A Survey on The Methods of Spatial Statistics

Hawri Hashm Sayed

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the Degree of

Master of Science
in
Applied Mathematics and Computer Science

Eastern Mediterranean University
June 2014
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Elvan Yılmaz
Director

I certify that this thesis satisfies the requirements as a thesis for the degree of Master of Science in Applied Mathematics and Computer Science.

Prof. Dr. Nazim Mahmudov
Chair, Department of Mathematics

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Applied Mathematics and Computer Science.

Assist. Prof. Mehmet Ali Tut
Supervisor

Examining Committee

1. Prof. Dr. Rashad Aliyev

2. Asst. Prof. Dr. Hüseyin Etikan

3. Asst. Prof. Dr. Mehmet Ali Tut

ABSTRACT

In this thesis the kriging approach is presented for interpolating spatial data points. The mathematical model of the kriging method is mentioned first. Then a small example is provided about the steps of this approach.

ArcGIS software package is an important tool for creating maps and also for performing geostatistics about large data points. A kriging example is also tried by this package and the predicted data points are listed.

Keywords: Geostatistics, spatial data, kriging, Arcgis.

ÖZ

Bu çalışmada, harita bilgileri (kordinatlar) üzerinde enterpolasyon uygulamalarında kullanılan kriging metodu bahsedilmektedir. Bu metodun matematiksel modeli yanında, seçilen küçük bir very örneği kullanılarak hesaplamalarla yaklaşık kestirme değerleri hesaplanmıştır. Bu analiz için geniş verilerin kullanılmasında en önemli yazılımlardan olan Arcgis paketi de kullanılarak seçilen verilerin analizi ve kestirim sonuçları eğrilerle birlikte sunulmuştur.

Anahtar Kelimeler: Coğrafik istyatistik, koordinat verileri, kriging, Arcgis.

ACKNOWLEDGMENTS

At the first I would like to thanks my God for help me and gave me a good health to finish my study.

I would like to express my deep gratitude to my thesis supervisor (Assist. Prof. Mehmet Ali Tut), I have learned many things when I became his student, I appreciation him to accept my topic, and he instructing me how I collect idea. During the period of two years, many instructors in my department and many friends gave color to my life. I cannot list all of them by limited sentences. I want to thank all my professors in Mathematics Department at EMU, I would also like to thank all (Galozy- school) s' teachers which support me in my academic life in this country. My advanced thanks is going to my parents which contribute me and I could graduate by their advice which always showed me the wright way in my life. My strong thanks is going to my brother (Ary) to the twin soul mate my way to his good heart and sincerity, and his wife (Zhino). In honor way, I would like to thank my honey sisters (Chwvin, Azhin, Hero, and sweety Hala)

Another thanks is faced to my dear uncle (Hemn) and his wife (Bnar).

I would like to thanks all my friends who have been with me and supported me since childhood in my city.

TABLE OF CONTENTS

LIST OF SYMBOLS / ABBREVIATIONS

u, u_α	Location vectors for estimation point and one of the neighboring data points, indexed by α
$n(u)$	Number of data points in local neighborhood used for estimation of $Z^*(u)$
$m(u), m(u_\alpha)$	Expected values (means) of $Z(u)$ and$Z(u_\alpha)$.
$\lambda_\alpha(u)$	Kriging weight assigned to datum $Z(u)$and $Z(u_\alpha)$for estimation location u same datum will receive different weight for different estimation location
$Z(u)$	Is treated as a random field with a trend component
$C(u)$	The covariance function
μ_{ok}	Lagrange parameter.
μ	Mean of spatial random variable
σ_ε^2	The error variance (estimator variance)
v	The estimation variance coordinate vector in R^n, n=1, 2 or 3.
$L(v) \sim N(m, \sigma^2)$	a second order stationary multivariate normal or (Gaussian) random function
σ^2	Population Variance
$Y(v) \sim N(0,1)$	Standard normal random function.

σ^2, m	The parameter are called the logarithmic mean and variance respectively of Z(v).
h	A displacement between two spatial locations Z(u) and Z(u+h).
D	A specific location which is including known data and unknown data.
γ	Semivariogram function.
2γ	Variogram function.
ρ	Correlation coefficient function
a	Range
$C(0)$	Apriori variance
$\gamma(\infty)$	Sill

Chapter 1

INTRODUCTION

1.1 Spatial Statistics

The statistician job is mainly concerns in the collection, the organization, summarization and analysis of data. The previous works lead us to the drawing of inferences about a body of data, when only a part of the data is observed. When it comes in modeling space, we talk about spatial statistics. This means spatial statistics is the collection and the analysis of spatial data. In other words, the spatial statistics is based on the collection of data and its use for pattern analysis, spatial association, scale and zoning, classification, geostatistics, spatial econometrics and spatial sample relationships and trends [1]. With the growth of computational of all science, we have some core software such as Geographic Information System (GIS) which is often used in spatial statistic.

1.1.1 A survey about what they did?

Spatial statistic / statistician helps us to well understand geographic phenomena. It can be used for drawing a map; which can later help us to locate the origin and the level of a particular phenomenon (disease, rain, sun) in some region. With spatial statistic, we can make a decision with a higher confidence level; this means we always use the fundamental statistical test theory to confirm our results at a certain level of signification. We can also remove the complexity and noise in our data before process them. This will enable us to get directly at the overall patterns [2].

This study examines the predication of random spatial process by using variogram function or covariance function for regionalized variable, as well as the prediction for this process by kriging. In this study we refer to the phenomenon of prediction by using the known data from region D to get the advantage of unknown data within region D the previous procedure is done by regression technique and kriging technique. Regression technique involves the using of generalized least square estimator. Kriging technique in spatial statistic is a technique which particularly is used. To predicate the phenomenon of location such as (under and surface of earth metals, underground water, pollution of environment, spread out of natural forests, also the prediction of the spread of diseases, disasters of nature and its use in the field of economy [3].

The kriging technique is used in any study if it is possible to define the phenomenon under study on the basis of the distance between data samples of this phenomenon. The kriging technique leads to reduce using regression technique for prediction. Explanatory variables while does not include kriging only know the distance between the views phenomenon, moreover the mean square error of prediction in the technique of kriging always smallest than the regression technique [4].

1.1.2 The Aim of Spatial Statistics

The basic aim of spatial statistics is to conduct the prediction on the phenomenon of spatial by kriging method. After studying the phenomenon under study spatial and examined in terms of stationary, and non-stationary property, and estimate the parameters of a variance function or a variogram function, in order to obtain an accurate prediction. Then when we get a prediction by kriging method. Since that's better than regression method if we compared with each other by variance calculation, and to find out which is better and has a best accuracy to be relied upon.

1.1.3 Types of Spatial Statistics

a) Geostatistics: Variogram and kriging

b) Lattice or areal data: lattice or Areal model have the aim of predicting Z(u), where u is an area instead of a point as in the geostatistical / point -suggested model case, Markov random field and Conditional auto regression model (CAM) [5].

c) Spatial point pattern: Complete spatial randomness (CSR) and K -function, L-function [5].

In this study kriging method will be discussed theoretically in chapter 2. In chapter 3 by using a sample points the kriging approach will be used for these interpolation. Chapter 4 is going to be the discussions and conclusion about the result.

Chapter 2

LITERATURE REVIEW

It is well known that there are a lots of words done about spatial data analysis by different authors. Most of these studies had involved ArcGIS software. Some selected paper will be mentioned in the following paragraphs to give an input about of such spatial statistics in different fields of science. The selected paper are for ESRI conference paper. ESRI is the owner of the software ArcGIS.

Stratified random sampling method was used to monitor water quality in Pinellas Country. At the beginning water quality parameters were tried to be determined by mean calculations. To see the spatial trends of the data Invers distance weight (IDW) was used. The quality of the water was shown that it is about watching that of the state water quality scores. The results were shared by Cynthia Meyer [6].

Health risk analysis is an important matter in Taiwan. The relationship between diffusion of environmental contamination and health of residents were analyzed by Hsiao-Hui Chen, et al. [7]. Several kriging models were used to obtain the best watching models. It was shown that old petroleum plants create pollution. Bedard, F., Savio, J., and Collins, K have prepared Canada statistics for crime prevention center of Canada. They focused on the relationship between crime density, land use, socioeconomic and demographic indicators, and travel pattern of the accused to the crime location by crime type. So that the office will be generated strategies for preventing crimes [8].

Bucciarelli, A., et al. (1994-2003) [9] have done a research about the annual mortality rate for accidental drug overdose in New York City. The data selected are between 1994- 2003. They were geocoded and mapped analyzed by ArcGIS. The important of geospatial technology has been understand by Barras, G [10] that search and Rescue (SAR) members can fight easily with the crimes. The data was displayed in ArcGIS.

Ward, B., Wells, B., Davenhall, B. (2011) conducted the data about spatial analysis of the basketball tournament. The study is collected the data to wonder either there is spatial correlation between distance of competing terms to their game sites [11].

According to Wetherbee, S. The spatiotemporal pattern salmonella sample is either non-random or random. If it is non spatial pattern spatiotemporal point pattern analysis will be used. For hypothesis testing spatial statistic from ArcGIS geoprocessing environment will be used. If non-random distribution are exploited, second question will be analyzed by using GIS. For example, what are relationships between salmonella distribution and the environmental geography of farms [12].

Deakin, R. (1990s) in his study examines that geospatial data is a fundamental building to grow the method which depend on internal mapping of defense infrastructure, flood plain shape and risk a respect. Within GIS, vector and tesseral data are linked and integrated to give input to novel probabilistic flood risk models, the output from which defend a national, web enabled flood map. Scenario development allowing is enabled by these methods to use by government to analyze potential investment policy options and response climate change [13].

Lemos, N., Batista, M., Silva, T., Nobrega, T. Their study presents an urban drainage performance indicator which is named IDU. IDU is regarded to GIS that does spatial decision confirm system possible. To calculating IDU, it is possible to consider that state of street in an urban sector neighborhood to values of IDU, a performance divisions is used. It is empirical action may be fruitful as a tool of urban planning and for pointing out infrastructural investment properties it was applied in the costal neighborhood of Joao Pessoa City, Brazil. The findings of this study indicated that the neighborhoods of Bessa and Aero club had the worse related performance urban draining quality, Cabo Branco, Tambau and Manaira had showed the best performance [14].

Le Roux, P. Discussed spatially a system which has been designed in city water serviced technical operation center. This is regard with business process and manages all in coming events generated in city to water service. This applications are divided, stored and resourced geo data base. There is a set of role within the system which guides what answering is standard for each application. The SOP's (standard operation procedure) makes events spatially to have roles to evaluate spatial patterns. Through a map service, the system is regard to city's exiting GIS architecture and harvest the data. It also GPS give ability to vehicles. One point of access for all application and part responses to events within the city will be ensured [15].

Chapter 3

PREDICTION BY USING KRIGING AND INVERSE DISTANCE WEIGHT IN SPATIAL STATISTICS

3.1 What is Kriging?

Kriging is an evaluation process that gives a perfect result with surface, and the good unbiased linear estimation either of each point value or block average [16].

Geostatistical technicality not just have the efficiency of producing a prediction surface however also supply a measure of certainty or precision [17].

