SPATIAL VARIABILITY MODELS AND PREDICTION ANALYSIS OF SOIL PROPERTIES USING GEOSTATISTICS

by

TEJO VIKASH BHEEMASETTI

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2014

Acknowledgements

This thesis would not have been possible without the help, support, and patience of my principal supervisor, Prof. Anand J. Puppala, not to mention his advice and unsurpassed knowledge in geotechnical engineering. His expertise in in-situ soil testing and his never-ending passion for this field were the driving factors in this research. It is highly impossible for me to thank him, in a few words or sentences, for his constant support and motivation.

The good advice, support and friendship of my second supervisor, Dr. Xinbao Yu, has been invaluable on both an academic and a personal level, for which I am extremely grateful. I would also like to thank Dr. Laureano Hoyos for the exceptional way that he explains concepts in his courses and for his willingness to serve on my committee. I also thank Dr. Chien-Pai Han, Dr. Shih-Ho Chao, and Dr. Mostafa Ghandehari for their willingness to serve on my examination committee. My special thanks to Dr. Guojun Cai and Dr. Bhaskar Chittoori for providing all the necessary help for this research.

I would like to extend my thanks to all my fellow students, Pinit Ruttanaporamakul (Tom) , Raju Acharya, Minh Le, Ranjan Rout, Naga Talluri, Ujwal Patil, Sadikshya Poudel, Spoorthi Reballi, Rathna Mothkuri, Justin Thomey, Hai-Feng Zou, Alejandro Pino, Asheesh Pradhan, Aritra Banerjee, Santiago Caballero, Ahmed Gaily, Nguyen Hai, Rodolfo Guerra, Minh Tran.

I would like to thank my friends Aravind Pedarla, Karthik Bora, Krishna Yellapu, Krishna Mohan, Sunil Kumar, Chandra Sekhar who are with me for the last eleven years and for being true friends. Special thanks to Aravind for encouraging me to come to USA and laying the first stone for pursuing a PhD degree.

My greatest thanks to Guru and Gauranga for providing their transcendental blessings. I also want to thank the Sarvajaya Madhava Das family and all the devotees of the Radha Kalachand ji temple for continuously enlightening me about the real purpose of my life. I am also thankful to my father, mother, and brother, Srinivas, Rathna and Srujan Bheemasetti for their unconditional love throughout my entire life.

Last, but not least, I would like to dedicate this thesis to my wife, Sravani, and my son Siddhan. A very special thanks to my wife, Sravani, for her unwavering support and encouragement and for all the sacrifices she made during my education. Without her, none of this work would have been possible.

November 14, 2014

Abstract

SPATIAL VARIABILITY MODELS AND PREDICTION ANALYSIS OF SOIL PROPERTIES USING GEOSTATISTICS

Tejo Vikash Bheemasetti, PhD

The University of Texas at Arlington, 2014

Supervising Professors: Anand J. Puppala and Xinbao Yu

Soils are composed of solid, water and air phases whose characteristics are highly variable. The interactions of these phases in the soil matrix can lead to different types of topographical formations and characteristics. Due to the uncertainty and complex interactions among these phases, studies on soils have always been a challenging problem for engineers. These variations and uncertainties make it necessary for engineers to adopt new techniques and methods to analyze soil properties in order to determine or interpret their generalized behaviors and patterns. Existing research in variability analysis tends to focus on the distribution of the soil properties, reliability-based design, and simulation of random fields. Despite an increase in the probabilistic and statistical analysis, many challenges remain in incorporating the spatial variability present in the soil properties into prediction analysis. In this research study, a framework was developed using univariate statistics and randomized random variable theory for analyzing the spatially-varied soil properties. The spatial variability present in the soil properties was modeled using the geostatistical tool, Variograms. The variability models were utilized to interpret the soil properties in three different studies in geotechnical engineering, encompassing natural soils, man-made soils, and natural soils rich with

chemicals such as sulfates. This research highlights the adaptability of the framework for analyzing the soil properties varying from low-to-high variability.

Table of Contents

List of Illustrations

List of Tables

Chapter 1

Introduction

1.1 General

Geological formations of the earth's crust date back to several millions of years. The top most layer of the earth's crust, where vegetation grows, is called soil, which is derived from a Latin word called Solum (Venkatramaiah, 2006). The formations of soils can be broadly classified into two categories, which are residual soils and transported soils. Residual soils are native soils which are derived from the disintegration of rock particles. The disintegrated particles, which are transferred from one location to another through weathering agencies like air, water, glaciers and other physical events or forces, are called transported soils. The various geological processes involved in soil formation, such as physical, chemical, and biological weathering and deposition, govern the physical and chemical characteristics of the soils, such as the shape of the particle, size of the particle, and mineralogical composition of the particle (Hyunki, 2005). These physiochemical characteristics govern the engineering properties of soils such as shear strength, compressive strength, consolidation, relative density, modulus, and permeability (Mitchell and Soga, 2005).

Soils are inherently heterogeneous in nature. Most of the geotechnical studies are based on the limited soil samplings performed in the field. Even though several extensive laboratory tests are available to evaluate the properties of soil, all are performed on limited soil samples. Therefore, uncertainty is always an issue in the geotechnical studies due to inherent spatial variability present in soils. Several statistical studies have been conducted to evaluate the uncertainty present in geotechnical studies (Spry et al., 1988; Soulie et al, 1990; Phoon et al., 1995). The statistical characteristics of a soil property, such as the mean and the variance, cannot be ascertained with precision

if the number of tests is circumscribed, nor can the distribution function associated with the given data be well established.

Existing statistical analysis in geotechnical projects makes use of univariate and bivariate analyses, which doesn't account for the spatial variability present in the soil properties. In earlier studies conducted by researchers using conventional statistics, the measured soil property values on the field were typically described by a normal distribution (Kuroda et al., 1983; Hannah et al., 2006; White et al., 2008b). However, this distribution only described the sample values measured at particular locations; it did not offer any information on which zones were likely to have high values and which zones had low values. The lack of incorporating spatial variability of the soil properties into the analysis brought inevitable uncertainties to the geotechnical designs (Einstein et al., 1982; Lacasse et al., 1996).

In this research study, geostatistics, a stochastic interpretation tool, was used for addressing and incorporating spatial variability into the geotechnical analysis. Geostatistics is the separate branch of statistics which deals with spatial analysis of the data sets. Spatial analysis or spatial statistics refers to the techniques which allow us to understand the data sets with respect to space, such as location of extreme values and the overall trend or degree of continuity (Isaaks and Srivastava, 1989). The use of geostatistics gained prominence in the 1960's with the work of French professor Georges Matheron on regionalized random variables, which originated from the work of Danie G. Krige in the 1950's (Davidovic et al., 2010).

The application of geostatistics is used mainly in the mining industry to predict the location of ore by describing the probability distribution of the existing ore locations (Krige, 1951). Due to its various successful applications, it is now widely adopted in various disciplines related to earth sciences, e.g., petroleum geology, hydrogeology,

hydrology, geography, soil sciences, forestry, oceanography, meteorology, landscape geology, geochemistry, and others (Holdaway, 2014). Its ability to deal with uncertainties and variations in material properties attracted geotechnical engineers to adopt this tool for better understanding the variability of soil (Hammah and Curran, 2006).

Numerous attempts were made from the late 1970's forward to apply geostatistics to the field of geotechnical engineering; still, geostatistics is not any part of geotechnical design procedures. However, in recent years the use of the geostatistics techniques has been emerging strongly and has been proven important, when compared with univariate statistics in the field of geotechnical engineering, in areas related to intelligent compaction of subgrades over a wider region (White et al., 2008; Vennapusa et al., 2010; White et al., 2011).

1.2 Research Objectives

The main objective of this research was to evaluate the variability of geotechnical properties in field conditions by using geostatistics and by utilizing the available spatial information. The geostatistical tool, Kriging, was used comprehensively, along with various variogram models developed to effectively deal with the high variability in geotechnical engineering projects. In order to accomplish this goal, two main objectives were formulated. These are:

1. Develop a framework for spatial variability analysis.
2. Evaluate and assess the developed framework in various geotechnical problems of interest, focusing on characterization of natural and man-made soil properties.

1.3 Thesis Organization

This thesis is comprised of seven sections: Introduction (Chapter 1), Literature review (Chapter 2), Formulation of framework (Chapter 3), Validation of spatial variability analysis framework on natural soils (Chapter 4), Validation of spatial variability analysis

framework on man-made materials (Chapter 5), Validation of spatial variability analysis framework on natural mineral deposits (Chapter 6) and Summary, conclusions and future research recommendations (Chapter 7).

Chapter 1 provides the introduction to spatial variability in soil properties, geostatistics, research scope, and thesis organization.

Chapter 2 provides a summary of the uncertainties of geotechnical parameters, spatial variability in soil properties, a summary of various statistical parameters in probability theory that account for the uncertainty, and a summary of various conventional estimation methods, followed by a summary of geostatistical estimation tools.

Chapter 3 presents the framework developed for incorporating the spatial variability in the prediction analysis. The framework was developed so that the geostatistical tools, such as variogram and kriging, in combination with univariate statistics could potentially be used in estimating the properties with minimum estimation error.

Chapters 4, 5 & 6 present the validation of the formulated framework using three geotechnical problems of interest: one that focuses on natural subsoils, another on artificially chemically-treated controlled low-strength material (CLSM) made of native clayey soil, and another on sulfate-rich natural soils.

In Chapter 4, a study using CPTU bore holes data was considered for analyzing the spatial variability in the strength parameter of soils for prediction analysis, using the formulated framework in Chapter 3.

In Chapter 5, the formulated framework was used in evaluating the uniform stiffness development in the CLSM material along the pipeline by incorporating the spatial variability into the data.

In Chapter 6, the formulated framework was used in evaluating the spatial variability of sulfate concentration levels and assessing the sulfate levels at unsampled locations.

Chapter 7 presents the summary of the research, conclusions, and recommendations for future research.

Chapter 2

Literature Review

2.1 Introduction

Uncertainty and variability in soil properties are two primary concerns that are intrinsically present in geotechnical engineering projects. The uncertainty in the soil properties remains undetermined due to constrained soil investigation (Amundaray, 1994). Various researchers in the past have implemented statistics and reliability-based methods to incorporate uncertainties of soil properties in geotechnical designs, which resulted in increased cost savings (Parsons et al., 2002). These methods enhanced the geotechnical design analysis by providing better understanding the acquired soil data (Lacasse et al., 1996). However, the soil properties vary from space to space and time to time.

The soil properties determined through laboratory and field tests represent only a sample set of data from which inferences are drawn on the population data or whole project area. Therefore, the statistical characterization of the spatial variability is extremely important and should be considered in geotechnical practice. In this chapter, a detailed literature review on the uncertainties and variabilities associated with soil properties, univariate statistics that are used to describe soil properties, spatial variability in soil properties, and a comprehensive summary of estimation methods that incorporate spatial variability are presented.

2.2 Uncertainties in Geotechnical Engineering

Uncertainties are present in every single extend that is associated with geotechnical engineering projects. Various past researchers have classified the uncertainties in geotechnical properties into three categories (Vanmarcke, 1977; Baecher, 1982; Tang, 1984; Baecher et al., 2003; Kulhawy et al., 1992; Phoon et al.,

1999). These are: inherent variability, in-situ measurement and transformation model. Details of each them are described below.

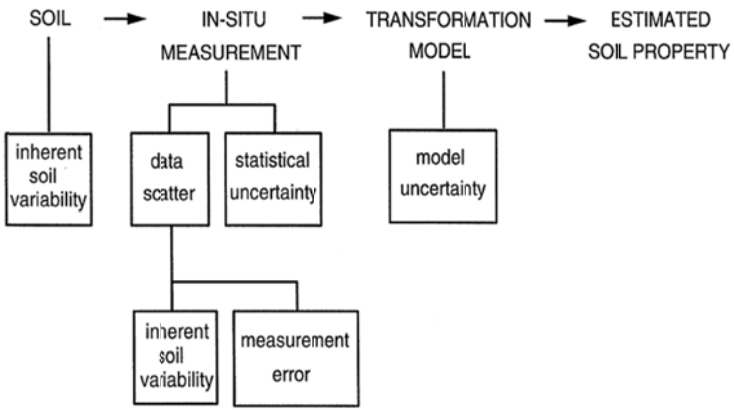


Figure 2-1 Uncertainty associated with soil property (Kulhawy, 1992)

2.2.1 Inherent Variability

The variability that is associated within the material is referred to as inherent variability. In soils, this variability could have evolved due to the influence of various geological processes, past stress history, and/or variations in moisture content (Phoon and Kulhawy, 1999). These physical processes influence the mineralogical composition of the particle and physiochemical characteristic of the particle. Different approaches were used by the researchers for modeling the inherent variability present in soil properties (Vanmarcke 1977; Baecher 1985; Spry et al. 1988; Filippas et al. 1988). Comprehensive studies were conducted on various soil properties by Phoon and Kulhawy (1999), where the inherent soil variability was modeled as a random field, using the scale of fluctuation and coefficient of variation as descriptive tools. Tables 2-1, 2-2 and 2-3 summarize the inherent variability present in strength properties, index properties, and field measurements for different types of soils and field tests (Phoon et al. 1995).

Table 2-1 Summary of inherent variability of strength properties (Phoon et al. 1995)

Property	Soil type	No. of data groups	No. of tests per group		Property value		Property COV (%)	
			Range	Mean	Range	Mean	Range	Mean
S_u (UC) (kN/m^2)	Fine Grained	38	2-538	101	6-412	100	6-56	33
S_u (UU) (kN/m^2)	Clay, silt	13	14-82	33	15-363	276	11-49	22
S_u (CIUC) (kN/m^2)	Clay	10	12-86	47	130-713	405	18-42	32
$\overline{\emptyset}$ (°)	Sand	7	29-136	62	35-41	37.6	5-11	9
$\overline{\emptyset}$ (°)	Clay, silt	12	5-51	16	9-33	15.3	0-50	21
$\overline{\emptyset}$ (°)	Clay, silt	9	-	-	17-41	33.0	4-12	9
$\tan\overline{\emptyset}\ (TC)$	Clay, silt	4	-	-	0.24-0.69	0.50	6-46	20
$\tan\overline{\emptyset}\ (DS)$	Clay, silt	3	-	-	-	0.61	6-46	20
$\tan\overline{\emptyset^b}$	Sand	13	6-111	45	0.65-0.92	0.74	5-14	9

* S_u, undrained shear strength; $\overline{\emptyset}$, effective stress friction angle; UC, Unconfined compression test; UU, unconsolidated-undrained triaxial compression test; CIUC, consolidated isotropic undrained triaxial compression test; DS, direct shear test; TC, triaxial compression test

Table 2-2 Summary of Inherent variability of index parameters (Phoon et al. 1995)

Property	Soil type	No. of data groups	No. of tests per group		Property value		Property COV (%)	
			Range	Mean	Range	Mean	Range	Mean
w_n (%)	Fine Grained	40	17-439	252	13-105	29	7-46	18
w_L (%)	Fine Grained	38	15-299	19	27-89	51	7-39	18
w_P (%)	Fine Grained	23	32-299	201	14-27	22	6-34	16
PI (%)	Fine Grained	33	15-299	12	1244	25	9-7	29
LI	Clay, Silt	2	32-118	75	-	0.094	60-8	74
γ (kN/m^3)	Fine Grained	6	5-3200	54	14-20	17.5	3-20	9
γ_d (kN/m^3)	Fine Grained	8	4-315	12	13-18	15.7	2-13	7
D_r (%)[c]	Sand	5	-	-	30-70	50	11-36	19
D_r (%)[d]	Sand	5	-	-	30-70	50	49-74	61

*w_n, natural water content; w_L, Liquid limit; w_P, plastic limit; PI, Plasticity Index; LI, Liquidity Index; γ, total unit weight; γ_d, dry unit weight; D_r, relative density; [c] Total variability for direct method of determination; [d] Total variability for indirect determination using standard penetration test (SPT) values.

Table 2-3 Summary of Inherent variability of field measurements (Phoon et al. 1995)

Test type	Property	Soil type	No. of data Groups	No. of tests per group		Property value		Property COV (%)	
				Range	Mean	Range	Mean	Range	Mean
CPTU	q_c (MPa)	Sand	57	10-2039	115	0.4-29.2	4.1	10-81	38
CPTU	q_c (MPa)	Silty Clay	12	30-53	43	0.5-2.1	1.59	5-40	27
CPTU	q_t (Mpa)	Clay	9	-	-	0.4-2.6	1.32	2-17	8
VST	S_u(kPa)	Clay	31	4-31	16	6-375	105	4-44	24
SPT	N	Sand	22	2-300	123	7-74	35	19-62	54
SPT	N	Clay, loam	2	2-61	2	7-63	32	37-57	44
DMT	A (ka)	Sand to SC	15	12-25	17	64-1335	512	20-53	33
DMT	A(kPa)	Clay	13	10-20	17	119-455	358	12-32	20
DMT	B (kPa)	Sand to SC	15	12-25	17	346-2435	1337	13-59	37
DMT	B (kPa)	Clay	13	10-20	17	502-876	690	12-38	20

Table 2.3 - *Continued*

Test type	Property	Soil type	No. of data Groups	No. of tests per group		Property value		Property COV (%)	
				Range	Mean	Range	Mean	Range	Mean
DMT	E_D(MPa)	Sand to SC	15	10-25	15	9.4-46.1	25.4	9-92	50
DMT	E_D(MPa)	Sand, silt	16	-	-	10.4-53.4	21.6	7-67	36
DMT	I_D	Sand to SC	15	10-25	15	0.8-84	2.85	16-130	53
DMT	I_D	Sand, silt	16	-	-	2.1-5.4	3.89	8-48	30
DMT	K_D	Sand to SC	15	10-25	15	0.8-8.4	2.85	16-130	53
DMT	K_D	Sand, silt	16	-	-	1.3-9.3	41	17-67	38
PMT	P_L (kPa)	Sand	4	-	17	1617-3566	228	23-50	40
PMT	P_L (kPa)	Cohesive	5	10-25	-	428-2779	104	10-32	15
PMT	E_{PMT} (MPa)	Sand	4	-	-	5.2-15.6	8.97	28-68	42

*CPTU, cone penetration test; VST, vane shear test; SPT, standard penetration test; DMT, dilatometer test; PMT, pressure meter test; q_c, CPTU tip resistance; q, corrected tip resistance; S_u, undrained shear strength; N, SPT blow count number; E_D, DMT modulus; I_D, DMT material index; K_D, DMT horizontal stress index; PMT limit stress; E_{PMT}, PMT modulus

2.2.2 In-Situ Measurement Error

The in-situ error is mainly caused during the soil testing operations. As shown in Figure 2-1, this variability is associated with data scatter, statistical uncertainty, measurement error, and inherent variability. Lumb (1971) and Orchant et al. (1988) quantified the variability in measured soil property into a summation of in situ property and measurement error. The measurement error could arise mainly because of two reasons: equipment error and procedural-operator error. Equipment error is caused by improper calibration of the equipment or damage in the testing tools, and the operator error is related to the skills of the operator. In the studies conducted by Phoon and Kulhawy (1999), it was observed that the variability due to measurement error for undrained shear strength ranges from 5 to 15%. Unlike the inherent variability, this error can be controlled by using extensive field measurement tools and skilled operators. In the studies conducted by Kulhway et al. (1992), the inherent soil variability and measurement error can be collectively summed up as a data scatter.

2.2.3 Transformation Uncertainty

Transformation uncertainty is associated with the transformation of field or laboratory measurements into soil properties by using various theoretical and empirical equations. The empirical correlations are usually developed based on testing of limited soils at a particular location. These soils are already accompanied with the measurement error variability and inherent variability. So, when the field or laboratory measurements are evaluated using these empirical equations, the uncertainty rises to a new level.

2.3 Univariate Statistics in Geotechnical Engineering

In order to apply the concepts of statistics to model the uncertainty and variability in soil properties, a basic understanding of the fundamental aspects of probability and statistics is required. In this section, a brief overview of the basic concepts is presented;

more detailed descriptions of these topics can be found in any introductory textbook on probability and statistics.

2.3.1 Random Variable

A random variable is a variable that can take any value in the data set. It is usually expressed as a real function, $Z(x)$, where x_i denotes a real number. The real number, x_i, will correspond to every outcome of an experiment; the function $\{Z(x) \leq x_i\}$ is an event for any real number x_i. The probability of an event to occur is described by $P\{Z(x) \leq x_i\}$, where the outcome takes a value between 0 and 1. A probability of 1 indicates a 100 % chance of that particular event occurring, and 0 refers to 0% chance of the event occurring.

Two types of random variables, discrete and continuous, exist in probability theory. The discrete random variable is one which takes on only a countable number of distinct values. In geotechnical engineering, the standard penetration test values (N-values) usually represent a discrete random variable. However, the continuous random variable takes on values from the continuous probability space. Friction angle and undrained shear strength of a soil layer usually represent a continuous random variable.

2.3.2 Probability Distributions

Probability distributions usually represent the probabilities of all the possible events that a random variable can take. Two types of the probability functions are generally utilized to describe the probability distribution; they are probability distribution function and cumulative distribution function.

2.3.2.1 Probability distribution function

The probability density function (pdf) is used for describing the probabilities associated with an event for a continuous random variable. The probability mass function (pmf) is used for describing the probability distribution of a discrete random variable

(Montgomery et al, 2010). The pmf and pdf are denoted as P(x) and f(x), respectively, where 'X' is the random variable itself and 'x' is the value that the random variable can take on. The probability density function f(x) of a continuous random variable is used to determine probability of areas as follows:

$$P\ (a\ <X\ <b) = \int_a^b f(x)\ dx \qquad (2.1)$$

The properties of probability density function are

(1) $f(x) \geq 0$

(2) $\int_{-\infty}^{\infty} f(x) = 1$

2.3.2.1 Cumulative distribution function

The cumulative distribution function (cdf) is another way of describing the probabilities associated with an event to occur. It is used to describe the probability distribution of a random variable that provides the probability that X is less than or equal to x.

$$F(x) = P(X\ \leq x) = \int_{-\infty}^{x} f(x)dx \qquad (2.2)$$

$$for\ -\infty\ <x\ <\ \infty$$

Since, the probabilities associated with it are cumulative, the probability of an event increases with an increase in the value of x, and finally, as x tends to ∞, $F(x) = P(X\ \leq x)$ tends to 1.

2.3.3 Elementary Statistical Parameters

The probability distribution function provides the description of probabilities associated with the random variables for different events to occur. In order to summarize the distribution of a random variable, three important statistical parameters, i.e., mean, variance and standard deviation, are necessary.

2.3.3.1 Mean

The mean or expected value of a random variable X is denoted by μ or *E*(x). It is the simplest description of a sample of data {x_1, x_2, x_3, x_4…….. x_n}, where mean $(\bar{X})$ is given as

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad (2.3)$$

Where, n is the number of observations in a sample data set

X_i is the i^{th} observation in a sample data set

From the above expression, it can be understood that the mean $\bar{X}$ provides equal weight factors to all the values. Similarly, the mean is also used to summarize a probability distribution function where μ or *E*(x) is given by

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)\ dx \qquad (2.3a)$$

Where, x is a random variable

f(x) is the distribution function of the random variable

2.3.3.2 Variance

Variance is another important parameter used to summarize the data or the distribution function. It is mainly used to describe the scatter of the data (Montgomery, 2010). The sample variance of a data set can be estimated by the expression,

$$Variance\ (X) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \qquad (2.4)$$

Where, X is the random variable

$\bar{X}$ is the sample mean

X_i is the ith observation of the random variable

n is the number of observations

The variance of a random variable X is a measure of dispersion or scatter, which is denoted as σ^2 or *V*(X).

$$\sigma^2 = V(X) = \int_{-\infty}^{\infty}(x - \mu)^2 f(x)dx \quad (2.4a)$$

Where, μ is the mean of the random variable

$f(x)$ is the probability density distribution function of the random variable

x is the random variable

The standard deviation of a random variable X is denoted by σ. The standard deviation is a measure of the deviation of the random variable from the mean or expected value. The standard deviation is obtained by taking the square root of the variance value.

2.3.4 Typical Probability Distributions

Most of the conventional statistical tests were developed based upon the normal distribution having a mean μ and variance σ^2. However, to model the distribution of geotechnical properties, the normal distribution may not be appropriate in all the cases. In studies conducted by Amundaray (1987), Harr (1987), Ang and Tang (1975), Hahn and Shapiro (1967) various distributions that can be used for modelling different geotechnical properties were presented. In this section, various probability distribution functions will be briefly discussed, along with the estimates of mean and variance. These estimates will provide the values that should be incorporated into geotechnical analysis.

2.3.4.1 Uniform Distribution

Uniform distribution refers to the probability distributions, where the probability of any event at a certain interval is most likely same. The uniform distribution function is expressed using the equation below

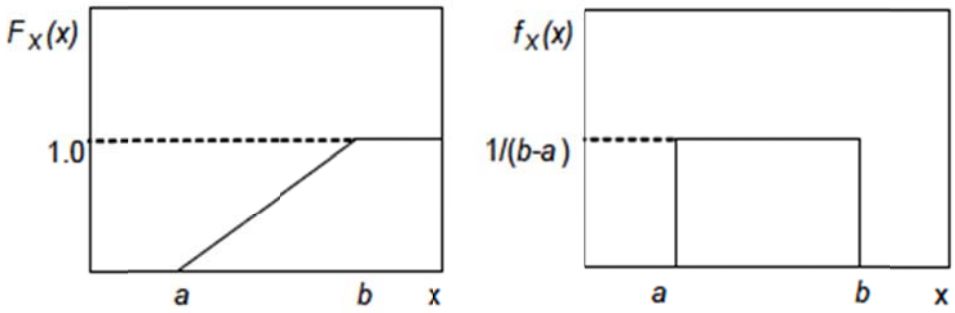
$$f(X) = \frac{1}{b-a}; \quad a \leq X \leq b \quad (2.5)$$

Where, a and b are real constants with a < b, and f (x) is equal to zero for X < a and X > b.

The probability density function and cumulative distribution function of a uniformly distributed random variable are shown in Figure 2-2. If a random variable is distributed uniformly, then the mean and variance of the uniform distribution is given using the expressions below

$$\bar{X} = \frac{(b-a)}{2} \tag{2.5a}$$

$$Variance\ (X) = \frac{(b-a)^2}{12} \tag{2.5b}$$



(a) (b)

Figure 2-2 Uniform distribution: (a) Probability density function, (b) Cumulative distribution function

If a geotechnical property is distributed uniformly, then the value of the property that has to be used in the geotechnical analysis can discovered by using the expressions for $\bar{X}$. This distribution can be used if the information about the measured data is poor and only the values in a certain region can be identified (Amundaray, 1994).

2.3.4.2 Normal distribution

Normal distribution, also called as Gaussian distribution, is the most commonly used distribution. The primary reason for this is the simplicity of defining the distribution

function with only two parameters mean (μ) and standard deviation (σ). The probability density function of a random variable is expressed as below,

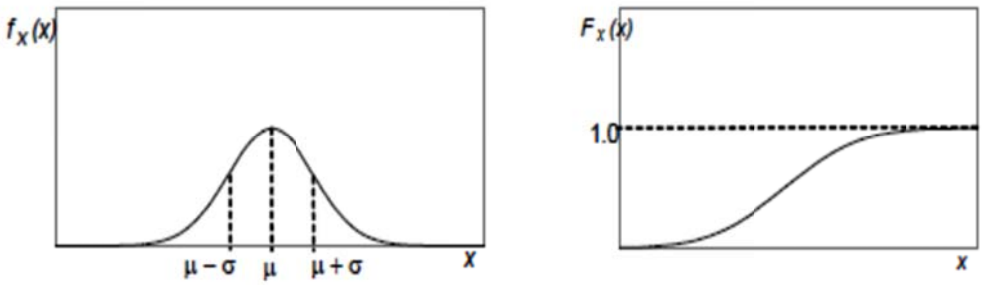
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad (2.6)$$

Where, $-\infty < x < \infty$

$-\infty < \mu < \infty$

$\sigma > 0$

The random variable X is said to be normally distributed with the parameters mean (μ) and variance (σ^2), and is denoted X ~ N(μ, σ^2). The probability density function of the cumulative distribution function for a normal distributed random variable is shown in Figure 2-3.



(a) (b)

Figure 2-3 (a) Probability density function and (b) cumulative distribution function for a normally distributed random variable

In the research studies conducted by Corotis et. al. (1975), Holtz & Krizek (1972) and Harr (1977), it was shown that most of the geotechnical properties were modeled using the normal distribution. If the number of observations was greater than 30, then the inherent variations of the geotechnical properties could be modelled using the normal distribution (Amundaray, 1994). The usage of the normal distribution enables the

researcher to learn the probability of any value by transforming the variable to a standard normal variable using the equation below,

$$z = \frac{(X - \mu)}{\sigma} \qquad (2.6a)$$

Where, z is a standard normal variable

μ is the mean of the random variable

σ is the standard deviation of the random variable

The standard normal variable has the properties mean, $\mu = 0$ and variance, $\sigma^2 = 1$. The probabilities associated with the standard normal variable are given in Figures 2-4.

z	.00	.01	.02	.03	.04	z
.0	.50000	.50399	.50798	.51197	.51595	0
.1	.53983	.54379	.54776	.55172	.55567	.1
.2	.57926	.58317	.58706	.59095	.59483	.2
.3	.61791	.62172	.62551	.62930	.63307	.3
.4	.65542	.65910	.66276	.66640	.67003	.4
.5	.69146	.69497	.69847	.70194	.70540	.5
.6	.72575	.72907	.73237	.73565	.73891	.6
.7	.75803	.76115	.76424	.76730	.77035	.7
.8	.78814	.79103	.79389	.79673	.79954	.8
.9	.81594	.81859	.82121	.82381	.82639	.9
1.0	.84134	.84375	.84613	.84849	.85083	1.0
1.1	.86433	.86650	.86864	.87076	.87285	1.1
1.2	.88493	.88686	.88877	.89065	.89251	1.2
1.3	.90320	.90490	.90658	.90824	.90988	1.3
1.4	.91924	.92073	.92219	.92364	.92506	1.4
1.5	.93319	.93448	.93574	.93699	.93822	1.5
1.6	.94520	.94630	.94738	.94845	.94950	1.6
1.7	.95543	.95637	.95728	.95818	.95907	1.7
1.8	.96407	.96485	.96562	.96637	.96711	1.8
1.9	.97128	.97193	.97257	.97320	.97381	1.9
2.0	.97725	.97778	.97831	.97882	.97932	2.0
2.1	.98214	.98257	.98300	.98341	.93882	2.1
2.2	.98610	.98645	.98679	.98713	.98745	2.2
2.3	.98928	.98956	.98983	.99010	.99036	2.3
2.4	.99180	.99202	.99224	.99245	.99266	2.4
2.5	.99379	.99396	.99413	.99430	.99446	2.5
2.6	.99534	.99547	.99560	.99573	.99585	2.6
2.7	.99653	.99664	.99674	.99683	.99693	2.7
2.8	.99744	.99752	.99760	.99767	.99774	2.8
2.9	.99813	.99819	.99825	.99831	.99836	2.9
3.0	.99865	.99869	.99874	.99878	.99882	3.0
3.1	.99903	.99906	.99910	.99913	.99916	3.1
3.2	.99931	.99934	.99936	.99938	.99940	3.2
3.3	.99952	.99953	.99955	.99957	.99958	3.3
3.4	.99966	.99968	.99969	.99970	.99971	3.4
3.5	.99977	.99978	.99978	.99979	.99980	3.5
3.6	.99984	.99985	.99985	.99986	.99986	3.6
3.7	.99989	.99990	.99990	.99990	.99991	3.7
3.8	.99993	.99993	.99993	.99994	.99994	3.8
3.9	.99995	.99995	.99996	.99996	.99996	3.9

Figure 2-4 Cumulative distribution function of the standard normal variable

(Montgomery et al. 2010)

z	.05	.06	.07	.08	.09	z
.0	.51994	.52392	.52790	.53188	.53586	.0
.1	.55962	.56356	.56749	.57142	.57534	.1
.2	.59871	.60257	.60642	.61026	.61409	.2
.3	.63683	.64058	.64431	.64803	.65173	.3
.4	.67364	.67724	.68082	.68438	.68793	.4
.5	.70884	.71226	.71566	.71904	.72240	.5
.6	.74215	.74537	.74857	.75175	.75490	.6
.7	.77337	.77637	.77935	.78230	.78523	.7
.8	.80234	.80510	.80785	.81057	.81327	.8
.9	.82894	.83147	.83397	.83646	.83891	.9
1.0	.85314	.85543	.85769	.85993	.86214	1.0
1.1	.87493	.87697	.87900	.88100	.88297	1.1
1.2	.89435	.89616	.89796	.89973	.90147	1.2
1.3	.91149	.91308	.91465	.91621	.91773	1.3
1.4	.92647	.92785	.92922	.93056	.93189	1.4
1.5	.93943	.90462	.94179	.94295	.94408	1.5
1.6	.95053	.95154	.95254	.95352	.95448	1.6
1.7	.95994	.96080	.96164	.96246	.96327	1.7
1.8	.96784	.96856	.96926	.96995	.97062	1.8
1.9	.97441	.97500	.97558	.97615	.97670	1.9
2.0	.97982	.98030	.98077	.98124	.98169	2.0
2.1	.98422	.98461	.98500	.98537	.98574	2.1
2.2	.98778	.98809	.98840	.98870	.98899	2.2
2.3	.99061	.99086	.99111	.99134	.99158	2.3
2.4	.99286	.99305	.99324	.99343	.99361	2.4
2.5	.99461	.99477	.99492	.99506	.99520	2.5
2.6	.99598	.99609	.99621	.99632	.99643	2.6
2.7	.99702	.99711	.99720	.99728	.99736	2.7
2.8	.99781	.99788	.99795	.99801	.99807	2.8
2.9	.99841	.99846	.99851	.99856	.99861	2.9
3.0	.99886	.99889	.99893	.99897	.99900	3.0
3.1	.99918	.99921	.99924	.99926	.99929	3.1
3.2	.99942	.99944	.99946	.99948	.99950	3.2
3.3	.99960	.99961	.99962	.99964	.99965	3.3
3.4	.99972	.99973	.99974	.99975	.99976	3.4
3.5	.99981	.99981	.99982	.99983	.99983	3.5
3.6	.99987	.99987	.99988	.99988	.99989	3.6
3.7	.99991	.99992	.99992	.99992	.99992	3.7
3.8	.99994	.99994	.99995	.99995	.99995	3.8
3.9	.99996	.99996	.99996	.99997	.99997	3.9

Figure 2-5 Cumulative distribution function of the standard normal variable

(Montgomery et al. 2010)

2.3.4.3 Log-Normal Distribution

The log-normal distribution is another way of representing the normal distribution of a random variable. The name follows by transforming the random variable (X) using natural logarithms (Ln). The probability density function for the log-normal distribution is expressed below,

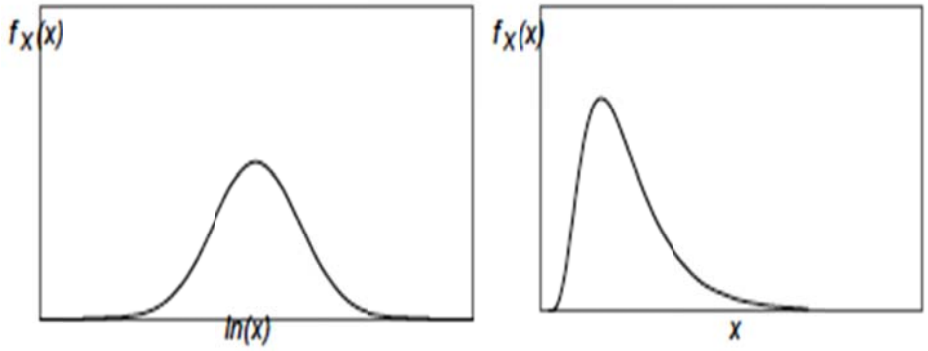
$$f(x) = \frac{1}{x\omega\sqrt{2\pi}}\ exp^{\left[-\frac{(ln(x)-\theta)^2}{2\omega^2}\right]} \qquad (2.7)$$

Where, $0 < x < \infty$

ω is the standard deviation of the random variable

θ is the mean of the random variable

The probability density function and cumulative distribution function of a random variable that is log-normally distributed are shown in Figures 2-5.



(a) (b)

Figure 2-6 (a) Probability density function and (b) cumulative distribution function for a log-normal distributed random variable

This distribution is particularly useful for large variations in the data, where the lower bound is close to zero and upper bound is infinite, as shown in Figure 2-6 (b). The geotechnical properties, like hydraulic conductivity, are well described using this

distribution (Amundaray, 1994). If the soil property of interest is distributed using log-normal distribution, then the value that needs to be used in further analysis is found by using the expression below,

$$E(X) = e^{\theta + \frac{\omega^2}{2}} \tag{2.7a}$$

2.3.4.4 Exponential Distribution

The exponential distribution is primarily used to model the number of flaws in a considered physical system. In geotechnical engineering, this distribution can be useful to model the discontinuities in the rock (Amundaray, 1994). If the parameter λ represents the number of discontinuities in a rock of certain length, then the probability density function for an exponential distribution can be expressed as below,

$$f(x) = \lambda e^{-\lambda x} \tag{2.8}$$

Where, $0 \leq x < \infty$

The probability density function and cumulative distribution function of a random variable that is exponentially modeled are shown in Figure 2-6.

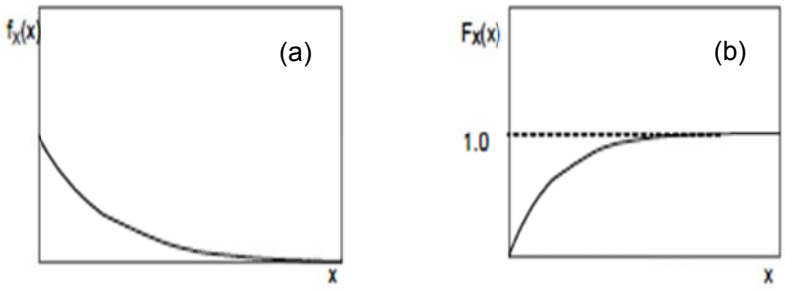


Figure 2-7 (a) Probability density function and (b) cumulative distribution function for an exponentially distributed random variable

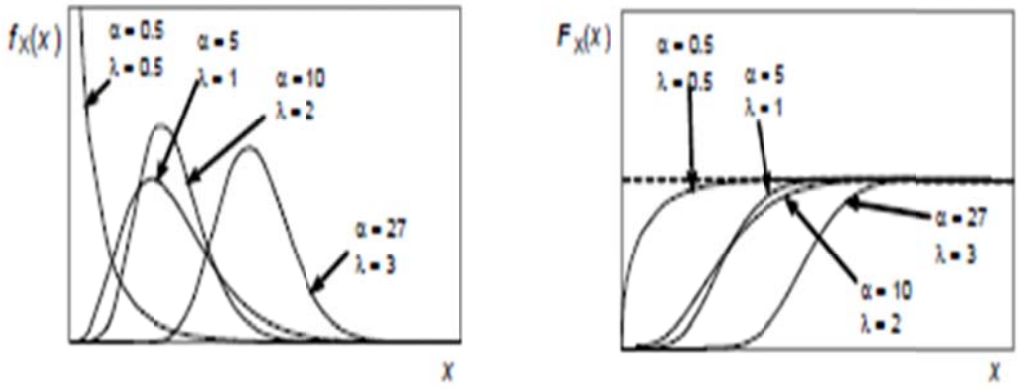
2.3.4.5 Gamma distribution

The gamma distribution is used to model a variety of random experiments (Montgomery, 2010). The probability density function of a gamma distribution function is expressed as below,

$$f(x) = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\int_0^\infty x^{r-1} e^{-x} dx} \quad (2.9)$$

Where, x > 0; λ > 0; and r > 0

The probability density function and cumulative distribution function of a random variable for gamma distribution are as shown in Figure 2-7.



(a) (b)

Figure 2-8 (a) Probability density function and (b) Cumulative distribution function for a random variable that has a gamma distribution

2.4 Characterization of Spatial Variability

Spatial variability refers to the measurements of a random variable that varies across the space. In the earlier sections, the uncertainties associated with the soil properties and various distributions to model the uncertainties were discussed. However, the univariate statistics that were discussed do not reflect the spatial correlation present in the soil properties. With the consideration of spatial correlation in the properties, more

realistic estimates and conclusions can be drawn. In this section, different approaches to characterizing the spatial variability present in the soil properties are presented.

2.4.1 Trend or Drift

Trend or drift refers to constant variation of the soil property along the subsurface profile. Computationally, the trend present in the random variable can be calculated as the weighted average of all the points within the neighborhood around the point (Davis, 1986). In the research study conducted by Cuba et al (2011), the expression for semi-variogram (tool in geostatistics for modeling spatial variability) had three components: (1) mean trend, (2) variance trend and (3) the stationary component. The mean and variance trend represents locally varying mean and variance values. The mean trend is obtained by calculating the expected value of the random variable in the original scale. Kanevski et al. (1996) utilized neural networks to obtain the mean trend present in the random variable. According to the studies conducted by Phoon et al. (1999), the soil property ($\xi(z)$) at any depth (z) can be modeled as the summation of deterministic trend function, deviation from trend, and measurement error, as shown in Figure 2-8.

$$\xi(z) = t(z) + w(z) + e(z) \tag{2.10}$$

Where, z is the depth along the subsurface;

$\xi(z)$ is the in-situ soil property;

$t(z)$ is deterministic trend function;

$w(z)$ is the deviation from trend;

$e(z)$ is the measurement error.

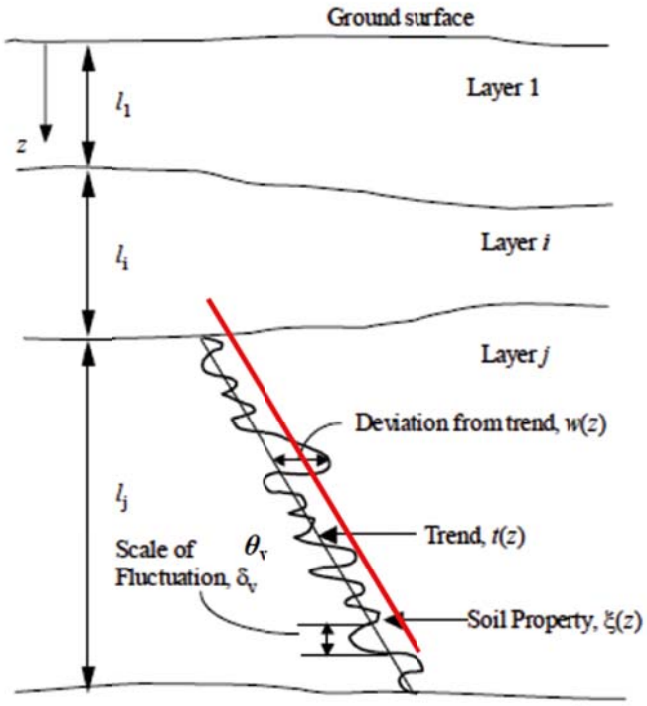


Figure 2-9 Spatial variation of soil properties (Davidovic et al., 2010)

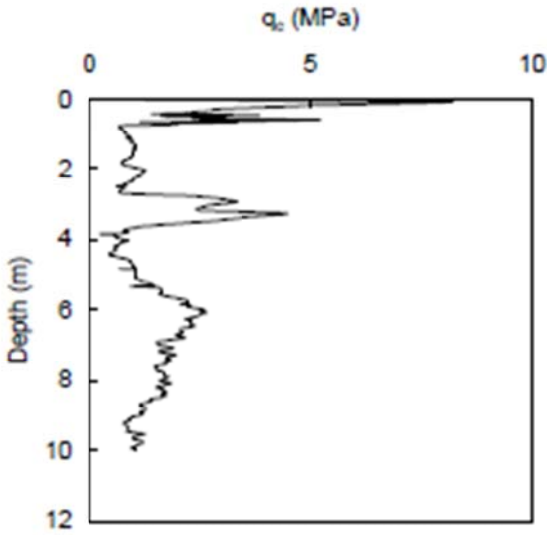
2.4.2 Scale of fluctuation

The scale of fluctuation refers to the fluctuations in the random variable values. VanMarcke (1984) introduced the scale of fluctuation (θ) as a descriptive tool for characterizing the variability present in the random field.

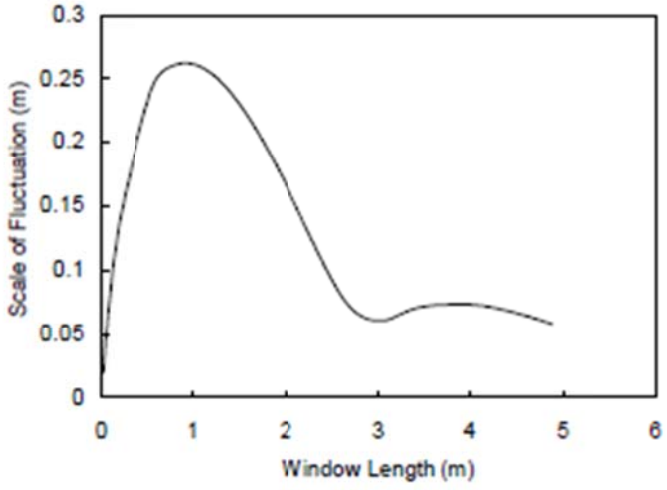
$$\theta = \int_{-\infty}^{+\infty} \rho(\tau) d(\tau) = 2 \int_{0}^{\infty} \rho(\tau) d(\tau) \qquad (2.11)$$

In Figure 2-8, the scale of fluctuation, denoted, as δ_v, refers to gentle variation in the soil property about the trend line (t (z)). The scale of fluctuation estimates the maximum distance over which a random variable shows strong correlation. In layman terms, if two points in a random field are separated by a distance greater than the scale of fluctuation implies, the two points are largely unrelated; and if two points are separated

by a distance smaller than the scale of fluctuation, a strong correlation is indicated (Fenton, 2002). In the research study conducted by Jones et al (2002), an example calculation was provided for estimating the scale of fluctuation for cone tip resistance (q_c) along the sub surface profile, as shown in Figure 2-9.



(a)



(b)

Figure 2-10 (a) Cone tip resistance along the depth, (b) Scale of fluctuation for corresponding window length

From the Figure 2-9 (b), it can be understood that for a window length of 1 meter, the scale of fluctuation reaches a peak value of 0.26m. So, at any distance greater than 0.26m, the tip resistance is not correlated. The detailed steps for calculation of the scale of fluctuation were explained in the technical report of Jones et al (2002). Various approaches, such as regression technique, spectral analysis, moment estimation and maximum likelihood estimation have been suggested by past researchers to discover the scale of fluctuation. DeGroot (1996) conducted and summarized extensive studies to obtain a scale of fluctuation for various in-situ tests, as shown in Table 2-4.

Table 2-4 Scale of fluctuation of soil properties in various in situ tests (Degroot, 1996)

Soil Property	Soil	Direction	Scale of Fluctuation (m)	Reference
SPT N Value	Dune Sand	Horizontal	40.00	Hilldale-Cunningham (1971)
	Alluvial Sand	Horizontal	33.40	DeGroot (1996)
DMT P_0	Varved Clay	Vertical	2.28	DeGroot (1996)
CPTU Cone Resistance	Sea Clay	Horizontal	60.00	Hoeg and Tang (1976); Tang (1979)
	Silty Clay	Horizontal	10.00 ~ 24.00	Lacasse and de Lamballerie (1995)
	Copper Tailings	Vertical	1.00	Baecher (1987)
	Sensitive Clay	Vertical	4.00 for q_c, f_s, and u_2	Chiasson et al. (1995)
	Silty Clay	Vertical	2.00	Lacasse and de Lamballerie (1995)
	Clay	Vertical	2.00	Vanmarcke (1977)
	Mexico Clay	Vertical	2.00	Alonzo and Krizek (1975)
	Clean Sand	Vertical	6.00	Alonzo and Krizek (1975)
	Clean Sand	Vertical	3.20	Kulatilake and Ghosh (1988)
	North Sea Sand	Horizontal	28.00 ~ 76.00	Keaveny et al. (1989)

Table 2.4 - *Continued*

Vane Shear Test Undrained Shear Strength	Clay	Vertical	2.00 ~ 6.00	Asaoka and A-Grivas (1982)
	Sensitive Clay	Vertical	2.00	Baecher (1987)
	Sensitive Clay	Horizontal	46.00	DeGroot and Baecher (1993)
Laboratory Undrained Shear Strength	Chicago Clay	Vertical	1.00 (Unconfined Compression Test)	Wu 91974)
	offshore Sites	Vertical	0.60 ~ 7.20 (Triaxial and DSS)	Keaveny et al. (1989)
Hydraulic Conductivity	Compacted Clay	Horizontal	1.00 ~ 4.00	Benson (1991)

2.4.3 Coefficient of Variation

The standard deviation gives an absolute measure of the dispersion in the data; whereas, the coefficient of variation is a relative measure of the spread or variability in the data. The coefficient of variation is calculated by expressing the standard deviation as a percentage of the mean.

$$COV = \frac{Standard\ devitaion\ (\sigma)}{Mean\ (\mu)} \tag{2.12}$$

Where, $Standard\ deviation\ (\sigma) = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$

$$Mean\ (\mu) = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Where, X is the random variable; $\bar{X}$ is the sample mean; X_i is the ith observation of the random variable, and n is the number of observations

Due to its simplicity, coefficient of variation is the most widely used statistical tool for characterizing the variability in civil engineering projects. The American Institute of

Concrete (1965) has suggested the coefficient of variation as a descriptive tool to categorize the degree of variability present in concrete material.

Coefficient of Variation < 10% - Excellent

10% < Coefficient of Variation < 15% - Good

15% < Coefficient of Variation < 20% - Satisfactory

Coefficient of Variation > 20% - Poor

The different stages of soil formation inherently produce high variability when compared to the brittle materials such as concrete. Therefore, in geotechnical engineering, if the coefficient of variation is less than 20%, the data can be regarded as low degree of variability. Rethati (1988) studied the effects of a number of samples on the coefficient of the variation of unconfined compressive strength. Figure 2-10 presents the results of this study, where UCS strength and coefficient of variation values were plotted against the number of samples. It can be clearly observed that there is a decrease in the unconfined compressive strength value, along with coefficient of variation until the number of samples is 30. The variation after 30 samples becomes more stable, following a constant trend. Amundaray (1994) recommended that at least 30 samples should be tested to obtain a realistic coefficient of variation of soil properties.

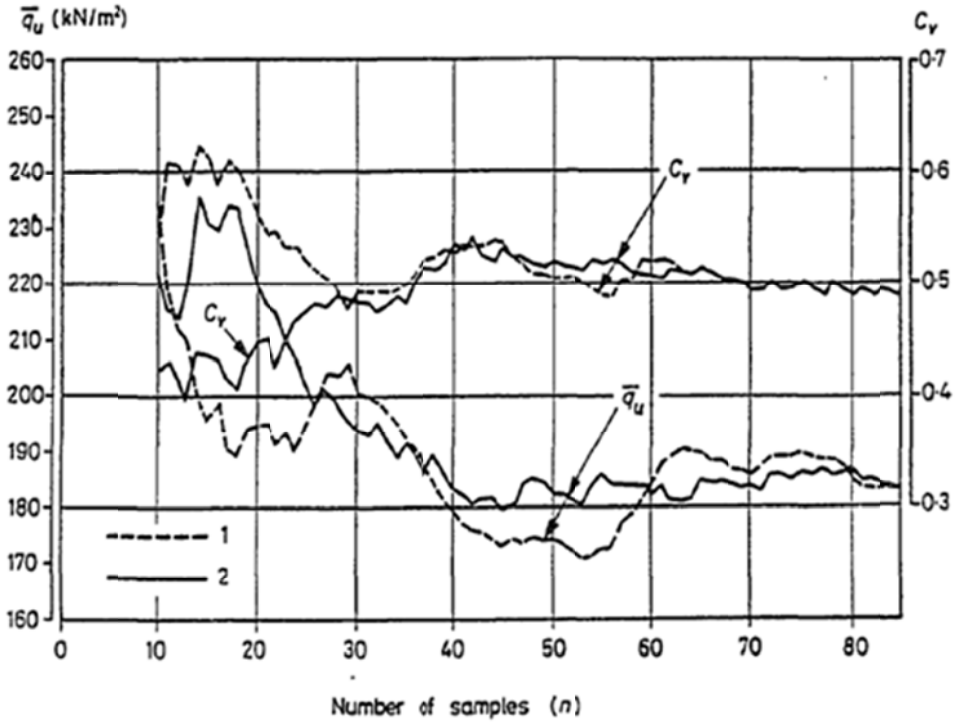


Figure 2-11 Coefficient of variation for unconfined compressive strength values

Several researchers in the past have studied the coefficient of variation of various soil properties obtained from field and laboratory measurements. Table 2-5 below summarizes the coefficient of variation values of different soil properties.

Table 2-5 Summary of coefficient of variation of different soil properties (Kim, 2011).

Soil Property	Soil	Direction	COV (%)	Reference
SPT N Value	Clay and Sand	Vertical	25.00 ~ 50.00	Phoon and Kulhawy (1996)
CPTU Cone Resistance	River Sand	Vertical	25.00 ~ 43.00	Reyna and Chameau (1991)
	River Sand	Vertical	25.00 ~ 43.00	Reyna and Chameau (1991)
	Clay	Vertical	20.00 ~ 40.00	Phoon and Kulhawy (1996)
	Clay	Vertical	20.00 ~ 40.00	Phoon and Kulhawy (1996)
	Sand	Vertical	20.00 ~ 60.00	Phoon and Kulhawy (1996)
CPTU Sleeve Friction	River Sand	Vertical	26.00 ~ 43.00	Reyna and Chameau (1991)

Table 2.5 - *Continued*

DMT A Reading	Clay	Vertical	10.00 ~ 35.00	Phoon and Kulhawy (1996)
	Sand	Vertical	20.00 ~ 50.00	Phoon and Kulhawy (1996)
DMT B Reading	Clay	Vertical	10.00 ~ 35.00	Phoon and Kulhawy (1996)
	Sand	Vertical	20.00 ~ 50.00	Phoon and Kulhawy (1996)
DMT Dilatometer Modulus	Sand	Vertical	20.00 ~ 60.00	Phoon and Kulhawy (1996)
	River Sand	Vertical	20.00 ~ 60.00	Reyna and Chameau (1991)
DMT Material Index	Sand	Vertical	20.00 ~ 60.00	Phoon and Kulhawy (1996)
	River Sand	Vertical	20.00 ~ 60.00	Reyna and Chameau (1991)
DMT Horizontal Stress Index	Sand	Vertical	20.00 ~ 60.00	Phoon and Kulhawy (1996)
	River Sand	Vertical	20.00 ~ 60.00	Reyna and Chameau (1991)
PMT Limit Pressure	Clay	Vertical	10.00 ~ 35.00	Phoon and Kulhawy (1996)
	Sand	Vertical	20.00 ~ 50.00	Phoon and Kulhawy (1996)
PMT Young's Modulus	Sand	Vertical	15.00 ~ 65.00	Phoon and Kulhawy (1996)
Vane Test Undrained Shear Test	Clay	Vertical	18.00 ~ 30.00	Asaoka and A-Gricas (1982)
	Clay	Vertical	10.00 ~ 40.00	Phoon and Kulhawy (1996)
Laboratory Relative Density	Sand	-	11.00 ~ 36.00	Haldar and Tang (1979)
Laboratory Natural Water Content	All Soil Types	-	9.00 ~ 32.00	Kulhawy et al. (1991)
Laboratory Liquid Limit	All Soil Types	-	3.00 ~ 19.00	Kulhawy et al. (1991)
Laboratory Plastic Limit	All Soil Types	-	7.00 ~ 17.00	Kulhawy et al. (1991)
Laboratory Void Ratio	All Soil Types	-	13.00 ~ 26.00	Kulhawy et al. (1991)
Laboratory Total Unit Weight	All Soil Types	-	2.00 ~ 12.00	Kulhawy et al. (1991)
Laboratory Effective Friction Angle	All Soil Types	-	6.00 ~ 21.00	Kulhawy et al. (1991)
Laboratory Compression Index	All Soil Types	-	26.00 ~ 48.00	Kulhawy et al. (1991)

2.4.4 h-scatter plots

Spatial continuity is a geostatistical tool through which the variation between two random variables can be discovered. If two data points are proximate to each other, they are more likely to have homogeneous values than two data that are far apart. H-scatter plots are used to describe the relationship between the value of one variable and the value of the same variable at nearby locations, which are separated by a distance 'h' (Isaaks and Srivastava, 1989). To illustrate further, consider a sample data set with 25 observations distributed in a grid area.

X_1	X_6	X_{11}	X_{16}	X_{21}
X_2	X_7	X_{12}	X_{17}	X_{22}
X_3	X_8	X_{13}	X_{18}	X_{23}
X_4	X_9	X_{14}	X_{19}	X_{24}
X_5	X_{10}	X_{15}	X_{20}	X_{25}

Figure 2-12 Sample data set in a grid area

In order to draw the h-scatter plot for the above case, label x-axis as X(i) and y-axis as X(i+h). If both the values that are separated by the distance 'h' are identical, they fall on the 45 degree straight line. Figure 2-12 represents an h-scatter plot for h (0, 0), where every individual is paired with the same value, thus falling on the straight line. Similarly, a separation distance of 2m gives a plot for h (0,2), where every individual is paired with other data value which is 2m apart. As the separation distance increases, the values become less similar, where the data points move away from the 45 degree straight line, as shown in Figure 2-13. The fatter the cloud of the h-scatter plot, the more dissimilarity

exists in the data values. Thus, h-scatter plot provides the spatial variation present in the random values. In order to quantify the spatial variability in the h-scatter plots, three important functions are used.

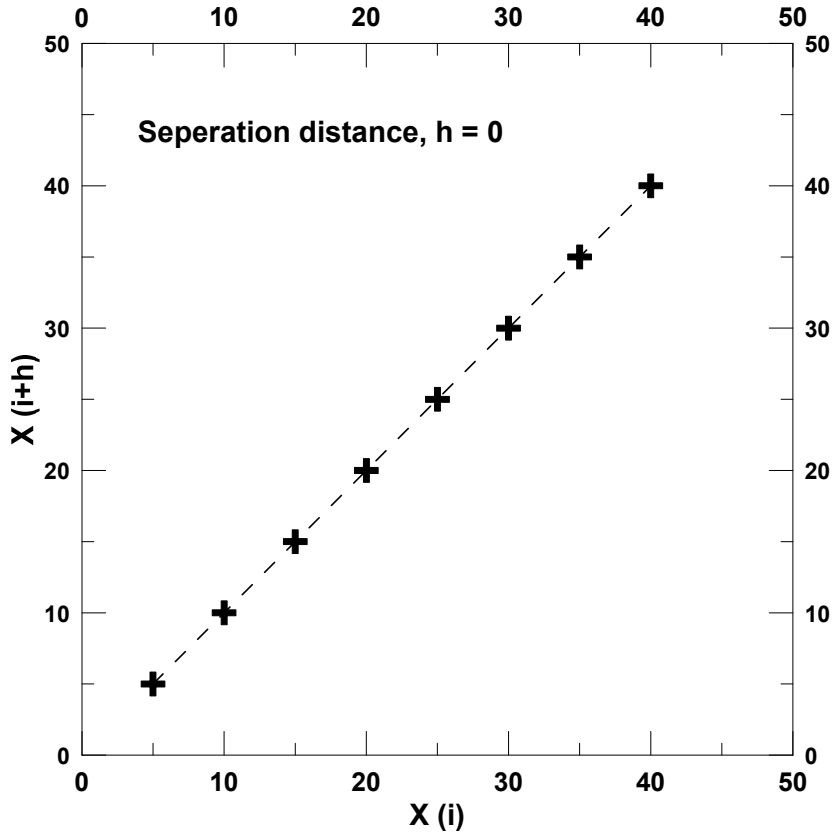


Figure 2-13 h-scatter plot for a separation distance of 0m

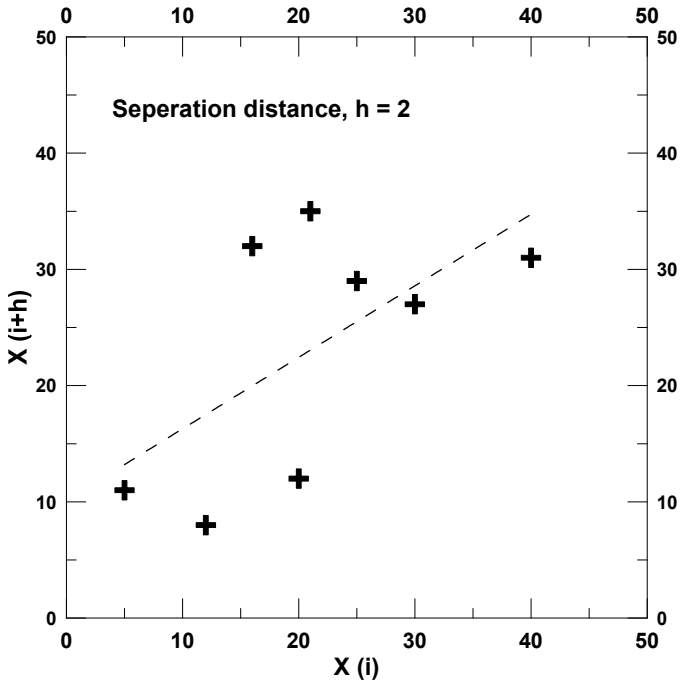


Figure 2-14 h-scatter plot for a separation distance of 2m

2.4.4.1 Covariance function

Covariance, in statistics, is used to find the similarity between two random variables. The absolute value of the covariance increases with the correlation between two variables. If the random variables are positively correlated, the covariance between them will be positive. On the contrary, if the covariance is negative, the random variables are negatively correlated.

$$Covariance\ (X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - m_x)\left(y_i - m_y\right) \qquad (2.13)$$

Where, n is total number of observations

x_i , y_i is i[th] observations of a random variables x and y

m_x, m_y is the mean of random variables x and y

The relationship between covariance of an h-scatter plot and h is called the covariance function (Isaaks and Srivastava, 1989). In geostatistics, the covariance

function is used to estimate the maximum lag distance that represents the spatial correlation in the data. The covariance is expressed as follows,

$$Cov(X_i, X_{i+h}) = \frac{[\sum X_i X_{i+h}] - \frac{1}{n-h}\sum X_i \frac{1}{n} X_{i+h}}{n-h-1} \quad (2.13a)$$

Where, h is the lag distance

X_i is the i[th] observation

X_{i+h} is the observation separated at a distance 'h' from X_i

n is the total number of observations

From the covariance function, it is evident that when the lag distance (h) is zero, the covariance function results in the variance of the random variable. As the lag distance increases, the covariance of the random variable decreases and reaches a constant value, depicting that the random variables are not correlated after that lag distance.

2.4.4.2 Correlation function

The relationship between the correlation coefficient of an h-scatter plot and separation distance (h) is called the correlation function or correlogram (Isaaks and Srivastava, 1989). The plot of correlogram provides the correlation between the two data values that are separated at a particular distance.

$$correlation\ coefficient\ (\rho) = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - m_x)(y_i - m_y)}{\sigma_x \sigma_y} \sim \frac{C(h)}{\sigma_{-h}.\sigma_{+h}} \quad (2.14)$$

Where, n is total number of observations

x_i , y_i is i[th] observations of a random variables x and y

m_x, m_y is the mean of random variables x and y

σ_x , σ_y is the standard deviation of the random variables x & y

Autocorrelation function is the relationship between the autocorrelation coefficient of a h-scatter plot and separation distance (h). It is obtained by normalizing the auto covariance with the variance of the random variable itself.

$$r_h = \frac{cov(X_i, X_{i+h})}{V(x)} \quad (2.15)$$

Where, r_h is the autocorrelation function

$cov(X_i, X_{i+h})$ is the covariance between random variable separated at h units apart

$V(x)$ is the variance of the random variable itself.

The autocorrelation function is used to find the scale of fluctuation of soil properties between two points. VanMarcke (1984) and Li et al. (1987) provided various autocorrelation function models between the scale of fluctuation (θ) and lag distance. VanMarcke (1984) and Li et al. (1987) provided various auto-correlation functions for measuring scale of fluctuation, as summarized in Table 2-6.

Table 2-6 Autocorrelation functions for measuring scale of fluctuation

Model No.	Autocorrelation Function
1	$\rho_t = 1 - \frac{\lvert\tau\rvert}{\theta}\ for\ \lvert\tau\rvert < \theta;$ $\rho_t = 0\ for\ \lvert\tau\rvert > \theta$
2	$\rho_t = e^{-2\lvert\tau\rvert/\theta}$
3	$\rho_t = e^{-\pi(\lvert\tau\rvert/\theta)^2}$
4	$\rho_t = e^{-4\lvert\tau\rvert/\theta}\left(1 + \frac{4\lvert\tau\rvert}{\theta}\right)$

The variability present in the soil properties, using lag distance and scale of fluctuation, can be characterized using the above models for autocorrelation,

2.4.4.3 Variogram

Variogram is another plausible index for characterizing the spatial variability present in the random variables (Amundaray, 1994) . The variogram or semi-variogram γ (h) is a traditional analysis tool used to describe the spatial continuity of the data in earth science application. It is defined as one-half of the average squared differences between the x and y coordinates of each pair of points in the h-scatter plot (Isaaks and Srivastava, 1989). The mathematical expression for calculating a variogram value is given as below:

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} [z(x_i + h) - z(x_i)]^2 \qquad (2.16)$$

Where, $z(x_i)$ = measurement taken at a location x_i

$z(x_i + h)$ = measurement taken at a location h distance away

n(h) = number of data pairs h units apart in the direction of the vector

h = lag distance

$\gamma(h)$ = variogram value

Using the above expression, an experimental variogram value can be obtained for each h-scatterplot, and a series of such values for different h-scatter plots gives rise to the variogram plot. Figure 2-14 shows a typical variogram, where each circle is represented by an experimental variogram value for an individual h-scatter plot.

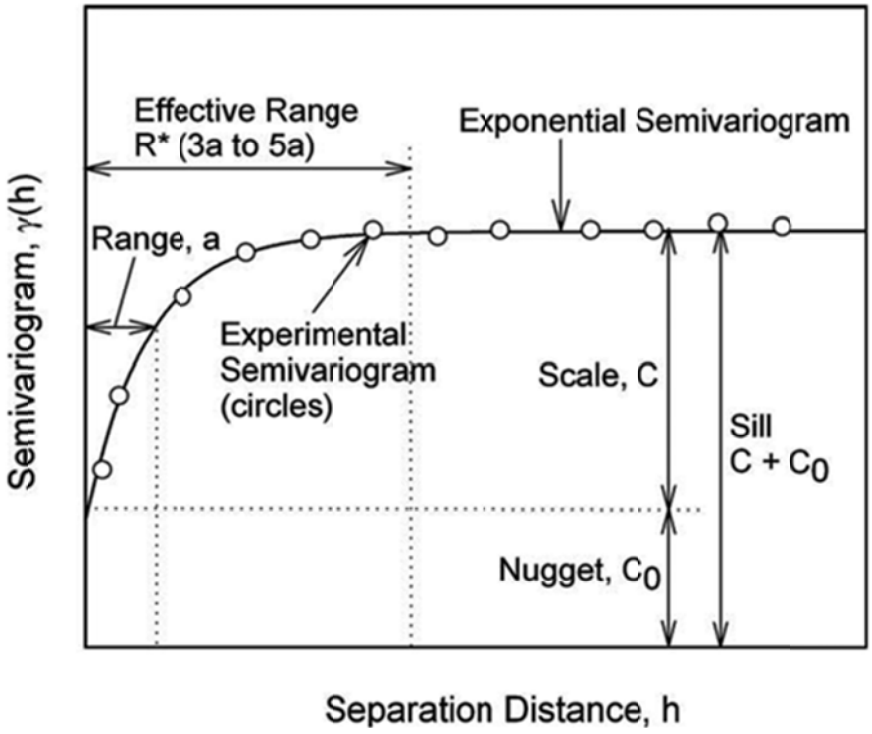


Figure 2-15 Typical Sample Variogram (Vennupusa et al., 2010)

The three main characteristics by which a variogram is often summarized are Range, Sill and Nugget (Isaaks and Srivastava, 1989).

Range (a): As the separation distance between pairs increases, the corresponding semivariogram value will also generally increase. Eventually, however, an increase in the distance no longer causes a corresponding increase in the semivariogram value and reaches a plateau. The distance at which the semivariogram reaches this plateau is called the range. The longer range value suggests greater spatial continuity (vennapusa, 2010).

Sill (C_0+C): The plateau that the semivariogram reaches at the range is called the sill. A semi-variogram (which is one-half of the variogram) generally has a sill that is approximately equal to the variance of the data (Srivastava 1996).

Nugget effect (C_0): Though the value of the semvariogram at h=0 is strictly zero, several factors, such as sampling error and very short scale variability, may cause sample values separated by extremely short distances to be quite dissimilar. This causes a discontinuity at the origin of the semivariogram and is called the nugget effect.

2.5 Simulations of Spatial Variability

In earlier sections, different statistical approaches to characterizing the uncertainty and spatial variability were discussed. Incorporating these statistical parameters into geotechnical analysis is often considered quite complex. Therefore, geotechnical engineers are sometimes required to perform conservative designs, which results in a significant increase in project cost. One alternative to this problem is to simulate the actual field conditions and study its behavior under different circumstances.

In geotechnical engineering, simulations are performed to replicate the actual field conditions. Simulations are not new in the field of geotechnical engineering; most of the slope stability analyses are performed by modeling the soil conditions in different layers. However, the layer properties are a result of the average mean value or the minimum value of the property of interest. Vanmarcke (1977) attempted to replicate actual field conditions by implementing the random field theory. In this section, various simulation techniques employed to represent the actual field conditions will be briefly discussed.

2.5.1 Random Field

Random field by name represents a field, where all the realizations z(x) are obtained from a random function Z(X). The random field theory is used in geotechnical engineering to simulate the field conditions by incorporating the spatial variability. Various statistical tools such as correlation distance, scale of fluctuation, and coefficient of variation were used to simulate a random field (Vanmarcke 1977; Fenton 1999a; Gui et

al. 2000; Elkateb et al.2003; Huang et al. 2010; Cho 2012; Zhu and Zhang 2012). The inferences obtained from the random field are used to solve the deterministic problems of interest. Fenton (1993) suggested that based on the first order moments, the random fields can be classified as Gaussian field and non-Gaussian field. The Non-Gaussian random fields can be created by using the nonlinear transformation of the data. Provided below are brief descriptions of various random field generators, where only the first two moments of the random fields, mean and covariance, are constant.

2.5.1.1 Moving average (MA) method

Moving average is the simplest method to generate a random field. The moving average technique constructs the random field Z(x) as a weighted average of a white noise process (Fenton, 1993)

$$Z(x) = \int_{-\infty}^{\infty} f(\xi - x) dW(\xi) \quad (2.17)$$

Where, $dW(\xi)$ is a zero mean incremental white-noise process with variance $d\xi$

f is a weighting function

In the studies conducted by Mignolet and Spanos (1992), it was mentioned that the accuracy of the MA method depends upon the pace of the program. The moving average technique is very time consuming because of the level of difficulty in finding the weighting function f.

2.5.1.2 Fourier transform method

The Fourier transform method is used to generate continuous random fields. This is performed based on the spectral representation of the mean square data for homogeneous random fields. Yaglom (1962) expressed continuous distribution function for generation of random field.

$$Z(x) = \int_{-\infty}^{\infty} e^{ix.\omega} W(d\omega) \quad (2.18)$$

Where, $W(d\omega)$ is an interval white-noise process with mean zero and variance $S(\omega)d\omega$

$S(\omega)$ is the spectral density function

The discrete Fourier transform is most generally used as the summation at a particular point, where in the actual Fourier transform method, the n-dimensional integration becomes n-dimensional sum (Gordon, 1993).

2.5.1.3 Fast Fourier transform method

The fast Fourier transform method is much more advanced and efficient than the actual Fourier transform method. Cooley and Tukey (1965) suggested this method as an alternative to the discrete Fourier transform method, where N^2 operations can be reduced to N ($\log_2 N$). A random field $Z(x)$can be expressed as (Gordon, 1993):

$$Z(x_j) = \lim_{k \to \infty} \sum_{k=-k}^{k} \left[A(\Delta\omega_k)cos(x_j\omega_k) + B(\Delta\omega_k)sin(x_j\omega_k)\right] \qquad (2.19)$$

Where, $\omega_k = k\pi/K$

$\Delta\omega_k$is an interval of length π/K centered at ω_k

$$A_k = \frac{1}{K}\sum_{j=0}^{k=1} Z_j \cos 2\pi\frac{jk}{K} = A_{k=k}$$

$$B_k = \frac{1}{K}\sum_{j=0}^{k=1} Z_j \sin 2\pi\frac{jk}{K} = -B_{k=k}$$

The detailed procedure and mathematical formulation of the above equation are detailed in the studies conducted by Gordon (Gordon, 1993).

2.5.1.4 Decomposition matrix method

Decomposition matrix is another often-used technique for generating homogeneous random fields. In this technique, the covariant matrix is decomposed into a lower triangular matrix and an upper triangular matrix. The decomposition matrix is expressed as:

$$K_m = L_m . U_m \quad (2.20)$$

Where, K_m is a covariance matrix

L_m is a lower triangular matrix

U_m is a upper triangular matrix

If K_m, the covariant matrix is positive definite, then the mean zero-discrete process $Z_i = Z(x_i)$ can be produced by:

$$Z = L_m V$$

Where, L_m is a lower triangular matrix and V is the unit vector

The decomposition matrix is only used for small fields, as the round-off error and time for programming increase with an increase in matrix size (Fenton, 1993).

2.5.1.5 Turning band method (TBM)

Matheron (1973) developed the turning band method for generation of random field in two or more dimensional spaces by using the existing one-dimensional techniques, such as fast Fourier transformation and decomposition matrix. The turning band method generates more efficient and fast random fields in multi-dimensional processes compared to the decomposition matrix and fast Fourier transform method (Fenton, 1994). Gordon (1993) compared the process time for generating the random field using three methods, as summarized in the table below.

Table 2-7 Comparison of processing time for generating random fields (Gordon, 1993)

Dimension	Fast Fourier Transform	Local Average Subdivision	Turning Band Method	
			16 lines	64 lines
1 Dimension	1.0	0.70	-	-
2 Dimension	1.0	0.55	0.64	2.6

Gordon (1993) outlined a detailed procedure for generating a random field, using a turning band method with reference to the below Figure 2-15. The random field-generating equation is expressed as:

$$Z(x_k) = 1/\sqrt{L} \sum_{i=1}^{L} Z_i(x_k * U_i) \quad (2.21)$$

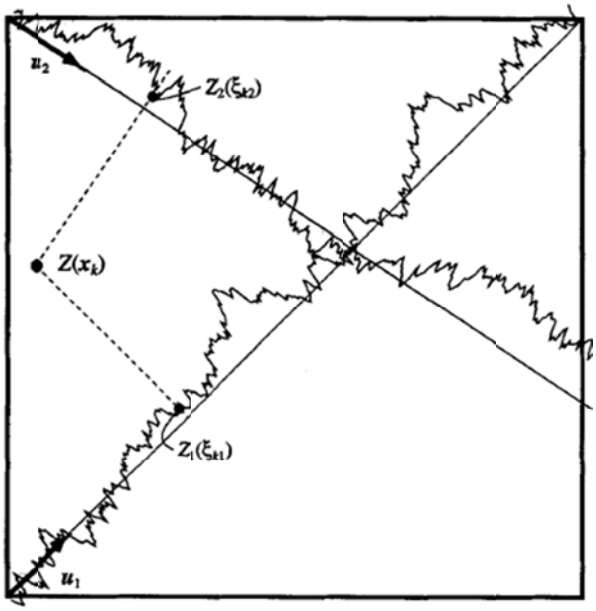


Figure 2-16 Random field generation using turning band theory (Gordon, 1993)

2.6 Predictions Using Conventional Estimation Methods

The spatial variability present in the random variables can be characterized using any of the techniques discussed in the earlier section. In any analysis, the distributions and statistical parameters of a property can be used to supplement the prediction analysis. For instance, consider that a cohesive strength value of a subsurface profile is fluctuating with a coefficient of variation of 0.32. In order to find the bearing capacity of a layer, the primary question is what value of cohesive strength should be considered. Always, as a worst case scenario, the minimum value of strength can be taken; however, that leads to uneconomical costs and unrealistic field conditions. Using probabilistic analysis, the raw estimates can be determined by evaluating the distribution of the

cohesive strength values and incorporating the parameters E(x) into further analysis. In this section, various conventional spatial estimation methods that have not been explored in geotechnical engineering will be discussed.

2.6.1 Polygonal Estimation

Polygonal estimation is a spatial estimation technique used to predict the values at unsampled point locations. This was developed from Voronoi diagrams and was later introduced to geophysics by Alfred H. Thiessen. One of the primary applications of this technique is to gauge the areal rainfall measurements (Thiessen, 1911). Later, this method was introduced for estimating the point measurements for various applications (Tabios and Salas, 1985; Dirks et al., 1998). In order to obtain the predictions, the boundaries of a polygon are created at a halfway distance between the known values. The polygonal area, as shown in Figure 2-16, represents the area of influence around the known value.

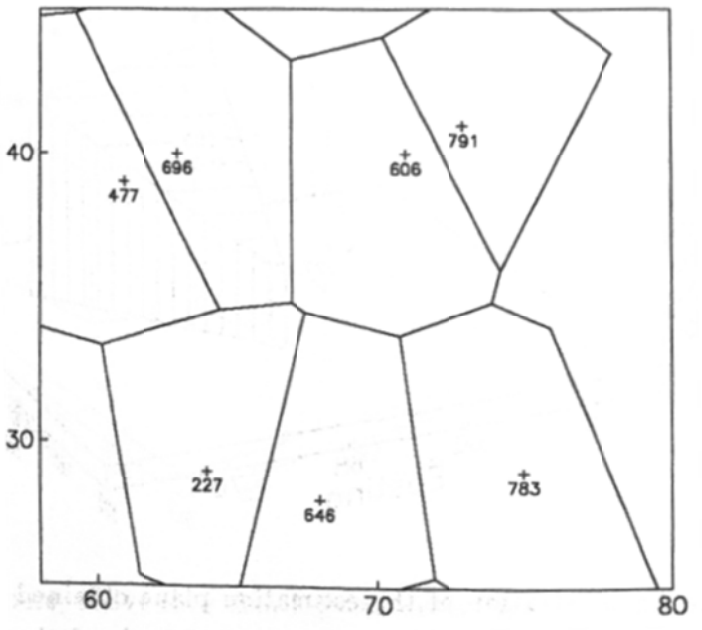


Figure 2-17 Polygonal area of influence around the known value

(Isaaks and Srivastava, 1989)

Using the polygonal technique, the estimate at any location will be the closest observation to the point which can expressed as (Goovaerts (2000) :

$$Z^*_{Pol}(u) = Z(u_{\alpha'}) \; with \; |u - u_{\alpha'}| < |u - u_{\alpha}| \; \forall_{\alpha} \neq \alpha' \qquad (2.22)$$

Goovaerts (2000) used the Thiessen polygon technique to determine the amount of annual rainfall, using 36 rain gage data. Figure 2-17 represents the location of the 36 gages and the estimated values of the areal rainfall over the entire area. Since the predictions were based on the polygonal figure, the effect of the stations and elevations were ignored, leading to large prediction errors.

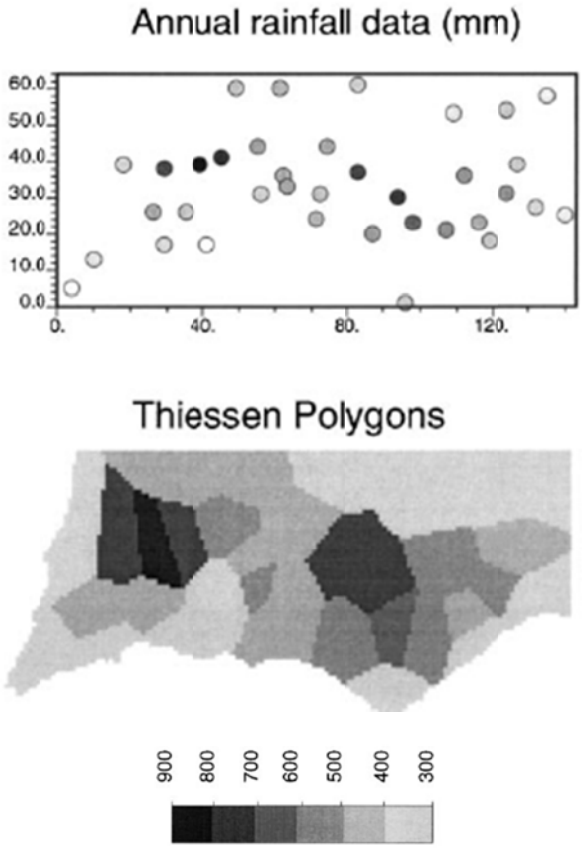


Figure 2-18 Thiessen polygonal estimation of rainfall data (Goovaerts, 2000)

2.6.2 Triangulation Estimation

Triangulation is a spatial estimation technique, named after Boris Delaunay, which is used for point estimations (Boris, 1934). A few estimation techniques, such as polygonal technique, are inherently associated with some discontinuities, as shown in Figure 2-18. By using the triangulation method, these discontinuities can be overcome by fitting a plane. The equation of the plane can be expressed as (Isaaks and Srivastava, 1989):

$$z = ax + by + c \tag{2.23}$$

Where, z is the value of the estimate

x and y are coordinates (such as easting/ northing)

c is constant.

