

# Multivariate Statistics and Geostatistical Analyses of Metal Elements in Soil of Waste Disposal Site in Khulna

by

**Sanjida Khair**



Department of Civil Engineering  
Khulna University of Engineering & Technology  
Khulna 9203, Bangladesh

**November 2017**

# **Multivariate Statistics and Geostatistical Analyses of Metal Elements in Soil of Waste Disposal Site in Khulna**

by

**Sanjida Khair**

A thesis submitted in partial fulfillment of the requirements for the Degree of  
Master of Science in Civil Engineering

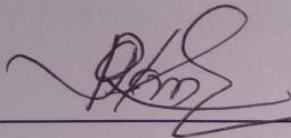


Khulna University of Engineering & Technology  
Khulna 9203, Bangladesh

**November 2017**

## Declaration


This is to certify that this thesis work entitled "Multivariate Statistics and Geostatistical Analyses of Metal Elements in Soil of Waste Disposal Site in Khulna" has been carried out by Sanjida Khair in the Department of Civil Engineering, Khulna University of Engineering & Technology, Khulna, Bangladesh. The above research work or any part of this work has not been submitted anywhere for the award of any degree or diploma.



---

Dr. Md. Rafizul Islam

Associate Professor



---

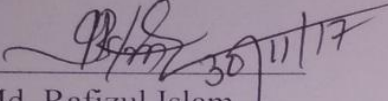
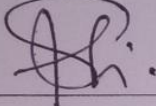
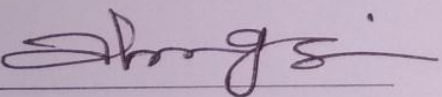
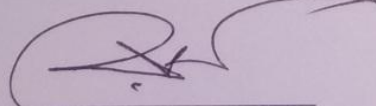
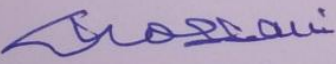
Sanjida Khair

Roll No. 1501558

## Approval

This is to certify that this thesis work submitted by *Sanjida Khair* entitled "*Multivariate Statistics and Geostatistical Analyses of Metal Elements in Soil of Waste Disposal Site in Khulna*" has been approved by the Board of Examiners for the partial fulfillment of the requirements for the degree of *Master of Science in Civil Engineering* in the Department of *Civil Engineering*, Khulna University of Engineering & Technology, Khulna, Bangladesh in November 2017.

### Board of Examiners

1.   
Dr. Md. Rafizul Islam  
Associate Professor  
Department of Civil Engineering  
Khulna University of Engineering & Technology, Khulna  
Chairman  
(Supervisor)
2.   
Dr. Md. Shahjahan Ali  
Professor and Head  
Department of Civil Engineering  
Khulna University of Engineering & Technology, Khulna  
Member
3.   
Dr. Muhammed Alamgir  
Professor  
Department of Civil Engineering  
Khulna University of Engineering & Technology, Khulna  
Member
4.   
Dr. Md. Rokonzaman  
Professor  
Department of Civil Engineering  
Khulna University of Engineering & Technology, Khulna  
Member
5.   
Dr. Md. Delwar Hossain  
Professor  
Department of Civil Engineering  
Bangladesh University of Engineering and Technology, Dhaka  
Member  
(External)



*Affectionately dedicated*

*To*

*my parents, husband, son  
for their love, support and encouragement  
and*

*Specially to my supervisor  
for his unconditional support and encouragement,  
without whom this research would not have been possible*

## **Acknowledgement**

Glory be to Almighty Allah (swt) for His endless care and enablement to start and conclude this thesis at this time. I would also like to thank my family for their unending encouragement and support throughout this research work. This research was made possible with their love, support and most of all patience.

I am very thankful to my supervisor Dr. Md. Rafizul Islam, Associate Professor, Department of Civil Engineering, KUET for providing valuable time, guidance and support in all the perspectives while doing this thesis work. I gained immensely from his fatherly advice and wealth of research experience without which I would not have been able to conclude this thesis timeously. It was a great privilege and honor to work and study under his guidance. His co-operation throughout the thesis periods have brought me to the point of successfully completing this thesis work. It would have been impossible for me to complete this thesis without his continuous support. It was a valuable experience to work under the guidance of Dr. Rafizul, a person having highly disciplined life style, well behaved even in adverse situations and capable of giving right and prompt decision with philosophical opinion.

Special thank goes to Prof. Dr. Md. Delwar Hossain, Department of Civil Engineering, BUET, Bangladesh for his constructive comments and corrections concerning my first draft made me realize just how far I was straying from the right path. Again I am also indebted to Dr. Delwar who spontaneously took on the responsibility of reviewing my thesis and also provided me with much good advices and suggestions.

I am very thankful to Prof. Dr. Muhammed Alamgir, Prof. Dr. Muhammad Harunur Rashid and Prof. Dr. Rokonuzzaman, Department of Civil Engineering, KUET, for being my thesis committee members as well as their encouragement and help for this thesis. They provided valuable comments about my work and sharing their great ideas. Their patience, understanding and insight have greatly influenced the outcome of this thesis.

Last but not least, special appreciation goes to my beloved husband Sm. Arifur Rahman for being patient while I was busy at work, his encouragement and moral support during the entire course. I thank to my son Jurayj Al Jarir for his patience and endurance throughout my absence during this research period.

Sanjida Khair

## Abstract

In disposal site, municipal solid waste (MSW) decomposes and produces three components of solid; liquid (leachate) and landfill gas. Open dumping facilities release huge quantity of metal elements into the surrounding water bodies, underlying soil layer and atmosphere. The main focus of this study was to identify the correlations of metal elements in soil, possible sources of their generation and contamination, distribution of metal elements spatially as well as the level of contamination of soil of a disposal site. To these endeavors, sixty soil samples were collected at a depth of 0-30 cm from the existing ground surface from a selected waste disposal site at Rajbandh, Khulna, Bangladesh. These study periods covered both the dry season (March to May, 2016) and rainy seasons (June to August, 2016). In the laboratory, the relevant metal elements of Aluminium (Al), Arsenic (As), Barium (Ba), Calcium (Ca), Cadmium (Cd), Cobalt (Co), Chromium (Cr), Copper (Cu), Iron (Fe), Mercury (Hg), Potassium (K), Manganese (Mn), Sodium (Na), Nickel (Ni), Lead (Pb), Antimony (Sb), Scandium (Sc), Strontium (Sr), Titanium (Ti), Vanadium (V) and Zinc (Zn) were measured through the standard test methods. Furthermore, the spatial distribution of metal elements in soil is necessitated to explore their extents. Implementation of interpolation techniques can provide better prediction of the distribution of metal elements in soil with least prediction errors.

To these attempts, conventional statistics such as K-S test, S-W test and normal QQ plot was performed using SPSS. Based on normal QQ plot, it was observed that almost all the metal elements were distributed normally except As in soil for both the dry and rainy seasons. The multivariate statistics such as Pearson's correlation, principal component analysis (PCA) and agglomerative hierarchical clustering (AHC) was performed using XLSTAT. Results of Pearson's correlation revealed that almost all the metal elements were strongly correlated with each other indicating these metal elements derived from the same generation sources. In addition, results of PCA and AHC depicted that almost all the metal elements in soil derived from anthropogenic/human activities; least number of metal elements from natural sources as well as from both the natural and anthropogenic sources.

In this study, Geostatistical analysis such as inverse distance weighting (IDW), local polynomial interpolation (LPI), radial basis functions (RBF) and ordinary kriging (OK) was performed using ArcGIS. Furthermore, the cross validations of IDW with power 1 to 5, LPI with order 1 to 3, RBF with five kernel functions as well as OK with eleven distinct models were performed to select the best fitted model for further assessing the performance of these interpolation techniques. Based on least value of MAPE, IDW with power 1 to 5 and RBF with different kernel functions showed comparatively more accurate prediction than that of LPI and OK. Based on RI, IDW1 showed best performance followed by OK. Lastly, based on all indices (MAPE, RI, etc.), IDW1 showed the best technique for all metal elements.

Moreover, produced prediction surface for all the interpolation techniques showed most of the contaminated hotspots was found near the centre of disposal site for all the metal elements. Semivariogram showed that almost all the metal elements were moderately correlated spatially and least number of metal elements were strongly and weakly correlated. In this study, a network model was developed by ANN to predict and depict the validity of observed concentration of metal elements obtained from laboratory based MSE and R-value. It was found that the predicted values from ANN were almost same as obtained from laboratory. Finally, it can be concluded that this study will so guide for more efficient prediction of spatial distribution of metal elements with their possible generation sources, and to remedial measures regarding the contamination of soil of the waste disposal sites all over the world.



## Contents

	<b>PAGE</b>
Title Page	i
Declaration	ii
Certificate of Research	iii
Acknowledgement	v
Abstract	vi
Contents	vii
List of Tables	xii
List of Figures	xiv
List of Illustrations	xv
Nomenclature	xvii
<b>CHAPTER I</b>	
<b>INTRODUCTION</b>	<b>1-13</b>
1.1 Background Information	1
1.2 Soil and Metal Element	6
1.3 Problem Statement and Justification	7
1.4 Objectives of the Study	9
1.5 Contribution of Knowledge	9
1.6 Significance of the Study	10
1.7 Scope and Limitations	10
1.8 Outline of the Thesis	11
<b>CHAPTER II</b>	
<b>LITERATURE REVIEW</b>	<b>14-44</b>
2.1 Introduction	14
2.2 Municipal Solid Waste	14
2.3 Dumping Facilities of MSW	15
2.3.1 Sanitary landfill	15
2.3.2 Open dumping	16
2.3.3 Overall impacts of MSW disposal site	17
2.4 Context of Heavy Metal	19
2.4.1 Providence of heavy metals in soil and environment	19

2.4.2	Behavior of heavy metals in soil	20
2.4.2.1	Accumulation	21
2.4.2.2	Solubility and mobility	22
2.4.2.3	Bioavailability	22
2.4.2.4	Toxicity	23
2.4.3	Consequences of MSW, soil and heavy metal in disposal site	23
2.5	Present Scenario of MSW Management and Disposal Facilities in Bangladesh	24
2.6	Distribution Pattern of Heavy Metals in Soil	26
2.7	Conventional Statistics	28
2.7.1	Normality test	28
2.7.1.1	Kolmogorov-Smirnov test	28
2.7.1.2	Shapiro- Wilk test	29
2.7.1.3	Normal Quantile-Quantile plot	29
2.7.2	Descriptive statistics	30
2.7.3	Case studies of conventional statistics in soil	30
2.8	Multivariate Statistics	30
2.8.1	Pearson's correlation	31
2.8.2	Principal component analysis	31
2.8.3	Agglomerative hierarchical clustering	32
2.8.4	Case studies of multivariate statistics for assessment of heavy metals in soil	32
2.9	Geostatistical Analysis	34
2.9.1	Inverse distance weighting	36
2.9.2	Local polynomial interpolation	36
2.9.3	Radial basis functions	37
2.9.4	Ordinary kriging	38
2.9.5	Semi-variogram and its estimation	38
2.9.6	Case studies of application of geostatistics of heavy metals in soil	40
2.10	Artificial Neural Network	43

<b>CHAPTER III</b>	<b>RESEARCH METHODOLOGY</b>	<b>45-64</b>
3.1	Introduction	45
3.2	Description of Study Site	47
3.3	Location and Soil Conditions of Waste Disposal Site	47
3.4	Soil Sampling	48
3.5	Laboratory Investigations	50
3.5.1	Acid digestion	50
3.5.2	Analysis of heavy metals with AAS	51
3.6	Descriptive Statistics	51
3.6.1	Normality test	51
3.6.1.1	Shapiro-Wilk test	51
3.6.1.2	Kolmogorov–Smirnov test	52
3.6.1.3	Normal QQ plot	53
3.6.2	Conventional statistics	53
3.7	Multivariate Statistics	53
3.7.1	Pearson correlation	54
3.7.2	Principal component analysis	54
3.7.3	Agglomerative hierarchical clustering	56
3.8	Geostatistical Analysis	56
3.8.1	Inverse distance weighting	57
3.8.2	Local polynomial interpolation	58
3.8.3	Radial basis functions	59
3.8.4	Ordinary kriging	60
3.9	Assessment of Method Performance	61
3.10	Artificial Neural Network	62
<b>CHAPTER IV</b>	<b>RESULTS AND DISCUSSION</b>	<b>65-141</b>
4.1	General	65
4.2	Descriptive Statistics	66
4.2.1	Normality test	66
4.2.2	Conventional statistics	69
4.2.2.1	Dry season	69
4.2.2.2	Rainy season	71

4.2.2.3	Seasonal comparison of the concentration of metal elements	72
4.3	Multivariate Statistics	72
4.3.1	Pearson correlation	75
4.3.2	Principal component analysis	76
4.3.3	Cluster analysis: Agglomerative hierarchical clustering	86
4.4	Cluster Analysis: Artificial Neural Network	89
4.5	Geostatistical Analysis	92
4.5.1	ArcGIS approach	93
4.5.2	Deterministic methods	93
4.5.2.1	Inverse distance weighting	93
4.5.2.2	Local polynomial interpolation	101
4.5.2.3	Radial basis functions	111
4.5.3	Geostatistical methods	119
4.5.3.1	Ordinary kriging	119
4.5.3.2	Semivariogram	124
4.6	Assessment of Method Performance	127
4.6.1	Mean absolute percentage error	127
4.6.2	Goodness of prediction	129
4.6.3	Relative improvement	130
4.7	Artificial Neural Network	131
4.7.1	Mean standard error	132
4.7.2	Regression coefficient	134
4.7.3	Error histogram	138
4.8	Concluding Remarks	140
<b>CHAPTER V</b>	<b>COMPARISON AND VALIDATION</b>	<b>142-148</b>
5.1	General	142
5.2	Comparison with previous studies	142
<b>CHAPTER VI</b>	<b>CONCLUSION AND RECOMMENDATIONS</b>	<b>149-150</b>
6.1	Conclusion	149
6.2	Recommendation for Further Studies	150

<b>REFERENCES</b>		151-155
<b>ANNEX</b>		
<b>Annex-A</b>	Screenshots of steps of all performed analysis	156-190
<b>Annex-B</b>	Representation of normal Q-Q plot and self-organizing map (SOM) of metal elements using XLSTAT	191-208
<b>Annex-C</b>	Cross validation results and spatial distribution of Metal elements using IDW	209-227
<b>Annex-D</b>	Cross validation results and spatial distribution of metal elements using LPI	228-261
<b>Annex-E</b>	Cross validation results and spatial distribution of metal elements using RBF's	262-281
<b>Annex-F</b>	Cross validation results and spatial distribution of metal elements using OK	282-300
<b>Annex-G</b>	Performance assessment of predicted data obtained ANN	301-314

## LIST OF TABLES

Table No.	Description	Page
2.1	Several researches on application of multivariate statistics for assessment of soil pollution	33
2.2	Several researches on application of geostatistics for assessment of soil pollution	42
4.1	Normality test of metal elements in soil of waste disposal site	66
4.2	Descriptive statistics of metal elements in soil for dry season (n=40)	70
4.3	Descriptive statistics of metal elements in soil for rainy season (n=20)	71
4.4	Correlation analysis and coefficients for the metal elements in dry season	73
4.5	Correlation analysis and coefficients for the metal elements in rainy season	74
4.6	PCA of metal elements in soil for dry and rainy seasons	76
4.7	All explained variables and factors derived using the orthogonal varimax rotation method of dry season	80
4.8	All explained variables and factors derived using the orthogonal varimax rotation method of rainy season	81
4.9	Results of cluster analysis for dry season	88
4.10	Results of cluster analysis for rainy season	89
4.11	Cross validation of IDW for Cd	94
4.12	Cross validation of IDW with different power for Ni, Pb and Zn	96
4.13	Cross validation of LP for Cd	102
4.14	Cross validation of LP for Ni	104
4.15	Cross validation of LP for Pb	107
4.16	Cross validation of LP for Zn	109
4.17	Results of cross validation of RBF with different kernal functions for Cd	113
4.18	Cross validation of RBF for Ni,Pb and Zn	115
4.19	Cross validation results of OK for different models of Cd in soil	120
4.20	Cross validation results of OK for different models of Ni in soil	120
4.21	Cross validation results of OK for different models of Pb in soil	123
4.22	Cross validation results of OK for different models of Zn in soil	123

4.23	Fitted parameters of the theoretical variogram model for heavy metal	124
4.24	Semivariogram parameters of best fitted models for metal elements	125
4.25	Performance assessment of interpolation methods based on MAPE	128
4.26	Performance assessment of interpolation methods based on G-value	129
4.27	Performance assessment of interpolation methods based on RI	130
4.28	Performance assessment of nntool model	135

## LIST OF FIGURES

Figure No.	Description	Page
2.1	Methods of MSW dumping facilities (a) Sanitary landfills (b) Open dumping	16
2.2	Sources of land pollution in MSW disposal site	17
2.3	Impacts of MSW disposal.	18
2.4	Contamination of soil from disposal site	21
2.5	Relation between MSW, soil and heavy metals in disposal site	24
2.6	Present scenario of MSW disposal site in Dhaka (Matuail landfill)	25
2.7	Amount of MSW collected in Khulna city (in ton/day)	26
2.8	Contaminant plume in (a) groundwater; (b) soil	27
2.9	Spatial prediction implies application of a prediction algorithm to an array of grid nodes (point 'a point spatial prediction). The results are then displayed using a raster map	35
2.10	Typical diagram of semivariogram	39
2.11	Typical ANN architecture model	43
3.4	SOM input space (3-D) and output space (2- D). Red dots signify the input patterns while the blue dots shows the connected SOM neurons	63
3.5	Artificial neural network, identical to the human neural network	63
5.1	Spatial distribution maps of soil heavy metal concentrations	144
5.2	Estimated Ordinary Kriging concentration maps for Pb, Cu, Cr, Zn, Mn, Cd, As and Hg (mg/kg)	146



## LIST OF ILLUSTRATIONS

Illustration No.	Description	Page
1.1	Outline and relations between five chapters of this study	13
3.1	Location map of Rajbandh at Khulna city of Bangladesh	46
3.2	Map showing of soil sampling locations in waste disposal site	48
3.3	Flow diagram of research methodology in this study	49
4.1	Normal QQ plots for dry season (a) Cd; (b) Ni; (c) Pb and (d) Zn	67
4.2	Normal QQ Plots in the rainy season (a) Cd; (b) Ni; (c) Pb and (d) Zn	68
4.3	Scree plot of the PCs (a) dry season; (b) rainy season	78
4.4	Correlation circle for metal elements in soil for dry season	82
4.5	Correlation circle for metal elements in soil for rainy season	83
4.6	Correlation plot of soil sampling locations of dry season	84
4.7	Biplot of soil sampling locations and concentrations of metal elements of dry season	84
4.8	Correlation plot of soil sampling locations of rainy season	85
4.9	Biplot plot of the locations for metal elements (rainy) in soil of waste disposal site	85
4.10	Dendrogram for metal elements in soil during (a) dry and (b) rainy season	87
4.11	Plot of SOM sample hits for metal elements in (a) dry season (b) rainy season	90
4.12	Plot of SOM weight plane for metal elements of (a) Cd, (b) Ni, (c) Pb and (d) Zn in dry season	91
4.13	Plot of SOM Weight Plane for metal elements of (a) Cd, (b) Ni, (c) Pb and (d) Zn, in rainy season	92
4.14	Spatial distribution of Cd in soil using IDW with power of 1- 5	95
4.15	Spatial distribution of Ni in soil using IDW with power of 1- 5	97
4.16	Spatial distribution of Pb in soil using IDW with power of 1- 5	98
4.17	Spatial distribution of Zn in soil using IDW with power of 1- 5	100
4.18	Spatial distribution of Cd in soil using LP with order of 1- 3	103
4.19	Spatial distribution of Ni in soil using LP with order of 1- 3	105

4.20	Spatial distribution of Pb in soil using LP with order of 1- 3	108
4.21	Spatial distribution of Zn in soil using LP with order of 1- 3	110
4.22	Spatial distribution of Cd in soil using RBF's	112
4.23	Spatial distribution of Ni in soil using RBF's	114
4.24	Spatial distribution of Pb in soil using RBF's	116
4.25	Spatial distribution of Cd in soil using RBF's	118
4.26	Spatial distribution of Cd, Ni, Pb and Zn in soil using OK for best fitted model	121
4.27	Semi variogram of metal elements	126
4.28	Performance plot of Cd	133
4.29	Performance plot of Ni	133
4.30	Performance plot of Pb	134
4.31	Performance plot of Zn	134
4.32	Regression coefficient of Cd in soil	136
4.33	Regression coefficient of Ni in soil	136
4.34	Regression coefficient of Pb in soil	137
4.35	Regression coefficient of Zn in soil	137
4.36	Histogram plot of Cd	139
4.37	Histogram plot of Ni	139
4.38	Histogram plot of Pb	139
4.39	Histogram plot of Zn	139

## Nomenclature

All the notation and symbols are defined where they first appear in the text or figures. For convenience, the more frequently used symbols and their meanings are listed below.

AAS	: Atomic absorption spectroscopy
AHC	: Agglomerative Hierarchical Clustering
Al	: Aluminum
ANN	: Artificial Neural Network
As	: Arsenic
ASPE	: Average Standard Prediction Error
ASTM	: American Society for Testing and Materials
Ba	: Barium
BCSIR	: Bangladesh Council of Scientific and Industrial Research
BH	: Borehole
Ca	: Calcium
CBO's	: Community Based Organizations
Cd	: Cadmium
Co	: Cobalt
Cr	: Chromium
CRS	: Completely Regularized Spline
CV	: Coefficient of Variation
DPHE	: Department of Public Health Engineering
EPA	: Environmental Protection Agency
ESRI	: Environmental Systems Research Institute
Fe	: Iron
GIS	: Geographic Information System
GPS	: Global Positioning System
G-value	: Goodness of Prediction
H <sub>2</sub> SO <sub>4</sub>	: Sulfuric acid
HCl	: Hydrochloric acid
HClO <sub>4</sub>	: Hypochlorite acid
Hg	: Mercury

HNO <sub>3</sub>	: Nitric acid
IDW	: Inverse Distance Weighting
IMQ	: Inverse multiquadric
JICA	: Japanese International Cooperation Agency
K	: Potassium
KCC	: Khulna City Corporation
Kg	: Kilogram
K-S test	: Kolmogorov–Smirnov test
KUET	: Khulna University of Engineering & Technology
LP	: Local Polynomial
MAPE	: Mean Absolute Percentage Error
mg	: Milligram
mL	: Milliliter
Mn	: Manganese
MPE	: Mean Prediction Error
MSE	: Mean Standard Error
MSPE	: Mean Standard Prediction Error
MSW	: Municipal solid waste
MQ	: Multiquadric
Na	: Sodium
Ni	: Nickel
NGO's	: Non-Government Organizations
nntool	: Neural Network tool
OK	: Ordinary Kriging
Pb	: Lead
PCA	: Principal Component Analysis
Q-Q plot	: Quantile Quantile plot
RBF's	: Radial Basis Functions
RI	: Relative Improvement
RMSE	: Root Mean Square Error
RMSSE	: Root Mean Standard Square Error
R-value	: Regression coefficient
Sb	: Antimony

Sc	: Scandium
SD	: Standard Deviation
SOM	: Self-Organizing Map
SPSS	: Statistical Package for the Social Science
ST	: Spline with tension
S-W test	: Shapiro Wilk test
Ti	: Titenium
TPS	: Thin Plate Spline
UNEP	: United Nations Environment Programme
UNFPA	: United Nations Fund for Population Activities
V	: Vanadium
Zn	: Zinc
°C	: Degree celsius
%	: Percentage
:	: is to

# **Multivariate Statistics and Geostatistical Analyses of Metal Elements in Soil of Waste Disposal Site in Khulna**

by

**Sanjida Khair**

A thesis submitted in partial fulfillment of the requirements for the Degree of  
Master of Science in Civil Engineering



Khulna University of Engineering & Technology  
Khulna 9203, Bangladesh

**November 2017**

## **Declaration**

This is to certify that this thesis work entitled “Multivariate Statistics and Geostatistical Analyses of Metal Elements in Soil of Waste Disposal Site in Khulna” has been carried out by Sanjida Khair in the Department of Civil Engineering, Khulna University of Engineering & Technology, Khulna, Bangladesh. The above research work or any part of this work has not been submitted anywhere for the award of any degree or diploma.

---

Dr. Md. Rafizul Islam

Associate Professor

---

Sanjida Khair

Roll No. 1501558

## Approval

This is to certify that this thesis work submitted by Sanjida Khair entitled “Multivariate Statistics and Geostatistical Analyses of Heavy Metals in Soil of Waste Disposal Site in Bangladesh” has been approved by the Board of Examiners for the partial fulfillment of the requirements for the degree of Master of Science in Civil Engineering, Khulna University of Engineering & Technology, Khulna, Bangladesh in the month of November 2017.

