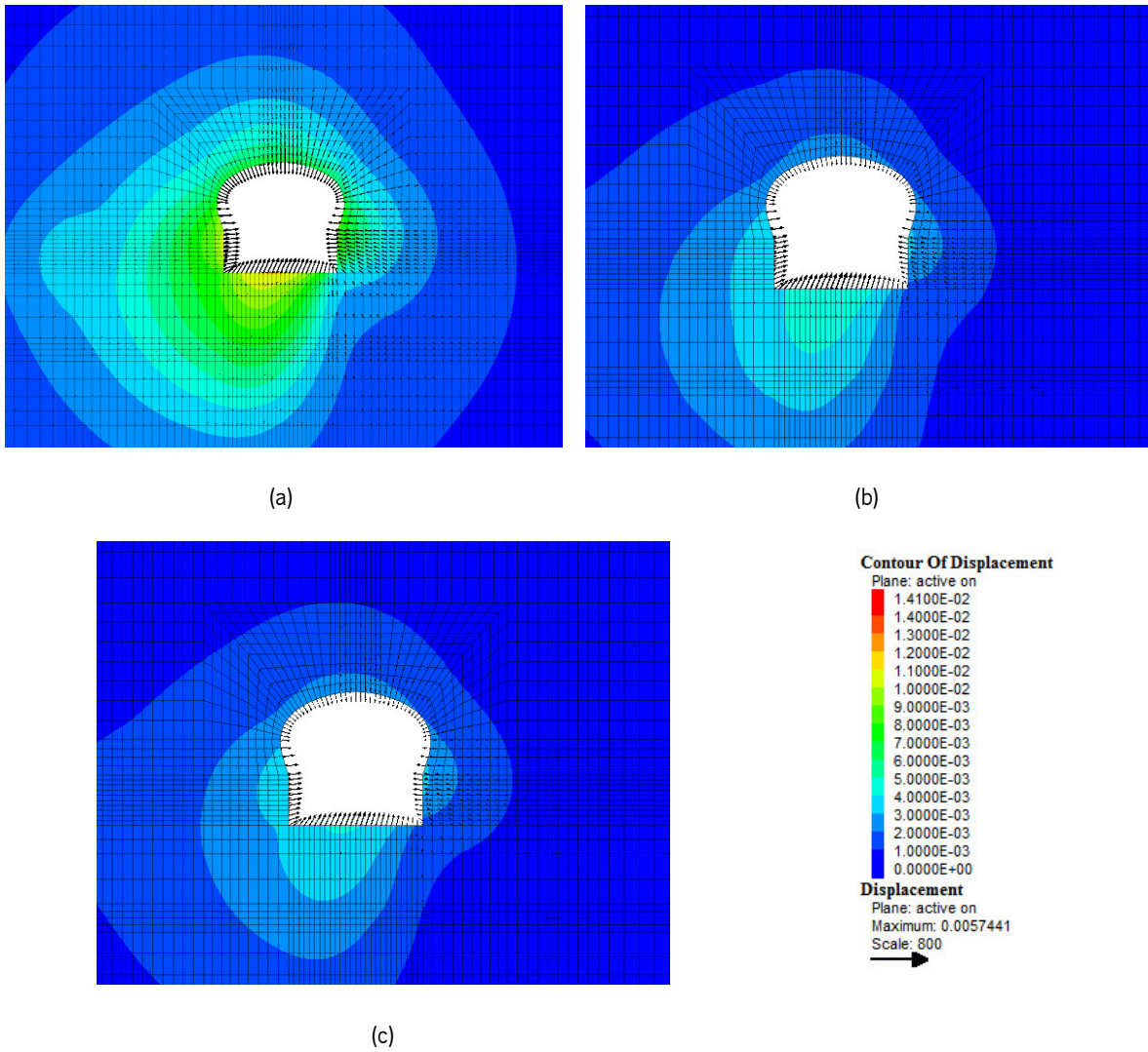


Figure 5.43 2D contours of the displacements magnitude (in m) and vectors for section B in the last excavation stage, for: a) the LNEC model (displacements with a scale of 400); b) cluster 3 of data set 3; and c) the Homogeneous model.

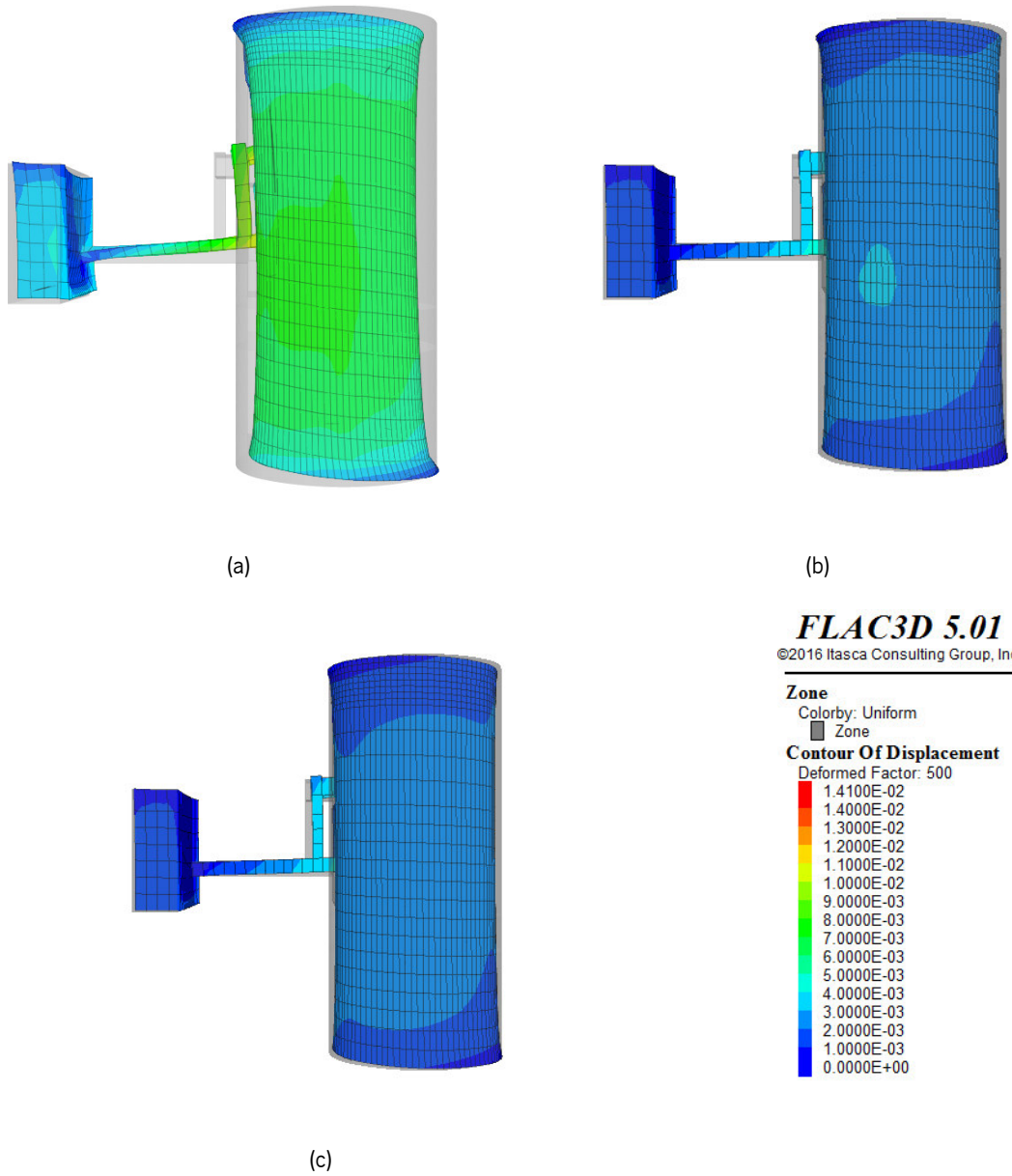


Figure 5.44 3D contours of the displacements magnitude (in m) with a deformed factor equal to 500 in the last excavation stage, for: a) the LNEC model; b) cluster 2 of data set 3; and c) the Homogeneous.

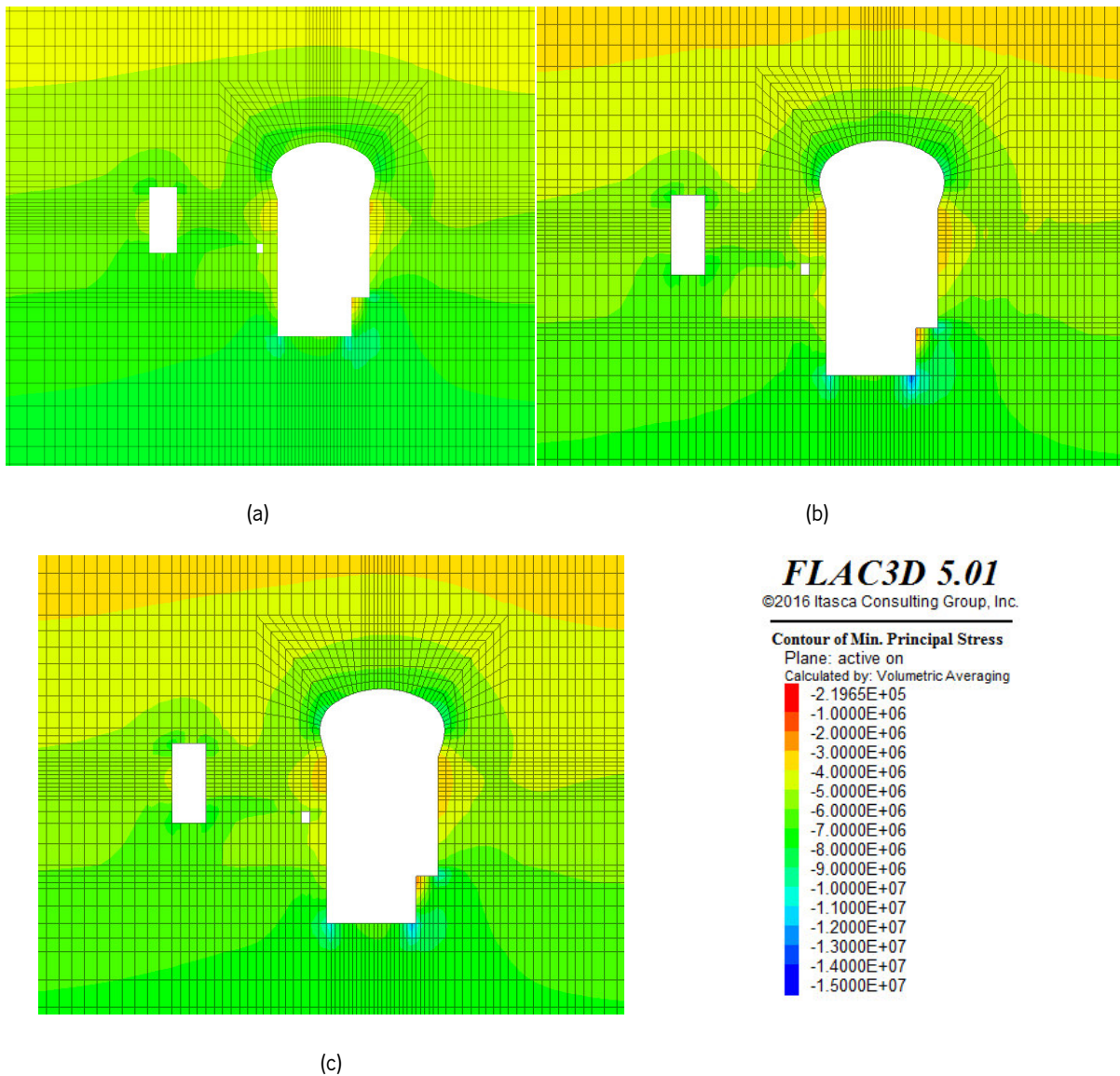


Figure 5.45 2D maximum principal stresses (in Pa) contours in section A at the last excavation stage, for: a) the LNEC model; b) cluster 1 of data set 3; and c) the Homogeneous model.

5.6.5. Discussion of the results

Due to the number of models considered for the numerical analysis, the discussion of the results will be focused on the main ones, that is data set 1, 2 and 3. Firstly, the numerical results will be commented for data set 1, highlighting the heterogeneity representation of the individual realisations when compared with a deterministic approach. Then, for data set 2 and 3 a comparison will be made with the numerical results obtained by LNEC (LNEC, 2013). Since an observational campaign was carried out during the cavern excavation, the displacements measured by extensometers will also be integrated for analysis. Finally, some evaluations will be made about the use of clusters to identify the heterogeneity and spatial variability of the rock mass of Salomonde II.

In what concerns the first point, and for data set 1, it was possible to observe from Table 5.21 that the adopted three clusters (individual realisations) provided good results in the definition of a lower and upper bound for the cavern displacements, as well as for the maximum principal stresses. It seems relevant to mention, that comparison with the Homogeneous model the displacements disparities reached 119%, with cluster 2 computing a displacement of 6.8 mm in the cavern south top and the Homogeneous model a smaller value of 3.1 mm. Analysing this difference from a geotechnical point of view, a deviation of almost 4.0 mm makes a relevant difference (in the millimetres' magnitude). Regarding the Mean model, as expected, the displacements and principal stresses values obtained agree with the ones from the Homogeneous model, confirming, once again the limitation of using the 100 realisations average values.

The displacements results, when compared with the measured displacements during the observational campaign, were not as good as expected, namely for the extensometers instated in section A; however, on most of the extensometers the qualitative trend is achieved and in the ones of section B the displacements magnitudes are achieved. In relation to the Homogeneous model, for most of the clusters were registered lower average MAD values. Overall, the significant differences between the clusters and the measured displacements happen, mainly in the first point, near the cavern surface, and can be explained by some disturbance provoked on the rock mass during the excavation.

With regard to the second point (data set 2 and 3), here assumed as the main data set, the numerical results from data set 2 are in line with the ones obtained for data set 1. However, and according with the rock mass empirical systems, the E_m value are higher for data set 3 and, consequently the displacements lower. In a general way, for data set 2 the displacements differences, in comparison with the Homogeneous model, are in all the clusters, above the displacements and principal stresses observed with a deterministic approach. Although the same does not happen in data set 3, where the Homogeneous model presents, in a semi part of the time, lower or higher values in comparison with the analysed clusters.

In LNEC model, the maximum displacement reached a magnitude of 14.0 mm in the left sidewall, while cluster number 3 shows a maximum value of 8.4 mm in a different point, the cavern arch. Similarly, in the cavern north top, while LNEC resulted in a 10.0 mm displacement, cluster number 1 showed a 16.2 mm magnitude, which is higher than expected.

For data set 3, with an increase of the rock mass deformation modulus for almost three times the value assumed in LNEC (2013), and since a linear elastic behaviour of the rock mass was assumed,

the expected displacements would be approximately one third of the ones obtained by LNEC; however, that effect did not happen in a linear way. Indeed, the maximum displacement in the north top is only 50% lower than the LNEC result (10.0 mm to 4.8 mm), which allows to conclude that the individual realisations can, by fact, identify and reproduce the spatial variability of the E_m . All the mentioned differences can be analysed in more detail in Figure 5.46, which includes the minimum mean and maximum displacements values in all analysed models, as well as the maximum displacement registered by LNEC(2013) model.

In terms of the comparison with the observed displacements for data set 2, the same qualitative trend is maintained, with a more succeeded adjustment in extensometers installed in section B. In the remaining extensometers, the results were worse, mainly in the first distances (0 m and 10 m). Regarding results obtained by LNEC, and since they used these observed displacements to obtain the E_m value by back-analysis techniques, it would be expected almost perfect adjustments in the extensometers (see Table 5.30); however, the obtained lines do not show the same qualitative trend as the one obtained using the clusters. Yet, analysing the average error between the considered extensometers this was the model with the lowest error, 0.348 mm. In what concerns data set 2 almost all clusters models, showed an average deviation closer to 2 mm, with the lowest value observed for clusters 1 and 8.

It is important to stress out that in underground works, a deviation of 1 or even 2 mm can be considered residual. Note, once again that the magnitude of the displacements (mm) is very small, which gives sufficient prominence to the extensometers electronical errors. Also, and because is a millimetre scale, temperature changes inside the rock mass can truly influence the displacements results, leading to measured higher values.

Table 5.30 Average values of the MAD mean values (in mm) obtain between the computed and the measured displacements for all the extensometers and the standard average values.

	Homo. 1/2	Data set 1	Data set 2	LNEC (2013)
Mean of μ	-2.919	-2.183	-2.645	-0.348
Mean of σ	-1.740	-0.357	-0.838	-1.000

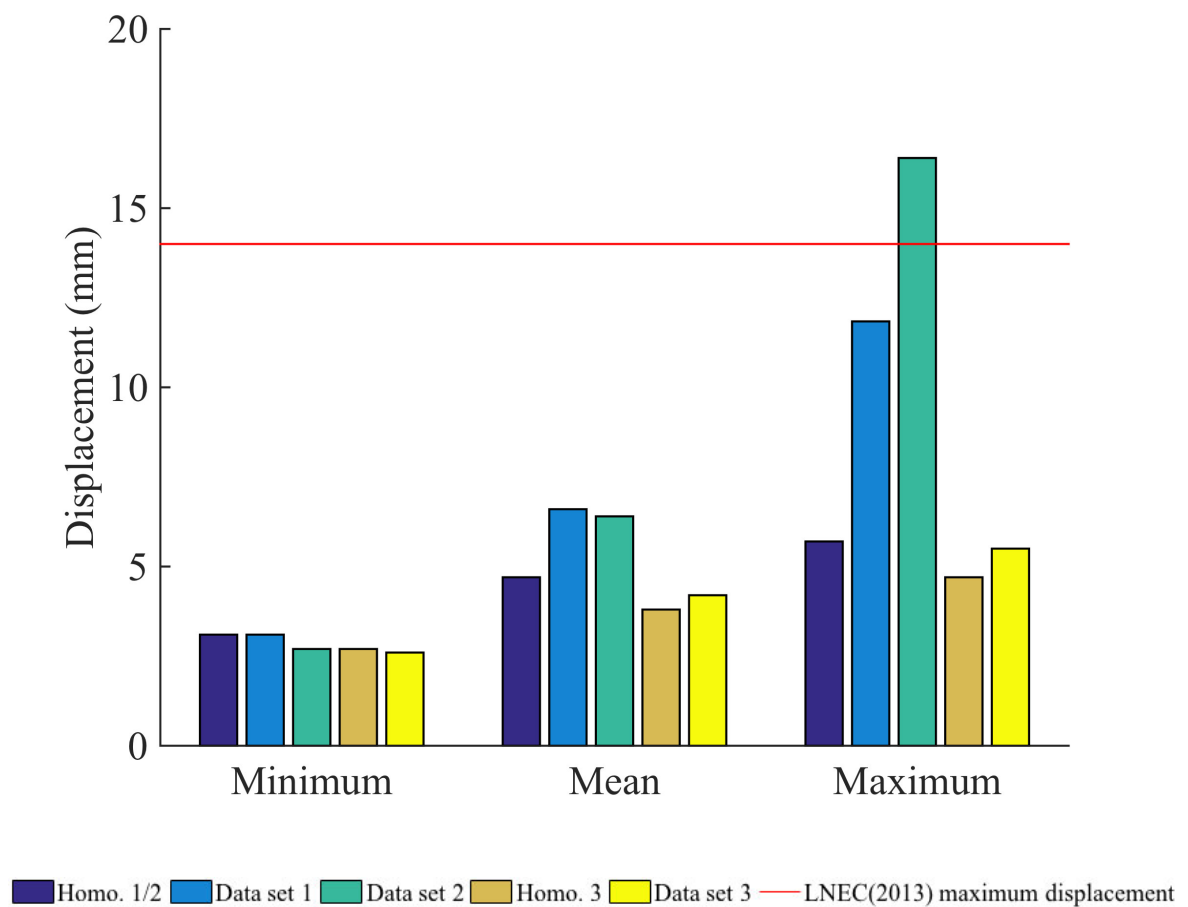


Figure 5.46 Graphical representation of the minimum, mean and maximum values of the maximum displacements registered in each analysed numerical model (data set 1, 2 and 3 and corresponding Homogeneous models). The red line represents the maximum displacement obtained by LNEC(2013) model for terms of compression.

5.7. CONCLUSIONS

In order to validate the numerical methodology proposed in Chapter 4, a complex underground work had to be considered. Therefore, the electrical power reinforcement of Salomonde was the selected case study to apply the developed methodology. This new complex of underground structures, are located near the Salomonde dam as denotes the name of Salomonde II. From all the underground structures the one chosen for numerical analysis was the powerhouse cavern and some auxiliary structures, namely the floodgate chamber and some access tunnels.

The reinforcement works started in 2009, where an initial campaign of characterisation of the rock mass was carried out (phase 1 of the information), mainly through mechanical boreholes executed from the surface. From there, it was possible to evaluate a total of 81 points with information of the E_i , 1570 with the information of W and 1568 with the information of RQD. In a second phase, and

during the cavern excavation more information was gathered from the cavern front. This time information about the RMR, Q and GSI, all empirical systems used to characterise the rock masses. Since both phases of information have different geotechnical parameters and quantity of information the geotechnical parameters were, firstly geostatistical simulated and then, using different empirical formulas, they were converted into the geomechanical parameters, namely the rock mass deformation modulus. At the end, and in order to take into account both phases of the information, three data sets of E_m values were created, data set 1 that used information from phase 1, data set 2 using the information from phase 1 plus some additional laboratory tests performed in the extensometers and piezometers logs, and finally, the third data set where the information regarding the empirical systems were added. Since the laboratory tests the front mappings give different type of geomechanical information, it seemed relevant to perform an individual analysis of data set 2 and 3.

A main step of the numerical methodology goes through the geostatistical simulation of the geomechanical or geotechnical information and, in this Chapter, it was done individually in several geotechnical parameters. The simulation was performed using as target points the Flac^{3D} mesh developed by Espada and Lamas (2014). Then, after the use of the empirical formulas, the 100 realisations of the E_m were reduced to a small number obtained after a scenario reduction methodology is applied. A lower bound of 3 clusters was defined for this methodology application, and the optimal number of clusters obtained for the third data set (2 clusters) was not used afterwards. As a consequence, a total of 3, 9 and 4 clusters were selected to represent the full realisation set of data set 1, 2 and 3, respectively. These scenarios were validated using the percentiles 10, 50 and 90, where quasi overlapping lines were obtained between the clusters configurations and the full realisations set. This good fitting can be attested by the E_m distribution of the values obtained using kernel smoothing function. At this point, with the scenario reduction results, it was possible to affirm that the rock mass of Salomonde II should be well characterised, represented by geostatistical individual realisations that allows the uncertainty reduction and consider the existence of variability in the E_m spatial distribution.

Finalised the clusters configurations selection and validation, the numerical analysis was performed in several models using Flac^{3D} (three for data set 1, nine for data set 2 and four for data set 3). Aiming to demonstrate the advantages of using this new numerical methodology, the numerical analysis of a Homogeneous model was carried out for the different data sets. Hence, and as the individual realisations that compose the clusters configurations give results that always bounced the homogeneous ones, it is possible to state the importance of having a wide range of possibilities to

use in underground works, mainly because the deterministic approach can compute displacements and principal stresses value that are underestimated.

A relevant point to stress out regards the numerical results from data set 1. The data set is composed by information gathered in a prospection phase and the numerical results are similar to the ones obtained when more information is added (data set 2). This can also be attested by the comparative analysis made with the extensometers measured displacements, reaching in each cluster better adjustments than the ones obtained with the deterministic approach. This fact shows the power of the developed methodology that, even with a small amount of data, is able to reach coherent results.

Moreover, it seems important to mention the advantages of using the individual realisations, mainly for the powerhouse cavern deformation analysis, where some irregularity zones can be identified and justified by the lower E_m values. This is from a great utility to identify zones more unstable in the rock mass.

Finally, note that considering the geotechnical data the estimated deformation modulus to consider into the numerical analysis would be closer to 30 GPa, although by using this methodology it was possible to obtain results that are closer to the ones obtained by LNEC that used the observational displacements of 12 extensometers to perform a back analysis of E_m . For this reason, and others already mentioned, it seems unmistakable to affirm that the findings of this Chapter show the utility and advantages of characterising the rock mass with the new proposed methodology, through which the heterogeneities/variability can be identified and reproduced. Even more, when compared with a deterministic approach, as proven by the Homogeneous model results.

Chapter 6

METHODOLOGY FOR OPTIMISATION OF BOREHOLES PLANS

6.1. INTRODUCTION

Prospection plans have the difficult task of ensuring a complete and rich characterisation of the rock mass providing a fair balance between costs optimisation and safety in geotechnical projects. In a desirable scenario, the number of boreholes to perform *in situ* should be as high as possible; however, in the reality of geotechnical works, this is not the case. The costs associated with boreholes execution in rock masses are considerably high, which limits their number. Moreover, due to the frequently large spacing between boreholes and the fragmentary nature of the obtained data, considerable uncertainties affect the geotechnical models, mostly in highly heterogeneous rock masses. Currently, boreholes location and depth are mainly defined based on experience and *know-how* of professionals, as such they are user-dependent.

Therefore, the search for more rational and balanced ways of planning the borehole locations, as they can provide higher quality data and decrease the uncertainties, is of utmost importance, essentially in large geotechnical projects, where the time and money available for rock mass model construction is short.

The geotechnical prospection plans in large geotechnical works are generally divided into two main phases: the initial phase, where a preliminary and confined characterisation of the rock mass is carried out, and a second phase, where the number of executed boreholes, as well as laboratory and *in situ* tests, are significantly higher and performed according to the necessary detail level. Thus, the application of a borehole plan optimisation methodology in geotechnical engineering is suitable in the second phase where the preliminary geotechnical information resulting from phase one can be used. After performing a literature review, a gap of consolidated methodologies for this purpose and to help professionals to optimise the boreholes position in the second phase of the prospection works by giving them information regarding the borehole quantity and depth was identified.

In this search, a few existing methodologies for boreholes optimisation combining different types of algorithms, in which the goal consists in minimising a wide range of uncertainty measures, were found. Some of the used measures are obtained by using geostatistical techniques, e.g., the variance of spatial prediction errors or other metrics obtained through kriging (McBratney et al., 1981; Scheck and Chou, 1983; Englund and Heravi, 1994; Van Groenigen et al., 1998; Marchant and Lark, 2005; Brus and Heuvelink, 2007; Soltani and Hezarkhani, 2009, 2013a, 2013b).

For detail, from all the mentioned works, authors like Olea (1984), Englund and Heravi (1994), Marchant and Lark (2005) and Soltani and Hezarkhani (2009, 2013a, 2013b) proposed interesting approaches to reduce the uncertainty associated with boreholes placement.

First, Olea (1984) presented a method to minimise the sampling requirements necessary to estimate a mappable spatial function at a specific level of accuracy based on the technique of kriging. This work uses the average standard error as global indices of sampling efficiency. As a result, an optimised pattern for the wells was obtained.

Subsequently, Englund and Heravi (1994) following Olea's work, applied the methodology to predict subunits of the site that need remediation (second phase) using as information the data obtained from remediation campaigns performed in a called phase one. At this stage, a cost-effectiveness optimisation was performed to assess the number of samples and phases required to remediate contaminated soil balanced with reasonable costs.

In 2005, Marchant and Lark developed an approach to optimise the sampling scheme used to build the variogram. The main goal was to understand what type of sampling scheme, of regular or iterative, could result in more accurate variograms to use in further simulations and, consequently, reduce the sampling costs. In each phase the information from previous phases is used to generate new information and to decide if a new phase is required. As an objective function, the authors minimised the developed expression to evaluate the uncertainty of variogram parameter estimation (e.g. sill, range, nugget, etc.).

Later, in 2009, Soltani and Hezarkhani proposed a more complex methodology upgrading to a *3D* application. In this work, the authors assumed as an objective function the maximisation of the kriging variance reduction, a metric calculated after dividing the kriging variance values obtained from the initial samples with the new additional boreholes. The goal was to find the optimal pattern for new boreholes to execute in gold and copper deposits using the Simulated Annealing optimisation algorithm.

In 2013, the same authors published two more works based on the previously proposed methodology, this time combining the simulated annealing algorithm and an objective function to assess the realistic value information that additional boreholes will bring for the deposit characterisation. So, the objective function considered was the range of reliability of each individual block resulting from their prediction.

In these recent papers, Soltani and Hezarkhani stated that, hitherto, the published works share the same limitation since they use the kriging variance value as the objective function and, therefore, they proposed the use of a different metric that in fact also results from the kriging prediction.

This fact leads to the identification of a weakness in this domain, mostly because the use of prediction only results in one possible “reality” for the random field and the obtained metrics do not reflect the local variability of the regionalised variable under consideration (Chilès and Delfiner, 2012). Thus, the replacement of kriging by geostatistical simulation that results in several realisations and, consequently, in a greater uncertainty reduction and accuracy in the spatial variability quantification of a random field, is a way to overcome those limitations. In what concerns the optimisation algorithm, almost all mentioned works used the Simulated Annealing; therefore, and since the goal is to search for the global minima, the stochastic algorithm was the best choice to use in the optimisation process of the proposed methodology. In Figure 6.1 a summary containing all the mentioned works, as well as the identification of all the methodologies limitations, is presented. The shown table served as a basis for the development of the boreholes plans optimisation methodology presented in this Chapter. The main purpose of the proposed methodology is to work as a helping tool in supporting the decision-maker when defining the prospection plans and to provide as main output the optimal position for additional boreholes, the gain in terms of geotechnical details and the minimum depth of each borehole.

This Chapter is organised as follows. Firstly, the chosen optimisation algorithm is presented in a general way in subsection 2. Then, the proposed methodology and all the detail regarding its application are exposed in subsection 3, followed by the application to a case study described in subsection 4. Discussion and conclusions follow in the last subsections.

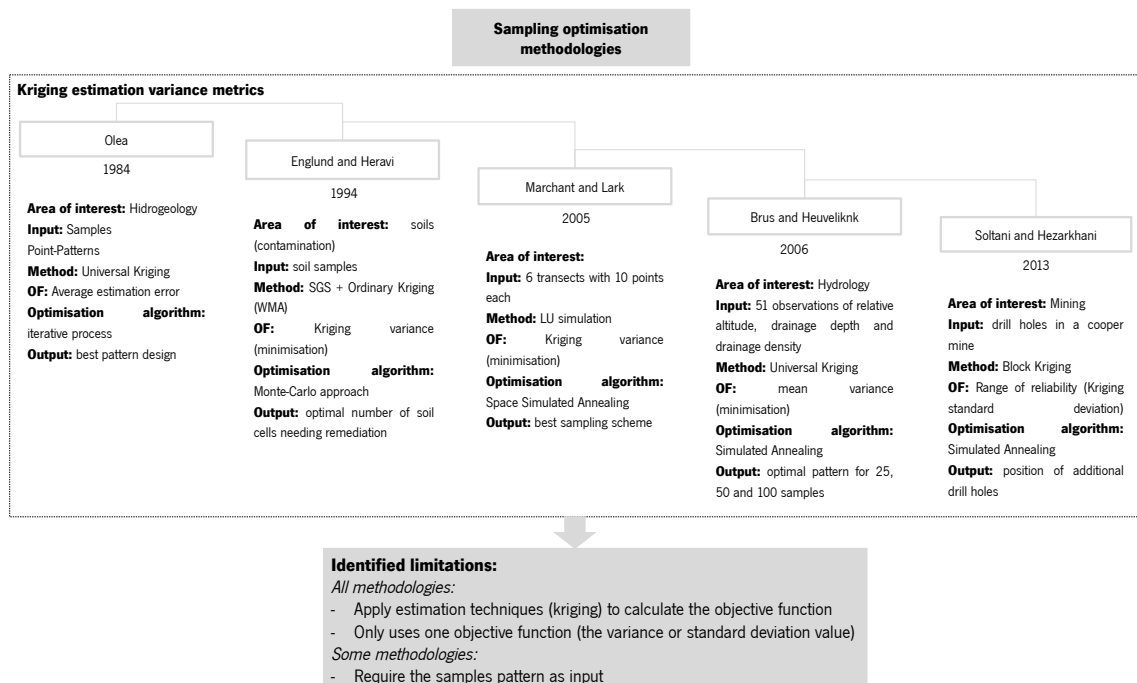


Figure 6.1 Identification of all the optimisation methodologies and their limitations along the years.