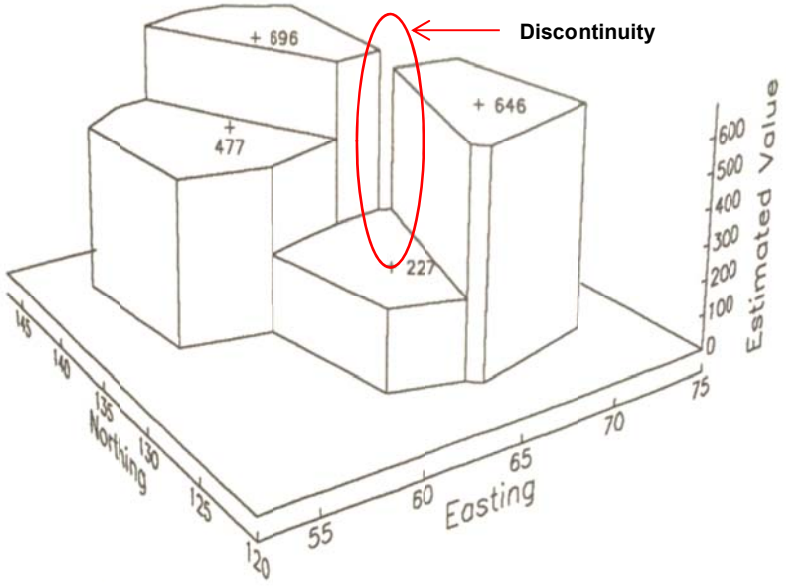


Figure 2-19 Inherent discontinuities present in polygonal estimation technique (Isaaks and Srivastava, 1989)

By knowing the values around the unknown location, a series of equations can be developed through which the constants a, b, and c are determined. Using the plane equation and constants, the unknown values are determined. Figure 2-19 from Isaaks and Srivastava (1989) provides the estimates, using the triangulation method for predicting the concentrations of a material in ppm.

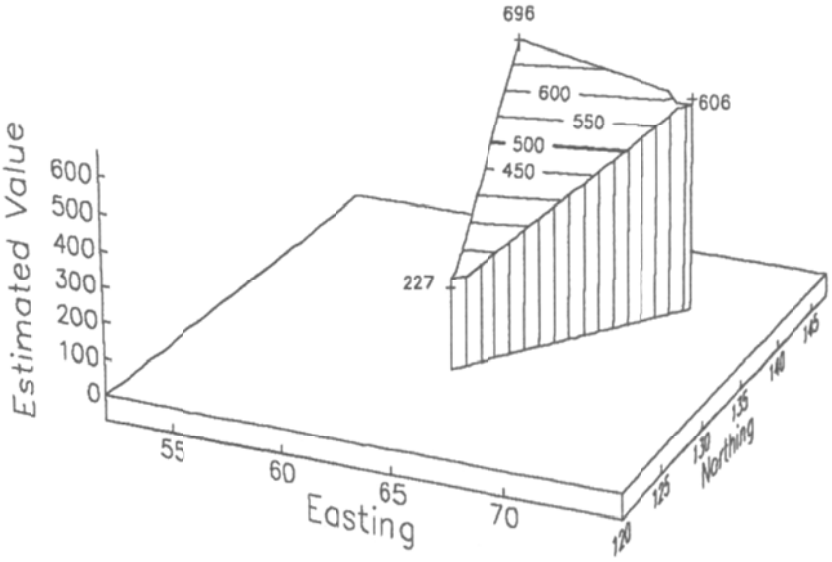


Figure 2-20 Triangulation estimation technique (Isaaks and Srivastava, 1989)

2.6.3 Inverse Distance Method

The estimate at any location is heavily influenced by the proximate values. The estimates made by using the polygonal technique make use of one closest value, and the triangulation method makes use of three proximate values (Isaaks and Srivastava, 1989). If the unknown value is surrounded by more than 3 points, the estimate at the unknown location is obtained by finding the local mean of the data. However, this technique provides equal weight factors to all the values and is mainly influenced by the extreme values. In order to accommodate this problem, the inverse distance method is utilized

where the distances from the unknown location to neighboring points are inversed and standardized with the mean distance.

The US National Weather Service (1972) estimated the rainfall depth using the inverse distance method (Bedient and Huber, 1992). The estimates using the inverse distance method are expressed as:

$$Z^*_{Inv}(u) = \frac{1}{\sum_{\alpha=1}^{n(u)} \lambda_\alpha(u)} \sum_{\alpha=1}^{n(u)} \lambda_\alpha(u) z(u_\alpha) \quad (2.24)$$

Where the weights, $\lambda_\alpha(u) = \frac{1}{|u - u_\alpha|^2}$

In the studies conducted by Goovaerts (2000), the area rainfall was estimated using the polygonal estimation method, inverse distance method, and kriging. It was observed that the estimates obtained from the kriging method resulted in minimal errors. followed by inverse distance method and polygonal method. Figure 2-20 shows the predictions obtained using the inverse distance method and kriging method.

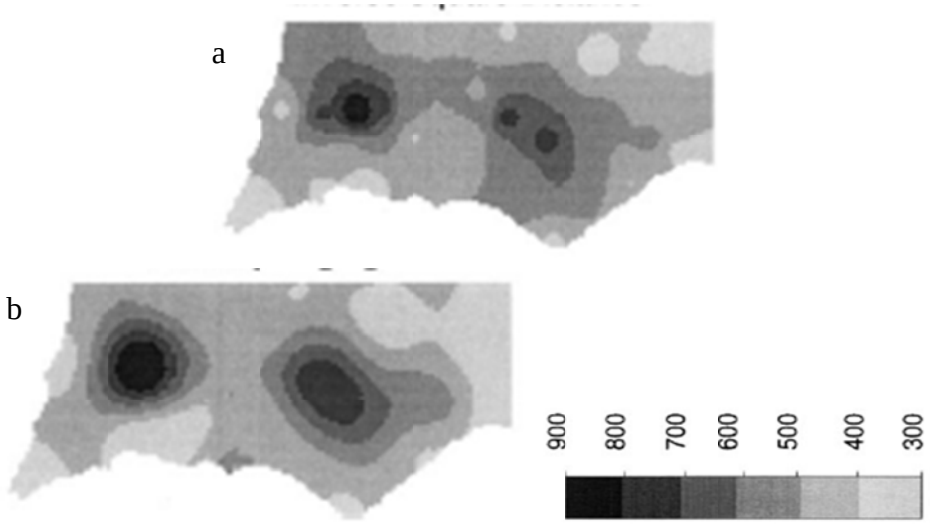


Figure 2-21 Estimations using (a) Inverse Distance Method (b) Ordinary Kriging (Goovaerts, 2000)

2.7 Predictions using Geostatistics

It is obvious that the estimate at any location is heavily influenced by the neighboring values. In conventional estimation methods, expected value (E (z)) at any point is found by providing equal weight to all the neighboring values or assigning the weights as per the polygonal or triangulation method. However, this is not true if the surrounding values are highly variable. Below is a hypothetical case presented to understand the above discussed scenario.

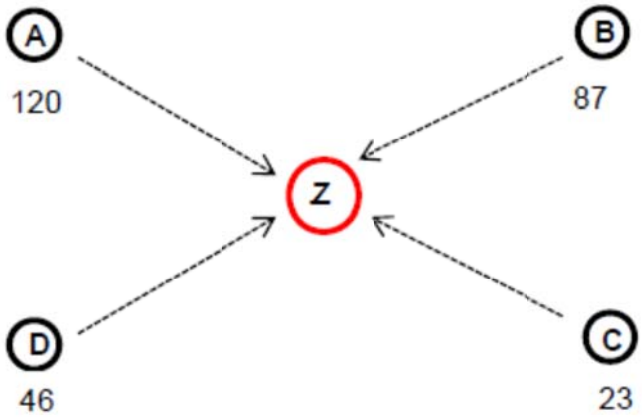


Figure 2-22 Hypothetical example for highly varying random variables

Consider a parameter with an unknown value in location 'z' which is surrounded by 4 different values at locations A, B, C and D. Through the univariate statistics, the expected value at 'z' will be an average of the neighboring values, which is 69. The expected value is given by:

$$E(Z) = \mu = \frac{1}{n}(w_i x_i) \tag{2.25}$$

The value of 69 is obtained by providing equal weight to all the neighboring values. However, the values around 'z' are varied, with a minimum value of 23 and maximum value of 120. So, providing equal weight to all the values would result in a rough and

misleading estimate. By applying geostatistics to the above study, the variability in the values is obtained by constructing and modelling a variogram. The weights of neighboring values are determined through the variogram model and the kriging algorithm.

Kumar et al (2006) summarized various applications of kriging in different fields, such as in the field of soil sciences (Burgess and Webster, 1980; Vieria et al., 1981; Berndtsson and Chen, 1994; Bardossy and Lehmann, 1998); hydrology (Goovaerts 2000; Creutin and Obled, 1982; Storm et al., 1988; Ahmed and de Marsily, 1989; Germann and Joss, 2001; Araghinejad and Burn, 2005); in depicting ground water levels (Delhomme, 1978; Volpi and Gambolati, 1978; Aboufirassi and Marino, 1983; Virdee and Kottegoda, 1984; Kumar, 1996); and in atmospheric science (Bilonick, 1988; Casado et al., 1994).

In this research study, Kriging, a best-linear, unbiased estimator (BLUE) in the geostatistics field was used for prediction analysis. Kriging is best because of its ability to reduce the error variance; linear because of the weighted linear combinations of data; unbiased as the data is considered stationary (Isaaks and Srivastava, 1989). In this section, the predictions based on geostatistical theory are discussed, with kriging used as the estimation tool. Various types of kriging methods and their applications are discussed in this section.

2.7.1 Simple Kriging

The kriging method is used to estimate the value of the unsampled locations by incorporating the spatial variability into the random variable (Miller et al. 2007). The spatial variability is obtained from the variogram model, where the maximum correlation distance between the random variable is obtained. Simple kriging is the simplest form of all kriging methods. This is mainly used in the mining industry, where the mean of the mining panels is a known value (Armstrong 1994). The main underlying assumption in

simple kriging is that the trend component is constant and the mean (m) is known, and the sample data set is expressed as a realization of random function (Thomey, 2013). Simple kriging is considered the least accurate of the various kriging methods, as it only assumes the first order moments constant (Olea 2009).

The estimate of simple kriging is expressed using the expression:

$$E[Z_{sk}^{*}(u)] = m = E[Z(u)] \quad (2.26)$$

Where, estimation error $Z_{sk}^{*}(u) - Z(u)$ is a linear combination of random variables.

2.7.2 Ordinary Kriging

Ordinary kriging is a type of kriging which is most widely when the first and second order moments are constant values, satisfying the second order stationarity (Haining et al. 2010; Olea 2009). In a research study conducted by Ahmed et al. (2012), the ordinary kriging was evaluated for mapping the salinity present in soils. The ordinary kriging was used to evaluate the depth of rock bed by incorporating the spatial variability present in the data values (Pijush, 2008). Using ordinary kriging, the estimate at any location is obtained by using the following equations:

$$\hat{Z}_X = \sum_{i=1}^{n} w_i x_i \quad (2.27)$$

Where, x_i is the neighboring value around the unsampled location

w_i are the weights of the corresponding neighbor values

$\hat{Z}_X$ is the estimate of the random variable

Since, kriging is a linear unbiased estimator the sum of all the weights corresponding to their neighboring locations is equal to 1.

$$\sum_{i=1}^{n} w_i = 1.0 \quad (2.27a)$$

Using the above estimation procedure will minimize the expected squared error between the true and predicted values. The expression below shows the expected squared error, e as:

$$Expected\ squared\ error, e = E\left\{\left[Z_x - \hat{Z}_{xi}\right]^2\right\} \quad (2.27b)$$

However, $E\left(\hat{Z}_{xi}\right) = \sum_{i-1}^{n} w_i E[Z(x_i)]$.

For an unbiased estimator, the estimated value should be equal to the true value.

$$E\left(\hat{Z}_{xi}\right) = E[Z(x_i)] \quad (2.27c)$$

Therefore, resulting sum of weights for the neighboring values equal to

$$\sum_{i=1}^{n} w_i = 1.0$$

The above can be accomplished by finding a set of values for which the differential equation for error with respect to each weight is 0.

$$\frac{\partial e}{\partial w_i} = 0 \quad (2.27d)$$

The solution to the above partial differential equation is obtained by applying the Lagrangian multipliers to the following set of simultaneous equations:

$$\sum_{j=1}^{n} a_j C_{ij} + \lambda = C_{xi}, \text{ and} \quad (2.27e)$$

$$\sum_{i=1}^{n} w_i = 1.0$$

Where, $C_{ij} = C(h)$ is the spatial correlation obtained from the variogram model

λ is a Lagrangian multiplier

$C_{xi} = C(h)$ all the data values and corresponding weights

In matrix form, the equations are:

$$\underbrace{\begin{bmatrix} C_{11} & C_{12}\ldots & C_{1n} & 1 \\ C_{n1} & C_{n2}\ldots & C_{nn} & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}}_{C} \underbrace{\begin{bmatrix} w_1 \\ w_n \\ \lambda \end{bmatrix}}_{\lambda} = \underbrace{\begin{bmatrix} C_{x1} \\ C_{xn} \\ 1 \end{bmatrix}}_{C_{xi}} \quad (2.27f)$$

Therefore, the solution which yields the weighing factors, w_i, is given by:

$$\widehat{W} = C^{-1}C_{xi} \quad (2.27g)$$

The weights of the neighboring values obtained using the above equation will be used for the estimation procedure.

2.7.3 Universal Kriging

Ordinary kriging, as discussed in the earlier section, is based on the condition where the mean is constant in the spatial process ($\sum_{j=1}^{n} a_j C_{ij} + \lambda = C_{xi}$). However, in some cases, the mean is not constant and is associated with the coordinates x and Y in various forms such as linear, quadratic or higher order trends (Isaaks and Srivastava, 1989). The below expressions are examples of linear and quadratic form.

$$Z(S_i) = \beta_0 + \beta_1 X_i + \beta_2 Y_i + \delta(s_i), \text{ linear} \quad (2.28)$$

$$Z(S_i) = \beta_0 + \beta_1 X_i + \beta_2 Y_i + \beta_3 X_i^2 + \beta_4 X_i Y_i + \beta_5 Y_i^2 + \delta(s_i), \text{ quadratic} \quad (2.29)$$

These trends in the data values effect the simple kriging estimations. Therefore, a new approach was developed, called universal kriging, where the weights of the neighboring values are estimated by accounting locally varying mean values. In universal kriging, unlike the simple or ordinary kriging, the intrinsic stationary condition has to be satisfied, accounting for trend in the mean as a deterministic component (Hohn, 1999; Olea, 2009). In case of the above deterministic trends, the predicted value at a location will again be expressed as a linear combination of the observed $Z(S_i)$, where i = 1,….,n values

$$\hat{Z}(S_0) = \omega_1 Z(s_1) + \omega_2 Z(s_2) + \cdots . + \omega_n Z(s_n) = \sum_{i=1}^{n} \omega_i Z_i \quad (2.29)$$

Where, $\sum_{i=1}^{n} \omega_i = 1.0$

In case of linear trend present in the data, the value $\hat{Z}(S_0)$ can be expressed as

$$\hat{Z}(s_0) = \beta_0 + \beta_1 \sum_{i=1}^{n} \omega_i X_i + \beta_2 \sum_{i=1}^{n} \omega_i Y_i + \sum_{i=1}^{n} \omega_i \delta(s_i) \quad (2.29a)$$

Comparing the above equation with the linear trend, the following conditions have to be satisfied.

$$\sum_{i=1}^{n} \omega_i X_i = X_0$$

$\sum_{i=1}^{n} \omega_i Y_i = Y_0$ and

$$\sum_{i=1}^{n} w_i = 1.0$$

2.7.4 Factorial Kriging

Factorial kriging was developed by Georges Matheron, based on a variogram filtering technique (Matheron, 1982). It is a multivariate geostatistical technique that is extensively used in the petroleum engineering field to reduce the noise present in the data (Magneron et al. 2009). Due to its ability to categorize the spatial components separately, it is also used to identify various metals present in soils (Queiroz et al., 2008; Jianshu et al., 2013; Benamghar et al., 2014)The main underlying assumption for factorial kriging is that the random function z(x) is modeled using two independent factors:

$$Z(x) = Z_1(x) + Z_2(x) \quad (2.30)$$

Where, $Z(x)$ is the random function

$Z_1(x)$ is the component of random variable due to noise

$Z_2(x)$ is the component of random variable due to signal

This technique is efficient in reducing the global noise; however, the limitation of this model is its inability to reduce the noise if the data is non-stationary.

2.8 Summary of Past Research Works

Variability in the soil properties is always a concern for geotechnical engineers. With high variability in soil properties, the complexity in choosing the appropriate design parameters escalates. Numerous studies were conducted to understand different types of variability associated with the soil properties. Univariate statistics were successfully utilized to analyze the distribution of the soil properties. However, spatial variability is one aspect which was not ascertained using univariate statistics. Different techniques such as random fields, Monte Carlo simulations, and turning band methods were used to simulate the real field conditions and describe the spatial variability present in the soil properties. However, the predictions made based on the random field theory had some limitations and resulted in error between true values and estimated values.

In this research study, geostatistics, developed from regionalized random variable theory, was implemented in various problems of interest. The spatial variability in the real field conditions was simulated by constructing a variogram, and an unbiased estimator kriging was applied for estimations.

Chapter 3

Formulation of Framework for Spatial Variability Analysis in Geotechnical Engineering

3.1 Introduction

Predictions in geotechnical engineering at unsampled locations are often based on univariate statistics, such as finding the expected value of all the observations and evaluating the distribution of its properties. The predictions obtained from the univariate statistics are highly sensitive due to its inability to capture the spatial variability that is distributed over the entire area. In this research study, geostatistics was used to incorporate the spatial variability present into the soil properties for prediction analysis.

In order for the predictions to be unbiased, with minimum error, certain steps had to be taken to ensure accurate simulation of the real field conditions. Any deviations or violations of the assumptions of the test procedures would lead to unrealistic results. In the current study, a framework was developed, using both the univariate statistics and geostatistics, so that the spatial variability present in the soil properties was incorporated into the prediction analysis. This chapter provides a detailed discussion of every step that is required to incorporate spatial variability into prediction analysis.

3.2 Data Acquisition

Data acquisition is the primary step in any engineering analysis. The type of data, quality of data, and quantity of the data play an important role in prediction analysis. In geotechnical engineering, type of data refers to the raw data obtained from the field, such as shear strength interpretations from field vane shear test or the data obtained from the laboratory studies. The geostatistical modeling, when conducted using field data, would give more accurate results when compared to the data obtained using empirical correlations. This could be due to the uncertainties involved with the model, in situ

measurement error, or calibration error. With the amount of uncertainty, the quality of data gets affected, resulting in misleading predictions. Another important parameter that affects the predictions is the quantity of data. A huge amount of data is usually more desirable, as it allows more accurate simulation of the real field conditions. With a minimum number of observations, the risk associated with the predictions usually increases.

3.2.1 Data Organization

Once the data is collected from the field or developed using empirical correlation, the next important step is to organize the data. In order to perform geostatistical modelling, the data should be organized in the Cartesian coordinate system. The x, y of the Cartesian system represents the spatial location of a particular data point in space; whereas, z represents the value of the variable of interest. Since geostatistics depend on the separation distance of the data points rather than location, the position of the data point can be given to our own individual coordinate system rather than depending on east, west, north or south or latitude and longitude. With the defined coordinates of the data points, the data is further used for univariate statistics and geostatistical modelling, as explained in the later sections.

3.3 Statistical Analysis

The statistical analysis of the data gives a fair understanding of the distribution of the data, which governs the type of statistical tests to be adopted. Even though the univariate statistics can't incorporate the spatial variability present into the soil properties, it helps in conducting several checks, such as constant mean value, constant variance, and Gaussian distribution of the data.

3.3.1 Histograms

Histograms provide a basic idea about the dispersion in the geotechnical data. A histogram is a statistical tool, which provides a graphical representation of the distribution of the data. Histograms are a plot of the frequency of the data against the class interval, as shown in Figure 3.1. The height of each bin represents the number of observations that fall in that particular bin size. The histogram plots are sensitive to the bin size and number of bins selected. A proper selection of bin size provides a histogram, where the frequency distribution of a continuous set of data and skewness of the data can be discovered.

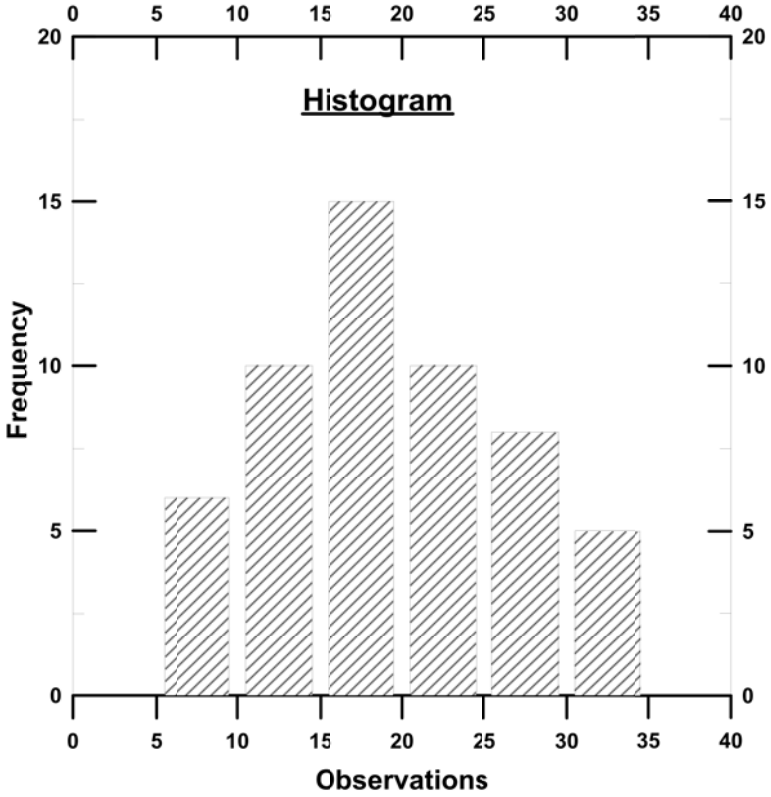


Figure 3-1 Plot of Histogram

In this research study, histograms were plotted to observe the distribution of the data and skewness present in the data. The skewness present in the data directly affects

the predictions analysis performed using geostatistics. The ideal condition, for the most accurate predictions with minimum error, results when the skewness is equal to zero. The number of class interval (CI) for plotting the histogram is found using Sturges (1926) Equation 3.1.

$$No.of\ Class\ Intervals = 1 + 3.3 \log_{10} N \quad (3.1)$$

Where, N is the total number of observations in the data set

Once the class interval (CI) is determined, the bin size is found by using Equation 3.2. The bin size obtained can be rounded off to the nearest decimal point while plotting the histogram.

$$Bin\ Size = \frac{Maximum\ Value - Minimum\ Value}{Number\ of\ Class\ Intervals} \quad (3.2)$$

3.3.2 Check for Normal Distribution of Data

Soils are formed due to natural geological processes; consequently, the distribution of the soil properties can vary from location to location. As most of the statistical tests are based on normal distribution, the soil properties, should be checked for having normal distribution of the data.

Statistical models are composed of a systematic component such as trend (deterministic or structural) and a random component (error) to capture the natural variation (Montgomery, 2010). Based on the central limit theorem, the random error component is normally distributed with mean value, μ and a standard deviation, σ. The data is said to be normally distributed if the probability of any random variable falls in real limits under the proper bell-shaped curve, with either of its ends approaching zero. The most desirable condition for conventional statistical analysis of a random variable is the when the data is normally distributed. Clark and Harper (2002), in their research study,

stated that the violation of normality affects the spatial correlation of the random variable and further impacts the final outcome of the predictions.

In this research study, the soil property under study was checked for the normality assumption. Histogram plots are the simplest way to check for normal distribution of the data; however, because of their sensitivity to the number of class intervals and bin size, they were not employed. Normal-Quantile plot (n-q plot) and Shapiro-Wilk test are two strong tests, among many, that were used in this study to check the normality of the data.

3.3.2.1 Normal – Quantile plot (n-q plot)

Normal-Quantile (n-q) plot is the most commonly-used graphical tool to assess how well the data fits the normal distribution. In an n-q plot, quantile values of a theoretical distribution are plotted against the normal values from the data set, as shown in Figure 3.2. Initially, the normal values for which the test was being performed were arranged in the ascending order. The ranks for all the values were assigned in the increasing order starting with 1. The rank proportion or theoretical quantile value was calculated for each value using below Equation 3.3.

$$q \sim \frac{i-0.5}{N} \tag{3.3}$$

Where, i is the rank of the observation

N is the total number of observations

The z-values or z-scores for all the theoretical quantile values were obtained using the normality tables. The obtained z-values, which are theoretical quantile values, were plotted against the normal values. Thereafter, the aim was to make a judgment as to whether the actual values were normally distributed. If this was the case, then the plotted points would create a straight diagonal line. Any systematic deviations from a straight

line, other than natural random fluctuations, suggest that the data cannot be considered to be normally distributed. Since, it is not always practically possible to obtain a straight line, the best fit line linear straight line is drawn for the plotted data. The coefficient of determination (R^2), which determines how well the data fits a statistical model, was determined using the equation below.

$$R^2 = \frac{SS_{res}}{SS_{tot}} \tag{3.4}$$

Where, SS_{res} is called residual sum of squares

SS_{total} is called total sum of squares

An assumption was made in this research, that if the R^2 value was greater than 80%, the data was assumed to be in good fit with the normal distribution.

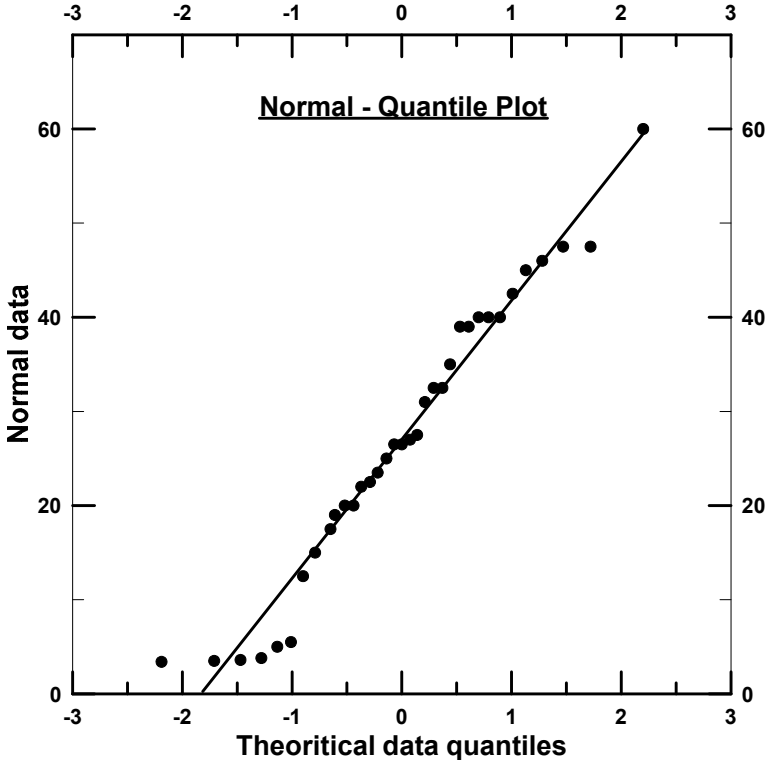


Figure 3-2 Normal-Quantile plot (n-q)

3.3.2.2 Shapiro-Wilk Test

The Shapiro-Wilk test is the most appropriate statistical test for evaluating the normal distribution of the data. It was developed by Samuel Sanford Shapiro and Martin Wilk in the year 1965. The test is based on the hypothesis testing that the sample data is normally distributed. The detailed statistical procedure involved in developing the test is given in the paper by Shapiro and Wilk (1965). The test statistic developed for verifying the null hypothesis, i.e., whether data is normally distributed, is given below.

$$W = \frac{b^2}{SS} \tag{3.5}$$

Where, b = $\sum_{i=1}^{n/2} a_i (x_{n+1-i} - x_i)$ (3.5a)

$$SS = \sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{3.5b}$$

x_i is the ith order statistic

$\bar{x}$ is the sample mean of the data

n is the number of observations

a_i is the weight for individual observation

In this research study, the Shapiro-Wilk test, along with normal quantile plots, was used for evaluating the normality of the geotechnical data of interest. Once the test statistic was calculated, the corresponding P-value was obtained, which is the probability that the sample average will take on a value that is at least as extreme as the observed value when the null hypothesis H_0 is true. If the p-value calculated for the above statistic is less than the significance level (α-level), then the null hypothesis is rejected, concluding that the data is normally distributed. If the p-value obtained is greater than the test statistics, it is concluded that there is not enough evidence to assess the distribution of the data. In this study, the significance level chosen for this statistical test is 0.05.

3.4 Stationarity in Data

Stationarity in geotechnical engineering is referred to as statistical homogeneity in soil parameters (Baecher and Christian, 2003). In an economic time series, stationarity is defined as the quality of a process in which the statistical parameters, such as mean and variance, do not change with the time (Challis and Kitney, 1991). In geostatistics, the stationarity refers to the data having same joint probability distribution over the space with constant mean and variance values. All pairs of random variables that are separated by a distance h depend upon the lag distance, but not on their location (Isaaks and Srivastava, 1989).

The stationarity is divided into two categories :strongly stationary process and intrinsic stationary process. The strongly stationary process, or truly stationary process, is called when all the higher-order moments are constant, including the variance and mean. In general terms, strongly stationary data can be described as data sets that have same constant mean, constant variance, and equal probability distribution. The intrinsic stationary process, or weak stationary process, refers to the data having a constant mean and variance throughout the space.

The geostatistical methods are optimal when the data is normally distributed and stationary (mean, variance, and joint probability distribution) do not vary significantly in space. The truly stationary processes are mainly theoretical and are discussed only for their mathematical properties, which are difficult to apply to practical problems. However, the intrinsic stationary has to be satisfied in order to conduct geostatistical analysis. Cuba et al (2011) discussed the various scenarios of the data with constant and varying statistical parameters.

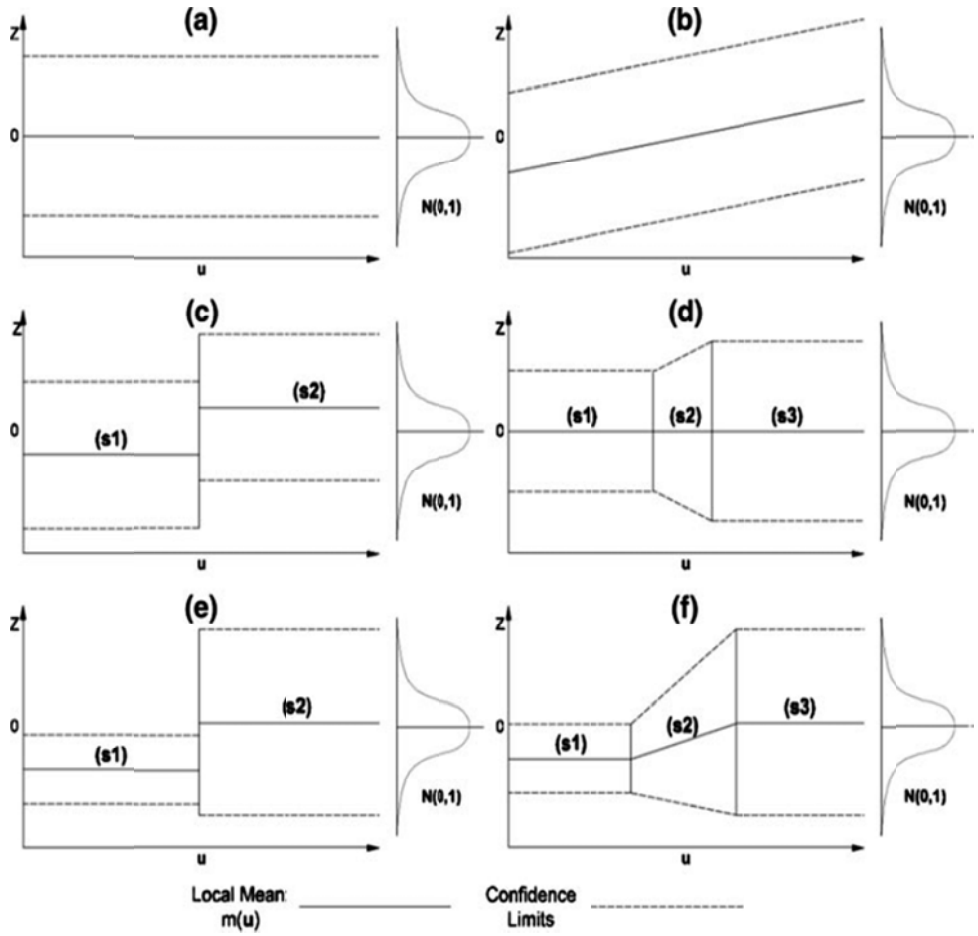


Figure 3-3 Schematic 1D Gaussian Environments by Miguel, Oy & Julian (2011)

The above Figure 3-3 provides various scenarios of the data distribution with: (a) constant local mean and local variance; (b) linear mean trend and constant variance; (c) two sub-regions of constant local variance, but different local means; (d) two regions of constant local mean, but different local variance with a transition zone; (e) two sub-regions of different local means and different local variances; (f) two regions of different local means and different variances with transition zone.

In this research study, the stationarity of the data was checked using different approaches. The basic and simplest way to check for stationarity is through the histograms. If the histograms are skewed to one side, it clearly depicts that the data does

not have a constant mean. However, as discussed in earlier sections, the histogram plot changes with the change in bin size and number of class intervals. Another standard procedure is to detect the stationarity through the experimental variogram. If the experimental variogram values continue to increase beyond the priori variance of the sample data set without reaching a sill, the data can be concluded as non-stationary. In this section, the intrinsic stationary process, i.e., constant mean and variance was checked using univariate statistics. In order to supplement this check, the stationary was also verified in the later sections by using experimental variogram.

3.4.1 Check for Stationarity in Data Using Univariate Statistics

In order to perform geostatistical analysis, the preliminary step is to check whether the data is stationarity or non-stationary. In this study, two statistical approaches, analysis of variance (ANOVA) and Bartlett's test, were employed to check for stationarity (constant mean and constant variance) in the data.

3.4.1.1 Check for Constant Mean using ANOVA

Analysis of variance (ANOVA) is a statistical test procedure developed by R.A. Fisher to analyze the difference in group means. In this research study, this test was used to evaluate the stationarity present in the data by testing whether the mean was constant in all the treatments/sections under consideration. This was accomplished by developing a statistical hypothesis, as shown below.

Null hypothesis: $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 \text{ ------- } \mu_n$

Alternative hypothesis: H_1: At least one mean is different

The difference in the mean values, or the variability present in the data values, was attributed to different sources of variations, such as variability within the treatments or variability between the treatments (Montgomery et al. 2010). The variability between the treatments is obtained through treatment sum of squares ($SS_{Treatment}$), and variability

within the treatments is obtained by the error sum of squares (SS_{error}). Therefore the total variability i.e., total sum of squares (SS_{Total}) will be equal to the treatment sum of squares and error sum of squares.

$$SS_{Total} = SS_{Treatment} + SS_{Error} \quad (3.6)$$

Where, $SS_{Total} = \sum_{i=1}^{a}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{..}\right)^2 = total\ sum\ of\ squares$ (3.6a)

$$SS_{Treatment} = n\sum_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2 = treatment\ sum\ of\ squares \quad (3.6b)$$

$$SS_{error} = \sum_{i=1}^{a}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.}\right)^2 = error\ sum\ of\ squares \quad (3.6c)$$

Once the variability in the groups and between the groups is assessed the hypothesis is tested using the statistics below. The detailed procedure of developing the test statistic (F_0) is discussed in design and analysis of experiments text books.

$$F_0 = \frac{MS_{Treatments}}{MS_{Error}} \quad (3.7)$$

Where, $MS_{treatments} = \frac{SS_{Treatment}}{a-1} = Treatment\ Mean\ Square$ (3.7a)

$$MS_{Error} = \frac{SS_{Error}}{a(n-1)} = \text{Error Mean square} \quad (3.7b)$$

The test statistic (F_0) is compared to the critical value obtained from a f-distribution table with (a -1) and (N – a) degrees of freedom. We reject the null hypothesis (H_0) that if the test statistic (F_0) is greater than $F_{\alpha,\ a-1,\ N-a,}$ the means are not equal in at least one section. Tables 3-1 and 3-2 show the typical data arrangement for a single factor ANOVA experiment and a summary of ANOVA test results.

Table 3-1 Typical data for single-factor ANOVA experiment

Treatment	Observations				Totals	Averages
1	Y_{11}	Y_{12}	…	Y_{1n}	$Y_{1.}$	$\bar{y}_{1.}$
2	Y_{21}	Y_{22}	…	Y_{2n}	$Y_{2.}$	$\bar{y}_{2.}$
…	…	…	…	…	…	…
A	Y_{a1}	Y_{a2}	…	Y_{an}	$Y_{n.}$	$\bar{y}_{a.}$
					$Y_{..}$	$\bar{y}_{..}$

Table 3-2 ANOVA for a single factor experiment, random effects model

Source of Variation	Sum of Squares (SS)	Degrees of Freedom)DF)	Mean Square (MS)	F_0
Treatments	$SS_{Treatmentes}$	a-1	$MS_{Treatments}$	$MS_{Treatments}/MS_E$
Error	SS_E	a(n-1)	MS_E	
Total	SS_T	an-1		

Where a = number of different levels of a single factor

n = number of observations in respective treatment

N = Total number of observations

3.4.1.1.1 Model Adequacy check

The main underlying assumption in the ANOVA method is that the residual errors (e_{ij}) are normally and independently distributed, and variance in all residual is structureless; that is, it should contain no obvious pattern.

$$e_{ij} = y_{ij} - \widehat{y_{ij}} \tag{3.8}$$

Where, e_{ij} is the residual error

y_{ij} is the observation in the ith treatment and jth row

$\hat{y}_{ij}$ is an estimate of the corresponding observation y_{ij}

3.4.1.2 Check for Constant Variance using Bartlett's test

Bartlett's test is a statistical test procedure named after Maurice Stevenson Bartlett to analyze 'k' samples from a population having equal variances or not. As a part of the stationary requirement, in this research study, this test was used to evaluate the constant variance present in the data in all the treatments/sections under consideration. This was accomplished by developing a statistical hypothesis, as shown below.

Null hypothesis: H_0: $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$ ------- σ_n^2

Alternative hypothesis: H_1: At least one variance is different

The basic procedure involved in Bartlett's test is computing a test statistic whose sample distribution can be approximated using a chi - square distribution with a-1 degrees of freedom. The test statistic is

$$\chi_0^2 = 2.3026\ \frac{q}{c} \tag{3.9}$$

Where, $q = (N-a)\ \log_{10} S_p^2 - \sum_{i=1}^{a}(n_i - 1)\log_{10} S_i^2$ (3.9a)

$$c = 1 + \frac{1}{3(a-1)}\ (\sum_{i=1}^{a}(n_i - 1)^{-1} - (N-a)^{-1}) \tag{3.9b}$$

$$S_p^2 = \frac{\sum_{i=1}^{a}(n_i - 1)\ S_i^2}{N-a} \tag{3.9c}$$

S_i^2 is the sample variance of the ith population

From the above statistic, it can be observed that the quantity q is large when the sample variances have high variability, and q is zero if the sample variances are equal. Therefore, the null hypothesis (H_0) shall be rejected when $X^2_0 > X^2_{\alpha,\ a-1}$, concluding that the variances are not equal.

3.5 Non-Stationary Data

The condition of stationarity in geotechnical engineering, which is when the soil properties have constant mean and variance at all locations, may not be applicable to all conditions. In such cases, the soil properties are said to be non-stationary in nature. It is extremely important to identify non-stationary behavior in soil parameters before performing geostatistical analysis. Non-stationarity refers to the data having locally varying means, variances, and covariance within the same population. Non-stationary data are unpredictable and cannot be modeled or forecasted. For geostatistics, the basic assumption is that the data is stationary.

In order to get reliable predictions in geotechnical engineering, the non-stationary data needs to be transformed to stationary data, using appropriate transformation methods. Before applying the transformation, it is important to identify different types of non-stationary processes present in the data. Figure 3-4 below shows the different types of non-stationary processes such as pure random walk; random walk with drift; trends (trends can be constant, positive or negative, independent of location); and a combination of any two that can present in the data.

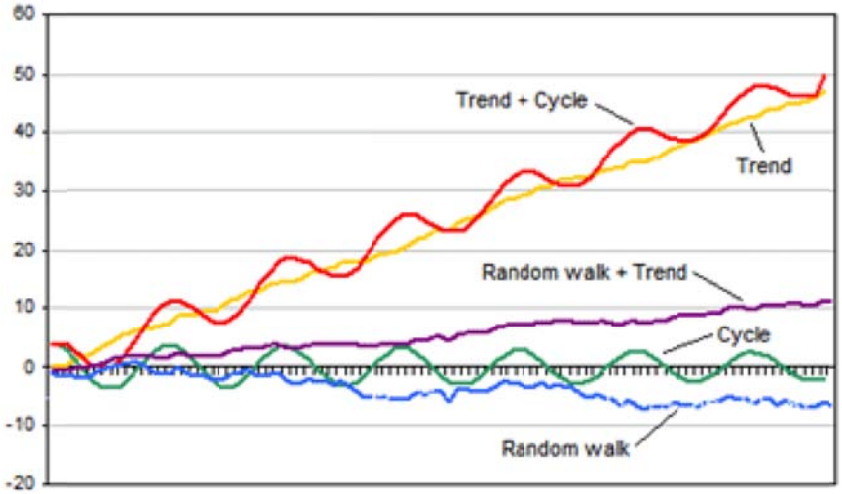


Figure 3-4 Different non-stationary processes (Investopedia.com)

3.5.1 Random Walk

Random walk refers to the arbitrary variation in the data when compared to the mean or previous value. The random walk can be sub-divided into two categories: pure random walk and random walk with a drift. Below are the two equations for random walk that explain the value at any location, with and without drift.

Pure Random Walk : $Y_t = Y_{t-1} + \varepsilon_t$ (3.10)

Random Walk with drift : $Y_t = \alpha + Y_{t-1} + \varepsilon_t$ (3.11)

Where, ε_t is the stochastic component

α is the drift

Y_t is the value at time t

Y_{t-1} is the value at time t-1

In pure random walk, the value at any location (Y_t) is equal to the previous value (Y_{t-1}) plus a stochastic (non-systematic) component (ε_t). Whereas, if the random walk is associated with the drift, the value at any location is governed by a constant or drift value (α), along with previous value and stochastic component.

3.5.2 Trend

The trend can be defined as low-frequency, large-scale variations (Olea, 1991; Cuba et al. 2011). In geostatistics, the trend refers to variation in the local mean and local variance. In the time series, the trend is given by the following equations:

Deterministic Trend : $Y_t = \alpha + \beta t + \varepsilon_t$ (3.12)

Random Walk with Drift and Deterministic Trend: $Y_t = \alpha + Y_{t-1} + \beta t + \varepsilon_t$ (3.13)

Where, βt is deterministic trend

In the deterministic trend, the value at any location is influenced by the stochastic component (local variation), drift, and the trend pattern of the data. In geostatistics, the trend present in the data develops a bad correlation between the variable values when

separated by large-lag distances. The trends present in the data can be identified by plotting the semi-variogram. If the empirical semi-variogram values continue to increase beyond the priori variance or global variance without reaching a sill, this often indicates that the spatial trend is present in the data.

Cuba et al (2011) presented the experimental semi-variogram plots, through which various trends that could present in the spatial data can be identified. Figure 3-5 shows the different conditions for the trends in the data by Cuba et al (2011). Figure 3-5(a) displays a condition where the semi-variogram values are below the priori variance, or global variance, reflecting no trend present in the data. In Figure 3-5(b), the semi-variogram values follow a linear trend from the origin and continue to increase linearly beyond the priori variance value. This linear increase, beyond the priori variance, reflects the linear trend present in the spatial data. Similarly, Figures 3-5 (c) and (d) reflect the parabolic and variance trend, where the semi-variogram values increase beyond the priori variance, reflecting the parabolic and variance trend present in the spatial data.

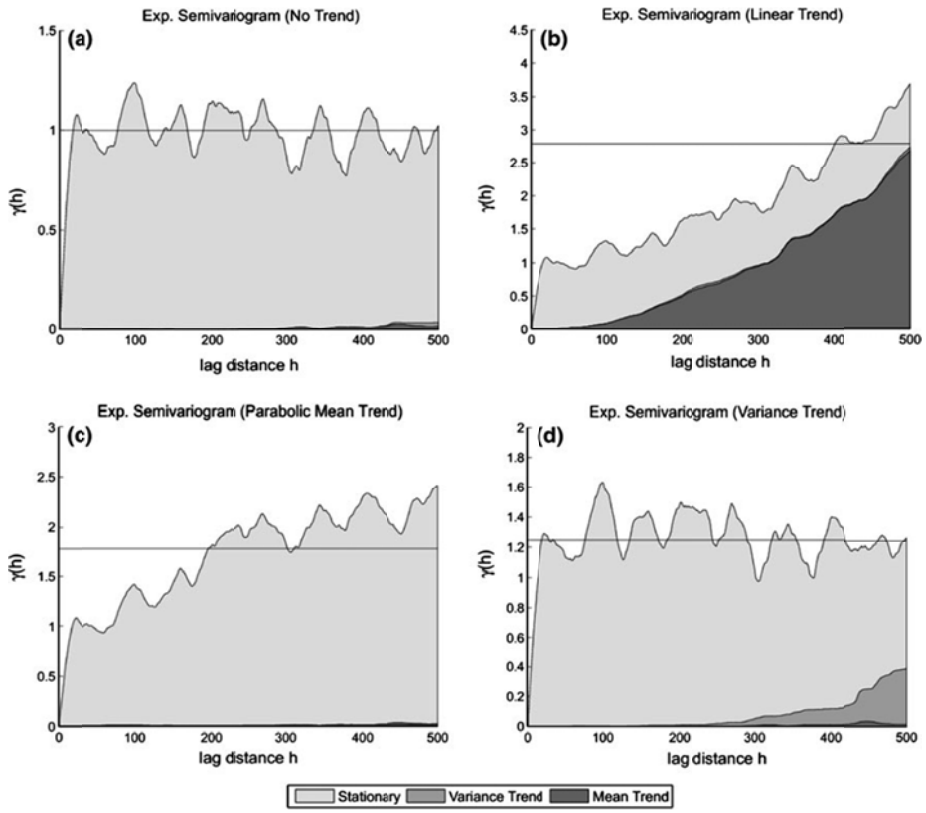


Figure 3-5 Trends present in the data by Miguel et al (2011)

3.6 Non-Stationary Data to Stationary Data

In order to perform geostatistical analysis, the primary requirement is to have stationary data. However, the heterogeneous soils are highly variable, with varying properties along the subsurface profile. This brings a limitation for applying geostatistics to the geotechnical engineering field. In such cases, one approach is to apply transformations to the data, so that the data behave close to stationary condition.

3.6.1 Transformations

Transforming geotechnical data is expressing the data in a different function. For example, the effective stress in a consolidation curve is represented on a logarithmic

scale instead of using a normal scale. Even though it is expressed in logarithmic function, the inherent meaning does not change, i.e., void ratio is dependent on the vertical stress.

Transformations play a vital role in mathematical and statistical analysis. As the name suggests, transformations are merely the expressions of the data in different forms. By transforming the data, many underlying assumptions of the statistical tests can be satisfied. The key for transformations is to use the right function. For example, by using the square root function, the upper bound of the data can be reduced more than the lower-bound data.

Howell (2007), in his study, suggested that the transformed data and the untransformed data should not differ in their basic assumptions about the data. Tabachick and Fidell (2007) provided the guidelines for selection of appropriate transformation methods based on the skewness present in the data.

Table 3-3 Transformation Method Guidelines by Tabachick and Fidell

Skewness in data	**Transformation method**
Moderately positive skewness	Square- Root $\bar{y}$ = SQRT (Y)
Substantially positive skewness	Logarithmic $\bar{y} = \text{Log}_{10}(Y)$
Substantially positive skewness (with zero values)	Logarithmic $\bar{y} = \text{Log}_{10}(Y + C)$
Moderately negative skewness	Square-Root $\bar{y}$ = SQRT (K-Y)
Substantially negative skewness	Logarithmic (Log_{10}) $\bar{y} = \text{Log}_{10}(K-Y)$

*Where C and K are constants, $\bar{y}$ is the transformed value

In this research study, transformations were performed to stabilize the variance in the ANOVA test and to convert the non-stationary data to stationary. Instead of using the empirical methods to select the appropriate transformation, a more formal statistical approach was employed.

3.6.1.1 Box- Cox Transformation

Douglas (2009) stated that the power family of transformations (i.e., $y^* = y^{\lambda}$) are extremely useful in stabilizing the variance, making the distribution normal and improving the fit of the model. Box-Cox (1964) is a statistical test procedure used in selecting the parameter 'λ' along with other model parameters, such as overall mean and treatment effects. The maximum likelihood estimate of λ is the value at which the error sum of squares is minimum. This is usually found by plotting a graph for the error sum of squares for different values of λ. In order to plot that graph, the data has to be transformed for different values of λ, using the equations below. Each value of λ produces a different set of data for which the error sum of squares is obtained.

$$y^{(\lambda)} = \begin{cases} \frac{y^{\lambda}-1}{\lambda \dot{y}^{\lambda-1}} & \lambda \neq 0 \\ \dot{y} \ln y & \lambda = 0 \end{cases} \quad (3.14)$$

Where, $\dot{y} = ln^{-1}[(1/n) \sum \ln y]$ is the geometric mean of the observations

In this research study, the Box-Cox procedure was utilized to select the appropriate transformation. The selected transformation was adopted for stabilizing the variance present in the data and to convert the non-stationary data into stationary.

3.6.2 Detrending

Detrending refers to a mathematical technique for removing a trend present in the data. The geotechnical models are often described as function of linear or non-linear trends, such as exponential power. These trends provide biased predictions, which affect the predictions at unsampled locations. The trends in the geotechnical data can be

identified by constructing an experimental variogram plot, as shown in Figure 3-5. Once the trend in the spatial data is determined, a trend surface is fitted using the least squares method. The trend surface values are subtracted from the original data, resulting in the residual values. These residual values are the detrended data that is used for geostatistical analysis. Kitanidis (1994) presented the equations that are used for detrending the data.

Linear Detrending: $Z_{new} = Z - [AX + BY + C]$ (3.15)

Quadratic Detrending: $Z_{new} = Z - [AX^2 + BY^2 + CXY + DX + EY + F]$ (3.16)

Where A, B, C, D, E, F are the constants and X, Y are spatial coordinates of the Z-value

Parabolic Detrending: $Z_{new} = Z - [A_0 + A_1X + A_2Y + A_3X^2 + A_4XY + A_5Y^2]$ by Vieira (2010) (3.17)

Where A_0, A_0, A_0, A_0, A_0 are constants and X, Y are spatial coordinates of the Z-value

3.7 Spatial Variability Analysis

The variability present in the geotechnical data with spatial distance is referred to as spatial variability. Spatial continuity is an important characteristic of geostatistics, through which the spatial variability in the data is captured. The three functions that are used to describe the spatial variability are covariance function, correlation function, or correlogram and semi-variance function or variogram.

In the time series or temporal analysis, the correlogram and covariance functions are often used. In geostatistics, semivariogram is the most commonly used tool to describe the spatial variability present in the variable of interest. The primary reason for using the variogram is because of its ability to filter the influence of spatially varying mean. Also, the second-order stationarity shall be satisfied when using the covariance function or correlation function, which is sometimes highly impractical to satisfy. Where

else, the variogram works well with the data that satisfies the basic intrinsic stationarity (constant mean, variance, joint probability distribution).

In this research study, variogram, a geostatistics tool is used to find the spatial variability of the soil properties. The captured spatial variability was incorporated into the prediction analysis through a stochastic interpretation tool called Kriging. Once the collected data was checked for Gaussian distribution, stationarity, and any trend in the data, it could be further used for spatial variability analysis. In this study, Surfer, a commercially available software was utilized for spatial variability modelling and prediction analysis.

3.7.1 Experimental Variogram

Variogram is defined as the average of the squared difference of the random variable for different lag distances. The mathematical definition of the variogram or semivariogram is

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} [z(x_i + h) - z(x_i)]^2 \qquad (3.18)$$

Where, $z(x_i)$ = measurement taken at a location x_i ;

$z(x_i + h)$ = measurement taken at a location h distance away;

n(h) = number of data pairs h units apart in the direction of the vector

h = lag distance

$\gamma(h)$ = variogram value

From the above equation, it can be clearly understood that the variogram value depends on the separation distance/lag distance (h) and on the number of data pairs. However, the number of pairs and separation distance are interrelated, which influences the variogram value. With the given number of observations, the number of pairs can be

computed using Equation 3-19. For example, if there are 30 observations with a distance of 3 meters apart, then 435 pairs can be formed using those 30 observations.

$$\text{Number of pairs} = n\,(n-1)/2 \qquad (3.19)$$

Where, n is the number of observations in a sample data.

The type of grid selected for constructing the variogram influences the number of pairs. In Surfer software, the variogram grid is divided based on the maximum lag distance, radial, and angular divisions, as shown in the Figure 3-6. The maximum lag distance governs the maximum spatial continuity that a variogram can capture. The angular division specifies the number of divisions that is made by the spokes in the variogram grid. The number of radial divisions specifies the area that is divided into a number of concentric circles. The angular divisions of 180 and radial divisions of 100 usually meet the requirements for almost any project due to its broad spectrum of layout.

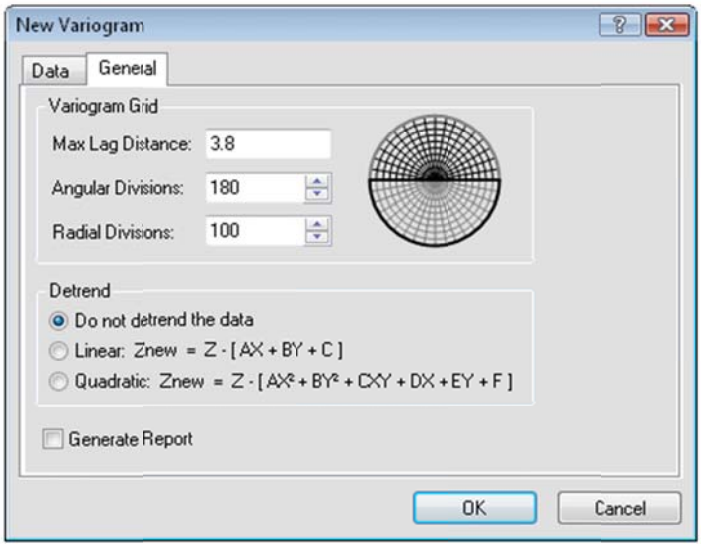


Figure 3-6 Grid Selection for constructing experimental variogram (surfer)

The Surfer software pre-calculates all the pairs based on the number of observations, and stores the information in the selected grid. Apart of the grid selected, the three most

important factors that heavily influence the variogram plot are lag distance, lag direction, and lag tolerance.

3.7.1.1 Lag distance (h)

Lag distance, or separation distance, refers to the distance between the two pairs of random variables. The mathematical equation for a variogram clearly indicates that as the lag distance increases, the corresponding number of pairs can increase or decrease and significantly influence variogram value. However, there are no standard procedures or rules employed for selecting the lag distance. In this research study, a trial and error procedure was adopted for selecting the lag distance. The lag distance that produced a minimum number of 30 pairs was selected as the desirable lag distance for the construction of a variogram. As an initial estimate, the minimum distance between two bore holes, or the distance between two test locations, can be given as the lag distance.

3.7.1.2 Lag Direction

Lag direction is another important factor that influences the variogram value and variogram plot. Lag direction signifies the direction in which the variogram should be constructed. For example, the direction of zero degrees (0°) indicates the pairs that are separated in x-direction, and direction of ninety degrees (90°) indicates the pairs that are separated in y-direction. In geotechnical engineering, the lag direction plays an important role. Through the lag direction, the horizontal and vertical variability in the soil properties can be modelled. In this research study, the variograms were constructed for different directions to capture the spatial variability in all the directions, as shown in Figure 3-7.

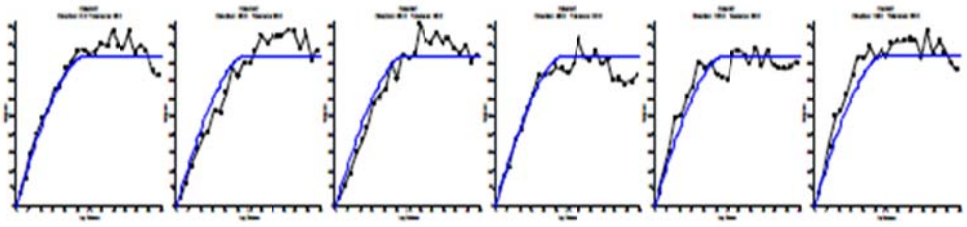


Figure 3-7 Directional variograms: starting from left 0°, 30°, 60°, 90°, 120°, 150°

(surfer software)

3.7.1.3 Lag Tolerance

Lag tolerance refers to the directional tolerance that needs to be considered for the construction of an experimental variogram. By incorporating the tolerance value, the nearby pairs of the random variables are also considered for the construction of the experimental variogram. For example, 0° (direction) + 45° (tolerance) considers all the random variable values that are placed at 45° in the x-direction, along with variables that are placed in 0° direction. The directional tolerance specifies the size of the angular window for the experimental variogram. The angular window can be given as:

Direction – Tolerance < Angle < Direction + Tolerance

In this research study, the experimental variograms were constructed for different directions, such as 0°, 30°, 45°, 60°, 90°. For every directional variogram, the tolerance limits were given from 0° to 90º.

3.7.2 Variogram Modeling

The important properties of the variogram, which define the spatial variability or spatial continuity present in the variable of interest, are range, sill, and nugget. These three properties affect the predictions performed, using geostatistical analysis. In order to model these properties, various standard variogram models should be used. Most of the standard variogram models are based on the assumption that the experimental

variogram will reach the sill value after certain lag distance. The global variance of the observations in a sample is an estimate of the sill value (David, 1977; Journel and Huijbregts, 1978). The distance at which the variogram value reaches sill value is called range, which indicates the spatial correlation of the random variables until that distance. Below are some standard models that are used to model a variogram with a sill.

3.7.2.1 Nugget Model

Nugget model, or nugget effect, is used to model the discontinuity at the origin. Strictly, the variogram value should be equal to zero when the lag distance is zero; however, due to small-scale variations and measurement variability, there will be specific variogram value at h = 0.

$$\gamma(h) = nC_0 \text{ , when h = 0} \quad (3.20)$$

Where, $\gamma(h)$ is the variogram value

h is the lag distance

C_0 is the Nugget

3.7.2.2 Linear Model

A simple linear model is the basic model to depict the spatial variability. In the linear variogram model, the variogram value increases linearly from the origin and reaches a constant value at the sill. Its variogram function is given by:

$$\gamma(h) = 0, \text{ when h = 0} \quad (3.21)$$

$$\gamma(h) = nC_0 + ph, \text{ when h > 0} \quad (3.21a)$$

Where, $\gamma(h)$ is the variogram value

h is the lag distance

C_0 is the Nugget

P is the slope of the line

3.7.2.3 Spherical Model

Spherical model is another standard model used to capture the spatial variability present in the variable of interest. The spherical model increases linearly from origin and reaches the sill value with the normal transition. The spherical variogram function is given by:

$$\gamma(h) = 0, \text{ when h = 0} \quad (3.22)$$

$$\gamma(h) = C_0 + C\left[\frac{3h}{2a} - \frac{h^3}{2a^3}\right], \text{ when 0 < h < a} \quad (3.22a)$$

Where, $\gamma(h)$ is the variogram value; h is the lag distance; C_0 is the Nugget; C is the scale of the variogram; a is the range of the variogram

3.7.2.4 Exponential Model

Exponential model is modelled using a specific function where the model reaches the sill asymptotically. Theoretically, the exponential model never reaches the sill value. However, the range using exponential model is defined as the lag at which the exponential variogram function reaches 95% of the sill value. The exponential function is given by:

$$\gamma(h) = 0, \text{ when h = 0} \quad (3.23)$$

$$\gamma(h) = C_0 + C\left[1 - exp\left(-\frac{h}{a}\right)\right] \; when \; h > 0 \quad (3.23a)$$

Where, $\gamma(h)$ is the variogram value; h is the lag distance; C_0 is the Nugget; C is the scale of the variogram;

a is the range of the variogram

3.7.2.5 Gaussian Model

Gaussian Model is similar to the exponential model, where the model reaches the sill asymptotically. The Gaussian model represents smoothly varying properties at the origin, with parabolic behavior at the origin. However, when the Gaussian model is used

without nugget effect, it can lead to numerical instabilities. The Gaussian function is given by:

$$\gamma(h) = 0, \text{ when h = 0} \quad (3.24)$$

$$\gamma(h) = C_0 + C\left[1 - exp\left(-\frac{h^2}{a^2}\right)\right] \; when \; h > 0 \quad (3.24a)$$

Where, $\gamma(h)$ is the variogram value; h is the lag distance; C_0 is the Nugget; C is the scale of the variogram; a is the range of the variogram

In this research study, all the experimental variograms were modelled using standard variogram models. All the models consisted of a finite sill, as shown in Figure 3-8; however, in practice, the variogram values continued to increase beyond the sill value. One has to be very careful, as every individual model presents different prediction values.

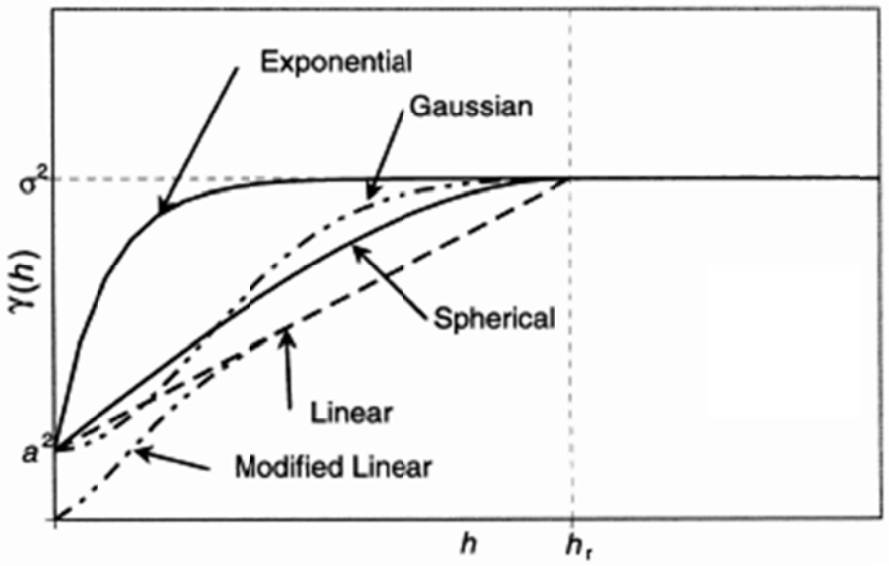


Figure 3-8 Variogram models (Jones et al., 2002)

3.7.2.6 Anisotropy

Anisotropy, by definition, refers to the soil properties in different directions. In geostatistics, anisotropy refers to the different spatial correlation structure in different directions. All the variogram models that were discussed in the earlier sections have

assumed the same spatial correlation structure. However, in practice, the properties can show the different spatial structure in different directions. One way to model anisotropy is to use geometric anisotropy, where the variogram value reaches the same sill in all directions. The most common approach to modelling geometric anisotropy is by finding the ranges (a_x, a_y, a_z) in three principal orthogonal directions and calculating the isotropic lag distance using the equation below (Bohling, 2005):

$$h = \sqrt{(h_x/a_x)^2 + \left(h_y/a_y\right)^2 + (h_z/a_z)^2} \quad (3.25)$$

Where, h_x is the lag distance in x- direction

h_y is the lag distance in y- direction

h_z is the lag distance in z- direction

a_x, a_y, a_z are the ranges in three principal directions

In this research study, the Surfer software was used to construct an experimental variogram and model the spatial variability. The geometric isotropy is a built-in option in the software, and was used to incorporate the anisotropy while modelling the variogram.

3.7.3 Kriging Analysis

This research study is aimed at incorporating the spatial variability present in the soil properties for prediction analysis. Surfer, commercially available software, was utilized for performing the prediction analysis, using kriging. Kriging is an estimation method that gives the best unbiased linear estimates to predict the values at unsampled locations (Armstrong, 1994). This is because of its ability to reduce the error variance of the predicted values.

Once the spatial variability is modelled using the accessible variogram models, kriging analysis is then performed utilizing the surfer program. The output of the kriging analysis is presented in the form of a contour map, where the estimates at unsampled

locations can be found, using a digitizing method. Before performing the kriging analysis, a grid should be created (this grid is different from the grid selected for constructing an experimental variogram) in x-direction and y-direction, as shown below in Figure 3-9.

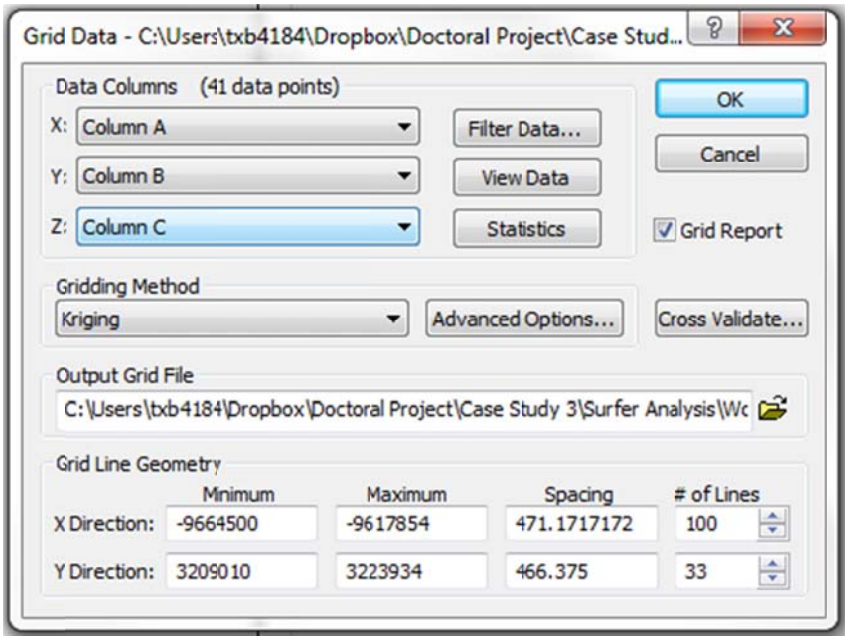


Figure 3-9 Grid selection using Surfer software

The Surfer software provides an option for performing advanced kriging analysis by incorporating the spatial variability model, as shown in Figure 3-10. If the model is not selected, the kriging analysis will be performed based on the default linear variogram.

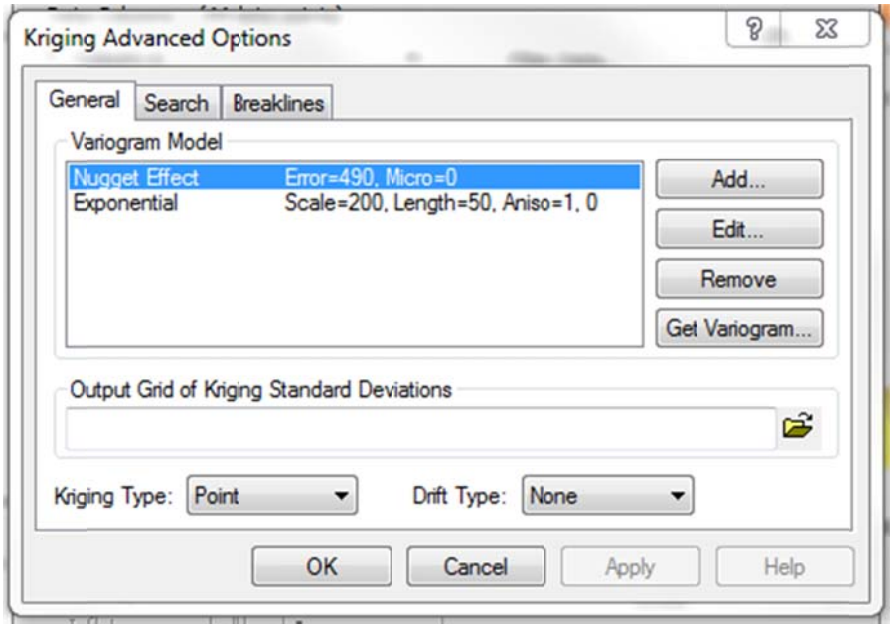


Figure 3-10 Selection of Variogram Model using surfer software

3.7.8 Cross Validation

Cross validation is a validation technique used to assess efficacy of the models developed using statistics. In geostatistics, cross validation refers to assessing and evaluating the estimation methods used for prediction analysis. In this research study, a cross validation exercise was performed to compare the true and estimated values, using only the information available in our sample data set.

In order to perform the cross validation, a few observations were deleted from the actual data set. Kriging analysis was performed using the new data set and contour maps that were generated. The estimated values were compared to the true sample values that were initially removed from the sample data set. A graphical plot was drawn to compare the true and estimated values, as shown in Figure 3-11. These values were plotted against the 45 degree line. If the true and estimated values are equal, the point will exactly fall on the straight line. Any deviation from the 45 degree straight line reflects the

error between the estimated and true value. Through this method, we can also judge whether the kriging method overestimates or underestimates the data.

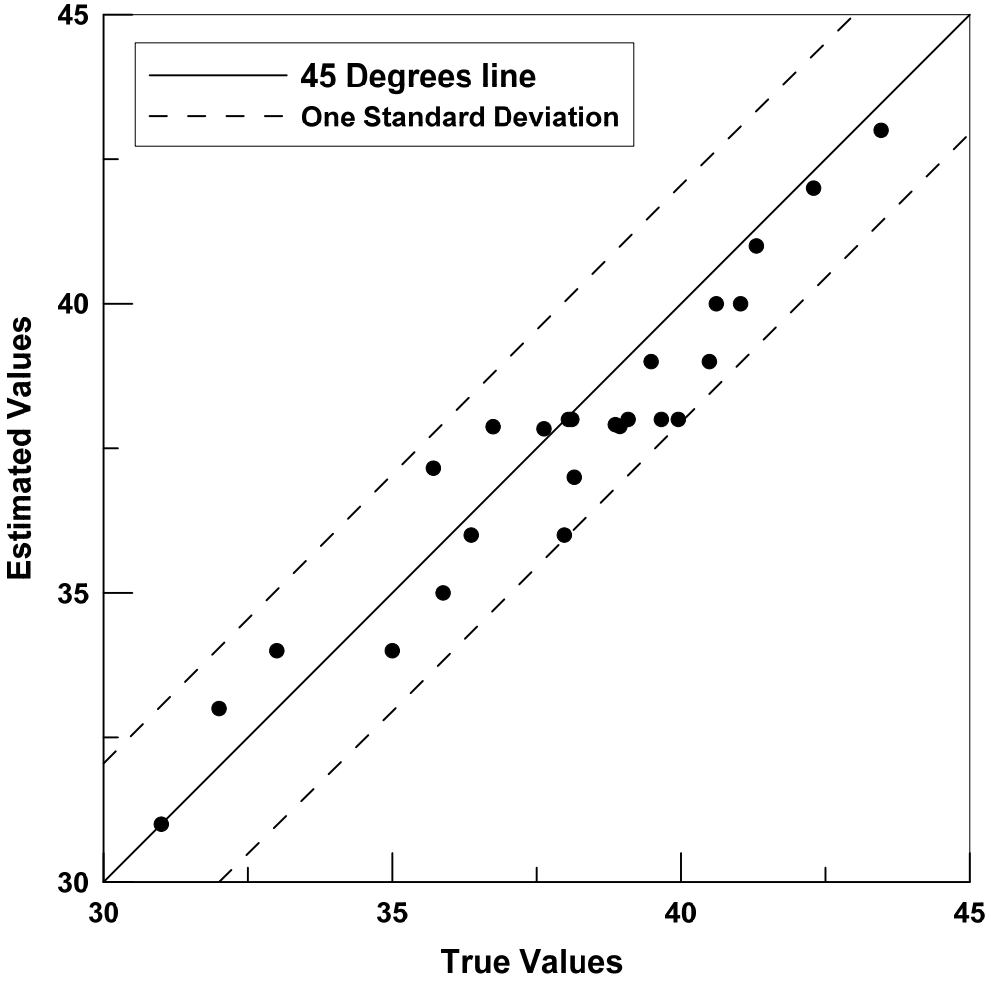


Figure 3-11 Cross Validation plot

3.7.9 Formulated guideline for Spatial Variability Analysis

The formulated guideline for conducting spatial variability analysis on soil properties is given below:

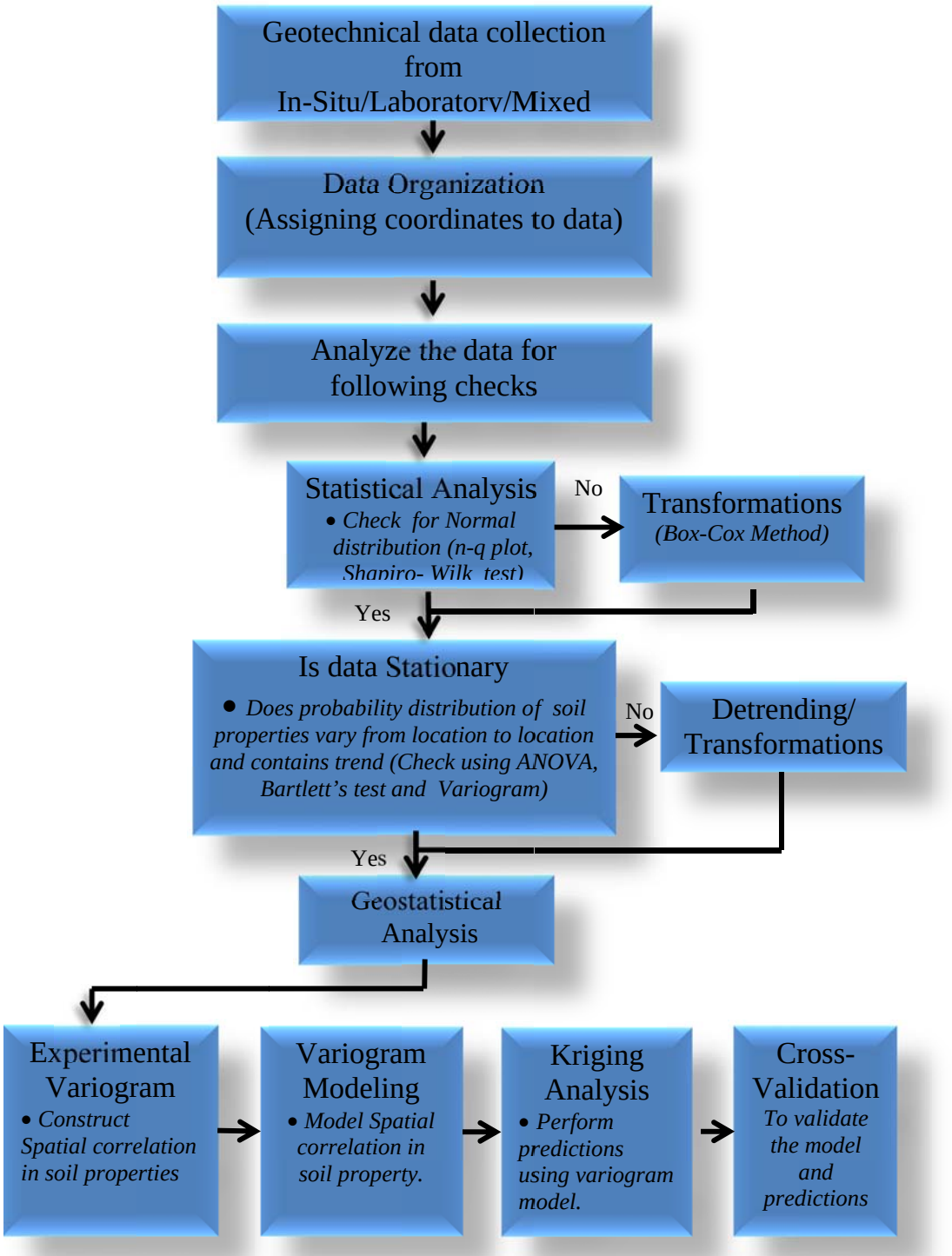


Figure 3-12 Formulated flow chart for Spatial Variability Analysis

3.8 Applications of spatial variability analysis flow chart

The spatial variability analysis flow chart provided earlier was utilized in this research study to address the variability present in the geotechnical engineering projects. The details of types of variability associated with soil properties are discussed in Chapter 2. However, the process of soil formation governs the variability present in the in-situ soils. Therefore, in this research, the spatial variability analysis was performed on three different types of materials which were formed under different conditions.

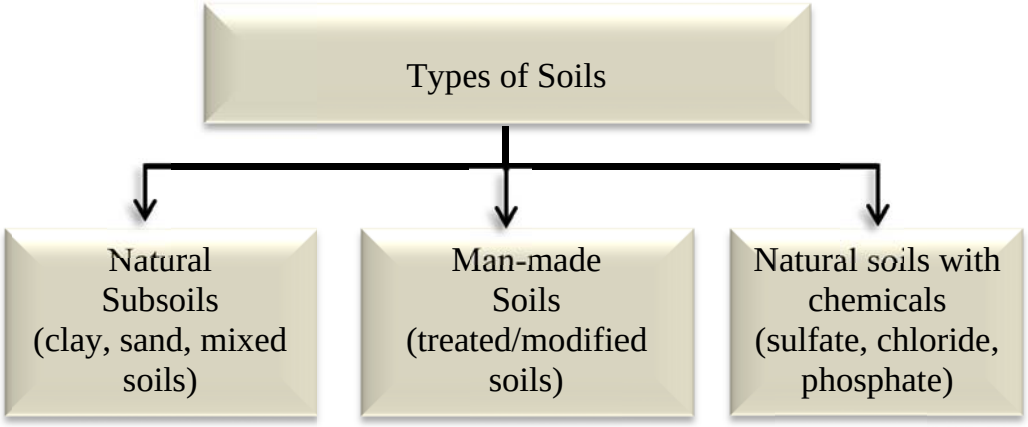


Figure 3-13 Types of materials associated with geotechntical engineering projects

The variability present in the natural soils is mainly governed by different processes involved in soil formation. Due to different geological processes, the soils are mostly heterogeneous in nature. The characterization of soil properties in heterogeneous soil is difficult to ascertain over the entire area with limited borehole data. Through this analysis, the spatial correlation of the known data can be established, using variograms. The variogram models will be utilized along with kriging to predict the soil properties over the entire area.

The man-made materials possess low variability when compared with the natural heterogeneous soils. This is due to the control over the materials and constituents of the

mixture. However, when the man-made materials are subjected to field conditions, it is difficult to assess the properties with time. Through the developed framework analysis, the spatial variability of in-situ properties can be modeled and the predictions can be performed at untested locations. The framework developed can serve in this case as a quality assurance and quality control tool.

The other application of this analysis is performed on the deposition of natural mineral pockets. The natural deposition of the minerals is governed by several factors such as subsurface conditions, ground water flow, and climatic conditions. With high scale variability in proximate areas, engineers often face challenges. Through this framework, the spatial variability in the mineral depositions can be captured, and the spatial variability models will be utilized to evaluate the distribution of high density zones.

3.9 Summary

The incorporation of spatial variability in the prediction analysis is often challenging in geotechnical engineering. In this chapter, a framework was developed to incorporate the spatial variability of soil properties into prediction analysis. The framework was developed by utilizing the concepts from univariate statistics and randomized random variable theory. The framework was broadly divided into three steps: geotechnical data collection, statistical analysis, and geostatistical analysis. The geotechnical data was comprised of in-situ soil properties for three different types of materials, as shown in Figure 3-13. The geostatistical analysis was performed using commercially available Surfer software. Figure 3-12 represents the flowchart of the analysis that was used in this study for developing spatial variability models and performing prediction analysis using geostatistics. The developed framework is evaluated and assessed in different geotechnical problems of interest in the following chapters.

Chapter 4

Spatial Variability Analysis of Natural Soil Properties Evaluated from Cone Penetration Test Data (Cptu)

4.1 Introduction

Cone penetration testing is the most extensively used test in geotechnical engineering for evaluating the subsurface profile. From the day of its invention in the 1930's to date, several types of cone penetrometers have been introduced (Mayne, 2007). Studies were conducted in developing correlations with the CPTU test parameters: tip resistance, shaft resistance, and pore water pressure. In the recent decade, researchers have focused on evaluating the variability of soil parameters obtained from the CPTU test due to its continuous profiling.

Kulhawy and Trautmann (1996) signified the sources of variability in CPTU test results where the type of cone (MCPTU or ECPTU) affects the test results significantly. Orchant et al. (1988) and Kulhawy et al. (1996 provided the summary of inherent measurement errors, with the CPTU test ranging from 5-25 (%) based upon the type of CPTU test. Phoon et al. (1995) described the variability of CPTU test parameters using coefficient of variation, where the tip resistance of a sandy soil varies from 10- 81 percent, silty clay layer from 5-40 percent and clayey soil from 2-17 percent.

Studies were conducted to address the variability in CPTU test parameters using the scale of fluctuation. Hegazy et al. (1996) provided the spatial correlation in the CPTU parameters using the auto-covariance distance (r_0) for different types of soils, where the maximum spatial correlation of tip resistance and sleeve friction is observed in the clayey soils ranging from 3.05 to 4.57 m and minimum correlation distance in mixed soils ranging from 0.34 to0.37m. So, the soil properties evaluated from the CPTU test results are inconclusive due to the variability.

In this research study, an attempt was made to incorporate the spatial variability present into the soil properties evaluated from the CPTU test data. Since, strength parameters of the soil layers govern the performance of a structure, spatial variability of a strength parameter was evaluated in this study. The framework developed and described in Chapter 3 was used to perform the spatial variability analysis, using univariate statistics along with geostatistics.

