COMPARATIVE DETERMINISTIC AND PROBABILISTIC MODELING IN GEOTECHNICS: APPLICATIONS TO STABILIZATION OF ORGANIC SOILS, DETERMINATION OF UNKNOWN FOUNDATIONS FOR BRIDGE SCOUR, AND ONE-DIMENSIONAL DIFFUSION PROCESSES

A Dissertation

by

NEGIN YOUSEFPOUR

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Zenon Medina-Cetina
Co-Chair of Committee,	Jean-Louis Briaud
Committee Members,	Stefan Hurlebaus
	Edgar Sanchez-Sinencio
	Francisco Gabriel Hernandez-Martinez
Head of Department,	John Niedzwecki

August 2013

Major Subject: Civil Engineering

ABSTRACT

This study presents different aspects on the use of deterministic methods including Artificial Neural Networks (ANNs), and linear and nonlinear regression, as well as probabilistic methods including Bayesian inference and Monte Carlo methods to develop reliable solutions for challenging problems in geotechnics. This study addresses the theoretical and computational advantages and limitations of these methods in application to: 1) prediction of the stiffness and strength of stabilized organic soils, 2) determination of unknown foundations for bridges vulnerable to scour, and 3) uncertainty quantification for one-dimensional diffusion processes.

ANNs were successfully implemented in this study to develop nonlinear models for the mechanical properties of stabilized organic soils. ANN models were able to learn from the training examples and then generalize the trend to make predictions for the stiffness and strength of stabilized organic soils. A stepwise parameter selection and a sensitivity analysis method were implemented to identify the most relevant factors for the prediction of the stiffness and strength. Also, the variations of the stiffness and strength with respect to each factor were investigated.

A deterministic and a probabilistic approach were proposed to evaluate the characteristics of unknown foundations of bridges subjected to scour. The proposed methods were successfully implemented and validated by collecting data for bridges in the Bryan District. ANN models were developed and trained using the database of bridges to predict the foundation type and embedment depth. The probabilistic Bayesian approach generated probability distributions for the foundation and soil characteristics and was able to capture the uncertainty in the predictions.

The parametric and numerical uncertainties in the one-dimensional diffusion process were evaluated under varying observation conditions. The inverse problem was solved using Bayesian inference formulated by both the analytical and numerical solutions of the ordinary differential equation of diffusion. The numerical uncertainty was evaluated by comparing the mean and standard deviation of the posterior realizations of the process corresponding to the analytical and numerical solutions of the forward problem. It was shown that higher correlation in the structure of the observations increased both parametric and numerical uncertainties, whereas increasing the number of data dramatically decreased the uncertainties in the diffusion process.

DEDICATION

To my father, *Ghorban*, who continuously followed up on this dissertation, even at his last moments: Baba, it's done!

And to my mother, *Nayer*, who has been a symbol of devotion, patience, and forgiveness in my life.

ACKNOWLEDGEMENTS

I would like to thank my advisors, Dr. Medina-Cetina and Dr. Briaud for their continuous guidance and support throughout the course of this research. Dr. Medina-Cetina's guidance and supports provided the opportunity for me to grow my knowledge and understanding in my chosen field. Also, Dr. Briaud's advices and supports were invaluable and encouraging throughout my PhD. I am also very grateful to my other committee members, Dr. Hernandez-Martinez, Dr. Hurlebaus, and Dr. Sanchez-Sinencio for their helps and guidance.

Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience. I also want to extend my gratitude to Kayvon Jahedkar and Anthony Garcia from Texas Department of Transportation at Bryan District, who accommodated the data collection for bridges and to Mark McClelland for his assistance with the bridge load calculation process. Also, the great help of Brittany Hanly in processing the database is highly appreciated.

At last but not least, I would like to extend my thanks to my mother and my father who have been my inspiration for learning and persistence in my life; to my sister who has been always caring and supportive; and to my husband who has been a great friend and a patient lover, since the beginning of this long journey till the end.

TABLE OF CONTENTS

Page

ABSTRACT…... ii

DEDICATION iv

ACKNOWLEDGEMENTS v

TABLE OF CONTENTS vi

LIST OF FIGURES ix

LIST OF TABLES xvi

CHAPTER I INTRODUCTION 1

CHAPTER II DETERMINISTIC METHODS 5

- Artificial Neural Networks (ANNs) 5
  - Radial Basis Function Networks 8
  - Multilayer Perceptron Networks 13
  - Generalized Regression Neural Networks (GRNNs) 20
  - Probabilistic Neural Networks (PNNs) 20
- Linear Regression (LR) and Nonlinear Regression (NLR) 20

CHAPTER III PROBABILISTIC METHODS 24

- Bayesian Inference 24
- Markov Chain Monte Carlo 26

CHAPTER IV PREDICTION OF THE STIFFNESS AND STRENGTH OF STABILIZED ORGANIC SOILS 28

- Introduction and Literature Review 28
- Approach 33
  - Experimental Database 33
  - Numerical Experiments 38
  - Models 43
- Results and Discussion 44
  - RBF Network Analysis 44
  - MLP Network Analysis 47
  - Stepwise Selection of Input Parameters for ANNs 52
  - Linear Regression (LR) Analysis 58

  Comparison of the Models 59
  Response Graphs and Sensitivity Analysis 60
Conclusions 68

CHAPTER V DETERMINATION OF UNKNOWN FOUNDATIONS FOR BRIDGE SCOUR: A DETERMINISTIC APPROACH 71

Introduction and Literature Review 71
Approach 76
  Database and Statistics 76
  Methodology 88
  Dead Load and Live Load Calculations 94
Deterministic Models 101
  Artificial Neural Networks 101
  Nonlinear Least Squares Method 110
Results and Discussion 110
  Foundation Type Classification 110
  Foundation Depth Prediction 114
Conclusions 128

CHAPTER VI DETERMINATION OF UNKNOWN FOUNDATIONS FOR BRIDGE SCOUR: A PROBABILISTIC APPROACH 130

Introduction 130
Approach 131
  Methodology 131
  Load and Bearing Capacity Calculations 136
  ANN Models for Bearing Capacity Prediction 138
  Bayesian Inference 140
Results and Discussion 145
  ANN Models for Bearing Capacity Prediction 145
  Validation of the Bayesian Probabilistic Method 148
Conclusions 170

CHAPTER VII UNCERTAINTY QUANTIFICATION FOR ONE-DIMENSIONAL DIFFUSION PROCESSES 172

Introduction and Literature Review 172
Problem Definition 177
  Data Synthesis 178
Methodology 186
Results and Discussion 188
  Bayesian Inversion, Unknown Parameters: $I, c, \Delta t$ 188
  Statistical Inference 211

Effect of Time Discretization on Numerical Uncertainty and Computational Cost 215
Full Bayesian Inversion 218
Conclusions 230

CHAPTER VIII CONCLUSIONS 232

Theoretical and Methodological Conclusions 232
Prediction of the Stiffness and Strength of Stabilized Organic Soils 234
Conclusions and Remarks 234
Recommendations for Future Research 234
Determination of Unknown Foundations for Bridge Scour 235
Conclusions and Remarks 235
Recommendations for Future Research 235
Uncertainty Quantification for One-Dimensional Diffusion Processes 236
Conclusions and Remarks 236
Recommendations for Future Research 236

REFERENCES 237

LIST OF FIGURES

Page

Figure 2-1. a) Biological Neuron, b) A Single Perceptron....6

Figure 2-2. Transfer Functions in ANNs....6

Figure 2-3. Architecture of the RBF Network....11

Figure 2-4. Architecture of the MLP Network....15

Figure 2-5. Architecture of the PNN Network....21

Figure 4-1. Histogram of the Unconfined Compression Tests....34

Figure 4-2. Histograms of the Control Variables in the Unconfined Compression Tests....36

Figure 4-3. Numerical Experiments: Group 1....39

Figure 4-4. Numerical Experiments: Group 2....40

Figure 4-5. Numerical Experiments: Group 3....42

Figure 4-6. RBF Network Architecture....43

Figure 4-7. MLP Network Architecture....44

Figure 4-8. Histograms of the R^2 Values, Corresponding to the 1000 Networks of the RBF-Tot Ensembles....47

Figure 4-9. Variation of the Average R^2 with the Number of Hidden Neurons, a) E , b) UCS....49

Figure 4-10. Histogram of the R^2 Values, Corresponding to the 1000 Networks in the MLP Ensemble....51

Figure 4-11. Predictions of E over a) Training, and b) Test Datasets, and Predictions of UCS for c) Training and d) Test Datasets Generated by a Network in the MLP Ensemble....52

Figure 4-12. Histogram of the R^2 Values Corresponding to the 1000 Networks of the LR Ensembles....58

Figure 4-13. Performances of Different Models over the Control Dataset....60

Figure 4-14. Response Graphs Depicting the Variation of *E* and *UCS* with respect to the Input Parameters....64

Figure 4-15. Response Graphs for Different Input Variables with respect to the Scale of Variation....68

Figure 5-1. A Typical Bridge Plan in the Bryan District, Showing the Original and Most Recent Ground Profile....77

Figure 5-2. Bridges in the Bryan District of TxDOT Collected for this Study, Showing the Scour-Critical Index, AADT, Year Built, and the Population Density over the Counties....79

Figure 5-3. Histograms for the Items and Parameters in the Database....82

Figure 5-4. A Typical Soil Boring Data Sheet....87

Figure 5-5. Skin Friction versus *TCP* for Left) *TCP* < *100* blows/ft, Right) *TCP* > *100* blows/ft (TxDOT 2006)....88

Figure 5-6. Deterministic Approach Flowchart....91

Figure 5-7. Total Load Calculation's Flowchart....92

Figure 5-8. Load versus Average Span Length for a Prestressed Concrete I-Beams Bridge with 24 (ft) Roadway Width....96

Figure 5-9. Standard AASHTO Design Loads and Maximum Reaction of an Interior Bridge Bent with Single Spans....99

Figure 5-10. Estimated Quantities for Bridge ID: 17-82-2144-01-002....100

Figure 5-11. Architecture of the MLP Classifier....102

Figure 5-12. Foundation Classification Using MLP-FT8, a) Predicted vs. Actual Scatter Plot, b) Histogram of Foundation Types, c) CA for each Foundation Type....112

Figure 5-13. Foundation Classification MLP-FT2....113

Figure 5-14. Pile Depth Predictions by the MLP Ensembles: a) MLPPL00, b) MLPPL01, c) MLPPL10-Conc, d) MLPPL11-Conc, e) MLPPL10-DrSh, f) MLPPL11-DrSh, g) MLPPL10-Steel, h) MLPPL11-Steel....116

Figure 5-15. Average Pile Depth Predictions by the MLP Ensembles: a) MLPPL00, b) MLPPL01, c) MLPPL10-Conc, d) MLPPL11-Conc,

e) MLPPL10-DrSh, f) MLPPL11-DrSh, g) MLPPL10-Steel, h) MLPPL11-Steel....118

Figure 5-16. 95% Prediction Interval and 10 Percentile of Pile Depth Predicted by a) MLPPL00, b) MLPPL01, c) MLPPL10-Conc, d) MLPPL11-Conc, e) MLPPL10-DrSh, f) MLPPL11-DrSh, g) MLPPL10-Steel, h) MLPPL11-Steel....120

Figure 5-17. Probability of Exceedance Curves for Pile Depth in the Normal and Log Scales a) MLPPL00, b) MLPPL01, c) MLPPL10-Conc, d) MLPPL11-Conc, e) MLPPL10-DrSh, f) MLPPL11-DrSh, g) MLPPL10-Steel, h) MLPPL11-Steel....122

Figure 5-18. Weights and Bias of the MLP Models: a) MLPPL00, b) MLPPL01, c) MLPPL10-conc, d) MLPPL11-conc, e) MLPPL10-DrSh, f) MLPPL11-DrSh, g) MLPPL10-steel, h) MLPPL11-steel....125

Figure 5-19. Comparison of NL and MLP Models....127

Figure 6-1. Diagram of the Two Main Steps Required for the Probabilistic Determination of Unknown Foundations....132

Figure 6-2. Flowchart of the Probabilistic Approach....133

Figure 6-3. (a) Point Bearing versus *TCP* for a) $TCP < 100$ blows/ft, and b) for $TCP > 100$ blows/ft (TxDOT 2006)....134

Figure 6-4. Histograms and Prior Distributions of a) *FT*, b) dp_{Dr}, c) f_u, and d) q_p....143

Figure 6-5. Empirical CDF of the Error Distributions for Left) MLP-BC10-DrSh, Right) MLP-BC11-DrSh....144

Figure 6-6. a) MLPBC00, b) MLPBC01, c) MLPBC10-Conc, d) MLPBC11-Conc, e) MLPBC10-DrSh, f) MLPBC11-DrSh, g) MLPBC10-Steel, h) MLPBC11-Steal, i) MLPBC10-Spread, j) MLPBC11-Spread (R^2 and *RMSE* Refer to the Average Predictions)....146

Figure 6-7. a) MCMC Samples for dp_{Dr}, b) Cumulative Mean, c) Cumulative Standard Deviation, and d) Autocorrelation Function of the Samples - Problem 11....149

Figure 6-8. CDF of Marginal Posterior Distribution of dp_{Dr} - Problem11....150

Figure 6-9. Prior and Marginal of the Posterior Distributions for dp_{Dr} - Problem*11*....151

Figure 6-10. Sample Chains for the Model Parameters - Problem10....152

Figure 6-11. CDF of the Marginal Posterior Distribution of dp_{Dr},f_u, and q_p - Problem10 153

Figure 6-12. Prior and Marginal Posterior Distributions for dp_{Dr},f_u, and q_p - Problem*10* 154

Figure 6-13. Prior (a) and Posterior (b) Joint Relative Frequency Density of dp_{Dr} vs. f_u - Problem *10* 155

Figure 6-14. Posterior Joint Relative Frequency Density of a) dp_{Dr} vs. q_p, *b)* f_u vs. q_p - Problem *10* 155

Figure 6-15. Sample Chains for the Model Parameters- Problem *01* 157

Figure 6-16. Prior (a) and Posterior (b) Distributions for *FT*- Problem01 157

Figure 6-17. Empirical CDF of the Marginal Posterior Distributions of the Foundation’s Dimensions for Conc, DrSh, Steal, and Spread – Problem*01* 158

Figure 6-18. Prior (a) and Posterior (b) Joint Relative Frequency Density of dp_{Dr} and f_u - Problem *01* 160

Figure 6-19. Prior (a) and Posterior (b) Joint Relative Frequency Density of dp_{Dr} vs. D_{Dr} - Problem *01* 160

Figure 6-20. Posterior Joint Relative Frequency Density of Foundation Dimensions for a) Conc, b) Steal, and c, d, e) Spread - Problem *01* 161

Figure 6-21. Sample Chains for the Model Parameters - Problem *00* 163

Figure 6-22. Empirical CDF of the Marginal Posterior Distributions of the Foundation Dimensions for Conc, DrSh, Steal, and Spread, and the Soil Parameters - Problem*00* 164

Figure 6-23. Prior and Marginal Posterior Distributions for a) dp_{Dr}, b) D_{Dr}, and c) *FT* - Problem*00* 166

Figure 6-24. Prior (a) and Posterior (b) Joint Relative Frequency Density of dp_{Dr} vs. D_{Dr} - Problem*00* 167

Figure 6-25. Posterior Joint Relative Frequency Density of a) Foundation Depth vs. Skin Friction, b) Foundation Depth vs. Point Bearing Capacity, c) Skin Friction vs. Point Bearing Capacity, for DrSh, and Foundation Depth vs. Foundation Size for d) Conc, e) Steal - Problem *00* 168

Figure 7-1. Flowchart of the Approach 177

Figure 7-2. Correlation Functions for Different Values of the Correlation Length 180

Figure 7-3. Synthetic Data and the Benchmark for Uncorrelated Observations, a) SO, b) MO 181

Figure 7-4. KL Synthesized Data for $\tau = 1$ 182

Figure 7-5. Eigenvalues of the Covariance Matrix 183

Figure 7-6. Eigenvectors of the Covariance Matrix 183

Figure 7-7. Correlation Function Representation 184

Figure 7-8. Empirical Covariance Function Obtained from the 10 KL Realizations 185

Figure 7-9. Empirical Correlation Function Obtained from the 10 KL Realizations 185

Figure 7-10. Design of the Numerical Experiments 187

Figure 7-11. Prior Probability Distribution of *Δt* 188

Figure 7-12. Cumulative Mean of the Sample Chains, a) Analytical and Numerical Model with *Δt=0.001*, and b) Numerical Model with *Δt=hyper-parameter* – SO 190

Figure 7-13. Posterior Joint PDF of *I* and *c* for, a) Analytical Model, b) Numerical Model with *Δt=0.001*, and c) Numerical Model with *Δt=hyper-parameter* – SO 191

Figure 7-14. Posterior CDFs of *I* and *c* – SO 192

Figure 7-15. Posterior Model Predictions, a) Analytical, b) Numerical Model with *Δt=0.001*, and c) Numerical Model with *Δt=hyper-parameter* - SO 193

Figure 7-16. Mean and Standard Deviation of the Posterior Model Predictions (Realizations) vs. Time – SO 194

Figure 7-17. Posterior Joint PDF of *Δt* with *I* and *c* - SO 195

Figure 7-18. Posterior CDF of *Δt* - SO 195

Figure 7-19. Cumulative Mean of the Sample Chains, a) Analytical and Numerical Model with *Δt=0.001*, and b) Numerical Model with *Δt=hyper-parameter* – MO 197

Figure 7-20. Posterior Joint PDF of *I* and *c*, a) Analytical, b) Numerical Model with $\Delta t=0.001$, and c) Numerical Model with *Δt=hyper-parameter* – MO 198

Figure 7-21. Posterior CDFs of *I* and *c* – MO 199

Figure 7-22. Posterior Model Predictions, a) Analytical, b) Numerical Model with $\Delta t=0.001$, and c) Numerical Model with *Δt=hyper-parameter* - MO 200

Figure 7-23. Mean and Standard Deviation of the Posterior Model Predictions vs. Time - MO 201

Figure 7-24. Posterior Joint PDF of Δt with *I* and *c* – MO 202

Figure 7-25. Posterior CDF of Δt - MO 202

Figure 7-26. Cumulative Mean of the Model Parameters, a) Analytical and Numerical Model with Δt=0.001, and b) Numerical Model with Δt=hyper-parameter -CO 203

Figure 7-27. Posterior Joint PDF of *I* and *c*, a) Analytical, b) Numerical Model with Δt=0.001, and c) Numerical Model with Δt=hyper-parameter – CO......204

Figure 7-28. Posterior CDFs of *I* and *c* - CO 205

Figure 7-29. Posterior Model Predictions, a) Analytical, b) Numerical Model with Δt=0.001, and c) Numerical Model with *Δt=hyper-parameter* – CO 207

Figure 7-30. Mean and Standard Deviation of Posterior Model Predictions vs. Time - CO 208

Figure 7-31. Posterior Joint PDF of Δt with *I* and *c* – CO 209

Figure 7-32. Comparing the Uncertainty Resulted from the Three Scenarios Using the Same Scales, a) SO, b) MO, c) CO......210

Figure 7-33. Posterior CDF of Δt - CO 211

Figure 7-34. Bias and Uncertainty of Model Parameters and Predictions for Different Observations Conditions, a) SO, b) MO, c) CO 213

Figure 7-35. Numerical Bias vs. Time Step for Different Observations Conditions, a) SO, b) MO, c) CO......216

Figure 7-36. Numerical Uncertainty vs. Time Step for Different Observations Conditions, a) SO, b) MO, c) CO 217

Figure 7-37. Computational Cost vs. Time Step Size for Different Observation Conditions 218

Figure 7-38. Cumulative Mean of the Samples for the Full Bayesian Inversion - SO (known σ) 219

Figure 7-39. Posterior CDFs of the Parameters for the Full Bayesian Inversion - SO 220

Figure 7-40. Joint Probability Distributions of the Parameters for the Full Bayesian Inversion - SO 221

Figure 7-41. Cumulative Mean of the Samples for the Full Bayesian Inversion - MO 222

Figure 7-42. Posterior CDFs of the Parameters for the Full Bayesian Inversion - MO 223

Figure 7-43. Joint Probability Distributions of the Parameters for the Full Bayesian Inversion - MO 224

Figure 7-44. Cumulative Mean of the Samples for the Full Bayesian Inversion - CO 226

Figure 7-45. Posterior CDF of the Parameters for Full Bayesian Inversion - CO 227

Figure 7-46. Joint Probability Distributions of the Parameters in the Full Bayesian Inversion – CO 228

LIST OF TABLES

Page

Table 4-1. Characteristics of the Organic Soils, Binders, and Mixtures 33

Table 4-2. List of Control and Response Variables and Their Levels/Range of Variation in the Database 35

Table 4-3. Characteristics of the Numerical Experiments 38

Table 4-4. Performances of the RBF Models Trained 45

Table 4-5. Performance of the RBF Model Trained Using the Whole Database 45

Table 4-6. Performances of the MLP Ensembles for *E* and *UCS* Prediction 50

Table 4-7. Stepwise Selection of Parameters for the RBF Network 55

Table 4-8. Stepwise Selection of Parameters for the MLP Network. 56

Table 4-9. Cross Correlation between the Input and Target Parameters 57

Table 4-10. Performance of the LR Ensemble for *E* and *UCS* Predictions 59

Table 4-11. Ranking of Parameters based on the Sensitivity Analysis 67

Table 5-1. Number of Unknown Foundation Bridges across the United States 72

Table 5-2. Percentage of Different Types of Substructures 78

Table 5-3. List of the Main Parameters in the Database Used in the Models 81

Table 5-4. Foundation Loads for 24 (ft) PSC I-Beam Bridge with a 15 (deg) Skew 96

Table 5-5. Load Estimation Coefficients for a Column 97

Table 5-6. Load Estimation Error for Five Sample Bridges 100

Table 5-7. Classification Results for PNN and MLP 103

Table 5-8. Comparison of Different ANN Models 105

Table 5-9. Input Parameters for Foundation Classification 107

Table 5-10. Input Parameters for Pile-Depth Prediction 108

Table 5-11. Ranking of the Parameters through FSS and BSE Methods........................111

Table 5-12. Results of Foundation Type Classification..113

Table 5-13. Results of MLP Ensembles for Pile Embedment Depth Prediction 115

Table 5-14. Results of the Best NL Model..127

Table 6-1. Input Parameters of the MLP Models for BC Prediction 139

Table 6-2. Problem Types and Unknown Parameters...141

Table 6-3. Definition of the Model Parameters and Their Corresponding Prior Distributions 142

Table 6-4. Performances of the MLP-BC Models for Prediction of Bearing Capacity 147

Table 6-5. Actual Values of Parameters for Bridge 17-166-0209-05-075 - Pier 2........148

Table 6-6. Statistics of the Posterior Distribution of the Model Parameter - Problem11...151

Table 6-7. Statistics of the Posterior Distribution of Model Parameters - Problem*10*...156

Table 6-8. Statistics of the Posterior Distributions of the Model Parameters - Problem*01*...162

Table 6-9. Statistics of Posterior Distributions of the Model Parameters - Problem*00*...169

Table 7-1. Definition of the Observational Scenarios...180

Table 7-2. Statistics of the Posterior Distributions of the Model Parameters................212

Table 7-3. Statistics of the Posterior Distributions of the Model Predictions................212

CHAPTER I

INTRODUCTION

Predicting the outcome of some measurements of a physical system is called the *modelization problem* or the *forward problem* (Tarantola 2005). Physical theories allow us to make such predictions, given the parameters that characterize the system. In contrast, the *inverse problem* uses the actual measurements to infer the values of the parameters characterizing the forward model. This can be mathematically represented as (Kaipio and Somersalo 2005; Tarantola 2005):

$$\mathbf{m} \mapsto \mathbf{d} = \mathrm{g}(\mathbf{m}) \tag{1-1}$$

where **d** is the vector of observations of the system, **m** is the vector of the model parameters, and *g* is the forward operator (usually a nonlinear operator). Inverse problems normally arise when direct measurement of the unknown quantity is impossible, and the measurements of other related quantities provide indirect access to infer the quantity of interest.

Inverse problems are often ill-posed, meaning that one of the conditions of existence, uniqueness, and stability are often violated in the solution (Marzouk et al. 2007). There are often several solutions for an inverse problem (the problem is over-determined), as there are several models that can fit through the noisy data. Thus, there is a significant level of uncertainty in the solution of inverse problems that initiates from the mathematical model, parameter estimations, measurements, and numerical approximations.

Deterministic methods often solve inverse problems based on the least-squares optimization. For instance, least-squares methods provide point estimates of the solution by minimizing the objective function which is the sum of squared errors (residuals). In order to address the ill-posed nature of the inverse problems, some deterministic methods add Tikhonov regularization terms to the objective function to treat strong ill-posed conditions (Tikhonov and Arsenin 1977).

On the other hand, probabilistic methods consider the model parameters as random variables and generate probability distributions for these parameters. Bayesian inference provides the possibility of taking into account current knowledge about a quantity of interest (prior probability distributions), and then update it based on new observations. Thus, Bayesian inference is a methodology to estimate the model parameters and to quantify the uncertainty associated with the estimations.

This study presents contrasting aspects on the use of deterministic methods including Artificial Neural Networks (ANNs), and linear and nonlinear regression models, as well as probabilistic methods including Bayesian inference and Monte Carlo to develop reliable solutions for inverse problems in geotechnics. This study addresses the theoretical and computational advantages and limitations of these methods in application to: 1) prediction of the stiffness and strength of stabilized organic soils, 2) determination of unknown foundations for bridge scour, and 3) uncertainty quantification for one-dimensional diffusion processes.

ANNs were implemented in this study to create nonlinear models to predict the unknown quantity of interest based on a number of observations of that quantity. ANNs were trained to learn from the available data, and during this process, their weights were adjusted to minimize the error function. Chapter II discusses the deterministic methods implemented in this study including ANNs, and linear and nonlinear regressions.

Bayesian inference was applied to quantify the uncertainty of the unknown parameters, given *a priori* information and new experimental observations. The uncertainty was then propagated from the model parameters to the model predictions in order to evaluate the uncertainty of the process. Chapter III describes the probabilistic methods implemented in this study.

Chapter IV presents the results of research on the stiffness and strength of stabilized organic soils. Organic soils are mainly soft and highly compressible; therefore, their strength properties are often improved by stabilization techniques including deep soil mixing. In contrast to mineral soils, only limited studies have investigated the mechanical behavior of stabilized organic soils. This study investigates the stiffness and strength of soils with different organic contents stabilized with various binder types and mixing proportions using ANNs. The predictive ANN models were developed using a database of laboratory experiments that were conducted at the University of Cambridge by Hernandez-Martinez (2006). Also, the effects of organic content, binder type and quantity, mixing proportions, aging, time, temperature, and other testing conditions on the stiffness and strength of stabilized organic soils were investigated through sensitivity analyses.

Chapter V and VI discuss the results of research on the determination of unknown foundations of bridges for scour evaluation. As of 2005, approximately 60,000 bridges throughout the United States were identified as having unknown foundations. Departments of Transportation (DOTs) have implemented different measures to evaluate the scour failure risk of over-river bridges with unknown foundations to monitor their safety on a regular basis. However, the problem of unknown foundations has not yet been properly addressed and there is an essential need for robust and more cost-effective solutions. This study introduces a deterministic and a probabilistic methodology for predicting the characteristics of unknown foundations of bridges, particularly their embedment depths, for scour evaluation. An evidence-based approach incorporating ANNs and the Bayesian inference method was implemented to generate single predictions for the type and embedment depth of foundations, and to evaluate the uncertainty of the predictions through probability distributions.

Chapter VII presents the results of research on evaluating the parametric and numerical uncertainties in one-dimensional diffusion processes under varying observation conditions. The ordinary differential equation (ODE) of diffusion has been widely applied to model various processes including the transport of contaminants in saturated soils and the movement of water (moisture) in unsaturated soil. There are uncertainties in the solution of the inverse diffusion problem, which initiate from various sources including the measurements and observations, model calibration, and numerical approximations. This study uses probabilistic methods including the Bayesian inference and Monte-Carlo to quantify the uncertainties that emerge from the model parameters and the numerical approximation of the one-dimensional diffusion problem. Finally, Chapter VIII presents the overall conclusions derived from the performed studies.

CHAPTER II

DETERMINISTIC METHODS

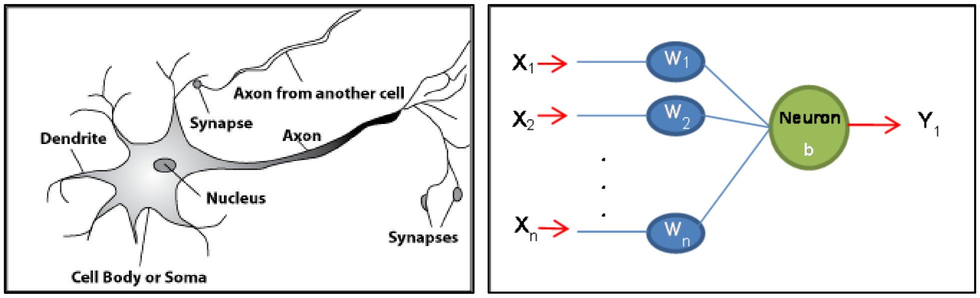
Artificial Neural Networks (ANNs)

Artificial Neural Networks (ANNs), first introduced by (McCulloch and Pitts 1943), are computational models consisting of parallel inter-connected processing units called neurons. ANNs have emerged from the idea of simulating the brain function with a computational system in terms of pattern recognition, learning, and generalization. Neurons in ANNs resemble the biological neurons, and connections between them resemble axons and dendrites through which biological neurons transmit and receive signals. Each connection between two neurons holds a value called weight, which represents the synapse at the intersection of the axon and dendrites between two biological neurons. The process of learning in ANNs, similar to human brain, is defined as adjustment of connections' weights (synapses in biological cells) by signals passing through them (Bishop 1995; Haykin 1999).

Figure 2-1 compares a biological neuron with a simple ANN model (perceptron). X_1 to X_n represent the input parameters, W_1 to W_n are weights of the connections, and b is bias of the neuron. Y_1 is the output of the neuron which can be written as:

$$\begin{aligned} v &= \sum_{i=1}^{n} W_i X_i + b \\ Y_1 &= \varphi(v) \end{aligned} \tag{2-1}$$

Function φ is defined as the neuron transfer function (activation function), which maps the inputs of a neuron to the desired output. There are a variety of possible transfer functions; the most common ones are: threshold, linear, Gaussian and sigmoid functions that are depicted in Figure 2-2.



a) **b)**

Figure 2-1. a) Biological Neuron, b) A Single Perceptron

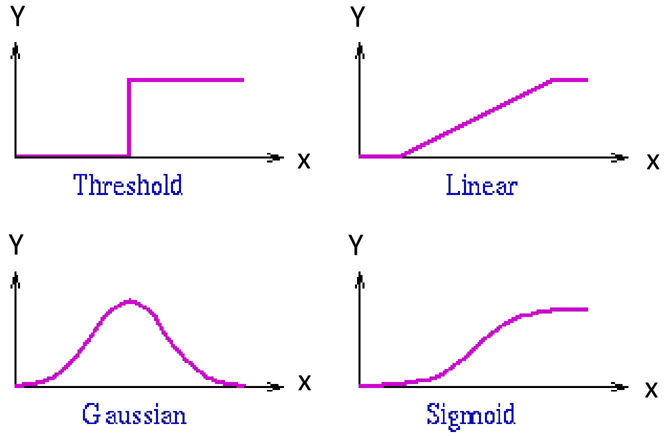


Figure 2-2. Transfer Functions in ANNs

ANNs are usually divided into two broad categories in terms of their architecture: feed-forward and recurrent or feedback networks. In feed-forward networks, the input of each unit in a layer is the output of the units in layer before, while in recurrent networks, the input of a neuron could come from units (neurons) in the latter layer or from itself. Feed-forward networks are memory-less in the sense that their response to an input is independent of the previous network state. Recurrent networks, on the other hand, are dynamic systems.

Learning in ANNs can be considered as updating network architecture and connections' weights so that the network can efficiently perform a task. The network is usually trained by presenting a set of input data. ANNs are able to learn from the underlying rules, like implicit relation between input-outputs, acting as an artificial intelligent model.

ANNs have been developed and modified over time to model complex problems in regression, pattern recognition, prediction, and optimization. Among various types, Radial Basis Functions (RBF) and Multilayer Perceptrons (MLP) are known as Universal Approximators and can be applied for the purpose of nonlinear regression modeling.

The performance of ANNs are usually evaluated using the coefficient of determination (R^2) between the actual or measured values (targets) and the predicted values by the model (outputs), and the root mean square of error (*RMSE*) in the model predictions. To avoid over-fitting, a fraction of the database is used to train the ANN model (training dataset), and the rest are used to evaluate and monitor the generalization ability of the model (validation and test datasets).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - d_i)^2}{\sum_{i=1}^{n}(d_i - \bar{d})^2} \tag{2-2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - d_i)^2} \tag{2-3}$$

where,

y_i = predicted values by a model (output)

d_i = actual or measured values (target)

$\bar{d}$ = average of the measured values over a dataset

All the computational models in this study including the ANN models were developed by programming in MATLAB (MathWorks 2013). Also, for developing the ANN models, the code modules in the Neural Network toolbox of MATLAB were utilized (Demuth et al. 2009).

Radial Basis Function Networks

RBFs are used both for complex pattern classification and approximation problems in high-dimensional spaces. Learning can be viewed as finding a surface in a multidimensional space that provides the best fit to the training data. RBF networks are designed as feed-forward networks consisting of three layers: input, hidden, and output as presented in Figure 2-3. According to Cover’s theorem, in interpolation problems the network performs a nonlinear mapping from the input space to the hidden space, and then a linear mapping from the hidden space to the output space (Haykin 1999).

The first layer consists of input nodes and applies an input vector containing a specific number of elements to the next layer (the hidden layer). The hidden layer is composed of nonlinear neurons that are connected directly to all of the neurons in the output layer. If applying the fixed center learning strategy, the centers of the hidden neurons are the input vectors in the training dataset, and therefore the number of hidden units is equal to the size of the training dataset. The number of nodes in the output layer depends upon the number of the elements in the target of the problem.

The (weighted) distances of the inputs from the hidden units' centers are computed and then passed through the radial basis function. In this sense, a neuron will have a maximum output when the new input exactly matches its center. The outputs of the hidden units are then multiplied by the weights of the second layer (output layer) and are transferred using a linear function. In this sense, the outputs of the network are superpositions of the activities of all the radial functions in the network (Poggio and Girosi 1990).

Learning Process

The hyper-surface maps an m-dimensional input space to the single dimensional output space. Having a set of data for N different point $\left\{\mathbf{X}_i \in \mathbb{R}^m | i = 1, 2, \ldots, N\right\}$ and corresponding set of N real numbers $\left\{d_i \in \mathbb{R}^1 | i = 1, 2, \ldots, N\right\}$ as the desired responses, a function $Y:\mathbb{R}^m \rightarrow \mathbb{R}^1$ can be found that satisfies the multivariable interpolation condition:

$$Y\left(\mathbf{X}_i\right) = d_i, \quad i = 1, 2, \ldots, N \tag{2-4}$$

where Y has the following form:

$$Y\left(\mathbf{X}\right) = \sum_{i=1}^{N} w_i \varphi(\left\|\mathbf{X} - \mathbf{X}_i\right\|) \tag{2-5}$$

φ is a radial-basis function, and **||.||** denotes a norm that is usually Euclidean. The radial-basis function widely used in RBFs' hidden neurons is the Gaussian function:

$$\varphi(\mathbf{X}) = e^{(-\frac{\left\|\mathbf{X}\text{-}\mathbf{C}\right\|^2}{2\sigma^2})} \tag{2-6}$$

where **X** is the input vector, **C** is the vector determining the center of the basis function, and σ is the parameter that specifies the spread of the basis functions and controls the smoothness of the interpolation.

The final output function of the network can be written as follows:

$$\mathbf{Y} = \mathbf{W}\, e^{-(\left\|\mathbf{X} - \mathbf{C}\right\| b_1)^2} + \mathbf{b}_2 \tag{2-7}$$

where,

Y = output matrix (model response)

W = weight matrix (model parameters)

X = input matrix (predictor variables)

C = centers of Gaussian functions

$b_1 = \dfrac{1}{\sigma\sqrt{2}}$ = spread of the Gaussian function (model parameter)

$\mathbf{b_2}$ = bias vector of the output layer (model parameter)

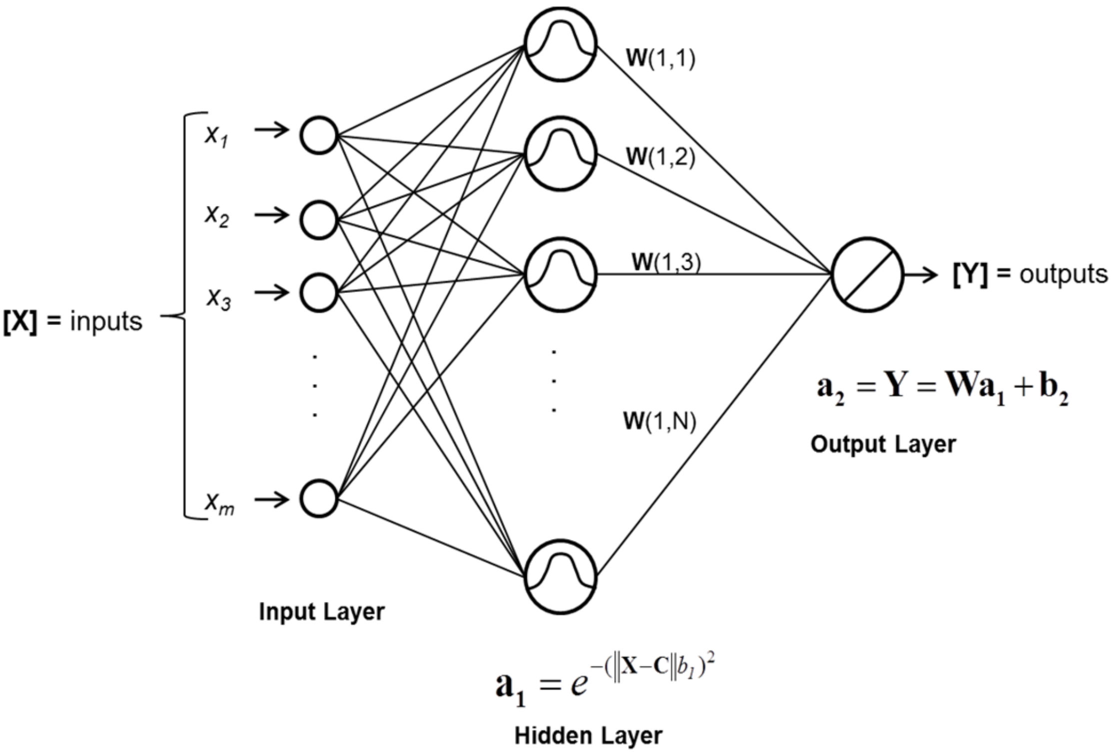


Figure 2-3. Architecture of the RBF Network

Considering $\mathbf{b_2=0}$, the RBF network provides an interpolating function which passes through all the data points. Such a function is over-fitted, and therefore highly oscillatory (Bishop 1995). Thus, the bias parameter was considered in order to smooth the interpolation and to improve the generalization of the model. The performance of the networks was evaluated for various values of b_1. The value that minimized the error on the test data set was eventually selected as the spread of the Gaussian functions.

The weights of the second layer were calculated so as to minimize the error function (the sum of squared errors) given by:

$$E = \frac{1}{2}\sum_{i=1}^{N}(\mathrm{Y}(\mathbf{X}_i) - d_i)^2 \tag{2-8}$$

Minimizing the error function of Eq. 2-8 leaded to solving a set of linear equations:

$$\begin{aligned} \mathbf{\Phi W}^{\mathrm{T}} &= \mathbf{d} \\ \mathbf{W}^{\mathrm{T}} &= (\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi})^{-1}\mathbf{\Phi}^{\mathrm{T}}\mathbf{d} \end{aligned} \tag{2-9}$$

where φ is the output of the hidden layer, and d is the target matrix containing d_i elements.

Training

An ensemble of networks were generated by the Random Subsampling method also known as Monte-Carlo cross-validation (Picard and Cook 1984). In this method, the RBF network is trained for a large number of iterations, and at each iteration data points are randomly selected and assigned to each subset of the database (the training and test subsets) at a 4:1 proportion. Therefore, a new model was trained and evaluated at each iteration that ultimately generated an ensemble of networks. The network learned from the data points in the training set and its generalization was examined over the test dataset. The performance of the model was considered to be the average performance of all the networks in the ensemble. Also the uncertainty in the model performance was assessed through the distribution of R^2 values.

Multilayer Perceptron Networks

MLPs are feed-forward neural networks with at least one hidden layer consisting of neurons with sigmoid activation functions. The learning procedure consists of adjusting the weights of connections in the network to minimize the objective or error function. The MLP network architecture is shown in Figure 2-4. MLPs apply either a log-sigmoid or a tan-sigmoid transfer function in the hidden units, and a linear function in the output units. The tan-sigmoid activation function for MLP networks appears as:

$$\varphi(\mathbf{X}) = \tanh(\mathbf{X}) = \frac{e^{\mathbf{X}} - e^{-\mathbf{X}}}{e^{\mathbf{X}} + e^{-\mathbf{X}}} \qquad \textbf{(2-10)}$$

and the log-sigmoid function appears as:

$$\varphi(\mathbf{X}) = \tanh(\mathbf{X}) = \frac{1}{1 + e^{-\mathbf{X}}} \qquad \textbf{(2-11)}$$

A primary analysis was performed to select the type of sigmoid function; tan-sigmoid and log-sigmoid functions resulted in relatively similar model performance. However, tan-sigmoid slightly outperformed the logistic function. In general, tan-sigmoid activation functions leads to faster convergence than log-sigmoid functions (Bishop 1995).

The output layer has one neuron associated with the response variable. The output function of the network can be written as:

$$\mathbf{Y} = \mathbf{W}_2 \frac{1 - e^{-2(\mathbf{W}_1\mathbf{X}+\mathbf{b}_1)}}{1 + e^{-2(\mathbf{W}_1\mathbf{X}+\mathbf{b}_1)}} + \mathbf{b}_2 \qquad \textbf{(2-12)}$$

where,

$\mathbf{Y}$ = output matrix (model response)

$\mathbf{W_1}$ = weight matrix of the hidden layer (model parameters)

$\mathbf{W_2}$ = weight matrix of the output layer (model parameters)

$\mathbf{X}$ = input matrix (predictor variables)

$\mathbf{b_1}$ = bias vector of the hidden layer (model parameters)

$\mathbf{b_2}$ = bias vector of the output layer (model parameters)

Learning Process

The weights are adjusted according to the Back Propagation (BP) algorithm. Error back-propagation learning consists of conducting two rounds: a forward pass and a backward pass. In the forward pass, an input vector is applied to the sensory nodes of the network, and its effect propagates through the network layer by layer. In a similar way to the Least Squares algorithm, the BP algorithm applies a correction proportional to the derivative of the error function with respect to weights, in the backward pass. The error function is minimized in this process by starting with any set of weights and repeatedly changing each weight by an amount proportional to the derivative of the error function (Haykin 1999).

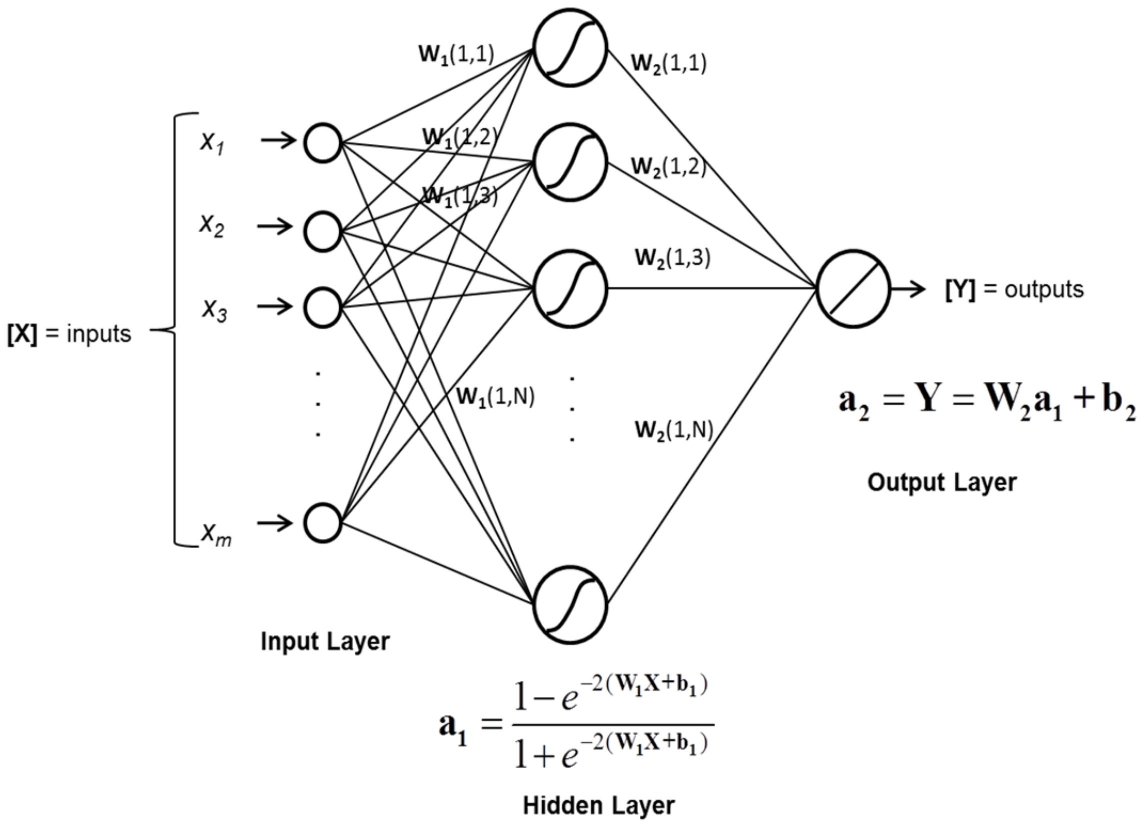


Figure 2-4. Architecture of the MLP Network

Forward Computation

This includes the computation of the local fields and function signal appearing at the output of all neurons, layer by layer through the network. The induced local field $v_j^{(l)}(n)$ for neuron j in layer l is:

$$v_j^{(l)}(n) = \sum_{i=0}^{m_1} w_{ji}^{l}(n)\, y_i^{(l-1)}(n) \tag{2-13}$$

where $y_i^{(l-1)}(n)$ is the output function of neuron *i* in layer *l-1* at iteration *n*, and $w_{ji}^l(n)$ is the synaptic weight of neuron *j* in layer *l* that is fed from neuron *i* in layer *l-1*. For *i* = 0, $y_0^{(l-1)}(n)=1$ and $w_{ji}^l(n)=b_j^l(n)$ where b_j^l is the bias applied to the neuron *j* in layer *l*.

For neuron *j* in layer *L* (output layer) the error signal is computed as:

$$\begin{aligned} y_j^L &= \varphi\left(v_j(n)\right) \\ e_j(n) &= d_j(n) - y^L{}_j(n) \end{aligned} \tag{2-14}$$

where $d_j(n)$ is the *j* element of the desired response vector $d(n)$.

Backward Computation

BP algorithm applies a correction to synaptic weights defined by the delta rule:

$$\Delta w_{ji} = \text{-}\eta \frac{\partial E(n)}{\partial w_{ji}(n)} \tag{2-15}$$

where η is called the leaning-rate parameter and *E* (*n*) is the error function:

$$E(n) = \frac{1}{2}\sum_{j=1}^{m_L} e_j(n)^2 \tag{2-16}$$

This method is based on the gradient decent algorithm for finding weights which leads to finding the minimum on the error surface. According to chain rule in calculus the gradient of the error function with respect to weight is given by:

$$\frac{\partial E(n)}{\partial w_{ji}{}^L(n)} = \frac{\partial E(n)}{\partial e_j{}^L(n)} \frac{\partial e_j{}^L(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)} = \frac{\partial E(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)} \tag{2-17}$$

where,

$$\frac{\partial E(n)}{\partial e_j(n)} = -e_j^{\,L}(n)$$

$$\frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} = \varphi'\left(v^L{}_j(n)\right) \tag{2-18}$$

$$\frac{\partial v_j(n)}{\partial w_{ji}(n)} = y_j^{\,L-1}(n)$$

Defining the local gradient for neuron *j* as:

$$\delta_j(n) = -\frac{\partial E(n)}{\partial v_j(n)} \tag{2-19}$$

This can be written as:

$$\delta^L{}_j(n) = \frac{\partial E(n)}{\partial v_j^{\,L}(n)} = e_j^{(L)}(n)\varphi'\left(v_j^{(L)}(n)\right)\text{: for neuron j in output layer L}$$

(2-20)

$$\delta^l{}_j(n) = \frac{\partial E(n)}{\partial v_j^{\,l}(n)} = \sum_k \frac{\partial E(n)}{\partial v_k^{\,l+1}(n)} \frac{\partial v_k^{\,l+1}(n)}{\partial v_j(n)} = \sum_k \delta_k^{(1+l)}(n) \frac{\partial\left(w_{ki}\, y_j^{\,l}(n)\right)}{\partial v_j(n)}$$

$$= \sum_k \delta_k^{(1+l)}(n) \frac{\partial\left(w_{ki}\varphi\left(v_j^{(l)}(n)\right)\right)}{\partial v_j(n)} = \tag{2-21}$$

$$\varphi'\left(v_j^{(l)}(n)\right)\sum_k w_{ki}\,\delta_k^{(1+l)}(n)\text{: for neuron j in hidden layer l}$$

Thus, the weight adjustment for connection between neuron *j* in layer *l* and neuron *i* in layer *l-1* can be written as:

$$w_{ji} = -\eta \frac{\partial E(n)}{\partial w_{ji}(n)} = \eta \delta_j^{(l)}(n)\, y_i^{(l-1)}(n) \tag{2-22}$$

Levenberg-Marquardt Algorithm

The most popular optimization algorithm for error back propagation is the gradient decent method. However, this method is often very slow, giving preference to high performance optimization algorithms that converge much faster, such as the Levenberg-Marquardt (LM), Conjugate Gradient, Quasi-Newton, and Bayesian Regularization among others. In this study, the LM optimization algorithm was applied to perform the nonlinear error minimizations.

The LM can be considered as an interpolation of the Gauss-Newton (GN) and the gradient decent methods. LM is known to be a more robust algorithm than GN because it can converge even when the initial guess is far from the optimum region; however, it is a bit slower than the GN algorithm. This algorithm solves the following matrix equation for the vector of weights **w**:

$$(\mathbf{J}^T\mathbf{J} + \lambda\, \mathrm{diag}(\mathbf{J}^T\mathbf{J}))\delta = \mathbf{J}^T[\mathbf{T} - \mathbf{Y}(\mathbf{X}, \mathbf{w})] \qquad \textbf{(2-23)}$$

where **Y** is the output of the network or the response of the fitted function at the parameter vector of **w**, and **w** represents the vector containing all the weights and biases of the network. **T** is the vector of the target values (observed or measured values), and **J** is the Jacobian matrix where J_{ij} is defined as:

$$J_{ij} = \frac{\partial \mathbf{Y}(x_i, w_j)}{\partial w_j} \qquad \textbf{(2-24)}$$

δ is the increment value that iteratively adjusts the model parameters (weights), and λ is the damping factor (also called the learning rate parameter) adjusted continuously throughout the process. If the objective function (the sum of the squared errors) decreases rapidly, a smaller value is chosen for λ, and the algorithm leans towards the Gauss-Newton. Whereas, for iteration that generates an insufficient reduction in the objective function, λ is increased that brings the algorithm closer to the gradient descent.

Training

For the purpose of training, the working database was split into three groups: training, validation, and test, with 3:1:1 proportions. The weights of the network were adjusted using training examples. The network performance was monitored during training over the validation dataset. The learning process continued until one of the stopping criteria was reached. The stopping criteria included the minimum error, maximum number of iterations, minimum gradient of error function, and when error over the validation dataset started to increase.

The random subsampling method was performed in order to have a more realistic estimate of the MLP models' performance and to evaluate the uncertainty in the measured performance of the models.

Also, in order to assess the optimum number of hidden neurons, an analysis was performed where MLP networks were created with different hidden neurons ranging from 5 to 100. As a result of this analysis, the best configuration for MLP networks was identified.

Generalized Regression Neural Networks (GRNNs)

GRNNs introduced by Specht (1991) have a similar architecture to RBFs; the only difference between them is in adjusting the second-layer weights. In RBF networks, the weights and biases in the second layer are computed by minimizing the difference in the *MSE*. This usually results in generating exact solutions or a zero error for training data. However, the weights in the second layer of the GRNN are not adjusted using error minimization. Instead, the weights are set equal to the value of the outputs (targets), so the hidden layer has the same number of neurons as the input/target pairs.

Probabilistic Neural Networks (PNNs)

PNNs are used for classification type of problems and have a similar architecture to GRNNs. However, the output layer for PNNs is a decision-making layer and has as many neurons as the number of classes (or categories) of the target parameter. The transfer function of the output neurons for PNNs is not linear as it is in GRNNs. This function (compete function) compares the inputs of all the neurons and outputs one, for the largest value and zero, for the others. Figure 2-5 presents the architecture of the PNN network used in this study.

Linear Regression (LR) and Nonlinear Regression (NLR)

Linear Regression (LR) and Nonlinear Regression (NLR) were performed and the performance of the regression models was compared with ANNs. The LR model was defined as:

$$\mathbf{Y} = \beta_0 + \beta_1 \mathbf{X}_1 + \cdots + \beta_n \mathbf{X}_n = \hat{\boldsymbol{\beta}}\mathbf{X} \quad \textbf{(2-25)}$$

where,

Y = output matrix (model response)

$\beta_0, \ldots, \beta_n$ = regression coefficients (model parameters)

X = input matrix (predictor variables)

The coefficients were computed using the linear least-squares method as:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \quad \textbf{(2-26)}$$

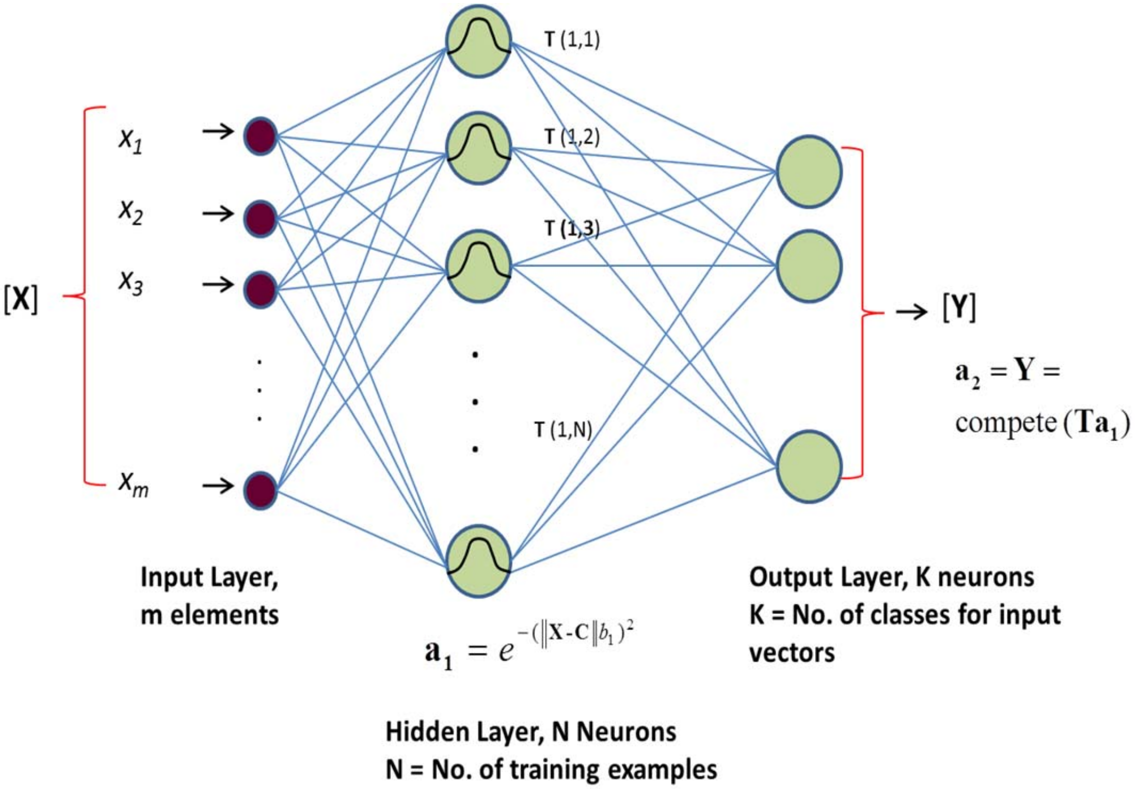


Figure 2-5. Architecture of the PNN Network

Also, a nonlinear model similar to the output function of MLPs was defined as:

$$\mathbf{Y} = \mathbf{LW}\ \mathrm{tansig}(\mathbf{IWX} + \mathbf{b}_1) + \mathbf{b}_2 \tag{2-27}$$

where,

$\mathbf{X}$ [$np \times ns$] = input data

$\mathbf{Y}$ [$no \times ns$] = target data

$\mathbf{IW}$ [$N \times np$] = weights of the first (hidden) layer

$\mathbf{LW}$ [$no \times N$] = weights of the second (output) layer

$\mathbf{b}_1$ [$N \times ns$] = bias of the first layer

$\mathbf{b}_2$ [$no \times ns$] = bias of the second layer

ns = number of training data

np = number of input variables

no = number of output variables

N = number of nonlinear functions (equal to number of hidden neurons in an MLP)

A nonlinear least square approach was then taken to compute the weights and biases. Similar to the MLPs, the LM algorithm was implemented to perform the optimization. The objective function to be minimized was the sum of the squared errors. The weights and biases were initialized by random selection in range [-1 1]. The LM algorithm adjusted the weights and biases at each iteration, starting with the initial values and continuing until the convergence or one of the stopping criteria was reached.

Similar to MLPs, the data was first normalized and 1,000 models were generated through the random subsampling of the data points for the training and test datasets. The main difference between the MLP and NLR models was the difference in computing the gradient of the objective function with respect to weights, which was performed by the backpropagation method for MLPs.

CHAPTER III

PROBABILISTIC METHODS

Bayesian Inference

Beliefs and state of knowledge about an unknown quantity can be mathematically represented as a probability distribution. Updating the current state of knowledge in the light of new evidence based on the Bayes' Theorem is called *Bayesian Inference* (Hoff 2009). In other words, Bayesian inference is a methodology to estimate the probability of a vector of variables when new evidence becomes available. According to the Bayes' paradigm, considering $\boldsymbol{\theta} = \{\theta_1, \ldots, \theta_n\}$, a vector of model parameters, and $\mathbf{d} = \{d_1, \ldots, d_n\}$, a vector of new observations about the process of interest, then $\mathrm{p}(\boldsymbol{\theta})$ represents the prior distribution, $\mathrm{p}(\mathbf{d}|\boldsymbol{\theta})$ represents the likelihood, and $\mathrm{p}(\boldsymbol{\theta}|\mathbf{d})$, the posterior distribution can be computed as (Robert 2007; Hoff 2009):

$$\mathrm{p}(\boldsymbol{\theta} \mid \mathbf{d}) = \frac{\mathrm{p}(\boldsymbol{\theta})\,\mathrm{p}(\mathbf{d} \mid \boldsymbol{\theta})}{\int \mathrm{p}(\boldsymbol{\theta})\,\mathrm{p}(\mathbf{d} \mid \boldsymbol{\theta})\,\mathrm{d}\,\boldsymbol{\theta}} \tag{3-1}$$

The denominator of the above equation is a normalizing constant; thus, the posterior distribution is proportional to the prior times the likelihood:

$$\mathrm{p}(\boldsymbol{\theta} \mid \mathbf{d}) \propto \mathrm{p}(\boldsymbol{\theta})\,\mathrm{p}(\mathbf{d} \mid \boldsymbol{\theta}) \tag{3-2}$$

Calculating the integral in the denominator is often mathematically complex, therefore $\mathrm{p}(\boldsymbol{\theta}|\mathbf{d})$ can be computed only up to the normalizing constant. Also estimating the statistics of the posterior such as mean or confidence intervals require solving complex mathematical integrations. In order to avoid solving difficult integrals, random sampling methods (Monte-Carlo) are often used to approximate different characteristics of posterior distributions by taking samples from p(**θ**|**d**).

The likelihood is defined as the probability distribution of the observations, which is formulated as the discrepancy between model predictions and observations. This discrepancy includes both the measurement error and the model discrepancy. Observations are usually assumed to be independent and identically distributed (i.i.d.) random variable following a Normal distribution, thus the likelihood function can be written as:

$$d_i - u(x_i, \boldsymbol{\theta}) \mid \sigma \sim \text{i.i.d. } N(0, \sigma^2) \qquad \textbf{(3-3)}$$

$$\mathrm{p}(\mathbf{d} \mid \boldsymbol{\theta}) = \mathrm{p}(d_1, \ldots, d_n \mid \boldsymbol{\theta}) = \prod_i N\left(d_i - u(x_i, \boldsymbol{\theta}), 0, \sigma^2\right) \qquad \textbf{(3-4)}$$

where $u(\mathrm{x}, \theta)$ is the forward model, and σ is the standard deviation of error.

If the statistics of the prior distribution or of the likelihood distribution are not known *a priori*, they can be defined as hyper-parameters or random variables themselves, and can be estimated as part of the Bayesian inversion. Assuming the parameters of the prior distributions and the standard deviation of error to be known *a priori* and all the model parameters to be independent, the following equation holds:

$$\mathrm{p}(\theta_1, \ldots, \theta_n \mid \mathbf{d}, \sigma) \propto \mathrm{p}(\theta_1) \ldots \mathrm{p}(\theta_n)\, \mathrm{p}(\mathbf{d} \mid \theta_1, \ldots, \theta_n, \sigma) \qquad \textbf{(3-5)}$$

Markov Chain Monte Carlo

The Markov-Chain Monte-Carlo method (MCMC) allows for generating random samples from a posterior probability distribution when only the kernel of the posterior distribution is available up to the normalizing constant. Theoretically, the simulated samples in the Markov-Chain will converge to a stationary distribution after a number of samples, irrespective of the starting point of the chain. The chain is assumed to converge to the target distribution when the cumulative mean and standard deviation of the samples become stationary. Therefore, the initial samples before convergence are discarded (burn-in value) and the rest of the samples are used to reproduce the posterior distributions. The MCMC allows for constructing the posterior joint probability distributions of the unknown parameters of a model given a set of observations from that model. Also, the posterior marginal distributions of the parameters (as well as the correlation between them) can be obtained using this method (Gamerman and Lopes 2006).

To select or reject samples during the sampling process, there exist different schemes such as the Metropolis-Hastings, the independence, and the Gibbs sampler (Hoff 2009). From all these, Metropolis-Hastings is the most widely used method. Considering $p(\theta)$ to be the target distribution (the posterior distribution), the pseudo code for Metropolis-Hasting's sampler can be written as:

```
- Select a starting point: $\theta^{(0)}$
- for $i=1:n$ (number of samples)
    - Sample a candidate point from an arbitrary proposal distribution: $\theta^{(i)} \sim q(\theta^{(i)} | \theta^{(i-1)})$
```

- Compute $\alpha = \min\left\{ \frac{\mathrm{p}(\theta^{(i)})\,\mathrm{q}(\theta^{(i-1)} \mid \theta^{(i)})}{\mathrm{p}(\theta^{(i-1)})\,\mathrm{q}(\theta^{(i)} \mid \theta^{(i-1)})}, 1 \right\}$

- Sample from the uniform distribution: $u \sim \mathrm{U}(0,1)$

- if $\alpha > u$, $\theta^{(i)} = \theta^{(i)}$,else, $\theta^{(i)} = \theta^{(i-1)}$

- end

The Metropolis algorithm was implemented assuming symmetric proposal distributions. In this study, the candidate samples were drawn from a multivariate normal distribution. The convergence of the Markov chains were evaluated using diagnostic plots, including plots of the cumulative mean and standard deviation of the samples. Also, the autocorrelation of the sample chain and the acceptance rate of the sampler were observed to ensure that the chain was not sticky and had a good mixing.

Note that the probabilistic methods in this study including the MCMC simulations were programmed and implemented in MATLAB.

CHAPTER IV

PREDICTION OF THE STIFFNESS AND STRENGTH OF STABILIZED ORGANIC SOILS

Introduction and Literature Review

Soils cover the vast majority of the lithosphere's surface, where soil formation processes take place due to interactions with the hydrosphere, atmosphere, and biosphere. Although both organic and mineral components are present in most soils, they are found in nature as either predominantly mineral or predominantly organic, depending upon their specific soil formation agents (i.e., parent material, climate, topography, bio-agents, and time). Organic soils usually originate in places with the continuous presence of water, since water prevents organic materials from degrading and turning into sediments. Organic soils are often considered soft and highly compressible, therefore, under high pressures, excessive settlement and instability may become of issue (Edil and Den Haan 1994; Kruse and Haan 2006). Thus, soil stabilization methods often are applied to improve the strength properties of organic soils.

The use of deep soil mixing to improve the stiffness and strength of organic soils is well established (Hampton and Edil 1998; Jelisic and Leppanen 2003; Lambrechts et al. 2003; McGinn and O´Rourke 2003; Hayashi and Nishimoto 2005). Extensive experimental tests and analytical studies have investigated the mechanical behavior of stabilized mineral soils (Mitchell 1981; Schnaid et al. 2001; Chew et al. 2004; Porbaha 2004; Kitazume 2005; Consoli et al. 2007; 2009; 2010). In comparison, limited studies are available in the literature that describes the strength properties and the mechanical behavior of stabilized organic soils (Clare and Sherwood 1956; Huttunen and Kujala 1996; Hampton and Edil 1998; Tremblay et al. 2002; Grubb et al. 2010a; 2010b).

This study models the stiffness and strength of cement-stabilized organic soils and analyzes the effects of various control variables on these parameters. The modeling and analyses were performed based upon a database including the results of unconfined compression tests performed on stabilized organic soils at the University of Cambridge (Hernandez-Martinez 2006).

The mechanical properties (stiffness and strength) of stabilized soils are often measured under various control variables when tested in a laboratory; therefore, modeling and interpreting the effects of these variables and their interactions with the mechanical properties of stabilized soils becomes increasingly complex as the number of control variables and treatment levels increase. In such high-dimensional systems, Artificial Neural Networks (ANNs) can provide a powerful nonlinear regression tool to capture the complex relationships between high-dimensional predictors (control variables) and the response variables (dependent variables).

The objective of this study was to demonstrate the ability of ANNs as robust models to estimate the stiffness (elasticity modulus, *E*) and unconfined compression strength (*UCS*) of organic soils. Also, this study aimed to investigate if ANNs can outperform the conventional empirical methods (e.g. linear regression) that consistently have been used in geotechnical practice for modeling the strength parameters of stabilized soils. Finally, it was intended to improve the understanding of the effects of the key factors, including organic and water content, binder type (Portland cement, blast furnace, pulverized fuel ash, lime, magnesium oxide, and gypsum), and mixing proportions (water to binder ratio, grout to soil ratio) on the mechanical behavior of stabilized organic soils under varying testing conditions, including specimen size, time, temperature, relative humidity, and carbonation.

The performance of ANN models were compared to Linear Regression (LR) models to emphasize the capability of ANN in performing high dimensional nonlinear regressions. Also, a Neural Network–based stepwise parameter selection method was proposed to identify the most influential parameters on the stiffness and strength of stabilized organic soils, and to generate an optimized model. A sensitivity analysis was performed by generating model response graphs in order to investigate the impact of each parameter on the prediction of *E* and *UCS*.

ANNs have been used widely in various geotechnical applications such as in the prediction of pile capacity (Goh 1994a, 1995, 1996; Lee and Lee 1996; Goh et al. 2005; Pal and Deswal 2008), constitutive modeling of soils behavior (Ellis et al. 1995; Penumadu and Jean-Lou 1997; Ghaboussi and Sidarta 1998; Zhu et al. 1998a, b; Penumadu and Zhao 1999), site characterization (Basheer et al. 1996; Rizzo et al. 1996), earth retaining structures (Goh et al. 1995), settlement of foundations (Goh 1994a; Sivakugan et al. 1998; Shahin et al. 2000), slope stability (Ni et al. 1996), design of tunnels and underground openings (Shi et al. 1998), liquefaction (Goh 1994b; Najjar and Ali 1998; Juang and Chen 1999; Young-Su and Byung-Tak 2006), soil permeability and hydraulic conductivity (Najjar and Basheer 1996), soil compaction (Basheer and Najjar

1995; Najjar and Basheer 1996; Najjar et al. 1996b), and soil swelling and classification of soils (Yudong 1995; Najjar et al. 1996a).

A number of studies have applied ANNs to model the stiffness and strength of stabilized soils and other cement-based products. For instance, Stegemann and Buenfeld (2002) performed a neural network analysis to predict the physical and environmental properties of cement-solidified wastes and wastes used in cement-based construction applications. The main purpose of this study was to investigate the effects of impurities on the quality of products made with Portland cement. Data from *UCS* tests of cement paste with pure metal compound additions were used to train a number of ANNs (Multilayer Perceptron). In a similar study, the effects of different types of industrial waste on the unconfined compression strength of Portland cement were investigated using ANNs. According to this study, ANNs were able to account for the nonlinear dependency of *UCS* and waste quantities (Stegemann and Buenfeld 2003). Stegemann and Buenfeld (2004) generated a database from the results of different laboratory tests on cement-stabilized wastes. ANNs were applied to predict the *UCS* and to investigate the effects of contaminants and hydraulic binders on the strength of stabilized wastes. It was found that ANNs were able to account for the nonlinear features in the behavior of stabilized wastes and that developing a more accurate model for practical applications would require more information about material composition, sample preparation, and testing conditions. Narendra et al. (2006) proposed a generic mathematical model using Multilayer Perceptron (MLP), Radial Basis Function (RBF), and Genetic Programming (GP). They used the data from experimental tests, conducted on stabilized CL and CH inorganic clays using different water contents, cement contents, and different curing periods. The *RMSE* computed for the proposed approaches showed that MLP had the best performance among all the methods. GP, RBF, and empirical models were ranked respectively in terms of model performance. Sebastia et al. (2003) used ANNs to predict the *UCS* of coal fly ash-cement mixtures. These researchers showed how neural network analysis could help to identify the variables that have the most significant influence on *UCS*, and compared the results of different MLP networks (with different combinations of input variables) with LR models. Das et al. (2011) utilized MLPs with the Levenberg-Marquardt, Bayesian Regularization, and Differential Evolution algorithms to predict the *UCS* and maximum dry density of cement-stabilized soils. The inputs for the ANN models included liquid limit, plasticity index, clay fraction, sand, gravel, moisture content, and cement content.

The following sections describe the experimental database, the predictive models generated in this study, and the inferences generated from the proposed approach. The performance of different models for prediction of *E* and *UCS* are discussed and compared. The main emphasis is on identifying patterns that can improve the overall understanding of the stiffness and strength of stabilized organic soils. The results of the stepwise parameter selection method are described, and the most influential parameters are identified for the different ANN models. Finally, the results of the sensitivity analysis are discussed based on the model response graphs presented for the different parameters, and the impact of each parameter on the model responses is illustrated.

Approach

Experimental Database

The experimental database includes the results of unconfined compression tests on three organic soils, including Irish moss peat (*Pt*) with 94% organic content, and two organic clays *OH-1* and *OH-2*, one with medium organic content (30%) and extremely high plasticity (*PL*=123%) and the other with low organic content (4%) and high plasticity (*PL*=34%), respectively. The soils were mixed with six mixtures made of six different binders, including Portland Cement (*PC*), Blast Furnace Slag (*BFS*), Pulverized Fuel Ash (*PFA*), Lime (*L*), Gypsum (*G*), and Magnesium Oxide Cement (*MgO-C*). Table 4-1 shows a summary of the characteristics of the organic soils, binders, and mixtures used to prepare the samples (Hernandez-Martinez 2006). A histogram of the number of experiments for each soil and binder is presented in Figure 4-1.