6.2. OPTIMISATION CONCEPTS

6.2.1. Simulated Annealing

In the metallurgical industry, a thermal process named annealing aims to forge iron in order to minimise the energy spent to cool and freeze the metal. The metal is heated to a maximum temperature able to change its physical properties (creating a particle disorder) and followed by slow cooling to guarantee that the final configuration of the solid is structurally superior. Simulated annealing (SA) (Kirkpatrick et al., 1983) is an iterative algorithm to solve combinatorial optimisation problems inspired in this process in order to find a balanced state for each temperature, this way minimising the internal energy of the process. In engineering problems, the use of SA has been increasing once it is an alternative to gradient-based methods or other local classical methods that can be trapped in local optima.

This algorithm starts by randomly generate a solution, the so-called new solution (j) that emerges after random changes in the parameters that generate the previous solution (i). Then, in the case of facing a minimisation problem, SA compares the objective function (OF) values for each solution. On the one hand, if $OF_j \leq OF_i$, solution j is automatically accepted and assumed as the temporary best solution. Under these conditions, the algorithm jumps to another iteration and new solutions are generated. On the other hand, if $OF_j > OF_i$ there is a possibility to accept solution j , even if it is a “worse” solution than solution i . By allowing these controlled uphill moves to counter the downhill moves, the algorithm is forced towards the global minimum that sometimes can be found near the worst solutions. This selection is made through the calculation of an acceptance probability (P_{accept}) that depends on a temperature parameter that decreases in a slow rhythm to avoid, once again, the algorithm to be trapped into a local minimum:

$$P_{accept} = \begin{cases} 1 & OF_j \leq OF_i \\ e^{\frac{-\Delta O}{T}} & otherwise \end{cases} \quad (6.1)$$

where ΔO represents the difference between the OF values in the states i and j ($\Delta O = OF_j - OF_i$) and T is the control temperature. In the case of lower temperature values, the probability of accepting worse solutions is also lower, allowing SA to converge faster (see Figure 6.2).

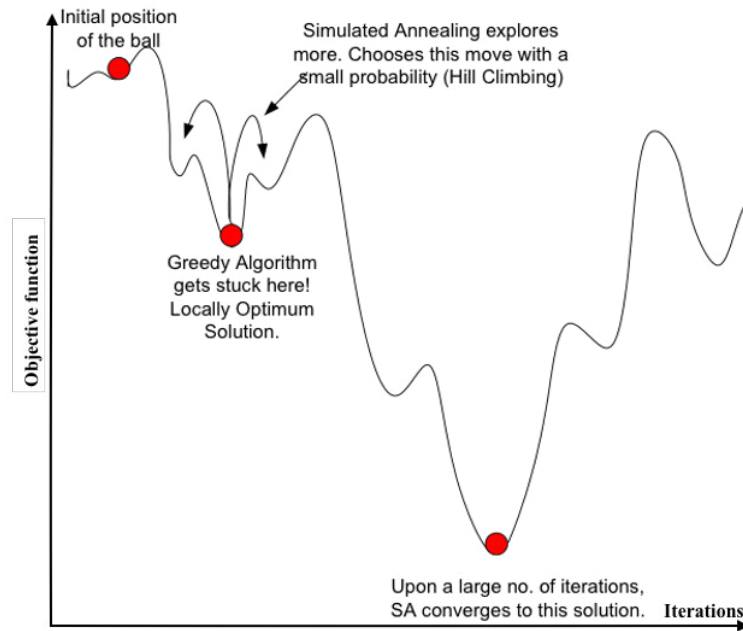


Figure 6.2 Simulated Annealing functioning method - hill climbing (adapated from Zhigljavsky, 1991).

As previously mentioned, the initial temperature to start the process must be high enough to allow SA exploring all the space of solutions. However, during the process, the temperature is progressively reduced until a threshold value defined by the user. This cooling should be slow in order to avoid rushing the stopping criteria of the algorithm, e.g.:

$$T_j = \alpha \times T_i \quad (6.2)$$

where T_i represents the temperature value used when solution i is generated and α represents the cooling constant, whose values usually range from 0.80 to 0.99 for a fast and slow cooling, respectively (Aarts and Korst, 1989).

Besides the previously mentioned parameters, other should be defined:

- A perturbation or transition kernel, which indicates the mechanism used to generate a new solution to be tested given a current solution;
- A maximum number of allowed moves for each temperature value. This number translates the number of times that SA generates new solutions before decreasing the temperature. Once reached this maximum number of moves, the temperature is decreased using the aforementioned cooling process. Additionally, the maximum number of accepted solutions given the same temperature value must also be defined.
- A stopping criterion: this criterion can be defined based on one or more key parameters. Many authors stop SA by defining a final value for the temperature, while others add more

criteria to the process (Yang, 2010; Brus and Heuvelink, 2007; Hernandez and Emery, 2009). In the present case, the stopping criterion will be composed of a temperature, an iteration number and a maximum number of rejections within a given temperature state.

The latter will allow stopping the algorithm if no progress is shown;

To sum up, the algorithm should be applied using the following general steps:

Step 1: The values of the above key parameters (initial temperature, maximum number of iterations, cooling constant and maximum number of rejections in a given temperature state) must be set. This can be defined after performing a sensitivity analysis to identify which parameter values allow the convergence of the algorithm to the global minima, or using as a reference existing values for similar optimisation problems.

Step 2: In the first iteration, an initial solution i should be randomly generated within the space limits. For the remnant process, the new generated solutions are based in random changes made to the current solution.

Step 3: The objective function is calculated for the new solution j and the decision of whether or not moving to the new solution is made.

Step 4: The temperature is decreased until reaching the thresholds defined for the stopping criterion, for which the algorithm gives the near-optimal solution for the optimisation problem.

In what concerns the algorithm advantages, it is important to highlight the fact that a mathematical model is not required, *i.e.* the algorithm only needs to perturb and evaluate the optimisation function without further details on the problem, which in real-life problems, where the amount of information is limited, is very useful.

6.2.2. Multi-criteria optimisation

With the growth in complexity and plurality of perspectives of engineering problems, the need to optimise more than one function or variable has increased. Therefore, the use of a multi-criteria approach allows a better understanding of the reality and supports the user in the decision because most of the functions in an optimisation process are in conflict, *i.e.* while some objective functions increase, a reverse movement may happen for the remaining. Since a unique and admissible solution does not exist for this type of problems, the approach main goal goes through finding the best solution that can, at the same time, satisfy all the objectives. In essence, a multi-criteria

approach entails an advantage not only in the problem conception, but also in the critical analysis once a range of solutions with different characteristics is obtained.

A wide range of methods can be applied to solve this multi-criteria optimisation problem, some more classical and others that appeared more recently. In the first group, the classical methods, the aim is to find the ideal solution by minimising an objective function $f(x)$, which is a vector containing all the decisions variables x . Here, the ideal solution should be the one corresponding to the minimum value in all the decisions variables. To help in the optimum identification, Pareto (1912) presented a concept to identify compromise solutions where each solution $x^* \in X$ (X represents the feasible space - set of constraints) is a Pareto optimum for a multi-objective problem if all other solutions $x \in X$ have:

$$f_i(x) \geq f_i(x^*) \text{ with } i \in I = [1, 2, \dots, k] \quad (6.3)$$

Or, at least there exists one objective function $i \in I$ such that,

$$f_i(x) > f_i(x^*) \quad (6.4)$$

In the previous formulations, the letter i denotes the number of objectives that compose the multi-objective problem and k the number of functions that composes the vector function I .

Examples of the classical methods that could be applied are:

- Weight Sum Method;
- Hierarchical Optimisation Method;
- Trade-off Method.

The Weight Sum Method proposed by Zadeh in 1963 transforms a set of objectives of an optimisation problem into a single and scalar objective function, where each decision variable receives a weight value (α_i) in accordance to its importance in the subject. This relation is given by:

$$f(x) = \sum_{i=1}^k \alpha_i f_i(x), \text{ where } \sum_{i=1}^k \alpha_i = 1 \text{ for } \alpha_i \geq 0 \quad (6.5)$$

Furthermore, to apply the method, it is important that all the decision variables are expressed in a similar unit or, if needed, standardise the set of objectives. This method represents the simplest and fastest way to transform a complex and hard multi-objective problem into an uni-objective one and, therefore, is widely used. As a disadvantage, the one most often pointed out, is the fact that for

mixed problems, like minimum and maximum problems only one type should be assumed; however, and since the present optimisation problem relies only on the objective function minimisation, this was the method adopted to perform the multi-criteria optimisation.

In what concerns the Hierarchical Optimisation Method (Osyczka, 1981) the decision variable is ordered in terms of its importance and, using this order, each objective function is minimised separately adding in each step a constraint to prevent the change in value of the previously minimised criterion. The complexity associated to this method can lead to a reduced use of it.

Regarding the Trade-off Method, also known as ϵ -Constraint Method, it is more linked to the economical field and consists in keeping only one of the objectives and restricting the remaining to the selected one as in the form of constraints. Details about the method can be found in Eskelinen and Miettinen (2012).

6.3. PROPOSED METHODOLOGY

6.3.1. Overview

To establish this methodology, it was necessary to combine two important techniques: the geostatistical simulation of the geotechnical variable of interest, conditionally to the available preliminary information, and an optimisation algorithm known as simulated annealing (SA).

In detail, the methodology starts by performing a geostatistical simulation from preliminary data (input), for which a wide range of geotechnical information can be obtained such as the Fracturing Degree (F), Weathering Degree (W), Rock Quality Designation (RQD) or the empirical rock mass classification systems, like Rock Mass Rating (RMR) (Bieniawski, 1989), Q-system (Barton, 1974) or Geological Strength Index (GSI) (Hoek et al., 2002). Geostatistical simulation is performed in a conditional way, thus guaranteeing that the preliminary data are reproduced in information and position, following the methodology presented by Pinheiro et al. (2016) (see Figure 6.3:Stage 1). As a result, a total of L realisations of the geotechnical information, which enable the determination of the objective function to use in the optimisation process, are obtained. In this case, two metrics will be used as objective functions: the local variance of the simulated values and the width of their 95% probability interval. Then, the SA algorithm is used to minimise the objective function values, which translate the uncertainty of the geotechnical model. In this sense, the output of the previously

explained process corresponds to the optimal length and position of additional boreholes to execute during the second phase of the prospection plans (see Figure 6.3:Stage 2).

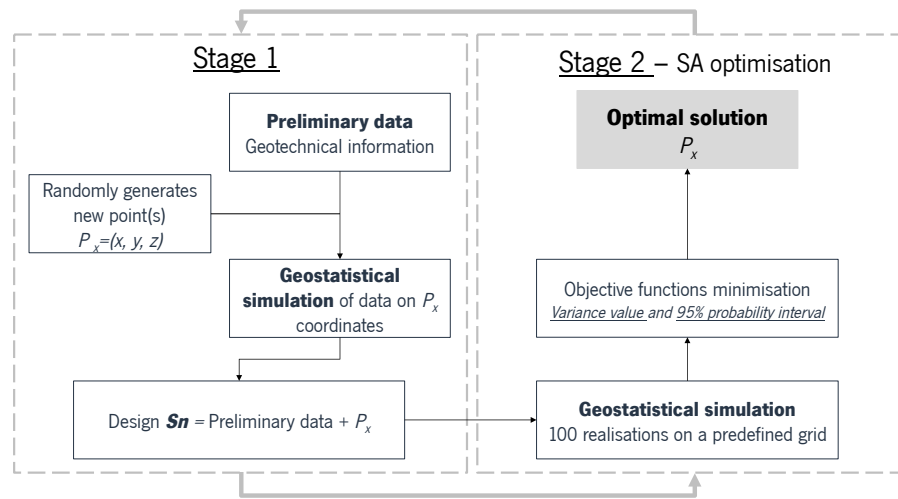


Figure 6.3 Diagram of the proposed methodology combining geostatistical simulation and simulated annealing.

6.3.2. Optimisation process

As already mentioned, the methodology requires geotechnical information to use as preliminary data. In general, this information is obtained from a limited number of boreholes already executed in the field, the location and depth of which are usually based on the *know-how* of the professionals and geological aspects. This geotechnical information should allow the calculation and fitting of a variogram of the measured regionalised variable (Chilès and Delfiner, 2012), which is an imperative tool to use in geostatistical simulation. The type of information used as initial data in this optimisation process can be diverse and directly derived from the borehole data.

In what concerns the optimisation process, it is initialised by randomly generating n points in a confined space used to represent the additional boreholes to execute. Then, conditional simulation is performed and a total of L realisations of the chosen variable at the n target points are obtained. At this stage, the turning bands algorithm is used to construct the realisations and the residual kriging approach is used to condition the realisations to the preliminary data values, *i.e.* to force the realisations to honour the information and position of the preliminary data (Emery and Lantuéjoul, 2006). The average of the L realisations is computed in order to represent each target point with a single value. After this process, each additional point has some simulated information comparable to the preliminary data and the remaining optimisation process can then be executed. The purpose of

this first stage is to investigate the contribution of each point in reducing the uncertainty of the geotechnical model by using the simulated values.

In the second stage, the n simulated values assigned to the additional boreholes to execute are joined to the preliminary data, resulting in a new sampling design (S_n). Consequently, using the points of the new design, once again a conditional simulation is executed, this time using as target points a predefined grid ($2D$ or $3D$) covering the region of interest. As an outcome of this simulation process, the objective function to be minimised, in order to find out the optimal solution for the presented problem, is obtained (see Figure 6.4). All the geostatistical routines and the SA algorithm have been programmed in a Matlab environment (Emery and Lantuéjoul, 2006; Lin and Fei, 2012; Yang, 2010), using a desktop computer with an Intel ® Core ™ i7-3610QM CPU @2.30GHz processor.

Considering that each borehole lists information at several depth ranges, *i.e.* each borehole contains the geotechnical information at several points, the optimisation is also made by considering alignments of points in the generation of the sampling design. Therefore, in this work, the developed methodology is also tested by assuming each borehole as a vertical alignment of points originated from an isolated point in the surface (header). This action allows the integration of all vertical information that can emerge from a borehole and best represents reality.

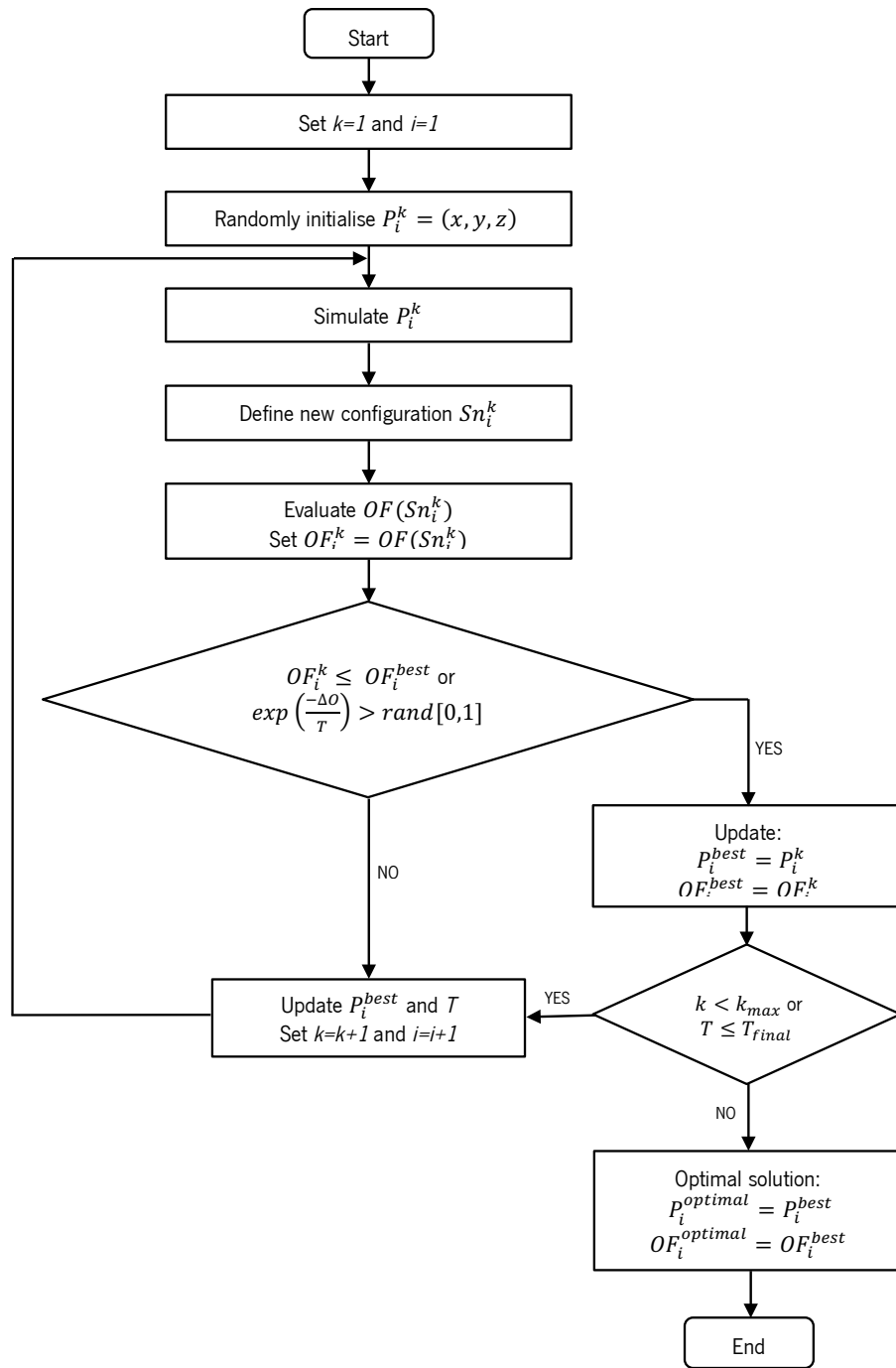


Figure 6.4 Simulated annealing workflow adapted to the proposed methodology.

6.3.3. Uni-objective approach

The bound constrained problem to be addressed here has the following form:

$$\begin{array}{ll} \min & OF(Sn) \\ \text{subject to} & lb \leq Sn \leq ub \end{array} \quad (6.6)$$

where $OF(Sn)$ represents the value of the objective function for design Sn . The bound constraints on the values of the variables Sn are related to the rock block dimensions, where the lower (lb) and upper (ub) bounds give the minimum and maximum values of the (x, y, z) coordinates, respectively.

Thereafter, two distinct objective functions are considered in the optimisation process: the variance of simulated values and the width of the 95% probability intervals of the simulated values. These two metrics are considered in order to provide more options to the decision-maker in choosing the metric that, from his/her point of view, is able to better represent the uncertainty of the geotechnical model. To simplify the notation, the argument of each objective function will be shown in the following sections.

The first objective function to be presented is the average width of the probability intervals obtained using the L realisations results, calculated as follows:

$$\overline{PI}_{95\%} = \text{mean}(PI_{95\%}) \quad (6.7)$$

where $PI_{95\%}$ is obtained for each target location from the set of L simulated values at this location, by calculating the difference between the percentiles of a lower limit and an upper limit given by $(1 - p)/2$ and $(1 + p)/2$, respectively, for a probability p equal to 0.95. After that, the average of the $\overline{PI}_{95\%}$ widths over all the target points is computed and used as an objective function. The $\overline{PI}_{95\%}$ final value should be as low as possible to decrease the uncertainty associated with the geotechnical property at unsampled locations.

The second objective function is based on the variance of the simulated values for each created design (Sn), calculated after the execution of two main steps: first, the variance of the L simulated values is calculated at each location; second, the average value of these variances over all the target locations is computed. Again, the lower this average variance, the lower the uncertainty at unsampled locations.

Table 6.1 shows details of the first step, where N represents the total number of grid nodes that compose each Sn design, L is the total number of geostatistical realisations, and $X_{p,L}$ represents the simulated value for point p at realisation L .

Table 6.1 First step to calculate the variance value.

Sn points	Realisation number			Local variance
	1	2	L	
1	$X_{1,1}$	$X_{1,2}$	$X_{1,L}$	$Var_1 = \frac{1}{L} \sum_{l=1}^L (X_{1,l} - \bar{X}_1)^2$
2	$X_{2,1}$	$X_{2,2}$	$X_{2,L}$	$Var_2 = \frac{1}{L} \sum_{l=1}^L (X_{2,l} - \bar{X}_2)^2$
N	$X_{N,1}$	$X_{N,2}$	$X_{N,L}$	$Var_N = \frac{1}{L} \sum_{l=1}^L (X_{N,l} - \bar{X}_N)^2$

Thus, the objective function based on the variance is given by:

$$Var = \frac{1}{N} \sum_{p=1}^N Var_p \quad (6.8)$$

where Var denotes the variance average between all the points and Var_p represented the variance value between all the L realisation for point p .

6.3.4. Multi-criteria approach

Taking into account the difficulty and plurality of possibilities associated with *in situ* decisions in geotechnics, a multi-criteria approach is also considered in this work. The goal is to facilitate future decisions and to analyse the effectiveness of the objectives (used metrics) when applying SA to boreholes optimisation.

Therefore, the two previously presented objective functions (criteria) are combined into a single function, since they can be complementary. Using this multi-criteria approach, the decision-maker will be able to use both objectives (metrics) and to identify one criterion that is dominated by the other by analysing the trade-off existing between them.

Hence, the multi-criteria problem to be optimised exhibits the following form:

$$\begin{array}{ll} \min & (OF_1(Sn), OF_2(Sn)) \\ \text{subject to} & lb \leq Sn \leq ub \end{array} \quad (6.9)$$

where,

$$OF_1 = \frac{\overline{PI}_{95\%}}{\overline{PI}_{95\%,initial}} \quad (6.10)$$

and

$$OF_2 = \frac{Var}{Var_{initial}} \quad (6.11)$$

Note that OF_1 and OF_2 are the objective functions normalised using the values of the objective functions of the preliminary data (initial data used for conditioning the geostatistical simulation). As mentioned before, the argument Sn is removed in order to simplify the notation. $\overline{PI}_{95\%}$ represents the average width of the 95% probability intervals and Var is the average variance obtained after the geostatistical simulation of the tested design (Sn). $\overline{PI}_{95\%,initial}$ and $Var_{initial}$ denote the average width of the 95% probability intervals and the average variance, respectively, obtained after geostatistical simulation on the predefined grid conditioned only to the preliminary data.

The Weight Sum Method, already presented in previous sections was the selected method to standardise the set of objectives into a single objective function by multiplying each one by a weight. This method is chosen among all the mentioned due to its simple use, as well as the assurance in finding the ideal and admissible space of solutions (Marler and Arora, 2009). The weight definition can be made using different approaches. However, in this work the weights were attributed manually since the weight and the solution vectors do not show a linear correspondence. The manual weight definition intends to simulate the decision-maker perspective, regarding the objective function. Hence, the problem to be minimised is converted into a uni-objective problem by aggregating the two objective functions into a single one, given by:

$$\begin{array}{ll} \min & (W_1 \cdot OF_1(Sn) + W_2 \cdot OF_2(Sn)) \\ \text{subject to} & W_1 + W_2 = 1 \\ & lb \leq Sn \leq ub \end{array} \quad (6.12)$$

where W_1 is the weight for objective function OF_1 and W_2 is the weight for objective function OF_2 .

6.4. THEORETICAL APPLICATION - CHILEAN DEPOSIT

6.4.1. Data presentation

The proposed methodology was applied using information of mechanical boreholes from an epithermal gold deposit located in the Andes mountain range, region of Atacama, northern Chile. Details regarding the information were already given in Chapter 3. For this application, the information gathered from the boreholes was the empirical Rock Mass Rating (RMR) system.

Considering the high quantity of the available data it was necessary to restrict the information to a confined area (block). Thus, using this information, two optimisation scenarios were built, one using isolated points and the second one using vertical alignments of points that better represent the reality of boreholes.

In what concerns the scenario composed by isolated points, 22 points were randomly selected from the available boreholes to represent the preliminary information to use in the geostatistical simulation and in the optimisation procedure. The number of 22 points was chosen having in mind a reasonable number of points to start for the proposed methodology validation. Regarding the scenario with vertical alignments of points, a total of 6 boreholes composed of 12 points each were chosen as preliminary information. Thus, each alignment to be found in the optimisation should be composed by 12 points evenly spaced along the depth. Each one of the 22 points (scenario 1) and 6 boreholes (scenario 2) contains information about the RMR within a rock block with the dimensions of 120 m × 440 m × 220 m along the east (X), north (Y) and vertical (Z) directions.

In this regard, to a better understating and perception of the two scenarios, maps with the 22 points and 6 chosen boreholes used as preliminary data are represented in Figure 6.5 and Figure 6.6, respectively.

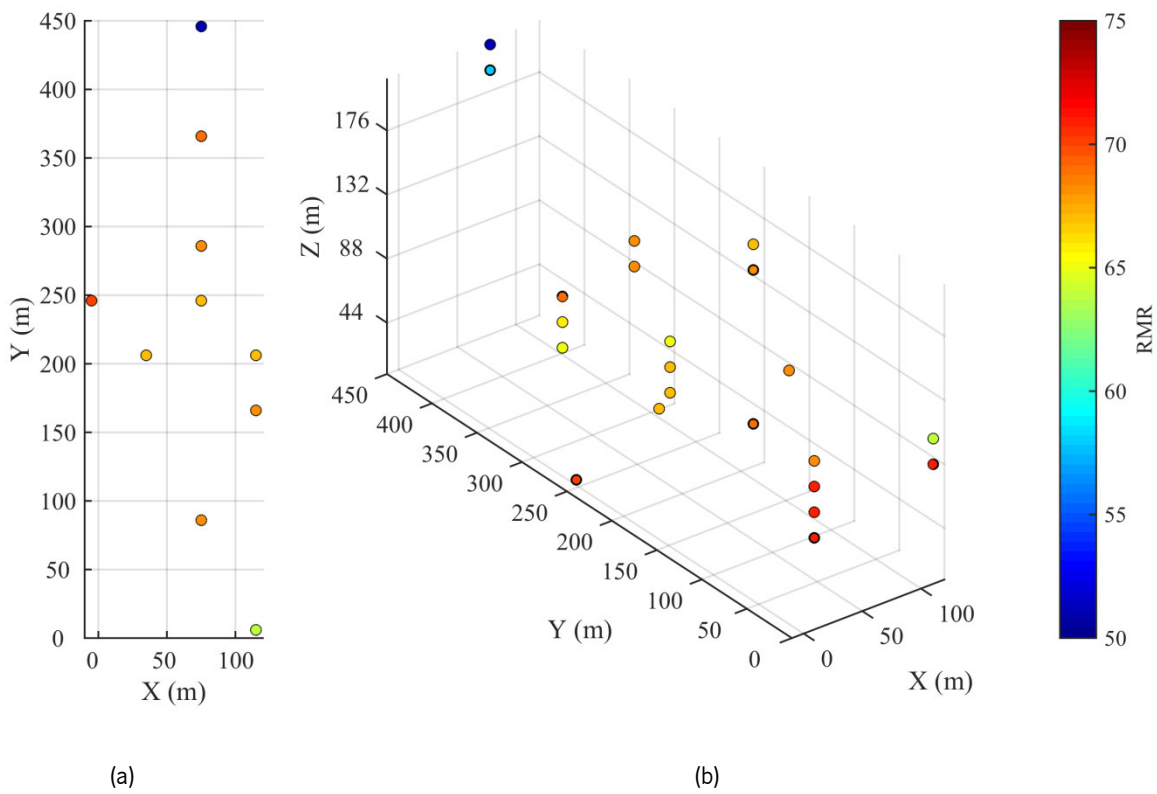


Figure 6.5 Mapping of the initial boreholes data used in the isolated points scenario in: a) XY plane; and b) XYZ perspective (X , Y and Z in meters).

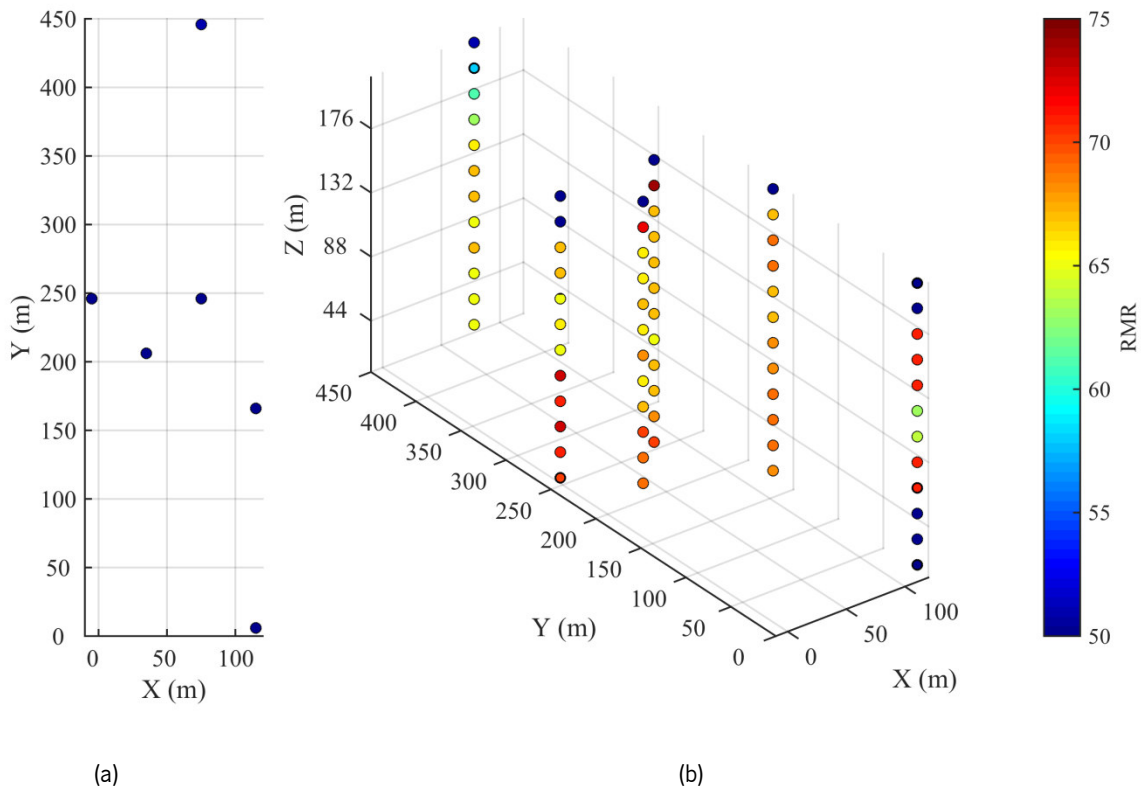


Figure 6.6 Mapping of the initial boreholes data used in the vertical alignments scenario in: a) XY plane; and b) XYZ perspective (X , Y and Z in meters).