3.2 Kriging methods

Kriging is a group of estimator involved to generate spatial data, this group contains ordinary kriging, simple kriging, universal kriging, co-kriging, and others. We have many types of kriging estimator method, but at most we are going to proof two of them:

1. Simple kriging $\boldsymbol{E}[\boldsymbol{Z}(\boldsymbol{u})] = \boldsymbol{m}$, where mean m is known, it is usually involved to practical applications because the mean is rarely known. It is sometimes used in large mines as mention in South Africa where the mean of every places is known since that locations has been mined for many years.

2. Ordinary kriging is the best method used in kriging method. It provide to estimate a value at a point of a location for which a variogram is known, using data in the neighborhood of the estimation location. Suppose stable of the first time of all random variable

$$\boldsymbol{E}[\boldsymbol{Z}^*(\boldsymbol{u})] = \boldsymbol{E}[\boldsymbol{Z}(\boldsymbol{u})] = \boldsymbol{m}$$

where mean $\boldsymbol{m}$ is unknown.

3. IRFK-Kriging suppose $\boldsymbol{E}[\boldsymbol{Z}(\boldsymbol{x})]$ to unbeknown polynomial in$\boldsymbol{x}$.
4. Indicator kriging usage instead of method itself, in order to evaluation transition probabilities, or the purpose of it, is used when it is covetable to estimate a distribution of values inside region instead of just a region the mean value of an region.
5. Multiple –indicator kriging (MIK).
6. Disjunctive kriging is a nonlinear generalization of kriging.
7. Lognormal kriging (logarithms): we need to limit for collectively hypothesis, the problem here is not treated support of change, but it is treat and support only estimation with multivariate lognormal distribution. It follows that: $\boldsymbol{Z}(\boldsymbol{v}) = \boldsymbol{e}^{\boldsymbol{m}+\boldsymbol{\sigma y}(\boldsymbol{v})} = \boldsymbol{e}^{\boldsymbol{L}(\boldsymbol{v})}$ [18].
8. Probability kriging.
9. Cokriging this type of kriging method is used, when math do estimate only one variable between two variable. Then that two variable called it co-variable, thus should have a good relationship should be define.
10. Universal kriging suppose (a public polynomial orientation model), and it is used to estimate spatial means when the data have a strong trend and the trend can be modeled by simple functions [19].

$$E[Z(x)] = m = \sum_{k=0}^{p} \beta_k f_k(x)$$

3.2.1 Simple kriging

In simple kriging, we suppose that the trend compound is a constant with mean known, so that:

$$Z_{sk}^*(u) = m + \sum_{\alpha=1}^{n(u)} \lambda_\alpha^{sk}(u)\left[Z(u_{\alpha)} - m\right] \qquad (3.1)$$

For unbiased: The estimate should be minimum variance and unbiased, so as to be unbiased the estimate error should be the expected value of zero. Or the expected error should be zero, so if m is equal to zero or the weights of kriging should be add up to one. Then the first condition the mean is known, that condition is leading us to simple kriging, if mean (m) is unknown, thus weights should be sum to 1.2.

$$E[Z(u_\alpha) - m] = 0 \text{ , (unbiased)} \qquad (3.2)$$

then

$$E[Z_{sk}^*(u)] = m = E[z(u)]$$

The mean of estimator' error of Z

$$[Z_{sk}^*(u) - Z(u)]$$

$$E[Z_{sk}^*(u) - Z(u)] = m + \sum_{\alpha=1}^{n(u)} \lambda_\alpha^{sk}(u)\left[E\left|Z(u_{\alpha)}\right| - m\right] - E[Z(u)]$$

$$= m + \sum_{\alpha=1}^{n(u)} \lambda_\alpha^{sk}(u)[m - m] - m = 0$$

Then the error variance is given by

$$\sigma_\varepsilon^2 = Var[Z^*(u) - Z(u)] = E((Z^*(u) - Z(u))^2 \qquad (3.3)$$

$$\sigma_\varepsilon^2 = E((Z^*(u))^2 + 2EZ^*(u)Z(u) - E\left(Z(u)\right)^2$$

Expanding this expression

$$E(Z^*(u))^2 = \sum_{\alpha=1}^{n(u)} \lambda_\alpha^{sk}\ C(u_\alpha - u)\ ,\ \ E(Z(u))^2 = \sum_{\beta=1}^{n(u)} \lambda_\beta^{sk}\, C(u_\beta - u)$$

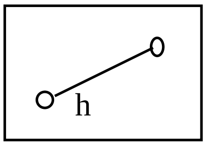
$$\sigma_\varepsilon^2 = \sum\nolimits_{\alpha=1}^{n(u)} \sum\nolimits_{\beta=1}^{n(u)}\ \lambda_\alpha^{sk}\, \lambda_\beta^{sk}\, C\big(u_\alpha - u_\beta\big) + C(u-u) - 2\sum_{\alpha=1}^{n(u)} \lambda_\alpha^{sk} C(u_\alpha - u)$$

$$= \sum_{\alpha=1}^{n(u)} \sum_{\beta=1}^{n(u)} \lambda_\alpha^{sk}\, \lambda_\beta^{sk} C\left(u_\alpha - u_\beta\right) + C\,(0) - 2\sum_{\alpha=1}^{n(u)} \lambda_\alpha^{sk}\ C\,(u_\alpha - u)$$

We can reduce the error variance, by take the derivative of our procedure in above steps.

$$Cov\big[\,Z\big(u_\alpha\big) - Z\big(u_\beta\big)\big] = C\big(u_\alpha - u_\beta\big), \alpha = 1, \dots \dots, n(u) \tag{3.4}$$

Since the spatial covariance depended just on the different vector between points:



$$1. E[Z(u+h)] = E[Z(u)]$$

$$2. Cov[Z(u+h), Z(u)] = C(h).$$

The covariance between any two points of each space depends just the vector h. The estimator variance minimal when, the first derivation equal to zero. A partial derivative of $\boldsymbol{\sigma_\varepsilon^2}$ with keep every of the weight$\boldsymbol{\lambda}$, are computed and set to zero. Then its leads us to equation that can been solved by the weights that minimize the estimate variance. But about the second derivative can be taken to show that in order to be minimum, but the maximum estimation variance is infinity. So any procedure of kriging must be a minimum [20].

$$\frac{\partial \sigma_{\varepsilon}^{2}}{\partial\ \lambda_{\alpha}^{sk}} = 0 \quad \text{where} \qquad \alpha = 1, \dots \dots, n(u).$$

$$A = \sum_{\alpha=1}^{n(u)} \sum_{\beta=1}^{n(u)} \lambda_{\alpha}^{sk} \lambda_{\alpha}^{sk} C\left(u_{\alpha} - u_{\beta}\right), \qquad B = -2 \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{sk}\ C\left(u_{\alpha} - u\right)$$

β replaced to α, since both of them are random variable as we mentioned before.

$$\sum_{\alpha=1}^{n} \lambda_{\alpha} = \lambda_{1} + \lambda_{2} + \cdots, \lambda_{n} \ = n\lambda$$

$$\frac{\partial B}{\partial \lambda_{\alpha}} = -2n\, C(u_{\alpha} - u) \quad , \quad \frac{\partial A}{\partial \lambda_{\alpha}} = \sum_{\beta=1}^{n(u)} \sum_{\alpha=1}^{n(u)} 2n\lambda_{\alpha}^{sk}\ C(u_{\alpha} - u_{\beta})$$

$$\frac{\partial B}{\partial \lambda_{\alpha}} = 2n \sum_{\beta=1}^{n(u)} \lambda_{\beta}^{sk}\, C\left(u_{\alpha} - u_{\beta}\right) - 2nC(u_{\alpha} - u) = 0$$

$$\sum_{\beta=1}^{n(u)} \lambda_{\beta}^{sk}\, C\left(u_{\alpha} - u_{\beta}\right) = C(u_{\alpha} - u) \tag{3.5}$$

In order to find the minimum or the maximum we are going to calculate the second derivative of the estimate variance.

$$\frac{\partial^{2} \sigma_{\varepsilon}^{2}}{\partial \lambda_{\alpha} \partial \lambda_{\beta}} = \sum_{\beta=1}^{n(u)} \lambda_{\beta}^{sk}\, C\left(u_{\alpha} - u_{\beta}\right) - C(u_{\alpha} - u)$$

$$\frac{\partial^{2} \sigma_{\varepsilon}^{2}}{\partial \lambda_{\alpha} \partial \lambda_{\beta}} = nC\left(u_{\alpha} - u_{\beta}\right) \tag{3.6}$$

If $C\left(u_{\alpha} - u_{\beta}\right) > 0$ minimum, or $C\left(u_{\alpha} - u_{\beta}\right) < 0$ maximum.

But in this case, the previous equation (3.5 shows us that the second derivative is always positive, thus the unique optimal weight comes from the first order condition [21].

Then the equation for simple kriging is written as:

$$\sum_{\beta=1}^{n(u)} \lambda_{\beta}^{sk}\, C\big(u_{\alpha} - u_{\beta}\big) = C(u_{\alpha} - u)\,, \alpha = 1, \dots \dots n(u).$$

Interpretation: The left side of the equation description the covariance between the locations, and the right side is the covariance between every location, and the location when an estimate is sought. The resolution of the system gives the best kriging weights λ.The procedure of simple kriging can be repeat at uniform interval moving each time the location u, a uniform grid of kriging estimator is obtained, which can be surrounds for representation as a map. Other important quantity is the best variance for every location u, it's obtained by substitution the first term of kriging system by the third term of our expression, of estimator variance σ_{ε}^{2} .Then this is the variance of simple kriging.

$$\sigma_{sk}^{2} = \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{sk}\, C(u_{\alpha} - u) + C(u\ - u) - 2\sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{sk}\, C(u_{\alpha} - u)$$

$$\sigma_{sk}^{2} = C(0) - \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{sk}\, C(u_{\alpha} - u). \qquad (3.7)$$

When sample region are scattered irregularly in space it is worthwhile to make a map of the kriging variance, same a complement to the map of kriged estimates. It gives a prediction of the varying accuracy of the kriged estimates in view of irregular effect in the information points [22].

3.2.2 Ordinary kriging

Ordinary kriging mean also be used to estimate a block value. With local second-order stationary, ordinary kriging implicit evaluates the mean in a moving neighborhood. To see this, first a kriging estimate of the local mean is set up, then a simple kriging estimator using this kriged mean is examined from the n neighborhood sample point u and add them linearly with weights λ_α.

$$Z_{ok}^{*}(u) = \sum_{\alpha=1}^{n(u)} \lambda_\alpha(u_\alpha) \qquad (3.8)$$

Clearly we must to sum up to one, since in the special case when all data value are a constant.