### Board of Examiners

1. \_\_\_\_\_  
Dr. Md. Rafizul Islam  
Associate Professor  
Department of Civil Engineering  
Khulna University of Engineering & Technology, Khulna  
Chairman  
(Supervisor)
2. \_\_\_\_\_  
Dr. Muhammad Harunur Rashid  
Professor and Head  
Department of Civil Engineering  
Khulna University of Engineering & Technology, Khulna  
Member
3. \_\_\_\_\_  
Dr. Muhammed Alamgir  
Professor  
Department of Civil Engineering  
Khulna University of Engineering & Technology, Khulna  
Member
4. \_\_\_\_\_  
Dr. Md. Rokonuzzaman  
Professor  
Department of Civil Engineering  
Khulna University of Engineering & Technology, Khulna  
Member
5. \_\_\_\_\_  
Dr. Md. Delwar Hossain  
Professor  
Department of Civil Engineering  
Bangladesh University of Engineering & Technology, Dhaka  
Member  
(External)



*Affectionately dedicated*

*To*

*my parents, husband, son  
for their love, support and encouragement  
and*

*Specially to my supervisor  
for his unconditional support and encouragement,  
without whom this research would not have been possible*

## **Acknowledgement**

Glory be to Almighty Allah (swt) for His endless care and enablement to start and conclude this thesis at this time. I would also like to thank my family for their unending encouragement and support throughout this research work. This research was made possible with their love, support and most of all patience.

I am very thankful to my supervisor Dr. Md. Rafizul Islam, Associate Professor, Department of Civil Engineering, KUET for providing valuable time, guidance and support in all the perspectives while doing this thesis work. I gained immensely from his fatherly advice and wealth of research experience without which I would not have been able to conclude this thesis timeously. It was a great privilege and honor to work and study under his guidance. His co-operation throughout the thesis periods have brought me to the point of successfully completing this thesis work. It would have been impossible for me to complete this thesis without his continuous support. It was a valuable experience to work under the guidance of Dr. Rafizul, a person having highly disciplined life style, well behaved even in adverse situations and capable of giving right and prompt decision with philosophical opinion.

Special thank goes to Prof. Dr. Md. Delwar Hossain, Department of Civil Engineering, BUET, Bangladesh for his constructive comments and corrections concerning my first draft made me realize just how far I was straying from the right path. Again I am also indebted to Dr. Delwar who spontaneously took on the responsibility of reviewing my thesis and also provided me with much good advices and suggestions.

I am very thankful to Prof. Dr. Muhammed Alamgir, Prof. Dr. Muhammad Harunur Rashid and Prof. Dr. Rokonuzzaman, Department of Civil Engineering, KUET, for being my thesis committee members as well as their encouragement and help for this thesis. They provided valuable comments about my work and sharing their great ideas. Their patience, understanding and insight have greatly influenced the outcome of this thesis.

Last but not least, special appreciation goes to my beloved husband Sm. Arifur Rahman for being patient while I was busy at work, his encouragement and moral support during the entire course. I thank to my son Jurayj Al Jarir for his patience and endurance throughout my absence during this research period.

Sanjida Khair

## Abstract

In disposal site, municipal solid waste (MSW) decomposes and produces three components of solid; liquid (leachate) and landfill gas. Open dumping facilities release huge quantity of metal elements into the surrounding water bodies, underlying soil layer and atmosphere. The main focus of this study was to identify the correlations of metal elements in soil, possible sources of their generation and contamination, distribution of metal elements spatially as well as the level of contamination of soil of a disposal site. To these endeavors, sixty soil samples were collected at a depth of 0-30 cm from the existing ground surface from a selected waste disposal site at Rajbandh, Khulna, Bangladesh. These study periods covered both the dry season (March to May, 2016) and rainy seasons (June to August, 2016). In the laboratory, the relevant metal elements of Aluminium (Al), Arsenic (As), Barium (Ba), Calcium (Ca), Cadmium (Cd), Cobalt (Co), Chromium (Cr), Copper (Cu), Iron (Fe), Mercury (Hg), Potassium (K), Manganese (Mn), Sodium (Na), Nickel (Ni), Lead (Pb), Antimony (Sb), Scandium (Sc), Strontium (Sr), Titanium (Ti), Vanadium (V) and Zinc (Zn) were measured through the standard test methods. Furthermore, the spatial distribution of metal elements in soil is necessitated to explore their extents. Implementation of interpolation techniques can provide better prediction of the distribution of metal elements in soil with least prediction errors.

To these attempts, conventional statistics such as K-S test, S-W test and normal QQ plot was performed using SPSS. Based on normal QQ plot, it was observed that almost all the metal elements were distributed normally except As in soil for both the dry and rainy seasons. The multivariate statistics such as Pearson's correlation, principal component analysis (PCA) and agglomerative hierarchical clustering (AHC) was performed using XLSTAT. Results of Pearson's correlation revealed that almost all the metal elements were strongly correlated with each other indicating these metal elements derived from the same generation sources. In addition, results of PCA and AHC depicted that almost all the metal elements in soil derived from anthropogenic/human activities; least number of metal elements from natural sources as well as from both the natural and anthropogenic sources.

In this study, Geostatistical analysis such as inverse distance weighting (IDW), local polynomial interpolation (LPI), radial basis functions (RBF) and ordinary kriging (OK) was performed using ArcGIS. Furthermore, the cross validations of IDW with power 1 to 5, LPI with order 1 to 3, RBF with five kernel functions as well as OK with eleven distinct models were performed to select the best fitted model for further assessing the performance of these interpolation techniques. Based on least value of MAPE, IDW with power 1 to 5 and RBF with different kernel functions showed comparatively more accurate prediction than that of LPI and OK. Based on RI, IDW1 showed best performance followed by OK. Lastly, based on all indices (MAPE, RI, etc.), IDW1 showed the best technique for all metal elements.

Moreover, produced prediction surface for all the interpolation techniques showed most of the contaminated hotspots was found near the centre of disposal site for all the metal elements. Semivariogram showed that almost all the metal elements were moderately correlated spatially and least number of metal elements were strongly and weakly correlated. In this study, a network model was developed by ANN to predict and depict the validity of observed concentration of metal elements obtained from laboratory based MSE and R-value. It was found that the predicted values from ANN were almost same as obtained from laboratory. Finally, it can be concluded that this study will so guide for more efficient prediction of spatial distribution of metal elements with their possible generation sources, and to remedial measures regarding the contamination of soil of the waste disposal sites all over the world.

# Contents

	<b>PAGE</b>
Title Page	i
Declaration	ii
Certificate of Research	iii
Acknowledgement	v
Abstract	vi
Contents	vii
List of Tables	xii
List of Figures	xiv
List of Illustrations	xv
Nomenclature	xvii
<b>CHAPTER I</b>	
<b>INTRODUCTION</b>	<b>1-13</b>
1.1 Background Information	1
1.2 Soil and Metal Element	6
1.3 Problem Statement and Justification	7
1.4 Objectives of the Study	9
1.5 Contribution of Knowledge	9
1.6 Significance of the Study	10
1.7 Scope and Limitations	10
1.8 Outline of the Thesis	11
<b>CHAPTER II</b>	
<b>LITERATURE REVIEW</b>	<b>14-44</b>
2.1 Introduction	14
2.2 Municipal Solid Waste	14
2.3 Dumping Facilities of MSW	15
2.3.1 Sanitary landfill	15
2.3.2 Open dumping	16
2.3.3 Overall impacts of MSW disposal site	17
2.4 Context of Heavy Metal	19
2.4.1 Providence of heavy metals in soil and environment	19

2.4.2	Behavior of heavy metals in soil	20
2.4.2.1	Accumulation	21
2.4.2.2	Solubility and mobility	22
2.4.2.3	Bioavailability	22
2.4.2.4	Toxicity	23
2.4.3	Consequences of MSW, soil and heavy metal in disposal site	23
2.5	Present Scenario of MSW Management and Disposal Facilities in Bangladesh	24
2.6	Distribution Pattern of Heavy Metals in Soil	26
2.7	Conventional Statistics	28
2.7.1	Normality test	28
2.7.1.1	Kolmogorov-Smirnov test	28
2.7.1.2	Shapiro- Wilk test	29
2.7.1.3	Normal Quantile-Quantile plot	29
2.7.2	Descriptive statistics	30
2.7.3	Case studies of conventional statistics in soil	30
2.8	Multivariate Statistics	30
2.8.1	Pearson's correlation	31
2.8.2	Principal component analysis	31
2.8.3	Agglomerative hierarchical clustering	32
2.8.4	Case studies of multivariate statistics for assessment of heavy metals in soil	32
2.9	Geostatistical Analysis	34
2.9.1	Inverse distance weighting	36
2.9.2	Local polynomial interpolation	36
2.9.3	Radial basis functions	37
2.9.4	Ordinary kriging	38
2.9.5	Semi-variogram and its estimation	38
2.9.6	Case studies of application of geostatistics of heavy metals in soil	40
2.10	Artificial Neural Network	43

<b>CHAPTER III</b>	<b>RESEARCH METHODOLOGY</b>	45-64
3.1	Introduction	45
3.2	Description of Study Site	47
3.3	Location and Soil Conditions of Waste Disposal Site	47
3.4	Soil Sampling	48
3.5	Laboratory Investigations	50
3.5.1	Acid digestion	50
3.5.2	Analysis of heavy metals with AAS	51
3.6	Descriptive Statistics	51
3.6.1	Normality test	51
3.6.1.1	Shapiro-Wilk test	51
3.6.1.2	Kolmogorov–Smirnov test	52
3.6.1.3	Normal QQ plot	53
3.6.2	Conventional statistics	53
3.7	Multivariate Statistics	53
3.7.1	Pearson correlation	54
3.7.2	Principal component analysis	54
3.7.3	Agglomerative hierarchical clustering	56
3.8	Geostatistical Analysis	56
3.8.1	Inverse distance weighting	57
3.8.2	Local polynomial interpolation	58
3.8.3	Radial basis functions	59
3.8.4	Ordinary kriging	60
3.9	Assessment of Method Performance	61
3.10	Artificial Neural Network	62
<b>CHAPTER IV</b>	<b>RESULTS AND DISCUSSION</b>	65-141
4.1	General	65
4.2	Descriptive Statistics	66
4.2.1	Normality test	66
4.2.2	Conventional statistics	69
4.2.2.1	Dry season	69
4.2.2.2	Rainy season	71

4.2.2.3	Seasonal comparison of the concentration of metal elements	72
4.3	Multivariate Statistics	72
4.3.1	Pearson correlation	75
4.3.2	Principal component analysis	76
4.3.3	Cluster analysis: Agglomerative hierarchical clustering	86
4.4	Cluster Analysis: Artificial Neural Network	89
4.5	Geostatistical Analysis	92
4.5.1	ArcGIS approach	93
4.5.2	Deterministic methods	93
4.5.2.1	Inverse distance weighting	93
4.5.2.2	Local polynomial interpolation	101
4.5.2.3	Radial basis functions	111
4.5.3	Geostatistical methods	119
4.5.3.1	Ordinary kriging	119
4.5.3.2	Semivariogram	124
4.6	Assessment of Method Performance	127
4.6.1	Mean absolute percentage error	127
4.6.2	Goodness of prediction	129
4.6.3	Relative improvement	130
4.7	Artificial Neural Network	131
4.7.1	Mean standard error	132
4.7.2	Regression coefficient	134
4.7.3	Error histogram	138
4.8	Concluding Remarks	140
<b>CHAPTER V</b>	<b>COMPARISON AND VALIDATION</b>	<b>142-148</b>
5.1	General	142
5.2	Comparison with previous studies	142
<b>CHAPTER VI</b>	<b>CONCLUSION AND RECOMMENDATIONS</b>	<b>149-150</b>
6.1	Conclusion	149
6.2	Recommendation for Further Studies	150

<b>REFERENCES</b>		151-155
<b>ANNEX</b>		
<b>Annex-A</b>	Screenshots of steps of all performed analysis	156-190
<b>Annex-B</b>	Representation of normal Q-Q plot and self-organizing map (SOM) of metal elements using XLSTAT	191-208
<b>Annex-C</b>	Cross validation results and spatial distribution of Metal elements using IDW	209-227
<b>Annex-D</b>	Cross validation results and spatial distribution of metal elements using LPI	228-261
<b>Annex-E</b>	Cross validation results and spatial distribution of metal elements using RBF's	262-281
<b>Annex-F</b>	Cross validation results and spatial distribution of metal elements using OK	282-300
<b>Annex-G</b>	Performance assessment of predicted data obtained ANN	301-314



## LIST OF TABLES

Table No.	Description	Page
2.1	Several researches on application of multivariate statistics for assessment of soil pollution	33
2.2	Several researches on application of geostatistics for assessment of soil pollution	42
4.1	Normality test of metal elements in soil of waste disposal site	66
4.2	Descriptive statistics of metal elements in soil for dry season (n=40)	70
4.3	Descriptive statistics of metal elements in soil for rainy season (n=20)	71
4.4	Correlation analysis and coefficients for the metal elements in dry season	73
4.5	Correlation analysis and coefficients for the metal elements in rainy season	74
4.6	PCA of metal elements in soil for dry and rainy seasons	76
4.7	All explained variables and factors derived using the orthogonal varimax rotation method of dry season	80
4.8	All explained variables and factors derived using the orthogonal varimax rotation method of rainy season	81
4.9	Results of cluster analysis for dry season	88
4.10	Results of cluster analysis for rainy season	89
4.11	Cross validation of IDW for Cd	94
4.12	Cross validation of IDW with different power for Ni, Pb and Zn	96
4.13	Cross validation of LP for Cd	102
4.14	Cross validation of LP for Ni	104
4.15	Cross validation of LP for Pb	107
4.16	Cross validation of LP for Zn	109
4.17	Results of cross validation of RBF with different kernal functions for Cd	113
4.18	Cross validation of RBF for Ni,Pb and Zn	115
4.19	Cross validation results of OK for different models of Cd in soil	120
4.20	Cross validation results of OK for different models of Ni in soil	120
4.21	Cross validation results of OK for different models of Pb in soil	123
4.22	Cross validation results of OK for different models of Zn in soil	123

4.23	Fitted parameters of the theoretical variogram model for heavy metal	124
4.24	Semivariogram parameters of best fitted models for metal elements	125
4.25	Performance assessment of interpolation methods based on MAPE	128
4.26	Performance assessment of interpolation methods based on G-value	129
4.27	Performance assessment of interpolation methods based on RI	130
4.28	Performance assessment of nntool model	135

## LIST OF FIGURES

Figure No.	Description	Page
2.1	Methods of MSW dumping facilities (a) Sanitary landfills (b) Open dumping	16
2.2	Sources of land pollution in MSW disposal site	17
2.3	Impacts of MSW disposal.	18
2.4	Contamination of soil from disposal site	21
2.5	Relation between MSW, soil and heavy metals in disposal site	24
2.6	Present scenario of MSW disposal site in Dhaka (Matuail landfill)	25
2.7	Amount of MSW collected in Khulna city (in ton/day)	26
2.8	Contaminant plume in (a) groundwater; (b) soil	27
2.9	Spatial prediction implies application of a prediction algorithm to an array of grid nodes (point ´a point spatial prediction). The results are then displayed using a raster map	35
2.10	Typical diagram of semivariogram	39
2.11	Typical ANN architecture model	43
3.4	SOM input space (3-D) and output space (2- D). Red dots signify the input patterns while the blue dots shows the connected SOM neurons	63
3.5	Artificial neural network, identical to the human neural network	63
5.1	Spatial distribution maps of soil heavy metal concentrations	144
5.2	Estimated Ordinary Kriging concentration maps for Pb, Cu, Cr, Zn, Mn, Cd, As and Hg (mg/kg)	146

## LIST OF ILLUSTRATIONS

Illustration No.	Description	Page
1.1	Outline and relations between five chapters of this study	13
3.1	Location map of Rajbandh at Khulna city of Bangladesh	46
3.2	Map showing of soil sampling locations in waste disposal site	48
3.3	Flow diagram of research methodology in this study	49
4.1	Normal QQ plots for dry season (a) Cd; (b) Ni; (c) Pb and (d) Zn	67
4.2	Normal QQ Plots in the rainy season (a) Cd; (b) Ni; (c) Pb and (d) Zn	68
4.3	Scree plot of the PCs (a) dry season; (b) rainy season	78
4.4	Correlation circle for metal elements in soil for dry season	82
4.5	Correlation circle for metal elements in soil for rainy season	83
4.6	Correlation plot of soil sampling locations of dry season	84
4.7	Biplot of soil sampling locations and concentrations of metal elements of dry season	84
4.8	Correlation plot of soil sampling locations of rainy season	85
4.9	Biplot plot of the locations for metal elements (rainy) in soil of waste disposal site	85
4.10	Dendrogram for metal elements in soil during (a) dry and (b) rainy season	87
4.11	Plot of SOM sample hits for metal elements in (a) dry season (b) rainy season	90
4.12	Plot of SOM weight plane for metal elements of (a) Cd, (b) Ni, (c) Pb and (d) Zn in dry season	91
4.13	Plot of SOM Weight Plane for metal elements of (a) Cd, (b) Ni, (c) Pb and (d) Zn, in rainy season	92
4.14	Spatial distribution of Cd in soil using IDW with power of 1- 5	95
4.15	Spatial distribution of Ni in soil using IDW with power of 1- 5	97
4.16	Spatial distribution of Pb in soil using IDW with power of 1- 5	98
4.17	Spatial distribution of Zn in soil using IDW with power of 1- 5	100
4.18	Spatial distribution of Cd in soil using LP with order of 1- 3	103
4.19	Spatial distribution of Ni in soil using LP with order of 1- 3	105

4.20	Spatial distribution of Pb in soil using LP with order of 1- 3	108
4.21	Spatial distribution of Zn in soil using LP with order of 1- 3	110
4.22	Spatial distribution of Cd in soil using RBF's	112
4.23	Spatial distribution of Ni in soil using RBF's	114
4.24	Spatial distribution of Pb in soil using RBF's	116
4.25	Spatial distribution of Cd in soil using RBF's	118
4.26	Spatial distribution of Cd, Ni, Pb and Zn in soil using OK for best fitted model	121
4.27	Semi variogram of metal elements	126
4.28	Performance plot of Cd	133
4.29	Performance plot of Ni	133
4.30	Performance plot of Pb	134
4.31	Performance plot of Zn	134
4.32	Regression coefficient of Cd in soil	136
4.33	Regression coefficient of Ni in soil	136
4.34	Regression coefficient of Pb in soil	137
4.35	Regression coefficient of Zn in soil	137
4.36	Histogram plot of Cd	139
4.37	Histogram plot of Ni	139
4.38	Histogram plot of Pb	139
4.39	Histogram plot of Zn	139

## Nomenclature

All the notation and symbols are defined where they first appear in the text or figures. For convenience, the more frequently used symbols and their meanings are listed below.

AAS	: Atomic absorption spectroscopy
AHC	: Agglomerative Hierarchical Clustering
Al	: Aluminum
ANN	: Artificial Neural Network
As	: Arsenic
ASPE	: Average Standard Prediction Error
ASTM	: American Society for Testing and Materials
Ba	: Barium
BCSIR	: Bangladesh Council of Scientific and Industrial Research
BH	: Borehole
Ca	: Calcium
CBO's	: Community Based Organizations
Cd	: Cadmium
Co	: Cobalt
Cr	: Chromium
CRS	: Completely Regularized Spline
CV	: Coefficient of Variation
DPHE	: Department of Public Health Engineering
EPA	: Environmental Protection Agency
ESRI	: Environmental Systems Research Institute
Fe	: Iron
GIS	: Geographic Information System
GPS	: Global Positioning System
G-value	: Goodness of Prediction
H <sub>2</sub> SO <sub>4</sub>	: Sulfuric acid
HCl	: Hydrochloric acid
HClO <sub>4</sub>	: Hypochlorite acid
Hg	: Mercury

HNO <sub>3</sub>	: Nitric acid
IDW	: Inverse Distance Weighting
IMQ	: Inverse multiquadric
JICA	: Japanese International Cooperation Agency
K	: Potassium
KCC	: Khulna City Corporation
Kg	: Kilogram
K-S test	: Kolmogorov–Smirnov test
KUET	: Khulna University of Engineering & Technology
LP	: Local Polynomial
MAPE	: Mean Absolute Percentage Error
mg	: Milligram
mL	: Milliliter
Mn	: Manganese
MPE	: Mean Prediction Error
MSE	: Mean Standard Error
MSPE	: Mean Standard Prediction Error
MSW	: Municipal solid waste
MQ	: Multiquadric
Na	: Sodium
Ni	: Nickel
NGO's	: Non-Government Organizations
nntool	: Neural Network tool
OK	: Ordinary Kriging
Pb	: Lead
PCA	: Principal Component Analysis
Q-Q plot	: Quantile Quantile plot
RBF's	: Radial Basis Functions
RI	: Relative Improvement
RMSE	: Root Mean Square Error
RMSSE	: Root Mean Standard Square Error
R-value	: Regression coefficient
Sb	: Antimony

Sc	: Scandium
SD	: Standard Deviation
SOM	: Self-Organizing Map
SPSS	: Statistical Package for the Social Science
ST	: Spline with tension
S-W test	: Shapiro Wilk test
Ti	: Titenium
TPS	: Thin Plate Spline
UNEP	: United Nations Environment Programme
UNFPA	: United Nations Fund for Population Activities
V	: Vanadium
Zn	: Zinc
°C	: Degree celsius
%	: Percentage
:	: is to



## CHAPTER I

### INTRODUCTION

#### 1.1 Background Information

The term “landfill” is a unit, designed and operated for the disposal of municipal solid waste (MSW) to protect the environmental receptors such as human, water, air, soil, etc. from the contaminants like metal elements presence in MSW stream (Alamgir et al., 2005). Open dumping is the main disposal method of MSW in developing countries like Bangladesh (Rafizul et al., 2012). In disposal sites, MSW decomposes and produces three components of solid (degraded waste); liquid (leachate that is infiltrating into the underlying layer) and landfill gas (Vaalgamaa and Conley, 2008). The huge quantities of MSW in open dumping sites; biological, physical and chemical process as well as due to the emission of toxic metal element from MSW, leachate and contaminated soil create vulnerable to the environmental components and nearby inhabitants (Nriagu and Pacyna, 1988). With the increasing demand for metals in industries and rapid urbanization in many parts of the world, contamination by metal elements in the terrestrial environment has become widespread in a global context. Increasing metal pollution has severely disturbed the natural geochemical cycling of the ecosystem. The sources of metal elements are either natural parent rock materials or anthropogenic/human activities (Tahir et al., 2007). The geological composition of parent materials primarily influences the nature of metal elements presence in soil. Human activities such as urban-industrial development, landfill management, vehicular emissions, fossil fuel combustion and agricultural practices also influence metal element contents in soil (Mireles et al., 2004).

The effects of metal elements are found to vary with the conditions prevailing in the dumping sites and its binding forms (Pebesma et al., 2007). In addition, the contamination of soil with metal elements draws prodigious consideration due to its impending threat to food safety and injurious belongings on the environmental components. Marina et al. (2003) indicated that non-biodegradability characteristics and elongated biological half-lives of metal elements for abolition, their accretion in nutrition chain will obligate a

substantial effect on humanoid well-being in elongated period. There has long been concern about the issue of contamination of soil by metal elements because of their toxicity for plant, animal and human beings as well as their lack of biodegradability (Zhuang et al., 2009). Interests arise among the researchers to find out the origin and consequences of these metal elements on the earth due to its toxic and detrimental effects in the environment (Jia et al., 2010).

Soil is one of the principal natural resources of the earth and a vital part of the natural environment. It mainly provides a habitat for a wide range of organisms and influences the plant species thus contribute to the diet of human and animal. So, soil is an indispensable to all living being in the world. Recently, open dumping for MSW is considered an increasing threat to the underlying soil layer and the surrounding environment of disposal sites. In addition, soil contamination as part of land degradation is caused by the presence of xenobiotic (human-made) chemicals or other alteration in the natural soil environment. It is typically caused by improper disposal of MSW, industrial activity and agricultural chemicals. Soil contamination occurs when the presence of toxic chemicals, pollutants or contaminants with high concentrations in soil and it has great risk to plants, wildlife, humans and of course for the soil itself (Jia et al., 2010). Materials that find their entry into the soil system persist and accumulate in toxic concentrations becoming sources of pollution in soil (Misra and Mani, 2009). The concentration of metal elements in soil and their impact on ecosystems can be influenced by many factors such as the parent rock, climate and anthropogenic activities (Jia et al., 2010). Among the pollutants that persist and accumulate in the soils include; inorganic toxic compounds for example fertilizers, organic wastes, organic pesticides and radio nucleides (Misra and Mani, 2009). The soil is thus becoming increasingly contaminated with chemicals and other pollutants which can reach the food chain, surface water or ground water and ultimately be ingested by man (Misra and Mani, 2009).