4.2 Data Acquisition

Data collection is the primary step for any statistical analysis. In this study, the CPTU test profiles were obtained from a soil exploration program conducted in China. The cone penetration tests were conducted at different locations to study the subsurface soils for the Suqian-Xinyi highway project. The layout of the CPTU bore hole locations are presented in the Figure 4-1, and the basic information is summarized in Table 4-1.

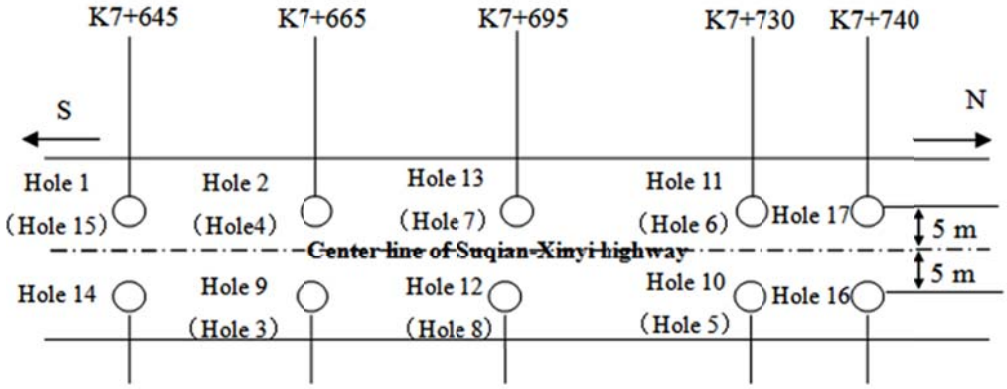


Figure 4-1 Layout of CPTU soundings in Suqian-Xinyi highway

Table 4-1 Basic information of CPTU soundings in Suqian-Xinyi highway

No.	Chainage Number	Penetration Depth (m)
Hole 1	K7+645	20.75
Hole 2	K7+665	21.00
Hole 9	K7+665	21.00
Hole 10	K7+730	20.00

Table 4.1 - *Continued*

Hole 11	K7+730	19.00
Hole 12	K7+695	21.00
Hole 13	K7+695	21.25
Hole 14	K7+645	20.00
Hole 16	K7+740	20.00
Hole 17	K7+740	20.75

In total, the CPTU test was conducted at 10 different locations in an area of 3100 square feet. Basic soil classification was performed on all the CPTU test data, using the Robertson et al. (1986) chart. The subsurface profile was predominantly composed of silty sand layers with clay lenses. The idealized soil profiles for all the CPTU test data is presented in Tables 4-2 to 4-12. The detailed CPTU profiles are provided in Appendix A. In order to conduct spatial variability analysis, the friction angle parameter was selected due to the dominance of silty sand layers throughout the subsurface profile.

Table 4-2 Idealized soil profile for CPTU data at bore hole-1

Cone Penetration Test at Hole 1 : k7+645						
Layer	Soil	From	To	Thickness (m)	Tip Resistance (MPa)	Friction Ratio (%)
1	SM- MS	0	2.2	2.2	3	1.5
2	MS-ML	2.2	5.2	3	1.9	0.9
3	MS-ML	5.2	8.3	3.1	4.4	1.2
4	SP-SM	8.3	10	1.7	7.8	1.4
5	SM- MS	10	11.6	1.6	5.3	1.4
6	SM- MS	11.6	14.3	2.7	6.2	1.4
7	SP-SM	14.3	16	1.7	9.2	1.6
8	SM- MS	16	19	3	7.8	1.8

Table 4-3 Idealized soil profile for CPTU data at bore hole-2

Cone Penetration Test at Hole 2 : k7+665						
Layer	Soil	From	To	Thickness (m)	Tip resistance (MPa)	Friction Ratio (%)
1	SM-MS	0	1	1	1	1.8
2	Clay	1	1.4	0.4	1.1	4
3	SM-MS	1.4	4.6	3.2	4	1.1
4	SP-SM	4.6	7.6	3	7	1.1
5	MS-ML	7.6	9.6	2	3.5	2.8
6	SP-SM	9.6	11.8	2.2	6.5	1.2
7	SP-SM	11.8	21	9.2	9	1.3

Table 4-4 Idealized soil profile for CPTU data at bore hole-9

Cone Penetration Test at Hole 9 : k7+665						
Layer	Soil	From	To	Thickness (m)	Tip resistance (MPa)	Friction Ratio (%)
1	ML	0	1.3	1.3	2.2	2.8
2	Clay/Org	1.3	1.6	0.3	0.3	13.5
3	ML	1.6	3	1.4	2.4	1.8
4	1	3	5	2	0.6	1.1
5	SM-MS	5	7.3	2.3	4.6	1.2
6	ML	7.3	11.6	4.3	2	1
7	SM-MS	11.6	21	9.4	3.8	1

Table 4-5 Idealized soil profile for CPTU data at bore hole-10

Cone Penetration Test at Hole 10 : k7+730						
Layer	Soil	From	To	Thickness (m)	Tip resistance (MPa)	Friction Ratio (%)
1	SM-MS	0	1	1	3.5	1.6
2	SM-MS	1	2.5	1.5	6	1.2
3	SM-MS	2.5	4.1	1.6	3.6	1
4	ML	4.1	6.6	2.5	2	0.76
5	SM-MS	6.6	12.5	5.9	3.3	0.9
6	SM-MS	12.5	21	8.5	5.5	1.4

Table 4-6 Idealized soil profile for CPTU data at bore hole-11

Cone Penetration Test at Hole 11 : k7+730						
Layer	Soil	From	To	Thickness (m)	Tip resistance (MPA)	Friction Ratio
1	SM-MS	0	5.2	5.2	4.5	1.3
2	SM-MS	5.2	7.4	2.2	5.4	1
3	SP-SM	7.4	9.2	1.8	8	1.1
4	Clay	9.2	9.7	1.3	1.2	2
5	SM-MS	9.7	20	10.3	8.5	1.5

Table 4-7 Idealized soil profile for CPTU data at bore hole-12

Cone Penetration Test at Hole 12 : k7+695						
Layer	Soil	From	To	Thickness (m)	Tip resistance (MPa)	Friction Ratio
1	SM-MS	0	0.5	0.5	6.4	1.6
2	ML	0.5	3.4	2.9	2.8	1.5
3	ML	3.4	6.5	3.1	1.9	0.7
4	SM-MS	6.5	10	3.5	4.6	1
5	SM-MS	10	12.2	2.2	2.6	0.9
6	SM-MS	12.2	21	8.8	5	1.2

Table 4-8 Idealized soil profile for CPTU data at bore hole-13

Cone Penetration Test at Hole 13 : k7+695						
Layer	Soil	From	To	Thickness (m)	Tip resistance (MPa)	Friction Ratio (%)
1	SM-MS	0	5.3	5.3	4	1.2
2	SP-SM	5.3	7.3	2	9.5	1
3	SM-MS	7.3	9.4	2.1	5.5	1.5
4	SM-MS	9.4	11.7	2.3	6.5	1.4
5	SM-MS	11.7	15.7	4	5.3	1.4
6	SP-SM	15.7	17.8	2.1	12	1.3
7	SM-MS	17.8	19.5	3.2	9	1.8
8	SM-MS	19.5	21	1.5	7.8	1.8

Table 4-9 Idealized soil profile for CPTU data at bore hole-14

Cone Penetration Test at Hole 14 : k7+645						
Layer	Soil	From	To	Thickness (m)	Tip resistance (MPa)	Friction Ratio (%)
1	ML	0	2	2	2.8	1.9
2	Clay	2	5.1	3.1	1.4	13
3	ML	5.1	8.2	3.1	3	1.7
4	SM-MS	8.2	9.6	1.4	5.4	1.3
5	SM-MS	9.6	11.5	1.9	5	1.5

Table 4-10 Idealized soil profile for CPTU data at bore hole-16

Cone Penetration Test at Hole 16 : k7+740						
Layer	Soil	From	To	Thickness (m)	Tip resistance (MPa)	Friction Ratio (%)
1	SM-MS	0	1.9	1.9	4.9	1.2
2	SM-MS	1.9	3.5	1.6	5.2	1
3	Clay	3.5	6.1	2.6	3.3	2.8
4	Clay	6.1	8.2	2.1	0.4	2.6
5	SM-MS	8.2	10.8	2.6	5.8	1.1
6	SP-SM	10.8	21	10.2	6	0.9

Table 4-11 Idealized soil profile for CPTU data at bore hole-17

Cone Penetration Test at Hole 17 : k7+740						
Layer	Soil	From	To	Thickness (m)	Tip resistance (MPa)	Friction Ratio (%)
1	SM-MS	0	2.5	2.5	5	1
2	ML	2.5	4.5	2	1.7	1.8
3	SM-MS	4.5	5.5	1	3.8	0.9
4	Clay	5.5	7.6	2.1	0.7	2.2
5	SM-MS	7.6	10.5	2.9	4.7	1.2
6	SP-SM	10.5	16.4	5.9	8	1
7	Clau	16.4	17.9	1.5	2.5	2.8
8	SP-SM	17.9	19.2	1.3	9.5	1.1

In order to perform spatial variability analysis of strength property, the friction angle values were determined, using the expression proposed by Kulhawy and Mayne (1990).

$$\phi' = 17.6^{\circ} + 11.0^{\circ} X \log(q_{t1}) \quad (4.1)$$

Where, q_{t1} $is\ normalized\ tip\ resistance\ given\ by$

$$q_{t1} = (q_t/\sigma_{atm})/((\sigma'_{v0}/\sigma_{atm})^{0.5}) \quad (4.2)$$

Where, σ_{atm} = 1 atm = 1 bar = 100 kPa ~ 1 tsf ~ 14.7 psi

σ'_{v0} = effective vertical overburden stress at corresponding depth

Thus, the friction angle values for all the available test data were calculated using the above expressions.

4.3 Data Organization

Due to the limitations involved in performing 3-dimensional variograms, only 5 of 10 bore holes were selected to analyze the spatial variability. The CPTU test data obtained from bore holes 1, 2, 11, 13 and 17 were used. As mentioned in Chapter 3, the data had to be organized in Cartesian coordinate system in order to perform spatial

variability analysis. In order to assign coordinates to the data, certain assumptions were made regarding the location of the bore holes. Bore hole 1 was assumed to be located at the (5,0) in (x,y) coordinate, where 5 represents 5 meters in horizontal direction, and 0 represents depth of the subsurface in vertical direction. Corresponding coordinates for bore hole 1 were obtained for rest of the bore holes. Figure 4-2 provides the schematic representation and coordinates of the bore holes considered for analysis.

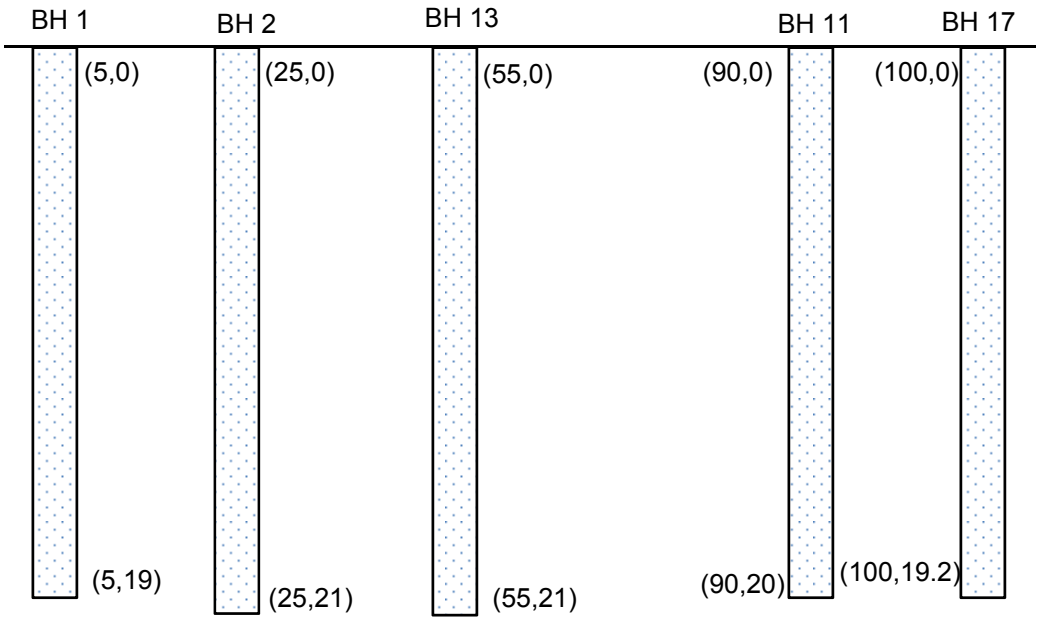


Figure 4-2 Bore hole layout with coordinates in Cartesian system

The friction angle values were calculated and represented in z-direction. For example, a set of coordinates (5, 15, 35) represents a friction angle of 35 degrees at a location 5 feet in horizontal direction and 10 feet in vertical direction. All the friction angle values with their corresponding coordinates in the Cartesian system are provided in Appendix A.

4.4 Statistical Analysis

Statistical analysis involves evaluating the distribution of friction angle values and checks, analyzing spatial variability in the data. The friction angle values were calculated using Equation 4-1 for all the tip resistance values. However, it is cumbersome to involve all the values to perform statistical tests; therefore, the random number generator tool was utilized in this research to randomly select the values of the friction angle in all the CPTU tests. The minimum and maximum values in a respective bore hole were given as lower and upper bound to generate random numbers. The table below provides the friction angle values generated using a random number generator that is used in statistical analysis.

Table 4-12 Friction angle values generated using a random number generator

Bore Hole	Friction Angle (Degrees)									
CPTU 1	38	38	43	38	35	41	38	33	35	39
CPTU 2	29	41	42	38	34	33	41	44	42	44
CPTU 11	37	33	44	40	42	33	36	38	42	43
CPTU 13	29	42	33	39	34	39	43	31	41	36
CPTU 17	46	30	43	31	36	46	33	38	36	39

4.4.1 Histograms

Histograms are plotted to observe the distribution of the friction angle values and skewness present in the data. The number of class intervals (CI) for plotting the histogram is found using Strugger's equation

$$No.of\ Class\ Intervals = 1 + 3.3 \log_{10} N = 6.6$$

Where, N is the total number of observations in the data set = 50

The bin size is found out using below equation:

$$Bin\ Size = \frac{Maximum\ Value - Minimum\ Value}{Number\ of\ Class\ Intervals} = 3.5$$

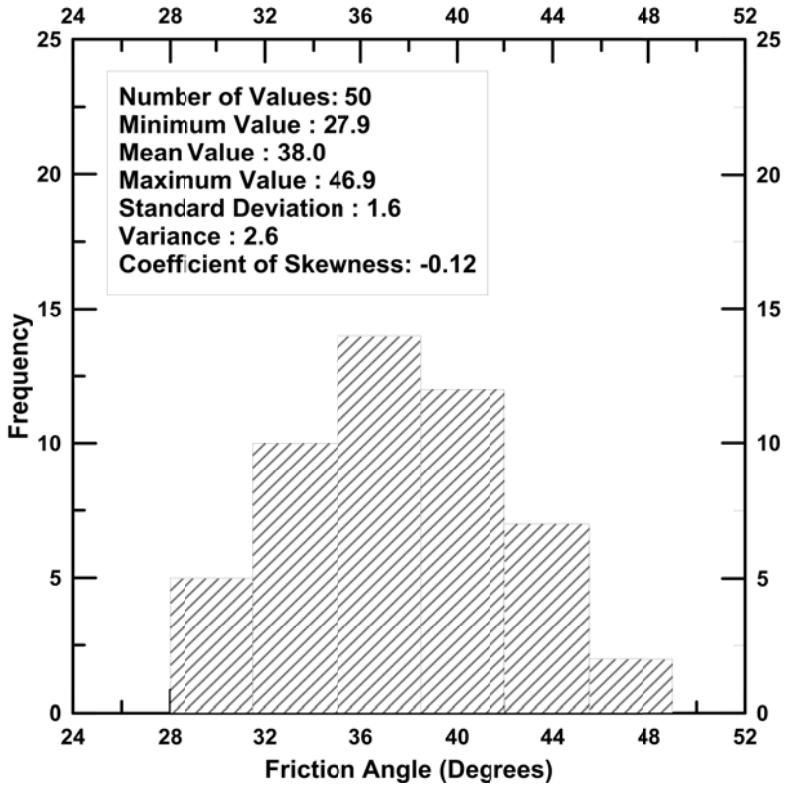


Figure 4-3 Histogram for Friction Angle values

From the above histogram plot, it can be inferred that the friction angle values are more or less distributed normally. The preliminary perspective of the distribution can

be evaluated using Gaussian or normal distribution. The coefficient of skewness of the friction angle values resulted in negative value, indicating that the data are skewed to one side.

4.4.2 Check for Gaussian distribution

The earlier histogram plot provides a visual overview of the distribution of the friction angle data. As most of the statistical tests rely on the normal distribution of the data, in this section two different approaches were used to find the distribution of the friction angle values.

4.4.2.1 Normal – Quantile plot

Normal-Quantile plots are the simplest way to evaluate how well the data fits a Gaussian distribution. The normality values of all the friction angle values were calculated using the expression stated in Chapter 3 and plotted against the corresponding friction angle values. Figure 4-4 presents the normal-quantile plot for the friction angle values, where the sample quantiles are plotted on the y-axis and theoretical quantile values are plotted on x-axis. From the figure, it can be inferred that the values fall follows a straight trend line. Using the regression approach, a best fit straight line was plotted for the data. The coefficient of determination (r^2), which determines the fit of a straight line to the data, was observed at 0.85. The r^2 value obtained and the data points close to the straight line depict that the values were normally distributed.

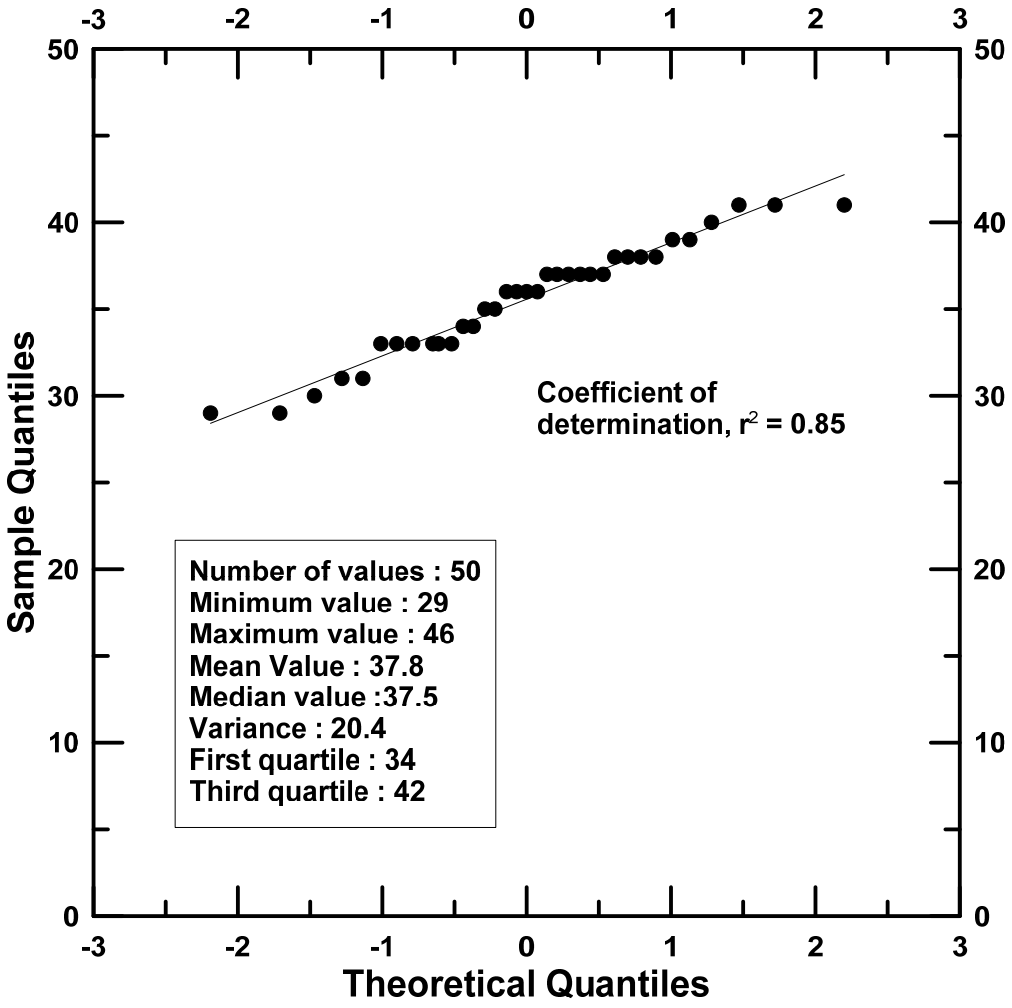


Figure 4-4 Normal-Quantile for Friction Angle values

4.4.2.2 Shapiro-Wilk test

Shapiro-Wilk test is a more formal approach used to check for the Gaussian distribution of the data. This test was conducted with a significance level alpha = 0.05. In order to conduct this test a hypothesis was developed as shown below:

H_0 : Sample data obtained from normally distributed population

H_1 : Sample data not from a normally distributed population

The above hypothesis was tested using the statistic; $W = b^2 / SS$.

Below table summarizes the Shapiro-Wilk test results:

Table 4-13 Summary of Shapiro-Wilk test

Parameter	**Result**
b^2	950.48
Sum of Squares (SS)	1002.72
Test Statistic ($W = b^2/SS$)	0.947
P-value	0.08
Critical Region	$\alpha = 0.05$
P- value from statistic > significance level	

From the above test results, it is observed that the probability value of the data was greater than the critical region. Hence, the null hypothesis could not be rejected, concluding that the sample data was derived from normally distribution.

4.5 Check for Stationarity in the data

The stationarity is the data is evaluated by conducting two tests: one for checking constant mean and another for checking constant variance in the data. As stated in Chapter 3, these are evaluated using ANOVA method and Bartlett's test.

4.5.1 Check for constant mean

The ANOVA test was used to check for constant mean in all the CPTU tests conducted in bore holes 1, 2, 11, 13 and 17. This was evaluated by constructing a hypothesis, as shown below:

$H_0 : \mu_1 = \mu_2 = \mu_{11} = \mu_{13} = \mu_{17}$

H_1 : At least one mean is different

The statistic that was used to evaluate the above hypothesis was $F_{0,}$ which is the ratio of mean square treatment to mean square error. This was compared with the value at a significance level of 0.05. Table 4-14 below provides the summary of the ANOVA test conducted for checking the constant mean in friction angle in all the bore holes.

Table 4-14 Summary of ANOVA results for checking constant mean

Analysis of Variance				
Source of Variation	Sum of Squares	DOF	Mean Square	F_0
Treatment	30.2	4	7.55	0.35
Error	971.8	45	21.60	
Total	1002	49		

From the above results, it can be inferred that the statistic was smaller than the critical region. i.e., f_0 (0.35) < $f_{critical}$ (2.57). Therefore, we did not reject the null hypothesis developed, concluding that the means of friction angle was constant in all the data. However to validate the conclusion, the assumptions of the model used to develop the ANOVA test had to be satisfied.

4.5.1.1 Model Adequacy check

The ANOVA test was developed based on the assumption that the error residual values were normally distributed and the residual variances in all the bore holes were constant. The residuals were calculated using Equation 3.8. The fitted value in Equation 3.8 was obtained by determining the mean of the friction angle values in the individual bore hole. Table 4.15 provides the residual values which were used for testing the model adequacy of the ANOVA model. Figures 4.5 and 4.6 depict the plots for checking the normality in the residual values and variance distribution.

Table 4-15 Residual values for the friction angle values

Bore hole	Residuals									
	1	2	3	4	5	6	7	8	9	10
CPTU 1	0.10	1.10	5.10	0.10	-2.90	3.10	0.10	-4.90	-2.90	1.10
CPTU 2	-9.80	2.20	3.20	-0.80	-4.80	-5.80	2.20	5.20	3.20	5.20
CPTU11	-1.80	-5.80	5.20	1.20	3.20	-5.80	-2.80	-0.80	3.20	4.20
CPTU 13	-7.70	5.30	-3.70	2.30	-2.70	2.30	6.30	-5.70	4.30	-0.70
CPTU 17	8.20	-7.80	5.20	-6.80	-1.80	8.20	-4.80	0.20	-1.80	1.20

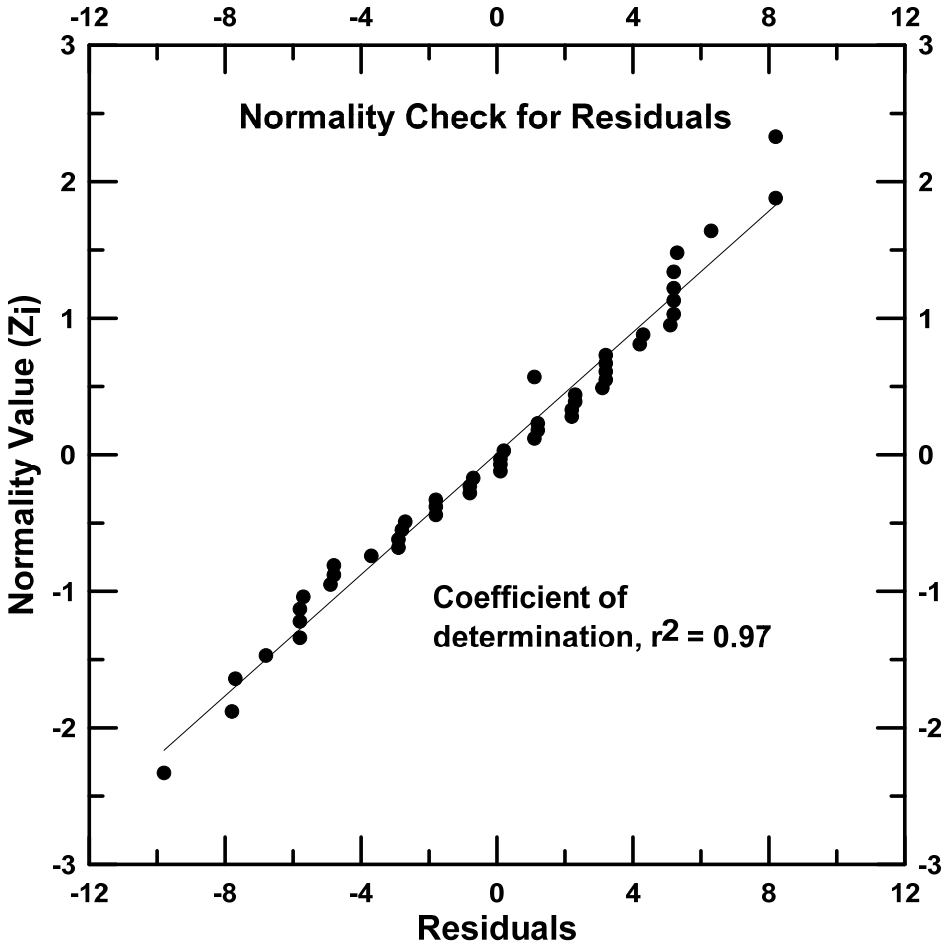


Figure 4-5 Residual normality plot for model adequacy check

From the above plot, it can be inferred that the residuals follow a straight line trend, where the coefficient of determination of the best fit line is observed to be 0.97. This depicts that the residual values are normally distributed.

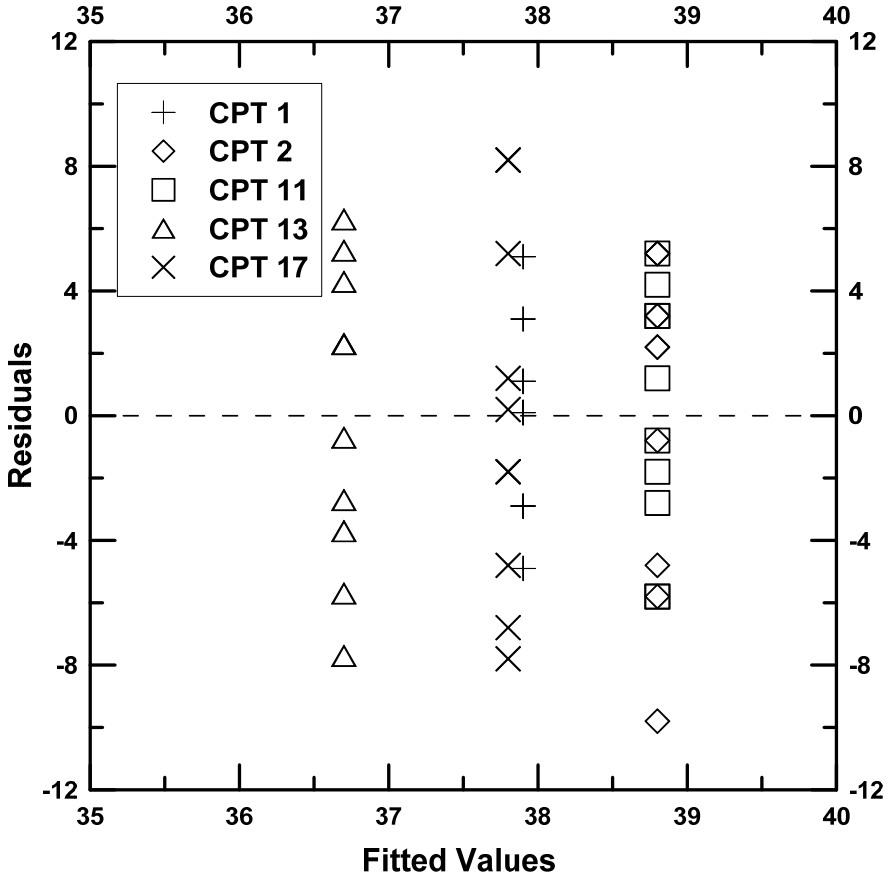


Figure 4-6 Model Adequacy check for constant variance in residual values

From the visual inspection of the above plot, it can be inferred that the residual variance in the friction angle values was constant. Thus, from the figures 4.5 and 4.6, the residual values satisfies the model adequacy check, validating the conclusion obtained from the ANOVA test.

4.5.2 Check for Constant Variance

Bartlett's test was used in evaluating the variation present in residual variance of the friction angle values obtained from all the bore holes. This was performed by developing a hypothesis, as shown below:

$$H_0 : \sigma^2_1 = \sigma^2_2 = \sigma^2_{11} = \sigma^2_{13} = \sigma^2_{17}$$

$$H_1 : \text{At least one variance is different}$$

The hypothesis was evaluated using the statistic expressed in Equation 3.9. The value obtained from the statistic was compared to a significance value 0.05. The table below provides the summary of the results obtained from the Bartlett's test.

Table 4-16 Summary of Bartlett's test results

Parameter	**Result**
Q	1.87
C	1.04
Test Statistic ($\chi_0^2 = 2.3026\ \frac{q}{c}$)	0.947
χ_0^2 = value	4.12
Critical Region	9.49
x^2 (4.12) < x^2_{critical} (9.49)	

Based on the test results, it can be inferred that the statistic value was smaller than the critical region, i.e., x^2 (4.12) < x^2_{critical} (9.49). Therefore, we do not reject the null hypothesis developed, concluding the variances of friction angle were constant in all the data.

Therefore, based on the ANOVA test and Bartlett's test, it was concluded that the friction angle data was stationary, having constant mean and variance.

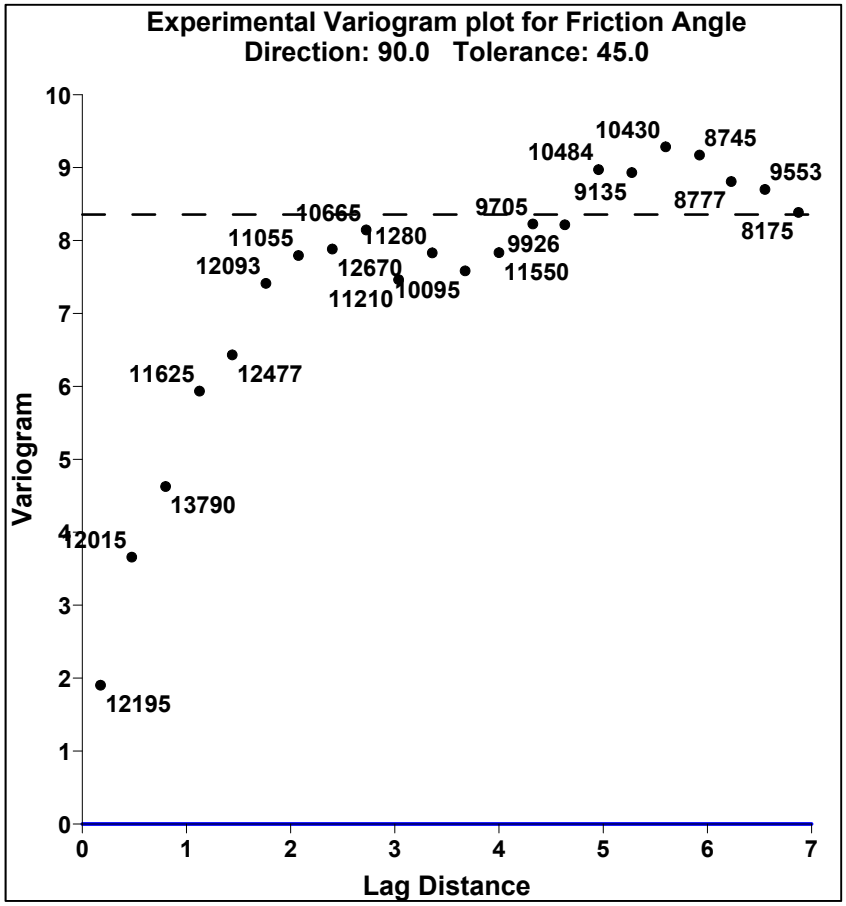
4.6 Geostatistical Analysis

After the statistical analysis and stationarity checks, spatial variability analysis was performed on the friction angle data, as shown in the guideline presented in Chapter 3. In order to perform the spatial variability analysis, the Cartesian coordinate system was given to all the friction angle values, as discussed in Section 4.3. The same coordinate system was followed throughout the geostatistical variability analysis.

4.6.1 Experimental Variogram

Variograms or semi-variograms were used for capturing the spatial variability present in the friction angle data. Figure 4-6 provides the experimental variogram plot for the friction angle data, where lag distance values were plotted on x-axis and variogram values were plotted on the y-axis. The variograms were calculated for different lag distances and different lag widths. The maximum lag distance of 7m, with a lag width of 0.28m provided promising results. An effort was also made to capture the spatial variability of the friction angle data, both horizontally and vertically. However, due to constraints in the CPTU data, only the variability in the vertical direction was captured.

The variogram values for all the lag distances were calculated using Equation 3.8. The direction of 90 degrees in the experimental variogram plot in Figure 4-6 indicates the variability captured in a vertical direction. A tolerance angle of 45 degrees was provided for calculating the variogram values. The number beside every experimental value indicates the number of pairs that was used to calculate the variogram value. The dashed line in the plot indicates the global variance of the friction angle values. From the plot it is evident, that the variogram values were around the global variance line, indicating that the data was stationary and there were no trends present in the data.



- Estimator type : Variogram
- Maximum lag distance : 7
- Number of lags : 25
- Lag Width : 0.28
- Direction of Variogram : 90 degrees (y-direction)
- Tolerance : 45 degrees

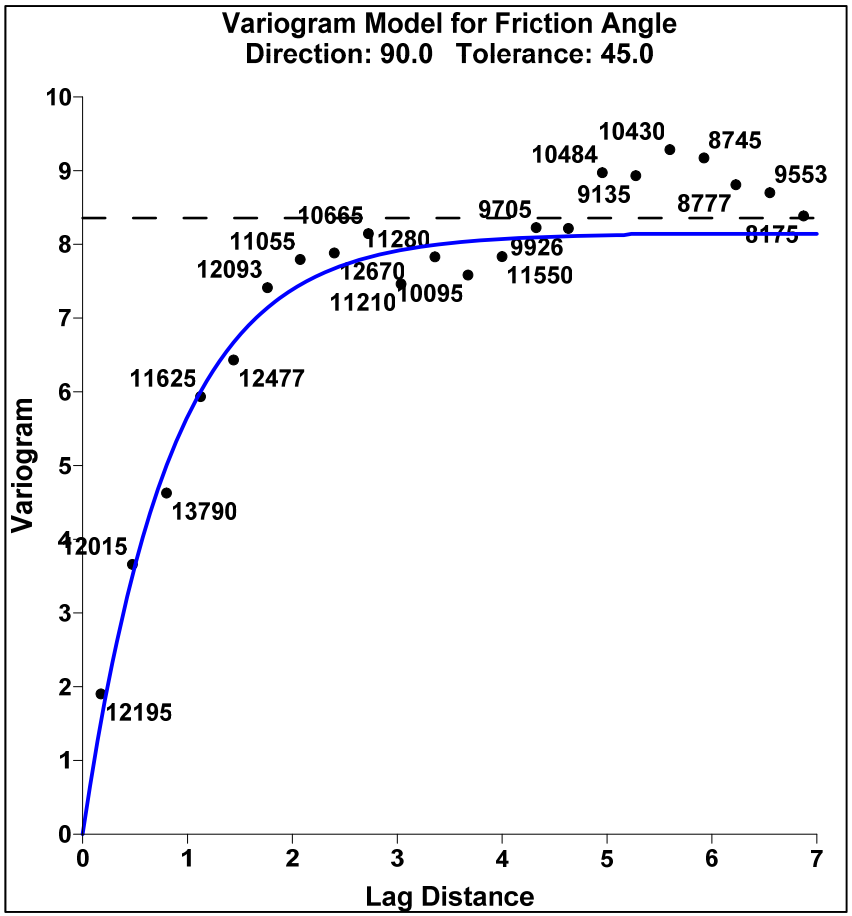
Figure 4-7 Experimental Variogram plot of friction angle data

4.6.2 Variogram Modelling of Friction Angle Data

From the experimental variogram plot in Figure 4-6, it can be observed that a certain spatial correlation existed between the friction angle values. This spatial correlation was depicted by modeling the characteristics of the variogram: range, sill, and nugget. The variogram values appear to have a zero intercept, indicating there wasn't any nugget effect. From Figure 4-6, it is observed that the experimental variogram value increased with an increase in lag distance.

The scale (C) or range (a) was used to model the spatial correlation distance, where the variogram value increased linearly with an increase in lag value to a distance of 0.86m. The effect of the range was observed from a distance of 2.58m (3a) to 4.3m (5a). That means that no spatial correlation existed after the lag distance of 4.3m. The sill indicates the vertical scale on the variogram plot, where the variogram value reached a constant value. In this case, the sill of the variogram plot was observed to be 8.14. By using these characteristics, the experimental variogram was modeled using an exponential model.

Figure 4-7 presents the variogram model for the experimental variogram in Figure 4-6. The blue line indicates the exponential model used to fit the experimental variogram values obtained for a maximum lag distance of 7m. Since the subsurface profile was mainly composed of silty sand layers, through this research, the spatial correlation distance for a silty sand soils in vertical direction can be defined as 2.58m to 4.3m, using the range value.



- **Model : Exponential**
- **Nugget (C_0) : 0**
- **Scale (C_0 +C) : 8.14**
- **Length (A) : 0.86**

Theoretical Equation:

$\gamma(h) = C_0 + C\,[1 - \exp(-h/a)]$ for $h > 0$

Model Equation :

$\gamma(h) = 8.14\,[1 - \exp(-h/0.86)]$ for $h > 0$

Figure 4-8 Variogram Modeling of friction angle data

4.6.3 Kriging Analysis

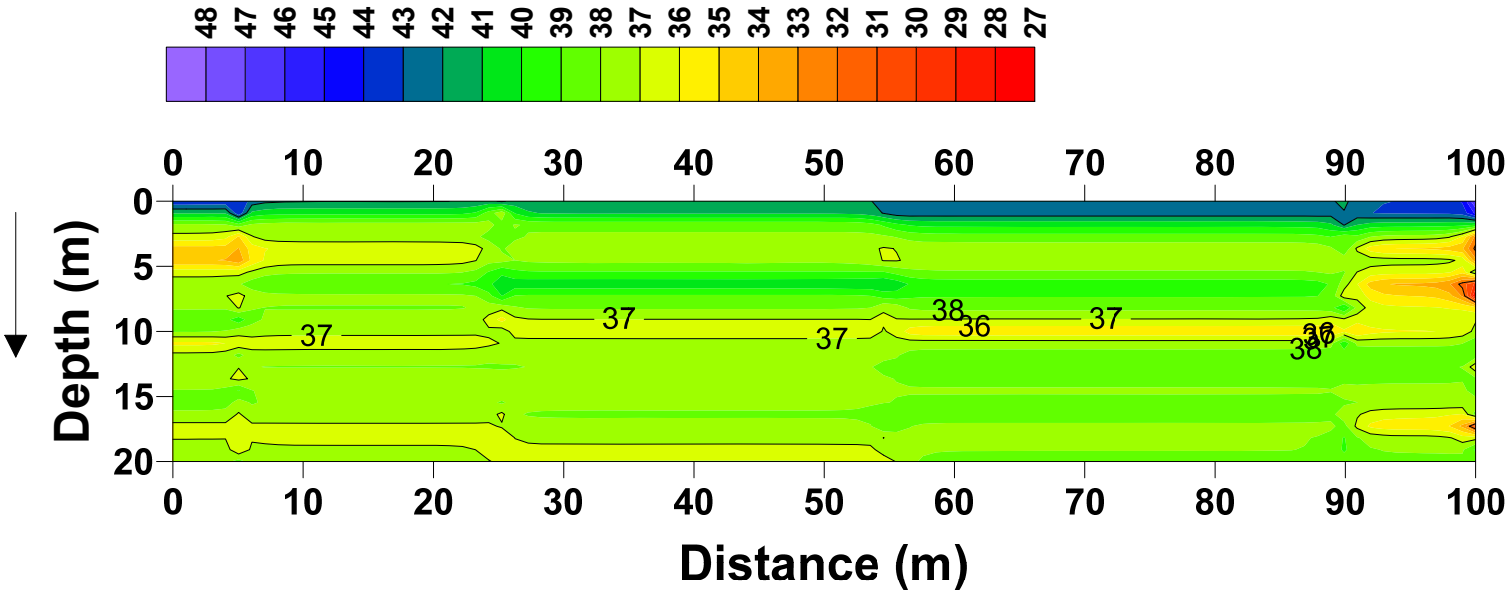
Prediction of soil properties at unsampled locations is often challenging. In this study, predictions of the friction angle values at unsampled locations were performed using kriging analysis. The kriging analysis incorporates the spatial variability in the available data, using the variogram model performed in the previous section. This model was used to calculate the weights of the known values around the unknown values. The output of the kriging analysis is provided in the form of contour maps, where all the friction angle values over the entire area 10m X 95m are mapped. In order to perform kriging analysis, the following grid was selected for plotting the contour map.

Table 4-17 Grid selected for Kriging analysis

Parameters selected	
Maximum lag distance	7
Angular divisions	180
Radial divisions	100

Figure 4-8 provides the contour map of the friction angle values that were obtained by performing kriging analysis. The x-axis on the contour map shows the horizontal distance of the layout, and the y-axis depicts the depth of the subsurface profile. It can be observed that the friction angle varies from 27 degrees to 48 degrees. The top 0.5 to 1 meter of subsurface constitutes a friction angle value of 43 to 48 degrees, below which there existed a layer of 7 to 9.5 meters thick silty sand layer with a friction angle of 6 to 39 degrees. This layer was followed by a thin clayey silt to silty sand

layer for a thickness of 0.5 to 1 meter with a friction angle of 33 to 36 degrees, which is underlain by a thick silty sand layer with a fiction angle value ranging from 35 to 37 degrees.



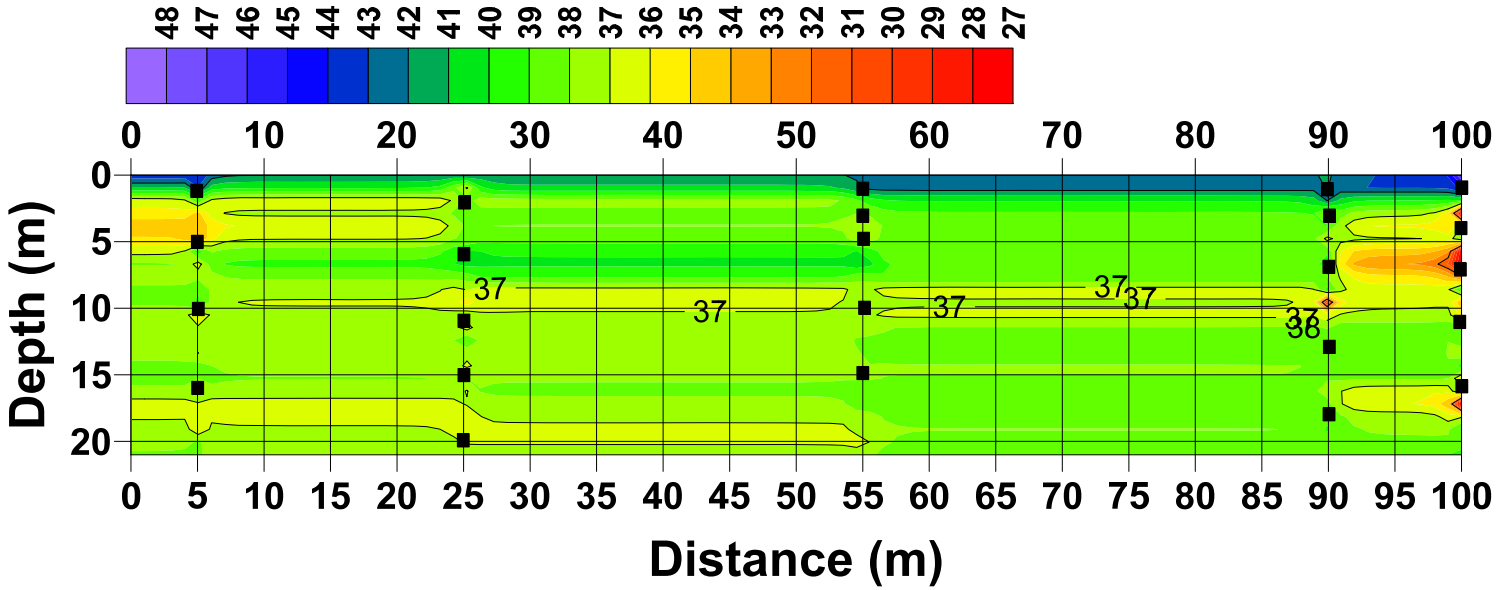
Contour Map of Friction Angle (Degrees)

Figure 4-9 Contour Map of Friction angle values obtained through kriging analysis

4.6.4 Cross Validation

A technique called cross validation was performed to validate the spatial variability model. This technique involves the deletion of the actual values and performing the kriging analysis to generate a new map over the entire area. The grid parameters presented in Table 4-17 and the exponential variability model were used for performing the kriging analysis.

In this study, 24 actual values chosen randomly were deleted from the original data set. The new set of data was used to generate the contour map of friction angle values, using the exponential model. Figure 4-9 represents the contour map of friction angle generated by a new set of data. From the figure, it is evident that the new map and original map depict almost the same range of friction angles over the depth of the subsurface profile. The black rectangular boxes in Figure 4-9 represent the locations where the actual values were deleted.



Contour Map of Friction Angle (Degrees)

Figure 4-10 Cross- Validation map generated using a new set of data

Using the digitize tool available in the Surfer software, the precise value of the friction angle was obtained at the deleted locations. Table 4-18 below provides the actual values and predicted values from the cross-validation map.

Table 4-18 Comparison of friction angle values

S.no	Actual Values	Predicted Values	S.no	Actual Values	Predicted Values
1	42.5	41.8	13	37.1	37.3
2	35.2	32.7	14	37.9	38.4
3	37.8	37.2	15	41.9	42.2
4	37.5	37.1	16	37.8	40.4
5	38.2	38.3	17	38.8	37.6
6	40.0	40.5	18	38.5	38.1
7	37.1	36.9	19	38.6	38.4
8	37.7	38.0	20	37.1	35.0
9	36.9	37.2	21	44.3	43.3
10	41.9	42.2	22	28.6	26.7
11	37.9	40.9	23	37.0	39.0
12	38.0	37.0	24	37.7	38.9

From the above table, it is evident that the predicted values were close to the actual values. Figure 4-10 presents the validation plot for the friction angle values, where actual values were plotted on x-axis and predicted values were plotted on y-axis. If both the predicted and actual values are same, the point exactly falls on the 45 degree line. In this case, all the 24 points were close to the 45 degree line and within 1 standard

deviation. This concludes that the spatial variability model chosen was apt for this study, and the spatial map generated using kriging analysis in Figure 4-8 was appropriate.

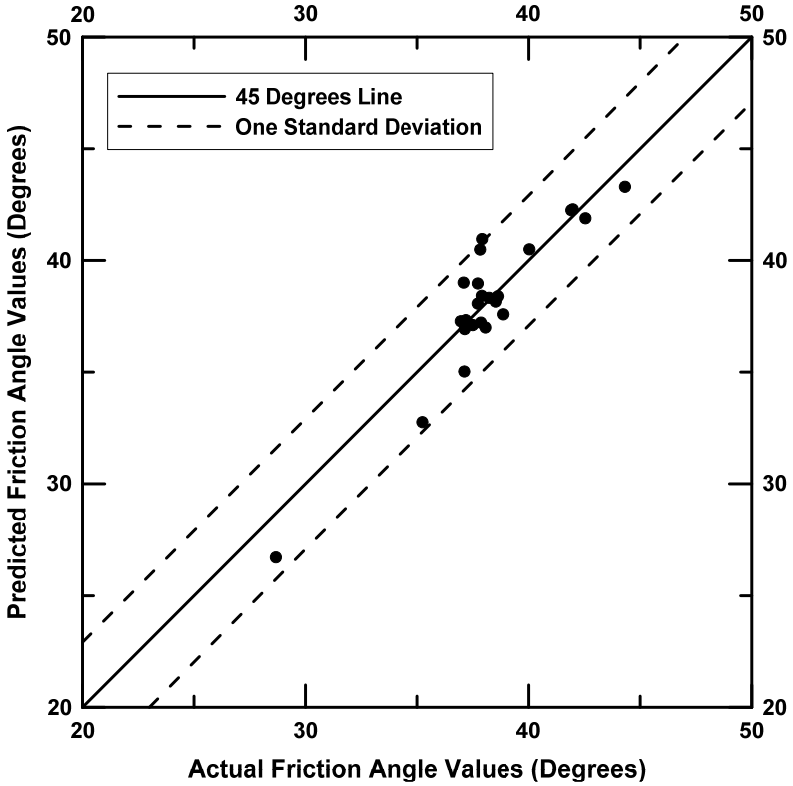


Figure 4-11 Cross- Validation plot for actual and predicted friction angle values

In view of the results above, the spatial variability of the friction angle was further used to optimize the number of bore holes. Considering the bore hole layout presented in Figure 4-1, it can be observed that the CPTU test conducted in bore holes 11 and 17 were close to each other. Therefore, the kriging analysis was performed by deleting the data obtained from CPTU test conducted in bore hole 11. Figure 4-11 presents a new contour map generated by using the spatial variability model developed for silty sand soils and the grid parameters that were presented in Table 4-17.

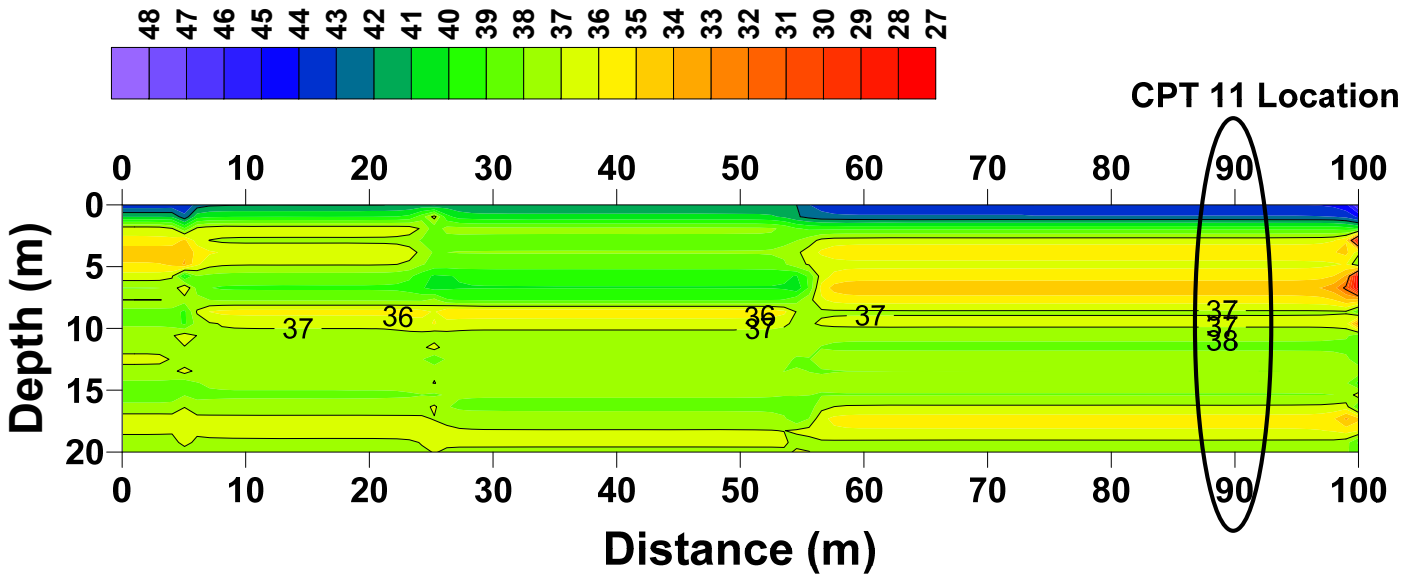


Figure 4-12 Contour Map of friction angle values without CPTU 11 data

The friction angle values at the bore hole 11 location were obtained from the above map, using the digitize tool available in the Surfer software. The validation plots in Figures 4-12 compare the predicted values to the original values. It was observed that the predicted values were close to the 45 degree line, and most of the values were within 1 standard deviation. Also, it can be observed that the predicted values were a little lower than the actual values, which is on the conservative side.

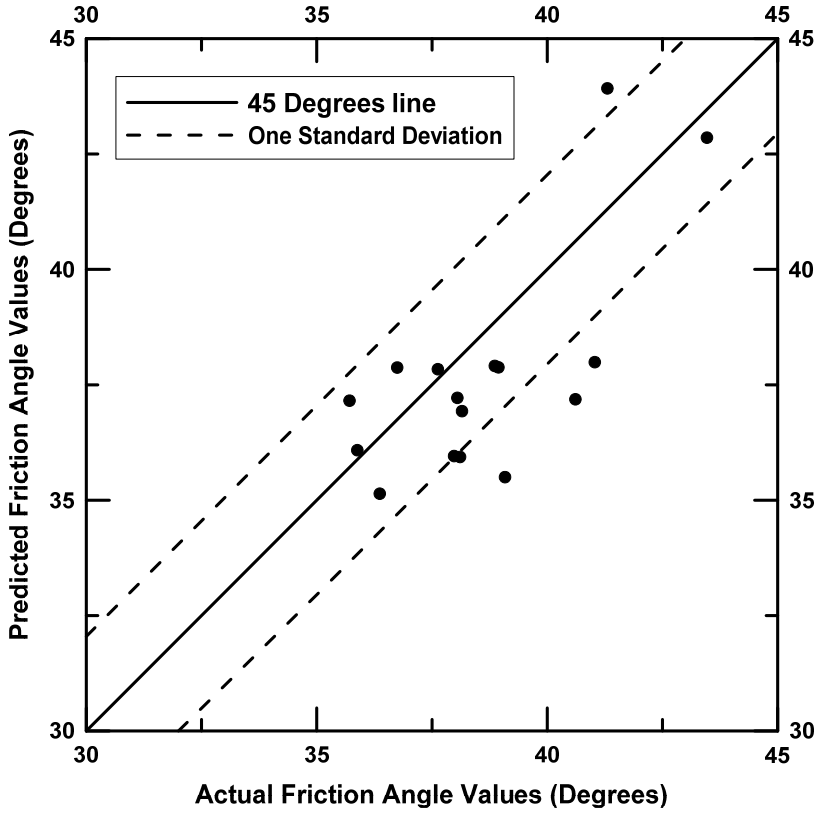


Figure 4-13 Comparison of predicted and actual values of bore hole location 11

4.7 Applications of the spatial variability model

In this section, a hypothetical example is provided where the spatial variability analysis is conducted for the selection of friction angle value. Consider a footing with a width of 10 meters, where the influence of the stress from the footing is distributed to a depth of 20 meters. Since the soil is silty sand, the friction angle value governs the

bearing capacity of the subsurface soil. Through the kriging analysis, the variation of friction angle values to a depth of 20 meters is evaluated for both 2:1 stress distribution and for an isobar of 0.1q.

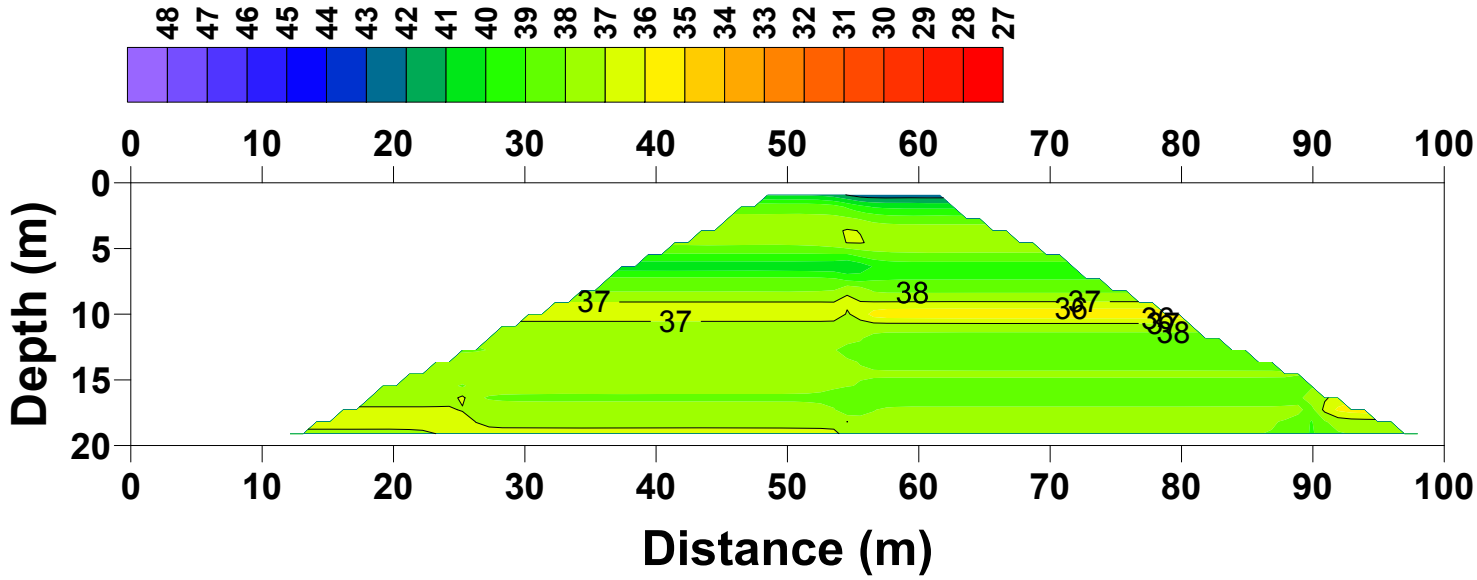
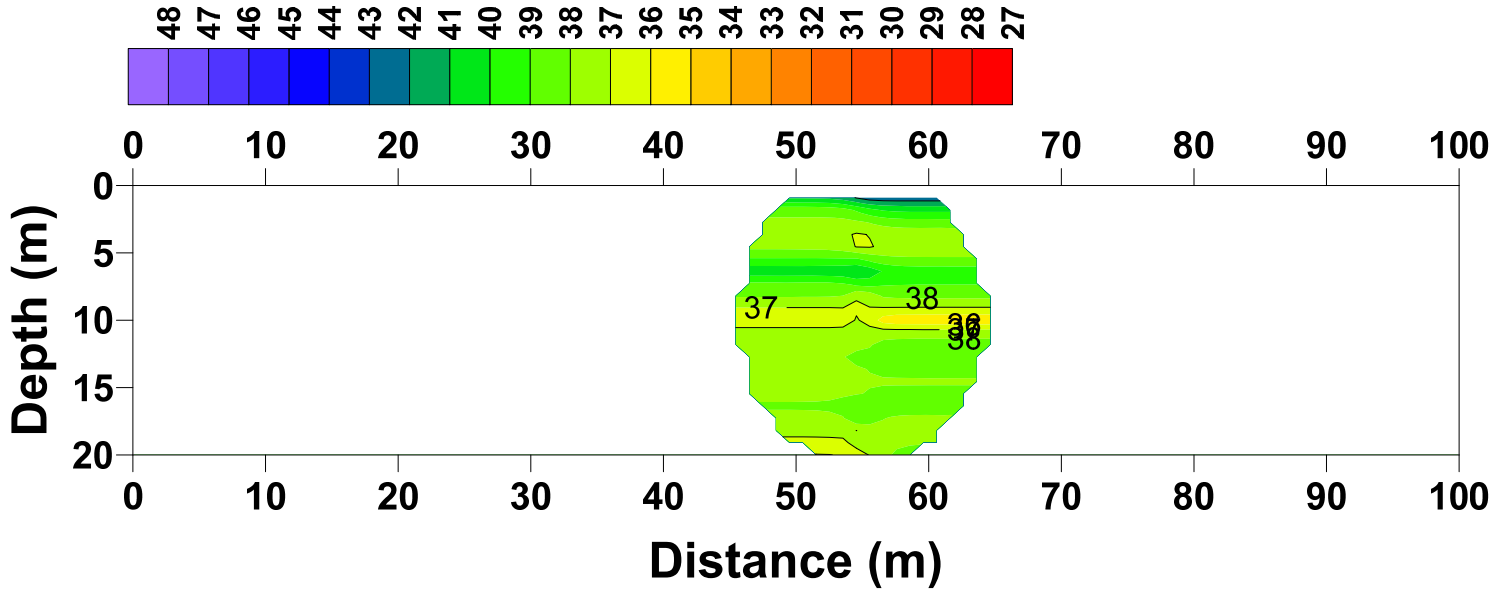


Figure 4-14 Friction Angle (Degrees) variation for 2:1 stress distribution

From the above chart, it can be observed that friction angle value varies from the 35.2 degrees to 42.6 degrees with a standard deviation value of 0.9.



Friction Angle (Degrees) for 0.1*q Isobar

Figure 4-15 Friction Angle (Degrees) variation for stress distribution using 0.1*q Isobar

From the above chart, it can be observed that friction angle value varies from the 35.2 degrees to 42.6 degrees, with a standard deviation value of 1.2.

4.8 Summary

In this study, an attempt was made to capture the spatial variability of field soil properties. Cone penetration tests were conducted in China as a part of a highway project. The basic soil classification was performed, and idealized subsurface profiles were provided for all the CPTU tests. Friction angle values of the soil were determined using the CPTU test parameters. Statistical analysis and geostatistical analysis were conducted according to the framework developed in the Chapter 3. Below are the important findings and conclusions from this study:

- The spatial correlation distance for a silty sand layer was found to be 0.86m, with an effective range from 2.58m to 4.3m.
- The spatial variability present in the silty sand layer is well explained with the exponential model.
- The maximum lag distance was found to be approximately one-third the total distance (i.e., 7m of maximum lag distance if the depth of bore hole is 21m).
- The friction angle design parameter evaluated from 2:1 stress distribution and pressure bulb was 38 degrees, with a standard deviation of 0.9 to 1.2; whereas, in the general case, it is found to be 37 degrees, with a standard deviation of 2.8.
- Mostly importantly, the framework developed in Chapter 3 was validated by using the CPTU data. So, this framework can be used to perform spatial variability analysis with any in-situ field data, provided the data is not less than 30 observations.

Chapter 5

Spatial Variability Analysis of Man-Made Treated Soils

5.1 Introduction

The stiffness of the soil in field conditions largely depends on the stress state of the soil, environmental changes, and internal and external conditions of the soil. Phoon and Kulhawy (1999) demonstrated that the stiffness of soils in field conditions varies from 15 - 65 percent in sandy soils, 9 - 92 percent in sand-to-clayey sand soils, and 7- 67 percent in silty soils. The influence of field conditions becomes extremely critical, especially in the case of buried pipelines. It is noted that external soil conditions are the primary cause of water pipeline breaks (Vipulanandan et. al., 2011). The most commonly occurring failure in pipelines is because of the circumferential stress developed due to the swelling or settlement of the bedding material (Seica et al. 2001; Rajani et al. 2004).

The difficulties involved in compacting the bedding material around the pipeline led to the use of controlled low strength material (CLSM) in the 1960's. CLSM is a self-compacted, cementitious material, used primarily as a backfill in lieu of compacted backfill and has become a popular material for projects such as void fill, foundation support, bridge approaches, and conduit bedding (Folliard et al., 2008). CLSM, with cement and fly ash additives, has been demonstrated, by many researchers, to be an effective bedding material for pipelines due to the material's self-compacting behavior and strength performance (Rajah et al., 2012; Boschert J. and Butler J., 2013).

A local water district has recently considered CLSMs as bedding material in the process of constructing a 150-mile water pipeline. This pipeline is intended to collect water from different lakes and bring additional water supplies to the Dallas/Fort Worth Metroplex. Figure 5-1 provides the layout area of the project, with existing TRWD

pipelines and a proposed new pipeline shown in yellow. The construction of the proposed pipeline, which is 9 ft. in diameter, is shown in Figure 5-2.

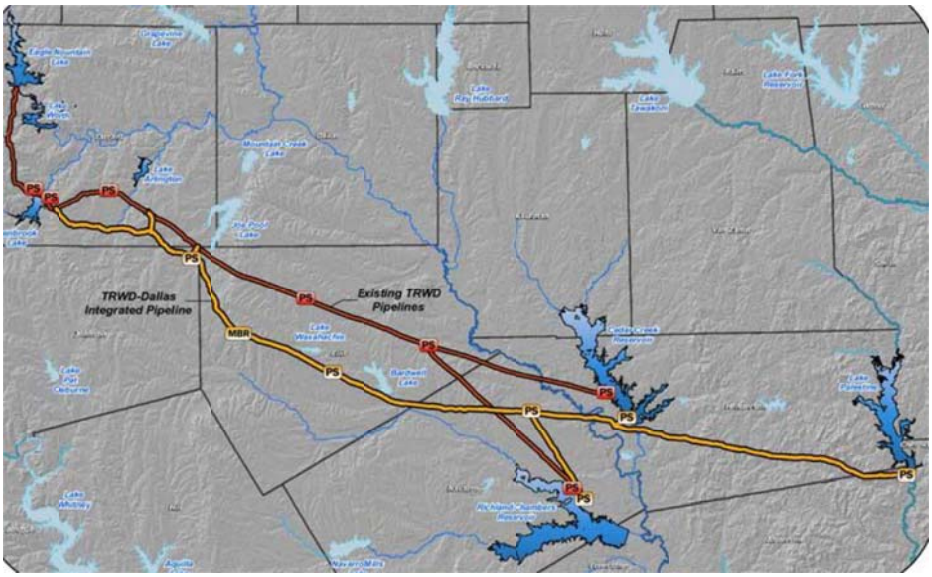


Figure 5-1 Integrated pipeline project layout (Source: TRWD)

Figure 5-2 Construction of the water pipeline at Line J

In this project, the bedding material, CLSM, was designed using native high plasticity soil. The proportions of the soil, cement, water used in preparing the mix design

were presented in the study conducted by Raavi (2012). The effective utilization of native high plasticity soil reduces the project costs and minimizes the negative impacts on the environment (Puppala and Hanchanloet, 1999; Abreu et al. 2008; Chittoori et al. 2012; Puppala et al. 2012a). The stiffness of the CLSM material in field conditions, at any particular stretch, depends upon the volume of CLSM occupied, which in turn depends on the soil conditions. One of the primary concerns relating to the CLSM is the development of uniform stiffness within the required time frame.

In this research study, a focus on geostatistics is used as a quality control tool to study the variations in CLSM stiffness properties with time. Stiffness measurements of the CLSM bedding material were determined using a non-destructive method in a 500ft. Prove-out test section. However, the stiffness measurements performed at certain intervals made it hard to predict the stiffness values throughout the pipeline or at untested locations. The framework developed in Chapter 3 was used in this study to obtain the stiffness values throughout the pipeline. The predictions were performed by incorporating the spatial variability present into the field stiffness measurements determined after a curing period of 1, 3, 7, 14 and 28 days. The spatial variability models developed, along with kriging algorithms, were used to predict the stiffness of the bedding material throughout the pipeline after 1, 3, 7, 14 and 28 days.

5.2 Data Acquisition

In this project, stiffness of the CLSM bedding material was determined using the spectral analysis of surface saves (SASW) technique. SASW is a seismic non-destructive method used to determine the small strain shear modulus of the materials. The detailed steps of performing the stiffness measurements for this project are presented in Mothkuri (2014).

An overview of the method is discussed in this section. The SASW technique was originally proposed and developed in the early 1980's, (Nazarian and Stokoe 1984; Stokoe et al. 1989). This method uses the dispersive characteristics of surface waves to determine the variation of the surface wave velocity of layered systems with depth. Figure 5-3 presents the tools that were used in this project to obtain stiffness measurements. The surface waves were generated using various types of hammers, depending upon the spacing between the two receivers. When an impact was created on the surface, several surface waves were generated with a wide spectrum of frequencies. At a known frequency, the phase difference recorded between the two receivers was used to calculate the travel time between two signals detected by the receivers, using Equation 5.1. With the known distance between the receivers and the travel time, the surface wave velocity at a given frequency was calculated using Equation 5.2. The shear wave velocity, which is correlated to surface wave velocity as shown in Equation 5.3, was used to determine the shear modulus of the material using equation 5.4

$$t(f) = \emptyset(f)/(360 \times f) \tag{5.1}$$

$Where,$ f = frequency, Hz

$t(f)$ = travel time for a given frequency,

$\emptyset(f)$ = phase difference in degrees for a given frequency

$$V_R(f) = D/t(f) \tag{5.2}$$

$Where, V_R(f)$ = surface wave velocity at a given frequency

D = distance between two receivers

$$V_s(f) = 0.95 * V_R(f) \tag{5.3}$$

$Where, V_s(f)$ = shear wave velocity at a given frequency

$$G = \rho V_s^2 \tag{5.4}$$

$Where, G =$ shear modulus

ρ = density of the material

V_s^2 = shear wave velocity

The stiffness measurements were determined along the 500 ft. pipeline which was comprised of 10 pipe sections of 50 ft. each, as shown in Figure 5-4. Measurements were taken at 85 test points (17 stations X 5 points each) along the seventeen established test stations, as shown in Figure 5-5. The cross-sectional figure of the pipeline with test points is shown in Figure 5-6.

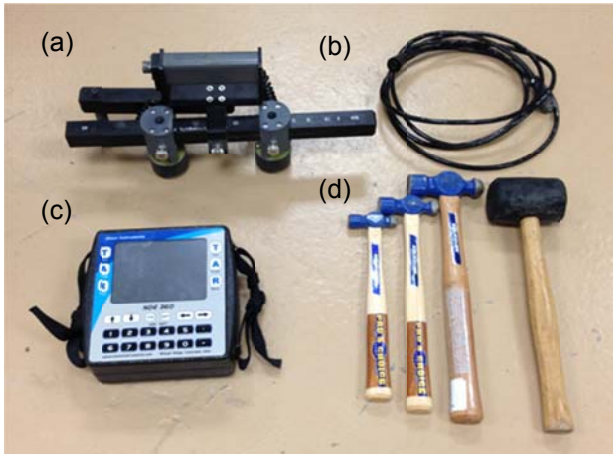


Figure 5-3 SASW tools: (a) SASW bar with geophone (b) Cables

(c) Data logger (d) hammers

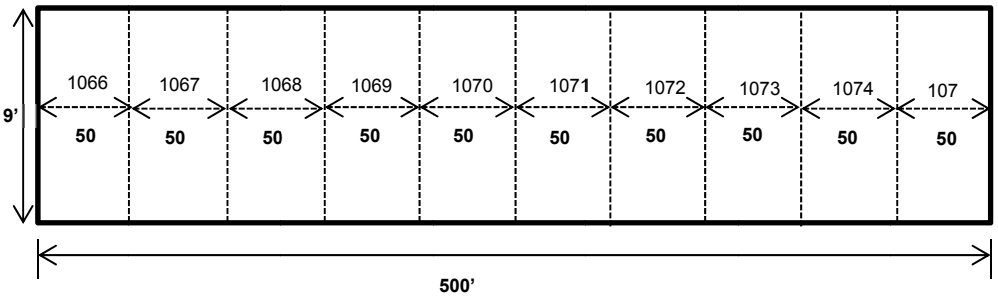


Figure 5-4 Top view of 500feet pipeline with 10 sections

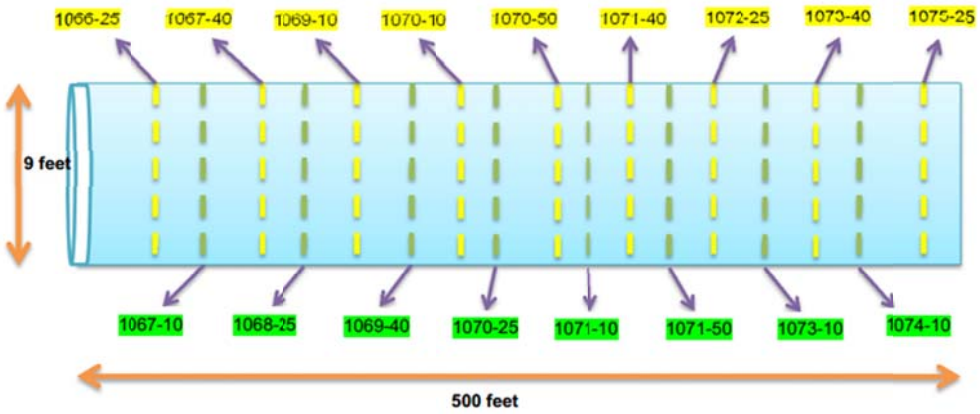


Figure 5-5 Seventeen test sections inside the pipeline

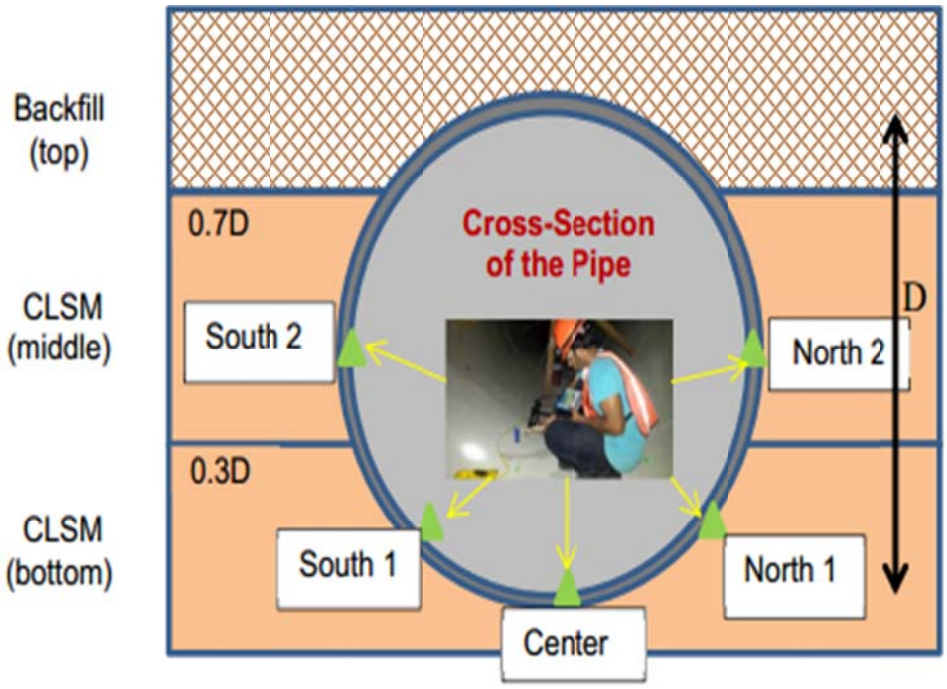


Figure 5-6 Cross-sectional view of the pipe and test points at each section

.

The bedding material, CLSM, was allowed to cure for a period of 28 days to reach its maximum strength. The stiffness measurements were taken continuously for the first 14 days, and then on day 28, readings were taken again to monitor the stiffness development of CLSM in field conditions. In this study, spatial variability analysis was conducted to predict the stiffness at unsampled locations and to evaluate the uniform stiffness development throughout the pipeline. Tables 5-1 to 5-5 provide the stiffness measurements for days 1, 3, 7, 14 and 28. Based on the laboratory tests on strength of CLSM, it was expected to achieve its maximum strength after 28 days of curing (Raavi, 2012).

Table 5-1 Stiffness measurements for Day

Section	Stiffness Observations in MPa					
	1	2	3	4	5	6
1066	264.4	238.1	281.1			
1067	330.3	254.4	261.0	245.3	308.7	270.7
1068	250.9	252.9				
1069	250.9	259.5	297.5	252.8	256.6	244.5
1070	261.1	258.9	294.7	202.2	251.0	275.4
1071	301.9	251.8	274.6	265.1	247.8	273.2
1072	217.7	248.6	273.4			
1073	280.2	254.3	244.4	270.4	255.0	319.0
1074	365.2	247.9	260.1			
1075	312.2	251.5	292.7			

Table 5-2 Stiffness Measurements Day 3

Section	Stiffness Observations in MPa									
	1	2	3	4	5	6	7	8	9	10
1066	276.5	275.1	272.7	289	271.9					
1067	271.4	303.3	266.8	261	256.7	285.8	274.3	309.0	346.0	274.6
1068	251.0	257.6	282.3	275	252.6					
1069	261.8	253.8	330.6	298	232.9	307.8	291.3	293.4	260.1	273.3
1070	293.9	261.1	261.2	295	276.2	276.1	268.7	275.4	275.0	264.5
1071	243.8	273.3	255.8	305	256.3	243.8	273.3	255.8	304.6	256.3
1072	275.8	260.0	293.7	277	255.4					
1073	217.4	290.8	285.9	270	298.0	258.5	277.5	280.0	300.7	269.0
1074	310.9	287.1	279.3	275	264.4					
1075	258.5	277.5	280.0	301	269.0					

Table 5-3 Stiffness Measurements of Day 7

Section	Stiffness Observations in MPa									
	1	2	3	4	5	6	7	8	9	10
1066	327.0	350.0	324.2	329.8	320					
1067	321.9	359.8	318.3	281.2	305	336.3	319.5	329.4	386.7	322.5
1068	301.4	302.8	333.9	295.4	300					
1069	312.3	299.0	382.1	345.8	281	358.2	336.5	344.9	300.8	321.1
1070	344.4	301.6	312.8	317.5	324	326.6	313.9	327.0	335.7	312.4
1071	294.3	318.5	307.4	345.3	304	294.3	318.5	307.4	345.3	304.2
1072	326.3	305.2	335.2	317.3	303					
1073	267.9	316.0	307.5	291.1	346	309.0	312.7	331.6	341.4	316.9
1074	361.3	332.3	310.8	315.7	312					
1075	309.0	312.7	331.6	341.4	317					

Table 5-4 Stiffness Measurements of Day 14

Section	Stiffness Observations in MPa									
	1	2	3	4	5	6	7	8	9	10
1066	398.4	388.0	398.8	386.5	391					
1067	370.0	398.3	373.7	397.9	341	395.3	408.3	371.1	410.1	389.1
1068	372.1	367.1	382.9	347.5	390					
1069	387.7	389.3	418.7	393.7	423	397.0	415.5	400.8	388.9	392.8
1070	390.5	393.6	399.1	397.1	407	361.9	356.0	407.3	380.7	386.0
1071	359.4	366.9	400.0	396.8	428	359.4	366.9	380.0	366.8	378.0
1072	377.5	394.8	385.3	379.6	415					
1073	368.5	383.9	357.3	399.1	387	383.8	381.4	399.4	381.4	406.8
1074	396.2	399.6	362.5	360.0	404					
1075	383.8	371.4	369.4	401.4	417					

Table 5-5 Stiffness Measurements of Day 28

Section	Stiffness Observations in MPa									
	1	2	3	4	5	6	7	8	9	10
1066	432.3	465.9	458.7	447.5	432					
1067	417.7	463.1	414.1	460.1	409	473.7	455.9	436.8	450.7	470.1
1068	455.5	407.6	437.3	410.2	446					
1069	442.2	421.0	440.6	417.5	440	420.0	455.9	443.0	454.3	427.6
1070	434.7	444.8	452.4	439.2	434	446.5	471.5	453.1	440.8	441.7
1071	424.6	414.2	453.2	440.8	421	424.6	414.2	413.2	440.8	420.9
1072	461.9	457.7	419.8	432.7	461					
1073	419.4	438.5	428.8	429.9	418	456.4	458.7	442.5	470.5	465.2
1074	449.5	478.2	420.2	468.1	450					
1075	456.4	458.7	442.5	470.5	465					

5.3 Data Organization

The 500 ft. pipeline is composed of ten sections, starting from 1066 to 1075. In order to conduct the spatial variability analysis of the stiffness measurements determined in 10 sections, the data had to be organized in the Cartesian coordinate system. Certain assumptions were made regarding the location of the test points. The test point at the center 'c', as shown in Figure 5-6 of section 1066-25, is assumed to be located at (25, 25) in (x, y) coordinate system. This point was taken as the reference point, and coordinates of the remaining test points were determined. Table 5-6 presents the total sections and distances of the test points with previous points.