For unbiasedness is guaranteed with unit sum weights.

$$E[Z^{*}(u) - Z(u)] = E\left[\sum_{\alpha=1}^{n(u)} \lambda_\alpha Z(u_\alpha) - Z(u) \times \sum_{\alpha=1}^{n(u)} \lambda_\alpha \right]$$

$$\sum_{\alpha=1}^{n(u)} \lambda_\alpha = 1$$

$$E[Z^{*}(u) - Z(u)] = \sum_{\alpha=1}^{n(u)} \lambda_\alpha E[Z(u_\alpha) - Z(u)] = 0 \qquad (3.9)$$

Since the expectation of this increments are zero. The estimation variance

$$\sigma_\varepsilon^2 = Var\,[Z^{*}(u) - Z(u)]$$

is the variance of linear combination.

$$\sigma_\varepsilon^2 = E\,[(Z^*(u) - Z(u))^2] \quad (3.10)$$

$$\sigma_\varepsilon^2 = C(u-u) + \sum_{\alpha=1}^{n(u)} \sum_{\beta=1}^{n(u)} \lambda_\alpha^{ok}\,\lambda_\beta^{ok}\,C\big(u_\alpha - u_\beta\big) - 2\sum_{\alpha=1}^{n(u)} \lambda_\alpha^{ok} C(u_\alpha - u)$$

By Lagrange parameter the mechanism of Lagrange parameter help us in modifying constrained minimization problem into an unconstrained. When we tackle the minimization of σ_ε^2, as we mentioned before it was an unconstrained problem, we go toward difficulties. Trying to solve the partial first derivative equal to zero will add one equation without adding any variable. In this case we have a system of$(\mathrm{n}+1)$ equations with only n variable. The solution of such an equation is not easy to find. To avoid the previous problem, we introduce a new variable called μ into our equation σ_ε^2. μ is the Lagrange parameter.

$$\sigma_\varepsilon^2 = C(u-u) + \sum_{\alpha=1}^{n(u)} \sum_{\beta=1}^{n(u)} \lambda_\alpha^{ok}\,\lambda_\beta^{ok}\,C\big(u_\alpha - u_\beta\big) - 2\sum_{\alpha=1}^{n(u)} \lambda_\alpha^{ok} C(u_\alpha - u) + 2\mu\Big(\sum_{\alpha=1}^{n(u)} \lambda_\alpha^{ok} - 1\Big) \quad (3.11)$$

Add a variable in an equation as we did previously is delicate. We should be sure that we haven’t change the sense of our equation. But we did it well, because the variable we added is zero at the end due to his unbiasedness condition.

$$\sum_{\alpha=1}^{n(u)} \lambda_\alpha^{ok} = 1$$

$$\sum_{\alpha=1}^{n(u)} \lambda_\alpha^{ok} - 1 = 0$$

$$2\mu(\sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{ok} - 1) = 0 \tag{3.12}$$

The new term we add is all we need to move from the constrained minimization problem to the unconstrained minimization problem.

Now for the error variance of the model, we have a function of (n+1) variable, these variables are: the n weights and the Lagrange parameter. Solving the (n+1) first partial derivative equal to zero, with respect to each of our variables will lead us to a system of (n+1) equations with (n+1) variables. The case where we set the partial first derivative equal to zero with the respect given to μ will give the unbiasedness condition.

$$\sigma_{\varepsilon}^{2} = C(u-u) + \sum_{\alpha=1}^{n(u)} \sum_{\beta=1}^{n(u)} \lambda_{\alpha}^{ok} \lambda_{\beta}^{ok} C(u_{\alpha} - u_{\beta})$$
$$- 2\sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{ok} C(u_{\alpha} - u) + 2\mu(\sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{ok} - 1)$$

$$\frac{\partial \sigma_{\varepsilon}^{2}}{\partial \mu} = 2\sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{ok} - 2 \tag{3.13}$$

When we set this equation to zero, we have the unbiasedness condition.

$$\sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{ok} = 1$$

The differentiation of σ_{ε}^{2} produces$(\mathrm{n}+1)$ equations that already includes the unbiasedness condition. Therefore the solution of those $(n+1)$equations generates the set of weights which minimizes σ_{ε}^{2} under the condition that the weights sum is 1.

By this solution, we will also obtain the value of μ that later will be useful for find the resulting minimized error variance.

Minimization of the Error Variance

Now we will calculate the $(n+1)$ first partial derivate of the previous equation and it will help us to minimize the error variance. So we have (equation below):

$$\sigma_\varepsilon^2 = C(u-u) + \sum_{\alpha=1}^{n(u)} \sum_{\beta=1}^{n(u)} \lambda_\alpha^{ok}\,\lambda_\beta^{ok}\,C\big(u_\alpha - u_\beta\big)$$

$$- 2\sum_{\alpha=1}^{n(u)} \lambda_\alpha^{ok} C(u_\alpha - u) + 2\mu(\sum_{\alpha=1}^{n(u)} \lambda_\alpha^{ok} - 1)$$

$$\frac{\partial\sigma_\varepsilon^2}{\partial\lambda_\alpha^{ok}} = 2\sum_{\beta=1}^{n(u)} \lambda_\beta^{sk}\,C\big(u_\alpha - u_\beta\big) - 2C(u_\alpha - u) + 2\mu$$

When we set this equation to zero, we have the unbiasedness condition:

$$2\sum_{\beta=1}^{n(u)} \lambda_\beta^{sk}\,C\big(u_\alpha - u_\beta\big) - 2C(u_\alpha - u) + 2\mu = 0$$

$$\sum_{\beta=1}^{n(u)} \lambda_\beta^{sk}\,C\big(u_\alpha - u_\beta\big) + \mu = C(u_\alpha - u) \qquad (3.14)$$

Previously we said that setting the first partial derivate to 0 with the respect toμ lead to the unbiasedness condition. The set of weights minimizing the error variance by the constraint that they sum to 1 is satisfying the following $(n+1)$ equations:

$$\left\{\begin{array}{l}\sum_{\beta=1}^{n(u)} \lambda_{\beta}^{ok} C\left(u_\alpha - u_\beta\right) + \mu \; = C(u_\alpha - u\,) \ldots for\ \alpha = 1, \ldots n \\ \sum_{\beta=1}^{n(u)} \lambda_{\beta}^{ok} \; = 1\end{array}\right\}$$

So we get a system of ordinary kriging, which can be written as follows

$$\begin{pmatrix} C(u_1 - u_1) & \ldots & C(u_1 - u_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ C(u_n - u_1) & \ldots & C(u_n - u_n) & 1 \\ 1 & \ldots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1^{ok} \\ \vdots \\ \lambda_n^{ok} \\ \mu^{ok} \end{pmatrix} = \begin{pmatrix} C(u_1 - u) \\ \vdots \\ C(u_n - u) \\ 1 \end{pmatrix}$$

A λ b

$$A\lambda = b$$

Ordinary kriging is the best interpolator in sense that, if u is conformable with a data location, then the estimate value is conformable with data value on that point [23].

$$Z^{*}(u) = Z(u), \qquad if\ \ u = u_\alpha$$

3.3 Inverse Distance Weight (IDW)

Is the other approach for interpolation data in geostatistics is predict easy to comprehend interpolator, while you adopt IDW, can be applying a "One size fits all" assumption to your sample points.

Inverse distance weight do best for dense evenly-spaced sample point sets. It doesn't think any direction in the data, so for example if actual surface value shift more in the north-south direction, then they work in the east-west trends(due to the of slope, wind, or similar other properties). The interpolated surface will arrange out this potential base more ever IDW interpolate thinks the value of the sample points and the distance different from them the estimate cell. Sample points nearer to the cell have a wider effect one the cell is evaluate value than sample points that have more distance them others.

3.4 Spatial Variable and Variogram Function

Statistics deals with spatial random variables differing from the normal random variables, since the statistical theory of spatial (or vacuum) based on the study of the differences between the spatial variables $\boldsymbol{Z}(\boldsymbol{u})$, $\boldsymbol{Z}(\boldsymbol{u}+\boldsymbol{h})$ which is separated between them by displacement of h. For each value of the variable values spatial coordinates represent the site that point whether on the surface of the earth in the level two-dimensional (2D) or in the ground or out of the ground in a vacuum the three dimensions (3D).

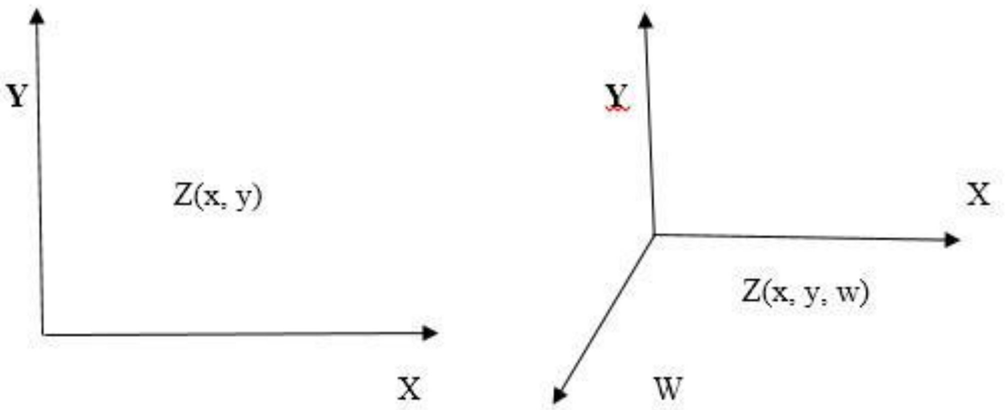


Figure 3-1: Two dimension and three dimension

Let $Z\ (u)$ be a spatial variable in a region u within D region in the Euclid: an space. Then $u \in D \subseteq R^k$ where $k = 2$ for two dimensional in space(x, y) [5], or $k = 3$ for three dimensional in space$(\boldsymbol{x}, \boldsymbol{y}, w)$. The covariance observation of spatial variable especially in spatial statistic, sometime be great or unknown which leads to Formation Values of correlation coefficients it is shown by the formula stated in [5];

$$\rho\big(Z(u), Z(u+h)\big) = \frac{\sigma_{u,u+h}}{\sigma_u \sigma_{u+h}} = \frac{Cov(Z(u), Z(\ u+h))}{\sigma_{Z(u)} \cdot \sigma_{Z(u+h)}}$$

While that value when we got it from the formula is small, then finally it gives to uncorrected interpretations and inaccurate results, and on the basis of this Kriging (1951) suggested semivariogram function given as:

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} [Z(u_i) - Z(u_i + h)]^2 \qquad (3.15)$$

[5].

Its mean square error between two existed locations or points or (Views spatial), which far from some of them to another it is displacement h. As it is known n(h) refine to pairs number points $Z(u_i) and\ Z(u_i + h)$ which makes separated between them displacement [5]. While multiply equation (3.15) by 2 change its name to variogram function, which it is written as in the following form:

$$2\gamma(h) = \frac{1}{n(h)} \sum_{i=1}^{n(h)} [Z(u_i) - Z(u_i + h)]^2$$

It is clarified by this figure.