Table 4-1. Characteristics of the Organic Soils, Binders, and Mixtures

Organic Soil	**Density (kg/m^3)**	***OC* (%)**	***w* (%)**
Irish Moss Peat (*Pt*)	294	94	210
	446		500
	1014		1000
Medium Organic Clay (*OH-1*)	1219	30	180
Low Organic Clay (*OH-2*)	1471	4	85

Binders	**Binder Mixtures**	**Ratio**
Portland Cement (*PC*)	*PC*	1
Blast Furnace Slag (*BFS*)	*PC+BFS*	1:2
Pulverized Fuel Ash *(PFA)*	*PC+PFA*	1:1
Lime (*L*)	*PC+PFA+L*	3:6:1
Magnesium Oxide Cement (*MgO-C*)	*PC+PFA+MgO-C*	1:3:1
Gypsum (*G*)	L-G-BFS	1:1:1

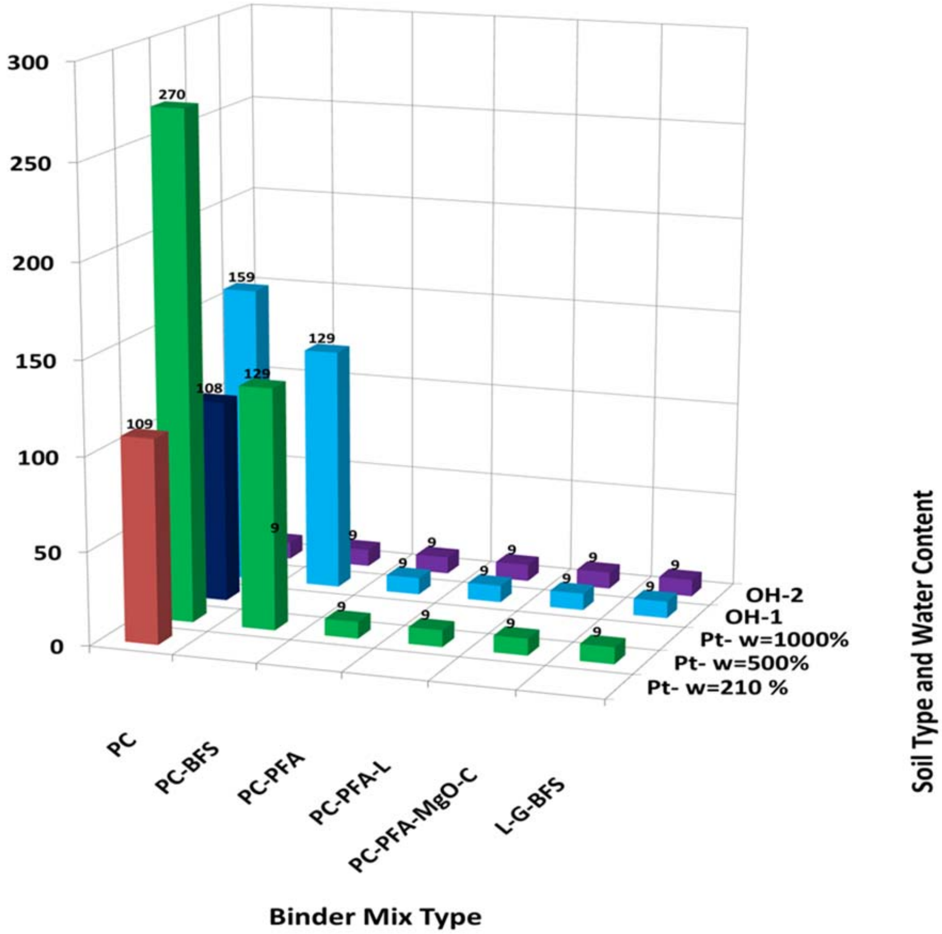


Figure 4-1. **Histogram of the Unconfined Compression Tests**

For preparing the testing specimens first, the organic soils were mixed with grout using a small, concrete-type mixer for ten minutes. Then the produced mix was placed in five layers in a split PVC. After the specimens were prepared, most were cured at the environment temperature (21°C). To accelerate the ageing, some of the specimens were placed in conventional ovens and in carbon dioxide (CO_2) incubators at three different temperatures (21°C, 45°C and 60°C) and at three different relative humilities (70%, 80%, and 90%). The samples were cured to different ages of 14, 28, 60, 90, 105, and 120 days before they were removed for testing. The experimental database includes the results of 1,030 tests, with total of 339 distinct tests after excluding the repetitive ones. For the purpose of modeling, the values of *E* and *UCS* were averaged for the repeated tests. Table 4-2 presents the control variables and their corresponding ranges of variation in the database. The histograms of the variables are provided in Figure 4-2.

Table 4-2. List of Control and Response Variables and Their Levels/Range of Variation in the Database

No.	**Variables**	**Range of Variation**	**Mean**	**Standard Deviation**
Control Variables (Input Parameters)				
1	Organic Content of Soil (*OC*)	4 (*OH-2*), 30 (*OH-1*), 94 (*Pt*) (%)	69.2 (%)	33.2 (%)
2	Water Content of Soil (*w*)	85, 180, 210, 500, 1000 (%)	400.2 (%)	258.6 (%)
3	Ratio of Binder for Portland Cement (*RB-PC*)	0, 0.2, 0.3, 0.33, 0.5, 1	0.749	0.339
4	Ratio of Binder for Blast Furnace Slag (*RB-BFS)*	0, 0.33, 0.67	0.180	0.292
5	Ratio of Binder for Pulverized Fuel Ash (*RB-PFA)*	0, 0.5, 0.6	0.045	0.154
6	Ratio of Binder for Lime (*RB-L*)	0, 0.1, 0.33	0.012	0.056
7	Ratio of Binder for Magnesium Oxide (*RB-MgO-C*)	0, 0.2	0.005	0.032
8	Ratio of Binder for Gypsum (*RB-G*)	0, 0.33	0.009	0.054
9	Quantity of Binder (*QB*)	100 - 500 (kg/m^3)	265.9 (kg/m^3)	76.6 (kg/m^3)
10	Grout to Soil Ratio (*G/S)*	0.14 - 3.38	0.856	0.568
11	Water to Binder ratio (*W/B*)	0.5, 0.8, 1	0.937	0.157
12	Specimen Diameter/Height (*D/H*)	0.5, 1	0.920	0.183
13	Time (*t*)	14 - 180 (days)	61 (days)	36 (days)
14	Temperature (*T*)	21, 45, 60 (ºC)	32.9 (ºC)	15.7 (ºC)
15	Relative Humidity (*RH*)	70, 80, 90 (%)	87.6 (%)	4.9 (%)
16	Carbonation (CO_2)	0, 20 (%)	2.9 (%)	7.1 (%)
Response Variables (Target Parameters)				
	Elastic Modulus (*E*)	0.83 - 214.62 (MPa)	34.03 (MPa)	31.44 (MPa)
	Unconfined Compression Strength (*UCS*)	0.04 - 2.09 (MPa)	0.50 (MPa)	0.40 (MPa)

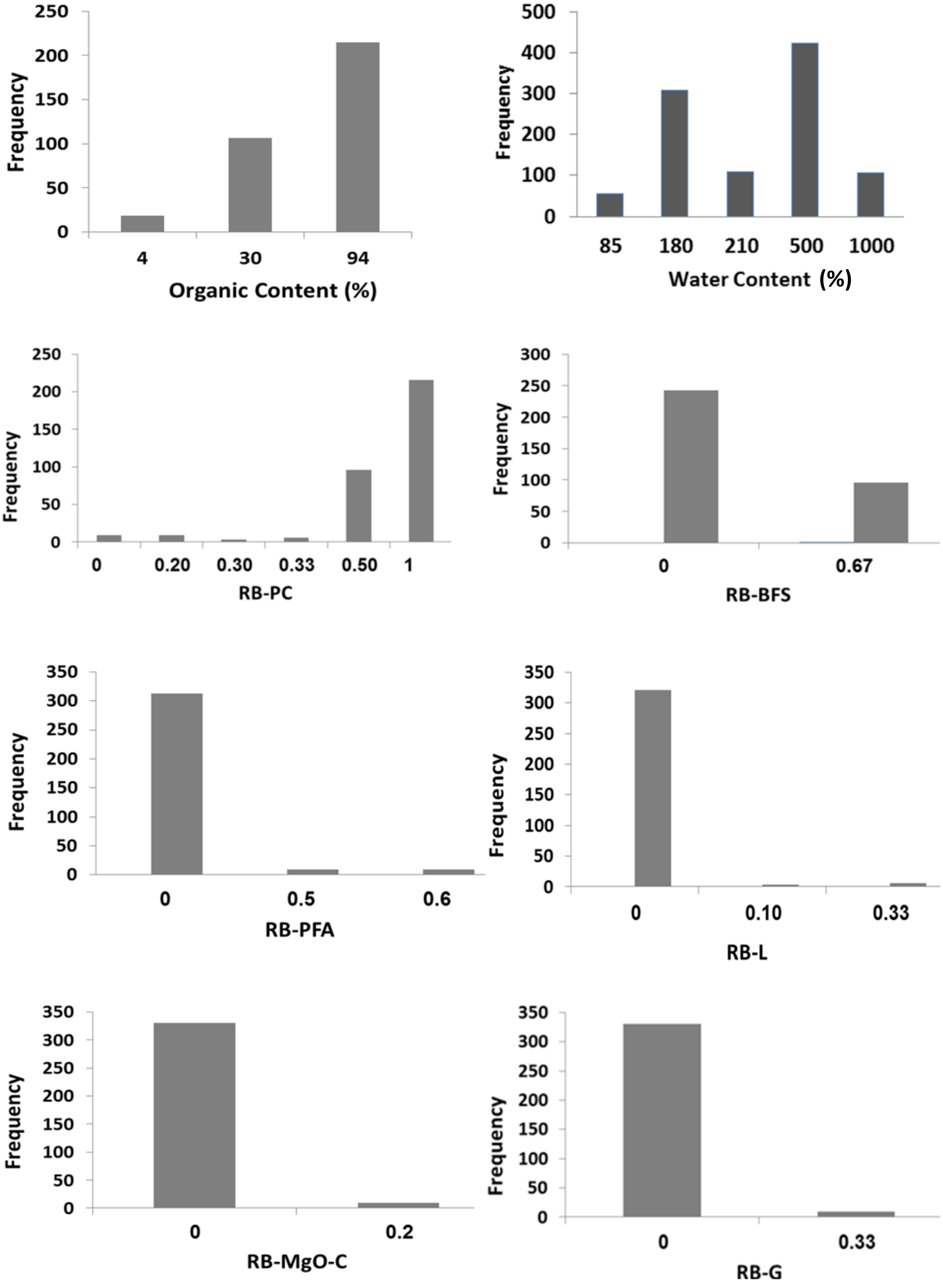


Figure 4-2. Histograms of the Control Variables in the Unconfined Compression Tests

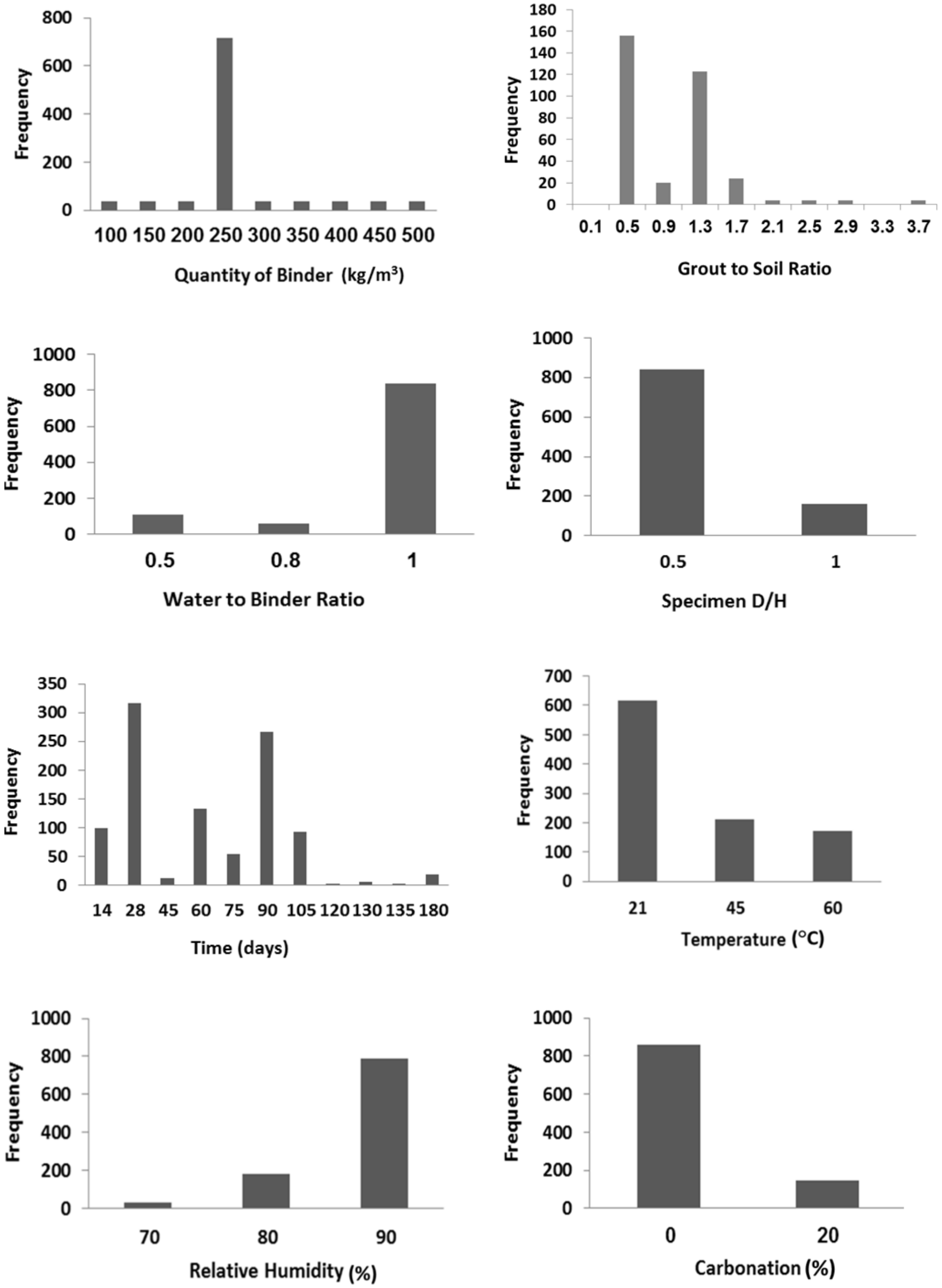


Figure 4-2. Continued

Numerical Experiments

The experimental database did not include a full factorial design. This means that only a fraction of the possible combinations of the different levels for the control variables were considered in the laboratory experiments. In order to study the effects of experimental design on the model performance, three subsets of the database were identified with a more complete factorial design that resulted in three numerical experiments.

In each numerical experiment, one or more specific factors (control variables) were varied in their ranges to investigate their effects, while other factors were held constant at their control values. Figure 4-3 to Figure 4-5 present the graphical representations of the three subsets associated with the numerical experiments. Table 4-3 presents a summary indicating the characteristics of each group according to the participating variables.

Table 4-3. Characteristics of the Numerical Experiments

	No. of Tests	**Varying Variables**	**Constant Variables**
Group 1	108	*w, W/B, QB, G/S, D/H, t*	*OC, RB-PC, RB-BFS, RB-PFA, RB-L, RB-MgO-C, RB-G, T, RH,* CO_2
Group 2	178	*OC, w, RB-PC, RB-BFS, T, t RH,* CO_2	*RB- PFA, RB-L, RB-MgO-C, RB-G, W/B , QB, G/S, D/H*
Group 3	53	*OC, w, RB-PC, RB-BFS, RB-PFA, RB-L, RB-MgO-C, RB-G, t*	*W/*B, *QB, G/S, D/H, T, RH ,* CO_2
Total	**339**		

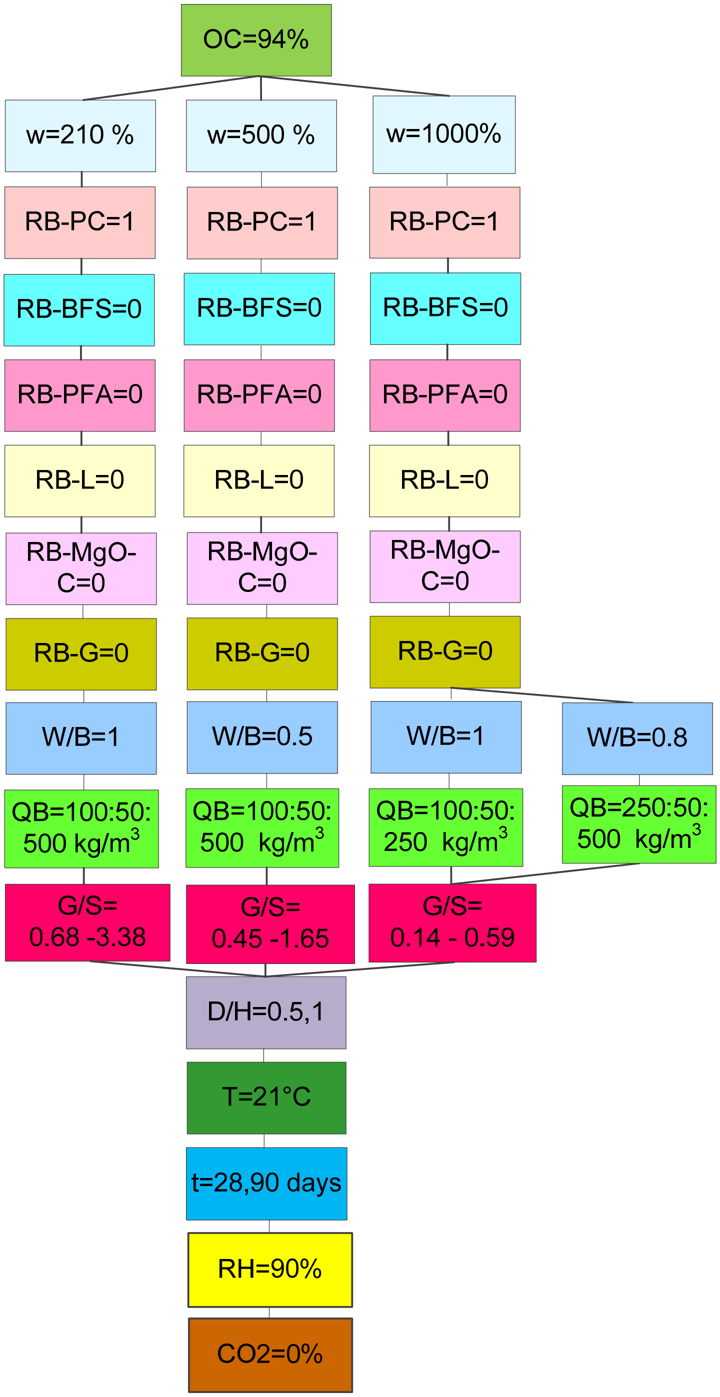


Figure 4-3. Numerical Experiments: Group 1

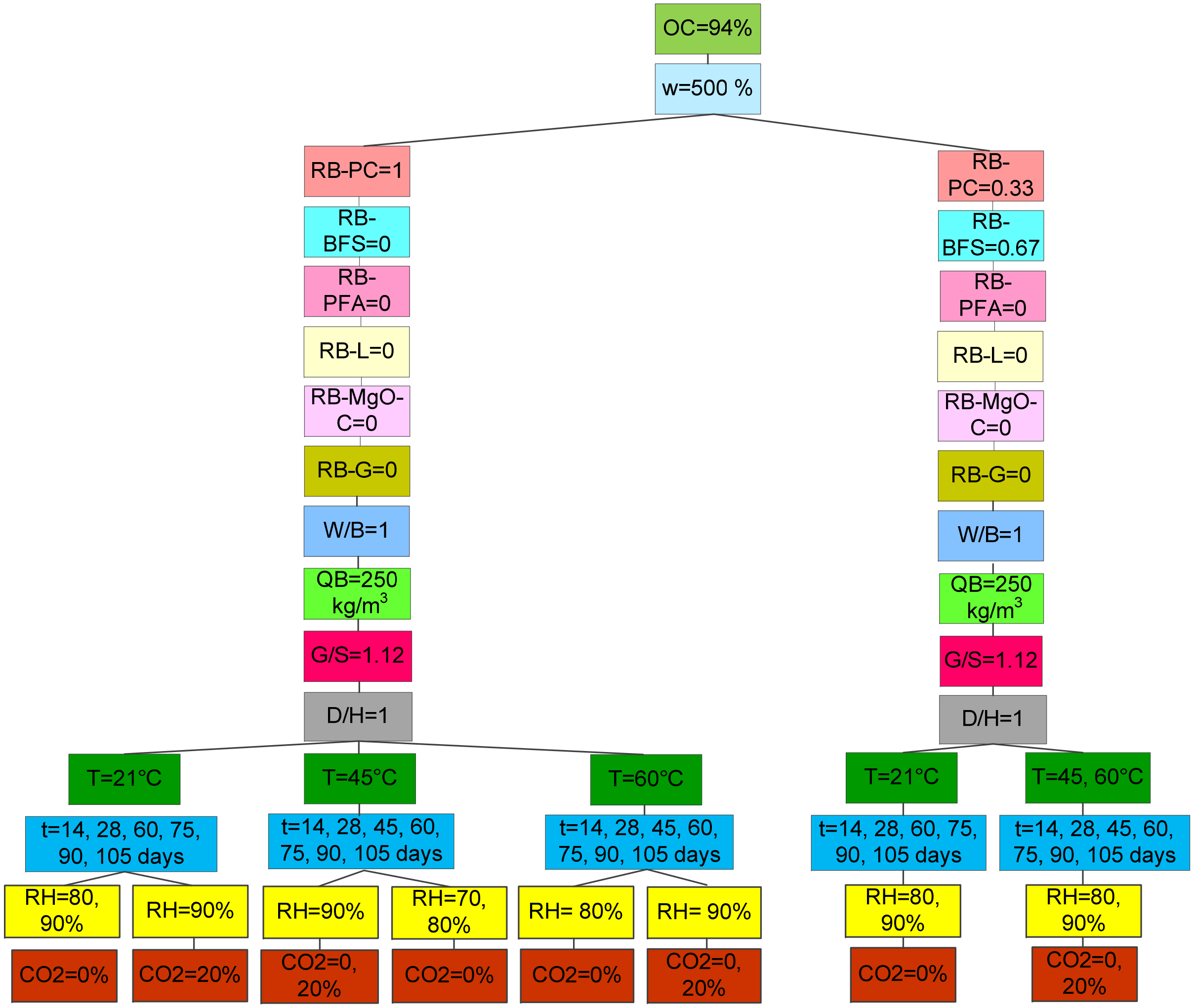


Figure 4-4. Numerical Experiments: Group 2

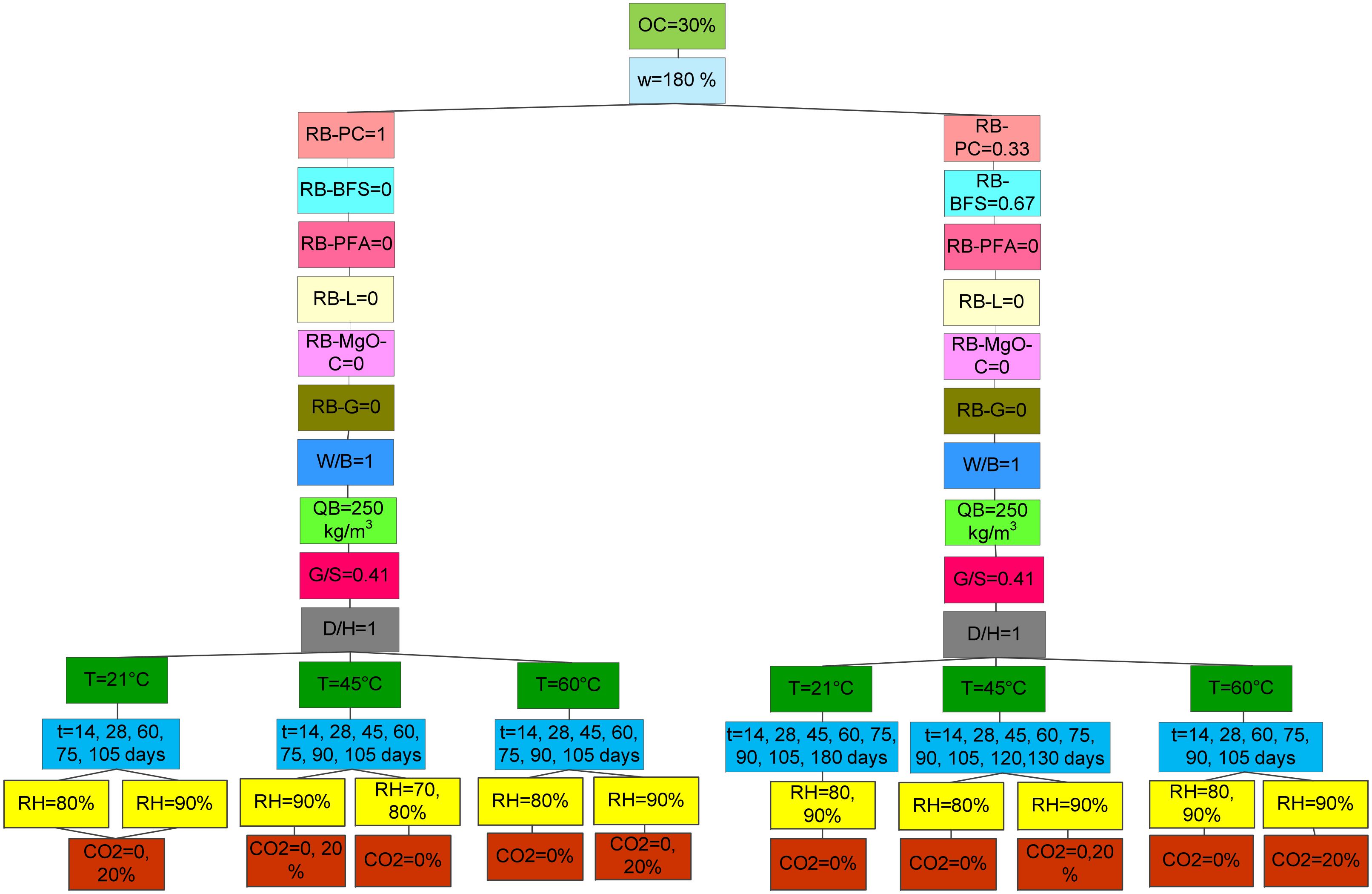


Figure 4-4. Continued

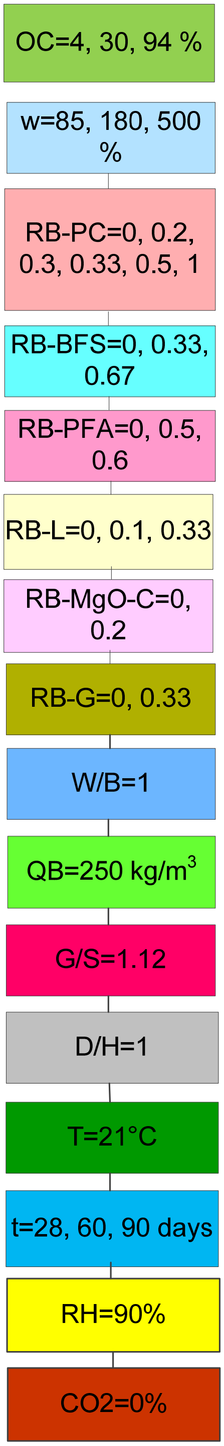


Figure 4-5. Numerical Experiments: Group 3

Models

RBF and MLP networks were developed to predict *E* and *UCS* as presented in Figure 4-6 and Figure 4-7. The performance of the networks were evaluated using the R^2 between the measured value in the laboratory (target) and the output of the model (output), and the *RMSE* of the model predictions. To avoid over-fitting, a fraction of the database was used to train the proposed ANN models (training dataset), and the rest was used to evaluate and monitor the generalization ability of the model (validation and test datasets). Ensembles of networks were generated by the random subsampling method for both RBF and MLP models, and the statistics of the networks' performances were evaluated. Also, a linear regression model was developed and trained in a similar manner as for the ANNs.

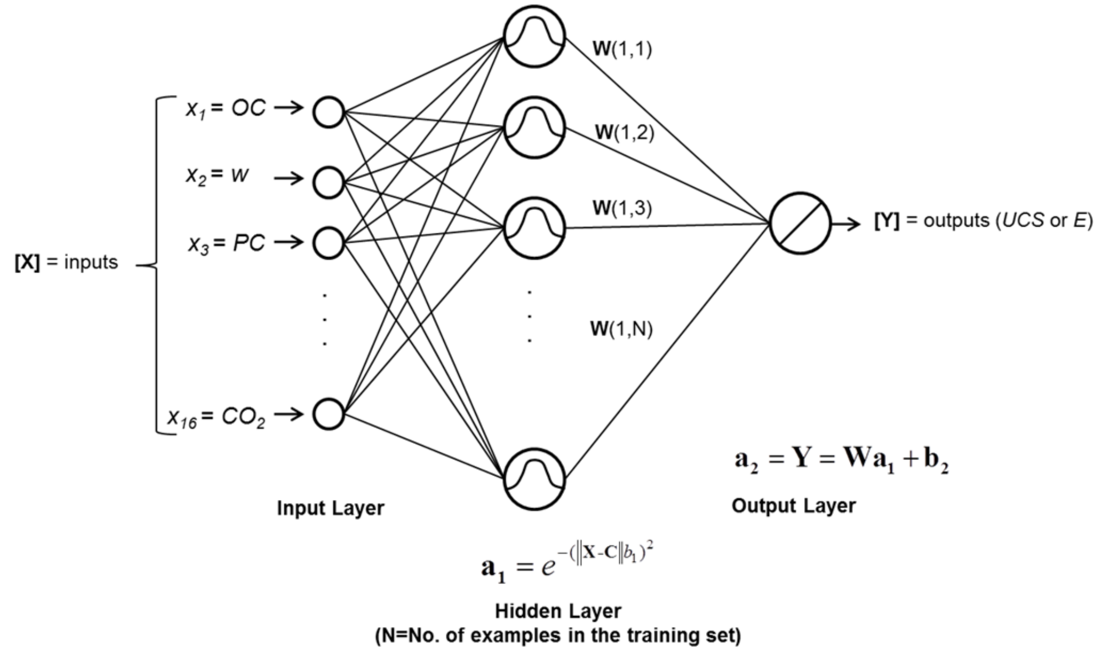


Figure 4-6. RBF Network Architecture

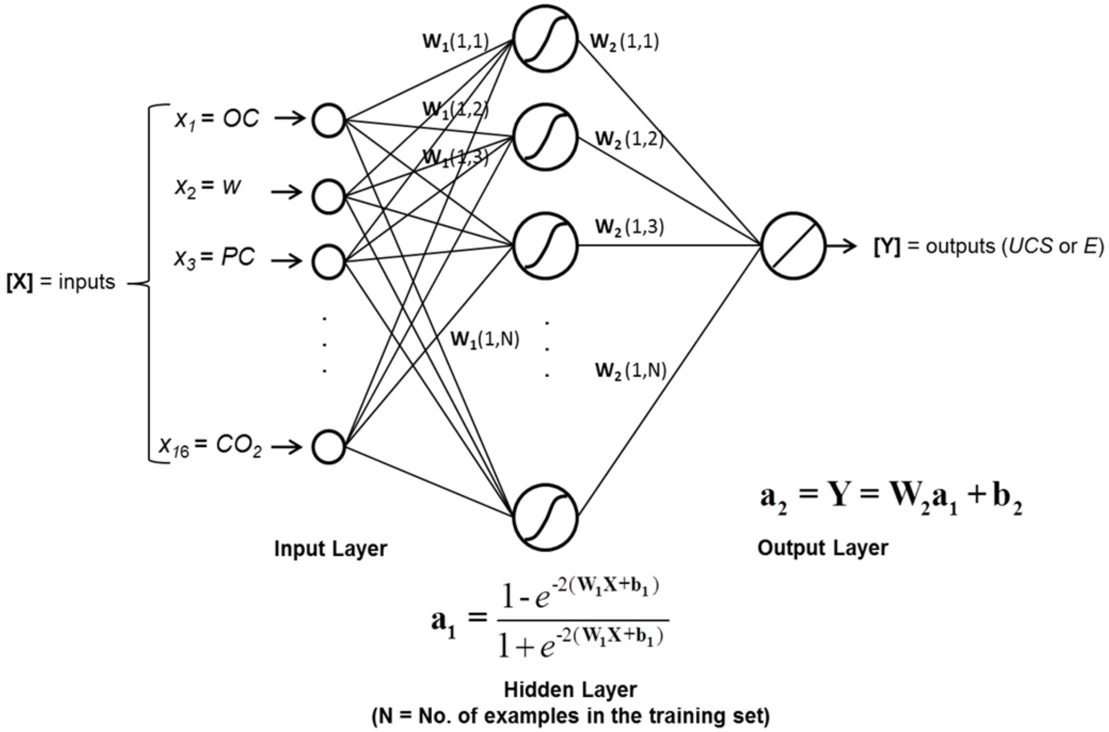


Figure 4-7. MLP Network Architecture

Results and Discussion

RBF Network Analysis

As discussed above, three numerical experiments were performed by defining three subsets of the database. This categorization of data provided three groups of experiments with more complete factorial designs (see Figures 4-3 through 4-5). The data was normalized prior to presenting to the network. The data normalization improved the performance of the RBF models, as the R^2 slightly increased over the test datasets. Table 4-4 and Table 4-5 present a summary of the prediction results for the three groups and for the total database for *E* and *UCS*.

Table 4-4. Performances of the RBF Models Trained with the Three Subsets

Model	**$RMSE_{ave}$ (Test)**	
	E	***UCS***
RBF-Tot	13.77	0.11
RBF-Gr1	12.59	0.11
RBF-Gr2	9.61	0.10
RBF-Gr3	25.09	0.14

Table 4-5. Performance of the RBF Model Trained Using the Whole Database

Model	**All**		**Training**		**Test**	
	R^2_{ave}	$RMSE_{ave}$	R^2_{ave}	$RMSE_{ave}$	R^2_{ave}	$RMSE_{ave}$
RBF-Tot-*E*	0.89	2.778	0.92	0.03	0.77	13.77
RBF-Tot-*UCS*	0.952	0.086	0.96	0.08	0.92	0.11

The performance of the models trained over the three different subsets and the model trained using the whole database was compared based on the average *RMSE* rather than R^2, as the data points that were used for training these models were different in terms of variability and number. It was observed that the $RMSE_{ave}$ over the test datasets for RBF-Tot was almost equal to that of RBF-Gr1 and RBF-Gr2 for prediction of *UCS*. However, for prediction of *E*, RBF-Gr1 and RBF-Gr2 slightly outperformed the RBF-Tot. Also, RBF-Gr3 showed significantly higher error comparing to other models including RBF-Tot, particularly for the prediction of *E*. The reason could be that Group 3 contained the least number of data points, which might have caused the RBF-Gr3 to be overfitted to the data points. On the contrary, it was observed that RBF-Gr2 that was trained with more data points (Group 2) showed the best performance among the three models, despite of the high variability in the levels of the input variables. The better performance of RBF-Gr2 comparing to RBF-Tot showed that clustering the database into subsets with more complete factorial designs could improve the performance of the trained models. However, in this study clustering the database did not improve the performance of all the models corresponding to the three subsets of the database. Also, for each Group, some of the variables were held constant; therefore, for sensitivity analysis either the three models should be used together, or the RBF-Tot, in order to be able to explore the effect of all the input parameters. Because the RBF-Gr1 and RNF-Gr3 did not outperform RBF-Tot, and to be able to have a comprehensive model that could take into account the variability of all the input parameters, RBF-Tot was selected for further analyses.

RBF-Tot demonstrated relatively better performance with $R^2_{ave}=0.92$ for prediction of *UCS* comparing to prediction of *E* with $R^2_{ave}=0.77$. The reason could be that *UCS* was the primary observation in the experiments obtained as the peak value of the stress-strain curve, whereas *E* not only depended on the peak value, but also on the level of deformation achieved. Therefore, there might be other factors, for instance the axial strain at failure, affecting *E* that were not considered as input variables for the models.

In general, the good performance of RBF-Tot shows that the RBF model learned the relationship between the inputs and the targets and successfully generalized the relation to the new data points. Figure 4-8 shows the distribution of the R^2 values that correspond to the 1,000 networks of the RBF-Tot ensemble. It is evident that the uncertainty of the predictions for the training dataset was significantly lower than for that of the test dataset. Also, one can observe that the uncertainty of the predictions for *E* was significantly higher than for *UCS*.

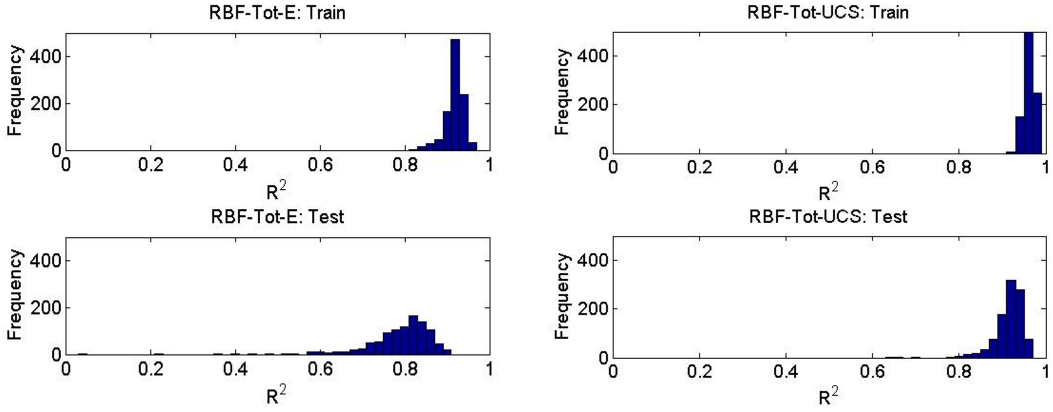


Figure 4-8. Histograms of the R^2 Values, Corresponding to the 1000 Networks of the RBF-Tot Ensembles

MLP Network Analysis

There are a number of rules of thumb for determining the number of hidden units. For instance, one rule is that one hidden layer usually suffices and the number of hidden neurons should be between the number of input elements and the number of output elements. Another rule indicates that the number of hidden neurons should not be greater than twice the number of input elements. However, there is no theory available yet which specifies the required number of hidden units for an MLP network. According to (Bishop 1995), in MLPs with continuous nonlinear hidden-layer activation functions, one hidden layer with an arbitrarily large number of units suffices for the "universal approximation" property. A more recent study (Trenn 2008) suggests that the minimum number of hidden units for a single hidden layer network to reach a certain order of approximation is:

$$n = \frac{\begin{pmatrix} N + n_0 \\ n_0 \end{pmatrix}}{n_0 + 2} \tag{4-1}$$

where,

n = Minimum number of hidden units

N = Order of approximation

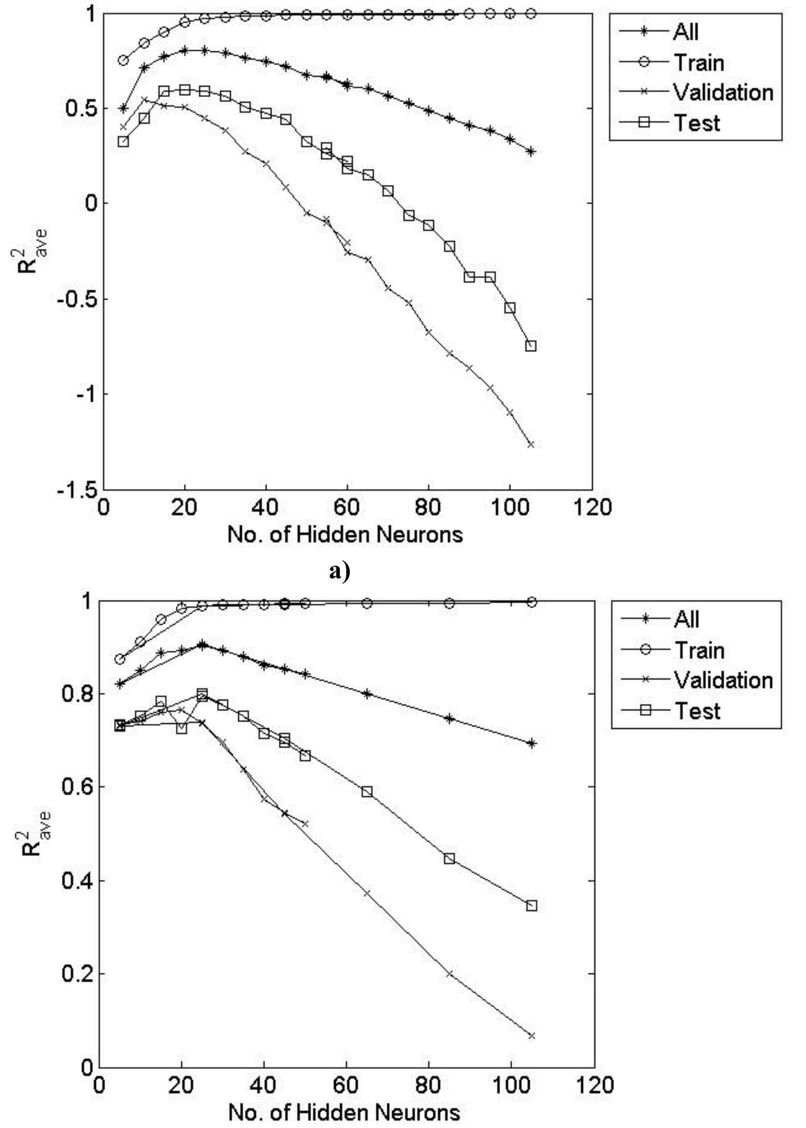
n_0 = Number of input variables

N refers to the degree of Taylor polynomials approximating the MLP function. For instance, in order to have an approximation to the order of 2 in this problem (with 16 input variables), at least nine hidden neurons is required.

Too few hidden neurons increase the prediction error and prevent the ANN model from having enough flexibility to fit the data points. On the other hand, too many hidden neurons make the model too complex and cause the model to be over-fitted to the data and fail to generalize to new data points.

In order to investigate the effect of the number of hidden neurons on the performance of the MLP networks, ensembles of MLP networks with different numbers of hidden neurons were created. Figure 4-9 presents the variation of the average R^2 for the MLP ensembles with the numbers of hidden neurons. According to this figure, the optimum number of hidden neurons for the MLP network was 20 for the prediction of E

and 25 for *UCS* predictions. The generalization ability of the network dropped significantly, as the number of hidden neurons exceeded the optimum value.



a)

b)

Figure 4-9. Variation of the Average R^2 with the Number of Hidden Neurons, a) *E*, b) *UCS*

Note that all of the MLP networks were trained using the whole database because the numerical experiments indicated that the division of the database did not enhance the performance of the models for all the three subsets.

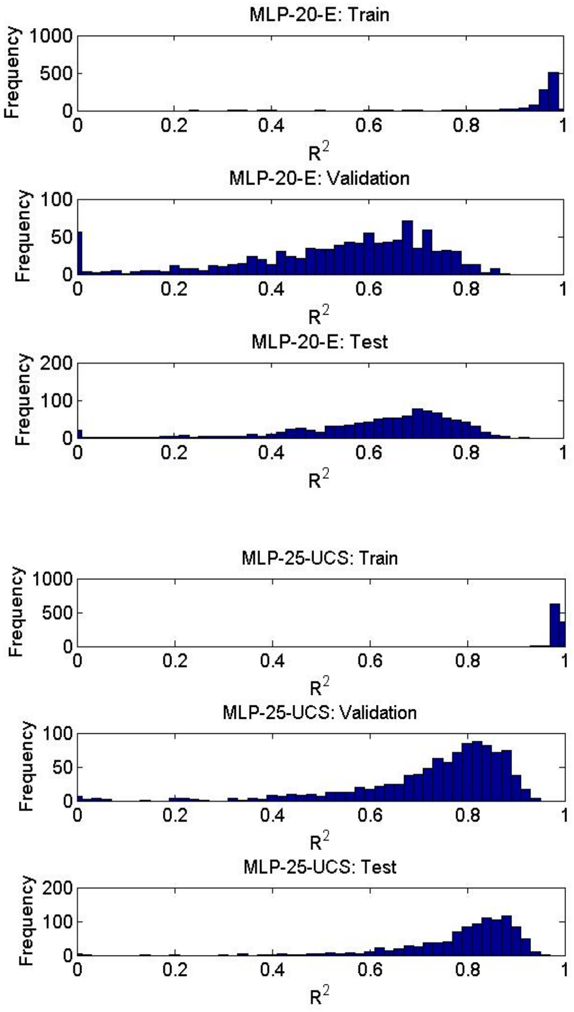
The prediction results of the MLP ensembles for *E* and *UCS* are summarized in Table 4-6. The average R^2 of the test dataset is 0.61 for *E* and 0.8 for *UCS*. Therefore, the RBF network seemed to outperform the MLP networks in this problem. Similar to RBF-Tot model, the MLP models showed higher R^2_{ave} for the prediction of *UCS* as compared to *E*.

Figure 4-10 presents the distribution of the R^2 associated with the 1,000 networks of the MLP ensemble. According to this figure, the uncertainty of the predictions over the training dataset was very small as compared to the other subsets. Also, the uncertainty of the prediction of *UCS* was less than *E*.

Figure 4-11 presents the predictions of *E* and *UCS* generated by one of the models in the MLP ensemble over the training and the test subsets.

Table 4-6. Performances of the MLP Ensembles for *E* and *UCS* Prediction

Model	R^2_{ave}				$RMSE_{ave}$			
	All	**Training**	**Validation**	**Test**	**All**	**Training**	**Validation**	**Test**
MLP-20-*E*	0.81	0.95	0.50	0.61	13.58	6.15	20.74	18.47
MLP-25-*UCS*	0.90	0.99	0.73	0.80	0.12	0.04	0.20	0.17



MLP-20-E: Train

MLP-20-E: Validation

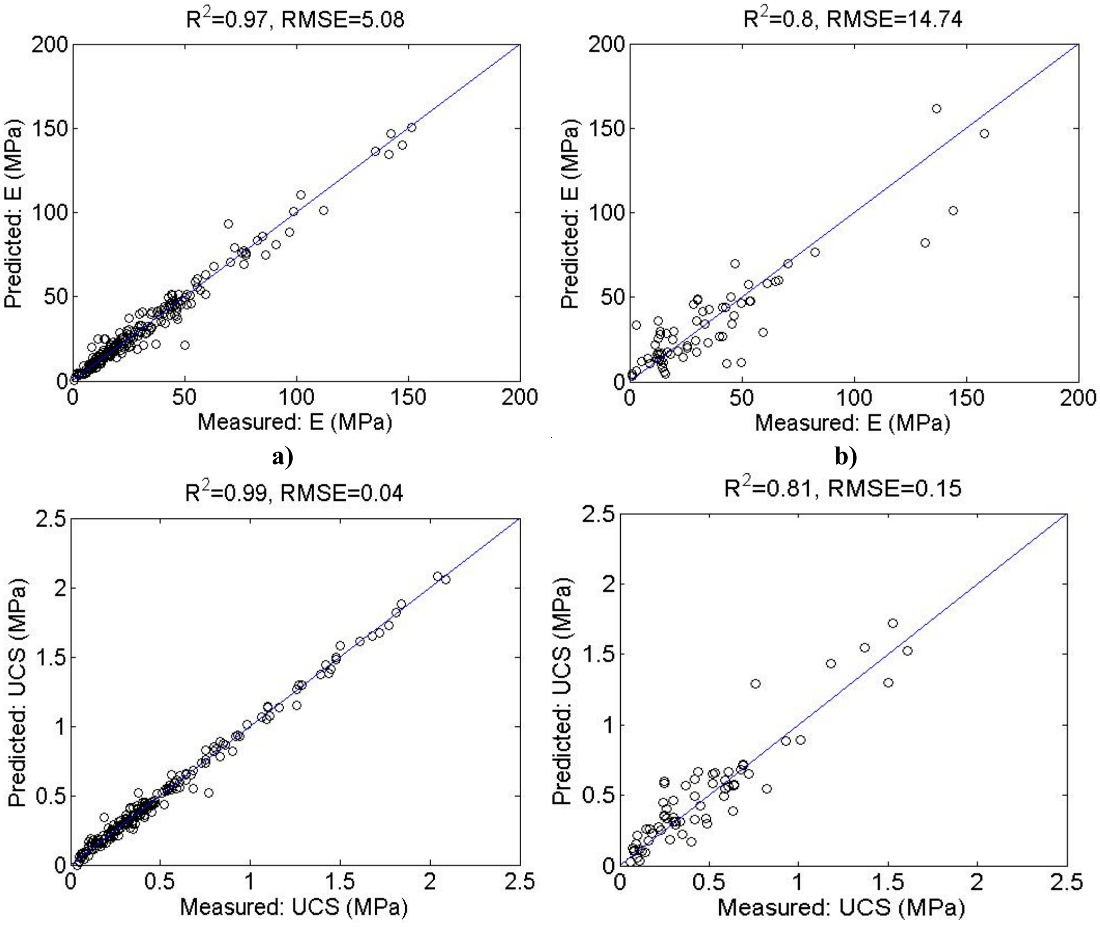
MLP-20-E: Test

MLP-25-UCS: Train

MLP-25-UCS: Validation

MLP-25-UCS: Test

Figure 4-10. Histogram of the R^2 Values, Corresponding to the 1000 Networks in the MLP Ensemble



a) b)

c) d)

Figure 4-11. Predictions of *E* over a) Training, and b) Test Datasets, and Predictions of *UCS* for c) Training and d) Test Datasets Generated by a Network in the MLP Ensemble

Stepwise Selection of Input Parameters for ANNs

In order to identify the most significant parameters for the strength properties of stabilized organic soils, a stepwise parameter selection method was implemented (Gevrey et al. 2003). The stepwise selection method for ANNs reduces the model dimensionality and optimizes the performance of the network. Highly correlated predictor variables can negatively affect the model performance. Thus, in order to optimize the performance of the model, it is necessary to remove the redundant parameters from the model.

This method includes Forward Stepwise Selection (FSS) and Backward Stepwise Elimination (BSE) of the input parameters, based on their significance to the performance of the network. At each step, an ensemble of networks is trained using a control subset of the database (including 80% of the whole database) and the selected input parameters. In forward stepwise selection, parameters compete with each other to be added to the model, one by one. A parameter is added to the model only if the adjusted R^2 for the model becomes larger than the adjusted R^2 of the reduced model.

In the first step, 11 ANN models were generated, each with only one of the parameters in the study. The parameter that had the greatest effect on the model performance (the largest adjusted R^2) was added to the model. Likewise, in the second step 10 models were generated with the selected parameter in step one, as well as each of the remaining parameters. This process continued until no other variable could be added to the model. In other words, the procedure stopped once adding the input parameters could no longer improve the model's performance.

In the backward elimination, a network containing all the parameters was created at first, and then the parameters were gradually removed one by one. The parameters competed with each other to be removed from the model. A parameter was removed when the adjusted R^2 for the reduced model became larger than the present model. In the first step 11 ANN models were generated, each with only one of the parameters missing. Likewise, in the second step 10 models were generated by removing another parameter from the 10 remaining parameters. This process continued until no other input parameter could be eliminated from the model.

Adjusted R^2 provided the possibility of penalizing the model for the number of parameters, meaning that R^2 was decreased in relation to the number of parameters in the model. Adjusted R^2 was calculated based on the following equation:

$$R^2_{adj} = \frac{(n-1)R^2 - p}{n-1-p} \tag{4-2}$$

where:

R^2_{adj} = Adjusted value

n = No. of data points presented to the network in the training dataset

$R^2 = R^2$ for the reduced model

p = No. of network's parameters (weights and biases) for the reduced model

Table 4-7 and Table 4-8 present the selected and removed parameters using the forward and backward stepwise methods for the RBF and MLP models, respectively. The parameters were ranked based on their sequence of addition to the model in the FSS process, which represented their significance to the model's performance. For the purpose of ranking, the process of adding the parameters continued even after the R^2_{adj} of the reduced model started to decrease. After this stage, the parameter that caused the least decrease in R^2_{adj} was added to the model at each step. This continued until all the parameters were added to the model.

Based on the FSS method, *w*, *OC*, *QB*, and *W/B* were parameters that could be removed from the RBF model for the prediction of *UCS*. However, based on the BSE method, only *W/B* could be eliminated. For the prediction of *E*, FSS and BSE selected the same parameters.

Table 4-7. Stepwise Selection of Parameters for the RBF Network

Order of Addition/ Elimination	**RBF-*E***			**RBF-*UCS***		
	FSS	**BSE**	**Ranking/ Sequence of Addition**	**FSS**	**BSE**	**Ranking/ Sequence of Addition**
1	*Binder Type*	*w*	*Binder Type*	*G/S*	*W/B*	*G/S*
2	*G/S*	*OC*	*G/S*	*Binder Type*		*Binder Type*
3	*t*	*QB*	*t*	*t*		*t*
4	*D/H*	*W/B*	*D/H*	*T*		*T*
5	*T*		*T*	*RH*		*RH*
6	*RH*		*RH*	CO_2		CO_2
7	CO_2		CO_2	*D/H*		*D/H*
8			*w*			*w*
9			*OC*			*OC*
10			*QB*			*QB*
11			*W/B*			*W/B*

Table 4-8. Stepwise Selection of Parameters for the MLP Network.

Order of Addition/ Elimination	MLP-*E*			MLP-*UCS*		
	FSS	**BSE**	**Ranking/ Sequence of Addition**	**FSS**	**BSE**	**Ranking/ Sequence of Addition**
1	*Binder Type*	*W/B*	*Binder Type*	*G/S*	*QB*	*G/S*
2	*G/S*		*G/S*	*Binder Type*		*Binder Type*
3	*t*		*t*	*t*		*t*
4	*D/H*		*D/H*	*T*		*T*
5	*T*		*T*	*RH*		*RH*
6	*RH*		*RH*	CO_2		CO_2
7	CO_2		CO_2	*W/B*		*W/B*
8	*QB*		*QB*	*D/H*		*D/H*
9			*w*	*QB*		*QB*
10			*OC*	*OC*		*OC*
11			*W/B*			*w*

The FSS method added all the parameters to the MLP model for *UCS* prediction, except for *w*. The BSE method removed only *QB*. For the prediction of *E*, the FSS method did not add *w*, *OC*, or *W/B* to the MLP. *W/B* was the only parameter eliminated from the model by the BSE method.

The final selected parameters were different for the RBF and MLP models. However, the parameters selected by both the FSS and BSE methods for the RBF model covered those selected by the MLP model. Also, the ranking of the parameters for the predictions of both *E* and *UCS* were the same for the MLP and RBF networks.

The *QB* is known to have a direct influence on the stiffness and strength of cemented soils. However, the stepwise selection method suggests that *QB* can be removed from the model. By looking at the cross correlation of all the input and target parameters (see Table 4-9), one can observe that there exists a high correlation between *G/S* and *QB* (*r=0.46*), i.e. the *G/S* variable could represent *QB* to some extent. Also, *OC* is a relevant parameter because organic matter reduces the strength of soil. However, the FSS method suggested removing *OC* from the models. This also could be due to the high correlation of *OC* with *G/S* (*r=0.61*). *W/B* and *w* were removed because of their limited significance to *E* and *UCS* prediction in the presence of other relevant variables. Also notice the high negative correlation (-0.81) between these two parameters (for lower soils with initial water content higher ratio of water to binder were used).

Table 4-9. Cross Correlation between the Input and Target Parameters

	Input Parameters																**Target Parameters**	
	OC	***w***	***RB-PC***	***RB-BFS***	***RB-PFA***	***RB-L***	***RB-MgO-C***	***RB-G***	***QB***	***G/S***	***W/B***	***D/H***	***t***	***T***	***RH***	***CO2***	***E***	***UCS***
OC	1.000	0.692	0.318	-0.175	-0.235	-0.165	-0.132	-0.132	0.156	0.606	-0.299	-0.326	-0.071	-0.158	0.107	-0.091	0.081	0.192
w		1.000	0.260	-0.166	-0.165	-0.116	-0.093	-0.093	0.137	0.003	-0.811	-0.286	-0.051	-0.172	0.115	-0.097	-0.191	-0.093
RB-PC			1.000	-0.792	-0.367	-0.416	-0.267	-0.364	0.154	0.183	-0.296	-0.323	-0.067	-0.126	-0.009	-0.020	0.177	0.298
RB-BFS				1.000	-0.181	0.050	-0.102	0.084	-0.129	-0.113	0.247	0.269	-0.002	0.330	-0.108	0.124	-0.069	-0.137
RB-PFA					1.000	0.131	0.599	-0.048	-0.061	-0.120	0.117	0.128	0.097	-0.222	0.144	-0.122	-0.160	-0.269
RB-L						1.000	-0.034	0.954	-0.043	-0.084	0.083	0.091	0.068	-0.158	0.102	-0.087	-0.128	-0.154
RB-MgO-C							1.000	-0.027	-0.034	-0.068	0.066	0.072	0.055	-0.125	0.081	-0.069	-0.081	-0.161
RB-G								1.000	-0.034	-0.068	0.066	0.072	0.055	-0.125	0.081	-0.069	-0.102	-0.109
QB									1.000	0.461	-0.276	-0.194	-0.011	-0.158	0.102	-0.087	0.404	0.592
G/S										1.000	0.224	-0.201	-0.047	-0.117	0.079	-0.067	0.450	0.563
W/B											1.000	0.372	0.020	0.303	-0.196	0.166	0.157	0.017
D/H												1.000	0.022	0.330	-0.214	0.181	0.108	-0.207
t													1.000	-0.035	0.017	-0.010	0.093	0.111
T														1.000	-0.311	0.353	-0.215	-0.291
RH															1.000	0.204	-0.019	0.051
CO2																1.000	-0.082	-0.091
E																	1.000	0.851
UCS																		1.000

Linear Regression (LR) Analysis

In order to compare with ANN models, the stiffness and strength of organic soils were modeled using a linear regression method as well. To that end, the database was divided into two subsets: training and test, with a 4:1 ratio. Similar to the ANN models, the random subsampling method was implemented to generate an ensemble of models (1,000 models) by randomly selecting data points for training and test subsets. The models were trained using the training dataset, and the generalization of the model was evaluated through the test dataset. Table 4-10 presents the performances of the LR ensembles. By comparing these results with the results of the ANNs, one can observe that R^2_{ave} of the LR model is significantly lower than the ANN models for both the training and test subsets. Figure 4-12 presents the distribution of R^2 values of the LR models for the training and test datasets.

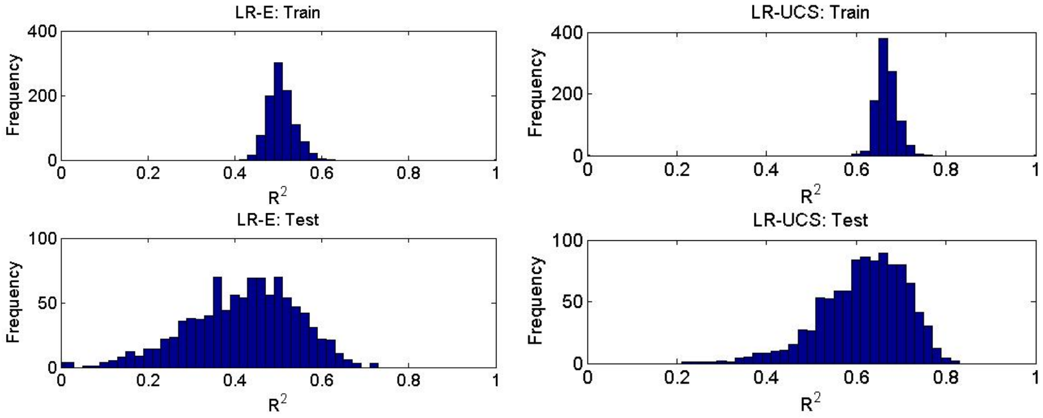


Figure 4-12. Histogram of the R^2 Values Corresponding to the 1000 Networks of the LR Ensembles

Table 4-10. Performance of the LR Ensemble for *E* and *UCS* Predictions

Model	R^2_{ave}			$RMSE_{ave}$		
	All	Training	Test	All	Training	Test
LR-*E*	0.49	0.51	0.42	22.33	22.04	23.48
LR-*UCS*	0.66	0.67	0.62	0.23	0.23	0.25

Comparison of the Models

The performance of the RBF, MLP, and LR models were evaluated and compared through the control database (used also for the stepwise parameter selection). Figure 4-13 presents the measured vs. predicted E and UCS for the three different models. R^2 value is one minus the normalized error of the model predictions with respect to the standard deviation of the data. So in this sense, it provides a unit-less measure for the goodness of a fit. Also, R^2 value can be considered as the fraction of data variance that is fully explained by the model. According to Figure 4-13, for RBF and MLP, R^2 is greater than 0.95, and $RMSE$ of E and UCS are considerably smaller comparing to the LR model. This illustrates that the ANN models outperformed the LR model to a great extent, because the ANN models were able to better capture the nonlinear relations between the input and the output variables.

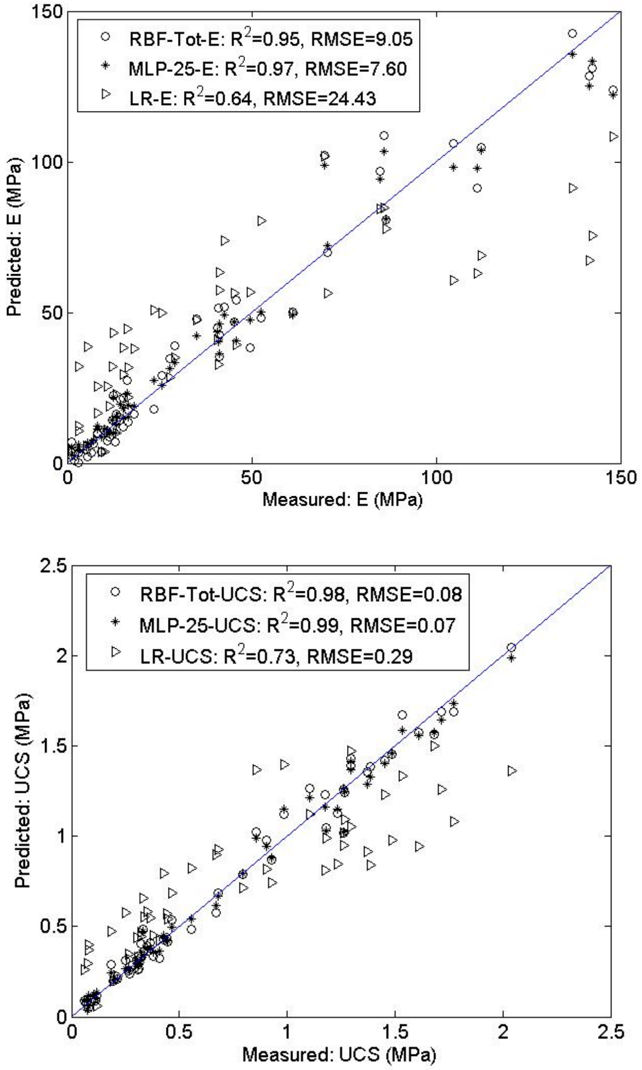


Figure 4-13. Performances of Different Models over the Control Dataset

Response Graphs and Sensitivity Analysis

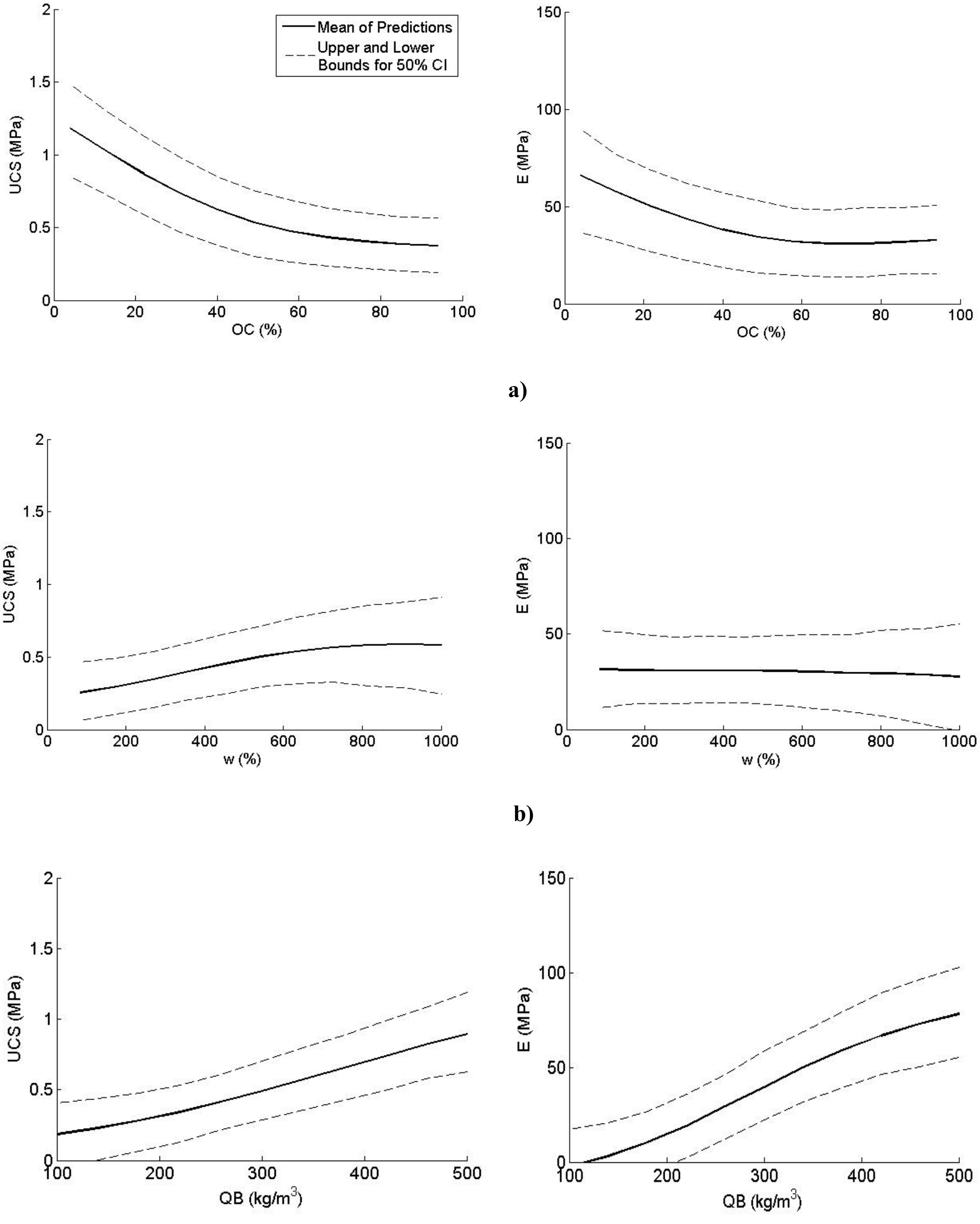
In addition to the stepwise selection method, the effects of input variables on the responses of the ANN models were evaluated using a sensitivity analysis method known as the profile method (Lek et al. 1996). This method not only illustrated the sensitivity of the model's responses to each input variable, but also provided profiles of the model response variations with respect to the input variables.

The MLP-20-*E* and MLP-25-*UCS* models were applied to generate the response graphs. For this purpose, a set of fictitious matrices were constructed corresponding to each of the input variables (except for the binder ratio). For each variable, all other input variables were held constant at their mean value, whereas the studied parameter varied over its range. The range of variation of each input variable was divided into 10 intervals and the model response was evaluated at each point. Note that in this analysis the effect of binder ratio was not explored, because it was not possible to vary one binder type ratio and keep the ratios of the other types constant at the same time (the sum of the binder ratios for the six binder types should be equal to 1). Thus, for each parameter, 11 input vectors were presented to the networks that resulted in 11 responses. It was by those responses that the response graphs were generated.

Figure 4-14 presents the response graphs showing the effects of *OC, w*, *QB*, *G/S*, *W/B*, *D/H*, *t*, *T*, *RH*, and CO_2 on *E* and *UCS*. The response graphs represent the mean prediction of the 1,000 networks included in the MLP ensembles. The upper and lower bounds for the response graphs are also provided, which correspond to the 50% Confidence Interval (CI) of the model predictions. In interpreting these graphs one should consider the mean values at which other input parameters (except for the parameter under analysis) were held constant (see Table 4-2).

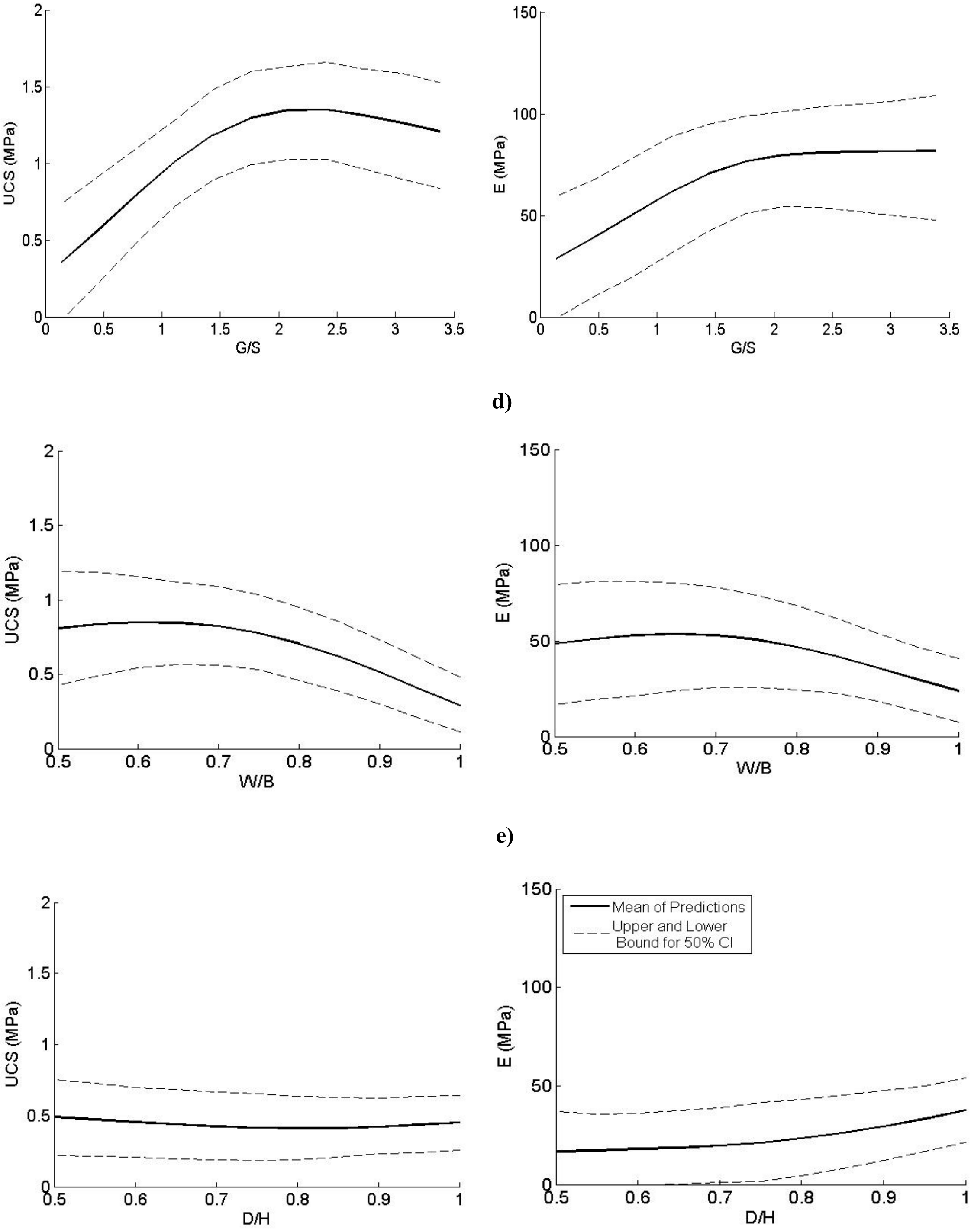
According to Figure 4-14.a, increasing *OC* significantly decreases the *E* and *UCS*. This agrees with the previous studies regarding the negative influence of organic matter on the stiffness and strength of soil (Edil and Den Haan 1994; Hernandez-Martinez 2006; Kruse and Haan 2006; Santagata et al. 2008). Figure 4-14.b indicates that *UCS* increases by *w* up *to w=800%*, and then stays constant up to *w=1000%*. This result is rather surprising because generally an increase in water content causes a decrease in strength. Exceptions to this general trend have occurred, primarily in the field when the initial water content is very low. Also, this can occur when adding large quantities of dry binders that prevent the proper mixing and hydration and result in lower strength (Åhnberg 2006). However, the results obtained herein could be due to different reasons such as the mixing ratio, binder quantity, or binder type. *E* showed a slight decrease as *w* increased. Figure 4-14.c illustrates the effect of *QB* on the *E* and *UCS*, indicating that the stiffness and strength of the cement-stabilized soils grew by increasing the quantity of binder, a trend which is well established. Figure 4-14.d illustrates the significant effect of *G/S* on the *E* and *UCS*, as was also demonstrated in the ranking of parameters. As expected, increasing the ratio of grout to soil developed a higher stiffness and strength. The optimum range for *G/S* was between 2 and 2.5, according to this figure. Figure 4-14.e indicates that the optimum range for *W/B* was between 0.6 to 0.7, and larger values of *W/B* caused a considerable decrease in *E* and *UCS*. Figure 4-14.f illustrates that *UCS* decayed and *E* grew slightly as *D/H* increased. Figure 4-14.g indicates that the age of the specimen is positively correlated with *E* and *UCS* for stabilized organic soils. It was observed that *E* and *UCS* grew fast during the first 90 days, the time in which the majority of the hydration process developed. After that, the rate of increase decayed. Figure 4-14.h indicates that the *E* and *UCS* of stabilized organic soils decayed by increasing the temperature (*T*). This result is completely opposite of the result obtained in common stabilized mineral soils where temperature accelerated the strength and improved the stiffness. The negative effect of temperature on the stiffness and strength of organic soils is related to several factors: i) increasing the temperature causes a gradual loss of the evaporable water initially present in the fresh mix, ii) it also causes dehydration in the chemical reactions of hardened cement paste and the conversion of calcium hydroxide into calcium oxide, iii) it causes changes in porosity, and iv) peat loses its evaporable water due to temperature increase (Hernandez-Martinez, 2006). Figure 4-14.i indicates that $RH=0.8$ was the optimum value for both *E* and *UCS*. According to Figure 4-14.j, carbonation did not have a significant effect on the stiffness and strength of the stabilized organic soils. However, a slight developing trend was observed, specifically for *UCS*, with respect to CO_2.