6.4.2. Geostatistical simulation

Regarding the geostatistical simulation conditioned to the preliminary data, some steps proposed by Pinheiro et al. (2016a, 2016b) needed to be implemented:

- First, the RMR preliminary data were analysed through the calculation of basic statistics (see Table 6.2). According to these statistics, the rock mass in consideration shows a good quality, because the RMR value ranges from 51 to 71, with a mean of 67 that can classify the rock mass as almost homogeneous;
- The preliminary data corresponding to the 22 points or to the 6 alignments were then transformed into data with a standard Gaussian distribution with a zero mean and a unit variance. Such a transformation was necessary for subsequent geostatistical modelling and simulation (Chilès and Delfiner, 2012);
- The experimental variograms of the Gaussian data were computed, using either the isolated points or the point alignments. These variograms were computed until a maximum distance of 100 m, beyond which the data values are found to exhibit a low correlation;
- The experimental variograms were fitted using isotropic spherical functions, as presented in Equations (6.13) and (6.14), where the distance written between brackets represents the correlation range and the number before the spherical structure denotes the adopted value for the sill (Chilès and Delfiner, 2012):

Scenario with isolated points (see Figure 6.7a):

$$\gamma = 0.495 \text{ Spherical } (40 \text{ m}) \quad (6.13)$$

Scenario with point alignments (see Figure 6.7b):

$$\gamma = 0.80 \text{ Spherical } (70 \text{ m}) \quad (6.14)$$

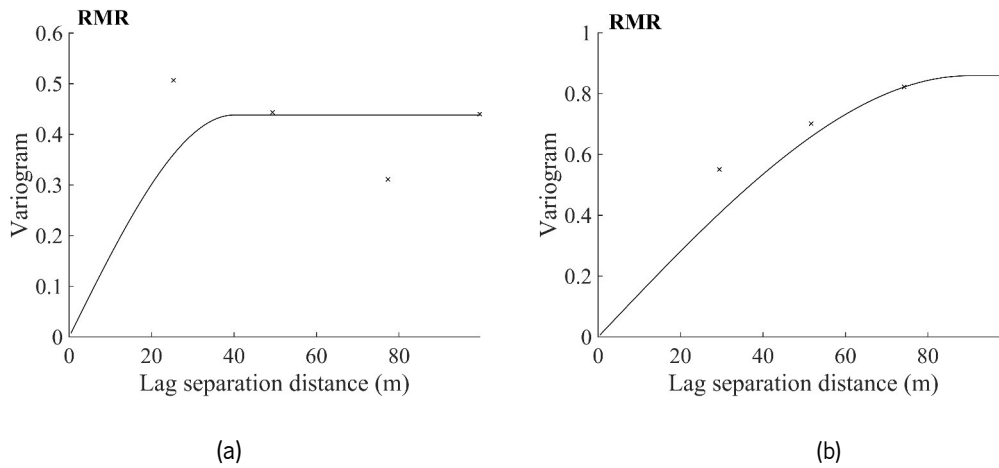


Figure 6.7 Experimental (crosses) and theoretical (solid line) variograms along the horizontal plane using: a) 22 isolated points; and b) 6 point alignments.

- A two-dimensional regular grid with a mesh of 5 m × 5 m and a total of 60 nodes along the east direction and 120 nodes along the north direction was defined to conditionally simulate the RMR and to calculate the objective functions;
- A Gaussian random field was simulated at the target grid nodes using the turning bands method (Emery and Lantuéjoul, 2006). The number of turning lines used to generate the random field was 1500, while the number of realisations was set to $L = 100$, so that the post-processing outputs (variances and probability intervals) could be calculated with a reasonable approximation (this number of realisations is commonly used in geostatistical applications as suggested by Chilès and Delfiner (2012)). Residual kriging was then used to condition the simulation to the preliminary data, with a unique neighbourhood implementation.
- The simulated Gaussian values for each realisation were finally back-transformed into their original scale (RMR).

Table 6.2 Basic statistics on RMR preliminary data and their Gaussian transforms.

	22 isolated points		6 point alignments	
	RMR original scale	Gaussian values	RMR original scale	Gaussian values
Number of points	22	22	62	62
Mean	66.77	0.00	67.29	0.00
Variance	19.99	0.94	12.27	0.98
Minimum	51.00	-2.00	51.00	-2.41
Maximum	71.00	2.00	74.00	2.41

To understand the differences between the preliminary data and the geotechnical improvements when new boreholes are added, the simulation conditioned only to the preliminary data was also performed on the predefined grid. As a result, the values obtained for the variance and width of 95% probability interval objective functions were $Var_{initial} = 9.00$ and $\overline{PI}_{95\%,initial} = 11.28$ when using the 22 isolated points and $Var_{initial} = 9.39$ and $\overline{PI}_{95\%,initial} = 11.36$ in the case of using the 6 point alignments. As an example, Figure 6.8 shows the first realisation of the RMR and the variance of 100 realisations, obtained conditioned to the preliminary data.

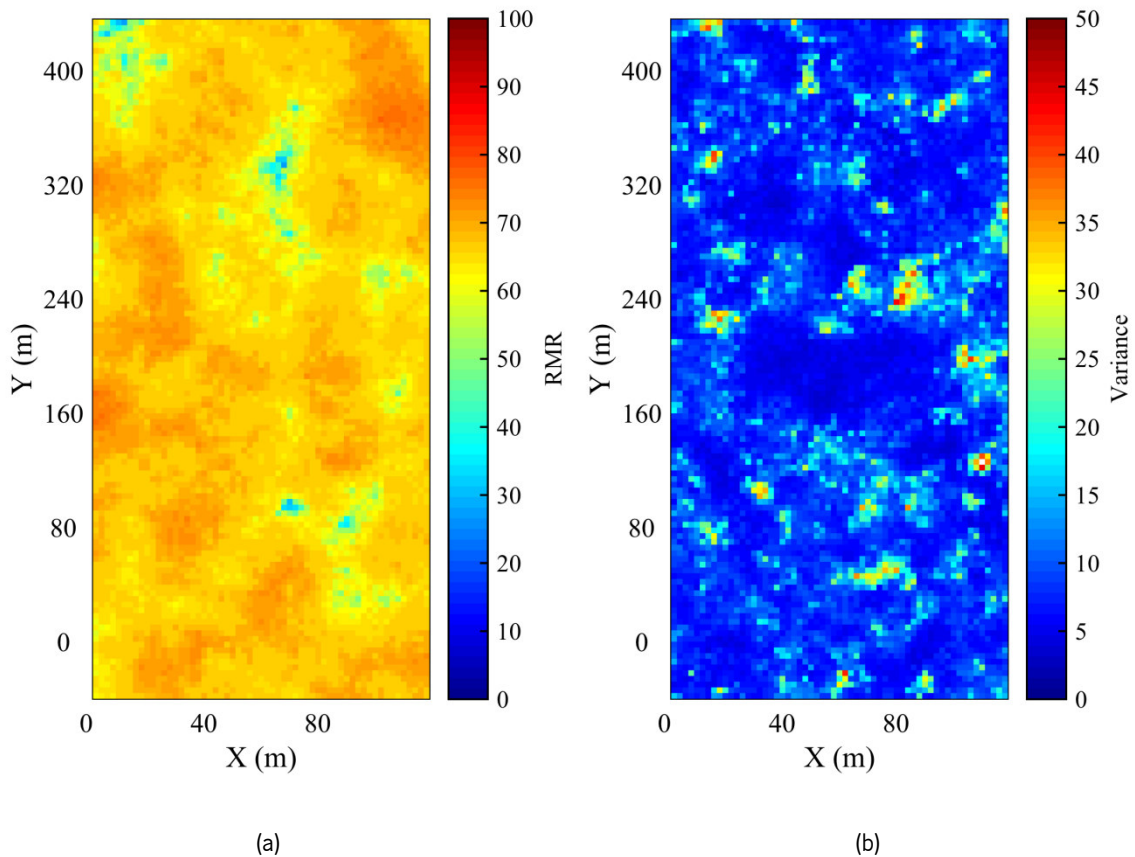


Figure 6.8 RMR preliminary data geostatistical simulation results: a) first realisation; and b) variance of 100 realisations.

6.4.3. Simulated Annealing internal parameters study

In order to execute the Simulated Annealing (SA) and to start the optimisation process, some internal parameters needed to be defined. Therefore, a sensitivity study was carried out, which enables to understand the influence of each parameter and find the best values for the algorithm. Several attempts were made, in this case for the optimisation of one additional point, firstly to comprehend the influence of the initial temperature value and then for the maximum number of rejections and runs within each temperature value. Table 6.3 displays the attempts made in this

sensitivity test as well as the SA internal parameters values. It is worth mentioning that in all the tests, a cooling factor equal to 0.80 was adopted in order to provide a slow cooling and allow the algorithm to better explore the searching space. This value represents the lower bound used for SA and commonly used in similar optimisations problems (Van Groenigen et al., 1998; Brus and Heuvelink, 2007).

Table 6.3 SA internal parameters values assumed for the sensitivity test performed.

Attempt n.	Cooling factor (α)	Initial temperature (°)	Final temperature (°)	Maximum number of rejections	Maximum number of moves	Maximum number of acceptances
1	0.80	1.00	0.10	20.00	20.00	5.00
2	0.80	2.00	0.10	30.00	20.00	5.00
3	0.80	2.00	0.01	40.00	30.00	10.00

More details regarding the optimisation results of all the attempts can be found in Appendix 2. Since the cooling factor remained the same in all the attempts ($\alpha = 0.80$), a value of 1.00° for the initial temperature was found to be too low, since the algorithm took too little time to converge to a solution and the searching space was too limited. Therefore, a value of 2.00° for the initial temperature was adopted. In what concerns the maximum number of rejections and moves, the tested values have proven insufficient and, therefore, new tests with higher values ($Max_{rejections} = 100.00$ and $Max_{moves} = 50.00$) were performed. From these results, it was possible to define the SA internal values to adopt in this optimisation case study as being the ones shown in Table 6.4.

In what concerns the computational time required by the geostatistical simulation, it could be divided in two. First, the necessary time to simulate L times the variable into the randomly generated coordinates, which in this case was approximately 30 minutes. Then, in a second stage, 100 realisations were simulated on a target grid to later compute the objective function value.

Table 6.4 SA internal parameters.

Cooling factor (α)	Initial temperature (°)	Final temperature (°)	Maximum number of rejections	Maximum number of moves	Maximum number of acceptances
0.80	2.00	0.01	100	50	10

6.5. OPTIMISATION RESULTS

6.5.1. Uni-objective results

Regarding the validation of the proposed optimisation methodology, a wide range of additional points and alignments of points was considered.

Concerning the isolated points scenario, the optimisation process starts by adding one point to the preliminary 22 points, so $n = 1$, and then by adding points consecutively one by one up to a total of 13 additional points. A number of points of 16, 20, 24 and 30 were also tested in order to accelerate the optimisation process and to understand the advantage of using a large number of additional points, *i.e.* the gain of using a large number of additional points to characterise the rock mass.

In the alignments scenario, the same logic as for the isolated point optimisation was followed. In this case, the optimisation also started by adding one alignment of points composed by a total of 12 points ($m = 12$), then 5 alignments were added one by one. In what concern the isolated points optimisation results, Table 6.5 and Table 6.6 show the number of evaluations performed by the algorithm, the final temperature and objective function values in the case of $\overline{PI}_{95\%}$ and Var functions, respectively.

Table 6.7 reports a summary with the optimal values of each objective function for every number of tested points. Moreover, to compare the geotechnical gain obtained by adding the points, a column represented by the symbol Δ shows the difference, in percentage, between the objective function original value for each additional point and the reference value obtained after the simulation using only the preliminary data (see section 6.4.2). In some cases, the difference in the objective function value was up to 80% (for the variance average value when 30 points are added). A graphical representation presented in Figure 6.9 gives a more practical view of the differences in the objective function value when a small or a high number of additional points is considered.

Table 6.5 Isolated points optimisation results when the 95% probability interval ($\overline{PI}_{95\%}$) was the objective function.

Number of additional points	Number of evaluations	T_{final}	Objective function value
1	143	0.088	10.460
2	149	0.090	10.330
3	147	0.009	10.400
4	152	0.087	10.220
5	155	0.087	9.730
6	155	0.087	8.880
7	155	0.009	8.360
8	155	0.087	7.740
9	145	0.087	7.400
10	153	0.087	7.030
11	153	0.087	6.750
12	148	0.087	6.450
13	151	0.087	6.350
16	160	0.087	5.960
20	156	0.087	5.560
24	152	0.009	5.160
30	152	0.009	4.610

Table 6.6 Isolated points optimisation results when the variance (Var) average value was the objective function.

Number of additional points	Number of evaluations	T_{final}	Objective function value
1	235	0.015	8.12
2	284	0.080	7.37
3	241	0.050	6.94
4	231	0.080	6.81
5	228	0.080	6.55
6	233	0.080	6.09
7	212	0.090	5.91
8	231	0.080	5.33
9	230	0.080	4.99
10	230	0.080	4.59
11	208	0.110	4.36
12	253	0.080	3.80
13	147	0.420	3.67
16	147	0.420	3.54
20	236	0.420	2.25
24	230	0.420	2.01
30	230	0.420	1.60

Table 6.7 Summary of the SA results for each additional point using two different objective functions ($\overline{PI}_{95\%}$ and Var).

Number of additional points	$\overline{PI}_{95\%}$			Var		
	Objective function values	Initial value (reference)	Δ	Objective function values	Initial value (reference)	Δ
1	10.46		-7%	8.12		-10%
2	10.33		-8%	7.37		-18%
3	10.40		-8%	6.94		-23%
4	10.22		-9%	6.81		-24%
5	9.73		-14%	6.55		-27%
6	8.88		-21%	6.09		-32%
7	8.36		-26%	5.91		-34%
8	7.74		-31%	5.33		-41%
9	7.40	11.28	-34%	4.99	9.00	-45%
10	7.03		-38%	4.59		-49%
11	6.75		-40%	4.36		-52%
12	6.45		-41%	3.80		-58%
13	6.35		-44%	3.67		-59%
16	5.96		-47%	3.54		-61%
20	5.56		-51%	2.25		-75%
24	5.16		-54%	2.01		-78%
30	4.61		-59%	1.60		-82%

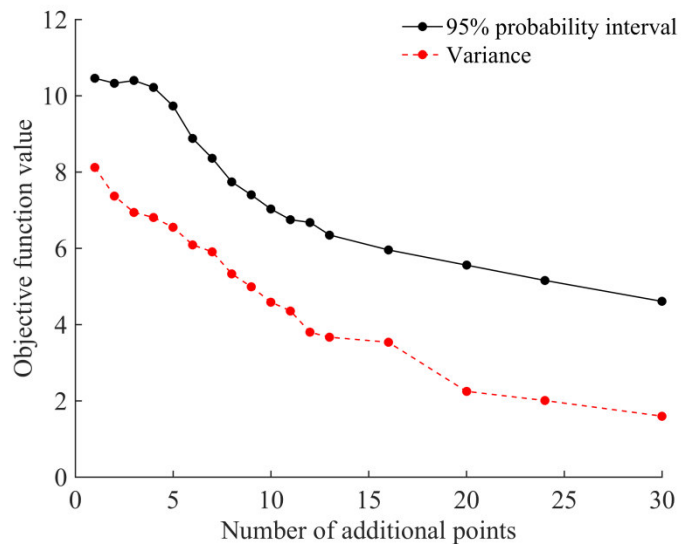


Figure 6.9 Convergence study for the number of additional points considering the average width of 95% probability intervals (black line) and the average variance (red line).

It was possible to observe, as expected, that both objective functions values decrease as the number of additional points increases. However, this decrease was not constant. In what concerns the

variance objective function, the decrease was significant until 12 additional points and also when the number of points increases from 16 to 20. After this value, the decrease in the objective function was modest. Analysing the objective function magnitudes with 20 additional points the variance value decreased to 25% of the initial value, while for the probability interval this reduction was more modest yet significant. In relation to the probability interval objective function, a significant reduction in the value was only observed starting from 5 additional points and up to 13. After this value, the decrease rate was smaller.

Since the number of additional points was considerably high, Figure 6.10 shows, as an example, a 3D representation with the optimised positions for 12 additional points at each objective function, along with the representation of the preliminary data (black points). A certain proximity between the points optimised with each objective function could be observed. However, their spatial locations were different, meaning that both objective functions behave differently in space. Also, it was interesting to notice that some of the points were aligned, which allow their simultaneous characterisation by performing non-vertical boreholes that reduce the execution costs. For this specific case, using the variance objective function results, the number of additional boreholes to perform is not 12, as the number of suggested points, but can be reduced to, approximately, 7 or 8 boreholes (see Figure 6.10a). Even so, in Figure 6.10b the number of boreholes to execute is lower (5 to 6 boreholes).

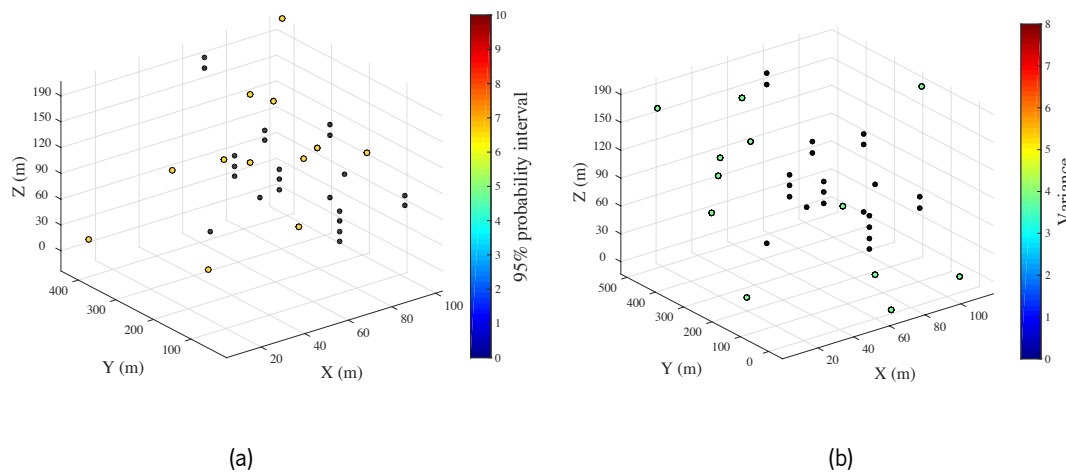


Figure 6.10 3D location of 12 additional points (colour points) with preliminary data represented by black points for: a) average width of 95% probability intervals; and b) average variance.

Concerning the SA output parameters, for the specific case of 12 points, the final temperature was 0.08 for both objective functions and these functions were evaluated a total of 400 times. Also, the computational time spent in all the process was of 9 hours approximately.

In addition, Figure 6.11 shows a 3D graph that intends to represent the search block used in the case of the 95% probability interval and variance functions. It was possible to observe that the algorithm was able to explore quite well within the block in order to find the optimal combination of 12 points, behaving in a more regular way in the case of the $\overline{PI}_{95\%}$ and in a more clustered way for the Var function.

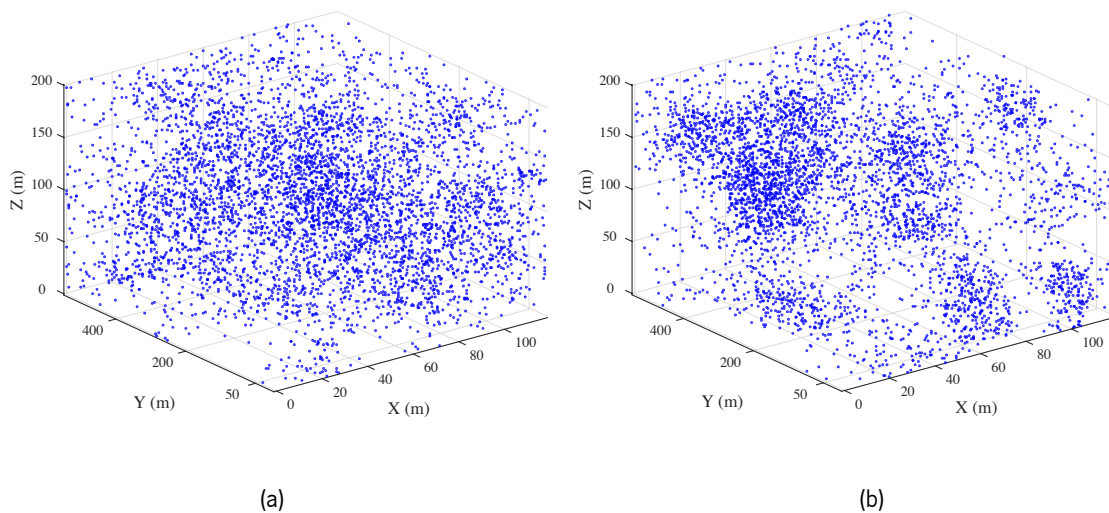


Figure 6.11 3D representation of the block used in the search when 12 points were added, for: a) average width of 95% probability intervals; and b) average variance.

Following the same logic as for the isolated point optimisation, the point alignment optimisation also started by adding one alignment of points composed by a total of 12 points ($n = 12$), then 5 alignments were added one by one. As a result, Table 6.8 and Table 6.9 displays the number of evaluations, final temperature and objective function values for each additional alignment of points when a $\overline{PI}_{95\%}$ and a Var functions was used, respectively. Likewise, Table 6.10 shows the obtained values for both objective functions, as well as the difference in relation to the objective values obtained using the 6 preliminary alignments. Likewise, Figure 6.12 presents the optimal value for each one of the objective functions for every alignment of points added to the preliminary data.

Table 6.8 Point alignments optimisation results when the 95% probability interval ($\overline{PI}_{95\%}$) was the objective function.

Number of additional alignments	Number of evaluations	T_{final}	Objective function value
1	273	0.110	9.910
2	398	0.029	8.620
3	307	0.056	8.400
4	373	0.036	7.830
5	377	0.029	7.480
6	421	0.023	7.110

Table 6.9 Point alignments optimisation results when the variance (Var) average value was the objective function.

Number of additional alignments	Number of evaluations	T_{final}	Objective function value
1	415	0.056	7.330
2	417	0.056	6.091
3	203	0.215	5.310
4	231	0.080	4.956
5	417	0.056	4.712
6	233	0.080	3.796

Table 6.10 Summary of the SA results for each additional alignment using two different objective functions ($\overline{PI}_{95\%}$ and Var).

Number of additional alignments	$\overline{PI}_{95\%}$			Var		
	Objective function values	Initial value (reference)	Δ	Objective function values	Initial value (reference)	Δ
1	9.91	11.36	-15%	7.33	9.39	-22%
2	8.62		-29%	6.09		-35%
3	8.40		-31%	5.31		-43%
4	7.83		-38%	4.96		-47%
5	7.48		-41%	4.71		-50%
6	7.11		-45%	3.80		-60%

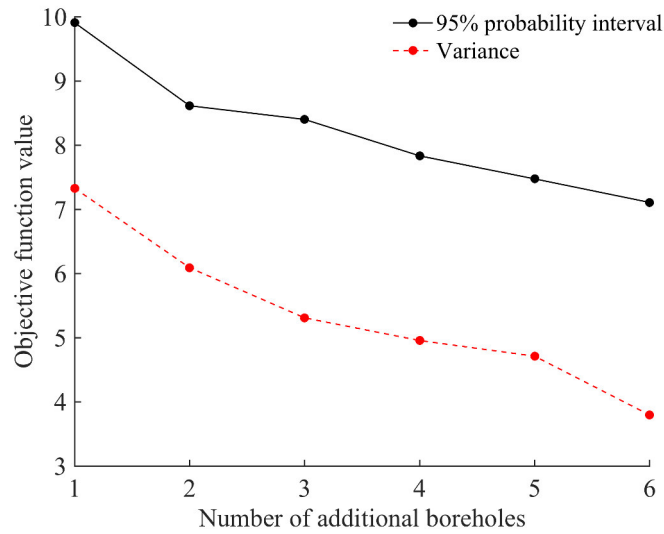


Figure 6.12 Convergence study for the number of additional alignments considering the average width of 95% probability intervals (black line) and the average variance (red line).

As observed in the isolated point optimisation, both objective functions decrease as the number of additional alignments increases. The decrease was more significant for the variance objective function when passing from 5 to 6 alignments, while the 95% probability function shows a more constant behaviour.

Figure 6.13 shows the 3D maps with the optimised positions for 3 additional alignments at each objective function along with the representation of the preliminary data (black points). In this case, and when compared to the isolated point optimisation, one notices that the obtained alignments were located closer to the preliminary alignments; however, the spatial locations of the alignments for both objective functions are considerably different, showing, once again, that both functions behave distinctly.

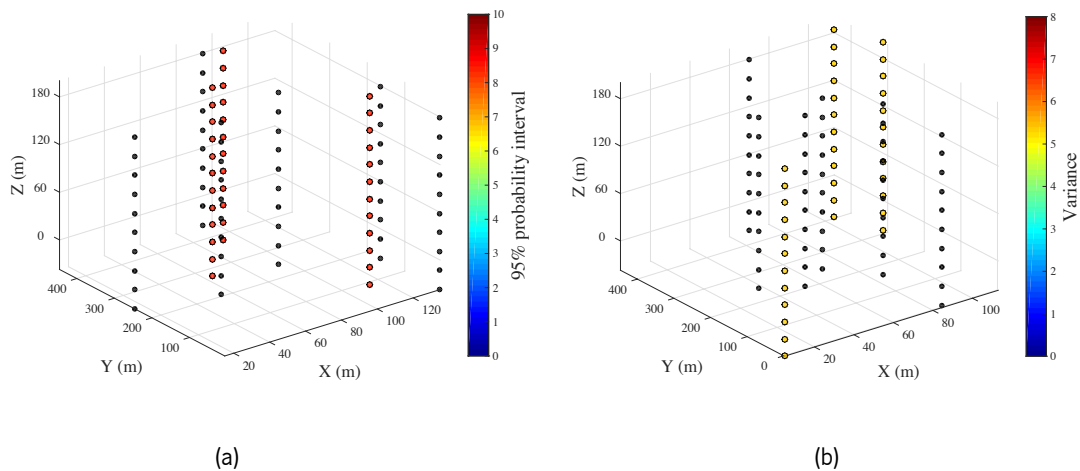


Figure 6.13 3D location of 3 additional point alignments (colour points) with preliminary data represent by black points for: a) average width of 95% probability intervals; and b) average variance.

In what concerns the algorithm convergence, a study was performed by running the algorithm twice for each one of the objective functions. In the case of the isolated points and point alignments scenarios, it was possible to state, by observing Figure 6.14, that the assumed values for SA internal parameters, namely the maximum number of iterations, were adjusted for the problem. This was confirmed by the existing sill between iterations 250 and 400 for both objective functions in the isolated points scenario (see Figure 6.14a and Figure 6.14c). Concerning the point alignments convergence results, the same trend was observed. Indeed, the minimum value for the objective function happens between iteration 300 and 400 (see Figure 6.14b and Figure 6.14d); however, in the case of the variance function second run, though the algorithm reached the 400 iterations the minimum value was registered much earlier (approximately at iteration number 110). Analysing Figure 6.14 closely, it was possible to observe that, for the variance of the points alignments and for the 95% probability interval of the isolated points, the algorithm stopped before 400 iterations, which allows to conclude that 400 was a satisfactory maximum to assume for the algorithm iterations.

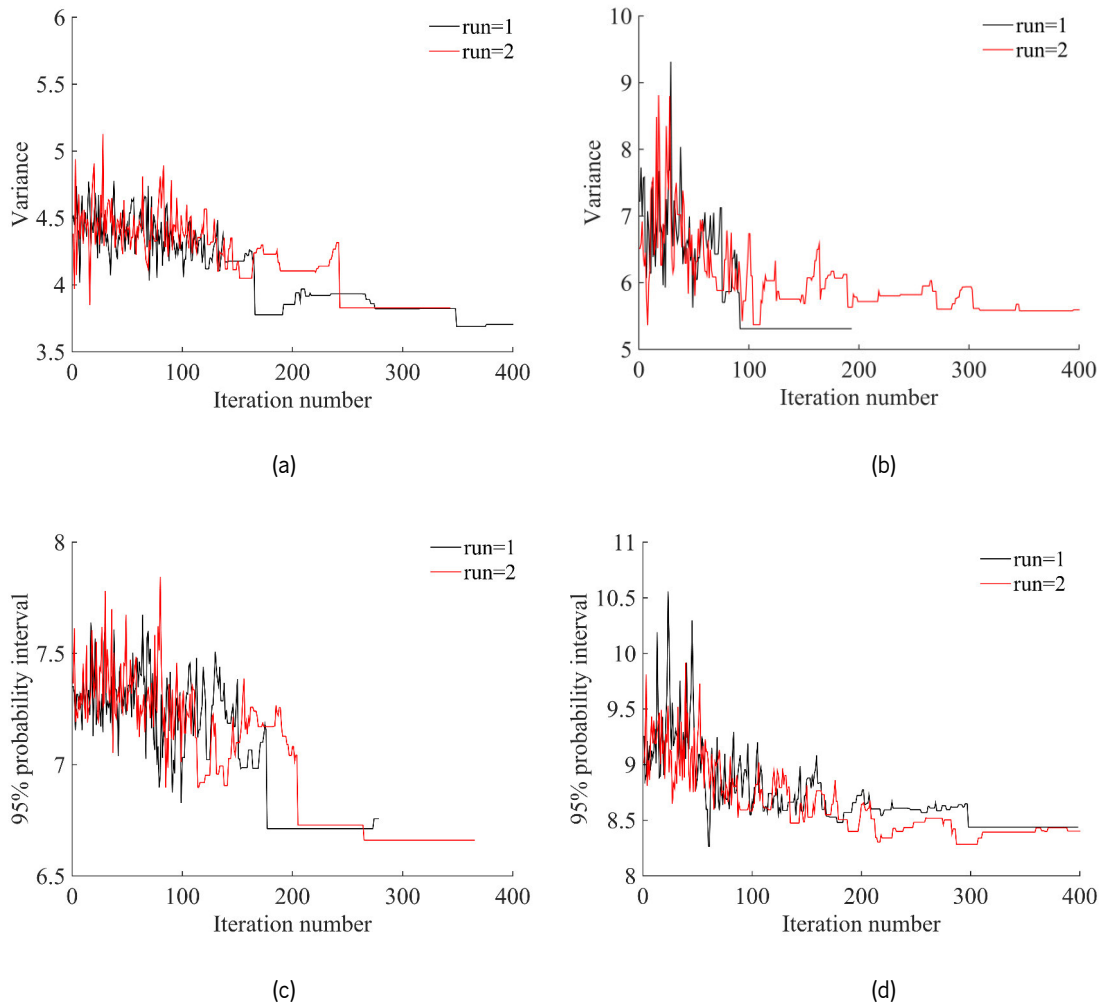


Figure 6.14 Convergence study relating the maximum number of iterations with the objective function value using: a) and b) variance as the objective function; c and d) 95% probability interval as the objective function obtained for 12 isolated points (a and c) and 3 point alignments (b and d).