Khulna is one of the fast growing commercial cities in Bangladesh. In this city, most of the MSWs are collected from door-to-door without any sorting and dumped in open disposal site at Rajbandh. Due to inadequate management practices of MSW in disposal site, the generated leachate percolates into the underlying soil, ground and surface water bodies and then contaminates the environmental components. The emission of toxic metal element

from deposited MSW, leachate and contaminated soil will be vulnerable to the environmental components (soil, air, water, etc.) and inhabitants. Moreover, soil is a crucial component of rural and urban environments, and in both places, land management is the key to safeguard the quality of soil (Bhuiyan et al, 2016). In present time, more concern about soil contamination by metal elements is necessary. Necessities arise to take steps for the proper disposal of MSW as well as maintenance of disposal site in Khulna city of Bangladesh. Moreover, to date, there is no comprehensive study to examine how the metal elements are correlated to each other as well as their possible sources of contamination such as anthropogenic or human activities and natural parent materials. In addition, there is no ready manual or guidebook related to this research from where one can easily get the information about quantitative distribution of metal elements spatially as well as the level of contamination of soil due to presence of metal elements in soil.

A statistical analysis of Hu et al. (2013) pointed that the metal element content in soil is usually a skewed normal distribution and is spatially auto correlated in case of soil pollution surveys. The incorporation of multivariate statistics and geostatistical analysis is a progressive method for ascertaining of pollution characteristics of metal elements in soil and distinctive their natural sources. The application of multivariate statistical approaches permit a better technique for classification, modeling and interpretation of soil monitoring data (Smith et al., 2007). In addition, geostatistical approaches can provide steadfast evaluations at unsampled positions providing that the sampling interval revolves the dissimilarity at the level of concentration (Kerry and Oliver, 2004). The technique of geographic information system (GIS) is enabling to manage large volume of spatial and non-spatial data that comes from various sources. Spatial distribution is essential for assessing the effect of metal elements in soil and to delineate contamination zones (Omran and Razek, 2012). Among statistical approaches, geostatistical kriging-based practices, including simple kriging and ordinary kriging, universal kriging as well as simple co-kriging have been continually used for three-dimensional analysis of soil data. Deterministic interpolation approaches, such as inverse distance weighting (IDW) method and its revisions are frequently applied (Chary et al., 2008) as well as Radial basis functions (RBF) and Local polynomial interpolation (LPI) through ArcGIS are also used to predict of values, standard error and condition number that are comparable to ordinary

kriging with measurement errors (Zhang, 2011). Creutin and Obled (1982) compared the performance of several interpolation techniques from produced prediction surfaces.

Artificial Neural Networks (ANN) is an information processing systems, which imitate the functioning of the human brain. Neurons are spread over several layers in the network to send each other information via given connections (Neuner, 2012). These connections are weighted and due to the change of weight they will be adapted in a way that the network is able to provide the best solution for a neural network has at least one input and output layer. ANN is considered as dependable and inexpensive techniques for data interpretation and prediction. The self-organizing map (SOM) is an unsupervised ANN used for data training to classify and effectively recognize patterns embedded in the input data space. Kohonen (2014) found that the application of SOM-ANN is useful for recognizing spatial patterns in contaminated zones by integrating chemical, physical, Eco toxicological and toxic kinetic variables in the identification of pollution sources and similarities in the quality of the samples. Remarkable SOM methods have been used in soil as well as water quality assessment for pollutant sources identification, pattern recognition and classification (Lee et al., 2006).

In this study, in total sixty soil samples were collected at a depth of 0-30 cm from the existing ground surface from different selected locations of waste disposal site at Rajbandh, Khulna, Bangladesh. The soil samples were collected in two different seasons; dry season (March to May, 2016) as well as rainy season (June to August, 2016). In the laboratory, the concentration of metals elements of Aluminium (Al), Arsenic (As), Barium (Ba), Calcium (Ca), Cadmium (Cd), Cobalt (Co), Chromium (Cr), Copper (Cu), Iron (Fe), Mercury (Hg), Potassium (K), Manganese (Mn), Sodium (Na), Nickel (Ni), Lead (Pb), Antimony (Sb), Scandium (Sc), Strontium (Sr), Titanium (Ti), Vanadium (V) and Zinc (Zn) was measured through the standard test methods. Therefore, this study was conducted to examine the spatial distribution and temporal variation of the concentration of metal elements in soil. In this study, the distribution of the concentration of metal elements was described using conventional statistics such as mean, maximum, minimum, median, standard deviation (SD), co-efficient of variation (CV), skewness and kurtosis by performing Statistical Package for the Social Sciences (SPSS).

The Pearson's correlation coefficient of metal elements was also conducted using SPSS to determine the accumulated concentrations of metal elements irrespective to their sources which provide an effective way to reveal the relationships between the multiple variables. Principal Component Analysis (PCA) and Agglomerative Hierarchical Clustering (HCA) were performed using XLSTAT to get the information about spatial distribution of metal concentrations in soil by comparing the linkage between different sampling sites. These PCA and HCA were performed also to know how the metal elements are correlated to each other as well as their possible sources of contamination such as anthropogenic or human activities and natural parent materials.

Furthermore, geostatistical analysis such as IDW, Ordinary Kriging, LP and RBF were performed using ArcGIS. These geostatistical analyses were performed to examine the quantitative distribution of metal elements spatially in soil as well as the level of contamination of soil due to presence of metal elements in soil. In addition, semivariogram parameters obtained ordinary kriging exhibited spatial dependence of metal elements in soil. The performance of interpolation techniques will assess by ascertaining the prediction error based on Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), Goodness of prediction (G-value) and Relative Improvement (RI). In this study, these prediction errors were used to select the best fitted model by performing cross validation of different models. In addition, ANN through MATLAB will perform to determine Mean Standard Error (MSE) and Regression factor (R-value) to check the validity and accuracy of the metal element concentration obtained from laboratory. Moreover, to date, there is no comprehensive study regarding the effect of metal element concentration from waste disposal site in Bangladesh. The main outcome of this study was to know the correlation of metal elements with each other in soil, possible sources of their contamination, distribution of metal elements spatially and the level of contamination of soil due to presence of metal elements in soil. Finally, this study will so guide for more efficient management practices of MSW and adoption of suitable technological solutions for the disposal of MSW in disposal sites not only the developing countries like Bangladesh but also all over the world.

## 1.2 Soil and Metal Element

Disposal site is one of the sources for the contamination of underlying soil, surface and ground water due to generation of leachate from MSW and its migration through MSW in and around of the disposal site (Wang et al., 2015). It is known that the impact of disposal sites can cause contaminate of all the environmental components (Ehab and Ahmed, 2015). The study of the impact of disposal sites on the environment dedicated to the scientific works of many scientists and researchers. In addition, land degradation or soil contamination caused by human activities has significant adverse effects on all the environmental components and ecosystems world-wide (Bai et al., 2013), and MSW is an important and emerging environmental problem.

Disposal sites containing hazardous materials are under critical observation today for potential hazards, resulting in the need for thorough risk analyses along with the soil, surface water and ground water that have been contaminated with metal elements leaching from disposal sites (Bai et al., 2013). Soil contamination by metal elements has become a critical environmental concern due to its potential adverse ecological effects (Li et al., 2013). The term “metal element” is generally used to describe a group of metals and metalloids with an atomic density greater than  $5.0 \text{ g/cm}^3$  and is toxic or poisonous even at low concentration. Their effects on living organisms generally results from contamination of either a biotic systems (soil, water and air). Metal elements occur naturally at low concentrations in soils; however, they are considered soil contaminants due to their widespread occurrence, as well as their acute and chronic toxicity. When metal elements are incorporated into the MSW generated regularly and these MSW are subsequently disposed of in disposal sites at their end of life or use, there is a high possibility that with time, the metal elements will be released into the surrounding ecosystems mainly soil and water (Chen et al., 2012). Metal elements also occur naturally, but rarely at toxic levels. Once metal elements released into the environment, they are often considered as a problematic environmental pollutant because of their well-known effects on living organisms.

In addition, metal elements constitute an ill-defined group of inorganic chemical hazards, and those most commonly found at contaminated disposal sites are Pb, Cr, As, Zn, Cd, Cu,

Hg, and Ni. Soil is a crucial component of rural and urban environments, and in both places, land management is the key to safeguard the quality of soil. Increased concern on soil contamination by metal elements has been shown in recent years. Soils are sources of substrate nutrients and are the basis of sustenance to livelihood. Thus, soils play an important role in ecological stability. Nevertheless, the quality of soil with regards to the concentrations of metal elements may be compromised (Nabulo et al., 2008). They are notorious when they bio-accumulate in soil and due to their long persistence time in the course of interaction with soil component, they consequently enter food chain through plants or animals. Thus plants grown on contaminated soils bio-accumulate these metal element contaminants which pose high risk to human health.

One of the major and still topical issues is the estimation of the effect of various factors and processes on the quality of soil. Unfavorable changes in the physical, chemical or biological properties of soil may result not only in a decrease of its fertility, but also the effects of they can even totally exclude it from production (Zhang et al., 2012). Several reports have been published which are documented only on the characterization of leachate and its effect on groundwater pollution in Bangladesh, but little information is available on the effect of disposal sites on soil contamination and its toxicological effects. Soil is the key part of the earth system as it control the hydrological, erosional, biological, and geochemical cycles. The soil system also offers goods, services and resources to humankind (Keesstra et al., 2012). This is why it is necessary to know how soils are contaminated by disposal sites and how metal elements are distributed spatially.

### **1.3 Problem Statement and Justification**

Unplanned and uncontrolled disposal of MSW has become a thread to all living element of the earth as they contain toxic metal elements. Consumption patterns of population have dramatically changed with passage of time. In consequence, the amount of waste generated from household, industrial and commercial field have increased than before. Although there are many alternative options of waste disposal, disposal site is the primary mean of MSW disposal in developing countries like Bangladesh (Samiul, 2016)

The only authorized MSW dumping site of Khulna is Rajbandh disposal site. The location of this site is not only for sore of eyes, but also causes hazard for environment. The soil in and around the dumping site is usually rich in toxic metal elements due to long retention and decomposition of MSW in this site. The metal elements get blended not only with the in-place soil, but also spread over surrounding soils of disposal site. This soil is used by the people living around the dumping site for planting vegetables and fruits. These plants bio- accumulates metal elements from the soil and when they are eaten by human beings and animals, the metal element accumulate in the body with serious health effects (UN, 2010). Leachate and rain water passes through MSW have also become a major source of contamination in water bodies surrounding the selected disposal site area. Consumption of this contaminated water may directly cause metal element poisoning to the consumer. In addition, hazardous MSW containing harmful chemicals and metal elements are openly disposed in the disposal site, which may even cause to death. The long-term input of heavy-metal elements could result in decreased defending capacity of soil, threatening the ecological environment (Murtaza, 2012). Despite of such terrible effects of metal contamination in soils of this disposal site, no proper step is taken to control or reduce the spread of metal elements in and around the soil of disposal site. Therefore it is essential to carry out an intensive study and monitor the nature and extent of such pollution on in and around the disposal site.

Disposal sites are supposed to be located away from residence because of the inherent environmental nuisance and poor aesthetic value associated with their operations. Rapid urbanization has resulted in existing dumping sites originally located at a safe distance outside the municipal boundaries now being increasingly encircled by settlements and housing estates. This research seeks to display the seasonal variability in concentration of metal elements in the soil in and around of the waste disposal site as well as the sources of their generation. The distribution of metal elements in soil of the surrounding areas of disposal site is also necessary to explore their extents. In this study, implementation of different interpolation techniques can provide better prediction of the distribution of metal elements spatially as well as the possible sources of contamination of metal elements in soil of the selected disposal site.



## **1.4 Objectives of the Study**

Soil contamination is a crucial part of land degradation caused by the presence of toxic chemicals or metal elements and other alteration in the natural soil environment. Soil contamination occurs when the presence of toxic chemicals, pollutants or contaminants with high concentrations in soil. There has long been concern about the issue of contamination of soil by metal elements because of their toxicity (Keesstra et al., 2012). Now a day, interest arise among the researchers to find out the origin and consequences of these metal elements on the earth due to its toxic and detrimental effects in the environment. Moreover, in this study, statistical and Geostatistical techniques were performed to know the information about correlation of metal elements in soil, possible sources of their contamination, quantitative distribution and the level of contamination of soil due to presence of metal elements in soil. The overall objectives can be summarized as follows:

1. To develop correlations of metal elements and their possible sources of contamination in soil of waste disposal site.
2. To predict the best fitted geostatistical model by assessing cross validation of different interpolation techniques.
3. To visualize the level of contamination and distribution of metal element spatially in the soil of the waste disposal site.
4. To develop a model and check the accuracy of the observed and predicted values of metal elements in soil using neural network modeling.

## **1.5 Contribution to Knowledge**

It is expected that this research will provide a guideline for performing various interpolation techniques to distribute metal elements spatially and their possible sources of contaminations. The multivariate statistics used in this study will contribute a clear knowledge about the origin of metal element, *i.e.*, from where these are generated either from natural parent material or anthropogenic/human activities. This study will provide an easy and simpler way to express the pattern of contamination level of metal elements in form of map in soil. Different colors attributed to different concentration levels indicated

the how intensely the region is contaminated and how immediate steps should be taken to reduce the contamination. Moreover, interpolation techniques used in this study will offer a reference to interpolate soil data for unsampled location provided with best interpolation technique with least prediction errors. Performance assessment of observed and predicted values will give chance to interpret the accuracy of assessment *i.e.*, how accurately the real field condition is presented in the produced map. Finally, this study can be used as a reference of guideline for future research related or similar to this one.

### **1.6 Significance of the Study**

The generation and measurement of the concentration of metal elements of Al, As, Ba, Ca, Cd, Co, Cr, Cu, Fe, Hg, K, Mn, Na, Ni, Pb, Sb, Sc, Sr, Ti, V and Zn will be used to aware the general population of Khulna region. In addition, multivariate statistical analysis will offer a clear conception about the sources of metal element generation. The study will also help to inform the authorities about the sources of spread as well as the distribution pattern of these metal elements in soil of the disposal site. So, necessary steps can be taken to control the spread of metal elements outside the disposal site area. The results from the study will also be used to find out the area needed to take under immediate remedial action to remove the metal elements where the levels were too high.

### **1.7 Scope and Limitations**

Rapid growth of population and industrialization surrounding the Khulna city tends to increase the generation of MSW, creates additional load to MSW management system and finally contaminated the environmental components and surrounding soil layer. Contaminations of metal element in and around of the Rajbandh disposal site possess a thread to the inhabitant of the study area. The spatial and temporal variation of metal elements concentration in and around the soil of waste disposal site was described using conventional statistics which shows the magnitude and pattern of contamination of metal elements in both the dry and rainy seasons. This study aims at categorize metal elements depending on its origin, *i.e.*, either the metal element comes from parent materials or originate due to human activities based on multivariate statistics included Pearson correlation, PCA and AHC. These analyses also indicated the correlation or similarity or

dissimilarity between the metal elements irrespective to their sources. Different interpolation techniques were implemented to show the spatial distribution of metal elements in and around the soil which offers mapping of the concentration of metal elements for different samples collected from different location as well as varying depth. Different interpolation techniques were implemented to produce prediction surface provided with prediction error to check the accuracy. The produced surface with least error offered more clear visualization of spread of metal elements in and around soil of waste disposal site. In this study, ANN presented a developed network to predict the unknown values of concentration of metal elements, if latitude, longitude and depth of sampling were set as input. This helps to evaluate the contamination of metal element at larger depth and the spatial distribution also can be displayed. Thus, it will be feasible to take necessary steps to manage the unintended disposal of waste to stop the spread of metal elements. Further study can also be carried out where it is needed to implement in future.

The main limitation of the study was no direct computation of index to assess the risk due to metal element contamination in and around the soil of waste disposal site. Some field conditions were ignored during the sampling procedure and the laboratory test values do not represent the in-situ condition of collected soil samples. More carefulness will help to get expected result.

## **1.8 Outline of the Thesis**

The thesis outline has been divided in five distinct chapters comprising different aspects of study. The outline and the relations between these five chapters are shown in Figure 1.1. The chapters portrayed the knowledge of origin and generation of metal elements and the sources of the contamination of soil of the selected waste disposal site. Consequences due to unplanned waste generation and disposal were described briefly. Distributions of metal elements in and around the soil were also represented for different interpolation techniques.

**Chapter I** represented a general knowledge on the background of waste disposal sites, MSW, contaminated soil, metal element and the possible sources of contamination of soil in disposal site. In addition, the problems associated to unplanned waste disposal and its

justification in context of world, objectives of the study and the scope and limitations of this study are also highlighted in this chapter.

**Chapter II** encompassed a literature review related to this topic. This chapter mainly deals about the MSW disposal facilities, MSW and impacts of disposal sites on the environmental components and surrounding areas. In addition, present scenario of MSW management and disposal facilities in Bangladesh, providence and distribution pattern of heavy metals in soil and its environmental impacts are also dealt in this chapter. This chapter deals about the review of previous studies conducted to assess spatial and temporal distribution of metal elements in soil, water as well as sediments in perspective of the world as well as in Bangladesh. This chapter also illustrates the conventional statistics, multivariate analysts and the geostatistical techniques used in this study.

**Chapter III** contained elaborate description of methodology of research. The sampling process, testing in the laboratory, principles of multivariate statistics, *i.e.*, Pearson's correlation, PCA, AHC and geostatistical analyses; *i.e.*, IDW, LPI, RBF and Ordinary kriging was described in this chapter. Prediction and performance assessment by ANN were also described. Stepwise implementations of different software's for analysis were also presented in this chapter.

**Chapter IV** depicted the concentration of metal elements present in collected sample from waste disposal site tested in the laboratory. Descriptive statistics including normality test *i.e.* Shapiro-Wilk test, K-S test and normal Q-Q plot was revealed for each metal element. Conventional statistics was described for the metal element concentration in the soil of the study. An incorporation of multivariate statistics with Pearson correlation, PCA and AHC were shown in this chapter. Geostatistical analysis was performed using GIS approach. A developed network of ANN was represented and performance assessment of the interpolation techniques was described based on literature indices as well as ANN network.

**Chapter V** represents the comparison and validation of the findings of the present with the results of published by various researches for similar cases of this study available in the literature.

**Chapter VI** comprises the conclusion and recommendations of the study based on logical experimental results and recommendations required for proper waste handling and management procedures with sufficient environmental and health safety and also recommendations for actions and studies to be required in future.

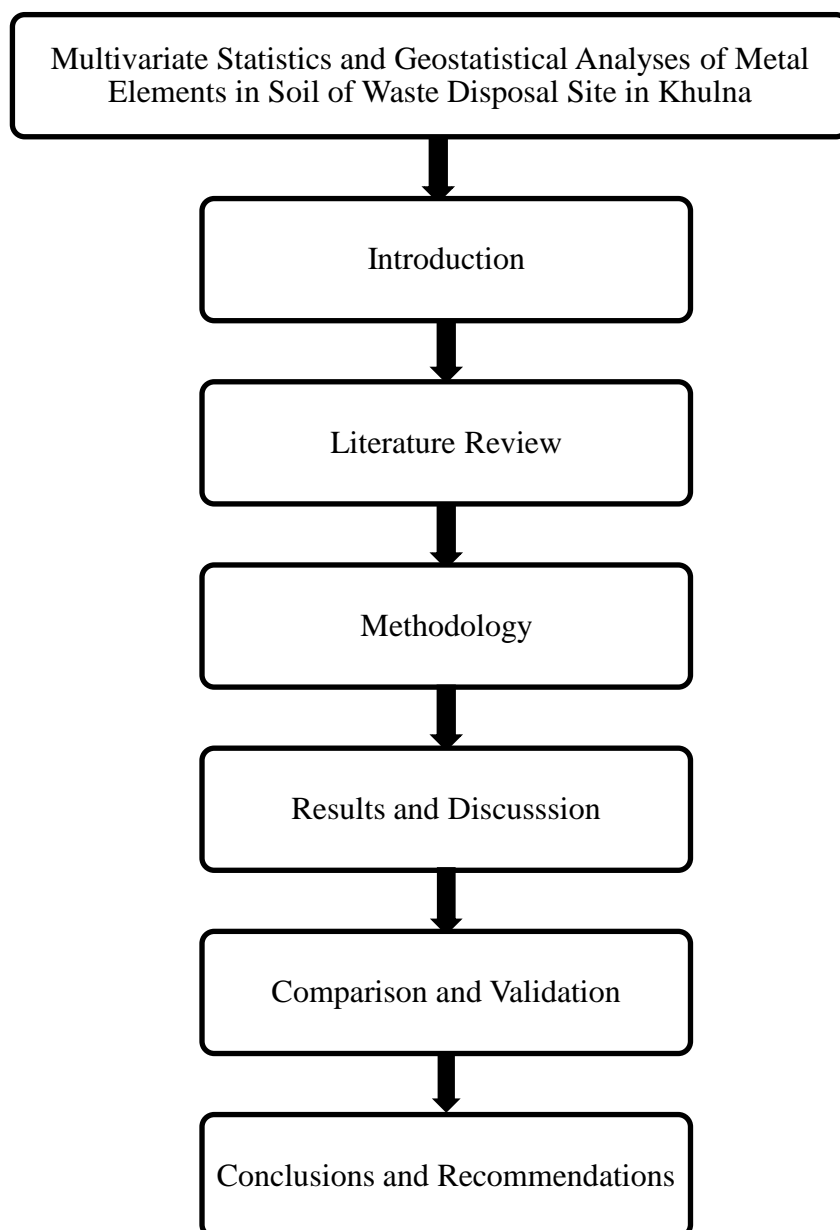


Figure 1.1: Outline and relations between six chapters of this study.

## **CHAPTER II**

### **LITERATURE REVIEW**

#### **2.1 Introduction**

This chapter deals with the information related to the contamination of soil due to the presence of heavy metals in the soil of disposal site. This chapter also illustrates the background of different multivariate and Geostatistical techniques as well as artificial neural network (ANN) used in this study. The literature review presented in this chapter was collected from available previous research reports and technical papers related to this topics. The organization of the literature review begins with an introduction to MSW and waste disposal facilities first, and is followed by the generation of MSWs and contamination of soil due to heavy metals. The interrelationship between soil, MSW and heavy metals is illustrated in this chapter. An assessment of multivariate statistics is also done to comprehend the generation sources of heavy metals as well as the correlation between the metals elements irrespective to their sources. After establishing the literature review related to heavy metal and soil, general discussion on conventional and multivariate statistical analysis is also discussed in this chapter.

#### **2.2 Municipal Solid Waste**

Municipal solid waste (MSW) can be defined as useless, unused, unwanted or discarded material available in solid or semisolid form (Qasim and Chiang, 1994). It normally includes all the community waste such as durable good, non-durable goods, containers and packaging, food waste, yard wastes and small amount of inorganic waste except industrial wastes. Many changes of MSW generation and composition have taken place due to urbanization including increase in the population. The rate of consumption has risen and the lifestyle of the people, too, has changed. Several studies have shown how increasing waste has affected society and the environmental components. The amount of waste production is a sign of the level of industrialization or a degree of development of a country or a city.

## **2.3 Dumping Facilities of MSW**

Municipal solid waste (MSW) is generated by natural and human activities and it has significant adverse effects on the environmental components (soil, water, etc.) and atmosphere worldwide (Li et al., 2013). The improper management of MSW raises public concern over potential harmful effects to the local communities as well as the environmental components. These concerns probably have become more pragmatic when recent intensive studies demonstrated increased human health risk caused by exposure to toxic chemicals and heavy metals in waste dumping sites (Agusa et al., 2003). MSW management simply means the collection, transport, processing, disposal, managing and monitoring of MSW materials to minimize its' consequences on humans and environment. There are several methods of managing of various types of MSW all over the world (Bai et al., 2013). Landfilling practices in developing countries like Bangladesh differ from that of the developed countries, which follow advanced landfilling practices such as sanitary landfill (engineered landfill) as opposed to open dumping practiced in the developing countries. There are two methods for the final disposal of MSW, one is open dumping and the other is sanitary landfill shown in Figure 2.1 (Rafizul et al., 2012). Open dump facilities cause additional harm to the environment, but not doing anything is not an option.

### ***2.3.1 Sanitary Landfill***

Sanitary landfill is one of the most widely used MSW management techniques; however, it needs high standard of environment protection in the operation of landfill (Oyeku and Eludoyin, 2010). The changing from dumping to high standard of environment protection needed time change technology, change in thinking and behavior (Pugh, 1999). In addition, sanitary landfill is an engineered technique for the disposal of MSW on the land by spreading them in thin layers followed by compacting them to the smallest practical volume before covering them with soil at regular intervals (Figure 2.1a). Sanitary landfilling involves placing MSW in lined pits with appropriate means of leachate and landfill gas control (Oyeku and Eludoyin, 2010). It is highly recognized as an environmentally and internationally desired technique of MSW disposal since it minimizes environmental damage and thus eliminates odors.