The response graphs for the different input variables are plotted versus the scale of variation (1 to 11) in Figure 4-15. Comparing the response graphs and the range of variation of *E* and *UCS* in the domain of each variable, one can calculate the ranking of the input parameters. This ranking represents the order of the input parameters with respect to their influence on the stiffness and strength of stabilized organic soils, as presented in Table 4-11.



c)

Figure 4-14. Response Graphs Depicting the Variation of *E* and *UCS* with respect to the Input Parameters

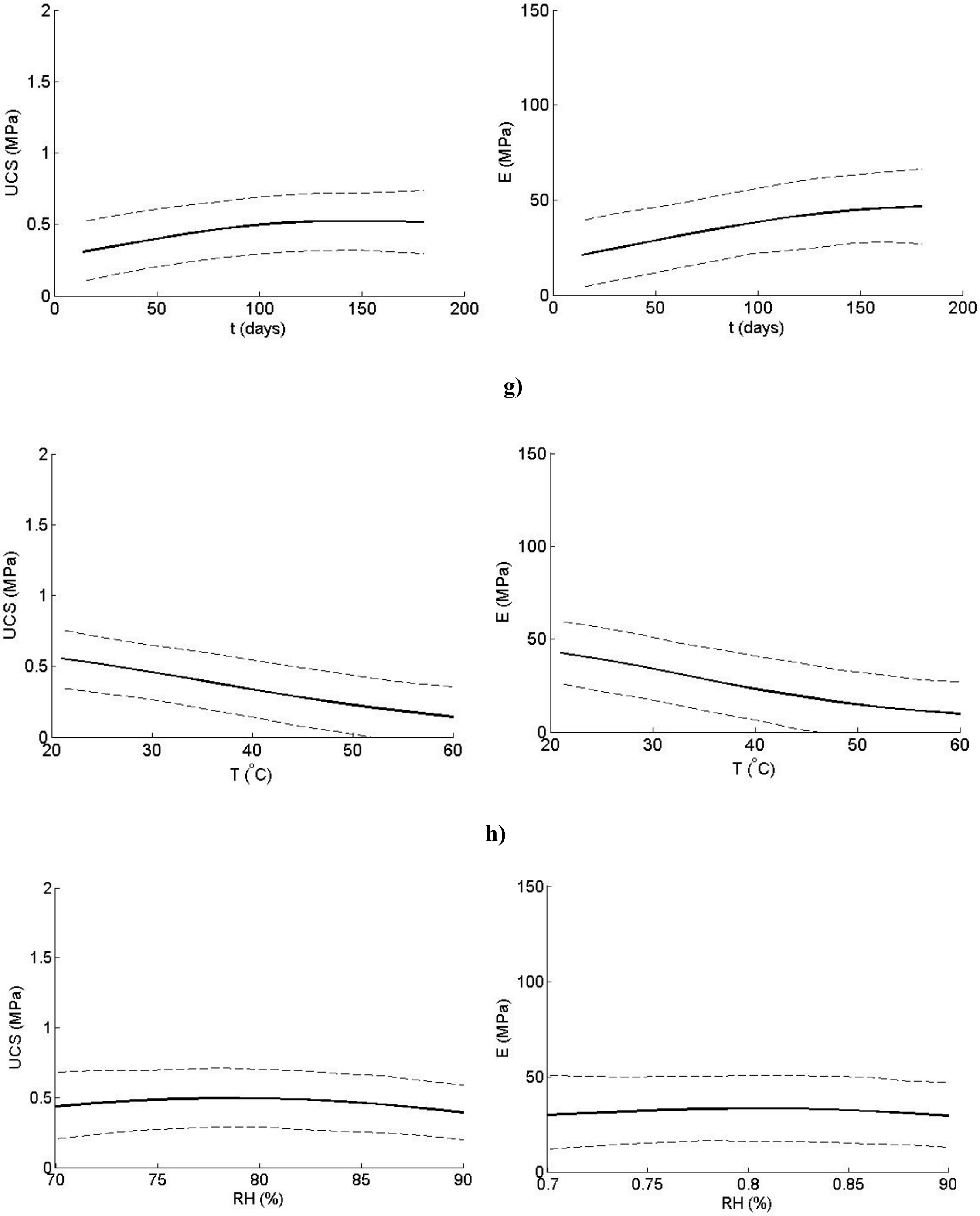


d)

e)

f)

Figure 4-14. Continued

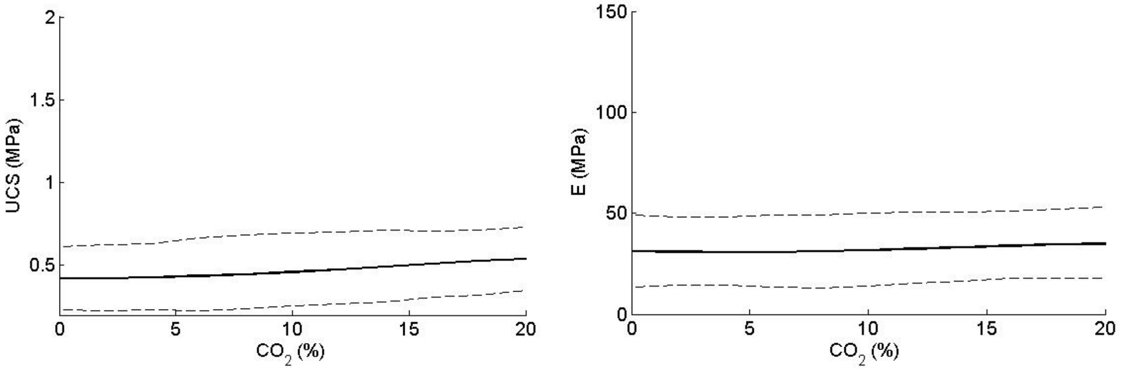


g)

h)

i)

Figure 4-14. Continued



j)

Figure 4-14. Continued

Table 4-11. Ranking of Parameters based on the Sensitivity Analysis

Ranking	Parameter	Max. *E* Variation (MPa)	Parameter	Max. *UCS* Variation (MPa)
1	*QB*	80.59	*G/S*	0.99
2	*G/S*	53.07	*OC*	0.81
3	*OC*	34.97	*QB*	0.71
4	*T*	32.91	*W/B*	0.55
5	*W/B*	29.93	*T*	0.41
6	*t*	25.61	*w*	0.33
7	*D/H*	20.88	*t*	0.22
8	*CO2*	4.25	*CO2*	0.12
9	*RH*	3.94	*RH*	0.10
10	*w*	3.74	*D/H*	0.08

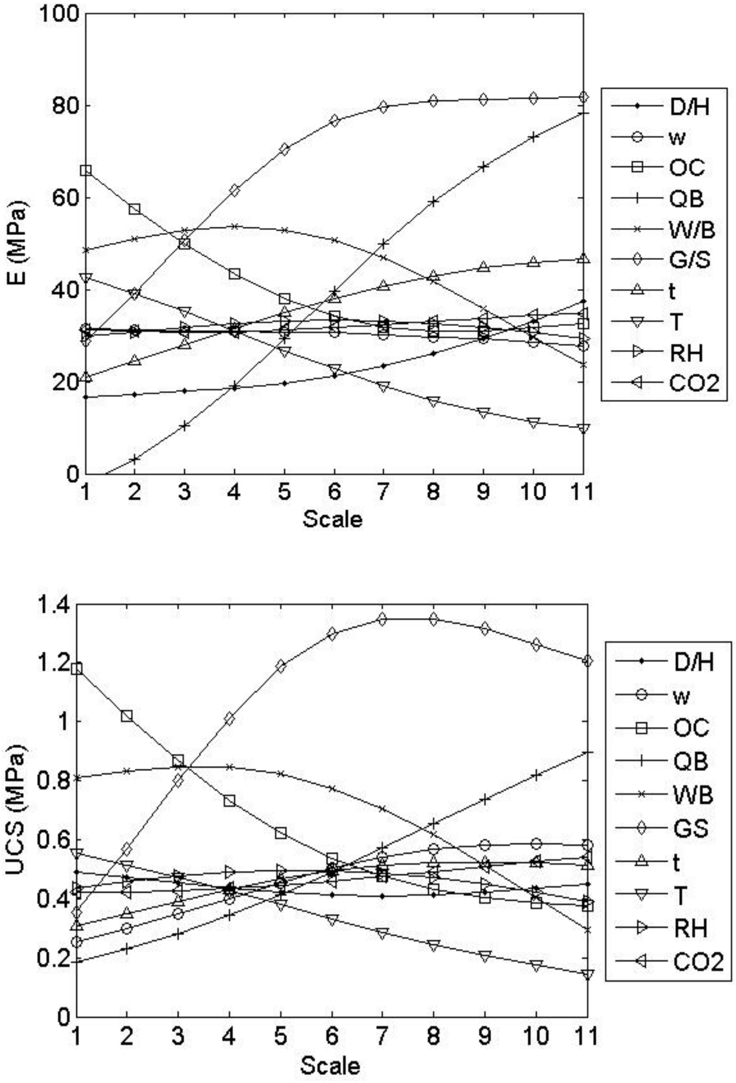


Figure 4-15. Response Graphs for Different Input Variables with respect to the Scale of Variation

Conclusions

This study shows how ANN modeling can successfully be applied to develop predictive nonlinear models of mechanical properties of stabilized organic soils. Two ANN models, RBF and MLP were generated to predict *E* and *UCS* for stabilized organic clays with different organic contents and a variety of binder mixes. A linear regression model was also developed to serve as a benchmark model. It was observed that the performance of the MLP model was slightly better than that of the RBF model over the control dataset. However the RBF, on average, showed better performance over the test dataset than the MLP. Also, both of the ANN models outperformed the LR model significantly. The uncertainty of the model predictions was assessed by generating an ensemble of networks using the random subsampling method.

ANN models predicted *E* with $R^2 > 0.8$ and *UCS with* $R^2 > 0.9$, after learning from the relationship between the input and the output parameters. Also, R^2 of the ANN models was higher for prediction of *UCS* than *E*. This was related to other relevant factors that might have affected *E* and have not been considered in our model (as a predictor variable), possibly factors controlling the level of deformation developed upon failure. Results of the RBF analysis for the three subsets of the database showed that the experimental design could significantly affect the performance of ANN models developed for the prediction of response variables.

The stepwise parameter selection method was implemented for the ANN models to identify the most relevant parameters and to eliminate the redundant parameters from the models. Also, a sensitivity analysis was performed using the MLP models that illustrated the effect of each input parameter on *E* and *UCS*. The generated response graphs demonstrated that ANNs can account for the nonlinear relations between the response and the input variables. Based on their significance to *E* and *UCS*, the input variables were ranked using both the stepwise selection and the sensitivity analysis methods.

According to the results, grout to soil ratio, quantity of binder, binder mix type, amount of organic matter, water to binder ratio, temperature, and time (aging) are the most influential parameters on *E* and *UCS*. Also, initial water content is relevant for *UCS* to some extent; however, it seems to have a negligible effect on *E*. The sample size barely affected *UCS*, whereas it had a considerable influence on *E*. Carbonation and relative humidity seemed to have relatively insignificant effects on both *E* and *UCS*.

CHAPTER V

DETERMINATION OF UNKNOWN FOUNDATIONS FOR BRIDGE SCOUR: A DETERMINISTIC APPROACH

Introduction and Literature Review

Following the FHWA requirement to inspect and monitor all bridges across the nation, the Departments of Transportation (DOTs) across the United States implemented different measures to determine bridges' vulnerability to failure (FHWA 2011). Through this process, a considerable number of bridges were found to have no available foundation information. Item 113 (Scour-Critical Bridges) of the National Bridge Inventory (NBI) uses a single digit to identify the current status of a bridge regarding its vulnerability to scour. This digit varies between 0 and 9, with 0 being the most critical and 9 the least. For Unknown Foundation Bridges (UFBs), the scour risk cannot be determined and a code of U is used (Weseman 1995). DOTs have implemented different measures to reclassify as many UFBs from the NBI as possible by updating Item 113.

As of 2005, approximately 60,000 bridges throughout the United States were identified as having unknown foundations. Most of these bridges were built between 1950 and 1980, which coincided with the construction of the interstate system. Majority of these bridges were classified as off-system bridges that counties and local agencies built and owned, and later were inherited by a DOT. Therefore, it was expected that UFBs would mainly be located on local roads. The numbers of UFBs across the United States, as well as their distribution over the years are provided in Table 5-1.

Table 5-1. Number of Unknown Foundation Bridges across the United States (Stein and Sedmera 2006)

	Rural Functional Classification						Urban Functional Classification						
	1	2	6	7	8	9	11	12	14	16	17	19	
State	**Principal Arterial — Interstate**	**Principal Arterial-Other**	**Minor Arterial**	**Major Collector**	**Minor Collector**	**Local**	**Principal Arterial — Interstate**	**Principal Arterial — Other**	**Minor Arterial**	**Major Collector**	**Minor Collector**	**Local**	**Totals**
Alabama	4	70	79	503	843	1,662	4	5	25	39	55	164	**3,453**
Alaska	7	29	4	19	23	46	1	0	1	7	1	13	**151**
Arizona	0	0	0	1	0	33	0	0	1	3	5	25	**68**
Arkansas	0	1	11	48	4	10	0	0	1	2	0	2	**79**
California	4	23	112	318	305	993	3	9	71	84	60	126	**2,108**
Colorado	1	2	9	4	3	8	0	0	0	1	1	0	**29**
Connecticut	0	0	0	0	0	0	0	0	0	0	0	0	**0**
Delaware	0	0	0	0	0	0	0	0	0	0	0	0	**0**
DC	0	0	0	0	0	0	0	0	0	0	0	8	**8**
Florida	3	110	111	224	188	837	13	27	74	136	280	444	**2,447**
Georgia	3	346	434	1,227	565	1,780	0	32	178	288	188	406	**5,447**
Hawaii	0	0	0	0	0	0	0	0	0	2	0	8	**10**
Idaho	0	1	1	71	74	318	0	0	3	6	9	14	**497**
Illinois	0	0	0	1	0	1	0	0	0	0	0	0	**2**
Indiana	0	1	0	140	263	828	0	0	42	101	75	156	**1,606**
Iowa	0	1	3	92	256	1,371	0	0	0	11	6	30	**1,770**
Kansas	0	0	0	0	1	5	0	0	0	0	0	0	**6**
Kentucky	0	0	0	0	0	1	0	0	0	0	0	1	**2**
Louisiana	17	13	180	527	488	2,963	12	1	30	84	58	401	**4,774**
Maine	6	2	1	4	3	76	2	0	0	2	5	4	**105**
Maryland	0	0	0	0	0	4	0	0	0	0	0	1	**5**
Massachusetts	2	0	10	25	16	70	0	1	42	95	45	52	**358**
Michigan	3	36	43	157	13	360	2	2	9	10	11	11	**657**
Minnesota	0	0	2	16	24	161	0	0	0	4	2	7	**216**
Mississippi	0	16	11	1,205	187	4,790	0	0	32	54	101	137	**6,533**
Missouri	0	0	1	6	1	0	0	0	0	0	0	0	**8**
Montana	2	1	5	1	429	1,244	0	0	0	1	0	2	**1,685**

Table 5-1. Continued

	Rural Functional Classification						Urban Functional Classification						
	1	2	6	7	8	9	11	12	14	16	17	19	
State	**Principal Arterial — Interstate**	**Principal Arterial — Other**	**Minor Arterial**	**Major Collector**	**Minor Collector**	**Local**	**Principal Arterial - Interstate**	**Principal Arterial — Other**	**Minor Arterial**	**Major Collector**	**Minor Collector**	**Local**	**Totals**
Nebraska	0	0	0	0	0	0	0	0	0	0	0	0	**0**
Nevada	0	0	2	1	3	24	0	0	1	10	1	3	**45**
New Hampshire	0	0	0	3	6	22	0	0	2	5	4	1	**43**
New Jersey	0	6	7	11	7	53	0	4	20	23	20	14	**165**
New Mexico	0	7	7	46	41	254	1	0	13	27	39	33	**468**
New York	0	0	0	1	1	13	0	2	7	9	4	12	**49**
North Carolina	0	29	95	464	700	3,949	0	2	30	81	77	379	**5,806**
North Dakota	0	0	3	210	0	1,780	0	0	0	5	3	7	**2,008**
Ohio	0	2	1	13	23	222	0	0	2	1	5	12	**281**
Oklahoma	0	0	9	1	1	9	1	2	5	0	0	0	**28**
Oregon	5	58	90	425	235	801	4	2	18	50	51	56	**1,795**
Pennsylvania	0	0	0	0	1	7	0	0	0	0	0	0	**8**
Rhode Island	0	0	0	0	0	0	0	0	0	0	0	0	**0**
South Carolina	21	49	125	592	443	1,904	6	0	20	49	96	144	**3,449**
South Dakota	0	0	0	1	0	0	0	0	0	0	0	0	**1**
Tennessee	6	8	32	74	252	654	0	0	8	27	24	73	**1,158**
Texas	9	18	40	199	190	6,524	2	4	205	463	319	1,140	**9,113**
Utah	0	0	0	1	0	4	0	0	1	0	0	2	**8**
Vermont	0	2	5	29	26	155	0	0	2	4	9	6	**238**
Virginia	0	0	0	0	2	16	0	0	0	0	0	0	**18**
Washington	0	0	1	47	39	102	0	0	5	4	3	5	**206**
West Virginia	0	0	0	0	0	0	0	0	0	0	0	0	**0**
Wisconsin	0	0	0	0	0	0	0	0	0	0	0	0	**0**
Wyoming	0	0	1	0	43	347	1	0	0	3	7	13	**415**
Puerto Rico	0	0	21	70	40	77	0	0	9	23	36	36	**312**
Totals	93	831	1,456	6,777	5,739	34,478	52	93	857	1,714	1,600	3,948	**57,638**

Texas has the largest population of UFBs in the United States (9,113). Among those, only 33 (0.4%) are located on principal arterials. By contrast, the state of Alaska has only 151 UFBs, but a significantly higher percentage of them are located on principal arterials (37 bridges or 25%). Although it is expected to find very few UFBs built in recent years, 69 UFBs located on principal arterials were built between 2000 and 2005 (Stein and Sedmera 2006).

Researchers have tackled the problem of unknown foundations in the United States with different approaches and, occasionally, with a combination of different methods. There are a number of studies on applying nondestructive testing (NDT) methods to obtain information about UFBs (Maser et al. 1998; Olson et al. 1998; Mercado and O'Neil 2003; Olson 2005; Robinson and Webster 2008), while only few studies have used statistical, computational, and numerical solutions to infer the foundation’s characteristics. Given the uncertainty of the results generated by these methods and the uncertain nature of scour, risk-based approaches seem appropriate for performing a more rational assessment of unknown foundations. National Cooperative Highway Research Program (NCHRP) (Stein and Sedmera 2006) and Florida DOT (Mclemore et al. 2010) implemented a risk-based approach to manage bridges with unknown foundations considered susceptible to scour. The proposed guideline was reported to identify successfully the required Plan of Actions, including investigation, countermeasures, monitoring, and rehabilitation or replacement of UFBs, based upon their scour evaluation.

Soft computing methods in general, and ANNs in particular, have been widely applied for infrastructure safety monitoring and management, including bridge safety and integrity assessment in the recent years (Elkordy et al. 1994; Kim et al. 2000; Adeli 2001; Kawamura et al. 2003; Arangio and Bontempi 2010). A number of studies have used ANNs to infer scour depth at bridge piers (Bateni et al. 2007; Lee et al. 2007; Zounemat-Kermani et al. 2009; Kaya 2010), but only a few have used ANN analyses to infer the type and dimension of the foundation (Mclemore et al. 2010; Sayed et al. 2011). Also, some studies have explored the use of ANNs to interpret the results of NDTs for determination of UFBs (Rix 1995).

Approximately 85% of UFBs in the state of Texas are categorized as off-system bridges. The bridge plans for most of these bridges are missing, and their foundation characteristics are unknown (i.e., type and dimensions of foundation). On-system bridges, on the other hand, are defined as bridges designed and constructed by DOTs whose design and construction records are available.

This study was conducted as part of a TxDOT sponsored project (Briaud et al. 2012), with the aim of providing a methodology for the deterministic evaluation of unknown foundations, which can facilitate the assessment of scour failure risk for UFBs. An evidence-based approach is proposed based on a database of bridges located in the TxDOT's Bryan District. This approach can be further modified and extended to be implemented in other districts in Texas, and even in other states. Notice that the proposed models were calibrated based on bridges' data collected from the Bryan District; therefore they cannot be applied to other regions due to site dependency, unless they get trained by corresponding bridges' data collected from those specific regions. The key hypothesis of this evidence-based approach was that the collected database was a representative sample of the Bryan District's population of bridges. It is assumed that the information retrieved for the known bridges in the database captured the design and construction criteria, as well as the soil characteristics in the region.

The deterministic method applies ANNs to learn from the inherent patterns in the foundation design for the bridges in the database, and further generalize these patterns to UFBs. This method allows engineers to make predictions regarding foundation type and embedment depth with a reasonable level of accuracy and avoid performing time consuming, costly experiments.

Approach

Database and Statistics

There are approximately 1,700 bridges in the Bryan District. Among those, 1,100 are built over rivers and streams and are exposed to scour failure. A set of meetings were held in the Bryan District office to investigate the possibility of collecting data for bridges in this district. Data collection was then started by finding bridge inspection folders and scanning the relevant documents such as bridge plans, bridge inspection sheets, etc. A typical bridge plan located in the Bryan district is presented in Figure 5-1.

Data were collected for approximately 40% (about 185) of the on-system bridges (511 bridges in total) built over rivers with known foundations. These bridges were sampled from four populations of bridges with major foundation types (drilled shafts

[DrSh], concrete piling [Conc], steel piling [Steel], and spread footing [Spread]) in compliance with their actual distributions throughout the Bryan District.

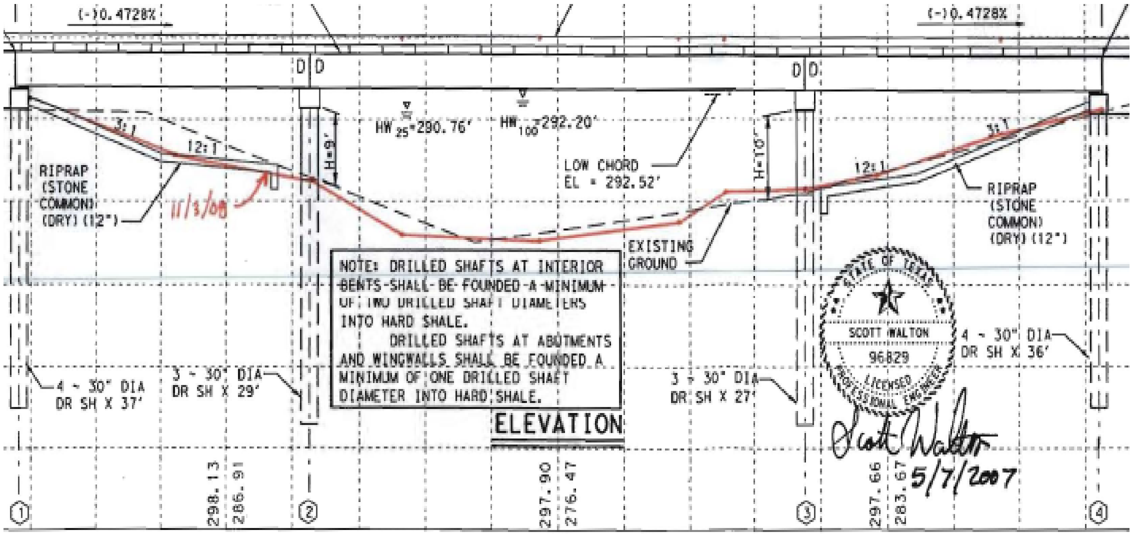


Figure 5-1. A Typical Bridge Plan in the Bryan District, Showing the Original and Most Recent Ground Profile

The statistics of the population of on-system bridges in the Bryan District showed that the most frequent type of bridge superstructure was concrete slab with girders (48.7%), followed by flat concrete slab (22.3%). It was found that the most common substructure below the ground (foundation) was concrete piling (54%), and above the ground was pile bent (66.3%). Table 5-2 shows the number and percentage of each type of substructures for on-system over-river bridges in the Bryan District. It can be observed that very few of the over-river bridges had spread footings, due to their susceptibility to scour. Also, it can be noticed that the number of driven piles (concrete piling) was more than twice of the number of drilled shafts.

Table 5-2. Percentage of Different Types of Substructures

Substructure Type Above Gr.	%	Substructure Type Below Gr.	%	Cap Type	%
Pile Bents	66.3	Steel Piling	12.1	Concrete Cap	95
Single Column Bent	0.0	Concrete Piling	54.0	Steel Cap	0
Multiple Column Bent	24.7	Timber Piling	0.6	Timber Cap	0
Concrete Column Bent with Tie Beam	0.4	Drilled Shaft	21.5	Masonry Cap	0
Concrete Column Bent Wall	0.2	Spread Footing	3.5	Other	0
Concrete Pier	2.7	Pile Cap on Steel Piling	2.0	NA	5
Masonry Pier	0.4	Pile Cap on Concrete Piling	0.8		
Trestle (all)	0.0	Pile Cap on Timber Piling	0.2		
Other	0.0	Other	0.0		
NA (Single Span Bridges)	5.3	NA	5.3		
Sum	**100**	**Sum**	**100**	**Sum**	**100**

In order to demonstrate the geographic risk for these bridges, a set of GIS maps were developed. Figure 5-2 demonstrates the location of the bridges in the Bryan District of the TxDOT included in this study, along with their Scour-Critical indices (Item 113), Average Annual Daily Traffic (AADT), year built, and population density. Note that the scour-critical index decreases as the scour vulnerability increases. Only three bridges were identified as scour critical, illustrated with red or yellow coloring in the map. The two most critical bridges were located in Burleson where the population density and therefore, the risk of life loss were relatively lower. Also, it was observed that the three scour-critical bridges were built before 1960. According to these maps, the higher the age of the bridge, the lower the scour-critical index and, thus, the more susceptible the bridge is to scour failure.

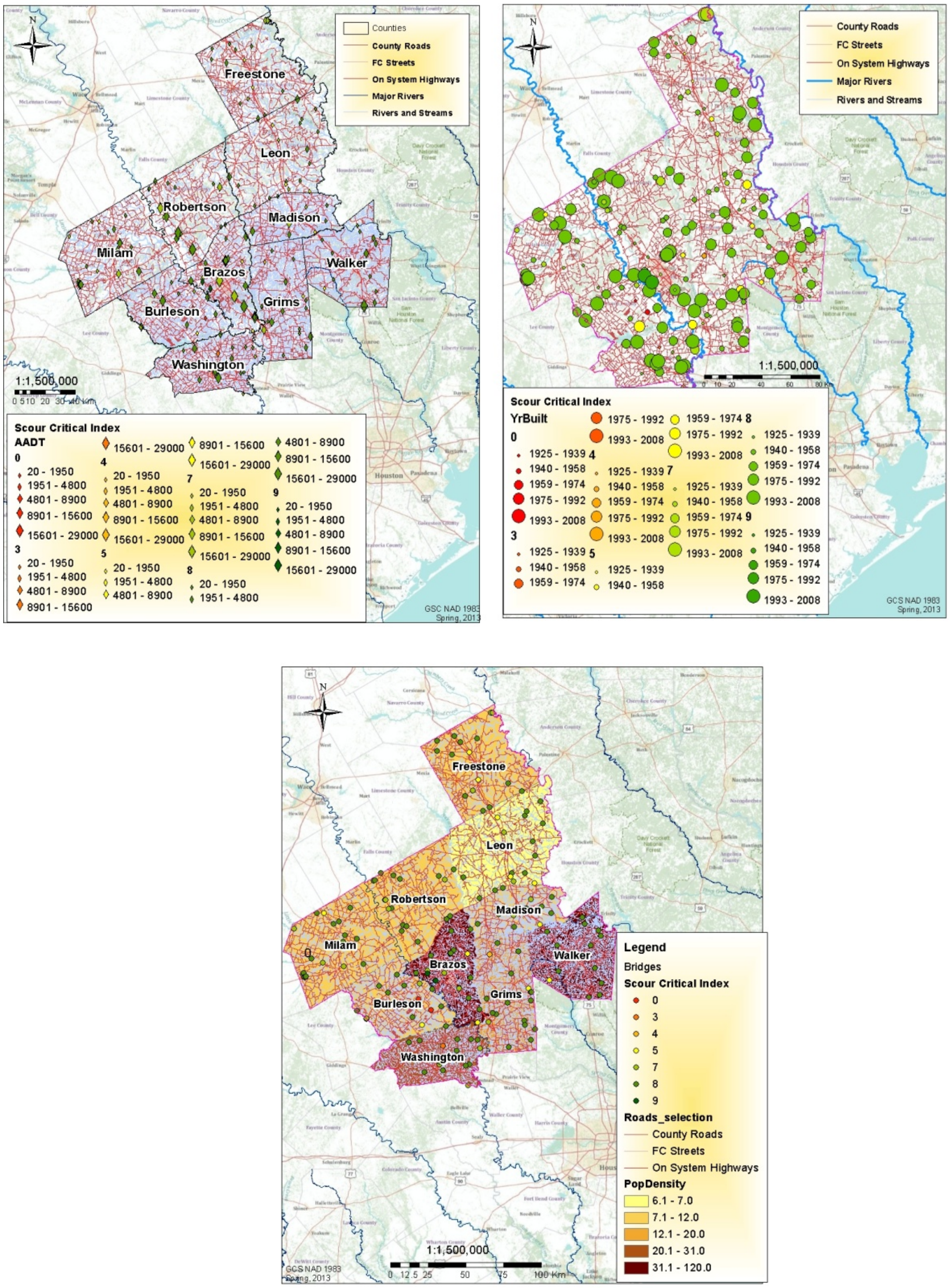


Figure 5-2. Bridges in the Bryan District of TxDOT Collected for This Study, Showing the Scour-Critical Index, AADT, Year Built, and the Population Density over the Counties

For each interior bent of the bridges, the relevant data for a representative pile was collected. The relevant information for inferring the type and embedment depth of each foundation was identified based on the design procedures in the AASHTO and TxDOT's bridge standard manuals (AASHTO 2002a, b; TxDOT 2006).

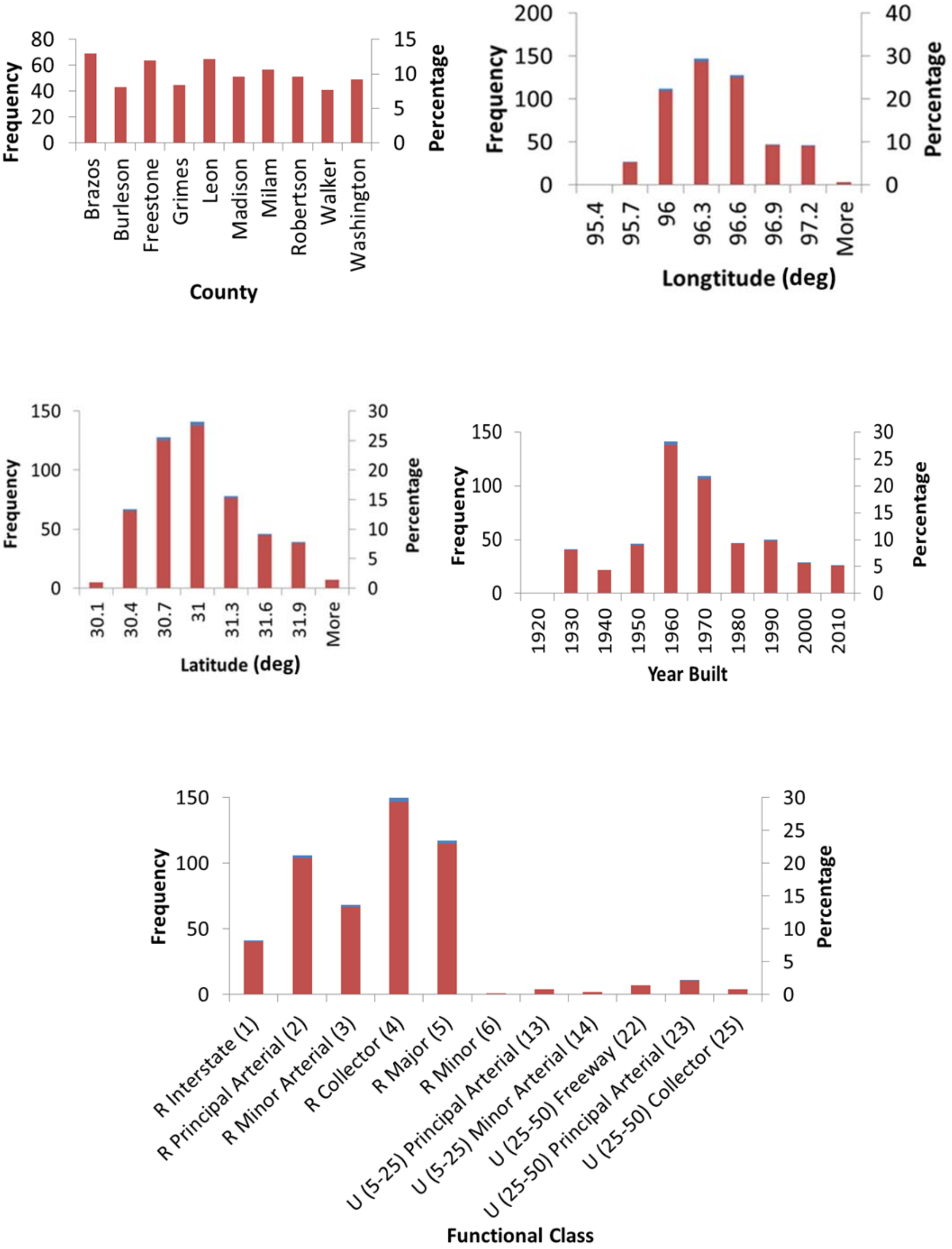
The data were found in the inspection folders and from the inspection database (also a part of NBI) provided by TxDOT. Some information was directly extracted from the design and construction plans found in the bridge inspection folders. The data that were collected for the 185 (out of the 511) on-system bridges in the Bryan District formed the dataset for foundations' embedment depth estimation. However, the whole database (511 bridges) was used for the purpose of foundation type prediction.

The populated database provided the parameters required for training the predictive models. These parameters include superstructure elements, substructure elements, load elements, and soil properties. Note that each of the bridges had one or more interior bents and thus the total number of records in the database was around 580. Table 5-3 presents the main items in the database. The histograms for some of these items are presented in Figure 5-3. Some of these items were used for pile embedment length estimation and some other for foundation type prediction, as will be explained in the following sections.

The soil boring data found in the bridge inspection folders provided the type and strength of soil based on results of Texas Cone Penetration (*TCP*) tests. A typical soil boring data sheet is provided in Figure 5-4. The ultimate skin friction (f_u) along the piles was computed from the *TCP* values using the correlation graphs (Figure 5-5) recommended by TxDOT (TxDOT, 2006). Note that for *TCP >100*, in the soil boring sheets the penetration depth for 100 blows was reported. However, for the purpose of modeling, the equivalent value of *TCP* for one foot of penetration was calculated. So the range of the *TCP* values in the database goes beyond 100.

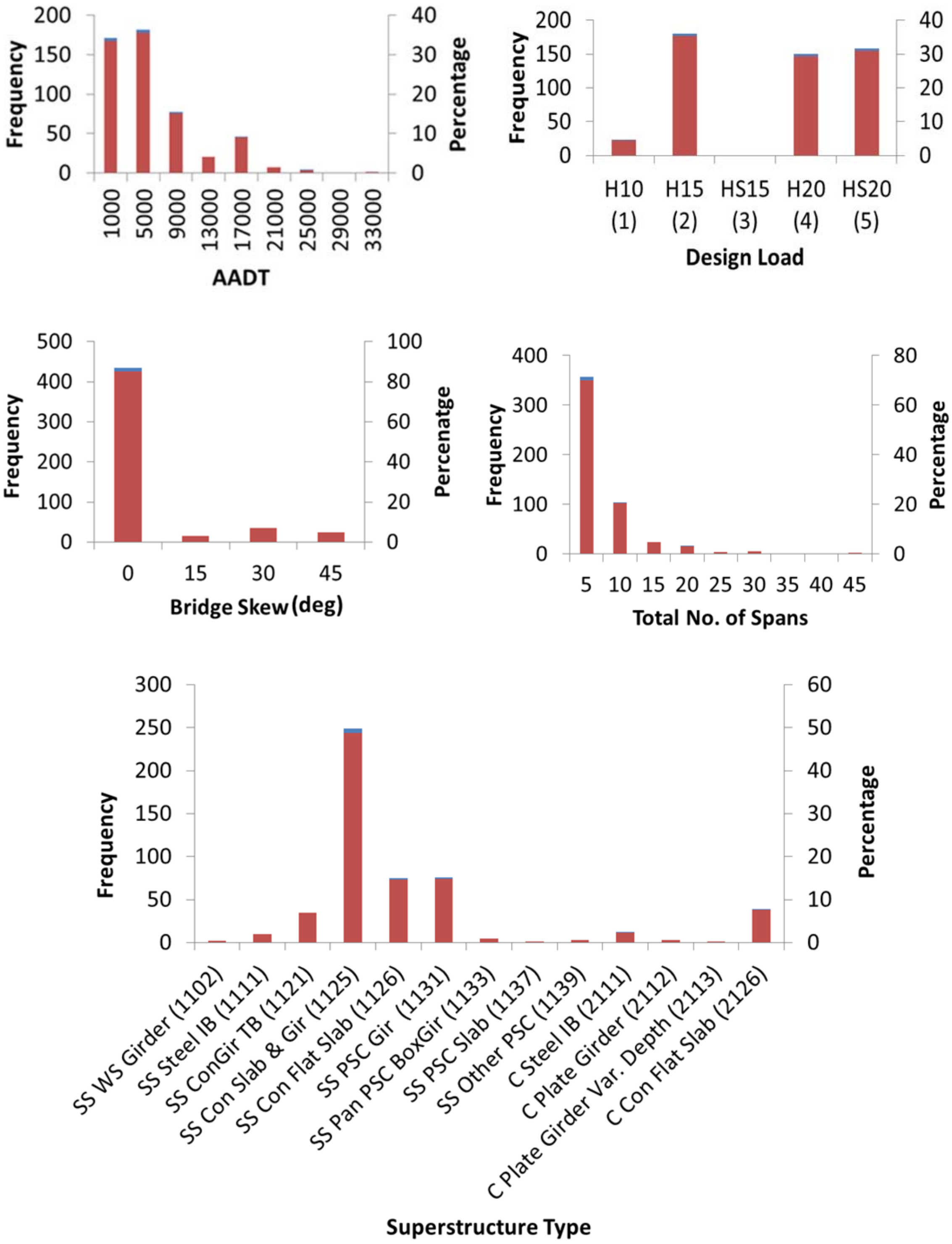
Table 5-3. List of the Main Parameters in the Database Used in the Models

NBI Item/Parameter	Min	Max	Mean	Median	Standard Deviation
Item 3: County	21	239	131.38	145	73.54
Item 16: Latitude (deg)	30.06	31.97	30.88	30.85	0.45
Item 17: Longitude (deg)	95.4	97.25	96.27	96.25	0.4
Item 26: Functional Class	1	25	4.15	4	3.8
Item 27: Year Built	1922	2009	1964	1961	20
Item 29: AADT	20	31070	4470.33	1900	5445.38
Item 31: Design Load	1	5	3.44	4	1.42
Item 34: Bridge Skew (deg)	0	45	4.59	0	12.2
Item 43-1: Main Span Type	1102	2126	1232.86	1125	309.03
Item 44-1-D1: Substructure Type above Ground for Main Span	1	7	1.71	1	1.22
Item 44-1-D2: Substructure Type below Ground for Main Span	1	8	2.58	2	1.29
Item 46: Total Number of Spans	1	44	5.4	4	4.93
Item 48: Max Span Length (m)	10	240	39.73	30	25.46
Item 51: Roadway Width (m)	19	90	34.45	33.5	12.29
Item 71: Waterway Adequacy	1.12	9	6.55	6	1.16
Item 107-1: Deck Type for Main Span	1	2	1.06	1	0.24
Average *TCP* along Pile (Avg. *TCP*) (blows/ft)	10.5	1414.75	134.47	98.7	141.81
***TCP* at Pile Tip (blows/ft)**	10	2400	291.36	228.57	299.91
Skin Friction (f_u) (MPa)	0.03	1.94	0.21	0.17	0.17
Point Bearing Capacity (q_p) (MPa)	0.10	24.59	2.88	2.14	2.99
Dead Load (kN)	62.41	10328.36	1997.12	1317.68	1888.62
Live Load (kN)	229.65	1324.35	474.41	411.53	250.55
Pile Embedment Depth (d_p) (m)	3.05	26	9.35	8.99	3.55



*R: Rural, U(X): Urban (X1000 population)

Figure 5-3. Histograms for the Items and Parameters in the Database



* C: Continous; Con:Concrete; PSC:Prestressed; SS: Single Span; Var: Variable; WS: Weathering Steel

Figure 5-3. Continued

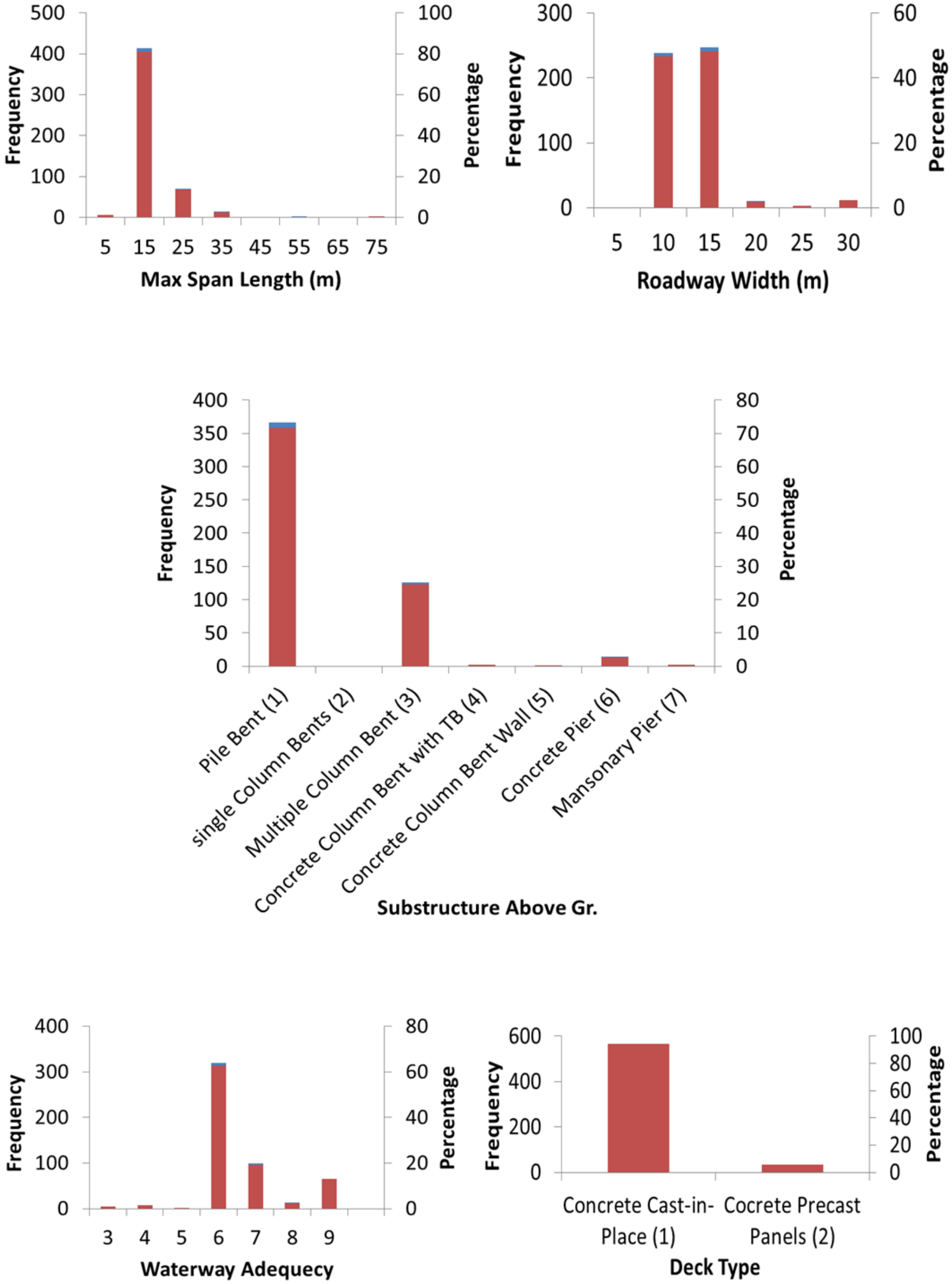


Figure 5-3. Continued

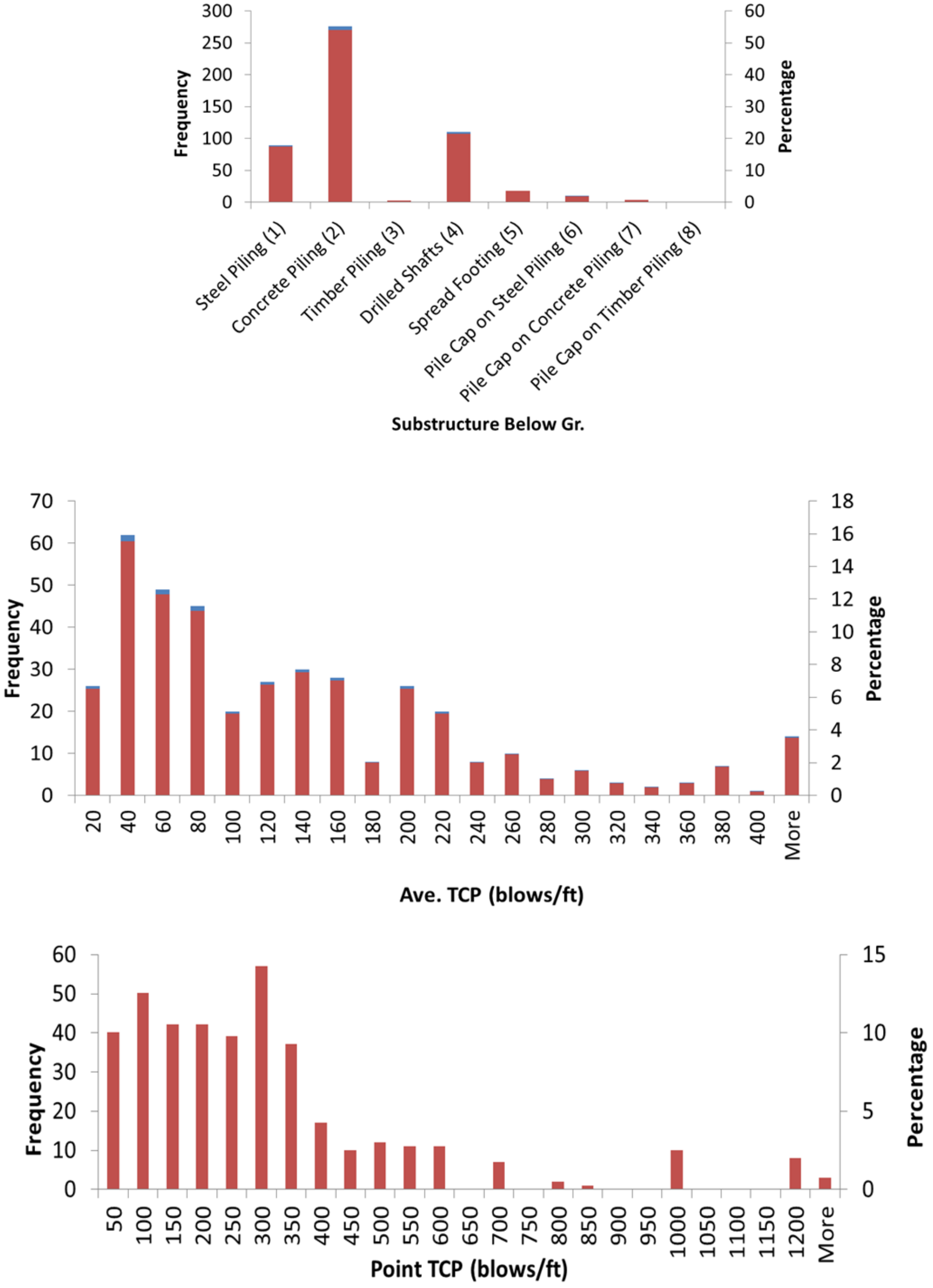


Figure 5-3. Continued

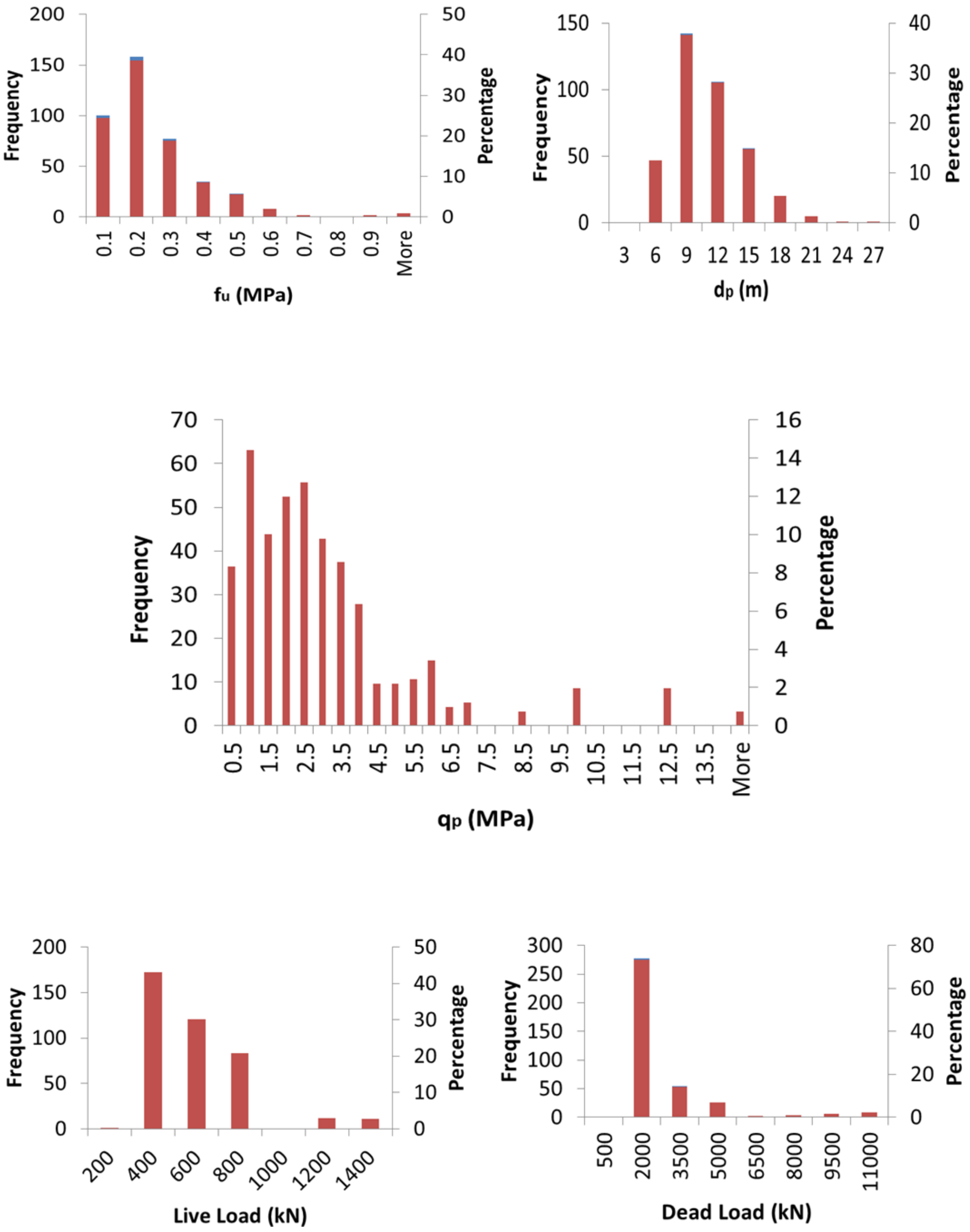


Figure 5-3. Continued

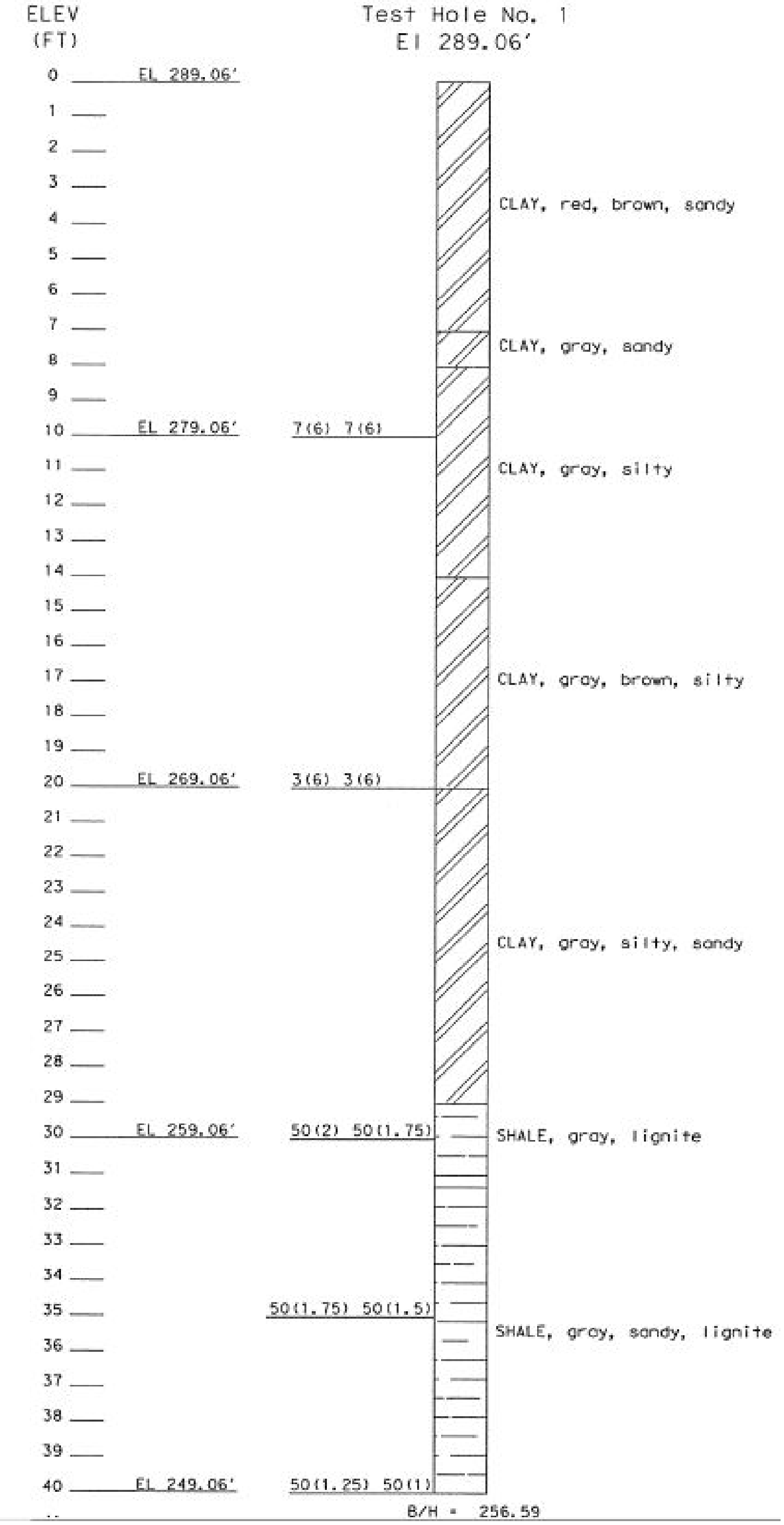


Figure 5-4. A Typical Soil Boring Data Sheet

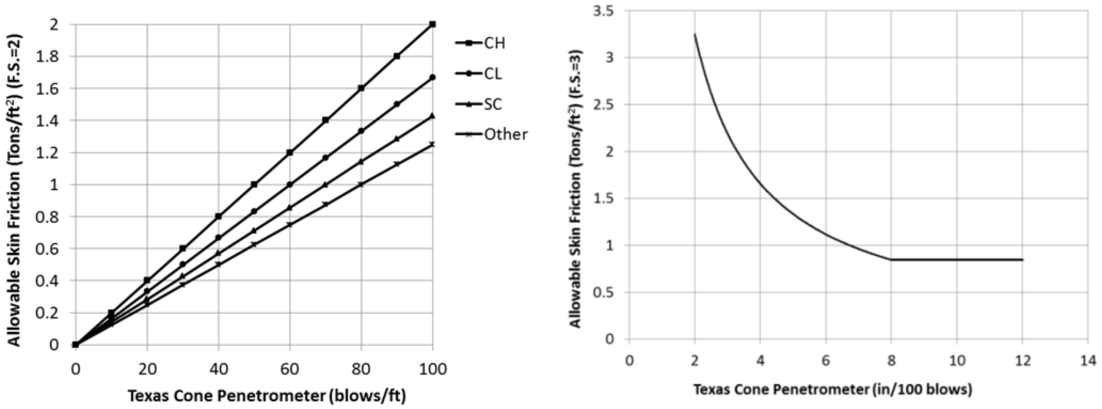


Figure 5-5. Skin Friction versus *TCP* for Left) *TCP < 100* blows/ft, Right) *TCP > 100* blows/ft (TxDOT 2006)

Methodology

Figure 5-6 presents the flowchart for the proposed deterministic method. The deterministic approach first determines the type of a foundation and whether the foundation is deep or shallow. It then follows to determine the embedment of the foundation, if it is classified as a deep foundation (i.e., for piling or drilled shafts). The method selects the appropriate ANN model (based on the foundation type and availability of soil boring data) to predict the depth of the foundation. Therefore, this method allows for calculating the scour risk of the foundation by determining if the foundation is shallow or otherwise determining the embedment depth of the foundation. The steps shown in Figure 5-6 are:

Step 1: Data Collection. Acquire the required parameters, as listed below. These items can be found in the NBI or in bridge inspection documents. More details on the description of each item can be found in the NBI's coding guide (Weseman 1995):

- Item 3: County

- Item 16: Latitude
- Item 17: Longitude
- Item 27: Year built
- Item 31: Design load
- Item 44.1-D1: Substructure type above ground for main spans
- Item 46: Total number of spans in bridge
- Item 48: Max span length
- Item 51: Roadway width
- Average TCP and skin friction along the piles (from soil boring data, if available)
- Average span length for the bent (from bridge inspection documents)

Step 2: Determining the Foundation Type Using ANNs. The type of foundation is predicted using two ANN classifier models. The ANN models were trained with a number of bridge examples to learn the existing patterns in the relation between bridges' input parameters and their foundation type. One of the proposed ANN classifiers was designed to predict any of the common foundation types, whereas the other one can be used solely to distinguish between deep or shallow foundations. The second model is recommended in a condition when an engineer can identify the foundation type from the substructure above ground with a high confidence level. This model also allows for identifying deep foundations without knowing exactly what the foundation type is (whether it is a concrete pile, steel pile, or a drilled shaft).

Step 3: Determining the Depth of the Hard Layer. If soil boring data is available adequately close to the foundation, then it is recommended to identify whether a hard layer (such as rock or shale) exists in the soil profile, and then determine the depth of that hard layer. Based on the TxDOT Geotechnical Manual (2006), drilled shafts should penetrate into a hard layer at least for one shaft diameter, if that hard layer is more than three shaft diameters below the surface. Otherwise, if the hard layer is close to surface, shafts are recommended to have a minimum length of three shaft diameters into the hard layer. However, in case of driven piles, they are normally stopped upon refusal when reaching a layer with $TCP > 100$ (blows/ft) (TxDOT 2006). The plans for over-river bridges in the Bryan District's database also show that driven piles and drilled shafts are generally extended up to the hard layer as a precaution and safety measure, when appropriate.

Step 4: Calculating Dead Load and Live Load. A simple, fast method is proposed in this study to estimate, with reasonable accuracy, the dead load and live load for bridge foundations. This procedure was designed based on the standard bridge design sheets that the TxDOT uses (TxDOT 2011). The total load for a bridge bent is obtained using the standard tables, depending on the roadway width and span length. The live load is calculated based on the standard design vehicles specified by the AASHTO-LRFD. The dead load is then calculated by subtracting the live load from the total load. Figure 5-7 presents the steps of the total load calculation procedure. The method is elaborated upon in the following sections.

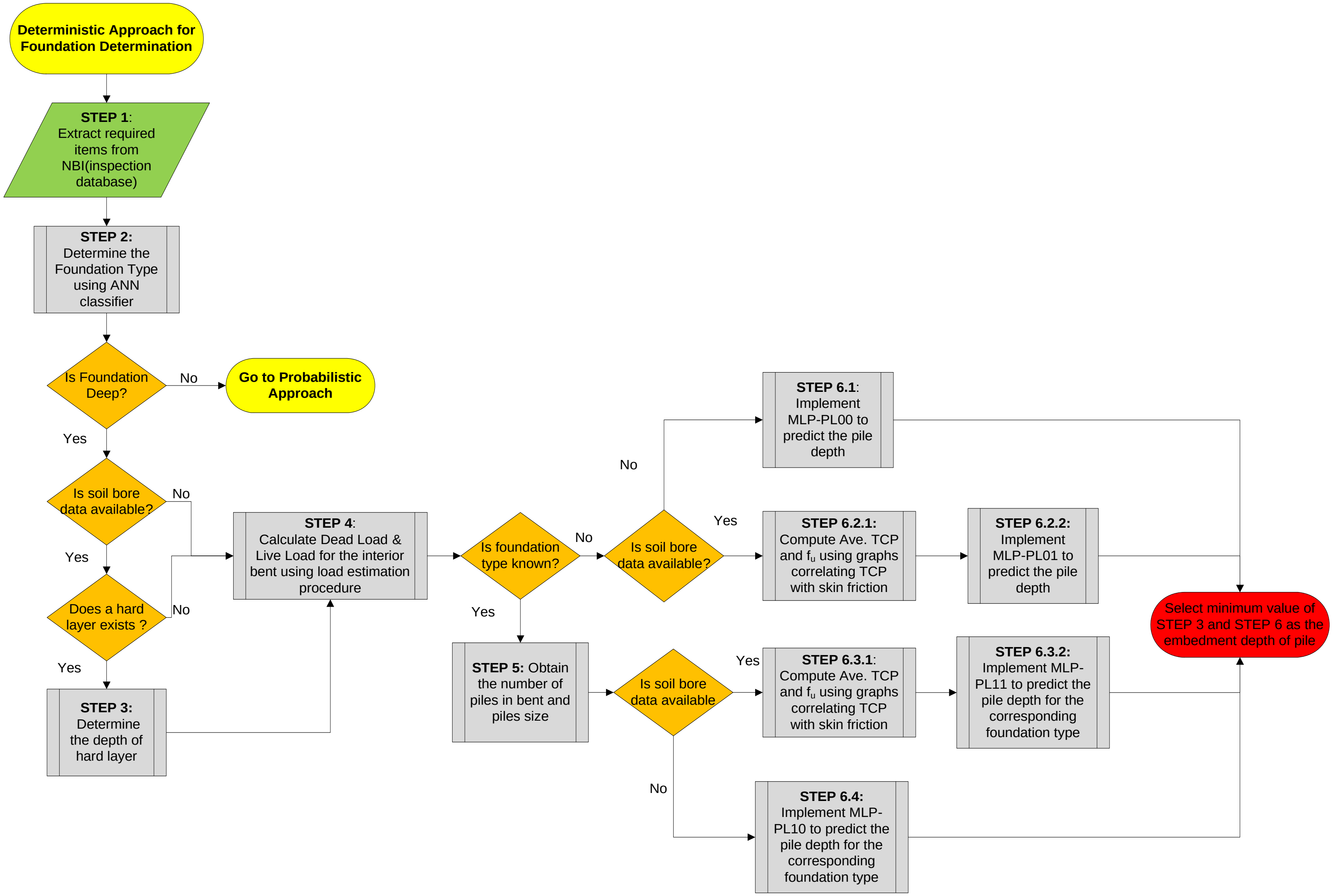


Figure 5-6. Deterministic Approach Flowchart

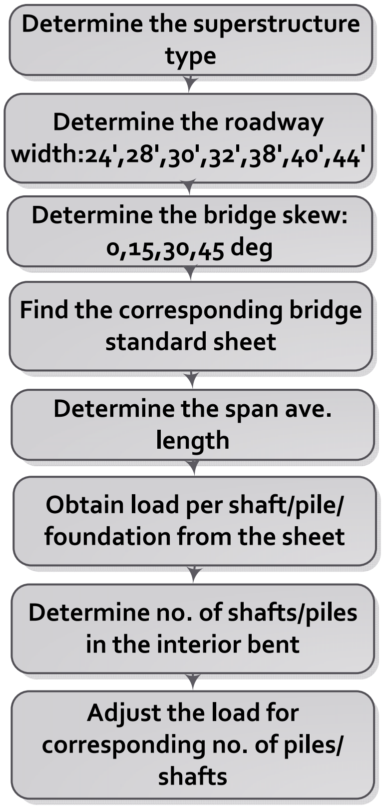


Figure 5-7. Total Load Calculation's Flowchart

Step 5: Obtaining the Foundation Properties. Acquiring those properties of the foundation that can be observed above the ground (such as size and number of piles for a bridge bent) can serve to enhance the performance of the ANN models.

Step 6: Predicting the Pile Depth Using ANN Models. Depending on the availability of soil boring data and foundation type information, four scenarios are defined in the following sub-steps:

Step 6-1: If the foundation type is known to be deep, but neither the foundation type nor the soil data are available, the ANN model *MLP-PL00* is recommended for predicting the foundation depth. This scenario assumes that the foundation type cannot be inferred, with the desired level of confidence, from the ANN classifier or by site investigation. More simply, it may also be used because the user is only interested in inferring the depth.

Step 6-2: This step addresses a scenario in which the foundation type is known to be deep, the exact foundation type is unknown (or cannot be inferred, with the desired confidence level, from the ANN classifier), but the soil boring data is available.

Step 6-2-1: Compute the average ultimate skin friction (f_u) based on the *TCP* values along the piles from the nearest soil data, using the graphs presented in Figure 5-5.

Step 6-2-2: Implement the ANN model *MLP-PL01* to predict the foundation (pile or shaft) embedment depth.

Step 6-3: This step addresses a scenario in which the foundation type is known and soil boring data is available.

Step 6-3-1: Compute the average soil skin friction (f_u) of piles.

Step 6-3-2: Implement *MLP-PL11* to predict the foundation embedment depth.

Step 6-4: If the foundation type is known or is inferred using the ANN classifier or by site investigation, but the soil boring data is not available, then implement the ANN model *MLP-PL10* to infer the foundation embedment depth.

The minimum value between the depth of the hard layer and the predicted foundation depth by the ANN model is considered to be the embedment depth.

Dead Load and Live Load Calculations

The dead load for a foundation in an interior bent is obtained by adding up all permanent loads that are sustained by the foundation from all components of the superstructure, including the deck, girders (beams), bent cap, column, and railing weights. However, this procedure requires the collection of a number of details from the superstructure, which makes the process very time-consuming. Therefore, another method is proposed to make the process faster. As mentioned earlier, TxDOT provides standard design sheets that can be used for the design of bridges. The total foundation load can be estimated with an acceptable level of accuracy using these sheets.

The load assessment process begins with identifying the superstructure type and finding the corresponding standard design sheet based on the bridge roadway width and skew. The different types of superstructures, as defined by the TxDOT, are: Prestressed I-Girder; Cast-In-Place Concrete Slab Span; Steel Beam; Prestressed Slab Beam; Prestressed Box Beam; Prestressed Concrete Double T-beam; and Concrete Slab and Girder (Pan Form).

The next step is to determine the roadway width from Item 51 of NBI. The standard sheets have been prepared for typical roadway widths varying from 7 m (24 ft) to 13 m (44 ft) (Note that all the items in NBI are in English units, therefore the typical roadway widths are in feet). However, the roadway widths among the bridges in the Bryan District vary between 5.8 m (19 ft) to 27 m (90 ft). The load estimates were extrapolated for widths greater than 13m (44 ft).

Finally, the bridge skew can be found in Item 34 of NBI. The bridge skews vary among 0, 15, 30, and 45 (deg). Once the superstructure type, the roadway width, and the bridge skew are defined, the corresponding TxDOT standard design sheet can be found for the bridge. The value provided by the bridge standard sheet for the load per pile/shaft should be adjusted according to the number of piles/shafts per bent for the desired bridge. The span length for each bent is obtained by averaging the left and right span lengths. The foundation load table obtained from a standard bridge design sheet for a 7.3 m (24 ft), prestressed concrete I-beam, 15 (deg) skewed is presented in Table 5-4.

A set of regression lines were developed based on the loads versus span lengths for both column (shaft) and pile bents given by the bridge standard sheets. Figure 5-8 presents the regression lines generated for a prestressed I-beam bridge with 7.3 m (24 ft) roadway width with 0, 15, 30, and 45 (deg) skews. Based on the regression results, a table of coefficients was developed that includes the regression coefficients associated with different typical values of roadway widths and bridge skews for each of the existing superstructure types. Table 5-5 shows the corresponding table of coefficients for a column bent of a prestressed concrete I-beam bridge versus typical roadway widths. A code was developed to automate the process of load assessment for the bridges in the database. This program inputs the table of coefficients, as well as a table containing the roadway widths, average span lengths, and the skews for a bridge bent, and then outputs the total load for the bent.