6.5.2. Multi-criteria results

To provide an insight into the most influential objective function in the optimisation results for each case, different weights values were assigned. These combinations of weights, W_1 and W_2 for the $\overline{PI}_{95\%}$ and Var functions, respectively, are presented in Table 6.11 along with the optimisation results for both objective functions when 5 points ($n = 5$) are added to the 22 preliminary points. This number of additional points was chosen based on the fact that a total of 5 additional points can be sufficient to increase the quality of the geotechnical model.

Since trying a large number of weight values could lead to a computational burden, a total of 11 most common combinations of weights was considered. Both objective functions corresponding weights are represented in Table 6.11 by the test number, while the $\overline{PI}_{95\%}$ and Var values are given in their normalised scale (ratio between the obtained value and the reference value presented

previously). The representation of the range of solutions in the objective space is depicted in Figure 6.15.

Table 6.11 Isolated points optimisation results for each combination of weights for both objective functions along with the combination of the two objective functions (Both).

Test number	W_1	W_2	$\overline{PI}_{95\%}$	Var	Both
T1	0.00	1.00	0.85	0.69	0.69
T2	0.10	0.90	0.87	0.72	0.73
T3	0.20	0.80	0.86	0.72	0.75
T4	0.25	0.75	0.84	0.69	0.73
T5	0.40	0.60	0.82	0.68	0.74
T6	0.50	0.50	0.86	0.71	0.79
T7	0.60	0.40	0.89	0.73	0.83
T8	0.75	0.25	0.90	0.67	0.85
T9	0.80	0.20	0.91	0.76	0.88
T10	0.90	0.10	0.82	0.67	0.81
T11	1.00	0.00	0.84	0.70	0.84

In Figure 6.15, the X -axis represents the $\overline{PI}_{95\%}$ values obtained for each weighted test and the Y -axis, likewise, represents the Var values. As expected, one can observe that both objectives are not conflicting since the points represented in Figure 6.15 result in a linear trend. In this sense, it was possible to detect that T5 and T10 result in the non-dominated points since the trade-off between the two objectives was softer, showing that the ideal solution should not exclude the variance criterion (T10). Also, T8 that valorises the probability interval in 75% gave a worse solution (a solution far from the diagonal). This information confirms that the variance criterion, even for lower weight values results in objective function values that are lower than the $\overline{PI}_{95\%}$, meaning that cannot be excluded of the optimisation process and should be seen as the main objective function. Besides that, the differences were not as high as expected and, in this case, the adoption of a multi-criteria approach gives similar results as the uni-objective one.

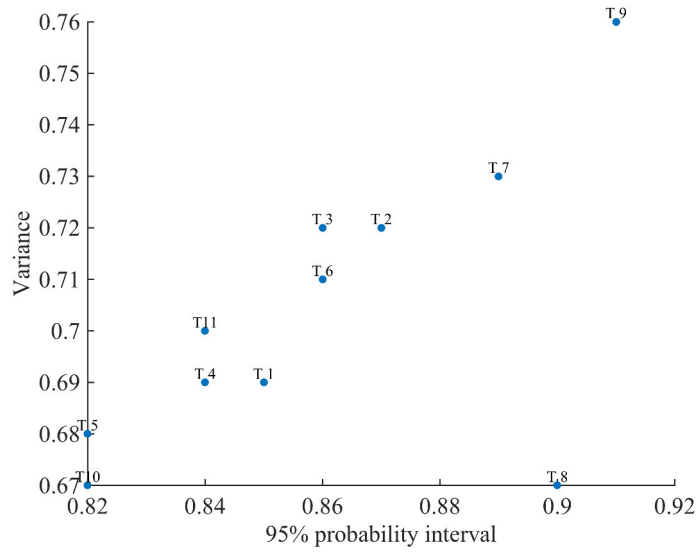


Figure 6.15 Isolated points objective function values for each weight combination test.

Furthermore, and in order to easily understand the contribution of each function in the decision of the number of boreholes to execute, Figure 6.16 presents the performance of each objective function individually and the aggregation of the two objective functions in the multi-criteria approach, for each weight combination test. From the graphical analysis, it was possible to observe that, for the first five tests, where the variance objective weights were higher (W_2 in Table 6.11), the combination of the two objective functions (blue line) showed the lowest values, which was in line with the state above about the variance as the main objective function.

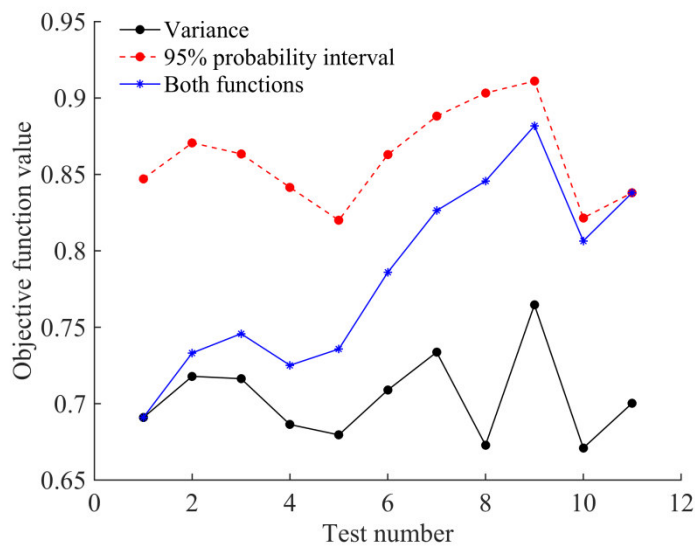


Figure 6.16 Representation of the objective functions values (95% probability interval, variance and both functions) obtained in the multi-criteria approach.

A multi-criteria approach was also performed in the scenario with point alignments and, similarly to the isolated points, this turn with the addition of 3 alignments of points. The results of both objective functions, when different combinations of weights were assumed, are displayed in Table 6.12.

Table 6.12 Point alignments optimisation results for each combination of weights for both objective functions along with the combination of the two objective functions (Both).

Test number	W_1	W_2	$\overline{PI}_{95\%}$	<i>Var</i>	Both
T1	0.00	1.00	0.72	0.53	0.53
T2	0.10	0.90	0.73	0.57	0.59
T3	0.20	0.80	0.72	0.53	0.57
T4	0.25	0.75	0.84	0.57	0.63
T5	0.40	0.60	0.74	0.58	0.64
T6	0.50	0.50	0.76	0.64	0.70
T7	0.60	0.40	0.74	0.59	0.68
T8	0.75	0.25	0.74	0.58	0.70
T9	0.80	0.20	0.74	0.59	0.71
T10	0.90	0.10	0.78	0.69	0.77
T11	1.00	0.00	0.74	0.61	0.74

In this case, the graphical analysis of both objective functions with the goal to detect conflicting solutions is presented in Figure 6.17. In the same way that the isolated points multi-criteria results, for this case, the linear trend between the solutions means that they were not in conflict. However, the worse solution was registered in test number four where the variance function was valorised in 75% (solution far from the diagonal). This result meets the conclusion stated for the isolated points approach confirming that the variance criterion should not be excluded (the worst value was obtained for the 95% probability interval criterion).

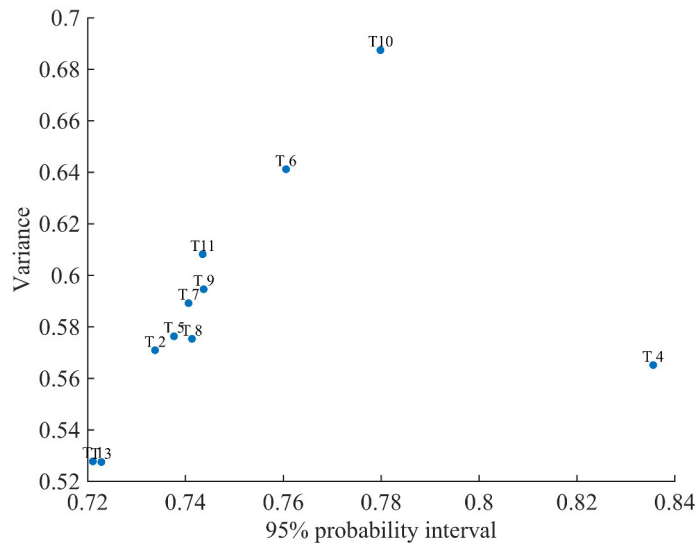


Figure 6.17 Point alignments objective function values for each weight combination test.

The relation between the weights combination tests and the objective functions normalised values is given in Figure 6.18. The aggregation of both objective functions is in the plot represented by the designation *both*. As expected, for test number four a peak value appeared in the $\overline{PI}_{95\%}$ line, highlighting once again the worst obtained solution. In terms of the function value evolution, the variation was considerably small in all performed tests. Regarding the variation criterion, was possible to perceive that, even in the tests where it receives the lowest weight values, the results are always the lowest ones, as can be observed by the both functions aggregations (blue line). These results agree with the ones obtained for the isolated points multi-criteria optimisation.

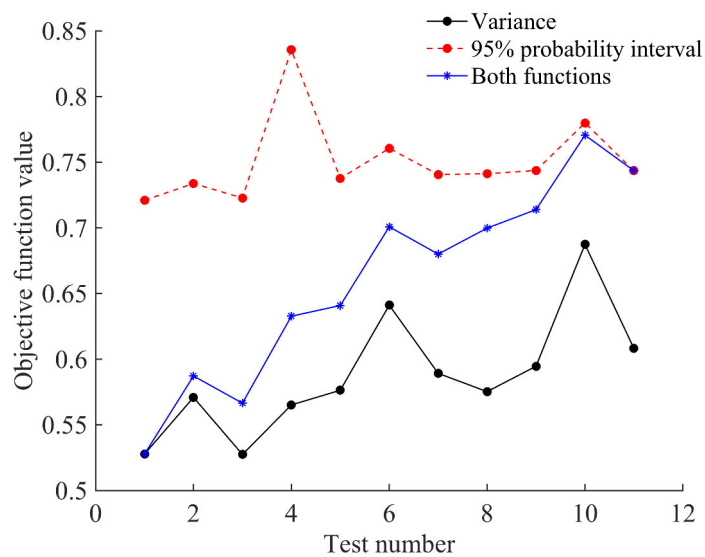


Figure 6.18 Representation of the objective functions values (95% probability interval, variance and both functions) obtained in the multi-criteria approach of the point alignments information.

6.5.3. Discussion of the results

Both approaches provide an important insight into how optimising the boreholes location increases the geotechnical detail of the rock mass and decreases the associated uncertainties. The uni-objective problem shows good results in giving the boreholes near-optimal locations according to two different objective functions. Figure 6.9 displays the resulting curve between the variance value and the number of additional points tested, and it was possible to observe a higher downhill compared with the 95% probability interval objective function. Also, it was possible to notice the differences between assuming the borehole as an isolated point or as a point alignment; however, they were not as significant as expected, once the total number of points used in the alignment optimisation was almost the triple as the one used in the point optimisation. In terms of values, the biggest difference was observed for *Var* function, where for one point *versus* one alignment the variance value was 10% lower for the last case, while the $\overline{PI}_{95\%}$ showed a reduction of only 5% compared with the results obtained from isolated point optimisation.

In this regard, Figure 6.15 shows that, when a higher weight was given to the probability interval function (T7 to T11), the $\overline{PI}_{95\%}$ value was, sometimes, higher and the variance was lower. This effect is called a trade-off (how much of a loss in one objective is one willing to sacrifice for a gain in another objective) and, in this case, the 95% probability interval was the objective function that sacrifices more, *i.e.* minimises slower when the goal is to give importance to the variance objective (see Figure 6.17). Identifying this trade-off may become difficult, once both functions are dimensionless (Marler and Arora, 2009). Nevertheless, this type of trade-off analysis can be helpful for the decision-maker in the weight values definition and in choosing the best metric to achieve the optimisation goals. In Figure 6.19 is showed a *XY* plane representation of 3 alignments optimised using the variance and 3 alignments optimised using the 95% probability interval as objective functions. For comparison, Figure 6.20 shows the same *XY* plane, this time showing the results of the multi-criteria approach when 3 points alignments are optimised. From the figures analysis, it was possible to observe that for T1 and T11 (corresponding to a unit weight for the variance and 95% probability interval function, respectively), the optimal locations are different from the ones obtained with the uni-objective approach (same colour points); however, in some of the points the locations are close of each other.

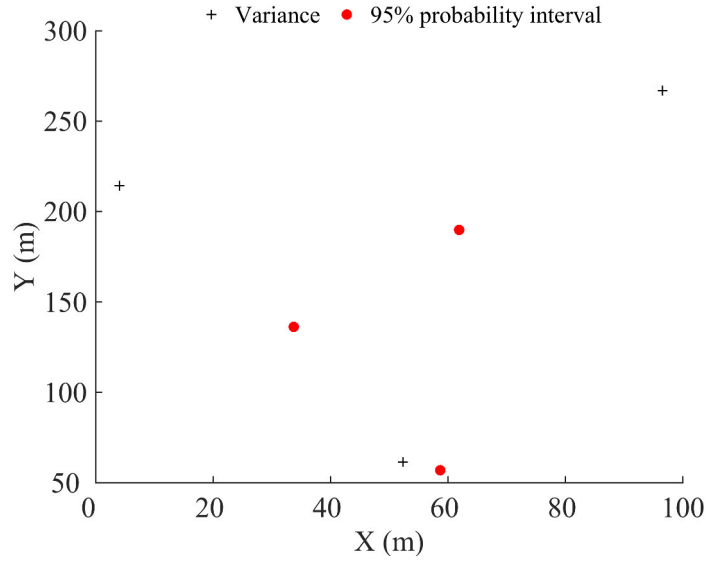


Figure 6.19 Spatial representation (XY plane) of the optimal points alignments (colour points) obtained for the variance (black plus) and 95% probability interval (red circles) as the objective functions.

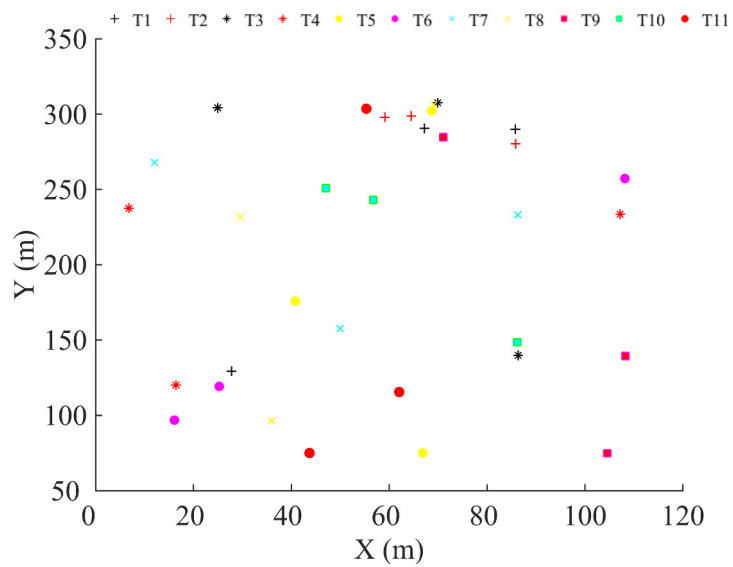


Figure 6.20 Spatial representation (XY plane) of the optimal points alignments (colour points) obtained in all the tests performed in a multi criteria optimisation.

6.6. CONCLUSIONS

The goal of this work was to define a methodology to optimise boreholes plans. The proposed methodology combines two methods: geostatistical conditional simulation, which results in the definition of the objective functions to measure the uncertainty at locations without data, and simulated annealing, used to perform the optimisation by minimising the defined objective functions. In this Chapter, boreholes information from a Chilean deposit was used, as an example used to validate and adjust the proposed methodology.

Two different objective functions were defined and tested in this work: the average variance of the simulated values and the average width of the 95% probability intervals of the simulated values over a region of interest. Moreover, not only a uni-objective problem was solved, but also a multi-criteria approach was carried out using the weight sum method. The methodology was applied using real data from the Chilean deposit and a different number of additional isolated points or point alignments (each alignment representing an additional borehole) were used in order to analyse the best spatial location for each isolated point or point alignment.

Compared with the multi-criteria approach, the uni-objective approach presents the advantage of needing a lower computational time (9 to 10 hours while the multi-criteria approach requires 18h to 22h), as the use of only one metric as an objective function simplifies the problem. On the other hand, the multi-criteria approach allows the decision-maker to give more importance to the metric (objective function) that is considered relevant, increasing the confidence in the optimisation results.

Regarding the isolated point and point alignment optimisations, the results are very different, mainly because of the total number of initial points used in both cases: in isolated point optimisation, a total of 22 points are used, while 62 points are used in point alignment optimisation. In the latter case, the objective functions are not as low as expected, compared with the isolated point optimisation. This fact comes to endorse that the XY location of the boreholes takes, in this case, an important role in geotechnical prospection, although the point alignment optimisation brings a more realistic output to represent boreholes. Also, by assuming a multi-criteria approach, throughout the weights implementation, the user of the methodology can have an additional feature that choose to apply according to the type of geotechnical detail needed.

It is important to point out that a point-by-point study allows the decision-maker to understand the geotechnical gain balanced with the economic costs. For example, will performing five boreholes instead of four bring a significant gain in the rock mass characterisation that justifies the difference in cost?

Therefore, one concludes that this methodology can contribute for a more rational approach in the formulation of prospection plans. Since the output is the measurement of the gain obtained from the addition of boreholes, including their spatial location, an indirect optimisation of costs can be expected as a consequence. Also, it is worth mentioning that the output information of this methodology should be seen as complementary information to be paired with expert knowledge from geotechnical engineers and other existing geotechnical information. One extra advantage associated with this methodology is the possibility to perform an updating in the information as the additional boreholes are executed.

Chapter 7

SALAMONDE II POWERHOUSE COMPLEX - BOREHOLE OPTIMISATION

7.1. INTRODUCTION

In the previous Chapter, a methodology to optimise boreholes plans was proposed. In essence, some geotechnical information regarding boreholes executed in a Chilean deposit was taken into consideration aiming to validate the methodology, mainly in terms of process; however, the optimal positions results could not be compared with really executed boreholes. In fact, an interesting exercise to carry out to truly validate the methodology would be the removal and then the optimisation of one or more boreholes that have already been performed. Thereby, the spatial location and geotechnical gain could be compared between the optimised positions and the one that was in fact adopted.

To perform the referred approach, a second case study was considered, this time using a complete and real area of interest. As mentioned, one and then two boreholes are removed from the data set and their location optimised and compared with the real ones. This type of analysis intends to measure the effectiveness of the proposed methodology in comparison to the traditional method, predominantly based on professional *know how*.

This Chapter starts, with a section where details are given regarding the used data set, as well as some scenarios considered for the optimisation tests. These scenarios were created by removing, firstly one borehole from the seven initial ones (scenario 1) and secondly, two boreholes (scenarios 2 and 3, where in each scenario the two removed boreholes were different).

Then, a third section presents the optimisation process, where the main differences in relation to the Chilean case study are highlighted. Subsequently, and because the variables of interest of this case study are different from the ones previously used, details concerning the geostatistical simulation are presented in a fourth section.

Subsequently, all the optimisation results are displayed in sections 5 and 6 along with the studies executed concerning the simulated annealing internal parameters values. They are extremely important to guarantee the accuracy and optimisation capability of the used methodology. Also, in this section the optimal positions obtained in all the three scenarios are compared with the real ones and the uncertainty reduction in using this methodology assessed. Finally, some discussion of the results and conclusions are drawn from the results.

7.2. DATA PRESENTATION

For the power reinforcement of Salamonde, new and larger underground structures had to be built, namely a powerhouse cavern. For this purpose, a detailed prospection plan was carried out to obtain some rock mass information regarding the area of interest. All the information regarding this underground work is presented in Chapter 5, namely the geotechnical and geometrical details required in this application.

Before the powerhouse construction, during the design phase, a total of 16 mechanical boreholes were executed from the surface, although only seven of them were located near the area of interest for this optimisation. These boreholes have a wide range of depths and orientations defined according to the *in situ* rock mass conditions.

The boreholes identification codes used in this application are: S1, S2, S5, S8, S9, S15 and S16. While some of them are located close to the powerhouse cavern, others were drilled in more distant places (see Figure 7.1). Due to the geotechnical complexity of the Salamonde II powerhouse complex, the number of boreholes executed was relatively large. For that reason and because in most of the geotechnical works the number of boreholes is not as high as in this case, some of them were not included in the initial set (boreholes S6 and S14).

Moreover, and to justify the removal of the mentioned boreholes, the selection criteria adopted were: 1) approximately the same number of boreholes in the north and south zones of the powerhouse cavern, that is why four boreholes are located in the north and northeast zones (S1, S2, S15 and S16) and three in the south zone (S5, S8 and S19); 2) because borehole S1 is very small, in the north part were selected four and three in south zone; 3) then, and because the geotechnical information regarding borehole S7 was not available for this study in due time, boreholes S14 and S6 had to be excluded from the data set selection, in order to avoid an unbalanced number of information in the north zone. This removal was made, not because the methodology is not able to consider unbalance data but because the goal was to test it with a more usual number of boreholes (reduced number); and 4) in the north zone, two sub zones were created, on the north and northeast of the powerhouse cavern, to obtain a wider range of geotechnical information (represented by blue and green colours in Figure 7.1).

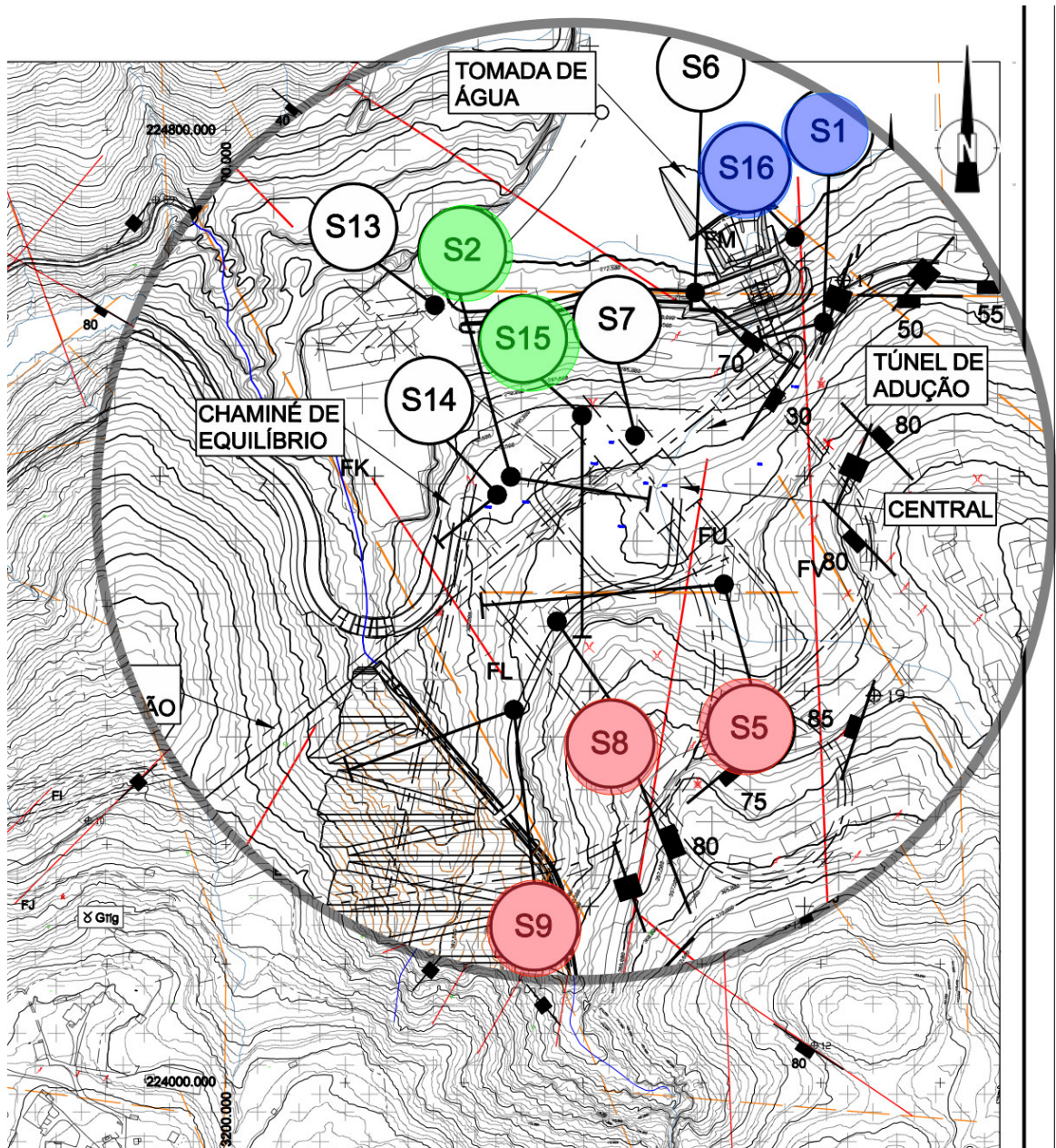


Figure 7.1 Section of the geological and prospection location plant with the identification of the selected boreholes (colour circles) (adapted from EDP (2009)).

Considering that each borehole lists the information at several depth ranges, *i.e.* each borehole allows obtaining numerous points containing geotechnical information, all these points must be represented and considered in the optimisation. Due to this fact, and similarly to the case study presented in Chapter 6 where a point alignment optimisation was performed, the SA algorithm had to be adapted and modified in order to replicate the Z coordinate (depth) to match the reality of the borehole geotechnical information.

Looking further into the details, the optimisation process adopted in this case study starts with the generation of n additional points. Note that these points have the goal of representing the removed boreholes. Then, for each randomly generated point, a total of m different Z coordinates was

created (the X and Y coordinates of the point remain the same). By doing this, the integration of all vertical information that can emerge from a borehole was contemplated. It is worth mentioning, that the optimisation methodology, at this stage, only allows the optimisation of vertical alignments.

In what concerns the geotechnical data used in this application, and differently from the Chilean deposit case study, the Rock Quality Design (RQD), the weathering degree (W) and the fracturing degree (F) were the variables chosen to perform the geostatistical simulation and, subsequently, used in the objective function definition. At this stage, it is important to mention the process used to withdraw and register all the boreholes information. Since the final logs only record something if the RQD, W and F values change along the depth, and for the sake of the geostatistical simulation, an almost regular grid was adopted for this log reading. Hence, for all the boreholes the geotechnical information was gathered at every 1 m in depth. In this way, some consistency was maintained in all the three variables.

Table 7.1 provides the basic statistics of the seven used boreholes composed by a total of 1036 points with information. According to these basic statistics, the rock mass shows an overall good quality, because the RQD has a mean of approximately 64 and is represented by a fracturing and weathering degree that is in average 3.

To better understand the location of each borehole, as well as the variation of the geotechnical information in depth, Figure 7.2 shows the spatial distribution of the seven boreholes for the three geotechnical variables considered, where it can be noticed that the area of interest is relatively well characterised. In addition, and in a geotechnical point of view, it is possible to observe from the RQD, F and W colour scales that the rock mass shows better quality for greater depths and a higher spatial variability for the first 100 m.

Table 7.1 Basic statistics of the initial data set composed by boreholes with codes S1, S2, S5, S8, S9, S15 and S16.

	RQD	F	W
Number of points with information	1087	1087	1087
Mean	64	3 ¹	3 ¹
Variance	964.4	1.1	0.8
Minimum	0	1	1
Maximum	100	5	5

¹ The mean value of the fracturing and weathering degree was computed to give a notion of the rock parameters averaging value, however, the most correct way to represent this type of variables (with a scale by classes) is computing the histograms.

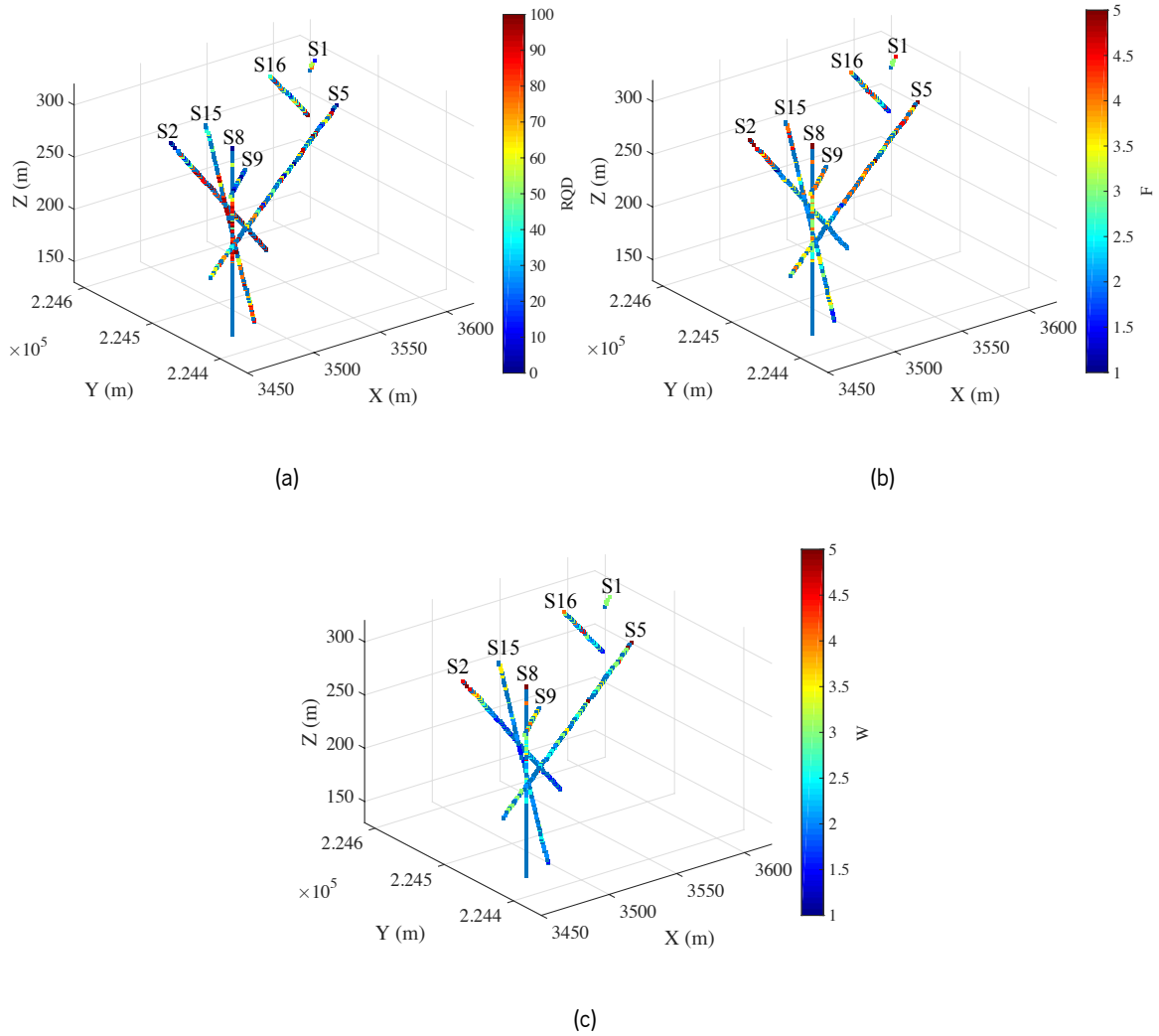


Figure 7.2 Spatial distribution of the boreholes with codes S1, S2, S5, S8, S9, S15 and S16 with a colour scale representing: a) RQD; b) F; and c) W.