### 2.3.2 Open Dumping

Open dump is a disposal practice without waste compaction during placement; neither compacts nor covers with soil. The land is used without preparation of engineering planning (Daniel and Koerner, 1995). Thus, the open dump site characteristics are unplanned heaps of uncovered waste, burning waste at the dump site, pools of standing polluted water, rat and fly infestation and waste scavenging at dump site (Pugh, 1999). Most widespread method of MSW disposal is open dumping practices. An open dumping is defined as a land disposal site at which MSW are disposed of in a manner that does not protect the environment, are susceptible to open burning, surface and ground water contamination (Figure 2.1b).

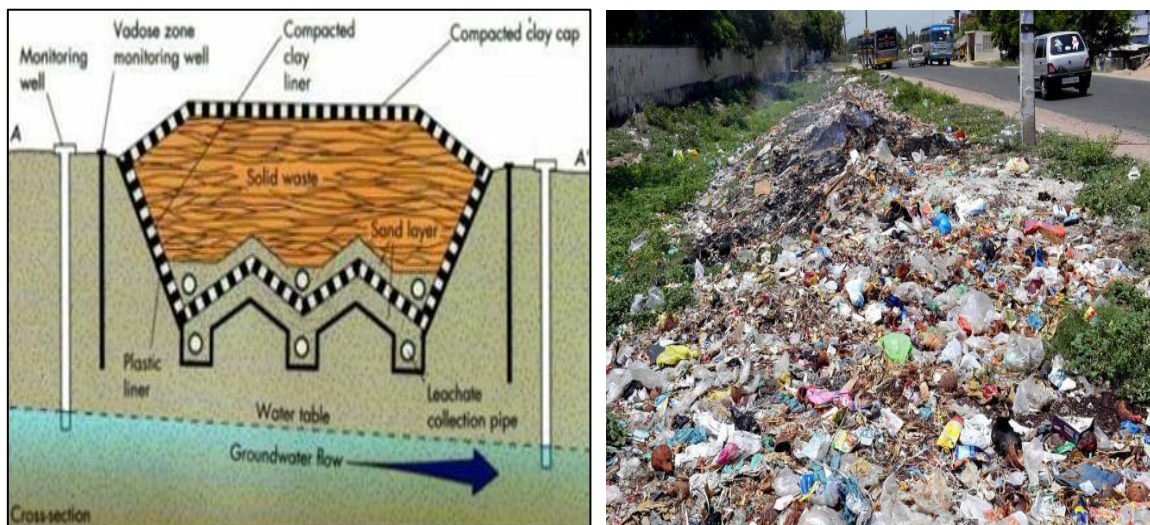


Figure 2.1: Methods of MSW dumping facilities (a) Sanitary landfills (b) Open dumping  
(Source: [http://earthsci.org/basicgeol/solid\\_waste/solid\\_waste.html](http://earthsci.org/basicgeol/solid_waste/solid_waste.html)).

In developed countries, landfills have historically been the primary method of MSW disposal because this method is the most convenient and because the threat of groundwater contamination was not initially recognized (Smith, 2007). The practice of landfill system as a method of MSW disposal in many developing countries is rarely practiced (Oyeku and Eludoyin, 2010). Internationally, almost 70% of MSW is disposed of at landfills. In landfill, deposited MSW generates leachate that constitutes a pollution source into the environment and water resources.



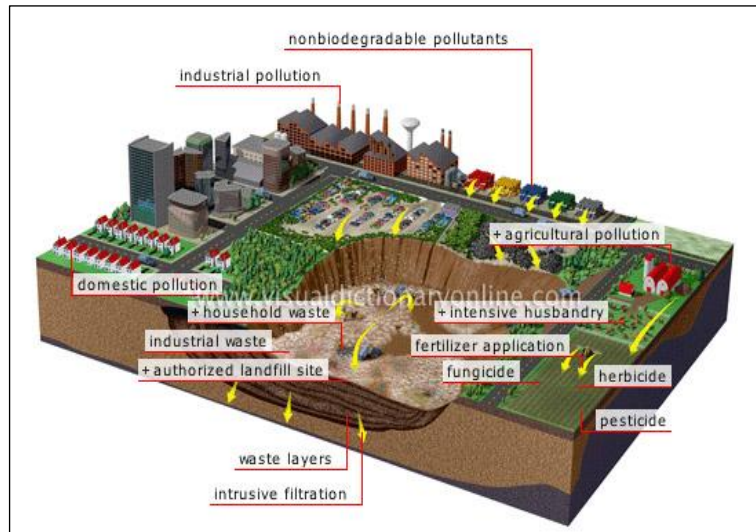


Figure 2.2: Sources of land pollution in MSW disposal site (Source: <http://imnh.isu.edu/digitalatlas/hydr/concepts/gwater/aquifer.html>).

In many towns, sorting is not done, and all the MSW (paper, food, diapers, glass etc.) is mixed up and deposited in the nearby landfill. In absence of proper unloading operation of landfill, it soon become full, smelly and unsafe for the environment due to disposal of all types of degradable and non-degradable materials of generated MSW (Chary et al.,2008). Thus, contamination of heavy metal occurs in and around the soil of MSW disposal site. Figure 2.2 illustrates a diagrammatical representation of the sources for land pollution due to unplanned and uncontrolled dumping of MSW in landfill. Sanitary landfills are lined at the bottom to minimize the leakage of soil pollutants and other toxins from getting into the water table and surrounding surface water bodies. This method is effective, but maintenance of landfill is expensive and difficult.

### 2.3.3 Overall Impacts of MSW Disposal Site

Pollutants can escape from improperly designed landfill in a variety of ways such as leachate and landfill gases. Leachate consists of a mixture of organic and inorganic compounds, many of which have a hazardous impact on the environment (Wang et al., 2015). According to Li et al., (2006), the greatest contamination threat to groundwater comes from the leachate generated in unlined landfill from the materials which most often contain toxic substances, especially when MSWs of industrial origins are dumping in

landfill. Li et al., (2006) stated that heavy metals such as Cd, As and Cr have been reported at excessive levels in groundwater due to landfills operation. Leachates contain a host of toxic and carcinogenic chemicals, which may cause harm to both humans and the environment. Furthermore, leachate-contaminated groundwater can adversely affect industrial and agricultural activities that depend on well water (Ashraf et al, 2013). Leachate impacts to groundwater may also present danger to aquatic species if the leachate-contaminated groundwater plume discharges to wetlands or streams.

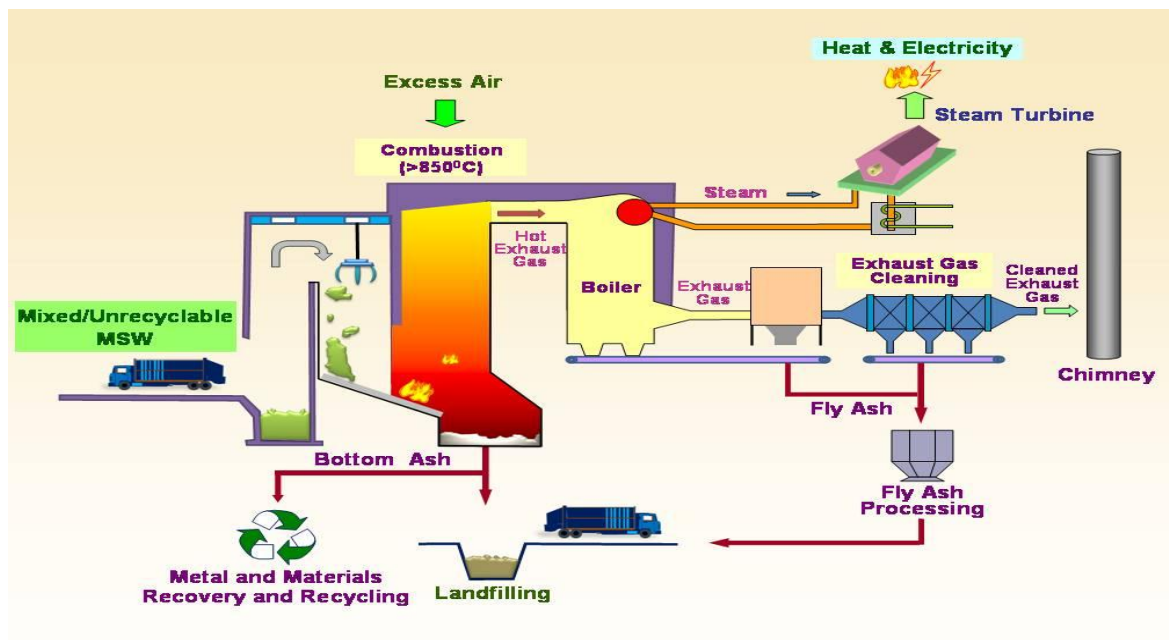


Figure 2.3: Impacts of MSW disposal (Source: St. Mary's County: Maryland, <http://www.co.saint-marys.md.us/dpw/recycleoverview.asp>).

Carbon dioxide, which comprises 40% to 60% of landfill gas, may pose specific asphyxiation hazard concerns. In assessing the public health issues of migrating landfill gas, environmental health professionals should investigate the presence of buried utility lines and storm sewers on or adjacent to the landfill (Agusa, 2003). But, hopefully there are some advantages of MSW disposal in landfill. While resource recovery and incineration both require extensive investments in infrastructure, and material recovery also requires extensive manpower to maintain, landfills have fewer fixed or ongoing costs, allowing them to compete favorably (Daniel and Koerner, 1995). Another advantage is having a specific location for disposal that can be monitored, where MSW

can be processed to remove all recyclable materials before tipping. Figure 2.3 demonstrates impacts of uncontrolled MSW disposal in landfill site.

## **2.4 Context of Heavy Metal**

Metals are defined as any element that has a silvery luster and is a good conductor of heat and electricity (McClean and Bledsoe, 1992). Theoretically, there are many terms used to describe and categorize metals, including trace metals, transition metals, micronutrients, toxic metals and heavy metals. A metal having a specific gravity more than  $5\text{gm/cm}^3$  is classified as heavy metal. In this study, to know the behavior and distribution pattern of metal elements in soil statistical and geostatistical analysis were performed. The prevalence of heavy metals in soil as well as environment and interrelationship between them is described in the following articles.

### **2.4.1 Prevalence of Heavy Metals in Soil and Environment**

The incapability to determine metal species in soils obstructs efforts to understand the mobility, bioavailability and fate of contaminant metals in the environmental systems together with the assessment of health risks posed by metal elements, and the development of methods to remediate metal contaminated sites. In soil, metals are found in one or more of the following several "pools" in the soil:

- i. dissolved in the soil solution;
- ii. occupying exchange sites in inorganic soil constituents;
- iii. specifically adsorbed in inorganic soil constituents;
- iv. associated with insoluble soil organic matter;
- v. precipitated as pure or mixed solids;
- vi. present in the structure of secondary minerals; and/or
- vii. present in the structure of primary minerals

However, in some natural soils, up to 30 to 60% of heavy metals can occur in unstable forms developed from parent materials rich in metal contamination as well as in the contaminated soils. Natural and anthropogenic sources are one of the root cause of heavy metal contamination which has caused widespread and variable the hazardous possibilities of environmental and health effect (Tahir et al., 2007). According to Ross (1994) the

anthropogenic sources of metal contamination can be divided into five main groups: (1) metalliferous mining and smelting (As, Cd, Pb and Hg); (2) industry (As, Cd, Cr, Co, Cu, Hg, Ni and Zn); (3) atmospheric deposition (As, Cd, Cr, Cu, Pb and Hg); (4) agriculture (As, Cd, Cu, Pb, Zn); and (5) MSW disposal (As, Cd, Cr, Cu, Pb, Hg and Zn). Vaalgamaa and Conley (2008) also stated a natural activity is another cause of heavy metal contamination. Industries such as plating, ceramics, glass, mining and battery manufacture are considered the main sources of heavy metals in local water systems causing the contamination of groundwater with heavy metals. Furthermore, heavy metals which are frequently found in high concentrations of waste landfill leachate are also a potential source of pollution for groundwater and underlying soil layer (Bai et al., 2013).

#### **2.4.2 Behavior of Heavy Metals in Soil**

Monitoring the endangerment of soil by heavy metals is of interest due to their influence on water bodies and underlying soil layers and also on plant life animals and humans shown in Figure 2.4. The chemical behavior of heavy metals in soils is controlled by a number of processes, including metal cation release from contamination source materials (e.g., fertilizer, sludge, smelter dust, ammunition, slag), cation exchange and specific adsorption onto surfaces of minerals and soil organic matter, and precipitation of secondary minerals. Increased anthropogenic inputs of Cu and Zn in soils have caused considerable concern relative to their effect on water contamination (Zhang et al., 2012). In addition, oxidizing conditions generally increase the retention capacity of metals in soil, while reducing conditions will generally reduce the retention capacity of metals (McClean and Bledsoe, 1992). Soil reduction has been shown to result in the coincident release of metals associated with minerals that are susceptible to reductive dissolution, in particular Mn and Fe oxides (Asante-Duah, 1998). Contaminants reaching the soil can be divided into two groups, namely micro pollutants and macro pollutants. Micro pollutants are natural or anthropogenic molecules, which are toxic at very low concentration. Macro pollutants are present in the environment locally and/or temporarily to a much higher degree than normal level. The main micro pollutants of soils are inorganic or organic compounds. (i) Inorganic micro pollutants such as Pb, Cd, Ni, Cr, Hg, Cu, Zn etc are mainly the toxic and potentially toxic heavy metals (ii) Organic micro pollutants include pesticides and certain non-pesticide organic molecules: e.g. aliphatic solvents, monocyclic

aromatics, halogenated aromatics, polychlorinated biphenyls (PCBs) and polycyclic aromatic hydrocarbons (PAHs), surfactants, plastifiers.

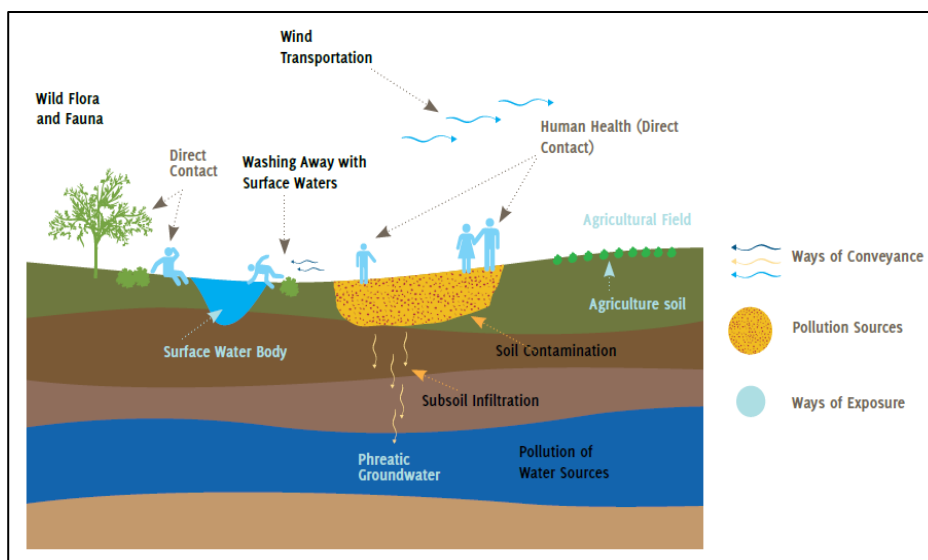


Figure 2.4: Contamination of soil from disposal site (Source: Environment and Climate Change: Canada, <https://www.ec.gc.ca/eau-water/default.asp?lang=En&n=6A7FB7>).

#### 2.4.2.1 Accumulation

Human activities have considerably increased the abundance of heavy metals in the environment. Their consequences in polluted soils is a subject of study as they exhibit direct potential toxic effects on biota and possess indirect threat to human health via the contamination of groundwater and accumulation in food crops (McLean and Bledsoe, 1992). Bioaccumulation characteristics of heavy metals make it more dangerous. Heavy metal pollution of soil enhances plant uptake causing accumulation in plant tissues and eventual phytotoxicity and change of plant community (Chary et al., 2008). In environments with high nutrient levels, metal uptake can be inhibited because of complex formation between nutrient and metal ions. Therefore, a better understanding of heavy metal sources, their accumulation in the soil and the effect of their presence in water and soil on plant systems seems to be a particularly important issue (Jia et al., 2010). Accumulation of heavy metals can also cause a considerable detrimental effect on soil ecosystems, environment and human health due to their mobilities and solubilities which determine their speciation (Kadar, 1995). Several studies have indicated that crops grown on soils contaminated with heavy metals have higher concentrations of heavy metals than

those grown on uncontaminated soil (Nabulo et al., 2006). Heavy metals accumulating in soil directly or indirectly through plants enter in food chains, thus endangering herbivores, indirectly carnivores and not least the top consumer humans (Kadar, 1995). Total levels of heavy metals have shown a trend relationship between metal concentration in soil and long term irrigation. Metals such as lead, arsenic, cadmium, copper, zinc, nickel, and mercury are continuously being added to our soils through various agricultural activities such as agrochemical usage and long-term. Application of urban sewage sludge in agricultural soils, industrial activities such as MSW disposal, MSW incineration and vehicle exhausts, together with anthropogenic sources. All these sources cause accumulation of metals and metalloids in our agricultural soils and pose threat to food safety issues and potential health risks due to soil to plant transfer of metals .

#### ***2.4.2.2 Solubility and Mobility***

Metals in the soil, surface and ground water can occur either as bonded to suspended or colloidal material (organic or inorganic) or dissolved in liquid. Soluble compounds are most common in soil and groundwater, while the suspended material can make a significant contribution in rivers and lakes. Soluble metals appear as cation complexes (e.g.  $\text{Pb}^{2+}$ ,  $\text{Zn}^{2+}$ ) with either water molecules (common water solution) or some inorganic or organic molecules coordinated, or as anions. Precipitation of metals can occur in the presence of e.g. sulphides, carbonates (at high pH), oxides, hydroxides, sulphates and phosphates. Partitioning coefficients can be used to correlate concentrations in different environmental compartments (Asante-Duah, 1998). The basic assumption is that the distribution of chemicals is driven by equilibrium. The partitioning coefficient can then be used to predict the concentration in the media of interest, such as soil, sediment, water or biota. Modeling can be used to assess the transportation and exposure of chemicals.

#### ***2.4.2.3 Bioavailability***

Bioavailability depends on biological parameters and on the physicochemical properties of metals, their ions, and their compounds. In addition, the bioavailability and mobility of metals in soil strongly depends on the extent of their sorption with solid phases. The bioavailability of some metals (Co, Ni, Cu, Zn) decreases of the soils with clay illuviation

due to their enhanced adsorption capacity. Agusa (2003) stated that Fe, Al and Mn and their insoluble compounds become toxic to plants and in addition, aluminum, iron and phosphorus.

#### **2.4.2.4 Toxicity**

Recently pollution of the general environment has gathered an increased global interest. In this respect, contamination of agricultural soils with heavy metals has always been considered a critical challenge in the scientific community. Heavy metals are generally present in agricultural soils at low levels. Due to their cumulative behaviour and toxicity, however, they have a potentially hazardous effect not only on crop plants but also on human health (Agusa et al., 2003). Even metals essential to plant growth, like Cu, Mn, and Zn can be toxic at high concentrations in the soil. Some elements not known to be essential to plant growth, such as As, Ba, Cd, Cr, Pb and Ni, also are toxic at high concentrations or under certain environmental conditions in the soil. Chemical hazards include chemical agents such as heavy metals, nutrients such as nitrogenous compounds, phosphorus compounds, minerals, insecticides, pesticides, fertilizers, fungicides, herbicides and organic hazards (Nabulo et al., 2008). Metals, unlike the hazardous organics cannot be degraded. Some metals such as Cr, As, Se, and Hg can be transformed to other oxidation states in soil, thus influencing their mobility and toxicity (Mclean and Bledsoe, 1992). Many of them (Hg, Cd, Ni, Pb, Cu, Zn, Cr, Co) are highly toxic both in elemental and soluble salt forms. High concentration of heavy metals in soils is toxic for soil organisms: bacteria, fungi and higher organisms. Short-term and long-term effects of pollution differ depending on metal and soil characters (Kadar, 1995). In the after-effect of heavy metal pollutions, the role of pollutant bounding or leaching increases which determines their bioavailability and toxicity.

#### **2.4.3 Consequences of MSW, Soil and Heavy Metal in Disposal Site**

The municipal solid waste (MSW) that disposed without proper planning and treatment are rich in different types of heavy metals. When MSW is dumped in a disposal site, they take a long time to disintegrate depending on their nature. For example, paper towel takes 2-4 weeks to disintegrate, whereas plastic bag takes 200-1000 years. The metals of the greatest concern due to their extensive use, their toxicity and their widespread distribution is Hg,

Pb, Cd, Cr and As (Baird et al., 2013). The toxicant may also cause effects on the microorganisms and soil fauna. However, the risk for further dispersion of the pollution to other recipients must also be considered. Figure 2.5 clearly demonstrates the relation between MSW, soil and heavy metals in disposal site.

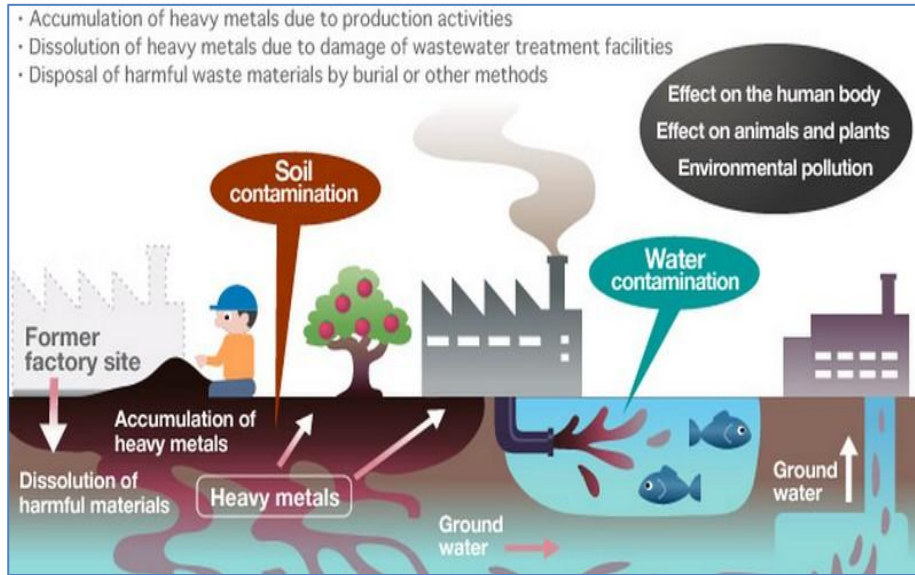


Figure 2.5: Relation between MSW, soil and heavy metals in disposal site (Source: Heavy metal treatment project, <http://www.sagasiki-kankyo.co.jp/en/jukinzoku/what/index.html>).

## 2.5 Present Scenario of MSW Management and Disposal Facilities in Bangladesh

Bangladesh is the ninth most populous country and twelfth most densely populated country in the world. There are six major cities in Bangladesh like Dhaka, Chittagong, Khulna, Rajshahi, Sylhet and Barisal. Currently, according to a UNFPA report, Dhaka is one of the most polluted cities in the world and one of the issues concerned is the management of municipal MSW. A study conducted by Samiul (2016) and found that Dhaka city produce more MSW in different ways, like- hospital, small industry, tannery and others heavy industry. Debnath (2015) investigated a study on the sources and collection process of MSW and existing management practice in Chittagong City Corporation. It has been found that total 122 tons of MSW of different type are being generated daily. The responsibility of the management is of them is mostly an authority in community system and a few on the house owner. Of the total quantity 45-55 percent is collected efficiently and the rest is left. It is broadly estimated that between 10-15 percent



of the total municipal budget is used for MSW management. The MSWs which are remain uncollected and dumped in open spaces, street and drains, clogging the drainage system that creates serious environmental degradation and health risks. The collected MSW is presently being disposed of mainly in a low-lying area. Few amount of MSW are being reclaimed or salvaged for recycling according to its market value.



Figure 2.6: Present scenario of MSW disposal site in Dhaka (Matuail landfill) (Source: <https://jessicamudditt.com/2011/02/04/dhakas-largest-waste-site-photographs-of-matuail-landfill/>).

A blog published by Jessica Mudditt (2011) informed about four thousand tonnes of MSW are deposited at Matuail Landfill, is the largest MSW site in Dhaka as it is responsible for 65 percent of the total MSW generated. Figure 2.6 shows the present scenario of matuail landfill site. However, according to a 15 month study conducted in 2003 by the Japanese International Cooperation Agency (JICA), only 44 percent of all MSW generated is collected. That means that 1,200 tons of garbage swamps Dhaka's public places every day. When garbage is illegally deposited into waterways, the fishermen lose their livelihood; and the fish, their habitat. JICA's urged Dhaka City Corporation, a self-governing corporation that is associated with the task of running the affairs of the city since 1864, to increase the scope of its MSW maintenance facilities in order to combat the odour, drain clogging, pollution and mosquitoes that afflicts many parts of Dhaka.