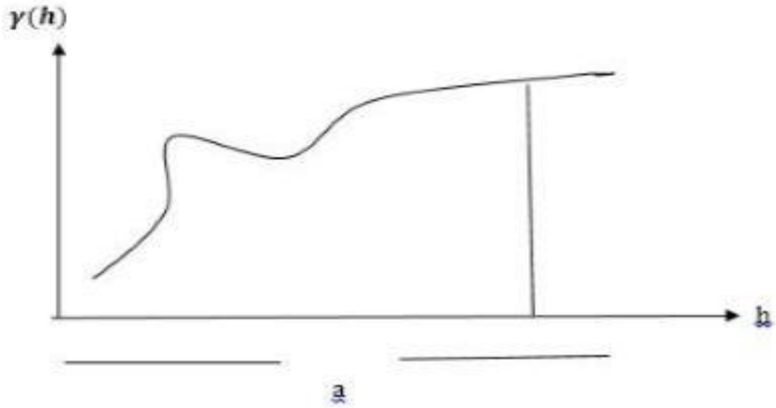


Figure 3-2: Curve of Variogram Function)

It's the displacement between points is increased then the variogram will become enlarge. This enlargement will be continuous until the height will be stable. In a certain displacement such as$(h = a)$. This displacement a is called the range then we notice the covariance, starts to be destroyed in variogram function. If is far apart the location (region) then it does not effect to variogram or it can be seen as a few amount. In the definition of semivariogram function it shows us, it is an increasing function with h and stable it will be equal to variance. It is characterized as:

1. $\gamma(0) = 0$

2. $\gamma(h) = \gamma(-h)$

Then variogram function is symmetric. Typically, as displacement h increases, mean square error between two variable $Z(u)$ and $Z(u+h)$ leads to increase too, it is called semivariogram function

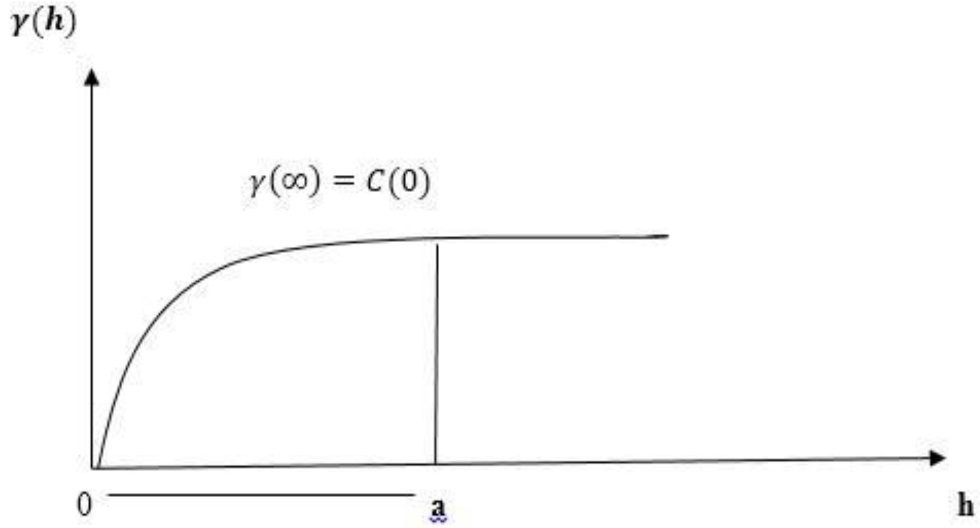


Figure 3-3: Increases Semivariogram Function

While the value $\gamma(\infty)$ is called "sill" in idiom of spatial statistics [24]. It means easily $C(0)$ is the apriori variance to random spatial function. It means $\lim_{h\to\infty}\gamma(h) = \gamma(\infty) = C(0)$ [5].

Also both covariance and variogram functions are made to specific by value of sill and range. It is called transferred and random spatial symmetric function. It is not only satisfying basic theory, but also satisfying stationary from second order: the methods of statistical for second order procedure will be considered in idiom of covariance or in mentioned that the advantage of procedure with variogram that, dislike covariance, since the mean couldn't have to be pre-estimate [25]. As in figure (4).

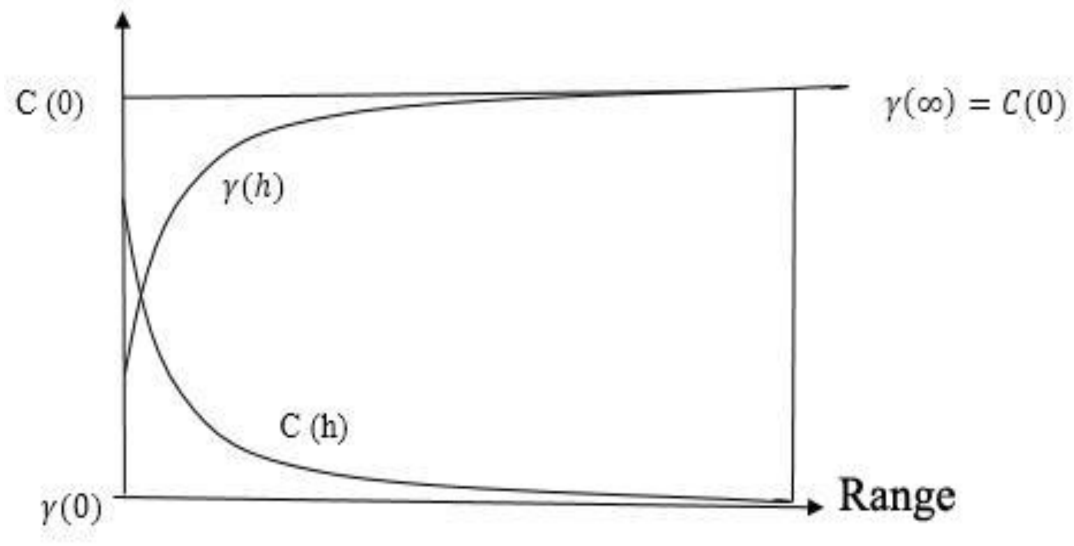


Figure 3-4: Covariance and Variogram Function

The formulation of general linear model in spatial statistics, if we assume that there is a spatial random procedure

$$[Z(u), u \in D], u = (x, y)^T, u \in R^2$$

If the spatial variable Z (u) is a linear, it can be written as;

$$Z(u) = \sum_{i=1}^{n(u)} f_i(u)\beta_i + e(u) = f^T(u) + e(u), for\ all\ u \in D$$

Where β_i unknown parameter and $f_i(u)$ known functions represents covariance location, then the spatial variable $Z(u)$ satisfies the following hypotheses as;

First hypotheses

$$E[Z(u)] = f^T(u)\beta, for\ all\ u \in D$$

Second hypotheses

$$E[Z(u+h) - Z(u)]^2 = 2\gamma(h), for\ all\ u, u+h\ \epsilon D$$

So 2γ is Isotropic 'which is mean the properties of the procedure are stable, although we apply a rotation on it'. It depends only on the distance h it isn't on the trend or the direction [26].

Third hypotheses

Covariance function exists and definite as [27].

$$Cov\ [Z(u), Z(u+h)] = C(h), for\ all\ u, u+h \in D.$$

We assume that there is n of pair spatial variable that's

$$Z(u_1), Z(u_2), Z(u_3), \dots \dots \dots \dots \dots \dots Z(u_n)$$

on the locations $u_1, u_2, u_3, \dots \dots \dots \dots . . u_n$

while the third model can be written in matrix following as:

$$Z = F\beta + e$$

$$\begin{pmatrix} Z(u_1) \\ Z(u_2) \\ \vdots \\ Z(u_3) \end{pmatrix} = \begin{pmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & y_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \\ \beta_2 \end{pmatrix} + \begin{pmatrix} e(u_1) \\ e(u_{2)} \\ \vdots \\ e(u_n) \end{pmatrix}$$

Chapter 4

KRIGING EXAMPL AND DATA ANALYSIS

4.1 Normal Distribution Data

The normal distribution is the most significant and most expand used distribution in statistics, it is often known like the bell curve, despite the tonal amount of such a bell could be fewer than pleasing. It is also called the Gaussian curve after the mathematician Karl Friedrich Gauss. The mentioned normal distribution is standard normal distribution whose mean is zero the unity variance.

4.1.1 Using Kriging

Data Normal? Check? If Not We Apply Transformation:

- Log-Normal
- Cox Box
- Arcsine

We are going to demonstrate Log-Normal, why? Because kriging needs? The density function of log normal distribution $LN[\mu, \sigma^2]$ is $\frac{1}{x\sqrt{2\pi}} e^{-(\log(x)-\mu)^2 / 2\sigma^2}$

Proof: Let us suppose that X is $LN[\mu, \sigma^2]$, then $x = e^y$ where y is $N[\mu, \sigma^2]$ then:

$$Prob(x < k) = Prob(e^y < k)$$

$$= p(y < \log(k))$$

$$= \int_{-\infty}^{\log(k)} \frac{1}{\sqrt{2\pi}\ \sigma} e^{-(y-\mu)^2 / 2\sigma^2} dy$$

Apply the transformation

$$x = e^y \quad , \quad y = \log(x) \quad , \quad dy = \frac{1}{x}\, dx$$

$$= \int_{-\infty}^{\log(k)} \frac{1}{\sqrt{2\pi}\ \sigma} e^{-(y-\mu)^2 / 2\sigma^2}\ dy$$

$$= \int_{-\infty}^{k} \frac{1}{x\sqrt{2\pi}\ \ \sigma}\ e^{-(\log(x)-\mu)^2 / 2\sigma^2}\ dx \tag{4.1}$$

[28].

We can apply a log-transformation given by the previous theorem to bring the data from a non-normal to normal distribution of this data

Table 4-1: The Values of Normal and Non-Normal Data

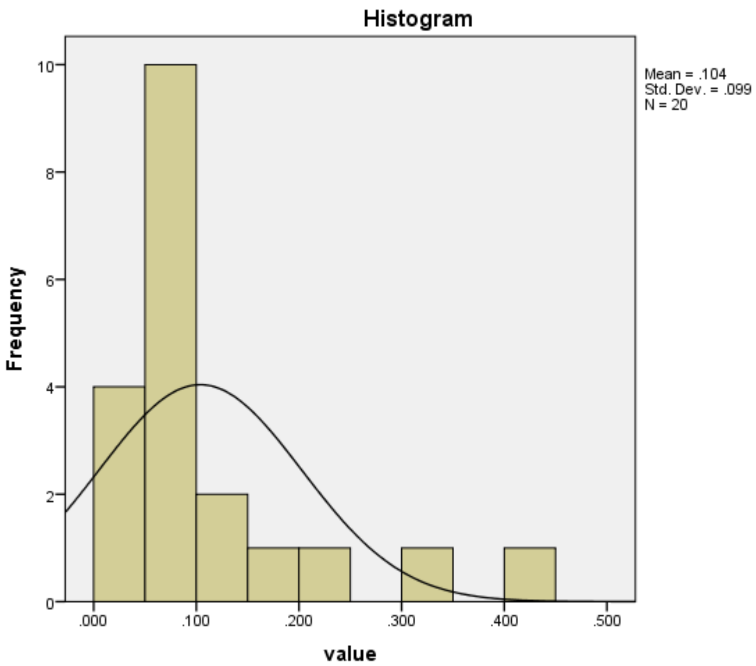
Value	logvalu
.400	-.40
.300	-.52
.230	-.64
.150	-.82
.130	-.89
.100	-1.00
.090	-1.05
.080	-1.10
.070	-1.15
.060	-1.22
.068	-1.17
.067	-1.17
.065	-1.19
.063	-1.20
.055	-1.26
.050	-1.30
.040	-1.40
.030	-1.52
.020	-1.70
.010	-2.00