Table 5-4. Foundation Loads for 24 (ft) PSC I-Beam Bridge with a 15 (deg) Skew

Foundation Loads				
Span Average (ft)	Drilled Shaft Loads (Tons/shaft)	Pile Load (Tons/Piles)		
		3 Pile Footing	4 Pile Footing	5 Pile Footing
30	96	35	27	22
35	106	39	30	24
40	115	42	32	26
45	125	45	34	28
50	134	48	37	30
55	143	51	39	32
60	152	54	41	34
65	161	57	43	35
70	169	60	46	37
75	178	63	48	39
80	187	66	50	41
85	196	68	52	42

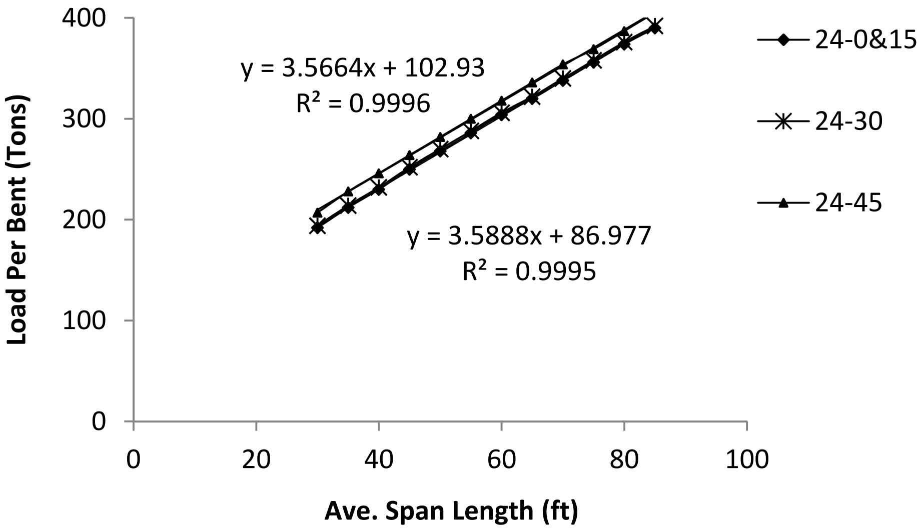


Figure 5-8. Load versus Average Span Length for a Prestressed Concrete I-Beams Bridge with 24 (ft) Roadway Width

Table 5-5. Load Estimation Coefficients for a Column Bent of a Prestressed Concrete I-Beam Bridge

Roadway (ft)	Skew (deg)	a	b
24	0	3.59	86.98
24	15	3.59	86.98
24	30	3.59	88.98
24	45	3.57	102.93
28	0	3.90	98.59
28	15	3.90	98.59
28	30	3.89	101.92
28	45	3.79	106.23
30	0	4.07	98.29
30	15	4.04	100.91
30	30	4.03	103.15
30	45	3.90	107.59
38	0	4.89	122.60
38	15	4.89	122.60
38	30	4.87	127.13
38	45	4.80	140.00
44	0	5.48	133.87
44	15	5.39	141.24
44	30	5.48	137.87
44	45	5.37	157.44

Prior to the adoption of the LRFD specifications, there were at least four different versions of AASHTO live-load standard computation methods. The H series modeled single unit trucks, while the HS series modeled semi-trucks. Note that the HS series has a variable dimension shown between the rear two axles. For the purpose of foundation load determination, this dimension will always be set at 4.26 m (14 ft), which serves to concentrate the maximum load over the interior bent. Figure 5-9 shows diagrams of the various load configurations. The lane loading (the uniform load) generally controls bridges with longer spans (> 21.3 m (70 ft)). Most structures on unknown foundations are locally owned with relatively short span lengths, so the truck loadings will generally control.

Since the LRFD design has been applied to structures built only in the past 10 years, there should not be as many LRFD-designed bridges on unknown foundations. However, the current TxDOT bridge standard sheets are all designed with the HL-93 live load. The live-load model for the LRFD code uses heavier loading conditions than the live-load models used in previous design specifications. Consequently, the resulting total number of loads will be higher. In order to adjust the total loads for the design vehicle and also to split the contributions for the dead and live loads, the following procedure is proposed:

- Assess the total load (live load + dead load) using the bridge's design standard sheets.
- Assess the live load based on the design vehicle HL93.
- Subtract the live load obtained from the previous step from the total load to obtain the dead load.
- Obtain the live load based on the real design vehicle of the bridge designated in Item 31 (Design Load) using the equations given in Figure 5-9.

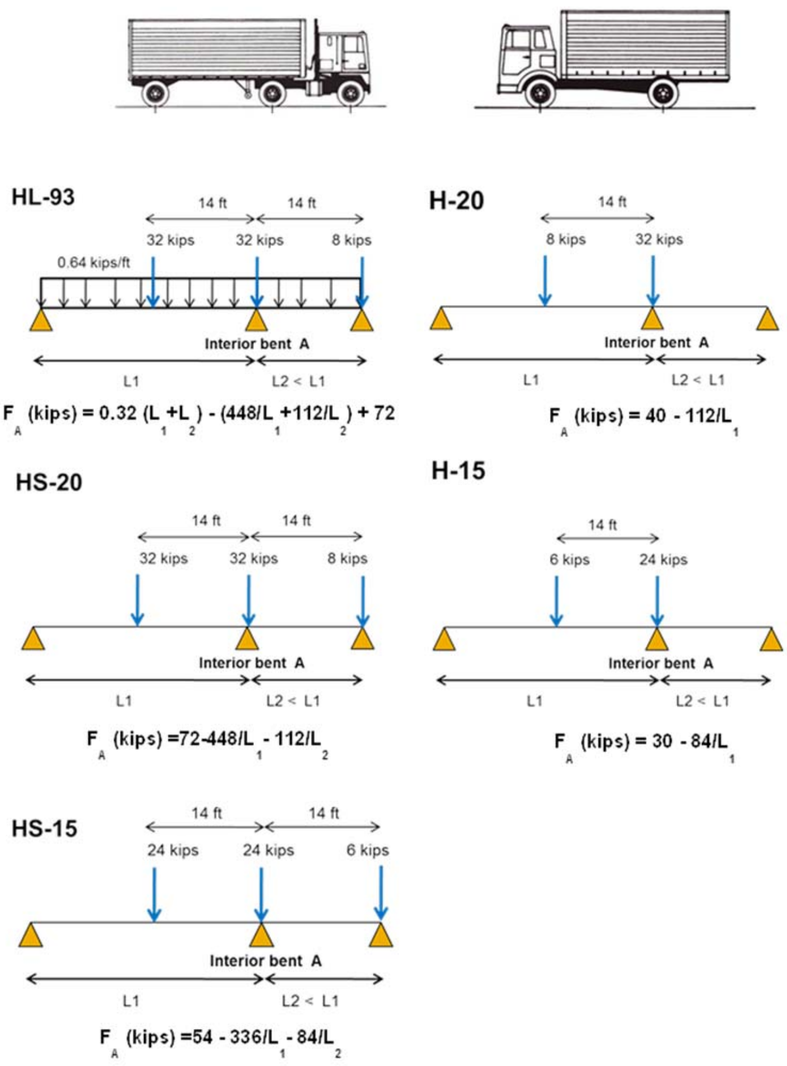


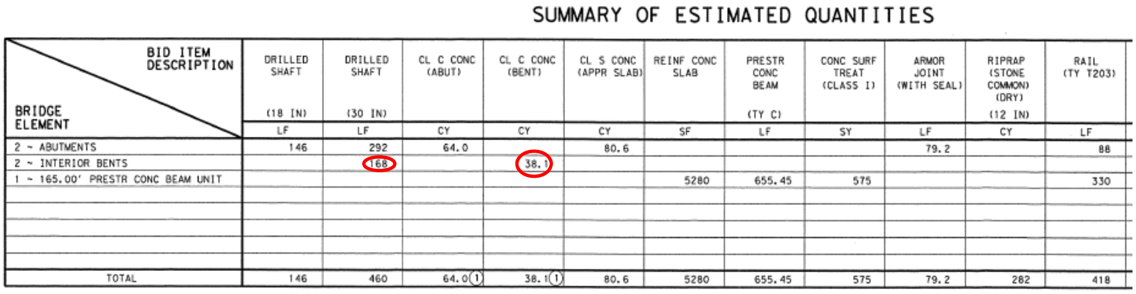
Figure 5-9. Standard AASHTO Design Loads and Maximum Reaction of an Interior Bridge Bent with Single Spans

In order to validate the proposed dead load estimation method, a number of bridges with different superstructure types were selected (see Table 5-6). The average error between the estimated values given by the proposed approximate method and the exact loads was about 11%.

Table 5-6. Load Estimation Error for Five Sample Bridges

Bridge	Error (%)
17-82-2144-01-002	2.56
17-21-0050-02-123	18.13
17-21-0050-01-001	8.33
17-239-0315-08-016	15.22
17-21-0049-09-43	12.80
Avg. Error	11.41

The exact dead loads for the interior bents of these bridges were calculated by summing the weights of different structural elements including the bent caps, slab/beams, shafts/piles, and railings. These details were found in the estimated table of quantities in the bridges' as-built plans. A sample calculation for the bridge 17-82-2144-01-002 is presented here (see Figure 5-10):



SUMMARY OF ESTIMATED QUANTITIES

BID ITEM DESCRIPTION / BRIDGE ELEMENT	DRILLED SHAFT (18 IN)	DRILLED SHAFT (30 IN)	CL C CONC (ABUT)	CL C CONC (BENT)	CL S CONC (APPR SLAB)	REINF CONC SLAB	PRESTR CONC BEAM (TY C)	CONC SURF TREAT (CLASS I)	ARMOR JOINT (WITH SEAL)	RIPRAP (STONE COMMON) (DRY) (12 IN)	RAIL (TY T203)
	LF	LF	CY	CY	CY	SF	LF	SY	LF	CY	LF
2 ~ ABUTMENTS	146	292	64.0		80.6				79.2		88
2 ~ INTERIOR BENTS		168		38.1							
1 ~ 165.00' PRESTR CONC BEAM UNIT						5280	655.45	575			330
TOTAL	146	460	64.0①	38.1①	80.6	5280	655.45	575	79.2	282	418

Sample Calculations:

L_1=75 ft (22.9 m), L_2=45 ft (13.7 m), RW=32 ft (9.8 m)
Estimated total load per shaft using the proposed method =228 kips (1014.5 kN)
Dead load = Total weights of the superstructure items:
Interior bent cap = 1/2 × (38.1 × 4.05) = 77.2 kips (343.54 kN)
Drilled shafts = 1/2 (168 × 0.15 × л × (30/12) ^{2}/4) = 61.9 kips (275.5 kN)
Concrete slab = 8/12 × 32 × (75 + 45)/2 × 0.15 = 192 kips (854.4 kN)
Railing = (75 + 45) × 0.38 = 45.6 kips (202.9 kN)
Beams = 4 × (75 + 45)/2 × 0.516 = 123.8 kips (550.9 kN)
Sum = 500.5 kips (2227.2 kN)

Figure 5-10. Estimated Quantities for Bridge ID: 17-82-2144-01-002

Live load-HL93:
2 × (0.32 (75 + 4) − (448/75 + 112/45) +72) = 203.9 kips (907.4 kN)
Total load:
500.5+203.9 = 704.4 kips/bent (3133.33 kN/bent) = 704.4/3 = 234.8 kips/shaft (1044.44 kN/shaft)
Error of estimation:
(234.8 − 228)/228 = 2.98%

Figure 5-10. Continued

Deterministic Models

Artificial Neural Networks

ANN models were developed to predict the type and embedment depth of unknown foundations. Radial Basis Function (RBF), Multi-Layer Perceptron (MLP), and Generalized Regression Neural Network (GRNN) were implemented to evaluate the foundation embedment depth, whereas the foundation type was predicted by Probabilistic Neural Networks (PNNs) and MLP classifiers.

Finding the best network type and configuration for the purpose of foundation type and depth prediction required considering different aspects of designing a neural network model. These aspects included the selection of: the network type and architecture, learning algorithm, input parameters, transfer functions, and the data-sampling method.

Neural Network Type and Algorithm

Foundation Type

ANN classifiers are powerful tools used to classify categorical data. Identifying the foundation type can be very helpful for determining the foundation embedment depth and scour vulnerability. Two ANN classifier models are proposed here for classifying a foundation based on the relevant information of the bridge. One determines the foundation type of bridges among all the common categories in the Bryan District. The other ANN classifier specializes only in determining if the foundation is either deep or shallow, because, for the proposed deterministic method in this study, it is important to know if the foundation is deep or not before making predictions about the foundation depth. Figure 5-11 shows the architecture of the MLP classifier used in this study.

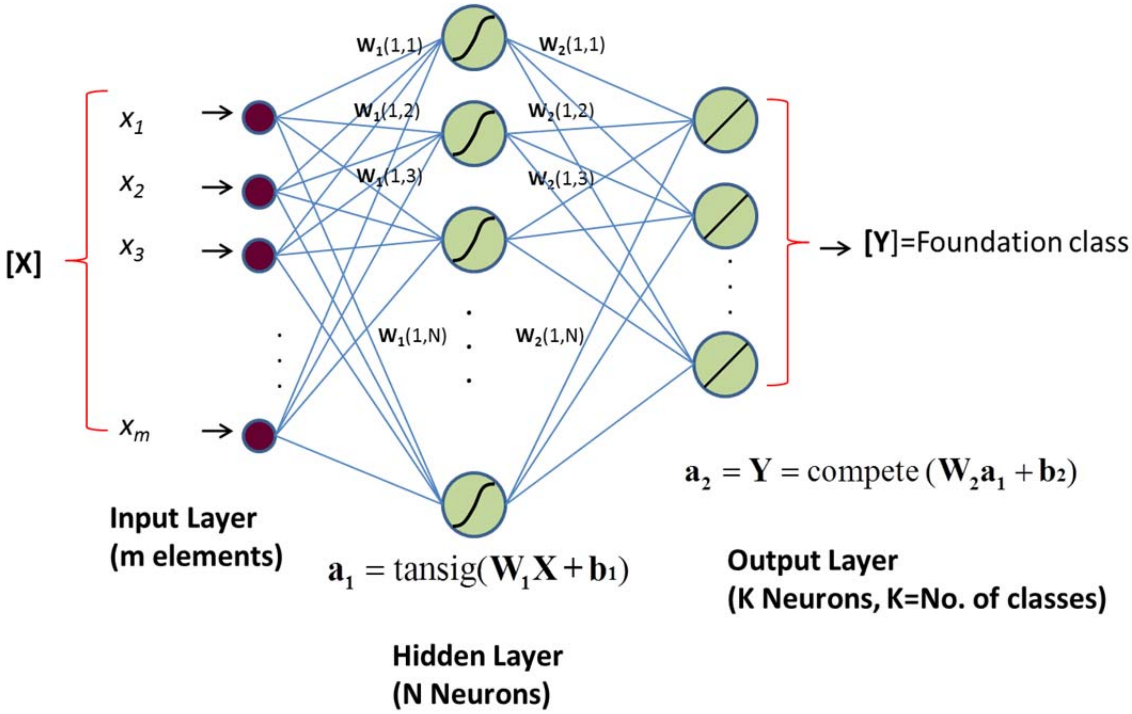


Figure 5-11. Architecture of the MLP Classifier

The classifier performance is measured by the classification accuracy (*CA*), which is defined as the ratio of the correctly classified examples over the total number of examples. A preliminary analysis demonstrated that MLP has greater classification accuracy than PNN. Table 5-7 shows the performance of each classifier.

$$CA = \frac{\text{No. of correctly classified examples}}{\text{Total no. of examples}} \tag{5-1}$$

Table 5-7. Classification Results for PNN and MLP

Model	PNN		MLP		
	Training	Test	Training	Validation	Test
CA	0.69	0.71	0.9	0.76	0.75

Foundation Depth

A preliminary analysis was performed to find the best network type using only a fraction of the database. MLP, RBF, and GRNN networks were developed and their performances were compared. The evaluation of the networks' performances was based on the root mean square error (*RMSE*) and the coefficient of determination (R^2) between the measured values (targets) and the predicted values (outputs) of the networks. The MLP networks were trained using two distinct optimization algorithms: Levenberg-Marquardt and Bayesian regularization, where the former is referred to as MLP-LM and the latter is referred to as MLP-BR. Before training, the database was pre-processed, meaning that the duplicated data points and the data points missing the pile depth value were eliminated. Also, both input and output data were normalized regarding the minimum and maximum values for each parameter, so that all parameters fell within the range [−1, 1].

In order to train the MLP-LM network, the data set was divided into three parts: 60% for training, 20% for validation, and 20% for test. The data points for training, validation, and test were sampled randomly. The network was trained using training data points and the performance of the network was monitored over the validation data points. The training was stopped when one of the following criteria was reached: the minimum mean squared error over the training data, maximum number of iterations, minimum gradient of the error function, and when the error over the validation dataset started to increase. The performance of the model was examined using the test data points, which had not been presented to the network.

The MLP-BR network was generated with the same architecture as the MLP-LM but trained using the Bayesian regularization algorithm. In this algorithm, weights and biases are assumed to be random variables, and the regularization parameters are associated with the variance of these distributions. Further, there is no validation error in this method, and the network will stop training when the effective number of weights and biases has converged (Demuth et al. 2009).

For RBF and GRNN networks, the working database was divided into training (70%) and test (30%) subsets. The weights and biases in the network were adjusted using the training set, and the generalization of the network was evaluated using the test set. Although both the RBF and GRNN networks showed relatively good performance for the training data, the small R^2 over the test dataset (particularly for the RBF network) indicated that these networks could not generalize well in this problem.

Table 5-8 presents the performances of different networks. The best generalization belongs to MLP-LM, followed by MLP-BR. The results show that both MLP networks outperformed the GRNN and RBF networks. Also, the GRNN network showed a better performance comparing to the RBF network, because only a small number of data points were presented to the models in the preliminary analysis. According to these results, the MLP network with the Levenberg Marquardt algorithm was selected for the prediction of pile depth.

Table 5-8. Comparison of Different ANN Models

	RBF		**GRNN**		**MLP-LM**		**MLP-BR**	
	Training	**Test**	**Training**	**Test**	**Training**	**Test**	**Training**	**Test**
R^2	1.00	0.11	0.60	0.38	0.76	0.79	0.90	0.66
RMSE	0.029	1800	8.29	10.26	7.56	9.99	4.00	8.08

Input Parameters

Foundation Type

The input parameters for foundation type prediction emerged from the factors participating in the decision making process for selecting the type of foundation. These factors included the location of the bridge, loading, soil type, depth of the hard layer, and flow characteristic under the bridge. A preliminary analysis was performed to select a set of items from the NBI that could contribute to the prediction of foundation types. Out of many items (more than 300), 12 input parameters were identified as presented in Table 5-9. The target parameter is Item 44-1-D2 of the NBI, which is the substructure type below ground for the main span.

A stepwise parameter selection method was implemented to select the relevant parameters that contributed the most to the foundation type classification. This method included Forward Stepwise Selection (FSS) and Backward Stepwise Elimination (BSE) of the input parameters based on their significance to the performance of the network. At each step, a number of networks were trained with the selected sets of input parameters. In FSS, parameters competed with each other to be added to the model, one by one. The process started with creating a network with only one input parameter and gradually adding the statistically significant parameters until no additional parameter could be added. A parameter was added to the model only if the adjusted R^2 for the model became larger than the adjusted R^2 of the reduced model. In the BSE, the process started with a network containing all the parameters and gradually removed parameters, one by one. The parameters competed with each other to be removed from the model. A parameter was removed when the adjusted R^2 for the reduced model became larger than the present model. This method also provides a ranking of the input parameter with respect to their significance to the prediction of the foundation type as will be explained later in Results section.

Table 5-9. Input Parameters for Foundation Classification

Input Parameters
Item 3: County
Item 26: Function Class
Item 27: Year Built
Item 29: AADT
Item 31: Design Load
Item 43-1: Main Span Type
Item 44-1-D1: Substructure Type above Ground for Main Span
Item 46: Total Number of Spans
Item 48: Max Span Length (m)
Item 51: Roadway Width (m)
Item 71: Water way adequecy
Item 107-1: Deck Type Main Span
Target Parameter
Item 44-1-D2: Substructure Type below Ground for Main Span

Foundation Depth

The criteria for selecting the relevant input parameters for pile depth prediction is based on the limit state design method of foundations which requires the foundation load and allowable bearing capacity to be equal. The parameters that contribute to the design of a bridge foundation include those associated with the superstructure, substructure, load, and location of the bridge. Several networks with different combinations of parameters were created and compared to determine the most relevant input parameters. Table 5-10 lists the final selected input parameters.

Table 5-10. Input Parameters for Pile-Depth Prediction

Input Parameters
Load
Dead Load per bent/ per pile (kN)
Live Load per bent/pile (kN)
Geometry
Item # 51: Roadway Width (m)
Soil Properties
Avg. *TCP* (blows/ft)
Skin Friction (MPa)
Time and Location
Item 3: County
Item 27: Year Built
Item 16: Latitude (deg)
Item 17: Longitude (deg)

Hidden Neurons and Transfer function

Based on a preliminary analysis on the number of hidden units, 20 hidden neurons were selected for this study. The transfer function, or activation function, maps the input of a neuron to its output in a computational network. Transfer functions can be linear, sigmoidal, Gaussian, or step functions. As explained earlier in Chapter II, MLPs apply either a log-sigmoid or a tan-sigmoid transfer function in the hidden neurons (Eq. 2-10 and Eq. 2-11) and a linear function in the output neurons.

In order to observe the influence of the transfer function on the performance of MLPs, a network with log-sigmoid and another with tan-sigmoid were trained. The results showed that the R^2 values for training, test, and validation sets were not significantly different for the two MLP networks. Based on this analysis, the MLP models for pile depth were designed with tan-sigmoid transfer functions in the hidden neurons.

Training

The working database was split into three groups: training, validation, and test, with 3:1:1 proportions. The weights of the network were adjusted using training examples. The network performance was monitored during training using validation examples. The learning process continued until one of the stopping criteria was met, as described earlier in Chapter II.

The way the data points are sampled for training, validation, and test datasets, or in other words, how the database is partitioned can affect the model performance. The only way to estimate a model's true error is to test the model error on an entire population, meaning that one would need to have access to an unlimited number of samples. However, this is not possible due to the costly process of data collection.

In order to make a reasonable estimate of model performance, the random subsampling method was implemented. In this method, data points were randomly selected for a fixed number of examples in training, validation, and test datasets. This process was performed iteratively a large number of times (10,000), and each time a new network is generated from scratch and the model performance is measured. As a result, an ensemble of networks were generated that can work in parallel to make predictions. The best estimation is the average of the predictions by all networks in the ensemble. The cross-validation method allows one to make use of all examples in the database for both training and testing purposes, and thus provides a more reasonable estimate of the true error. It also helps reduce the bias of predictions by providing ensembles of networks trained by different subsamples of the same population.

In this study, the ANN models in the ensemble with R^2 *(Test)* < 0.7 were removed from the ensemble. This allows for the forming of an ensemble of networks with best-performances to be used as a predictive model.

Nonlinear Least Squares Method

In order to compare ANNs with a nonlinear least square approach, a Nonlinear (NL) model similar to MLP output function was defined as described in Eq. 2-27. A nonlinear least square approach was then taken to compute the weights and biases. Similar to the MLP model, the LM algorithm was implemented to perform the optimization, as was previously explained.

Results and Discussion

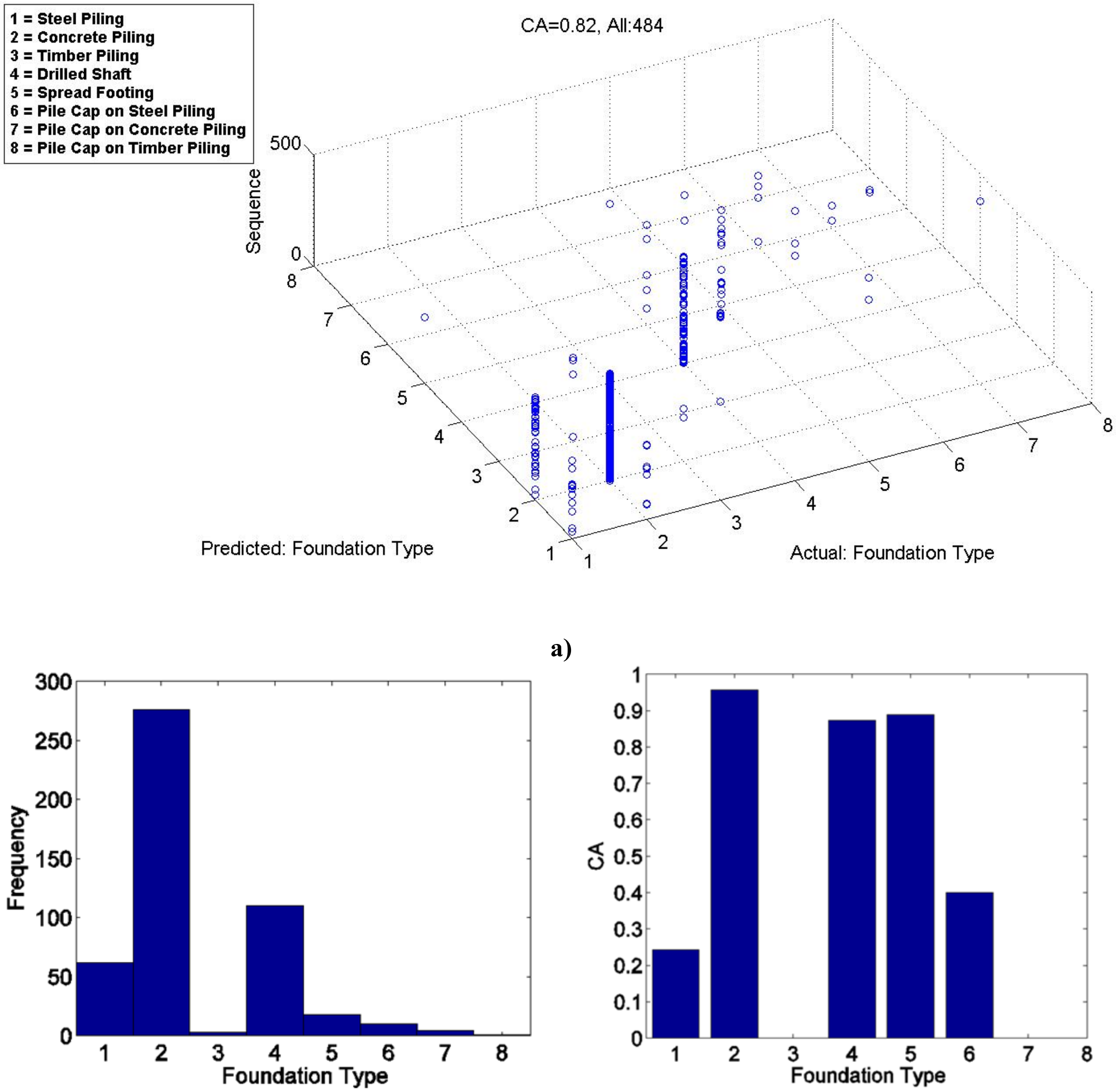
Foundation Type Classification

The FSS and BSE procedures were repeated fifty times to ensure the repeatability of the ranking of the input parameters. Based on this experiment, the input parameters were ranked by their significance in predicting the foundation type. Table 5-11 presents the ranking of the input parameters. The parameters that were eliminated through FSS and BSE are highlighted. Both the reduced (7 input parameters) and full (12 input parameters) models were implemented and their prediction accuracies were compared.

Table 5-11. Ranking of the Parameters through FSS and BSE Methods (the highlighted parameters are eliminated from the model)

FSS		BSE	
Ranking	**Item No.**	**Ranking**	**Item No.**
1	44_1	1	44_1
2	27	2	27
3	48	3	48
4	46	4	51
5	31	5	3
6	51	6	46
7	3	7	31
8	26	8	71
9	29	9	43-1
10	71	10	107-1
11	43-1	11	29
12	107-1	12	26

Figure 5-12 and Figure 5-13 show the predicted versus actual foundation types generated by one of the reduced MLP classifiers in the ensemble. Figure 5-12 presents the classifications among eight possible foundation types (FT8) and Figure 5-13 shows the classifications between deep and shallow foundations (FT2). In addition, Figure 5-12 shows the number of records in the working database and the classification accuracy for different types of foundation. Table 5-12 presents the total classification accuracy for the full and reduced MLPFT8 and MLPFT2 models.



a)

b) c)

Figure 5-12. Foundation Classification Using MLP-FT8, a) Predicted vs. Actual Scatter Plot, b) Histogram of Foundation Types, c) CA for each Foundation Type

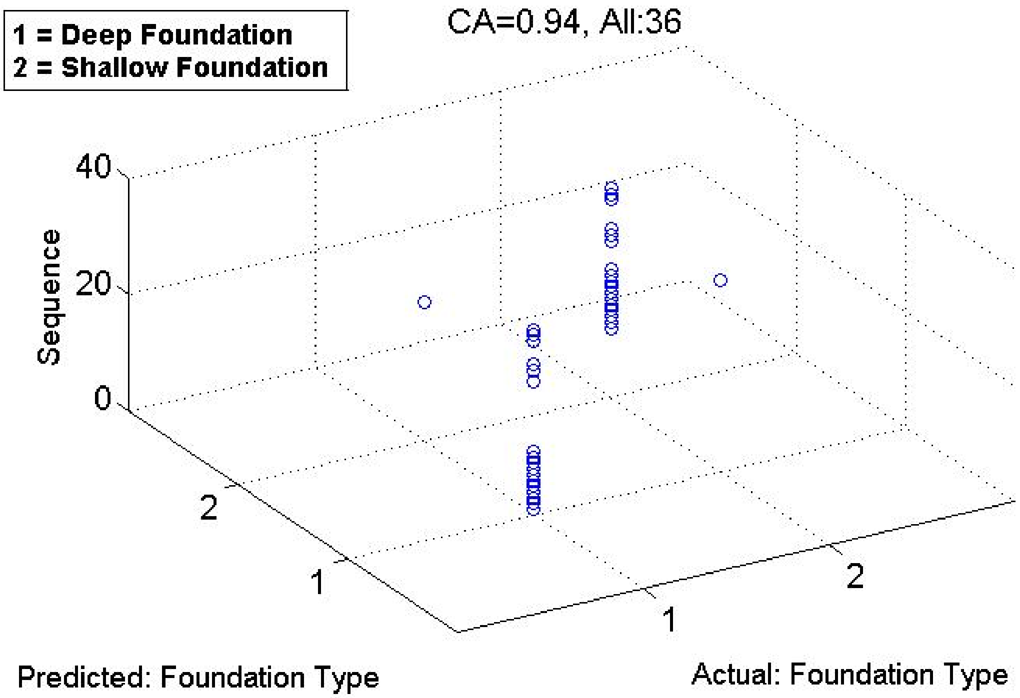


Figure 5-13. Foundation Classification MLP-FT2

Table 5-12. Results of Foundation Type Classification

Model	**MLP-FT2**			**MLP-FT8**		
Subset	**Training**	**Validation**	**Test**	**Training**	**Validation**	**Test**
No. of data points	22	7	7	290	97	97
Ave. *CA* - Reduced Model	1	0.91	0.90	0.91	0.75	0.75
Ave. *CA* - Full Model	1	0.9	0.89	0.93	0.71	0.70

It can be observed that both the MLP-FT8 and MLP-FT2 were able to classify the foundations with a fairly good level of accuracy. The *CA* over the test subset was 0.90 and 0.75 for MLP-FT2 and MLP-FT8, respectively. This result shows that it is significantly more accurate to classify a foundation as deep or shallow rather than to identify the exact foundation type. Also, by comparing the full and reduced models, it was observed that the proposed parameter selection method improved the classification accuracy of the model. This improvement was more significant for MLP-FT8.

Foundation Depth Prediction

ANN Prediction Results

Eight different MLP models were developed to predict pile depth based on the availability of soil data and foundation type information. The coding of each of these models was described earlier in the Approach section. Note that for the scenarios assuming that the soil data is not available the soil strength parameters were eliminated from the list of input parameters, leaving seven input parameters to be used. For the scenarios assuming that the soil data is available, all the nine input parameters were used to train the models.

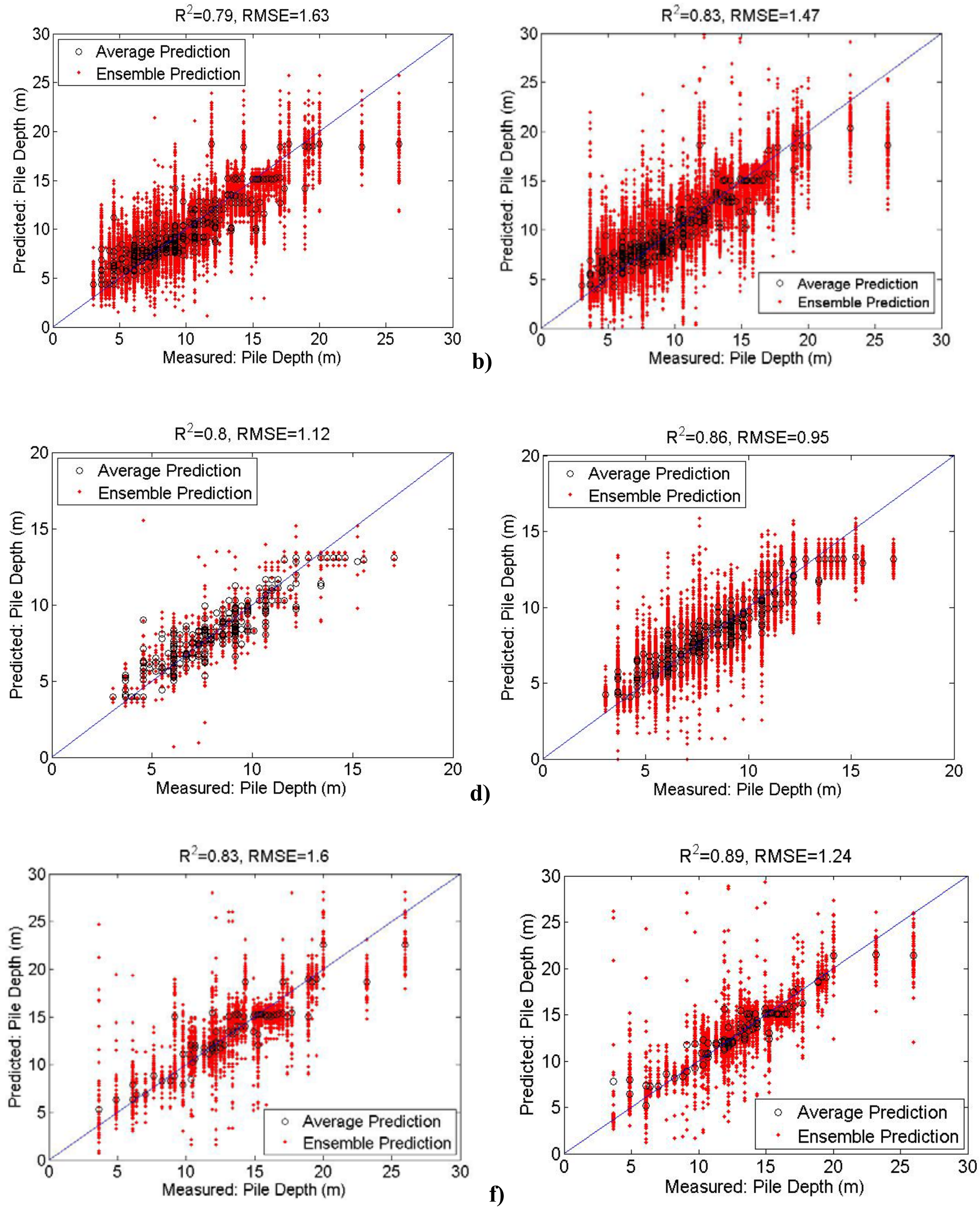
Table 5-13 presents the performances of MLPPL models for the eight different scenarios. Note that R^2_{ave} and $RMSE_{ave}$ present the average performance of the best networks selected. Therefore, the number of the networks in the ensembles varied for different scenarios. Figure 5-14 shows the MLP ensemble predictions for all the data points in the corresponding databases. This figure shows each network's prediction as well as the average predictions over all the networks in the ensemble. R^2 and $RMSE$ for the average ensemble predictions are also presented in this figure. The performances of the models can be compared through the R^2 of the average predictions. After comparing

MLPPL00 with MLPPL01 and MLPPL10 with MLPPL11, it was observed that adding the soil strength parameters as inputs to the model significantly improved the performance of the models.

It can be observed that for all eight MLP models, except for steel piling, the order of magnitude of R^2 of the average predictions, for the examples in the working data set, was 0.8 or above. Due to the limited amount of data for steel piling, the steel piling models could not perform as good as the two other types of piling.

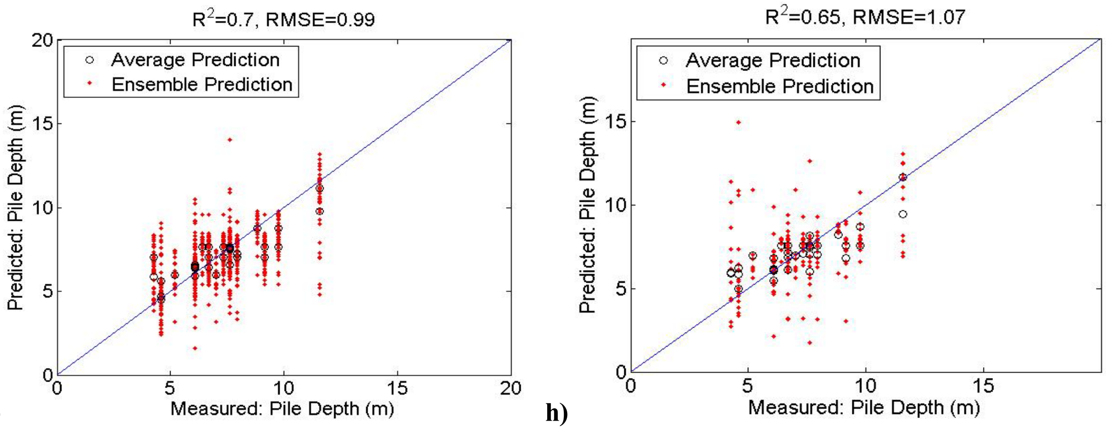
Table 5-13. Results of MLP Ensembles for Pile Embedment Depth Prediction

Model \ Subset	No. of Data	Training		Test	
		R^2_{ave}	$RMSE_{ave}$	R^2_{ave}	$RMSE_{ave}$
MLP-PL00	378	0.84	1.41	0.72	1.95
MLP-PL01	378	0.88	1.18	0.71	1.92
MLP-PL10-Conc	217	0.88	0.85	0.71	1.39
MLP-PL10-DrSh	80	0.87	1.39	0.76	1.93
MLP-PL10-Steel	35	0.75	0.85	0.77	0.94
MLP-PL11-Conc	259	0.93	0.64	0.72	1.42
MLP-PL11-DrSh	91	0.98	0.54	0.75	1.97
MLP-PL11-Steel	35	0.84	0.65	0.79	0.89



a) b) c) d) e) f)

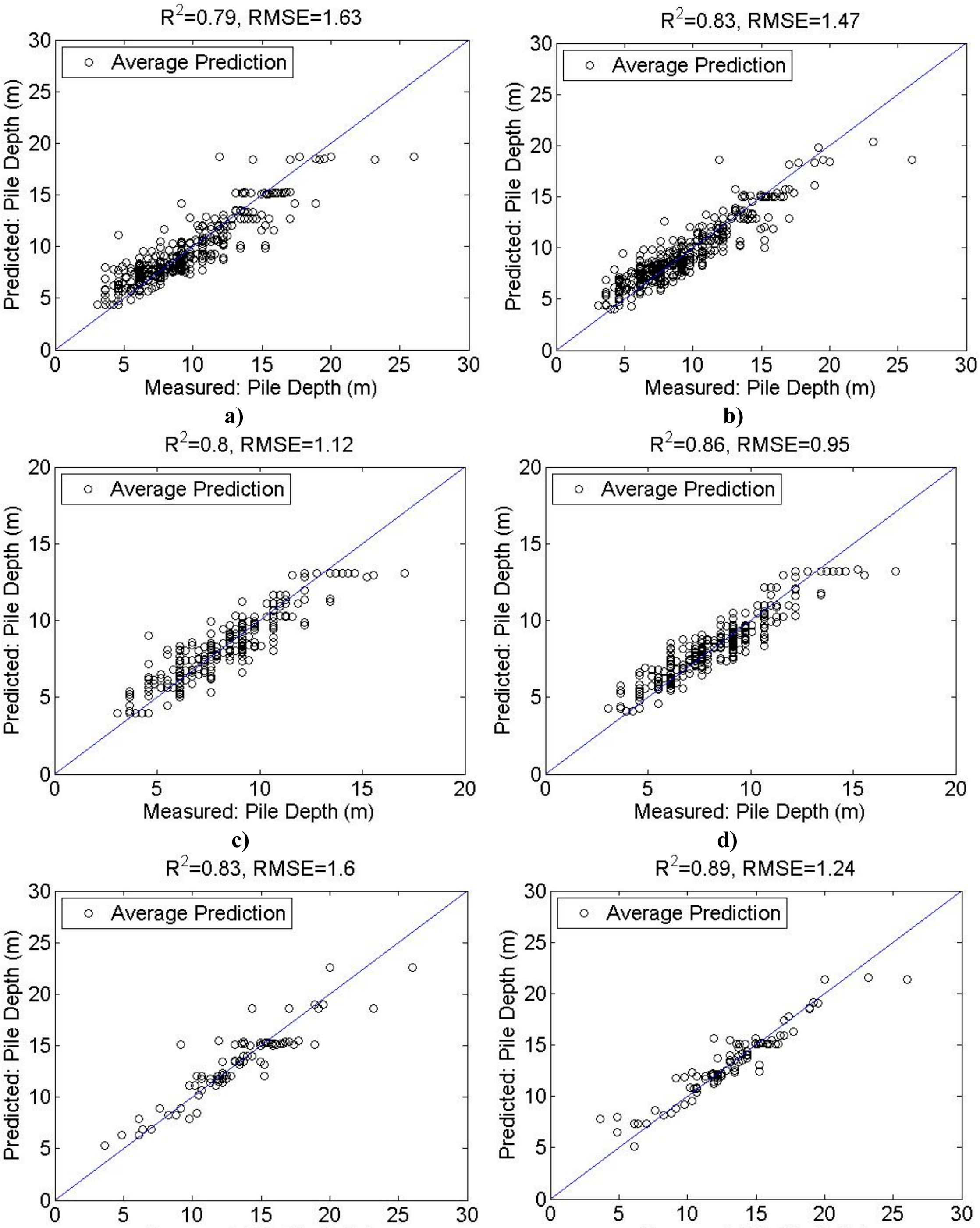
Figure 5-14. Pile Depth Predictions by the MLP Ensembles: a) MLPPL00, b) MLPPL01, c) MLPPL10-Conc, d) MLPPL11-Conc, e) MLPPL10-DrSh, f) MLPPL11-DrSh, g) MLPPL10-Steel, h) MLPPL11-Steel



g) h)

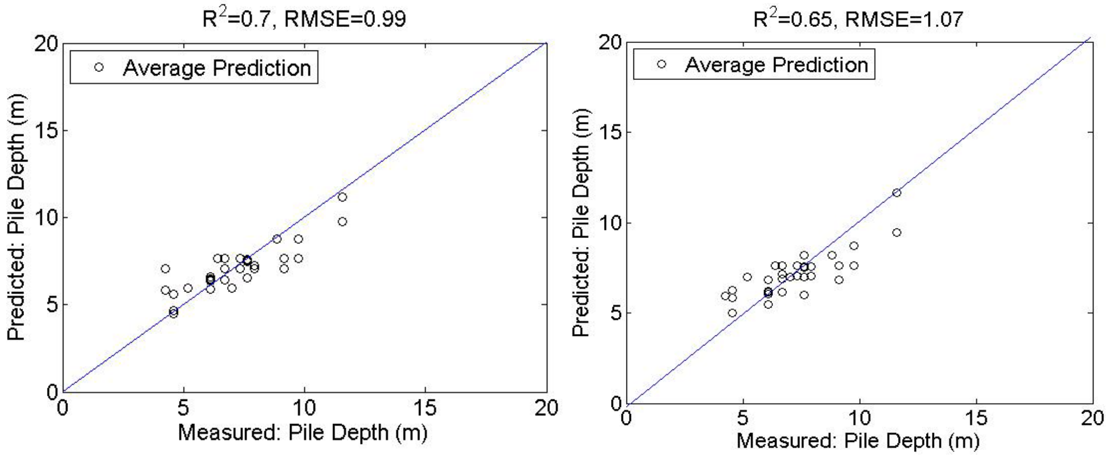
Figure 5-14. Continued

Figure 5-15 shows the average predictions of the networks in the ensembles for the eight MLP models. Also, Figure 5-16 presents the average of the ensemble predictions along with the 95% prediction intervals and the 10th percentiles of the pile depth distributions generated for each example in the database. The prediction interval provides a range that the predicted pile depth would be within that range with 95% probability. Also, the 10th percentile provides a value for which the predicted pile depth would be less than that value with a 10% probability. This could provide a conservative value to compute the probability of failure due to scour.



a) b) c) d) e) f)

Figure 5-15. Average Pile Depth Predictions by the MLP Ensembles: a) MLPPL00, b) MLPPL01, c) MLPPL10-Conc, d) MLPPL11-Conc, e) MLPPL10-DrSh, f) MLPPL11-DrSh, g) MLPPL10-Steel, h) MLPPL11-Steel



g) **h)**

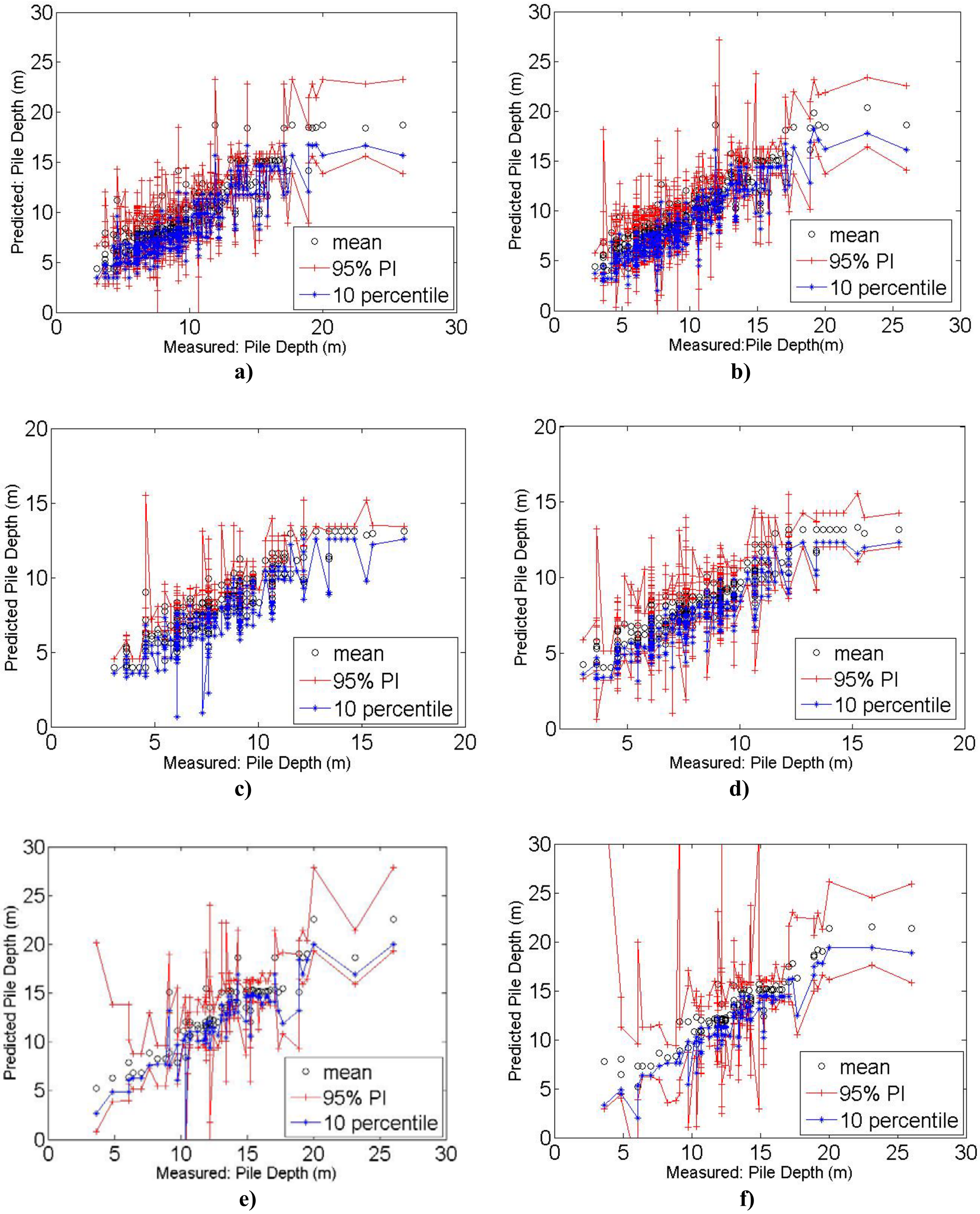
Figure 5-15.Continued

Figure 5-17 shows the Probability of Exceedance (*PoE*) curves obtained from the predicted pile depths by each MLP models versus the actual pile depths. The horizontal axis corresponds to a correction factor *C* that was multiplied by the predicted values of the pile depths:

$$C = \frac{d_p(Corrected)}{d_p(\text{Predicted})} \qquad (5\text{-}2)$$
$$PoE = \Pr(d_p(\text{Corrected}) > d_p(\text{Actual}))$$

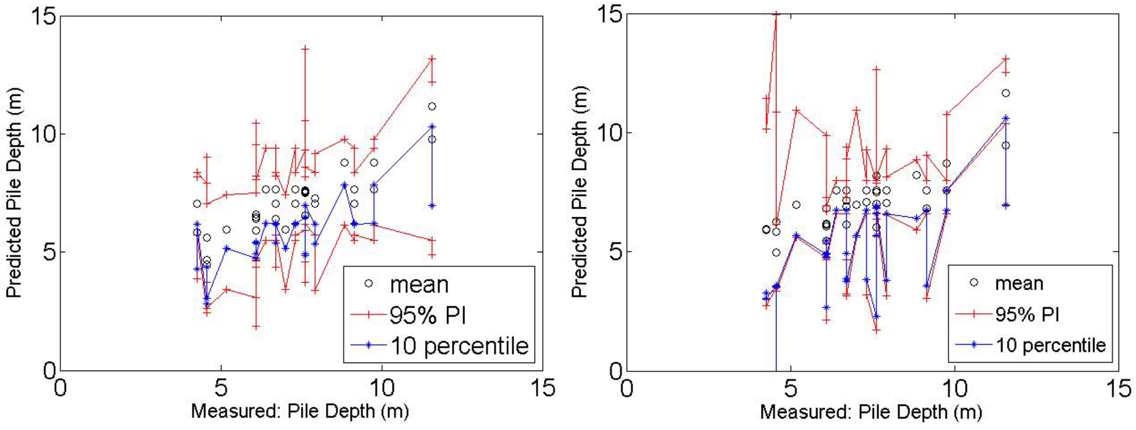
The probability of exceedance graphs for each model was obtained by taking the following steps:

- Taking a range for the correction factor *C* between 0 to 5
- For each value of *C* multiply the predicted pile depths by the correction factor
- Count the number of data points that fall above the 1:1 line (considering the y axis to represent the predicted values)



a) b) c) d) e) f)

Figure 5-16. 95% Prediction Interval and 10 Percentile of Pile Depth Predicted by a) MLPPL00, b) MLPPL01, c) MLPPL10-Conc, d) MLPPL11-Conc, e) MLPPL10-DrSh, f) MLPPL11-DrSh, g) MLPPL10-Steel, h) MLPPL11-Steel



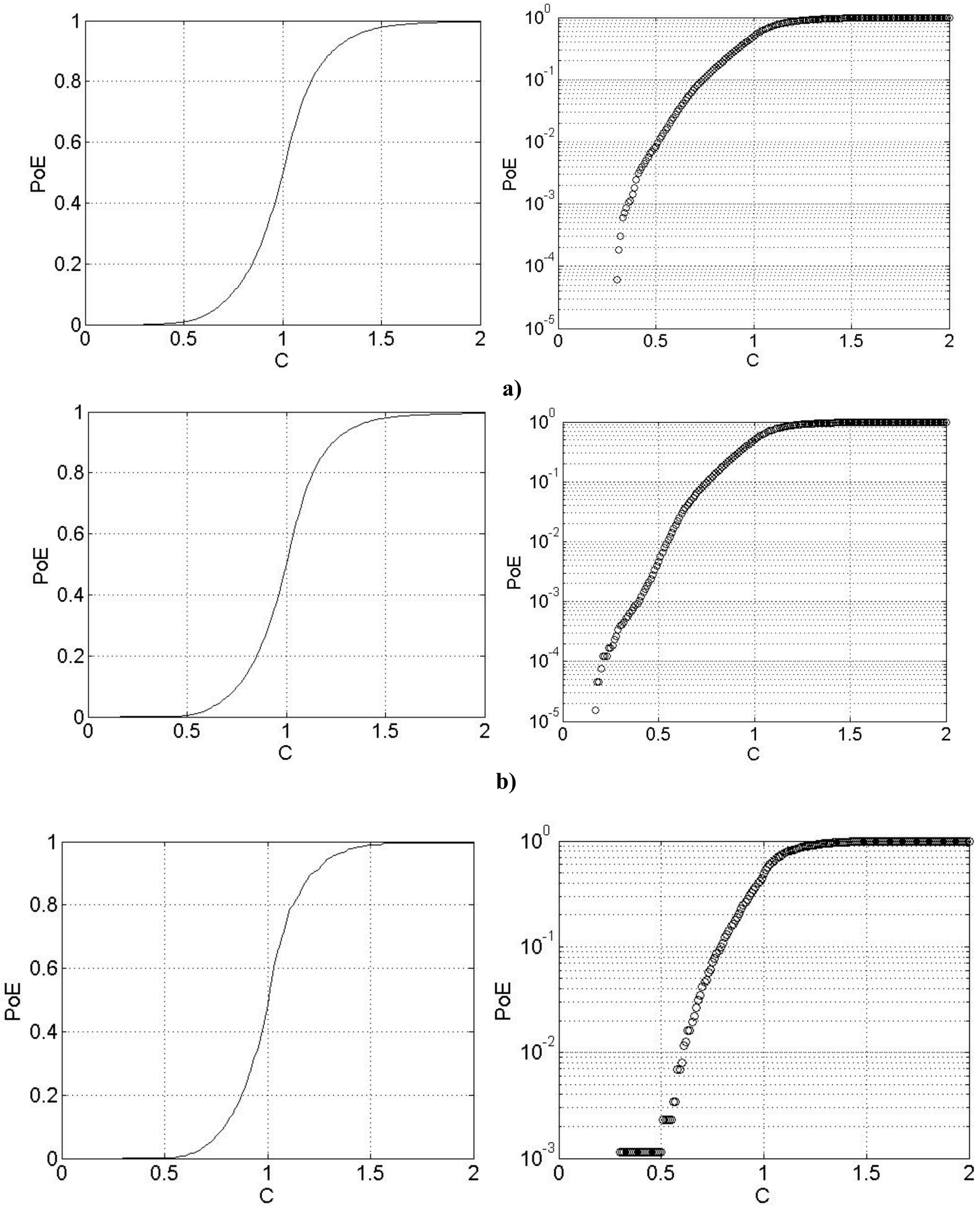
g) h)

Figure 5-16. Continued

- Divide the number of points above the line over the total number of data to obtain the probability that the predicted pile depth will be more than the actual pile depth (probability of exceedance)

- Plot the PoE values versus the corresponding correction factors

These graphs allows for estimating the pile depth with the desired level of confidence from the proposed MLP models. For a desired PoE one can use the graphs in normal or log scale to determine the appropriate correction factor. The predicted pile depth by the model needs to be multiplied by this factor to satisfy the desired probability of exceedance. It is evident that the lower probability of exceedance, the smaller correction factor is required for a specific model.

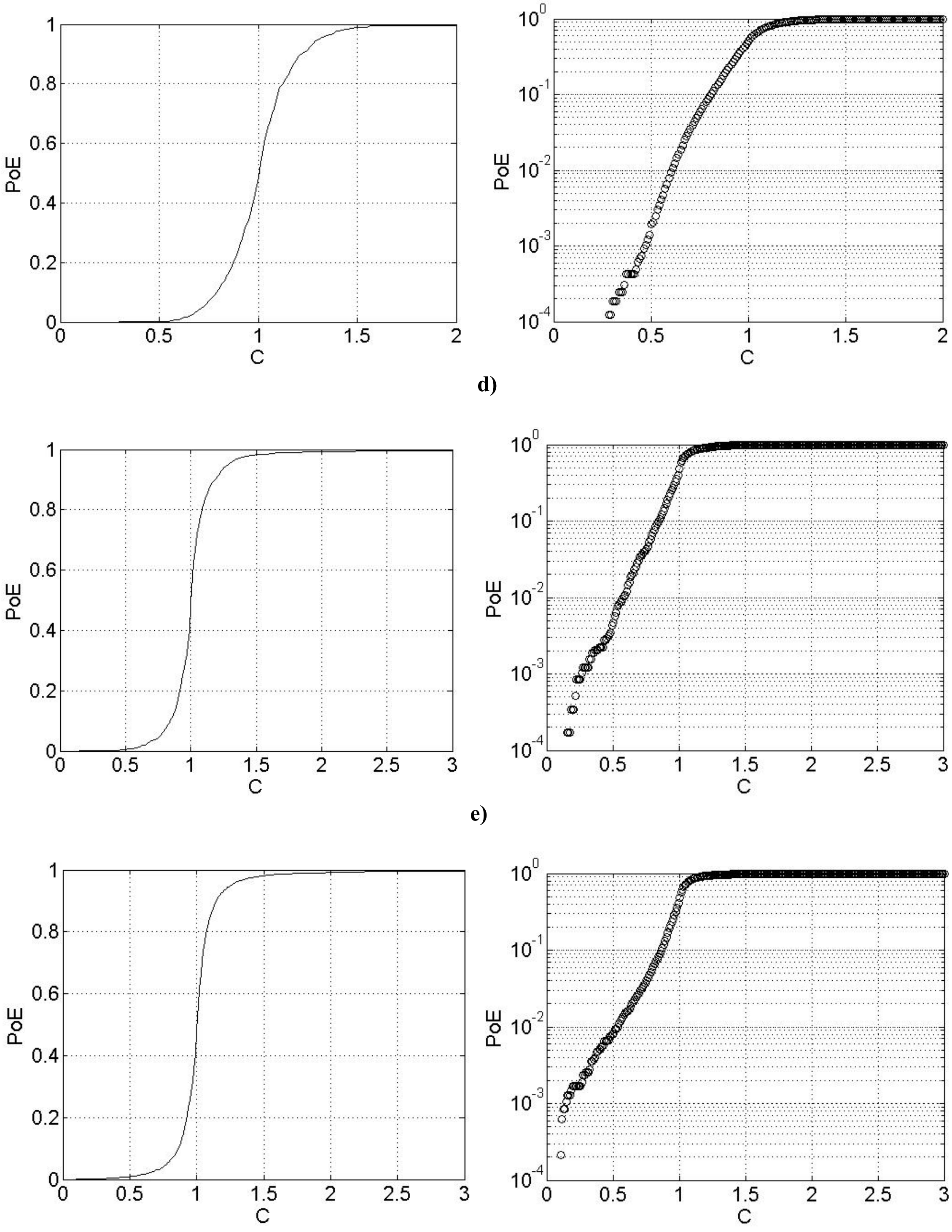


a)

b)

c)

Figure 5-17. Probability of Exceedance Curves for Pile Depth in the Normal and Log Scales a) MLPPL00, b) MLPPL01, c) MLPPL10-Conc, d) MLPPL11-Conc, e) MLPPL10-DrSh, f) MLPPL11-DrSh, g) MLPPL10-Steel, h) MLPPL11-Steel

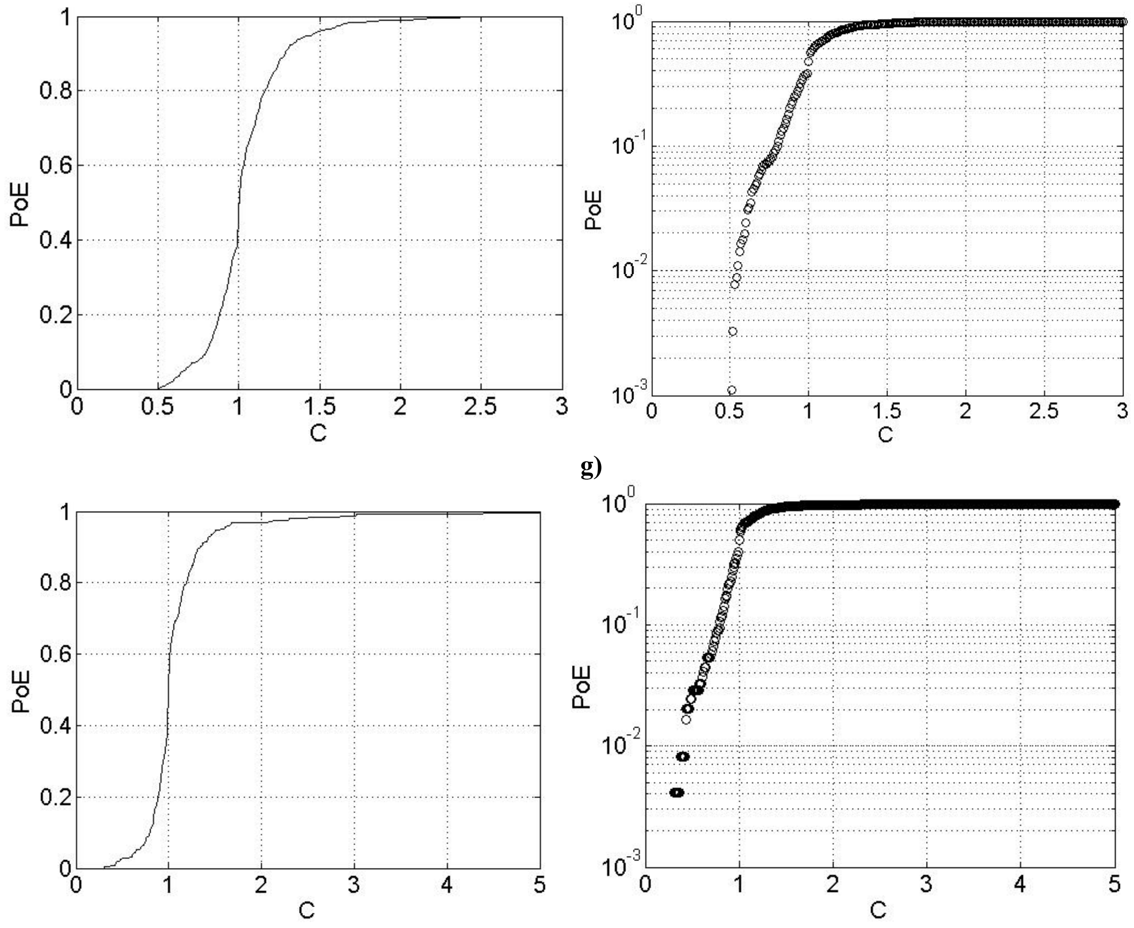


d)

e)

f)

Figure 5-17. Continued

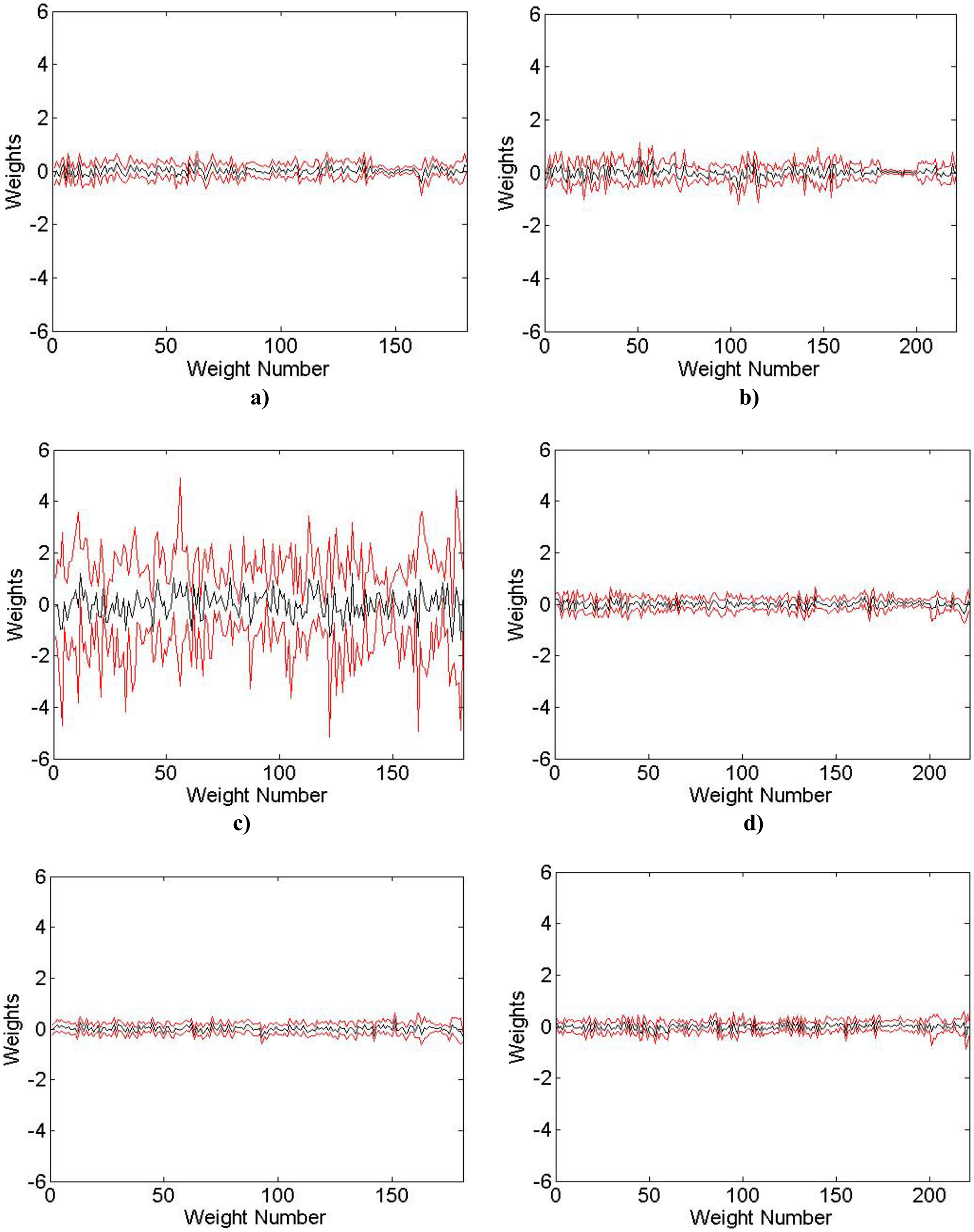


g)

h)

Figure 5-17.Continued

Figure 5-18 presents the average of the weights and biases of each ensemble, along with the 95% confidence interval for these parameters. The weights and biases of all connections are sorted in a vector presented as Weights. The total number of weights and biases is 221 for models with soil parameters and 181 for models without soil parameters. Note that the last 21 of the weights presented in this figure corresponds to the biases of the 20 neurons in the hidden layer and the output neuron. It seems that most of the weights and biases are between -1 and 1. Also, biases are relatively larger than weights.



a)

b)

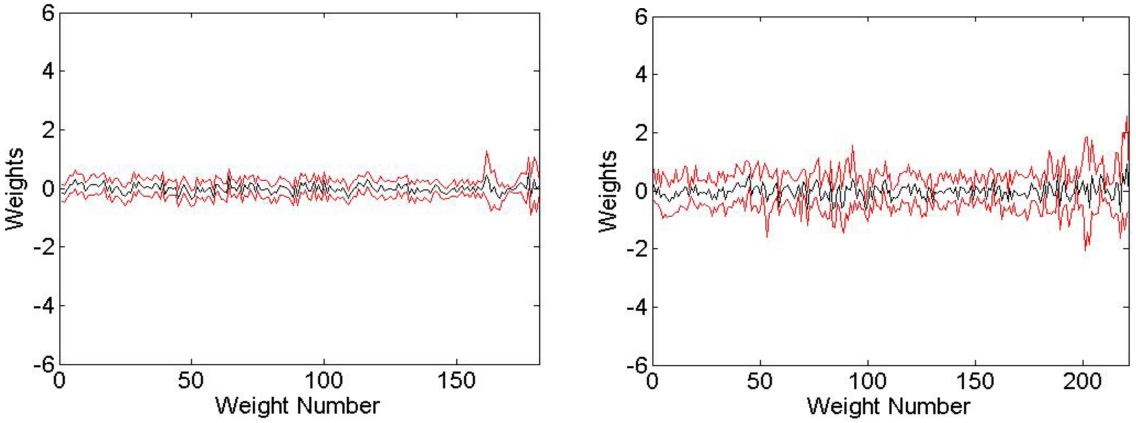
c)

d)

e)

f)

Figure 5-18. Weights and Bias of the MLP Models: a) MLPPL00, b) MLPPL01, c) MLPPL10-conc, d) MLPPL11-conc, e) MLPPL10-DrSh, f) MLPPL11-DrSh, g) MLPPL10-steel, h) MLPPL11-steel



g) **h)**

Figure 5-18. Continued

Nonlinear Least Squares Results

The NL model was trained using the whole database. Similar to the MLP model, 1,000 models were generated by the random subsampling of data points for training and test subsets. Also, the data was normalized before training the model. According to the results, none of the models had an R^2 greater than 0.7 for the test dataset. Table 5-14 provides the performance of the best NL model. The results of the nonlinear least square method for prediction of the pile embedment depth is presented and compared with the MLP in Figure 5-19.

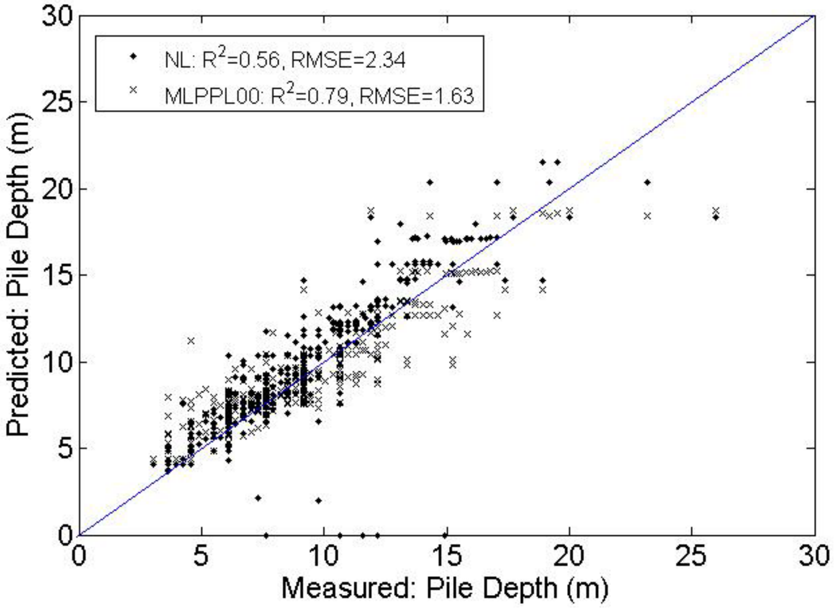


Figure 5-19. Comparison of NL and MLP Models

Table 5-14. Results of the Best NL Model

Subset / **Model**	**No of Data**	**Training**		**Test**	
		R^2	***RMSE***	***R^2***	***RMSE***
NL	378	0.63	0.0978	0.68	0.11

Although the function of both MLP and NL models were the same and the same optimization method (Levenberg-Marquardt) was implemented for the both models, the MLP network highly outperformed the NL model. This indicated that the BP algorithm was more successful in finding the minimum of the objective function.

Conclusions

The deterministic approach proposed in this study allows an engineer to make predictions about the substructure type and embedment depth of unknown foundations with a reasonable level of accuracy without performing time-consuming, costly *in situ* experiments. The predicted embedment depths can be used to evaluate the scour vulnerability of bridges; therefore, the proposed methodology can assists DOTs to reclassify the U coded bridges in the NBI by updating Item 113 (the scour-critical index).

The proposed approach first determines the type of foundation, and then, if the foundation is recognized to be deep (either concrete, steel piling, or drilled shafts), it selects the appropriate ANN model based on the foundation type and the availability of soil data to predict the embedment depth. Finally the minimum between the predicted pile depth and the depth of the hard layer under the bridge is selected as the pile embedment depth. According to the results, the foundation type was successfully determined using ANN classifiers.

A stepwise parameter selection method was implemented to optimize the model performance by eliminating the redundant input variables from the model. The ANN classifier accommodates the categorization of an unknown foundation as either deep or shallow. It also can accurately classify the foundation among different possible types.

The ANN models developed in this study approximated the pile embedment depth based on the bridge load, soil properties, location, and year built. The R^2 of predicted versus actual pile depth values for the ANN models was greater than 0.8.

Clustering the data based on the type of foundation resulted in better model performance for concrete and drilled shaft types, for which there existed a sufficient number of data points. Adding skin friction and average *TCP* as input parameters to the ANN models significantly improved the performance of the models. This result confirms the significant effect of soil type and soil strength on pile embedment depth. Also, comparing the results of the nonlinear least square method with the MLP backpropagation algorithm indicates that BP outperforms the nonlinear least squares algorithms of optimization. The ANN ensembles proposed in this study generate distributions for pile depth that take into account the randomness in the predictions, and in this sense provide a prediction band for pile depth in addition to point estimates.

Although the present study was based on a working database populated in the TxDOT Bryan District, the proposed methodology can be applied to make predictions of any unknown foundation in any other district, as long as the models are trained using the evidence about known foundations characteristics in that district.

CHAPTER VI

DETERMINATION OF UNKNOWN FOUNDATIONS FOR BRIDGE SCOUR: A PROBABILISTIC APPROACH

Introduction

The predictions generated by the deterministic method do not address the uncertainties in the foundation characteristics initiated from the soil properties, bearing capacity calculations, and load estimations. Considering also the uncertain nature of scour, there is an inherent need to define a probabilistic approach that can provide the level of uncertainty in the prediction of unknown foundations' characteristics, by making use of available evidence of known foundations.