Table 5-6 Coordinates of the test points

Section	Station No.	Distance from previous station	Coordinates in (x,y) system for center points
1066	25'	0 feet -Reference point	(25,25)-Reference point
1067	10'	35 feet	(60,25)
	40'	30 feet	(90, 25)
1068	25'	35 feet	(125. 25)
1069	10'	35 feet	(160, 25)
	40'	30 feet	(190, 25)
1070	10'	20 feet	(210, 25)
	25'	15 feet	(225, 25)
1071	10'	35 feet	(260, 25)
	40'	30 feet	(290, 25)
1072	25'	35 feet	(325, 25)

Table 5.6 - *Continued*

1073	10’	35 feet	(360, 25)
	40’	30 feet	(390, 25)
1074	10’	20 feet	(410, 25)
1075	25’	65 feet	(475, 25)

Similarly, the coordinates of the test points at locations on either side of the center were calculated. The horizontal distance from center to south 1 was 2.78 ft. and to south 2 was 4.38 ft.; similarly, from center to north 1 was 2.78 ft. and to north 2 was 4.38 ft. To determine the spatial variability, the ‘z’ coordinate was given as the stiffness value. For example, 25, 25, 238.1 shows that the stiffness of CLSM at section 1066-25 was 238.1 MPa. The coordinates for all the stiffness measurements utilized for the analysis are presented in Appendix B.

5.4 Statistical Analysis

In this section, the statistical analysis for the stiffness values obtained on days 1, 3, 7, 14 and 28 was performed. The elementary statistical parameters such as mean, variance, and standard deviation were evaluated to supplement the spatial variability analysis.

5.4.1 Histograms

Histogram, in this analysis, was used to identify the distribution associated with the stiffness values. It provided the visual aid for the distribution of the stiffness values, along with the summary of elementary statistical parameters. The number of class intervals that were required to construct a histogram plot were determined using Equation 3-1. The class intervals obtained were used in Equation 3-2 to determine the bin size.

The table below summarizes the number of class intervals and bin sizes obtained for stiffness values on days 1, 3, 7, 14 and 28.

Table 5-7 Summary of parameters used to construct histograms

Day	Observations	Number of Class Intervals	Bin Size
1	44	6.4	25
3	75	7.2	18
7	75	7.2	17
14	75	7.2	12.5
28	75	7.2	9.5

The above parameters were used to construct the histograms. Figures 5-7 to 5-11 provided the histogram plot for the stiffness values determined on days 1, 3, 7, 14 and 28. The stiffness of CLSM values were plotted on x-axis and y-axis, representing frequency of the observations. From the plots, it can be inferred that the distribution of stiffness on 1, 3, 7 and 14 days were more closely related to Gaussian distribution, with a little skewness on either side. However, the histogram plot for day 28 is more uniformly distributed. This can be because of reaching the maximum strength in almost all the test points. The maximum frequency on all the plots depicts the stiffness reached in most of the test point locations. On day 1, the maximum test points reached a stiffness value ranging from 250 – 275 MPa; on day 3, the maximum test points reached stiffness value of 270 – 280 MPa; on day 7, the maximum test points had a stiffness value of 300 – 320 MPa; on day 14, the stiffness of the maximum test points was 380 – 400 MPa; and on day 28, the maximum test points reached a stiffness value of 410- 480 MPa. The

increase in mean stiffness value from day 1 to day 28 was observed to be 39.5 percent; however, the standard deviation of the stiffness values on all days varied from 18- 29 percent.

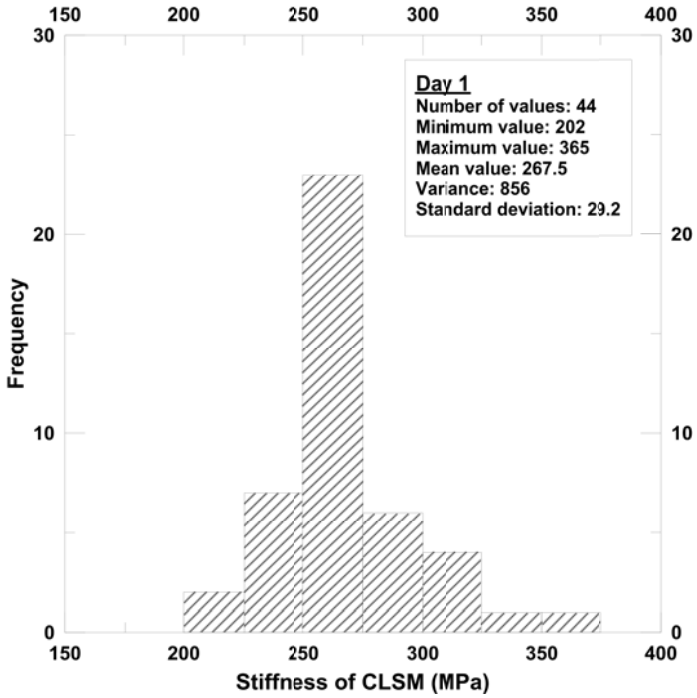


Figure 5-7 Histogram plot for stiffness values determined on day 1

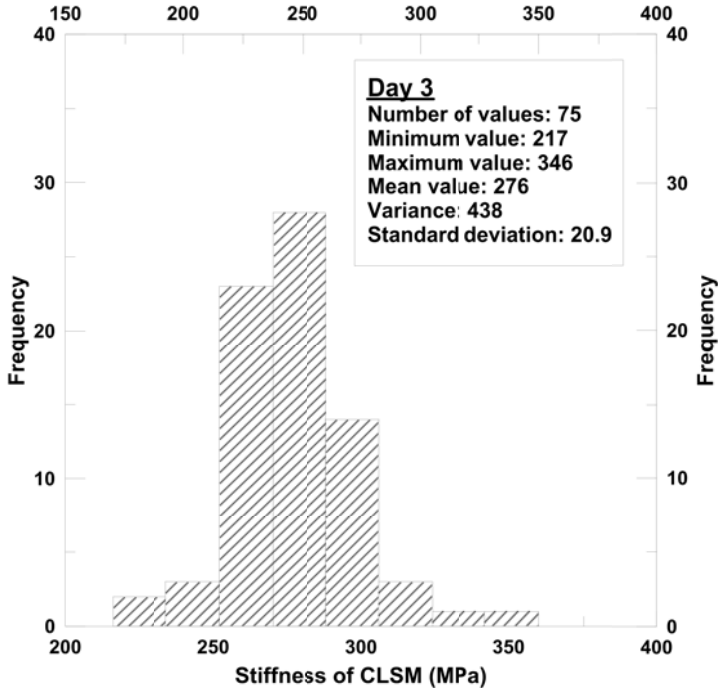


Figure 5-8 Histogram plot for stiffness values determined on day 3

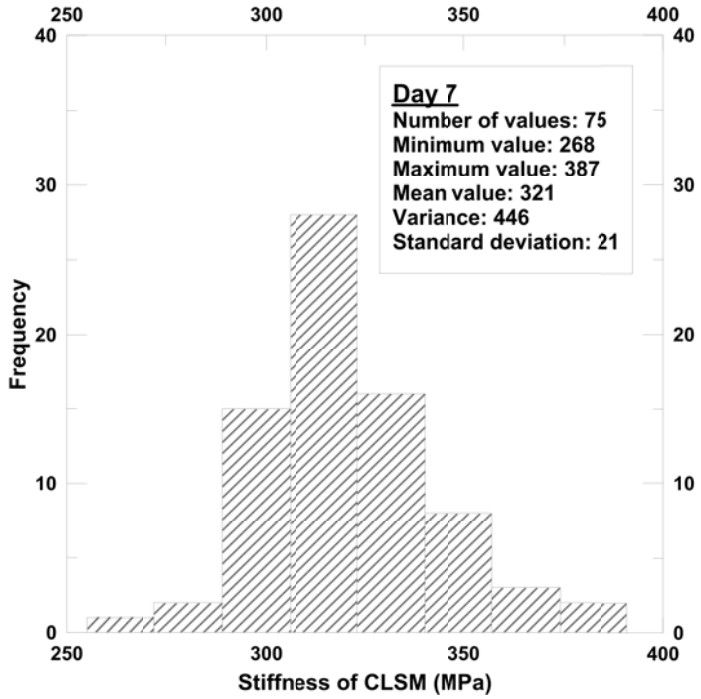


Figure 5-9 Histogram plot for stiffness values obtained on day 7

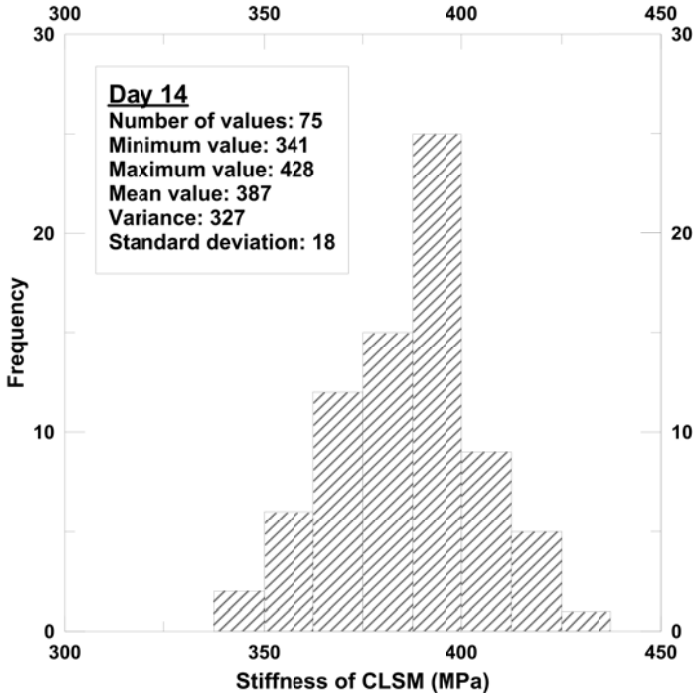


Figure 5-10 Histogram plot for stiffness values obtained on day 14

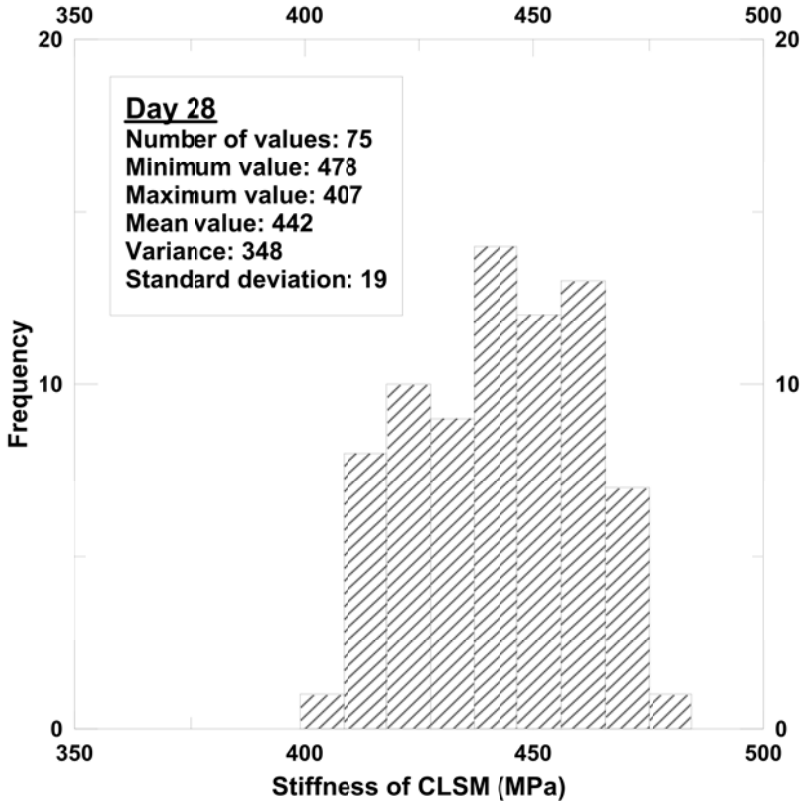


Figure 5-11 Histogram plot for stiffness values obtained on day 28

5.4.2 Check for Gaussian Distribution

The histogram plots presented earlier provided the basic information on the distribution of data. As most of the statistical tests rely on the Gaussian distribution, in this section, the stiffness values determined for days 1, 3, 7, 14 and 28 were used to check for Gaussian distribution. The Shapiro-Wilk test is a more formal approach to checking for Gaussian distribution; however, the number of observations is limited to 50. Hence, in this study, the normal-quantile plots were used to check for Gaussian distribution in the data.

5.4.2.1 Normal- Quantile plot

Figures 5-12 to 5-16 present the normal-quantile plots for the stiffness values obtained on days 1, 3, 7, 14 and 28, respectively. The theoretical quantiles for

corresponding stiffness values were calculated using Equation 3-3. The theoretical quantile values were plotted on the y-axis against the corresponding stiffness quantile values on the x-axis. It can be inferred from the plots that the data values followed a linear trend. Using the regression approach method, a best fit trend line was modelled for the plotted data values. The coefficient of determination was calculated using Equation (3-4) to measure the fit of the trend line. Table 5-8 below summarizes the coefficient of determination values obtained for all the plots.

Table 5-8 Coefficient of determination for normal-quantile plots

Day	Coefficient of determination (r^2)
1	0.92
3	0.96
7	0.97
14	0.99
28	0.97

The coefficient of determination values were more than 0.9 in all the plots, demonstrating that the linear trend provided the best fit for the available data points. This leads to the conclusion that the stiffness of all the data points were distributed normally or had a Gaussian distribution.

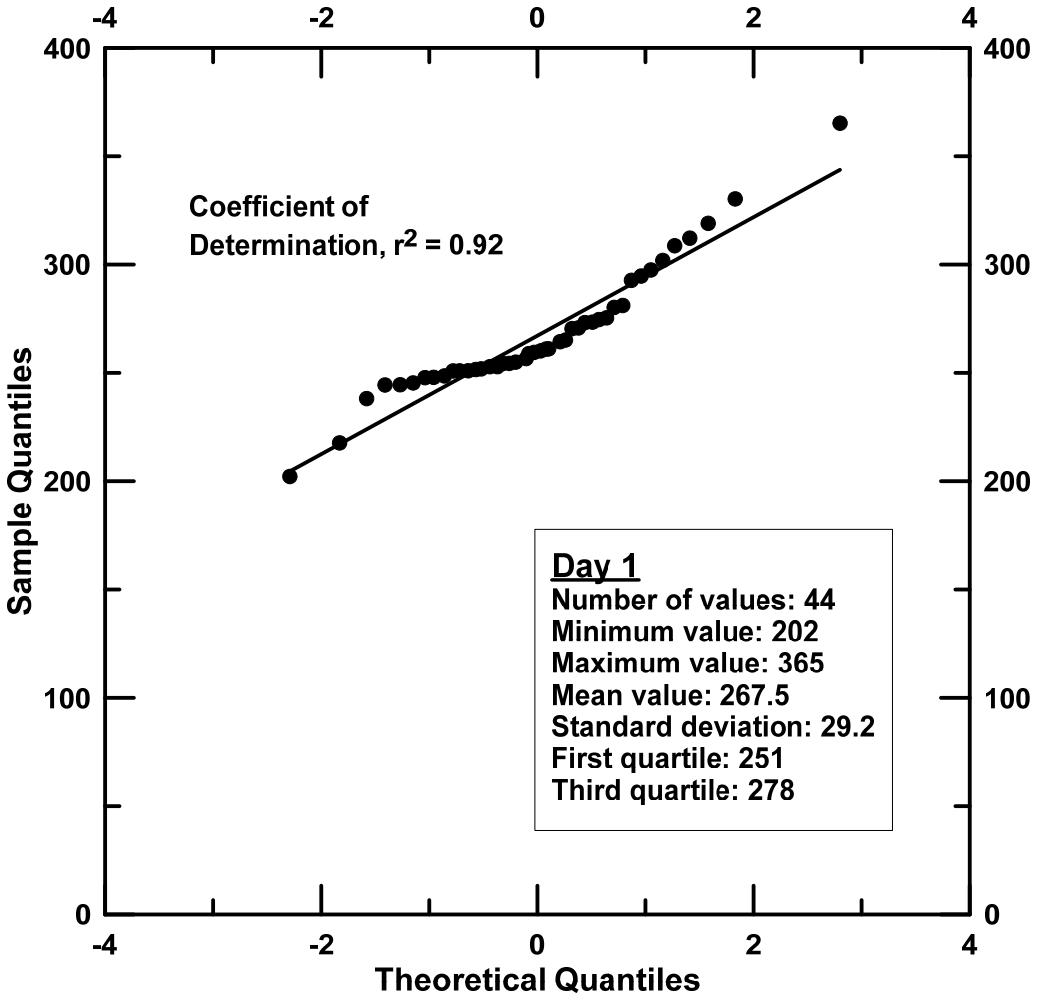


Figure 5-12 Normal- Quantile plot for stiffness values of day 1

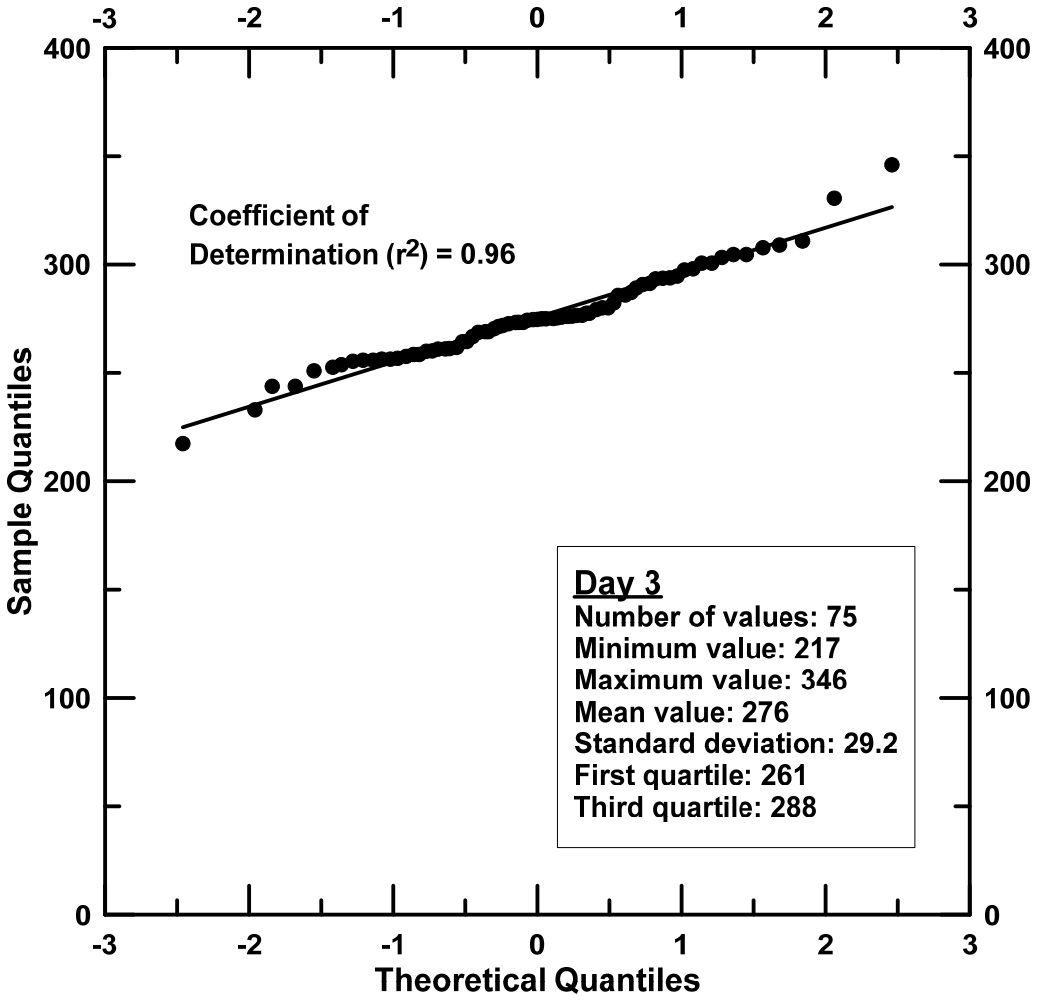


Figure 5-13 Normal- Quantile plot for stiffness values of day 3

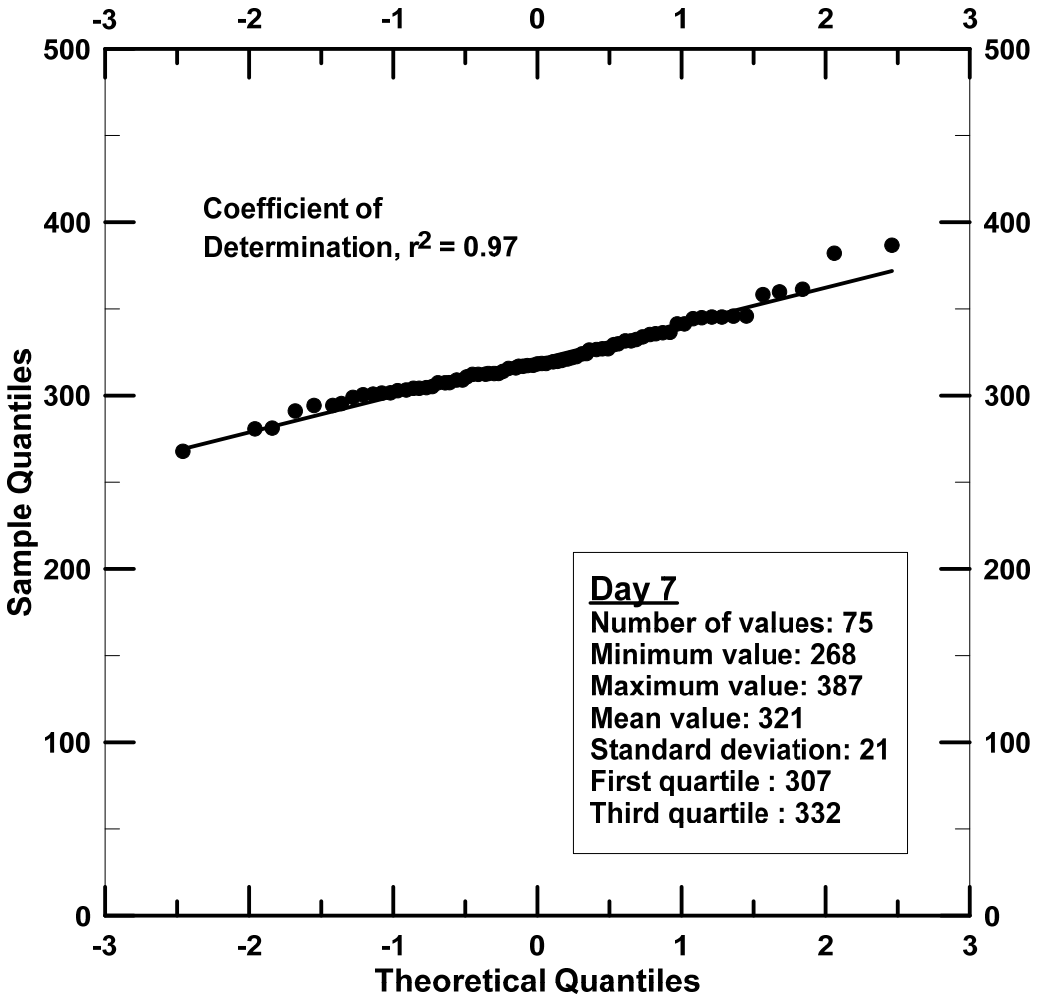


Figure 5-14 Normal- Quantile plot for stiffness values of day 7

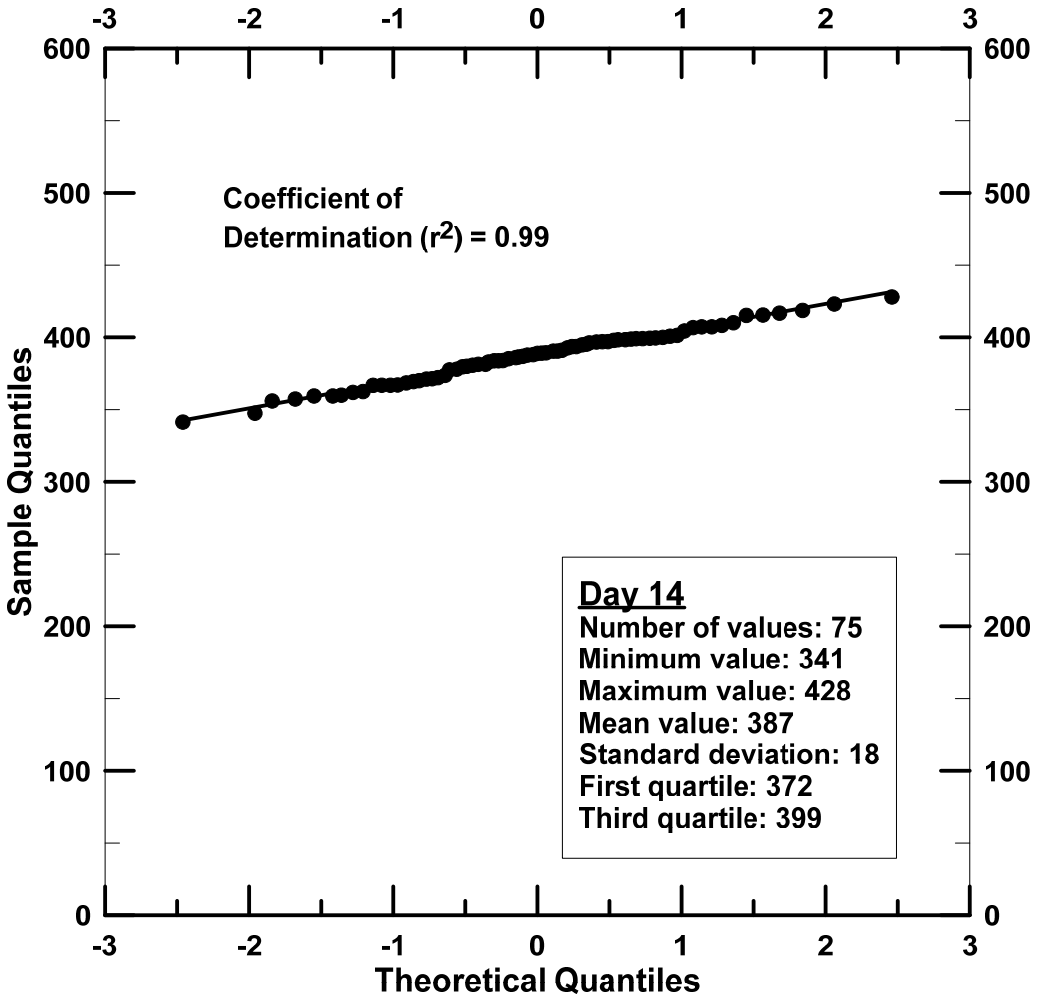


Figure 5-15 Normal- Quantile plot for stiffness values of day 14

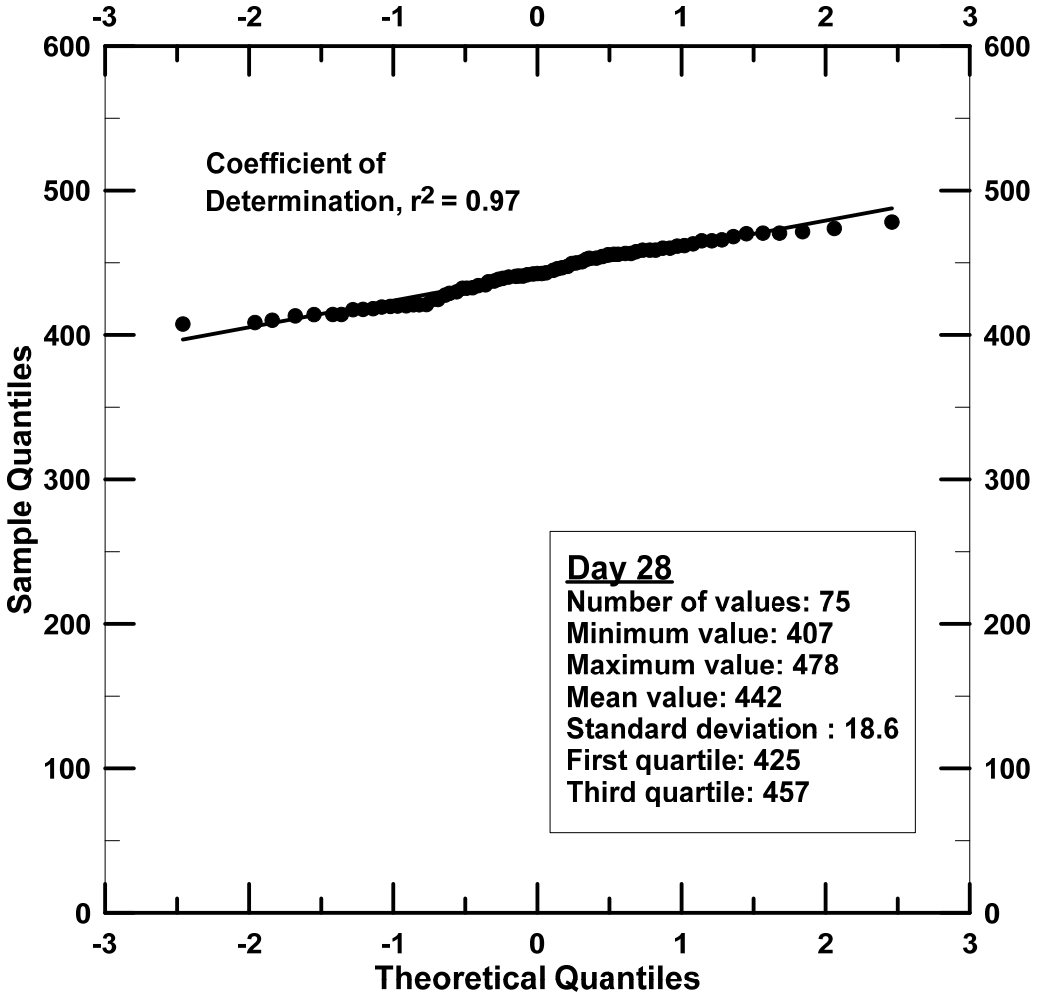


Figure 5-16 Normal- Quantile plot for stiffness values of day 28

5.5 Check for Stationarity

As mentioned in Chapter 3, in order to perform geostatistical analysis, the data has to be stationary. The stationarity in the data was evaluated by conducting two tests: one for checking constant mean value and another for checking constant variance. The ANOVA test method was used to determine the constant mean in the stiffness values, and Bartlett's test was used to evaluate the constant variance in stiffness values for days 1, 3, 7, 14 and 28. In order to perform ANOVA and Bartlett test, the stiffness values obtained from 10 sections were analyzed and compared.

5.5.1 Check for Constant Mean Value

The ANOVA test was used to check for constant mean values in all the stiffness measurements obtained in the 10 sections. This was performed by comparing mean stiffness values in each section. Below is the hypothesis constructed for evaluating constant mean in the data:

$H_0 : \mu_{1066} = \mu_{1067} = \mu_{1068} = \mu_{1069} = \mu_{1070} = \mu_{1071} = \mu_{1072} = \mu_{1073} = \mu_{1074} = \mu_{1075}$

H_1 : At least one mean is different

The statistic that was used to evaluate the above hypothesis is $F_{0,}$ which is the ratio of mean square treatment to mean square error. This was compared to the critical value at a significance level of 0.05. Tables 5-9 to 5-13 below provide the summary of the ANOVA test results conducted for checking the constant mean in stiffness values in all sections after 1, 3, 7, 14 and 28 days of curing.

Table 5-9 Summary of ANOVA results for day 1 stiffness values

Analysis of Variance				
Source of Variation	Sum of Squares	DOF	Mean Square	F_0
Treatment	6295.5	9	699.5	0.77
Error	30504.5	34	897.1	
Total	36800.1	43		

*DOF = Degrees of freedom

Table 5-10 Summary of ANOVA results for day 3 stiffness values

Analysis of Variance				
Source of Variation	Sum of Squares	DOF	Mean Square	F_0
Treatment	2958.2	9	328.6	0.72
Error	29454.4	65	453.1	
Total	32412.6	74		

*DOF = Degrees of freedom

Table 5-11 Summary of ANOVA results for day 7 stiffness values

Analysis of Variance				
Source of Variation	Sum of Squares	DOF	Mean Square	F_0
Treatment	3650.1	9	405.5	0.87
Error	30208.5	65	464.7	
Total	33858.6	74		

*DOF = Degrees of freedom

Table 5-12 Summary of ANOVA results for day 14 stiffness values

Analysis of Variance				
Source of Variation	Sum of Squares	DOF	Mean Square	F_0
Treatment	3804.5	9	422.7	1.34
Error	20403.9	65	313.9	
Total	24208.5	74		

*DOF = Degrees of freedom

Table 5-13 Summary of ANOVA results for day 28 stiffness values

Analysis of Variance				
Source of Variation	Sum of Squares	DOF	Mean Square	F_0
Treatment	5758.3	9	639.8	2.11
Error	19679.0	65	302.7	
Total	25437.3	74		

*DOF = Degrees of freedom

The statistic $F_{0,}$ that was calculated for the stiffness values, were compared with critical value ($f_{crit)}$. The f_{crit} value was obtained by using the f-distribution table at a significance level (α) of 0.05. The f_{crit} for day 1 was 2.57 and for rest of the days was 2.16, as the number of observations for days 3, 7, 14 and 28 are equal. From the above ANOVA test results for all days, it can be inferred that the statistic (F_0) was smaller than the critical region. Therefore, we do not reject the null hypothesis developed, concluding the means of the stiffness values were constant in all the sections for all days.

5.5.1.1 Model Adequacy check

In order to validate the conclusions obtained from the ANOVA tests, the basic assumptions of the ANOVA model have to be met. The assumptions are: the residual error values are normally distributed and the residual variances are structureless. The

residuals of the stiffness values were calculated using Equation 3.8. The fitted value in Equation 3.8 was obtained by determining the mean of the stiffness value in each section. Tables 5-14 to 5-19 provide the residual values calculated for the stiffness values of all days that were used for testing the model adequacy of the ANOVA model.

Table 5-14 Residual values for the day 1 stiffness values

Section	Residuals					
	1	2	3	4	5	6
1066	-0.69	7.83	-7.14			
1067	-0.48	45.38	-5.20	-48.26	10.96	-2.39
1068	-11.33	37.67	-26.33			
1069	27.25	-24.75	-15.64	-9.36	-74.75	97.25
1070	249.67	-42.33	-47.33	-70.33	-13.69	-75.97
1071	7.72	-60.28	-78.28	-46.28	123.86	53.24
1072	140.20	-48.56	-91.64			
1073	-65.85	9.32	-118.68	-124.06	379.96	-80.68
1074	-43.02	173.15	-130.14			
1075	-18.48	85.74	-67.25			

Table 5-15 Residual values for day 3 stiffness values

Section	Residuals									
	1	2	3	4	5	6	7	8	9	10
1066	-0.5	-1.9	-4.3	12.0	-5.1					
1067	-13.4	18.4	-18.1	-23.8	-28.1	0.8	-10.6	24.1	61.1	-10.2
1068	-12.6	-6.0	18.6	11.0	-11.0					
1069	-18.4	-26.4	50.3	17.2	-47.3	27.5	11.0	13.1	-20.0	-6.9
1070	19.1	-13.6	-13.4	20.0	1.5	1.4	-5.9	0.7	0.2	-10.1
1071	-22.9	6.5	-10.9	37.8	-10.4	-22.9	6.5	-10.9	37.8	-10.4
1072	3.5	-12.2	21.3	4.3	-16.9					
1073	-57.4	15.9	11.1	-4.4	23.1	-16.3	2.7	5.2	25.8	-5.7
1074	27.5	3.8	-4.0	-8.3	-18.9					
1075	-18.6	0.3	2.8	23.5	-8.1					

Table 5-16 Residual values for day 7 stiffness values

Section	Residuals									
	1	2	3	4	5	6	7	8	9	10
1066	-3.1	19.8	-5.9	-0.3	-10.4					
1067	-6.1	31.8	-9.7	-46.7	-23.4	8.2	-8.5	1.3	58.7	-5.5
1068	-5.3	-3.9	27.0	-11.4	-6.3					
1069	-15.9	-29.2	53.9	17.6	-47.3	30.0	8.3	16.7	-27.3	-7.0
1070	22.7	-19.9	-8.8	-4.1	2.4	5.0	-7.6	5.4	14.0	-9.2
1071	-19.6	4.5	-6.5	31.3	-9.7	-19.6	4.5	-6.5	31.3	-9.7
1072	8.8	-12.2	17.7	-0.1	-14.2					
1073	-46.1	2.0	-6.5	-22.8	31.9	-5.0	-1.2	17.5	27.3	2.9
1074	34.8	5.8	-15.6	-10.8	-14.2					
1075	-13.3	-9.5	9.2	19.0	-5.4					

Table 5-17 Residual values for day 14 stiffness values

Section	Residuals									
	1	2	3	4	5	6	7	8	9	10
1066	5.9	-4.5	6.3	-6.1	-1.5					
1067	-15.5	12.8	-11.8	12.4	-44.2	9.7	22.8	-14.4	24.6	3.6
1068	0.1	-5.0	10.9	-24.5	18.5					
1069	-13.0	-11.4	17.9	-7.1	22.4	-3.7	14.7	0.1	-11.8	-8.0
1070	2.5	5.7	11.2	9.1	19.3	-26.1	-32.0	19.4	-7.2	-1.9
1071	-20.8	-13.4	19.8	16.6	47.8	-20.8	-13.4	-0.2	-13.4	-2.2
1072	-13.0	4.3	-5.2	-10.9	24.7					
1073	-16.4	-0.9	-27.5	14.3	2.1	-1.0	-3.5	14.6	-3.5	21.9
1074	11.6	15.1	-22.0	-24.6	19.9					
1075	-4.7	-17.2	-19.1	12.8	28.2					

Table 5-18 Residual values for day 28 stiffness values

Section	Residuals									
	1	2	3	4	5	6	7	8	9	10
1066	-15.0	18.5	11.3	0.1	-15.0					
1067	-27.4	18.0	-31.0	15.0	-36.5	28.6	10.8	-8.3	5.6	25.0
1068	24.2	-23.7	6.0	-21.1	14.6					
1069	6.0	-15.2	4.4	-18.7	3.9	-16.2	19.6	6.8	18.1	-8.6
1070	-13.1	-3.1	4.6	-8.6	-13.6	-1.3	23.7	5.3	12.3	-6.1
1071	-2.2	-12.6	26.5	14.1	-5.8	-2.2	-12.6	-13.5	14.1	-5.8
1072	15.2	11.0	-26.9	-14.0	14.8					
1073	-23.5	-4.3	-14.0	-12.9	-24.6	13.6	15.9	-0.3	27.7	22.4
1074	-3.7	25.0	-33.0	14.9	-3.3					
1075	-2.2	0.1	-16.2	11.8	6.5					

The residual values provided in the tables above were used to check for the model adequacy (i.e., e_{ij} ~) in the ANOVA model. The normal distribution of the residual values was evaluated using the normality plots, and the variance was evaluated by plotting the residual values against the fitted values for all the sections. Figures 5-17 to 5-21 present the model adequacy plots for the residual values presented in the previous tables.

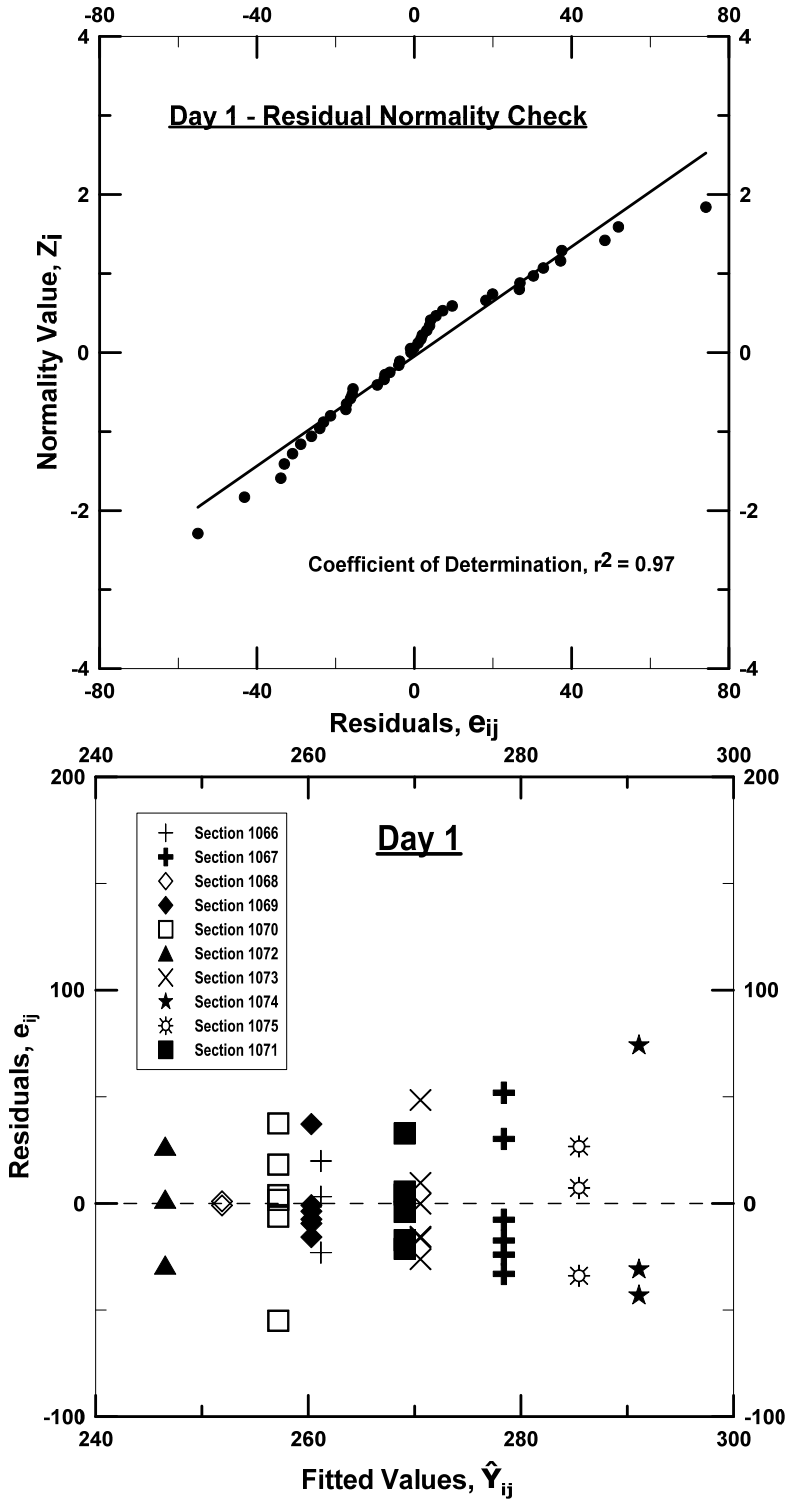


Figure 5-17 Model Adequacy plots for day 1 residual values

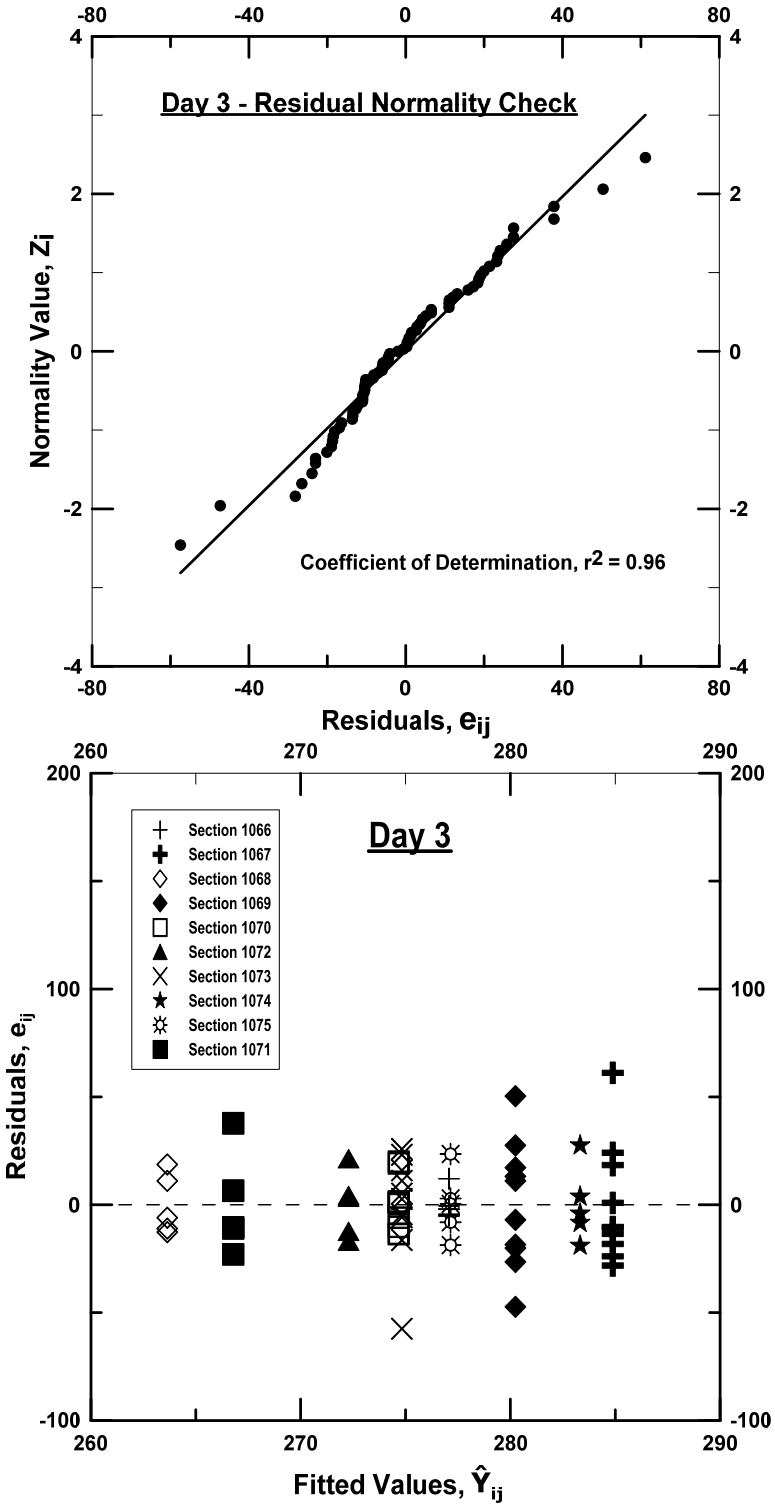


Figure 5-18 Model Adequacy plots for day 3 residual values

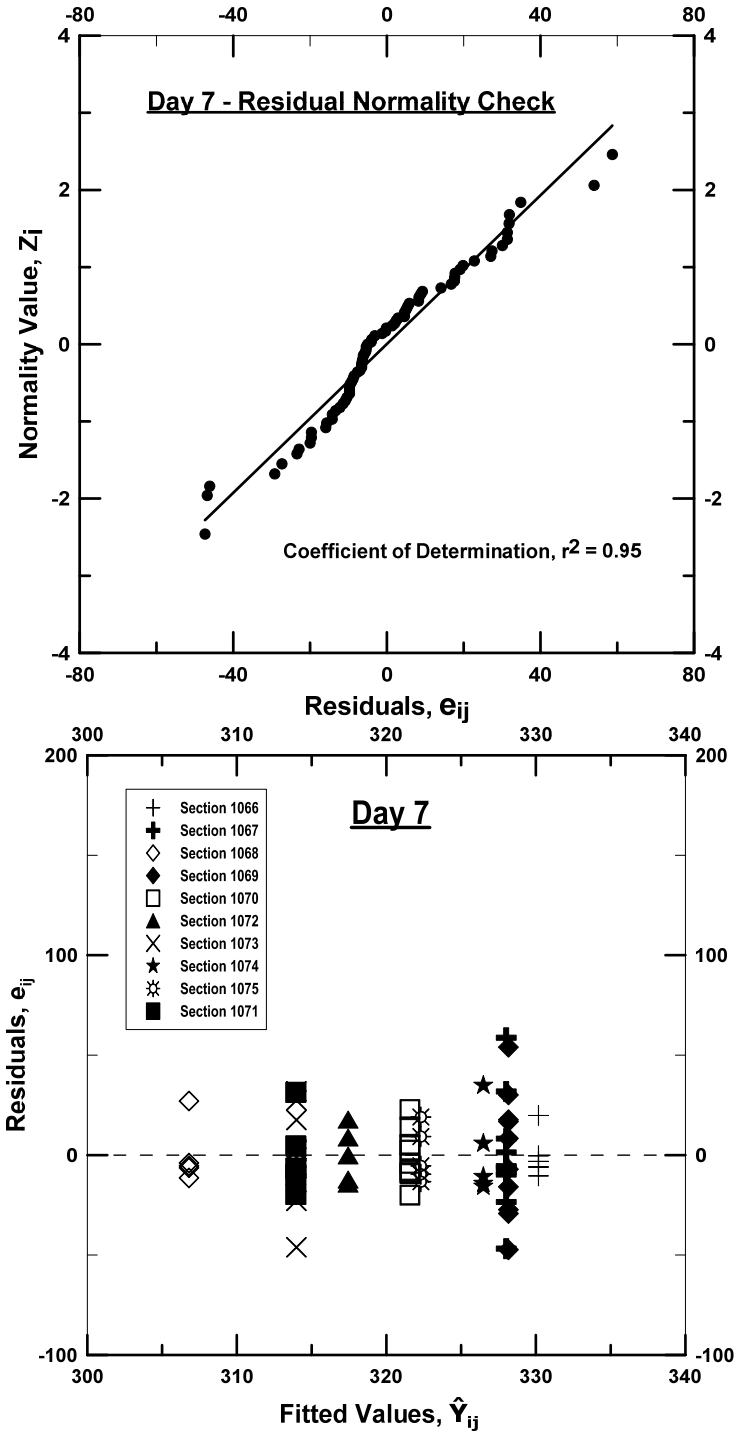


Figure 5-19 Model Adequacy plots for day 7 residual values

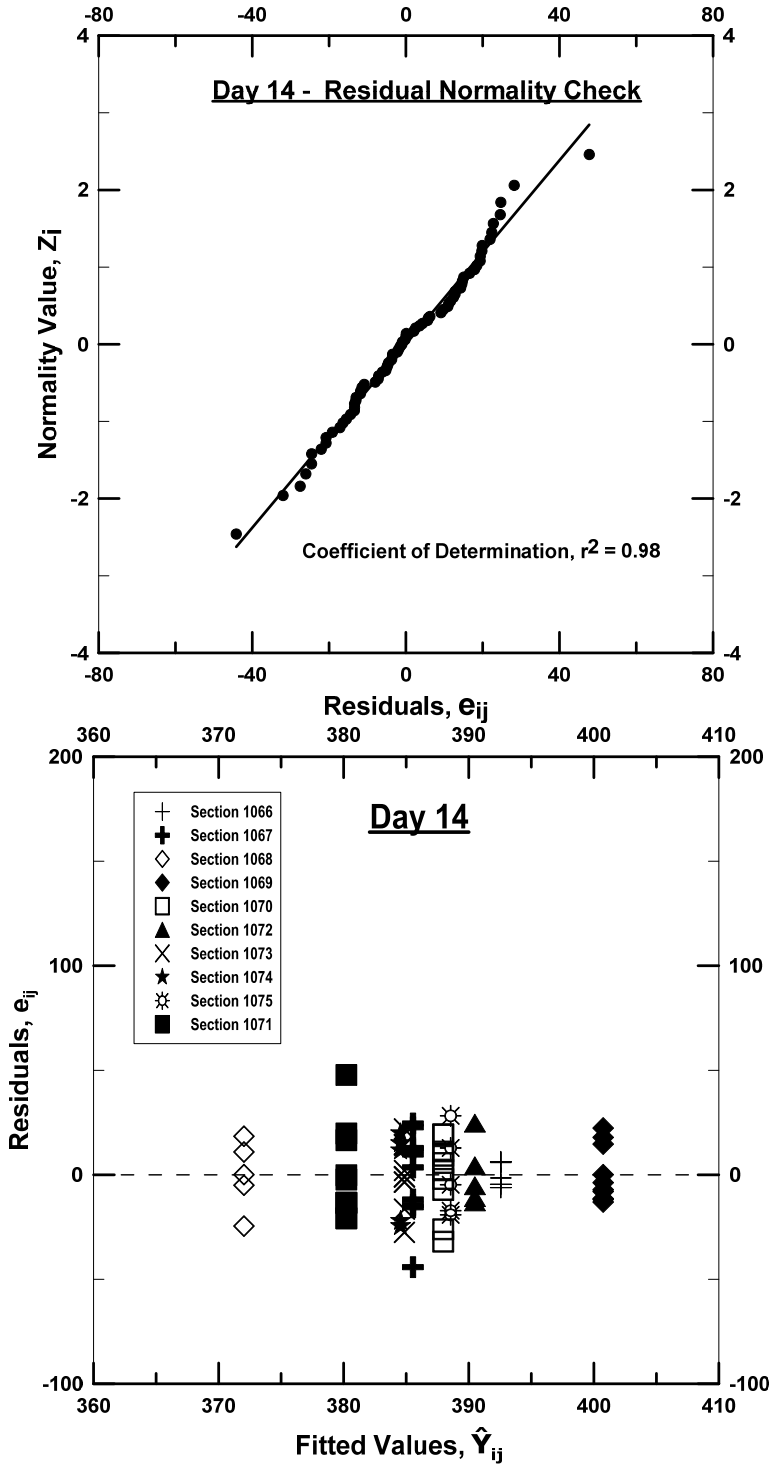


Figure 5-20 Model Adequacy plots for day 14 residual values

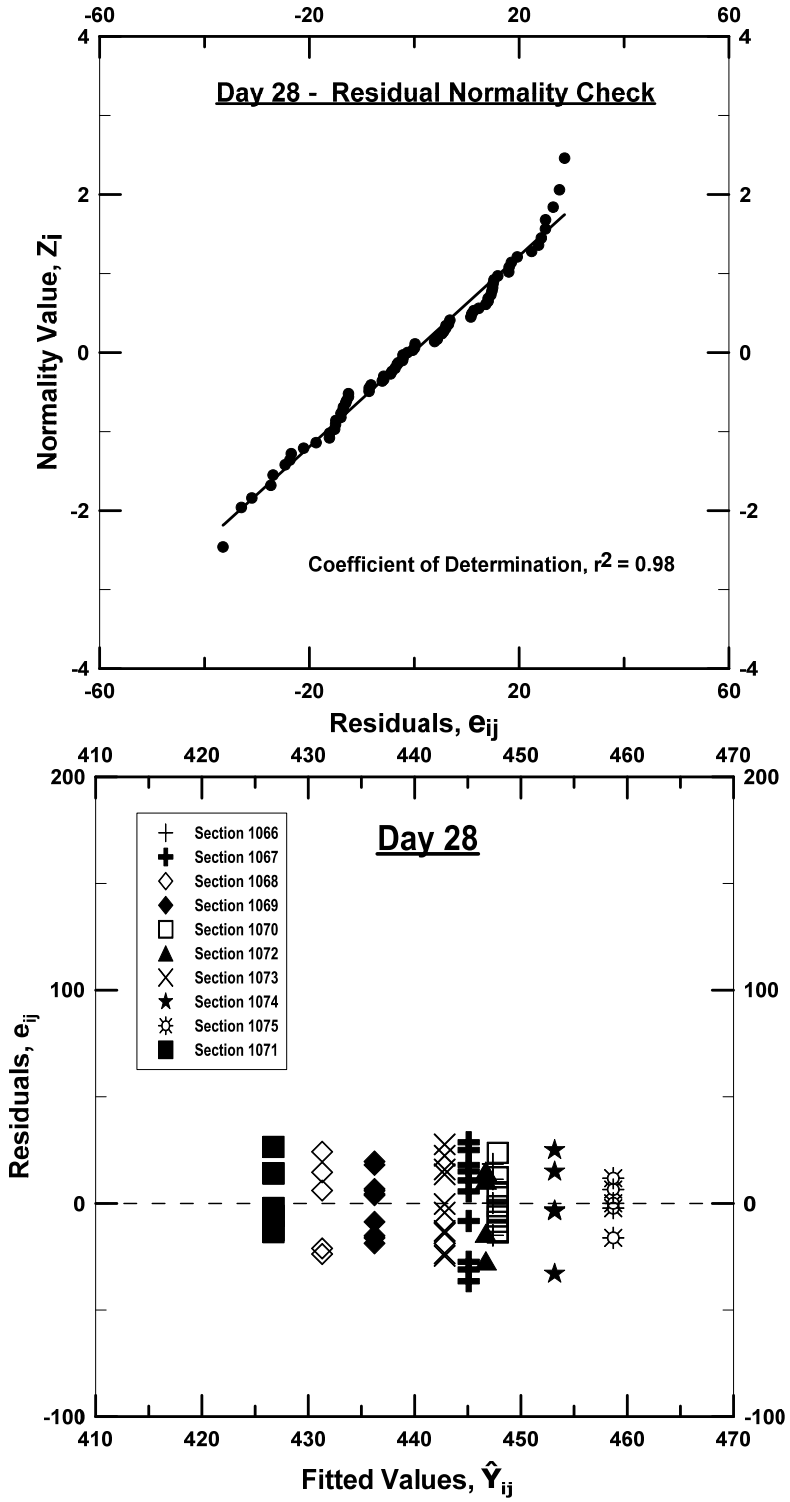


Figure 5-21 Model Adequacy plots for day 28 residual values

From the model adequacy plots, it can be inferred that the residuals in all the normality plots followed a linear trend. A best fit line was modeled using the linear regression approach, where the coefficient of determinations (r^2) for all the normality plots were greater than 0.90. Also, from visual inspection of the above plots, it can be inferred that the dispersion range of residuals in all the sections did not deviate much. Therefore, the residuals of all the stiffness values were normally and independently distributed, with no variance trends. This validated the conclusions obtained from the ANOVA test results.

5.5.2 Check for Constant Variance Value

The Bartlett's test was used to check for constant variance present in the stiffness values obtained from all the sections. This was performed by developing a hypothesis, as shown below:

$H_0 : \sigma^2_{1066} = \sigma^2_{1067} = \sigma^2_{1068} = \sigma^2_{1069} = \sigma^2_{1070} = \sigma^2_{1071} = \sigma^2_{1072} = \sigma^2_{1073} = \sigma^2_{1074} = \sigma^2_{1075}$

H_1 : At least one variance is different

The above hypothesis was evaluated using the statistic expressed in Equation 3.9. The value obtained from the statistic was compared to the critical value obtained at a significance value (α) 0.05. The chi-square distribution table was used to determine the critical value. Tables 5-19 to 5-23 provided the summary of the results obtained from the Bartlett's test in evaluation of constant variance for days 1, 3, 7, 14 and 28.

Table 5-19 Summary of Bartlett's test results for day 1 stiffness values

Parameter	**Result**
Q	5.65
C	1.14
Test Statistic ($\chi_0^2 = 2.3026\ \frac{q}{c}$)	11.3
Critical Region	16.9

Table 5.19 - *Continued*

x_1^2 (11.3) < $x_{1\ critical}^2$ (16.9)

Table 5-20 Summary of Bartlett's test results for day 3 stiffness values

Parameter	**Result**
Q	7.03
C	1.06
Test Statistic ($\chi_0^2 = 2.3026\ \frac{q}{c}$)	15.2
Critical Region	16.9
x_1^2 (15.2) < $x_{1\ critical}^2$ (16.9)	

Table 5-21 Summary of Bartlett's test results for day 7 stiffness values

Parameter	**Result**
Q	6.3
C	1.06
Test Statistic ($\chi_0^2 = 2.3026\ \frac{q}{c}$)	13.6
Critical Region	16.9
x_1^2 (13.6) < $x_{1\ critical}^2$ (16.9)	

Table 5-22 Summary of Bartlett's test results for day 14 stiffness values

Parameter	**Result**
Q	4.1
C	1.06

Table 5.22- *Continued*

Test Statistic ($\chi_0^2 = 2.3026\ \frac{q}{c}$)	8.91
Critical Region	16.9
x_i^2 (8.91) < $x_{i\ critical}^2$ (16.9)	

Table 5-23 Summary of Bartlett's test results for day 28 stiffness values

Parameter	**Result**
Q	4.2
C	1.06
Test Statistic ($\chi_0^2 = 2.3026\ \frac{q}{c}$)	9.10
Critical Region	16.9
x_i^2 (9.1) < $x_{i\ critical}^2$ (16.9)	

Based on the test results summarized in above Tables 5-19 to 5-23, it can be inferred that the statistic value (χ_0^2) was smaller than the critical region ($x_{i\ critical}^2$). Therefore, the null hypothesis was not rejected, concluding that the variances of stiffness values weree constant in all sections for days 1, 3, 7, 14 and 28. Hence, based on the ANOVA test and Bartlett's test, the stiffness values were stationary, having constant mean and variance.

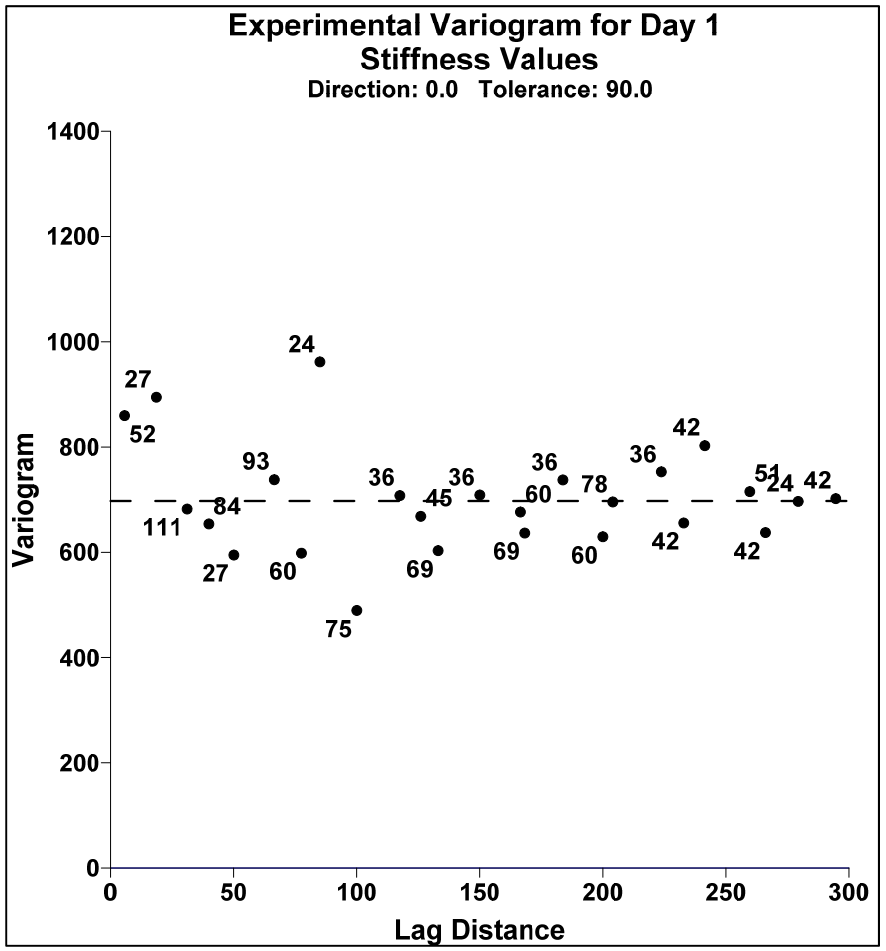
5.6 Geostatistical Analysis

Geostatistical analysis was performed in this study to predict the stiffness values of the CLSM at undetermined locations. This analysis included construction of

experimental variogram values, modelling of spatial variability in stiffness values for all days, and use kriging to perform predictions based on the spatial variability model.

5.6.1 Experimental Variogram

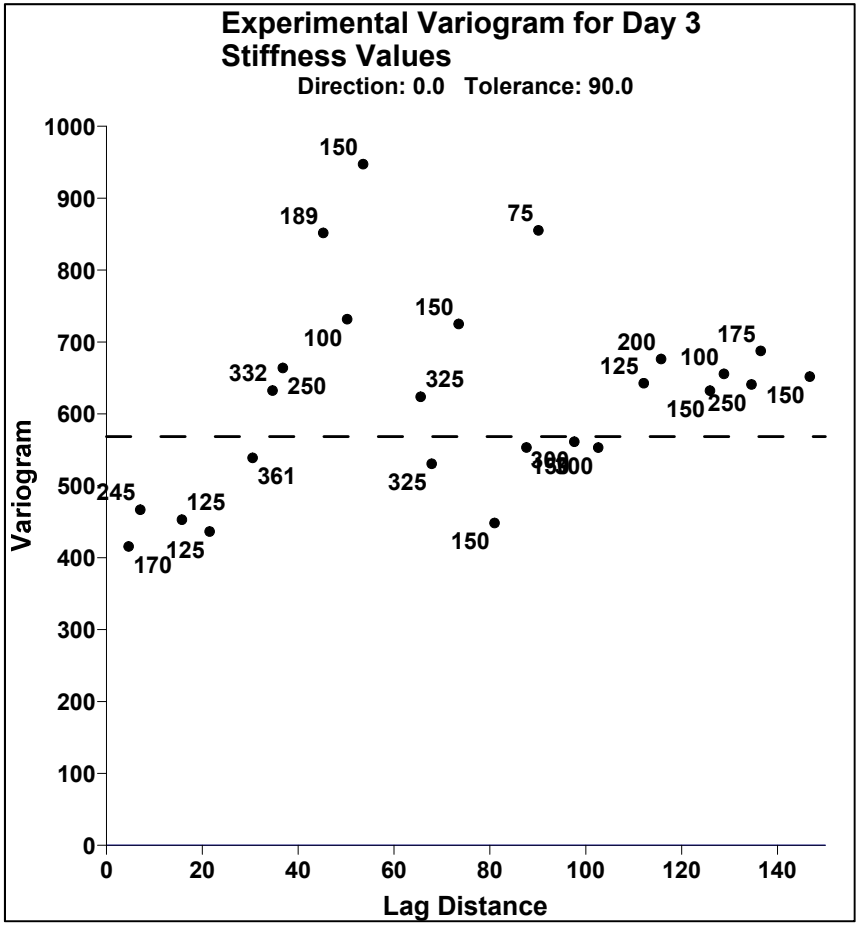
The stiffness values that were obtained in the 10 sections of the pipeline were used to construct the experimental variogram. The experimental variograms were constructed to identify the spatial correlations in the stiffness values in the bedding material. The grid parameters were selected based on the trial and error procedure until the maximum number of pairs was greater than or equal to 30.The semi-variogram or variogram values were calculated using Equation 3-18, which were plotted against the lag distance. Figures 5-22 to 5-26 present the experimental variogram plots for the stiffness values obtained on days 1, 3, 7, 14 and 28, respectively. The variogram values were plotted on the y-axis, and the corresponding lag distance value was plotted on the x-axis.



Grid Parameters for day 1

- **Estimator type : Variogram**
- **Maximum lag distance : 300**
- **Number of lags : 25**
- **Lag Width :20**
- **Direction of Variogram :0 degrees (x-direction)**
- **Tolerance : 90 degrees**

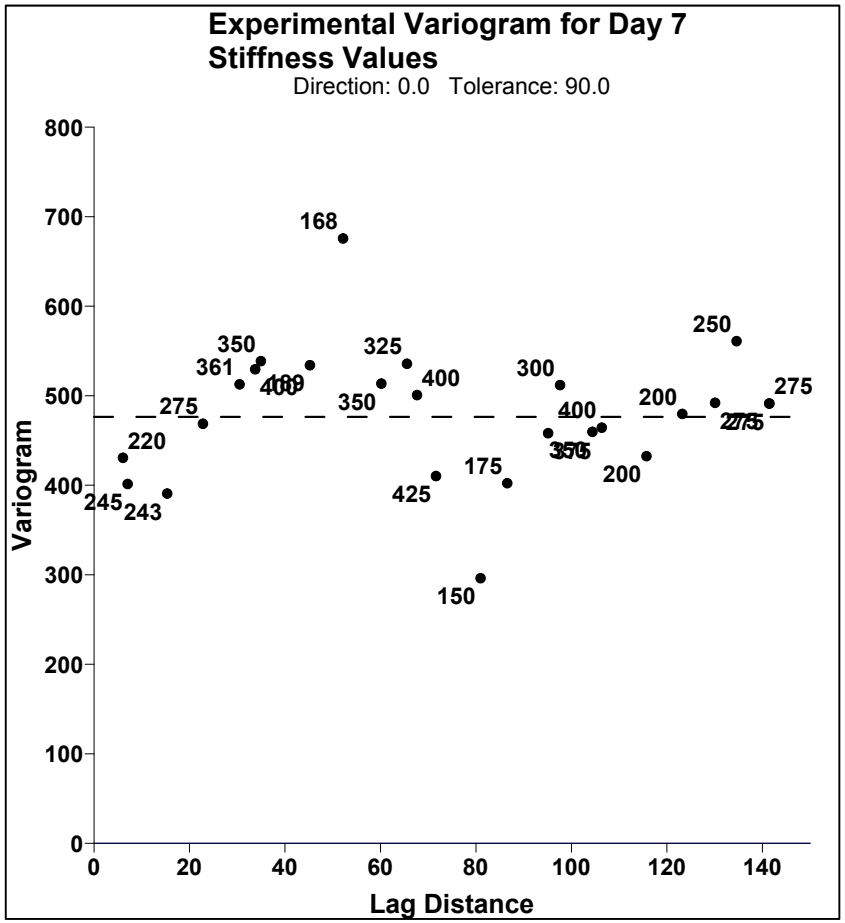
Figure 5-22 Experimental variogram value for day 1 stiffness values



Grid Parameters for day 3 :

- **Estimator type : Variogram**
- **Maximum lag distance : 150**
- **Number of lags : 25**
- **Lag Width :20**
- **Direction of Variogram : 0 degrees (x-direction)**
- **Tolerance : 90 degrees**

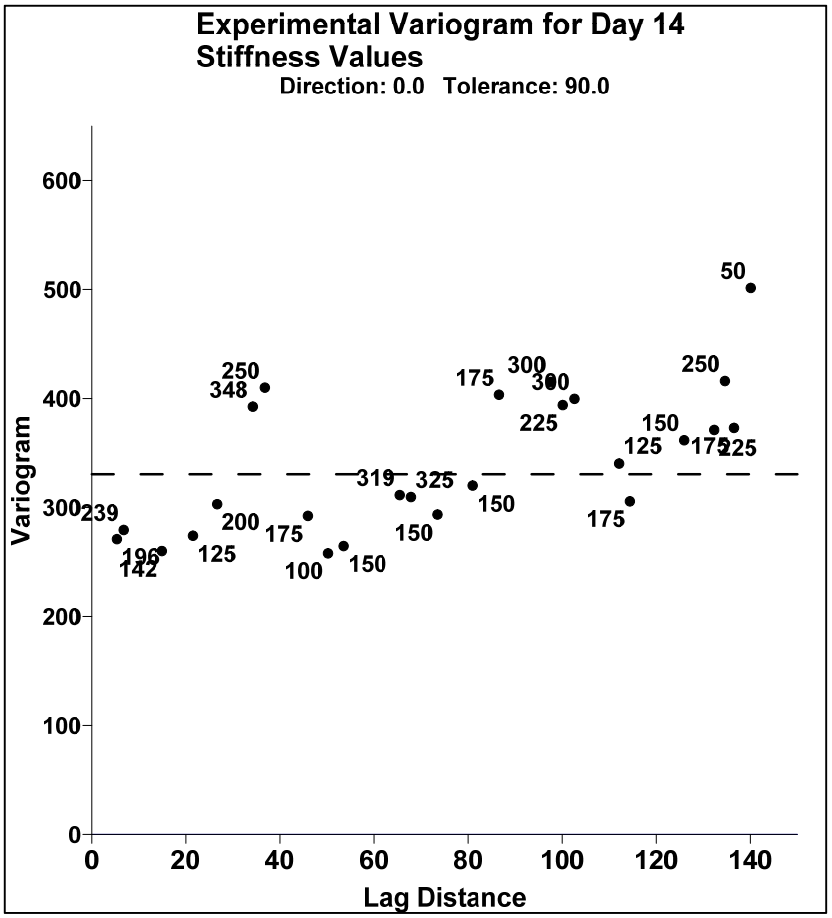
Figure 5-23 Experimental variogram plot for day 3 stiffness values



Grid Parameters for day 7:

- **Estimator type : Variogram**
- **Maximum lag distance : 150**
- **Number of lags : 25**
- **Lag Width :20**
- **Direction of Variogram : 0 degrees (x-direction)**
- **Tolerance : 90 degrees**

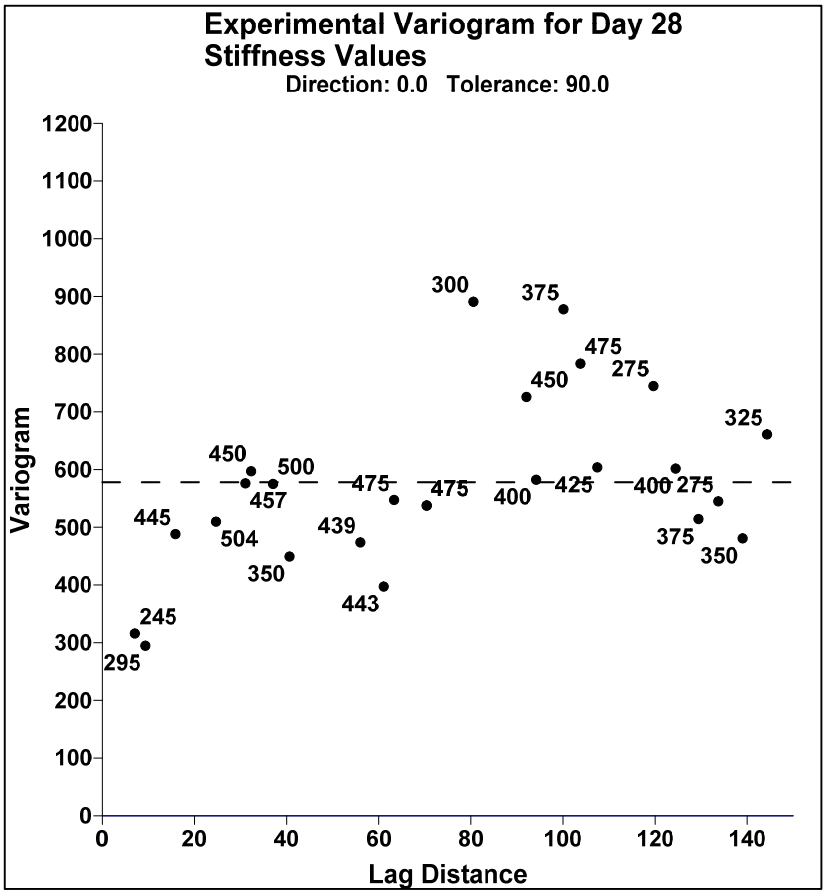
Figure 5-24 Experimental variogram plot for day 7 stiffness values



Grid Parameters for day 14:

- **Estimator type : Variogram**
- **Maximum lag distance : 150**
- **Number of lags : 25**
- **Lag Width :15**
- **Direction of Variogram : 0 degrees (x-direction)**
- **Tolerance : 90 degrees**

Figure 5-25 Experimental variogram plot for day 14 stiffness values



Grid parameters for day 28:

- **Estimator type : Variogram**
- **Maximum lag distance : 150**
- **Number of lags : 25**
- **Lag Width :30**
- **Direction of Variogram : 0 degrees (x-direction)**
- **Tolerance : 90 degrees**

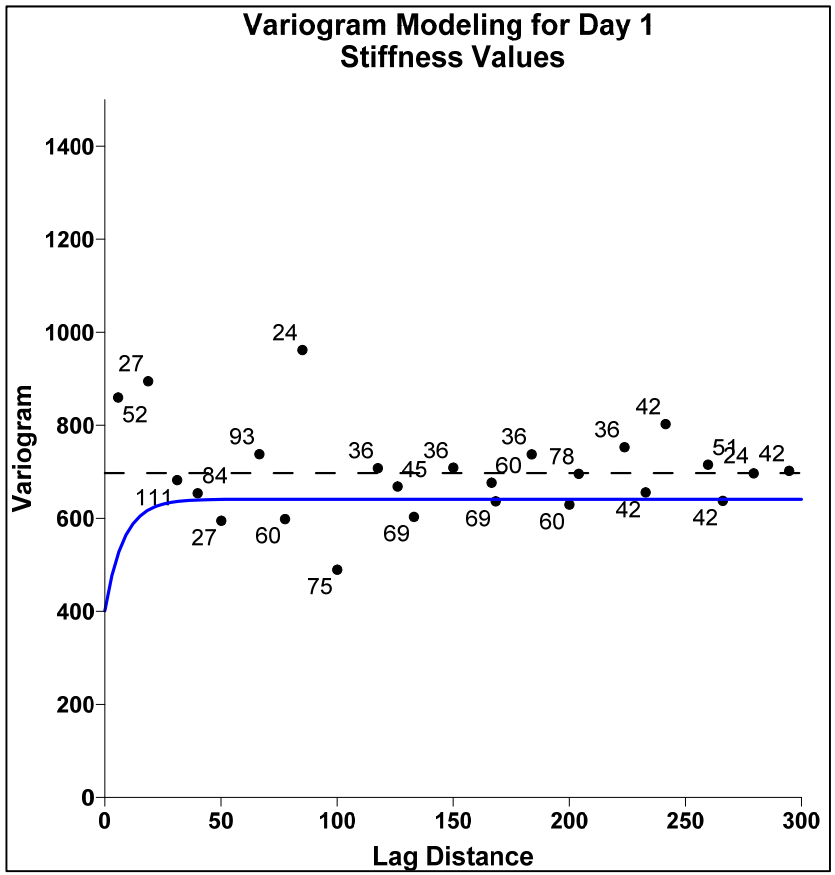
Figure 5-26 Experimental variogram plot for day 28 stiffness values

The above experimental variogram plots were used to identify the spatial correlation in stiffness values after a curing period of 1, 3, 7, 14 and 28 days. The grid parameters selected in the plots were almost the same, satisfying the number of pairs for calculating an experimental variogram value. It was assumed that the x-direction in this study represented the longitudinal distance, i.e., 500 ft., and the y-direction represented the transverse distance, i.e., 9 ft., which was the diameter of the pipe used at the site. Since, the measurements obtained through the SASW technique were in longitudinal distance, the direction of the variograms was selected as 0 degrees. The maximum lag distance that was used in constructing all the variograms was 150m except for day 1. This is because the CLSM had not developed uniform stiffness throughout the pipeline, which was also evident from the standard deviation of 29.2 MPa from the histogram plots. Hence, to identify a spatial correlation the experimental variogram plots were constructed for a lag distance of 300m.

5.6.2 Variogram Modeling

In order to capture the spatial variability in stiffness values, the experimental variograms were modelled using various models, as mentioned in section 3.7.2. From the experimental variogram plots, it can be observed that the variogram values did not show any specific trend. This can be attributed to low variability in stiffness values, even with an increase in the lag distance. However, all the plots reached sill value, which is lower than the global variance. This behavior was modelled in this study.

Also, at a lag distance of 0 m, the variogram values were not equal to zero, which resembled a nugget effect model present in the stiffness values. Figures 5-27 to 5-31 present the variogram models that were used to define the spatial variability present in the stiffness values for days 1, 3, 7, 14 and 28, respectively.