Non-Normal Distribution

Statistics

value

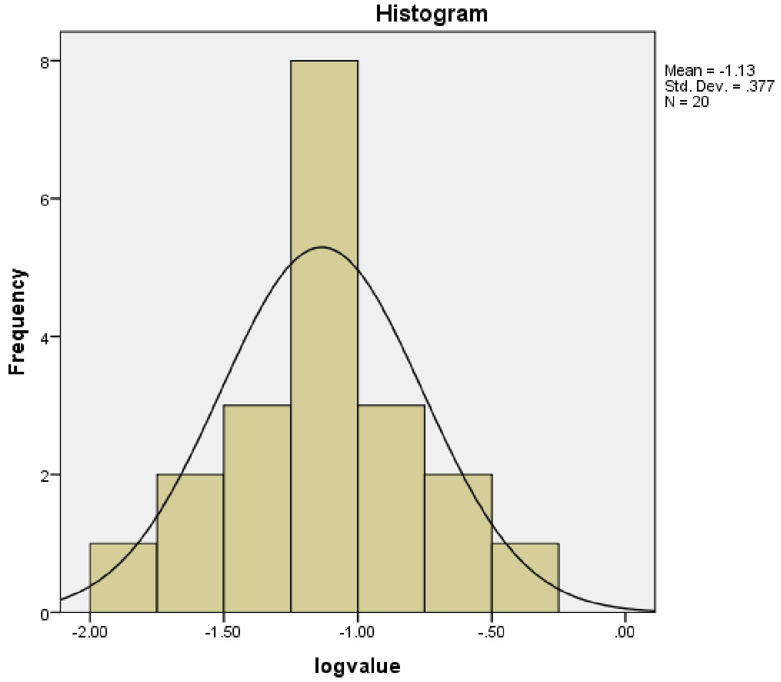
N	Valid	20
	Missing	0
Mean		.10390
Std. Error of Mean		.022078
Median		.06750
Mode		.010[a]
Std. Deviation		.098736
Variance		.010
Skewness		1.999
Std. Error of Skewness		.512
Kurtosis		3.777
Std. Error of Kurtosis		.992
Range		.390
Minimum		.010
Maximum		.400
Sum		2.078



Normal Distribution

Statistics

		Log value
N	Valid	20
	Missing	0
Mean		-1.1349
Std. Error of Mean		.08426
Median		-1.1707
Mode		-2.00^{a}
Std. Deviation		.37680
Variance		.142
Skewness		-.121
Std. Error of Skewness		.512
Kurtosis		.760
Std. Error of Kurtosis		.992
Range		1.60
Minimum		-2.00
Maximum		-.40
Sum		-22.70



4.2 Using Kriging Example

An Example

The examples which comes below includes the estimating of oil table known the elevation at three particular points. The charts that comes below reveals the three known wells and their elevation in meter. The unknown points is labeled A.

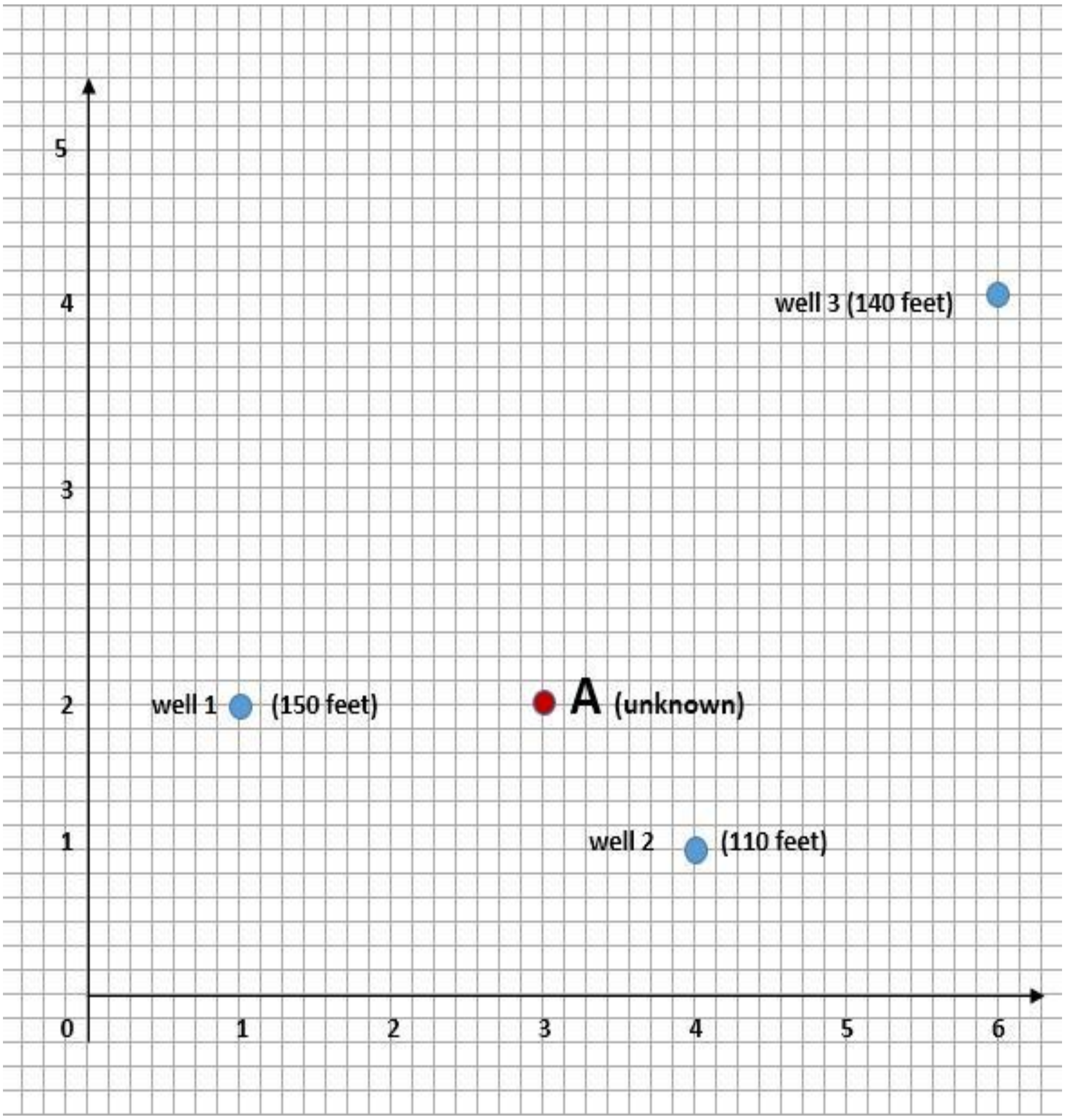


Figure 4-1: Wells Elevation with Point A

The provided information is given in the table which comes below:

Wells	X- coordinate	Y- coordinate	Oil table elevation
1	1	2	150
2	4	1	110
3	6	4	140
A	3	2	Unknown

This distances between wells and point A

	1	2	3	A
1	0	3.16	5.39	2
2	3.16	0	3.61	1.41
3	5.39	3.61	0	3.61

From an analysis of the area under testing, using a linear semivariance must be showed, yet to clarify a linear semivariance can be used there is a slop of (4) m^2 per km and a Y-intercept of 0 within a neighborhood of$(20\ \mathrm{km})$, the semivariance of distance between wells and point A, were collected$(\text{semivariance} = 4 \times \text{ distance})$.

	1	2	3	A
1	0	12.64	21.56	8
2	12.64	0	14.44	5.64
3	21.56	14.44	0	14.44

To find the weight, this equation should be solved as following example

$$
\begin{aligned}
&w_1\gamma(\mathrm{h}_{11}) + w_2\gamma(\mathrm{h}_{12}) + w_3\gamma(\mathrm{h}_{13}) + \lambda = \gamma(\mathrm{h}_{1A}) \\
&w_1\gamma(\mathrm{h}_{21}) + w_2\gamma(\mathrm{h}_{22}) + w_3\gamma(\mathrm{h}_{23}) + \lambda = \gamma(\mathrm{h}_{2A}) \\
&w_1\gamma(\mathrm{h}_{31}) + w_2\gamma(\mathrm{h}_{32}) + w_3\gamma(\mathrm{h}_{33}) + \lambda = \gamma(\mathrm{h}_{3A}) \\
&w_1 + w_2 + w_3 + 0 = 1
\end{aligned}
$$

So in matrix from these equations become:

$$
\begin{bmatrix} \gamma(h_{11}) & \gamma(h_{12}) & \gamma(h_{13}) & 1 \\ \gamma(h_{21}) & \gamma(h_{22}) & \gamma(h_{23}) & 1 \\ \gamma(h_{31}) & \gamma(h_{32}) & \gamma(h_{33}) & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \lambda \end{bmatrix} = \begin{bmatrix} \gamma(h_{1a}) \\ \gamma(h_{2A}) \\ \gamma(h_{3A}) \\ 1 \end{bmatrix}
$$

The inverse of the left hand matrix has to showed so that the weight will be determined and in the example, the multiply of inverse matrix by the right hand matrix can be show the weight, this example is shown as.

$$
\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \lambda \end{bmatrix} = \begin{bmatrix} 0.3805 \\ 0.4964 \\ 0.1232 \\ 0.9319 \end{bmatrix}
$$

The calculation of the estimate unknown value $\sigma_{E.A}$ can show as.

$$
\sigma_E = \sum_{i=1}^{n} w_i \sigma_i
$$

$$\sigma_{E.A} = w_1\sigma_1 + w_2\sigma_2 + w_3\sigma_3 \tag{4.2}$$

$$\sigma_{E.A} = 0.3805 \times 150 + 0.4964 \times 110 + 0.1232 \times 140 = 128\ meters$$

The estimation variance σ_E^2 can also be calculated

$$\sigma_E^2 = w_1\gamma(h_{1A}) + w_2\gamma(h_{2A}) + w_3\gamma(h_{3A}) + \lambda \tag{4.3}$$

$$\sigma_E^2 = 0.3805 \times 8.0 + 0.4964 \times 5.64 + 0.1232 \times 14.44 - 0.9319 \times 1.0$$

$$= 6.70\ \ \mathrm{m}^2$$

The square root of the estimation variance is the standard error s_e of the estimate is and equals:

$$s_e = \sqrt{\sigma_E^2} = \sqrt{6.70} = 2.59\ \text{meter}$$

The standard error can be adopt like confidence interval be circled to the true value if it is supposed which errors of estimation are distributed. Therefor the probability which the true elevation is inside one standard error up or down the estimated value is 68% two standard errors a way could provide a confidence of 95%. For this example the oil table elevation at point A is

$$Y_A = \sigma_{E.A} \pm s_e \times 2 \quad (4.4)$$

$$= 128.9 \pm 5.18 = [134.08,123.72] \text{ meters with 95\% probability}$$

The same procedure is used with the new unknown coordinates if the elevation of another location B has to be pointed out. Thus, the coordinates of location A (3, 2) would be changed to location B$(4, 4)$.