A probabilistic methodology and modeling approach are proposed in this study, incorporating Artificial Neural Networks (ANNs) and the Bayesian inference method. The equations of Bearing Capacity (BC) of foundations are considered as the forward models, as they control the physics of bridge stability by satisfying the static equilibrium between foundation load and allowable bearing capacity. The loads are first calculated for a bridge, and then the BC of the foundation is approximated using ANN models based on load, soil resistance (if known), location, and year the bridge was built. Given the BC and loads, the inverse problem can be solved for the unknown parameters of the bridge using the Bayesian solution. This method provides posterior joint probability distributions for the unknown foundation dimensions and soil characteristics that allow for generating point estimations and determining the uncertainty of the estimations.

The proposed methodology was implemented using the data collected for bridges located in the TxDOT's Bryan District (Briaud et al. 2011). The characteristics of the database were explained in Chapter V. This method can further be extended to be used in other districts in Texas, as well as in other states. The proposed method was validated by a case study performed on a bridge in the Bryan District.

Approach

Methodology

Figure 6-1 depicts the basic components of the evaluation of unknown foundations using the proposed probabilistic approach. This method is composed of two stages, where each stage includes two basic procedures. Figure 6-2 presents a flowchart of the probabilistic method. The steps are described below:

Step 1: Data Extraction. The items required for implementing the probabilistic approach are retrieved from the NBI. These parameters are used as input data for the ANN models to estimate the BC:

- Item 3: County (code).
- Item 16: Latitude (to be converted to degrees).
- Item 17: Longitude (to be converted to degrees).
- Item 27: Year built.

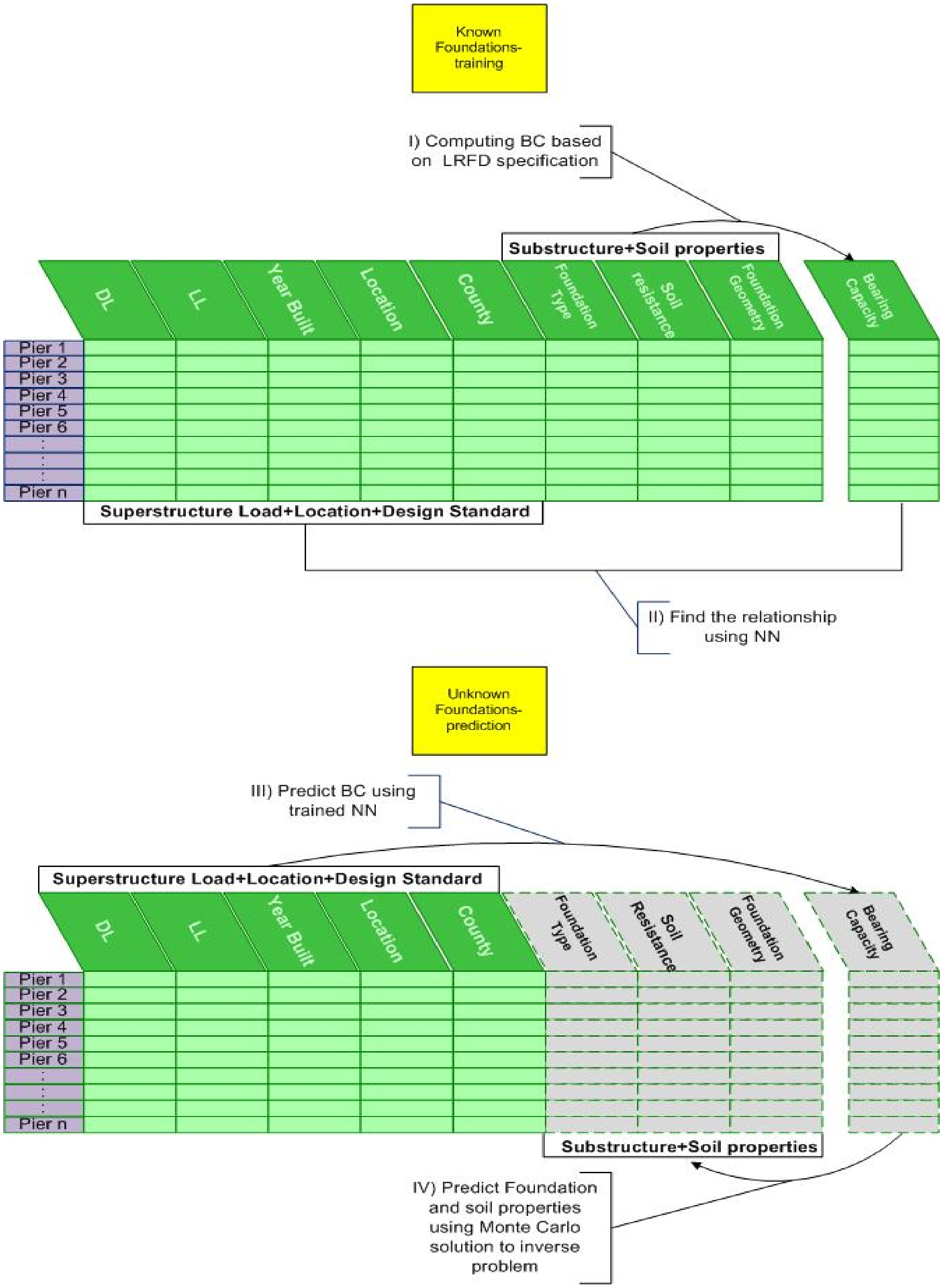


Figure 6-1. Diagram of the Two Main Steps Required for the Probabilistic Determination of Unknown Foundations

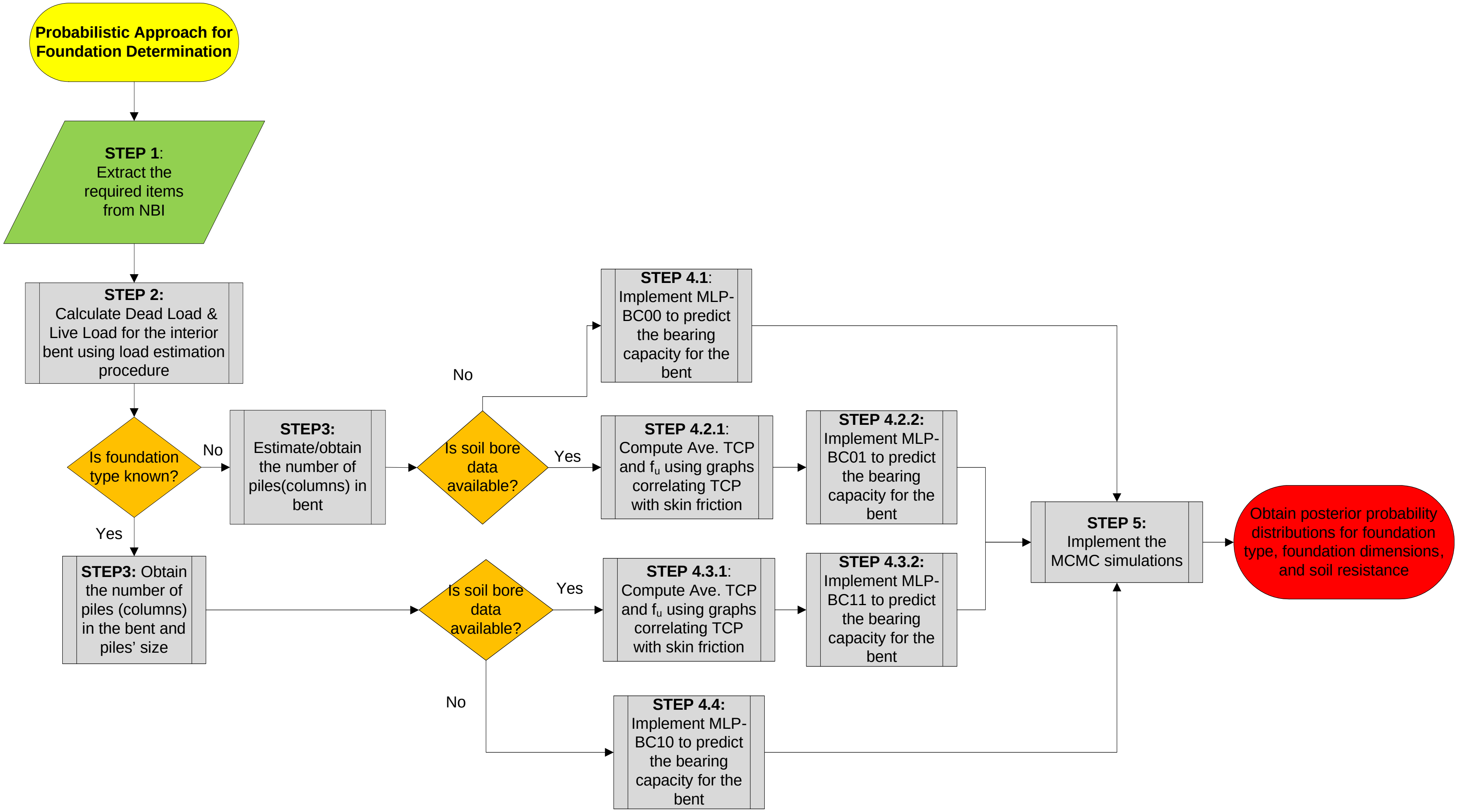
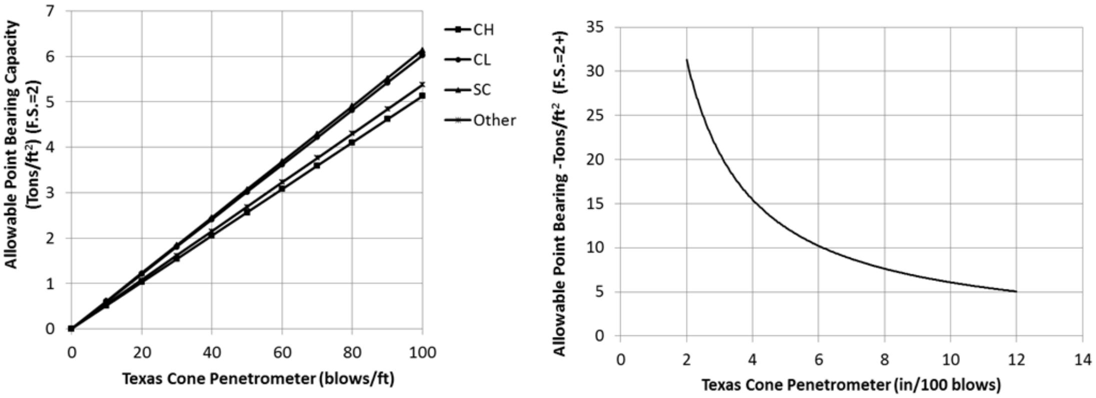


Figure 6-2. Flowchart of the Probabilistic Approach

The soil boring data were found in the bridge inspection folders and bridge plans, and provides the type and strength of soil based on the results of the Texas Cone Penetration (*TCP*) test. The ultimate skin friction (f_u) and point bearing (q_p) of the piles were computed from the *TCP* values using the correlation graphs (see Figure 5-5 and Figure 6-3) provided by TxDOT Geotechnical Manual.



a) **b)**

Figure 6-3. (a) Point Bearing versus *TCP* for a) *TCP < 100* blows/ft, and b) for *TCP > 100* blows/ft (TxDOT 2006)

Step 2: Dead Load and Live Load Calculation. The method proposed to estimate with reasonable accuracy the dead loads and live loads for bridge foundations was earlier explained in Chapter V (Briaud et al. 2012). This procedure was designed based on the standard bridge design sheets that the TxDOT uses (TxDOT 2011). The total load for a bridge bent was obtained using the standard tables, depending on roadway width and span length. The live load was calculated based on the standard design vehicles specified by the AASHTO-LRFD. The dead load was then calculated by subtracting the live load from the total load.

Step 3: Obtaining Foundation Properties. In a situation where foundation type is known, the size and number of piles (columns) for a bridge bent are required. If the foundation type is unknown, at least the number of piles per bent needs to be estimated based on the roadway width of the bridge or by other means such as visiting the bridge.

Step 4: Predicting Bearing Capacity Using the ANN Models. Depending on the soil boring data and foundation type information availability, one of the following sub-steps is implemented:

Step 4-1: If the foundation type is unknown and soil boring data is not available, the ANN model MLP-BC00 is recommended to predict the BC for the interior bent of the bridge. It is assumed here that the foundation type cannot be inferred with the desired level of confidence using a classifier model or by visiting the site.

Step 4-2: This step addresses a situation where the foundation type is unknown but the soil boring data is available:

Step 4-2-1: Compute the average ultimate skin friction (f_u) and the average *TCP* value along the piles from the nearest borehole data.

Step 4-2-2: Implement the ANN model, MLP-BC01, to predict the BC of the bridge interior bent.

Step 4-3: This step addresses a situation where the foundation type is known and the soil bore data is available.

Step 4-3-1: Compute the average ultimate skin friction (f_u) and the average *TCP* value along the piles from the nearest borehole data.

Step 4-3-2: Implement MLP-BC11 to predict the BC of the bridge interior bent.

Step 4-4: If the foundation type is known but the soil bore data is not available, then one can implement the ANN model MLP-BC10 to predict the BC of the bridge interior bent.

Step 5: Implementing the Bayesian Paradigm. The probabilistic solution to the inverse problem is implemented by the use of the Bayesian paradigm and MCMC simulations. This process finds samples from predefined distributions for the unknown variables (prior distributions) and generates posterior distributions by keeping only those samples that minimize the difference between BC function and the ANN-predicted value. Finally, this method provides posterior probability distributions for the foundation type, foundation dimensions, and soil resistance, depending on what the unknowns are.

Load and Bearing Capacity Calculations

The load and BC of the foundations of the bridges included in the working database were calculated based on the standards of AASHTO and the TxDOT Geotechnical Manual. As explained earlier, a simple, efficient method was proposed by the authors to obtain the dead and live loads of the bridges based on the standard bridge sheets provided by the TxDOT. As a result of this process, the dead load, live load, and BC were calculated for the known bridges in the database. This database served to train the ANNs to predict the BC.

As of 2000, Load Factor Design (LFD) has been used for on-system bridges in Texas; however, either Allowable Stress Design (ASD) or LFD may have been used for off-system bridges (AASHTO 2002a; TxDOT 2002). The BCs of deep foundations were calculated based on the standards given by the TxDOT's Geotechnical Manual.

However, this manual does not include any standards for the design of spread footings; therefore, the AASHTO LRFD standards (AASHTO 2002b) were followed.

The allowable bearing capacity equations used in this study for the four foundations types are presented below:

$$Q_a = GN \times 0.5(f_u dd_{Cn}(4D_{Cn}) + q_p D_{Cn}{}^2): \text{Concrete Piling} \quad \textbf{(6-1)}$$

$$Q_a = GN \times 0.5(0.7 f_u dd_{Dr}(\pi D_{Dr}) + q_p(\pi D_{Dr}{}^2/4)): \text{Drilled Shaft} \quad \textbf{(6-2)}$$

$$Q_a = GN \times 0.5(f_u dd_{St}(\frac{A_{St}}{3}) + q_p A_{St}): \text{Steel Piling} \quad \textbf{(6-3)}$$

$$\begin{aligned} Q_a = {} & GN \times 0.5 \times B \times L[7.5c(1+0.2\frac{B}{L}) + \frac{1}{2}\gamma' B(e^{\pi \tan\phi_f}\frac{1+\sin\phi_f}{1-\sin\phi_f}+1)\tan\phi_f(1-0.4\frac{B}{L}) \\ & +0.5\gamma' D_f e^{\pi \tan\phi_f}\frac{1+\sin\phi_f}{1-\sin\phi_f}(1+\frac{B}{L}\tan\phi_f)]: \text{ Spread Footing} \end{aligned} \quad \textbf{(6-4)}$$

where,

Q_a (kN) : Allowable Bearing Capacity

GN : Number of Piles/Shafts in a Bent

q_p (kPa) : Point Bearing Capacity of Piles

f_u (kPa) : Skin Friction of Piles

D_{Cn} (m) : Dimension of Squared Concrete Piles

D_{Dr} (m) : Diameter of Drilled Shafts

A_{St} (m^2) : Area of Steel Piles Considering Soil Plug

dd_{Cn} (m) = dp_{Cn}-3 : Discounted Concrete Piles Embedment Depth

dd_{Dr} (m) = dp_{Dr}-3 : Discounted Drilled Shafts Eembedment Depth

dd_{St} (m) = dp_{St}-3 : Discounted Steel Piles Embedment Depth

c (kPa) : Cohesion

B (m) : Width of Spread Footings

L (m) : Length of Spread Footings

D_f (m) : Spread Footings Embedment Depth

ϕ (deg) : Soil Friction Angle

γ' (kN/m^3) : Soil Buoyant Density

A resistance factor of 0.5 was applied to obtain the allowable bearing capacity. For drilled shafts, in addition to the resistance factor, a reduction factor of 0.7 was applied to the skin friction to take into account the soil disturbance during the drilling. Also, for drilled shafts and piles in interior bents of river crossing bridges, the first 3 m of piles were disregarded in the BC calculations. This procedure is due to the potential erosion from scour, future excavation, and seasonal soil moisture variations. Also, the tip area for the steel piling was calculated assuming that a soil plug was formed at the tip. The friction angles of cohesion less soils were obtained using the graphs recommended by the TxDOT's Geotechnical Manual based on the *TCP* value at the bottom of the spread footings. Also, the density of soil was obtained based on the *TCP* values and their relation to the SPT values (Touma and Reese 1972).

ANN Models for Bearing Capacity Prediction

A set of MLP networks were developed and trained using the working database to predict the BC of a bridge's foundation based on the relevant features of that bridge. The networks included one hidden layer containing 20 neurons with tan-sigmoid activation functions. Table 6-1 presents the input parameters for the MLP models. For those cases where the soil properties were not available, the average *TCP* and skin friction were removed from the input parameters. The input parameters selected for BC approximation were those that participated in the design of the bridge foundation, including dead load, live load, soil resistance, year built, and location. The longitude and latitude of the bridges correlated with the soil characteristics and were particularly informative when the soil bore data was not available. Also, the year built and county of location took into account the typical construction and design criteria of bridges in different counties throughout the years.

Table 6-1. Input Parameters of the MLP Models for BC Prediction

Parameter	Min	Max	Mean	Std. Dev.	Median
Load Elements					
Dead Load (per bent) (kN)	562.26	10324.80	1824.22	1687.44	1312.40
Live Load (per bent) (kN)	229.53	1323.90	449.05	223.89	409.91
Soil Properties					
Ave. *TCP* (blows/ft)	10.50	1414.75	140.83	149.86	107.40
Skin Friction (kPa)	28.73	1942.97	207.54	170.28	166.62
Location and Time					
County	21	239	141.58	69.82	154
Year Built	1925	2008	1968	25	1963
Latitude (deg)	30.06	31.97	30.87	0.45	30.83
Longitude (deg)	95.42	97.25	96.24	0.46	96.16

The MLP networks weights and biases were adjusted by BP algorithm using the LM optimization method. The performances of the networks were evaluated using R^2 and *RMSE*. The random subsampling method was performed by iteratively sampling data points for training, validation, and test data sets, generating a new network each time. The networks were trained using the training datasets, and their performances were monitored using the validation datasets. The training was stopped following the criteria described earlier in Chapter II. The performances of the networks were examined using the test datasets that had not been presented to the networks. Ten thousand networks were generated using the random subsampling method to form an ensemble of MLP networks for the prediction of the BC. The networks with $R^2 < 0.8$ over the test datasets were removed from the ensemble.

Ten ensembles of the MLP networks were developed to predict the BC based on the availability of soil and foundation type information, named as MLP-BC *11*, *10*, *01* (Conc, DrSh, Steal, Spread), *00* (Conc, DrSh, Steal, Spread). The first index represented the information regarding foundation type and the second index represented the soil bore data availability. An index value of *0* implied that the information was unavailable for the corresponding parameter or that it was unknown, whereas an index of *1* implied the opposite case.

Bayesian Inference

As explained earlier in Chapter III, Bayesian inference is a methodology used to estimate the probability of a vector of variables when new evidence becomes available. In this study, Bayesian inference method was coupled with MCMC method to generate posterior probability distributions for the unknown parameters of a bridge foundation. The data collected from the bridges in the Bryan District served to construct prior probability distributions for the unknown parameters, and the BC of the foundation predicted using the ANN models was considered as a measurement or observation.

The Metropolis algorithm was implemented and the candidate samples were drawn from a multivariate normal distribution. The convergence of the Markov chains were evaluated using diagnostic plots, including plots of the samples' cumulative mean and standard deviation with respect to the sample sequence. Also, the autocorrelation of the sample chain and the acceptance rate of the sampler were observed to ensure that the chain was not sticky and had a good mixing.

Model Parameters and Prior Distributions

Four types of problems were defined based on possible scenarios of the foundation type and soil characteristics information, including problems *11, 10, 01,* and *00*. For each of the four scenarios, the unknown parameters associated with the bridge foundation and soil were considered to be model parameters and defined as random variables (see Table 6-2).

Table 6-2. Problem Types and Unknown Parameters

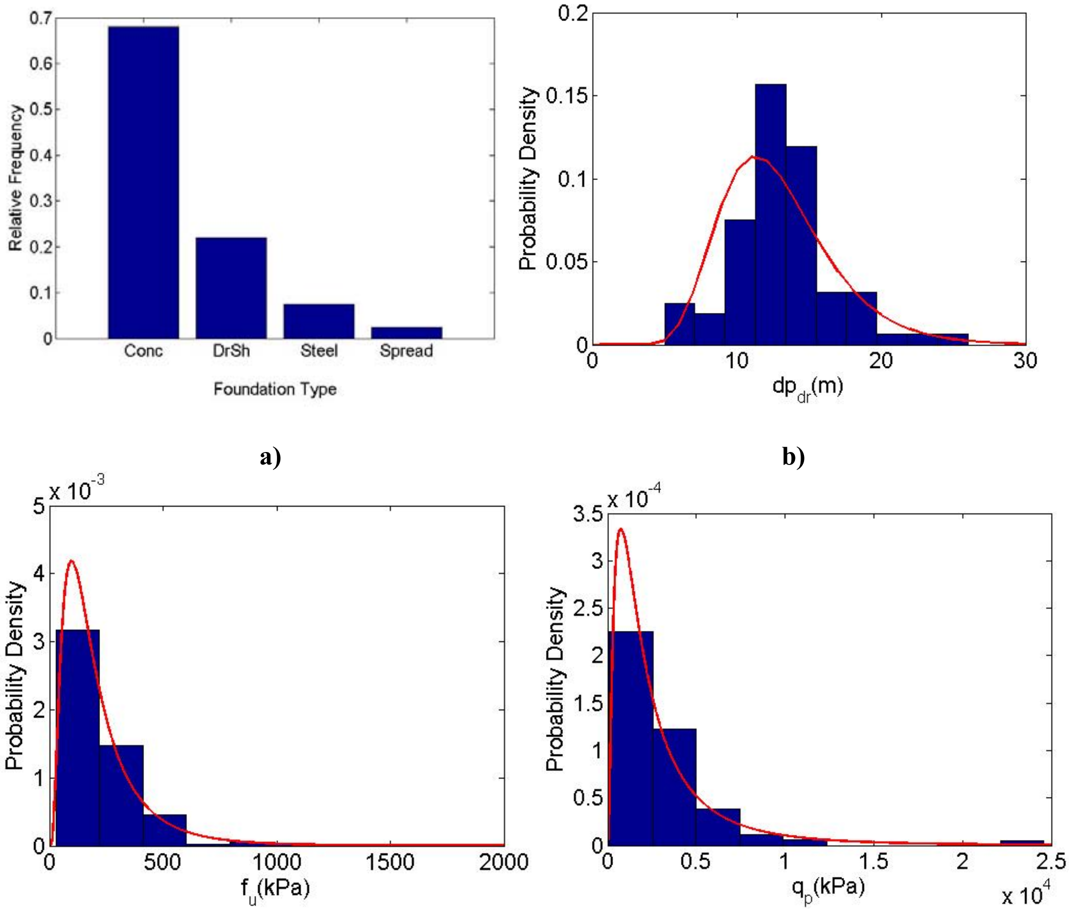
Problem	**Model Parameters**
Problem11-Conc	dp_{Cn}
Problem11-DrSh	dp_{Dr}
Problem11-Steel	dp_{St}
Problem11-Spread	B, L, D_f
Problem10-Conc	dp_{Cn}, f_u, q_p
Problem10-DrSh	dp_{Dr}, f_u, q_p
Problem10-Steel	dp_{St}, f_u, q_p
Problem10-Spread	B, L, D_f, C, φ
Problem01	$^*FT, dp_{Cn}, D_{Cn}, dp_{Dr}, D_{Dr}, dp_{St}, A_{St}, B, L, D_f$
Problem00	$FT, dp_{Cn}, D_{Cn}, dp_{Dr}, D_{Dr}, dp_{St}, A_{St}, B, L, D_f, f_u, q_p, \varphi$

*FT: Foundation Type

The lognormal distribution was considered to be the prior distribution of the model parameters (except for foundation type) according to their data distributions. Also, for all problem scenarios, the unknown parameters were assumed to be independent. Table 6-3 presents the prior distribution of the unknown parameters. The histograms and the fitted prior distributions of some of the model parameters are presented in Figure 6-4. The histograms were created using the database of bridges in the Bryan district.

Table 6-3. Definition of the Model Parameters and Their Corresponding Prior Distributions

Parameter	**Explanation**	**Prior**
Foundation Type		
FT	1 = Conc, 2 = DrSh, 3 = Steel, 4 = Spread	Discrete distribution
Concrete Piling		
dp_{Cn} *(m)*	pile depth for concrete piles	Lognorm (2.07,0.30)
D_{Cn} *(cm)*	dimension of pile (cross section)	Lognorm (3.64,0.06)
Drilled Shafts		
dp_{Dr} *(m)*	pile depth for a drilled shafts	Lognorm (2.53,0.30)
D_{Dr} *(cm)*	diameter of drilled shafts	Lognorm (4.41,0.17)
Steel Piling		
dp_{St} *(m)*	pile depth for steel piles	Lognorm (1.97,0.26)
A_{St} (cm^2)	cross section area for steel piles	Lognorm (6.80,0.08)
	Spread Footing	
B *(m)*	width of a spread footing	Lognorm (0.45,0.20)
L *(m)*	length of a spread footing	Lognorm (2.09,1.13)
D_f *(m)*	embedment depth for spread footings	Lognorm (0.77,0.86)
Soil Resistance		
f_u *(kPa)*	skin friction	Lognorm (5.13,0.74)
q_p *(kPa)*	point-bearing capacity	Lognorm (7.58,0.98)
φ *(deg)*	friction angle of soil (for spread footings)	Lognorm (3.70,0.21)



a) b)

c) d)

Figure 6-4. Histograms and Prior Distributions of a) *FT*, b) dp_{Dr}, c) f_u, and d) q_p

Likelihood Function

Likelihood was defined as the probability distribution of the model prediction error, and error was defined as the difference between the predicted BC by the ANN models and the estimated BC from the analytical equations (Q_a). The Probability Density Function (PDF) of the error was considered to be a normal distribution. Therefore, the likelihood function can be written as:

$$
\begin{aligned}
&Err = BC_{ANN} - Q_a \\
&Err \sim N(\mu, \sigma) \\
&P(D \mid \theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\left[(BC_{ANN} - Q_a) - \mu\right]^2}{2\sigma^2}}
\end{aligned}
\qquad \textbf{(6-5)}
$$

where μ is the mean and σ is the standard deviation of the error distribution. Note that the error function depends upon the BC equation; thus, four different likelihood functions were incorporated for the four types of foundation. Figure 6-5 shows the Cumulative Density Function (CDF) of the prediction error generated by MLP-BC10-DrSh and MLP-BC11-DrSh models.

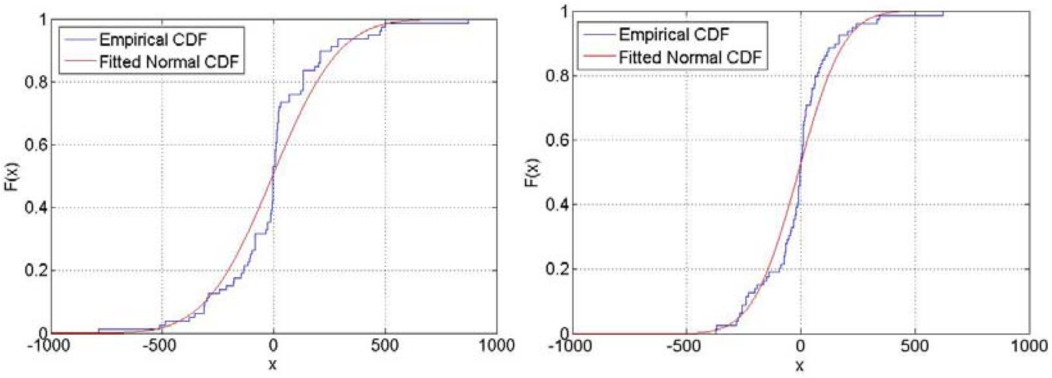


Figure 6-5. Empirical CDF of the Error Distributions for Left) MLP-BC10-DrSh, Right) MLP-BC11-DrSh

Results and Discussion

ANN Models for Bearing Capacity Prediction

The ensemble predictions, as well as the average predictions for the bridge piers in the database, are shown in Figure 6-6 for the ten MLP-BC models. Note that the R^2 and *RMSE* presented for each model correspond to the average predictions. Also, the BC represents the allowable bearing capacity of a bridge interior prier (bent). Table 6-4 summarizes the average performances of the MLP-BC ensembles for the train and test datasets. The performances of the models can be compared through the R^2 of the average predictions presented in Figure 6-6. Comparing MLP-BC10 with MLP-BC11, it was observed that adding the soil strength parameters as inputs to the models significantly improved the performances of the models for the four foundation types.

It can be observed that for all ten MLP-BC models, except for MLP-BC-Spread, the order of magnitude of R^2 of average predictions was around 0.8 or above. Due to the limited amount of data for spread footing, the MLP-BC-Spread could not perform as good as the other models.

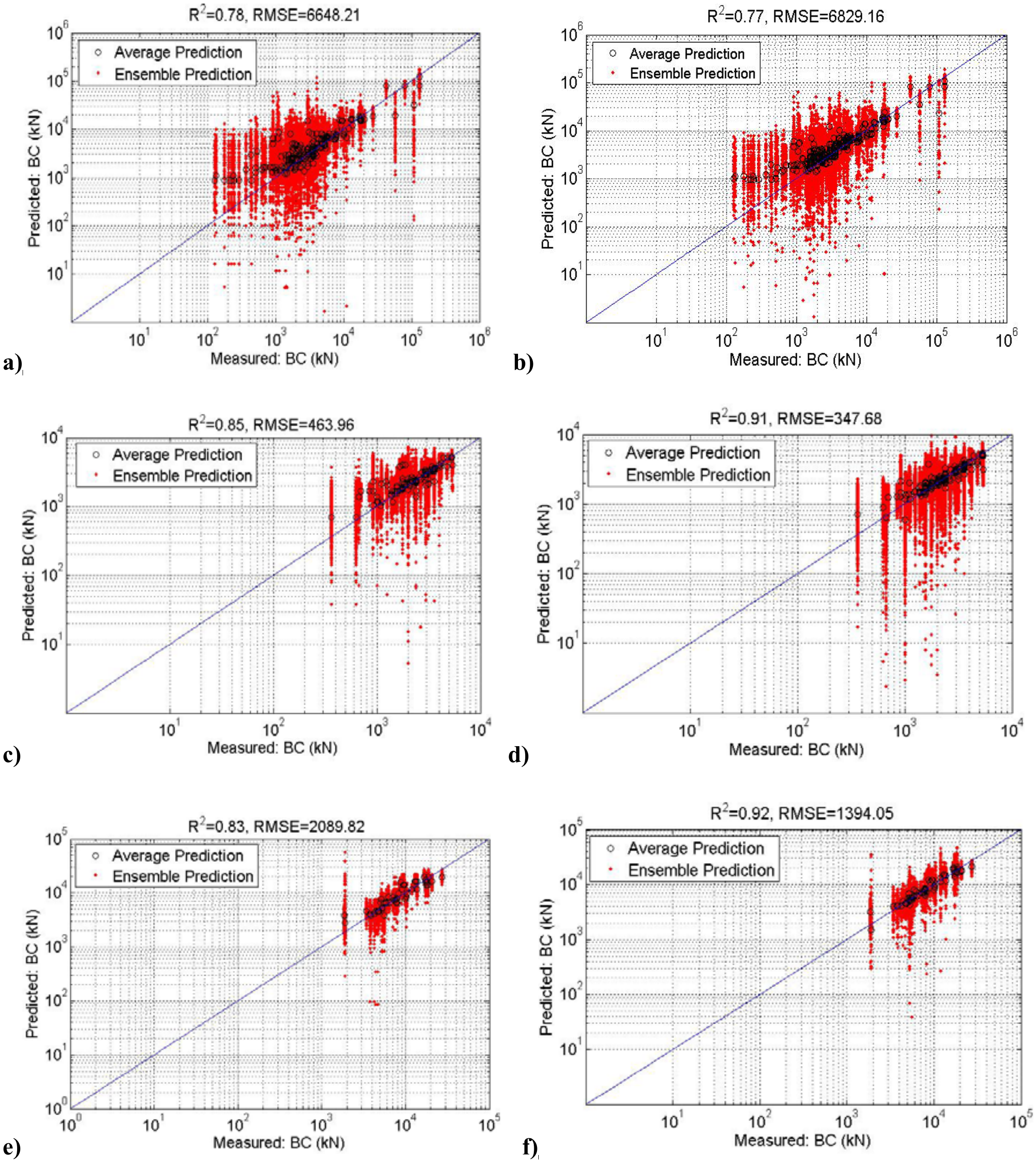


Figure 6-6. a) MLPBC00, b) MLPBC01, c) MLPBC10-Conc, d) MLPBC11-Conc, e) MLPBC10-DrSh, f) MLPBC11-DrSh, g) MLPBC10-Steel, h) MLPBC11-Steal, i) MLPBC10-Spread, j) MLPBC11-Spread (R^2 and *RMSE* Refer to the Average Predictions)

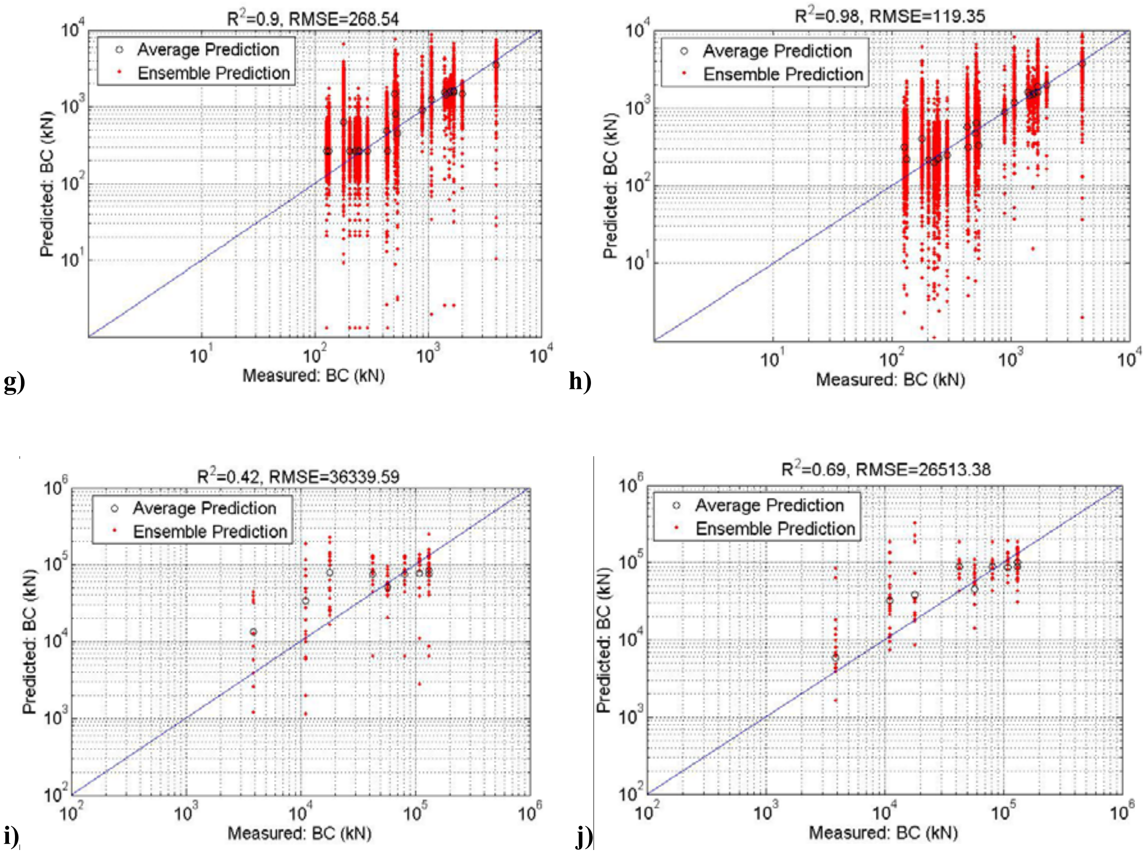


Figure 6-6. Continued

Table 6-4. Performances of the MLP-BC Models for Prediction of Bearing Capacity

Subset Model	No of Data	Train		Test	
		R_{ave}^2	$RMSE_{ave}$	R_{ave}^2	$RMSE_{ave}$
MLP-BC00	279	0.87	4712.11	0.87	5334.54
MLP-BC01	279	0.90	4119.76	0.87	5002.68
MLP-BC10-Conc	278	0.89	439.64	0.88	539.40
MLP-BC11-Conc	278	0.96	265.21	0.88	534.95
MLP-BC10-DrSh	79	0.85	1923.00	0.84	1943.83
MLP-BC11-DrSh	79	0.98	658.56	0.85	2503.87
MLP-BC10-Steel	30	0.92	183.86	0.91	224.45
MLP-BC11-Steel	30	1.00	21.89	0.91	213.59
MLP-BC10-Spread	9	0.78	17250.08	0.89	14433.74
MLP-BC11-Spread	9	0.76	17756.91	0.90	14146.91

Validation of the Bayesian Probabilistic Method

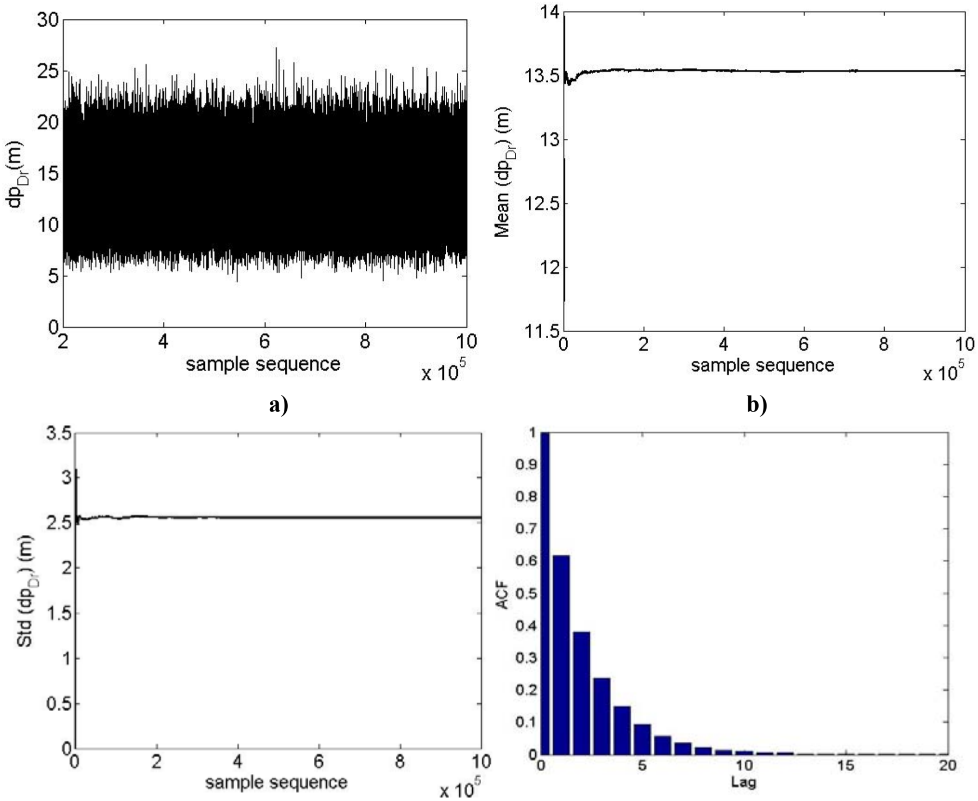
In order to validate the proposed probabilistic approach a case study was performed on a bridge in the Bryan District. The bridge was analyzed under the four possible scenarios of Problems *11*, *10*, *01* and *00*. Pier 2 of the bridge with ID 17-166-0209-05-075 was selected for this purpose. This bridge founded on drilled shafts over Big ELM creek and is located one mile south of FM 485. Table 6-5 shows the actual values of the bridge's parameters. The BC for this bridge was first predicted by the ANN models, and then the foundation and soil characteristics were estimated through the MCMC simulations and compared with the actual values.

Table 6-5. Actual Values of Parameters for Bridge 17-166-0209-05-075 - Pier 2

Parameter	Actual Value
FT	1 (DrSh)
D_{Dr}	76 cm (30 in)
dp_{Dr}	11.28 m (37 ft)
GN	3
q_p	3,255.9 kPa (34 tsf)
f_u	157.1 kPa (1.64 tsf)
BC_{ANN}	6,638.3 kN (746.17 tons)
Q_a	5475.9 kN (615.51 tons)
γ	24.72 kN/m^3 (0.075 tcf)
φ	30 deg

Problem 11: Known Foundation Type and Soil Resistance Parameters

The MCMC simulations were performed for 1,000,000 iterations. Figure 6-7 shows the sample chain for the pile depth (*dp_{Dr}*), along with the cumulative mean and standard deviation of the samples (convergence analysis plots). The autocorrelation function (*ACF*) for the chain is also presented.



a) **b)** **c)** **d)**

Figure 6-7. a) MCMC Samples for *dp_{Dr}*, b) Cumulative Mean, c) Cumulative Standard Deviation, and d) Autocorrelation Function of the Samples - Problem 11

The convergence analysis plots demonstrated that the Markov chain achieved a stationary state, as the samples had a good mixing over the range of the parameter, and the cumulative mean and standard deviation became stationary after a number of initial samples. Also, it was observed that the autocorrelation between the samples was fairly small indicating that the samples were independent and could represent the posterior probability distribution. Figure 6-8 shows the posterior marginal empirical Cumulative Density Function (CDF) of dp_{Dr}. The CDF provides a confidence interval for the prediction of depth. It also allows for determining the probability of exceeding a specific minimum foundation depth (e.g., calculated depth of scour) which is equivalent to the probability that the scour vulnerability of the bridge foundation will be more than a threshold value. Figure 6-9 compares the prior and marginal posterior distributions of the pile depth. It was observed that the effect of the new evidence (foundation BC) on the mean and the uncertainty of the foundation depth were remarkable.

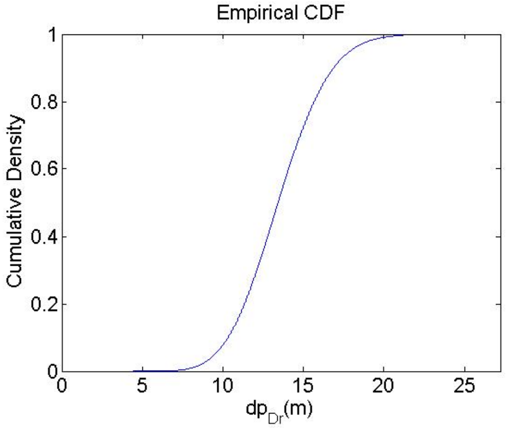


Figure 6-8. CDF of Marginal Posterior Distribution of dp_{Dr} - Problem11

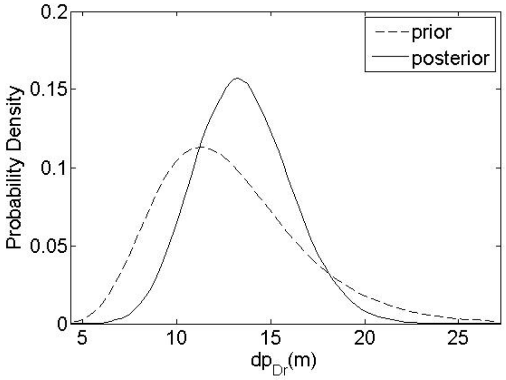


Figure 6-9. Prior and Marginal of the Posterior Distributions for dp_{Dr} - Problem*11*

Table 6-6 presents the expected value (*E*), standard deviation (*Std*), 95% confidence interval ($CI_{95\%}$), mode value (i.e., the most frequently occurring value), and the 10th quantile of the posterior distribution for dp_{Dr}. It was observed that the mean of the marginal posterior distributions for dp_{Dr} were close to the actual pile depth. According to these statistics, the expected value (mean) of the pile depth was 13 m; considering the 10th quantile of the distribution, the probability that the pile depth would be less than 10 m was only 10 %. Thus, 10 m would be a sufficiently conservative estimation for the embedment depth.

Table 6-6. Statistics of the Posterior Distribution of the Model Parameter - Problem11

Statistics	Value
dp_{Dr} (m)	
E	13.53
Std	2.56
$CI_{95\%}$	[8.86 18.84]
Mode	12.90
10^{th} quantile	10.32

Problem 10: Known Foundation Type and Unknown Soil Resistance Parameters

The MCMC simulations were performed for 1,000,000 iterations for the same bridge pier, this time considering the scenario of Problem *10*. Figure 6-10 shows the simulations for all the model parameters (dp_{Dr}, f_u, and q_p).

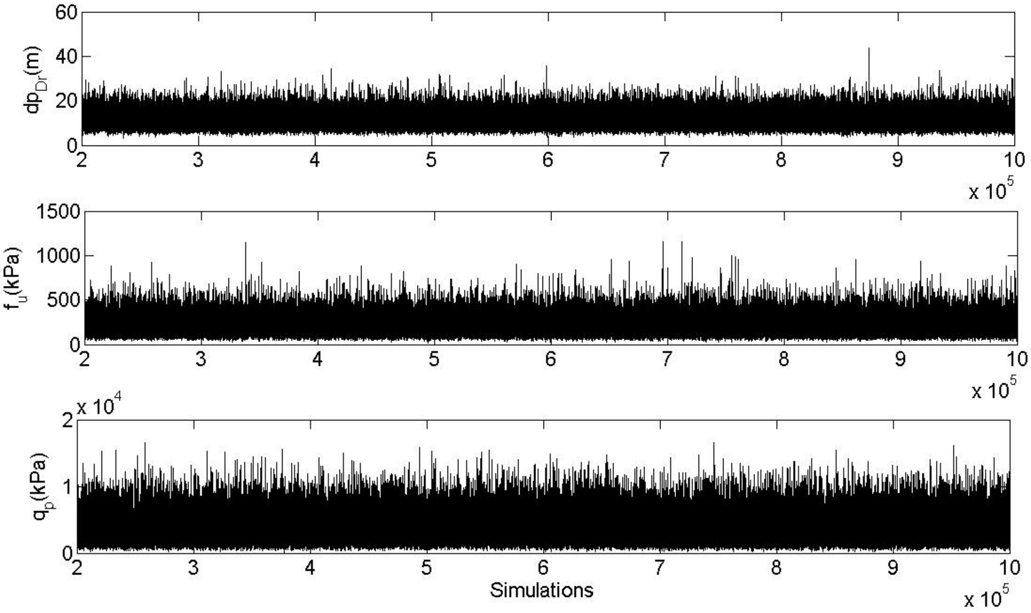


Figure 6-10. Sample Chains for the Model Parameters - Problem10

The convergence analyses indicated that the sample chain converged to the target distribution as the mean, and standard deviation became constant after some point in the sample chain. The posterior CDF of the drilled shaft depth as well as the skin friction and point bearing capacity at the tip of shaft are presented in Figure 6-11. A comparison between the priors and the marginal of the posterior distributions of all the model parameters are presented in Figure 6-12. It was observed that the uncertainty of the parameters estimations was reduced by introducing the BC of the foundation. Also, the maximum likelihood value (mode of the posterior distribution) approached the actual value for all three parameters.

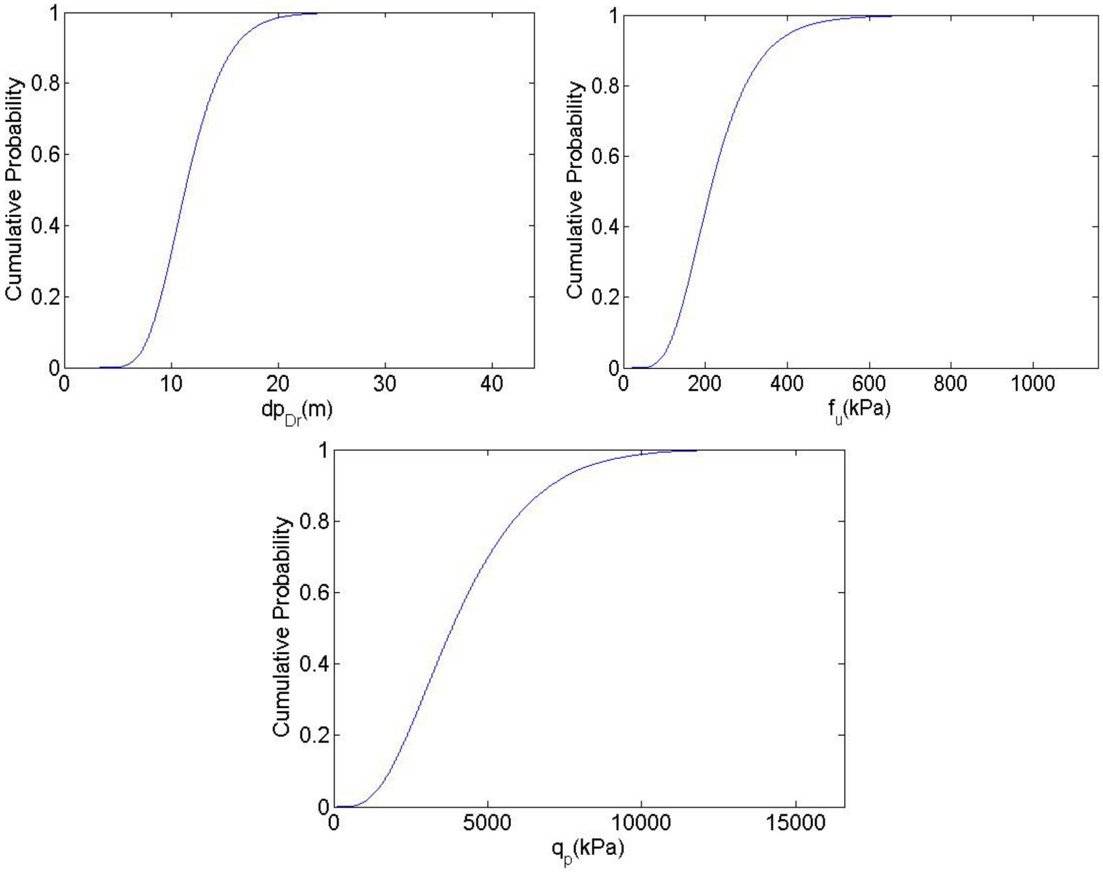


Figure 6-11. CDF of the Marginal Posterior Distribution of dp_{Dr},f_u, and q_p - Problem10

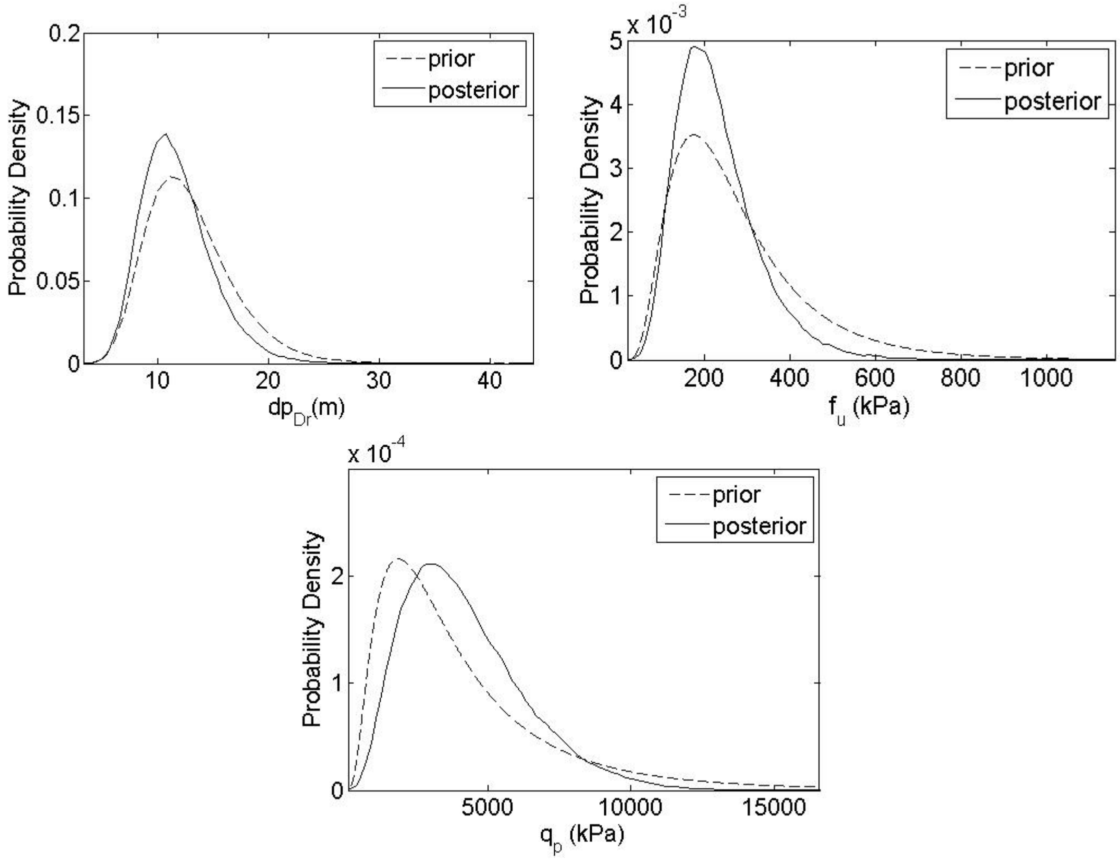
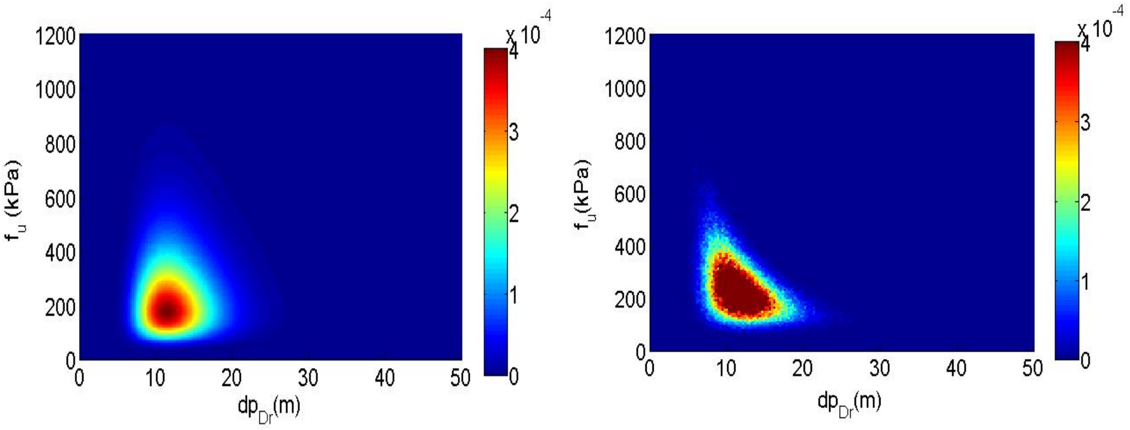


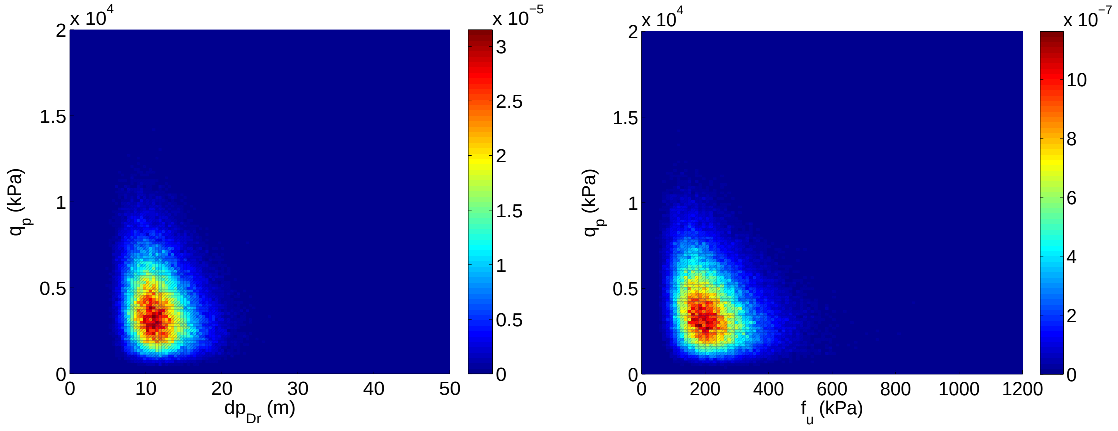
Figure 6-12. Prior and Marginal Posterior Distributions for dp_{Dr}, f_u, and q_p - Problem*10*

Figure 6-13 shows the prior and posterior joint relative frequency density for dp_{Dr} and f_u. The prior joint distribution confirmed that the two parameters were assumed independent; whereas the posterior joint distribution illustrated that there was actually a negative nonlinear correlation between them. Figure 6-14 shows the posterior joint relative frequency for dp_{Dr} versus q_p, and f_u versus q_p.



a) b)

Figure 6-13. Prior (a) and Posterior (b) Joint Relative Frequency Density of *dp_{Dr}* vs. *f_u* - Problem *10*



a) b)

Figure 6-14. Posterior Joint Relative Frequency Density of a) *dp_{Dr}* vs. *q_p*, *b) f_u* vs. *q_p* - Problem *10*

Table 6-7 provides the statistics of the posterior distribution. It was observed that the mean of the marginal posterior distributions for dp_{Dr}, f_u, and q_p were close to the actual values.

Table 6-7. Statistics of the Posterior Distribution of Model Parameters - Problem*10*

Parameter / Statistics	dp_{Dr} *(m)*	f_u *(kPa)*	q_p *(kPa)*
E	11.72	228.87	4194.29
Std	3.16	97.68	2087.57
***CI** $_{95\%}$*	[6.72 18.98]	[90.97 463.48]	[1179.76 9166.15]
Mode	12.56	189.60	3297.97
10th quantile	8.05	123.53	1826.14

Problem 01: Unknown Foundation Type and Known Soil Resistance Parameters

As for Problems *11* and *10*, the MCMC simulations were performed considering the scenario of Problem *01*. The MCMC simulations were performed for 2,000,000 iterations. Figure 6-15 shows the simulations for the ten model parameters. Note that in this scenario, the foundation type (*FT*) was a discrete variable that took values 1, 2, 3 and 4, corresponding to DrSh, Conc, Steal, and Spread, respectively. The value of the *FT* sample at each iteration of the MCMC simulations determined the likelihood function to be used among the four possible functions.

The convergence analysis plots (i.e., the cumulative mean and standard deviation) showed that the chains achieved a stationary state. Figure 6-16 shows the prior and posterior distributions for *FT*. Notice that the relative frequency for DrSh (which is the actual foundation type) significantly increased as compared to the other foundation types, which indicated that the foundation was a drilled shaft.

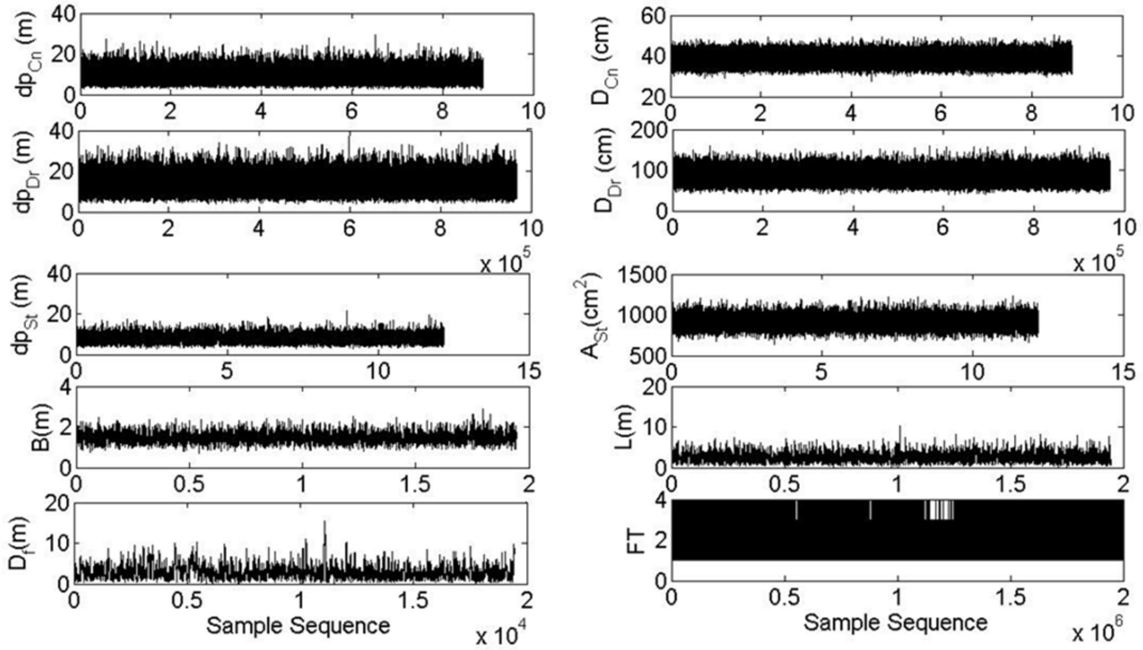
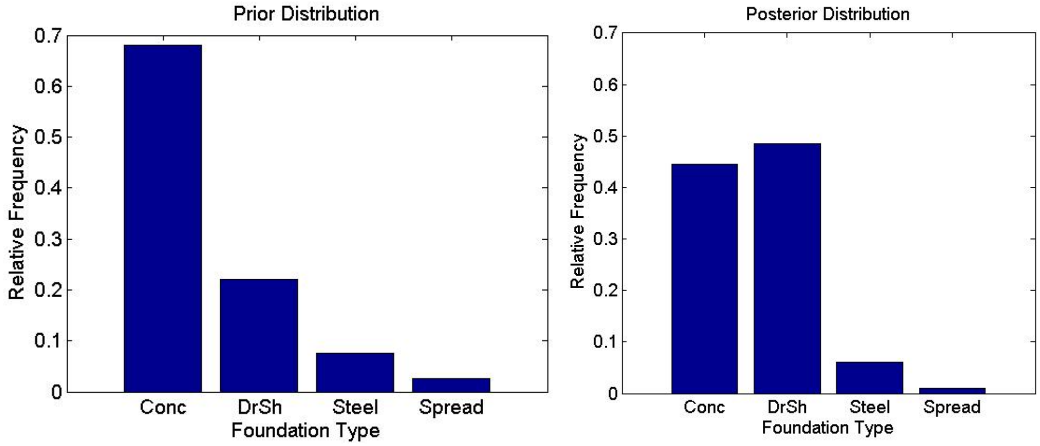


Figure 6-15. Sample Chains for the Model Parameters- Problem *01*



a) **b)**

Figure 6-16. Prior (a) and Posterior (b) Distributions for *FT*- Problem01

The empirical CDF of the posterior distributions for the unknown parameters including the dimensions of the foundations (*dp*$_{Cn}$, *dp*$_{Dr}$, *dp*$_{St}$, and *D*$_f$) are shown in Figure 6-17.

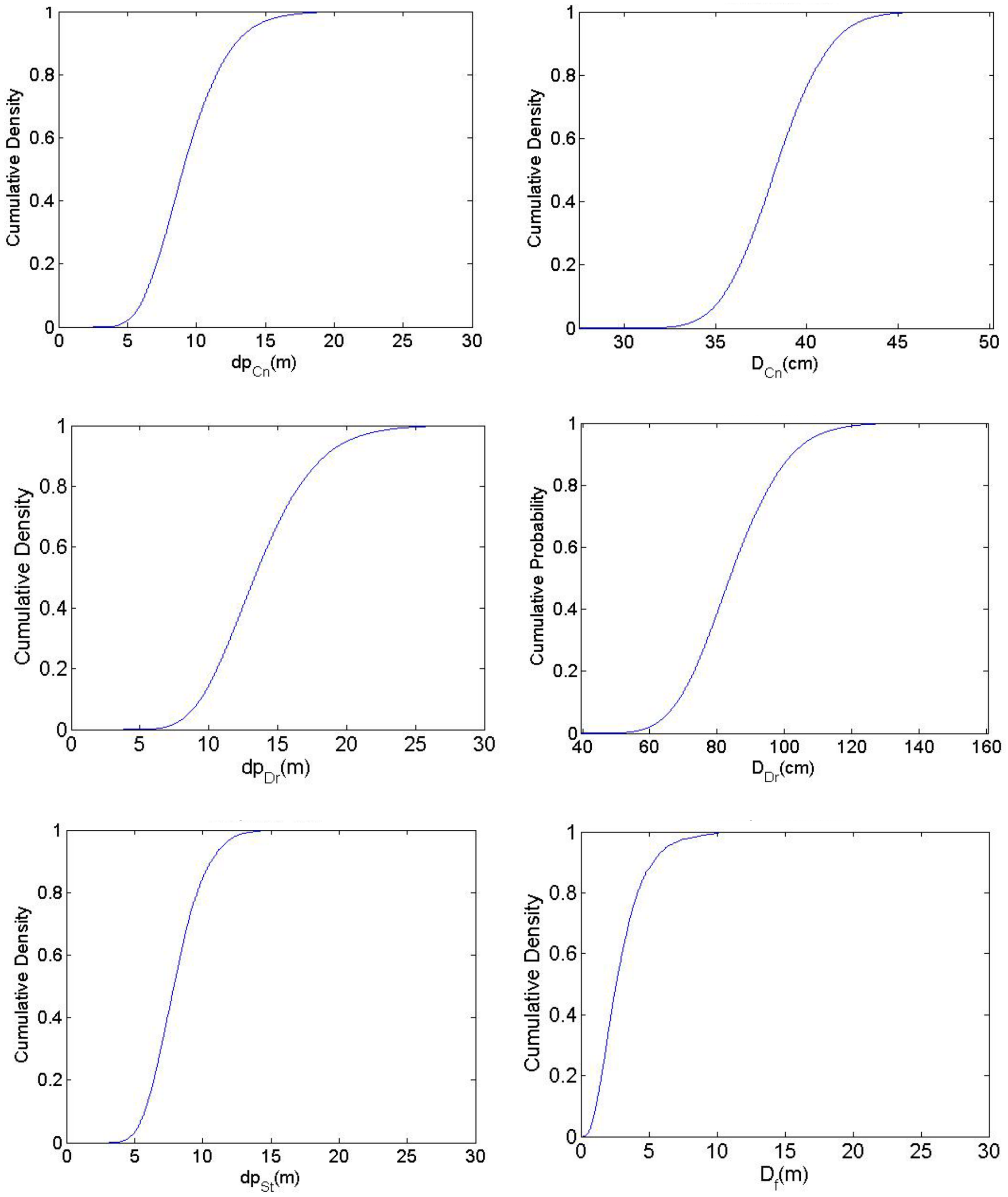


Figure 6-17. Empirical CDF of the Marginal Posterior Distributions of the Foundation's Dimensions for Conc, DrSh, Steal, and Spread – Problem*01*

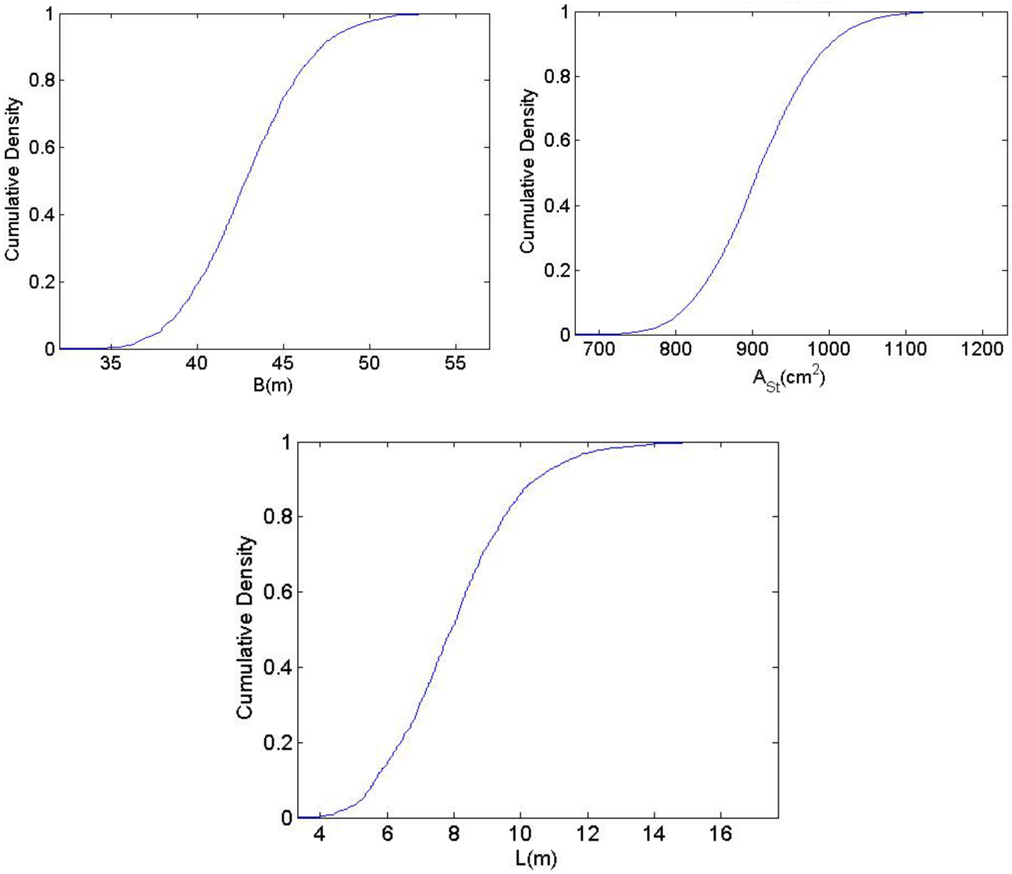


Figure 6-17. Continued

Figure 6-18 shows the prior and the marginal of the posterior PDFs for the depth and size of the drilled shaft. Similar plots were generated for all other model parameters. It was observed that the posterior distributions were more certain (lower standard deviation) compared to the prior distributions. However, the uncertainty reduction was not as significant as those obtained for Problems *11* and *10*. The reason for this could be due to the fact that there was only one data point used as evidence, while there were ten unknown parameters. Also, the large standard deviation of the error distribution made the posterior lean more on the prior distribution. The same simulation with a small standard deviation of error distribution was performed, and a significant decrease was observed in the uncertainty of the parameters.