Variogram model for day 1:

- **Model : Nugget Effect + Exponential**
- **Scale : 240**
- **Length (A) : 8**
- **Nugget (C_0) : 400**

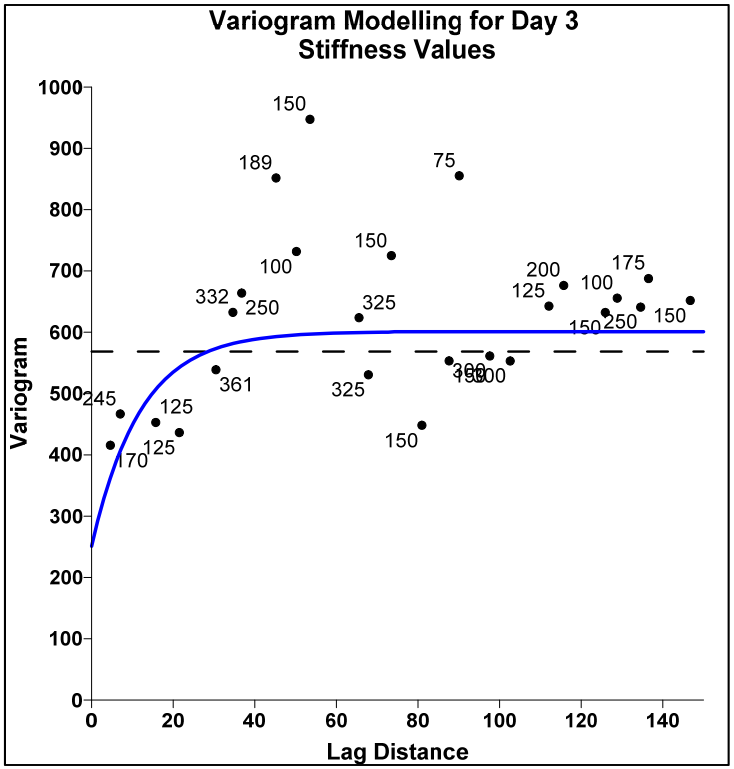
Theoretical Equation:

$\gamma(h) = C_0 + C\,[1 - \exp(-h/a)]$ for $h > 0$

Model Equation :

$\gamma(h) = 400 + 240\,[1 - \exp(-h/8)]$ for $h > 0$

Figure 5-27 Variogram model for day 1 stiffness values



Variogram model for day 3:

- **Model : Nugget Effect + Exponential**
- **Scale : 350**
- **Length (A) : 12**
- **Nugget (C_0) : 250**

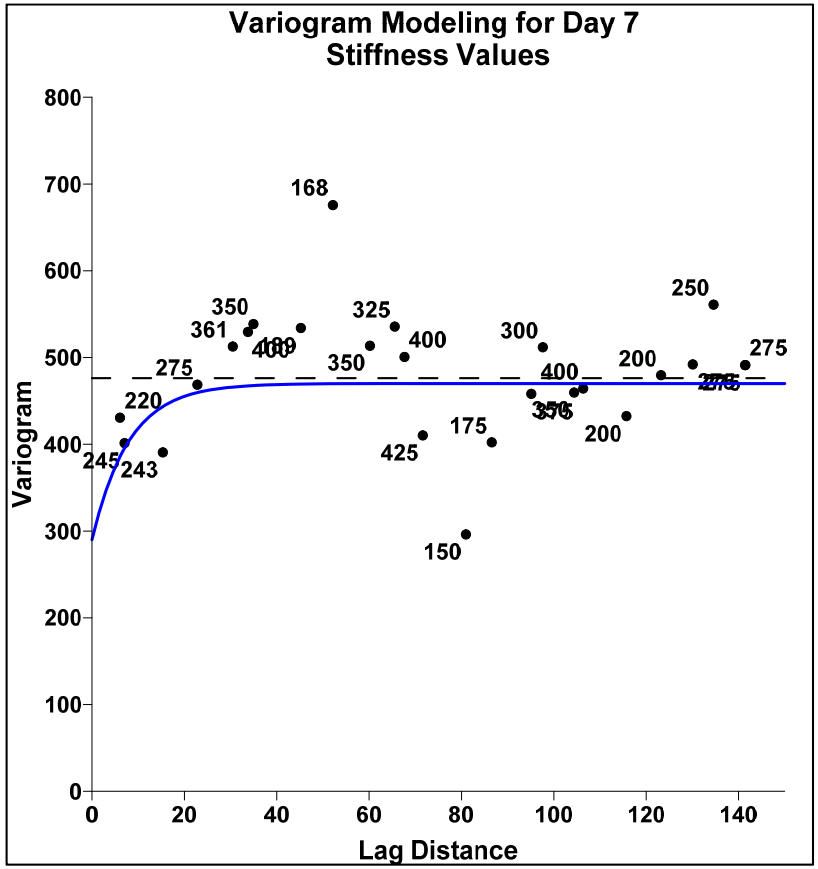
Theoretical Equation:

$\gamma(h) = C_0 + C\,[1 - \exp(-h/a)]$ for $h > 0$

Model Equation :

$\gamma(h) = 250 + 350\,[1 - \exp(-h/12)]$ for $h > 0$

Figure 5-28 Variogram model for day 3 stiffness values



Variogram model for day 7:

- **Model : Nugget Effect + Exponential**
- **Scale : 180**
- **Length (A) : 14**
- **Nugget (C_0) : 290**

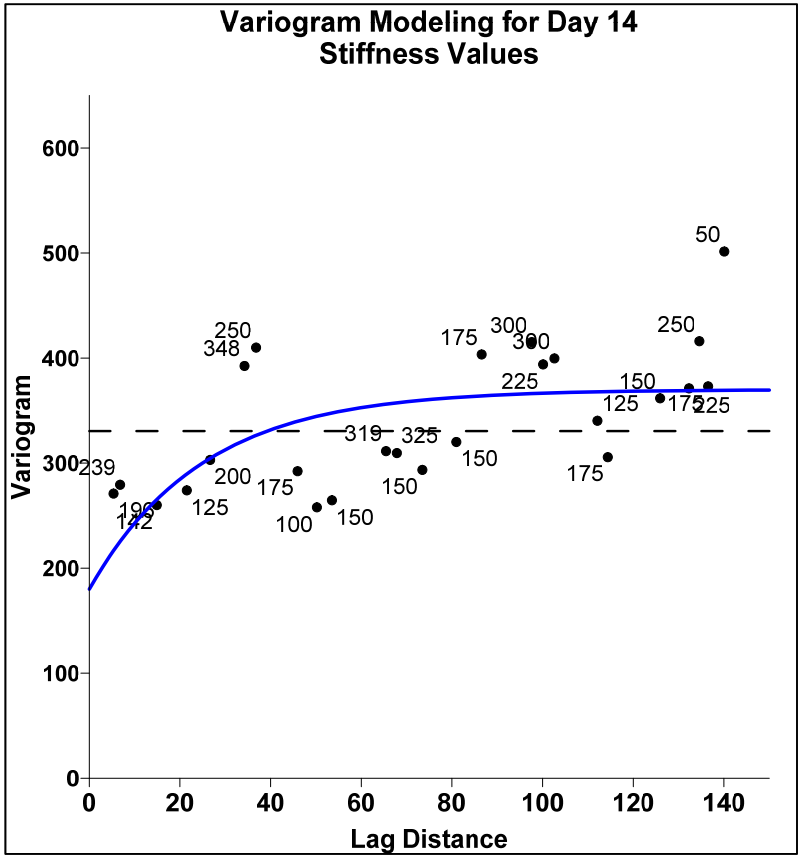
Theoretical Equation:

$\gamma(h) = C_0 + C\ [\ 1 - \exp(-h/a)]$ for $h > 0$

Model Equation :

$\gamma(h) = 290 + 180\ [\ 1 - \exp(-h/14)]$ for $h > 0$

Figure 5-29 Variogram model for day 7 stiffness values



Variogram model for day 14:

- **Model : Nugget Effect + Exponential**
- **Scale : 190**
- **Length (A) : 23**
- **Nugget (C_0) : 180**

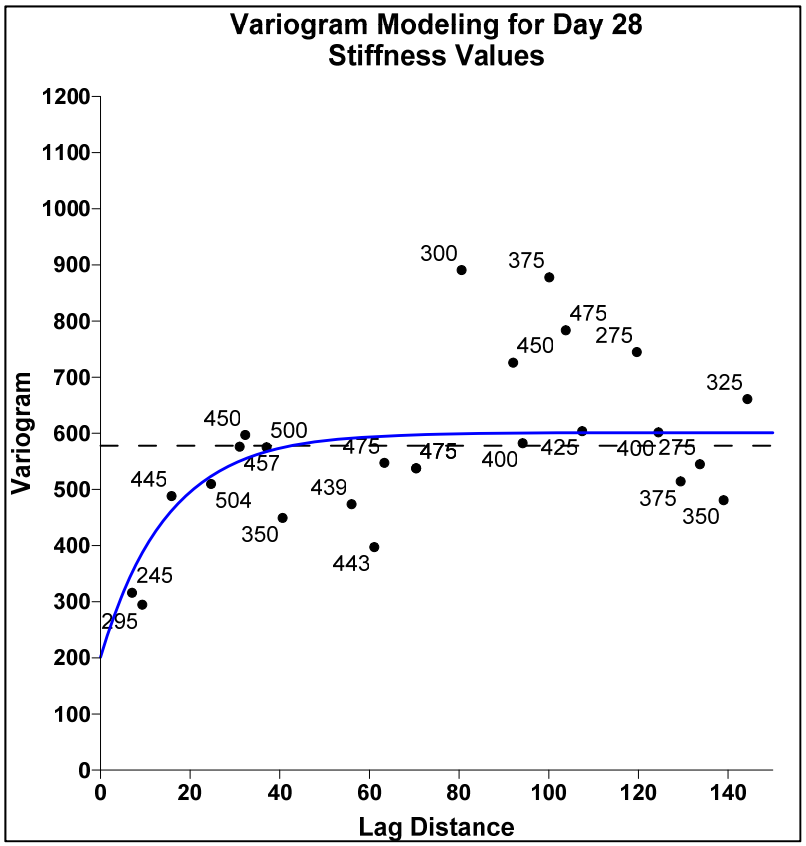
Theoretical Equation:

$$\gamma(h) = C_0 + C\,[1 - \exp(-h/a)] \text{ for } h > 0$$

Model Equation :

$$\gamma(h) = 180 + 190\,[1 - \exp(-h/23)] \text{ for } h > 0$$

Figure 5-30 Variogram model for day 14 stiffness values



Variogram model for day 28:

- **Model : Nugget Effect + Exponential**
- **Scale : 400**
- **Length (A) : 24**
- **Nugget (C_0) : 200**

Theoretical Equation:

$\gamma(h) = C_0 + C\,[1 - \exp(-h/a)]$ for $h > 0$

Model Equation :

$\gamma(h) = 200 + 400\,[1 - \exp(-h/24)]$ for $h > 0$

Figure 5-31 Variogram model for day 28 stiffness values

The spatial variability in the stiffness values for days 1, 3, 7, 14 and 28 were modeled as shown in figures above. With an increase in lag distance the experimental variogram increased and approached the sill value. The nugget effect was modelled to describe the variation of CLSM stiffness at close distances, such as south 1, south 2 and north 1, north 2.

It was observed that the exponential models with a nugget effect were the best fit models for describing the spatial variability in stiffness values for all days. The minimum spatial correlation distance was observed on day 1 with 8m, and the maximum correlation distance was observed on day 28, with a distance of 24m. Therefore, it can be generalized that the correlation distance in stiffness values varied from 8-24m in the CLSM bedding material. However, this might differ from project to project, depending on the field conditions.

5.6.3 Kriging Analysis

The histogram plots shown earlier depicted significant variations in stiffness values, with a standard deviation ranging from 18 to 30 MPa. The dispersion in the stiffness values in the individual sections was high enough that the normal prediction results would be erratic. In this study, the variability in stiffness values were captured by modeling the variogram as presented in Section 5.6.2. The spatial variability models developed for all the days were used to predict the stiffness values at unsampled locations. The predictions were performed using the geostatistical tool ‘kriging’. The kriging uses the spatial correlation distance, ranging from 8-25m, obtained from the spatial variability model.

Using the correlation distance and kriging algorithm, the weights of the neighboring values were determined. The grid parameters that were used for the kriging

analysis were provided in Table 5-24, and the spatial variability model parameters were summarized in Table 5-25.

Table 5-24 Grid parameters selected for kriging analysis

Parameters selected					
Days	1	3	7	14	28
Maximum lag distance	300	150	150	150	150
Angular divisions	180	180	180	180	180
Radial divisions	100	100	100	100	100

Table 5-25 Spatial variability model parameters

Parameters selected					
Days	1	3	7	14	28
Correlation distance	8	12	14	23	24
Nugget effect	400	250	290	180	200

The above grid and spatial variability parameters were used to predict the stiffness values for days 1, 3, 7, 14 and 28, respectively. Figures 5-31 to 5-35 present the stiffness contour maps predicted using kriging analysis.

Day 1: Kriging analysis

Spatial variability model : **γ (h) = 400 + 240 [1- exp (-h/8)] for h > 0**

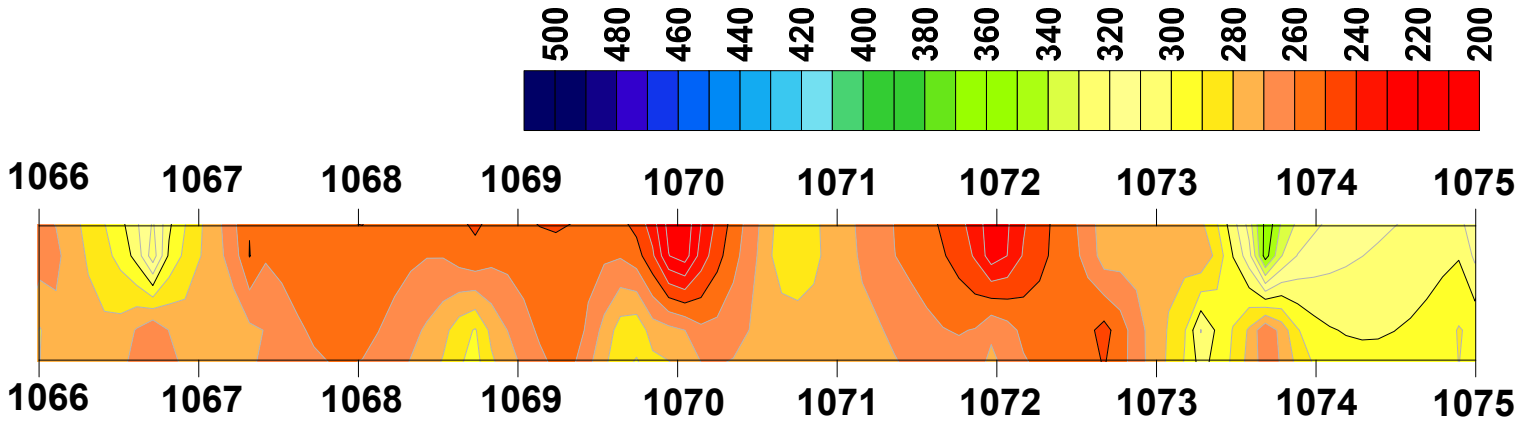


Figure 5-32 Kriging analysis map for stiffness (MPa) of CLSM after 1 day curing period

Day 3: Kriging analysis

Spatial variability model : **γ (h) = 250 + 350 [1- exp (-h/12)] for h > 0**

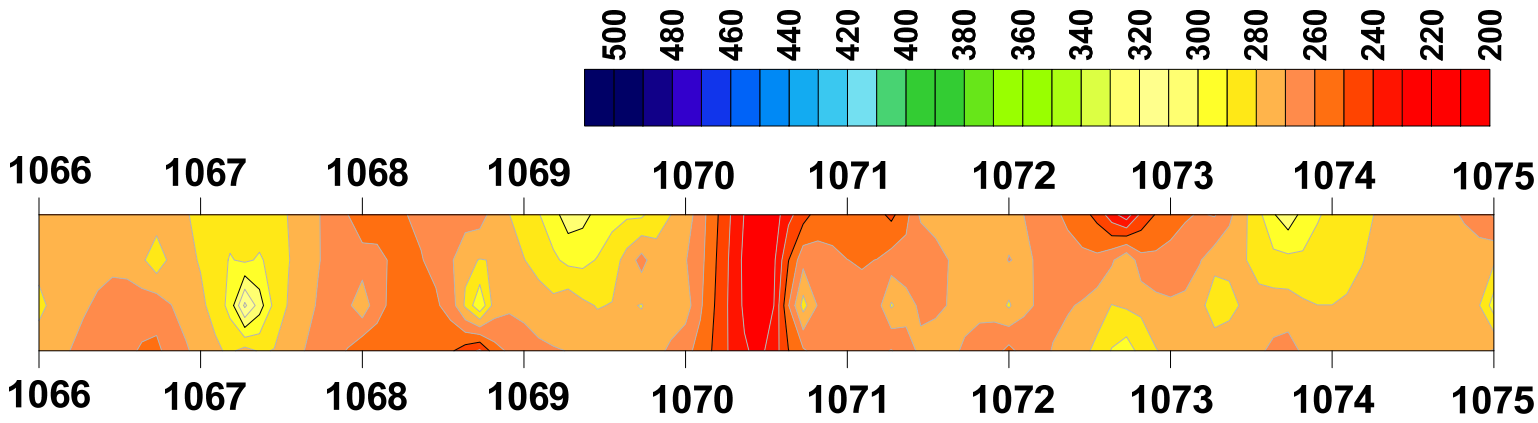


Figure 5-33 Kriging analysis map for stiffness (MPa) of CLSM after 3 days curing period

Day 7: Kriging analysis

Spatial variability model : **γ (h) = 290 + 180 [1- exp (-h/14)] for h > 0**

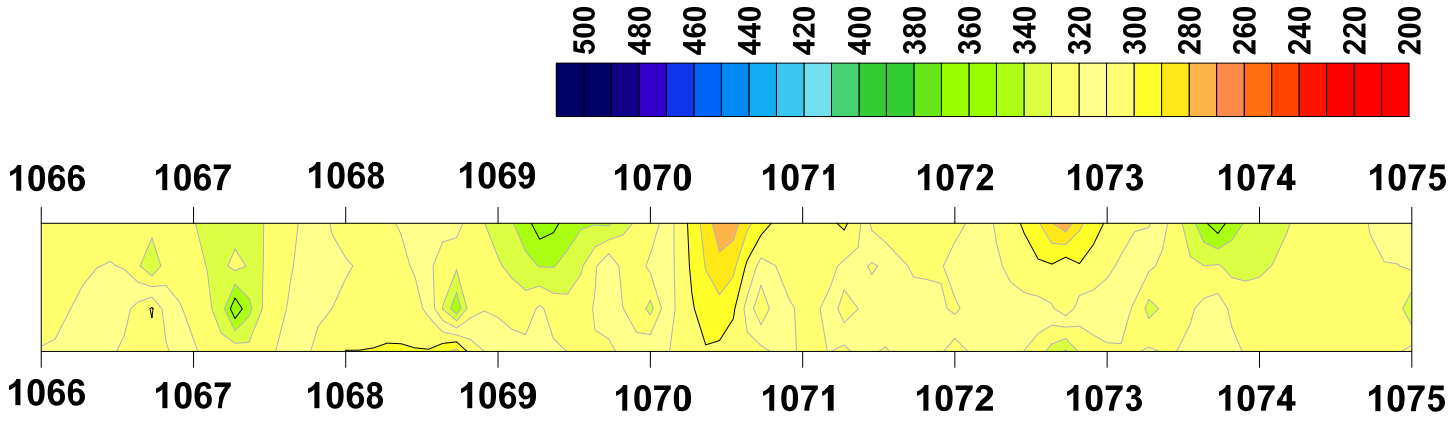


Figure 5-34 Kriging analysis map for stiffness (MPa) of CLSM after 7 days curing period

Day 14: Kriging analysis

Spatial variability model : **γ (h) = 180 + 190 [1- exp (-h/23)] for h > 0**

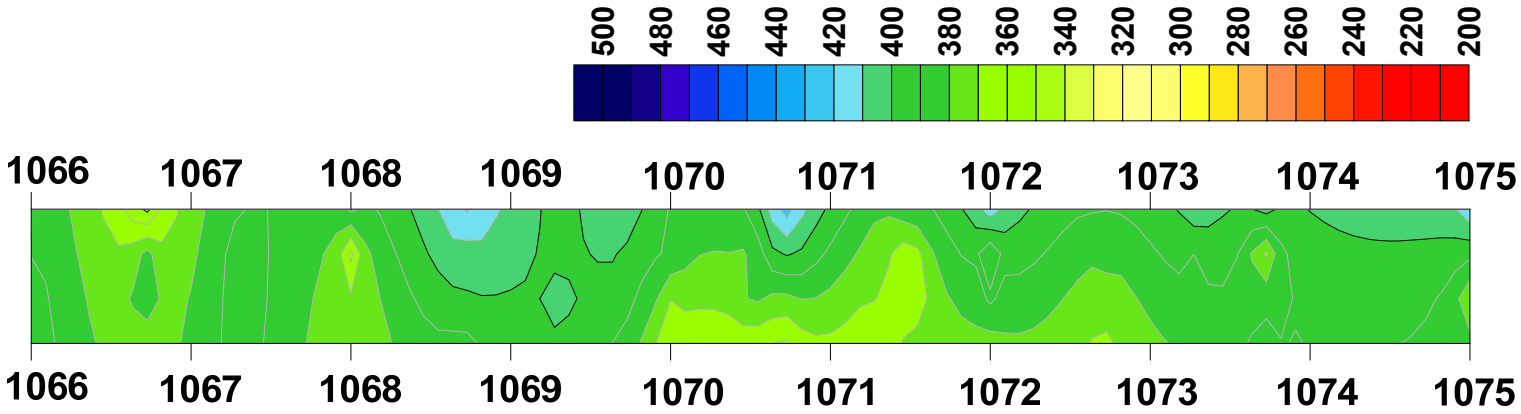


Figure 5-35 Kriging analysis map for stiffness (MPa) of CLSM after 14 days curing period

Day 28: Kriging analysis

Spatial variability model : **γ (h) = 200 + 400 [1- exp (-h/24)] for h > 0**

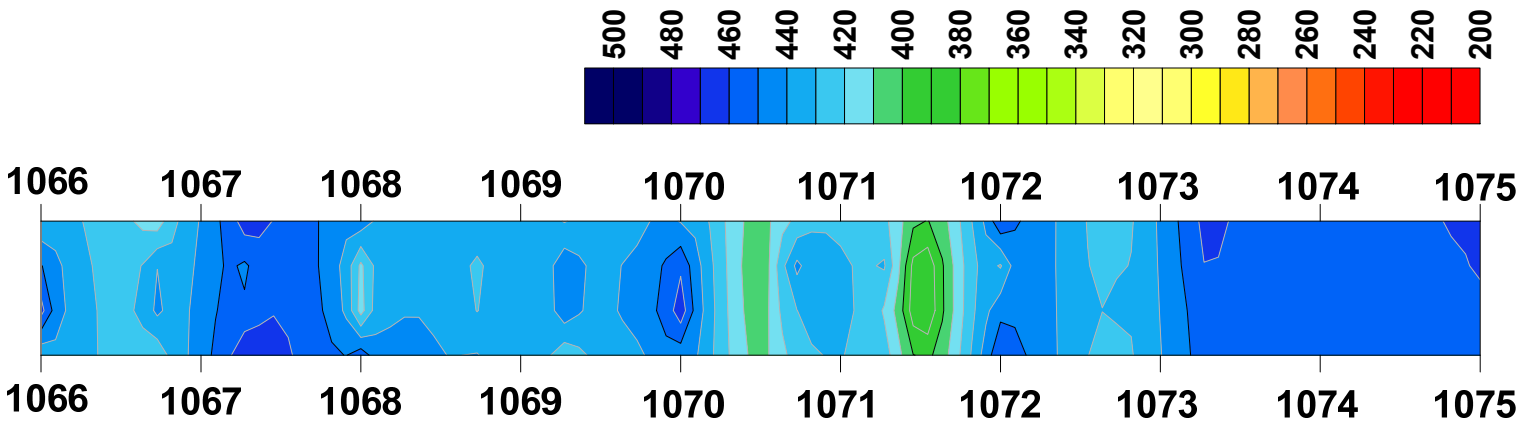


Figure 5-36 Kriging analysis map for stiffness (MPa) of CLSM after 28 days curing period

The above kriging maps show the stiffness values of CLSM along the 500 ft., pipeline section for days 1, 3, 7, 14 and 28, respectively. The maps provided represent the top view of the pipeline with a diameter 9 ft., which is in transverse direction and 500 ft. in length, which is in longitudinal direction, a total of 10 sections. All the predictions were unbiased and obtained by capturing the spatial variability of the stiffness values.

In order to observe increase in stiffness of CLSM, the color scale for the maps generated for days 1, 3, 7, 14 and 28 was kept the same. For day 1, the contour map depicts the stiffness values in the range of 200-250 MPa; and for day 3, the CLSM achieved a stiffness range of 260–280 MPa. For day 7, the stiffness value ranged from 300-320 MPa; for day 14, the stiffness values increased and ranged from 340-400 MPa; and by day 28, the CLSM achieved a stiffness value of 400-480 MPa. From the above maps, it can be inferred that there was uniform increase and development of the stiffness values along the pipeline for all the days.

5.6.4 Cross Validation

The contour maps generated in the earlier sections provided good insight into the spatial variation of stiffness of CLSM. However, to ensure that the spatial variability model generated was correct, cross validation was performed. Cross validation refers to the process of deleting the original values and predicting them again with the spatial variability model. In this study, a few stiffness values were randomly deleted on each day, and kriging analysis was the performed again, using the corresponding spatial variability model. The stiffness of CLSM on new kriging maps at any specific location was determined using a digitizing tool available on Surfer software. Then, the originally deleted values were compared with the predicted values on a 45 degree validation comparison plot. Figures 5-36 to 5-40 represent the new stiffness contour map

generated. The black rectangular boxes in the maps represent the locations of the deleted value.

Day 1: Cross-validation using kriging Analysis

Spatial variability model : **γ (h) = 400 + 240 [1- exp (-h/8)] for h > 0**

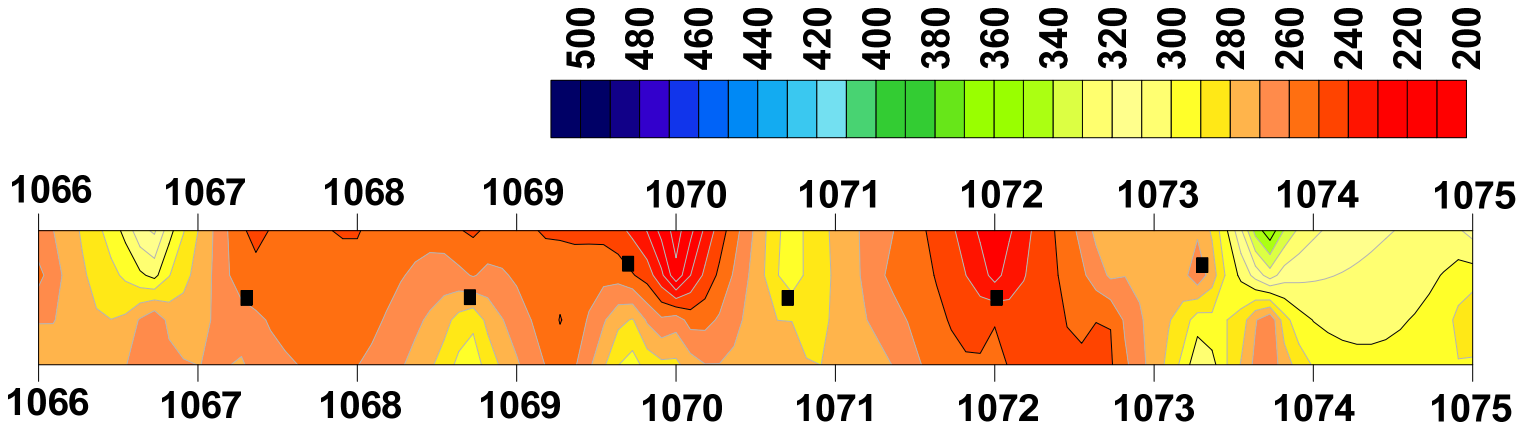


Figure 5-37 Cross validation map for stiffness (MPa) of CLSM after 1 day curing period

Using the digitize tool available in the Surfer software, the precise value of the stiffness values were obtained at the deleted locations. Table 5-26 and Figure 5-38 below provide the comparison of actual and predicted values from the cross-validation map.

Table 5-26 Comparison of actual and predicted Stiffness values

S.no	Actual Values (MPa)	Predicted Values (MPa)
1	262.4	258.4
2	271.3	273.7
3	258.6	248.4
4	281.5	286.6
5	258.5	237.8
6	276.2	269.0

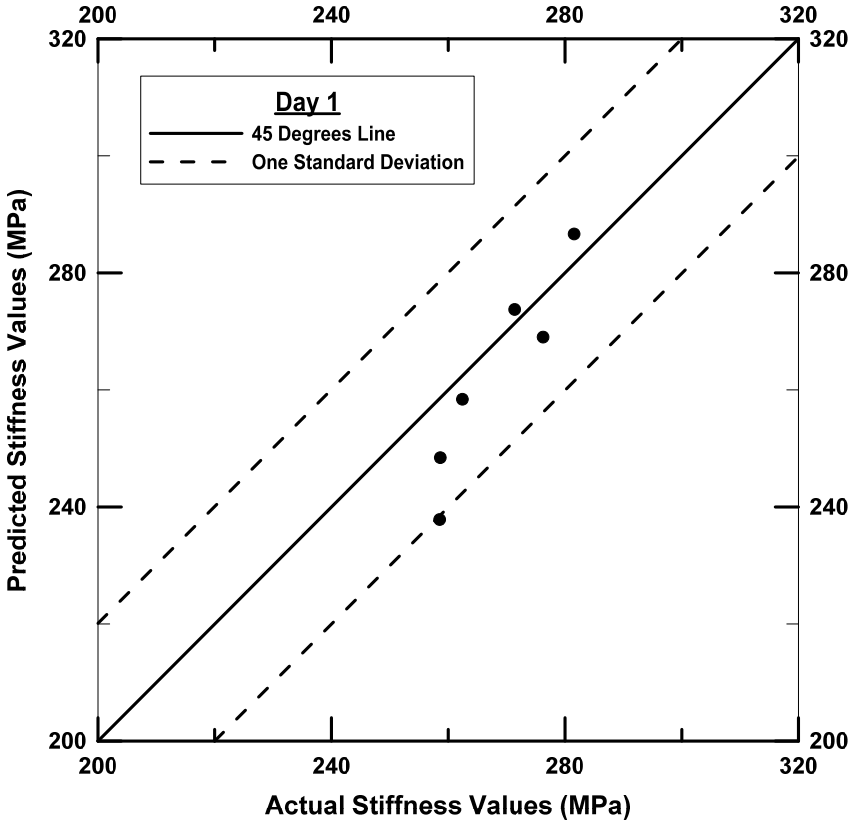


Figure 5-38 Comparison of actual and predicted stiffness values on day 1

Day 3: Cross-validation using kriging Analysis

Spatial variability model : **γ (h) = 250 + 350 [1- exp (-h/12)] for h > 0**

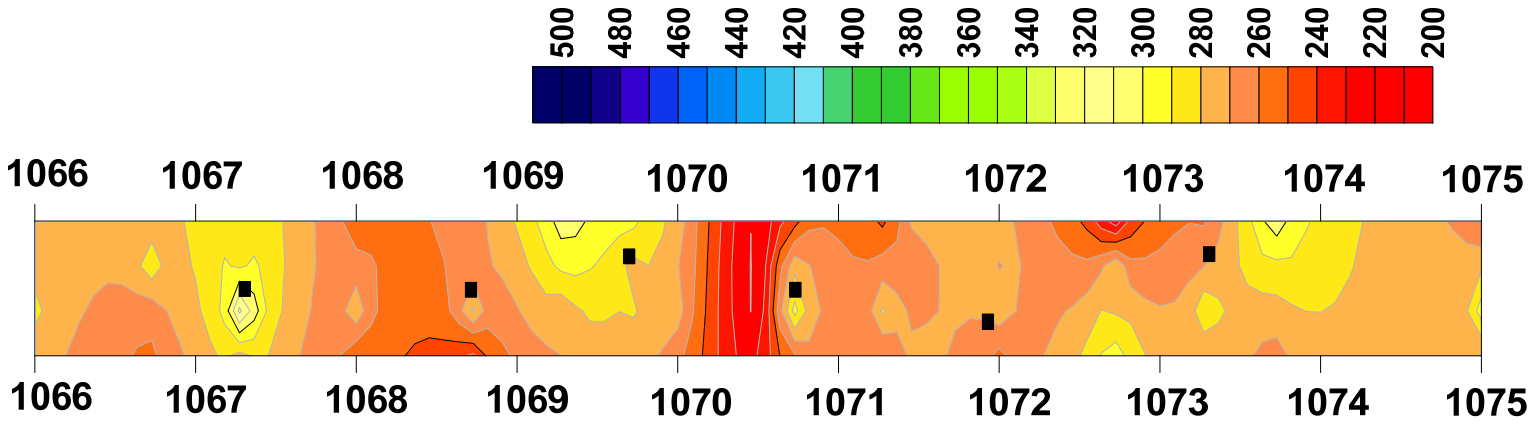


Figure 5-39 Cross validation map for stiffness (MPa) of CLSM after 3 days curing period

The precise value of the stiffness values was obtained using the digitize tool available in the Surfer software at the deleted locations. Table 5-27 and Figure 5-40 below provide the comparison of actual and predicted values from the cross-validation map.

Table 5-27 Comparison of actual and predicted Stiffness values for day 3

S.no	Actual Values (MPa)	Predicted Values (MPa)
1	301.7	302.9
2	288.2	271.7
3	272.4	281.6
4	275.2	285.3
5	272.6	270.1
6	272.9	268.8

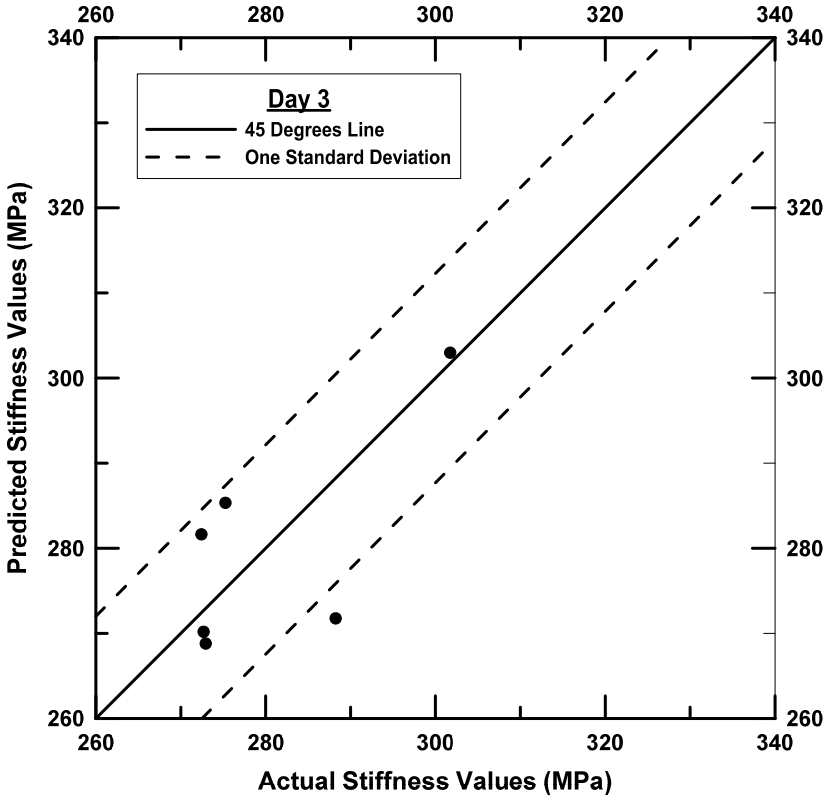


Figure 5-40 Comparison of actual and predicted stiffness values on day 3

Day 7: Cross-validation using kriging Analysis

Spatial variability model : **γ (h) = 290 + 180 [1- exp (-h/14)] for h > 0**

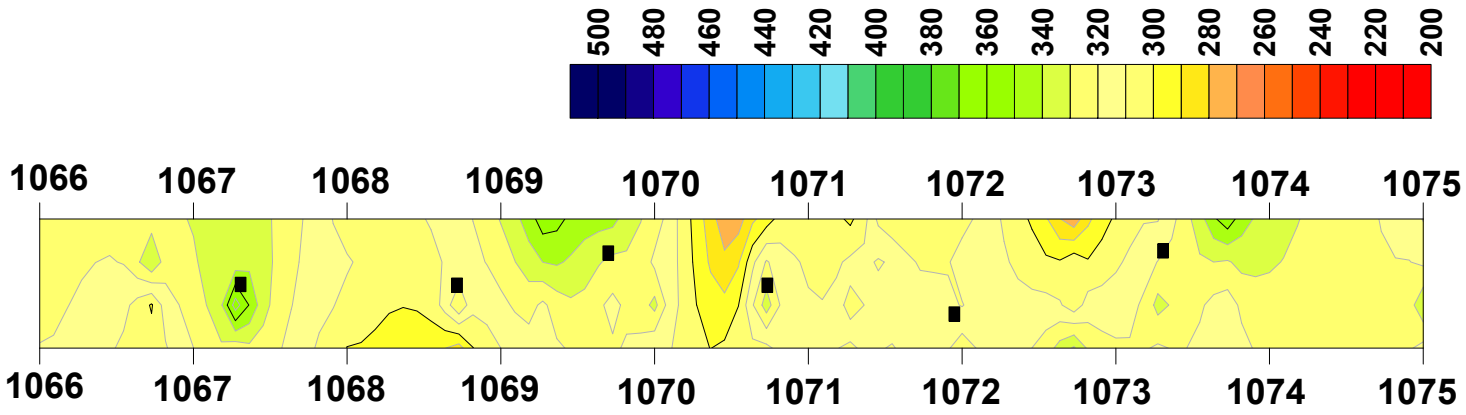


Figure 5-41 Cross validation map for stiffness (MPa) of CLSM after 7 days curing period

The precise value of the stiffness values were obtained using the digitize tool available in the Surfer software at the deleted locations. Table 5-28 and Figure 5-42 below provide the comparison of actual and predicted values from the cross-validation map.

Table 5-28 Comparison of actual and predicted Stiffness values for day 7

S.no	Actual Values (MPa)	Predicted Values (MPa)
1	339.5	346.2
2	336.1	319.0
3	321.7	331.8
4	319.2	326.6
5	317.4	317.0
6	316.7	319.8

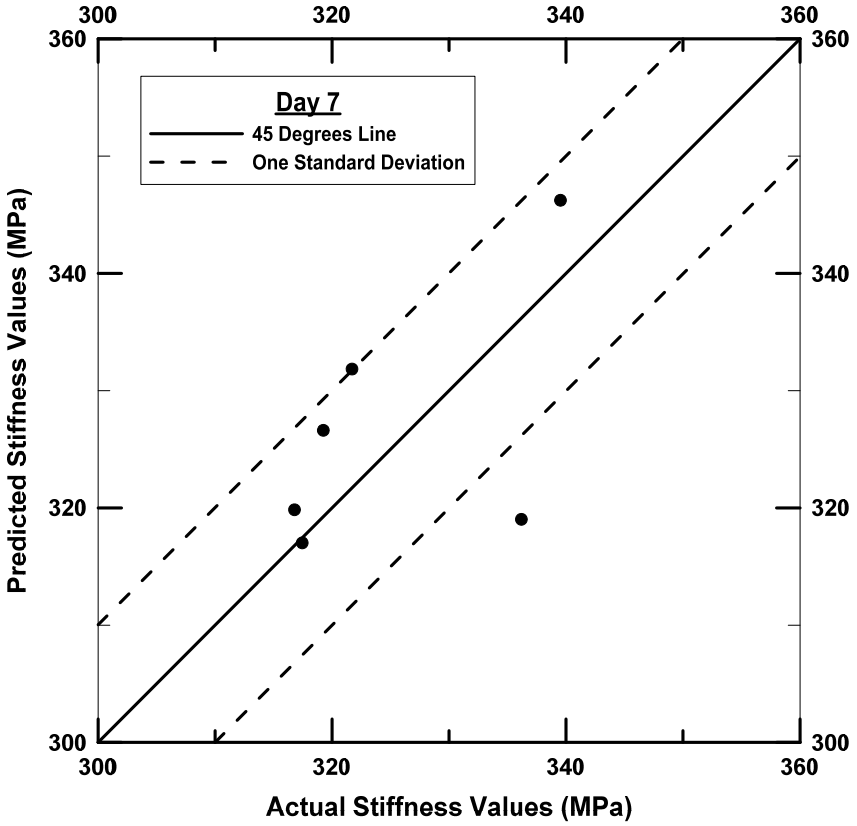


Figure 5-42 Comparison of actual and predicted stiffness values on day 7

Day 14: Cross-validation using kriging Analysis

Spatial variability model : **γ (h) = 180 + 190 [1- exp (-h/23)] for h > 0**

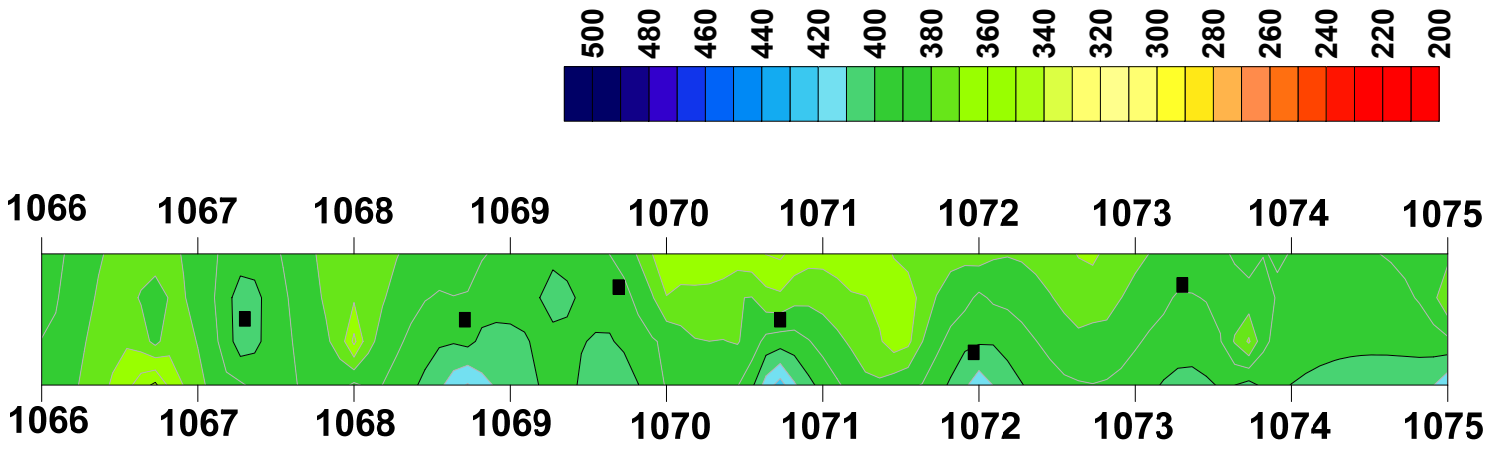


Figure 5-43 Cross validation map for stiffness (MPa) of CLSM after 14 days curing period

The precise value of the stiffness values were obtained using the digitize tool available in the Surfer software at the deleted locations. Table 5-29 and Figure 5-44 below provide the comparison of actual values predicted values from the cross-validation map.

Table 5-29 Comparison of actual and predicted Stiffness values for day 14

S.no	Actual Values (MPa)	Predicted Values (MPa)
1	393.8	405.1
2	402.5	393.9
3	392.9	392.5
4	391.9	383.8
5	395.0	401.7
6	386.0	388.6

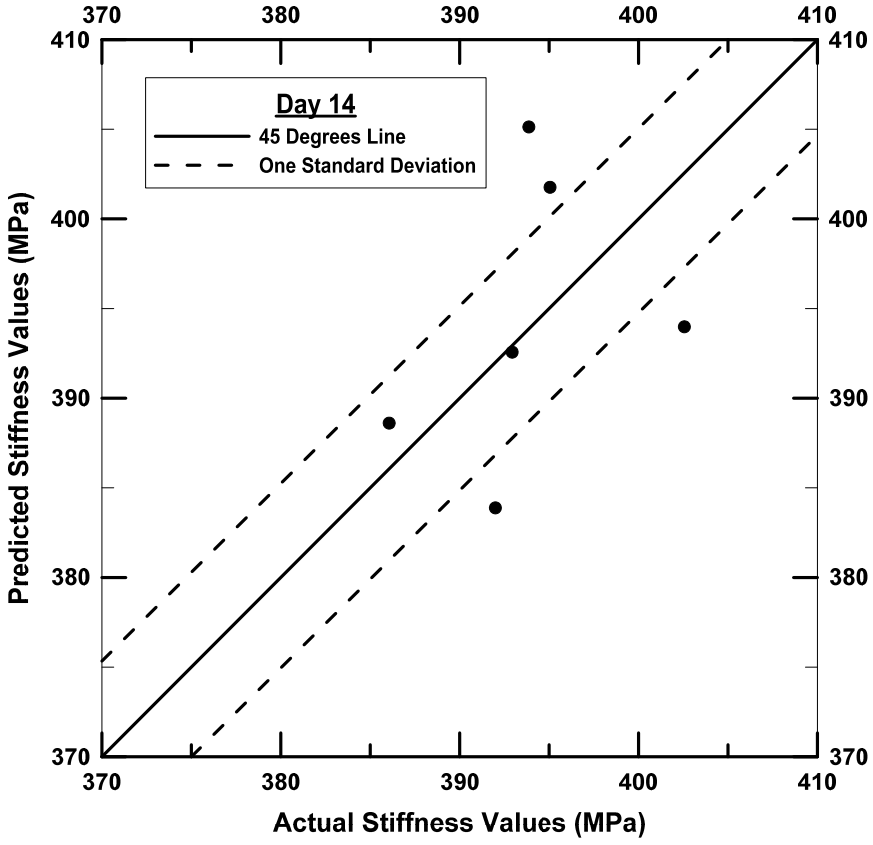


Figure 5-44 Comparison of actual and predicted stiffness values on day 14

Day 28: Cross-validation using kriging Analysis

Spatial variability model : **γ (h) = 200 + 400 [1- exp (-h/24)] for h > 0**

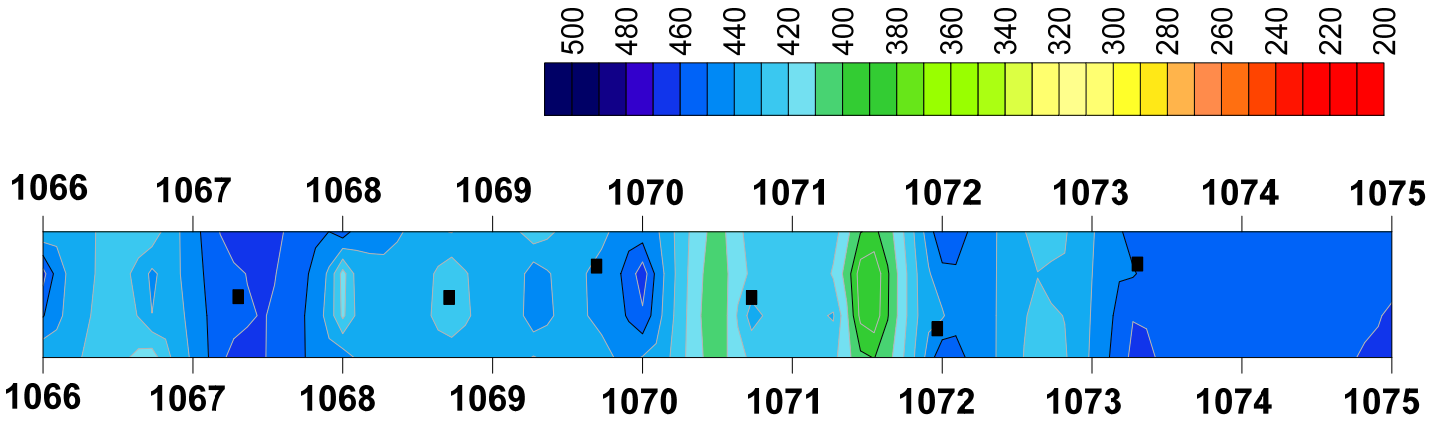


Figure 5-45 Cross validation map for stiffness (MPa) of CLSM after 28 days curing period

The precise value of the stiffness values were obtained using the digitize tool available in the Surfer software at the deleted locations. Table 5-30 and Figure 5-46 below provide the comparison of actual values predicted values from the cross-validation map.

Table 5-30 Comparison of actual and predicted Stiffness values for day 28

S.no	Actual Values (MPa)	Predicted Values (MPa)
1	452.2	456.1
2	429.0	422.4
3	442.0	442.4
4	435.6	429.2
5	436.6	442.5
6	454.2	451.5

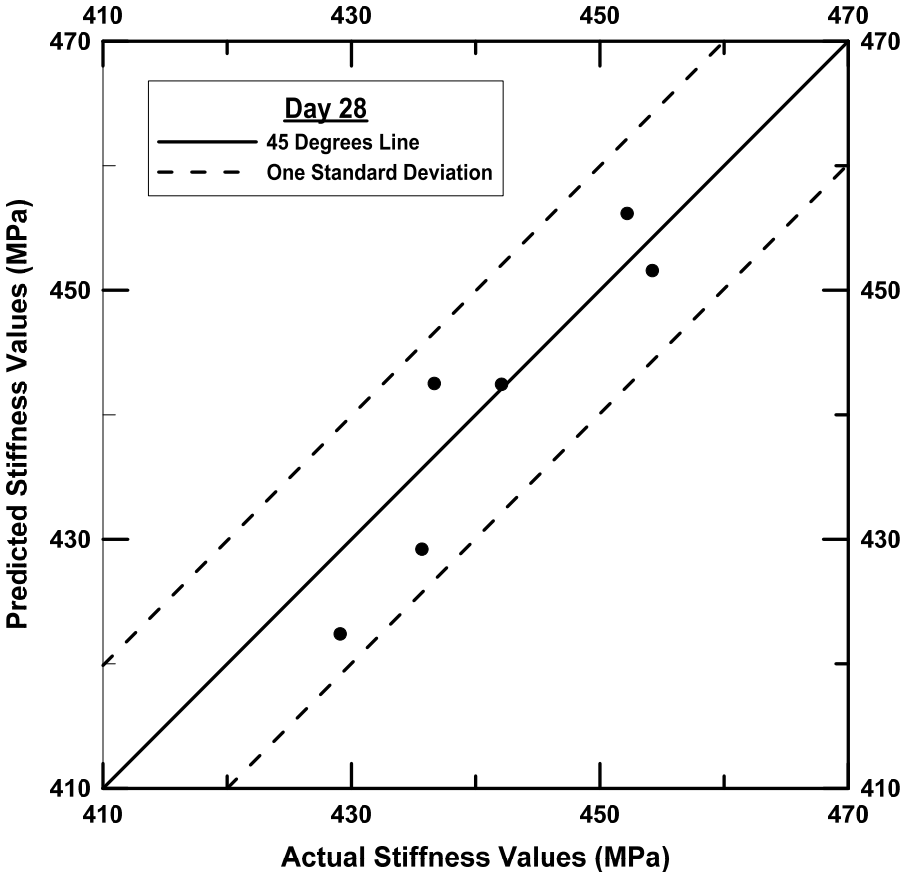


Figure 5-46 Comparison of actual and predicted stiffness values on day 28

The above figures present the cross-validation plots and comparison of actual and predicted values, using a 45 degree line plot. The actual values were plotted on the x-axis, and the predicted values are represented in the y-axis. If the actual and predicted values, are the same they should fall on the 45 degree line. It can be inferred from the plot that all the predicted values obtained through the digitize technique were within 1 standard deviation to the actual deleted values. This affirms that the spatial variability models developed for individual days were more than sufficient to provide predictions, and the spatial maps for the stiffness values for 1, 3, 7, 14 and 28 days were appropriate.

5.7 Summary

Controlled Low Strength Material (CLSM) was used as the bedding material in a water pipeline project to support a large pipeline system. The uniform development of stiffness of CLSM was of primary concern due to the novel mix design approach followed while preparing CLSMs, using native clays. The stiffness measurements of CLSM along the pipeline section were obtained using non-destructive seismic method (SASW). The statistical analysis was performed to understand the distribution present in the stiffness measurements. The geostatistical analysis was performed to predict the stiffness of the CLSM throughout the pipeline section by developing the spatial variability model with the known values. Below are important findings determined from this study:

- The framework developed in Chapter 3 was validated for performing the spatial variability analysis of the CLSM data in a pipeline construction project.
- The spatial correlation distance of the stiffness values ranged from 8 to 24 m. However, after a 28-day curing period, the correlation distance for CLSM stiffness values could be 24 m. The increase in spatial correlation distance from day 1 to day 28 signifies the increase in homogeneity in stiffness values.

- The nugget effect and exponential model best describe the spatial variability present in the stiffness values for all days.
- The spatial variability analysis can be used as a quality assurance tool by mapping the stiffness values over the entire pipeline.
- Through this analysis, the uniform stiffness development of CLSM, or any material, can be ensured.

Chapter 6

Spatial Variability Analysis of Sulfate-Rich Natural Soils

6.1 Introduction

Expansive soils containing sulfates which have been chemically stabilized are known as man-made expansive soils (Puppala et. al., 2012). The calcium-based stabilizers react with natural sulfates present in the soil, leading to the formation of new minerals, ettringite and thaumasite. The heaving mechanics that occur due to hydration of these minerals are referred to as sulfate-induced heave (Sherwood 1962). Several countries across the globe have reported heaving due to the presence of sulfates in soils (Hawkins 1987; Little 1989; Wimsatt 1999; Chen et al. 2005; Mingyu 2006; Rollings et al. 2006; Zhiming 2008; Adams 2008; Bagley et al. 2009; Puppala et al. 2010). In the United States, several pavement failures have been recorded due to sulfate-induced heave (Perrin, 1992; Dermatas, 1995; Puppala et al., 2006).

The damages induced due to sulfate heave results in millions of dollars spent in repair costs annually (Hunter, 1988; Petry, 1994; Kota et al. 1996). In the last decade, extensive importance has been given to the determination of sulfate concentrations in soils. The modified UTA method is the accurate and rapid test method used to determine the soluble sulfates present in soils. The detailed procedure and steps involved in the modified UTA method are presented in Puppala et al. (2002). Several studies have been conducted to categorize the sulfate concentration levels that cause problems (Petry et al. 1992; Berger et al. 2002). Puppala et al. (2003) provided the level of risk involved in lime stabilization with an increase of sulfate content in soils. However, the variability of sulfate concentrations in subsoil profiles provides uncertainty to the level of risk that is associated with chemical stabilization process. In the research study conducted by Puppala et al (2013), very high sulfate concentrations of 44,000ppm were recorded in the

Childress District. Figure 6-1 shows the presence of sulfate soils and Eagleford shale in Texas, USA.

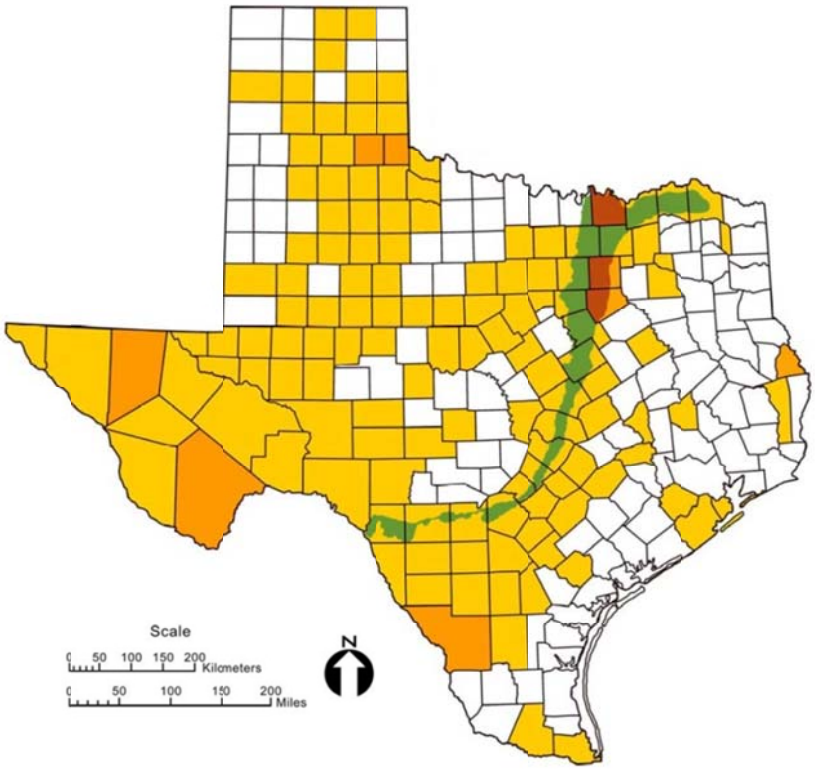


Figure 6-1 Geology of Texas with sulfate presence in soils (Harris, 2005)

The varying sulfate contents often lead to ambiguity in interpreting the sulfate concentration at unsampled locations. This uncertainty leads to the misinterpretation of the stabilizer dosage content required for chemical stabilization of the entire region. As sulfates can be present in pockets, which may affect the sulfate contour maps. In this research study, an attempt was made to predict the sulfate concentration at unsampled locations and assess how the geostatistical modeling of sulfates could provide comprehensive knowledge of sulfate contours in the region.

The high sulfate content present in the Texas soils led to the determination of sulfate contents in a pipeline project. This is the same pipeline considered in Chapter 5[however, in Chapter 5, only 500 ft. was considered for spatial variability in stiffness values. In this analysis, the 150-mile long pipeline was considered for the determination of sulfate content present in soils.

In another research study conducted by Thomey (2013), the quantification of the sulfate contents was performed using the modified UTA method, and mapping of sulfate content was performed using geostatistics. This mapping was performed based on the assumptions that the data is stationary with constant mean and variance, and spatial variability models are linear.

In this research study, the sulfate content determined by Thomey (2013) was used to capture the spatial variability, using the variograms and kriging analysis in accordance with the framework developed in Chapter 3. The main goal of this analysis was to explain how this information can benefit the engineering community in extracting the information from the limited knowledge of soil testing data. This analysis will also bring out any limitations in the study.

6.2 Data Collection

The pipeline considered for this analysis was placed through six geological formations in the North Texas region. In order to ascertain the safety of the pipeline against sulfate issues, the sulfate concentrations at all these formations were determined at different depths. Figure 6-2 displays the entire layout of the pipeline, highlighting the six formations.

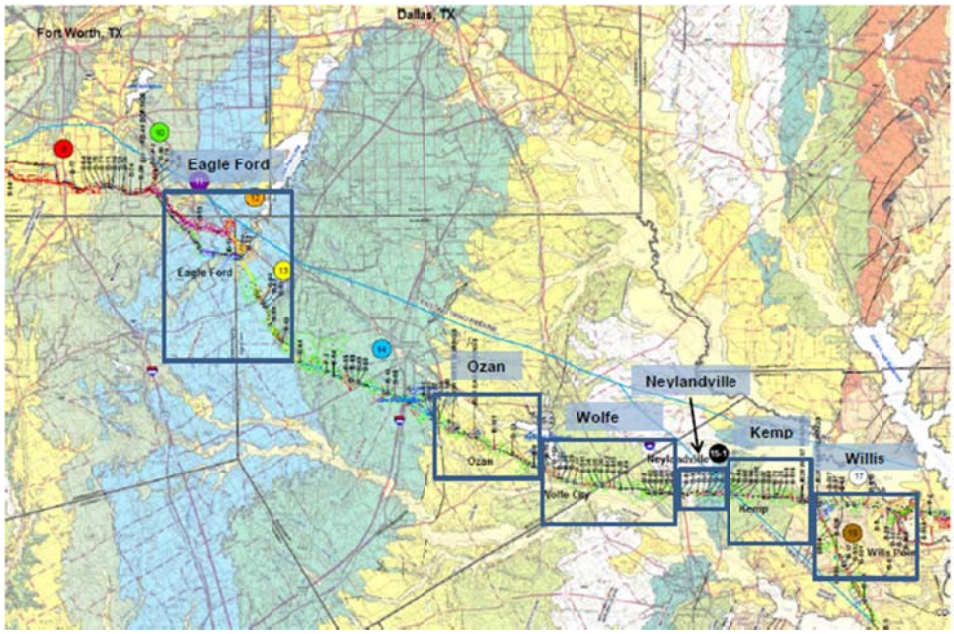


Figure 6-2 Pipeline layout through 6 geological formations

(Source: Fugro Consultants Inc.)

Considering the 9 ft. diameter of the pipeline and placement of it below the ground surface, the soil samples were obtained at depths of 5, 10, 15 and 20 ft. The soils sampled from these depths were brought to the University of Texas at Arlington laboratory for determination of the sulfate content in them. The modified UTA method was used to determine the sulfate contents in the soils. The details of the test procedure and comprehensive test results were presented in Thomey (2013). In this analysis, the measured sulfate content results were utilized to perform the spatial variability analysis.

The sulfate concentration levels at depths of 5, 10, 15 and 20 ft. are presented in Tables 6-1 to 6-4. The number of observations indicates the number of boreholes where the samples were obtained. The sulfate concentration was represented with units of "ppm," which represents parts per million. This unit resembles the measurement of the

mass of the sulfate content per unit volume of water. This can be also expressed as milligrams per liter (mg/L). It should be noted that,

1 parts per million = 1 mg/L

The number of observations in each formation indicates the number of boreholes drilled to collect the soil samples. It can be observed from the tables that the number of observations considered is not equal, and this is attributed to the provision of the pipeline in that particular formation. In total, 301 sulfate measurements were obtained along the pipeline alignment. The detailed information of the boreholes in each formation, along with the sulfate concentration levels at 5, 10, 15 and 20 ft., are provided in Appendix C.

Table 6-1 Sulfate concentration at 5 ft. depth for 6 geological formations

Geological Formation	Sulfate Concentration (ppm)									
	1	2	3	4	5	6	7	8	9	10
Kemp	180	1880	1270	11600	2120	260	115	1065	1400	3215
Neylandville	975	535	90	460						
Wolfe	1200	50	310	20	125	670	1065	25	20	65
Eagleford	540	185	130	8080	750	18,450	19,620	15,260	9360	4200
Wills	260	105	55	650	90	45	350	315	750	240
Ozan	1130	16000	280	290	17400	16760	950	115	17110	16700

Geological Formation	Sulfate Concentration (ppm)								
	11	12	13	14	15	16	17	18	19
Kemp									
Neylandville									
Wolfe	80	390							
Eagleford	7000	5530	14,050	75	2200	3340	16,000	1320	1005
Wills	300	280	135	200	415				
Ozan	50	1100	570	17,300	1,015				

Table 6-2 Sulfate concentration at 10 ft. depth for 6 geological formations

Geological Formation	Sulfate Concentration (ppm)									
	1	2	3	4	5	6	7	8	9	10
Kemp	280	17000	255	16100	8420	255	300	690	840	1070
Neylandville	1500	690	135	180						
Wolfe	750	785	800	1050	200	100	10600	115	120	45
Eagleford	500	1320	525	185	200	2250	18,300	1620	3620	560
Wills	150	55	390	670	325	20	300	380	230	100
Ozan	2820	1055	275	200	1235	1435	755	110	2345	15,620

Geological Formation	Sulfate Concentration (ppm)										
	11	12	13	14	15	16	17	18	19	20	21
Kemp											
Neylandville											
Wolfe	115	310									
Eagleford	890	1050	15,100	15,200	6300	370	180	105	11,725	280	1090
Wills	150	40	200	10	90						
Ozan	1270	1050	415	1750	25						

Table 6-3 Sulfate concentration at 15 ft. depth for 6 geological formations

Geological Formation	Sulfate Concentration (ppm)									
	1	2	3	4	5	6	7	8	9	10
Kemp	325	1100	215	1360	10000	90	220	700	850	970
Neylandville	605	790	110	210						
Wolfe	90	700	700	1000	480	1255	135	45	85	105
Eagleford	1270	300	140	1915	715	1160	16,160	640	15,670	3800
Wills	200	30	185	575	365	30	650	220	20	270
Ozan	1500	1100	275	205	1255	1315	690	100	1510	15,940

Geological Formation	Sulfate Concentration (ppm)								
	11	12	13	14	15	16	17	18	19
Kemp									
Neylandville									
Wolfe	470								
Eagleford	17,500	1560	530	60	1030	275	17,100	290	1095
Wills	25	60	70	115	70	35			
Ozan	1370	975	620	11,310	20				

Table 6-4 Sulfate concentration at 20 ft. depth for 6 geological formations

Geologic Formation	Sulfate Concentration									
	1	2	3	4	5	6	7	8	9	10
Kemp	100	2400	470	2430	18080	250	200	620	1170	15120
Neylandville	710	450	315	840						
Wolfe	60	620	625	775	335	1460	330	130	75	350
Eagleford	540	185	130	8080	750	18,450	19,620	15,260	9360	4200
Wills	510	20	200	675	510	20	400	350	5	200
Ozan	1130	16,000	280	290	17,400	16,760	950	115	17,110	16,700

Geological Formation	Sulfate Concentration (ppm)								
	11	12	13	14	15	16	17	18	19
Kemp									
Neylandville									
Wolfe	14100								
Eagleford	7000	5530	14,050	75	2200	3340	16,000	1320	1005
Wills	15	40	250	230	20				
Ozan	50	1100	570	17,300	1,015				

6.3 Data Organization

The sulfate content results summarized in the above tables had to be organized in the Cartersian coordinate system before performing spatial variability analysis. The northing and easting of the boreholes were used as x and y coordinates. The sulfate content value at a corresponding depth in a borehole was represented by z coordinate. Table 6-5 provides the coordinates of Kemp formation at 15 ft. depth.

Table 6-5 Coordinates for Kemp formation at 15 feet depth

Boring	X-coordinate (Easting)	Y-coordinate (Northing)	z-coordinate (Sulfate Concentration)
B-007	-9641723	3221603	325
B-025	-9634307	3221148	1100
B-068	-9649225	3221794	215
B-069	-9635347	3221152	1360
B-070	-9634595	3221168	10000
B-094	-9641409	3221516	90
B-095	-9640995	3221472	220
B-096	-9639983	3221385	700
B-097	-9636216	3221141	850
B-180	-9639943	3221362	970

Similarly, the coordinates of all the sulfate concentration levels were determined using the northing and easting of the boreholes data. The complete coordinate details of

all the boreholes considered in this study at 5, 10, 15 and 20 ft. depths are provided in Appendix C.

6.4 Statistical Analysis

Statistical analysis was performed on the sulfate concentrations (ppm) obtained at all depths. This analysis will help in evaluating the basic statistical parameters and supplement the geostatistical analysis attempted here. In this section, the statistical analysis, comprised of histograms, evaluated the data for Gaussian distribution, ANOVA test, Bartlett's test, and Box-Cox transformations.

6.4.1 Histograms

Histograms were plotted for sulfate concentration observations at depths of 5, 10, 15, and 20 ft., respectively. The number of class intervals and bin size required to construct a histogram were determined using Equations 3-1 and 3-2. Table 5-7 below summarizes the number of class intervals and bin sizes determined for constructing histograms for sulfate observations at these depths.

Table 6-6 Summary of parameters used to construct histograms

Depth	Observations	Number of Class Intervals	Bin Size
5	75	7.18	2722
10	77	7.22	2540
15	75	7.18	2431
20	74	7.16	2431

Histogram plots were constructed using the above parameters. Figures 6-3 to 6-6 provided the histogram plots for the sulfate concentration values determined at various depths. The sulfate concentration values were plotted on x-axis and y-axis, and these represented the frequency of the observations. The distribution of sulfate concentration in all the plots was scattered with a maximum number of observations below 3,000 ppm at all depths. However, in all plots there was a long tail of sulfate observations, indicating positive skewness in the data. The standard deviation of sulfate values of various regions ranged from 4285 to 6369, representing a large variability in the data values. The high standard deviation values can be attributed to the large sulfate concentration values obtained at all the depths. From the visual inspection of plots and standard deviation, the sulfate values obtained at a depth of 15m provided least variability when compared to sulfate values at depths of 5, 10 and 20 m, respectively. All the histogram plots depict the exponential distribution of the data.

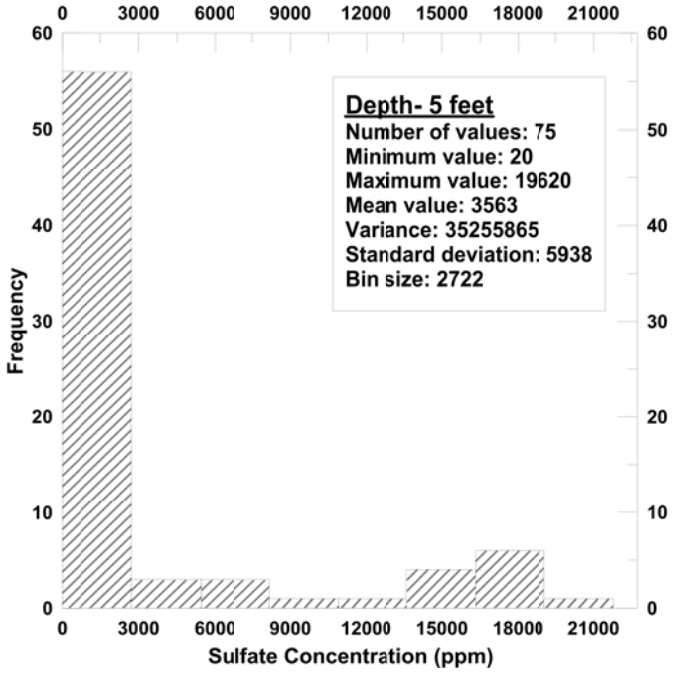


Figure 6-3 Histogram plot sulfate concentration values at 5 ft. depth

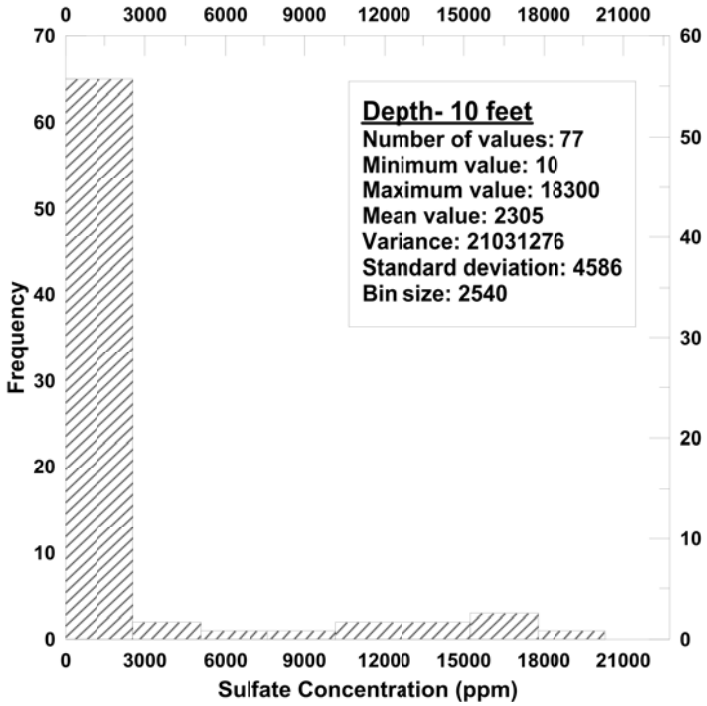


Figure 6-4 Histogram plot sulfate concentration values at 10 ft. depth

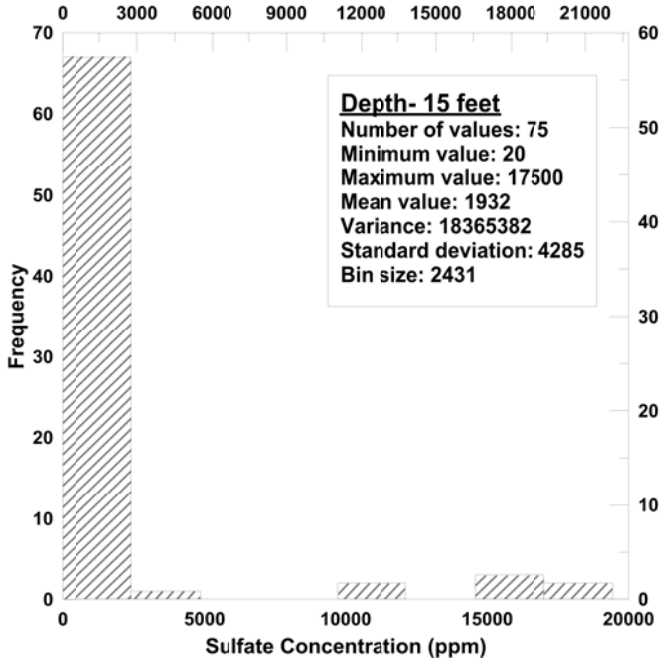


Figure 6-5 Histogram plot sulfate concentration values at 15 ft. depth

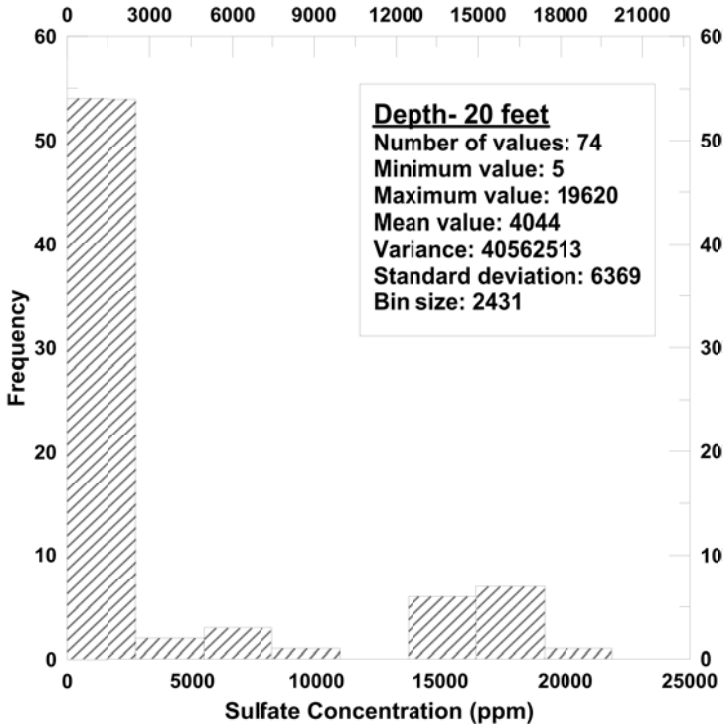


Figure 6-6 Histogram plot sulfate concentration values at 20 ft. depth

6.4.2 Check for Gaussian Distribution

From the earlier histogram plots, it was clear that the data values were not Gaussian-distributed. In this section the data values are evaluated for Gaussian distribution, using normal-quantile plots based on the framework discussed in Chapter 3. As stated in Chapter 5, the Shapiro-Wilk test was not used due to the limited number of observations for conducting this test.

6.4.2.1 Normal-Quantile plot

Normal-Quantile plots were constructed for the sulfate observations obtained at all depths, as shown in Figures 6-7 to 6-10. At a particular depth, the theoretical quantiles were calculated from corresponding sulfate concentration values, using Equation (3-3). The theoretical quantile values were plotted on the y-axis, while the corresponding sulfate concentration quantiles were plotted on the x-axis. Using the regression approach method, a best fit trend line was modelled for the plotted data values. The coefficient of determination was calculated using Equation 3-4 to measure the good fit of the trend line. Table 6-7 below summarizes the coefficient of determination values obtained for all the plots.

Table 6-7 Coefficient of determination for normal-quantile plots

Depth (feet)	Coefficient of Determination (r^2)
5	0.62
10	0.51
15	0.44
20	0.64

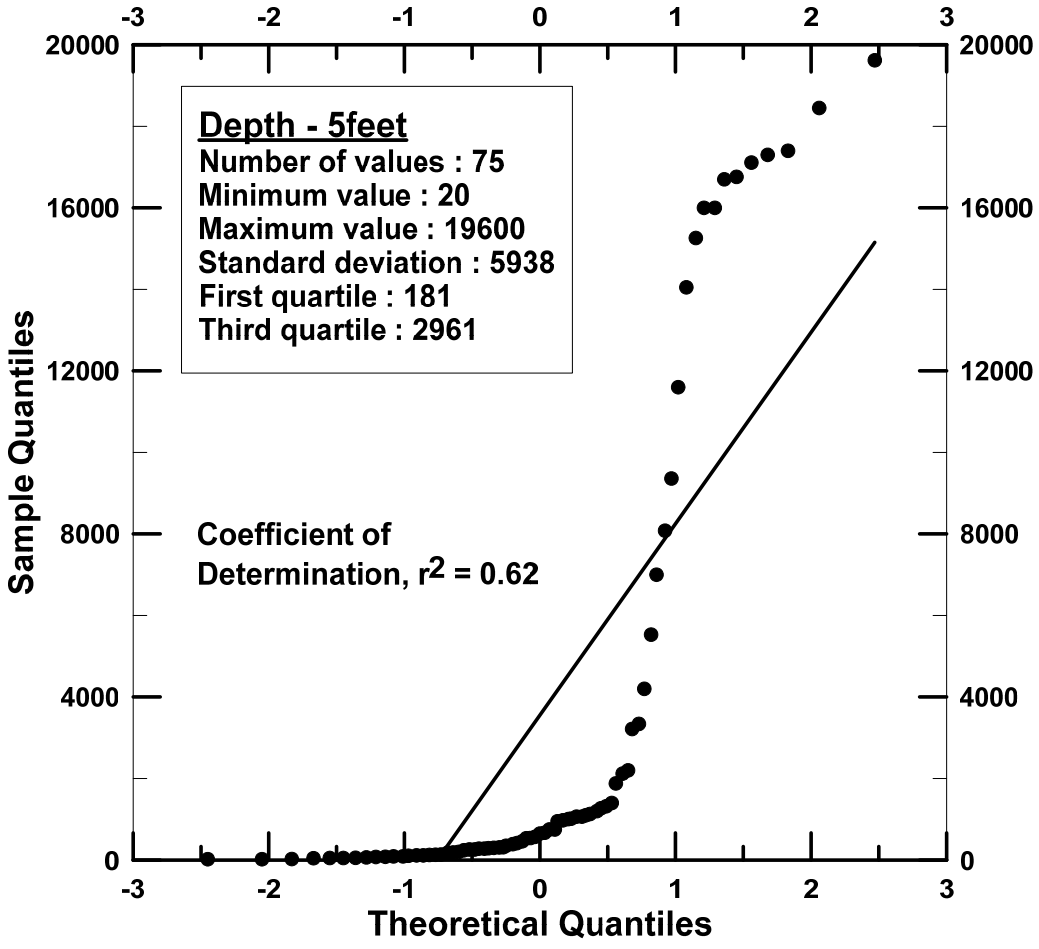


Figure 6-7 Normal-Quantile plot for sulfate concentration values at 5 ft. depth

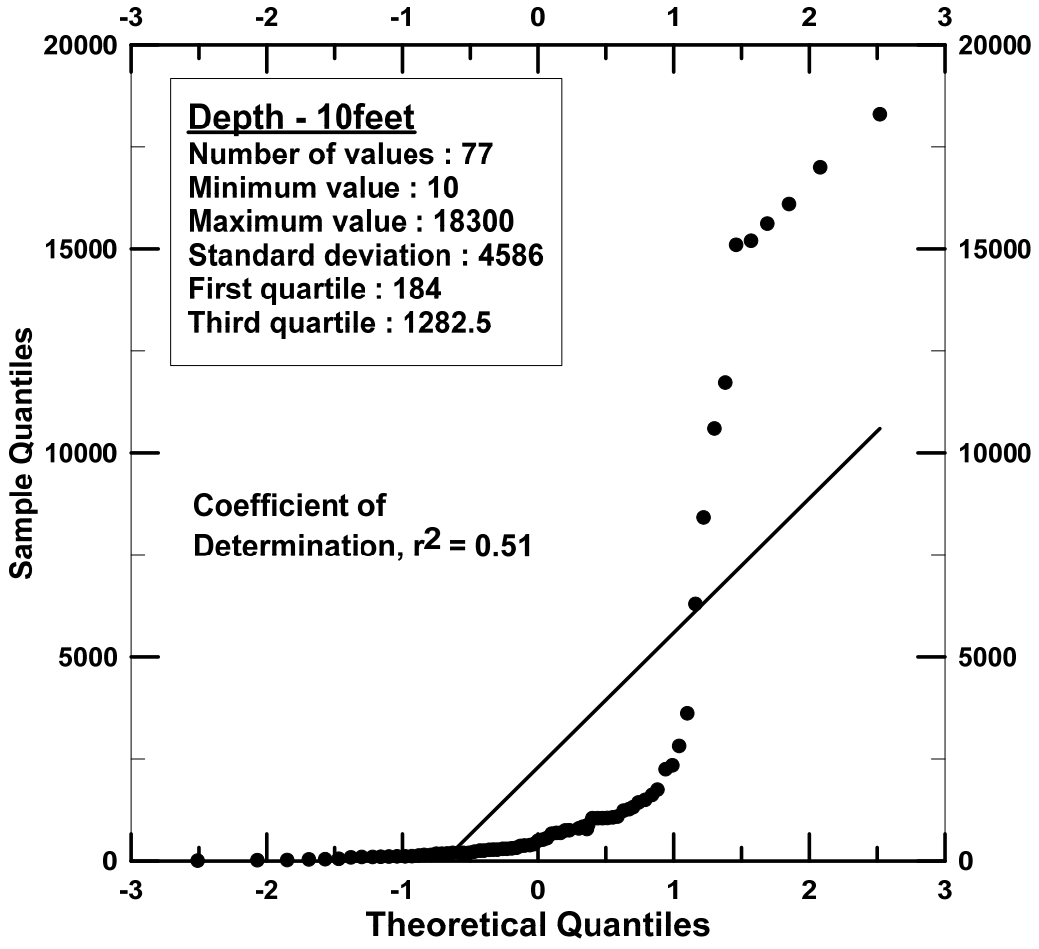


Figure 6-8 Normal-Quantile plot for sulfate concentration values at 10 ft. depth

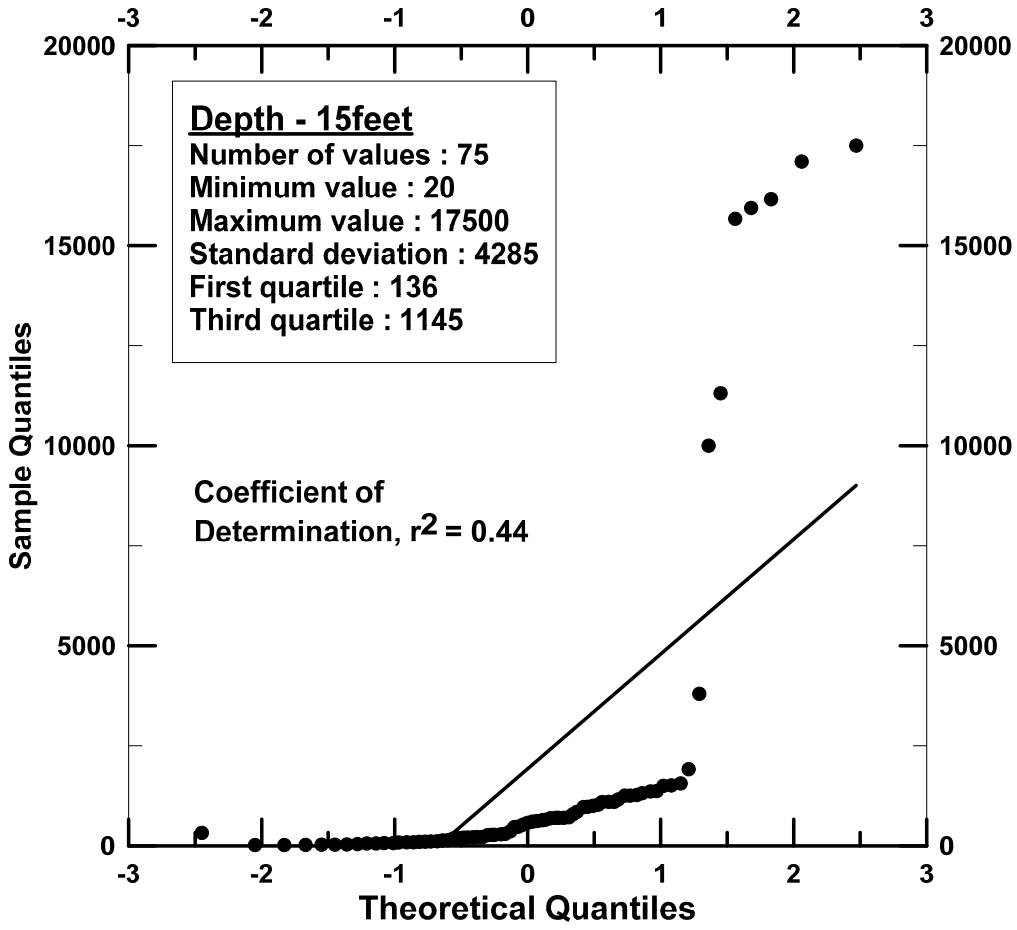


Figure 6-9 Normal-Quantile plot for sulfate concentration values at 15 ft. depth

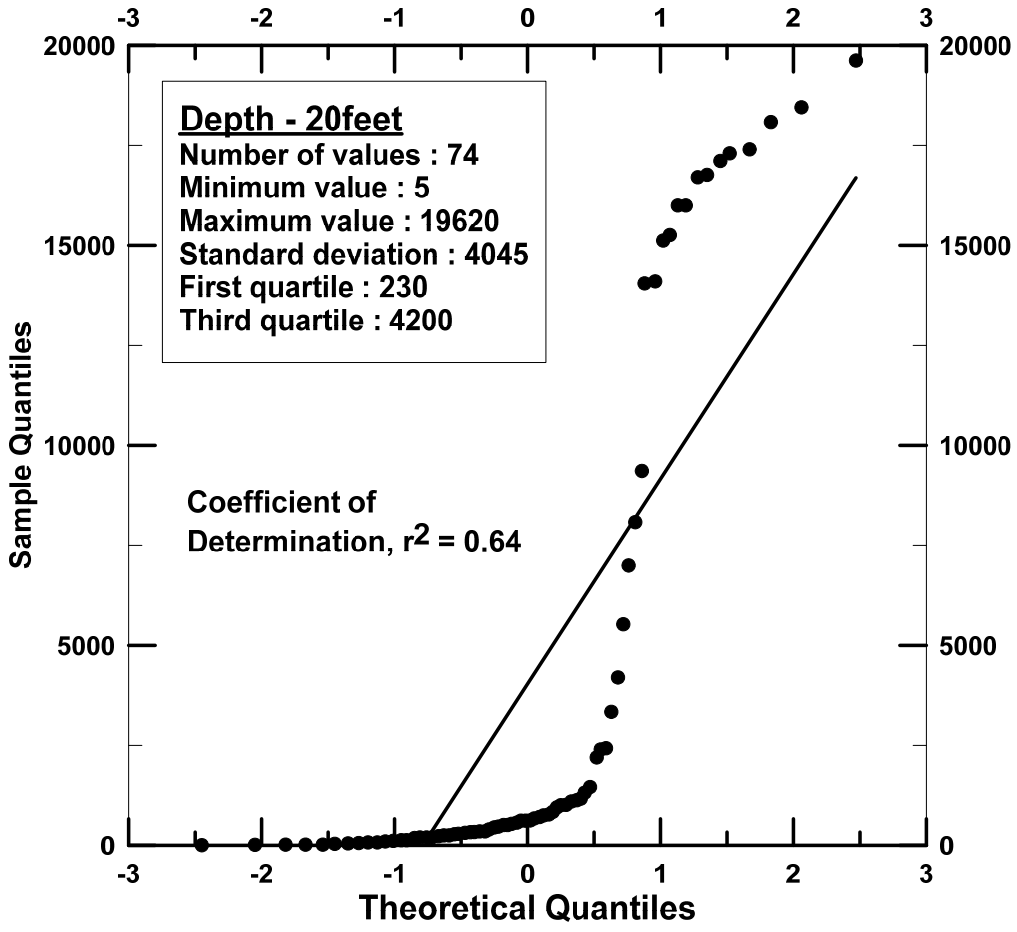


Figure 6-10 Normal-Quantile plot for sulfate concentration values at 20 ft. depth

From all the normal-quantile plots, it is clearly evident that the data points do not follow the linear trends. Also, the coefficient of determination was very low, signifying that the data was not Gaussian-distributed. Most of the statistical tests, along with geostatistics, performs well if the data is Gaussian-distributed. However, it might not always be practically possible. In such cases, one way to approach it is by expressing the data in different ways by using transformations, which are described in the following section.