7.3. OPTIMISATION PROCESS

Since the variables chosen to perform the geostatistical simulation of this optimisation process are different from the ones used in the Chilean case study, the resulting objective function to use in SA optimisation is consequently different. As mentioned, the data obtained from Salamonde II comprise the information of three geotechnical variables, RQD, F and W, and in order to use the information of all of the three variables, a tri-objective function was built. This was achieved after combining the variance of the geostatistical individual simulation of the three variables. As a consequence, the following minimisation problem was defined,

$$\begin{aligned} \min & (OF_1(Sn), OF_2(Sn), OF_3(Sn)) \\ \text{subject to} & lb \leq Sn \leq ub \end{aligned} \quad (7.1)$$

where,

$$OF_1 = \frac{Var(RQD)}{Var(RQD)_{initial}} \quad (7.2)$$

$$OF_2 = \frac{Var(F)}{Var(F)_{initial}} \quad (7.3)$$

$$OF_3 = \frac{Var(W)}{Var(W)_{initial}} \quad (7.4)$$

Note that OF_1 , OF_2 and OF_3 are the objective functions normalised using the values resulting from the geostatistical simulation of the preliminary data (data used for conditioning the geostatistical simulation). In this case, all the reference values are represented by the variance values of the three considered variables ($Var(RQD)_{initial}$, $Var(F)_{initial}$ and $Var(W)_{initial}$).

Since the RQD system and the fracturing degree classification give a similar type of information, *i.e.* the qualitative classification of the rock mass in terms of fractures, it was found relevant to reduce the overall contribution of both variables in the optimisation process. Hence, considering the Weight Sum Method already presented (see Chapter 6), it was possible to transform the three variables to minimise into a single objective function (see equation (7.5)).

$$\begin{aligned} \min & \quad (W_1 \cdot OF_1(Sn) + W_2 \cdot OF_2(Sn) + W_3 \cdot OF_3(Sn)) \\ \text{subject to} & \quad W_1 + W_2 + W_3 = 1 \\ & \quad lb \leq Sn \leq ub \end{aligned} \quad (7.5)$$

where W_1 , W_2 and W_3 represent the weight values for objective functions OF_1 , OF_2 and OF_3 , respectively. In what concerns the weights values assumed for all the three variables, the following trios were taken into consideration:

$$T_1 = \begin{bmatrix} 0.15 \\ 0.15 \\ 0.70 \end{bmatrix} \quad T_2 = \begin{bmatrix} 0.20 \\ 0.20 \\ 0.60 \end{bmatrix} \quad T_3 = \begin{bmatrix} 0.25 \\ 0.25 \\ 0.50 \end{bmatrix} \quad T_4 = \begin{bmatrix} 0.30 \\ 0.30 \\ 0.40 \end{bmatrix} \quad T_5 = \begin{bmatrix} 0.35 \\ 0.35 \\ 0.15 \end{bmatrix}$$

Going into details about the optimisation process followed, three different scenarios were considered. In the first scenario, the information of one borehole, namely S5, was removed, remaining only the geotechnical information of 6 boreholes. It is important to stress out that this first

scenario was the one chosen to test all the previously presented weight trios. Then, in the second scenario, two boreholes, specifically S8 and S15, were removed from the total of 7 boreholes. These were chosen since they are in different zones of the area of interest (upper and bottom zones). Finally, the third scenario considers the removal of boreholes S5 and S9, chosen due to their proximity and to counter the assumption made in the second scenario

At this point it is important to mention that scenarios 2 and 3 were considered, mainly to execute some assessments regarding the SA input parameters values to adopt in the optimisation process of scenario 1.

The optimisation of each scenario starts by removing the selected boreholes (one or two) and, then, by combining the SA algorithm with geostatistical simulation techniques, their location is optimised. As a matter of fact, the main goal of this application is to understand if, using the developed methodology, one or two of the seven boreholes could be in different positions, bringing more quality information and reducing the error in the area to geomechanically characterise, when compared with the original positions (7 initial boreholes). As a consequence, to validate the proposed methodology the value of the OF obtained for the additional boreholes is compared with the value obtained after the geostatistical conditional simulation of the 7 initial boreholes in the same predefined grid and using the same techniques and conditions as the ones used before. This type of analysis was imperative to truly quantify the exact gain of using the proposed methodology in comparison with the actual reality of the borehole plans. In Figure 7.3 one can find the 3D graphical representation of the boreholes remaining in all the three scenarios using the RQD as an example variable.

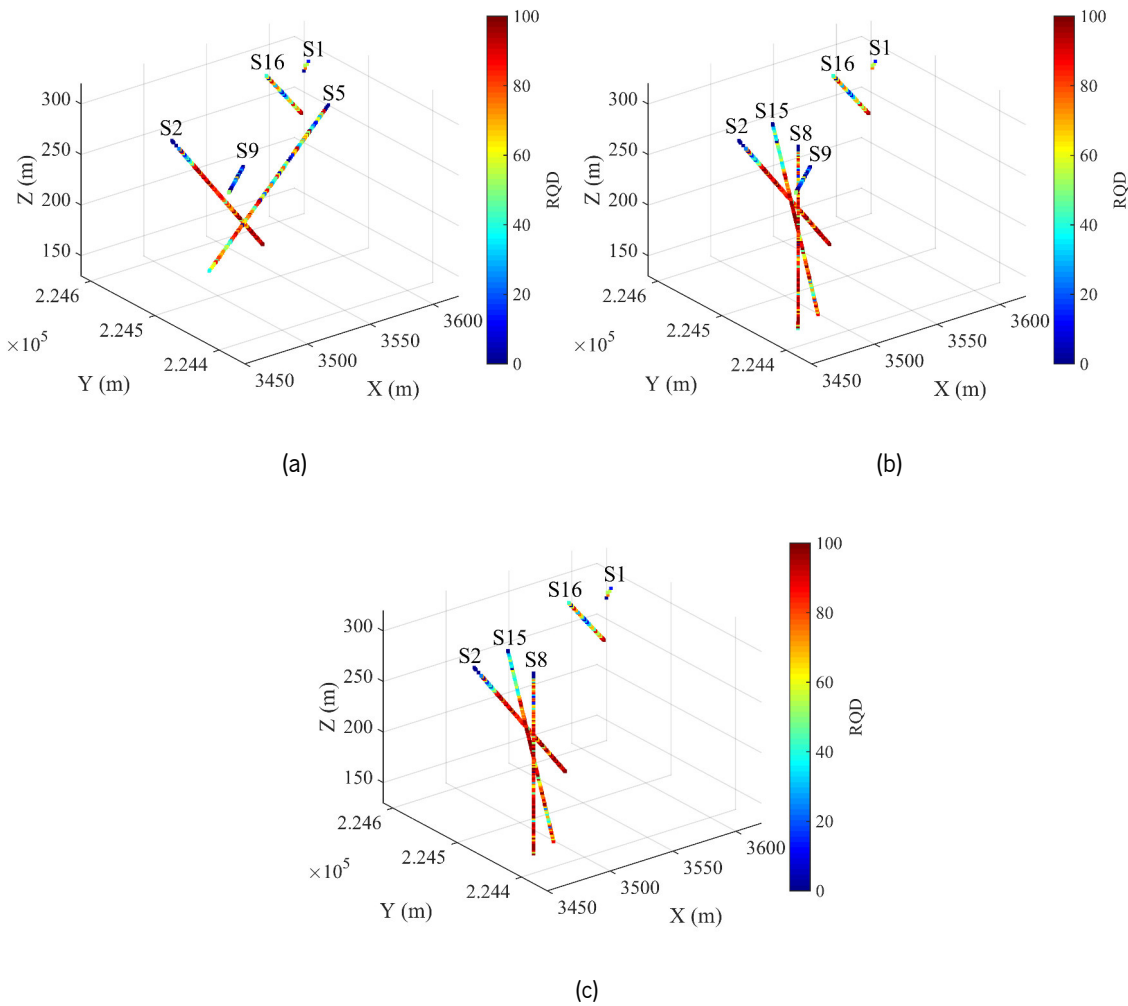


Figure 7.3 Spatial distribution of the remaining boreholes with a colour representing the RQD in: a) scenario 1; b) scenario 2; and c) scenario 3.

7.4. GEOSTATISTICAL SIMULATION

7.4.1. Scenario 1

To perform the geostatistical simulation the main steps adopted in the first case study are the same adopted here, and are based on the ones presented in Pinheiro et al. (2016a, 2016b). So, first, the RQD, F and W selected data composed of only 6 boreholes (since one borehole was removed to execute the optimisation) had to be analysed through the calculation of basic statistics (see Table 7.2). Based on these results, and since the variation ranges of the three variables are wide, the rock mass under consideration can be classified as slightly heterogeneous in all the area.

Table 7.2 Basic statistics of boreholes with codes S1, S2, S8, S9, S15 and S16 (6 boreholes).

	RQD	F	W
Number of information points	855	855	855
Mean	65	3	3
Variance	1012.8	1.0	0.8
Minimum	0	1	1
Maximum	100	5	5

In a second stage, the selected data were transformed into data with a standard Gaussian distribution with a zero mean and a unit variance. From this point, a subsequent step was the analysis of existing correlations between the three parameters. This checking will allow confirm if they are not (or are weakly) cross-correlated. Otherwise, the separate variable simulation should be replaced by a joint simulation (cosimulation), which would make the model quite more complex but accurate (Emery and Cornejo, 2010). Table 7.3 displays the correlation coefficient between the three variables computed using the Pearson product-moment.

Table 7.3 Correlation matrix between RQD, F and W.

	RQD	F	W
RQD	1.000	-0.690	-0.605
F	-0.690	1.000	0.600
W	-0.605	0.600	1.000

Accordingly to Table 7.3 results, it was possible to notice that the correlation between the RQD and the other two variables is inverse, which hinders the possibility of joint simulation. On the other hand, the weathering (W) and fracturing (F) degrees are strongly correlated (>0.5), indicating, at a first glance, that a joint simulation between both variables should be performed. However, a joint simulation between variables W and F was performed but the results, in terms of realisations average, were similar to the ones of the individual simulations. Indeed, for variable W the cross validation performed showed slight worst results when the joint simulation of both variables is considered. In this sense, an individual simulation of the three variables was adopted. A following step concerns the computation of the experimental variograms to a maximum distance of 100 m, beyond which the data values are found to exhibit a low correlation. To fit the experimental variograms, isotropic spherical and Gaussian functions were used. Equations (7.6), (7.7) and (7.8) present the functions used in the variogram fitting, where the distance written between brackets represents the correlation range and the number before the spherical and Gaussian structures

denotes the adopted value for the sill (Chilès and Delfiner, 2012). In Figure 7.4 the experimental and theoretical variograms of the three variables using the information of 6 boreholes are displayed. From figure analysis, it was possible to detect the existence of nugget effect in all three variables, proving some spatial discontinuity of the variables.

$$\text{RQD: } \gamma = 0.36 + 0.31 \text{ Spherical (40 m)} + 0.22 \text{ Gaussian (100 m)} \quad (7.6)$$

$$\text{F: } \gamma = 0.19 + 0.64 \text{ Spherical (20 m)} \quad (7.7)$$

$$\text{W: } \gamma = 0.18 + 0.65 \text{ Spherical (50 m)} \quad (7.8)$$

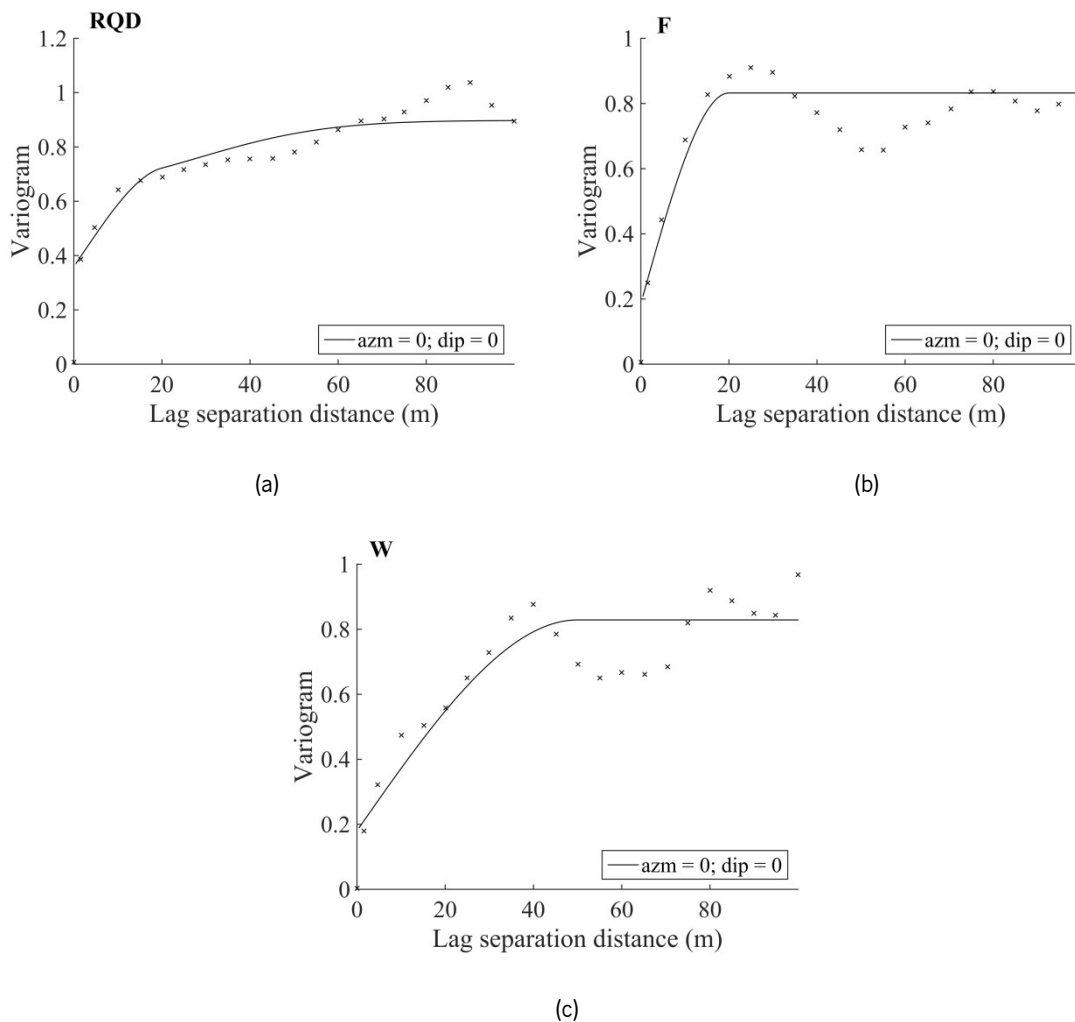


Figure 7.4 Experimental (crosses) and theoretical (solid line) variograms along the horizontal plane using the information of 6 boreholes, for: a) RQD; b) F; and c) W.

A two-dimensional regular grid with a mesh of 5 m × 5 m and a total of 50 nodes along the east direction and 70 nodes along the north direction was defined to conditionally simulate the three variables and to calculate the objective function.

To perform the geostatistical simulation at the target grid nodes, the turning bands method (Emery and Lantuéjoul, 2006) was used. Alike to all geostatistical simulations performed in this thesis, the number of turning lines used was 1500 and the number of realisations was set to 100. To condition the simulation an ordinary kriging was used with a *moving neighbourhood* implementation. The radius chosen for the neighbourhood was 400 m with a total of 100 points per octant. As a final step, the simulated Gaussian values were back-transformed into their original scale to map and post-process the output values.

To understand the differences between the initial data and the geotechnical improvements when the removed boreholes are optimised (preliminary data), the simulation conditioned to the initial data (7 boreholes) was also performed on the predefined grid. All the details can be consulted in the following section. Moreover, to facilitate any type of analysis, the reference values used to normalise the results were the variance of the 100 realisations average obtained for the simulation of the 6 boreholes in the predefined grid (see Table 7.4). So, the selected data set is the one that should serve as a basis for future comparisons.

Table 7.4 Basic statistics using the 6 boreholes data set to perform the geostatistical simulation of variables RQD, F and W.

	Average of simulated RQD	Average of simulated F	Average of simulated W
Total of grid points	3500	3500	3500
Mean	76	3	2
Variance (reference values)	764.4	1.0	0.5
Maximum	90	4	3
Minimum	46	1	1

7.4.2. Initial data

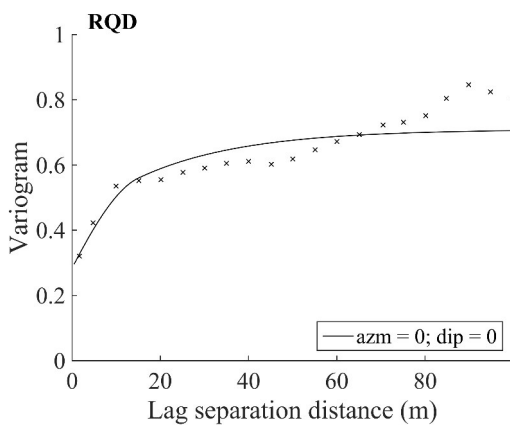
Regarding the initial set experimental variograms, the scenario 1 omnidirectional assumption was made here. For the variogram models, the chosen mathematical functions are presented in Equations (7.9), (7.10) and (7.11), where the distance written between brackets represents the correlation range and the number before the spherical and exponential structures denotes the adopted value for the sill.

In Figure 7.5 the obtained variograms and the corresponding models for all the three variables (RQD, F and W) are represented. Through the variogram analysis, it can be noticed that all three variables variograms exhibit a nugget effect, which implies some spatial discontinuity of the variables, possible due to some error measurement in the data recovery; however, and comparing with scenario 1 variogram models, the nugget effect in all variable is, as expected, smaller since the number of data had increased significantly (the addition of 1 borehole represents the addition of a considerable number of points with information).

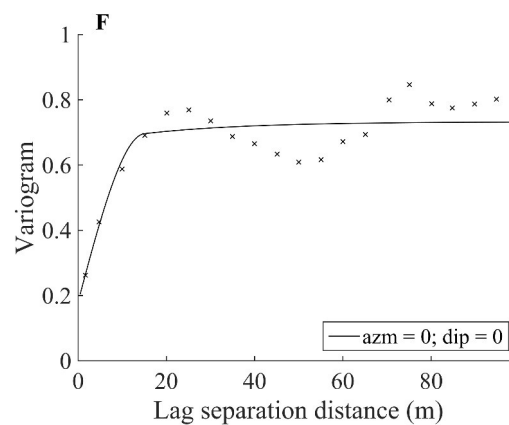
$$\text{RQD: } \gamma = 0.28 + 0.14 \text{ Spherical (15 m)} + 0.28 \text{ Exponential (70 m)} \quad (7.9)$$

$$\text{F: } \gamma = 0.18 + 0.48 \text{ Spherical (20 m)} + 0.07 \text{ Exponential (70 m)} \quad (7.10)$$

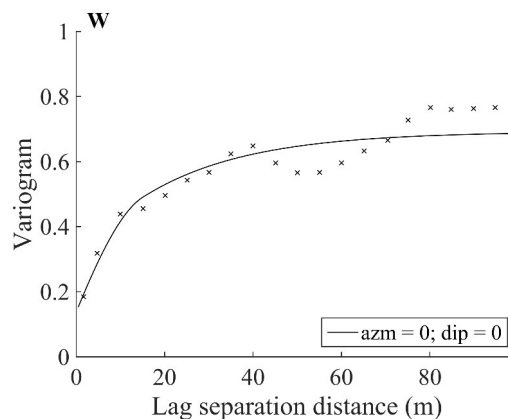
$$\text{W: } \gamma = 0.14 + 0.17 \text{ Spherical (20 m)} + 0.38 \text{ Exponential (70 m)} \quad (7.11)$$



(a)



(b)



(c)

Figure 7.5 Experimental (crosses) and theoretical (solid line) variograms using the initial data set of 7 boreholes, for: a) RQD; b) F; and c) W.

For the conditional simulation, the turning bands method was chosen with a total of 1500 turning lines, while the total number of realisations was set to 100. The remaining simulation assumptions were the same as the ones adopted for scenario 1. As a result, Table 7.5 shows the basic statistics of the 7 original boreholes simulation, including the normalised values of all the three variables using as a reference the variance obtained from the 6 boreholes simulation.

Table 7.5 Basic statistics using the 7 boreholes data set to perform the geostatistical simulation of variables RQD, F and W.

	Average of simulated RQD	Average of simulated F	Average of simulated W
Total of grid points	3500	3500	3500
Mean	79	3	2
Variance	630.6	0.9	0.50
Variance normalised	0.8	1.0	0.9
Maximum	99	4	3
Minimum	44	1	1

7.5. SIMULATED ANNEALING INPUT PARAMETERS STUDY

7.5.1. Introduction

Before presenting any optimisation results, it is important to refer that an additional SA parameter study had to be carried out. A similar study was performed for the case study presented in Chapter 6, mostly to assess the temperature values and rejections inside each temperature state. Even though, since the problem under consideration is different, additional parameters had been tuned. Therefore, the study was divided into two parts: 1) along with some changes of the maximum number of rejections and moves, the main goal was to assess the step size value to use in the function responsible for the new points generations. In this regard, a total of seven tests was performed where the step size and maximum number of rejections and acceptances were varied. In this first test only the information of scenario 2 was used; 2) then, and after finding the best step size value, the points generating function had to be chosen out of two. This was an important step to guarantee that the algorithm truly explores all the space in study; therefore, in this second test not only scenario 2 was tested but also scenario 3, which assumes different removed boreholes. This third scenario inclusion diminishes a possible influence of the boreholes positions in the points generating function choice.

7.5.2. Step size value (Scenario 2)

To perform the step size study, the information of scenario 2 (without boreholes S8 and S15) was used in the optimisation. As such, a total of seven tests was executed where different step values varying from 1 to 0.01 were assumed. For the initial temperature, a higher value of 4 was adopted so that the algorithm could explore all the space. Also, from test number 1 to test number 3, the function used to generate the points coordinates was Equation (7.12), while from test number 4 to test number 7 the function was the one presented in Equation (7.13). The main difference between both equations is in the range of random values used to increment in the generated points; in the first the increment is lower and the algorithm performs closer searches, whereas in the second equation the algorithm is allowed to go further.

$$P_{new} = P_{old} + s \times (ub - lb) \times rand[-1; 1] \quad (7.12)$$

$$P_{new} = P_{old} + s \times (ub - lb) \times rand[-2; 2] \quad (7.13)$$

The new point P_{new} is generated using a previously generated point, P_{old} , an advance step size value (s) and the difference between the lower bound (lb) and the upper bound (ub) of the searching space, and the random part of this generation is given by the product of the difference with a range of random values.

In this regard, Table 7.6 displays the values adopted in all the tests for the SA input parameters. The main goal was to tune the step size (s) value and the number of iterations that should be executed to result in the minimum objective function value.

Table 7.6 Input parameters values for all the seven tests performed using information from scenario 2 to study the step size value.

Test n.	Cooling factor (α)	Initial Temperature (°)	Final Temperature (°)	Max. number of rejections	Max. number of moves	Max. number of accepted points	Step size (s)
1	0.8	4	0.1	300	50	15	0.01
2	0.8	4	0.1	300	50	15	0.10
3	0.8	4	0.1	300	50	15	1.00
4	0.8	4	0.1	300	50	15	0.10
5	0.8	4	0.1	500	100	30	1.00
6	0.8	4	0.1	500	100	30	0.10
7	0.8	4	0.1	400	150	20	0.10

The results of all the seven optimisation tests are represented in Figure 7.6, from where it is possible to observe that the best results were obtained for tests number 2 and number 4, having both adopted a step size of 0.10. These two tests gave the lowest value for the objective function and both converged before reaching 700 iterations. However, both tests have different functions to generate the random points, meaning that an additional study has to be carried out to choose the best point generating function.

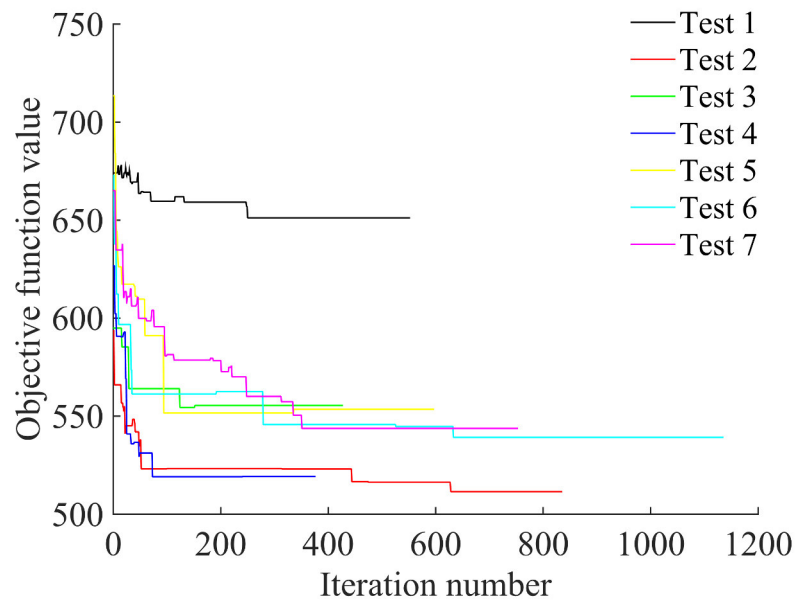


Figure 7.6 Convergence study for a total of seven tests using as information the scenario 2.

7.5.3. Points generating function study (Scenarios 2 and 3)

As mentioned before, tests number 2 and 4 were the ones that showed the best results. Hence, the SA input parameters values adopted in this function study were the same as the ones used in the previous study. Since test number 4 had stopped before the 400 iterations, this was the value used here. The number of iterations was set constant since the convergence study was not carried out at this stage (see Table 7.7). In this subsection, a study to choose the best function to generate new points was carried out, considering the data from scenarios 2 and 3.

Table 7.7 Input parameters values for tests number 2 and 4 using information from scenarios 2 and 3 to study the points generating function.

Test n.	Cooling factor (α)	Initial Temperature (°)	Final Temperature (°)	Max. number of rejections	Max. number of iterations	Max. number of moves	Max. number of accepted points	Step size (s)
2a	0.8	4	0.1	300	400	50	15	0.1
4a	0.8	4	0.1	300	400	50	15	0.1

In the following tables the main results for the optimisation tests performed using the data from scenarios 2 and 3 are represented. Each table contains the number of evaluations, which corresponds to the number of times that the OF was executed, the final temperature value, working here as one of the SA stopping criteria, the maximum number of iterations, the OF minimum value, which should be the lowest possible, and the time spent in each run. Furthermore, the optimal solution coordinates are also given. Note that these coordinates represent the starting point location for the borehole. The results from all the four runs executed assuming the input parameters value from test 2a are represented in Table 7.8 and 7.9, while for test 4a they are displayed in Table 7.10 and 7.11 for scenarios 2 and 3, respectively.

Table 7.8 Scenario 2 optimisation results of test number 2a.

Run n.	Number of evaluations	Final Temperature (°)	Objective function value	Time	Optimal points		
					X (m)	Y (m)	Z (m)
1	409	0.7	540.4	17h20	3576	224389	122
					3506	224519	227
2	409	0.7	517.1	17h20	3389	224362	126
					3510	224436	131
3	409	0.7	544.0	17h20	3628	224565	187
					3520	224502	152
4	409	0.7	544.1	17h10	3491	224446	271
					3449	224540	150

Table 7.9 Scenario 3 optimisation results of test number 2a.

Run n.	Number of evaluations	Final Temperature (°)	Objective function value	Time	Optimal points		
					X (m)	Y (m)	Z (m)
1	409	0.7	353.0	16h35	3489	224390	337
					3521	224534	338
2	409	0.7	348.6	16h35	3626	224582	141
					3524	224519	133
3	409	0.7	345.0	16h35	3577	224300	217
					3456	224422	174
4	409	0.7	349.0	16h35	3514	224491	141
					3612	224577	132

Table 7.10 Scenario 2 optimisation results of test number 4a.

Run n.	Number of evaluations	Final Temperature (°)	Objective function value	Time	Optimal points		
					X (m)	Y (m)	Z (m)
1	409	0.7	540.4	17h20	3390	224355	150
					3518	224435	285
2	409	0.7	510.4	17h20	3497	224464	162
					3389	224365	111
3	409	0.7	538.4	17h20	3558	224423	130
					3514	224518	359
4	409	0.7	532.4	17h10	3524	224533	353
					3556	224424	113

Table 7.11 Scenario 3 optimisation results of test number 4a.

Run n.	Number of evaluations	Final Temperature (°)	Objective function value	Time	Optimal points		
					X (m)	Y (m)	Z (m)
1	409	0.7	336.4	17h16	3523	224524	168
					3531	224332	233
2	409	0.7	376.6	17h16	3612	224596	99
					3622	224621	101
3	409	0.7	341.1	17h16	3523	224343	224
					3523	224530	361
4	409	0.7	336.4	17h16	3523	224524	168
					3531	224332	233

In order to analyse the main differences in the convergence between the two tests, 2a and 4a, including their 4 runs, a convergence representation was built. This graphic relates the number of performed iterations for each test in each run with the corresponding objective function value (see Figure 7.7). It was possible to observe that, for scenario 2, test 4a was the one registering the lowest

OF value in run number 2 (see Figure 7.7a). However, the subsequent best test was test 2a, also in run number 2, which confirms that both tested functions can give similar results. In addition, it can be stated that apart from run number 3 of test 2a and 4a, all the other tests and runs stabilised before reaching the maximum number of iterations, meaning that the chosen number of 400 iterations is enough to guarantee a suitable optimisation of the two additional boreholes. Regarding scenario 3 convergence, the lowest OF value was reached in run number 4 for test 4a (see Figure 7.7b). Even though, the difference between both tests are almost residual, confirming, once again, that both tested point generating functions show a good potential in exploring the search space. Moreover, it is worth noticing that almost all the runs have stabilised in the OF minimum value before reaching 400 iterations, with exception of runs number 1 and 4 in test 2a where it was possible to observe a small stabilisation sill.

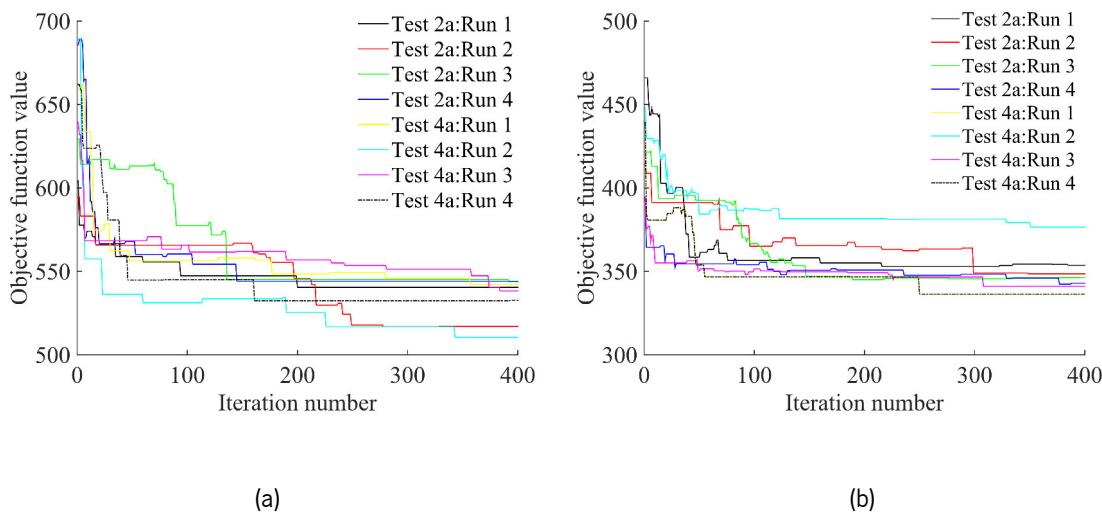


Figure 7.7 Convergence analysis comparing the number of iterations with the OF value for tests number 2a and 4a in each run for: a) scenario 2; and b) scenario 3.