In Khulna, the waste disposal site located at Rajbandh, with an area of 5 acres, is 8 km far from the city centre i.e. Royal and Castle Salam Square of Khulna city and situated along

the North-side of Khulna-Satkhira highway. Khulna is one of the fast growing commercial cities in Bangladesh with a population of 1.9 million which produce about 450-520 tons of MSW per day.

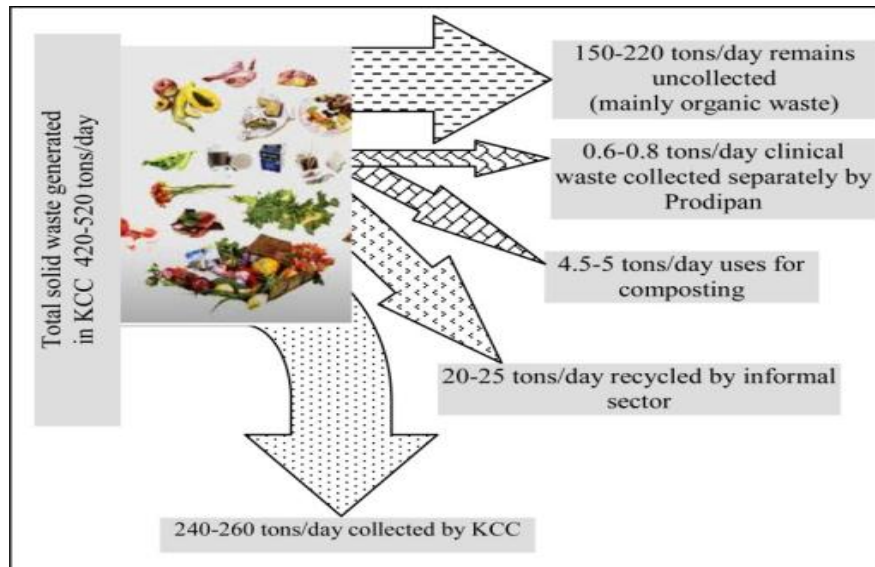


Figure 2.7: Amount of MSW collected in Khulna city (in ton/day) (Source: Aborjona and Paribesh, <http://www.wasteconcern.org/newsletters/issue5/issue5.html>).

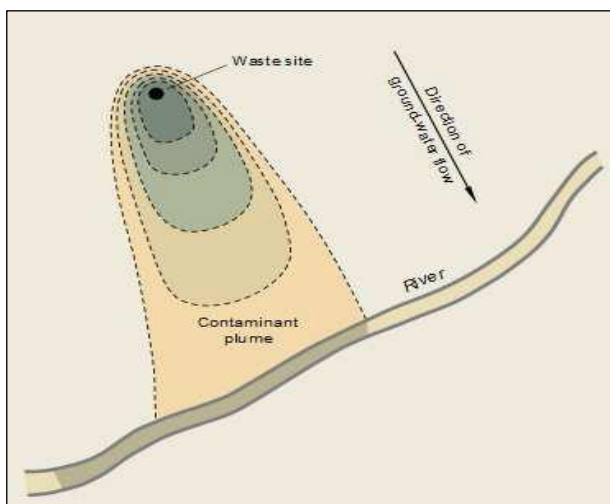
Khulna City Corporation (KCC), community based organizations (CBOs) and non-government organizations (NGOs) are taking care of only 42% of the total MSW generated while the rest of them are unattended. In fact, most of the MSWs are collected from door-to-door without any sorting and either dumped in open space or improperly landfilled which is likely to contaminate the air and ground water. Figure 2.7 shows the amount of collected MSW in Khulna city by both the government and non-government organizations. Due to inadequate management practices and the low standard sanitary landfill, the leachate percolates into the groundwater and contaminates the groundwater source which is a potential threat to next generation. Necessities arise to take steps for proper disposal of MSW as well as maintenance of MSW disposal site in different cities of Bangladesh.

## 2.6 Distribution Pattern of Heavy Metals in Soil

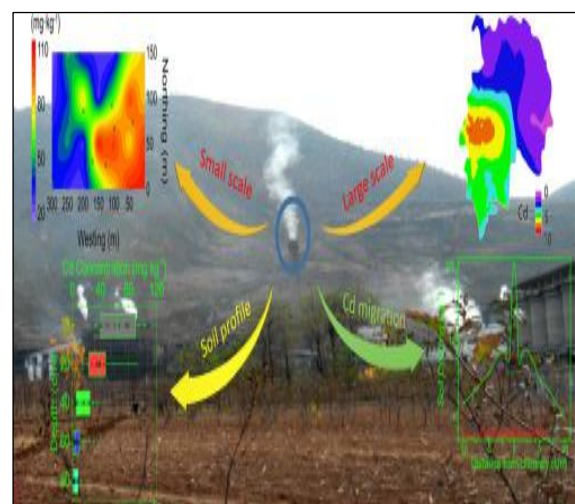
The interpolated map using geostatistics provides the best and simpler way to comprehend the risk of contamination zone depending on the metal concentrations plotted with the optimal interpolation model exhibits the peak distribution for metal contamination. Soil

maps convey information about the soil to land users. In the United States, soil map units are grouped based on the underlying principles of Soil Taxonomy (Soil Survey Staff, 1999). Known information about a map unit can be used to extrapolate those properties to similar map units across landscapes and regions. Soil properties are generally considered to be continuous across a landscape (Isaake and Srivastava, 1989). Within map units, there is a varying degree of heterogeneity, can lead to difficulties in using the map to predict properties at any point within or along the boundary of a soil map unit. Despite these problems, soil maps provide a readily available source of soil property information across a landscape. These maps can be used to predict and model properties that were not explicitly classified in the map itself.

Figure 2.8 shows the distribution pattern of pollutants in a river and landfill site (Smith et al., 2007). Figure 2.8 (a) shows a typical diagram of distribution pattern of extent of heavy metals in groundwater layer due to presence of MSW site near a river. The deep ash color represents high contamination of pollutants, whereas with the increased distance the effect of pollutant decreases. Thus the map shows light brown color to indicate less contamination. Figure 2.8 (b) shows a diagram of contamination of cadmium in soil due to incineration of MSWs dumped in a landfill, where different color zone is assigned for different contamination level, which provide a simple knowledge of amount of contamination in soil layer. The red color represents contamination hotspots and blue color represents less contamination zone.



(a)



(b)

Figure 2.8: Contaminant plume in (a) groundwater; (b) soil (Source: US.EPA, 1993).

Several studies were performed to show distribution pattern of heavy metal concentration in soils. A case study of Beijing, China by Zou et al. (2015) found that that the peak distribution of heavy metal was in the northwest for Cu, Cd, Pb and Zn; in the southeast for As; mainly on the urban fringe for Hg; and mainly in the southwest for Cr and Cd. A study of heavy metal contamination of urban soil was done by Li et al. (2012) to analyze the impacts of land use on the heavy metal pollution of soils and development of the maps of the spatial distribution of heavy metal concentrations in an old industrial city (shenyang) in northeast china.

## **2.7 Conventional Statistics**

Statistical theory delivers a guide to compare methods of data collection, where the problem is to generate informative data using optimization and randomization while measuring and controlling for observational error. In this study, conventional statistics includes normality test to determine whether the data is normally distributed or not. In addition descriptive test offers quantitative descriptions in a manageable form.

### ***2.7.1 Normality Test***

Normality tests are used to verify whether a data set is well-modeled by a normal distribution and to compute how probable it is for a random variable underlying the data set to be distributed normally. The choice of a non-parametric test is important as in the real world we are mostly faced with data which fails to meet the assumptions of normality and stationarity underlying parametric tests. In this study, normality test was performed by K-S test and Shapiro-wilk test using SPSS. Nevertheless, for more accuracy normal QQ plot was also plotted to check the normality and hence discussed in the following articles.

#### ***2.7.1.1 Kolmogorov-Smirnov Test***

The Kolmogorov-Smirnov (K-S) test is a hypothesis test procedure for determining if two samples of data are from the same distribution. The K-S test is presented for distinguishing of distribution of forecasts errors and identifying the model with the lower stochastic error. The two-sample one-sided K-S test aims at ascertaining whether the forecast with the

lowest error according to some loss function also has a stochastically smaller error in comparison to the competing forecast and thereby enables the comparison of the predictive accuracy of forecasts. A study conducted by Yasrebi et al. (2009) and performed K-S test to check either the heavy metal parameters are normally distributed or not. Lu et al. (2012) also used K-S test in a study of multivariate and geostatistical analyses of the spatial distribution and origin of heavy metals in the agricultural soils in Beijing, China to check normality of the data points and it was normally distributed by applying log transformation. In the present study, normality test was performed using SPSS to test the accuracy of the statistical analysis.

#### ***2.7.1.2 Shapiro-Wilk Test***

According to Razali and Wah (2011), the Shapiro-Wilk (SW-test) test is based on the correlation between the data and the corresponding normal scores and provides better power than the K-S test. Power is the most frequent measure of the value of a test for normality that the ability to detect whether a sample comes from a non-normal distribution (Conover, 1999). Some researchers recommend the S-W test as the best choice for testing the normality of data. On the other hand, Shapiro and Wilk (1965) recommended that the limitation of this test is what feature of the distribution is non-normal. In the present study, S-W test was also performed to test the normality.

#### ***2.7.1.3 Normal Quantile-Quantile Plot***

The normal quantile-quantile (QQ) plot is implemented to find the linearity of statistical data where the S-W test does not utilize deviations from the theoretical distribution function. According to Razali and Wah (2011) research, based on a visual inspection in a Q-Q plot, a sample is therefore considered to be consistent with a normal distribution if the empirical and theoretical quantiles fall close to the line representing the theoretical distribution. This decision is helped additionally by an assessment of whether the points fall inside the envelope of 95% pointwise confidence intervals. Anderson and Darling (1954) also recommended new graphical methods and test statistics for testing composite normality. In the present study, normal QQ plot was plotted for all metal elements using XLSTAT.

### ***2.7.2 Descriptive Statistics***

Heckman (1976) gave an idea about the statistical terms. The median of a data set is the value in the middle when the data items are arranged in ascending order. A study of Stuart and Ord, 1987 found that whenever a data set has extreme values, the median is the preferred measure of central location. It is often desirable to consider measures of variability (dispersion), as well as measures of location. The range of a data set is the difference between the largest and smallest data values. It is the simplest measure of variability. The variance is a measure of variability that utilizes all the data. The standard deviation of a data set is the positive square root of the variance. The coefficient of variation indicates how large the standard deviation is in relation to the mean.

### ***2.7.3 Case Studies of Application of Conventional Statistics in Soil***

Lu et al. (2012) performed an analysis on descriptive statistics of the heavy metal concentrations in the agricultural soils for As, Cd, Cu, Hg, Pb and Zn, respectively. Another case study of Yasrebi (2009) showed to identify the variability of different soil properties. Furthermore, numerous studies were performed for statistical analysis. In this study, conventional statistics including normality test and descriptive statistics were performed to see the concentration distribution as well as temporal variation of heavy metal contamination in and around the landfill site.

## **2.8 Multivariate Statistics**

Multivariate analysis is the area of statistics that deals with observations made on many variables. The main objective of this analysis is to study how the variables are related to one another, and how they work in combination to distinguish between the cases on which the observations are made. In this study, multivariate statistics including Pearson's correlation, principal component analysis (PCA) and agglomerative hierarchical clustering (AHC) were performed using XLSTAT and discussed in the following articles.

### ***2.8.1 Pearson's Correlation***

In statistics, Pearson's correlation is a parametric measurement developed by Karl Pearson from a related idea introduced by Francis Galton in 1880s. The bivariate Pearson's correlation produces a sample correlation coefficient,  $r$ , which measures the strength and direction of linear relationships between pairs of continuous variables.

The bivariate Pearson's Correlation is commonly used to measure the following:

- I. Correlations among pairs of variables.
- II. Correlations within and between the sets of variables.

Pearson's correlation should be used only when there is a linear relationship between variables. It can be a positive or negative relationship, as long as it is significant. Correlation is used for testing in within groups studies. In case of heavy metal contamination in soil, several studies were performed to find the association between metal elements and the similarity of their contamination sources. Zou et al. (2015) studied on sources of heavy metals in farmland soils of Beijing suburbs, China based on Pearson's correlation. It was found that some metal elements such as Cr was moderately correlated with Cd and Zn, whereas, Hg was only correlated with As and not with the other elements as the pair. Another study by Mahmoudabadi on accuracy assessment of geostatistical methods for zoning of metal elements in soil of urban-industrial areas also found that for the elements Cd, Cu, Pb and Zn were derived from different sources due to lack of high correlation. In this study, the principal of Pearson's correlation was performed using XLSTAT to examine the association of metal elements in soil irrespective to their sources.

### ***2.8.2 Principal Component Analysis***

According to the definition given by Abdi and Williams (2010), multivariate statistics is a subdivision of statistics encompassing the simultaneous observation and analysis of more than one outcome variable. The application of multivariate statistics is multivariate analysis. Analysis performed by Anderson and Darling (1954) interpreted that multivariate statistics concerns understanding the different aims and background of each of the different forms of multivariate analysis, and how they relate to each other. Principal component analysis (PCA) estimates those components that contribute most to the variation in the data sets. Large-scale investigations of metal concentrations have been

conducted in China, mainly in the areas with natural background levels and minor anthropogenic pollution sources (Liu et al., 2003). According to Abdi and Williams (2010), the goals of PCA are:

- (1) to extract the most important information from the data table;
- (2) to compress the size of the data set by keeping only this important information;
- (3) to simplify the description of the data set; and
- (4) to analyze the structure of the observations and the variables.

### ***2.8.3 Agglomerative Hierarchical Clustering***

Although not substantially different from PCA, cluster analysis could be used as an alternative method to confirm results and provide grouping of variables. *Statistician manual* refers to cluster analysis comprises a group of statistical techniques that measures similarity or distance between the objects. Agglomerative hierarchical clustering (AHC) accomplishes successive fusions of data into clusters where each object initially starts out as its own cluster. It also differs to the extent that different measures are employed to measure the distance between clusters. A study conducted by Zhang et al. (2012) for heavy metals in agricultural soils in northern china and it showed the correlation between heavy metals by construction of dendrogram.

### ***2.8.4 Case Studies of Multivariate Statistics for Assessment of Heavy Metals in Soil***

The application of multivariate statistical approaches to the problem allows a better classification, modeling, and interpretation of soil monitoring data. This strategy makes it possible to detect relationships between the chemical pollutants and specific soil parameters, between sampling sites and, therefore, to achieve a stratification of the pollution (Zou et al., 2015). Further, it becomes possible to identify possible pollution sources and to construct apportioning models allowing the determination of the contribution of each identified source to the formation of the total pollutant mass (Einax and Soldt, 1995).



Table 2.1: Several researches on application of multivariate statistics for assessment of soil pollution

Authors	Title of the Research	Location	Objectives of the Research	Findings of the Research
Maria et al. (2010)	Soil Contamination Interpretation by the Use of Monitoring Data Analysis	Bulgaria	to identify possible pollution sources.	Two major clusters were found to explain the sampling site locations according to soil composition i.e. one cluster for coastal and mountain and another—for typical rural and industrial sites
Lu et al. (2012)	Multivariate and geostatistical analyses of the spatial distribution and origin of heavy metals in the agricultural soils in Shunyi, Beijing, China	China	to determine possible sources of heavy metals in agricultural land in Shunyi	The analysis suggested the heavy metals of Cd, Cu and Zn derived from agricultural practices, whereas As and Pb from parent materials and Hg from atmospheric deposits.
Xiaoyu et al. (2013)	Heavy metal contamination of urban soil in an old industrial city (Shenyang) in Northeast China	China	to identify the relationship between urban land use and the heavy metal contamination of urban soil	Industrial District was polluted by heavy metals and that the high concentrations of Pb, Cu, Zn, Cd and Hg were not only distributed in industrial areas, but were also widely distributed in residential areas and parks.
Zou et al. (2015)	Analysis of Spatial Variations and Sources of Heavy Metals in Farmland Soils of Beijing Suburbs	China	To provide comprehensive understanding of heavy metal distributions in the farmland soils of the Beijing	The Cu, Cd, Pb and Zn pollutant levels were mainly affected by the road distribution and land use status, Hg derived from industrial MSW, Ni had no significant effects.
Lee et al.(2006)	Metal contamination in urban, suburban, and country park soils of Hong Kong: A study based on GIS and multivariate statistics	Hong Kong	to assess and compare metal contamination in soils of urban, suburban and country park areas of Hong Kong	The peak distribution was in the northwest for Cu, Cd, Pb and Zn; in the southeast for As; mainly on the urban fringe for Hg; and mainly in the southwest for Cr and Cd.
Lu et al. (2010)	The spatial distribution and sources of metals in urban soils of Guangzhou, China	China	to identify possible sources for these metals	Fe, Ni and Mn are derived from natural sources; As, Cu, Hg, Pb and Zn from anthropogenic sources and Cd from both sources.
Candeias et al. (2011)	The use of multivariate statistical analysis of geochemical data for assessing the spatial distribution of soil contamination by potentially toxic elements in the Aljustrel mining area	Portugal	To identify possible sources of contamination that can explain the spatial patterns of soil pollution in the area;	The source of heavy metals originated from geogenic or anthropogenic origin
Fahad et al. (2016)	Spatial distribution and source identification of heavy metal pollution in roadside surface soil: a study of Dhaka	Bangladesh	to assess the concentration and distribution patterns of heavy metals in urban soil	high positive loading is found near the Boilapur-Amin Bazar landfill site (PC1), near Boilapur (PC2), and near the Noyahati-Amin Bazar landfill site (PC3). CA formed three major clusters for both water parameters and sampling site.

Multivariate analytical techniques give more details on the data structure collected from environmental media than univariate methods since multiple parameters are measured during soil, sediment, water and air quality analysis. In soil analytical studies, a variety of geostatistical tools with a wide range of applications exist (Facchinelli et al., 2001). Zamani et al. (2012) made a research on multivariate statistical assessment of heavy metal pollution sources of groundwater. The results illustrated PC1 with Pb and Cu are important byproducts of lead industries indicating its anthropogenic sources. PC2 reveals 24.4% of the total variances are positively loaded with Zn and negatively loaded with Co. PC3 shows that 17.4% of total variance is positively loaded with Ni. PC4 explains 15.1% of total variance, is positively loaded with Cd and Fe.

Multivariate statistical analysis has been widely used for source apportionment of metals in soil and water in different parts of the world. In context of Bangladesh, a research was done by Bhuiyan et al. (2016) to assess the groundwater quality of lakshimpur district of Bangladesh using water quality indices, geostatistical methods and multivariate analysis. The results of principal component/factor analysis indicate that anthropogenic (agrogenic, surface runoff, and domestic sewage) and natural/geogenic sources (weathering of source rock) are responsible for variation in physicochemical parameters and metal contents in groundwater systems at the southeastern coastal region of Bangladesh. Table 2.1 provides a summary of several research cases on application of multivariate statistics for the assessment of contamination of soil. In this research, Pearson's correlation offers the correlation between the metal elements irrespective to their sources. In addition, PCA and AHC were performed to evaluate the origin of metal elements presence in soil and correlation between these metal elements. The results obtained by multivariate statistics can be used as a guideline to reduce the roots of metal elements spread over the disposal site area.

## **2.9 Geostatistical Analysis**

Geostatistics is a subset of statistics specialized in analysis and interpretation of geographically referenced data. Moreover, Geostatistics is a class of statistics used to analyze and predict the values associated with spatial or spatiotemporal phenomena. It

incorporates the spatial (and in some cases temporal) coordinates of the data within the analyses. Many geostatistical tools were originally developed as a practical means to describe spatial patterns and interpolate values for locations where samples were not taken (Nielsen and Wendroth, 2003). Those tools and methods have since evolved to not only provide interpolated values, but also measures of uncertainty for those values. The measurement of uncertainty is critical to informed decision making, as it provides information on the possible values (outcomes) for each location rather than just one interpolated value.

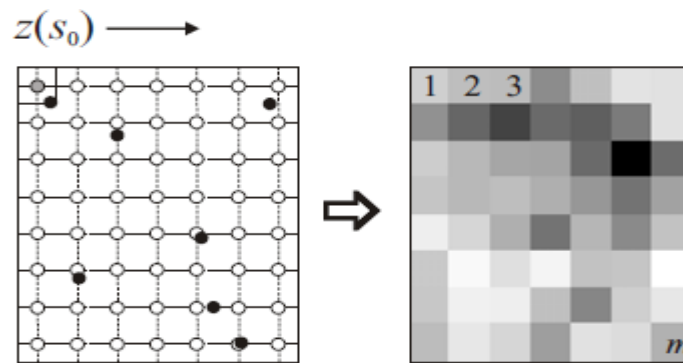


Figure 2.9: Spatial prediction implies application of a prediction algorithm to an array of grid nodes (point ‘a point spatial prediction). The results are then displayed using a raster map (Source: Heckman, 1976).

Geostatistical analysis has also evolved from uni- to multivariate and offers mechanisms to incorporate secondary datasets that complement a (possibly sparse) primary variable of interest, thus allowing the construction of more accurate interpolation and uncertainty models. According to the bibliographic research of Nielsen and Wendroth, 2003, the top 10 application fields of geostatistics are: (1) geosciences, (2) water resources, (3) environmental sciences, (4) agriculture and/or soil sciences, (5/6) mathematics and statistics, (7) ecology, (8) civil engineering, (9) petroleum engineering and (10) limnology. One of the main uses of geostatistics is to predict values of a sampled variable over the whole area of interest, which is referred to as spatial prediction or spatial interpolation. An important distinction between geostatistical and conventional mapping of environmental variables is that the geostatistical prediction is based on the application of quantitative and statistical techniques. A guideline by EPA defines spatial distribution as the graphical display of an arrangement of a phenomenon across the earth, is an important tool in

geographical and environmental statistics. Geostatistical mapping represents analytical production of maps by using field observations, auxiliary information and a computer program that calculates values at locations of interest. Figure 2.9 shows the mechanism of establishing interpolation map of predicting unsampled points using geostatistics. The geostatistics and ArcGIS are useful tools for the identification of pollution sources, the assessment of pollution trends and the potential risks of heavy metals. In this study, the most widely used geostatistical interpolation techniques such as inverse distance weighting (IDW), local polynomial interpolation (LPI), radial basis function (RBF) and ordinary kriging (OK) were used to predict the spatial distribution of metal elements with best fitted model.

### ***2.9.1 Inverse Distance Weighting***

Inverse distance weighting (IDW) is a type of deterministic method for multivariate interpolation with a known scattered set of points. The assigned values to unknown points are calculated with a weighted average of the values available at the known points. The name given to this type of methods was motivated by the weighted average applied, since it resorts to the inverse of the distance to each known point, when assigning weights (Pebesma, et al., 2007). Greater weighting values are assigned to values closer to the interpolated point. As the distance increases the weight decreases (ESRI, 2001) and weighting power that decides how the weight decreases as the distance increases. Numerous researches were performed to show the spatial distribution of heavy metals in soil and assessment of accuracy of interpolation techniques using IDW through GIS technique. In this study, the technique of IDW with power 1 to 5 was performed to find out the best interpolation technique to provide spatial distribution of metal elements in soil of waste disposal site to understand the actual field condition.

### ***2.9.2 Local Polynomial Interpolation***

Local polynomial interpolation (LPI) is a moderately quick deterministic interpolator that provides prediction, prediction standard error and condition number surfaces that are comparable to ordinary kriging with measurement errors (Smith et al., 2007). It is more flexible than the global polynomial method, but there are more parameter decisions. Local

polynomial methods do not allow you to investigate the autocorrelation of the data, making it less flexible and more automatic than kriging. There are no assumptions required of the data (ESRI, 2001).

LPI fits the local polynomial using points only within the specified neighborhood instead of all the data. A first-order global polynomial fits a single plane through the data.; a second –ordedr global polynomial fits a surface with a bend in it, allowing surfaces representing valleys; a three-order global polynomial allows for two bends. Order of polynomial greater than 3 is not generally recommended for most situations due to its high critical spatial condition number thershold value which indicate less reliability. Then the neighborhoods can overlap, and the surface value at the center of the neighborhood is estimated as the predicted value. Numerous studies were done to examine the performance of LPI for spatial distribution of heavy metals in soils. In this study, the technique of LPI with order 1 to 3 was performed to find out the best interpolation technique to provide spatial distribution of metal elements in soil of waste disposal site to understand the actual field condition.