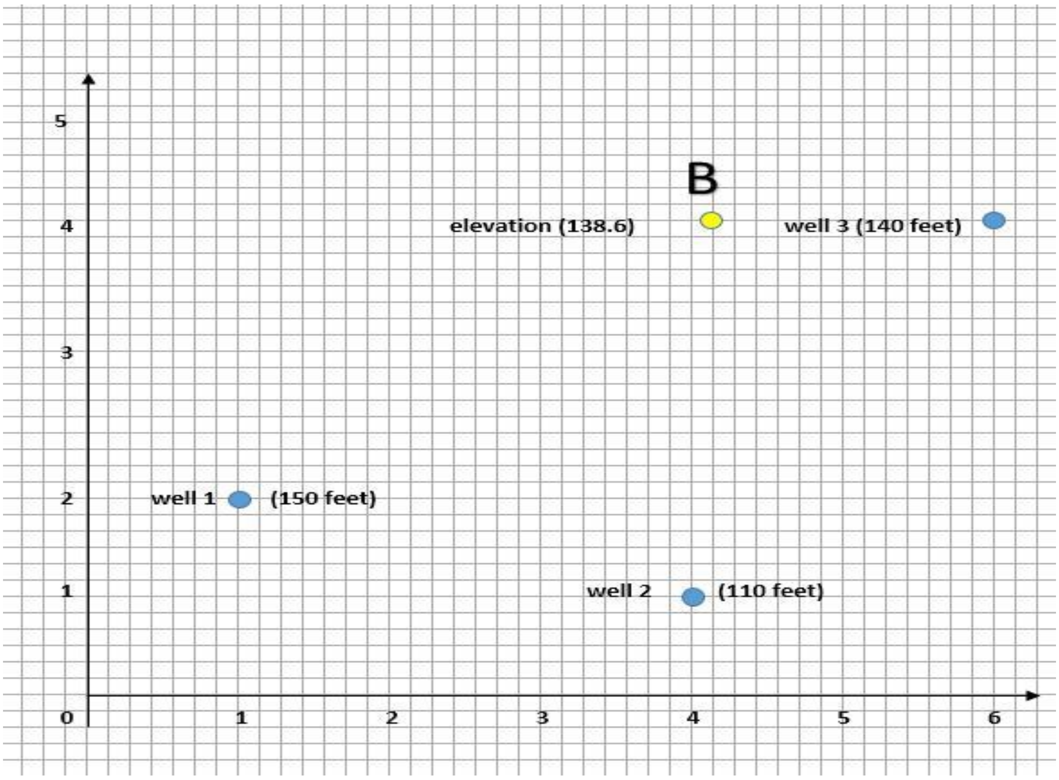


Figure 4-2: Wells Elevation with Point B

The adaptation of the new unknown location can be show as below, while the weights assigned to the three known sample

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \lambda \end{bmatrix} = \begin{bmatrix} 0.2762 \\ 0.1381 \\ 0.5857 \\ 0.0589 \end{bmatrix}$$

Using the similar procedure as previous, the oil table elevation at point B is

$$\sigma_{E.B} = w_1\sigma_1 + w_2\sigma_2 + w_3\sigma_3 \tag{4.5}$$

$$= 0.2762 \times 150 + 0.1381 \times 110 + 0.5857 \times 140 = 138.61$$

To find the estimation variance, there must be to calculate the distance between each of points, then standard error can be computing. Finally, the oil table elevation at location B is

$$Y_B = \sigma_{E.B} \pm s_e \times 2 \tag{4.6}$$

$$= 138.6 \pm 6.45 = [145.05,132.15] meters, wth\ 95\%\ probability$$

Because of the location of point B can be appeared as demonstrate below, the weight changes at each well. When point B pointing out, the elevation of the area toward which the arrow point, where the thickness of arrow shows the weight assigned for that specific well [29].

4.3 Using Inverse Distance Weight (IDW) Example

An Example:

Mathematical Form

$$Total\ weight(w) = \sum_{i=1}^{3} \lambda_i\,(u) \tag{4.7}$$

A general formula of finding an interpolated value u at a given point x based on samples $u_i = u(x_i)$ for $i = 1, 2, ..., N$ using IDW is an interpolating function.

$$u(x) = \begin{pmatrix} \frac{\sum_{i=1}^{N} \lambda_i(u) * Z_i(u)}{\sum_{i=1}^{N} \lambda_i} & ..if\ d(u, u_i) \neq 0\ for\ all\ i \\ Z_i(u) & .....if\ d(u, u_i) = 0\ for\ some\ i \end{pmatrix} \tag{4.8}$$

$$A = 100 \quad , \quad B = 160 \quad , \quad C = 200$$

$$\lambda_i = \frac{1}{d(u, u_i)^p} \tag{4.9}$$

$$\lambda_A = \frac{1}{4^2} = 0.0625$$

$$\boldsymbol{\lambda}_B = \frac{1}{3^2} = 0.1111$$

$$\boldsymbol{\lambda}_C = \frac{1}{2^2} = 0.2500$$

By equation (4.7)

$$w = 0.4236$$

$$Z(u) = \sum_{i=1}^{3} \lambda_{A,B,C} * Z_{A,B,C}$$

$$Z(u)_{A,B,C} = \sum 0.625 * 100 + 0.1111 * 160 + 0.2500 * 200$$

$$= 6.25 + 17.76 + 50.00 = 74.01$$

$$u(x) = \frac{Z(u)}{w} = \frac{74.01}{0.4236} = 175$$

IDW: closest 3 and neighbors p= 2

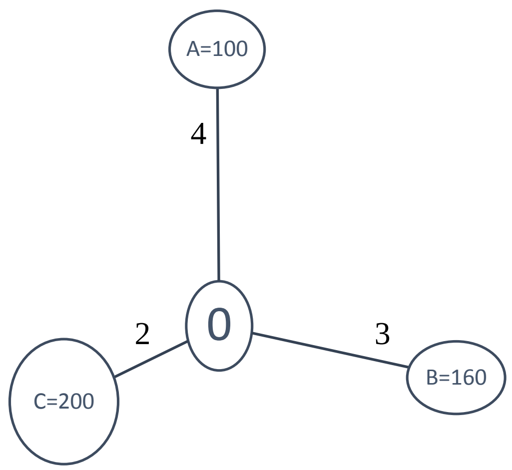
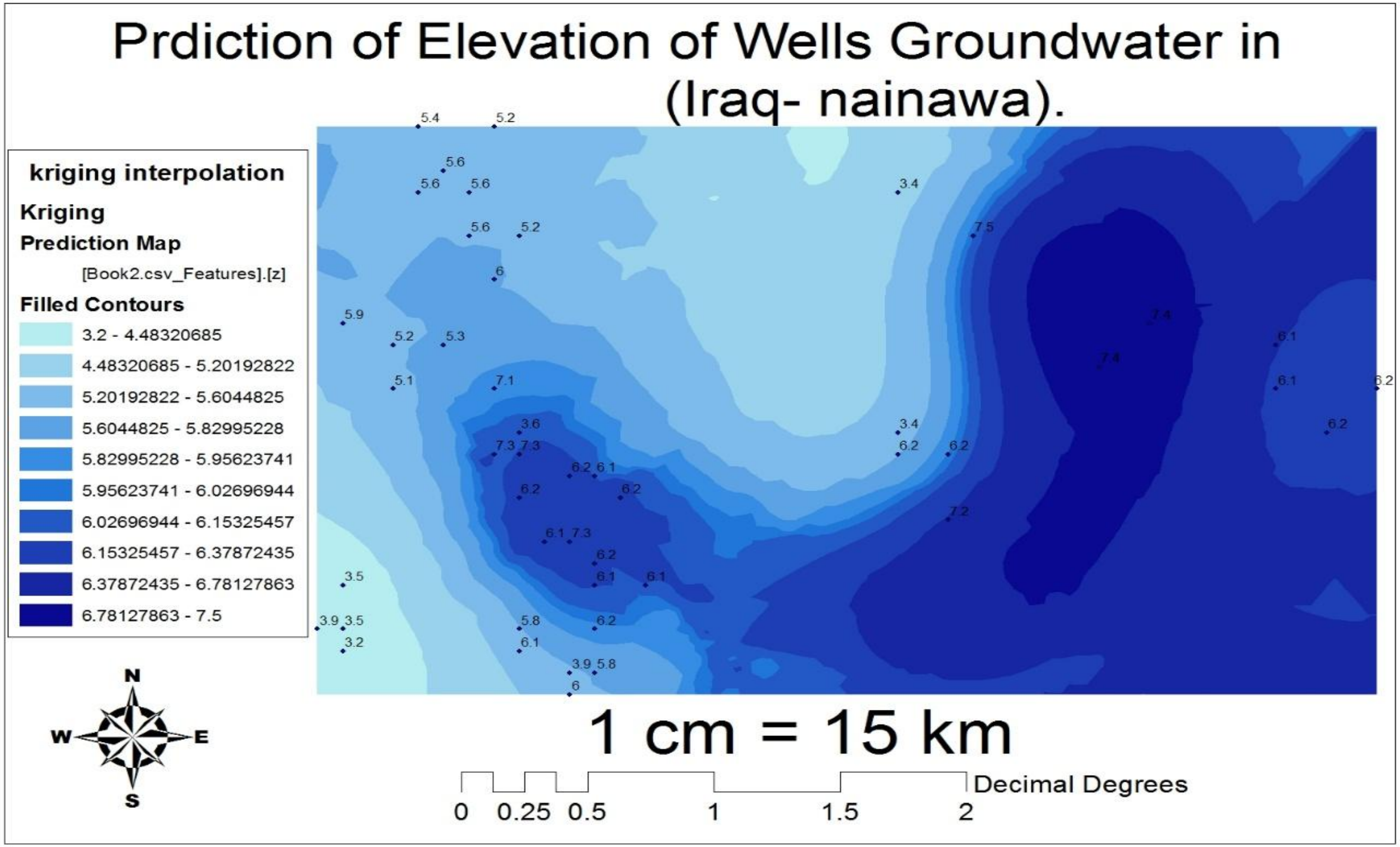
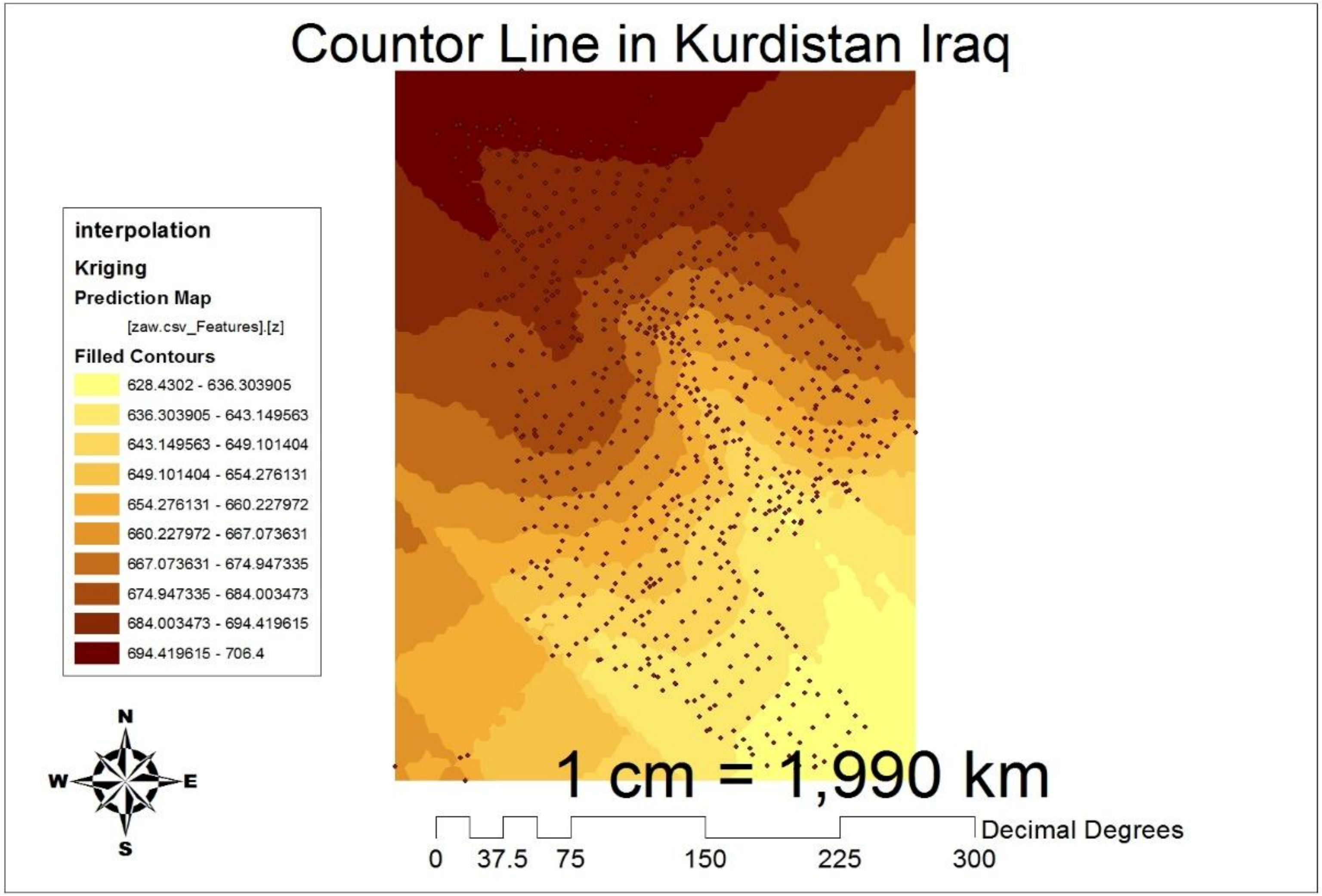