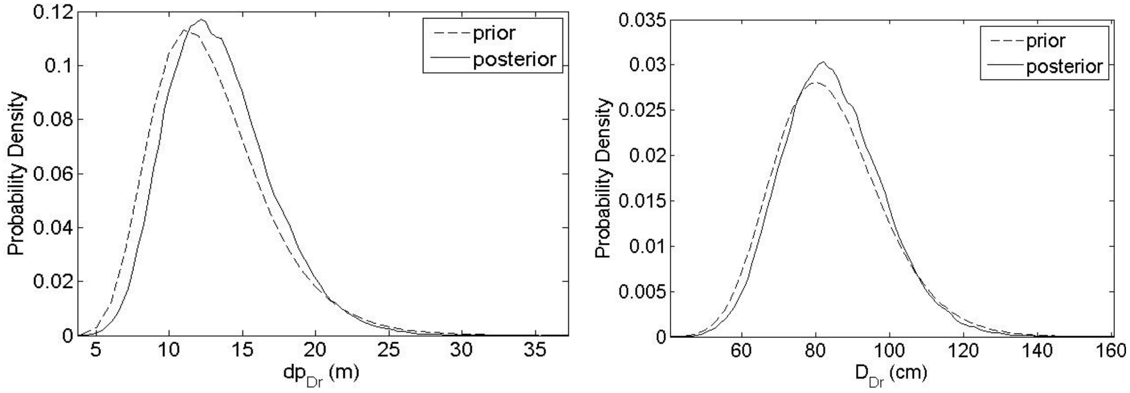
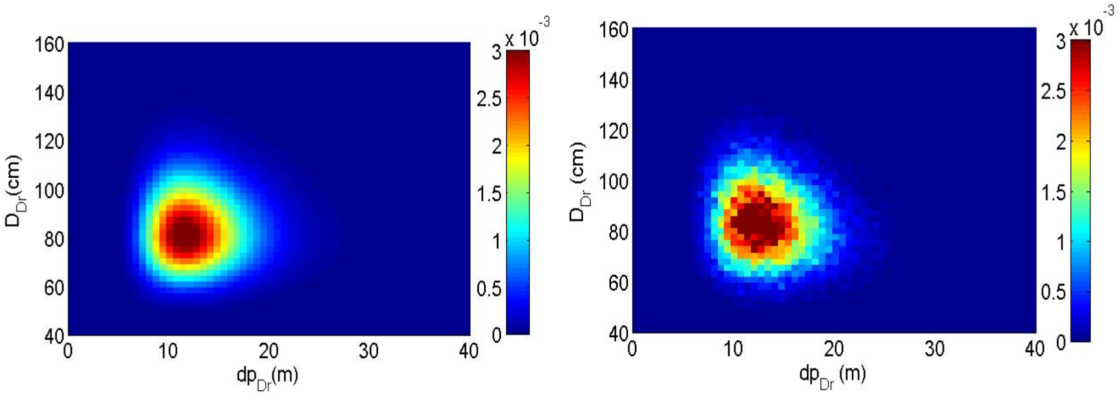


Figure 6-18. Prior (a) and Posterior (b) Joint Relative Frequency Density of dp_{Dr} and f_u - Problem *01*

Figure 6-19 presents the prior and posterior joint relative frequency densities for dp_{Dr} and D_{Dr}. The prior joint distribution demonstrates the independence of the two model parameters, while a slight negative correlation was observed between the dp_{Dr} and D_{Dr} in the posterior joint distribution.



a) **b)**

b)

Figure 6-19. Prior (a) and Posterior (b) Joint Relative Frequency Density of dp_{Dr} vs. D_{Dr} - Problem *01*

Figure 6-20 shows the posterior joint relative frequency density of the foundation dimensions for Conc, Steal, and Spread footing.

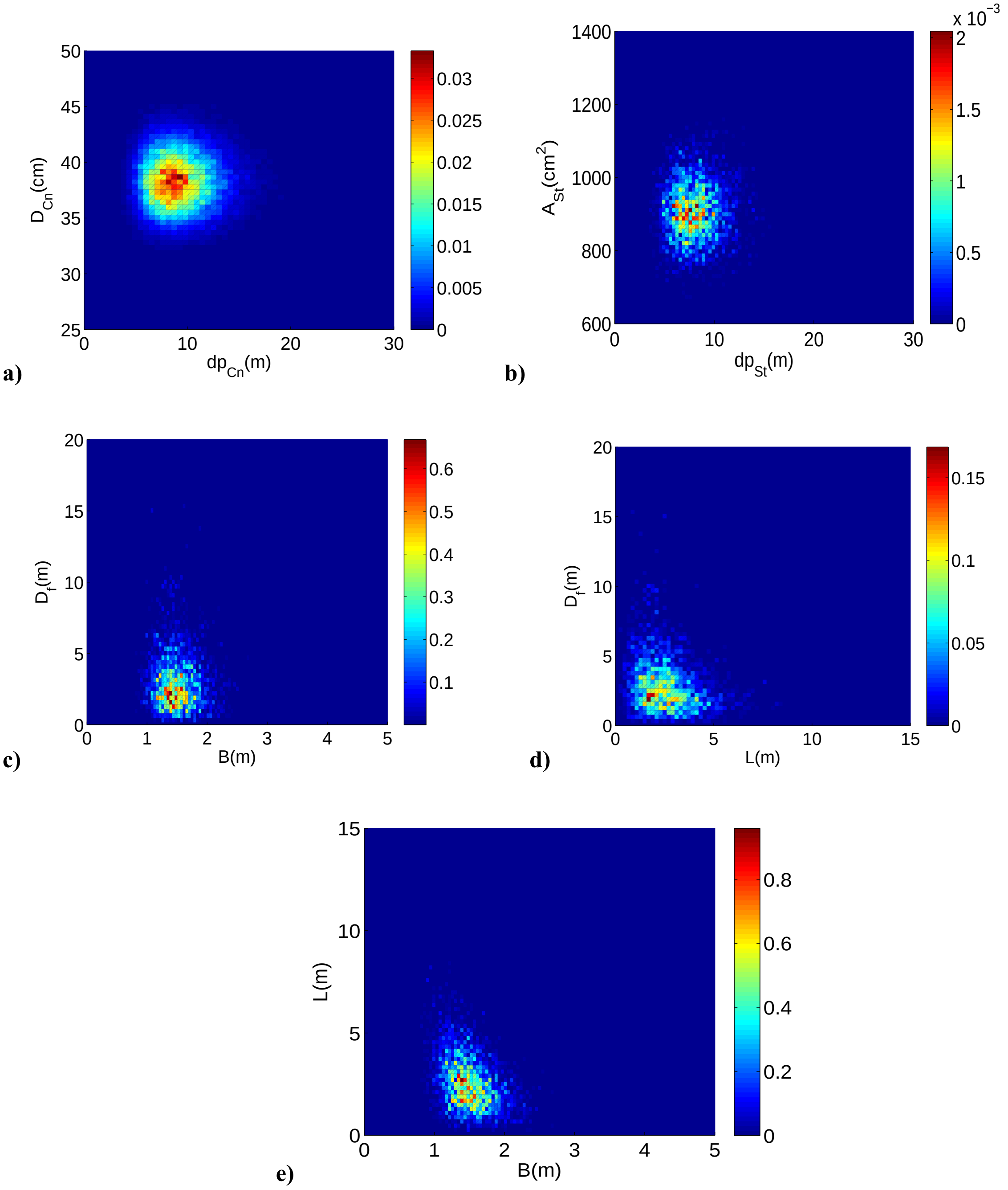


Figure 6-20. Posterior Joint Relative Frequency Density of Foundation Dimensions for a) Conc, b) Steal, and c, d, e) Spread - Problem *01*

Table 6-8 summarizes the statistics of the posterior distributions for the model parameters. It was observed that the mean of the marginal posterior distributions for the dp_{Dr} and D_{Dr} were close to the actual values.

Table 6-8. Statistics of the Posterior Distributions of the Model Parameters - Problem*01*

Parameter / **Statistics**	***FT=1 (Conc)***		***FT=2 (DrSh)***		***FT=3 (Steal)***		***FT=4 (Spread)***		
	dp_{Cn} ***(m)***	D_{Cn} ***(cm)***	dp_{Dr} ***(m)***	D_{Dr} ***(cm)***	dp_{St} ***(m)***	A_{St} ***(cm^2)***	B ***(m)***	L ***(m)***	D_f ***(m)***
E	9.22	38.35	13.47	84.66	7.93	909.68	1.48	2.53	2.89
Std	2.60	2.34	3.53	13.49	1.88	71.16	0.26	1.17	1.78
CI $_{95\%}$	[5.06 15.13]	[33.99 43.15]	[7.67 21.38]	[60.81 113.46]	[4.80 12.05]	[778.45 1056.26]	[1.01 2.07]	[0.74 5.22]	[0.63 7.27]
Mode	9.70	38.38	10.63	80.21	6.25	886.58	1.65	2.57	3.86
10th quantile	6.16	35.38	9.26	68.02	5.66	819.93	1.16	1.17	1.09

Problem 00: Unknown Foundation Type and Unknown Soil Resistance Parameters

Herein, the worst case scenario was assumed where there was no information about the foundation type or soil. 1M samples were generated by MCMC simulations. Figure 6-21 shows the sample chains for the 14 model parameters. The convergence analysis plots showed that the chains for all the model parameters converged to stationary distributions, except for the spread footing parameters. By looking at the *FT* simulations it was observed that only a limited number of samples of *FT=4* (spread footing) were accepted through the Markov chain sampler which indicates that the foundation was very unlikely to be a spread footing.

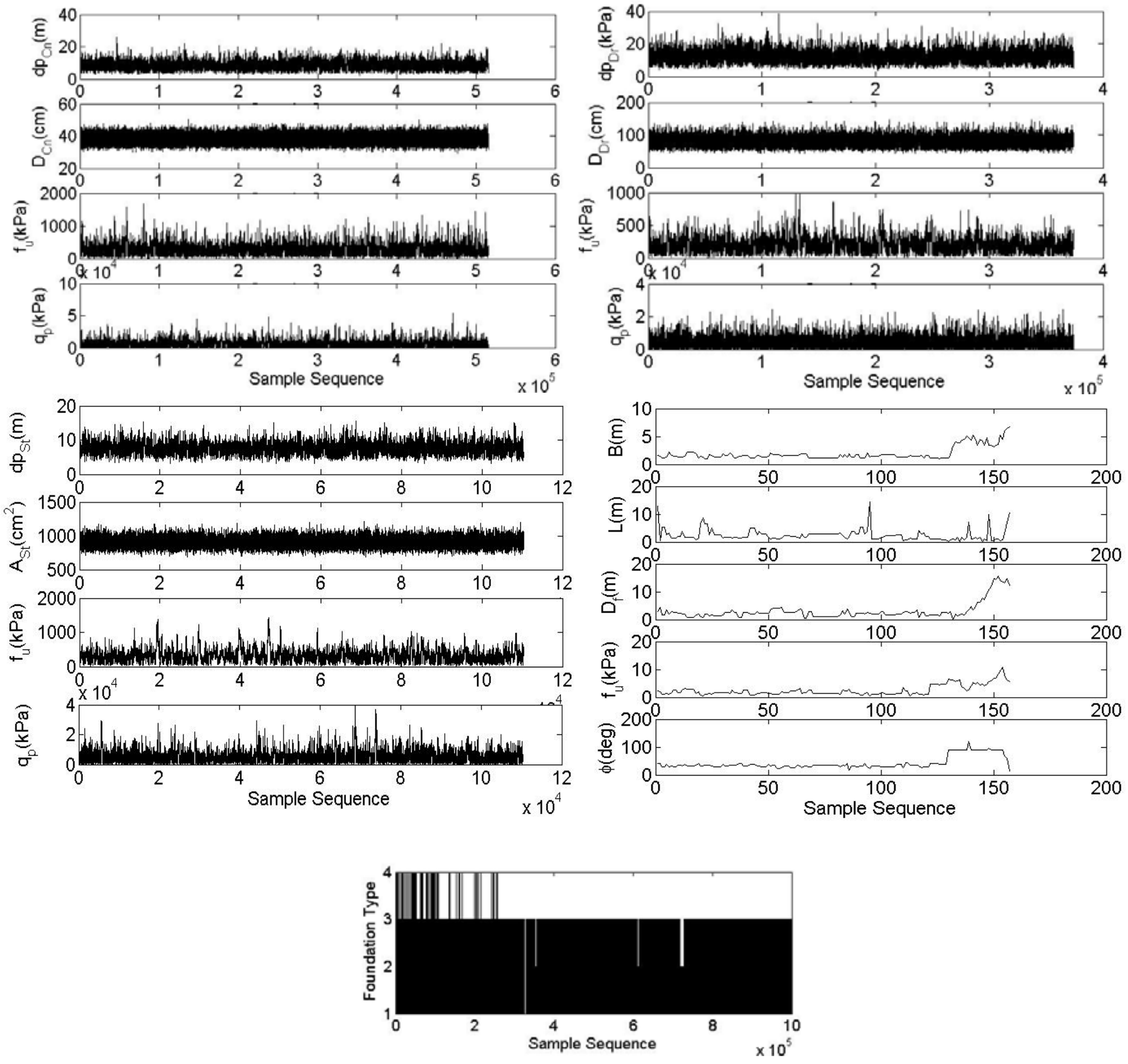


Figure 6-21. Sample Chains for the Model Parameters - Problem *00*

The posterior CDFs of the unknown parameters are provided in Figure 6-22 including the foundation's dimensions for Conc, DrSh, Steal, and Spread, as well as the point bearing capacity and skin friction for DrSh, and the friction angle for Spread.

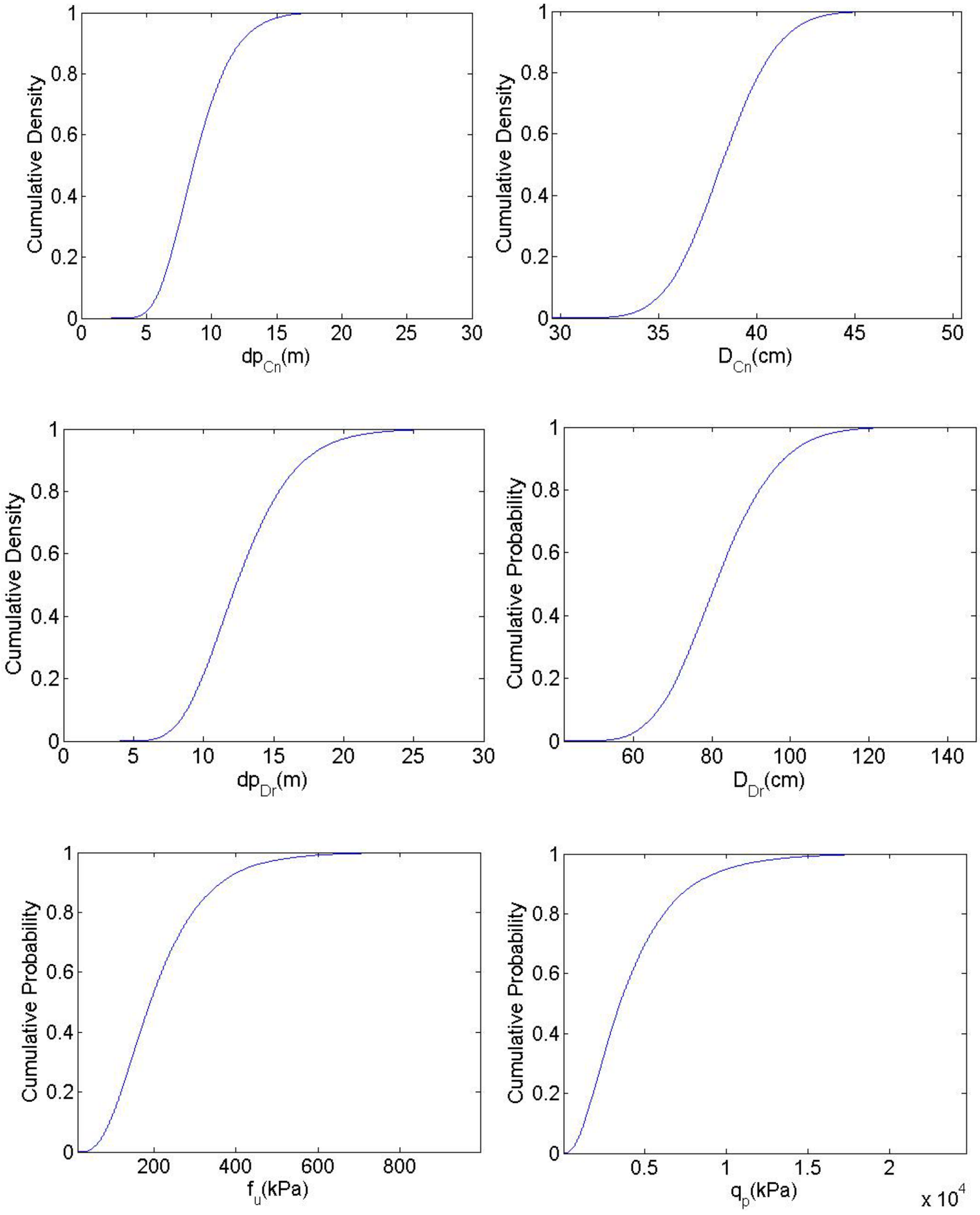


Figure 6-22. Empirical CDF of the Marginal Posterior Distributions of the Foundation Dimensions for Conc, DrSh, Steal, and Spread, and the Soil Parameters – Problem*00*

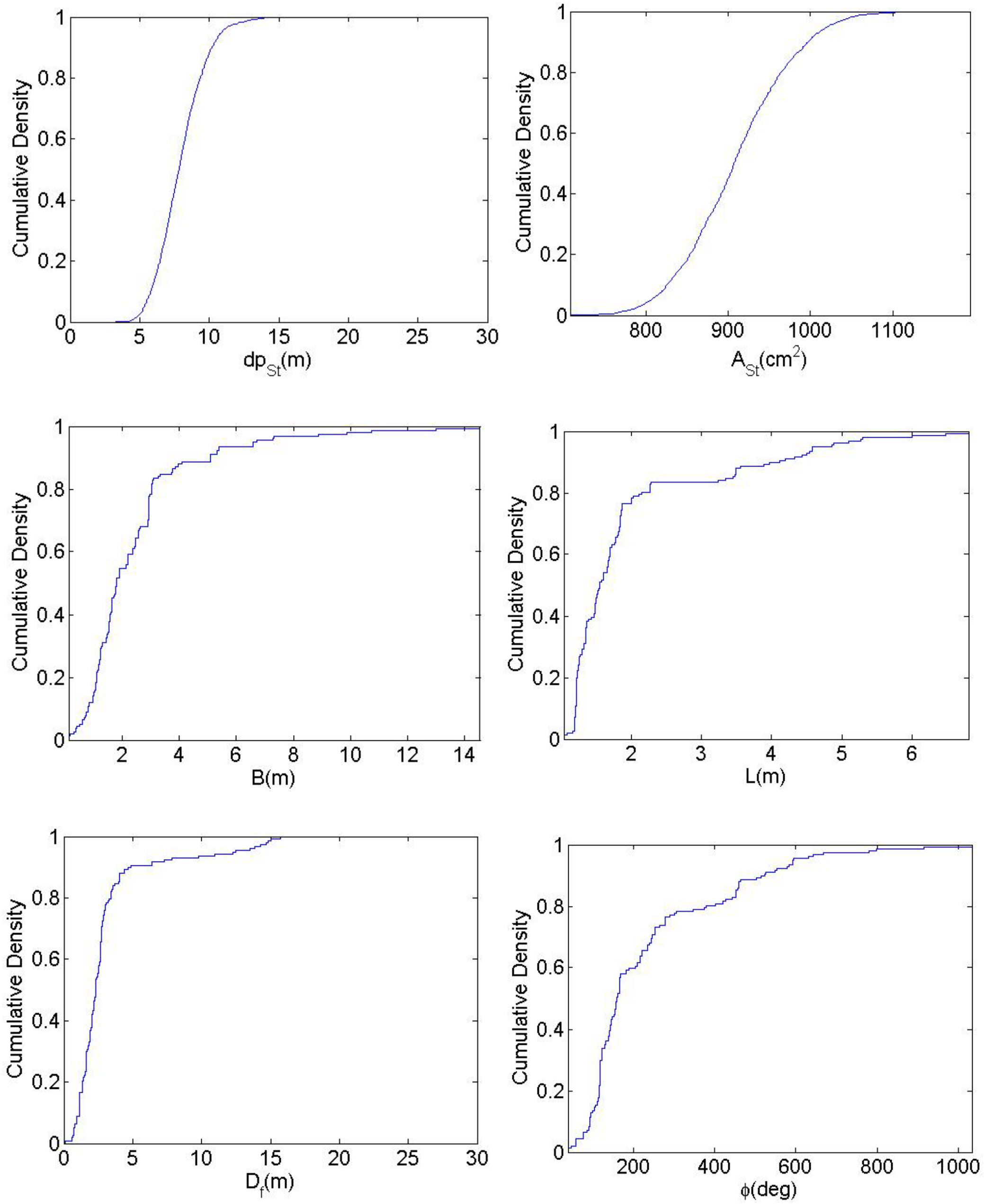
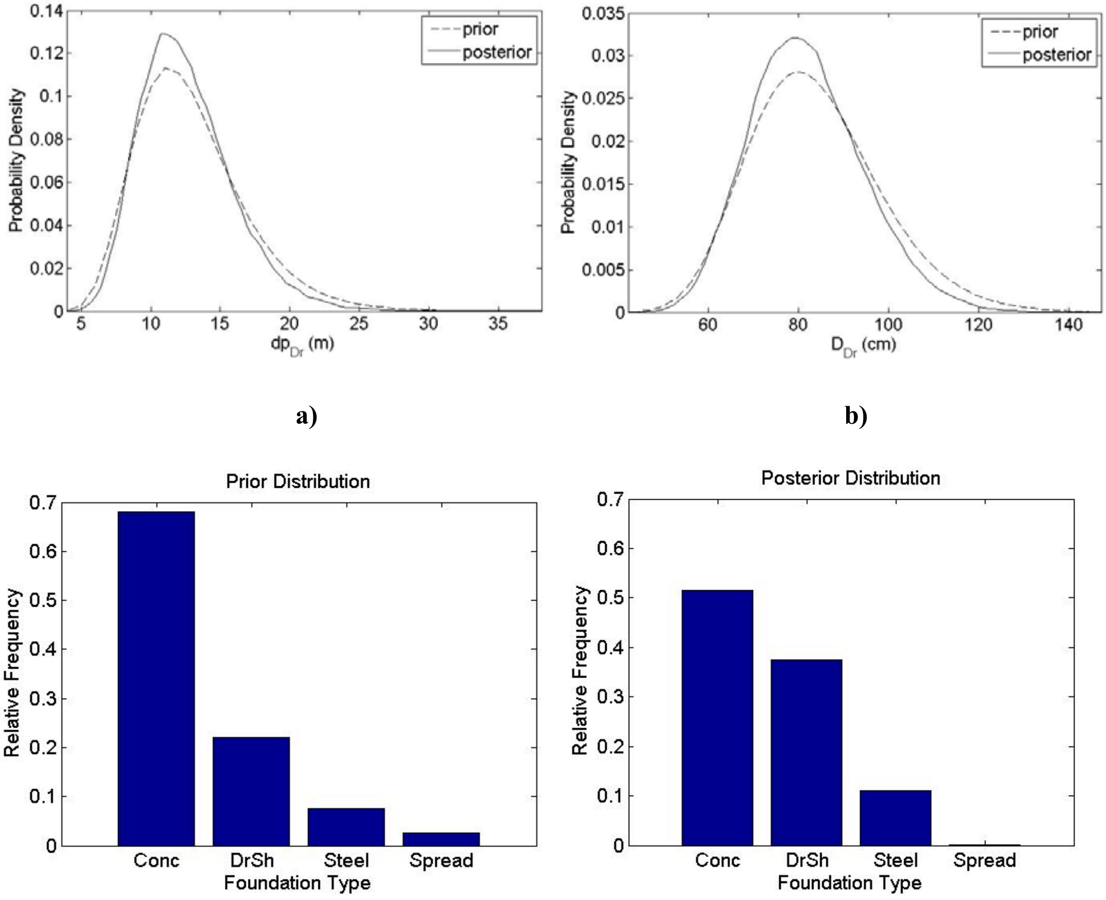


Figure 6-22. Continued

Figure 6-23 compares the prior and posterior marginal distributions for dp_{Dr}, D_{Dr} and *FT*. The increase in the probability of having DrSh in the posterior distribution of the *FT* indicated the actual foundation type. Also, the results of this problem reflected only a small difference between the prior and posterior distributions, due to the total lack of information about the substructure characteristics of this bridge, as compared to other problem types.



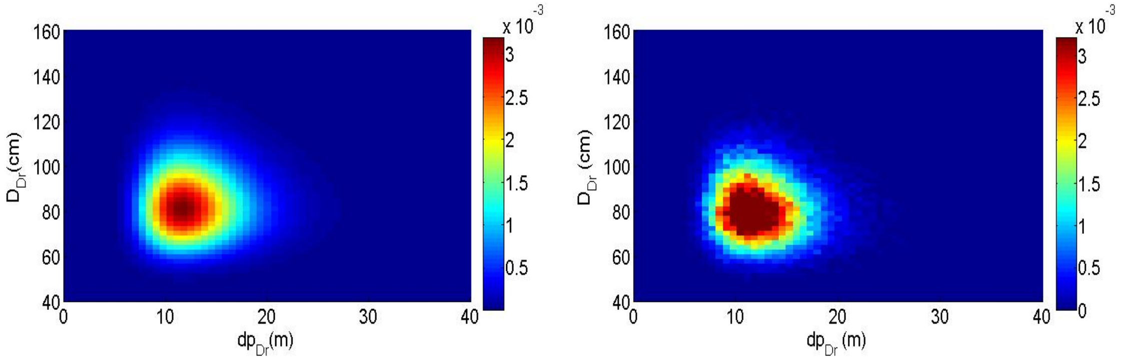
a) **b)**

c)

Figure 6-23. Prior and Marginal Posterior Distributions for a) dp_{Dr}, b) D_{Dr}, and c) *FT* - Problem*00*

The joint relative frequency distributions of the prior and posterior distributions for the D_{Dr} and dp_{Dr} are presented in Figure 6-24. It was observed from this figure that the uncertainty of the parameters was reduced in the posterior, as compared to the prior distributions. Also, a slight negative correlation between the depth and size of the drilled shafts was observed in the joint posterior distribution. Figure 6-25 presents the posterior joint probability distributions for the foundation depth versus foundation size and soil resistance parameters for DrSh, Conc, and Steal. A negative correlation was observed between the depth of the drilled shaft, and the skin friction and point bearing capacity. A similar trend was also observed between the f_u and q_p, which confirmed that in softer layers of soil near surface, foundations need to go deeper to reach harder layers for gaining more BC.

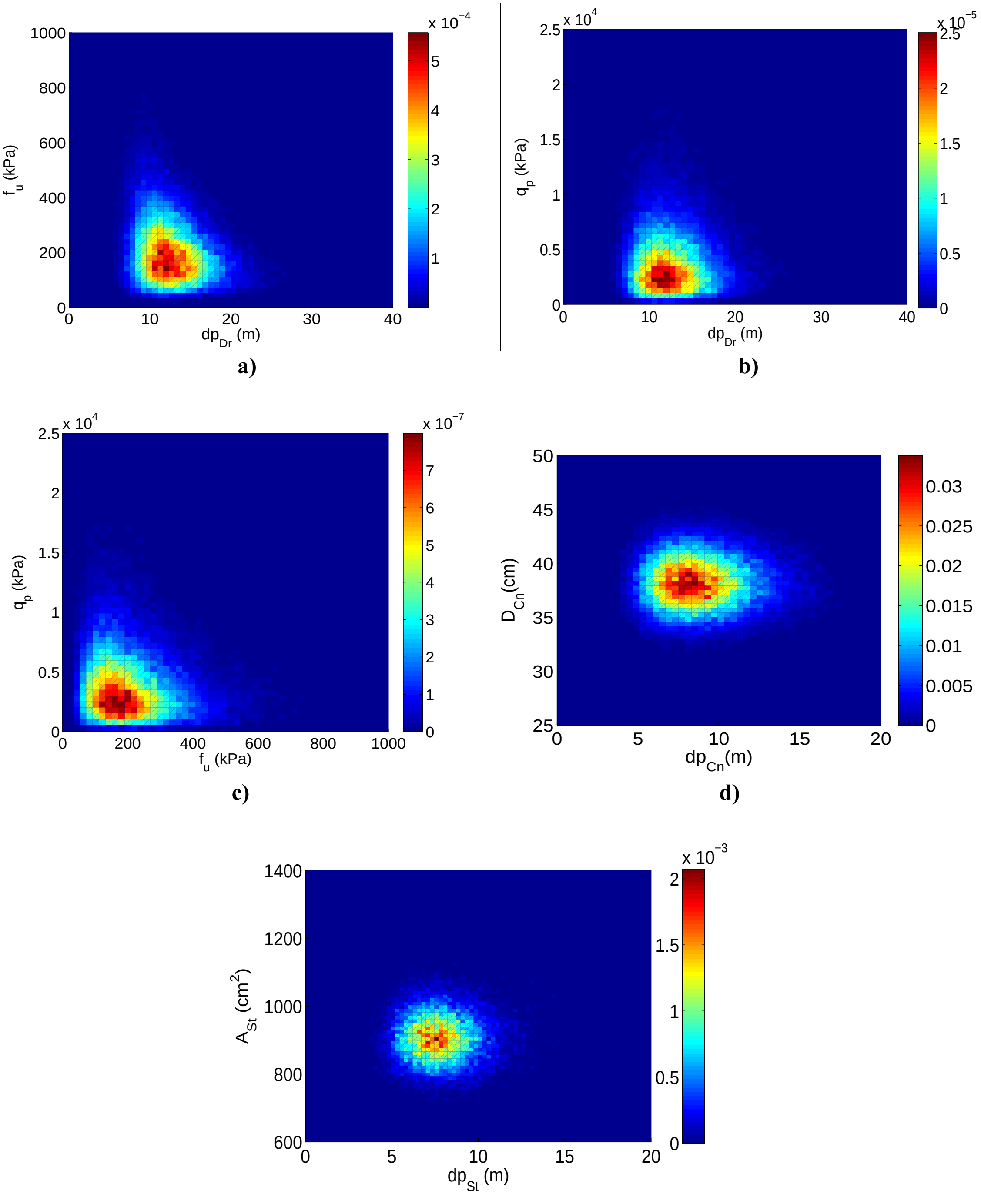
Table 6-9 provides the statistics for the model parameters obtained from the corresponding posterior marginal distributions. Notice that the expected values of the dp_{Dr}, D_{Dr}, f_u, and q_p were fairly close to the actual values for the given bridge.



a)

b)

Figure 6-24. Prior (a) and Posterior (b) Joint Relative Frequency Density of dp_{Dr} vs. D_{Dr} - Problem*00*



e)

Figure 6-25. Posterior Joint Relative Frequency Density of a) Foundation Depth vs. Skin Friction, b) Foundation Depth vs. Point Bearing Capacity, c) Skin Friction vs. Point Bearing Capacity, for DrSh, and Foundation Depth vs. Foundation Size for d) Conc, e) Steal - Problem *00*

Table 6-9. Statistics of Posterior Distributions of the Model Parameters - Problem*00*

Parameter Statistics	dp_{Cn} (m)	D_{Cn} (cm)	f_u (kPa)	q_p (kPa)
		FT=1, Conc		
E	8.75	38.25	310.26	6591.16
Std	2.38	2.26	186.73	5558.87
CI 95%	[4.98 14.3]	[34.04 42.85]	[75.65 788.10]	[874.29 21628.35]
Mode	10.68	35.84	229.82	1691.12
10th quantile	5.97	35.41	118.74	1629.84

FT=2, DrSh

Parameter Statistics	dp_{Dr} (m)	D_{Dr} (cm)	f_u (kPa)	q_p (kPa)
E	12.67	81.81	211.63	4261.32
Std	3.36	12.65	113.00	3245.31
CI 95%	[7.33 20.36]	[59.82 109.09]	[60.33 495.08]	[727.78 12030.33]
Mode	16.94	93.98	119.70	6735.76
10th quantile	8.74	66.29	90.97	1314.78

FT=3, Steel

Parameter Statistics	dp_{St} (m)	A_{St} (cm^2)	f_u (kPa)	q_p (kPa)
E	7.90	907.29	312.18	5640.26
Std	1.66	65.23	160.88	4502.64
CI 95%	[5.10 11.55]	[786 1042.19]	[80.44 710.54]	[817.79 16958.14]
Mode	6.85	907.48	242.27	3354.47
10th quantile	5.86	824.13	130.23	1535.03

FT=4, Spread

Parameter Statistics	*B* (m)	*L* (m)	D_f (m)	f_u (kPa)	φ (deg)
E	1.96	2.45	3.05	238.44	42.33
Std	1.17	2.16	3.05	183.86	22.26
CI 95%	[1.14 5.21]	[0.35 9.30]	[0.65 14.29]	[59.37 730.65]	[22.27 89.98]
Mode	1.19	2.89	1.13	118.74	34.46
10th quantile	1.18	0.80	1.13	93.84	27.84

Conclusions

A probabilistic method incorporating ANNs and the Bayesian inference method was proposed in this study. A step by step guideline also was provided that can be used to infer the unknown characteristics of a bridge foundation based on the bridge's superstructure, load, soil, location, and year built. This method can generate point estimations, as well as confidence interval estimations, for both the foundation and soil characteristics of the bridge. Thus, this method can be used by DOTs to predict the foundation type and foundation embedment depth to determine the scour vulnerability of bridges with unknown foundations in order to reclassify these bridges.

The ANNs proved to be successful in estimating the BCs of bridges with various foundation types based on load, soil resistance, location, and year built. The Bayesian paradigm and the MCMC sampling method were incorporated to quantify the uncertainty in the unknown foundation determination process. This was achieved by conditioning estimates of the characteristics of the unknown foundation on the superstructure characteristics and the estimate of the corresponding BC. The generated marginal posterior CDFs for the unknown parameters allow for the making of probabilistic inferences for these parameters. For instance, using the CDFs it is possible to determine the probability of having an embedment depth less than the calculated depth of scour, and consequently to assess the scour failure risk.

The proposed method was cross-validated by a case study on a bridge located in the Bryan District. The results showed that the actual values of the bridge parameters (assumed to be unknown) were close to the predicted mode and the expected values of the corresponding posterior distributions.

The results of this study showed that the proposed methodology and guidelines could be used to evaluate unknown bridge foundations with a reasonable level of accuracy and a very low cost and therefore, allowing for making better decisions regarding countermeasures and plans of action. This method could be implemented by DOTs in any district so long as the models are trained using the evidence from known foundations characteristics in that region.

CHAPTER VII

UNCERTAINTY QUANTIFICATION FOR ONE-DIMENSIONAL DIFFUSION PROCESSES

Introduction and Literature Review

As discussed in Chapter I, a physical system or phenomena is described through a forward model, which is often formulated in the form of an ordinary differential equation (ODE) or a partial differential equation (PDE). An inverse problem arises when one is interested in estimating the parameters of the forward model given the observations or measurements of the system. Given a value for the vector of model parameters (**m**), the solution of a forward problem is unique. However, given a vector of observations (**d**), often the solutions for the inverse problem is not unique, as many values of the parameters may fit the model through the observations (Kaipio and Somersalo 2005; Tarantola 2005).

In this sense, the uncertainties in the solution of an inverse problem initiates from the measurement (observational) uncertainties and the modeling imperfections (modelization uncertainties) (Tarantola 2005). In a more general categorization, these uncertainties can be either aleatoric or epistemic. Aleatoric uncertainties are defined as the inherent variability in outcome of experiments that cannot be reduced by more accurate measurements. On the contrary, epistemic uncertainties are not inevitable and can be reduced by more accurate models and measurements and therefore, include the uncertainties related to the definition and calibration of forward models.

In this study, the uncertainties are defined by a Gaussian distribution imposing a covariance operator. Thus, the modelization uncertainty is described as (Tarantola 2005):

$$\mathrm{p}(\mathbf{d}\,|\,\mathbf{m}) = \alpha \exp\left(-\frac{1}{2}(\mathbf{d}-\mathbf{g}(\mathbf{m}))^{T}\mathbf{C_T}^{-1}(\mathbf{d}-\mathbf{g}(\mathbf{m})) \right) \tag{7-1}$$

where $\mathbf{C_T}$ is the covariance matrix representing the modelization uncertainty, and α represents the normalizing constant. Likewise, considering $\mathbf{C_d}$ as the covariance matrix, the measurements (observational) uncertainty can be represented as:

$$\mathrm{p}_D(\mathbf{d}) = \alpha \exp\left(-\frac{1}{2}(\mathbf{d}-\mathbf{d}_{obs})^{T}\mathbf{C_d}^{-1}(\mathbf{d}-\mathbf{d}_{obs}) \right) \tag{7-2}$$

where $\mathbf{d_{obs}}$ is the vector of measured (observed) values, and $\mathbf{d}$ represents the true value of the measurements. Using the Bayesian paradigm, as described earlier in Chapter III, the uncertainty of the model parameters can be represented by a posterior distribution that can be written as:

$$\mathrm{p}(\mathbf{m}\,|\,\mathbf{d}) = \alpha\, \mathrm{p}_M(\mathbf{m}) \exp\left(-\frac{1}{2}(\mathbf{g}(\mathbf{m})-\mathbf{d}_{obs})^{T}\mathbf{C_D}^{-1}(\mathbf{g}(\mathbf{m})-\mathbf{d}_{obs}) \right) \tag{7-3}$$

where $\mathbf{C_D} = \mathbf{C_d} + \mathbf{C_T}$, a combination of model and observational uncertainties, and $\mathrm{p}_M(\mathbf{m})$ is the prior distribution over the model parameters' space *M*. It can be shown, mathematically, that this is valid even for nonlinear forward problems (Tarantola 2005).

As opposed to frequentist methods, Bayesian inference allows for incorporating the current beliefs and state of knowledge about the unknown parameters through prior probability distributions. Besides, it is the only method allowing for conditioning on the observations by updating the current beliefs when new information becomes available (Robert 2007). The prior distributions are often selected in an ad-hoc manner or as to simplify the computations. Nevertheless, even if a prior distribution does not exactly reflect our prior information, the corresponding posterior distribution can still provide useful information about the parameter for situations where limited amount of observations are available (Hoff 2009).

Computing the posterior statistics often require solving difficult integrals over the corresponding density function, thus numerical methods are applied to approximate the posterior quantities. These methods include asymptotic, deterministic (e.g. quadrature), and random sampling (Monte-Carlo) methods. However, for non-conjugate prior distributions, the posterior distributions are often computed up to the normalizing constant (due to the complicated integral evaluation), which makes the direct sampling impossible (Hoff 2009). Also, for high-dimensional posterior distributions, Monte Carlo methods are often inefficient. In such cases, the Markov Chain Monte Carlo (MCMC) method is utilized to generate samples and construct the posterior probability distributions of the model parameters (Tarantola 2005; Gamerman and Lopes 2006; Marzouk et al. 2007). This method has been widely applied to solve the inverse problems in variety of fields including geophysics and geotechnical engineering (Wang and Zabaras 2004; Medina-Cetina 2006; Gallagher et al. 2009; Briaud et al. 2011).

The forward model estimation is an essential part in the Bayesian solution of inverse problems. In many problems, especially those described by partial differential equations, the analytical solution is not available. In such cases, numerical solutions are applied to approximate the forward model and give rise to the numerical uncertainty. There are various techniques to propagate the uncertainty from the model parameters to the forward model predictions, including Monte-Carlo (MC) simulation and the stochastic Polynomial Chaos (PC) and Karhunen-Loeve (KL) expansion methods (Marzouk et al. 2007). For complex forward models running the MC uncertainty propagation method can be computationally intensive, as it requires the forward model to be evaluated for each sample. An alternative approach to avoid evaluating the forward model for each sample is to represent the stochastic process using series analogues to Fourier-type such as the KL and the PC expansion methods. The KL method, first derived by Karhunen (1946) and Loeve (1963), expands the process function by the spectral decomposition of the corresponding covariance function of the process. Wiener (1938) first introduced the theory of the Homogenous Chaos and Ghanem and Spanos (2003) implemented the Polynomial Chaos method into the finite element context. This method suggests the spectral expansion of random variables and stochastic processes using a set of orthogonal polynomials and has been successfully applied for uncertainty assessment in various fields including structural mechanics (Ghanem et al. 2000), flow and transport in porous media (Ghanem 1998, 1999; Yang et al. 2004; Dostert et al. 2008), fluid dynamics (Mathelin et al. 2005; Najm 2009), and diffusion problems (Xiu and Em Karniadakis 2002; Elman et al. 2011).

The purpose of this study is to evaluate the epistemic uncertainties initiated from model (parameters) calibration, also known as parametric uncertainty, and the numerical uncertainty (associated with the numerical solution of the forward problem) in the solution of one-dimensional diffusion processes, under varying conditions of observations. The effects of the number of measurements (observations) and the correlation between the measurements on minimizing the uncertainty in the solution of the inverse diffusion problem were investigated.

The ordinary differential equation (ODE) of diffusion equation has been widely applied to model various processes such as radioactive decay, heat transfer, and chemical reactions. In geotechnical engineering in particular, the transport of contaminants in saturated soils is modeled by the diffusion equation, which controls the design of waste containment barriers (Goodall and Quigley 1977; Gillham et al. 1984; Quigley and Rowe 1986; Daniel and Shackelford 1988; Desaulniers and Cherry 1989; Johnson et al. 1989; Shackelford 1990; Shackelford and Daniel 1991a, b; Shackelford 1993) . Also, the movement of water (moisture) in unsaturated soils is described by a diffusion-type water-flow equation (Gardner and Mayhugh 1958a; Gardner and Mayhugh 1958b; Bresler and Hanks 1969; Warrick et al. 1971; Bresler 1973; Goodall and Quigley 1977; Quigley and Rowe 1986; Daniel and Shackelford 1988; Desaulniers and Cherry 1989; Shackelford 1990, 1993).

Three scenarios depicting three different observations conditions were defined, focusing on the effect of number of data and the correlation structure. The data was synthesized using a benchmark model. The inverse problem was then solved based upon both the analytical and numerical solution of the forward model. The numerical uncertainty was evaluated by comparing the corresponding results of the two solutions.

Upon solving the inverse problem, the uncertainty was propagated using the MC method. Figure 7-1 presents the steps of the proposed approach. correlated

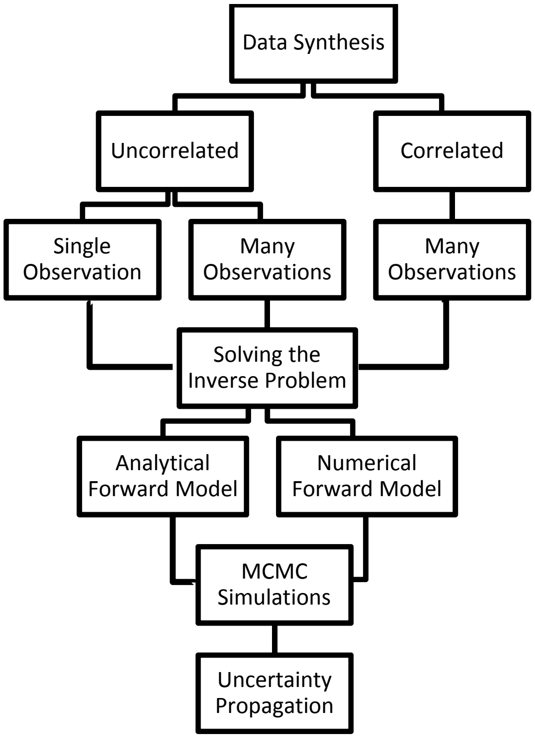


Figure 7-1. Flowchart of the Approach

Problem Definition

The one-dimensional diffusion (decay) process is modeled with the following ODE which serves as the forward model in this study:

$$u'(t) = -cu(t), c > 0 \quad \textbf{(7-4)}$$

$$u(0) = I, I \in R \quad \textbf{(7-5)}$$

where I is the initial condition and c is the coefficient of diffusion. The analytical solution for this ODE can be written as:

$$u(t;I,c) = Ie^{-ct} \tag{7-6}$$

A numerical solution based on the Forward Euler method is introduced to estimate $u_{i+1} = u(t_{i+1})$:

$$\begin{aligned} u(0) &= I \\ u_{i+1} &= u_i - c\Delta t u_i \end{aligned} \tag{7-7}$$

where $t_i = i\Delta t$ and Δt is the time step size, which can be defined as a hyper parameter. Based on this solution, u_k for an arbitrary time step k can be written as:

$$u_k = I(1 - c\Delta t)^k \tag{7-8}$$

The total error between the analytical estimation of the process and the actual value consists of two different error parts (assuming zero measurement error):

$$e_{total} = u_k - U_0 = e_F + e_N \tag{7-9}$$

1. Numerical Error: $u_k - u(T_0;I,c) = e_N$
2. Fitting Error: $u(T_0;I,c) - U_0 = e_F$

where U_0 is the actual or measurement value.

The inverse problem is defined as solving the forward problem given a set of observations $(u_1, \cdots, u_n)$ for *I* and *c*.

Data Synthesis

In this study, the following benchmark model is considered and a set of observations were synthesized based on this model:

$$\begin{aligned} \overline{u} &= 1e^{-0.2t} \\ E(u_i) &= \overline{u} \end{aligned} \tag{7-10}$$

The following domains are assumed for both generating the data and later on, for implementing the Bayesian solution of the inverse problem:

$$\begin{aligned} &0 \le I \le 10 \\ &0 \le c \le 10 \\ &0 \le t \le 5 \\ &0 \le u \le I \end{aligned} \tag{7-11}$$

Table 7-1 presents the definition of the three scenarios of observations conditions that were considered for the assessment of uncertainty. The uncertainty in the solution of the inverse problem were compared for the case with only limited data available (Single Observation) with the case with substantial amount of data (Many Observations). Also, the effect of data correlation on the inverse problem solution was explored when all other conditions remain the same (Correlated Observations). N denotes the number of data per location, T denotes the number of locations in the t domain, L is the location of data (t), and ρ denotes the correlation structure of the data.

In order to investigate the effect of data correlation, the correlation structure of the data is first defined arbitrarily as the following function:

$$\rho(\Delta) = e^{-(1/\tau)\Delta} \tag{7-12}$$

where τ is called the correlation length and Δ is the lag distance for a stationary random process. The correlation function is presented in Figure 7-2 for $\tau = 0.2,\ 1,\ 5$, which represent slightly, moderately and highly correlated structure respectively.

Table 7-1. Definition of the Observational Scenarios

Observational Scenario	Definition
Single Observation (SO)	$N=1$, $T=1$, $L=1.25$
Many Observations (MO)	$N=10$, $T=10$, $0\le L\le 5$, $\tau\approx 0$
Correlated Observations (CO)	$N=10$, $T=10$, $0\le L\le 5$, $\tau=1$

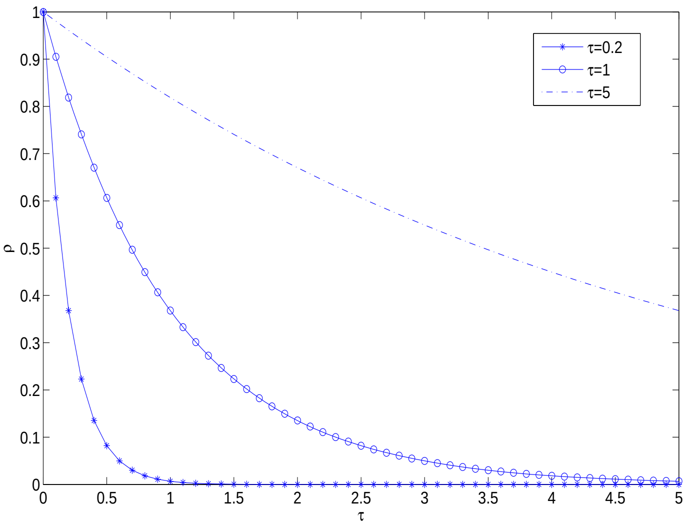


Figure 7-2. Correlation Functions for Different Values of the Correlation Length

Uncorrelated Observations

Data is first generated by drawing independent samples from a normal distribution considering the expected value to be defined by the benchmark model at the specific *t* locations. For cases with multiple data locations, the locations (*t*) are randomly selected from a uniform distribution. The variance of the normal distribution at each location is considered to be equal to 0.01:

$$\{u_1,\cdots,u_n\}\ i.i.d. \sim N(e^{-0.2t_i},0.1) \tag{7-13}$$

Figure 7-3 presents the synthetic data with the benchmark model for the SO and MO cases.

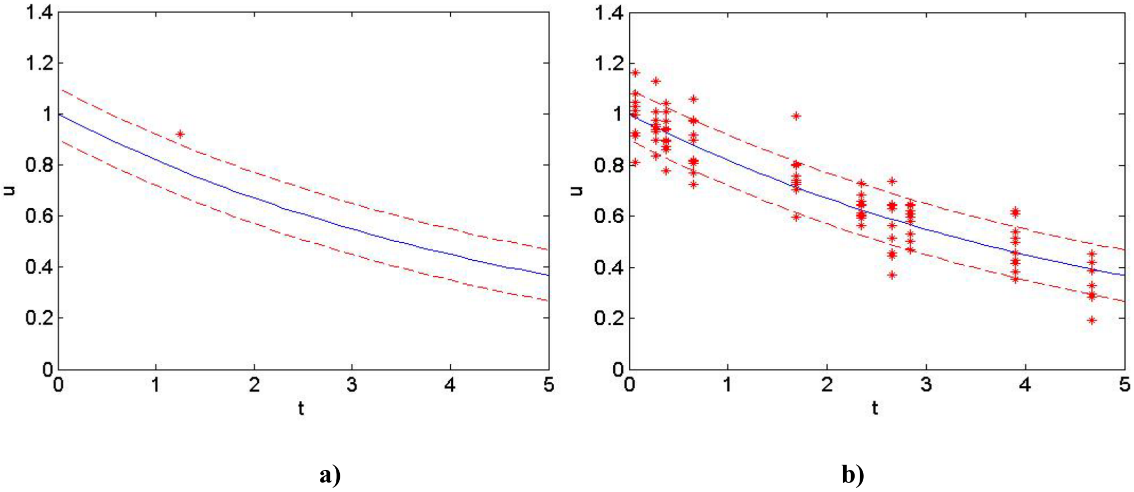


Figure 7-3. Synthetic Data and the Benchmark for Uncorrelated Observations, a) SO, b) MO

Correlated Observations

The data is synthesized using the Karhunen-Loeve (KL) expansion which admits the following equation:

$$u(t,\theta) = \bar{u}(t) + \sum_{n=0}^{\infty} \xi_n(\theta)\sqrt{\lambda_n} f_n(t) \tag{7-14}$$

where $\{\xi_n\}$ is a set of random variables that are selected using Monte-Carlo sampling from a standard normal distribution, λ_n and $f_n(t)$ are the eigenvalues and eigenvectors of the covariance kernel, respectively. In order to compute the covariance matrix and its eigenvalues and eigenvectors, the *t* domain was finely discretized and 1000 KL realizations were generated along the domain of *t*. Afterwards, 10 KL realizations were sampled at each of the specific locations (*t*) defined earlier as in the MO case. The synthesized data generated in this way are presented in Figure 7-4. The eigenvalues and eigenvectors corresponding to the covariance function are shown in Figure 7-5 and Figure 7-6, respectively. The surface representation of the correlation function is also presented in Figure 7-7.

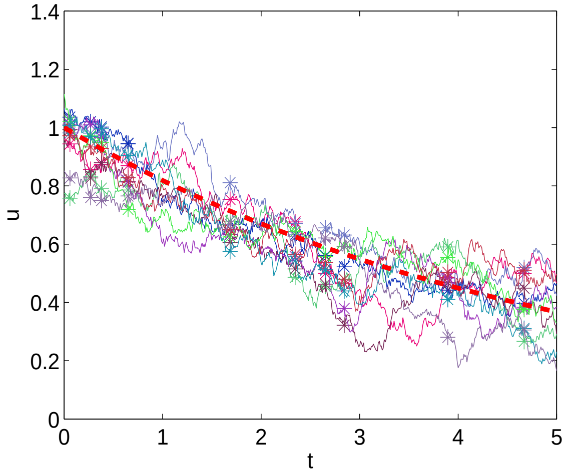


Figure 7-4. KL Synthesized Data for $\tau = 1$

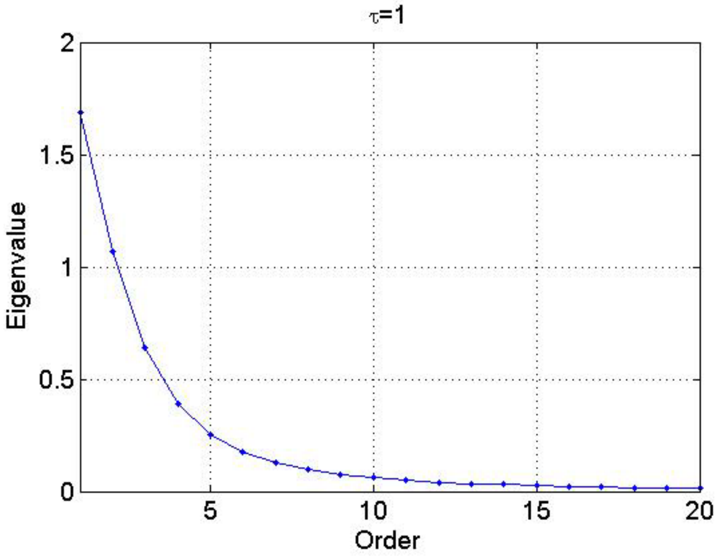


Figure 7-5. Eigenvalues of the Covariance Matrix

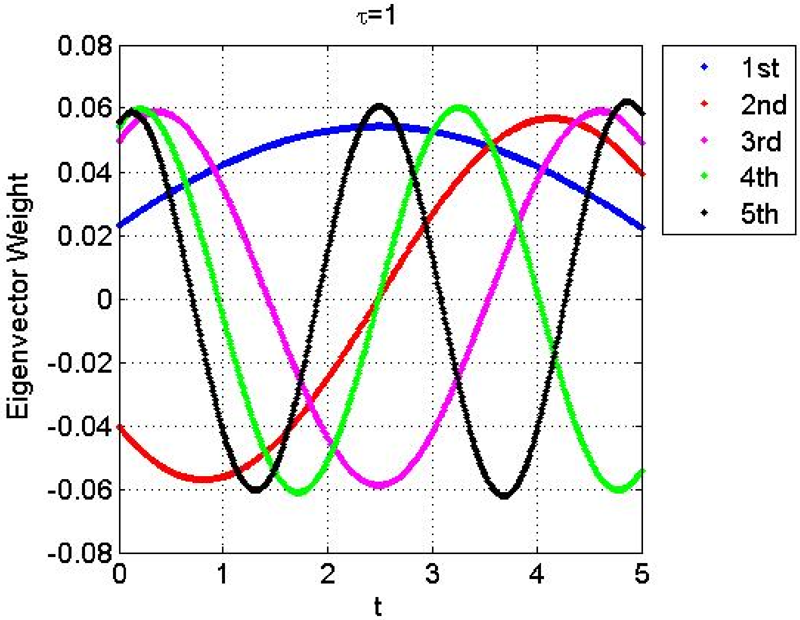


Figure 7-6. Eigenvectors of the Covariance Matrix

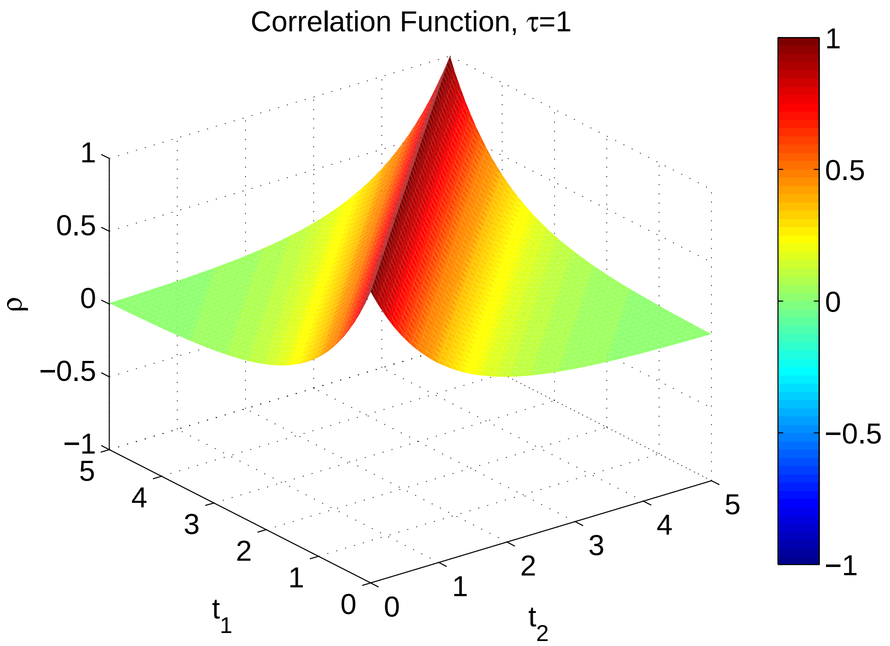


Figure 7-7. Correlation Function Representation

For the purpose of the Bayesian inversion using the MCMC method, the covariance of the likelihood distribution was initially defined as the empirical covariance matrix computed from the synthetic data. Figure 7-8 and Figure 7-9 represent the empirical covariance and correlation of the synthetic data. The empirical covariance matrix was not positive definite, therefore computing the likelihood of the data was not possible. In order to address this problem, the covariance matrix was computed based on the actual correlation function defined earlier by Eq. 7-12.

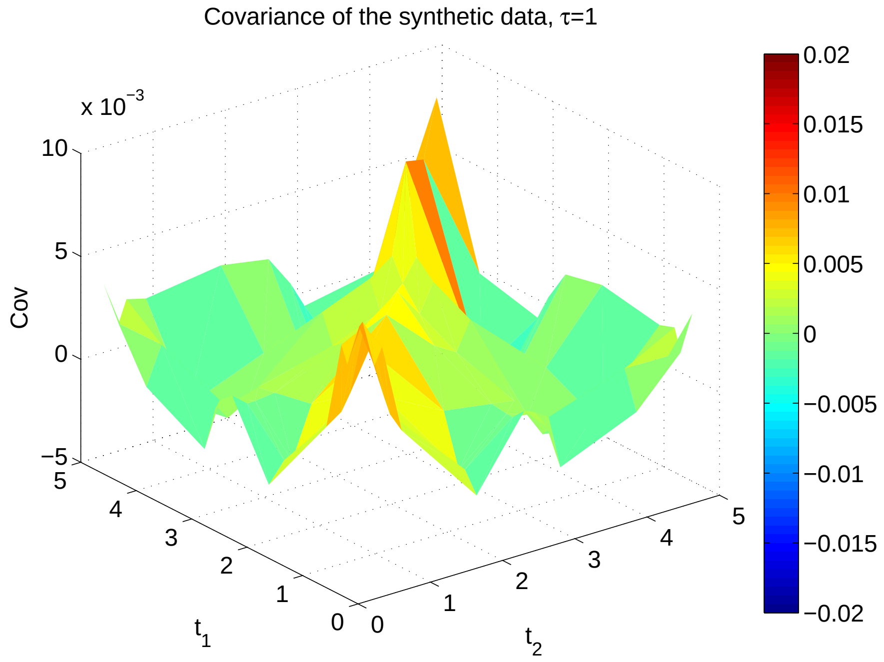


Figure 7-8. Empirical Covariance Function Obtained from the 10 KL Realizations

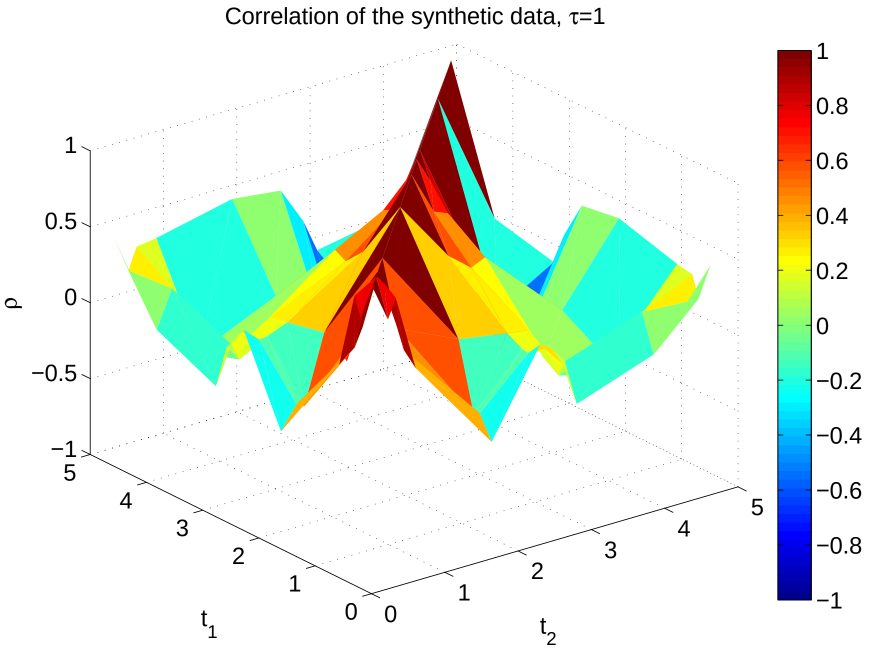


Figure 7-9. Empirical Correlation Function Obtained from the 10 KL Realizations

Methodology

The parametric uncertainty was evaluated through the Bayesian inversion method. Considering $\boldsymbol{\theta} = \{I, c\}$, the vector of model parameters and $\mathbf{d} = \{u_1, \cdots, u_m\}$, the vector of new observations of the process then $\mathrm{p}(\boldsymbol{\theta})$ represents the prior distribution and $\mathrm{p}(\mathbf{d} \mid \boldsymbol{\theta})$ represents the likelihood. According to the Bayes' theorem, the posterior probability density of model parameters, given the observed data, can be written as:

$$\mathrm{p}(\boldsymbol{\theta} \mid \mathbf{d}) = \frac{\mathrm{p}(\boldsymbol{\theta})\mathrm{p}(\mathbf{d} \mid \boldsymbol{\theta})}{\int \mathrm{p}(\boldsymbol{\theta})\mathrm{p}(\mathbf{d} \mid \boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}} \tag{7-15}$$

The Markov Chain Monte Carlo (MCMC) method is used to sample from the posterior distribution and approximate the corresponding statistics.

The standard deviation of the error (σ) and the step size (Δt) are two parameters that can be assumed to be known *a priori*; or can be considered as hyper-parameters and evaluated through the Bayesian inference. Assuming the prior distributions to be known and all the model parameters to be independent *a priori*, the following equation holds:

$$\mathrm{p}(I, c, \Delta t, \sigma \mid \mathbf{d}) \propto \mathrm{p}(I)\,\mathrm{p}(c)\,\mathrm{p}(\Delta t)\,\mathrm{p}(\sigma)\,\mathrm{p}(\mathbf{d} \mid I, c, \Delta t, \sigma) \tag{7-16}$$

The effects of the numerical error (associated with domain discretization) on the overall uncertainty in the solution of inverse problem were studied by performing three groups of numerical experiments:

1) Analytical solution of the ODE (*Ana*),

2) Numerical solution of the ODE with constant $\Delta t = 0.001$ (*Const*), and

3) Numerical solution of the ODE, assuming that Δt is a hyper-parameter (*Var*).

These numerical experiments were performed considering the three observations scenarios: SO, MO, and CO.

The numerical uncertainty in the solution of an inverse problem initiates from the numerical solution of the forward problem, therefore it depends on the size of domain discretization. In this study, the impact of the time step size on the uncertainty was measured by two quantities: the Numerical Bias (*NB*) and the Numerical Uncertainty (*NU*) given by the following equations:

$$\begin{aligned} NB &= \left\| E_{num}\left(\{I,c,u\}\right) - \{1, 0.2, e^{-0.2t}\} \right| - \left| E_{ana}\left(\{I,c,u\}\right) - \{1, 0.2, e^{-0.2t}\} \right\| \\ NU &= \left| Std_{num}\left(\{I,c,u\}\right) - Std_{ana}\left(\{I,c,u\}\right) \right| \end{aligned} \tag{7-17}$$

The inverse problem was solved for different values of step size $\Delta t = 0.1, 0.01, 0.001, 0.0001$, and *NB* and *NC* were measured for each numerical experiment. These experiments were performed twice to ensure repeatability. Figure 7-10 presents the design of numerical experiments performed in this study.

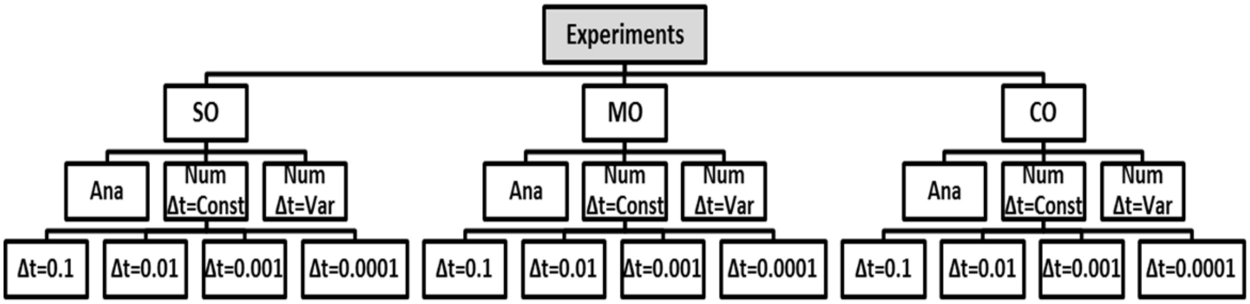


Figure 7-10. Design of the Numerical Experiments

Results and Discussion

Bayesian Inversion, Unknown Parameters: $I, c, \Delta t$

This section presents the results of the MCMC simulations for the three experiments explained in the previous section. As the variance of the synthetic data is known to be 0.01, σ was considered to be known *a priori*. Also non-informative priors were considered for parameters I and c. However for the cases where Δt was considered as a hyper-parameter, a lognormal prior (see Figure 7-11) was introduced with mean and variance equal to 0.0001:

$$\begin{aligned} &\Delta t \sim lognorm(\mu, \sigma) \\ &\mu = -13.82 \\ &\sigma = 3.03 \end{aligned} \tag{7-18}$$

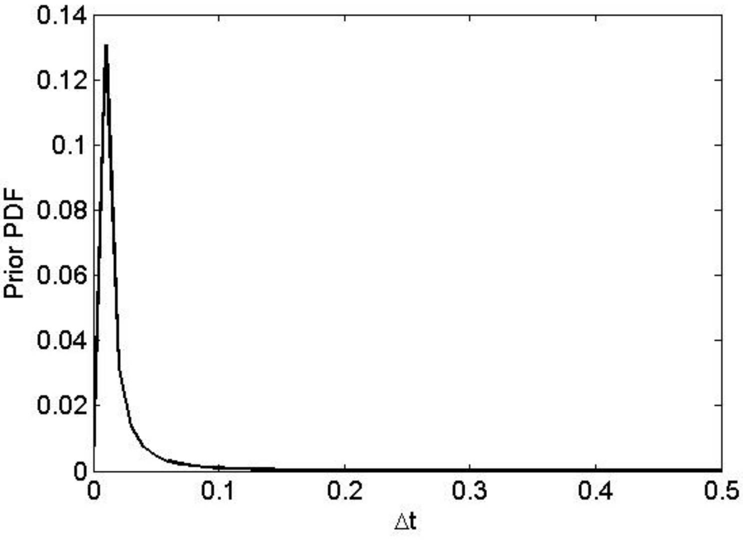


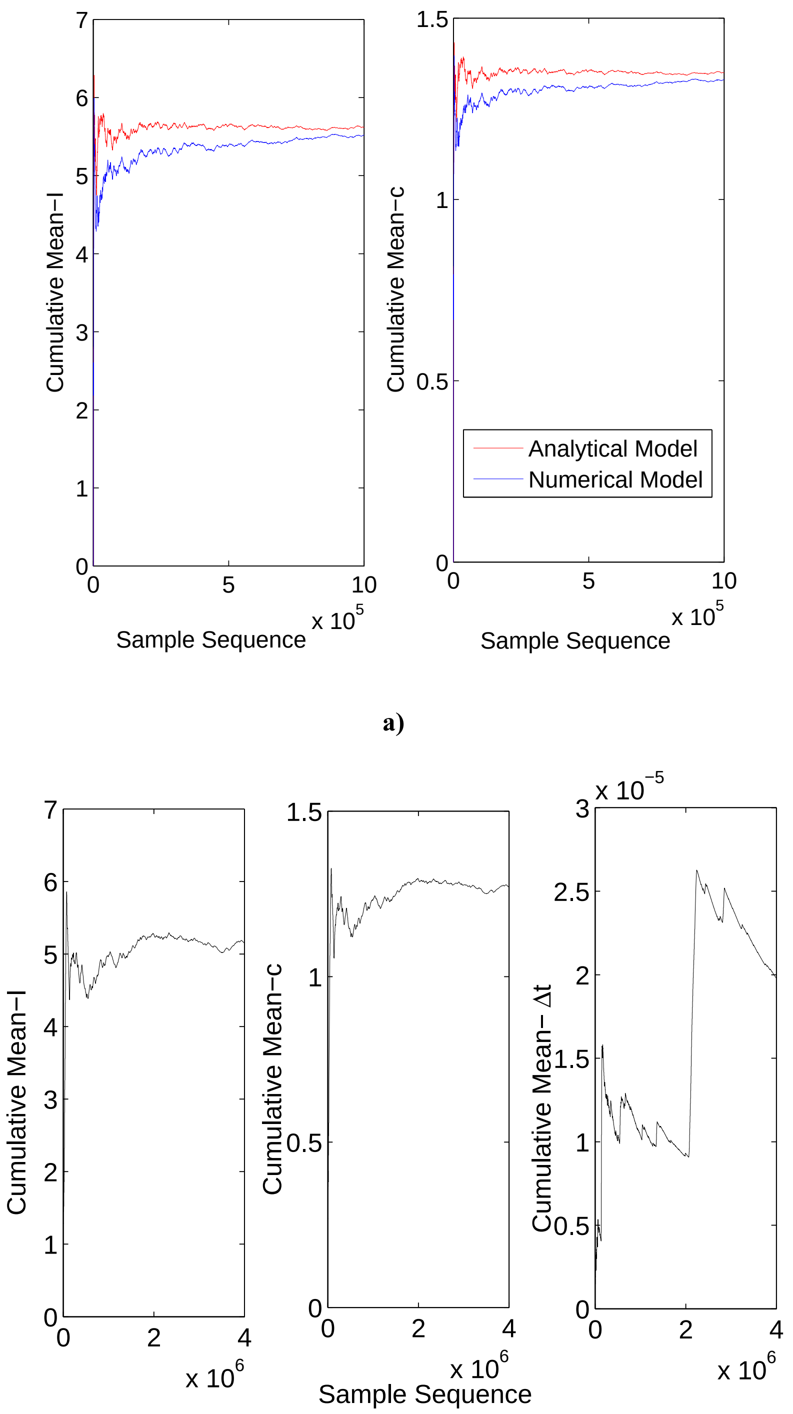
Figure 7-11. Prior Probability Distribution of *Δt*

Single Observation

A single observation at $t=1.25$ (one fourth of the t domain) was synthesized following the process explained previously, and the inverse problem was solved using the Bayesian inference method.

Figure 7-12 shows the cumulative mean of the simulated Markov-Chains and indicates that the chains converged to the stationary distributions. However, the mean of the chains did not converge to the true values of I and c (i.e. 1 and 0.2). The reason could be the insufficient amount of observations and the fact that the problem was under-determined (there was only one data, but more than one parameter to infer). For the case that Δt was introduced as a hyper parameter, the chain for Δt did not converge, even after 3 M simulations.

The posterior distributions of the model parameters were constructed from the generated MCMC samples after the burn-in point. As described earlier in Chapter III, the burn-in point is referred to the sample in a Markov chain where the sample chain converges to a stationary distribution. Figure 7-13 shows the posterior joint distributions of the model parameters. It was observed that the inversion results were very similar for the different experiments, indicating that the numerical error did not have significant impact on the solution of the inverse problem in this case. Also, Figure 7-14 shows the empirical CDFs of the marginal posterior distributions. Comparing the experiment incorporating constant Δt with the case where Δt was considered as a hyper-parameter, it was observed that the constant Δt resulted in more accurate parameterization and less uncertain posterior model predictions.