Moreover, Figure 7.8 depicts the starting locations (XY plane) of all the optimal points obtained in both tests and scenarios, as well as the original removed boreholes positions (starting coordinates) to facilitate any further comparison. By analysing Figure 7.8a, it was possible to observe for scenario 2 that almost all the tests provide a similar solution, supported by the density of optimal points located near each other. Also, it is worth mentioning that three clusters (density of points) could be defined, the first near one of the removed boreholes (upper zone), a second one in the right side of the other removed borehole and, finally, a third cluster defined far from each one of the original boreholes coordinates in the left bottom part of the search square. These three clusters intend to translate possible solutions if two new additional boreholes had to be executed. In what concerns scenario 3, from Figure 7.8b analysis, it is possible to detect two main clusters of points, one in the

upper right part of the area and the other one in the central left part; however, none of them near the original boreholes points. However, two of the obtained points are located near one of the original boreholes (test 2a in runs number 1 and 3). Similarly to scenario 2, the identified clusters represent the most probable zones of the rock mass where two new additional boreholes should be executed.

Furthermore, in order to understand the moves from run to run in each test of scenarios 2 and 3, a new $2D$ representation was added (see Figure 7.9 and 7.10). From there, it was possible to confirm that for a total of 4 runs the algorithm tends to jump between the same solutions, *i.e.* from run to run the first and second points exchange positions, which confirms that those are, indeed, the optimal positions. The mentioned effect can be detected, more easily, between the runs of test 4a (see Figure 7.9b and 7.10b).

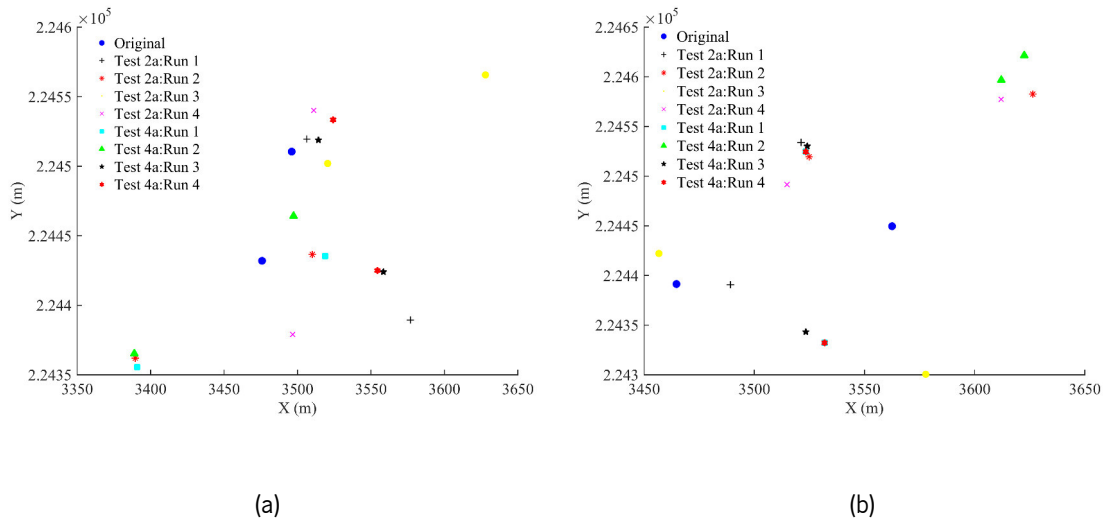


Figure 7.8 Spatial representation (XY plane) of the optimal points obtained for test number 2a and 4a in all the runs along with the original locations of the two removed boreholes (blue circles), for: a) scenario 2; and b) scenario 3.

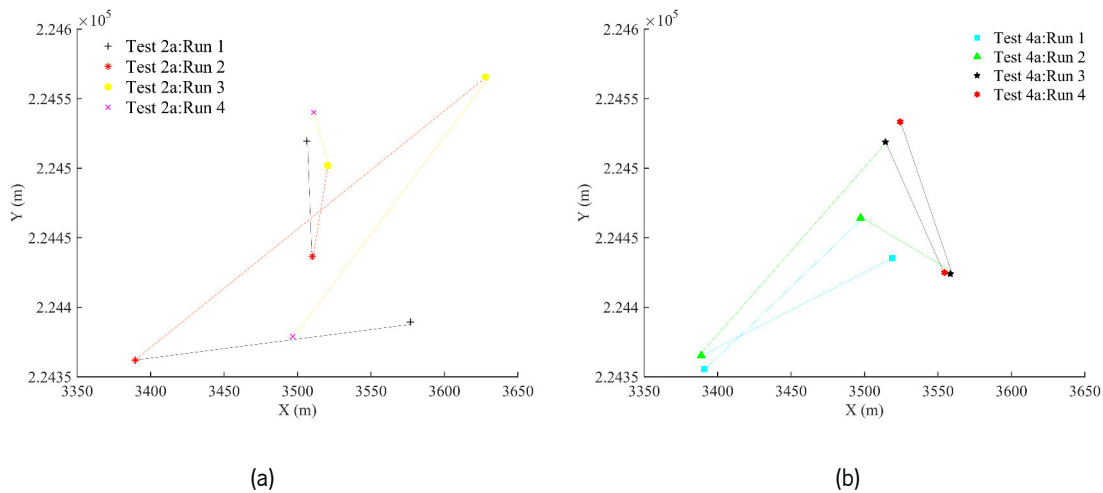


Figure 7.9 Scenario 2 spatial representation (XY plane) of the optimal points movements (coloured arrows) obtained from run to run, in: a) test 2a; and b) test 4a.

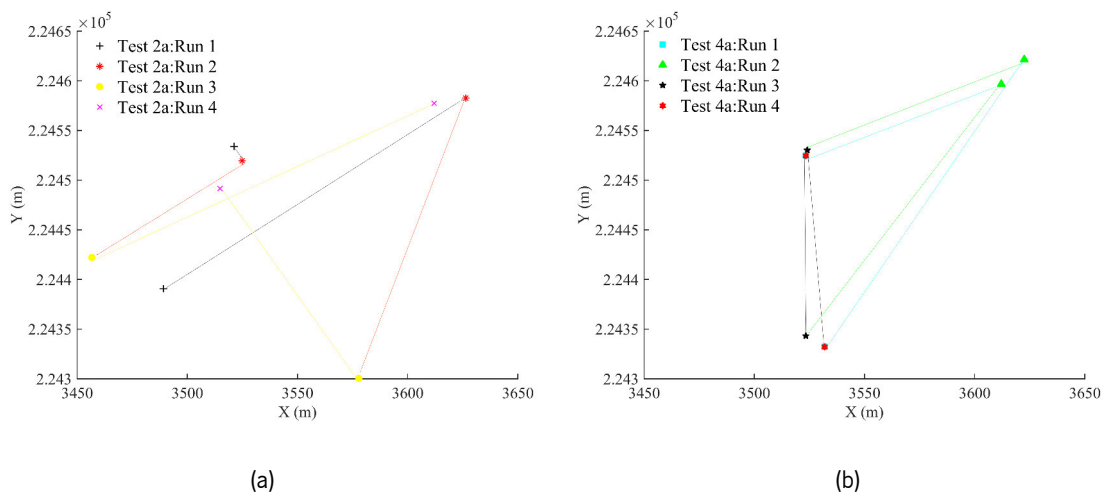


Figure 7.10 Scenario 3 spatial representation (XY plane) of the optimal points movements (coloured arrows) obtained from run to run using data from scenario 2, in: a) test 2a; and b) test 4a.

7.6. OPTIMISATION RESULTS

7.6.1. Scenario 1

From the previous section results, it was possible to observe that the most adjustable step size value was 0.10, while the most exploratory point generating function was the one presented in Equation (7.13). In addition, using the results of the SA internal parameters study it was possible to define the range of values to adopt for the number of iteration threshold. The optimisation tests executed using the data from scenario 1 considers a threshold of 400 and 500 iterations. The adopted values for the SA input parameters are represented in Table 7.12, being the maximum number of iterations the only difference between tests number 1 and number 2. Note that all the assumptions made for the SA input parameters study are valid here, once the number of removed boreholes is the single changeable feature between scenarios.

Furthermore, for this first scenario, to better represent the boreholes information in depth, the average of points in the Z -axis of all the six considered boreholes was computed. As such, an average number of 150 points per alignment (borehole) was considered in the point generation by the SA algorithm.

Table 7.12 Input parameters values adopted to perform the optimisation using the data from scenario 1.

Test n.	Cooling factor (α)	Initial Temperature (°)	Final temperature (°)	Max. number of rejections	Max. number of iterations	Max. number of runs	Max. number of accepted points	Step size
1	0.8	4	0.1	300	400	50	10	0.1
2	0.8	4	0.1	300	500	50	10	0.1

In the following tables the main results of all the runs completed in the optimisation of scenario 1 are presented. Each table contains the number of objective function evaluations, the final temperature value, which represents one of the stopping criteria, the minimum OF value and the total time spent in the process. Besides, the optimal solution, meaning the X , Y and Z coordinates, are also represented in the tables, as well as the final coordinate Z , which gives the maximum depth to consider in this new borehole execution. Each trio was tested 4 times, 2 runs for a limit of 400 iterations and 2 times for a limit of 500.

The optimisation results obtained using the weights from trio number 1 are displayed in Table 7.13. As can be observed, in all the tests the objective function value was almost the same (0.82). This

fact confirms that the maximum number of iterations, for trio 1 weights, is not as relevant as expected. Figure 7.11a shows the spatial location of the optimal points, from where it is possible to observe that they are located near each other which validate the optimisation methodology.

Table 7.13 Optimisation results using the weight values from trio number 1.

Test n.	Run n.	Number of evaluations	Final temperature (°)	Objective function value	Time	Optimal point			Final Z cord. (m)
						X (m)	Y (m)	Z (m)	
1	1	400	0.3	0.82	72h	3505	224346	300	102
	2	400	0.3	0.81	72h	3516	224377	300	100
2	1	500	0.3	0.82	76h	3511	224390	300	101
	2	500	0.3	0.82	76h	3515	224389	300	104

If the weights values from trio number 2 are assumed, the obtained results are quite similar, not only in the final objective function value but also in the points coordinates (see Table 7.14). As a matter of fact, some of the optimal points are almost in the same position as the ones obtained in trio 1, as can be observed in Figure 7.11b. In what concerns the computational time, 72 hours and 76 hours (approximately 3 days) were the registered times from a maximum number of 400 and 500 iterations, respectively. Note that all the SA runs were performed using a desktop computer with an Intel ® Core™ i7-3610QM CPU @2.30GHz processor.

Table 7.14 Optimisation results using the weight values from trio number 2.

Test n.	Run n.	Number of evaluations	Final temperature (°)	Objective function value	Time	Optimal point			Final Z cord. (m)
						X (m)	Y (m)	Z (m)	
1	1	400	0.3	0.82	72h	3525	224396	300	101
	2	400	0.3	0.82	72h	3512	224393	300	100
2	1	500	0.3	0.82	76h	3510	224369	300	101
	2	500	0.3	0.82	76h	3505	224366	300	100

Regarding trio 3, the optimisation results were similar to the previous ones. Indeed, the objective function values were the same as obtained before (see Table 7.15). Even though the optimal points are still in the same zones as for trio 1 and 2, in this case they are located more closely to each other, reinforcing, once again, the lower influence that the number of iterations assumes (see Figure 7.11c).

Likewise, assuming the weight values from trio number 4 (see Table 7.16), the obtained results are quite similar to the ones obtained in trio 3, not just in the objective function value but also in their

locations, as can be perceived by comparing Figure 7.11c and 7.11d. Finally, the results from trio number 5 are displayed in Table 7.17, which are in line with all the previous trio results.

Table 7.15 Optimisation results using the weight values from trio number 3.

Test n.	Run n.	Number of evaluations	Final temperature (°)	Objective function	Time	Optimal point			Final Z cord. (m)
						X (m)	Y (m)	Z (m)	
1	1	400	0.3	0.82	72h	3511	224389	300	100
	2	400	0.3	0.81	72h	3513	224374	300	102
2	1	500	0.3	0.82	76h	3517	224366	300	100
	2	500	0.3	0.81	76h	3509	224372	300	102

Table 7.16 Optimisation results using the weight values from trio number 4.

Test n.	Run n.	Number of evaluations	Final temperature (°)	Objective function	Time	Optimal point			Final Z cord. (m)
						X (m)	Y (m)	Z (m)	
1	1	400	0.3	0.82	72h	3524	224356	300	102
	2	400	0.3	0.82	72h	3514	224371	300	100
2	1	500	0.3	0.82	76h	3509	224365	300	101
	2	500	0.3	0.82	76h	3525	224362	300	102

Table 7.17 Optimisation results using the weight values from trio number 5.

Test n.	Run n.	Number of evaluations	Final temperature (°)	Objective function	Time	Optimal point			Final Z cord. (m)
						X (m)	Y (m)	Z (m)	
1	1	400	0.3	0.82	72h	3505	224357	300	102
	2	400	0.3	0.82	72h	3511	224367	300	100
2	1	500	0.3	0.82	76h	3519	224381	300	101
	2	500	0.3	0.82	76h	3538	224346	300	102

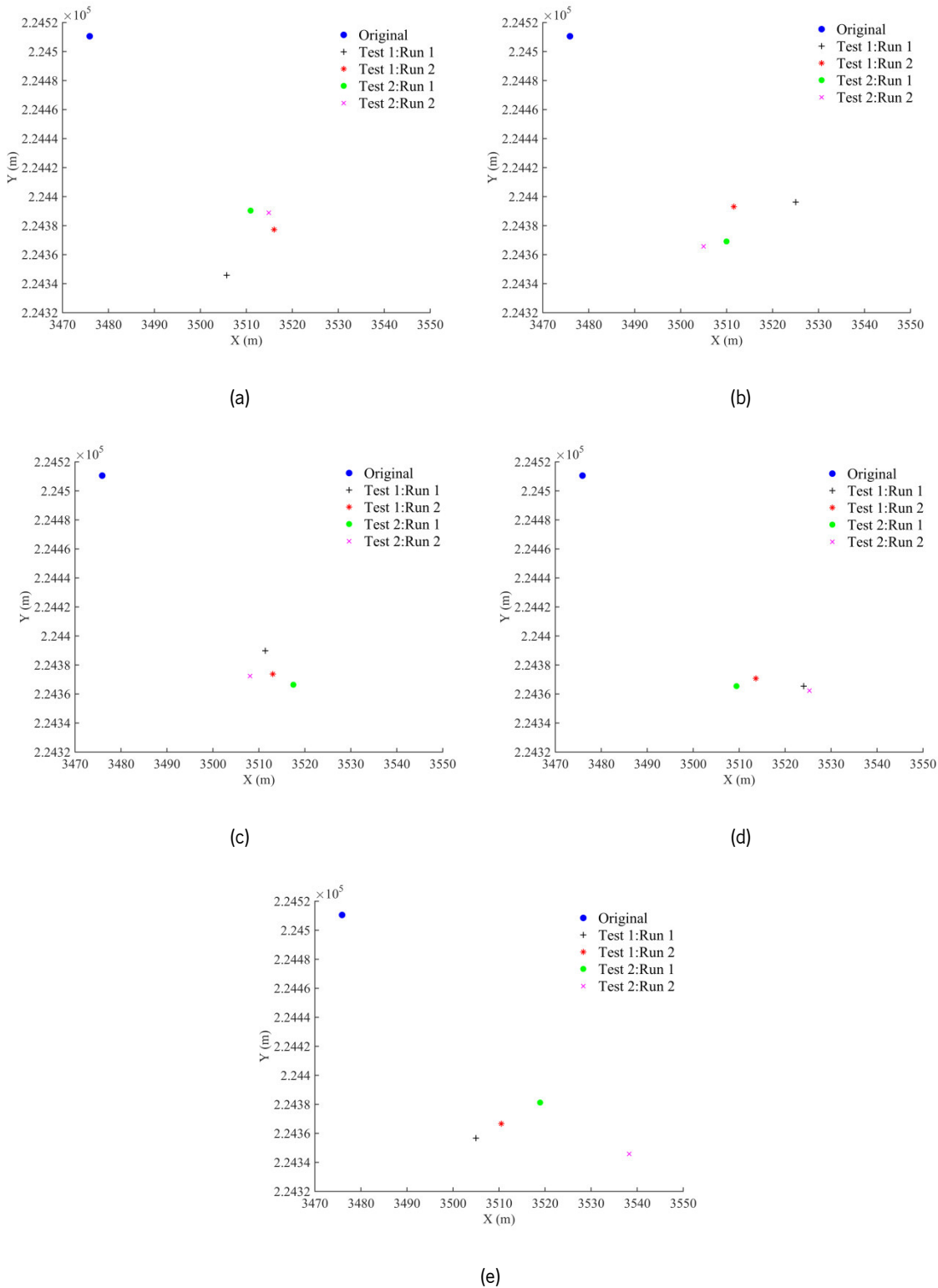


Figure 7.11 Spatial representation (XY plane) of the optimal points (colour points) obtained in all the tests, for: a) trio 1; b) trio 2; c) trio 3; d) trio 4; and e) trio 5.

The convergence study of all the trios that were analysed and the graphical results are displayed in Figure 7.12. The behaviour of each run for each trio was quite similar, having all of them found the minimum value for the OF before reaching the 400 iterations and, more precisely, in run number 2

of test 1. Even though, the remaining runs resulted in values near the optimum. The difference between the optimal solution and the worst solution was of 10% in all the tests.

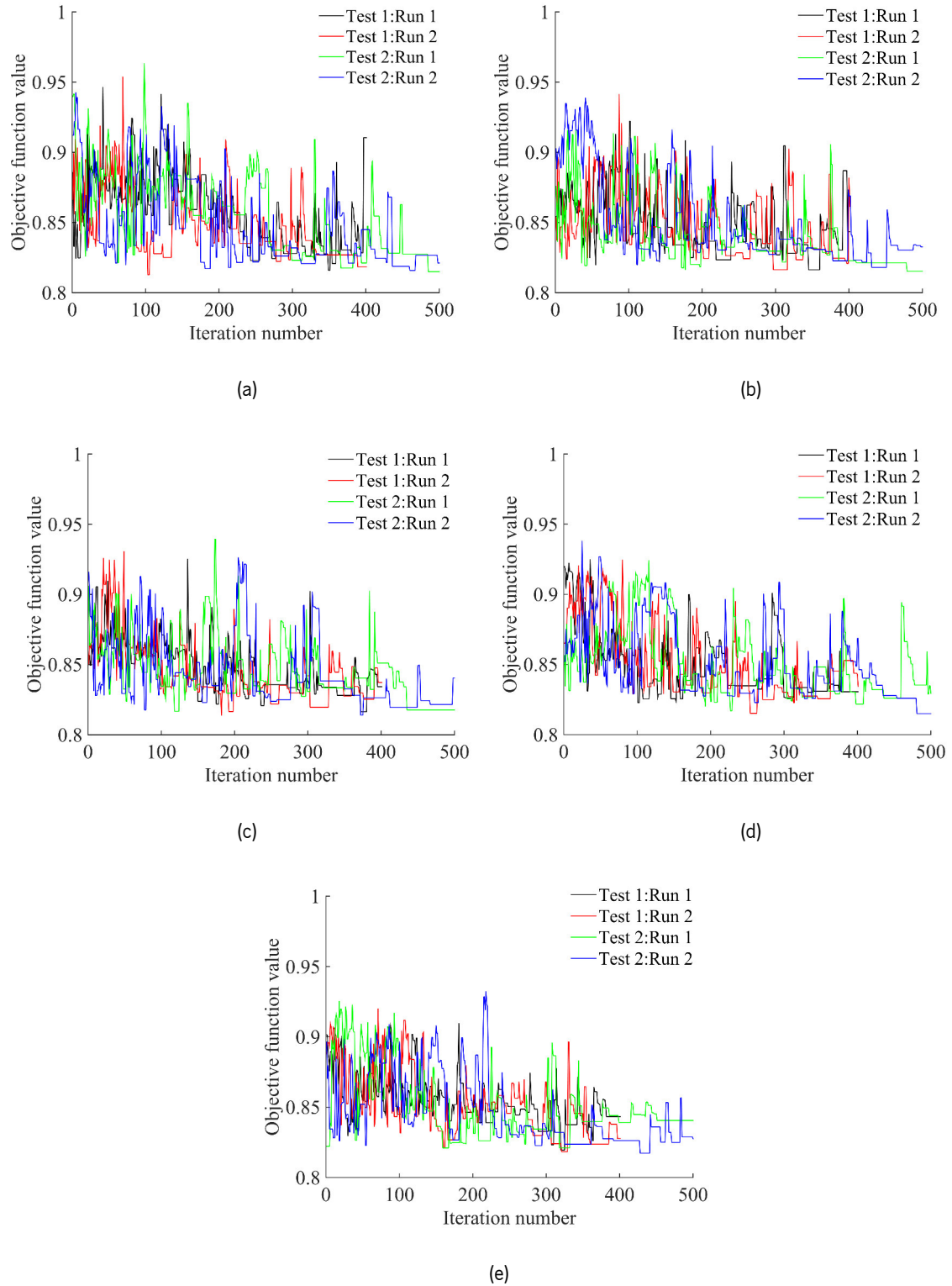


Figure 7.12 Convergence study for all the tests performed assuming the weights values from: a) trio 1; b) trio 2; c) trio 3; d) trio 4; and e) trio 5.

7.6.2. Analysis of the results

All the performed tests provide relevant information into how using this optimisation methodology is possible to gain some geotechnical detail and accuracy in the numerical model. Regarding scenario 1 optimisation, all the five trios were tested four times, each with the expectation of getting different results between them, since different weights were assigned to the three variables into consideration: however, the obtained differences between the trios were almost residual. In this regard, Table 7.18 contains the gain in each run of all the five trios in comparison with the OF value obtained after simulating the initial set. In there, it is possible to observe that almost every trio show a 11% improvement in the OF values. Even if apparently small, a difference of 11% in geotechnical engineering can be interesting and important, mostly, in a scope of 7 boreholes. Moreover, because in this case the resulting optimal points in each tested trio were located far from the original borehole position, as can be seen in Figure 7.13. In detail, a single cluster was formed from this optimisation case study, confirming the assertiveness of this optimisation methodology since, in every run and for every tested weights trios, the result is almost the same. Logically, even if the optimal borehole location was far from the original one, an evaluation from the professional should be employed in order to understand the advantages in performing the borehole in that place and probably drill an inclined borehole in the optimal rock zone limited by [3505, 3538; 224346, 224389].

Regarding the weights trios tested in this scenario, contrarily to expected, the differences were insignificant. This leads to the conclusion that, in this case, the distinction made in the three variables in terms of the global geotechnical gain were not relevant, but can exist and influence further analysis if the professional so decides.

Table 7.18 Results comparison between the optimised objective function values and the initial ones considering the weight values from all the five trios.

Test n.	Run n.	Initial OF value (7 boreholes)	Δ (%)				
			Trio 1	Trio 2	Trio 3	Trio 4	Trio 5
1	1	0.9	-11%	-11%	-11%	-10%	-11%
	2		-11%	-11%	-11%	-11%	-11%
2	1		-11%	-11%	-11%	-11%	-11%
	2		-11%	-11%	-11%	-11%	-11%

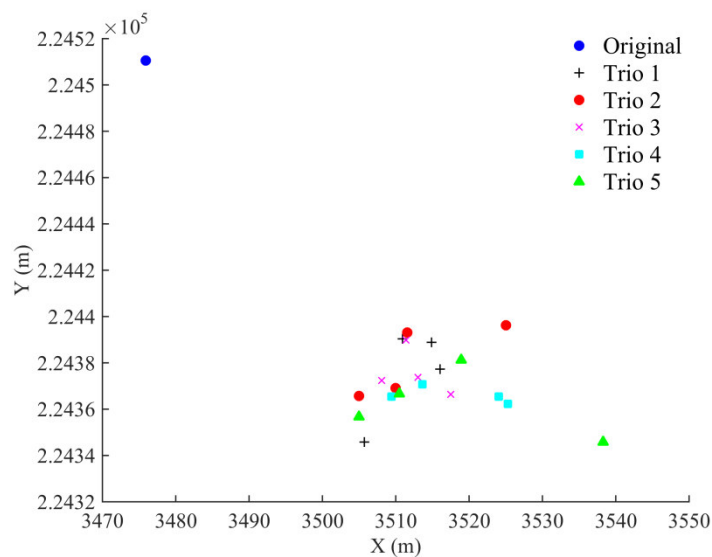


Figure 7.13 Spatial representation (XY plane) of the optimal points (colour points) obtained in all the tests for all the five trios.

In what concerns scenarios 2 and 3, both provide an important insight into how this methodology increases the geotechnical detail of the rock mass and decreases the associated uncertainties comparing to the traditional positioning method. This can be stated after analysing the existing differences of OFs' between both scenarios and the one obtained considering the initial boreholes data set (7 boreholes). As already mentioned, the initial data set was simulated into the same target grid as used in this optimisation. Using the OF value from the geostatistical simulation using the initial set (OF=779.28), percentage differences were calculated and the outcomes are presented in Table 7.19. In detail, for scenario 2 the geotechnical gain, here measured by comparing the average variance resulting from the conditional simulation, was, in average, of 32%. For scenario 3 the calculated gains were higher, reaching a value of 55% in average. These values can be translated in an uncertainty decrease in the geotechnical model if the two removed boreholes, for both scenarios, were executed in the optimal positions instead of their original locations.

Table 7.19 Comparative analysis using the optimisation results of scenarios 2 and 3 with the OF value obtained from the initial set simulation.

Test n.	Run n.	Optimal OF value		Initial OF value (7 boreholes)	Δ (%)	
		Scenario 2	Scenario 3		Scenario 2	Scenario 3
2a	1	540.4	353.0	779.3	-31%	-55%
	2	517.1	348.6		-34%	-55%
	3	544.0	345.0		-30%	-56%
	4	544.1	349.0		-30%	-55%
4a	1	540.4	336.4		-31%	-57%
	2	510.4	376.6		-35%	-52%
	3	538.3	341.1		-31%	-56%
	4	532.4	336.4		-32%	-57%

7.7. CONCLUSIONS

The goal of this work was to apply a new methodology developed to optimise borehole plans. The methodology combines the geostatistical simulation, core technique of this PhD thesis, with the optimisation algorithm SA. All the details about the methodology were given in Chapter 6. In this application, a real case study was used to truly assess and validate the optimisation methodology. Therefore, the data resulting from the design phase of the power reinforcement of Salamonde dam were taken into consideration. The main goal was to use the information of previously executed boreholes, mainly the RQD, W and F geotechnical information, remove some of them and then perform their optimisation. As results, the optimal coordinates could be compared with the real ones and some uncertainty assessment was made comparing the objective function of the optimal and real data.

In detail, the objective function was defined after combining the variance average value of the mentioned geotechnical variables, the RQD empirical system and the qualitative evaluation of the rock mass fracturing and weathering degrees. To build this single objective function, a weight sum method was adopted. The goal was defining different weight values according the three variables importance in the geotechnical detail desired by the user. As input, the information of seven boreholes was used. These boreholes compose the prospection plan of an important underground work built to reinforce the electrical power of the already built Salamonde dam. Contrarily to the Chilean deposit case study, in this application only the alignments optimisation was performed. Also, the main goal was somewhat different since it was not to add new boreholes and quantify the decrease in terms of the OF value, but to remove the original boreholes, use the remaining ones to

perform the optimisation and compare the OF value with the original location. In this regard, three different scenarios were tested: the first considered the removal of one borehole and works as the main tested scenario where the weight values trios were tried, then in the second and third scenarios two distinct boreholes were removed and their locations were optimised.

Compared with the first scenario, the measured geotechnical gain of scenarios 2 and 3, here translated by the difference between the OF value of the optimised scenarios and the original one, was considerably higher, reaching around 55% decrease in the average of the simulation variance. In turn, scenario 1 achieved only 11% of difference, which in geotechnical engineering can still be relevant. In particular, through the analysis of all the tests and runs, it was possible to observe that the optimal point locations were close to one another, which in comparison to the original location of the removed borehole were considerably far (30 m in the X –direction and 70 m Y -direction). This fact leads to the conclusion that: first, the optimisation methodology is coherent in all the tests and runs, providing in almost all cases the same optimal solution, and, secondly, the original boreholes, if executed near the optimal zone had decreased the uncertainty of the geotechnical model. From a geological point of view, scenario 1 results should be faced as trustworthy and a useful help in the decision on where to drill the borehole (since all the runs and trios gave very close results). In detail, the removed borehole should be drilled in a rock zone between 3505 m and 3525 m in the X direction and between 224346 m and 224380 m in the Y - direction, that is a 20 m x 35 m block. Also, and according to all optimisation tests results, the borehole must have a minimum length of approximately 200 m. Note that at this stage the methodology only optimises vertical boreholes and their inclination is not one of the obtained features.

By assuming a multi-criteria approach, throughout the weights implementation, the user of the proposed methodology can have an additional feature to apply according to the type of geotechnical detail needed.

In essence, it is possible to conclude the validity of the methodology to optimise boreholes positions if some initial information exists. It is important to stress out the need in having preliminary borehole information to allow the variogram computation.

As observed by the optimisation results, the uncertainty reduction can be significant if the methodology is used as a helping tool for the professionals at the time of the prospection plan elaboration. As a matter of fact, the time spent in the plan definition can considerably decrease and the level of geotechnical detail increase, which from an engineer point of view is, nowadays, an essential point to optimise not only the work costs but also to respect the deadlines.

Chapter 8

CONCLUSIONS AND FUTURE WORKS

8.1. SUMMARY AND MAIN CONTRIBUTIONS

This thesis addressed one of the major issues of the geotechnical engineering, namely in underground works, the heterogeneity and spatial variability of rock masses. The existing characterisation methodologies are limited to a rock mass modelling as a homogeneous mean, even if the geotechnical community knows that in most cases the rock masses are heterogeneous. The influence of this simplification has relevant impacts in geotechnical works, not only in an economical point of view but also in the safety role.

Seen in these terms, the main goal of this thesis in developing a numerical methodology that are able to reduce uncertainty associated to the geomechanical parameters assessment, their spatial variability and at the same time identify the heterogeneities existing in the rock masses was achieved. Besides the new numerical methodology, another relevant problem, such as the boreholes plan definition, was raised and a new helping methodology was developed.

These two new methodologies were established after an intensive literature review of probabilistic methodologies applied in geotechnics and after performing a first analysis on the robustness and viability of the geostatistics as the core technique of both.