### ***2.9.3 Radial Basis Functions***

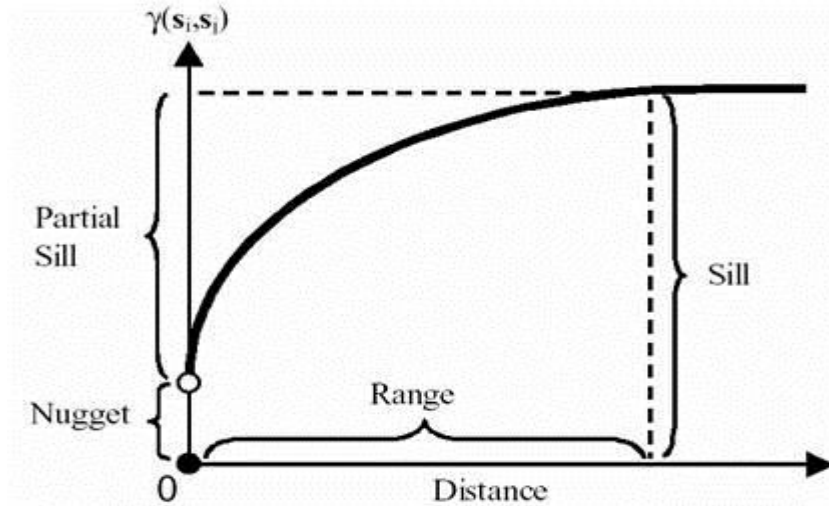
Radial Basis Functions (RBF's) are moderately quick deterministic interpolator . This method provides prediction surfaces that predicts values identical with those measured at the same point and the generated surface requires passing through each measured point. Xie et al. (2011) suggested that the predicted values can vary above the maximum or below the minimum of the measured values. A research of Johnston et al. (2001) identified RBF fits a surface through the measured sample values while minimizing the total curvature of the surface. RBF is ineffective when there is a dramatic change in the surface values within short distances (Xie et al., 2011). In this study, the technique of RBFs were performed to find out the best interpolation technique to provide spatial distribution of metal elements in soil of waste disposal site to understand the actual field condition.

#### ***2.9.4 Ordinary Kriging***

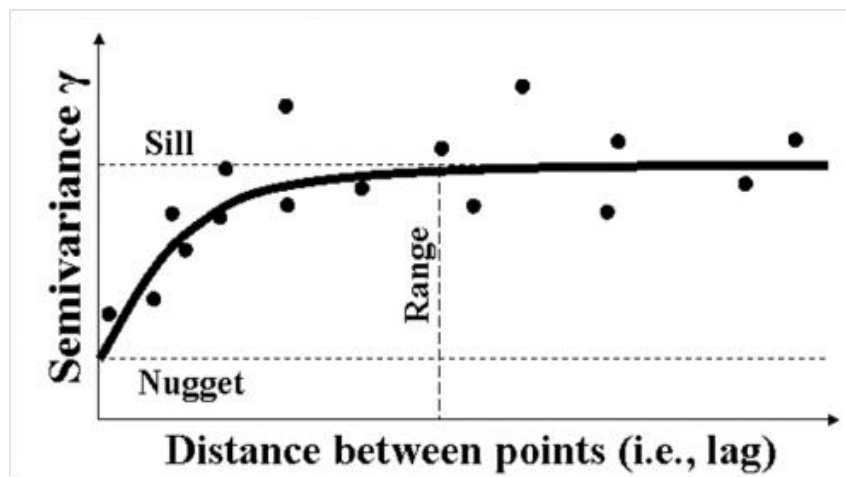
Kriging is deterministic interpolation method that focuses on estimation that gives the best unbiased (minimum variance) linear estimate of point values or block averages. The method was created by Dr. D.G. Krige, for the South African mine fields and was improved by Professor G. Matheron into the techniques used today (Zhang, 2011). Ordinary kriging (OK) is the original form of Kriging, is the most widely used, and the most robust form (Zhang, 2011). Ordinary Kriging is usually referenced with the acronym B.L.U.E. meaning best linear unbiased estimator. Ordinary kriging is “linear” because its estimates are weighted linear combinations of the data used in analysis; it is “unbiased” since it tries to have the mean residual error equal zero; is “best” because it aims at minimizing the variance of the errors associated with the analysis (Isaaks and Srivastava 1989). The whole idea behind ordinary kriging is to find the optimal weights that minimize the mean square estimation error (Yan et al., 2015). The assumptions made in Ordinary Kriging are: 1) the sample is the partial realization of a random function  $Z(s)$ , where  $(s)$  denotes the spatial location; 2)  $Z(s)$  is second order stationary or can honor the intrinsic hypothesis; 3) the mean is constant but unknown (no trend). In this study, the technique of ordinary kriging was performed to find out the best fitted model to provide spatial distribution of metal elements in soil of waste disposal site to understand the actual field condition.

#### ***2.9.5 Semi-variogram and its Estimation***

The spatial variability of the soil heavy metal pollution was described using semivariogram methodology, which reflects the variations between two observed values at different distances. The variogram is the basic geostatistical tool for measuring spatial autocorrelation of a regionalized variable (Isaaks and Srivastava 1989). Semi- variogram models are the spatial structure of the regionalized variable and provides weighting information to the kriging algorithm for interpolation (Yan et al., 2015). Kriging uses the semi-variogram to predict values at unobserved location using minimization of errors (Krige, 1951). These principles have been applied to soil science for over two decades (Burgess and Webster, 1980).



(a)



(b)

Figure 2.10: Typical diagram of semivariogram (Source: Heckman, 1976).

Upon completion of the semivariogram plot a mathematical function is then applied to fit the data. A properly fitted mathematical model then allows for a computer program to calculate linear estimates that reflect the spatial extent and orientation of autocorrelation in the variable to be mapped. All variograms that have mathematical functions fitted to them have a range and a sill by which predictions are no longer valid due to a lack of correlation between data points (Isaaks and Srivastava 1989). This sill is the limiting value of the mathematical model and the range is the lag distance from the origin of the variogram to sill. Figure 2.10 is a typical example of a semi variogram indicating different parameters to give a clear conception for interpreting these terms.

Variograms that reach a sill and have a finite range are named as bounded and variograms that do not reach a sill are named as unbounded. In many cases mathematical model fitting can be a difficult task to achieve. In those cases it may be necessary to change the lag intervals by increasing or shortening them to achieve a more distinct trend (Candeias et al.,2011)). A study of Yang et al. (2013) described semivariogram results nugget to sill ratios representing nature of heterogeneity. The nugget effect may be caused by random factors highlighted the stronger spatial correlation.

#### ***2.9.6 Case Studies of Application of Geostatistics of Heavy Metals in Soil***

New computer applications have allowed more wide-spread development and use of geostatistical techniques. However, there are still many applications of geostatistics that have not yet been explored. The purpose of this study is to use geostatistics to evaluate and compare the spatial distribution of soil properties in a native landfill. Numerous studies were performed to assess the spatial distribution metal elements present in soil. Yang et al. (2013) examined ability of kriging interpolation by creating heavy metals maps at a high level of spatial resolution in soil of Norway. Xie et al. (2011) showed the spatial distribution of heavy metal in soil Beijing using different interpolation methods, and to investigate the relationship between the accuracy of prediction and local variation in soil heavy metal content. Accuracy of geostatistical interpolation techniques such as kriging, cokriging and IDW, was also examined by Mahmoudabadi et al. (2012) for zoning of heavy metals in soils of urban-industrial areas of Iran.

A study of Xie et al. (2011) gave a detailed discussion on different types of interpolation techniques. According to his statement, IDW is an example of a gradual, exact, mathematical interpolator in which points closer to the measured data receives more weight in the averaging formula. He also stated that RBF can predict values above the maximum or below the measured number. Global polynomial interpolation accounts for bends in the data; whereas Surfaces that do not display a series of bends, however, such as one that increases, flattens out, and increases again, can be better represented using local polynomial interpolation. The variability in the kriging estimates is less than the variability of the unobserved, true spatial process. He stated that spatial variation is analyzed using variograms, which plot the variance of paired sample measurements as a function of



distance between samples. The spatial distribution and source identification of heavy metal pollution in roadside surface soil in Dhaka Aricha highway. Table 2.2 enlisted some of previous studies of application of geostatistics with their aims and goals obtained. In this study, spatial distribution of heavy metals in and around the soil of studied landfill site is represented in a better and simplified way. The maps produced from spatial analysis provide knowledge of pattern of contamination level of heavy metals in form of map. Different colors attributed to different concentration levels indicate the how intensely the region is contaminated and how immediate steps should be taken to reduce the contamination.

Table 2.2: Several researches on application of geostatistics for assessment of soil pollution

Authors	Title of the Research	Location	Objectives of the Research	Findings of the Research
Maria et al. (2010)	Soil Contamination Interpretation by the Use of Monitoring Data Analysis	Bulgaria	to get information about some spatial distribution of the soil pollutants.	Accurate shaded maps were presented according to surface and subsurface soil layers.
Lu et al. (2012)	Multivariate and geostatistical analyses of the spatial distribution and origin of heavy metals in the agricultural soils in Shunyi	China	to determine the spatial distribution of heavy metals in agricultural land in Shunyi	Contaminants were displayed in form of geostatistical maps.
Xiaoyu et al. (2013)	Heavy metal contamination of urban soil in an old industrial city (Shenyang) in Northeast China	China	to assess the concentration and distribution patterns of heavy metals in urban soil	The estimated maps of Cr, Zn, Mn, Cd, As and Hg are presented several hotspots of high metal concentration; Pb and Cu showed a very similar spatial pattern, indicating from same sources.
Zou et al. (2015)	Analysis of Spatial Variations and Sources of Heavy Metals in Farmland Soils of Beijing Suburbs	China	To provide comprehensive understanding of heavy metal distributions in the farmland soils.	The Cu, Cd, Pb and Zn pollutant levels were higher than Hg; Ni had no significant effects.
Lee et al. (2006)	Metal contamination in urban, suburban, and country park soils of Hong Kong: A study based on GIS and multivariate statistics	Hong Kong	to evaluate the relationship between heavy metals and their possible sources using GIS spatial analysis	The urban and suburban soils were highly enriched with metals such as Cu, Pb, and Zn, in comparison with the country park soils. The urban soils were found to be more contaminated than the suburban soils.
Lu et al. (2010)	The spatial distribution and sources of metals in urban soils of Guangzhou, China	China	to determine concentrations of metals including variability and spatial distribution patterns	The spatial distribution maps of As, Cd, Cu, Hg, Pb and Zn concentrations displayed several hotspots of heavy metal pollution.
Candeias et al. (2011)	The use of multivariate statistical analysis of geochemical data for assessing the spatial distribution of soil contamination by potentially toxic elements in the Aljustrel mining area	Portugal	to determine the associations between the different toxic elements and their spatial distribution	it is possible to conclude that soils around Algaes/Feitais tailing deposits, Este´reis and A´guas Claras mine dams and S. Joa˜o mine showed severe contamination.

## 2.10 Artificial Neural Network

A study conducted by Shahin et al. (2004) and stated that over the last few years or so, the use of artificial neural network (ANN) has increased in many areas of engineering. In particular, ANN has been applied to many geotechnical engineering problems and have demonstrated some degree of success. A review of the literature reveals that ANN has been used successfully in pile capacity prediction, modeling of soil behaviour, site characterization, earth retaining structures, settlement of structures, slope stability, design of tunnels and underground openings, liquefaction, soil permeability and hydraulic conductivity, soil compaction, soil swelling and classification of soils (Rooki et al.,2011). Figure 2.11 shows a typical architectural model of ANN.

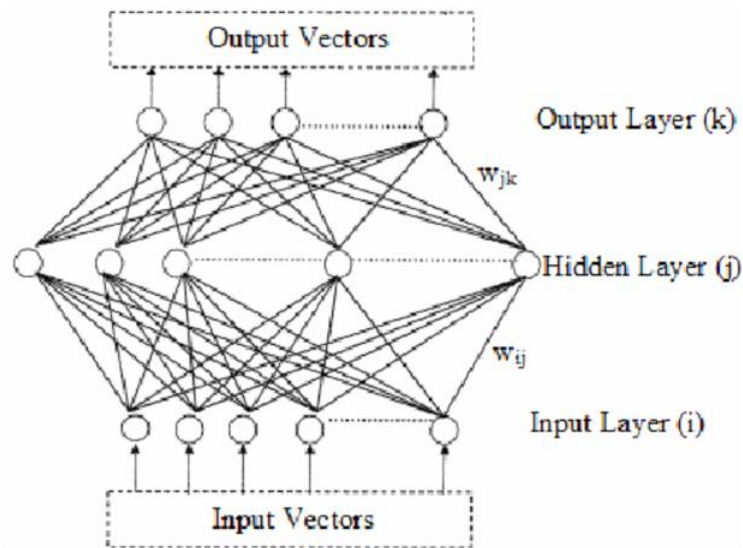


Figure 2.11: Typical ANN architecture model (Source: Alkaiem et al., 2016).

Back propagation is the best known training algorithm for neural networks and has the best performance. Shahin et al. (2004) found that the Marquardt-Levenberg method, based on Gauss-Newton's equations can be used for training the network with a neuron in the output layer to minimize the square sum of the nonlinear objective function. A studies conducted by Yun and Uchimura (2007) used the SOM algorithm for spatial clustering of health and census data. Standard SOM algorithms have been applied to remote-sensing studies for classifications and spatial clustering. Rooki et al.(2011) made a research on prediction of heavy metal in acid mine drainage using ANN. Alkaiem et al. (2016) used of artificial

intelligence techniques to predict distribution of heavy metals in groundwater of Lakan lead-zinc mine in Iran. In that study, the ANN was developed to estimate the heavy metals concentrations in groundwater using  $\text{SO}_4$ , Cl, and TDS as input parameters, and Fe, Mn, Pb, and Zn as output parameters and the performance of ANN was satisfactory. Mehidi et al. (2014) used ANN process to assess risk in cement industries in Bangladesh and it was found that ANN is most flexible and useful tool for predicting risk in any industry (Yun and Uchimura., 2007). Furthermore, more studies as needed to understand this problem. In the present study, ANN was used to predict concentration of higher depth that has not been tested in laboratory as well as the accuracy of predicted values.

## CHAPTER III

### RESEARCH METHODOLOGY

#### 3.1 Introduction

This chapter deals with the overall research methodology in this study. It includes the information about the study area and soil condition. In this study, total sixty disturbed soil samples were collected at a depth of 0-30 cm from the existing ground surface from different selected locations within the waste disposal site in both the dry season (March to May, 2016) and rainy season (June to August, 2016). The method of soil sampling was also highlighted in this chapter. In the laboratory, the concentration of metal elements such as aluminium (Al), arsenic (As), barium (Ba), calcium (Ca), iron (Fe), mercury (Hg), potassium (K), manganese (Mn), sodium (Na), nickel (Ni), lead (Pb), antimony (Sb), scandium (Sc), strontium (Sr), titanium (Ti), vanadium (V) and zinc (Zn) in soil were measure and monitor through standard test methods. Then, the measured concentrations of metal elements were used to perform the descriptive conventional statistics using Statistical Package for the Social Sciences (SPSS) 16 to assess the basic features of soil data in a simpler, but meaningful way.

Firstly, the normality test was performed through Kolmogorov–Smirnov test using XLSTAT. The Pearson correlation coefficient was accomplished to measure the linear correlation between metal elements in soil using SPSS. The principle component analysis (PCA) was implemented to explain the variance-covariance structure of variables through linear combinations using XLSTAT. In addition, the cluster analysis through agglomerative hierarchical analysis (AHC) was also performed using XLSTAT to classify the metal elements on the basis of dissimilarity between sets of observations. To find out the spatial distribution of metal elements in soil GIS was performed and also highlighted in this chapter. The different geostatistical interpolation techniques such as inverse distance weighting (IDW), local polynomial (LP) interpolation, radial basis function (RBF) interpolation and ordinary kriging (OK) were implemented to predict the accuracy of spatial distribution of the concentrations of metal elements in soil and the possible sources

of contamination of soil. The performance of these interpolation techniques were assessed on the basis of indices such as mean absolute percentage error (MAPE), relative improvement (RI) and goodness of prediction (G-value). Furthermore, a model was developed using nntool of artificial neural network (ANN) to predict the concentration of metal elements presence in soil of the waste disposal site using MATLAB. For evaluating predicted values of metal elements in soil using ANN, the value of mean standard error (MSE) and regression coefficient (R-value) were considered to check the validity or accuracy of obtained results from laboratory and also described in this chapter.

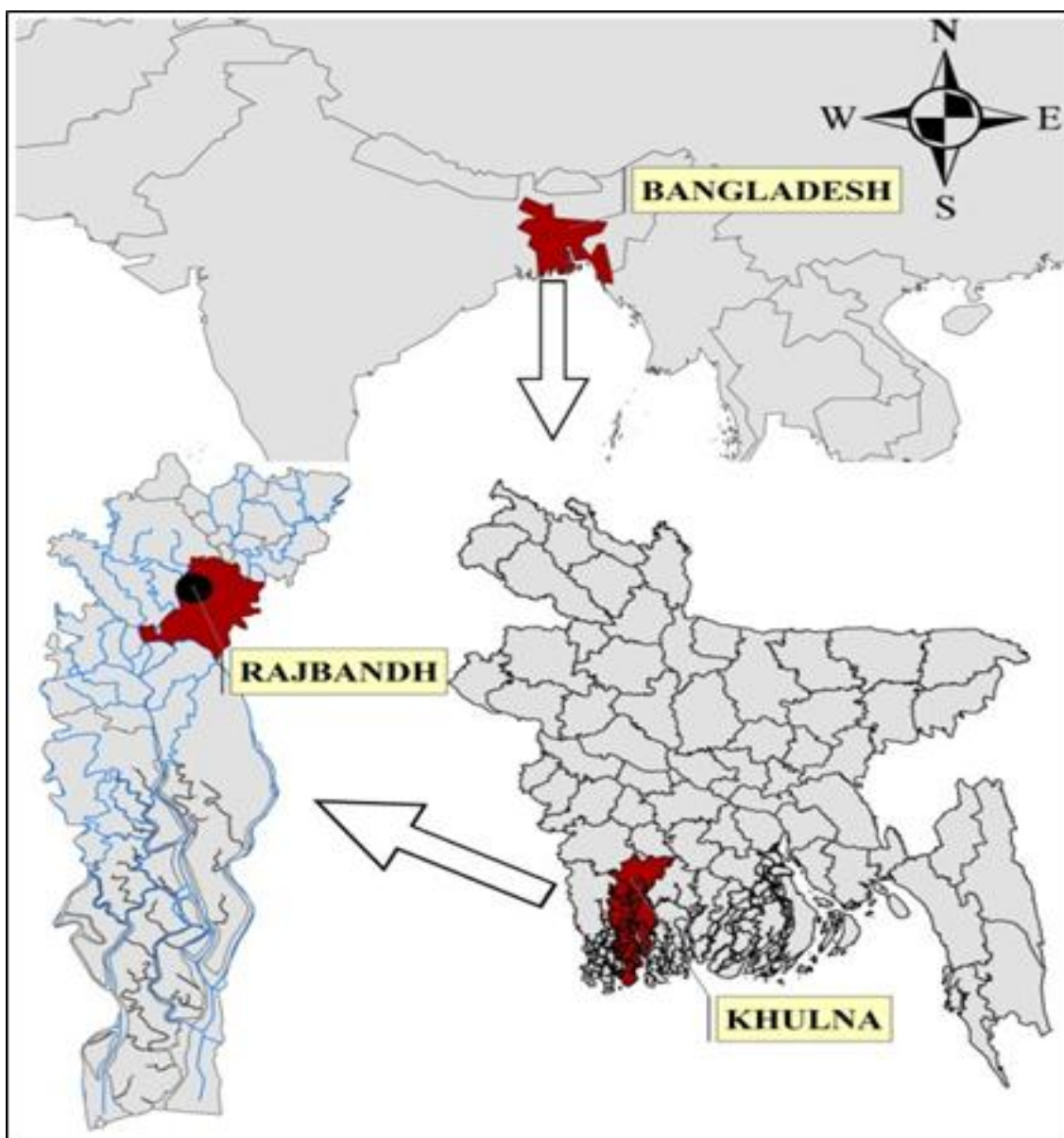


Figure 3.1: Location map of Rajbandh at Khulna city of Bangladesh (Source: Aborjona and Paribesh, <http://www.wasteconcern.org/newsletters/issue5/issue5.html>).

### **3.2 Description of Study Site**

Khulna is the third established metropolitan city of Bangladesh. It is located in the Khulna Division. It has an area of 4394.45 km<sup>2</sup> and is bordered on the north by the Jessore district and the Narail district, on the south by the Bay of Bengal, on the east by the Bagerhat District, and on the west by the Satkhira district. The geological location of Khulna is 22.35<sup>0</sup>N and 89.30<sup>0</sup>E, surrounded by Rupsa, Arpangachhia, Shibsas, Pasur, and the Koyra. Urban development is dribbling into neighboring zones to the North and West results a huge amount of waste generation. The areas of KCC and KCPA are 45.65sq.km and 69.50sq.km, respectively. The increasing population in Khulna city tends to dispose increasing amount of municipal solid waste (MSW) as well as liquid waste termed as leachate. These MSW contain a large amount of metal elements which get direct contact to the environment. This may result a great threat to the environment and human health. The selected waste disposal site, Rajbandh is the only certified waste dumping site of Khulna shown in Figure 3.1. Based on aforementioned authenticities, it has become inevitable of comprehensive study of distribution of metal elements in soils ascends in the vicinity of the Rajbandh waste disposal site.

### **3.3 Location and Soil Conditions of Waste Disposal Site**

The waste disposal site located at Rajbandh, Khulna with an area of 5 acres, is 8km far from the city centre i.e. Royal & Castle Salam Square of Khulna city and situated along the North-side of Khulna-Satkhira highway. The percolation and seepage capacity of leachate from MSW in disposal sites depends mainly on the basic characteristics of the soils (Daniel and Koerner, 1995). It is therefore important to know the physical and mechanical properties of underlying soils as thoroughly as possible before assessing their physico-chemical or hydro-mechanical behavior. Based on these concepts, in laboratory through standard ASTM (2004) methods, some relevant physical and mechanical properties of soil from this selected disposal site were determined by Rafizul (2011). Soil moisture content, plastic limit, liquid limit, plasticity index and shrinkage limit of clay soil used as CCL were found 22, 22, 43, 21 and 16 %, respectively. In addition, the percentages of soil constituents were found as sand, silt and clay of 10, 56.6 and 33.4%,

respectively. Then the value of soil pH, optimum moisture content, maximum dry density and coefficient of permeability were found 6.7, 18 %, 16 kN/m<sup>3</sup> and 1.90x10<sup>-7</sup> cm/sec, respectively.

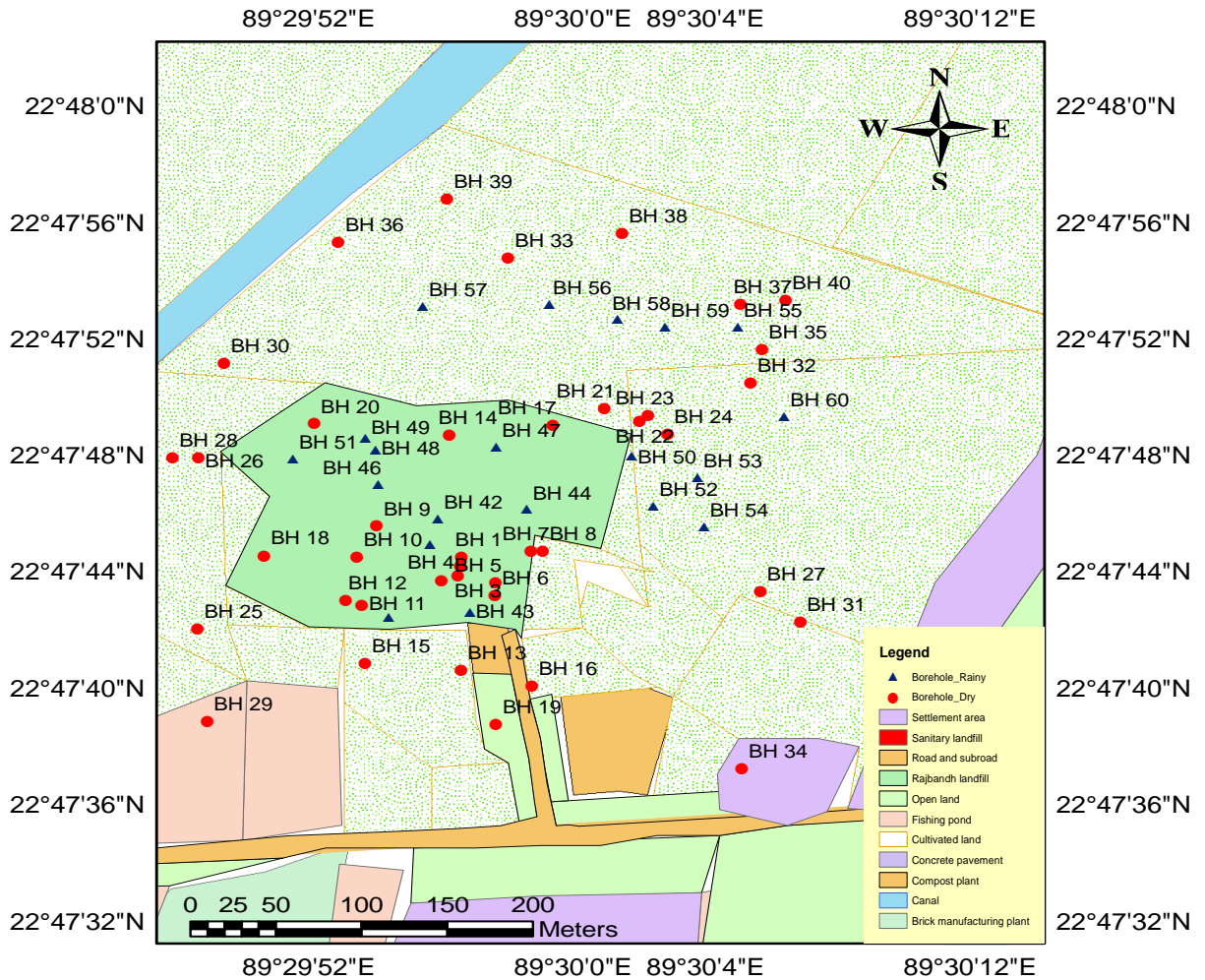


Figure 3.2: Map showing of soil sampling locations in waste disposal site.