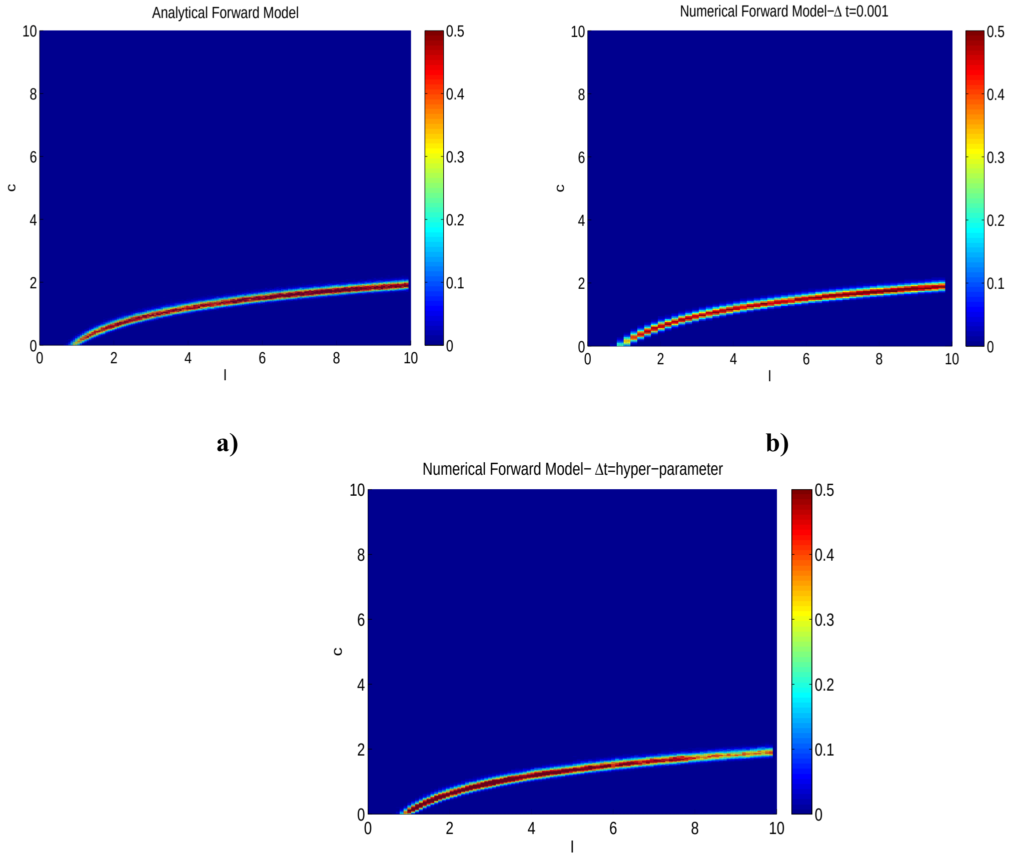


a)

b)

Figure 7-12. Cumulative Mean of the Sample Chains, a) Analytical and Numerical Model with *Δt=0.001*, and b) Numerical Model with *Δt=hyper-parameter* – SO

The uncertainty was propagated from the model parameters to the model predictions by taking a large number of samples from the posterior probability distributions of the model parameters using the MC method, and then running the forward model for each sample. As a result of this process 100 K posterior model predictions or realizations of the process were generated. The first and second moments (expected value and variance) were then computed to evaluate the uncertainty in the model predictions.


Analytical Forward Model

Numerical Forward Model−Δ t=0.001

Numerical Forward Model− Δt=hyper−parameter

a)

b)

c)

Figure 7-13. Posterior Joint PDF of *I* and *c* for, a) Analytical Model, b) Numerical Model with *Δt=0.001*, and c) Numerical Model with *Δt=hyper-parameter* – SO

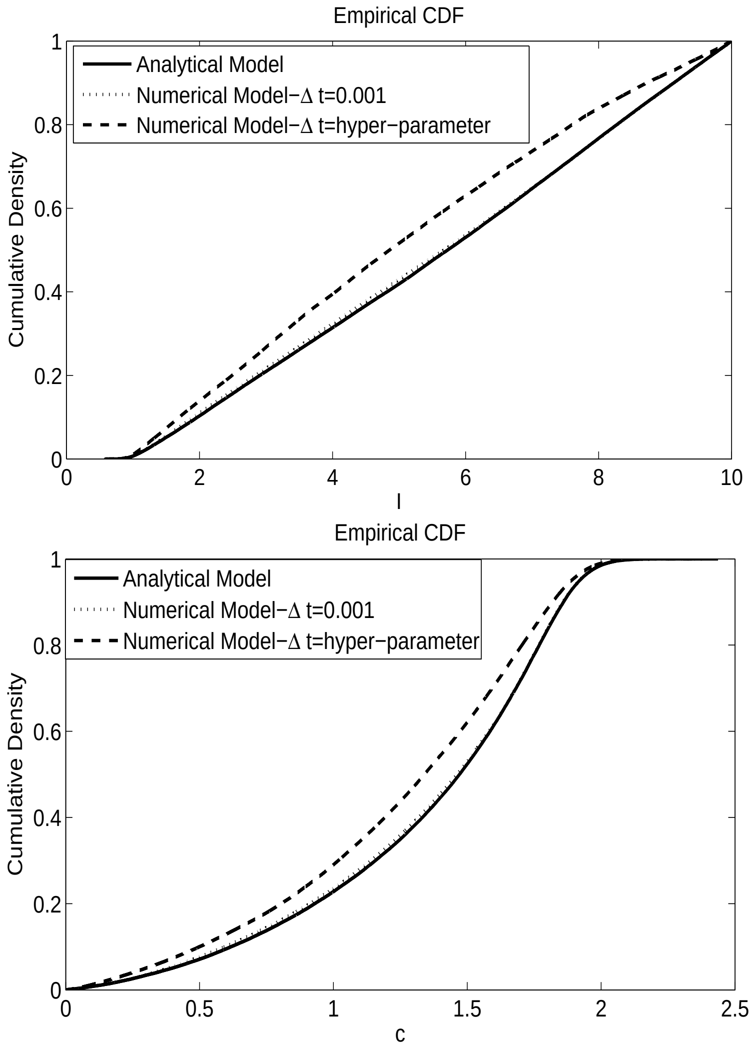
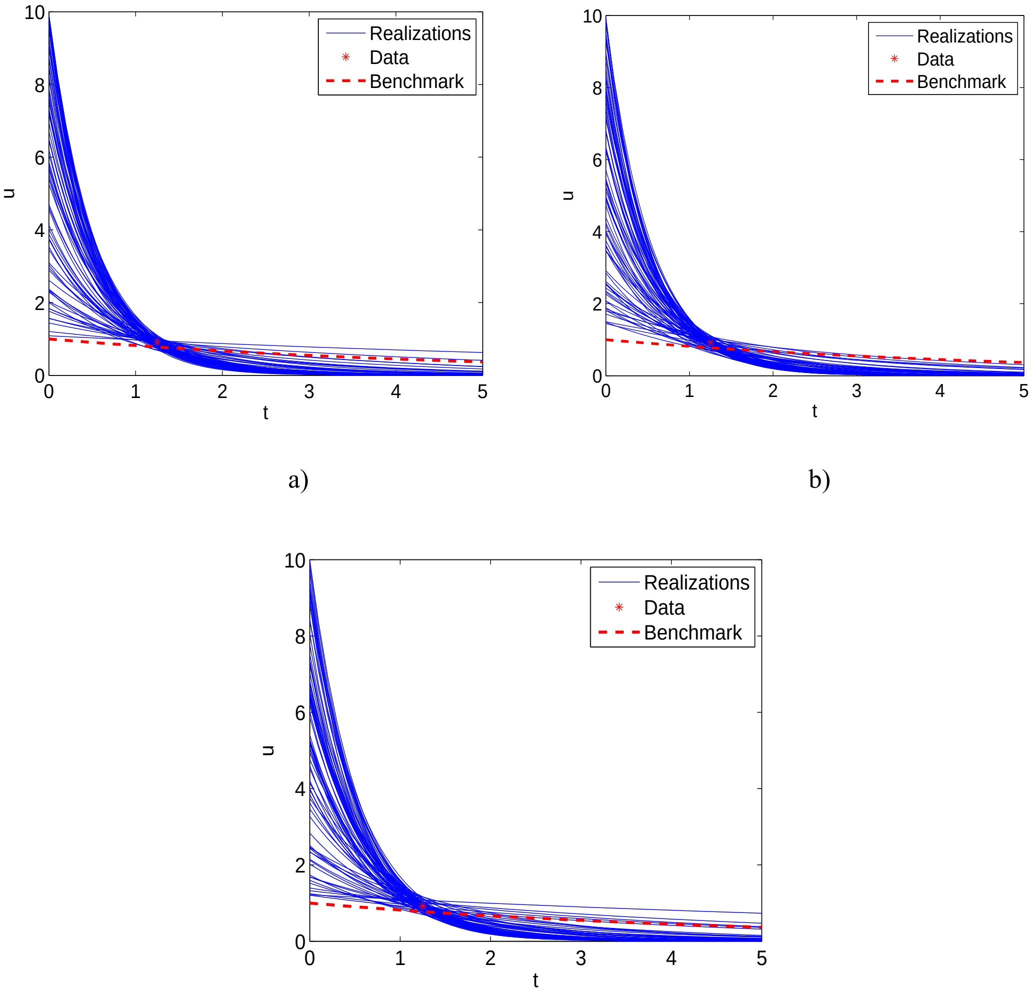


Figure 7-14. Posterior CDFs of *I* and *c* – SO

Figure 7-15 presents a sample of realizations of the process. Figure 7-16 presents the variation of the mean and the standard deviation of the process with respect to ***t*.** Also, Figure 7-17 presents the joint PDF of *Δt* with *I* and *c*. It was observed that the model parameters were independent from the time step, as no correlation was observed between them. Also Figure 7-18 presents the posterior CDF of the time step. The mean is around 2×10^{-5}, the standard deviation is 5.3×10^{-5}, and the 95% Confidence Interval is $[0\ 0.2\times10^{-5}]$.



a)

b)

c)

Figure 7-15. Posterior Model Predictions, a) Analytical, b) Numerical Model with *Δt=0.001*, and c) Numerical Model with *Δt=hyper-parameter* - SO

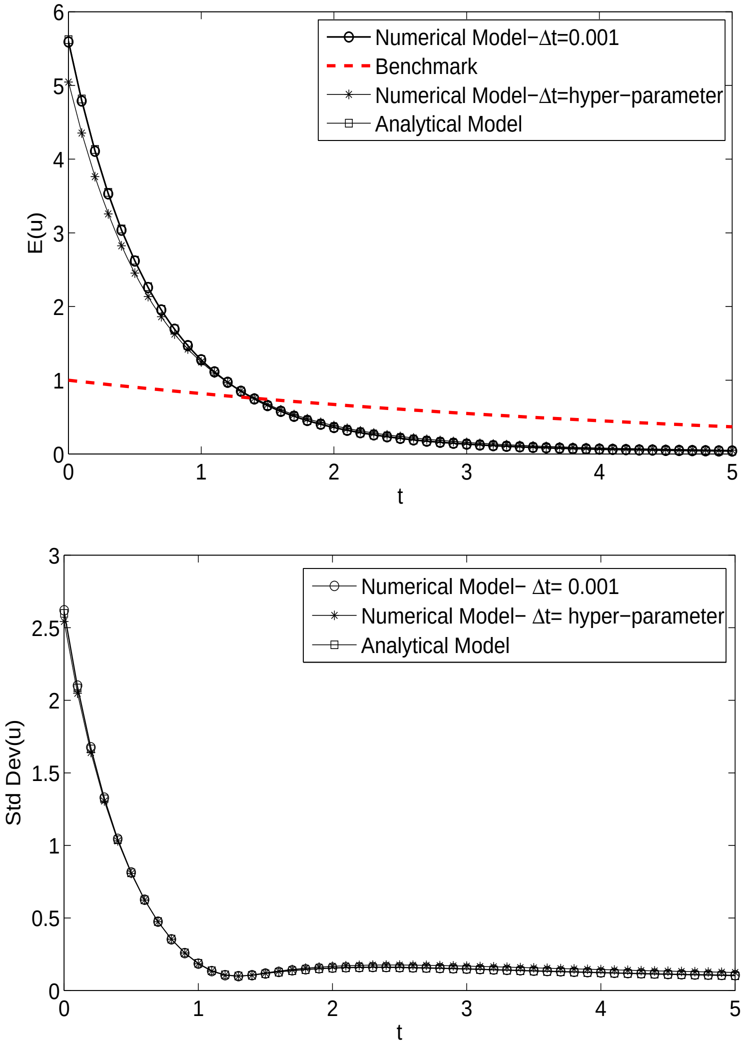


Figure 7-16. Mean and Standard Deviation of the Posterior Model Predictions (Realizations) vs. Time – SO

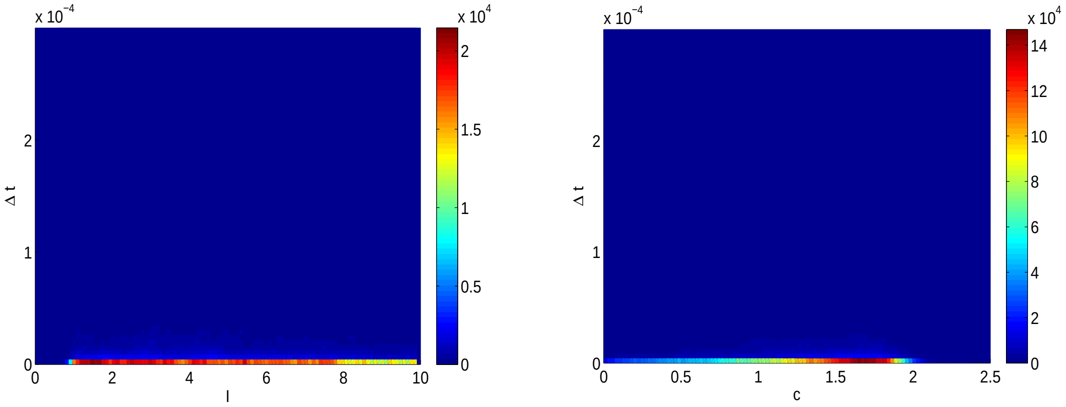


Figure 7-17. Posterior Joint PDF of *Δt* with *I* and *c* - SO

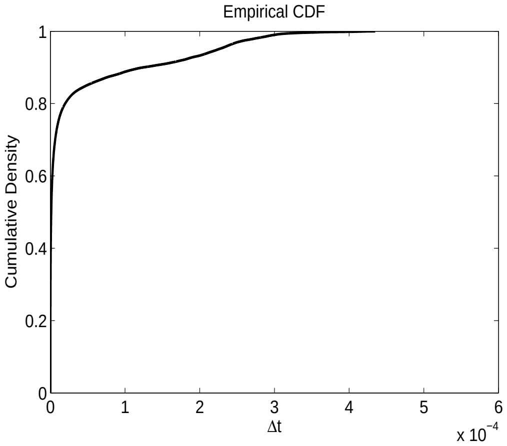


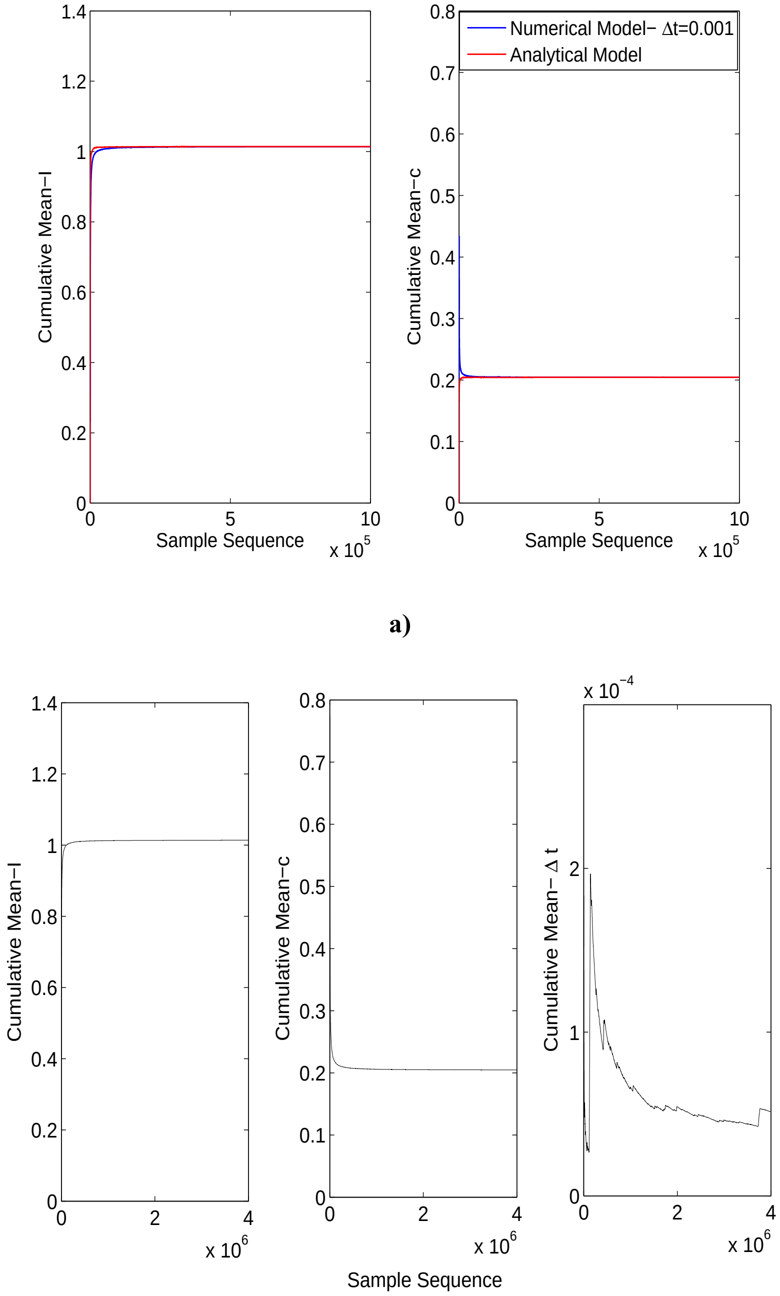
Figure 7-18. Posterior CDF of *Δt* - SO

Many Observations

Considering the MO experiment the inverse problem was solved through Bayesian inversion, this time having a substantial amount of observations available. This obviously makes the system over-determined as opposed to SO experiment.

Figure 7-19 shows the cumulative mean of the chains and indicates that the chains converged to the stationary distributions. It was observed that the mean of the chains converged to the true values of *I* and *c* (i.e. 1 and 0.2), as opposed to the SO case. For the experiment that *Δt* was introduced as a hyper-parameter, the chain became stationary at around 4 M iterations; however, the other two model parameters converged to the true values rather quickly.

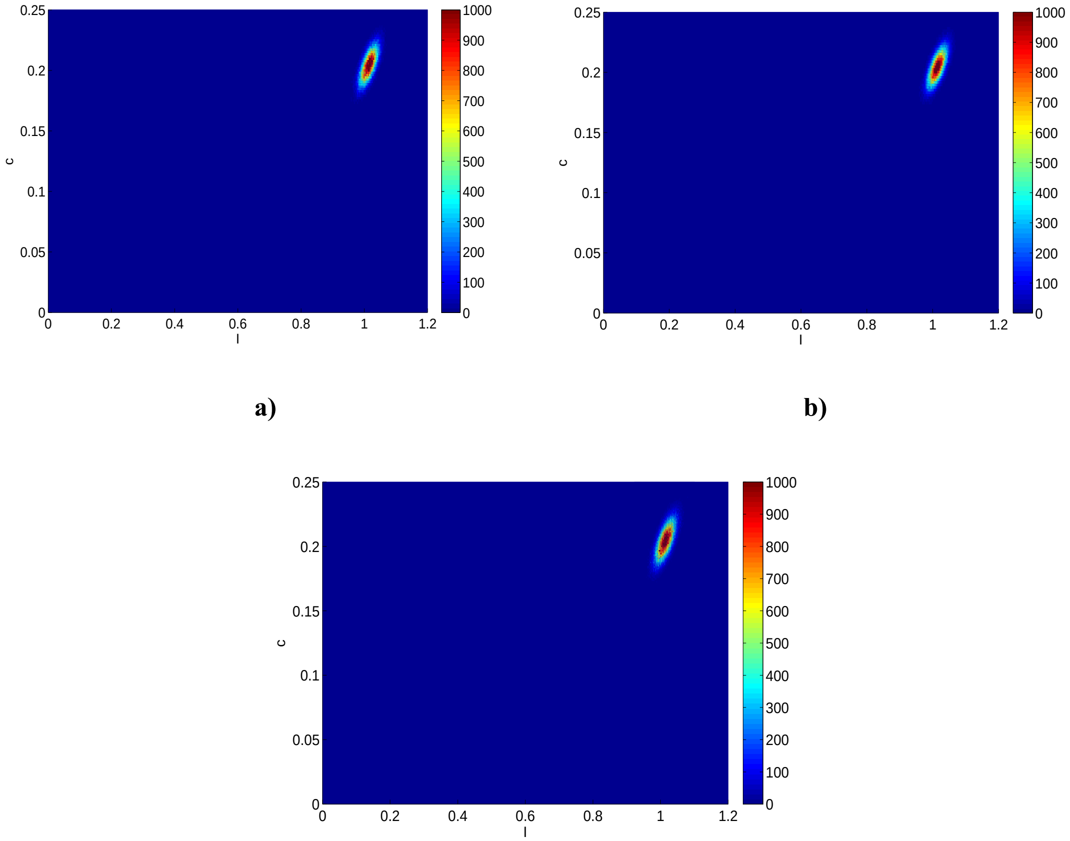
Figure 7-20 presents the posterior joint distributions of the model parameters. The inversion results were very similar for the different experiments. This indicated that the numerical error did not have a significant impact on the solution of the inverse problem when there was a substantial amount of data. Also, Figure 7-21 shows the empirical posterior CDFs of the model parameters. The posterior CDFs of the model parameters for the three experiments showed a slight difference as oppose to the SO case. Figure 7-22 presents a sample of realizations of the process. Figure 7-23 presents the variation of the mean and the standard deviation of the process with respect to *t*. Figure 7-24 presents the posterior joint PDF of *Δt* versus *I* and *c*. It was observed that, similar to SO case, the model parameters were independent from the hyper-parameter. Figure 7-25 presents the posterior CDF of *Δt*. The mean value of *Δt* is around 5×10^{-5}, the standard deviation is 1.7×10^{-4}, and the 95% Confidence Interval is [0 0.5×10^{-3}].



a)

b)

Figure 7-19. Cumulative Mean of the Sample Chains, a) Analytical and Numerical Model with *Δt=0.001*, and b) Numerical Model with *Δt=hyper-parameter* – MO



a)

b)

c)

Figure 7-20. Posterior Joint PDF of *I* and *c*, a) Analytical, b) Numerical Model with *Δt=0.001*, and c) Numerical Model with *Δt=hyper-parameter* – MO

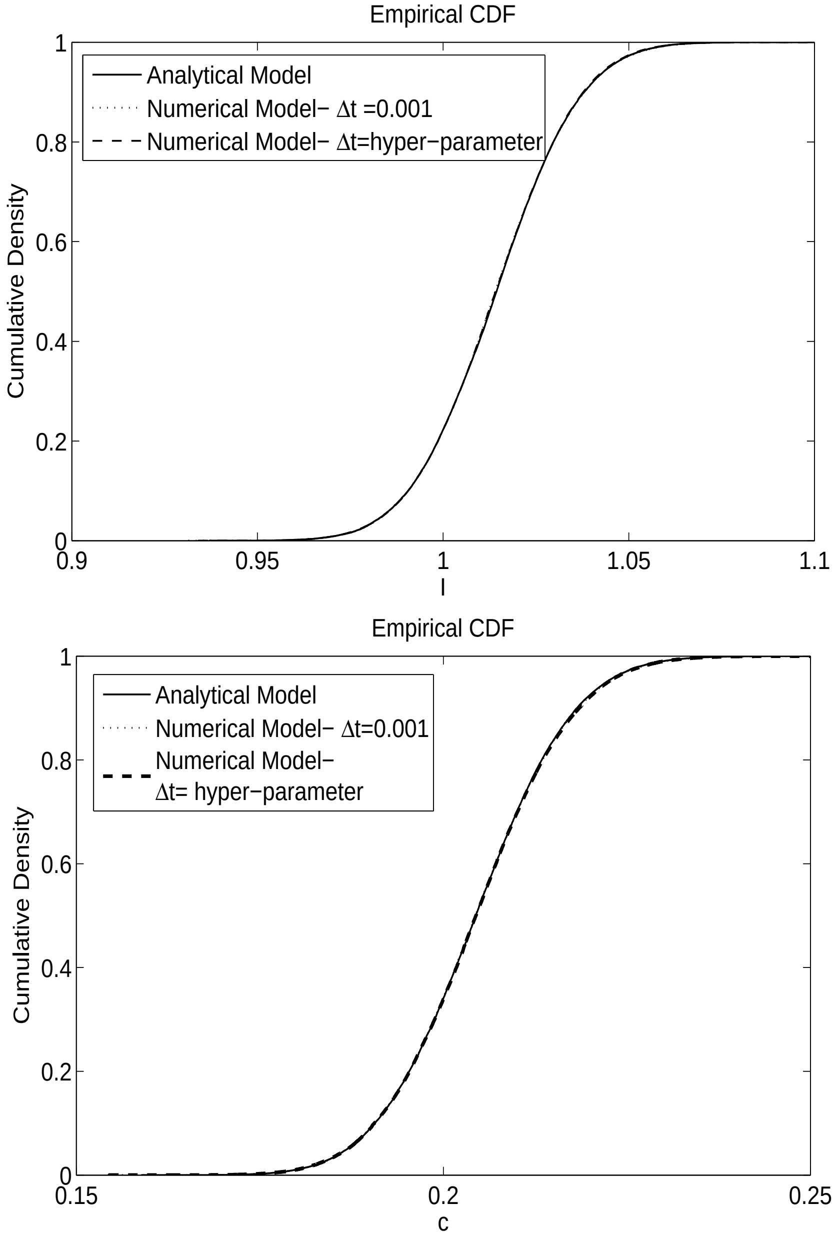
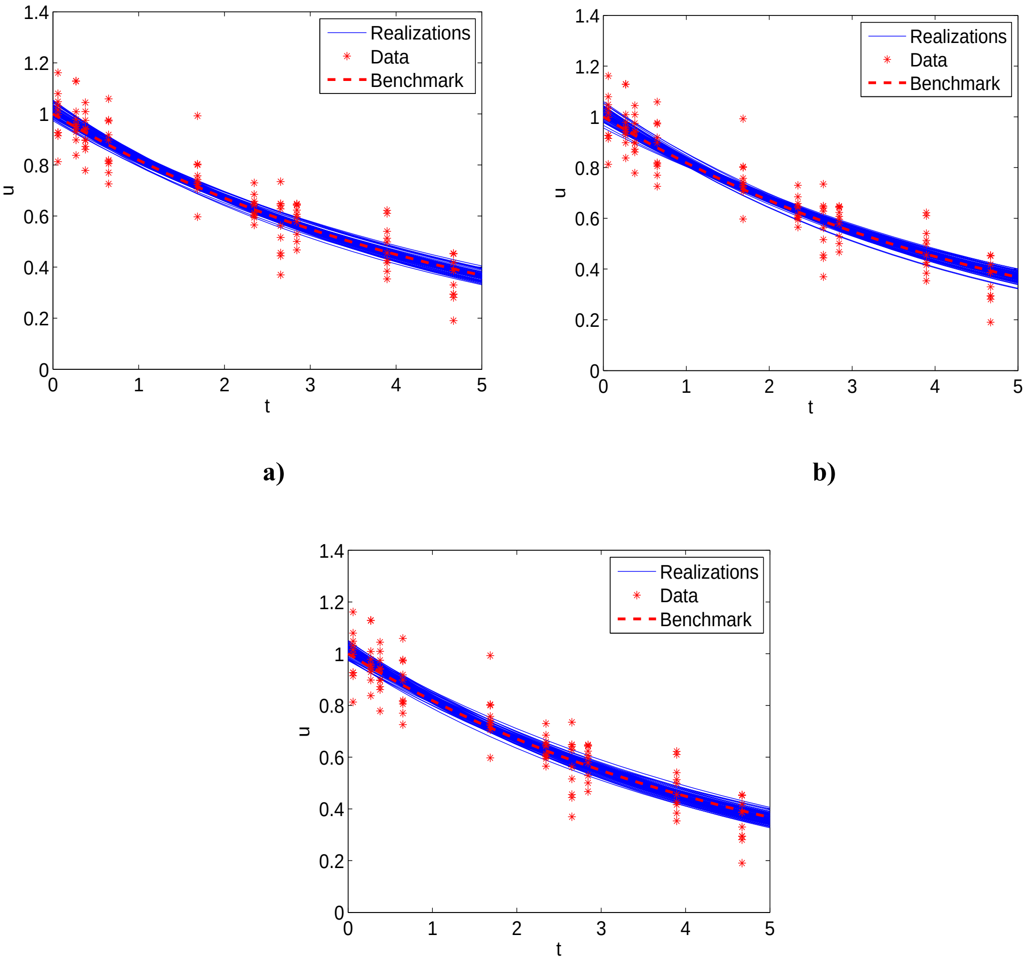


Figure 7-21. Posterior CDFs of *I* and *c* – MO



a)

b)

c)

Figure 7-22. Posterior Model Predictions, a) Analytical, b) Numerical Model with *Δt=0.001*, and c) Numerical Model with *Δt=hyper-parameter* - MO

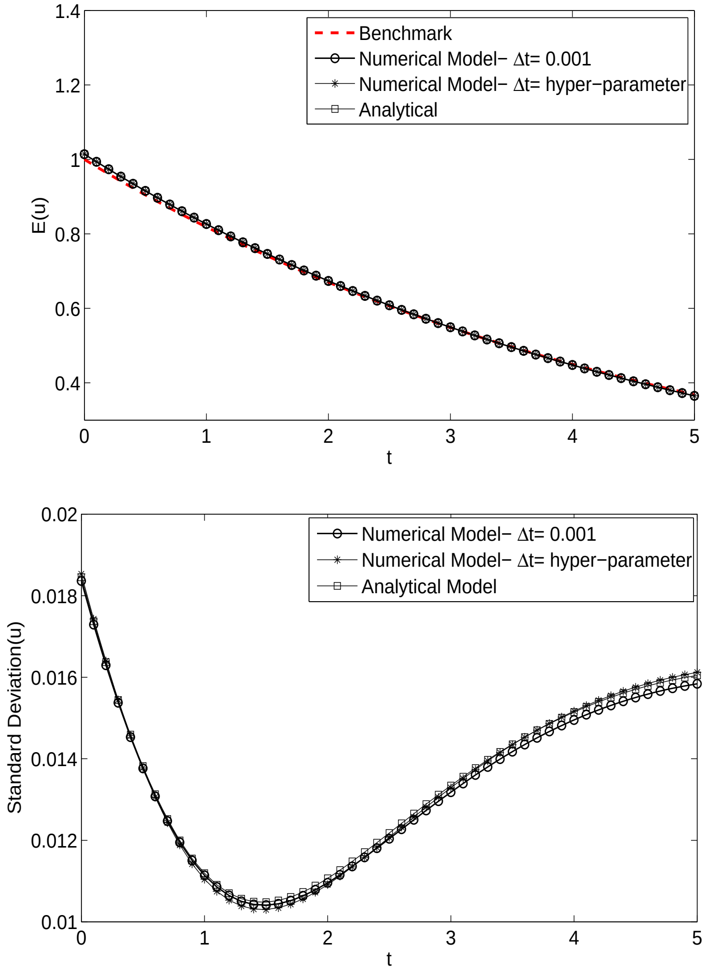


Figure 7-23. Mean and Standard Deviation of the Posterior Model Predictions vs. Time - MO

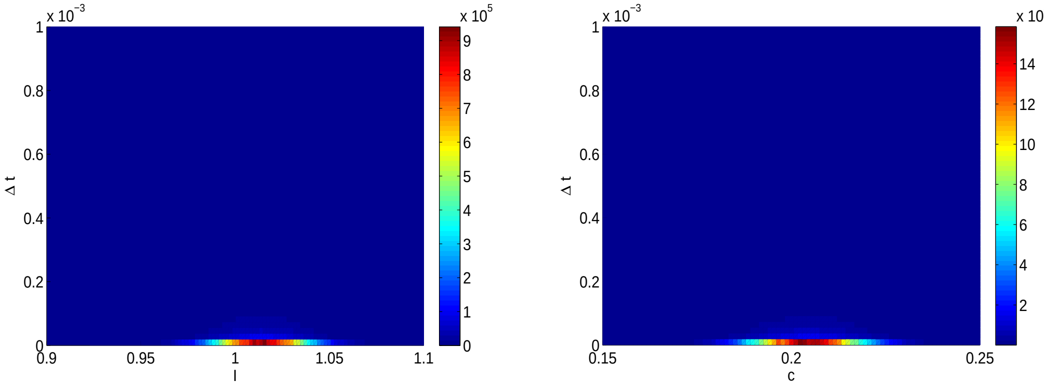


Figure 7-24. Posterior Joint PDF of Δt with I and c – MO

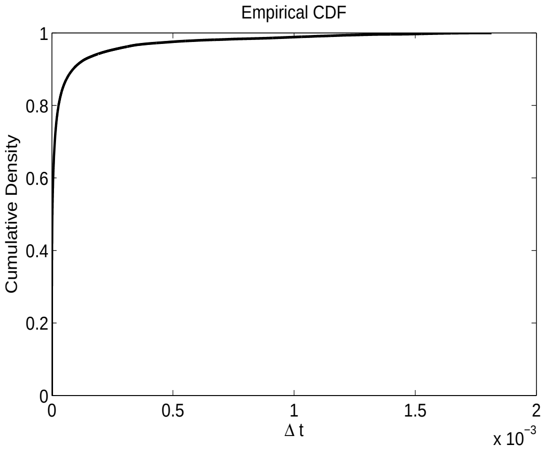


Figure 7-25. Posterior CDF of Δt - MO

Correlated Observations

The inverse problem was solved considering CO scenario of the observations. Figure 7-26 shows the cumulative mean of the chains for I and c. It was observed that the chain for c converged to its true value (0.2); however, the chain for I did not converged to a value slightly less than the actual value (1). Also, the stationary distributions were very similar for both the analytical and the numerical models.

However, similar to MO case, the chain associated with the *Δt* showed slow convergence.

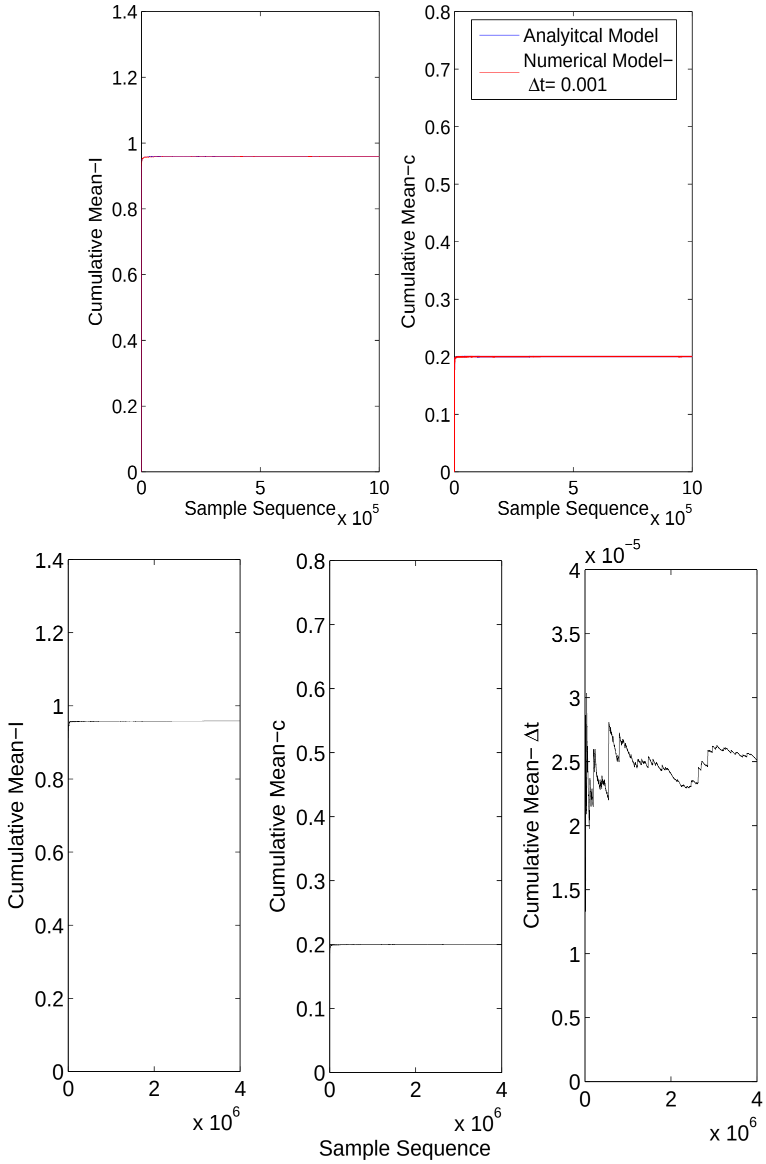
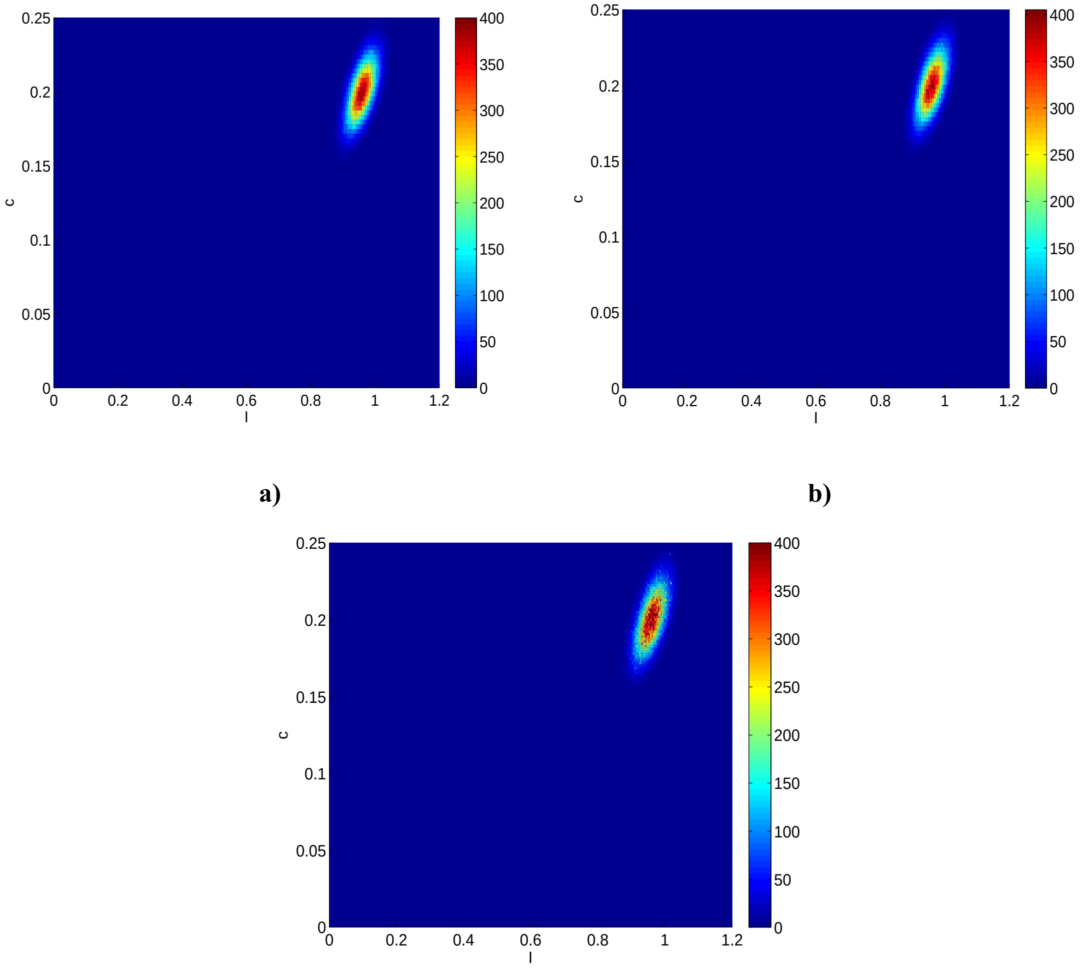


Figure 7-26. Cumulative Mean of the Model Parameters, a) Analytical and Numerical Model with *Δt*=0.001, and b) Numerical Model with *Δt*=hyper-parameter – CO

Figure 7-27 shows the posterior joint distributions of *I* and *c* and Figure 7-28 shows the posterior empirical CDFs of the model parameters. It was observed that the posterior CDFs of the model parameters were not significantly affected by the numerical uncertainty.



a)

b)

c)

Figure 7-27. Posterior Joint PDF of *I* and *c*, a) Analytical, b) Numerical Model with *Δt*=0.001, and c) Numerical Model with *Δt*=hyper-parameter – CO

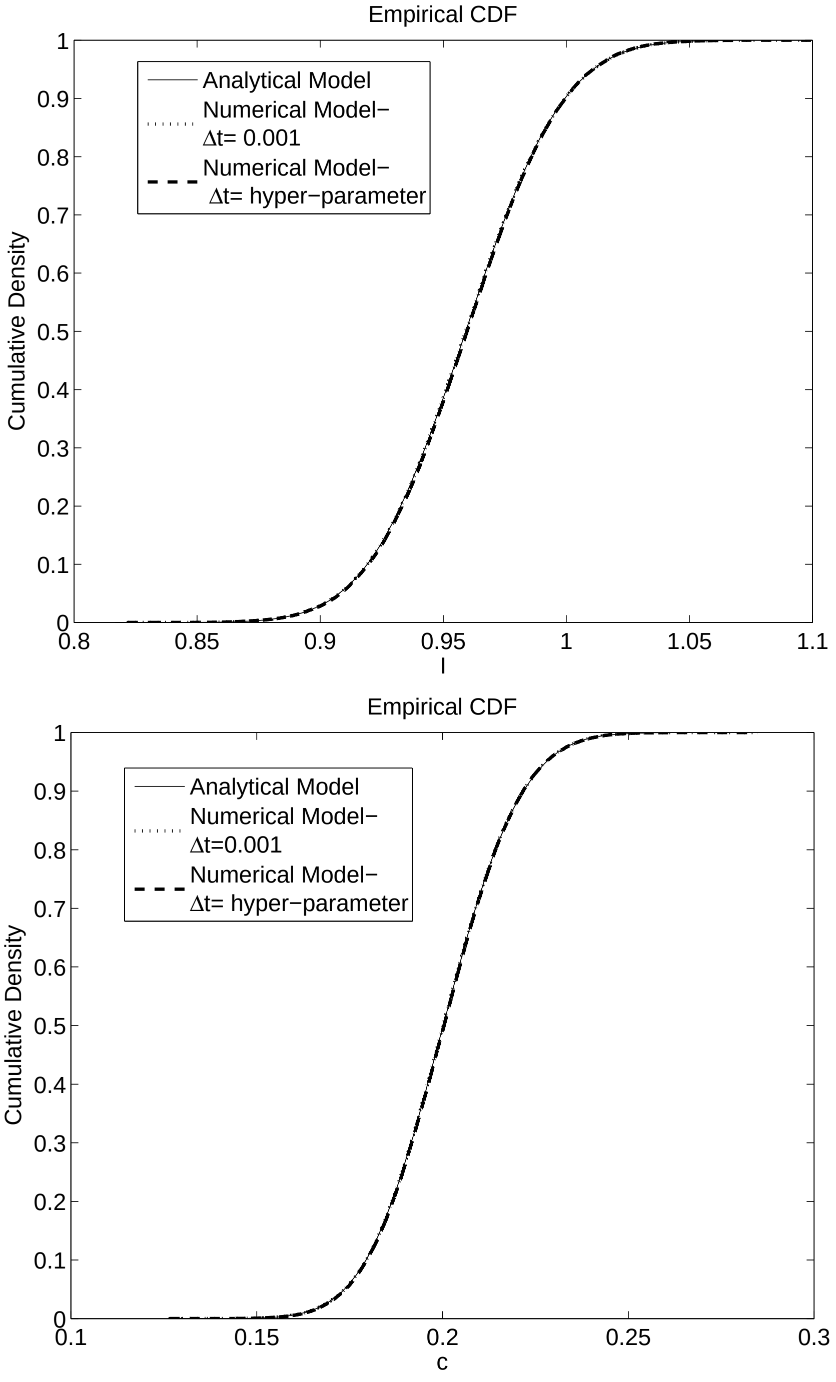
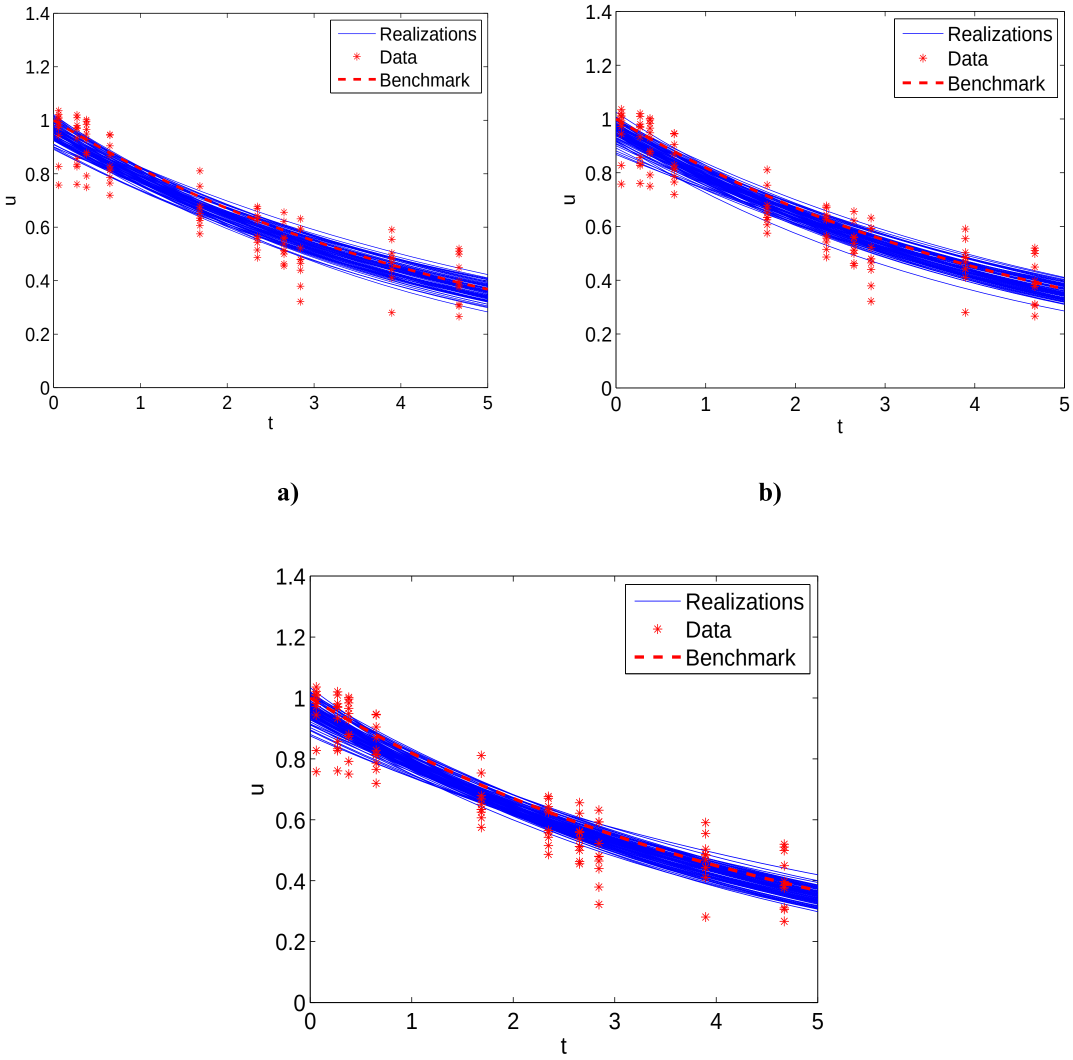


Figure 7-28. Posterior CDFs of *I* and *c* - CO

100K realizations of the process were generated (Figure 7-29) by sampling from the posterior distributions of *I* and *c*. Figure 7-30 presents the variation of the mean and standard deviation of the posterior predictions versus time. It was observed that the correlation in the data caused a bias in the estimation of the initial value (*I*) of the process (note the discrepancy between the trend of the process and the benchmark). It was shown that the numerical error and uncertainty due to the time discretization were insignificant even under correlated data structure, as long as the time step is substantially small.

As it is observed in Figure 7-31 the joint distributions of Δt versus *I* and *c* are very similar to the case with uncorrelated observations, and the model parameters show independency with the hyper-parameter. Figure 7-33 presents the posterior empirical CDF of the time step that shows slightly smaller mean, comparing to the uncorrelated case. The mean value of the posterior distribution for Δt is 2.5×10^{-5}, the standard deviation is 7.2×10^{-5}, and the 95% CI is $[0\ 0.2\times10^{-3}]$.

Figure 7-32 presents the joint posterior distribution for the model parameters comparing the uncertainty over the three scenarios of observations. It was observed that the uncertainty decreased significantly as more observations become available, also correlated observations resulted in more uncertainty in the solution of the inverse problem.



a)

b)

c)

Figure 7-29. Posterior Model Predictions, a) Analytical, b) Numerical Model with *Δt*=0.001, and c) Numerical Model with *Δt=hyper-parameter* – CO

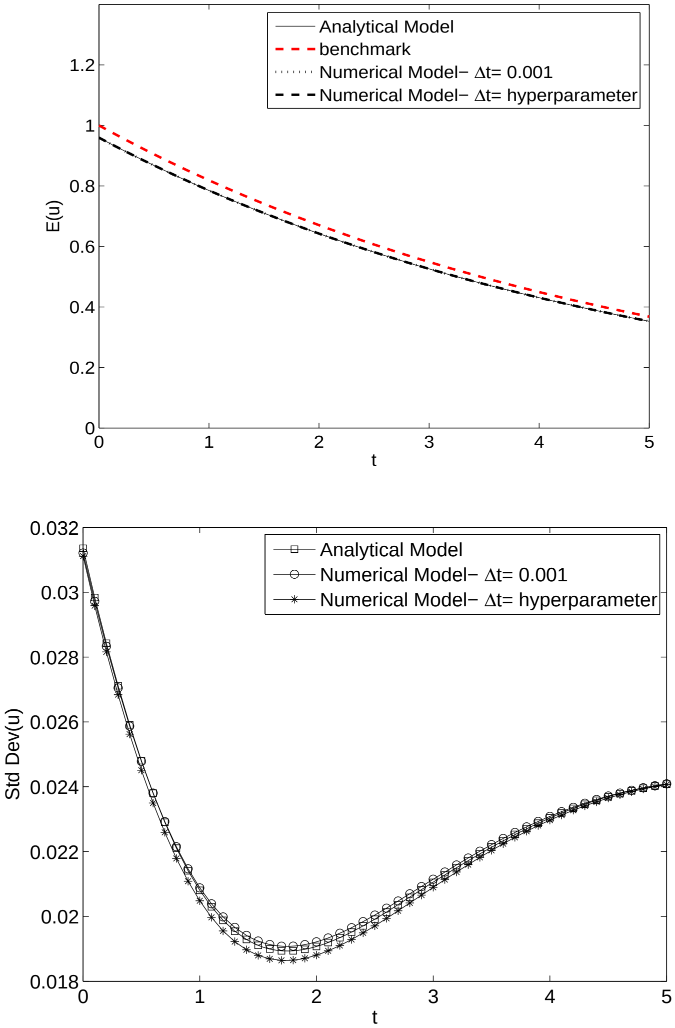


Figure 7-30. Mean and Standard Deviation of Posterior Model Predictions vs. Time - CO

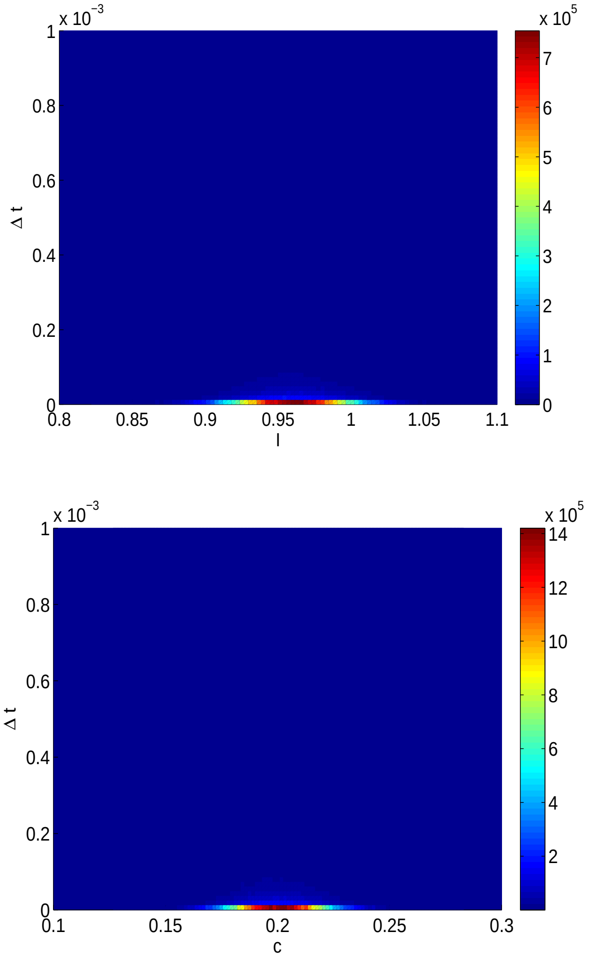
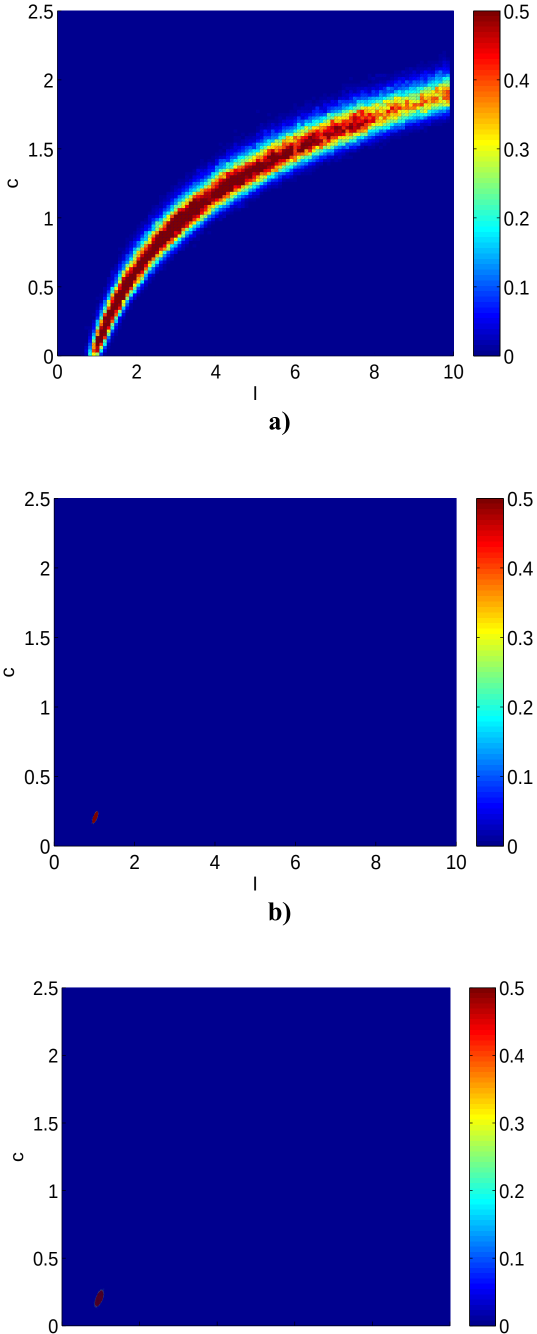


Figure 7-31. Posterior Joint PDF of *Δt* with *I* and *c* – CO



a)

b)

c)

Figure 7-32. Comparing the Uncertainty Resulted from the Three Scenarios Using the Same Scales, a) SO, b) MO, c) CO.

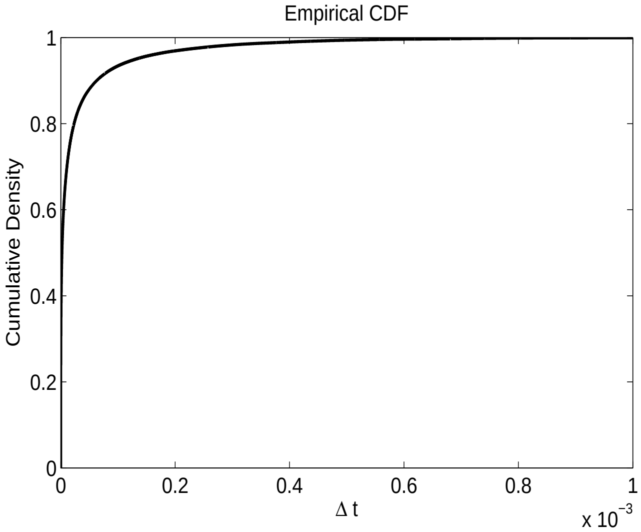


Figure 7-33. Posterior CDF of *Δt* - CO

Statistical Inference

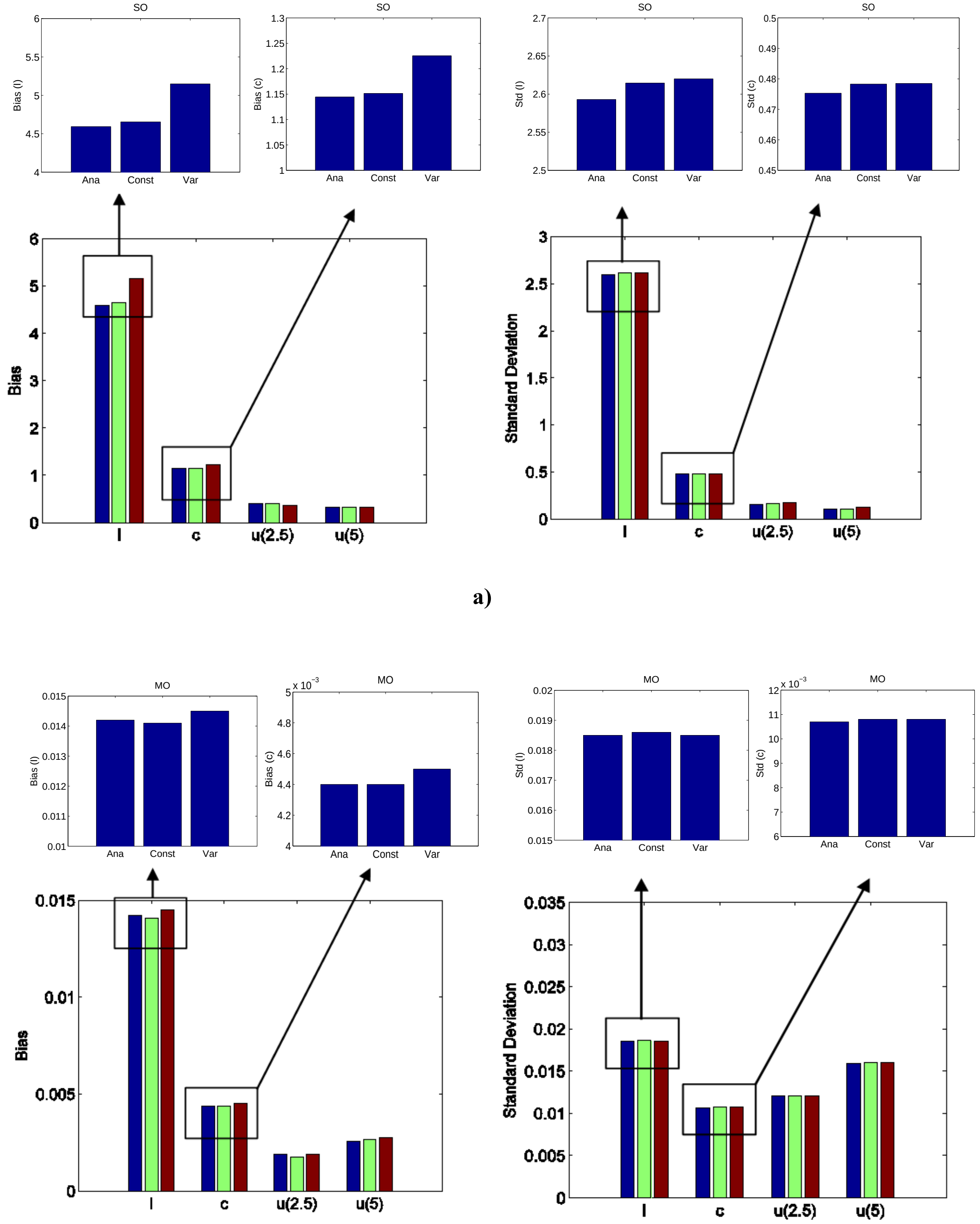
Table 7-2 presents a summary of the statistics for the posterior PDFs of the model parameters. Also, Table 7-3 presents the statistics associated with the model predictions at the one-half ($t=2.5$) and the end of the domain ($t=5$). This allows for a better analysis of the error and uncertainty changes due to varying evidence conditions. Also, Figure 7-34 illustrates the bias (error between the true value and the mean of the posterior PDFs) in the estimation of I, c, and u, and the total uncertainty, i.e. the standard deviation of their posterior PDFs for the three observations scenarios.

Table 7-2. Statistics of the Posterior Distributions of the Model Parameters

	E (I)	***Std (I)***	***E (c)***	***Std(c)***	***Corr(I,c)***
SO-Ana	5.5939	2.5930	1.3444	0.4753	0.9473
SO-Const	5.6551	2.6145	1.3515	0.4783	0.9473
SO-Var	6.1497	2.6201	1.4254	0.4785	0.9458
MO-Ana	1.0142	0.0185	0.2044	0.0107	0.6732
MO-Const	1.0141	0.0186	0.2044	0.0108	0.6762
MO-Var	1.0145	0.0185	0.2045	0.0108	0.6748
CO-Ana	0.9590	0.0314	0.2003	0.0166	0.5989
CO-Const	0.9591	0.0315	0.2003	0.0165	0.6006
CO-Var	0.9591	0.0311	0.2005	0.0165	0.5946

Table 7-3. Statistics of the Posterior Distributions of the Model Predictions

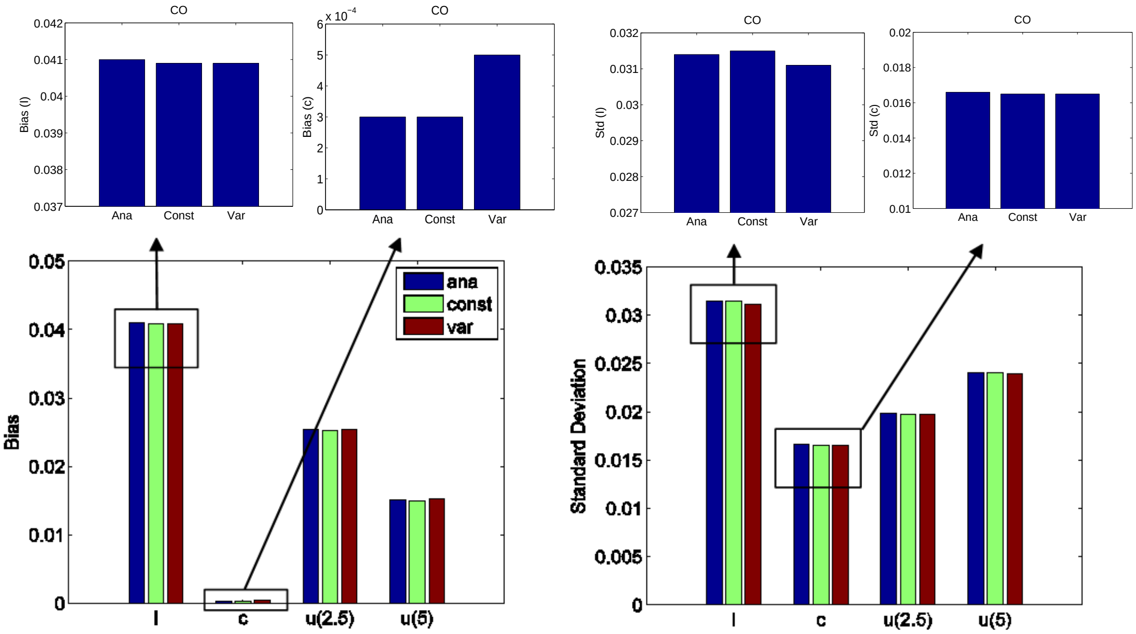
	E(u(2.5))	***Std(u(2.5))***	***E(u(5))***	***Std(u(5))***
SO-Ana	0.2082	0.1593	0.0376	0.1038
SO-Const	0.2069	0.1601	0.0378	0.1040
SO-Var	0.2360	0.1763	0.0509	0.1245
MO-Ana	0.6084	0.0121	0.3653	0.0159
MO-Const	0.6083	0.0121	0.3652	0.0160
MO-Var	0.6084	0.0121	0.3651	0.0160
CO-Ana	0.5812	0.0198	0.3528	0.0241
CO-Const	0.5813	0.0197	0.3529	0.0240
CO-Var	0.5812	0.0197	0.3526	0.0239



a)

b)

Figure 7-34. Bias and Uncertainty of Model Parameters and Predictions for Different Observations Conditions, a) SO, b) MO, c) CO



c)

Figure 7-34. Continued

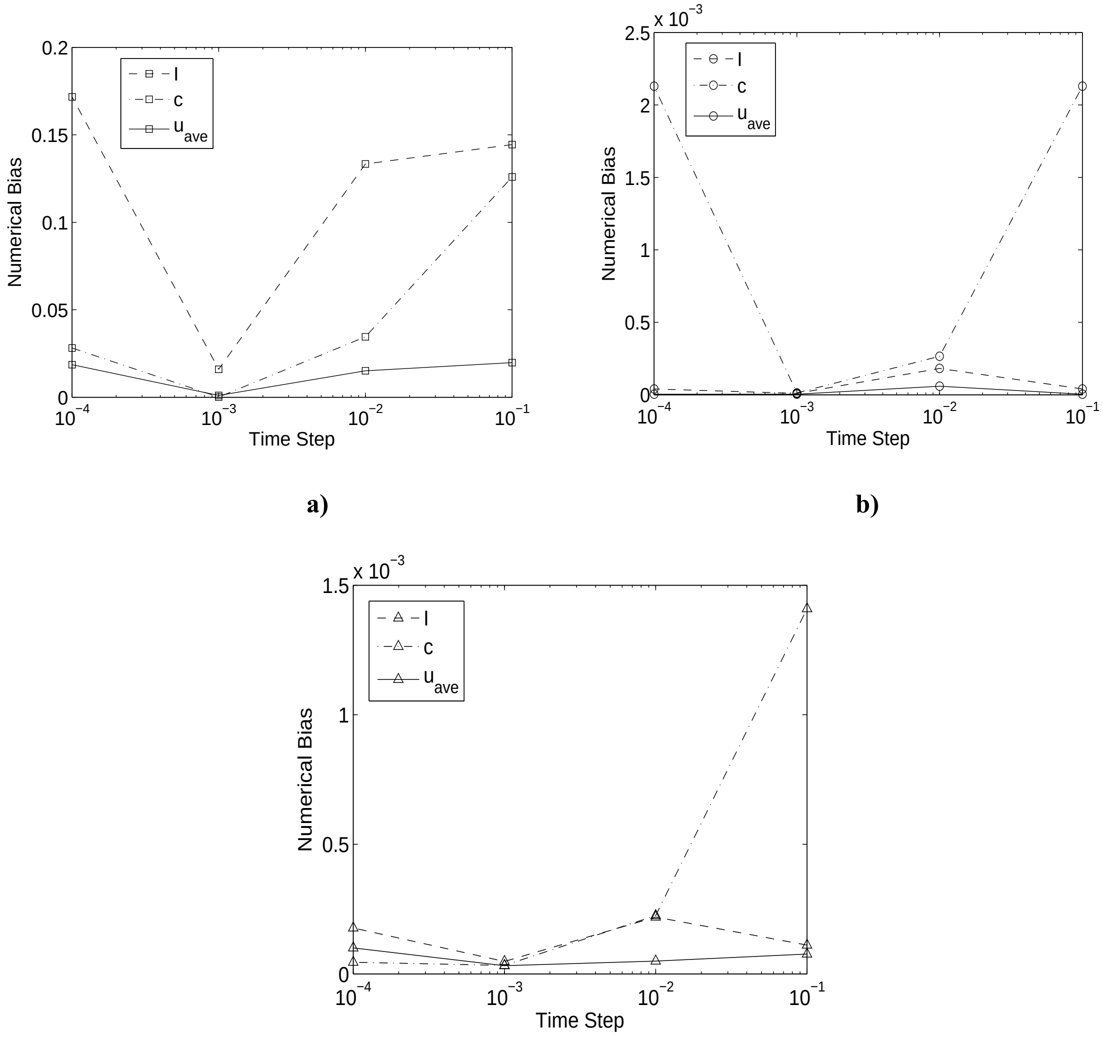
According to the results, the most accurate estimations of the model parameters belonged to the case with many but uncorrelated Observations (MO), followed by the case with correlated data structure (CO), and finally the single observation case (SO). This can be observed by comparing the discrepancy between *E(I)* and *E(c)* with 1 and 0.2, respectively, that represents the bias or systematic uncertainty of the solution. Also, the difference in *Std(I)*, *Std(c)*, and *Std(u)* for the three scenarios of observations shows that correlation in the data structure increases the uncertainty of the model parameters and predictions. It was observed that increasing the number of observations significantly reduces the uncertainty. Also, increasing the correlation and intensity of the observations reduced the correlation between the model parameters *I* and *c*.

The effects of numerical uncertainty can be better observed through comparing the results of the analytical model (*Ana*), numerical model with $\Delta t = 0.001$ (*Const*), and numerical model with *Δt=hyper-parameter* (*Var*). It was observed that, conditioned on having substantial amount of observations, the bias and the uncertainties in the solution of the inverse problem were not highly affected by the numerical error. However, to achieve this, the step size needs to be sufficiently small or assessed as a random variable through the Bayesian inversion. Also, it was noticed that the constant time step resulted in less bias and uncertainty. In order to control the Monte-Carlo error, all the statistics were computed with the same number of samples.

Effect of Time Discretization on Numerical Uncertainty and Computational Cost

Figure 7-35 and Figure 7-36 present the variation of the numerical bias and uncertainty with respect to the time step size. It was observed that the numerical bias reached its minimum at $\Delta t = 0.001$. Also, the numerical bias for the SO case was much higher than the MO and CO cases. The numerical uncertainty for the SO case decreased significantly by reducing the time, whereas for the MO and CO cases [0.001 0.01] was the optimum range.

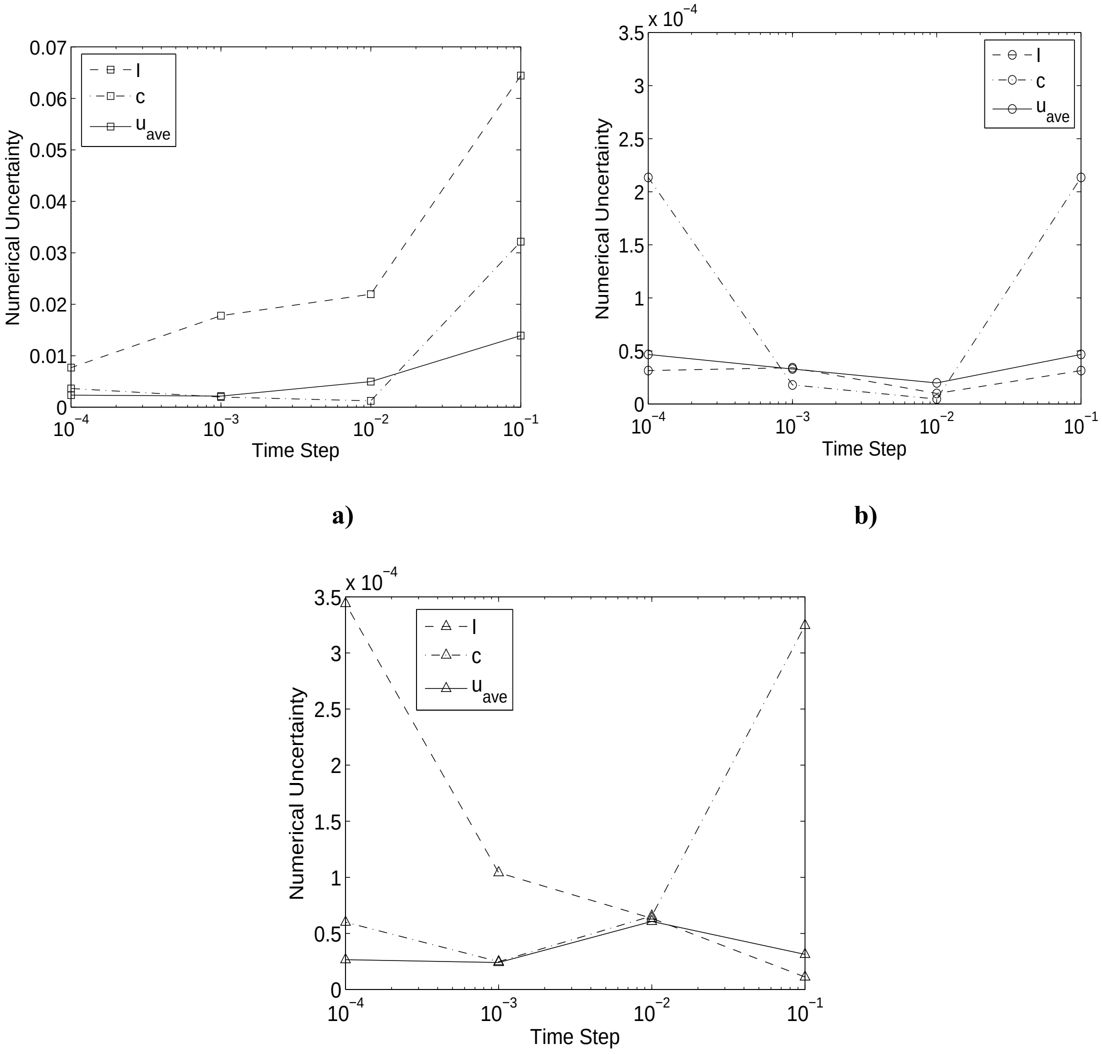
The computational cost was measured by the average time required for executing a single MCMC simulation. The average time of simulations vs. the time step size is presented in Figure 7-37. These simulations were carried out on a Quad-core CPU running at 2.4 GHz, with 64 GB RAM. As it was observed, the computation cost is almost the same for the different time steps. However, the cost for the MO experiment was almost 10^7 times more than SO. Also, the cost of computation for CO was slightly more than MO.



a) **b)**

c)

Figure 7-35. Numerical Bias vs. Time Step for Different Observations Conditions, a) SO, b) MO, c) CO



a)

b)

c)

Figure 7-36. Numerical Uncertainty vs. Time Step for Different Observations Conditions, a) SO, b) MO, c) CO

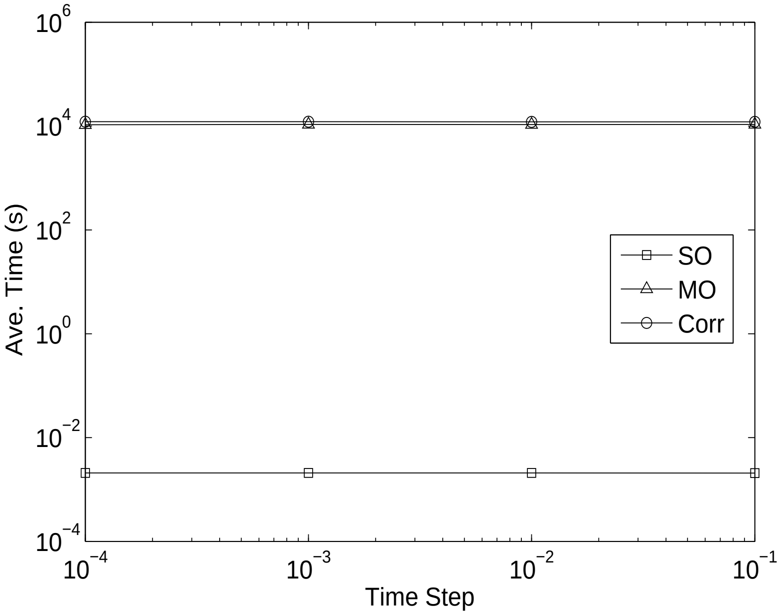


Figure 7-37. Computational Cost vs. Time Step Size for Different Observation Conditions

Full Bayesian Inversion

This section presents the results of the full Bayesian inversion where all the previous assumptions about the value of error standard deviation (σ), the prior PDF parameters for the time step, i.e. the mean and standard deviation of the normal prior distribution (μ and η), and the correlation length (τ) were relaxed. Therefore, these parameters were introduced as hyper-parameters and inferred through the MCMC simulations.