The organisational scheme followed in the thesis included each Chapter conclusions, therefore, in this section, a summary and reference to the main and original contributions of this thesis, as well as some future work, are presented next.

Methodology 1: Numerical methodology to model heterogeneous rock masses

Before starting any kind of original work, an updating and summary of the most recent works, where probabilistic and stochastics techniques were applied in geotechnical works was carried out. At the same time, a list of empirical formulas to be used in order to obtain the geomechanical parameters was made. Moreover, some reference to methodologies used to select the geostatistical realisations was presented aiming to help in the development of a scenario reduction methodology to use after the geomechanical parameters conditional simulation.

In what concerns to the the first methodology original accomplishments worth mentioning, they were the following:

- A first test on the geostatistical techniques applied to geotechnical parameters was carried out. In there, some different approaches (as categorical and as continuous variable) were formulated to simulated the RMR empirical system and their main differences analysed. By using empirical formulas was possible to spatially represent the rock mass deformation modulus in $2D$.
- Individual realisations, realisations average and probability maps were obtained from the geostatistical simulation that can be used for different purposes, since risk analysis to geomechanical parameters assessment.
- A first known attempt to import to the finite difference software, Flac^{3D} the “heterogeneity information” was made. To do that, a new *FISH* code was written to allow the importation of the geomechanical information to each zone centroid, separately.
- Development of a new and innovative methodology that models the rock mass as a heterogeneous mean starting with the preliminary geotechnical information, passing through the conditional geostatistical simulation, followed by a scenario reduction step and finally, the exportation to a finite differences software to perform the rock mass mechanical behaviour analysis.
- The development of a new scenario reduction methodology based on clusters and fitted for the geomechanical parameters geostatistical realisations, as well as the metrics used to assess the statistical representability of the selected realisations in comparison with the full set of realisations.
- Enhanced the need of this new numerical methodology, by comparing the results of the used realisations with the considered deterministic model (shaped to simulate the reality of the nowadays geotechnical models).
- The application of the developed methodology to a real case study, in this case a powerhouse complex of Salamonde II, where a detailed and extensive analysis was made using the developed methodology and a deterministic approach. In there, information obtained in two distinct phases of the works were considered, having the data set of the first phase resulting in overall in better values and adjustments when compared with a second and third data sets composed by a higher number of preliminary information. This fact proves the power of the developed methodology and its advantage since requires a low number of information to give accurate results.
- The increase in security and cost optimisation are some of the main advantages of this methodology, which does not translate into greater *pre*, calculation and *pos* - processing time when compared with the normally used methodologies.

Methodology 2: Borehole plans optimisation methodology

This second methodology is related to the boreholes optimisation through which results the optimal location and the minimum depth to be consider at the time that new additional boreholes are executed. This methodology intends to function as a helping tool to be used by the professionals jointly with yours *know how* in geotechnical prospection. As performed in the thesis first chapter, an updating and summary of existing optimisation methodologies was carried out. From there was possible to assess the need in developing this methodology that combines the geostatistical technique with the optimisation algorithm, Simulated Annealing.

The main conclusions to withdraw from the last two chapters of the thesis, are:

- The formulation of the optimisation problem to be minimised and the associated constrains. To do that, was necessary to combine several techniques, like geostatistics, with the optimisation algorithm.
- The use of two different metrics as the objective function of the optimisation problem. This fact allows the professional to decide which statistical metric is more fitted and reliable according to the geotechnical parameters used in the optimisation.
- The combination of both tested metrics in a multicriteria approach, to understand the influence and behaviour of each objective function in the optimisation process.
- The integration of an alignments of points optimisation instead of individual points, which are closer to the reality since for each borehole the information is gathered at several depths.
- Application of the developed methodology to a Chilean gold deposit, where the information of the RMR empirical systems was used from some randomly selected boreholes and the optimisation was performed one by one until a certain level (13 new boreholes) and in a sparser way until a total of 30 new additional boreholes was achieved. From this analysis was possible to understand the geotechnical gain and uncertainty reduction that results from each additional borehole and decide the best number to execute according to the existing budget.
- The developed methodology was applied to a second case study, this turn the powerhouse complex used before, to which was randomly removed, firstly one borehole and then, two boreholes. The optimisation methodology was applied to assess the removed boreholes optimal locations and only one metric was used as objective function, however, a multicriteria approach with three different geotechnical parameters was carried out in order to understand their influence in the optimisation process.

- Highlighted the conclusions of this second application, from which it was possible to state that by using the developed methodology, the geotechnical gain (error reduction) can reach 10% if one additional borehole is executed and more than 50% in the case of two additional boreholes. Also, the computational time need in not as relevant as other methodologies, and the outputs can be obtained after a day of calculations.

In sum, the main innovative contributions of this thesis were the following:

- The development and validation of a new numerical methodology to use in the rock masses characterisation. The methodology as the advantage of identifying the rock mass heterogeneities, of modelling the geomechanical parameters spatial variability and to reduce the uncertainty associated to its characterisation. The differences in comparison with the deterministic approach commonly adopted are significant, enhance the need in perform this type of numerical analysis in complex underground works.
- The application of innovative tools to simulate the geomechanical parameters, namely the geostatistics that had been proving robust and adequate to use in geotechnical engineering.
- The development of a borehole plans optimisation methodology that was validated as a relevant tool to help the professional in assessing the best location to performed additional boreholes, as well as the economic viability associated to a certain number of additional boreholes.

8.2. FUTURE WORKS

Besides the step forward in applying geostatistical technique to use in the rock mass characterisation, the proposed methodologies are not enclosed to improvements. Both methodologies still have some steps that need optimisation and more thinking in order to improve the quality and range of application for them. Therefore, future works should be focus in the mentioned problem and each methodology step evaluated and tested for different types of formations, namely a shale rock type that adds some complexity in the characterisation due to the existence of the foliation. Regarding the first characterisation methodology the points that need further development are:

- A very important improvement should be performed at the geotechnical data collection since the accuracy of this information will determinate the accuracy of the simulated rock mass models. If the *in situ* and laboratory tests are performed only in rock mass zones that show

the better quality, and since the geostatistical simulation uses the interaction between the points, the simulation will translate and mimic the input values. Therefore, it is of utmost importance gather realistic preliminary information of the rock mass.

- Apply the proposed methodology as it is, in different rock masses, where the presence of heterogeneities is more accentuate and their effects on the rock mass mechanical behavior assessed. Although the rock mass of the Salamonde II powerhouse complex used in this thesis could be classified as homogeneous, however, had permitted to prove some points and advantages of this proposed methodology.

Regarding the second proposed methodology destined to borehole plans optimisation some relevant improvement can be performed, such as:

- In the first part of the optimisation methodology, the chosen variable (geotechnical parameters) was conditionally simulated on the randomly generated points coordinates and the average of the 100 realisations was used to characterise this point; however, and has been said throughout the thesis, the average value tends to smooth the extreme values. For that reason, the scenario reduction methodology developed in the characterisation methodology can be used in this second methodology to select some relevant realisations to use in the remaining optimization process. By doing this, the uncertainty is reduced even more, since a higher number of possibilities is used.
- If the mentioned step is included, an additional step must be added to help deciding the optimal locals for new boreholes, once each one of the selected realisations will be optimised.
- Finally, and because boreholes are not always vertical, a relevant contribution will be the integration of inclined boreholes. To do that, will be required to performed a large amount of calculation since in each randomly generated coordinates a 360° search have to be performed.

Chapter 9

REFERENCES

- Aarts, E. H. L. and Korst, J. H. M. (1989). *Simulated Annealing and Boltzmann Machines*. Wiley, Chichester.
- Ang, A. H.-S. and Tang, W. H. (1984). *Probability Concepts in Engineering Planning and Design, Vol. I: Basic Principles*. John Wiley & Sons.
- Armstrong, M., Galli, A., Beucher, H., Le Loc'h, G., Renard, D., Doligez, B., Eschard, R., and Geffroy, F. (2011). *Plurigaussian Simulations in Geosciences (2nd edn.)*. Springer, Berlin, 176 p.
- Armstrong, M., Ndiaye, A., Razanatsimba, R. and Galli, A. (2013). Scenario reduction applied to geostatistical simulations. *Mathematical Geosciences*, 45, 165-182.
- Baecher, G. B., Lanney, N. A. and Einstein, H. H. (1977). Statistical description of rock properties and sampling. *Proceedings, The 18th U.S. Symposium on Rock Mechanics (USRMS)*, 22-24 June, Golden, Colorado, 1-8.
- Barton, N. (1983). Application of Q-system, index tests to estimate shear strength and deformability of rock masses. *Proceedings of the International Symposium on Engineering Geology and Underground Construction*, Lisbon, Portugal, I (II), 51-70.
- Barton, N. and Quadros, E. (2002). Engineering and hydraulics in jointed rock masses. In ISRM (editor), "EUROCK 2002 - Course A," Funchal, Portugal.
- Barton, N., Lien, R. and Lunde, J. (1974). Engineering classification of rock masses for the design of tunnel support. *Rock Mechanics and Rock Engineering*, 6, 189-236.
- Bieniawski, Z. T. (1976). Rock mass classification in rock engineering. *Proc. Symp Explor Rock Eng*, 1, 97-106. Cape Town: Balkema
- Bieniawski, Z. T. (1989). *Engineering Rock Mass Classifications*. Wiley, New York, USA, 272 p.
- Borghi, A., Renard, P. and Jenni, S. (2012). A pseudo-genetic stochastic model to generate karstic. *Journal of Hydrology*, 414-415, 516-529.
- Bowman, A. W. and Azzalini, A. (1997). *Applied Smoothing Techniques for Data Analysis*. Oxford, Oxford University Press.

- Brus, D. J. and Heuvelink, G. B. M. (2007). Optimization of sample patterns for universal kriging of environmental variables. *Geoderma*, 138, 86-95.
- Cai, M., Kaiser, P. K., Tasaka, Y. and Minami, M. (2007). Determination of residual strength parameters of jointed rock masses using GSI system. *International Journal of Rock Mechanics and Mining Sciences*, 44, 247-265.
- Carvalho, J. (2004). Estimation of rock mass modulus. Personal communication.
- Chilès, J. P. and Delfiner, P. (2012). *Geostatistics modelling spatial uncertainty*. John Wiley & Sons, New York, 699 p.
- Choi, S. Y. and Park, H. D. (2002). Comparison among different criteria of RMR and Q-system for rock mass classification for tunnelling in Korea. *Tunnelling and Underground Space Technology* 17 (4), 391–401.
- Chong, T., Phoon, K. K. and Soubra, A. –H. (2013). Stability analysis of a pressurized tunnel face in a spatially varying sand using the statistical approach of limit analysis. Proceedings of the 11th International Conference on Structural Safety and Reliability, ICOSSAR 2013.
- Christian, J. T. (1996). Reliability methods for stability of existing slopes. Uncertainty in the Geologic Environment: from Theory to Practice, Proceedings of Uncertainty '96, Madison, Wisconsin, USA, ASCE Geotechnical Special Publication, 58, 409–418.
- Christian, J. T. and Baecher, G. B. (2003). *Reliability and Statistics in Geotechnical Engineering*, John Wiley & Sons Ltd, 618 p.
- David, M. (1988). *Handbook of applied advanced geostatistical ore reserve estimation*. Developments in Geomathematics 6. Elsevier Science Publishers B. V., Amsterdam, 216 p.
- Deere, D. U. (1968). Geological consideration. *Rock Mechanics in Engineering Practice*. John Wiley & Sons Ltd.
- Deisman, N., Khajeh, M. and Chalaturnyk, R. J. (2013). Using geological strength index (GSI) to model uncertainty in rock mass properties of coal for CBM/ECBM reservoir geomechanics. *International Journal of Coal Geology*, 112, 76-86.
- Deutsch, C. V. (1989). DECLUS: a fortran 77 program for determining optimum spatial declustering weights. *Computers & Geosciences*, 15 (3), 325-332.

Deutsch, C. V. (1998). Fortran programs for calculating connectivity of three-dimensional numerical models and for ranking multiple realizations. *Computers and Geosciences*, 24, 69-76.

Deutsch, C. V. (1999). Reservoir modeling with publicly available software. *Computers and Geosciences*, 25, 355-363.

Deutsch, C. V. and Journel, A. G. (1998). *GSLIB: Geostatistical Software Library and user's guide*. 2.ed. New York, Oxford University Press, 369 p.

Deutsch, C. V. and Srinivasan, S. (1996). Improved reservoir management through ranking stochastic reservoir models. In: *SPE/DOE Improved Oil Recovery Symposium*, 21-24 April 1996, Tulsa, USA. doi: 10.2118/35411-MS.

Deutsch, C.V. (2002). *Geostatistical Reservoir Modelling*. New York, Oxford University Press, 376 p.

Deutsch, C.V. (2015). Cell Declustering Parameter Selection. Geostatistical Lessons, University of Alberta

Dubost, J. (2009). *Variabilité et Incertitudes en Géotechnique: De leur estimations à leur prise en compte*. Ph.D. thesis, University of Bordeaux I, Bordeaux, France, 356 p.

Dupacova, J., Grove-Kuska, N. and Romisch, W. (2003). Scenario reduction in stochastic programming: an approach using probability metrics. *Mathematical Programming*, 95, 493–511. doi:10.1007/s10107-002-0331-0

EDP (2009). Power reinforcement of Salamonde- Salamonde II. Project- Volume VIII – Stability studies. COBA, TOMO VIII-B: CENTRAL, 82 p.

EDP (2009). Power reinforcement of Salamonde- Salamonde II. Project- Volume V, Geotechnical and Geological Characterisation studies. COBA, TOMO V-A: Geotechnical and Geological studies, 98 p.

Egaña, M. and Ortiz, J. (2013). Assessment of RMR and its uncertainty by using geostatistical simulation in a mining project. *Journal of GeoEngineering*, 8 (3), 83-90.

El-Ramly, H. M. (2001). *Probabilistic analysis of landslide hazards and risks, bridging theory and practice*. M.Sc. thesis, University of Alberta.

Ellefmo, S. L. and Eidsvik, J. (2009). Local and spatial 476 joint frequency uncertainty and its application to rock mass characterisation. *Rock Mechanics and Rock Engineering*, 42(4), 667-68.

- Emery, X. (2007). Conditioning simulations of Gaussian random field by ordinary kriging. *Mathematical Geology*, 39, 607-623. doi:10.1007/s11004-007-9112-x
- Emery, X. and Cornejo, J. (2010). Truncated Gaussian simulation of discrete-valued, ordinal coregionalized variables. *Computers and Geosciences*, 1325-1338.
- Emery, X. and Lantuéjoul, C. (2006). TNSIM: A computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. *Computer and Geosciences*, 1615-1628.
- Emery, X., Ortiz, J. M. and Cáceres, A. M. (2008). Geostatistical modelling of rock type domains with spatially varying proportions: application to a porphyry copper deposit. *Journal of the Southern African Institute of Mining and Metallurgy*, 108, 284-292.
- Englund, E. J. and Heravi, N. (1994). Phased sampling for soil remediation. *Environmental and Ecological Statistics*, 1, 247 – 263.
- Eskelinen, P. and Miettinen, K. (2012). Trade-off Analysis Approach for Interactive Nonlinear Multiobjective Optimization. *OR Spectrum*, 34(4), 803-816.
- Espada, M. and Lamas, L. (2014). Monitored and calculated deformational behaviour of a large cavern during excavation. In Alejano, Perucho, Olalla & Jiménez (eds), *Rock Engineering and Rock Mechanics: Structures in and on Rock Masses*; Proc. Eurock 2014, Vigo, Espanhã, 983-988. Taylor & Francis.
- Etminan, A. H. and Seifi, A. (2008). An improved model for geostatistical simulation of fracture parameters and their effect on static and dynamic models. *The Open Petroleum Engineering Journal*, 1, 47-57.
- Exadaktylos, G. and Stavropoulou, M. (2008). A specific upscaling theory of rock mass parameters exhibiting spatial variability: Analytical relations and computational scheme. *International Journal of Rock Mechanics & Mining Sciences*, 45, 1102-1125.
- Fenton, G. A. (1997). Probabilistic methods in geotechnical engineering. In: *Workshop presented at ASCE GeoLogan'97 conference, Logan, Utah*.
- Ferrari, F., Apuani, T. and Giani, G. P. (2014). Rock Mass Rating spatial estimation by geostatistical analysis. *International Journal of Rock Mechanics and Mining Sciences*, 70, 162-176.

- Galera, M., Álvarez, M. and Bieniawski, Z. T. (2007). Evaluation of the deformation modulus of rock masses using RMR. Comparison with dilatometer tests. *Proceedings of the ISRM Workshop W1*. Madrid, Spain.
- Goovaerts, P. (2001). Geostatistical modelling of uncertainty in soil science. *Geoderma*, 103, 3-26.
- Gravanis, E., Pantelidis, L. and Griffiths, D. V. (2014). An analytical solution in probabilistic rock slope stability assessment based on random fields. *International Journal of Rock Mechanics and Mining Sciences*, 71, 19–24.
- Griffiths, D.V., Fenton, G.A. and Tveten, D.E. (2002). Probabilistic geotechnical analysis: How difficult does it need to be?. *UEF Conference on Probabilistics in Geotechnics*.
- Gringarten, E. and Deutsch, C. V. (2001). Variogram interpretation and modelling. *Mathematical Geology*, 33 (4), 507-535.
- Heitsch, H. and Romisch, W. (2003). Scenario reduction algorithms in stochastic programming. *Computational Optimization and Applications*, 24, 187–206. doi:10.1023/A:1021805924152
- Heitsch, H. and Romisch, W. (2009). Scenario tree modelling for multistage stochastic programs. *Mathematical Programming*, 118, 371-406. doi:10.1007/s10107-007-0197-2
- Hernandez, J. and Emery, X. (2009). A geostatistical approach to optimize sampling designs for local forest inventories. *Canadian Journal of Forest Research*, 39, 1465 -1474.
- Hoek, E. and Brown, E. (1997). Practical estimates of rock mass strength. *International Journal of Rock Mechanics and Mining Science*, 34, 1165–1186.
- Hoek, E. and Brown, E. T. (1997). Practical estimates of rock mass strength. *International Journal of Rock Mechanics and Mining Sciences*, 34(8),1165–1186.
- Hoek, E. and Diederichs, M. S. (2006). Empirical estimation of rock mass modulus. *International Journal of Rock Mechanics and Mining Sciences*, 43, 203–215.
- Hoek, E., Carranza-Torres, C. and Corkum, B. (2002). Hoek-Brown failure criterion-2002 Edition. *Proceedings of NARMS- TAC conference*, Toronto, 1, 267–273.
- Hu, Q. F. and Huand, H. W. (2007). Risk analysis of soil transition in tunnel works. *Underground Space – the 4th Dimension of Metropolises – Bartak, Hrdina*.

- Huber, M. and Hicks, M. A. (2010). Probabilistic calculation of differential settlement due to tunnelling. Guema, van Gelder & Proske: Proceedings of the 8th International Probabilistic Workshop, Szczecin, Stuttgart.
- Imanzadeh, S. (2013). *Effects of uncertainties and spatial variation of soil and structure properties on geotechnical design. Cases of continuous spread footings and buried pipes*. Ph.D. thesis. University of Bordeaux, Bordeaux, France, 279 p.
- Itasca (2012). *FLAC - Fast Lagrangian Analysis of Continua*. Version 5.0 User's manual, Minneapolis, USA.
- Jaksa, M. B., Brooker, P. I. and Kaggwa, W. S. (1997). Inaccuracies associated with estimating random measurement errors. *ASCE journal of Geotechnical and Geoenvironmental Engineering*, 123 (5), 393-401.
- Jeon, S., Hong, C. and You, K. (2009). Design of tunnel supporting system using geostatistical methods. In: Huang, Liu (Ed.), *Geotechnical Aspects of Underground Construction in Soft Ground*, 781–784.
- Journel, A. G. (1974). Geostatistics for conditional simulation of ore bodies. *Economic Geology*, 69, 673-680.
- Journel, A. G. (1977). "Géostatistique minière". Centre de Géostatistique, Fontainebleau, France.
- Kirkpatrick, S., Gelatt, C. D. and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220, 671-680.
- Krige, D. G. (1951). A statistical approach to some basic mine evolution problems on the Witwatersrand. *Chemical Metallurgical and Mining Society of South Africa.*, 52, 119-39.
- Krige, D. G. (1966). Two-dimensional weighted moving average trend surface for one evaluation. *South Africa Institute of Mining and Metal*, 66, 13-38.
- Krige, D. G., Guarascio, M., Camisani-Calzolari, F. A. (1989). Early South African geostatistical techniques in today's perspective, in Armstrong, M., (ed.) *Geostatistics*, pp. 1-19.
- Lantuéjoul, C. (2002). *Geostatistical simulation*. 1st ed. Springer-Verlag Berlin Heidelberg, 256 p. doi: 10.1007/978-3-662-04808-5

- Le Goc, R., Darcel, C., Davy, P., Pierce, M. and Brossault, M. A. (2014). Effective elastic properties of 3D fractured systems. *In: Proceedings of the 1st International Conference on Discrete Fracture Network Engineering, Paper ID 142, Vancouver, Canada.*
- Lin, L. and Fei, C. (2012). The Simulated Annealing Algorithm Implemented by the MATLAB. *International Journal of Computer Science Issues*, 9 (6), 357-360.
- LNEC (2009). Power reinforcement of Salamonde. Tests with a three-dimensional deformer to determine the stress state (in portuguese). Technical report, LNEC, Report 403/09: Lisboa, Portugal.
- LNEC (2012). Power reinforcement of Salamonde. Interpretative model of the initial stress state (in portuguese). Technical report, LNEC, Report 348/2012: Lisboa, Portugal.
- LNEC (2013). Power reinforcement of Salamonde. Analysis of the deformation behaviour of the underground power station. Technical report, LNEC, Report 450/2013: Lisboa, Portugal.
- Luo, Z., Atamturktur, S., Juang, C. H., Huand, Hongwei and Lin, P-S. (2011). Probability of serviceability failure in a braced excavation in a spatially random field: Fuzzy finite element approach. *Computers and Geotechnics*, 38, 1031–1040.
- Madani, N. and Asghari, O. (2013). Fault detection in 3D by sequential Gaussian simulation of Rock Quality Designation (RQD). *Arabian Journal of Geosciences*, 12 (10), 3737-3747.
- Manchuk, J. G., Leuangthong, O. and Deutsch, C. V. (2009). The proportional effect. *Mathematical Geoscience*, 41, 799–816.
- Mao, S. and Journel, A. (1999). Generation of a reference petrophysical/seismic data set: the Stanford V reservoir. 12th Annual Report Stanford Center for Reservoir Forecasting, Stanford USA.
- Marchant, B. P. and Lark, R. M. (2005). Adaptive sampling and reconnaissance surveys for geostatistical mapping of the soil. *European Journal of Soil Science*, 57, 831-845.
- Marler, R. T. and Arora, J. S. (2009). The weighted sum method for multi-objective optimization: new insights. *Structures Multidiscipline Optimization*, 41, 853-862.
- Matheron, G. (1963). The intrinsic random functions and their applications. *Advances in Applied Probability*, 5, 439–468.

Matheron, G. (1971). *The Theory of Regionalized Variables and its Applications*, Ecole Nationale Supérieure des Mines de Paris, Fontainebleau, 212 p.

Matheron, G. (1973). The intrinsic random functions and their applications. *Advances in Applied Probability*, 5, 439-468.

McBratney, A. B., Webster, R. and Burgess, T. M. (1981). The design of optimal sampling schemes for local estimation and mapping of regionalized variables. *Computers & Geosciences*, 7, 331-334.

McLennan, J. A. and Deutsch, C. V. (2005). Ranking geostatistical realizations by measures connectivity. In: *SPE/PS-CIM/CHOA, international thermal operations and heavy oil symposium*, Calgary, Canada.

Metropolis, N. and Ulam, S. (1949). The Monte Carlo Method. *Journal of the American Statistical Association*, 44 (247), 335-341.

Mitri, H. S., Edrissi, R. and Henning, J. (1994). Finite element modeling of cable bolted stopes in hard rock ground mines. *Presented at: SME Annu Meet, Albuquerque, NM, USA*, 94–116.

Mohammadi, H. and Rahmancejad, R. (2010). The estimation of rock mass deformation modulus using regression and artificial neural network analysis. *The Arabian journal for science and engineering*, 1A, 205-217.

Nicholson, G. A. and Bieniawski, Z. T. (1990). A nonlinear deformation modulus based on rock mass classification. *International Journal of Mining Geological Engineering*, 8, 181–202.

Oh, S., Chung, H. and Kee Lee, D. (2004). Geostatistical integration of MT and boreholes data for RMR evaluation. *Environmental Geology*, 46, 1070-1078.

Olea, R. A. (1984). Sampling design for spatial functions. *Mathematical Geology*, 16, 369-392.

Osyczka, A. (1981). An approach to multi-criterion optimization for structural design. In: *Proceedings of the 11th ONR Naval Structural Mechanics Symposium on Optimum Structural Design*, Tucson, AZ.

Ozturk, C. A. and Nasuf, E. (2002). Geostatistical assessment of rock zones for tunneling. *Tunneling and Underground Space Technology*, 17, 275-285.

Ozturk, C. A. and Simdi, E. (2014). Geostatistical investigation of geotechnical and constructional properties in Kadikoy–Kartal subway, Turkey. *Tunnelling and Underground Space Technology*, 4, 35–45.

Pareto, V. (1906). *Manuale di Economia Politica*. Società Editrice Libreria, Milan.

Phoon, K. K. and Kulhawy, F. H. (1999). Characterization of geotechnical variability. *Canadian Geotechnical Journal*, 1999, 36(4), 612-624. doi:10.1139/t99-038

Pinheiro, M., Emery, X., Miranda, T. and Vallejos, J. (2016a). Truncated Gaussian Simulation to Map the Spatial Heterogeneity of Rock Mass Rating. *Rock Mechanics and Rock Engineering*, 49(8), 1-6. doi: 10.1007/s00603-016-0928-x

Pinheiro, M., Vallejos, J., Miranda, T. and Emery, X. (2016b). Geostatistical simulation to map the spatial heterogeneity of geomechanical parameters: A case study with rock mass rating. *Engineering Geology*, 205, 93-103. doi: 10.1016/j.enggeo.2016.03.003

Popescu, R., Deodatis, G. and Nobahar, A. (2005). Effects of random heterogeneity of soil properties on bearing capacity. *Probabilistic Engineering Mechanics*, 20, 324-341. doi: 10.1016/j.probengmech.2005.06.003

Rahim, S. (2015). Reservoir Geological Uncertainty Reduction and Its Applications in Reservoir Development Optimization. M.Sc. thesis, University of Alberta, Canada, 143 p.

Read, S. A. L., Richards, L. R. and Perrin, N. D. (1999). Applicability of the Hoek–Brown failure criterion to New Zealand greywacke rocks. In: *Proceedings of the Ninth International Congress on Rock Mechanics*, 2, 655–660. Paris, France.

Rosenbaum, M.S., Rosen, L. and Gustafson, G. (1997). Probabilistic models for estimating lithology. *Engineering and Geology*, 47, 43–55.

Rosenblueth, E. (1975). Point estimates for probability moments. *Proc Natl Acad Sci U S A.*, 72(10), 3812-3814.

Rouhani, S., Srivastava, R. M, Desbarats, A. J., Cromer, M. V. and Johnson, A. I. (1996). Description: This Special Technical Publication of the ASTM contains an excellent collection of papers that provide an overview of geostatistics and describe applications to environmental and

geotechnical engineering. *Geostatistics for Environmental and Geotechnical Applications*, ASTM STP, 1283.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65. doi: 10.1016/0377-0427(87)90125-7

Ryu, D. W., Kim, T. K. and Heo, J. S. (2003). A study on geostatistical simulation technique for the uncertainty modelling of RMR. *Tunnel and Underground*, 13, 87–99.

Scheck, D. and Chou, D. R. (1983). Optimum locations for exploratory drill holes. *International Journal of Mining Engineering*, 1(4), 343–355.

Scheidt, C. and Caers, J. (2009a). Uncertainty quantification in reservoir performance using distances and Kernel methods- application to a West Africa deep-water turbidite reservoir. *SPE Journal*, 14/4, 680-692.

Scheidt, C. and Caers, J. (2009b). Representing spatial uncertainty using distances and kernels. *Mathematical and Geosciences*, 41, 397–419.

Serafim, J. L. and Pereira, J. P. (1983). Considerations of the Geomechanics Classification of Bieniawski, *Proc. Int. Sym Eng. Geol. Under. Const.*, 1, II.33-II.42.

Shi, J. and Malik, J. (2000). Normalized cuts and image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 22 (8), 888-905. doi: 10.1109/34.868688

Singh, A. P., Maucec, M., Carvajal, G. A., Mirzadeh, S., Knabe, S. P., Al-Jasmi, A. K. and El Din, I. H. (2014). Uncertainty quantification of forecasted oil recovery using dynamic model ranking with application to a ME carbonate reservoir. In: *International Petroleum Technology Conference*, 20-22 January 2014, Doha, Qatar. doi: IPTC 17476-MS

Singh, B., Viladkar, M., Samadhiya, N. and Mehrotra, V. (1997). Rock mass strength parameters mobilized in tunnels. *Tunnelling and Underground Space Technology*, 12, 47–54.

Soltani, S. and Hezarkhani, A. (2009). Additional exploratory boreholes optimization based on three-dimensional model of ore deposit. *Archives of Mining Science*, 54(3), 495–506.

Soltani, S. and Hezarkhani, A. (2013a). A simulated annealing-based algorithm to locate additional drillholes for maximizing the realistic value of information. *Natural Resources Research*, 22(3).

- Soltani, S. and Hezarkhani, A. (2013b). Proposed algorithm for optimization of directional additional exploratory drill holes and computer coding. *Arabian Journal of Geosciences*, 6(2), 455–462.
- Sonmez, H., Gokceoglu, C. and Ulusay, R. (2004). Indirect determination of the modulus of deformation of rock masses based on the GSI system. *International Journal of Rock Mechanics*, 1, 849–857.
- Stavropoulou, M., Exadaktylos, G. and Saratsis, G. (2007). A combined three-dimensional geological-geostatistical numerical model of underground excavations in rock. *Rock Mechanics and Rock Engineering*, 40 (3), 213-243.
- Struyf, A., Hubert, M., Rousseeuw, P. J. (1997). Clustering in an Object-Oriented Environment. *Journal of Statistical Software*, 1(4). doi: 10.18637/jss.v001.i04
- Thacker, B. H. and Painter, S. L. (2000). Modeling spatial heterogeneity using multi-Scale Random fields. *Proceedings of the 8th ASCE Joint Specialty Conference on Probabilistic Mechanics and Structural Reliability*, July 24-26, South Bend, Indiana.
- Tibshirani, R., Walther, G. and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society Series B*, 63 (2), 411–423.
- Ulam, S., Richtmyer, R. D. and Von Neumann, J. (1947). Statistical methods in neutron diffusion. Los Alamos Scientific Laboratory report LAMS–551.
- Van Groenigen, J. W., Siderius, W. and Stein, A. (1999). Constrained optimisation of soil sampling for minimisation of the kriging variance. *Geoderma*, 87, 239-259.
- Verman, M. (1993). *Rock mass – tunnel support interaction analysis*, Ph.D. Thesis, University of Roorkee, Roorkee, India, p. 267.
- Wackernagle, H. (2006). *Multivariate geostatistics: An introduction with applications*. 3rd edition, Springer, New York, 388 p. doi:10.1007/978-3-662-05294-5
- Yang, X. S. (2010). *Engineering Optimization: An Introduction with Metaheuristic Applications*. John Wiley & Sons, NJ: University of Cambridge, 347 p.
- You, K. H. (2003). An estimation technique of rock mass classes for a tunnel design. *Geotechnical Engineering*, 19, 319–326 (in Korean with English abstract).