6.5 Transformations

Transformation refers to expressing the same data in different terms. The key to transformations is selecting the right transformation function. Transformations in general can be used in various situations. However, in this study the transformations were performed to check the Gaussian distribution in data, using new function and converting the non-stationary data to stationary. It should be noted that by using the transformed data, it is not guaranteed that the data will turn into Gaussian; however, if the data comes close to a Gaussian distribution, it shall be employed.

In this study, the Box-Cox method was used to select the appropriate transformation function. Using the function, an attempt was made, on all the sulfate concentration data, to look for nearly-Gaussian behavior trends. The methodology of the Box-Cox method is presented in Chapter 3. In this section, the results of the Box-Cox method for the sulfate concentration values at various depths are discussed. For different values of 'λ', a new set of data was generated, using Equation 3-14. The error sum of square values was determined from the corresponding data generated. Tables 6-8 to 6-11 represent the error sum of squares determined for different values of 'λ' for sulfate concentration values at all depths. Corresponding graphs for 'λ' versus error sum of squares (SS_{Error}) are presented in Figures 6-11 to 6-14.

From the plots, it is evident that the least error sum of squares for sulfate concentrations at all depths was observed when λ-value was equal to 0. For a λ-value of 0, the corresponding transformation function to be applied is the natural logarithm (Ln). The transformed data using the natural logarithmic (Ln) function at all depths is presented in the Tables 6-12 to 6-15, which are considered for further analysis.

Table 6-8 Box-Cox results for sulfate concentration values at 5 ft. depth

Λ	$SS_E(\lambda)$
-1	1.5E+09
-0.75	4.8E+08
-0.5	1.9E+08
-0.25	1.06E+08
0	9.0E+07
0.25	1.3E+08
0.5	2.5E+08
0.75	6.5E+08
1	1.9E+09

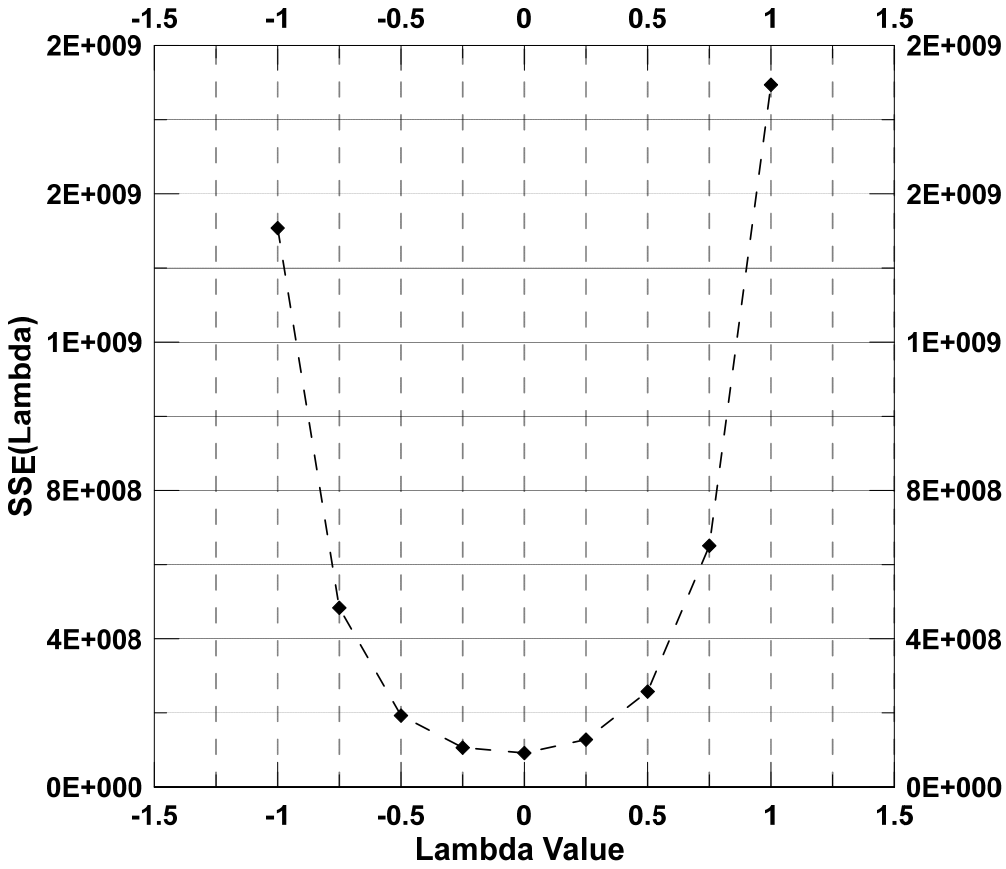


Figure 6-11 Box-Cox plot for selecting the most likelihood λ-value at 5 ft. depth

Table 6-9 Box-Cox results for sulfate concentration values at 10 ft. depth

Λ	$SS_E(\lambda)$
-1	1118431700
-0.75	298495738
-0.5	103531842.1
-0.25	54329432.25
0	47457890.4
0.25	74393532.09
0.5	162172368.3
0.75	443979352.9
1	1405685528

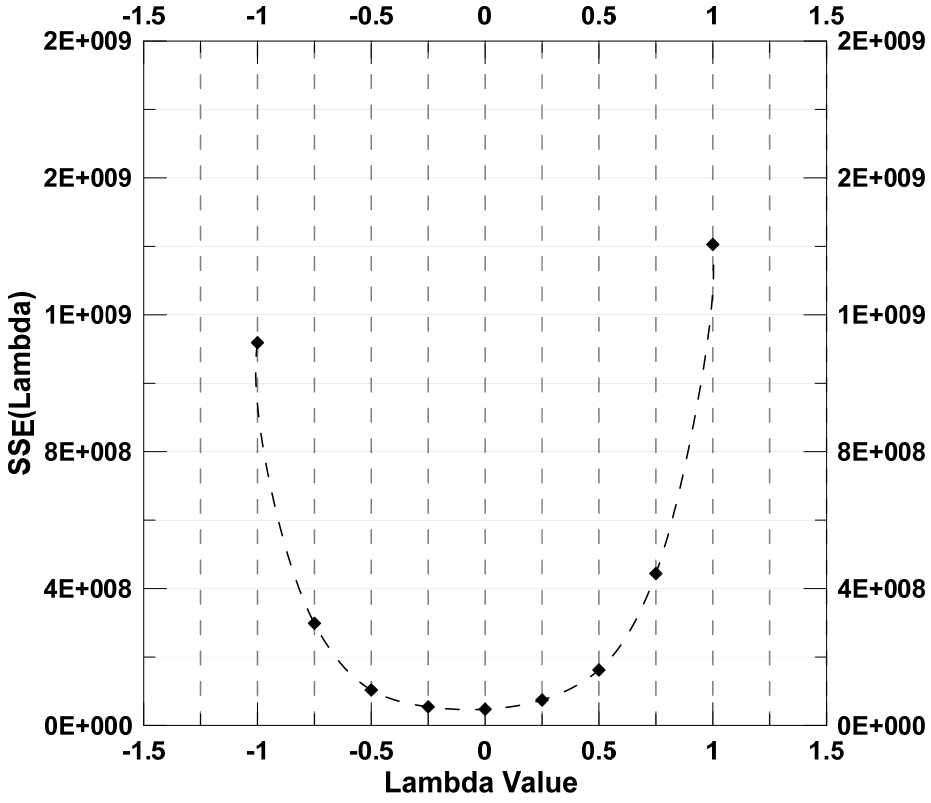


Figure 6-12 Box-Cox plot for selecting the most likelihood λ-value at 10 ft. depth

Table 6-10 Box-Cox results for sulfate concentration values at 15 ft. depth

Λ	SSE(λ)
-1	301771635.9
-0.75	116510143.4
-0.5	53600335.12
-0.25	32336206.18
0	35697453.74
0.25	45410757.3
0.5	107161564.7
0.75	328212310.4
1	1166107458

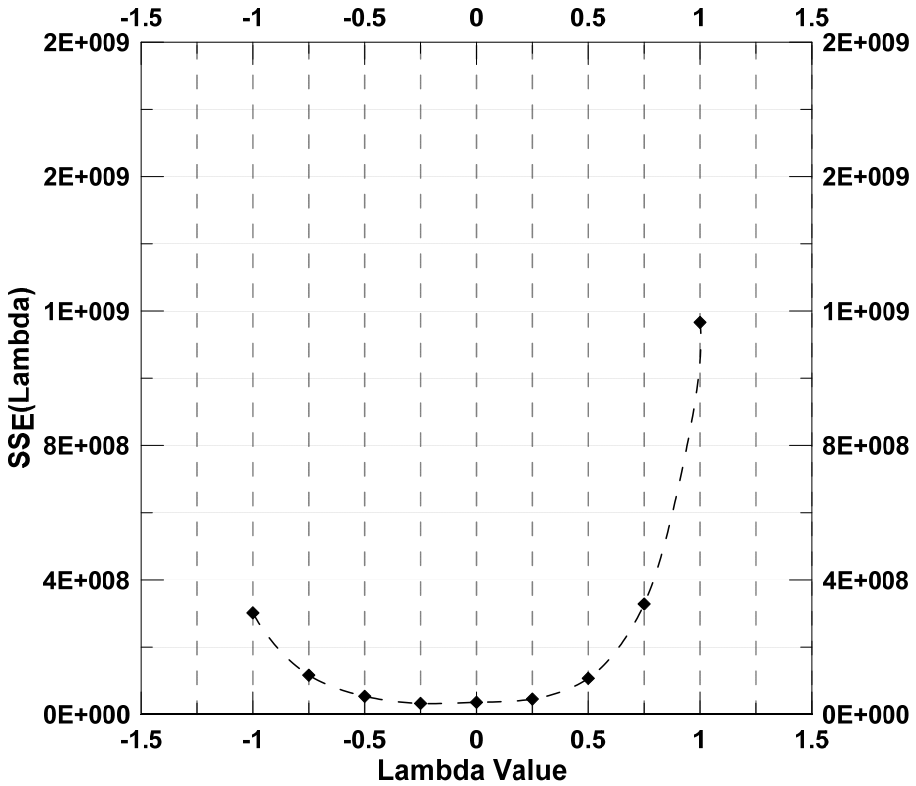


Figure 6-13 Box-Cox plot for selecting the most likelihood λ-value at 15 ft. depth

Table 6-11 Box-Cox results for sulfate concentration values at 20 ft. depth

Λ	SSE(λ)
-1	15377496368
-0.75	2397855141
-0.5	506858122.2
-0.25	178489623.3
0	127659026.1
0.25	172162368.2
0.5	341375780.8
0.75	839852497.5
1	2360400858

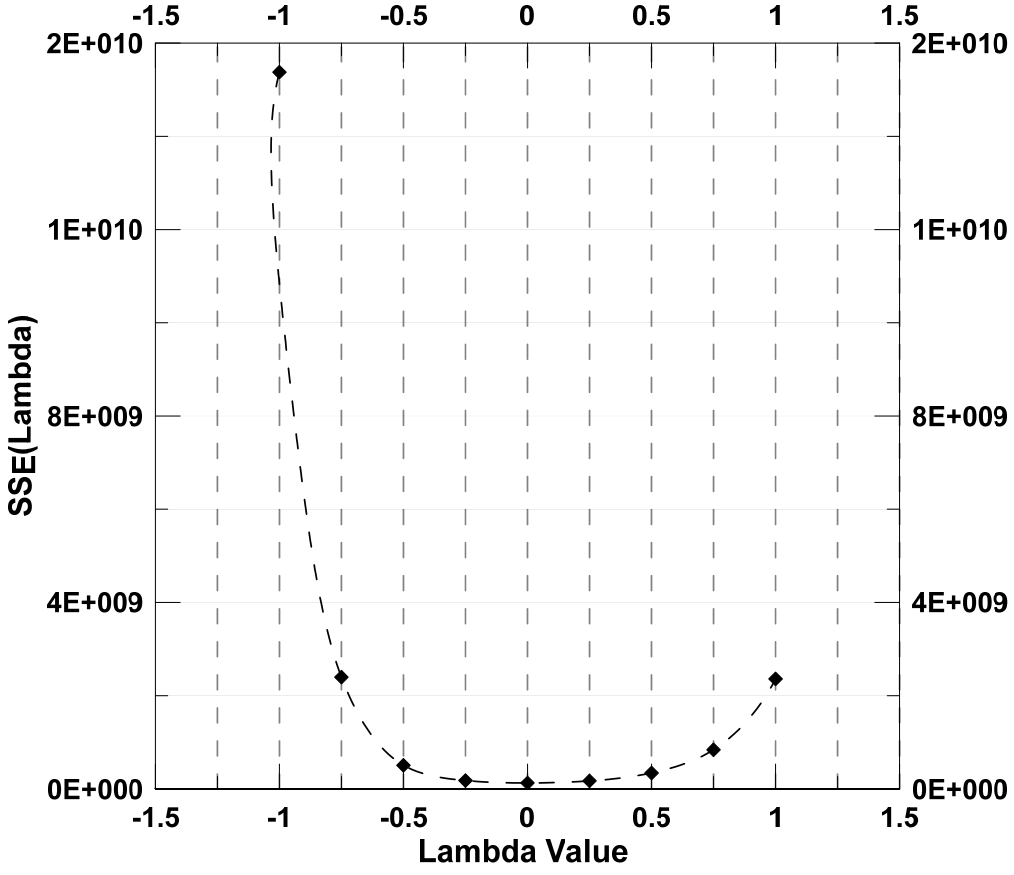


Figure 6-14 Box-Cox plot for selecting the most likelihood λ-value at 20 ft. depth

Table 6-12 LN-Transformed Sulfate concentration values at 5 ft. depth

Geological Formation	Transformed Sulfate Observations																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Kemp	5.2	7.5	7.1	9.4	7.7	5.6	4.7	7.0	7.2	8.1									
Neylandville	6.9	6.3	4.5	6.1															
Wolfe	7.1	3.9	5.7	3.0	4.8	6.5	7.0	3.2	3.0	4.2	4.4	6.0							
Eagleford	6.3	5.2	4.9	9.0	6.6	9.8	9.9	9.6	9.1	8.3	8.9	8.6	9.6	4.3	7.7	8.1	9.7	7.2	6.9
Wills	5.6	4.7	4.0	6.5	4.5	3.8	5.9	5.8	6.6	5.5	5.7	5.6	4.9	5.3	6.0				
Ozan	7.0	9.7	5.6	5.7	9.8	9.7	6.9	4.7	9.7	9.7	3.9	7.0	6.3	9.8	6.9				

Table 6-13 LN-Transformed Sulfate concentration values at 10 ft. depth

Geological Formation	Transformed Sulfate Concentration																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Kemp	5.6	9.7	5.5	9.7	9.0	5.5	5.7	6.5	6.7	7											
Neylandville	7.3	6.5	4.9	5.2																	
Wolfe	6.6	6.7	6.7	7.0	5.3	4.6	9.3	4.7	4.8	4	4.7	5.7									
Eagleford	6.2	7.2	6.3	5.2	5.3	7.7	9.8	7.4	8.2	6	6.8	7.0	9.6	9.6	8.7	6	5	5	7	6	7
Wills	5.0	4.0	6.0	6.5	5.8	3.0	5.7	5.9	5.4	4	5.0	3.7	5.3	2.3	4.5						
Ozan	7.9	7.0	5.6	5.3	7.1	7.3	6.6	4.7	7.8	9	7.1	7.0	6.0	7.5	3.2						

Table 6-14 LN-Transformed Sulfate concentration values at 15 ft. depth

Geological Formation	Transformed Sulfate Concentration values																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Kemp	5.8	7.0	5.4	7.2	9.2	4.5	5.4	6.6	6.7	6.9									
Neylandville	6.4	6.7	4.7	5.3															
Wolfe	4.5	6.6	6.6	6.9	6.2	7.1	4.9	3.8	4.4	4.7	6.2								
Eagleford	7.1	5.7	4.9	7.6	6.6	7.1	9.7	6.5	9.7	8.2	9.8	7.4	6.3	4.1	6.9	5.6	9.7	5.7	7.0
Wills	5.3	3.4	5.2	6.4	5.9	3.4	6.5	5.4	3.0	5.6	3.2	4.1	4.2	4.7	4.2	3.6			
Ozan	7.3	7.0	5.6	5.3	7.1	7.2	6.5	4.6	7.3	9.7	7.2	6.9	6.4	9.3	3.0				

Table 6-15 LN-Transformed Sulfate concentration values at 20 ft. depth

Geological Formation	Transformed Sulfate Concentration Values																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Kemp	4.6	7.8	6.2	7.8	9.8	5.5	5.3	6.4	7.1	9.6									
Neylandville	6.6	6.1	5.8	6.7															
Wolfe	4.1	6.4	6.4	6.7	5.8	7.3	5.8	4.9	4.3	5.9	9.6								
Eagleford	6.3	5.2	4.9	9.0	6.6	9.8	9.9	9.6	9.1	8.3	8.9	8.6	9.6	4.3	7.7	8.1	9.7	7.2	6.9
Wills	6.2	3.0	5.3	6.5	6.2	3.0	6.0	5.9	1.6	5.3	2.7	3.7	5.5	5.4	3.0				
Ozan	7.0	9.7	5.6	5.7	9.8	9.7	6.9	4.7	9.7	9.7	3.9	7.0	6.3	9.8	6.9				

The above transformed values represent the sulfate concentration values, which are expressed using natural logarithmic function. Figures 6-15 to 6-18 represent the normal-quantile plots for the transformed data at all depths, comprised of 6 geological formations.

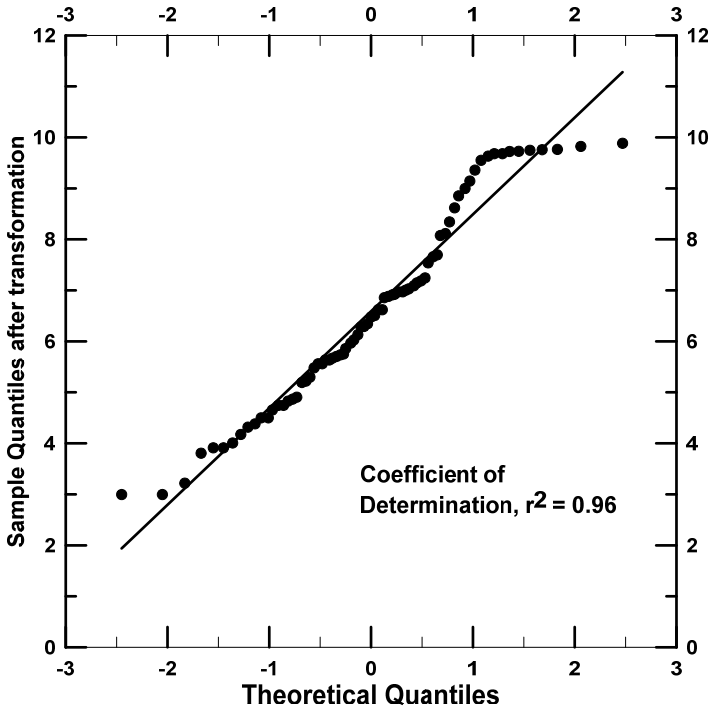


Figure 6-15 Normal-Quantile plot for transformed data at 5 ft. depth

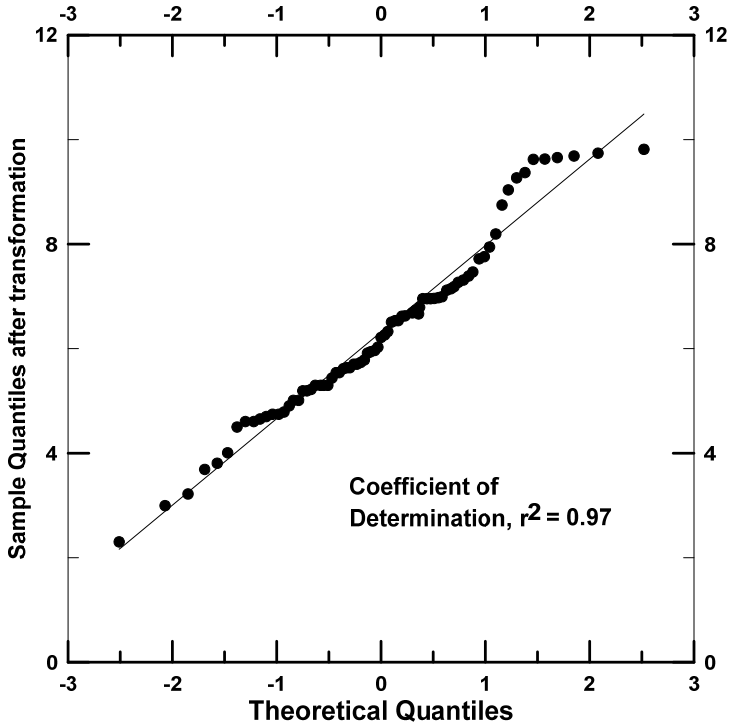


Figure 6-16 Normal-Quantile plot for transformed data at 10 ft. depth

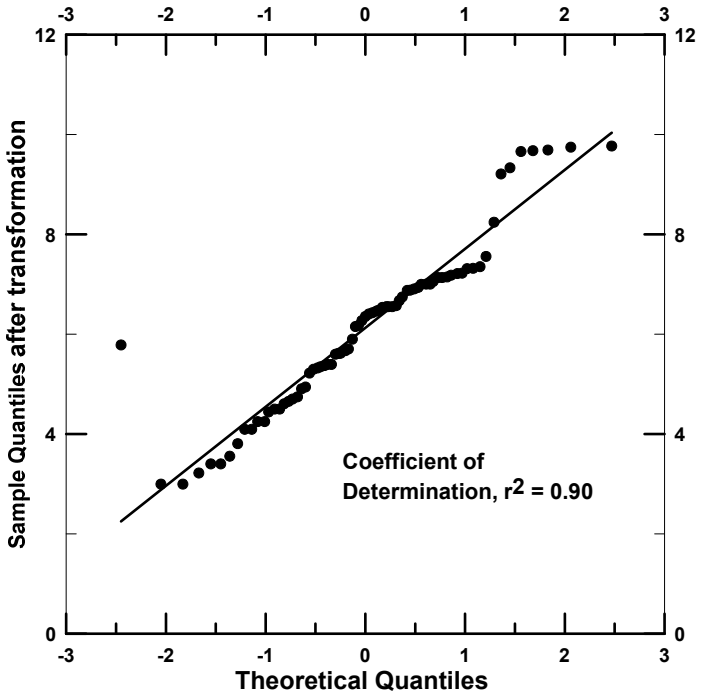


Figure 6-17 Normal-Quantile plot for transformed data at 15 ft. depth

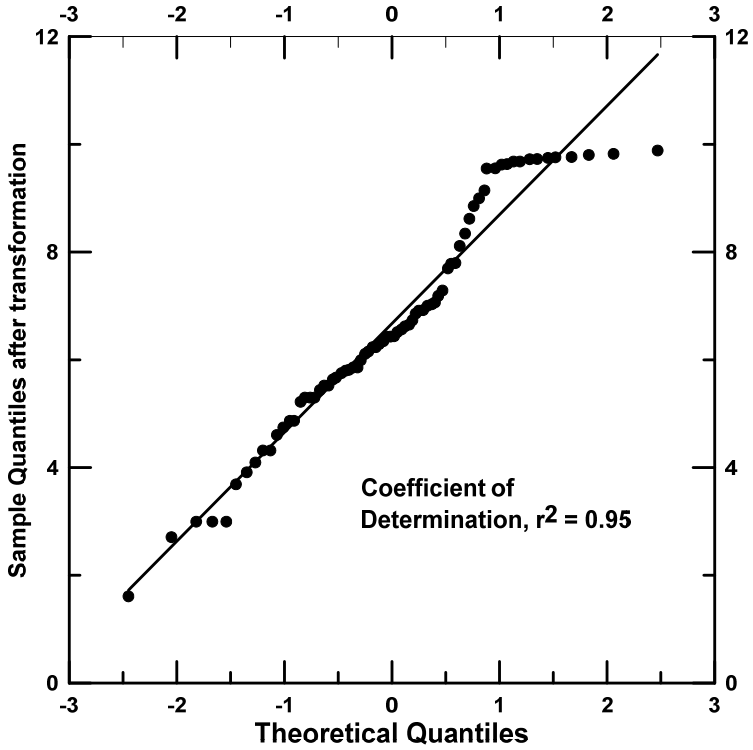


Figure 6-18 Normal-Quantile plot for transformed data at 20 ft. depth

From the above plots, it can be observed that there is significant improvement in the normal-quantile plots with transformed data. All the data points fall close to the straight line. The coefficient of determination of the plots before and after the transformation is given in below Table 6-16. From the results, it can be concluded that the transformed data values depict Gaussian distribution behavior.

Table 6-16 Coefficient of determination for normal-quantile plots before and after transformation

Depth (feet)	Before Transformation-(r^2)	After Transformation-(r^2)
5	0.62	0.96
10	0.51	0.97
15	0.44	0.90
20	0.64	0.95

6.6 Check for Stationarity

Stationarity in the data refers to constant mean and variance, with equal joint probability distribution. In this case, stationarity refers to constant mean and variance in sulfate concentration values determined at six geological formations, at different depths. This was evaluated using the ANOVA test, Bartlett's test, and experimental variogram. In this section, the stationarity evaluated for the transformed data through ANOVA and Bartlett's test is discussed.

6.6.1 Check for Constant Mean

The constant mean of the sulfate concentration values obtained at different depths were evaluated using the ANOVA test. The transformed sulfate concentration values presented in Tables 6-12 to 6-15 were used for the analysis. In order to check for constant mean, a hypothesis was constructed for the data, as shown below:

$H_0 : \mu_{kemp,depth} = \mu_{wills,depth} = \mu_{neylandville,depth} = \mu_{ozan,depth} = \mu_{eagleford,depth} = \mu_{wolfe,depth}$

H_1 : At least one mean is different

The statistic that was used to evaluate the above hypothesis is f_0, which is the ratio of mean square treatment to mean square error. This was compared to the critical value at a significance level of 0.05. Tables 6-17 to 6-20 below provide the summary of the ANOVA test results conducted for checking the constant mean in sulfate concentration values at all depths.

Table 6-17 Summary of ANOVA results for sulfate concentrations at 5 ft. depth

Analysis of Variance				
Source of Variation	Sum of Squares	DOF	Mean Square	f_0
Treatment	104.4	5	20.8	8.4
Error	171.7	69	2.5	
Total	276.1	74		

*DOF = Degrees of freedom

Table 6-18 Summary of ANOVA results for sulfate concentrations at 10 ft. depth

Analysis of Variance				
Source of Variation	Sum of Squares	DOF	Mean Square	f_0
Treatment	53.0	5	10.6	4.89
Error	153.6	71	2.16	
Total	206.6	76		

*DOF = Degrees of freedom

Table 6-19 Summary of ANOVA results for sulfate concentrations at 15 ft. depth

Analysis of Variance				
Source of Variation	Sum of Squares	DOF	Mean Square	f_0
Treatment	64.3	5	12.8	6.33
Error	140.1	69	2.0	
Total	204.4	74		

*DOF = Degrees of freedom

Table 6-20 Summary of ANOVA results for sulfate concentrations at 20 ft. depth

Analysis of Variance				
Source of Variation	Sum of Squares	DOF	Mean Square	f_0
Treatment	106.2	5	21.2	7.11
Error	203.0	68	2.98	
Total	309.3	73		

*DOF = Degrees of freedom

The statistic f_0 that was calculated for the sulfate values at various depths was compared with associated critical values ($f_{crit)}$. The f_{crit} value was obtained using the f-distribution table at a significance level (α) of 0.05 for different degrees of freedom. Due to the close number of observations, the f_{crit} value for all the cases was about 2.34. From the results summarized in the above tables, it can be inferred that the transformed sulfate concentration values obtained at all depths were greater than the critical values (i.e., 8.4 > 2.34; 4.89 > 2.34; 6.33 > 2.34; 7.11 > 2.34), resulting in unequal mean values. This could be due to the high sulfate concentration values present in the Eagleford geological formation.

6.6.1.1 Model Adequacy Check

The above conclusion obtained from ANOVA test results is viable only if the basic assumptions of the ANOVA model were satisfied. The assumptions were that the error residuals were normally distributed and structureless. The residuals for the transformed sulfate concentration values were calculated using the Equation 3.8. The true values in the equation are the transformed sulfate concentration values, and the fitted values are the mean values obtained for a particular geological formation at a specific depth. Tables 6-21 to 6-24 present the residual values calculated for all the sulfate concentration values at different depths.

Table 6-21 Residual values for sulfate concentration values at 5 ft. depth

Geologic Formation	Residuals									
	1	2	3	4	5	6	7	8	9	10
Kemp	-1.8	0.6	0.2	2.4	0.7	-1.4	-2.2	0.0	0.3	1.1
Neylandville	0.9	0.3	-1.4	0.2						
Wolfe	2.2	-1.0	0.8	-1.9	-0.1	1.6	2.1	-1.7	-1.9	-0.7
Eagleford	-1.6	-2.7	-3.0	1.1	-1.3	1.9	2.0	1.8	1.3	0.5
Wills	0.2	-0.7	-1.3	1.1	-0.9	-1.5	0.5	0.4	1.3	0.1
Ozan	-0.5	2.2	-1.9	-1.8	2.3	2.2	-0.6	-2.8	2.2	2.2

Geologic Formation	Residuals								
	11	12	13	14	15	16	17	18	19
Kemp									
Neylandville									
Wolfe	-0.5	1.1							
Eagleford	1.0	0.7	1.7	-3.6	-0.2	0.2	1.8	-0.7	-1.0
Wills	0.4	0.3	-0.4	-0.1	0.7				
Ozan	-3.6	-0.5	-1.2	2.3	-0.6				

Table 6-22 Residual values for sulfate concentration at depth 10 ft.

Geological Formation	Residuals									
	1	2	3	4	5	6	7	8	9	10
Kemp	-1.5	2.6	-1.6	2.6	1.9	-1.6	-1.4	-0.6	-0.4	-0.1
Neylandville	1.3	0.5	-1.1	-0.8						
Wolfe	0.8	0.8	0.9	1.1	-0.5	-1.2	3.4	-1.1	-1.0	-2.0
Eagleford	-0.8	0.2	-0.7	-1.8	-1.7	0.7	2.8	0.4	1.2	-0.7
Wills	0.2	-0.8	1.1	1.7	0.9	-1.9	0.9	1.1	0.6	-0.2
Ozan	1.3	0.3	-1.0	-1.4	0.5	0.6	0.0	-2.0	1.1	3.0

Geological Formation	Residuals										
	11	12	13	14	15	16	17	18	19	20	21
Kemp											
Neylandville											
Wolfe	-1.1	-0.1									
Eagleford	-0.2	0.0	2.6	2.6	1.8	-1.1	-1.8	-2.3	0.0	-1.4	0.0
Wills	0.2	-1.2	0.4	-2.5	-0.4						
Ozan	0.5	0.3	-0.6	0.8	-3.4						

Table 6-23 Residual values for sulfate concentration at depth 15 ft.

Geological Formation	Residuals									
	1	2	3	4	5	6	7	8	9	10
Kemp	-0.7	0.5	-1.1	0.8	2.7	-2.0	-1.1	0.1	0.3	0.4
Neylandville	0.6	0.9	-1.1	-0.4						
Wolfe	-1.1	0.9	0.9	1.3	0.6	1.5	-0.7	-1.8	-1.2	-1.0
Eagleford	0.0	-1.4	-2.2	0.4	-0.6	-0.1	2.6	-0.7	2.5	1.1
Wills	0.7	-1.2	0.6	1.7	1.3	-1.2	1.8	0.8	-1.6	1.0
Ozan	0.6	0.3	-1.1	-1.4	0.4	0.5	-0.2	-2.1	0.6	3.0

Geological Formation	Residuals								
	11	12	13	14	15	16	17	18	19
Kemp									
Neylandville									
Wolfe	0.5								
Eagleford	2.6	0.2	-0.9	-3.0	-0.2	-1.5	2.6	-1.5	-0.1
Wills	-1.4	-0.5	-0.4	0.1	-0.4	-1.1			
Ozan	0.5	0.2	-0.3	2.6	-3.7				

Table 6-24 Residual values for sulfate concentration at depth 20 ft.

Geological Formation	Residuals									
	1	2	3	4	5	6	7	8	9	10
Kemp	-2.4	0.8	-0.9	0.8	2.8	-1.5	-1.7	-0.6	0.1	2.6
Neylandville	0.3	-0.2	-0.5	0.4						
Wolfe	-2.0	0.3	0.3	0.6	-0.3	1.2	-0.3	-1.2	-1.8	-0.2
Eagleford	-1.6	-2.7	-3.0	1.1	-1.3	1.9	2.0	1.8	1.3	0.5
Wills	1.6	-1.6	0.7	1.9	1.6	-1.6	1.4	1.2	-3.0	0.7
Ozan	-0.5	2.2	-1.9	-1.8	2.3	2.2	-0.6	-2.8	2.2	2.2

Geological Formation	Residuals								
	11	12	13	14	15	16	17	18	19
Kemp									
Neylandville									
Wolfe	3.5								
Eagleford	1.0	0.7	1.7	-3.6	-0.2	0.2	1.8	-0.7	-1.0
Wills	-1.9	-0.9	0.9	0.8	-1.6				
Ozan	-3.6	-0.5	-1.2	2.3	-0.6				

The residual values presented above were used to validate the assumptions in the ANOVA model. Figures 6-11 to 6-14 present the model adequacy plots.

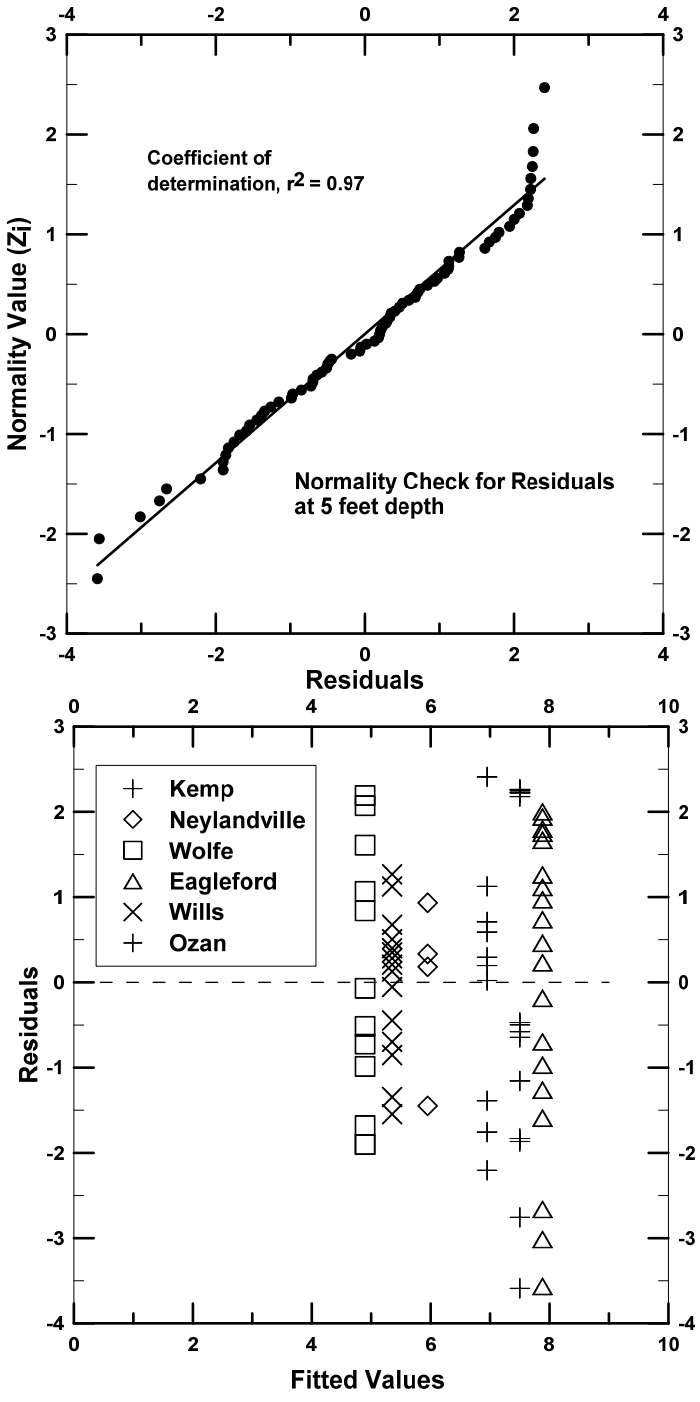


Figure 6-19 Model Adequacy plot for residual values at 5 ft. depth

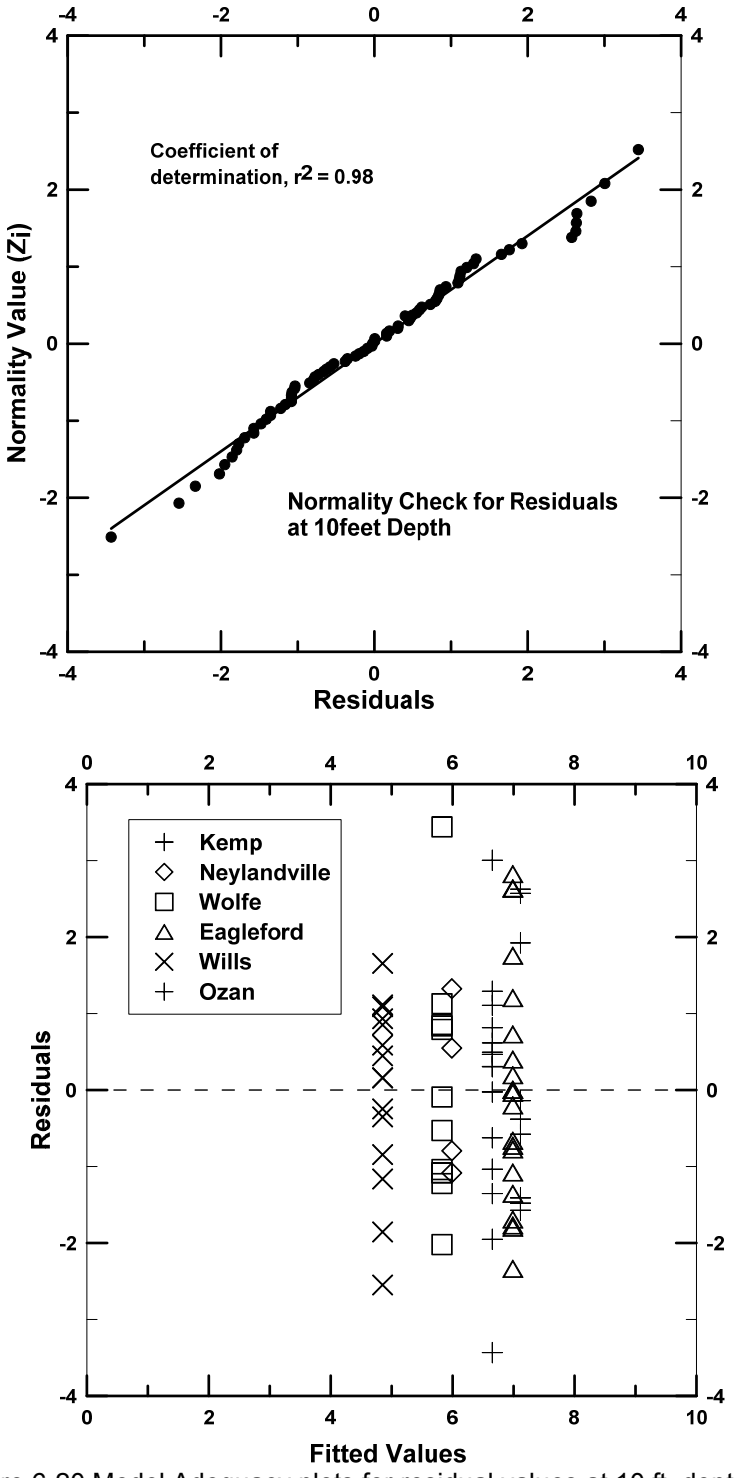


Figure 6-20 Model Adequacy plots for residual values at 10 ft. depth

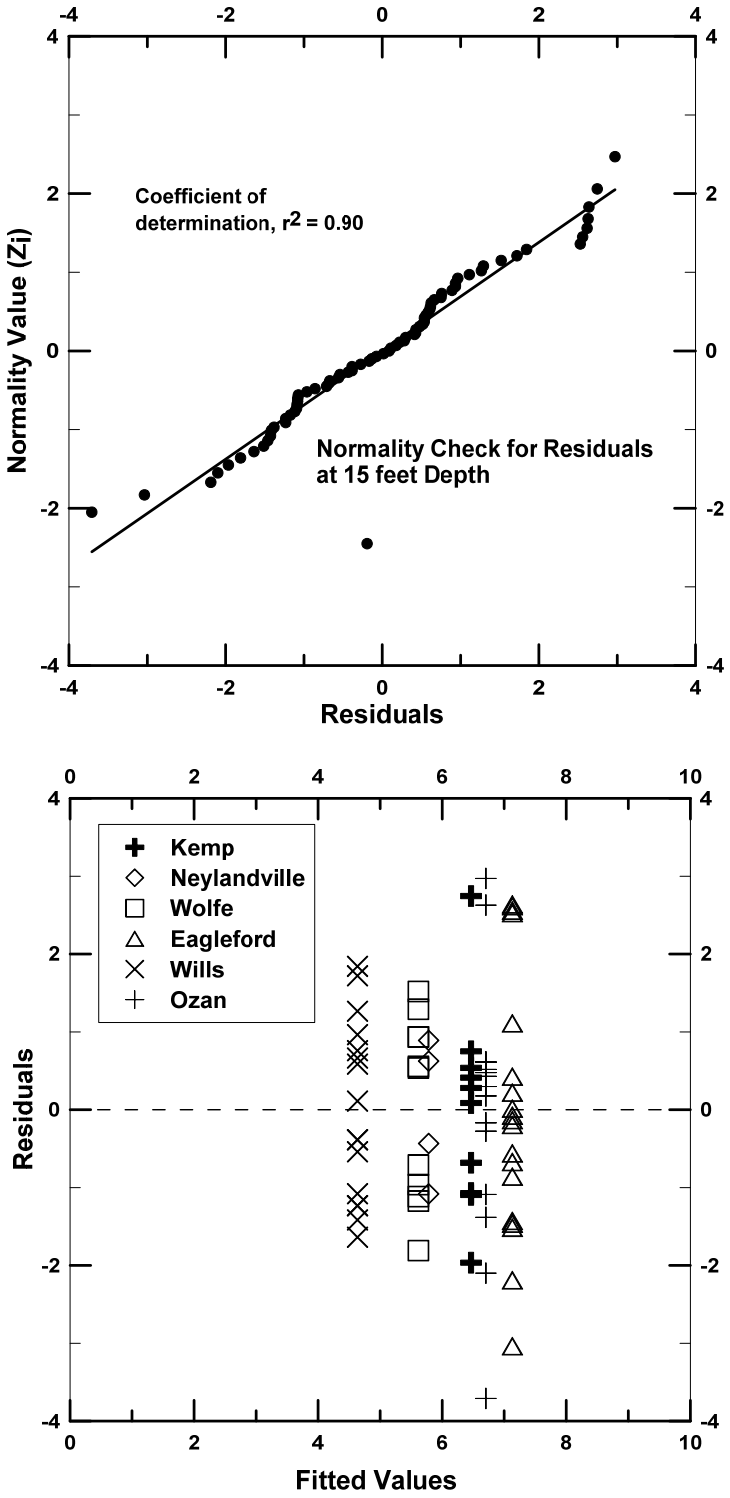


Figure 6-21 Model Adequacy plots for residual values at 15 ft. depth

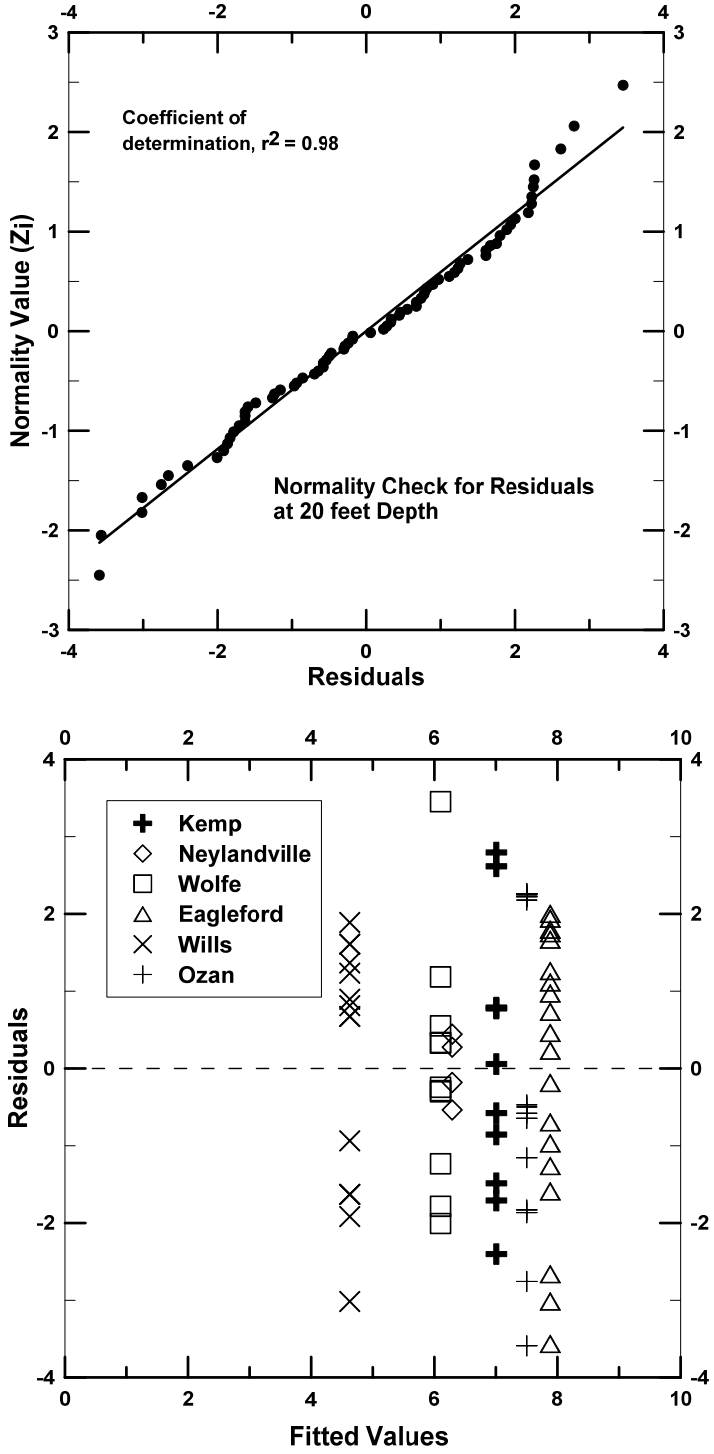


Figure 6-22 Model Adequacy plots for residual values at 20 ft. depth

From the residual normality plots, it's evident that the residual values obtained at all depths lie on a straight line, signifying that the data was normally distributed. The coefficient of determination values obtained for all the plots were higher than 0.90, depicting that the fitted model was appropriate. The residual values were also plotted with the corresponding fitted/mean values to enable observation of the structure in the residual values. In all the plots, the residuals most likely had the same variability, depicting that there were no trends. Therefore, the conclusions obtained were validated, resulting in unequal mean values in the data.

6.6.2 Check for constant variance

As a part of stationary requirement, the data had to have to equal variance` In this study, the transformed sulfate concentration values were used to evaluate for stationary in the data. Bartlett's test was used to check for constant variance present in the sulfate concentration values at all depths. This was performed by developing the following hypothesis:

$$H_0 : \sigma^2_{kemp} = \sigma^2_{wills} = \sigma^2_{neylandville} = \sigma^2_{wolfe} = \sigma^2_{ozan} = \sigma^2_{eagelford}$$

H_1 : At least one variance is different

The above hypothesis was evaluated using the statistic expressed in Equation 3.9. The value obtained from the statistic was compared to the critical value obtained at a significance value (α) 0.05. The chi-square distribution table was used to determine the critical value. Tables 6-16 to 6-19 provide the summary of the results obtained from the Bartlett's test in evaluation of constant variance at all depths.

Table 6-25 Summary of Bartlett's test results for sulfate values at 5 ft. depth

Parameter	**Result**
q	1.18
c	1.04
Test Statistic ($\chi_0^2 = 2.3026\ \frac{q}{c}$)	2.61
Critical Region	9.49
x^2 (2.61) < x^2_{critical} (9.49)	

Table 6-26 Summary of Bartlett's test results for sulfate values at 10 ft. depth

Parameter	**Result**
q	0.91
c	1.04
Test Statistic ($\chi_0^2 = 2.3026\ \frac{q}{c}$)	2.03
Critical Region	9.49
x^2 (2.03) < x^2_{critical} (9.49)	

Table 6-27 Summary of Bartlett's test results for sulfate values at 15 ft. depth

Parameter	**Result**
q	2.03
c	1.04
Test Statistic ($\chi_0^2 = 2.3026\ \frac{q}{c}$)	4.48
Critical Region	9.49
x^2 (4.48) < x^2_{critical} (9.49)	

Table 6-28 Summary of Bartlett's test results for sulfate values at 20 ft. depth

Parameter	**Result**
Q	2.97
C	1.04
Test Statistic ($\chi_0^2 = 2.3026\ \frac{q}{c}$)	6.55
Critical Region	9.49
x^2 (6.55) < x^2_{critical} (9.49)	

From the results summarized in the above tables, it was observed that the test statistic calculated was less than the critical value obtained at a significance level of 0.05. This shows that the variances in all the formations at all depths are equal.

The stationarity in the transformed sulfate data was evaluated in this study using the ANOVA test and Bartlett's test for constant mean and variance. From the results and discussions presented in Sections 6.6.1 and 6.6.2, it can be inferred that the sulfate data didn't have constant mean, but contained constant variance. This was expected, as the sulfate concentration values in the Eagleford formation enhanced the mean value at all the depths. The alternative hypothesis (H_1) that was constructed for evaluating the constant mean indicated that the mean value in the Eagleford formation was clearly different from the rest of the mean values.

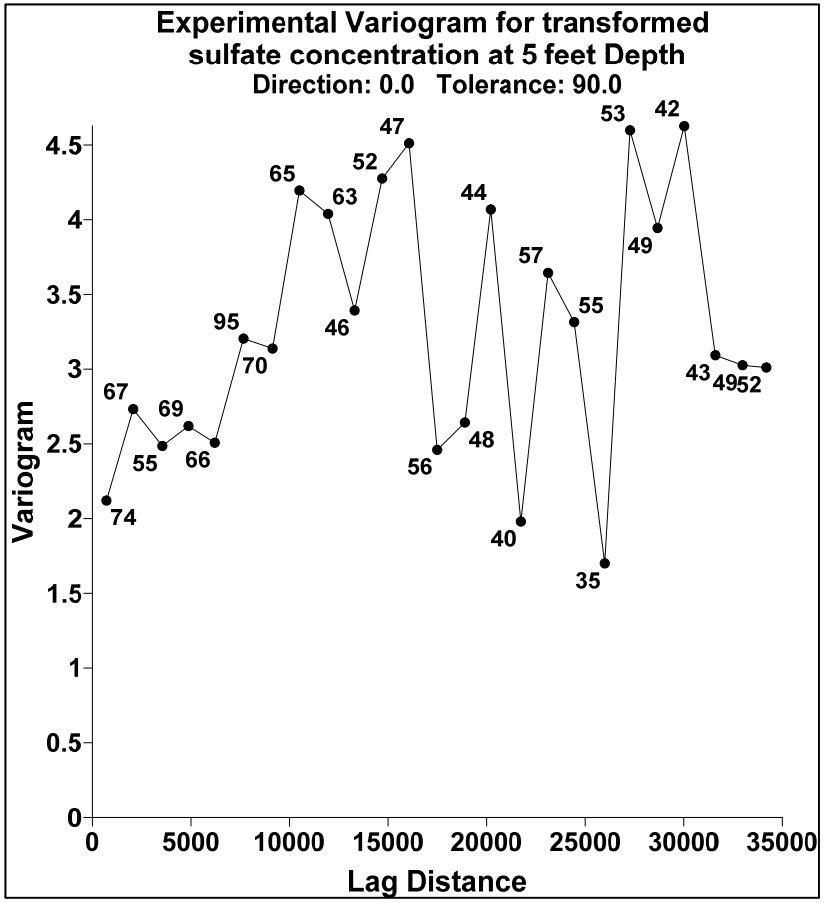
Therefore, from the results of the ANOVA test (Section 6.5.1), it can be concluded that the data was non-stationary, but the results of the Bartlett's test (Section 6.5.2) showed that the data was stationary. For further investigation, the stationarity in the data was also evaluated using the variogram. It is assumed in this study that if at least one method meets the requirement of stationarity, the spatial variability analysis can be performed.

6.7 Geostatistical Analysis

Geostatistical analysis is performed in this study to capture the variability in the sulfate concentration values. In the earlier case studies, the raw data was used directly in performing the analysis; however, in this case, the transformed data using natural logarithm was used to conduct the analysis. Using this analysis, the spatial variability in the sulfate concentrations at a particular depth in different geological formations was modeled. The developed models were used for interpretation of sulfate concentration values at unsampled locations.

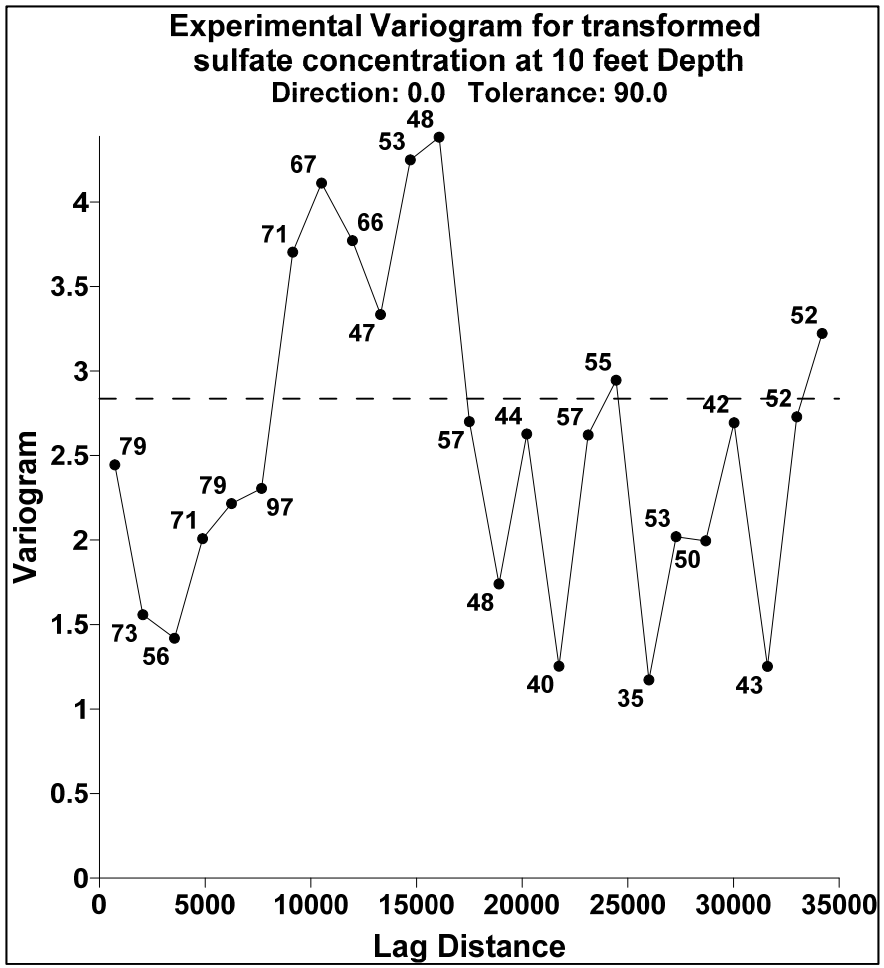
6.7.1 Experimental Variograms

Variograms are a plot of variogram values with lag distance. The variogram values were calculated using Equation 3-18. The variogram values were plotted on the y-axis, with corresponding lag distance values on the x-axis. In this section, experimental variograms were constructed using various grid parameters, so that the number of pairs used in computing a variogram value was not less than 30. As seen in the previous sections, the stationary conditions were not satisfied by conducting the ANOVA test and Bartlett's test. The ANOVA test resulted in unequal mean values in the sulfate concentration values. Therefore, the constructed variograms were not only used to capture the variability, but also used to check the stationarity of the transformed data. Figures 6-23 to 6-26 represent the experimental variograms constructed using the transformed sulfate concentration data. Along with the experimental variogram plots, the grid parameters used for constructing experimental variograms were also presented in the figures.



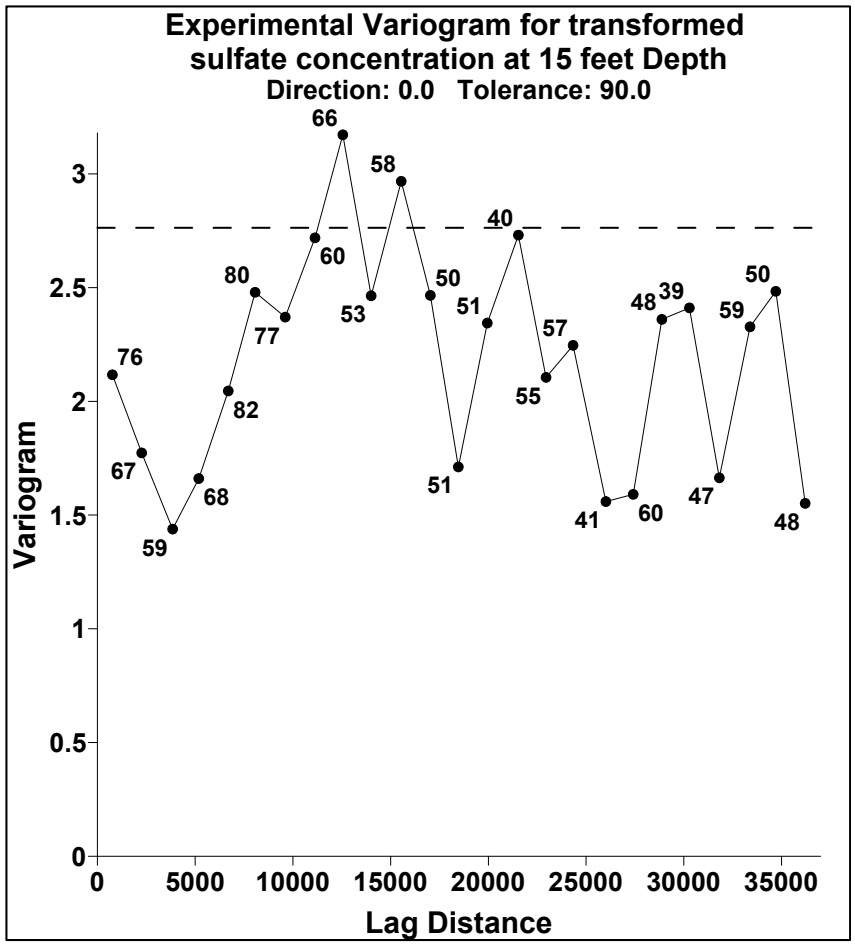
- **Estimator type : Variogram**
- **Maximum lag distance : 35000**
- **Number of lags : 25**
- **Lag Width : 1400**
- **Direction of Variogram : 0 degrees (x-direction)**
- **Tolerance : 90 degrees**

Figure 6-23 Experimental Variogram plot at 5 ft. depth



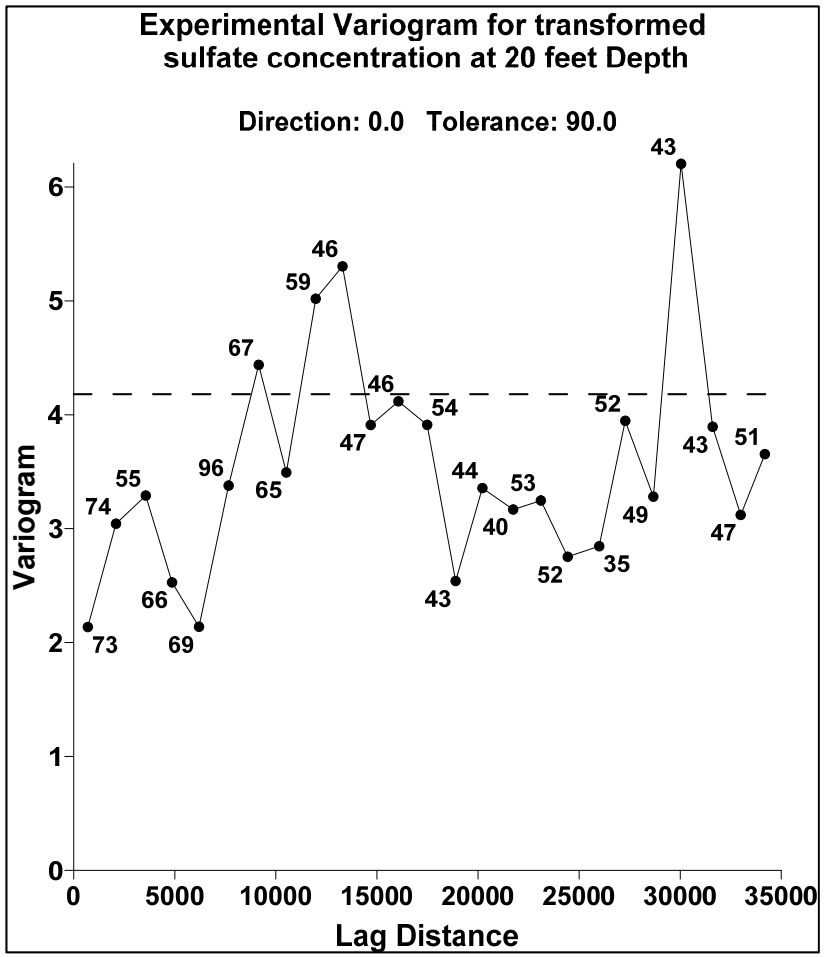
- **Estimator type : Variogram**
- **Maximum lag distance : 35000**
- **Number of lags : 25**
- **Lag Width : 1400**
- **Direction of Variogram : 0 degrees (x-direction)**
- **Tolerance : 90 degrees**

Figure 6-24 Experimental Variogram plot at 10 ft. depth



- **Estimator type : Variogram**
- **Maximum lag distance : 35000**
- **Number of lags : 25**
- **Lag Width : 1400**
- **Direction of Variogram : 0 degrees (x-direction)**
- **Tolerance : 90 degrees**

Figure 6-25 Experimental Variogram plot at 15 ft. depth



- **Estimator type : Variogram**
- **Maximum lag distance : 35000**
- **Number of lags : 25**
- **Lag Width : 1400**
- **Direction of Variogram : 0 degrees (x-direction)**
- **Tolerance : 90 degrees**

Figure 6-26 Experimental Variogram plot at 20 ft. depth

Several trial and error procedures were adopted in the selection of the grid parameters to construct the variogram plots. Due to the similar scale of variation in all the sulfate concentration values at all depths, the same grid parameters were reflected in all the variogram plots. It can be inferred from the plots, that the variogram values at all depths increased with an increase in the lag distance. There was a sudden drop of variogram values after lag distance of 15000m. Also, it can be inferred that the variogram values are varying about the global variance value and there is no indication of trends present in the data, as the variogram values are reaching to a definite sill with an increase in lag distance. However, the variogram values in all the plots seemed to vary within a short lag distance. This can be attributed to the high variability present in the sulfate concentration values. In order to capture the spatial variability present in the transformed sulfate concentrations, the semi-variogram value was normalized by the global variance. This type of variogram is called standardized variogram.

$$\gamma(h) = \frac{\frac{1}{2n(h)} \sum_{i=1}^{n(h)} [z(x_i + h) - z(x_i)]^2}{S^2} \qquad (6.1)$$

In the standardized variogram, i.e., when the semi-variogram values are normalized by the lag variance, the influence of the local means can be reduced. The same result was reflected in the ANOVA test, that, at all the depths, the means were not constant. However, by normalizing the variogram values, the varying local mean values of sulfate concentration can be subsided. The standardized variogram plots were constructed using the same grid parameters. Figures 6-27 to 6-30 present the standardized variogram plots for the transformed sulfate concentration at all depths.

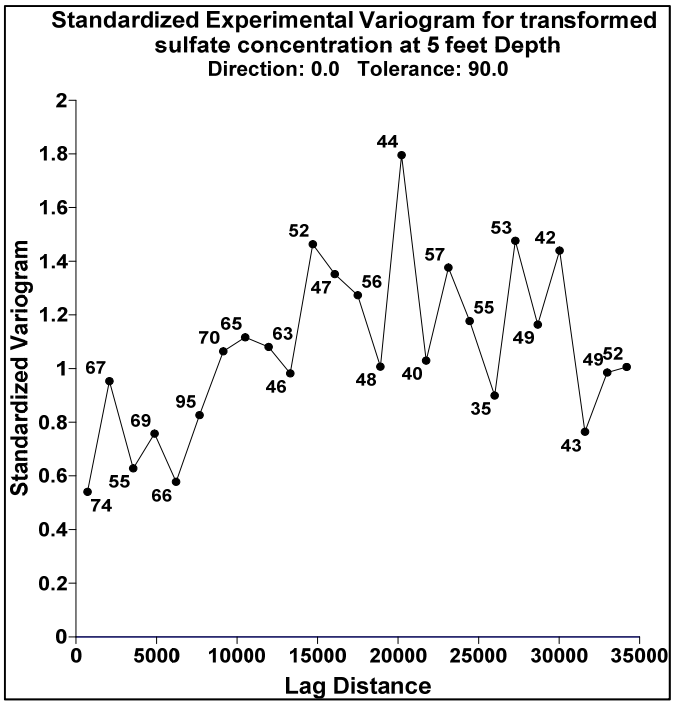


Figure 6-27 Standardized variogram plot at 5ft depth

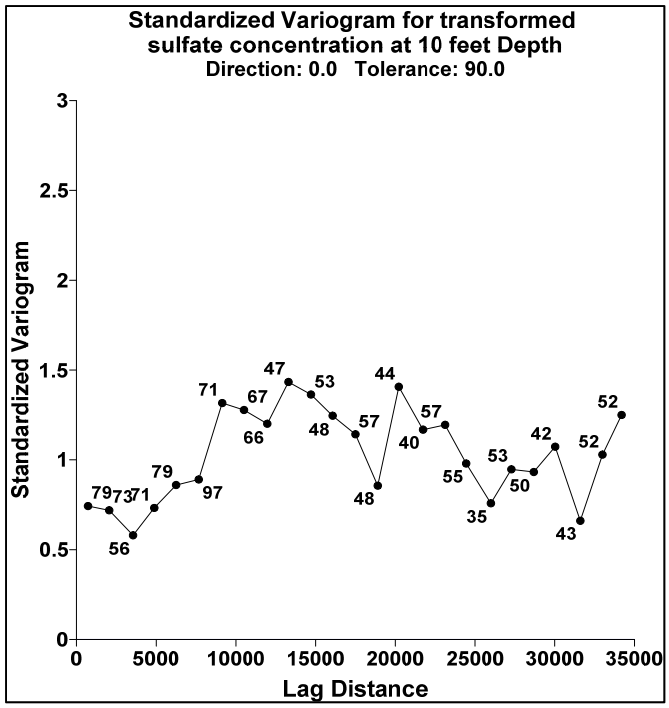


Figure 6-28 Standardized variogram plot at 10 ft. depth

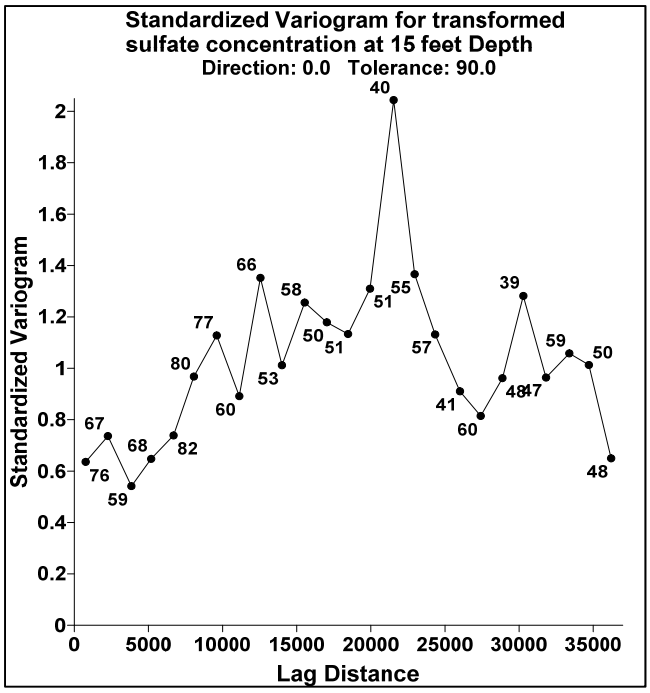


Figure 6-29 Standardized variogram plot at 15 ft. depth

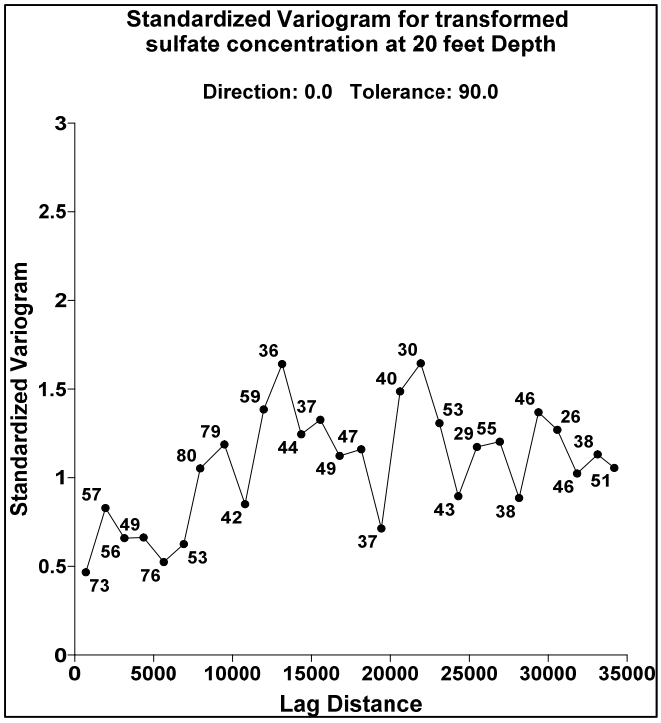


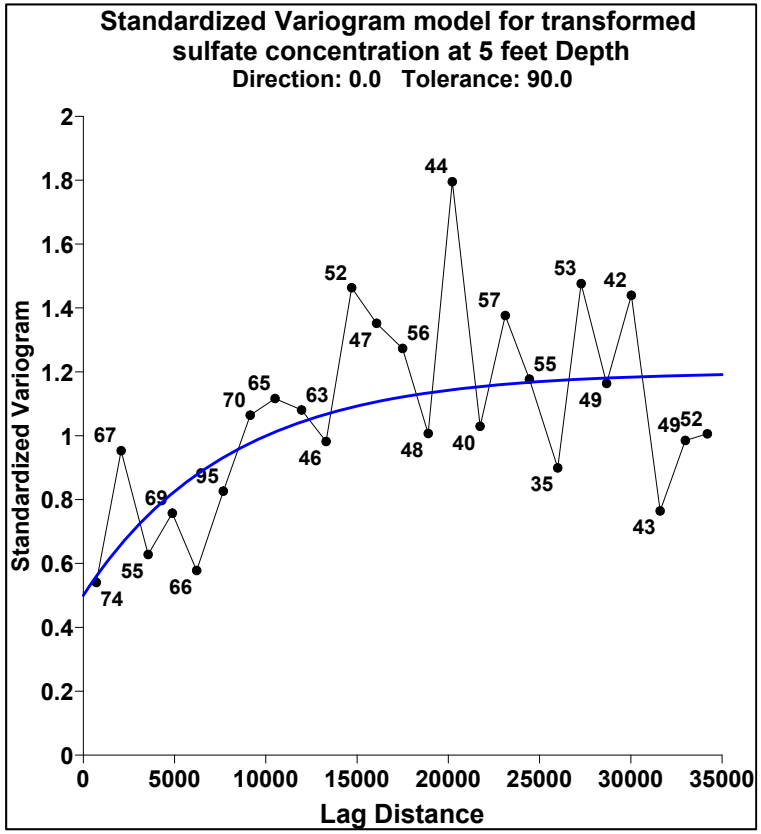
Figure 6-30 Standardized variogram plot at 20 ft. depth

The figures above represent the standardized variogram plots, where the standardized variogram value was represented on the y-axis and the corresponding lag distance on the x-axis. A significant improvement was observed in the variogram plot after standardizing the variogram values. At all depths, the standardized variogram value increased with an increase in the lag distance and reached a sill value. These plots were used in further analysis for developing the spatial variability model.

6.7.2 Variogram Model

The standardized variogram plots constructed earlier were used to capture the spatial variability present in the transformed sulfate concentration values. The different models presented in Section 3.7.2 were used to model the spatial variability. The three characteristics of a variogram plot (range, sill, and nugget) were modeled in this section. It was observed that at a lag distance of '0,' the variogram values in all the plots were not equal to zero, depicting the nugget effect present in the sulfate concentration values.

Several models such as Gaussian, spherical, and exponential were used to reflect the variability present in the data. Of all the models, the exponential function depicted the best model that can be ascribed to the data. Figures 6-31 to 6-34 present the spatial variability models that were developed for sulfate concentration values obtained at all depths.