### 3.4 Soil Sampling

In this study, total sixty soil samples were collected from the distinct locations of the waste disposal site (Figure 3.2). All the samples were collected at a depth of 0-30 cm from the existing ground surface of waste disposal site. The latitude and departure of all the soil sampling locations was recorded using GPS device, which were later imported into a geographic information system (ArcGIS 10.1). In total sixty soil samples, forty samples were collected in dry season (April, 2016) then rest twenty samples were collected in rainy season (June, 2016). In dry season, the sampling points were selected maintaining gradual



addition of about 10 m distance from the 1st borehole (BH-1) by the subsequent boreholes.

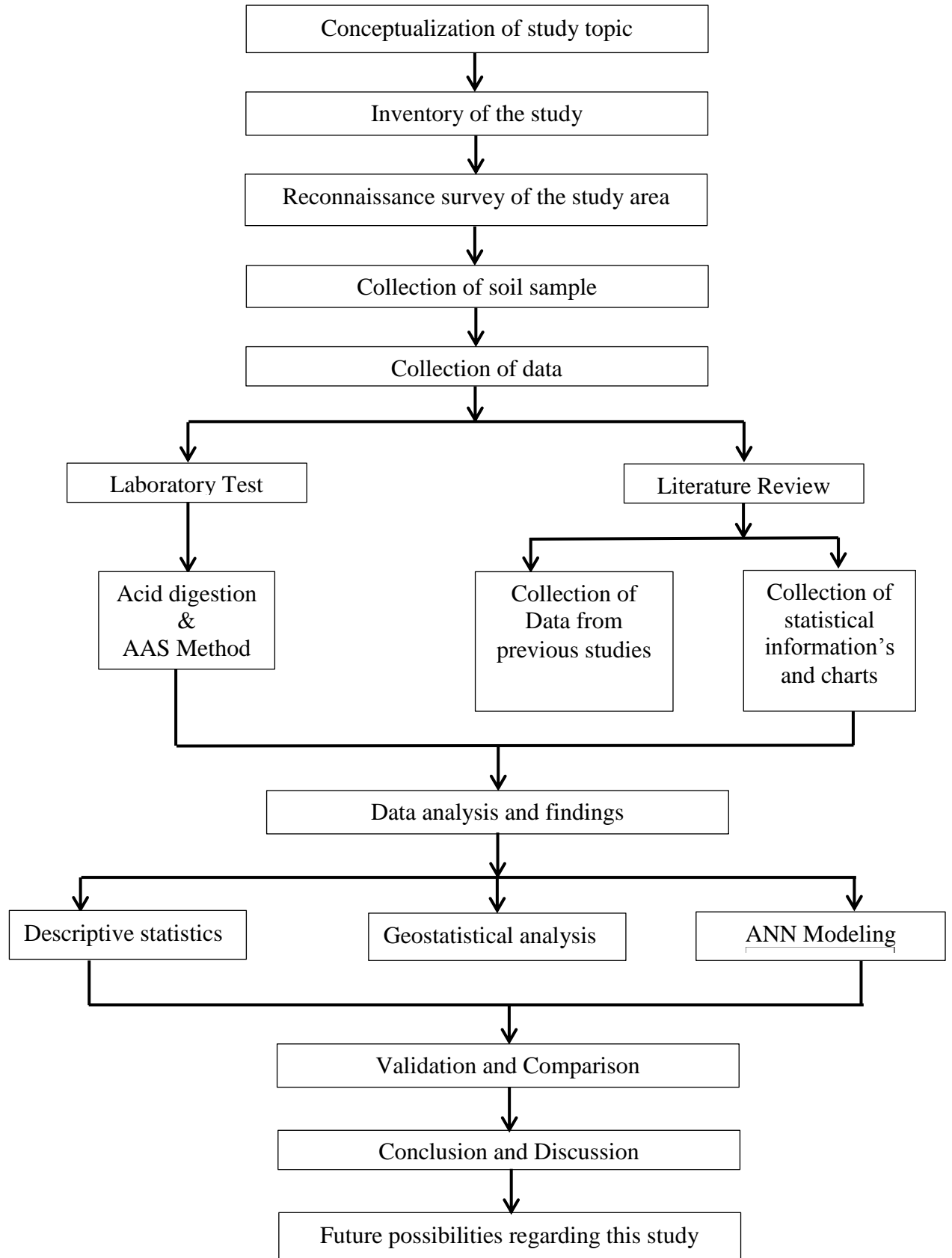


Figure 3.3: Flow diagram of research methodology in this study.

The first sampling point, BH-1 is located at the centre of the waste disposal site. On the other hand, the first borehole of rainy season (BH-41) is about 30 m apart from BH-1 which is the centre of the site and maintains a gradual addition of about 15 m in selecting other following boreholes. Proper care was taken to remove any loose material, debris, coarse aggregates from the bottom of the excavated pit. The soil samples were collected from the bottom of the borehole by excavating the ground manually by using hand shovels. Samples were taken in large polythene bags and eventually transported to the laboratory. Figure 3.2 depicted the soil sampling locations in waste disposal site at Rajbandh, red circles indicated sampling points in dry season and blue triangles indicated sampling points in rainy season. The overall research methodology of this study is illustrated in Figure 3.3.

### **3.5 Laboratory Investigations**

The soil samples were carried in the laboratory to measure the concentration of metal elements of Al, As, Ba, Ca, Fe, Hg, K, Mn, Na, Na, Pb, Sb, Sc, Sr, Ti, V and Zn in the soil samples. Moreover, the values of some metal element concentrations were collected from secondary sources. The procedure of acid digestion and atomic absorption spectrophotometer (AAS) analysis are described in the following articles.

#### **3.5.1 Acid Digestion**

To measure the concentration of metal elements in soil, laboratory work was done following the standard test method. In laboratory investigation, at first 10 g of each soil sample was taken into a 100 mL conical flask. Already, the flask had been washed with deionized water prepared by adding 6 mL HNO<sub>3</sub>/HClO<sub>4</sub> acid in ratio 2:1 and left overnight. Each sample was kept into the temperature of 150°C for about 90 minutes. Later, temperature was raised to 230°C for 30 minutes. Subsequently, HCl solution was added in ratio 1:1 to the digested sample and re-digested again for another 30 minutes. The digested sample was washed into 100 mL volumetric flask and mixture obtained was cooled down to room temperature.

### 3.5.2 Analysis of Metal Elements with AAS

After performing the digestion procedure, metal element concentrations in this digested solution were determined using atomic absorption spectrophotometer (AAS) and the amount of each heavy metal was deduced from the calibration graph. The concentration of the metal elements of Al, Fe, Mn, Cr, Cu, Pb, Zn, Ni, Cd, As, Co, Sb, Sc and Hg in mg/kg were measured in the laboratory.

### 3.6 Descriptive Statistics

The descriptive statistics including normality test such as Shapiro-Wilk test and K-S test was performed. In normality test, the normal quantile-quantile (QQ) plot was also schemed to check the distribution of data points more accurately. Moreover, the conventional statistics was also performed to see the variability of measured concentrations of metal elements and hence discussed in the following articles.

#### 3.6.1 Normality Test

In statistics, normality tests are used to determine if a data set is well-modeled by a normal distribution and to compute how likely it is for a random variable underlying the data set to be normally distributed depending on one's interpretations of probability. In descriptive statistics, one measures a goodness of fit of a normal model to the data (Royston, 1991). In this study, the normality test was analyzed and hence discussed in the following articles.

##### 3.6.1.1 Shapiro-Wilk Test

The Shapiro–Wilk test was published in 1965 by Samuel Sanford Shapiro and Martin Wilk, is a test of normality in frequentist statistics designed to detect all departures from normality. In this study, the deviation from normality,  $W$  was calculated using the following equation considering the dataset of soil samples as  $x_1, x_2, \dots, x_n$ .

$$W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{(\sum_{i=1}^n x_i - \bar{x})^2} \dots \dots \dots (3.1)$$

Where, the  $x_{(i)}$  is the ordered sample value i.e. concentration of metal elements corresponding to boreholes  $x_{(1)}$  is the concentration of first borehole point) and the  $a_{(i)}$  are constants generated from the means, variances and covariances of the order statistics of a sample of size  $n$  from a normal distribution. The test rejects the hypothesis of normality when the significance value ( $p$ ) is less than or equal to 0.05. Failing the normality test allows to state with 95% confidence the data does not fit the normal distribution.

### 3.6.1.2 Kolmogorov–Smirnov test

A nonparametric test of Kolmogorov–Smirnov test (K–S test), named on Andrey Kolmogorov and Nikolai Smirnov is used to compare a sample with a reference probability distribution (one-sample K–S test), or two samples (two-sample K–S test) to equalize of continuous and one-dimensional probability distributions. In this study, the following formula for the computation of the Kolmogorov-Smirnov goodness of fit statistic was used.

$$D = \max\left(F(Y_i) - \frac{i-1}{N}, \frac{i}{N} - F(Y_i)\right) \quad ; (1 \leq i \leq N) \dots\dots\dots(3.2)$$

Where,  $F$  is the theoretical cumulative distribution of the distribution being tested which must be a continuous distribution, and it must be fully specified (i.e., the location, scale, and shape parameters cannot be estimated from the data. In this study, the null distribution of K–S test was calculated under the null hypothesis that the sample was drawn from the reference distribution. Here, it can be noted that If the significance value ( $p$ ) of the K-S test is greater than 0.05, the data is normal. If it is below 0.05, the data significantly deviate from a normal distribution. The K-S test was performed on all metal element considering the significance level as 0.05. The hypothesis of this test are as follows:

$H_0$ : The data are normally distributed

$H_a$ : The data are not normally distributed

So, when testing for normality:

Probability  $> 0.05$ ; the data are normally distributed.

Probability  $< 0.05$ ; the data are not normally distributed.

The steps of normality test performed by SPSS is given in Figure A.1 in Annex-A.

### **3.6.1.3 Normal QQ Plot**

The quantile-quantile (Q-Q) plot is a probability plot, which is a graphical technique to assess if a set of data probably came from some theoretical distribution such as a Normal or exponential. If a set of intervals for the quantiles is chosen, a point (x, y) on the plot corresponds to one of the quantiles of the second distribution (y-coordinate) plotted against the same quantile of the first distribution (x-coordinate) (ESRI,2001). The points plotted in a Q-Q plot are always non-decreasing when viewed from left to right. The steps of construction of normal QQ plot using XLSTAT is given in Figure A.2 provided in Annex-A. Though the Q-Q plot follows the 45° line  $y = x$ , if the two distributions agree after linearly transforming the values in one of the distributions, then the Q-Q plot follows some line, but not necessarily the line  $y = x$ . Q-Q plots are commonly used to compare a data set to a theoretical model (Engineering Statistics Handbook). An assessment of "goodness of fit" that is graphical, rather than reducing to a numerical summary. In this study, the normal QQ Plot was drawn for each metal element in soil to verify either its normal distributed or not.

### **3.6.2 Conventional Statistics**

The researchers of Heckman (1976) gave an idea about the desire to measure the variability, as well as the location. In a research of Yasrebi et al. (2009) were also computed the conventional statistical parameters to describe the distribution of metal element concentration. In this study, through SPSS, the conventional statistical parameters in terms of mean, maximum, minimum, median, SD, CV, skewness and kurtosis for two seasons (i.e. dry and rainy) was analyzed to show the seasonal variation of metal elements in soil. Many other researchers also concluded. The steps of computing conventional statistics are illustrated in Figure A.3 in the Annex-A.

### **3.7 Multivariate Statistics**

In this research work, the multivariate statistical analysis of Pearson correlation, principal component analysis (PCA) and agglomerative hierarchical clustering (AHC) was performed and hence discussed in the following articles.

### 3.7.1 Pearson Correlation

The bivariate Pearson correlation produces a sample correlation coefficient,  $r$ , which measures the strength and direction of linear relationships between pairs of continuous variables. In this study, the value of correlation coefficient,  $r$ , was computed using the following equation considering one dataset  $\{x_1, \dots, x_n\}$  containing  $n$  values and another dataset  $\{y_1, \dots, y_n\}$  containing  $n$  values.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \dots \dots \dots (3.3)$$

Correlation varies in the range  $[-1, 1]$ . The sign of the correlation coefficient indicated the direction of the relationships, while the magnitude of the correlation indicates the strength of the relationship (Rogers and Nicewander, 1988). A perfectly negative linear relationship indicated by the value of  $-1$ ,  $0$  values indicated no relationship and  $1$  indicated perfectly positive linear relationship. In this study, Pearson correlation coefficient was performed for both the dry and rainy seasons. The null hypothesis ( $H_0$ ) and alternative hypothesis ( $H_1$ ) of the significance test for correlation was expressed depending on a two-tailed test. Two-tailed significance test are as follows.

$H_0: r = 0$  (“the correlation coefficient is 0, there is no association”)

$H_1: r > 0$  (“the correlation coefficient is not 0, a nonzero correlation could exist”)

As all the correlations,  $r > 0$ , a nonzero correlation existed between the metal elements in soil of waste disposal site. The steps of computing Pearson Correlation coefficient,  $r$  is illustrated in Figure A.4 in the Annex-A.

### 3.7.2 Principal Component Analysis

The principal component analysis (PCA) is probably the most popular multivariate technique that is used to analyze a dataset of inter-correlated quantitative dependent variables. Fundamentals of PCA involves the transformation of a set of multivariate data containing analytical constituents (variables) into a new orthogonal set by allocating total variance to uncorrelated variables (principal components – PCs) using the correlation

matrix, whereby the individual variable represent the linear combination of the initial data variables. The PCs are in decreasing order based on factor loading, having the PCs with the largest variance occupying the first PC (PC1) and this follows successively to the PC with the smallest variance (PCn).

A large amount of quantitative analysis relies on Principal Component Analysis (PCA). This is usually referred to in tandem with eigenvalues, eigenvectors and lots of numbers. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible. The eigenvalue for a given factor measures the variance in all the variables which is accounted for by that factor. A factor's eigenvalue may be computed as the sum of its squared factor loadings for all the variables to measure the amount of variation in the total sample accounted for by each factor. The factor loadings are the correlation coefficients between the variables and factors .To get the percent of variance in all the variables accounted for by each factor, add the sum of the squared factor loadings for that factor. This is the same as dividing the factor's eigenvalue by the number of variables. In this study, the following equation was used through XLSTAT to compute the variance.

$$\left. \begin{aligned} PC1 &= a_1x_1 + a_2x_2 + \dots + a_nx_n \\ PCn &= \sum_{j=1}^n a_{1j}x_j \end{aligned} \right\} \dots\dots\dots (3.4)$$

where;

$a_{1j}$  = eigenvectors obtained from the correlation matrix

$x_j$  = input variables

In this study, the PCA method was performed sequentially, first by information extraction in the input space (with n-dimensions) to determine the directions of which the input variables display the most substantial variability. The PC coefficients and the eigenvalues ( $\lambda_i > 0, i = 1,2, \dots, n$ ) 0 for the correlation matrix ( $C = E\{xx^T\}$ ) with respect to their eigenvectors ( $e_i > 0, i = 1,2, \dots, n$ ) 0, is called the loadings were then calculated which gives a new set of variables that explains the variability in the original dataset; the first PCs retained a greater proportion of the total variance, consequently leading to effective and practical dimensionality reduction exercise. The steps of PCA are illustrated in Figure A.5 in the Annex-A.

### 3.7.3 Agglomerative Hierarchical Analysis

AHC is a statistical modeling of input data used for multivariate analysis. The main goal of the AHC analysis is to spontaneously classify the data into groups of similarity (clusters) searching objects in the n-dimensional space located in closest neighborhood and to separate a stable cluster from other clusters. In case of metal element concentrations, the sampling sites were considered as objects for classification, each one determined by a set of variables. In order to achieve the goal this series of procedures was maintained in the present study:

1. Normalization of the raw input data to dimensionless units in order to avoid the influence of the different range of chemical dimensions (concentration);
2. Determination of the distance between the objects of classification by application of some similarity measure, e.g., Euclidean distance or correlation coefficient;
3. Performing appropriate linkage between the objects by some of the cluster algorithms like single, average or centroid linkage;
4. Plotting the results as dendrogram;
5. Determination of the clustering pattern;
6. Interpretation of the clusters both for objects and variables.

Figure A.6 in the Annex-A showed the steps of AHC performed by XLSTAT software. Dendrograms were the output of AHC which display the cluster hierarchy and the distances at which the clusters were joined which is helpful to select an appropriate number of clusters for the dataset. Cluster was selected by cutting the dendrogram where there is a significant jump in the distance of the cluster joins which is equivalent to selecting the knee point in a k-Means curve as established in the research of Anderberg (1973). Terzano et al. (2007) made an assessment on the origin and fate of Cr, Ni, Cu, Zn, Pb, and V in industrial polluted soil by combined microspectroscopic techniques and bulk extraction methods.

### 3.8 Geostatistical Analysis

Spatial interpolation is widely used when data are collected at distinct locations (e.g. soil profiles) for producing continuous information. According to Burgess and Webster (1980),



geostatistics has been applied in case of spatial interpolation for more than 20 years. A research conducted by Zhang et al. (2011) also verified this statement. In the present investigation of metal elements in waste disposal site, deterministic (i.e., create surfaces from measured points) and geostatistical (i.e., utilize the statistical properties of the measured points) interpolation techniques were used. A selection of deterministic interpolation techniques was made based on the extent of similarity such as IDW and LPI. In addition, an Interpolation technique based on degree of smoothing such as RBF was also selected. Geostatistical interpolation, namely ordinary kriging (OK) was used to generate the spatial distribution of metal elements in soil of the study area. The common steps for all type of interpolation techniques of geostatistical analysis using ArcGIS were illustrated in Figure A.7 in the Annex-A.

### ***3.8.1 Inverse Distance Weighting***

The inverse distance weighting (IDW) is one of the commonly applied deterministic interpolation technique which explicitly implements the assumption that things that are close to one another are more alike than those that are farther apart. To predict a value for any unmeasured location, IDW will use the measured values surrounding the prediction location (ESRI,2001). In IDW method, it is expected substantially that the rate of correlations and similarities between neighbors is proportional to the distance between them which is defined as a distance reverse function of every point from neighboring points. Isaake and Srivastava, 1989 recommended that the value of the power parameter is one of the important factor of IDW interpolation method . It is established that the default value of power,  $p=2$ , is called the inverse distance squared weighted interpolation. There is no thoretical justification to prefer this value of others, and the effect of changing  $p$  should be investigated by previewing the output and examining the cross validation result statistics (ESRI,2001). It was proposed that the weights assigned to the interpolating points are the inverse of its distance from the interpolation point. Consequently, the close points are made-up to have more weights (so, more impact) than distant points and vice versa. As an estimation with closest sampled points was given more importance, it is just considered integer values of parameter, because the values lower than one are closest to a simple average estimation (Isaake and Srivastava, 1989).

$$Z_0 = \frac{\sum_{i=1}^N z_i d_i^{-n}}{\sum_{i=1}^N d_i^{-n}} \dots\dots\dots(3.5)$$

Where,  $Z_0$  is the estimation value of variable  $z$  in point  $i$ ,  $Z_i$  is the sample value in point  $i$ ,  $d_i$  is the distance of sample point to estimated point,  $N$  is the coefficient that determines weigh based on a distance.,  $n$  is the total number of predictions for each validation case. In this study, estimations were made using different integer powers of 1 to 5 as mentioned in the research of Yasrebi et al., 2009 proposed the smaller the root mean square prediction error (RMSPE), the smaller will be the error of the prediction surface. Figure A.8 in the Annex-A indicated the steps of IDW interpolation techniques.

### 3.8.2 Local Polynomial Interpolation

LPI is a moderately quick deterministic interpolator that provides prediction, prediction standard error and condition number surfaces that are comparable to ordinary kriging with measurement errors. According to ESRI, 2001 it was stated that it is more flexible than the global polynomial method, but there are more parameter decisions. Local polynomial methods do not allow you to investigate the autocorrelation of the data, making it less flexible and more automatic than kriging (Xie et al., 2011). There are no assumptions required of the data. According to Hani and Abari, 2011, LPI fits the local polynomial using points only within the specified neighborhood instead of all the data. According to ESRI (2001) a first-order global polynomial fits a single plane through the data.; a second-order global polynomial fits a surface with a bend in it, allowing surfaces representing valleys; a three-order global polynomial allows for two bends. Order of polynomial greater than 3 is not generally recommended for most situations due to its high critical spatial condition number threshold value which indicate less reliability. Then the neighborhoods can overlap, and the surface value at the center of the neighborhood is estimated as the predicted value. LPI is capable of producing surfaces that capture the short range variation. Interpolation quality was judged by examining the RMSPE of the crossvalidation and the validation: the smaller the RMSPE's are, the better the interpolation is (Brovelli et al., 2011). Figure A.9 showed the steps of LP interpolation techniques in the Annex-A.

### 3.8.3 Radial Basis Function

RBF are moderately quick deterministic interpolator . This method provides prediction surfaces that predicts values identical with those measured at the same point and the generated surface requires passing through each measured point. Xie et al., (2011) suggested that the predicted values can vary above the maximum or below the minimum of the measured values .A study of Yao et al.,(2013) radial basis function (RBF) is a real-valued function whose value depends only on the distance from the origin, so that  $\phi(x) = \phi(\|x\|)$ ; or alternatively on the distance from some other point  $c$ , called a center, so that  $\phi(x, c) = \phi(\|x - c\|)$ ; Any function  $\phi$  that satisfies the property  $\phi(x) = \phi(\|x\|)$  is a radial function. The norm is usually Euclidean distance, although other distance functions are also possible. ESRI 2001 recommended RBF method is a family of five deterministic exact interpolation techniques: thin-plate spline (TPS), spline with tension (SPT), completely regularized spline (CRS), multi-quadratic function (MQ) and inverse multi-quadratic function (IMQ).

Commonly used types of radial basis functions include (writing  $r = \|x - x_i\|$ ).

$$\text{Multiquadric: } \phi(r) = \sqrt{1 + (\epsilon r)^2} \dots \dots \dots (3.6)$$

$$\text{Inverse Multiquadric: } \phi(r) = \frac{1}{1 + (\epsilon r)^2} \dots \dots \dots (3.7)$$

$$\text{Thin Plate Spline: } \phi(r) = r^2 \ln(r) \dots \dots \dots (3.8)$$

Johnston et al. (2001) identified RBF fits a surface through the measured sample values while minimizing the total curvature of the surface. RBF is ineffective when there is a dramatic change in the surface values within short distances (ESRI, 2001). The most widely used RBF that is IMQ was selected in this study. Brovelli (2011) stated that once a specific kernel function was selected, by modifying the parameters in the control panel and the power of the interpolation, it is possible to alter the neighborhood search strategy. Interpolation quality was judged by examining the RMSPE of the cross validation and the validation: the smaller they are, the better the interpolation is (Brovelli, 2011). Figure A.10 in the Annex-A showed the steps of RBF interpolation using ArcGIS software.

### 3.8.4 Ordinary Kriging

The kriging approach is used to express the spatial continuity (autocorrelation). Ordinary kriging method is the most general and widely used of the kriging methods, incorporates statistical properties of the measured data (spatial autocorrelation). It delivers estimation at an unobserved location of variable  $z$ , based on the weighted average of adjacent observed sites within a given area. An assessment by the ordinary kriging analyst at an unsampled site  $z(s_0)$  is defined by (Yasrebi et al., 2009):

$$Z(s_0) = \sum_{i=1}^n \lambda_i Z(s_i) \dots \dots \dots (3.9)$$

where,  $\lambda$  are the weights assigned to each of the observed samples. Summation of these weights equal to unit value so that the analyst offers an unbiased estimation (Yasrebi et al., 2009):

$$\sum_{i=1}^n \lambda_i = 1 \dots \dots \dots (3.10)$$

ESRI 2001 provides with ordinary kriging which have functions such as Circular, Spherical, Tetraspherical, Pentaspherical, Exponential, Gaussian, Rational quadratic, Hole effect, K-Bessel, J-Bessel, Stable from which to choose for modeling the empirical semivariogram. Geostatistical analyses were initiated with semivariogram to evaluate for each soil property as follow (Isaake and Srivastava, 1989):

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i + h) - z(x_i)]^2 \dots \dots \dots (3.11)$$

Where,  $\gamma(h)$  = The experimental semivariogram value at distance interval  $h$ .