Zadeh, L. A. (1963). Optimality and non-scalar-valued performance criteria. *IEEE Trans Automat Contr AC*, 8, 59–60.

Zhang, Y. (2011). Introduction to Geostatistic – Course notes. Department of Geology and Geophysics, University of Wyoming, USA.

Zhigljavsky, A. A. (1991). Theory of Global Random Search, Kluwer Academic Publishers. *Mathematics and Its Applications (Soviet Series)*, 65. ISBN 978-94-011-3436-1

Appendix 1

APPENDIX TO CHAPTER 5

A1.1.	Boreholes main characteristics	304
A1.2.	BASIC statistics from phase 1	306
A1.3.	Information from phase 2 – recovery method	307
A1.4.	Cell size study to use in declustering.....	309
A1.5.	Geostatistical simulation of data set 1 and 2.....	310
A1.6.	empirical formulas used in the geotechnical parameters.....	311
A1.7.	Deterministic approaches	312

A1.1. BOREHOLES MAIN CHARACTERISTICS

Table A1.1 Summary of the boreholes main characteristics (EDP, 2009).

Borehole code	Initial coordinate (m)	Azimuth (°)	Angle with the horizontal (°)	Length (m)	Water level depth (m)
S 1	294.7	256	60	60.0	10.0
S 2	297.8	099	60	160.0	16.0
S 3	NOT PERFORMED				
S 4	152.0	-	90	30.0	9.0
S 5	311.4	256	45	160.0	13.4
S 6	273.9	136	58	65.0	1.0
S 7	298.3	-	90	190.0	22.5
S 8	310.53	-	90	190.0	13.4
S 9	308.6	250	51	130.0	23.8
S 10	356.1	116	50	180.0	17.5
S 11	371.8	040	50	80.0	0.5
S 12	281.2	104	50	185.0	31.0
S 13	146.4	034	70	48.3	0.0
S 14	298.3	231	80	200.0	3.0
S 15	297.7	180	55	180.0	0.8
S 16	272.1	107	60	60.0	6.0

Table A1.2 Summary of the uniaxial compressive strength tests performed on the borehole

Borehole code	Length (m)	Weathering degree W	Length of the used sample (mm)	Diameter \emptyset (mm)	σ_c (MPa)	E_i (GPa)	ν	α^1 (°)
S2	144.7 – 145.0	W1-W2	125.7	50.7	104.8	38.0	0.18	60
	149.6 – 150.0	W1-W2	125.7	50.7	120.6	51.3	0.16	60
	154.4 – 154.7	W1-W2	125.3	50.7	81.8	55.6	0.16	60
	158.0 – 158.2	W1-W2	125.6	50.8	71.7	55.1	0.13	60
	159.6 – 160.0	W1-W2	125.2	50.7	117.8	49.2	0.19	60
S5	149.3 – 149.7	W3	125.5	50.6	47.9	-	-	45
	154.2 – 154.6	W3	125.6	50.6	69.7	-	-	45
	158.4 – 158.7	W2	125.4	50.6	92.9	-	-	45
S6	23.9 – 24.3	W3	125.0	50.6	21.8	-	-	58
	55.0 – 55.3	W1-W2	125.8	50.4	102.2	-	-	°8
S7	107.6 – 108.0	W1-W2	125.8	50.5	129.4	44.3	0.18	90
	121.6 – 122.1	W1	125.7	50.4	139.3	57.6	0.25	90
	148.5 – 148.9	W2	125.7	50.6	113.7	51.7	0.24	90
	165.6 – 166.0	W1	125.7	50.7	119.7	47.7	0.23	90
	187.1 – 187.5	W1	126.0	50.6	100.7	56.6	0.17	90
S11	17.8 – 18.0	W3	125.6	50.3	49.3	-	-	50
	58.0 – 58.3	W3	125.8	50.0	84.7	-	-	50
	69.6 – 70.0	W3	125.6	50.5	92.6	-	-	50
S12	145.1 – 145.6	W1-W2	125.5	50.8	114.2	71.3	0.23	50
	161.9 – 162.1	W1-W2	125.8	50.8	98.0	62.9	0.29	50
S14	177.8 – 178.1	W1-W2	125.7	50.7	114.0	69.0	0.23	50
	42.1 – 42.4	W2-W3	126.3	50.6	89.1	34.1	0.15	80
	62.2 – 62.7	W3	126.3	50.4	86.1	50.7	0.16	80
	66.2 – 66.7	W2-W3	126.2	50.6	97.5	45.8	0.18	80
	120.8 – 121.5	W3	124.4	50.7	99.0	47.6	0.19	80
S16	195.0 – 195.6	W3	126.5	50.7	84.5	41.9	0.14	80
	7.6 – 7.9	W3	125.2	50.7	43.4	-	-	60
	11.5 – 11.9	W3	124.3	50.7	37.6	-	-	60
	17.5 – 17.8	W3	125.1	50.8	61.0	-	-	60

¹ Angle between the sample longitudinal axis and the horizontal axis.

A1.2. BASIC STATISTICS FROM PHASE 1

Table A1.3 Statistical analysis of the geotechnical information that composes data set 1 of phase 1 (RQD, F and W).

Variable	Number of samples	Minimum	Maximum	Mean	Variance
RQD (%)	1479.00	0.00	100.00	64.00	964.40
W	1479.00	1.00	5.00	2.60 ²	0.70
F	1479.00	1.00	5.00	2.90 ²	1.50

Table A1.4 Statistical analysis of the geotechnical information that composes data set 2 of phase 1 (E_i , W, σ_c and ν).

Variable	Number of samples	Minimum	Maximum	Mean	Variance
E_i (GPa)	81	18.10	62.10	56.40 ³	101.10
W	81	1.00	3.50	1.90 ²	0.30
σ_c (MPa)	28	21.80	138.30	87.60	902.70
ν	79	0.08	0.42	0.21	0.003

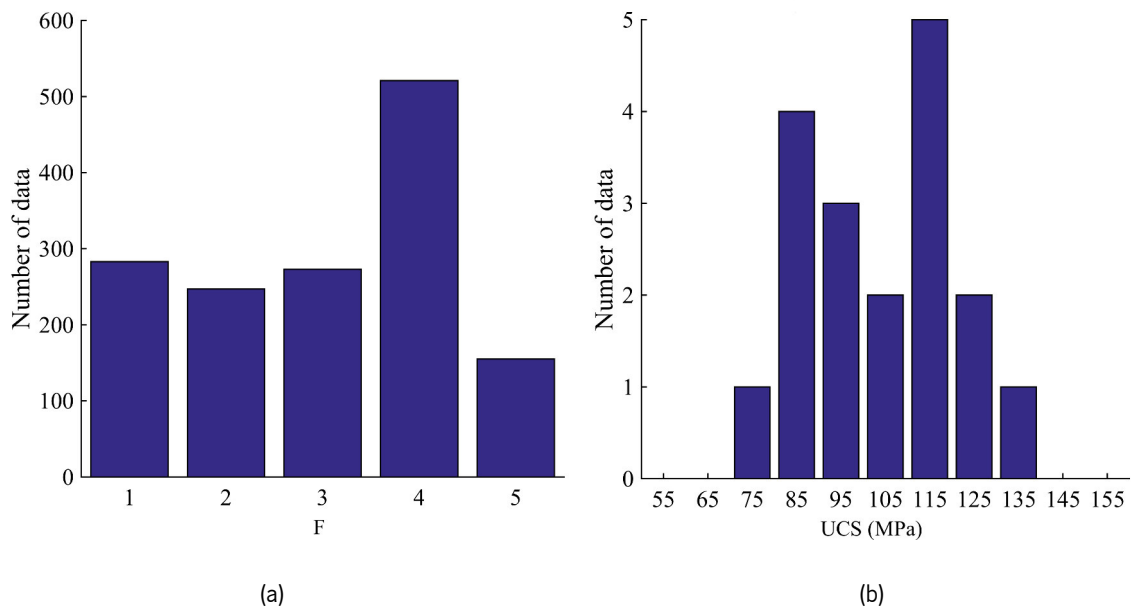


Figure A1.1 Histograms obtained using all the phase 1 information of Salomonde II, for: a) F; and b) σ_c .

² The mean values of the weathering and fracturing degrees were computed to give a notion of the averaging value, however, the most correct way to represent this type of variables is by computing their histograms.

³ For the deformation modulus of the rock intact, the computed mean was harmonic.

A1.3. INFORMATION FROM PHASE 2 – RECOVERY METHOD

Before going into details concerning the method adopted to recover the geotechnical mapping information, it is important to present all the zones that were considered in this process. Therefore, beside the powerhouse cavern, other underground works were taken into account (see Figure A1.2 and Table A1.5).

Table A1.5 List of all the structures considered in this phase 2 information recovery.

Powerhouse cavern (C)	Other structures
Right side wall of the cavern (C);	Access tunnel to the cavern (TAC);
Left side Wall of the cavern (C);	Attack tunnel to the cavern roof (TAAC);
North end of the cavern (C);	Access tunnel to the floodgate chamber (TACC);
South end of the cavern (C);	Tailrace tunnel – only in zones near the cavern (TR);
	Headrace tunnel – only in zones near the cavern (TAD).

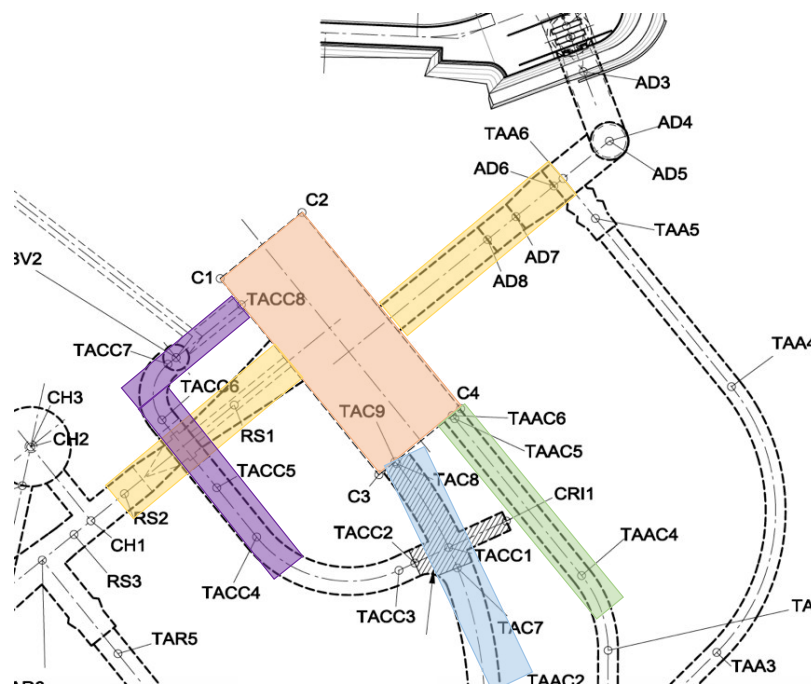


Figure A1.2 Representative scheme identifying all the zones from where the geotechnical mappings were recovered (adapted from EDP (2009)).

In Figure A1.2 the points represented by the letter *C* correspond to the four edge points of the cavern from which all the coordinates calculation was performed. The coordinates of these points can be found in Table A1.6. Thereafter, the mapping registered was made using the sequence presented in Table A1.7 and was named as step 1. Then, and because the numerical model has different coordinates, it was necessary to transform the *M*, *P* and *Z* coordinates into global coordinates like *X*, *Y* and *Z*. It is important to mention that a range of 20 m was adopted for

coordinate P in the power house cavern mappings (direction y of the cavern). However, and since the advance step used to build the mappings was higher than 20 m, some information had to be looped to maintain a regular grid of information. For the remaining structures, the adopted step for the register was the same as shown in the mapping documents (see Table A1.7).

Finally, it was possible to organise all the underground works geotechnical information in the layout presented in Table A1.8 for further analysis.

Table A1.6 Information regarding the coordinates of the powerhouse cavern edge points for an elevation equal to 126 m.

Point ID	M (m)	P (m)	Z (m)
C1	3591.47	224522.72	126.00
C2	3512.06	224539.38	126.00
C3	3531.46	224473.26	126.00
C4	3552.06	224489.92	126.00

Table A1.7 Details about the mapping registered for the left side wall of the cavern.

Elevation (m)	PK initial (m)	PK final (m)	Document page
138.50	0+65.85	0+38.50	11
	0+38.50	0+18.00	10
	0+18.00	0+05.00	10
	0+05.00	0+0.00	10
135.50	0+28.00	0+65.50	9
	0+0.00	0+28.00	9
	0+21.50	0+31.00	8
125.50	0+31.00	0+55.00	7
	0+0.00	0+21.50	7
...

Table A1.8 Final layout of the recovered information for the left side Wall of the cavern.

X (m)	Y (m)	Z (m)	RMR	RQD	Q	GSI
3512.08	224539.38	138.50	67.00	90.00	15.00	65.00
3512.08	224519.38	138.50	67.00	90.00	15.00	65.00
3512.08	224499.38	138.50	71.00	95.00	15.83	70.00
3512.08	224479.38	138.50	72.00	95.00	15.83	70.00
3512.08	224473.53	138.50	72.00	95.00	15.83	70.00

A1.4. CELL SIZE STUDY TO USE IN DECLUSTERING

According to Deutsch (1989) the size of the cell used to perform the data declustering is an essential step to achieve good results in the variable uncertainty assessment and simulation. Declustering techniques organise the data based on their proximity and to avoid redundancy of the information, since close data get a reduced weight and the contrary for data located far apart (Deutsch, 2015).

Several types of declustering can be used (e.g. polygonal, cell, kriging, inverse distance weighting, etc.), however the most common and adjusted to all data types are the cell-based type, which in opposition with a polygonal type is more fitted for declustering in $3D$.

In a cell based type a grid of equal volume is placed over the data domain and the number of occupied cells is counted (n_{occ}), then to each occupied cell a same weight value is attributed to the data according to the number of samples inside the cell ($1/n_{occ}$). In order to find the best cell size a plot of the declustered mean should be constructed from where a minimum or a maximum mean value can be chosen as the optimal cell size. In this particular case, the first reached minimum mean value was set as the selection criteria, mainly so that the declustered mean is not very different from the original data mean. Therefore, from Figure A1.3 analysis it was possible to detect the minimum mean value registered for a cell size of 7 m x 7 m x 7 m for data set 1 and 8 m x 8 m x 8 m for data set 2.

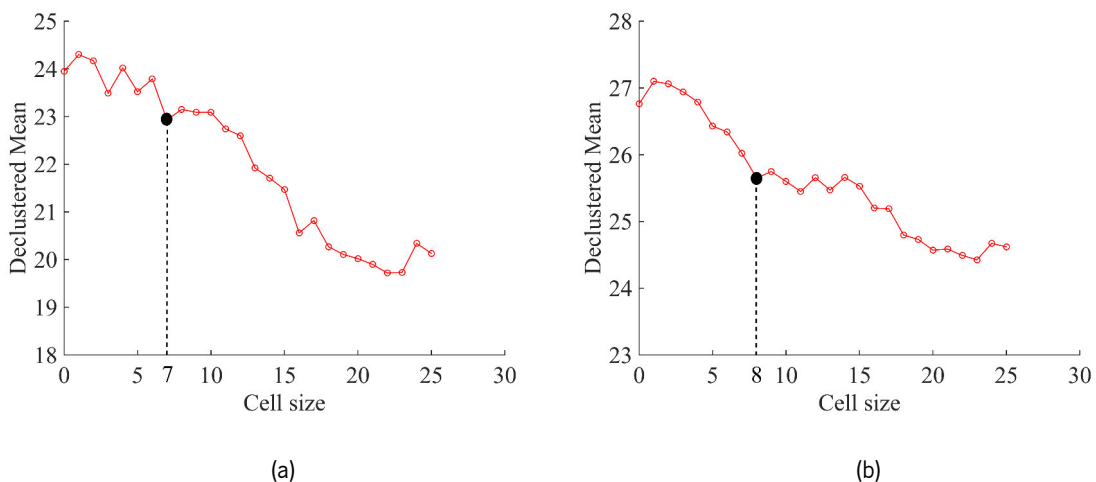


Figure A1.3 The cell declustering mean versus 25 cell sizes (each value represents a $3D$ square in m) for: a) data set 1; and b) data set 2.

A1.5. GEOSTATISTICAL SIMULATION OF DATA SET 1 AND 2

The anamorphosis is a function that relates the original E_m values with the associated Gaussian values and can be computed using the available data. To fit this function, a piecewise linear interpolation between the empirical points and exponential functions for tail extrapolation is used (Emery and Lantuéjoul, 2006). This adjustment should be as continuous as possible to give good simulation results of the variable. Figure A1.4 displays the anamorphosis function used for the both data sets simulation and as can be observed, both show a good continuity and shape.

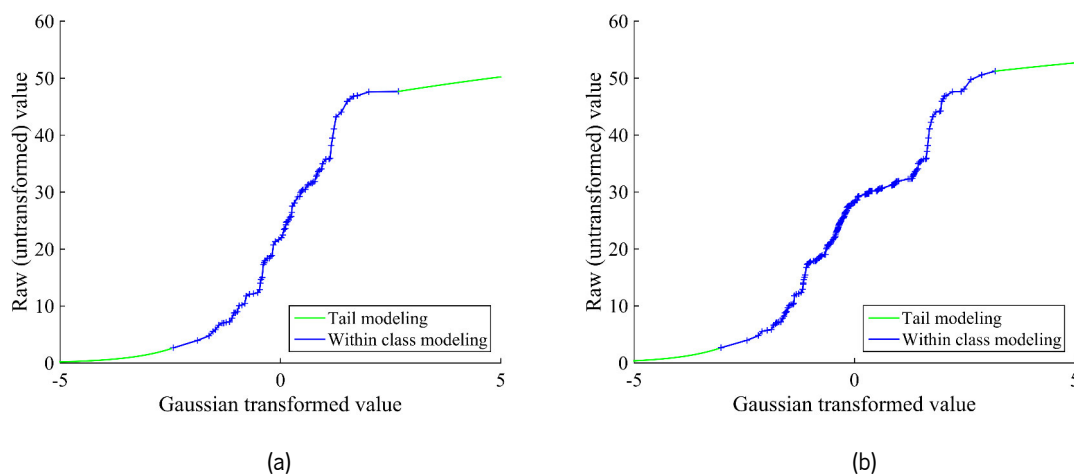


Figure A1.4 Anamorphosis function with the E_m values in GPa (ordinate) and the associated Gaussian value (abscissa) for: a) data set 1; and b) data set 2.

Table A1.9 Calculated proportions for W of data set 1 with the corresponding Gaussian thresholds.

Category	Cumulative proportion	Lower threshold	Upper threshold
1	0.253	$-\infty$	-0.665
2	0.584	-0.665	+0.225
3	0.874	+0.225	+1.150
4	0.960	+1.150	+1.980
5	1.000	+1.980	$+\infty$

Table A1.10 Calculated proportions for W of data set 2 with the corresponding Gaussian thresholds.

Category	Cumulative proportion	Lower threshold	Upper threshold
1	0.26	$-\infty$	-0.645
2	0.62	-0.645	+0.305
3	0.89	+0.305	+1.235
4	0.97	+1.235	+1.885
5	1.00	+1.885	$+\infty$

A1.6. EMPIRICAL FORMULAS USED IN THE GEOTECHNICAL PARAMETERS

Table A1.11 Empirical expressions used to obtain E_m and the corresponding authors for both phases of geotechnical information.

Data	Author(s)	Require parameters	Limitations	Equation (E_m in GPa)
Data set 1 and 2	Kayabasi et al. (2003)	RQD, E_i and W	-	$E_m = 0.135 \times \left(\frac{E_i \times (1 + RQD/100)}{W} \right)^{1.1811}$
Data set 3	Read et al. (1999)	RMR	-	$E_m = 0.1 \times \left(\frac{RMR}{10} \right)^3$
	Barton et al. (1983)	Q	Q>1	$E_m = 25 \times \log(Q)$
	Hoek and Diederichs (2006)	D, GSI	-	$E_m = 100000 \times \left[\left(1 - \frac{D}{2} \right) / \left(1 + e^{\frac{75+25D-GSI}{11}} \right) \right]$
	Kayabasi et al. (2003)	RQD, E_i and W	-	$E_m = 0.135 \times \left(\frac{E_i \times (1 + RQD/100)}{W} \right)^{1.1811}$
	Nicholson and Bieniawski (1990)	E_i , RMR	-	$E_m = (E_i/100) \times (0.0028RMR^2 + 0.9e^{\frac{RMR}{22.82}})$

In order to understand the main differences between each one of the formulas presented in Table A1.9, the basic statistics using data from phase 1 and phase 2 are computed. It is possible to observe that the formulas resulting in lower E_m value are Kayabasi et al. (2003) and Nicholson and Bieniawski (1990), while the one using the empirical systems show higher E_m values.

Table A1.12 E_m (in GPa) basic statistics after applying the empirical formulas selected for data set 3.

	Read et al. (1999)	Barton et al. (1983)	Hoek and Diederichs (2006)	Kayabasi et al. (2003)	Nicholson and Bieniawski (1990)
Mean	32.26	26.30	33.86	19.23	17.48
Variance	23.59	23.13	108.43	165.38	40.43
Standard deviation	4.86	4.81	10.41	12.86	6.35
Minimum	12.67	0.43	7.89	0.14	0.80
Maximum	42.05	32.58	48.93	58.16	33.38

A1.7. DETERMINISTIC APPROACHES

Contrarily to the heterogeneous approaches where the geotechnical parameters were geostatistical simulated to obtain the deformation modulus of the rock mass, for the deterministic approach the methodology is slightly different.

Since the goal of this Chapter is to prove the value and enrichment in using the geostatistical simulation in comparison with the deterministic and more traditional approach, the geotechnical data obtained in both phases is used, directly and without performing any type of geostatistical simulation. Although, the empirical formulas applied are the same as the one used for the heterogeneous data sets.

Therefore, and performing the same division as before, the information from both phases were treated separately.

The results of the deterministic model 1 and 2 are similar and, therefore, only one homogeneous model was used to compare both data sets.

Table A1.13 Statistical analysis of E_m (in GPa) obtained values for the deterministic approaches to use for comparison with data sets 1, 2 and 3.

	Deterministic 1	Deterministic 2	Deterministic 3
Total of points with information	81	102	307
Mean	21.68	21.87	29.00
Variance	172.97	174.95	107.85
Standard deviation	13.15	13.23	10.39
Minimum	2.00	2.00	2.00
Maximum	49.25	49.25	49.25

Appendix 2

APPENDIX TO CHAPTER 6

A2.1. Simulated Annealing internal parameters study	314
---	-----

A2.1. SIMULATED ANNEALING INTERNAL PARAMETERS STUDY

A2.1.1 Initial values

In Table A2.1 the input parameters values adopted for all the tests performed using as an objective function the 95% probability interval and the variance average, are presented.

Table A2.1 Input parameters for all the tests.

Attempt n.	Cooling factor (α)	Initial temperature (°)	Final temperature (°)	Maximum number of rejections	Maximum number of moves	Maximum number of acceptances
1	0.80	1.00	0.10	20	20	5
2	0.80	2.00	0.10	30	20	5
4	0.80	4.00	0.10	40	30	10

A2.1.2 Test 1

In test number one, SA algorithm performed a total of 34 evaluations of the objective function and the minimum value encountered was 11.69 for the 95% probability interval (see Figure A2.1a). Considering the variance as an objective function, the algorithm was executed 46 times resulting in a minimum value of 7.76 (see Figure A2.1b). The final temperature was lower than the predefined limit (0.09). To illustrate the behaviour of the algorithm in this first test, Figure A2.1 shows a colour surface obtained using the $v4$ type of interpolation (bi-harmonic spline interpolation) existing in software Matlab (Matlab, 2015a). As observed, in both objective functions case the algorithm performed the hill climbing movement where higher values for the objective function were accepted. Since the total number of evaluations was quite reduced, a second test with different values for the input parameters had to be carried out.

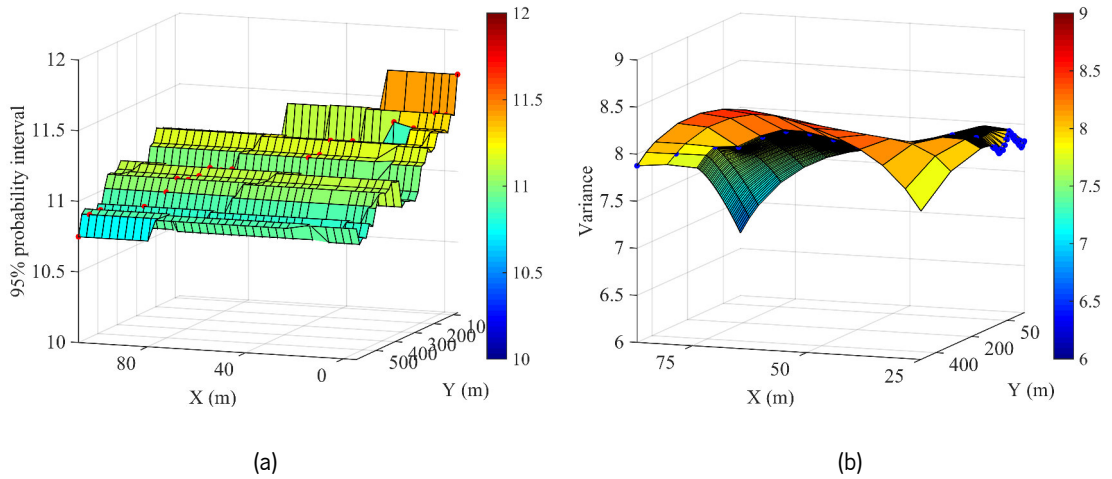


Figure A2.1 Spatial plot for test number 1 using the 'v4' type of interpolation to obtain the colour surface for; a) $\overline{PI}_{95\%}$; and b) Var .

A2.1.3 Test 2

For this second test the algorithm increased the number of evaluations with a total of 44 and 235 for the 95% probability interval and variance average, respectively. Regarding the objective functions values both registered a decrease to 10.48 and 6.70, as expected (see Figure A2.2). However, for the $\overline{PI}_{95\%}$ optimisation, the number of evaluations was insufficient to allow a good search of the space by the SA, leading to conclude that, similarly to test number 1, the input parameters value showed to be adjusted. Moreover, in this test the final temperature value was equal and inferior to the limit and the interpolation used for both objective functions was a v4 type.

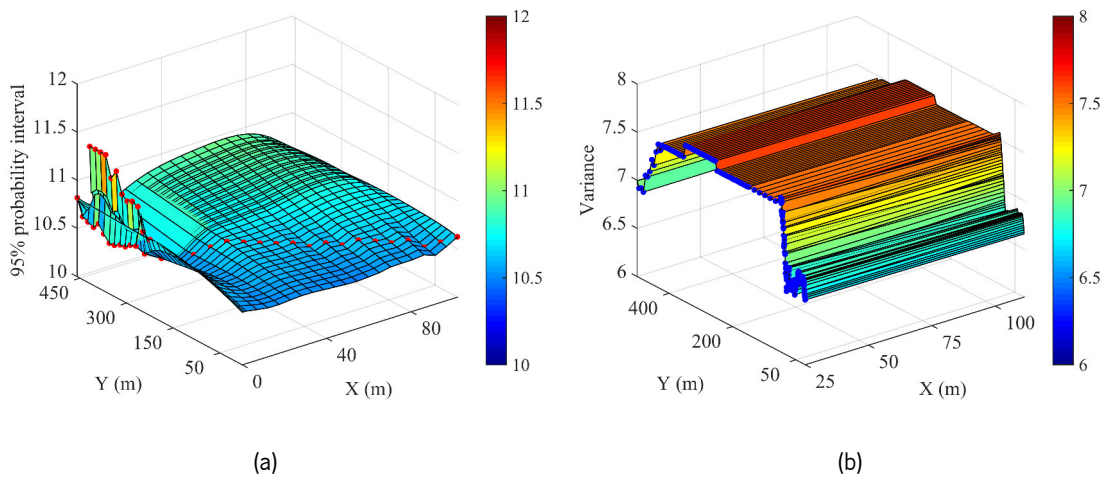


Figure A2.2 Spatial plot for test number 2 using the 'v4' type of interpolation to obtain the colour surface for; a) $\overline{PI}_{95\%}$; and b) Var .

A2.1.4 Test 3

In this final test, a lower value for the final temperature was adopted in order to expect a considerable increase in the number of the objective function evaluations. Although, in the case of the variance objective function the increase was not very meaningful, rising from 235 to 252. In contrast, the $\overline{PI}_{95\%}$ function had an increased to a total of 95 evaluations. Nonetheless, the number of evaluations is still lower, which allows to conclude that the input parameters that should be increased are the maximum number of rejections, maximum number of moves and maximum number of acceptances.

In what concerns the objective functions minimisation, a value of 10.46 was obtained for the $\overline{PI}_{95\%}$ function, while contrarily to the expected, the Var function shows a minimum of 10.00. Regarding the surface interpolation, a *nearest* type was chosen for the $\overline{PI}_{95\%}$ and a *v4* in the case of the variance function (see Figure A2.3).

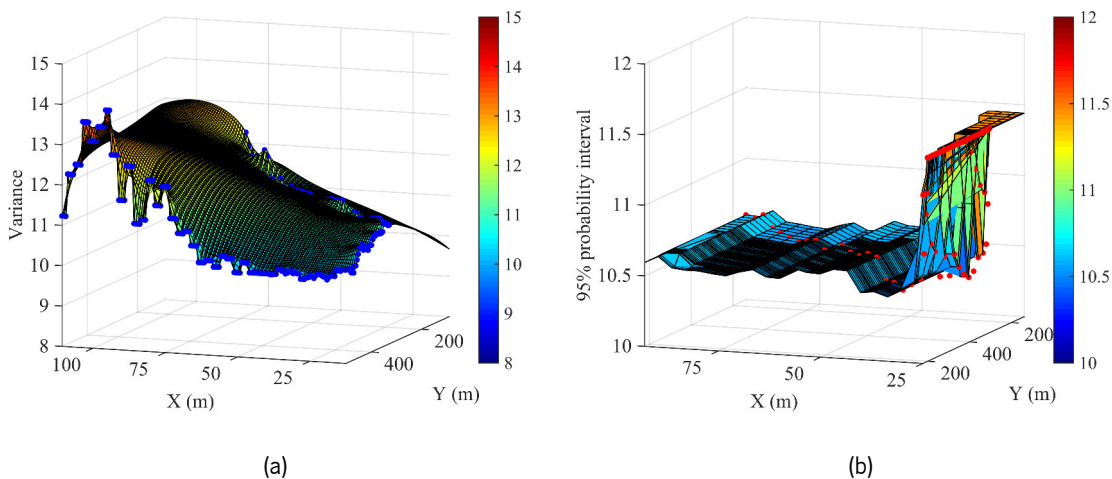


Figure A2.3 Spatial plot for test number 3 using the 'nearest' and 'v4' type of interpolation to obtain the colour surface, respectively for; a) $PI_{95\%}$; and b) Var .