Single Observation

For this experiment, initially, the two model parameters and the four hyper-parameters were considered as unknown quantities. The chains did not show convergence even after 4M simulations (It was observed earlier that considering the standard deviation of the error distribution to be constant (0.1), stationarity was achieved). In order to remedy this, another inversion was performed considering a normal prior distribution for σ with a mean of 0.1 and standard deviation of 1. Although, the chains did not converge even with an informative prior for σ, the sample chain showed a higher convergence rate. The third inversion was performed considering σ equal to 0.1 (its true value). Figure 7-38 shows the cumulative mean of the samples obtained from 8M simulations for this experiment. It was observed that considering a constant standard deviation for the error made the convergence significantly faster. Figure 7-39 presents the posterior CDFs of the parameters. The posterior joint PDFs for some of the parameters are presented in Figure 7-40.

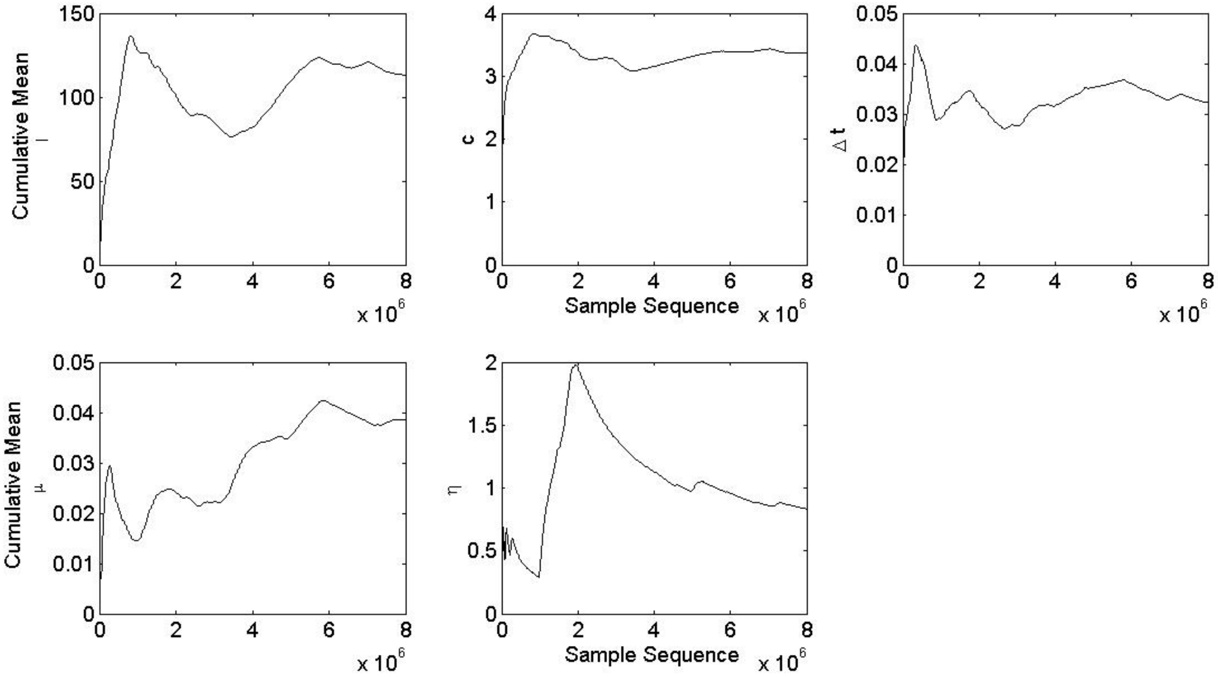


Figure 7-38. Cumulative Mean of the Samples for the Full Bayesian Inversion - SO (known σ)

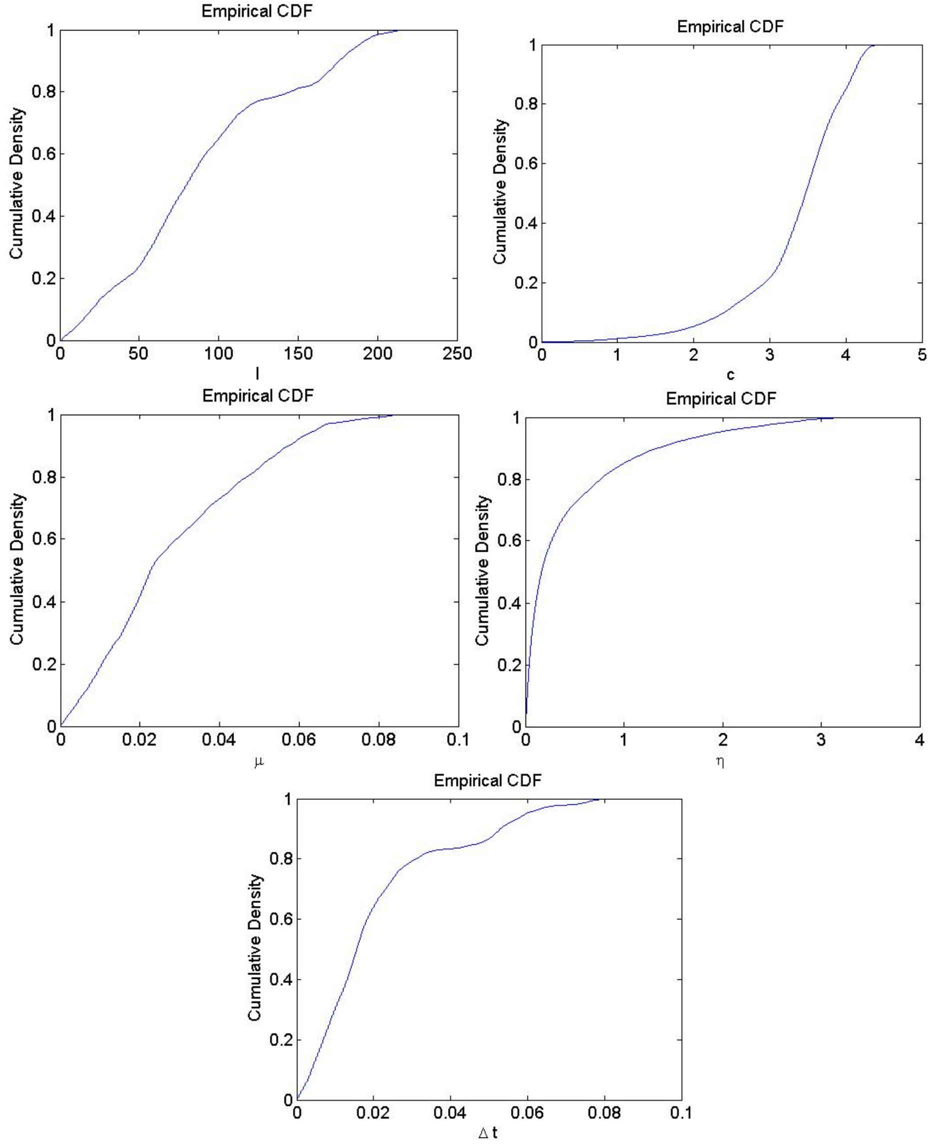


Figure 7-39. Posterior CDFs of the Parameters for the Full Bayesian Inversion - SO

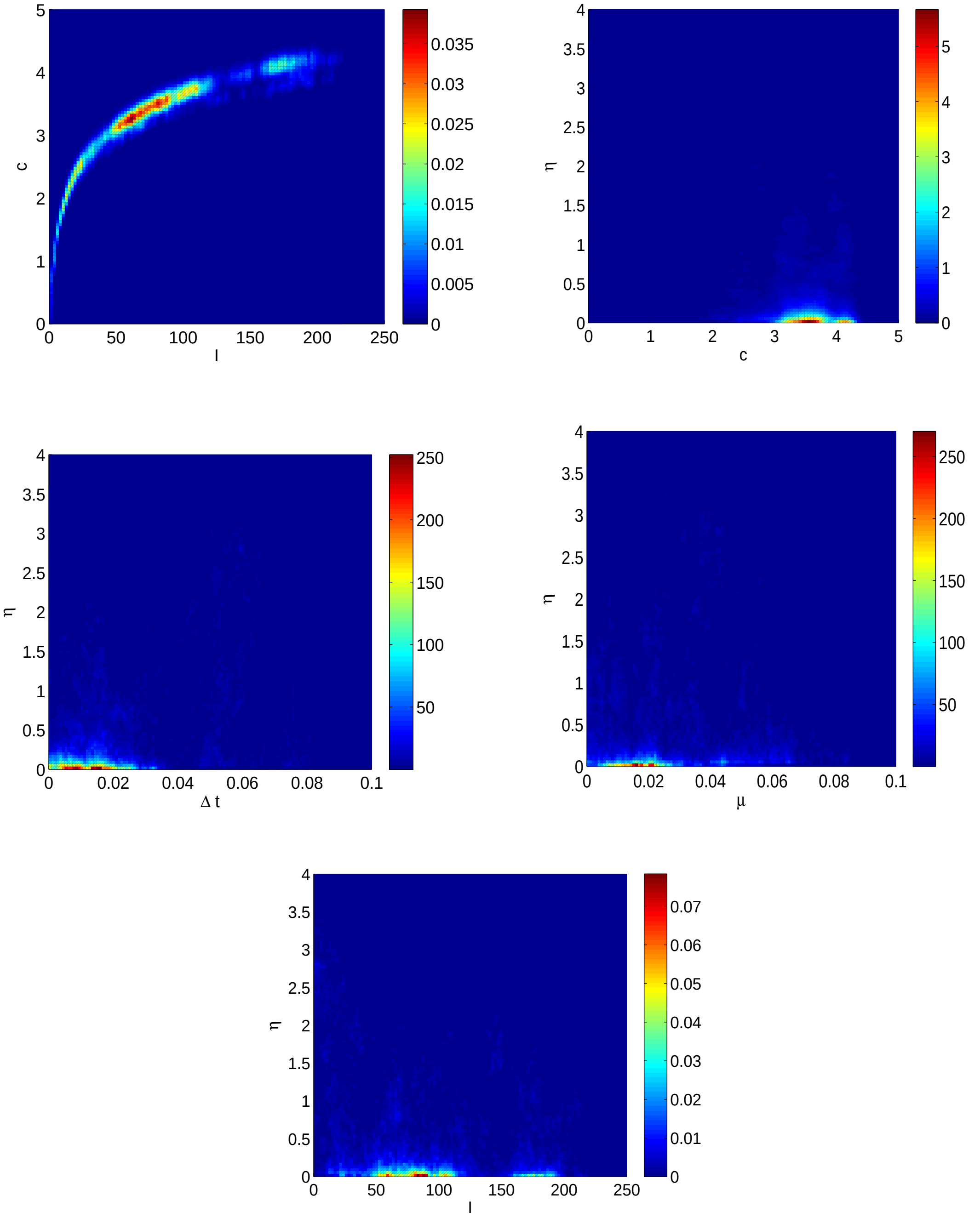


Figure 7-40. Joint Probability Distributions of the Parameters for the Full Bayesian Inversion - SO

Many Observations

Figure 7-41 shows the cumulative mean of the samples obtained from the full Bayesian inversion for the MO scenario. Comparing this with the cumulative mean in Figure 7-19, one can observe that the mean of the samples converged to the same values for the model parameters, I and c up to three digits precision. Also, the sample chain for σ converged to its true value (0.1). However, the sample chains of other hyper-parameters did not achieve stationarity. The cumulative mean of Δt varied in a range between 0.5×10^{-4} and 6×10^{-4}. This range is slightly higher than the value that the sample converged to, in the case where the prior distribution parameters were known a priori. Figure 7-42 presents the posterior CDFs of the unknown parameters. Also, Figure 7-43 presents the posterior joint distributions of the unknown parameters. It was observed that the posterior distributions for both Δt and μ were multi-modal.

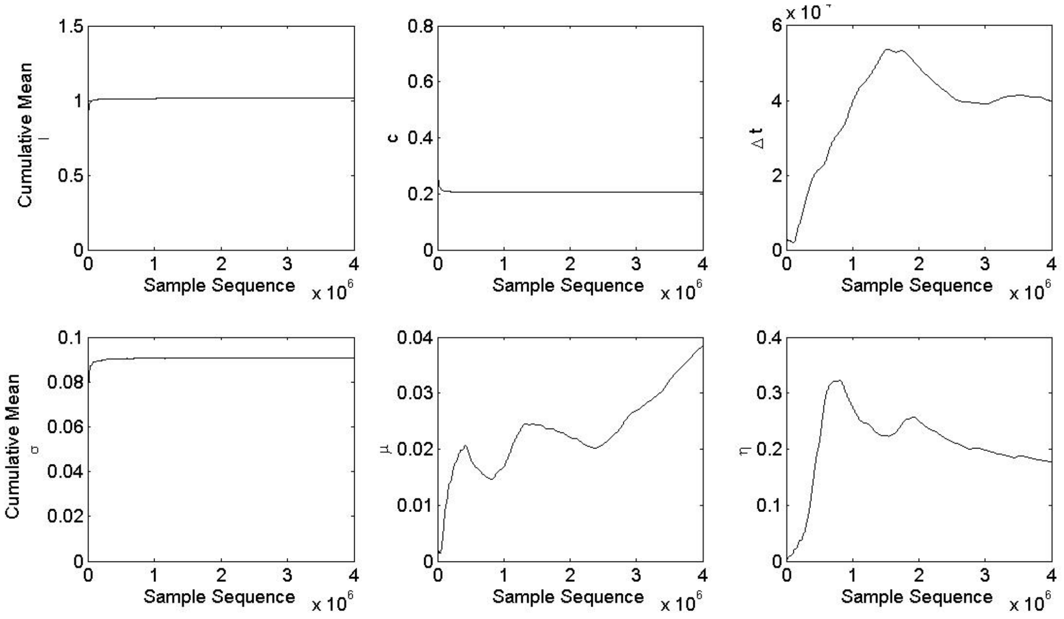


Figure 7-41. Cumulative Mean of the Samples for the Full Bayesian Inversion – MO

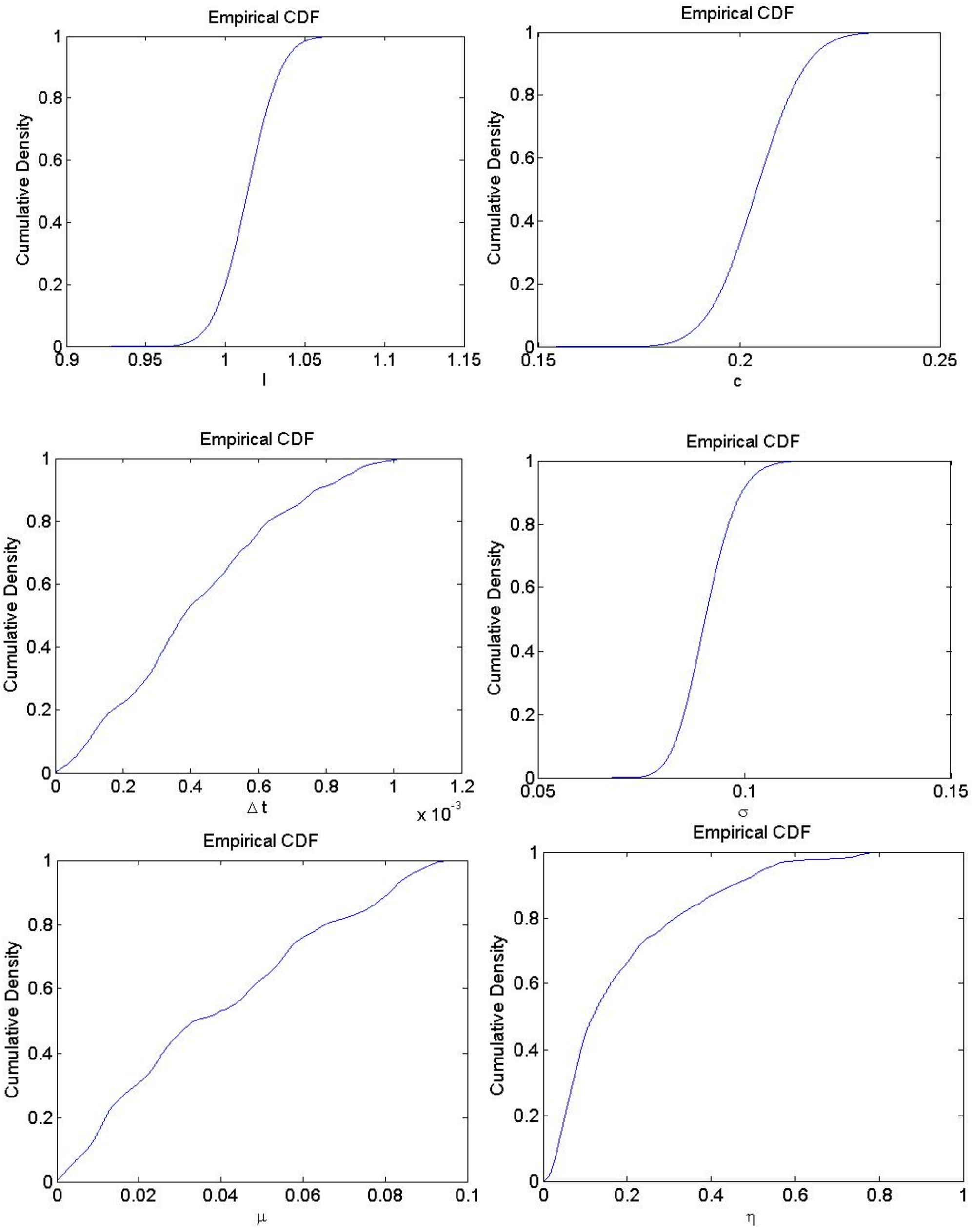


Figure 7-42. Posterior CDFs of the Parameters for the Full Bayesian Inversion - MO

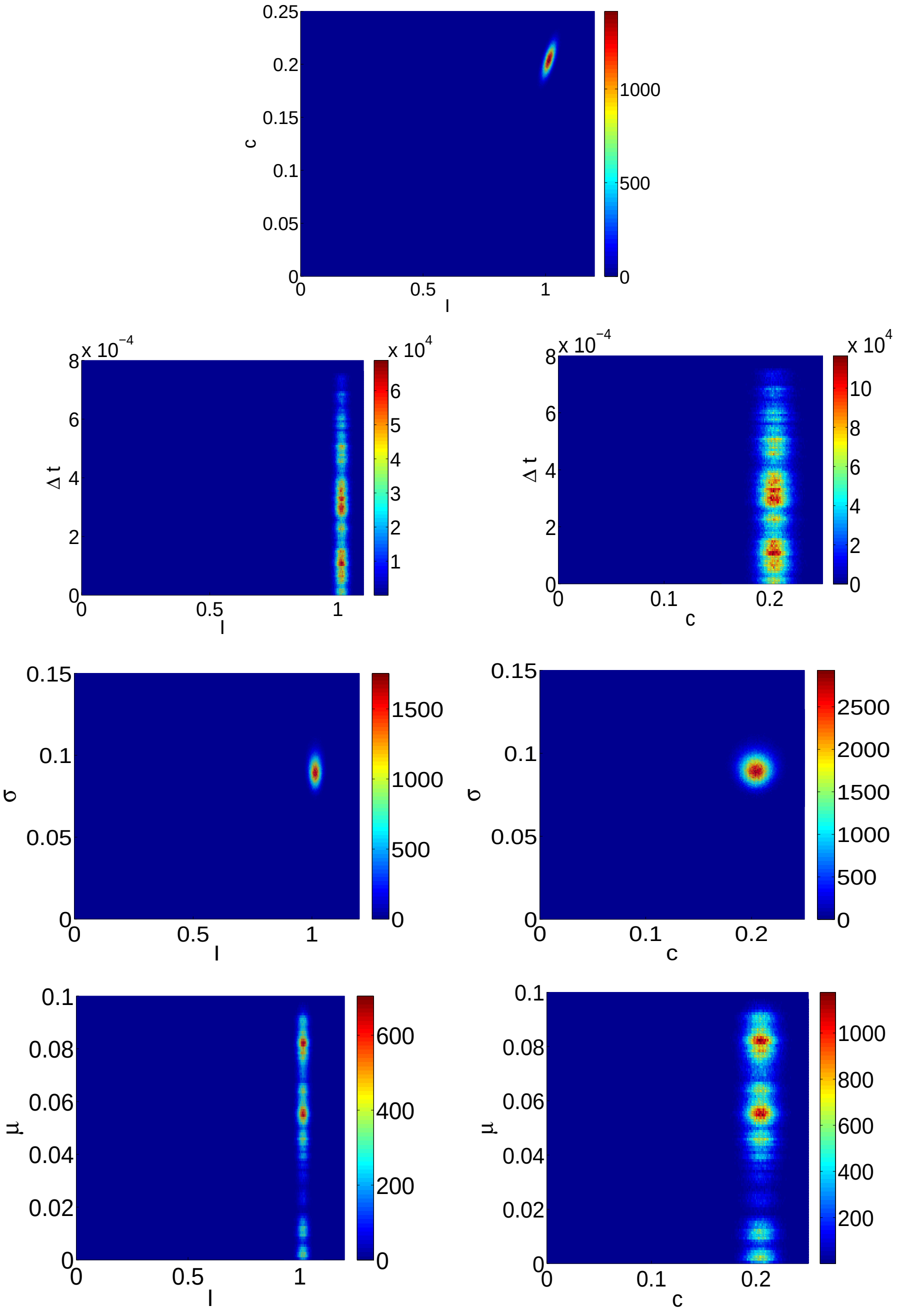


Figure 7-43. Joint Probability Distributions of the Parameters for the Full Bayesian Inversion - MO

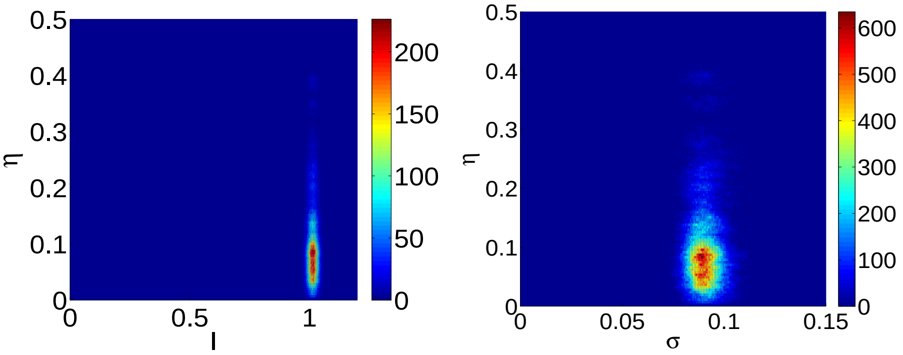


Figure 7-43. Continued

Correlated Observations

As indicated earlier, the correlation structure of the observations was modeled based on the function given by Eq. 7-12, where τ represents the correlation length parameter that controls the amount of correlation between the observations. In this analysis, for the sake of simplicity, the inverse of τ was considered and denoted by $\delta = \frac{1}{\tau}$.

Figure 7-44 shows the cumulative mean of the samples obtained from the full Bayesian inversion based on the CO scenario. The cumulative mean of the sample chains for the model parameters I and c converged to the same value (up to two digits) as for the case where σ, μ, and η were assumed known *a priori* (see Figure 7-26). The cumulative mean for Δt converged to a slightly higher value, relative to the case where Δt was the only hyper-parameter. The true values for the hyper-parameters σ and δ are 0.1 and 1, respectively. The mean of the sample converged to 0.0844 for σ and to 1.53 for δ that are very close to the actual values. However, the chain did not achieve stationarity for other hyper-parameters. Figure 7-45 presents the posterior CDF and Figure 7-46 presents the joint probability distributions of the unknown parameters. Δt, μ

and to some extent *η* show multi-modal joint distributions. The model parameters *I* and *c* show positive correlation, whereas *δ* and *σ* showed negative correlation. This negative correlation indicates that as the correlation between the observations increase, the standard deviation of the error (the uncertainty in the model predictions) increases as well.

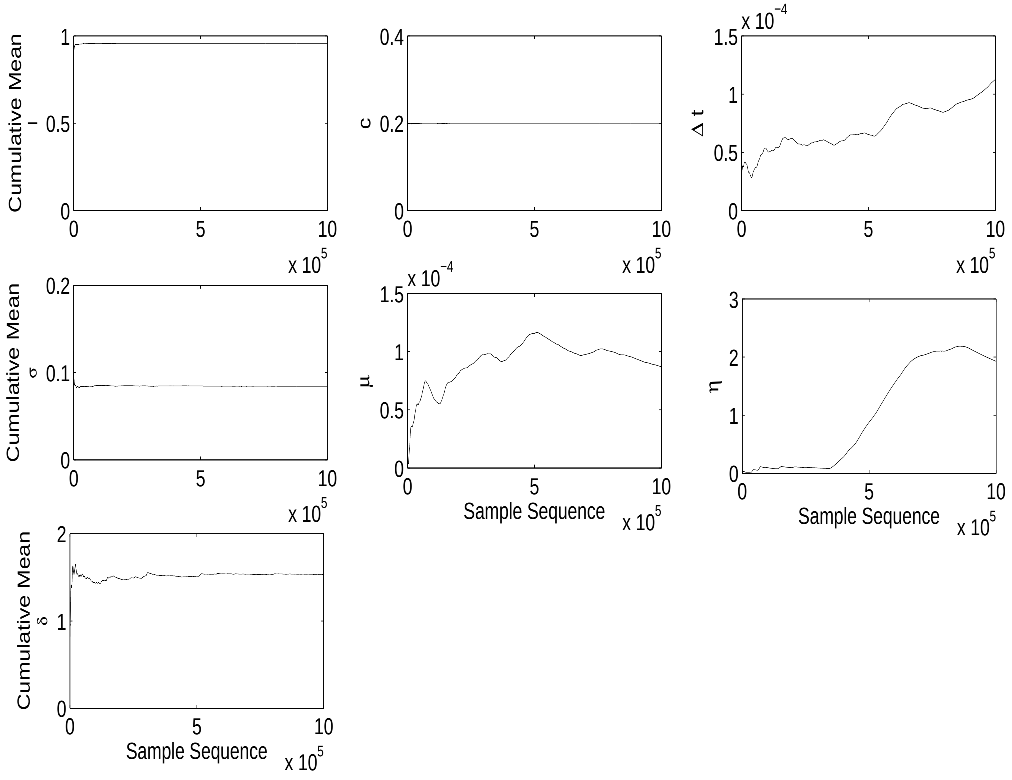


Figure 7-44. Cumulative Mean of the Samples for the Full Bayesian Inversion - CO

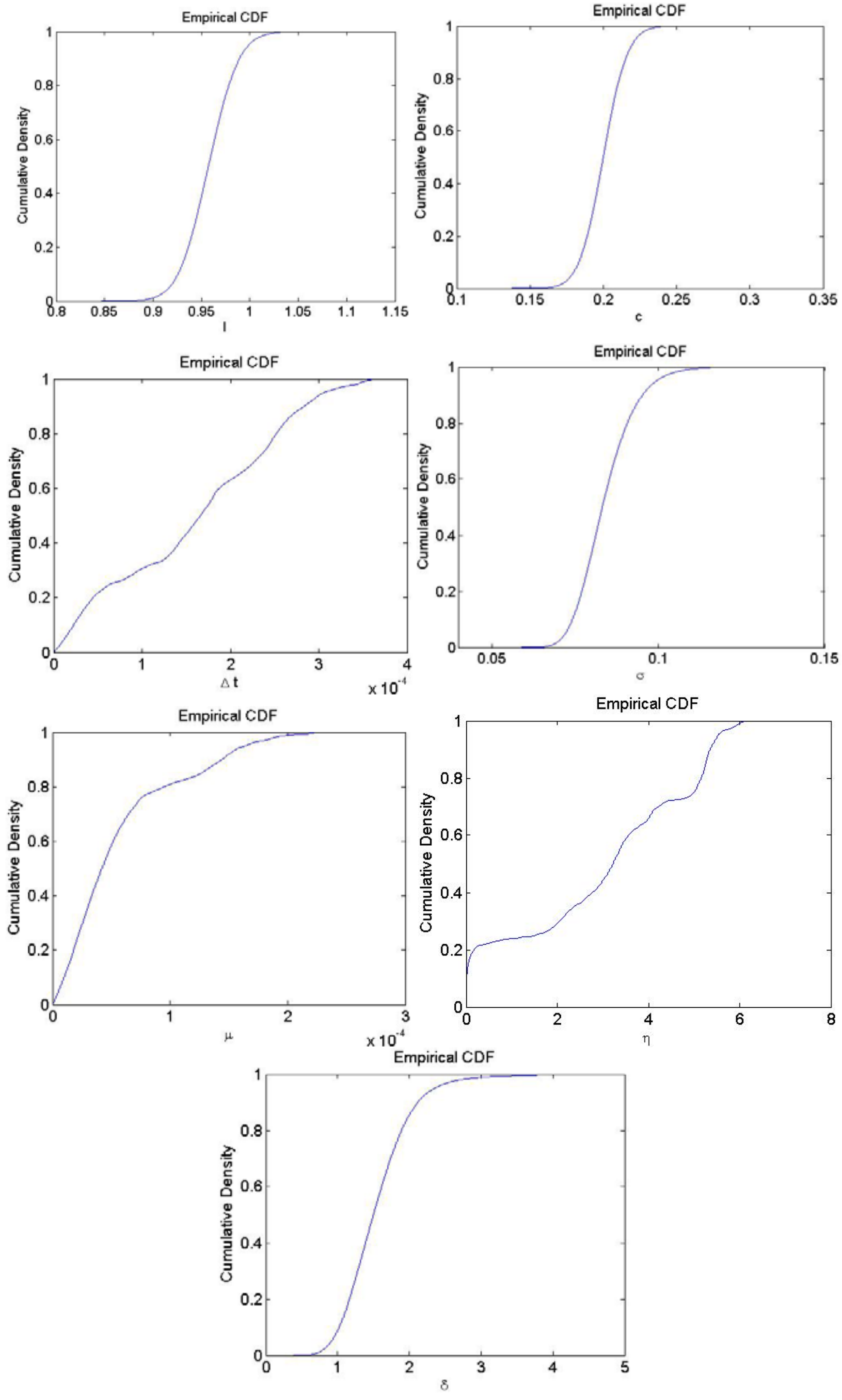


Figure 7-45. Posterior CDF of the Parameters for Full Bayesian Inversion - CO

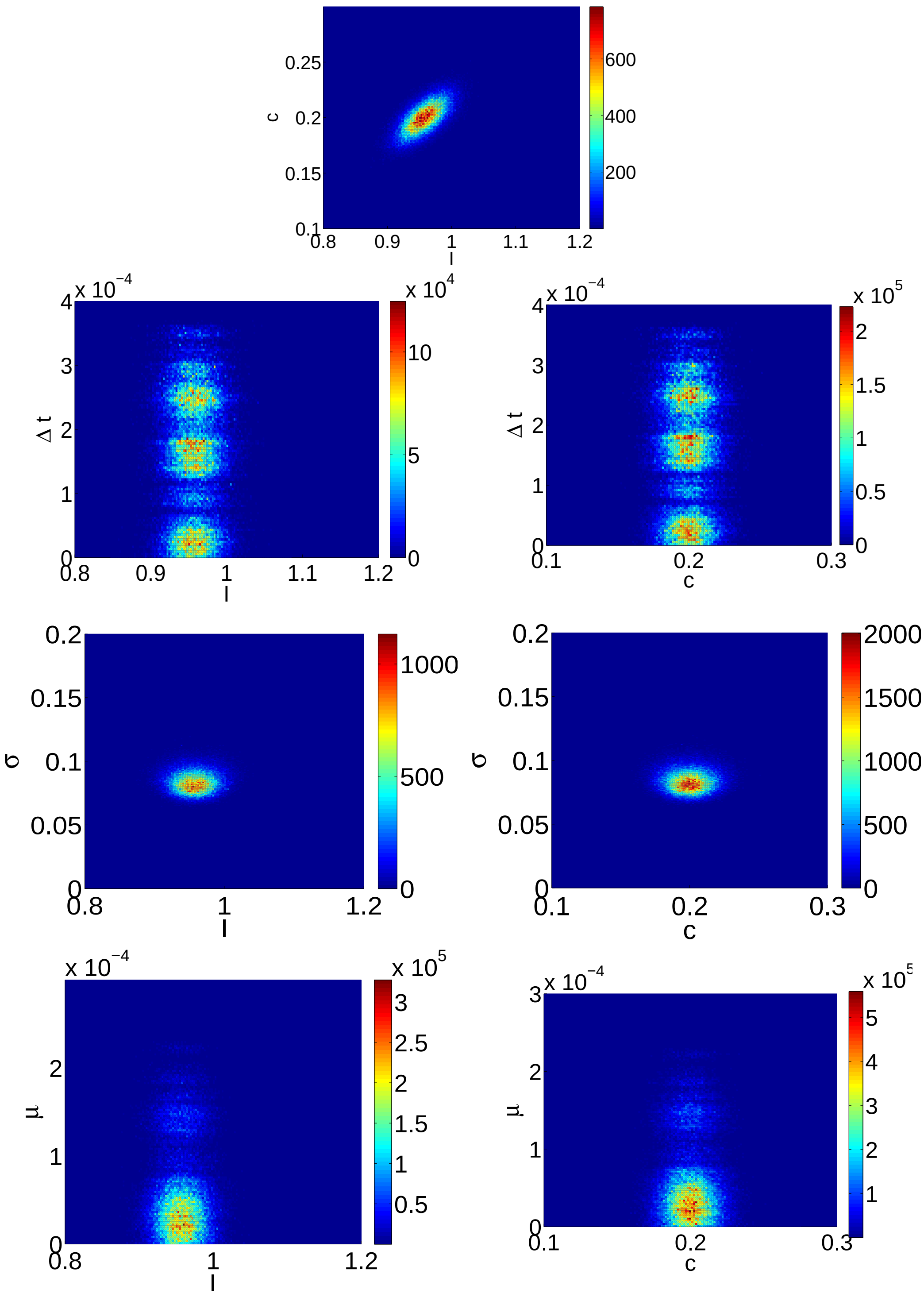


Figure 7-46. Joint Probability Distributions of the Parameters in the Full Bayesian Inversion – CO

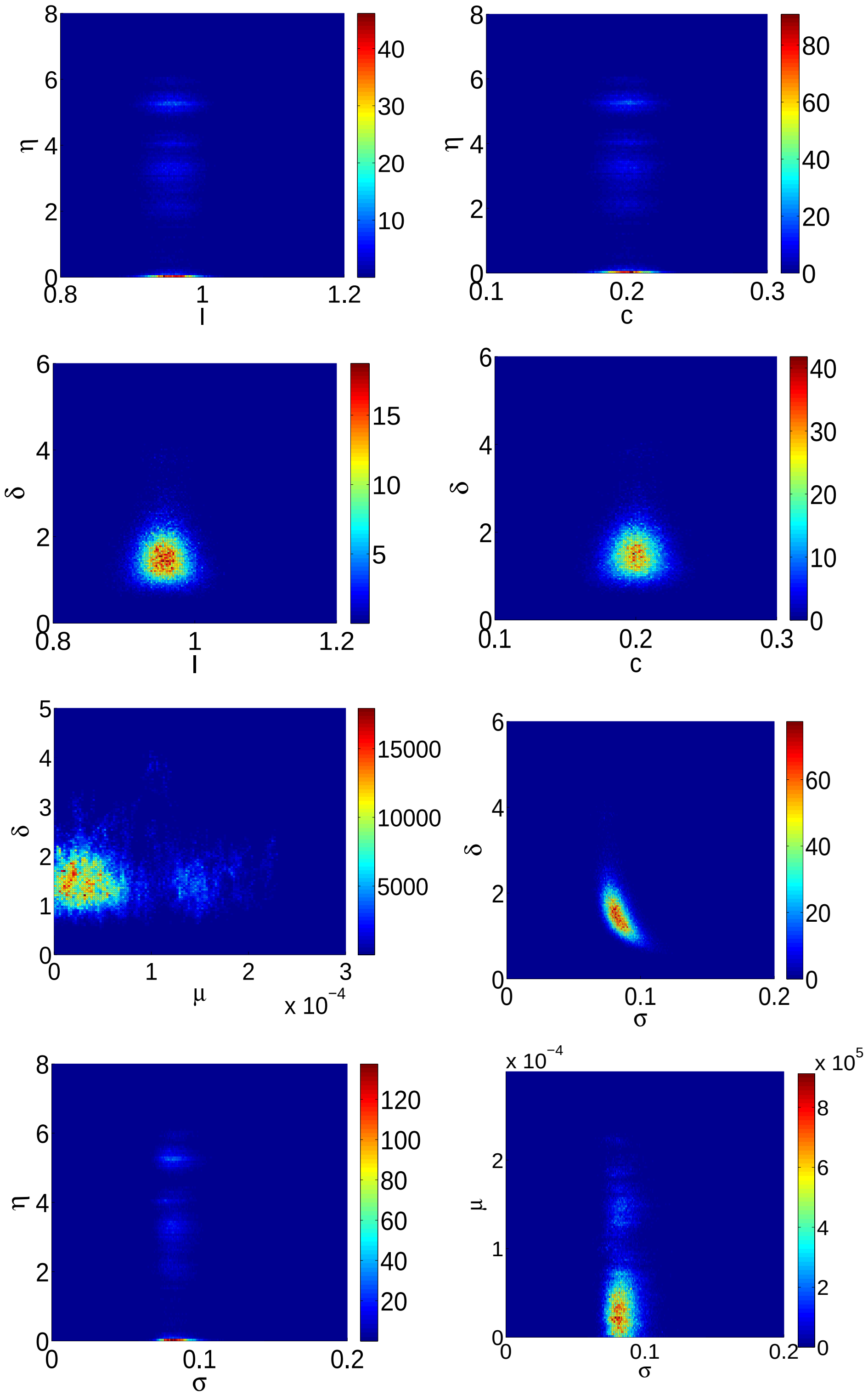


Figure 7-46. Continued

Conclusions

Both parametric and numerical uncertainties in the solution of inverse problems were explored using the Bayesian inference method. This study specifically focused on the uncertainty induced by the numerical solution of the ODE. The numerical uncertainty was evaluated through series of experiments performed under three scenarios of observations. The effects of observations' intensity and correlation, the discretization (step) size, and the prior and likelihood formulations were explored.

It was shown that higher correlation in the structure of the data can increase both the parametric and numerical uncertainty in the solution of the inverse problem, whereas, increasing the data intensity caused drastic decrease in the uncertainty.

The numerical uncertainty evaluated was insignificant for the experiments with many observations, considering the step size sufficiently small or as a hyper-parameter. Comparing the experiment incorporating constant Δt with the cases where Δt was considered as a hyper-parameter, it was found that the constant Δt resulted in the more accurate parameterization and less uncertain posterior model predictions. However, for the case with a single observation the numerical uncertainty was significantly higher and decreased by reducing the step size. For the experiments with many observations, the numerical uncertainty reached its minimum in the range [0.001 0.01] of the step size, whereas, the optimum value for the time step was 0.001 in terms of numerical bias.

The computational cost for the MCMC simulations was measured for the different experiments. The computational cost did not change much by decreasing the time step for this problem. However, the computational cost of simulation for the case of many observations was 10^7 larger than the single observation case. Also, the cost for the case with correlated observations was higher than the uncorrelated data.

CHAPTER VIII

CONCLUSIONS

Theoretical and Methodological Conclusions

Different aspects of developing deterministic and probabilistic solutions for inverse problems in geotechnics were explored in this study. A variety of computational methods including linear and nonlinear regression, ANNs, Bayesian inference, and the MCMC method were implemented. The theoretical and computational merits and limitations of these methods were discussed in application to prediction of stiffness and strength of organic soils, determination of unknown foundations for bridges subjected to scour, and uncertainty quantification for one-dimensional diffusion processes.

This study showed how ANN modeling can successfully be applied to develop nonlinear predictive models for high-dimensional problems, where many factors impact the response variables. The generated response graphs demonstrated that ANNs can account for the nonlinear relations between the response and the input variables. The ANN-based sensitivity analysis techniques that were implemented in this study allowed for ranking the input parameters, in terms of their statistical significance, and eliminating the redundant parameters.

Various types of ANN models including MLP, RBF, GRNN, and PNN were implemented for both approximation and classification of the target variable. It was found that MLP outperforms RBF and GRNN for high-dimensional problems where there is limited training data. In classification problems, MLP outperformed PNN in terms of the classification accuracy. The uncertainty of the model predictions was assessed by generating an ensemble of networks, using the random subsampling method (also known as Monte-Carlo cross-validation). It also allowed for obtaining PDFs and making interval estimations for the quantity of interest.

The Levenberg-Marquardt algorithm proved to be an effective optimization technique for MLPs. A comparative analysis between the nonlinear least square optimization and the BP algorithm showed that BP outperforms the regular nonlinear least square method. Also, the ANN models showed significantly better performances as compared to the benchmark LR models.

The Bayesian probabilistic solution developed for the inverse problems allowed for incorporating the current knowledge about the unknown quantities (in terms of prior probability distributions) and updating that based on the new observations. This resulted in narrower posterior probability distributions for the parameters and reducing the uncertainty in the predictions. The Bayesian inference method generated joint probability distributions of the unknown parameters. The joint distributions not only allowed for making probabilistic inferences (for instance, a value for the embedment of a foundation that the actual depth would be less than that with 10% probability), but also made it possible to evaluate the correlation between the unknown quantities.

Prediction of the Stiffness and Strength of Stabilized Organic Soils

Conclusions and Remarks

- ANN modeling was successfully applied to develop nonlinear models for the mechanical properties of stabilized organic soils with different organic contents and a variety of binder mixes.
- According to the stepwise parameter selection and the sensitivity analysis method, grout to soil ratio, quantity of binder, binder mix type, amount of organic matter, water to binder ratio, temperature, and time (aging) were the most influential parameters, respectively for the prediction of *E* and *UCS*.

Recommendations for Future Research

- Investigating the impact of other influential parameters on stiffness and strength of the stabilized organic soils that were not considered in this study, for instance soil type and grain distribution.
- Developing a probabilistic predictive model by implementing a Bayesian paradigm in computing the weights of the neural networks**.**
- Implementing other feature selection methods to identify the most significant parameters and developing more effective reduced models.

Determination of Unknown Foundations for Bridge Scour

Conclusions and Remarks

- The developed ANN classifiers facilitated the categorization of the unknown foundations as either deep or shallow. They were also able to accurately classify the foundations among the eight different common types.
- The results of the predicted versus the actual pile depth showed that the ANN models were able to predict the embedment depth of deep foundations with $R^2_{ave} > 0.80$.
- The generated posterior CDFs for the foundation type, the foundation dimensions, and the soil resistance parameters allowed for making probabilistic inferences for these parameters, including point estimations and interval estimations.
- The estimated 10% quantile of the probability distributions for the embedment depth of the foundations provided a conservative estimation to calculate the scour failure risk.

Recommendations for Future Research

- Incorporating the data from geophysical experiments on unknown foundations of bridges, such as the resistivity imaging and induced polarization methods, into the ANN models to develop more reliable predictive models.
- Performing finite element analyses to predict the depth of the unknown foundations by modeling the superstructure and substructure of bridges.

Uncertainty Quantification for One-Dimensional Diffusion Processes

Conclusions and Remarks

- The time step size Δt was introduced as a hyper-parameter and was evaluated as a random variable through the Bayesian inversion. Except for the single observation scenario, the MCMC simulations for Δt converged to stationary distributions.
- The numerical uncertainty decreased by reducing the step size for the experiment with a single observation. However, for the case with many observations the numerical uncertainty reached its minimum in the range [0.001 0.01] of the step size. The optimum value for the time step was 0.001 in terms of numerical bias.
- According to results, higher correlation in the structure of the data increased both the parametric and numerical uncertainties in the solution of the inverse problem, whereas, increasing the data intensity caused drastic decrease in the uncertainties.

Recommendations for Future Research

- Investigating the effect of both parametric and numerical uncertainties in the solution of the partial differential equation for diffusion, under different observation scenarios.
- Investigating the computational efficiency of stochastic surrogate models in representing the two-dimensional diffusion processes in comparison to the MC method.

REFERENCES

AASHTO (2002a). "AASHTO Standard Specifications for Highway Bridges." American Association of State Highway and Transportation Officials, Washington, DC.

AASHTO (2002b). "LRFD Highway Bridge Design Specifications." American Association of State Highway and Transportation Officials, Washington, DC.

Adeli H. (2001). "Neural Networks in Civil Engineering: 1989–2000." *Computer-Aided Civil and Infrastructure Engineering*, 16(2), 126-142.

Ahnberg H. (2006). "Strength of Stabilised Soils: A Laboratory Study on Clays and Organic Soils Stabilised with Different Types of Binder." Department of Construction Sciences Lund University, Lund, Sweden.

Arangio S. and Bontempi, F. (2010). "Soft Computing Based Multilevel Strategy for Bridge Integrity Monitoring." *Computer-Aided Civil and Infrastructure Engineering*, 25(5), 348-362.

Basheer I. A. and Najjar, Y. M. (1995). "A Neural-Network for Soil Compaction." *5th Int. Symp. Numerical Models in Geomechanics*, G. N. Pande and S. Pietruszczak, eds., Roterdam: Balkema.

Basheer I. A., Reddi, L. N. and Najjar, Y. M. (1996). "Site Characterization by Neuronets: An Application to the Landfill Siting Problem." *Ground Water*, 34(4), 610-617.

Bateni S. M., Borghei, S. M. and Jeng, D. S. (2007). "Neural Network and Neuro-Fuzzy Assessments for Scour Depth around Bridge Piers." *Engineering Applications of Artificial Intelligence*, 20(3), 401-414.

Bishop C. M. (1995). *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford University Press, Oxford.

Bresler E. (1973). "Simultaneous Transport of Solutes and Water under Transient Unsaturated Flow Conditions." *Water Resources Research*, 9(4), 975-986.

Bresler E. and Hanks, R. (1969). "Numerical Method for Estimating Simultaneous Flow of Water and Salt in Unsaturated Soils." *Soil Science Society of America Journal*, 33(6), 827-832.

Briaud J.-L., Medina-Cetina, Z., Hurlebaus, S., Everett, M., Tucker, S., Yousefpour, N. and Arjwech, R. (2012). "Unknown Foundation Determination for Scour." Report No. 0-6604-1, FHWA.

Chew S. H., Kamruzzaman, A. H. M. and Lee, F. H. (2004). "Physicochemical and Engineering Behavior of Cement Treated Clays." *Journal of Geotechnical and Geoenvironmental Engineering*, 130(7), 696-706.

Clare K. E. and Sherwood, P. T. (1956). "Further Studies on the Effect of Organic Matter on the Setting of Soil-Cement Mixtures." *Journal of Applied Chemistry*, 6(8), 317-324.

Consoli N. C., Cruz, R. C., Floss, M. F. and Festugato, L. (2010). "Parameters Controlling Tensile and Compressive Strength of Artificially Cemented Sand." *Journal of Geotechnical and Geoenvironmental Engineering*, 136(5), 759-763.

Consoli N. C., Da Fonseca, A. V., Cruz, R. C. and Heineck, K. S. (2009). "Fundamental Parameters for the Stiffness and Strength Control of Artificially Cemented Sand." *Journal of Geotechnical and Geoenvironmental Engineering*, 135(9), 1347-1353.

Consoli N. C., Foppa, D., Festugato, L. and Heineck, K. S. (2007). "Key Parameters for Strength Control of Artificially Cemented Soils." *Journal of Geotechnical and Geoenvironmental Engineering*, 133(2), 197-205.

Daniel D. and Shackelford, C. (1988). "Disposal Barriers that Release Contaminants only by Molecular Diffusion." *Nuclear and Chemical Waste Management*, 8(4), 299-305.

Das S., Samui, P. and Sabat, A. (2011). "Application of Artificial Intelligence to Maximum Dry Density and Unconfined Compressive Strength of Cement Stabilized Soil." *Geotechnical and Geological Engineering*, 29(3), 329-342.

Demuth H., Beale, M. and Hagan, M. (2009). "MATLAB Neural Network Toolbox User's Guide." The MathWorks.

Den Haan E. and Kruse, G. (2007). "Characterisation and Engineering Properties of Dutch Peats." *Proceedings of the Second International Workshop of Characterisation and Engineering Properties of Natural Soils,* Singapore.

Desaulniers D. E. and Cherry, J. A. (1989). "Origin and Movement of Groundwater and Major Ions in a Thick Deposit of Champlain Sea Clay near Montreal." *Canadian Geotechnical Journal*, 26(1), 80-89.

Dostert P., Efendiev, Y. and Hou, T. (2008). "Multiscale Finite Element Methods for Stochastic Porous Media Flow Equations and Application to Uncertainty

Quantification." *Computer Methods in Applied Mechanics and Engineering*, 197(43), 3445-3455.

Edil T. N. and Den Haan, E. J. (1994). "Settlement of Peats and Organic Soils Settlement." *ASCE Geotechnical Special Publication*, 4(2), 1543-1572.

Elkordy M. F., Chang, K. C. and Lee, G. C. (1994). "A Structural Damage Neural Network Monitoring System." *Computer-Aided Civil and Infrastructure Engineering*, 9(2), 83-96.

Ellis G. W., Yao, C., Zhao, R. and Penumadu, D. (1995). "Stress-Strain Modeling of Sands Using Artificial Neural Networks." *Journal of Geotechnical Engineering*, 121(5), 429-435.

Elman H. C., Miller, C. W., Phipps, E. T. and Tuminaro, R. S. (2011). "Assessment of Collocation and Galerkin Approaches to Linear Diffusion Equations with Random Data." *International Journal for Uncertainty Quantification*, 1(1), 19-33.

FHWA (2011). "Unknown Foundations." US Department of Transportation, <http://www.fhwa.dot.gov/unknownfoundations/>.

Gallagher K., Charvin, K., Nielsen, S., Sambridge, M. and Stephenson, J. (2009). "Markov Chain Monte Carlo (MCMC) Sampling Methods to Determine Optimal Models, Model Resolution and Model Choice for Earth Science Problems." *Marine and Petroleum Geology*, 26(4), 525-535.

Gamerman D. and Lopes, H. F. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman & Hall/CRC, New York.

Gardner W. and Mayhugh, M. (1958a). "Solutions and Tests of the Diffusion Equation for the Movement of Water in Soil." *Soil Science Society of America Journal*, 22(3), 197-201.

Gardner W. R. and Mayhugh, M. S. (1958b). "Solutions and Tests of the Diffusion Equation for the Movement of Water in Soil." *Soil Sci. Soc. Am. J.*, 22(3), 197-201.

Gevrey M., Dimopoulos, L. and Lek, S. (2003). "Review and Comparison of Methods to Study the Contribution of Variables in Artificial Neural Network Models." *Ecological Modeling*, 160(3), 249-264.

Ghaboussi J. and Sidarta, D. E. (1998). "New Nested Adaptive Neural Networks (NANN) for Constitutive Modeling." *Computers and Geotechnics*, 22(1), 29-52.

Ghanem R. (1998). "Probabilistic Characterization of Transport in Heterogeneous Media." *Computer Methods in Applied Mechanics and Engineering*, 158(3), 199-220.

Ghanem R. (1999). "Stochastic Finite Elements with Multiple Random Non-Gaussian Properties." *Journal of Engineering Mechanics*, 125(1), 26-40.

Ghanem R., Red-Horse, J. R. and Sarkar, A. (2000). "Modal Properties of a Spaceframe with Localized System Uncertainties." *8th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability,* Kareem A, Haldar A, Spencer Jr BF, Johnson EA (eds). ASCE: New York, 2000. Number PMC200-269.

Ghanem R. G. and Spanos, P. D. (2003). *Stochastic Finite Elements: a Spectral Approach*. Courier Dover Publications, Mineola, New York.

Gillham R. W., Robin, M. J. L., Dytynyshyn, D. J. and Johnston, H. M. (1984). "Diffusion of Nonreactive and Reactive Solutes through Fine-Grained Barrier Materials." *Canadian Geotechnical Journal*, 21(3), 541-550.

Goh A. T. C. (1994a). "Nonlinear Modelling in Geotechnical Engineering Using Neural Networks." *Australian Civil Engineering Transactions, CE*, 36(4), 293-297.

Goh A. T. C. (1994b). "Seismic Liquefaction Potential Assessed by Neural Networks." *Journal of Geotechnical Engineering*, 120(9), 1467-1480.

Goh A. T. C. (1995). "Empirical Design in Geotechnics Using Neural Networks." *Empirical Design in Geotechnics Using Neural Networks*, 45(4), 709-714.

Goh A. T. C. (1996). "Pile Driving Records Reanalyzed Using Neural Networks." *Journal of Geotechnical Engineering.*, 122(6), 492.

Goh A. T. C., Kulhawy, F. H. and Chua, C. G. (2005). "Bayesian Neural Network Analysis of Undrained Side Resistance of Drilled Shafts." *Journal of Geotechnical and Geoenvironmental Engineering*, 131(1), 84-93.

Goh A. T. C., Wong, K. S. and Broms, B. B. (1995). "Estimation of Lateral Wall Movements in Braced Excavations Using Neural Networks." *Canadian Geotechnical Journal*, 32(6), 1059-1064.

Goodall D. C. and Quigley, R. (1977). "Pollutant Migration from Two Sanitary Landfill Sites near Sarnia, Ontario." *Canadian Geotechnical Journal*, 14(2), 223-236.

Grubb D. G., Chrysochoou, M., Smith, C. J. and Malasavage, N. E. (2010a). "Stabilized Dredged Material. I: Parametric Study." *Journal of Geotechnical and Geoenvironmental Engineering*, 136(8), 1011-1024.

Grubb D. G., Malasavage, N. E., Smith, C. J. and Chrysochoou, M. (2010b). "Stabilized Dredged Material. II: Geomechanical Behavior." *Journal of Geotechnical and Geoenvironmental Engineering*, 136(8), 1025-1036.

Hampton M. B. and Edil, B. (1998). "Strength Gain of Organic Ground with Cement-Type Binders." *Proceedings of Sessions of Geo-Congress 98, Soil Improvement for Big Digs,* Geotechnical Special Publication, ASCE.

Hayashi H. and Nishimoto, S. (2005). "Strength Characteristics of Stabilized Peat Using Different Types of Binders." *Proceedings of the International Conference on Deep Mixing*, Stockholm, Sweden.

Haykin S. (1999). *Neural Networks: A Comprehensive Foundation*. Prentice Hall International Inc., Upper Saddle River, NJ.

Hernandez-Martinez F. G. (2006). "Ground Improvement of Organic Soils Using Wet Deep Soil Mixing." Ph.D. Thesis, University of Cambridge, Cambridge, UK.

Hoff P. D. (2009). *A First Course in Bayesian Statistical Methods*. Springer, New York, NY.

Huttunen E. and Kujala, K. (1996). "On the Stabilization of Organic Soils." *Proceeding of the Second International Conference on Ground Improvement Geosystems*, Tokyo.

Jelisic N. and Leppanen, M. (2003). "Mass Stabilisation of Organic Soils and Soft Clay." *Proceedings of the Third International Conference on Grouting and Ground Treatment*, Geotechnical Publication No. 120, ASCE.

Johnson R. L., Cherry, J. A. and Pankow, J. F. (1989). "Diffusive Contaminant Transport in Natural Clay: A Field Example and Implications for Clay-lined Waste Disposal Sites." *Environmental Science & Technology*, 23(3), 340-349.

Juang C. H. and Chen, C. J. (1999). "CPT-Based Liquefaction Evaluation Using Artificial Neural Networks." *Computer-Aided Civil and Infrastructure Engineering*, 14(3), 221-229.

Kaipio J. P. and Somersalo, E. (2005). *Statistical and Computational Inverse Problems*. Springer, New York, NY.

Karhunen K. (1946). *To the Spectral Theory of Stochastic Processes* [in German]. Finnish Academy of Sciences, Helsinki, Finland.

Kawamura K., Miyamoto, A., Frangopol, D. M. and Kimura, R. (2003). "Performance Evaluation of Concrete Slabs of Existing Bridges Using Neural Networks." *Engineering Structures*, 25(12), 1455-1477.

Kaya A. (2010). "Artificial Neural Network Study of Observed Pattern of Scour Depth around Bridge Piers." *Computers and Geotechnics*, 37(3), 413-418.

Kim S.-H., Yoon, C. and Kim, B.-J. (2000). "Structural Monitoring System Based on Sensitivity Analysis and a Neural Network." *Computer-Aided Civil and Infrastructure Engineering*, 15(4), 189-195.

Kitazume M. (2005). "Technical Session 2a: Ground Improvement." *Proceedings of the 16th International Conference on Soil Mechanics and Geotechnical Engineering,* Vol 5, Osaka, Japan.

Lambrechts J. R., Ganse, M. A. and Layhee, C. A. (2003). "Soil Mixing to Stabilize Organic Clay for I-95 Widening." *Proceedings of the Third International Conference in Grouting and Ground Treatment,*, Geotechnical Publication No. 120, ASCE.

Lee I.-M. and Lee, J.-H. (1996). "Prediction of Pile Bearing Capacity Using Artificial Neural Networks." *Computers and Geotechnics*, 18(3), 189-200.

Lee T. L., Jeng, D. S., Zhang, G. H. and Hong, J. H. (2007). "Neural Network Modeling for Estimation of Scour Depth around Bridge Piers." *Journal of Hydrodynamics, Ser. B*, 19(3), 378-386.

Lek S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J. and Aulagnier, S. (1996). "Application of Neural Networks to Modelling Monlinear Relationships in Ecology." *Ecological modelling*, 90(1), 39-52.

Loeve M. (1963). *Probability Theory*. Graduate Texts in Mathematics, Springer-Verlag, New York, NY.

Marzouk Y. M., Najm, H. N. and Rahn, L. A. (2007). "Stochastic Spectral Methods for Efficient Bayesian Solution of Inverse Problems." *Journal of Computational Physics*, 224(2), 560-586.

Maser K. R., Sanqeui, M., Lichtenstein, A. and Chase, S. B. (1998). "Determination of Bridge Foundation Type from Structural Response Measurements." *Proceeding of the Nondestructive Evaluation Techniques for Aging Infrastructure and Manufacturing Conference*. San Antonio, SPIE, 55-67.

MathWorks. (2013). < http://www.mathworks.com/>.

Mathelin L., Hussaini, M. Y. and Zang, T. A. (2005). "Stochastic Approaches to Uncertainty Quantification in CFD Simulations." *Numerical Algorithms*, 38(1-3), 209-236.

McCulloch W. and Pitts, W. (1943). "A Logical Calculus of the Ideas Immanent in Nervous Activity." *Bulletin of Mathematical Biology*, 5(4), 115-133.

McGinn A. J. and O´Rourke, T. D. (2003). "Performance of Deep Mixing Methods at Fort Point Channel." *Report to Massachussetts Turnpike Authority, Federal Highway Administration Bechtel/Parsons Brinckerhoff*, Cornell University.

Mclemore S., Zendegui, S., Whites, J., Sheppard, M., Gosselin, M., Demir, H., Passe, P. and Hayden, M. (2010). "Unknown Foundation Bridges Pilot Study." *Report to Federal Highway Administration & Florida Department of Transportation*.

Medina-Cetina Z. (2006). "Probabilistic Calibration of a Soil Model." A Ph.D. Dissertation, Johns Hopkins University, Baltimore.

Mercado E. J. and O'Neil, M. W. (2003). "Methods to Measure Scour Depth and the Depth of Unknown Foundations." *3rd International Conferences on Applied Geophysics*, Orlando, FL.

Mitchell J. K. (1981). "Soil Improvement Methods and Their Applications in Civil Engineering." Dept. of Civil Engineering, North Carolina State University, Raleigh, NC.

Najjar Y., Basheer, I. and McReynolds, R. (1996a). "Neural Modeling of Kansas Soil Swelling." *Transportation Research Record: Journal of the Transportation Research Board*, 1526(-1), 14-19.

Najjar Y. M. and Ali, H. E. (1998). "CPT-based Liquefaction Potential Assessment: A Neuronet Approach." *Geotechnical Special Publication, ASCE*, 1, 542-553.

Najjar Y. M. and Basheer, I. A. (1996). "Utilizing Computational Neural Networks for Evaluating the Permeability of Compacted Clay Liners." *Geotechnical and Geological Engineering*, 14(3), 193-212.

Najjar Y. M., Basheer, I. A. and Naouss, W. A. (1996b). "On the Identification of Compaction Characteristics by Neuronets." *Computers and Geotechnics*, 18(3), 167-187.

Najm H. N. (2009). "Uncertainty Quantification and Polynomial Chaos Techniques in Computational Fluid Dynamics." *Annual Review of Fluid Mechanics*, 41, 35-52.

Narendra B. S., Sivapullaiah, P. V., Suresh, S. and Omkar, S. N. (2006). "Prediction of Unconfined Compressive Strength of Soft Grounds Using Computational Intelligence Techniques: A Comparative Study." *Computers and Geotechnics*, 33, 196-208.

Ni S. H., Lu, P. C. and Juang, C. H. (1996). "A Fuzzy Neural Network Approach to Evaluation of Slope Failure Potential." *Computer-Aided Civil and Infrastructure Engineering*, 11(1), 59-66.

Olson L. (2005). "Dynamic Bridge Substructure Evaluation and Monitoring." Report No. FHWA-RD-03-089, FHWA.

Olson L., Jalinoos, F. and Aouad, M. F. (1998). "Determination of Unknown Subsurface Bridge Foundations." NCHRP 21-5 Interim Report Summary, USDOT Geotechnical Engineering Notebook, FHWA.

Pal M. and Deswal, S. (2008). "Modeling Pile Capacity Using Support Vector Machines and Generalized Regression Neural Network." *Journal of Geotechnical and Geoenvironmental Engineering*, 134(7), 1021-1024.

Penumadu D. and Jean-Lou, C. (1997). "Geomaterial Modeling Using Artificial Neural Networks." *Artificial Neural Networks for Civil Engineers: Fundamentals and Applications, ASCE,* 160-184.

Penumadu D. and Zhao, R. (1999). "Triaxial Compression Behavior of Sand and Gravel Using Artificial Neural Networks (ANN)." *Computers and Geotechnics*, 24(3), 207-230.

Picard R. R. and Cook, R. D. (1984). "Cross-Validation of Regression Models." *Journal of the American Statistical Association*, 79(387), 575-583.

Poggio T. and Girosi, F. (1990). "Regularization Algorithms for Learning That Are Equivalent to Multilayer Networks." *Science*, 247(4945), 978-982.

Porbaha A. (2004). "Design Aspects and Properties of Treated Ground." *Proceedings of the Deep Mixing Workshop, Geo-Trans 2004*, Los Angeles, CA.

Quigley R. M. and Rowe, R. K. (1986). "Leachate Migration through Clay below a Domestic Waste Landfill, Sarnia, Ontario, Canada: Chemical Interpretation and Modelling Philosophies." *Hazardous and Industrial Solid Waste Testing and Disposal, ASTM STP*, 933, 93-103.

Rix G. J. (1995). "Interpretation of Nondestructive Tests on Unknown Bridge Foundations Using Artificial Neural Networks." *Nondestructive Evaluation of Aging Structures and Dams*, 2457, 102-112.

Rizzo D. M., Lillys, T. P. and Dougherty, D. E. (1996). "Comparisons of Site Characterization Methods Using Mixed Data." *Proceeding of Uncertainty in the Geologic Environment: From Theory to Practice Conference,* Geotechnical Special Publication No 1-58, ASCE.

Robert C. (2007). *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. Springer, New York, NY.

Robinson B. and Webster, S. (2008). "Successful Testing Methods for Unknown Bridge Foundations." *Fifth Highway Geophysics - NDE Conference*, Charlotte, NC.

Santagata M., Bobet, A., Johnston, C. T. and Hwang, J. (2008). "One-Dimensional Compression Behavior of a Soil with High Organic Matter Content." *Journal of Geotechnical and Geoenvironmental Engineering*, 134(1), 1-13.

Sayed S. M., Sunna, H. and Moore, P. R. (2011). "Re-Classifying Bridges with Unknown Foundations." *5th International Conference on Scour and Erosion*, San Francisco, CA.

Schnaid F., Prietto, P. D. M. and Consoli, N. C. (2001). "Characterization of Cemented Sand in Triaxial Compression." *Journal of Geotechnical and Geoenvironmental Engineering*, 127(10), 857-868.

Sebastia M., Fernandez Olmo, I. and Irabien, A. (2003). "Neural Network Prediction of Unconfined Compressive Strength Coal Fly Ash–Cement Mixtures." *Cement and Concerte Research*, 33, 1137-1146.

Shackelford C. D. (1990). "Transit-Time Design of Earthen Barriers." *Engineering Geology*, 29(1), 79-94.

Shackelford C. D. (1993). "Contaminant Transport." *Geotechnical Practice for Waste Disposal*, Chapman & Hall, London, 33-65.

Shackelford C. D. and Daniel, D. E. (1991a). "Diffusion in Saturated Soil I: Background." *Journal of Geotechnical Engineering*, 117(3), 467-484.

Shackelford C. D. and Daniel, D. E. (1991b). "Diffusion in Saturated Soil II: Results for Compacted Clay." *Journal of Geotechnical Engineering*, 117(3), 485-506.

Shahin M. A., Jaksa, M. B. and and Maier, H. R. (2000). "Predicting the Settlement of Shallow Foundations on Cohesionless Soils Using Back-Propagation Neural Networks." Research Report No. R 167, The University of Adelaide, Australia.

Shi J., Ortigao, J. A. R. and Bai, J. (1998). "Modular Neural Networks for Predicting Settlements during Tunneling." *Journal of Geotechnical and Geoenvironmental Engineering*, 124(5), 389-395.

Sivakugan N., Eckersley, J. D. and Li, H. (1998). "Settlement Predictions Using Neural Networks." *Australian Civil Engineering Transactions, CE*, 40, 49-52.

Specht D. F. (1991). "A General Regression Neural Network." *IEEE Transactions on Neural Networks*, 2(6), 568-576.

Stegemann J. A. and Buenfeld, N. R. (2002). "Prediction of Unconfined Compressive Strength of Cement Paste with Pure Metal Compound Additions." *Cement and Concrete Research*, 32, 903-913.

Stegemann J. A. and Buenfeld, N. R. (2003). "Prediction of Unconfined Compressive Strength of Cement Paste Containing Industrial Wastes." *Waste Management*, 23(4), 321-332.

Stegemann J. A. and Buenfeld, N. R. (2004). "Mining of Existing Data for Cement-Solidified Wastes Using Neural Networks." *Journal of Environmental Engineering-Asce*, 130(5), 508-515.

Stein S. and Sedmera, K. (2006). "Risk-Based Management Guidelines for Scour at Bridges with Unknown Foundations." National Cooperative Highway Research Program (NCHRP), Project 24-25, Springfield, VA.

Tarantola A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial & Applied Mathematics (SIAM), Philadelphia, PA.

Tikhonov A. N. and Arsenin, V. Y. (1977). *Solutions of Ill-posed Problems*. Halsted Press, New York, NY.

Touma F. T. and Reese, L. C. (1972). "The Behavior of Axially Loaded Drilled Shafts in Sand." FHWA.

Tremblay H., Duchesne, J., Locat, J. and Leroueil, S. (2002). "Influence of the Nature of Organic Compounds on Fine Soil Stabilization with Cement." *Canadian Geotechnical Journal*, 39(3), 535-546.

Trenn S. (2008). "Multilayer Perceptrons: Approximation Order and Necessary Number of Hidden Units." *IEEE Transactions on Neural Networks*, 19(5), 836-844.

TxDOT (2002). "Bridge Inspection Manual." Texas Department of Transportation, Austin.

TxDOT (2006). "Geotechnical Manual." Texas Department of Transportaion, Austin.

TxDOT (2011). <http://www.dot.state.tx.us/insdtdot/orgchart/cmd/cserve/standard/b ridge- e.htm#PanForm>.

Wang J. and Zabaras, N. (2004). "A Bayesian Inference Approach to the Inverse Heat Conduction Problem." *International Journal of Heat and Mass Transfer*, 47(17), 3927-3941.

Warrick A., Biggar, J. and Nielsen, D. (1971). "Simultaneous Solute and Water Transfer for an Unsaturated Soil." *Water Resources Research*, 7(5), 1216-1225.

Weseman W. A. (1995). "Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges." Office of Engineering, Washington DC.

Wiener N. (1938). "The Homogeneous Chaos." *American Journal of Mathematics*, 60(4), 897-936.

Xiu D. and Em Karniadakis, G. (2002). "Modeling Uncertainty in Steady State Diffusion Problems via Generalized Polynomial Chaos." *Computer Methods in Applied Mechanics and Engineering*, 191(43), 4927-4948.

Yang J., Zhang, D. and Lu, Z. (2004). "Stochastic Analysis of Saturated–Unsaturated Flow in Heterogeneous Media by Combining Karhunen-Loeve Expansion and Perturbation method." *Journal of Hydrology*, 294(1), 18-38.

Young-Su K. and Byung-Tak, K. (2006). "Use of Artificial Neural Networks in the Prediction of Liquefaction Resistance of Sands." *Journal of Geotechnical and Geoenvironmental Engineering*, 132(11), 1502-1504.

Yudong C. (1995). "Soil Classification by Neural Network." *Advances in Engineering Software*, 22(2), 95-97.

Zhu J.-H., Zaman, M. M. and Anderson, S. A. (1998a). "Modeling of Soil Behavior with a Recurrent Neural Network." *Canadian Geotechnical Journal*, 35(5), 858-872.

Zhu J.-H., Zaman, M. M. and Anderson, S. A. (1998b). "Modelling of Shearing Behaviour of a Residual Soil with Recurrent Neural Network." *International Journal for Numerical and Analytical Methods in Geomechanics*, 22(8), 671-687.

Zounemat-Kermani M., Beheshti, A. A., Ataie-Ashtiani, B. and Sabbagh-Yazdi, S. R. (2009). "Estimation of Current-Induced Scour Depth around Pile Groups Using Neural Networks and Adaptive Neuro-Fuzzy Inference System." *Applied Soft Computing*, 9(2), 746-755.