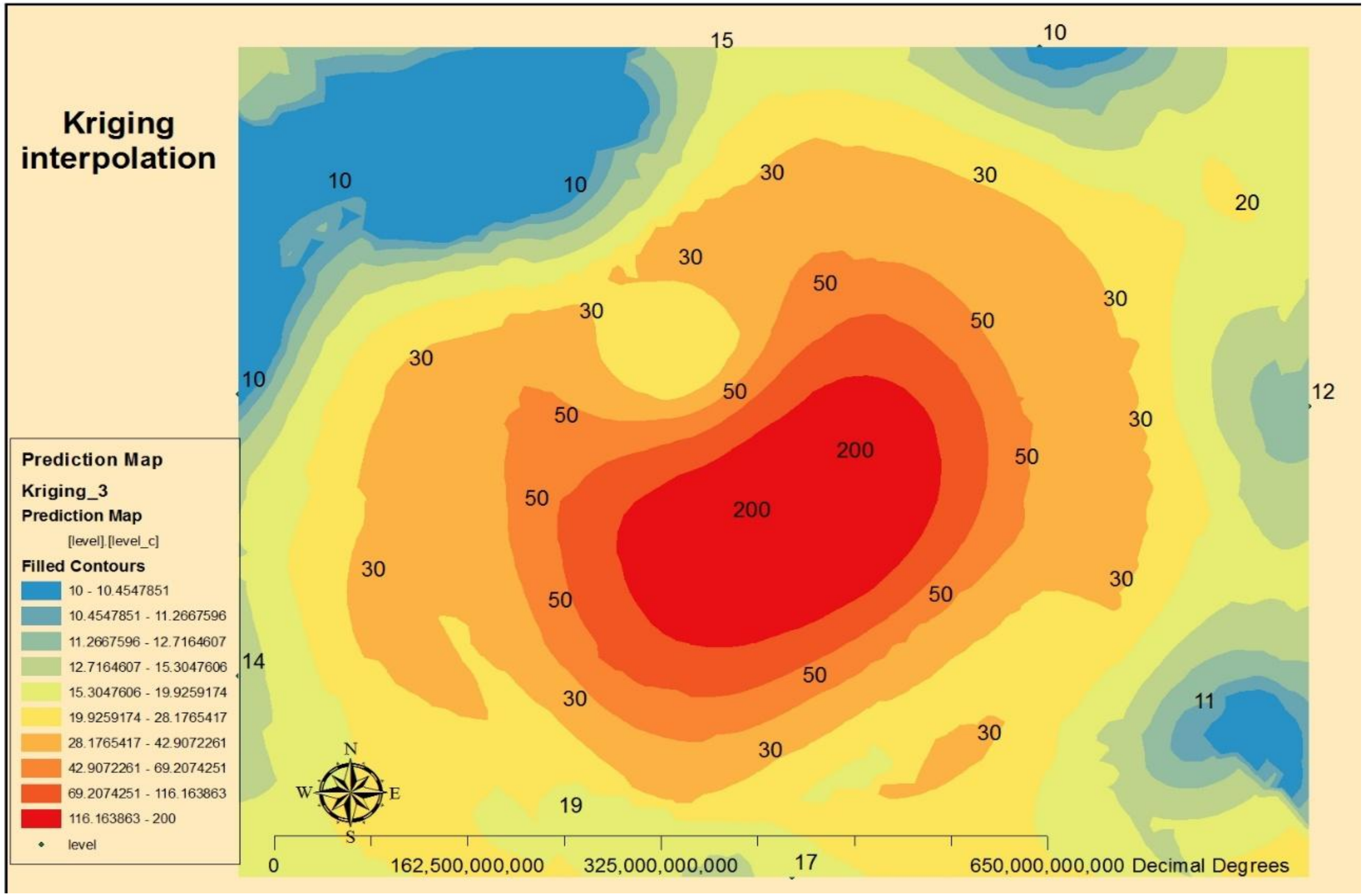
Figure 4-3: IDW: Closest 3 and Neighbors p= 2