Variogram model at 5 feet depth:

- **Model : Nugget effect + Exponential**
- **Nugget (C_0) : 0.5**
- **Scale (C) : 0.7**
- **Range/ Length (A) : 8000**

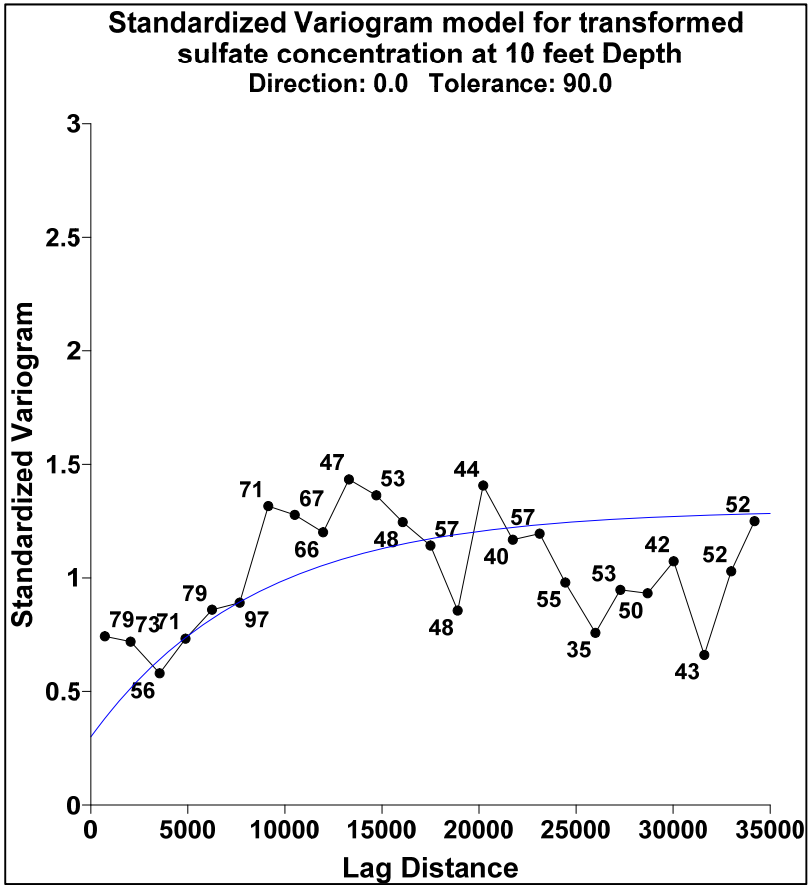
Theoretical Equation:

$$\gamma(h) = C_0 + C\,[1 - \exp(-h/a)] \text{ for } h > 0$$

Model Equation :

$$\gamma(h) = 0.5 + 0.7\,[1 - \exp(-h/8000)] \text{ for } h > 0$$

Figure 6-31 Standardized Variogram model at 5 ft. depth



Variogram model at 10 feet depth:

- **Model : Nugget effect + Exponential**
- **Nugget (C_0) : 0.3**
- **Scale (C) : 1.0**
- **Range/ Length (A) : 8500**

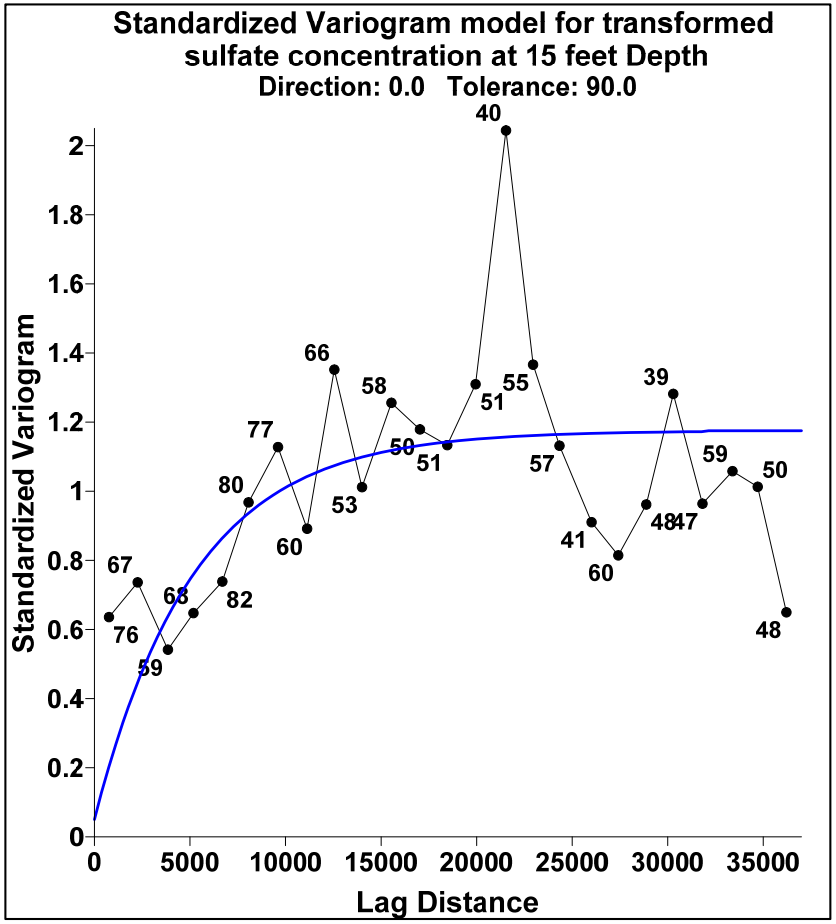
Theoretical Equation:

$\gamma(h) = C_0 + C\,[1 - \exp(-h/a)]$ for $h > 0$

Model Equation :

$\gamma(h) = 0.3 + 1.0\,[1 - \exp(-h/8500)]$ for $h > 0$

Figure 6-32 Standardized Variogram model at 10 ft. depth



Variogram model at 15 feet depth:

- **Model : Nugget effect + Exponential**
- **Nugget (C_0) : 0.05**
- **Scale (C) : 1.12**
- **Range/ Length (A) : 5200**

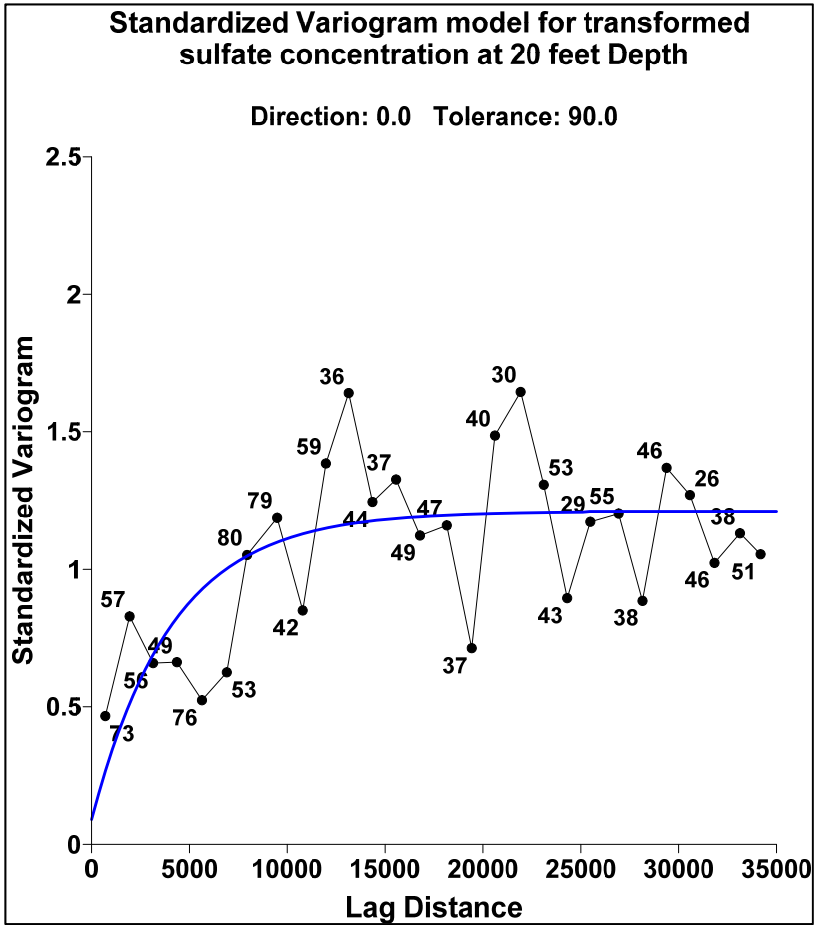
Theoretical Equation:

$\gamma(h) = C_0 + C\,[1 - \exp(-h/a)]$ for $h > 0$

Model Equation :

$\gamma(h) = 0.05 + 1.12\,[1 - \exp(-h/5200)]$ for $h > 0$

Figure 6-33 Standardized Variogram model at 15 ft. depth



Variogram model at 20 feet depth:

- **Model : Nugget effect + Exponential**
- **Nugget (C_0) : 0.09**
- **Scale (C) : 1.12**
- **Range/ Length (A) : 4100**

Theoretical Equation:

$\gamma(h) = C_0 + C\,[1 - \exp(-h/a)]$ for $h > 0$

Model Equation :

$\gamma(h) = 0.09 + 1.12\,[1 - \exp(-h/4100)]$ for $h > 0$

Figure 6-34 Standardized Variogram model at 20 ft. depth

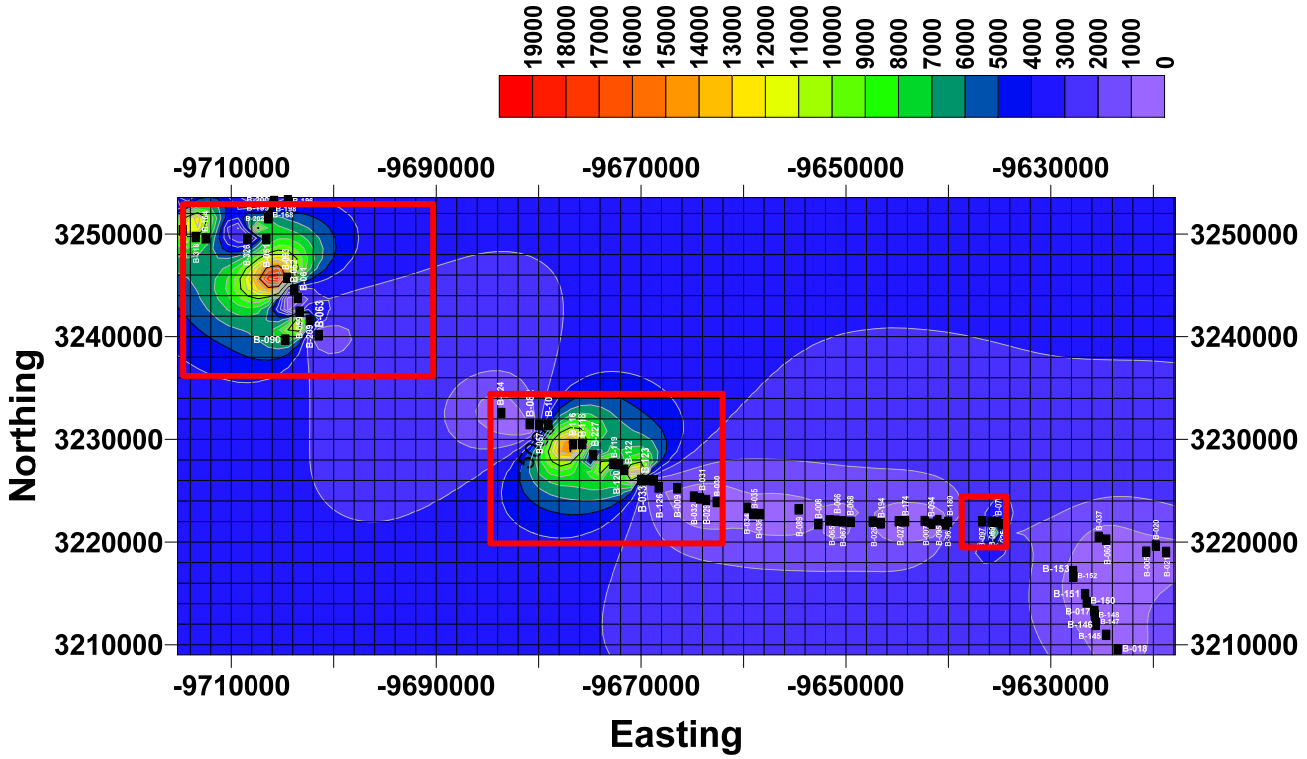
The exponential function, along with the nugget effect, was considered In order to model the spatial variability in sulfate concentrations in different models. The nugget effect, which is an interception of the variogram curve towards the y-axis, varied from 0.05 to 0.5. The range, i.e., spatial correlation distance obtained through the experimental variogram plot, varied from 4100-8500. This range value was too high, due to the limited observations obtained at larger intervals with larger variability. The influence of the neighboring values around the unsampled location was determined using the exponential and nugget models.

6.7.3 Kriging Analysis

Kriging analysis was performed to predict the sulfate concentrations at unknown locations. The spatial variability models developed at different depths in the earlier section were used along with kriging algorithm. The predictions obtained were produced in the form of a contour map, as shown in Figures 6-35 to 6-38. The grid parameters that were used to produce the contour maps for all the depths were kept the same, and these details are presented in Table 6-29 below.

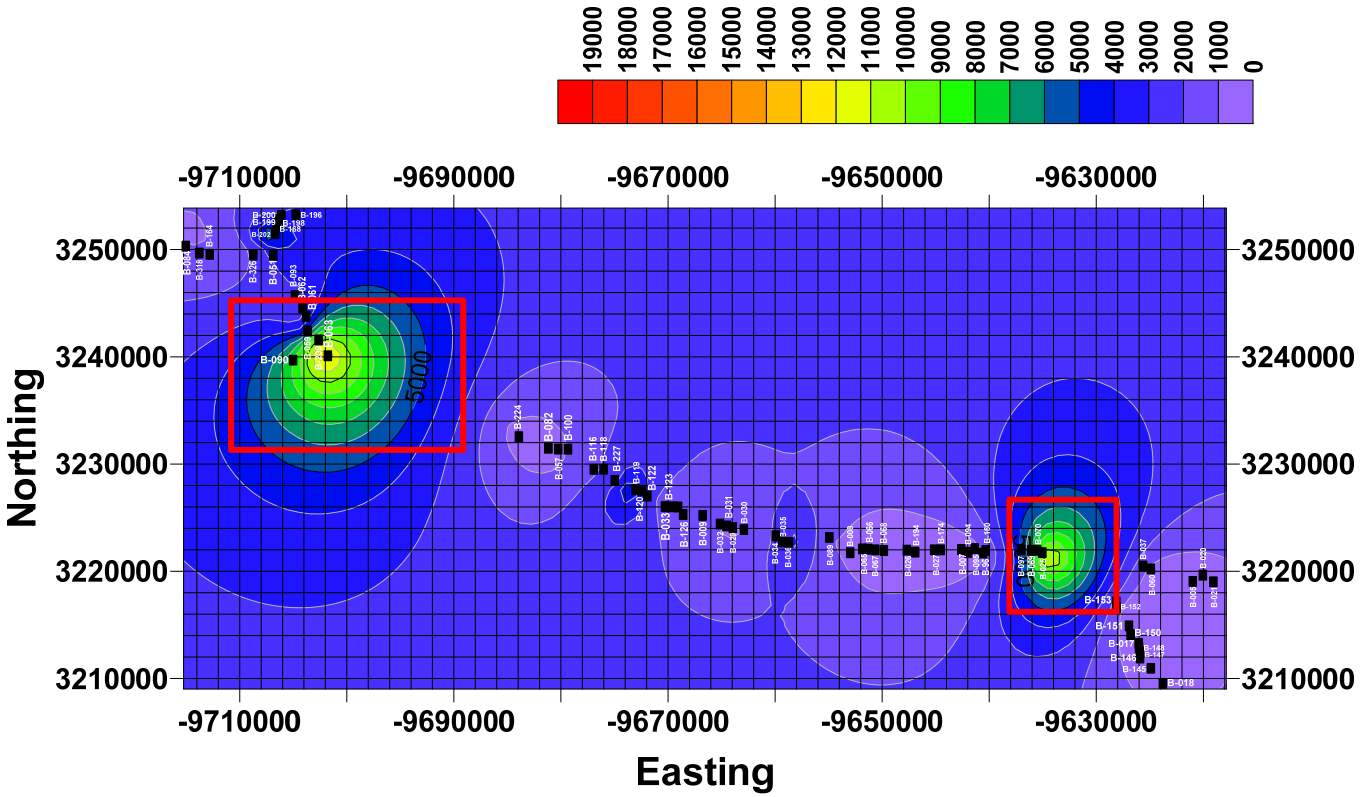
Table 6-29 Grid parameters selected for kriging analysis

Parameters selected				
Depth	5	10	15	20
Maximum lag distance	35000	35000	37000	35000
Angular divisions	180	180	180	180
Radial divisions	100	100	100	100



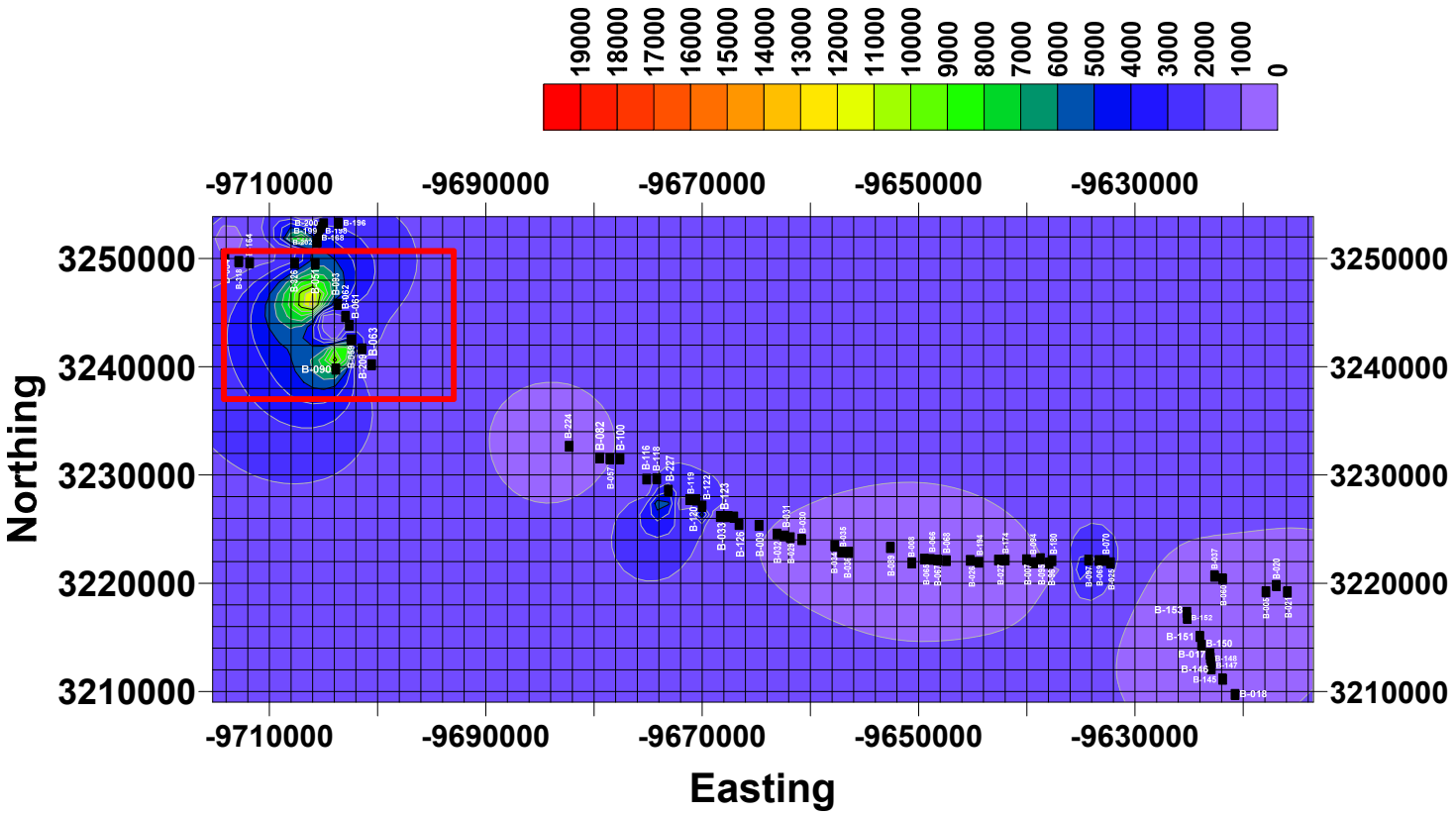
Contour Map for Sulfate concentration (ppm) at 5 feet depth

Figure 6-35 Contour map for sulfate concentration (ppm) at 5 ft. depth



Contour Map for Sulfate concentration (ppm) at 10 feet detpth

Figure 6-36 Contour map for sulfate concentration (ppm) at 10 ft. depth



Contour Map for Sulfate concentration (ppm) at 15 feet detpth

Figure 6-37 Contour map for sulfate concentration (ppm) at 15 ft. depth

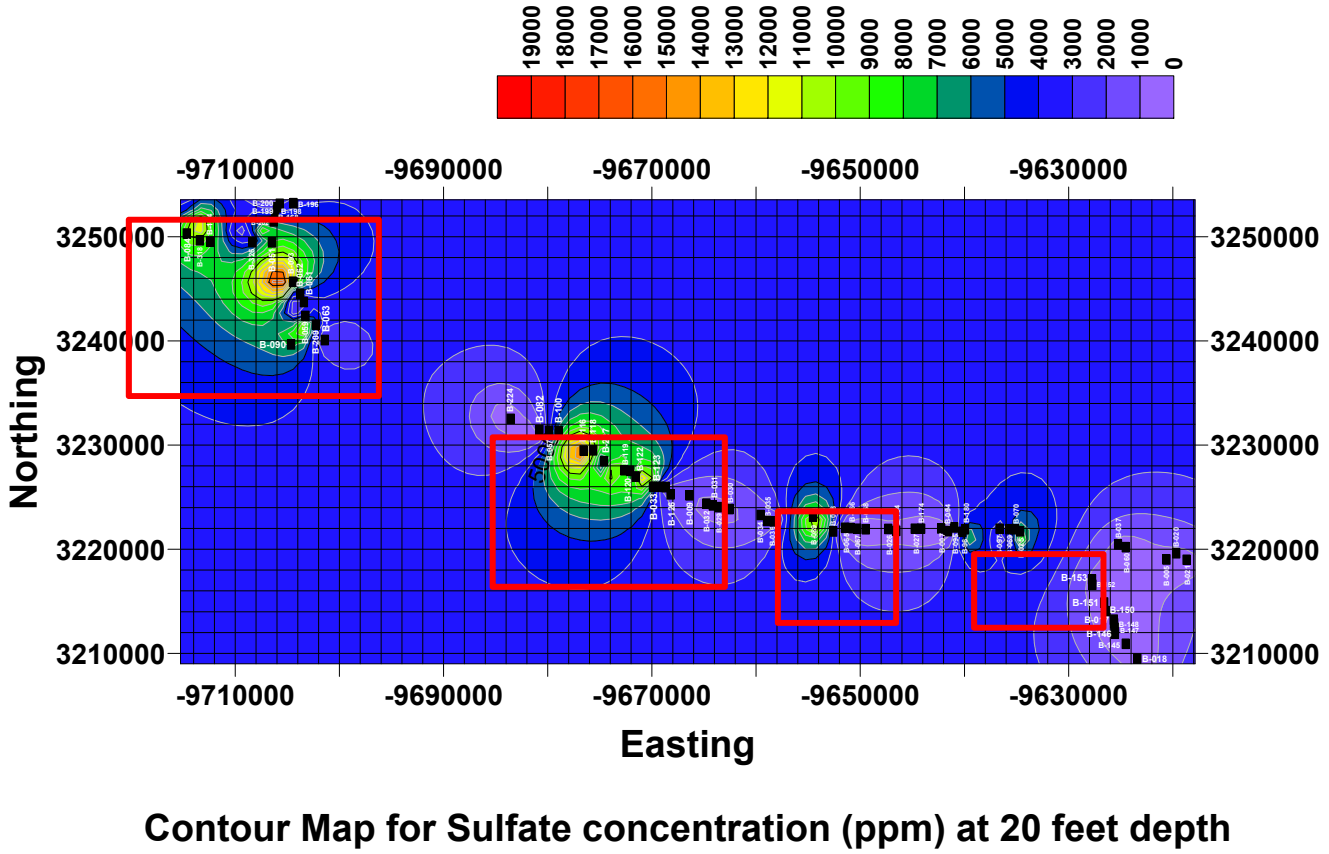


Figure 6-38 Contour map for sulfate concentration (ppm) at 20 ft. depth

The above plots provide the sulfate concentrations at different depths. The color scale of the contour maps indicates sulfate concentration in parts per million (ppm). The color scale is kept constant, so that the location of the high sulfate content at any depth can be easily identified. The x-axis in the contour map indicates Northing and y-axis indicates Easting. The black rectangular boxes indicate the boreholes conducted along the new proposed pipeline.

Due to the limited number of observations and correlation distance specified in the model, the sulfate concentrations, after a certain distance, remain constant. In the contour map produced at 5 ft. depth, it can be observed that high sulfate concentrations are present in Eagleford and Wolfe formations, which are highlighted in red rectangular boxes. In the contour map produced at 10 ft. depth, high sulfate concentrations are observed at Eagleford and Kemp formations.

At a 15 ft. depth, only a few locations in the Eagleford formation have high sulfate concentrations. This is in accordance with the results obtained from the histograms, where the maximum number of observations was well below 3000 ppm of sulfate content. From the contour map produced at a 20 ft. depth, it can be observed that the high sulfate concentration is present at Eagleford, Wolfe and Kemp geological formations. In order to check the appropriateness of the prediction maps produced using the kriging analysis, and in accordance with variability models, the cross-validation technique was performed.

6.7.4 Cross Validation

In the process of cross validation, the sulfate concentration values at different locations were deleted and the contour maps were produced, again using the developed spatial variability model. From the contour maps produced, using the digitization tool available in the surfer software, the predicted values were determined. The predicted values were compared with the original values. Thus, the effectiveness of the spatial

variability model is was validated. Figures 6-39 to 6-42 represent the cross-validation contour maps.

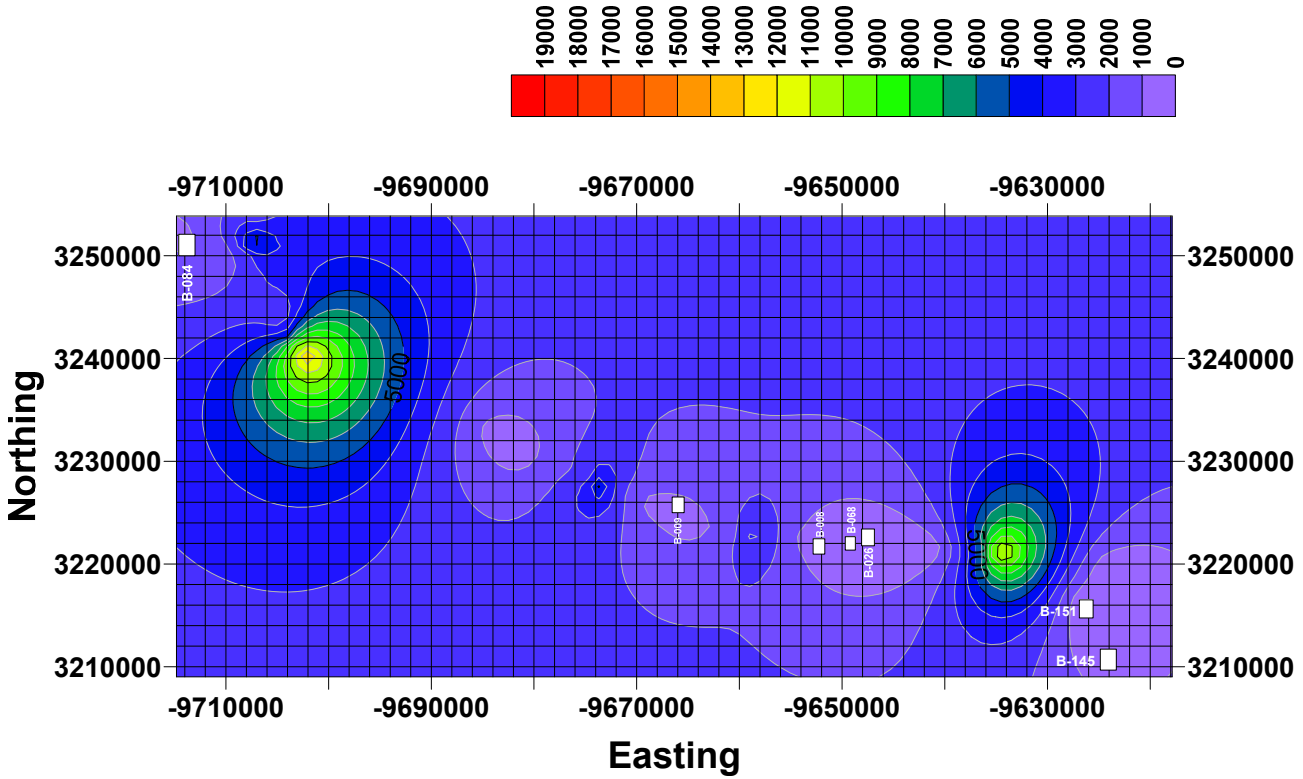


Figure 6-39 Cross validation map at 5 ft. depth

The rectangular white boxes represent the locations of the deleted data. Using the digitize tool, the sulfate concentrations at the deleted locations from the new contour maps were determined and compared to the original data. Table 6-30 and Figure 6-40 present the results of the sulfate concentration values.

Table 6-30 Comparison of actual and predicted sulfate concentration values

S.no	Actual Values (ppm)	Predicted Values (ppm)
1	315	370
2	280	305
3	1200	998
4	1270	990
5	4200	5100
6	1130	1200
7	975	950

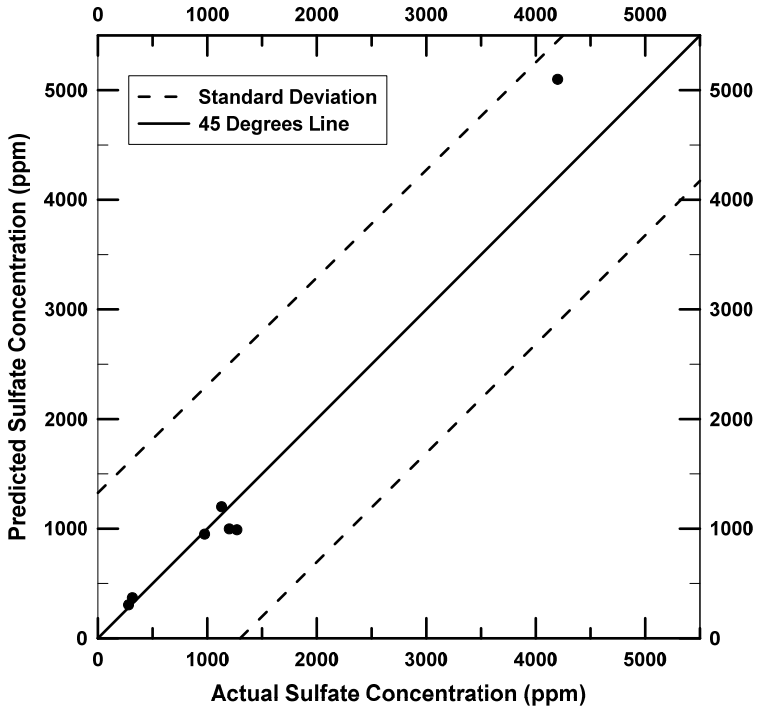


Figure 6-40 Comparison of actual and predicted sulfate concentrations at 5 ft. depth

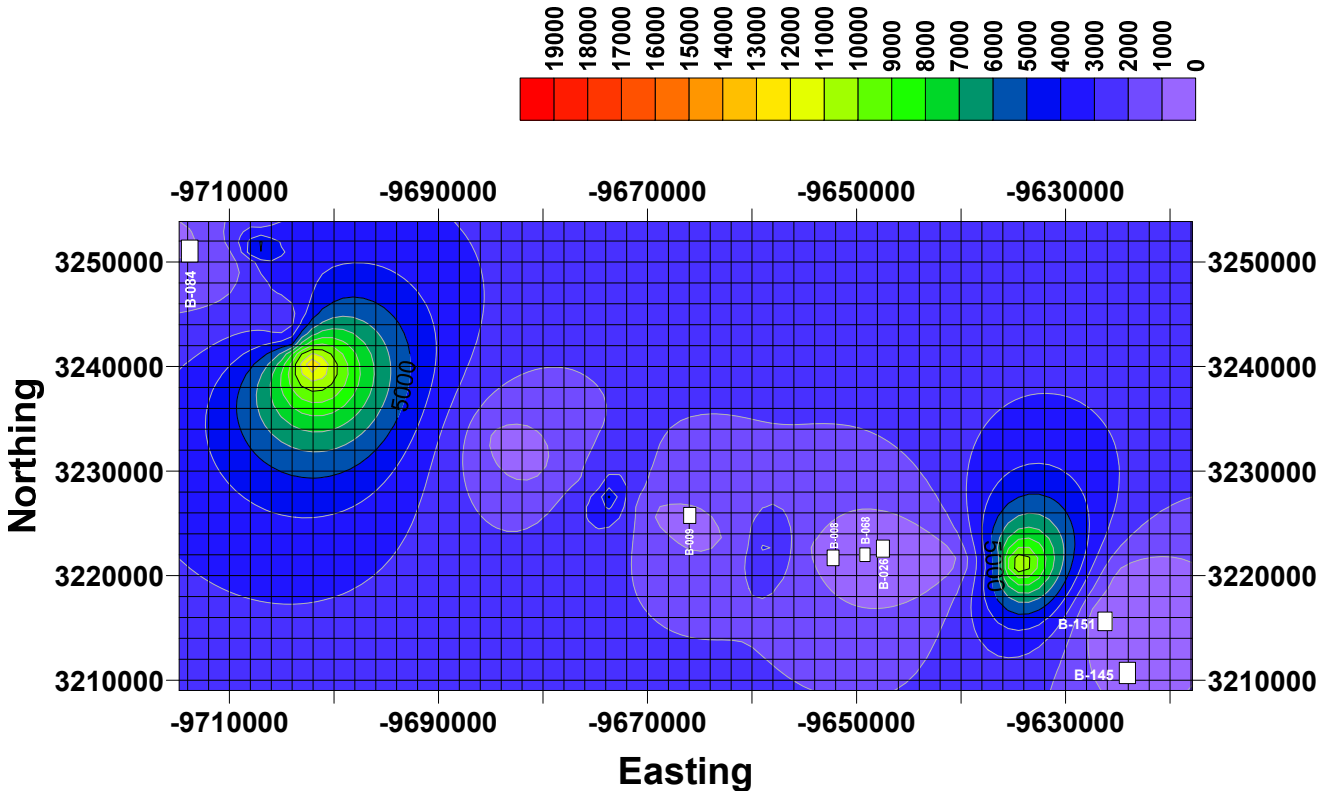


Figure 6-41 Cross validation map at 10 ft. depth

Table 6-31 and Figure 6-42 present the results of the sulfate concentration values at 10 ft. depth for evaluating the appropriateness of the model developed.

Table 6-31 Comparison of actual and predicted sulfate concentration values

S.no	Actual Values (ppm)	Predicted Values (ppm)
1	380	399
2	200	252
3	750	683
4	255	308
5	1050	974
6	2820	1914
7	1500	997

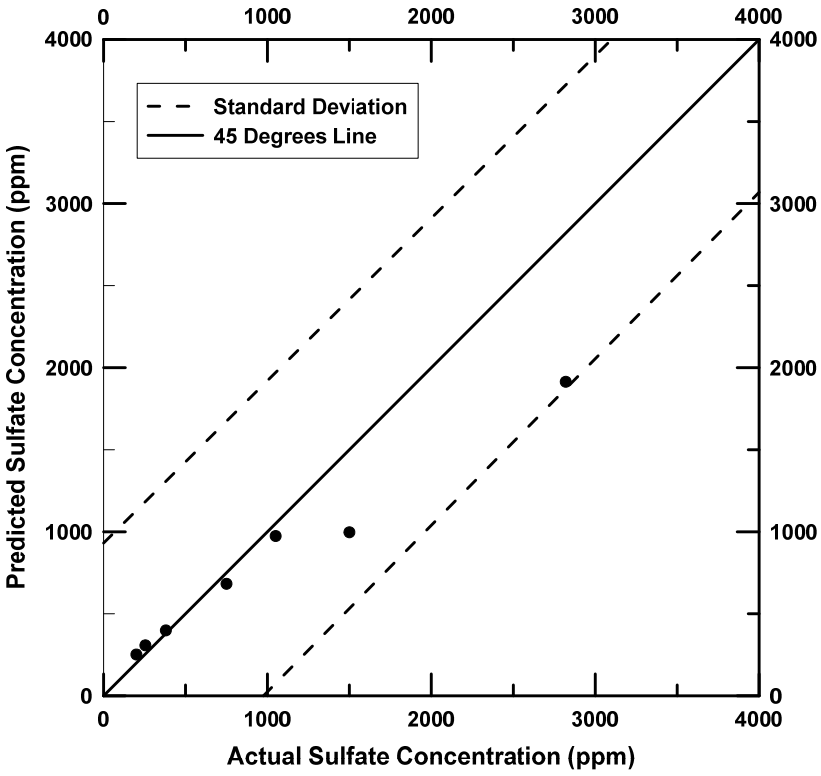


Figure 6-42 Comparison of actual and predicted sulfate concentrations at 10ft. depth

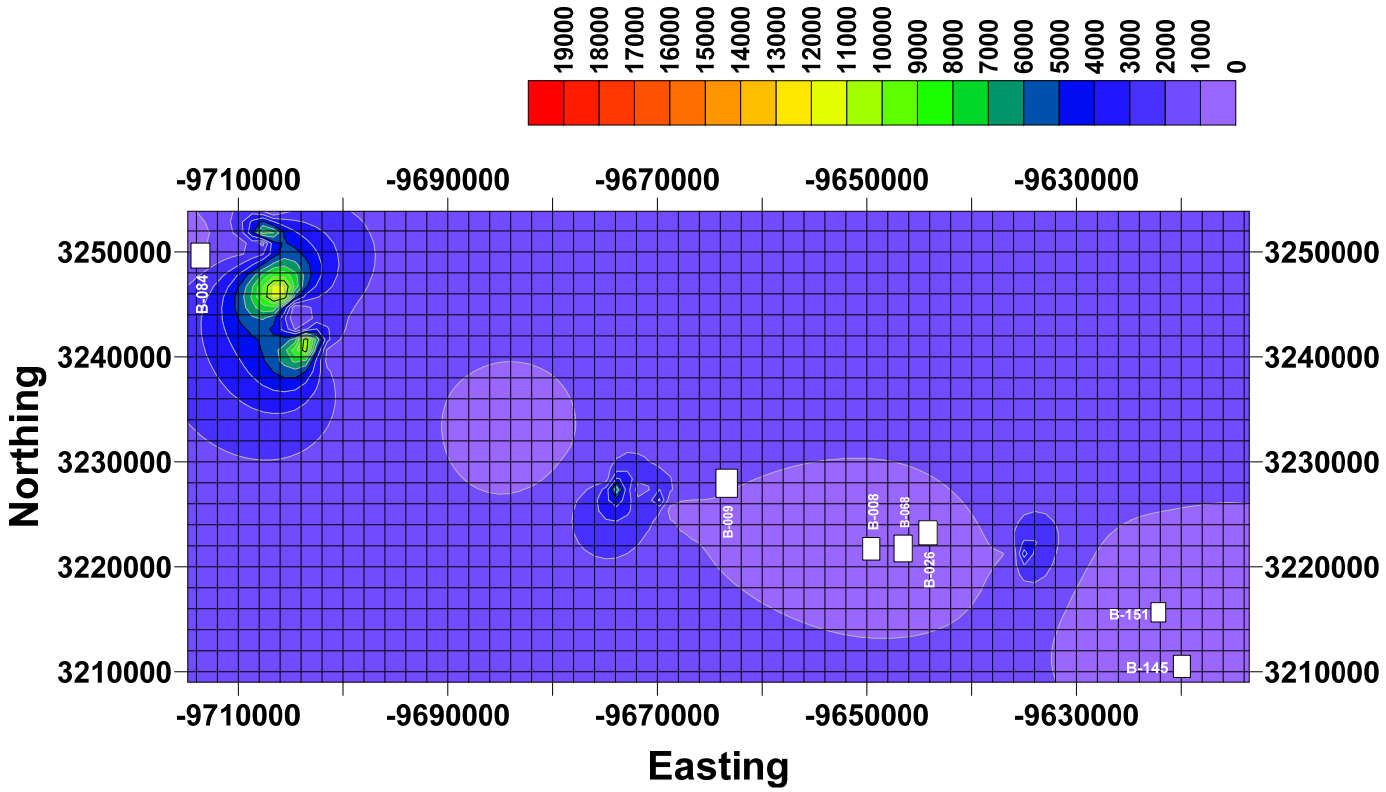


Figure 6-43 Cross validation map at 15 ft. depth

Table 6-32 and Figure 6-44 present the results of the sulfate concentration values at a 15 ft. depth for evaluating the appropriateness of the model developed.

Table 6-32 Comparison of actual and predicted sulfate concentration values

S.no	Actual Values (ppm)	Predicted Values (ppm)
1	220	280
2	70	130
3	90	177
4	215	277
5	3800	1900
6	1500	1200
7	605	565

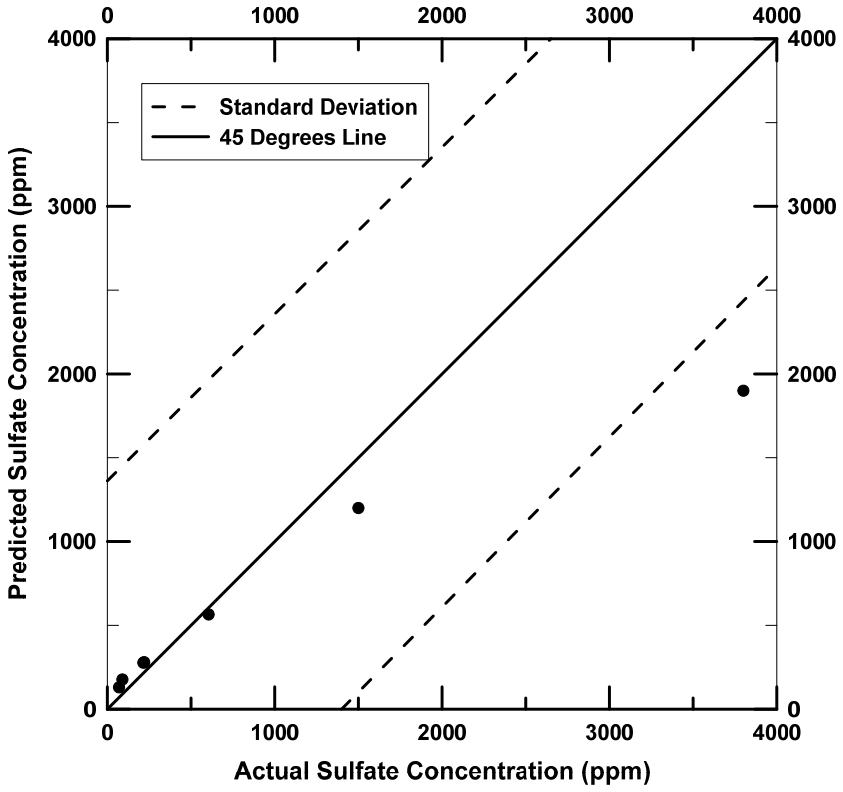


Figure 6-44 Comparison of actual and predicted sulfate concentrations at 15 ft. depth

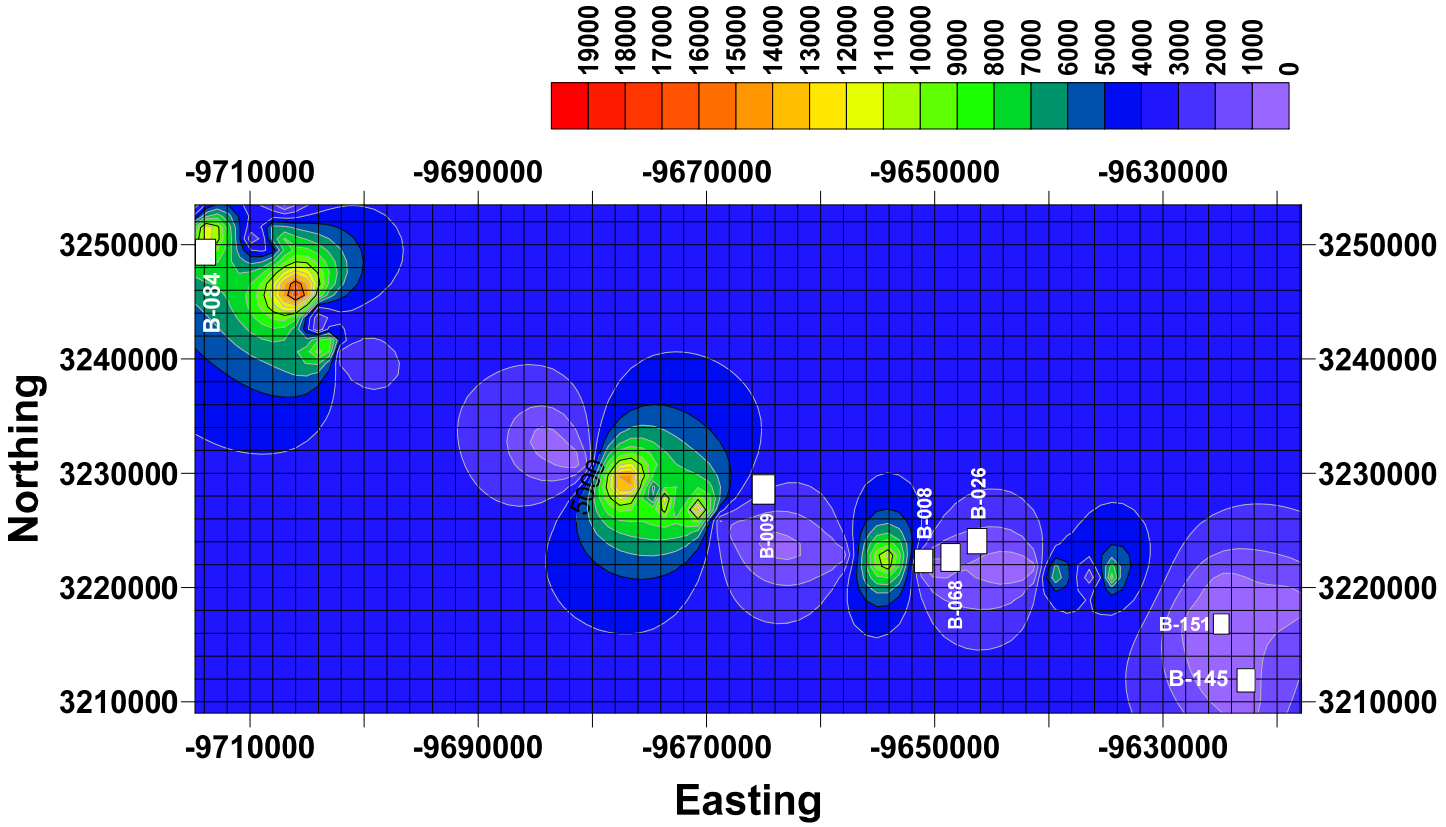


Figure 6-45 Cross validation map at 20 ft. depth

Table 6-33 and Figure 6-46 present the results of the sulfate concentration values at 20 ft. depth for evaluating the appropriateness of the model developed.

Table 6-33 Comparison of actual and predicted sulfate concentration values

S.no	Actual Values (ppm)	Predicted Values (ppm)
1	350	390
2	250	210
3	60	180
4	470	590
5	4200	5800
6	1130	1220
7	710	920

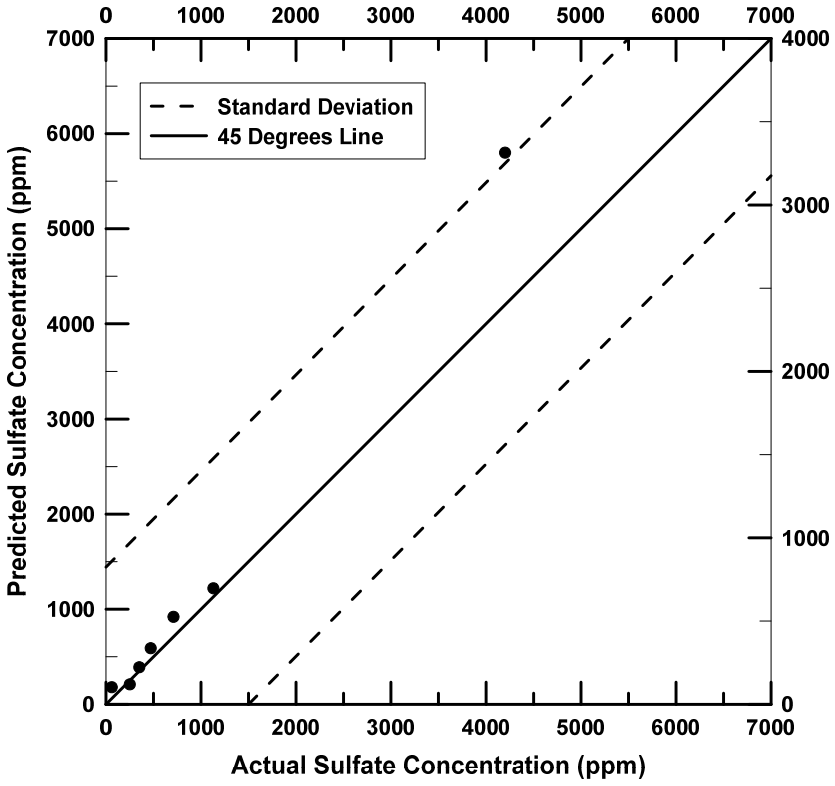


Figure 6-46 Comparison of actual and predicted sulfate concentrations at 20 ft. depth

Figures 6-39 to 6-46 presented above represent the cross-validation maps and comparison of actual and predicted values using a 45 degree line plot. In the 45 degree line plot, the actual values were plotted on the x-axis and the predicted values on the y-axis. If the actual and predicted values are the same, they should fall on the 45 degree line. The comparison plots depict that most of the predicted values are close to the 45 degree line. However, at high sulfate concentrations, the predictions are either close to or outside the 1-standard deviation line. This is because of the large variability in short distances in the actual data. However, most of the points fall close to the 45 degree line, indicating the spatial variability models developed are appropriate.

6.8 Summary

The presence of high sulfates in soils is becoming a huge concern to civil infrastructure in projects that require ground improvement tasks with calcium-based stabilizer treatments. Keeping in mind the sulfate heaving mechanism, a comprehensive laboratory study was conducted to determine sulfates present in the soils at different depths for a pipeline project. The sulfate concentrations obtained were used to conduct the spatial variability analysis. The spatial variability models were developed by constructing standardized variograms. The kriging analysis was performed to determine the sulfate concentration at unsampled locations. Below are some of the findings of this study.

- The framework developed in Chapter 3 was validated by predicting the sulfate concentration at unsampled locations.
- The spatial variability present in the sulfate concentration values was modelled using the nugget and exponential functions.

- The spatial correlation distance obtained in this study for sulfate concentration ranged from 4100 to 8500. The high correlation distance was obtained due to limited bore holes available at large distances.
- The natural logarithmic transformation function implemented in this study successfully transformed the non-Gaussian data to Gaussian.
- In order to reduce the effect of the locally varying means, the standardization of variogram values seems appropriate.

Chapter 7

Summary, Conclusions and Recommendations

7.1 Summary and Conclusions

The variability present in soil properties is a growing concern to geotechnical engineers. The variability studies are often confined to textbooks, where limited information is provided on how to incorporate the variability in analyzing the soil properties. In this research, an attempt was made to provide a framework for incorporating the spatial variability present in soil properties into prediction analysis. The framework developed was validated by analyzing three different soil properties in the different studies.

In Chapter 4, spatial variability analysis was performed on the natural soils, where the friction angle parameter of soils was evaluated from the CPTU test data. The CPTU data was obtained as a part of an exploration program for a national highway project conducted in China. Data from five CPTU bore holes was considered for the analysis. The soil classification was performed on the CPTU data, and it was observed that the subsurface profile was predominantly silty sand soils.

The friction angle parameter of silty sand soils was interpreted and utilized for evaluating the statistical characteristics and spatial variability analysis. The friction angle values followed the normal distribution, which was evaluated by employing the Shapiro-Wilk test and normal-quantile plots. The spatial variability analysis was performed in accordance with the framework developed in Chapter 3. It was observed that the friction angle values followed an exponential model with non-linear behavior. With distance, the exponential model reached an asymptotic value, indicating that the friction angle values were widely dissimilar at a spatial correlation distance ranging from 2.6 m to 4.3 m. This

explains the variability of friction angle values in a vertical direction, indicating that the stratification of the soil layers is necessary after a distance of 4.3 m.

The kriging analysis performed, using the spatial variability models, provided the unbiased estimates of the friction angle values. The predictions were validated using the cross-validation technique. It was shown that geostatistics can effectively work with fewer bore holes, optimizing the project costs. This concludes that the framework developed was successful in incorporating the spatial variability present in the natural soils in the prediction analysis.

In Chapter 5, spatial variability analysis was performed on the stiffness measurements of the pipeline bedding material. The CLSM was used as a bedding material, which was prepared using a native high plasticity soil. The stiffness measurements were obtained after a curing period of 1, 3, 7, 14 and 28 days, respectively, using the SASW non-destructive method. In this study, geostatistics was used to study the variation of stiffness measurements with time.

The statistical distributions of the stiffness measurements were more likely normally distributed. The spatial variability analysis conducted in accordance with Chapter 3 revealed that the spatial correlation distance increased from 8m to 24m, with an increase in the curing period. The kriging maps produced, using the spatial variability models, depict the uniform stiffness development on all days. Therefore, it can be concluded that the CLSM prepared using native high plasticity soil can be effectively used as bedding material for pipelines.

In Chapter 6, spatial variability analysis was performed on the sulfate concentrations present in the soils. When sulfates present in the soils react with calcium-based stabilizers, the soil behaves as an expansive soil, which causes distress to the civil infrastructure. In this study, sulfates concentrations present in soils, along the pipeline

alignment, were determined using the modified UTA method. Geostatistics was used in this study to perform the spatial variability modeling and to provide the distribution of sulfates at unsampled locations.

The statistical distributions of the sulfate concentrations revealed an exponential trend present in the values, which explains a large scatter present in the sulfate concentration values. This created a limitation on using the geostatistical modeling. However, this problem was approached by transforming the data, using the Box-Cox transformations method. The spatial variability analysis conducted on the transformed data reflects the non-linear behavior in the spatial correlation distance, which ranged from 4100 to 8500. The exponential function, along with nugget effect, were used to model the spatial variability present in the sulfate concentrations. This is because, the exponential function describes the non-linear behavior, and the nugget effect captures the large variation in short scale measurements. Through this, the sulfate contour maps were produced, using kriging analysis in all six geologic formations, which will be helpful in determining the stabilizer concentrations for ground improvement works.

The above three geostatistical studies presented in Chapters 4, 5 and 6 evaluated the variability present in the in-situ soil properties in field conditions using geostatistics. Kriging was comprehensively used, along with the variogram models, which effectively dealt with the properties having low and high variability. The framework developed successfully incorporated the spatial variability present into the in-situ soil properties in prediction analysis.

7.2 Limitations of the Framework

The framework developed in Chapter 3 was successful in analyzing the spatial variability present in soil properties. This can be used with any field in-situ properties to

model the variability and to perform predictions. However, there are few limitations that need to be considered before ascertaining the results:

- The assumption of Gaussian distribution is highly inappropriate for performing geostatistical analysis using the developed framework.
- The Box-Cox transformation cannot be used for transforming both non-gaussian and non-stationary data.
- The minimum number of pairs needed for computing the variogram value is 30.
- The stationarity of the data should be evaluated before performing the geostatistical analysis. The non-stationarity in the data could lead to biased predictions.
- The quality and quantity of the data that is being considered for the analysis is extremely important. Using the raw field properties is more efficient than using the data from empirical correlations.

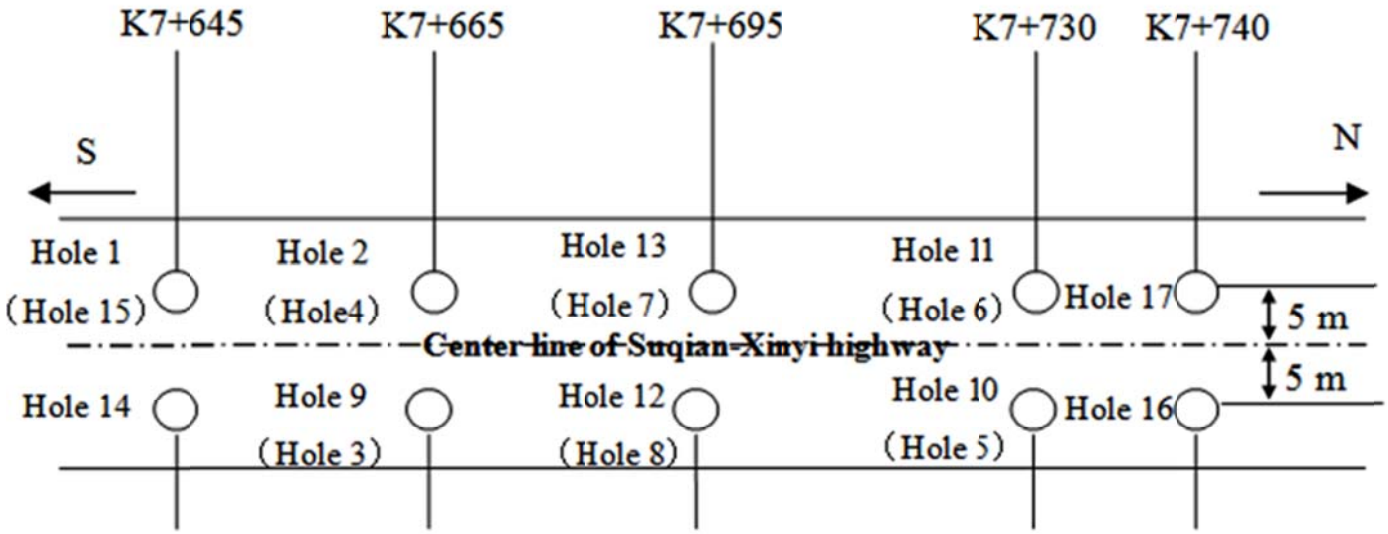
7.3 Recommendations for Future Research

With careful consideration of the limitations mentioned earlier, the framework developed can be extended to the analysis of different parameters in geotechnical engineering. The spatial variability models can be developed for different soil properties for performing prediction analysis. In this study, the predictions were performed based on spatial distance between the observations. However, this analysis framework can be extended to performing spatio-temporal analysis, i.e., the models can be developed with respect to distance and another variable time. The spatio-temporal modeling facilitates understanding of the behavior of properties with distance and time. Also, comprehensive reliability and risk assessment studies can be studied, based on the prediction results obtained from the spatial variability analysis.

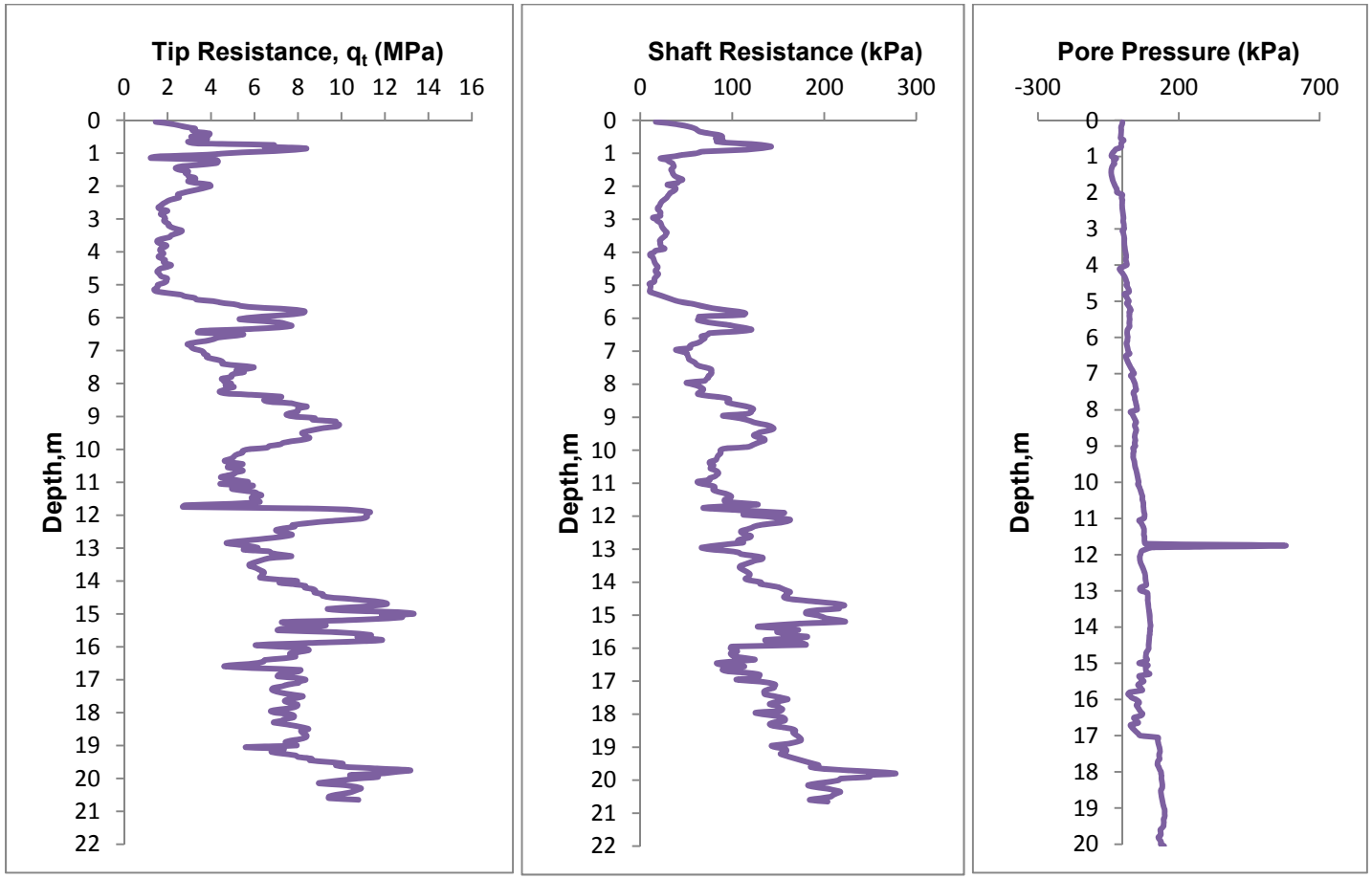
Appendix A

Spatial Variability Analysis of Natural Soil Properties Evaluated from Cone Penetration Test Data (CPTU)

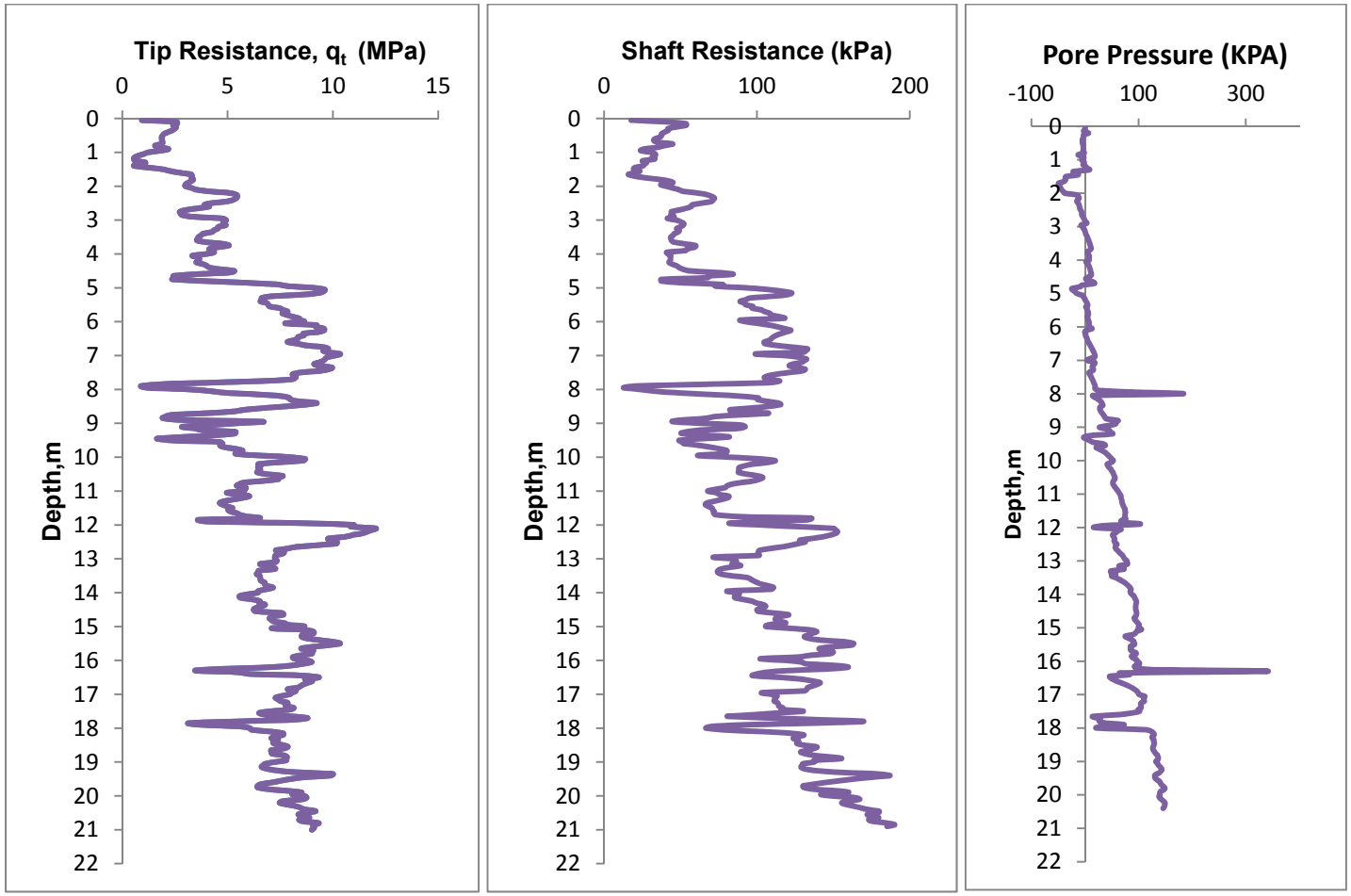
Cone Penetration Test Layout and Profiles for all the Bore Holes from China



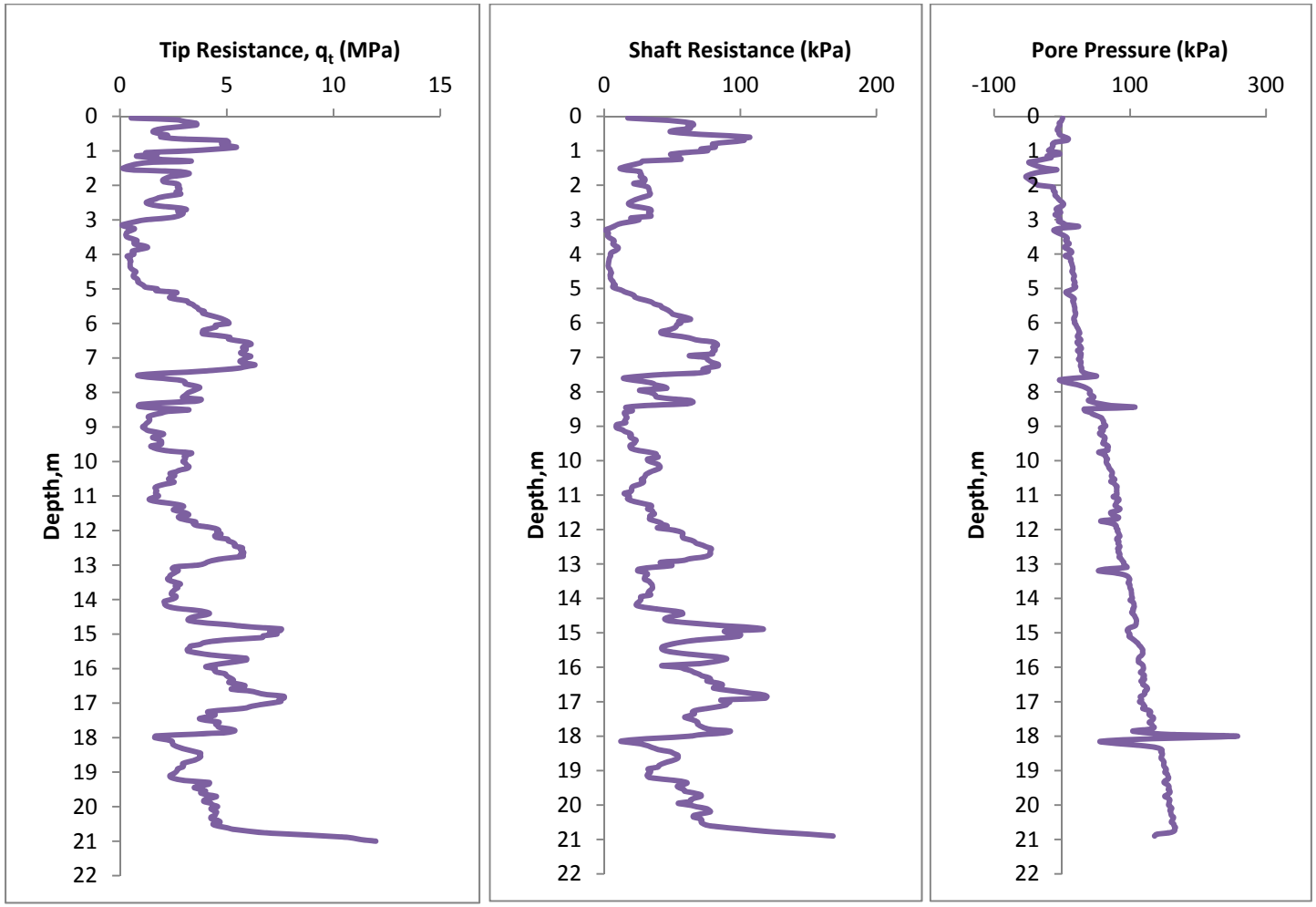
CPTU test parameters in bore hole 1



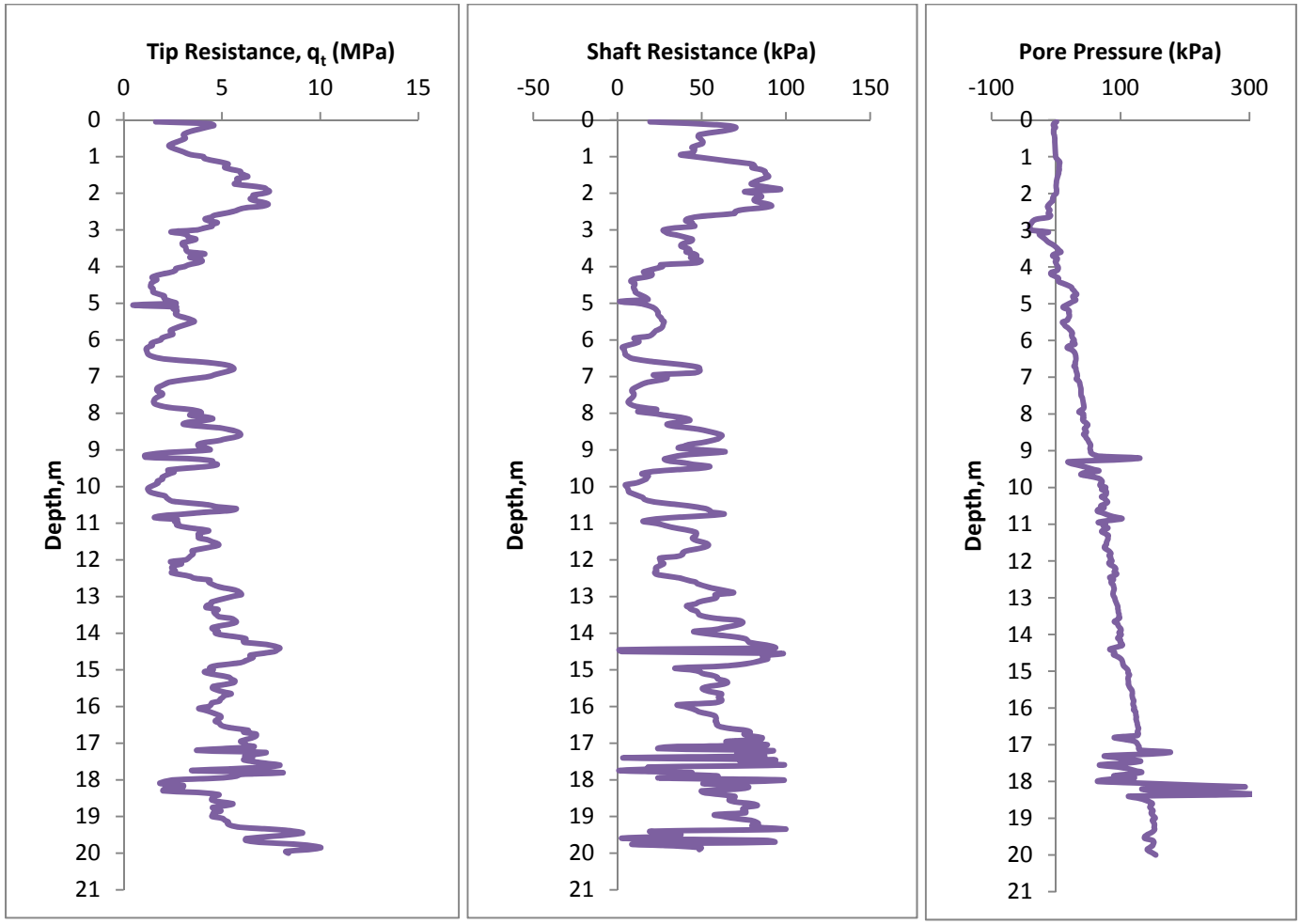
CPTU test parameters in bore hole 2:



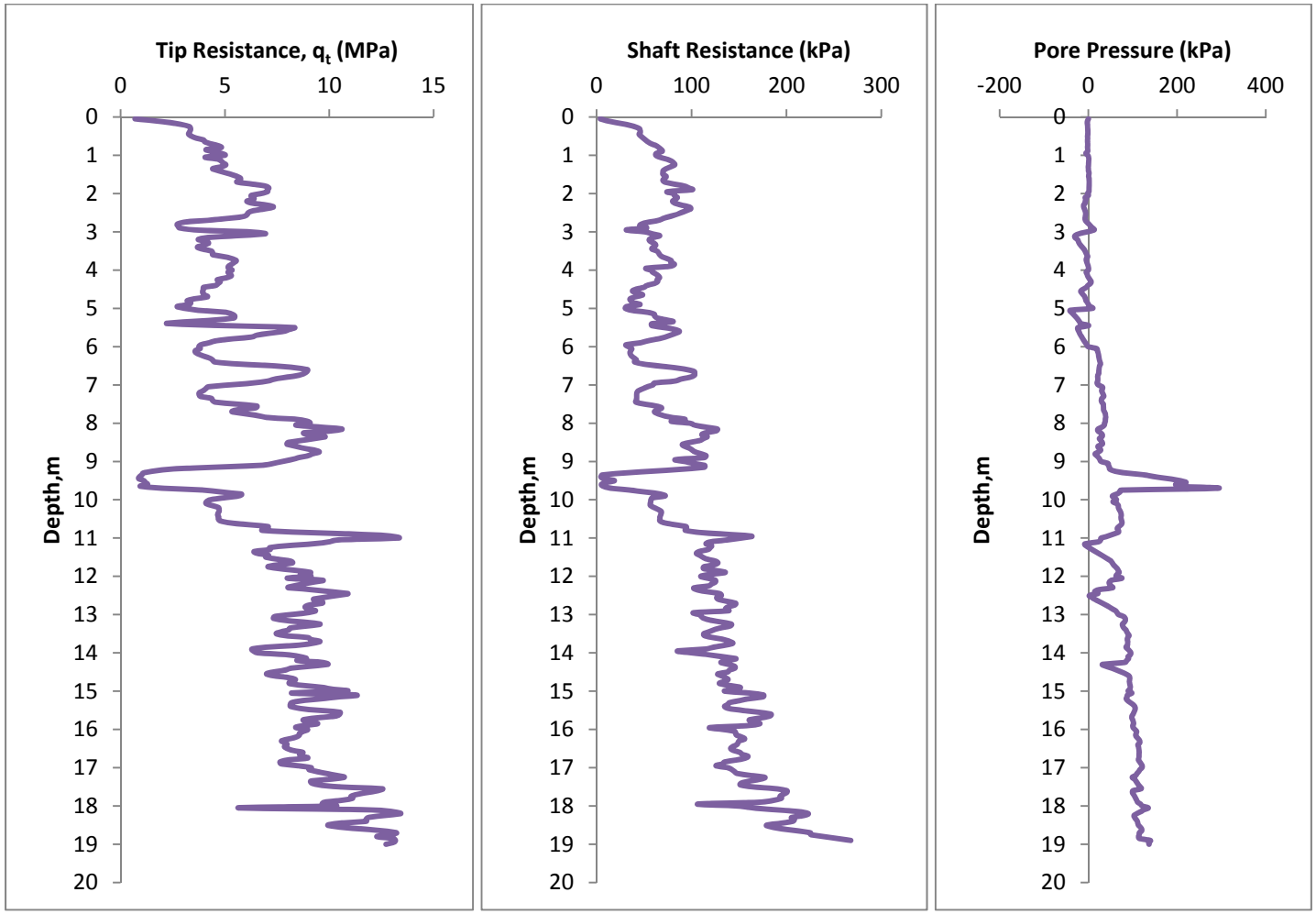
CPTU test parameters in bore hole 9:



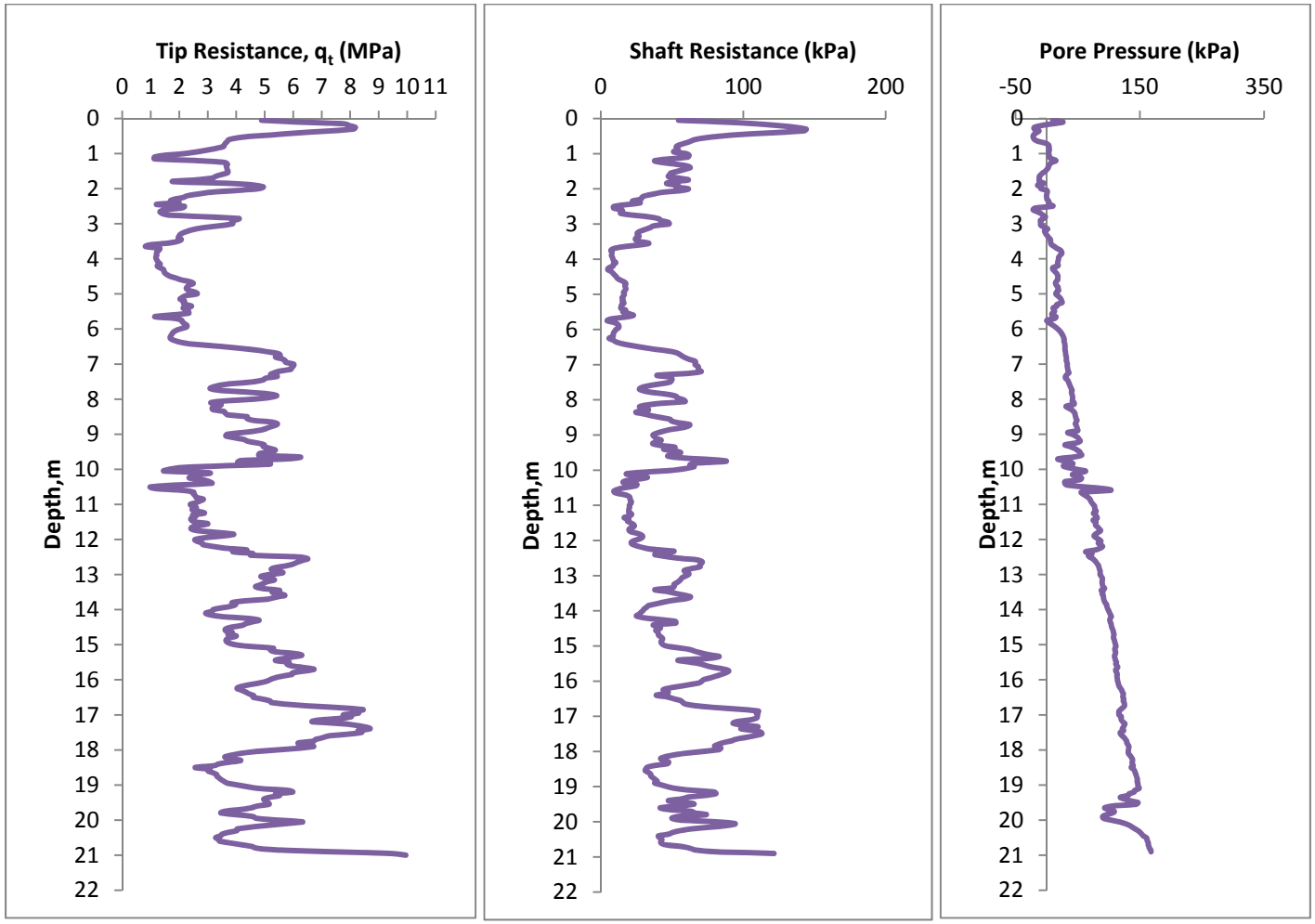
CPTU test parameters in bore hole 10:



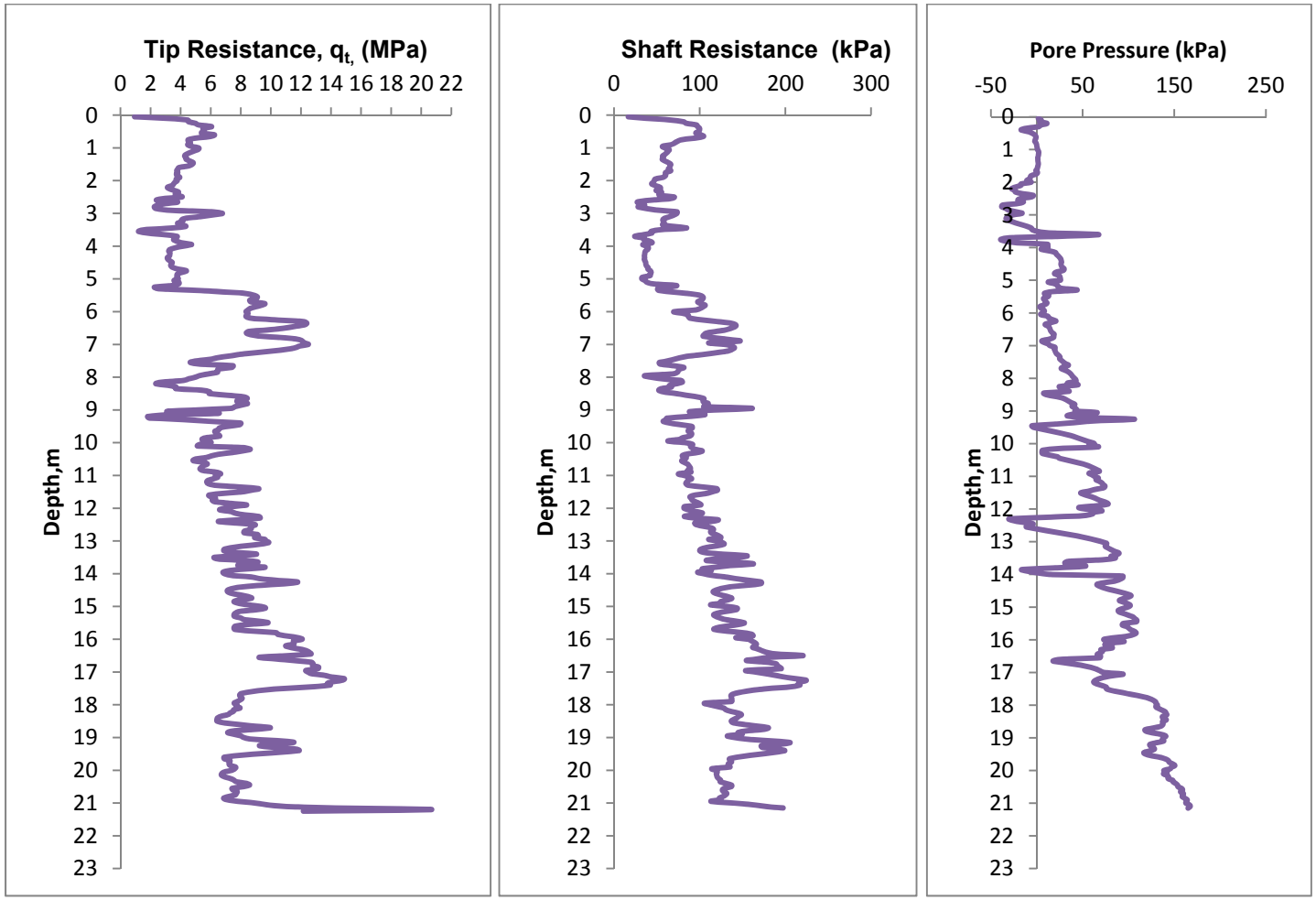
CPTU test parameters in bore hole 11:



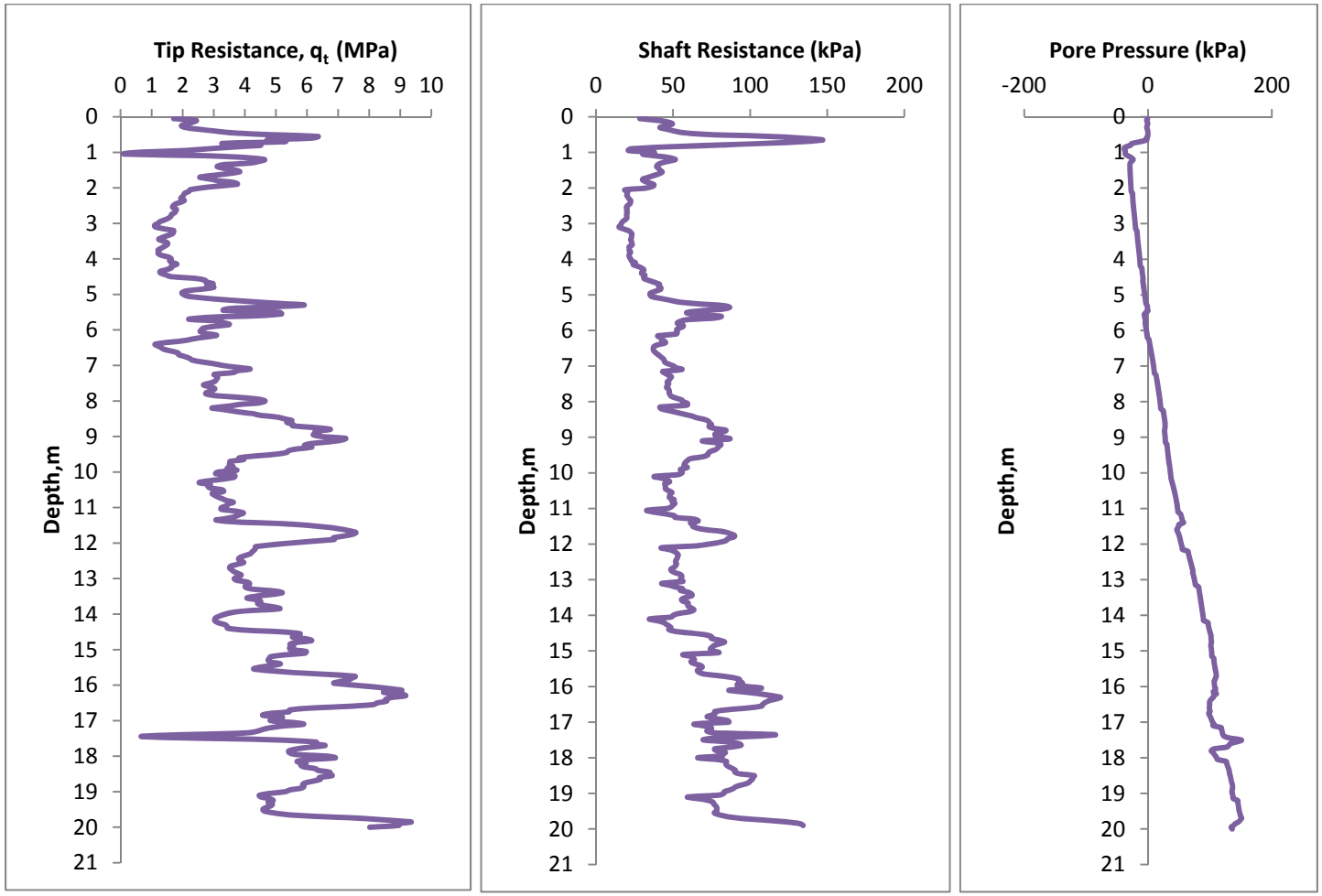
CPTU test parameters in bore hole 12:



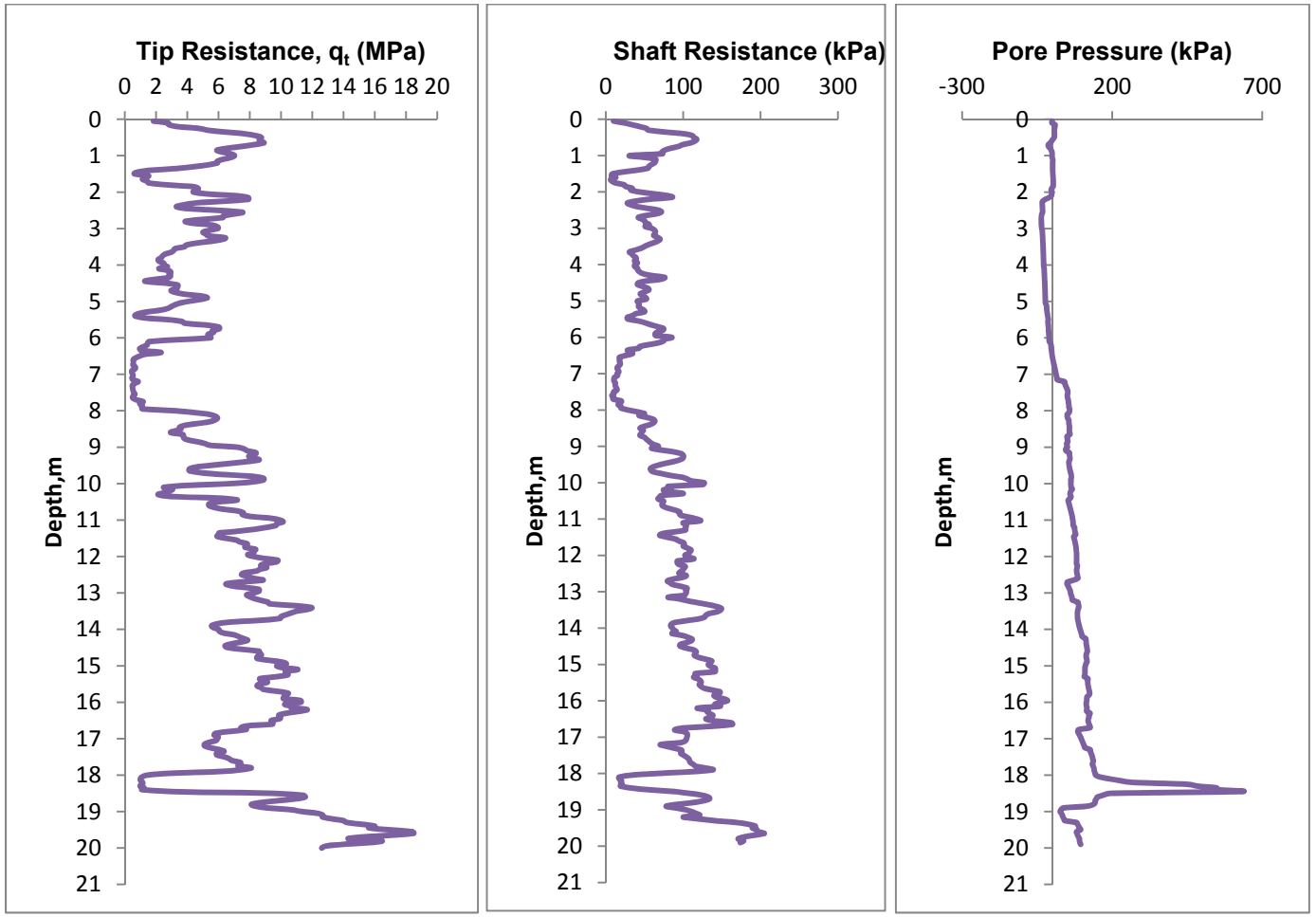
CPTU test parameters in bore hole 13:



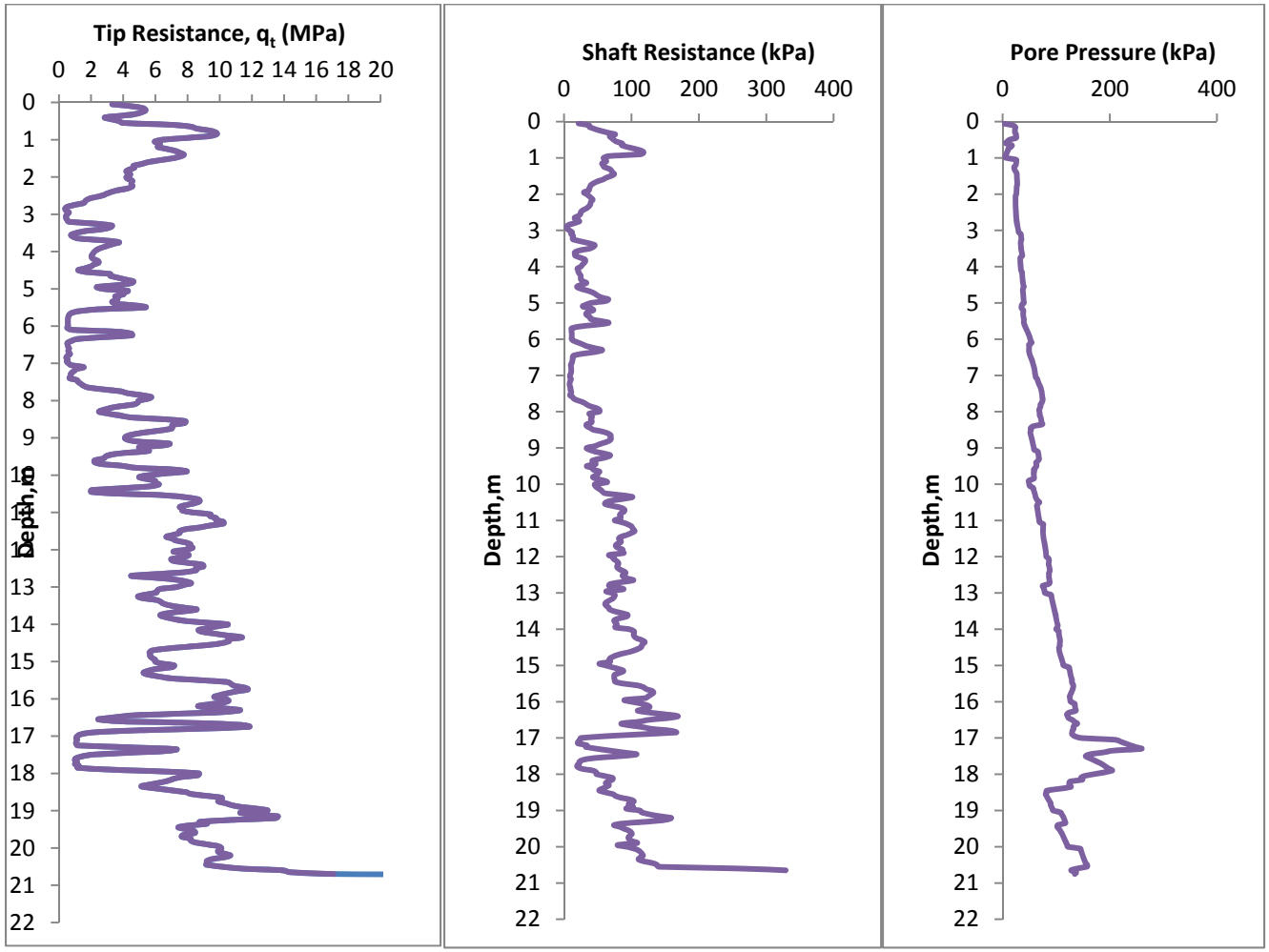
CPTU test parameters in bore hole 14:



CPTU test parameters in bore hole 16:



CPTU test parameters in bore hole 17:



Determination of Friction Angle Values for All CPTU Profiles

CPTU -1

Layer	Soil	Depth (m)		Thickness (m)	Friction Angle (Degrees)
		From	To		
1	SM- MS	0	2.2	2.2	40.7
2	MS-ML	2.2	5.2	3	34.4
3	MS-ML	5.2	8.3	3.1	37.4
4	SP-SM	8.3	10	1.7	39.1
5	SM- MS	10	11.6	1.6	36.6
6	SM- MS	11.6	14.3	2.7	37.5
7	SP-SM	14.3	16	1.7	38.8
8	SM- MS	16	19	3	37.3

CPTU -2

Layer	Soil	Depth (m)		Thickness (m)	Friction Angle (Degrees
		From	To		
1	SM-MS	0	1	1	36.7
2	Clay	1	1.4	0.4	34.9
3	SM-MS	1.4	4.6	3.2	38.6
4	SP-SM	4.6	7.6	3	39.5
5	MS-ML	7.6	9.6	2	35.2
6	SP-SM	9.6	11.8	2.2	37.6
7	SP-SM	11.8	13	1.2	39.3
8	SP-SM	13	21	8	37.3

CPTU -13

Layer	Soil	Depth (m)		Thickness (m)	Friction Angle (Degrees)
		From	To		
1	SM-MS	0	5.3	5.3	38.7
2	SP-SM	5.3	7.3	2	40.8
3	SM-MS	7.3	9.4	2.1	37.4
4	SM-MS	9.4	11.7	2.3	37.6
5	SM-MS	11.7	15.7	4	35.8
6	SP-SM	15.7	17.8	2.1	39.3
7	SM-MS	17.8	19.5	1.7	37.6
8	SM-MS	19.5	21	1.5	38.5

CPTU – 11

Layer	Soil	Depth (m)		Thickness (m)	Friction Angle (Degrees)
		From	To		
1	SM-MS	0	5.2	5.2	39.3
2	SM-MS	5.2	7.4	2.2	38
3	SP-SM	7.4	9.2	1.8	39.2
4	Clay	9.2	9.7	1.3	29.1
5	SM-MS	9.7	20	10.3	37.7

CPTU – 17

Layer	Soil	Depth (m)		Thickness (m)	Friction Angle (Degrees)
1	SM-MS	From	To	2.5	41.6
2	ML	2.5	4.5	2	33.9
3	SM-MS	4.5	5.5	1	37.0
4	Clay	5.5	7.6	2.1	27.7
5	SM-MS	7.6	10.5	2.9	36.7
6	SP-SM	10.5	16.4	5.9	38.2
7	Clay	16.4	17.9	1.5	31.
8	SP-SM	17.9	19.2	1.3	38.1

Assigning the Coordinates to CPTU Layers for Predicting the Friction Angle

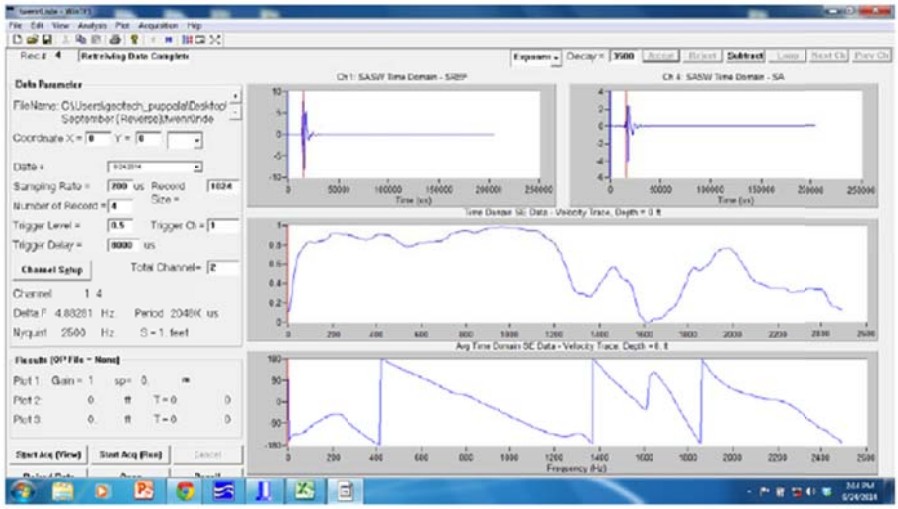
Values along the Profile:

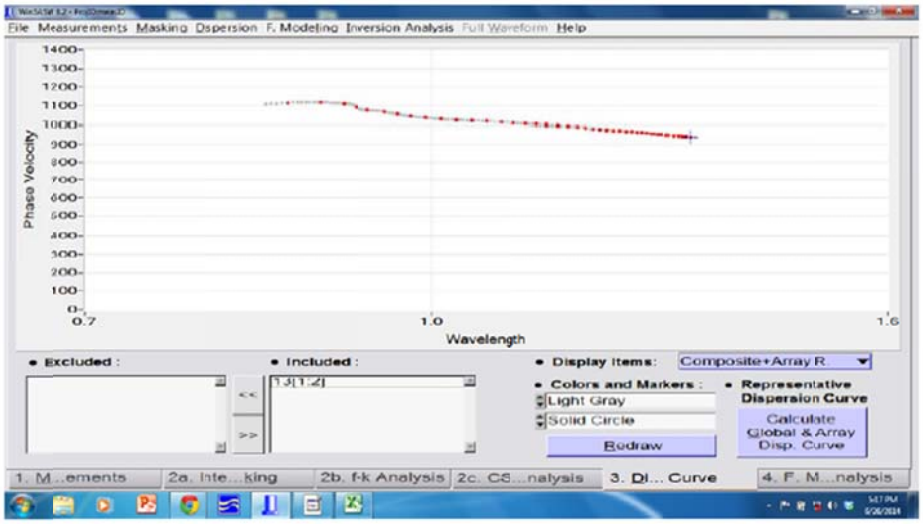
CPTU 1			CPTU 2			CPTU 13			CPTU 11			CPTU 17		
5	0	39	25	0	37	55	0	39	90	0	39	100	0	41
5	2.2	39	25	1	37	55	5.3	39	90	5.2	38	100	2.5	41
5	5.2	34	25	1.4	35	55	7.3	41	90	7.4	39	100	4.5	34
5	8.3	37	25	4.6	39	55	9.4	37	90	9.2	29	100	5.5	37
5	10	30	25	7.6	39	55	11.7	37	90	9.7	37	100	7.6	28
5	11.6	36	25	9.6	35	55	15.7	35				100	10.5	37
5	14.3	37	25	11.8	37	55	17.8	39				100	16.4	38
5	16	38	25	13	39	55	19.5	37				100	17.9	31
5	19	37	25	21	37	55	21	38				100	19.2	37

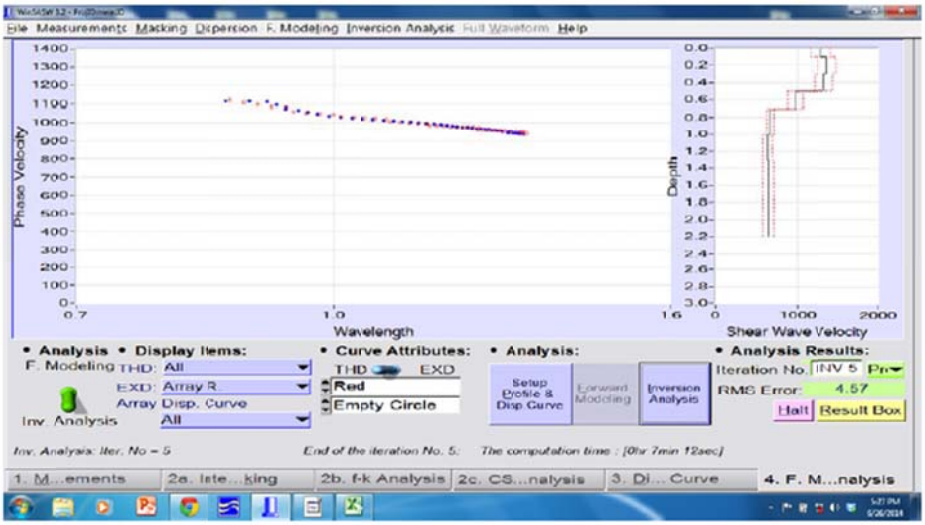
Appendix B

Spatial Variability Analysis of Man-Made Treated Soils

SASW Wave Propagation Signal:






File Measurements Masking Dispersion F. Modeling Inversion Analysis Full Waveform Help
Phase Velocity
Wavelength
Depth
Shear Wave Velocity
• Analysis • Display Items:
F. Modeling
THD: All
EXD: Array R
Array Disp. Curve
Inv. Analysis
All
• Curve Attributes:
THD EXD
Red
Empty Circle
• Analysis:
Setup Profile & Disp Curve
Forward Modeling
Inversion Analysis
• Analysis Results:
Iteration No. INV 5
RMS Error: 4.57
Halt
Result Box
Inv. Analysis: Iter. No = 5
End of the iteration No. 5:
The computation time : [0hr 7min 12sec]
1. M...ements
2a. Ste...king
2b. f-k Analysis
2c. CS...nalysis
3. Di... Curve
4. F. M...nalysis

Coordinates for Spatial Variability Analysis

Day 1 Stiffness Measurements:

Section	Test point	X	Y	Z		Section	Test point	X	Y	Z
1066	A	25	20.62			1071-10	A	260	20.62	
	B	25	22.22	264.4			B	260	22.22	301.9
	C	25	25	238.1			C	260	25	251.8
	D	25	27.78	281.1			D	260	27.78	274.6
	E	25	29.38				E	260	29.38	
1067-10	A	60	20.62			1071-40	A	290	20.62	
	B	60	22.22	330.3			B	290	22.22	265.1
	C	60	25	254.4			C	290	25	247.8
	D	60	27.78	261.0			D	290	27.78	273.2
	E	60	29.38				E	290	29.38	
1067-40	A	90	20.62			1072-25	A	325	20.62	
	B	90	22.22	245.3			B	325	22.22	217.7
	C	90	25	308.7			C	325	25	248.6
	D	90	27.78	270.7			D	325	27.78	273.4
	E	90	29.38				E	325	29.38	
1068-25	A	125	20.62			1073-10	A	360	20.62	
	B	125	22.22	250.9			B	360	22.22	280.2
	C	125	25	252.9			C	360	25	254.3
	D	125	27.78				D	360	27.78	244.4
	E	125	29.38				E	360	29.38	
1069-10	A	160	20.62			1073-40	A	390	20.62	
	B	160	22.22	250.9			B	390	22.22	270.4
	C	160	25	259.5			C	390	25	255.0
	D	160	27.78	297.5			D	390	27.78	319.0
	E	160	29.38				E	390	29.38	
1069-40	A	190	20.62			1074-10	A	410	20.62	
	B	190	22.22	252.8			B	410	22.22	365.2
	C	190	25	256.6			C	410	25	247.9
	D	190	27.78	244.5			D	410	27.78	260.1
	E	190	29.38				E	410	29.38	
1070-10	A	210	20.62			1075-25	A	475	20.62	
	B	210	22.22	261.1			B	475	22.22	312.2
	C	210	25	258.9			C	475	25	251.5
	D	210	27.78	294.7			D	475	27.78	292.7
	E	210	29.38				E	475	29.38	
1070-25	A	225	20.62							
	B	225	22.22	202.2						
	C	225	25	251.0						
	D	225	27.78	275.4						
	E	225	29.38							

Day 3 Stiffness Measurements

Section	Test point	X	Y	Z
1066	A	25	20.62	276.5
	B	25	22.22	275.1
	C	25	25	272.7
	D	25	27.78	289.1
	E	25	29.38	271.9
1067-10	A	60	20.62	271.4
	B	60	22.22	303.3
	C	60	25	266.8
	D	60	27.78	261.0
	E	60	29.38	256.7
1067-40	A	90	20.62	285.8
	B	90	22.22	274.3
	C	90	25	309
	D	90	27.78	346.0
	E	90	29.38	274.6
1068-25	A	125	20.62	251.0
	B	125	22.22	257.6
	C	125	25	282.3
	D	125	27.78	274.7
	E	125	29.38	252.6
1069-10	A	160	20.62	261.8
	B	160	22.22	253.8
	C	160	25	330.6
	D	160	27.78	297.5
	E	160	29.38	232.9
1069-40	A	190	20.62	307.8
	B	190	22.22	291.3
	C	190	25	293.4
	D	190	27.78	260.1
	E	190	29.38	273.3
1070-10	A	210	20.62	293.9
	B	210	22.22	261.1
	C	210	25	261.2
	D	210	27.78	294.7
	E	210	29.38	276.2
1070-25	A	225	20.62	276.1
	B	225	22.22	268.7
	C	225	25	275.4
	D	225	27.78	275.0
	E	225	29.38	264.5

Section	Test point	X	Y	Z
1071-10	A	260	20.62	243.8
	B	260	22.22	273.3
	C	260	25	255.8
	D	260	27.78	304.6
	E	260	29.38	256.3
1071-40	A	290	20.62	243.8
	B	290	22.22	273.3
	C	290	25	255.8
	D	290	27.78	304.6
	E	290	29.38	256.3
1072-25	A	325	20.62	275.8
	B	325	22.22	260.0
	C	325	25	293.7
	D	325	27.78	276.6
	E	325	29.38	255.4
1073-10	A	360	20.62	217.4
	B	360	22.22	290.8
	C	360	25	285.9
	D	360	27.78	270.4
	E	360	29.38	298.0
1073-40	A	390	20.62	258.5
	B	390	22.22	277.5
	C	390	25	280.0
	D	390	27.78	300.7
	E	390	29.38	269.0
1074-10	A	410	20.62	310.9
	B	410	22.22	287.1
	C	410	25	279.3
	D	410	27.78	275.0
	E	410	29.38	264.4
1075-25	A	475	20.62	258.5
	B	475	22.22	277.5
	C	475	25	280.0
	D	475	27.78	300.7
	E	475	29.38	269.0

Day 7 Stiffness measurements

Section	Test point	X	Y	Z
1066	A	25	20.62	327.0
	B	25	22.22	320.3
	C	25	25	324.2
	D	25	27.78	329.8
	E	25	29.38	319.8
1067-10	A	60	20.62	321.9
	B	60	22.22	359.8
	C	60	25	318.3
	D	60	27.78	281.2
	E	60	29.38	304.6
1067-40	A	90	20.62	336.3
	B	90	22.22	319.5
	C	90	25	329.4
	D	90	27.78	386.7
	E	90	29.38	322.5
1068-25	A	125	20.62	301.4
	B	125	22.22	302.8
	C	125	25	333.9
	D	125	27.78	295.4
	E	125	29.38	300.5
1069-10	A	160	20.62	312.3
	B	160	22.22	299.0
	C	160	25	382.1
	D	160	27.78	345.8
	E	160	29.38	280.8
1069-40	A	190	20.62	358.2
	B	190	22.22	336.5
	C	190	25	344.9
	D	190	27.78	300.8
	E	190	29.38	321.1
1070-10	A	210	20.62	344.4
	B	210	22.22	301.6
	C	210	25	312.8
	D	210	27.78	317.5
	E	210	29.38	324.1
1070-25	A	225	20.62	326.6
	B	225	22.22	313.9
	C	225	25	327.0
	D	225	27.78	335.7
	E	225	29.38	312.4

Section	Test point	X	Y	Z
1071-10	A	260	20.62	294.3
	B	260	22.22	318.5
	C	260	25	307.4
	D	260	27.78	345.3
	E	260	29.38	304.2
1071-40	A	290	20.62	294.3
	B	290	22.22	318.5
	C	290	25	307.4
	D	290	27.78	345.3
	E	290	29.38	304.2
1072-25	A	325	20.62	326.3
	B	325	22.22	305.2
	C	325	25	335.2
	D	325	27.78	317.3
	E	325	29.38	303.2
1073-10	A	360	20.62	267.9
	B	360	22.22	316.0
	C	360	25	307.5
	D	360	27.78	291.1
	E	360	29.38	345.9
1073-40	A	390	20.62	309.0
	B	390	22.22	312.7
	C	390	25	331.6
	D	390	27.78	341.4
	E	390	29.38	316.9
1074-10	A	410	20.62	361.3
	B	410	22.22	332.3
	C	410	25	310.8
	D	410	27.78	315.7
	E	410	29.38	312.2
1075-25	A	475	20.62	309.0
	B	475	22.22	312.7
	C	475	25	331.6
	D	475	27.78	341.4
	E	475	29.38	316.9

Day 14 Stiffness Measurements

Section	Test point	X	Y	Z
1066	A	25	20.62	398.4
	B	25	22.22	388.0
	C	25	25	398.8
	D	25	27.78	386.5
	E	25	29.38	391.1
1067-10	A	60	20.62	370.0
	B	60	22.22	398.3
	C	60	25	373.7
	D	60	27.78	397.9
	E	60	29.38	341.4
1067-40	A	90	20.62	395.3
	B	90	22.22	408.3
	C	90	25	371.1
	D	90	27.78	410.1
	E	90	29.38	389.1
1068-25	A	125	20.62	372.1
	B	125	22.22	367.1
	C	125	25	382.9
	D	125	27.78	347.5
	E	125	29.38	390.5
1069-10	A	160	20.62	387.7
	B	160	22.22	389.3
	C	160	25	418.7
	D	160	27.78	393.7
	E	160	29.38	423.1
1069-40	A	190	20.62	397.0
	B	190	22.22	415.5
	C	190	25	400.8
	D	190	27.78	388.9
	E	190	29.38	392.8
1070-10	A	210	20.62	390.5
	B	210	22.22	393.6
	C	210	25	399.1
	D	210	27.78	397.1
	E	210	29.38	407.3
1070-25	A	225	20.62	361.9
	B	225	22.22	356.0
	C	225	25	407.3
	D	225	27.78	380.7
	E	225	29.38	386.0

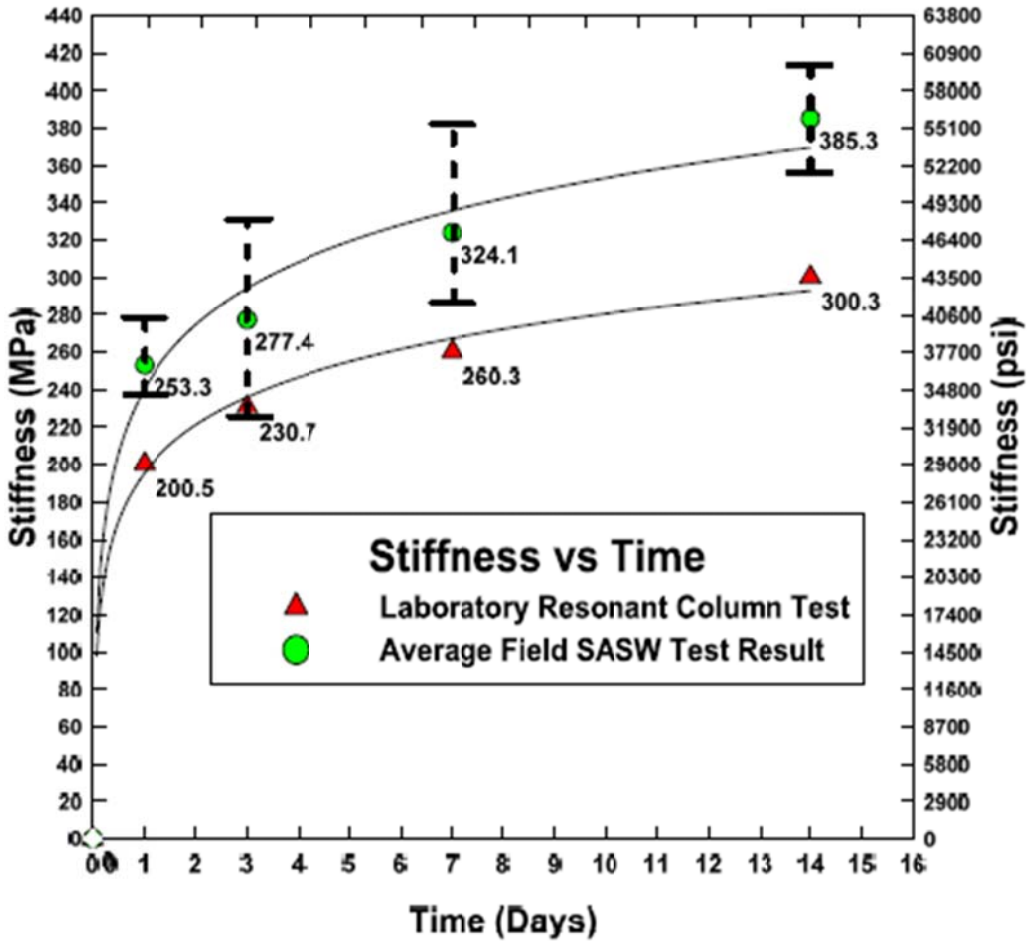
Section	Test point	X	Y	Z
1071-10	A	260	20.62	359.4
	B	260	22.22	366.9
	C	260	25	400.0
	D	260	27.78	396.8
	E	260	29.38	428.0
1071-40	A	290	20.62	359.4
	B	290	22.22	366.9
	C	290	25	380.0
	D	290	27.78	366.8
	E	290	29.38	378.0
1072-25	A	325	20.62	377.5
	B	325	22.22	394.8
	C	325	25	385.3
	D	325	27.78	379.6
	E	325	29.38	415.2
1073-10	A	360	20.62	368.5
	B	360	22.22	383.9
	C	360	25	357.3
	D	360	27.78	399.1
	E	360	29.38	386.9
1073-40	A	390	20.62	383.8
	B	390	22.22	381.4
	C	390	25	399.4
	D	390	27.78	381.4
	E	390	29.38	406.8
1074-10	A	410	20.62	396.2
	B	410	22.22	399.6
	C	410	25	362.5
	D	410	27.78	360.0
	E	410	29.38	404.4
1075-25	A	475	20.62	383.8
	B	475	22.22	371.4
	C	475	25	369.4
	D	475	27.78	401.4
	E	475	29.38	416.8

Day 28 Stiffness Measurements

Section	Test point	X	Y	Z
1066	A	25	20.62	432.3
	B	25	22.22	465.9
	C	25	25	458.7
	D	25	27.78	447.5
	E	25	29.38	432.4
1067-10	A	60	20.62	417.7
	B	60	22.22	463.1
	C	60	25	414.1
	D	60	27.78	460.1
	E	60	29.38	408.6
1067-40	A	90	20.62	473.7
	B	90	22.22	455.9
	C	90	25	436.8
	D	90	27.78	450.7
	E	90	29.38	470.1
1068-25	A	125	20.62	455.5
	B	125	22.22	407.6
	C	125	25	437.3
	D	125	27.78	410.2
	E	125	29.38	445.9
1069-10	A	160	20.62	442.2
	B	160	22.22	421.0
	C	160	25	440.6
	D	160	27.78	417.5
	E	160	29.38	440.1
1069-40	A	190	20.62	420.0
	B	190	22.22	455.9
	C	190	25	443.0
	D	190	27.78	454.3
	E	190	29.38	427.6
1070-10	A	210	20.62	434.7
	B	210	22.22	444.8
	C	210	25	452.4
	D	210	27.78	439.2
	E	210	29.38	434.2
1070-25	A	225	20.62	446.5
	B	225	22.22	471.5
	C	225	25	453.1
	D	225	27.78	460.1
	E	225	29.38	441.7

Section	Test point	X	Y	Z
1071-10	A	260	20.62	424.6
	B	260	22.22	414.2
	C	260	25	453.2
	D	260	27.78	440.8
	E	260	29.38	420.9
1071-40	A	290	20.62	424.6
	B	290	22.22	414.2
	C	290	25	413.2
	D	290	27.78	440.8
	E	290	29.38	420.9
1072-25	A	325	20.62	461.9
	B	325	22.22	457.7
	C	325	25	419.8
	D	325	27.78	432.7
	E	325	29.38	461.5
1073-10	A	360	20.62	419.4
	B	360	22.22	438.5
	C	360	25	428.8
	D	360	27.78	429.9
	E	360	29.38	418.2
1073-40	A	390	20.62	456.4
	B	390	22.22	458.7
	C	390	25	442.5
	D	390	27.78	470.5
	E	390	29.38	465.2
1074-10	A	410	20.62	449.5
	B	410	22.22	478.2
	C	410	25	420.2
	D	410	27.78	468.1
	E	410	29.38	449.9
1075-25	A	475	20.62	456.4
	B	475	22.22	458.7
	C	475	25	442.5
	D	475	27.78	470.5
	E	475	29.38	465.2

Comparision of Stiffness (G_{max}) from in situ SASW test and Laboratory Resonant Colum Test conducted on CLSM:



Appendix C

Spatial Variability Analysis of Sulfate-Rich Natural Soils

Coordinates at 5 ft. depth

Easting	Northing	Sulfate Concentration (ppm)	Boring	Formation
-96.4172	32.21603	180	B-007	Kemp
-96.3431	32.21148	1880	B-025	
-96.4923	32.21794	1270	B-068	
-96.3535	32.21152	11600	B-069	
-96.346	32.21168	2120	B-070	
-96.4141	32.21516	260	B-094	
-96.41	32.21472	115	B-095	
-96.3998	32.21385	1065	B-096	
-96.3622	32.21141	1400	B-097	
-96.3994	32.21362	3215	B-180	
-96.468	32.21516	975	B-026	Neylandville
-96.4392	32.2167	535	B-027	
-96.4431	32.21675	90	B-174	
-96.4633	32.21772	460	B-194	
-96.1994	32.1866	260	B-005	Wills
-96.2537	32.12988	105	B-017	
-96.2283	32.0901	55	B-018	
-96.1895	32.19346	650	B-020	
-96.1785	32.18559	90	B-021	
-96.2477	32.20023	45	B-037	
-96.2385	32.19983	350	B-060	
-96.2425	32.10732	315	B-145	
-96.2469	32.11703	750	B-146	
-96.2475	32.1186	240	B-147	
-96.2531	32.12788	300	B-148	
-96.2586	32.13917	280	B-150	
-96.2623	32.14661	135	B-151	
-96.2686	32.16087	200	B-152	
-96.2726	32.1664	415	B-153	
-96.5192	32.21961	1200	B-008	Wolfe
-96.6332	32.23484	50	B-029	
-96.6237	32.23276	310	B-030	

-96.6404	32.23674	20	B-031	
-96.645	32.23934	125	B-032	
-96.5927	32.22919	670	B-034	
-96.5853	32.22828	1065	B-035	
-96.5817	32.22782	25	B-036	
-96.5101	32.2186	20	B-065	
-96.5052	32.21919	65	B-066	
-96.5021	32.21925	80	B-067	
-96.5423	32.22429	390	B-089	
-96.6781	32.25739	1130	B-009	Ozan
-96.7063	32.26723	16,000	B-033	
-96.813	32.30978	280	B-057	
-96.8152	32.31101	290	B-082	
-96.7678	32.29217	17,400	B-118	
-96.7334	32.27537	16,760	B-120	
-96.7258	32.2735	950	B-122	
-96.844	32.33065	115	B-224	
-96.7729	32.29347	17,110	B-116	
-96.7338	32.27541	16,700	B-119	
-96.6996	32.26382	50	B-124	
-96.7468	32.28229	1100	B-227	
-96.8077	32.30712	570	B-100	
-96.7	32.26406	17,300	B-123	
-96.6912	32.25882	1,015	B-126	
-97.0791	32.50168	540	B-051	Eagle Ford
-97.0406	32.42314	185	B-059	
-97.0458	32.44405	130	B-061	
-97.0488	32.44807	8080	B-062	
-97.0214	32.40384	750	B-063	
-97.05	32.44884	18,450	B-090	
-97.0595	32.45993	19,620	B-093	
-97.1362	32.51262	15,260	B-164	
-97.0768	32.51596	9360	B-168	
-97.1525	32.53552	4200	B-84	
-97.074	32.51824	7000	B-198	
-97.0768	32.51327	5530	B-201	
-97.0768	32.50699	14,050	B-202	

-97.0685	32.53477	75	B-196	
-97.075	32.51746	2200	B-199	
-97.0754	32.51716	3340	B-200	
-97.035	32.41299	16,000	B-209	
-97.1482	32.529	1320	B-318	
-97.0972	32.5055	1005	B-326	

Coordinates at 10 ft. depth

Easting	Northing	Sulfate Concentration (ppm)	Boring	Formation
-96.4172	32.21603	280	B-007	Kemp
-96.3431	32.21148	17000	B-025	
-96.4923	32.21794	255	B-068	
-96.3535	32.21152	16100	B-069	
-96.346	32.21168	8420	B-070	
-96.4141	32.21516	255	B-094	
-96.41	32.21472	300	B-095	
-96.3998	32.21385	690	B-096	
-96.3622	32.21141	840	B-097	
-96.3994	32.21362	1070	B-180	
-96.468	32.21516	1500	B-026	Neylandville
-96.4392	32.2167	690	B-027	
-96.4431	32.21675	135	B-174	
-96.4633	32.21772	180	B-194	
-96.1994	32.1866	150	B-005	Wills
-96.2537	32.12988	55	B-017	
-96.2283	32.0901	390	B-018	
-96.1895	32.19346	670	B-020	
-96.1785	32.18559	325	B-021	
-96.2477	32.20023	20	B-037	
-96.2385	32.19983	300	B-060	
-96.2425	32.10732	380	B-145	
-96.2469	32.11703	230	B-146	
-96.2475	32.1186	100	B-147	

-96.2531	32.12788	150	B-148	
-96.2586	32.13917	40	B-150	
-96.2623	32.14661	200	B-151	
-96.2686	32.16087	10	B-152	
-96.2726	32.1664	90	B-153	
-96.5192	32.21961	750	B-008	Wolfe
-96.6332	32.23484	785	B-029	
-96.6237	32.23276	800	B-030	
-96.6404	32.23674	1050	B-031	
-96.645	32.23934	200	B-032	
-96.5927	32.22919	100	B-034	
-96.5853	32.22828	10600	B-035	
-96.5817	32.22782	115	B-036	
-96.5101	32.2186	120	B-065	
-96.5052	32.21919	45	B-066	
-96.5021	32.21925	115	B-067	
-96.5423	32.22429	310	B-089	
-96.6781	32.25739	2820	B-009	Ozan
-96.7063	32.26723	1055	B-033	
-96.813	32.30978	275	B-057	
-96.8152	32.31101	200	B-082	
-96.7678	32.29217	1235	B-118	
-96.7334	32.27537	1435	B-120	
-96.7258	32.2735	755	B-122	
-96.844	32.33065	110	B-224	
-96.7729	32.29347	2345	B-116	
-96.7338	32.27541	15,620	B-119	
-96.6996	32.26382	1270	B-124	
-96.7468	32.28229	1050	B-227	
-96.8077	32.30712	415	B-100	
-96.7	32.26406	1750	B-123	
-96.6912	32.25882	25	B-126	
-97.0795	32.50392	500	B-013	Eagle Ford
-97.1349	32.53871	1320	B-015	
-97.0791	32.50168	525	B-051	
-97.0406	32.42314	185	B-059	
-97.0458	32.44405	200	B-061	

-97.0488	32.44807	2250	B-062	
-97.0214	32.40384	18,300	B-063	
-97.05	32.44884	1620	B-090	
-97.0595	32.45993	3620	B-093	
-97.1362	32.51262	560	B-164	
-97.0768	32.51596	890	B-168	
-97.1525	32.53552	1050	B-84	
-97.074	32.51824	15,100	B-198	
-97.0768	32.51327	15,200	B-201	
-97.0768	32.50699	6300	B-202	
-97.0685	32.53477	370	B-196	
-97.075	32.51746	180	B-199	
-97.0754	32.51716	105	B-200	
-97.035	32.41299	11,725	B-209	
-97.1482	32.529	280	B-318	
-97.0972	32.5055	1090	B-326	

Coordinates at 15 ft. depth

Easting	Northing	Sulfate Concentration (ppm)	Boring	Formation
-96.4172	32.21603	325	B-007	Kemp
-96.3431	32.21148	1100	B-025	
-96.4923	32.21794	215	B-068	
-96.3535	32.21152	1360	B-069	
-96.346	32.21168	10000	B-070	
-96.4141	32.21516	90	B-094	
-96.41	32.21472	220	B-095	
-96.3998	32.21385	700	B-096	
-96.3622	32.21141	850	B-097	
-96.3994	32.21362	970	B-180	
-96.468	32.21516	605	B-026	Neylandville
-96.4392	32.2167	790	B-027	
-96.4431	32.21675	110	B-174	

-96.4633	32.21772	210	B-194	
-96.1994	32.1866	200	B-005	Wills
-96.2537	32.12988	30	B-017	
-96.2283	32.0901	185	B-018	
-96.1895	32.19346	575	B-020	
-96.1785	32.18559	365	B-021	
-96.2477	32.20023	30	B-037	
-96.2385	32.19983	650	B-060	
-96.2425	32.10732	220	B-145	
-96.2469	32.11703	20	B-146	
-96.2475	32.1186	270	B-147	
-96.2531	32.12788	25	B-148	
-96.2586	32.13917	60	B-150	
-96.2623	32.14661	70	B-151	
-96.2686	32.16087	115	B-152	
-96.2726	32.1664	70	B-153	
-96.1351	32.18983	35	P-002	
-96.5192	32.21961	90	B-008	Wolfe
-96.6332	32.23484	700	B-029	
-96.6237	32.23276	700	B-030	
-96.6404	32.23674	1000	B-031	
-96.5927	32.22919	480	B-034	
-96.5853	32.22828	1255	B-035	
-96.5817	32.22782	135	B-036	
-96.5101	32.2186	45	B-065	
-96.5052	32.21919	85	B-066	
-96.5021	32.21925	105	B-067	
-96.5423	32.22429	470	B-089	
-96.6781	32.25739	1500	B-009	Ozan
-96.7063	32.26723	1100	B-033	
-96.813	32.30978	275	B-057	
-96.8152	32.31101	205	B-082	
-96.7678	32.29217	1255	B-118	
-96.7334	32.27537	1315	B-120	
-96.7258	32.2735	690	B-122	
-96.844	32.33065	100	B-224	

-96.7729	32.29347	1510	B-116	
-96.7338	32.27541	15,940	B-119	
-96.6996	32.26382	1370	B-124	
-96.7468	32.28229	975	B-227	
-96.8077	32.30712	620	B-100	
-96.7	32.26406	11,310	B-123	
-96.6912	32.25882	20	B-126	
-97.1349	32.53871	1270	B-015	Eagle Ford
-97.0406	32.42314	300	B-059	
-97.0458	32.44405	140	B-061	
-97.0488	32.44807	1915	B-062	
-97.0214	32.40384	715	B-063	
-97.05	32.44884	1160	B-090	
-97.0595	32.45993	16,160	B-093	
-97.1362	32.51262	640	B-164	
-97.0768	32.51596	15,670	B-168	
-97.1525	32.53552	3800	B-84	
-97.074	32.51824	17,500	B-198	
-97.0768	32.51327	1560	B-201	
-97.0768	32.50699	530	B-202	
-97.0685	32.53477	60	B-196	
-97.075	32.51746	1030	B-199	
-97.0754	32.51716	275	B-200	
-97.035	32.41299	17,100	B-209	
-97.1482	32.529	290	B-318	
-97.0972	32.5055	1095	B-326	

Coordinates at 20 ft. depth

Easting	Northing	Sulfate Concentration (ppm)	Boring	Formation
-96.4172	32.21603	100	B-007	Kemp
-96.3431	32.21148	2400	B-025	
-96.4923	32.21794	470	B-068	
-96.3535	32.21152	2430	B-069	
-96.346	32.21168	18080	B-070	

-96.4141	32.21516	250	B-094	
-96.41	32.21472	200	B-095	
-96.3998	32.21385	620	B-096	
-96.3622	32.21141	1170	B-097	
-96.3994	32.21362	15120	B-180	
-96.468	32.21516	710	B-026	Neylandville
-96.4392	32.2167	450	B-027	
-96.4431	32.21675	315	B-174	
-96.4633	32.21772	840	B-194	
-96.1994	32.1866	510	B-005	Wills
-96.2537	32.12988	20	B-017	
-96.2283	32.0901	200	B-018	
-96.1895	32.19346	675	B-020	
-96.1785	32.18559	510	B-021	
-96.2477	32.20023	20	B-037	
-96.2385	32.19983	400	B-060	
-96.2425	32.10732	350	B-145	
-96.2469	32.11703	5	B-146	
-96.2475	32.1186	200	B-147	
-96.2531	32.12788	15	B-148	
-96.2586	32.13917	40	B-150	
-96.2623	32.14661	250	B-151	
-96.2686	32.16087	230	B-152	
-96.2726	32.1664	20	B-153	
-96.5192	32.21961	60	B-008	Wolfe
-96.6332	32.23484	620	B-029	
-96.6237	32.23276	625	B-030	
-96.6404	32.23674	775	B-031	
-96.645	32.23934	335	B-032	
-96.5853	32.22828	1460	B-035	
-96.5817	32.22782	330	B-036	
-96.5101	32.2186	130	B-065	
-96.5052	32.21919	75	B-066	
-96.5021	32.21925	350	B-067	
-96.5423	32.22429	14100	B-089	
-96.6781	32.25739	1130	B-009	Ozan
-96.7063	32.26723	16,000	B-033	

-96.813	32.30978	280	B-057	
-96.8152	32.31101	290	B-082	
-96.7678	32.29217	17,400	B-118	
-96.7334	32.27537	16,760	B-120	
-96.7258	32.2735	950	B-122	
-96.844	32.33065	115	B-224	
-96.7729	32.29347	17,110	B-116	
-96.7338	32.27541	16,700	B-119	
-96.6996	32.26382	50	B-124	
-96.7468	32.28229	1100	B-227	
-96.8077	32.30712	570	B-100	
-96.7	32.26406	17,300	B-123	
-96.6912	32.25882	1,015	B-126	
-97.0791	32.50168	540	B-051	
-97.0406	32.42314	185	B-059	
-97.0458	32.44405	130	B-061	
-97.0488	32.44807	8080	B-062	
-97.0214	32.40384	750	B-063	
-97.05	32.44884	18,450	B-090	
-97.0595	32.45993	19,620	B-093	
-97.1362	32.51262	15,260	B-164	
-97.0768	32.51596	9360	B-168	
-97.1525	32.53552	4200	B-84	Eagle Ford
-97.074	32.51824	7000	B-198	
-97.0768	32.51327	5530	B-201	
-97.0768	32.50699	14,050	B-202	
-97.0685	32.53477	75	B-196	
-97.075	32.51746	2200	B-199	
-97.0754	32.51716	3340	B-200	
-97.035	32.41299	16,000	B-209	
-97.1482	32.529	1320	B-318	
-97.0972	32.5055	1005	B-326	

References

Abreu, D. G., Jefferson, I., Braithwaite, P. A., and Chapman, D, N. (2008). “Why is Ssustainability Important in Geotechnical Engineering?” Proc., Geocongress 2008, ASCE, Reston, VA, 821-828.

Aboufirassi, M., and Marino, M. A. (1983) Kriging of Water levels in the Souss Aquifer Morocco, Math. Geol., vol.15, 537-551.

Adams, A.G., Dukes, O.M., Cerato, A.B., and Miller, A.G. (2008). “Sulfate Induced Heave in Oklahoma Soils due to Lime Stabilization.” Geotechnical Special Publication, n 179, p 444-451, (2008), Proceedings of Sessions of GeoCongress 2008 - GeoCongress 2008: Characterization, Monitoring, and Modeling of GeoSystems, GSP 179.

Ahmed A. Eldeiry., Luis A. Garcia. (2012). “Evaluating the Performance of Ordinary Kriging in Mapping Soil Salinity.” Journal of Irrigation and Drainage Engineering., Vol. 138, No.12, ISSN 0733-9437/2012/12; pp.1046-1059.

Ahmed, S., and de Marsily, G. (1989) Cokriged Estimates of Transmissivities using Jointly Water Level Data, In: M. Armstrong (ed.), Geostatistic, Kluwer Academic Pub., 2: 615-628.

Amundaray, J. I. (1994). Modeling Geotechnical Uncertainty by Bootstrap Resampling. PhD thesis, Purdue University, West Lafayette, IN.

Ang, A.H-S., and Tang,W .H . (1975): Probability Concepts in Engineering Planning and Design, Vol. 1: Basic Principles, John Wiley & Sons.

Araghinejad, S., and Burn, D.H. (2005). Probabilistic Forecasting of Hydrological Events using Geostatistical Analysis. Hydrological Sciences Journal, vol. 50, 837-856.

Armstrong, M., 1994. Is Research in Mining Geostats as Dead as a Dodo? In: Dimitrakopoulos, R. (Ed.)., Geostatistics for The Next Century, Kluwer Academic, Dordrecht, pp. 303–312.

Baecher, G. B. (1982). Simplified Geotechnical Data Analysis. In Reliability Theory and its Applications in Structural and Soil Mechanics (ed. P. Thoft-Christensen) 257–277. Dordrecht: D. Reidal Publishing.

Baecher, G.B. (1985) “Geotechnical Error Analysis” Special Summer Course 1.60s, Massachusetts Institute of Technology, Cambridge.

Baecher, G. B. & Christian, J. T. (2003). Reliability and Statistics in Geotechnical Engineering. Chichester: John Wiley & Sons

Bagley, A.D., Cesare, J.A. (2009). “Case History: Sulfate Induced Heave in Lime-Treated Soils Beneath a Structure in Western Colorado.” ASCE, Proceedings 5th Congress on Forensic Engineering, pp. 234-243.

Bardossy, A., and Lehmann, W. (1998) Spatial Distribution of Soil Moisture in a Small Catchment. Part I: Geostatistical Analysis. J. of Hydro., vol. 206, 1-15.

Bedient, P.B., Huber, W.C., 1992. Hydrology and Floodplain Analysis. 2nd ed, Addison-Wesley, Reading, MA.

Benamghar, A., Gomez-Hernandez, J.Jaiime. (2014) “Factorial Kriging of a Geochemical Dataset for Heavy-Metal Spatial-Variability Characterization: The Wallonian Region”. Environmental Earth Sciences, V 71, n 7, pp 3161-3170.

Berger, E., Little, D. N. and Graves, R. (2001). “Technical Memorandum: Guidelines for Stabilization of Soils Containing Sulfates.” http://www.lime.org/publications.html Accessed September, 2010.

Berndtsson, R., and Chen, H. (1994) Variability of Soil Water Content along a Transect in a Desert Area. J. of Arid Environ., vol.27,127-139.

Bilonick, R. A. (1988) Monthly Hydrogen Ion Deposition Maps for the North Eastern US from July 1982 to Sept. 1984. Atmos. Environ., vol.22,1909-1924.

Bohling, G. (2005) “Data Analysis in Engineering and Natural Science Course.” people.ku.edu/~gbohling/cpe940/Variograms.pdf

Boschert J., and Butler J.,(2013) “CLSM as a Pipe Bedding: Computing Predicted Load using the Modified Marston Equation.” ASCE Pipelines 2013, pp. 1202-123.

Box, George E. P.; Cox, D. R. (1964). "An Analysis of Transformations." Journal of the Royal Statistical Society, Series B 26 (2): 211–252

Burgess, T.M., and Webster, R. (1980) Optimal Interpolation and Isarithmic Mapping of Soil Properties, I: The Semivariogram and Punctual Kriging. J. of Soil Sci., Vol. 31, 315-331.

Casado, L.S., Rohani, S., Cardelino, C.A., and Ferrier, A.J. (1994) Geostatistical Analysis and Visualization of Hourly Ozone Data... Atmos. Environ., vol .28, 2105-2118

Challis, R.E., and Kitney, R.I. (1991). “Biomedical Signal Processing(in four parts). Part 1 Time-Domain methods.” Medical & Biological Engineering & Computing, 28, 509-524.

Chen, D. H., Harris, P., Scullion, T., and Bilyeu, J. (2005). “Forensic Investigation of a Sulfate-Heaved Project in Texas”. Journal of Performance of Construction Facilities, ASCE, November 2005, pp. 324-330.

Chittoori, B., Puppala, A. J., Reddy, R., and Marshall, D. (2012). “Sustainable Reutilization of Excavated Trench Material.” Geocongress 2012, ASCE, Reston, VA, 4280-4289.

Chittoori, B., Puppala, A. J., Raavi, A., (2014). “Strength and Stiffness Characterization of Controlled Low-Strength Material Using Native High-Plasticity Clay.” J. Mater. Civ. Eng., 10.1061/(ASCE)MT.1943-5533.0000965, 04014007, 1-8.

Cho, S.E. 2012. Probabilistic Analysis of Seepage that Considers the Spatial Variability of Permeability for an Embankment on Soil Foundation. Engineering Geology, 133: 30–39. doi:10.1016/j.enggeo.2012.02.013.

Clark, I., Harper, W.V. (2002). Practical Geostatistics 2000. Ecosse North America Lic, Columbus, Ohio, USA

Cooley .W. and Tukey J.W. 1965)."An Algorithm for the Machine Calculation of Complex Fourier Series." Mathematics of Computation, 19(90), 297-301, 1965.

Corotis, R.B, Azzouz, A.S., and Krizek, R.J. (1975) "Statistical Evaluation of Soil Index Properties and Constrained Modulus", Proc. of the 2nd. Int. Conf. on Applications of Statistics and Probability to Soil and Structural Engineering, Aachen, Germany.

Creutin, J.D., and Obled, C. (1982) Objective Analysis and Mapping Techniques for Rainfall Fields: An Objective Comparison. Water Resour. Res., vol.18,413-431.

Cuba Miguel A.., Oy Leuangthong., & Julian M.O. (2011), "Detecting and Quantifying Sources of Non-stationarity via Experimental Semivariogram Modeling." Stocj Environ Res Risk Assess (2012) 26:247-260.

David, M. (1977). "Geostatistical Ore Reserve Estimation." Elsevier Scientific Publishing, New York, 364 p.

Davidović . N and Prolović. V (2010). "Modeling of Soil Parameters Spatial Uncertainty by Geostatistics." Facta Univ. Ser. Archit. Civil Eng., 8 (2010), pp. 111–118

Davis, J.C. (1986). Geostatistics and Data Analysis in Geology. John Wiley & Sons.

Delaunay, Boris: "Sur la Sphère Vide. A la Mémoire de Georges Voronoï." Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathématiques et na, 1934, no. 6, 793–800.

Delhomme, J.P. (1978) Kriging in the Hydroscience. Adv. in Water Resour., vol.1, 251-266.`

DeGroot, D. J. (1996). "Analyzing Spatial Variability of In Situ Soil Properties." (invited paper). *Uncertainty in the Geologic Environment, From Theory to Practice*, Proceeding of Uncertainty'96, Geotechnical Special Publication No. 58.

Dermatas, D. (1995). "Ettringite-induced Swelling in Soils: State-of-the-art." Applied Mechanics Rev, Vol. 48, pp. 659-673. Douglas C. Montgomery, 2009, Design and Analysis of Experiment, 7th edition.

Dirks, K.N., Hay, J.E., Stow, C.D., Harris, D., 1998. High-Resolution Studies of Rainfall on Norfolk Island Part II: Interpolation of Rainfall Data. J. Hydrol. 208 (3-4), 187–193.

Einstein, H.H. and Baecher, G.B. (1982), "Probabilistic and Statistical Methods Engineering Geology, I. Problem Statement and Introduction to Solution." Rock Mechanics and Rock Engineering, Vol. 12, pp. 47-61 of Soil Heterogeneity: Quantification and Implications on Geotechnical Field Problems. Canadian Geotechnical Journal, 40(1): 1–15. doi:10.1139/t02-090.

Fenton, G.A., and Griffiths, D.V., Statistics of Block Conductivity through a Simple Bounded Stochastic Medium, *Water Resources Research*, 29(6), 1825--1830, 1993.

Fenton, G.A (1994)." Error Evaluation of Three Random Field Generators", *ASCE Journal of Engineering Mechanics,* 120(12), 1994, 2478-2497.

Fenton, G.A., and Vanmarcke, E. H. (1998). "Spatial Variation in Liquefaction Risk." Geotechnique, Vol.48, No.6, pp. 819-831.

Fenton G.A (1999a). "Estimation for Stochastic Soil Models." *ASCE Journal of Geotechnical and Geoenvironmental Engineering,* 125(6), 1999a, 470-485.

Fenton, G.A.(2002). Risk Assessment and Management, in Lecture Notes, Department of Engineering Mathematics, Dalhousie University, Canada, 2002.

Filippas, O.B., Kulhawy, F.H., and Grigoriu, M.D. (1988) “Reliability-based Foundation Design for Transmission Line Structures: Uncertainties in Soil Property Measurement.”

Folliard, K. J., Trejo, David, Sabol, S. A., and Leshchinsky, D. (2008) “Development of a Recommended Practice for Use of Controlled Low-Strength Material in Highway Construction.” NCHRP 597 Report, 5-39.

Goovaerts, P. (2000). “Geostatistical Approaches for Incorporating Elevation into the Spatial Interpolation of Rainfall.” Journal of Hydrology 228 (2000), 113-129.

Germann, U., and Joss, J. (2001) Variograms of Radar Reflectivity to Describe the Spatial Continuity of Apline Precipitation. J. of Applied Meteorology, vol.40,1042-1059.

Gui, S., Zhang, R., Turner, J.P., and Xue, X. 2000. Probabilistic Slope Stability Analysis with Stochastic Soil Hydraulic Conductivity. Journal of Geotechnical and Geoenvironmental Engineering, **126**(1): 1–9. doi:10.1061/(ASCE)1090-0241(2000)126:1(1).

Haining, R.P. et al., (2010). “Geography, Spatial Data Analysis and Geostatistics: An Overview.” Geographical Analysis. Vol. 42, pp.7-31.

Hammah, R.E., and Curran, J.H. (2006). “Geostatistics in Geotechnical Engineering: Fad or Empowering?” GeoCongress 2006*:* Geotechnical Engineering in the Information Technology Age, v 2006, p 102.

Hahn, G.J., and Shapiro, S.S. (1967): Statistical Models in Engineering, John Wiley & Sons, New York.

Hawkins, A B; Pinches, G M (1987) “Expansion due to Gypsum Growth.” Proceedings

6th International Conference on Expansive Soils, New Delhi, 1-4 December 1987 P183- 187.

Harr, M.E. (1977): Mechanics of Particulate Media: A Probabilistic Approach. McGraw-Hill, New York.`

Harris, P., Voldt, J. V., Sebesta, S., and Scullion, T. (2005). "Recommendations for Stabilization of High-Sulfate Soils in Texas.. Technical Report 0-4240-3, Texas Transportation Institute.

Hegazy, A.H., Mayne, P.M., and Rouhani, S. (1996). "Geostatistical Assessment of Spatial Variability in Piezocone Tests." *Uncertainty in the Geologic Environment, From Theory to Practice*, Proceeding of Uncertainty '96, Geotechnical Special Publication No. 58.

Hohn, M.E., "Geostatistics and Petroleum Geology 2nd Ed". Kluwer Academic Norwell, MA.

Holdaway, K., "Harness Oil and Gas Big Data with Analytics. Optimize Exploration and Production with Data-Driven Models." John Wiley & Sons, ISBN: 9781118910955.

Holtz, R.D., and Krizek, R.J. (1972) "Statistical Evaluation of Soil Test Data." Proceedings of the First International Conference on Applications of Statistics and Probability to Soil and Structural Engineering, Hong Kong.

Howell, D.C. (2007). Statistical Methods for Psychology, 6th edition

Huang, J.S., Griffiths, D.V., and Fenton, G.A. 2010. System Reliability of Slopes by RFEM. Soils and Foundations, 50 (3): 343‐353. doi:10.3208/sandf.50.343.

Hunter, D. (1989). "The Geochemistry of Lime-Induced Heave in Sulfate-Bearing Clay Soils." (Dissertation) University of Nevada, Reno, 1989.

Hyunki Kim. (2005). “Spatial Variability in Soils: Stiffness and Strength.” PhD. Thesis, Georgia Institute of Technology, 201.

Isaaks, E. H., and Srivastava, R. M. (1989). “ An Introduction to Applied Geostatistics”, Oxford University Press, New York.

Jianshu, Lv., Liu, Yang., Zhang, Zulu., Dai, Jierui. (2013) “Factorial Kriging and Stepwise Regression Approach to Identify Environmental Factors Influencing Spatial Multi-Scale Variability of Heavy Metals in Soils.” Journal of Hazardous Materials, V 261, pp 387-397.

Jones, A.L., Kramer. S.L. and Arduino, P. (2002), “Estimation of Uncertainty in Geotechnical Properties for Performance-based Earthquake Engineering,” PEER report 2002/16. Journel, A. G., and Huijbregts, C. J. (1978). “Mining Geostatistics,” Academic Press, London, 600 p.

Journel, A.G., Huijbregts, C.J., 1978. Mining Geostatistics. Academic Press, New York.

Kanevski M, Arutyunyan R, Bolshov L, Demyanov V, Maignan M (1996) “Artificial Neural Networks and Spatial Estimations of Chemobyl Fallout.” Geoinformatics 7(1-2): 5-11.

Kim, J. H. (2011). Improvement of Geotechnical Site Investigations via Statistical Analyses and Simulation. PhD thesis, Georgia Institute of Technology.

Kitanidhis, Peter (1997). “Introduction to Geostatistics – Applications in Hydrogeology.” Cambridge University Press, New York, 249 pp., ISBN 0-521-58747-6.

Kota, P. B. V. S., Hazlett, D., and Perrin, L. (1996) “Sulfate-Bearing Soils: Problems with Calcium-Based Stabilizers.” Transportation Research Record 1546, TRB, National Research Council, Washington, D.C., 1996, pp. 62-69.

Krige, Danie G. (1951). “A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand”. J. of the Chem., Metal. And Mining Soc. of South Africa 52 (6): 119-139.

Kulhawy, F.H., Birgisson, B., and Grigoriu, M.D. (1992). Reliability-based Foundation Design for Transmission Line Structures: Transformation Models for In Situ Tests. Electric Power Research Institute, Palo Alto, Calif., Report EL-5507(4).

Kulhawy, F.H., and Phoon, K. (1996). “Engineering Judgment in the Evolution from Deterministic to Reliability-based Foundation Design (plenary).” *Uncertainty in the Geologic Environment, From Theory to Practice*, Proceeding of Uncertainty ’96, Geotechnical Special Publication No.58.

Kulhawy, F.H., and Trautmann, C.H. (1996). ”Estimation of In Situ Test uncertainty in the Geologic Environment (GSP 58)”. Edited by C.D. Shackelford, P.P. Nelson, and M.J.S. Roth. American Society of Civil Engineers, New York, pp. 269– 286.

Kumar, V. (1996) Space Time Modelling of Ground Water with Assistance of Remote Sensing, Ph.D., Indian Institute of Technology, New Delhi, India.

Kumar, V. and Remadevi (2006).”Kriging of Groundwater levels- A Case Study.” Journal of Spatial Hydrology. Vol.6, No.1,81-92.

Kuroda, K., Yamada, M., Kani, S. (1983). “Statistical Decision in Design & Construction of Geotechnical Structures based on Probabilistic & Fuzzy Soil Information.” Pitagora Editrice, V 1, p 561-572.

Lacasse, S. and Nadim, F. (1996), “Uncertainties in Characterizing Soil Properties.” in Uncertainty in the Geologic Environment (GSP 58), edited by Shackelford, C.D., Nelson, P.P., and Roth, M.J.S., ASCE, New York, pp. 49-75

Li. K.S., White W. (1987). “Probabilistic Characterization of Soil Profiles.” Res. Report 19, Canberra: UNSW, ADFA (1987).

Little, D.N., Deuel, L. (1989), “Evaluation of Sulfate–Induced Heave at Joe Pool Lake.” Chemical Lime Company, June.

Lumb, P. 1971. Precision and Accuracy of Soil Tests. *In* Proceedings of the 1st International Conference on Applications of Statistics and Probability in Soil and Structural Engineering, Hong Kong, pp. 329–345.

Magneron, Cedric., Deraisme, Jacques., Jeannee, Nicolas (2009). “Noise Reduction by M-Factorial Kriging.” IAMG’09- pp 23-28

Matheron G. (1973). “The Intrinsic Random Functions and their Applications.” Adv. In Appl. Probab., 5, 439-468, 1973.

Matheron G. (1982). “Pour une Analyse Krigeante des Donnees Regionalisees.” Report N-732, Centre de Geostatistique, ENSMP, Fontainebleau, France.

Mayne, P.W. (2007). “Cone Penetration Testing: A Synthesis of Highway Practice.” Project 20-5. Transportation Research Board, Washington, D.C. NCHRP synthesis 368.

Mingyu, H., Fumei, L., Mingshu, T., (2006), “The Thaumasite Form of Sulfate Attack in Concrete of Yongam Dam.” Cement and Concrete Research, Vol. 36, Iss. 10, pp. 2006-2008.

Miller, J., Franklin, J., and Aspinall, R. (2007). “Incorporating Spatial Dependence in Predictive Vegetation Models.” Ecol. Model., 202 (3-4), 225-242.

Mignolet M.P. and Spanos P.D. (1992). Simulation of Homogenous Two-Dimensional Random Fields: Part I- AR and ARMA Models. ASME J. Appl. Mech. 59, S260-S269, 1992.

Mitchell J.K., Soga, K (2005). “Fundamentals of Soil Behavior.” 3rd edition, Wiley Publishers, ISBN: 978-0-47-46302-3.

Montgomery., Runger and Hubele (2010)."Engineering Statistics." 4th edition, SBN-13:978*0-40-52694-1.

Mothukori. P. R., (2014) "Non-Destructive Sasw Evaluation of Controlled Low Strength Material as a Pipeline Bedding Material." M.S. thesis, Univ. of Texas, Arlington, TX, 110.

Nazarian, S., and Stokoe, K. H., II (1984). "In Situ Shear Wave Velocities from Spectral Analysis of Surface Waves." Proceedings of the 8th World Conference on Earthquake Engineering, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, Vol. III, 31-38.

Olea, R.A. (2009). "A Practical Primer on Geostatistics." USGS, Open File Report. 2009-1103.

Orchant, C.J., Kulhawy, F.H., and Trautmann, C.H. (1988). Reliability-based Foundation Design for Transmission Line Structures: Critical Evaluation of In Situ Test Methods. Electric Power Research Institute, Palo Alto, Calif., Report EL-5507(2).

Parsons, R.L., Frost, J.D., 2002. Evaluating Site Investigation Quality using GIS and Geostatistics. Journal of Geotechnical and Geoenvironmental Engineering 128, 451–461.

Perrin, L. (1992). "Expansion of Lime-Treated Clays Containing Sulfates." Proceedings of 7th International Conference on Expansive Soils, Vol. 1, ASCE Expansive Soil Research Council, New York, 1992, pp. 409-414.

Petry, T. M. and Little, D. N., (1992), "Update on Sulfate Induced Heave in Treated Clays: Problematic Sulfate Levels." Transportation Research Board, TRR 1362, Washington, DC.

Phoon, K.K., Kulhawy, F.H. and Grigoriu, M.D. (1995), "Reliability-based Design of Foundation for Transmission Line Structure," Report TR-105000, Electric Power

Research Institute, Palo Alto.

Phoon, K., and Kulhawy, F.H. (1999). "Characterization of Geotechnical Variability." Canadian Geotechnical Journal, Vol. 36, pp. 625-639.

Puppala, A, J., and Hanchanloet, S. (1999). “Evaluation of Chemical Treatment Method (sulphuric acid and lignin mixture) on Strength and Resilient Properties of Cohesive Soils.” Proc., 78th Transportation Research Board Annual Meeting, Transportation Research Board, Washington, DC.

Puppala, A. J., Viyanant, C., Kruzic, A. P., and Perrin, L., (2002). Evaluation of a Modified Soluble Sulfate Determination Method for Fine-Grained Cohesive Soils. Geotechnical Testing Journal, Volume 25, Issue 1, pp. 85–94.

Puppala, A. J., Wattanasanticharoen, E. and Punthutaecha, K. (2003). “Experimental Evaluations of Stabilization Methods for Sulphate-rich Expansive Soils.” Ground Improvement Vol. 7, No. 1, 2003. pp. 25-35.

Puppala, A. J., Kadam, R., Madhyannapu, R., and Hoyos, L. R. (2006). “Small-Strain Shear Moduli of Chemically Stabilized Sulfate-Bearing Cohesive Soils.” Journal of Geotechnical and Geoenvironmental Engineering, ASCE, March 2006, pp.322-336.

Puppala, A.J., et al., (2010) “Forensic Investigations to Evaluate Sulfate–Induced Heave Attack on a Shotcrete Tunnel Liner.” Journal of Materials in Civil Engineering, ASCE, Vol 22, pp. 914-922.

Puppala, A.J., Naga S. Talluri., Bhaskar S. Chittoori and Ahmed Gaily. (2012). Lessons Learned from Sulfate-Induced Heaving Studies in Chemically-Treated Soils. Proceedings of the International Conference on Ground Improvement and Ground Control. Research Publishing, Vol. 1, November, 2012, pp.85-98.

Puppala, A. J., Saride, S., and Williammee, R. (2012a). “Sustainable Reuse of Limestone Quarry Fines and RAP in Pavement Base/Subbase Layers.” J. Mater. Civ. Eng., 10.1061/(ASCE)MT.1943-5533.0000404, 418-429.

Puppala, A.J., Naga S. Talluri., Ahmed Gaily, and Bhaskar S. Chittoori. (2013). Heaving Mechanisms in High Sulfate Soils. Proceedings of the 18th International Conference on Soil Mechanics and Geotechnical Engineering, Paris 2013.

Queiroz, Joaquim C. B., Sturaro, Jose R., Saraiva, Augusto C. F., Barbosa Landim, Paulo M. (2008) “Geochemical Characterization of Heavy Metal Contaminated Area using Multivariate Factorial Kriging”. Environmental Geologu, V 55, n 1, pp 95-105.

Raavi, A. (2012). “Design of Controlled Low Strength Material for Bedding and Backfilling Using High Plasticity Clay.” M.S. Thesis, Univ. of Texas, Arlington, TX, 118.

Rajani, B., and Tesfamariam S. (2004). Uncoupled Axial, Flexural, and Circumferential Pipe-Soil Interaction Analyses of Partially Supported Jointed Water Mains. Canadian Geotechnical Journal, 41, 997-1010.

Rajah S., McCabe M., and Plattsmier J., (2012) “Classification and Specification of Bedding and Backfill for Buried Pipelines.” ASCE Pipelines 2012: Innovations in Design, Construction and Maintenance. pp:940-951.

Randal J. Barnes (1991). “The Variogram Sill and the Sample Variance”. Mathematical Geology, Vol.23, No.4, 673-678.

Rethati, L. (1988): Probabilistic Solutions in Geotechnics, Developments in Geotechnical Engineering Vol. 46, Elsevier, Budapest.

Rollings, R.S., Rollings, M.P., Poole, Toy., Wong, G.S., and Gutierrez, Gene (2006).“Investigation of Heaving at Holloman Air Force Base, New Mexico.”

Journal of Performance of Constructed Facilities, v 20, n 1, p 54-63, February 2006.

Samui, Pijush., Sitharam, T.G. (2008). “A Comparative Study of Ordinary Kriging and Support Vector Machine Models for the Spatial Variability of Rock Depth in Banglore.” Geotechnical Special Publication, n 179, pp934-941.

Seica, M. V. and Packer, J. A. (2001) “Properties and Strength of Aged Cast Iron Pipes.” Journal of Materials in Civil Engineering, ISSN 0899-1561.

Shapiro, S. S.; Wilk, M. B. (1965). "An Analysis of Variance Test for Normality (complete samples)." *Biometrika* 52 (3–4): 591–611.

Sherwood, P. T. (1962). “Effect of Sulfates on Cement and Lime Treated Soils.” Highway Research Board, Bulletin 353, pp. 98-107.

Soulie, M., Montes, P. and Silvestri, V. (1990), “Modelling of Spatial Variability of Soil Parameters.” Canadian Geotechnical Journal, Vol. 27, pp. 617-630

Spry, M.J., Kulhawy, F.H., and Grigoriu, M.D. (1988) “Reliability-based Foundation Design for Transmission Line Structures: Geotechnical Site Characterization Strategy. Electric Power Research Institute, Palo Alto, Calif., Report EL-5507(1).

Stokoe, K. H. II, Rix, G. J., and Nazarian, S. (1989). “In Situ Seismic Testing with Surface Waves.” Proceedings of the 12th Int. Conf. on Soil Mechanics and Foundation Engineering, Rio De Janiero, 331-334.

Storm, B., Jenson, K.H., and Refsgaard, R.C. (1988) Estimation of Catchment Rainfall Uncertainty and its Influence on Runoff Prediction. Nordic Hydrology, vol.19,79-88.

Sturges, H. A. (1926). "The Choice of a Class Interval." *Journal of the American Statistical Association*: 65–66.

Tabachnick, B.G., & Fidell, L. S. (2007), Using Multivariate Statistics. 5th edition.

Tabios, G.Q., Salas, J.D., 1985. A Comparative Analysis of Techniques for Spatial Interpolation of Precipitation. Water Resour. Bull. 21 (3), 365–380.

Tang, W.H. 1984. Principles of Probabilistic Characterization of Soil Properties. *In* Probabilistic Characterization of Soil Properties: Bridge between Theory and Practice. *Edited by* D.S. Bowles and H-Y. Ko. American Society of Civil Engineers, Atlanta, pp. 74–89.

Thiessen, A.H., 1911. Precipitation Averages for Large Areas. Monthly Weather Rev. 39 (7), 1082–1084.

Thomey, J. (2013) "Quantification and Geostatistical Mapping of Soluble Sulfates in Soils along a Pipeline Alignment." M.S. thesis, Univ. of Texas, Arlington, TX, 145.

Vieira, S.R., Carvalho,J.R., Ceddia, B, M., Ganzalez, P. A.(2010). "Detrending Non-Stationary Data for Geostatistical Applications." Bragantia vol.69, 1-9.

Vipulanandan, C., Qiao, W., Hovsepian, H., (2011) "Case Studies on Water Pipeline Failures in the Active Zone." Proceedings of Geo-Frontiers 2011 Conference, n211 GSP, p 2474-2483

VanMarcke. E. H. (1984) Random Fields: Analysis and Synthesis. M.I.T. Press, Cambridge.

Vanmarcke, E.H. (1977) "Probabilistic Modeling of Soil Profiles." Journal of the Geotechnical Engineering Division, ASCE, 103(GT11, pp 1227-1246. Electric Power Research Institute, Palo Alto, Calif., Report EL-5507(3).

Venkatramaiah.C (2006)."Geotechnical Engineering." Third Edition, New Age International Publishers, ISBN :81-224-1793.

Vennapusa, K. R. P., White, D. J., and Morris, M.D. (2010). "Geostatistical Analysis for Spatially Referenced Roller-integrated Compaction Measurement." *Journal of*

Geotechnical and Geoenvironmental Engineering, American Society of Civil Engineers, 136(6): 813-822.

Vieira, S.R., Nielsen, D.R., and Biggar, J.W. (1981) Spatial Variability of Field Measured Infiltration Rate. Soil Sci. Soc. Am. J., vol.45,1040-1048.

Vipulanandan, C., Qiao, W., Hovsepian, H., (2011) "Case Studies on Water Pipeline Failures in the Active Zone." Proceedings of Geo-Frontiers 2011 Conference, n211 GSP, p 2474-2483.

Virdee, T.S., and Kottegoda, N.T. (1984) A Brief Review of Kriging and its Application to Optimal Interpolation and Observation Well Selection. Hydro. Sci. J., vol.29,367-387.

Volpi, G., and Gambolati, G. (1978) On the Use of Main Trend for the Kriging Technique in Hydrology. Adv. Water Resour., vol.1,345-349.

White, D. J., Thompson, M. J., Vennapusa, P., and Siekmeier, J. (2008). "Implementing Intelligent Compaction Specification on Minnesota TH-64 Synopsis of Measurement Values, Data Management, and Geostatistical Analysis." Transportation Research Record, *Journal of the Transportation Research Board*, 2045, 1–9.

White, D. J., Vennapusa, P., and Gieselman, H. (2008b). "Roller-integrated Compaction Monitoring Technology: Field Evaluation, Spatial Visualization, and Specifications." Pro., 12th Intl. Conf. of Intl. Assoc. for Computer Methods and Advances in Geomechanics (IAC-MAG).

White, D. J., Vennapusa, P., and Gieselman, H. (2011). "Field Assessment and Specification Review for Roller-integrated Compaction Monitoring Technologies." *Advances in Civil Engineering*, 2011,1-15.

Yaglom A.M. (1962). “An Introductory to the Theory of Stationary Random Functions.” Dove, Mineola, NY 1962.

Zhiming, S. (2008). “Forensic Investigation of Pavement Premature Failure Due to Soil Sulfate-Induced Heave.” Journal of Geotechnical and Geoenvironmental Engineering, v134, n 8, p 1201-1204, August 2008.

Zhu, H., Zhang, L.M., Zhang, L.L., and Zhou, C.B. 2013. Two-dimensional Pprobabilistic Infiltration Analysis with a Spatially Varying Permeability Function. Computers and Geotechnics, **48**: 249–259. doi:10.1016/j.compgeo.2012.07.010.

Biographical Information

Tejo Vikash Bheemasetti was born and brought up in Visakhapatnam which is located in southern part of India. He completed his Bachelor's of Technology in Civil Engineering in the year 2007 from GMR Institute of Technology, which is affliated to JNTU, Hyderabad. After completing his bachelor's degree he attended Indian Institute of Technolgy, Kharagpur to pursue a Master's degree. He received his Master of Technology in Civil Engineering in May, 2009 with specialization in Geotechnical Engineering. Upon completion of his Master's degree, Tejo Vikash worked as geotechnical design engineer for 2 years 7 months in geotechnical engineering firms.

In the Spring of 2012, he was accepted and enrolled into the University of Texas at Arlington's graduate program. He worked under the guidance of Dr. Anand J Puppala and Dr. Xinbao Yu on variability of soil properties in geotechnical studies. He successfully defended his dissertation in November 2014. During his course of study he submitted several technical papers and book chapters and co-author for geotechnical conferences.