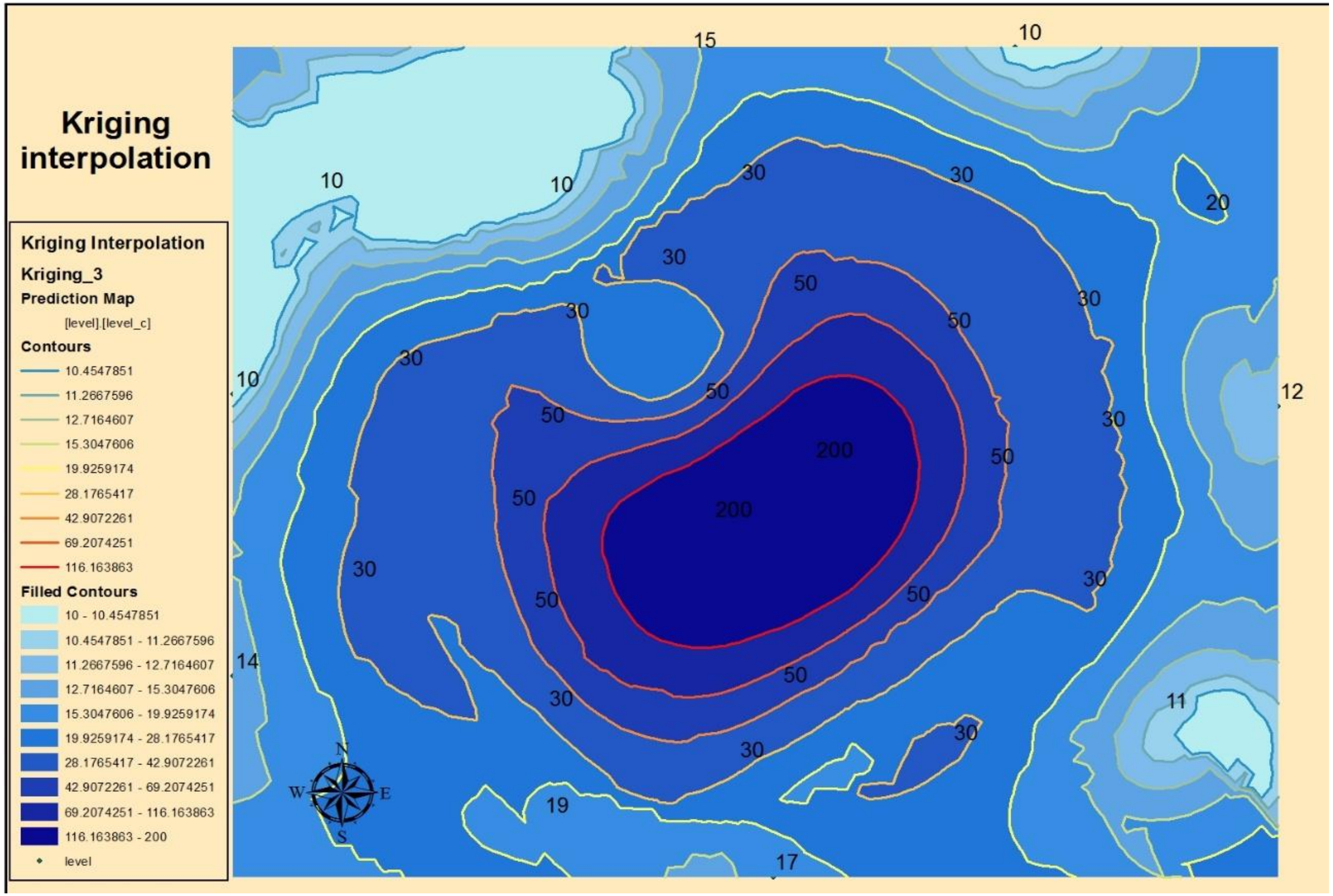
Map 4-1: Kriging Interpolatio



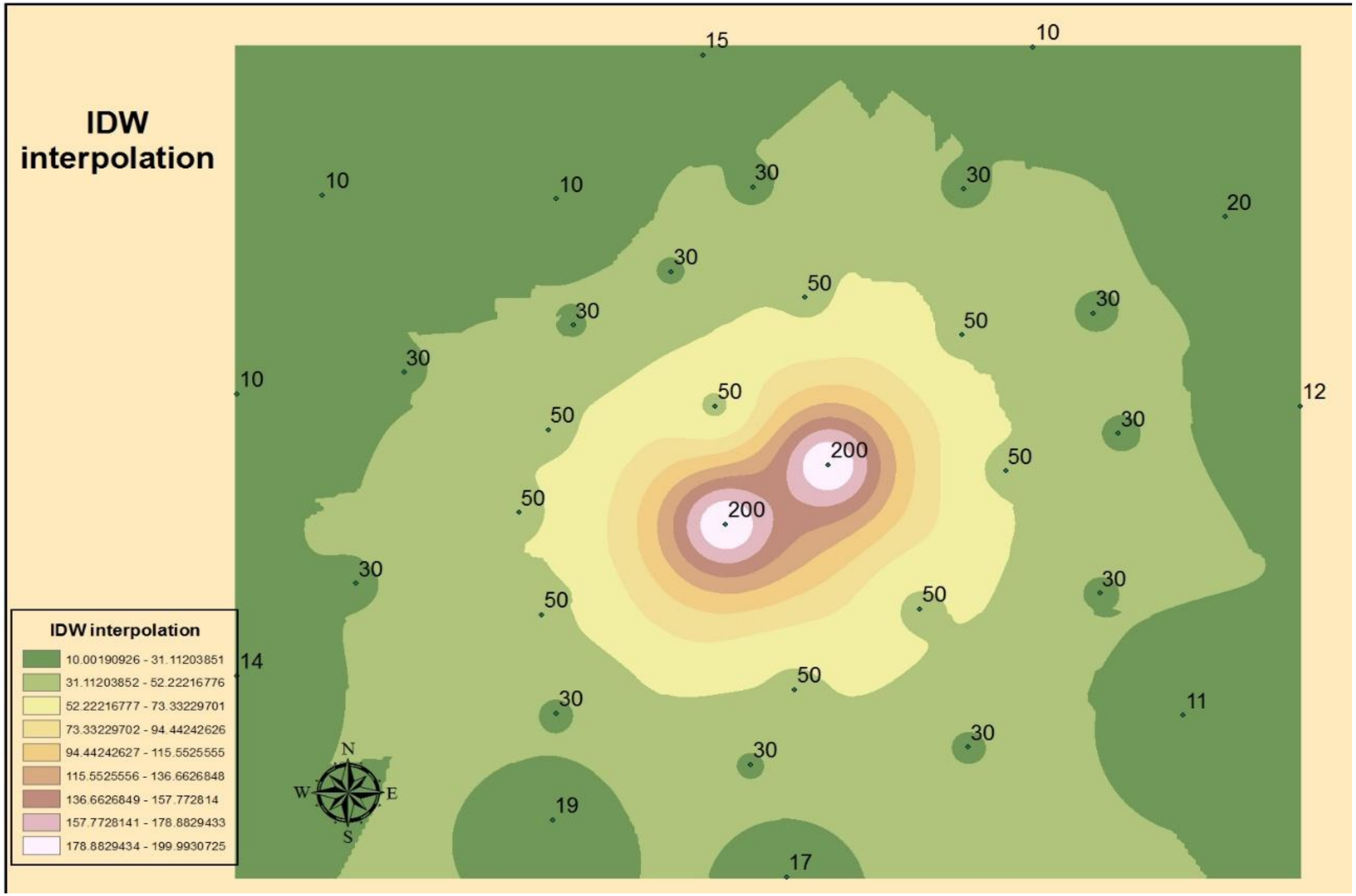
Map 4-2: Kriging Interpolation of Contours Line



Map 4-3: Kriging Interpolation of Our Spatial Random



Map 4-4: Kriging Interpolation of Our Spatial Random Points with



Map 4-5: Invers distance Weight Interpolation of Our Spatial Random Points

Table 4-2: The Value's Prediction of Our Random Spatial Points for Kriging

OBJECTID *	Shape *	Id	level_c	Included	Predicted	Standard Error
1	Point	0	15	Yes	14.999481	0.403082
2	Point	0	10	Yes	10.013628	0.269205
3	Point	0	20	Yes	19.988346	0.53714
4	Point	0	12	Yes	12.012663	0.322874
5	Point	0	11	Yes	11.018021	0.296158
6	Point	0	17	Yes	17.029174	0.457199
7	Point	0	19	Yes	19.002503	0.509536
8	Point	0	14	Yes	14.010167	0.37669
9	Point	0	10	Yes	10.013483	0.269186
10	Point	0	10	Yes	10.003487	0.268964
11	Point	0	10	Yes	10.024661	0.268897
12	Point	0	30	Yes	30.099768	0.802855
13	Point	0	30	Yes	29.930445	0.80295
14	Point	0	30	Yes	29.997181	0.800955
15	Point	0	30	Yes	29.979433	0.80031
16	Point	0	30	Yes	29.903781	0.801683
17	Point	0	30	Yes	29.959281	0.804794
18	Point	0	30	Yes	29.92215	0.794926
19	Point	0	30	Yes	30.203719	0.796347
20	Point	0	30	Yes	29.934753	0.803322
21	Point	0	30	Yes	29.924388	0.802967
22	Point	0	30	Yes	30.089558	0.78853
23	Point	0	30	Yes	29.701174	0.78478
24	Point	0	50	Yes	49.084058	1.244537
25	Point	0	50	Yes	51.160309	1.357355
26	Point	0	50	Yes	49.742734	1.325341
27	Point	0	50	Yes	50.200114	1.337228
28	Point	0	50	Yes	50.130914	1.335332
29	Point	0	50	Yes	50.210495	1.342582
30	Point	0	50	Yes	50.135282	1.332187
31	Point	0	50	Yes	49.550201	1.284665
32	Point	0	50	Yes	50.935166	1.290196
33	Point	0	200	Yes	198.039155	5.289897
34	Point	0	200	Yes	198.519845	5.299688

Table 4-3: The Value's Prediction of Our Random Spatial Points for Invers Distance Weights

OBJECTID *	Shape *	Id	level_c	Included	Predicted	Error
1	Point	0	15	Yes	15	0
2	Point	0	10	Yes	10	0
3	Point	0	20	Yes	20	0
4	Point	0	12	Yes	12	0
5	Point	0	11	Yes	11	0
6	Point	0	17	Yes	17	0
7	Point	0	19	Yes	19	0
8	Point	0	14	Yes	14	0
9	Point	0	10	Yes	10	0
10	Point	0	10	Yes	10	0
11	Point	0	10	Yes	10	0
12	Point	0	30	Yes	30	0
13	Point	0	30	Yes	30	0
14	Point	0	30	Yes	30	0
15	Point	0	30	Yes	30	0
16	Point	0	30	Yes	30	0
17	Point	0	30	Yes	30	0
18	Point	0	30	Yes	30	0
19	Point	0	30	Yes	30	0
20	Point	0	30	Yes	30	0
21	Point	0	30	Yes	30	0
22	Point	0	30	Yes	30	0
23	Point	0	30	Yes	30	0
24	Point	0	50	Yes	50	0
25	Point	0	50	Yes	50	0
26	Point	0	50	Yes	50	0
27	Point	0	50	Yes	50	0
28	Point	0	50	Yes	50	0
29	Point	0	50	Yes	50	0
30	Point	0	50	Yes	50	0
31	Point	0	50	Yes	50	0
32	Point	0	50	Yes	50	0
33	Point	0	200	Yes	200	0
34	Point	0	200	Yes	200	0

Chapter 5

CONCOLUSION

In chapter three selected data were used for interpolation in both Invers Distance weight (IDW) method and also in simple kriging method. It was shown by these analysis that the kriging approach is the interpolation method, as mentioned in the theory. The error in kriging is almost zero compared the error in Invers Distance weight (IDW). Where the error is the difference between actual and the predicted data. The ArcGIS software is very powerful for such analysis by using its tools.

REFERENCES

[1] P. P. D. Ribeiro, A Package for Geostatistical Analysis, Springer, 2001, 2001.

[2] B. D. Ripley, Spatial statistics, New Jersey: John Wiley & Sons, Inc., Hoboken, , 2004.

[3] P. J. R. Peter J. Diggle, Model-based Geostatistics, Brazil: Springer, March. 26.2007.

[4] N. A. C. Cressie, Statistics for spatial data, J. Wiley .1993, 20 Nov 2007.

[5] N. Cressie, Statistic, for Spatial Data, Second Edition., New York: John Wiley & Sons, 1993.

[6] C. Meyer, "Evaluating Water Quality using Spatial Interpolation Methods, Pinellas County, Florida, U.S.A.," in ESRI International User Conference, San Diego, California, 2006.

[7] P.-H. C. H.-L. C. I.-F. M. D. P. H. H. Hsiao-Hui Chen, "Spatial Interpolation for Identifying Soil Contamination Area," in The Esri International User Conference, 2006.

[8] J. S. K. C. Frederic bedsrd, "Spatial Analysis of Crime Data-City of Montreal, Canada," in The Esri International User Conference, 2006.

[9] K. M.-P. S. G. K. T. D. V. Angela Bucciarelli, "Spatial Analysis of Drug Overdose Deaths, New York City, 1994-2003," in Esri International User Conference, 2006.

[10] G. Barras, "Spatial Data Convicts North Korean Drug Traffickers," in The Esri International User Conference, 2006.

[11] B. W. B. D. Brian Ward, "A Spatial Analysis of the NCAA Basketball Tournament," in The Esri International User Conference, 2006.

[12] S. Wetherbee, "Spatial Data Analysis of Salmonella in Dairy Farms," in The Esri International User Conference, 2006.

[13] R. Resources, "Spatial Data Modeling to Support National Flood Risk Assessment," in The Esri International User Conference, 2006.

[14] M. b. N. L. T. S. T. N. Niedja Lemos, "A Spatial Decision Support System for Urban Draining Systems," in The Esri International User Conference, 2006.

[15] P. L. Roux, "A Spatial Enable Event Management System, City of Cape Town," in The Esri International User Conference, 2006.

[16] D. M. Armstrong, Basic linear geostatistics, Fontainebleau / France: Springer-Verlag Berlin Heidelberg 1998, May 1998.

[17] P. A. Burrough, Principles of Geographical Information Systems for Land Resources Assessment, New York: Oxford University Press. , 1986.

[18] C. Roth2, "Mathematical Geology," Is Lognormal Kriging Suitable for Local estimation, vol. Vol. 30, no. No. 8, 1998.

[19] K. D. E. B. Ganguli, "SRS.FS," 21 08 updated in 2007. [Online]. Available: http://webcam.srs.fs.fed.us/impacts/ozone/spatial/kriging.shtml.

[20] C. V. D. Michael J. Pyrcz, Geostatistical Reservoir Modeling, USA: Oxford uneversity press, 2014.

[21] P. D. P. G. M. F. Alan E. Gelfand, Handbook of Spatial Statistics, USA: Chapman & Hall/CRC Handbooks of Modern Statistical Methods, 2010.

[22] D. Wackernagel, Multivariate Geostatistics, Franc-paris: Springer-Verlag Berlin Heidelberg , 2003.

[23] E. H. I. Srivastava, "Chapter 12 ,Ordinary kriging," in An introduction to Applied Geostatistics, Oxford University Press, 1989, pp. 278-332.

[24] O. Dubrule, Geostatistics for seismic data integration in earth models, America: Paul Weimer, 2003.

[25] X. G. Carlo Gaetan, Spatial Statistics and Modeling, Springer Series in Statistics, 2009.

[26] A. G. Journel, Geostatistics: Models and tools for the earth sciences, Stanford, California: Kluwer Academic Publishers-Plenum Publishers,Mathematical Geology,Volume 18, Issue 1, 1986-01-01.

[27] P. Goovaerts, Geostatistics for Natural Resources Evaluation, Oxford: Oxford University Press, 1997.

[28] J. Norstad, " DEFINITIONS AND SUMMARY OF THE PROPOSITIONS," in The Normal and LognormalDistributions, j-norstad@northwestern.edu, February 2, 1999 , Updated: November 3, 2011, p. http://www.norstad.org.

[29] J. C. Davis, Statistics and Data Analysis in Geology, John Wiley & Sons, 1973.

[30] D. M. Lane, "Introduction to Normal Distributions," in Mathematical Reviews, American Mathematical Society, 1987.