$N(h)$  = Number of sample pairs within the distance interval  $h$ .

$z(x_i), z(x_i + h)$  = Sample values at two points separated by distance interval  $h$ .

The semivariogram measures the strength of the statistical correlation as a function of lag distance. Nugget variance ( $C_0$ ) as the variance at nil distance, range ( $A$ ) as the distance at which the spatial correlation vanishes and the sill ( $C+C_0$ ) corresponds to the maximum variability in the absence of spatial dependence. The coefficient of determination ( $R^2$ ) was employed to determine goodness of fit. Spatial dependence for

the soil variables were assessed by the ratio between the nugget semivariance and the total semivariance (Cambardella et al., 1994). The objective of cross-validation is to make an informed decision about which model provides the most accurate predictions. The predictions are unbiased, by a mean prediction error close to 0 and standard errors are accurate, indicate by a root-mean-square-standardized prediction error close to 1 (Brovelli et al., 2011). A study of Xie et al. (2011) found that the prediction do not deviate much from the measured values, indicate by root-mean-square error and average standard error that are as small as possible. Figure A.11 in the Annex-A showed the steps of RBF interpolation using ArcGIS software.

### 3.9 Assessment of Method Performance

The precision of interpolation techniques was measured based on several literature indices. MAPE is defined as the percentage errors between predicted and observed value. Yao et al., (2013) found that small values of MAPE represents a model with less errors and more accurate predictions. The MAPE is computed by the equation given below (Yao et al., 2013; Yasrebi et al., 2009):

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{(p_i - o_i)}{o_i} \dots \dots \dots (3.12)$$

The effectiveness of the models was evaluated using a goodness of prediction statistic (G), also known as the coefficient of determination, is computed by (Yao et al., 2013; Yasrebi et al., 2009):

$$G = 1 - \left[ \frac{\sum_{i=1}^n (p_i - o_i)^2}{\sum_{i=1}^n (o_i - \bar{o})^2} \right] \dots \dots \dots (3.13)$$

Where n is the number of validation points,  $p_i$  is the predicted value at point i,  $o_i$  is the observed value at point i, and  $\bar{o}$  is the sample arithmetic mean. A study of Yasrebi et al., (2009) showed G-value equal to 1 indicates perfect prediction, a more reliable model than if the sample mean had been used, a negative value specifies a less reliable model than if the sample mean had been used, and a value of zero indicates that the sample mean should be used (Yao et al., 2013).

Finally, the relative improvement (RI) of the best method compared with the others is calculated with equation (Yasrebi et al., 2009):

$$RI = \frac{100 |RMSE_{Best} - RMSE_{Current}|}{RMSE_{Best}} \dots\dots\dots(3.14)$$

In order to analyze the effect of model parameters on pollution assessment, ordinary kriging, IDW with power 1 to 5, LP with order 1 to 3 and RBF five kernel functions of CRS, IMQ, MQ, ST and TPS were selected.

### 3.10 Artificial Neural Networks

Silipo et al., (1999) stated Artificial Neural Networks (ANNs) are intelligent methods in advanced computing that quantitatively analyze information by learning, the same way as human intelligent systems. The composition of the ANN includes an array of several units (neurons) connected together that function as a processing system which can be used for specialized problem solving by classifying data and also for recognizing patterns in a data set (Neuner, 2012). The ANN is characterized by a group of models that can carry out dimensionality reduction tasks by transforming non-linear input data into neurons with a lower dimensionality structure (Figure. 3.16 and Figure 3.17). The fundamental application of the self-organizing map (SOM) consists of cooperation, competition and updating mechanisms in accomplishing unsupervised learning, with the goal of reducing the dimensions found in any data set. These are presented by mapping the high dimensions in the data to a 2- dimensional lattice feature map which also retains the topology of the data, presented as a discrete map.

According to Kohonen, 2001, the input space is the environment where the original data are prior to the application of the SOM technique, while after the data have been treated, it forms a grid of neurons that represents the output space, which is often in 2 dimensions and with a rectangular or hexagonal grid of neurons. The SOM output space can also be in either 1-dimensional or even 3-dimensional (Kohonen, 2014).

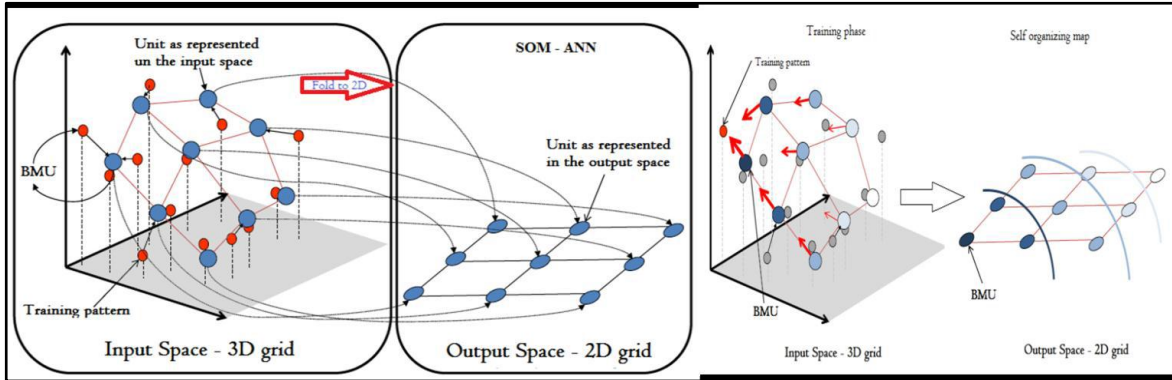


Figure 3.16: SOM input space and output space. Red dots signify input patterns while blue dots show the connected SOM neurons (Source: Olawoyin et al., 2012).

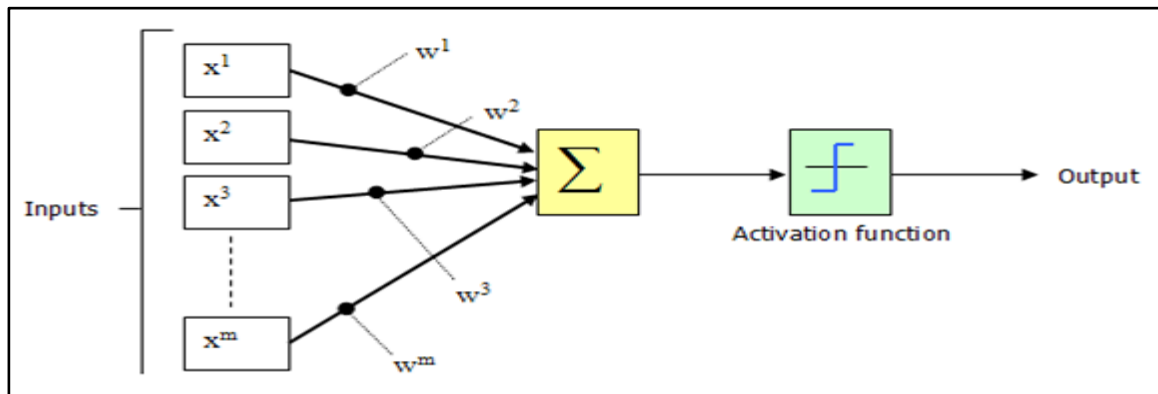


Figure 3.17: Artificial neural network, identical to the human neural network (Source: Olawoyin et al., 2012).

In this study, the connection patterns of metal elements were presented in soil of waste disposal site. The weight planes for each element of the input vector that connect each input to each of the neurons were visualized. According to Kohonen (2014), darker colors represented larger weights and correlation between the metal elements. The early-stopping-approach was used to generalize the network. The available learning data in the network is divided into three sets: training data, validation data and test data. The training data set was used to train the neural network through weight adaption. The errors of the output of the validation data set in relation to the training data set were recorded during the training phase and diminish at the beginning of the training. When the neural network tends to overfit, this error increases and the weights are determined on the basis of the minimum error (Rooki et al., 2011). The data was divided like this: 70% training data, 20% validation data and 10% test data. According to Shahin et al., (2004) statement, there

were no hard rules applied for data division. It was collected depending on the complexity of the problem and on the amount and nature of the learning data. Alkaiem et al. (2016) recommended the most important indicator for the quality of the network generalization is its mean square error (MSE) in MATLAB. MSE measured the average of the squares of the errors or deviations. It is always non-negative, and values closer to zero are better. The error histogram plotted for training data were to provide additional verification of network performance which shows the distribution of the residuals between targets and network output (Shahin et al., 2004). On the basis of research of Neuner (2012) stated smaller error indicates perfect prediction. The R-value measures the correlation between observed and predicted values. R value of 1 and 0 means a close and random relationship, respectively (Alkaiem et al., 2016). The R-value between the predicted and the actual values of metal elements present in contaminated soil for training, validation, testing and the whole datasets and errors for training cities was determined to find the accuracy of nntool of ANN. Figure A.12 showed the steps of nntool modeling of ANN provided in the Annex-A.



## CHAPTER IV

### RESULTS AND DISCUSSION

#### 4.1 General

This chapter deals with the spatial distribution and temporal variation of the concentration of metal elements such as aluminium (Al), arsenic (As), barium (Ba), calcium (Ca), iron (Fe), mercury (Hg), potassium (K), manganese (Mn), sodium (Na), nickel (Ni), lead (Pb), antimony (Sb), scandium (Sc), strontium (Sr), titanium (Ti), vanadium (V) and zinc (Zn) in soil of waste disposal site. The statistical analysis of the concentration of metal elements was described in terms of mean, maximum, minimum, median, standard deviation (SD), co-efficient of variation (CV), skewness and kurtosis using Statistical Package for the Social Sciences (SPSS). The Pearson's correlation of metal elements was also performed using SPSS to examine the accumulation of the concentration of metal elements irrespective to their sources. In this chapter, Principal Component Analysis (PCA) and Agglomerative Hierarchical Clustering (AHC) were performed using XLSTAT to get the information about the spatial distribution of metal elements and to identify the possible sources of contamination of soil. In this chapter, Geostatistical analysis such as ordinary kriging (OK), inverse distance weighting (IDW), local polynomial interpolation (LP) and radial basis function (RBF) were performed using ArcGIS to find out the best fitted model and to distribute the metal elements spatially in soil. Semivariogram parameters obtained from ordinary kriging (OK) exhibited spatial dependence of metal elements in soil and also discussed in this chapter. The performance of various interpolation techniques were assessed by ascertaining the error in the predictions based on Mean Absolute Percentage Error (MAPE), Goodness of prediction (G-value) and Relative Improvement (RI). Accuracy of interpolation was also assessed by determining and plotting local CV with mean relative error (MRE). In this chapter, the validity and accuracy of the metal elements concentration obtained from laboratory were checked in terms of Mean Standard Error (MSE) and Regression factor (R-value) based on Artificial Neural Network (ANN) using MATLAB.

## 4.2 Descriptive Statistics

Descriptive statistics analysis was performed to examine the quantitative distribution of metal elements presence in soil of waste disposal site in a manageable form. Firstly, the normality of metal elements were performed using the Kolmogorov- Smirnov (K-S) test, Shapiro-Wilk (S-W) test and Normal QQ Plot. Statistical characteristics of metal elements were analyzed using SPSS (Reza et al., 2010). In this study, the normal QQ plot for each metal element was done by XLSTAT and hence discussed in the following articles.

### 4.2.1 Normality Test

Table 4.1: Normality test of metal elements in soil of waste disposal site

Metals	Dry season						Rainy season					
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk			Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.	Statistic	df	Sig.	Statistic	df	Sig.
Al	0.11	40	0.200*	0.954	40	0.104	0.157	20	0.200*	0.934	20	0.18
As	0.201	40	0	0.891	40	0.001	0.354	20	0	0.711	20	0
Ba	0.17	40	0.005	0.898	40	0.002	0.122	20	0.200*	0.952	20	0.403
Ca	0.123	40	0.132	0.928	40	0.013	0.154	20	0.200*	0.951	20	0.38
Cd	0.095	40	0.200*	0.972	40	0.421	0.206	20	0.026	0.884	20	0.021
Co	0.151	40	0.022	0.942	40	0.042	0.16	20	0.194	0.896	20	0.035
Cr	0.173	40	0.004	0.922	40	0.009	0.285	20	0	0.791	20	0.001
Cu	0.245	40	0	0.751	40	0	0.138	20	0.200*	0.92	20	0.101
Fe	0.128	40	0.097	0.949	40	0.068	0.208	20	0.023	0.865	20	0.01
Hg	0.148	40	0.028	0.904	40	0.002	0.239	20	0.004	0.734	20	0
K	0.12	40	0.148	0.933	40	0.02	0.208	20	0.023	0.853	20	0.006
Mn	0.179	40	0.002	0.816	40	0	0.179	20	0.091	0.881	20	0.018
Na	0.202	40	0	0.85	40	0	0.179	20	0.093	0.916	20	0.084
Ni	0.111	40	0.200*	0.94	40	0.035	0.118	20	0.200*	0.944	20	0.285
Pb	0.159	40	0.013	0.842	40	0	0.19	20	0.058	0.942	20	0.263
Sb	0.084	40	0.200*	0.95	40	0.076	0.094	20	0.200*	0.952	20	0.398
Sc	0.115	40	0.199	0.958	40	0.143	0.101	20	0.200*	0.974	20	0.828
Sr	0.122	40	0.133	0.94	40	0.035	0.107	20	0.200*	0.957	20	0.477
Ti	0.095	40	0.200*	0.94	40	0.035	0.091	20	0.200*	0.964	20	0.628
V	0.149	40	0.026	0.949	40	0.071	0.097	20	0.200*	0.971	20	0.77
Zn	0.182	40	0.002	0.921	40	0.008	0.153	20	0.200*	0.952	20	0.404

<sup>a</sup>Lilliefors Significance Correction;

\*This is a lower bound of the true significance.

From K-S test, for an alpha level of 0.05, a set of data with a significance value ( $p$ ) of less than 0.05 would reject the null hypothesis that the datasets were form of normally distributed. The output obtained from both the nonparametric normality test of K-S and S-W for all the studied metal elements in soil for both the dry and rainy seasons is provided in Table 4.1. Based on Table 4.1, the highest  $p$  value for Al, Ni, Sb and Ti (0.200) and lowest one is for As (0.000) in soil was observed for both the dry and rainy seasons in case of K-S test. In addition, the metal elements of Al, Ca, Cd, Fe, K, Ni, Pb, Sb, Sc, Sr and Ti in soil for dry season were normally distributed which indicated that the null hypothesis was accepted at the significance value ( $p$ ) greater than 0.05.

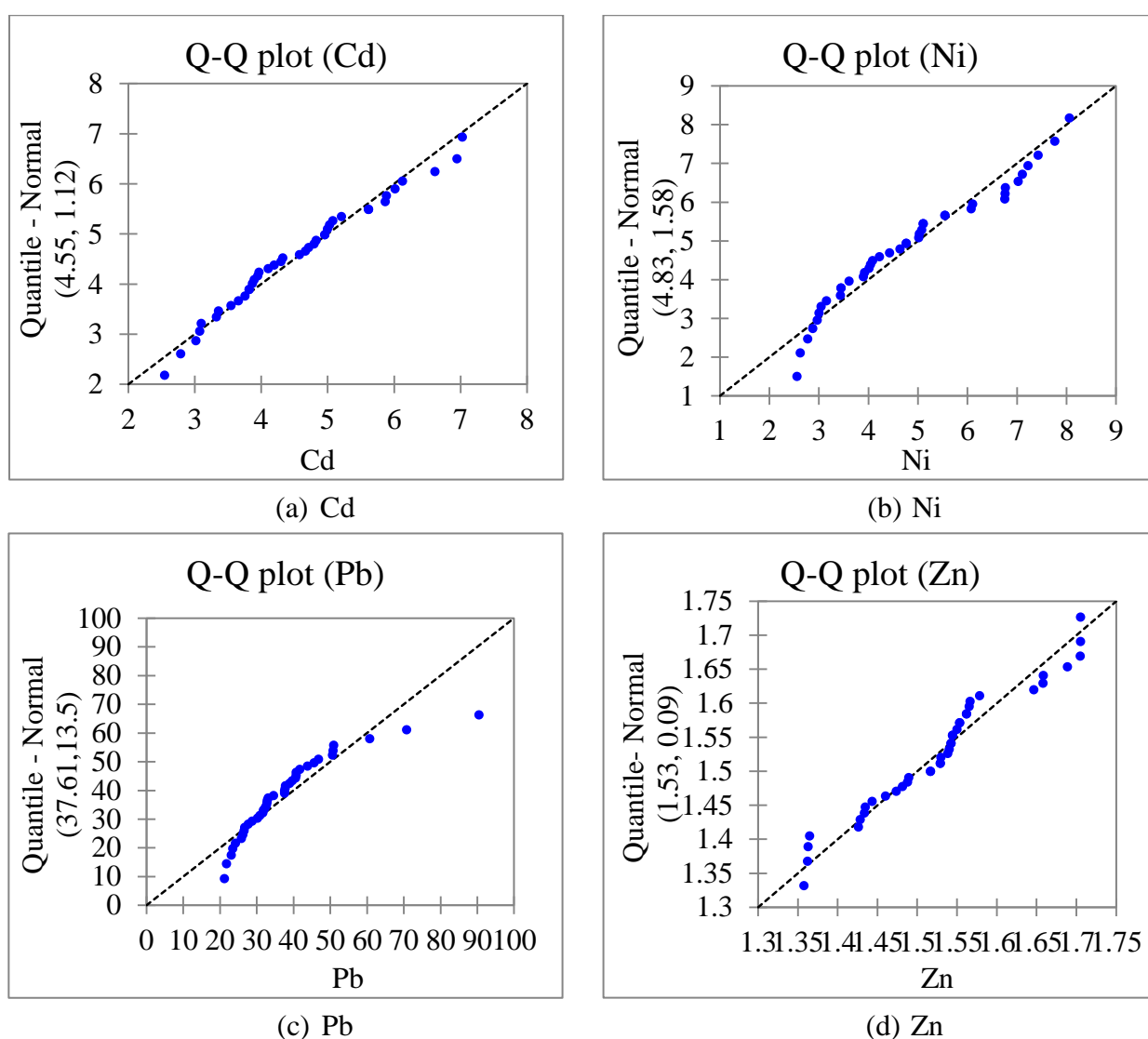


Figure 4.1: Normal QQ plots for dry season (a) Cd; (b) Ni; (c) Pb and (d) Zn.

Results also showed that the other metal elements of As, Ba, Co, Cr, Cu, Hg, Mn, Na, V and Zn in soil were not normally distributed and log transformation was applied to distribute of these metal elements normally. In rainy season, the metal elements of Al, Ba, Ca, Co, Cu, Mn, Na, Ni, Pb, Sb, Sc, Sr, Ti, V and Zn in soil were normally distributed as the significance  $p$ -value was greater than 0.05. However, in this rainy season, some metal elements were not satisfied the significance level, then the normally distribution was performed by log transformation system. Moreover, in S-W test, the metal elements of As, Ba, Ca, Co, Cr, Cu, Hg, K, Mn, Na, Ni, Sr, Ti and Zn in soil were normally distributed in dry season. In rainy season, all the metal elements were normally distributed except As, Cd, Co, Cr, Fe, Hg and Mn as the significance  $p$ -value was less than 0.05.

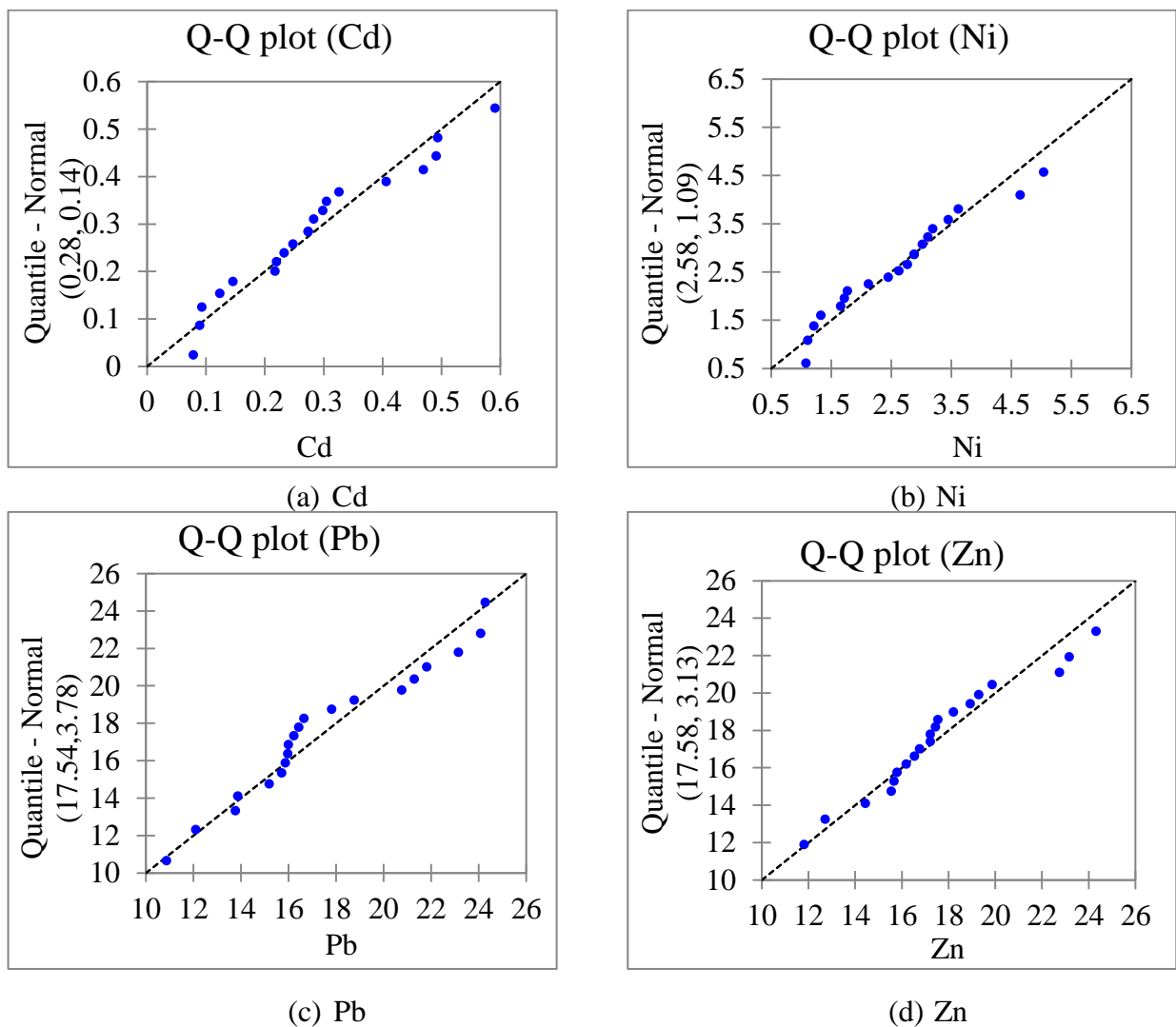


Figure 4.2: Normal QQ Plots in the rainy season (a) Cd; (b) Ni; (c) Pb and (d) Zn.

As normality tests were too sensitive to sample size, it would be better to avoid using them unless there were a large number of inspections to do and automate the process. Graphical methods are a better alternative to evaluate normality; in particular QQ plots (Royston,1991). The normal QQ plot for the metal elements of Cd, Ni, Pb and Zn in soil for both the dry and rainy season is depicted in Figure 4.1 and Figure 4.2, respectively. In this analysis, the cumulative distribution values obtained from cumulative distribution plot and cumulative distribution values obtained from standard normal distribution plot were considered as *ordinate* and *abscissa*, respectively for plotting of normal QQ.

Based on Figure 4.1 and Figure 4.2, it can be interpreted that the metal elements of Al, Ba and Ca in soil were normally distributed because all points fall very close to “45<sup>0</sup> reference line”. Moreover, it was observed that the data points for metal element of As for rainy season, were deviated from “45<sup>0</sup> reference line” referred to the data points were not normally distributed. Moreover, based on results obtained from K-S test, S-W test and normal QQ Plot, it can be concluded almost all the metal elements in soil for both the dry and rainy seasons were distributed normally except As. Thus, log transformation was applied to As for normal distribution. In contrast, the normal QQ plot for other metal elements of Cd, Co,Cr, Cu, Fe, Hg, K, Mn, Na, Ni, Pb, Sb, Sc, Sr, Ti, V and Zn, is provided in Figure B.1 to Table B.17 and B. 18 to Table B.34 in Annex-B for both the dry and rainy season, respectively.

#### **4.2.2 Conventional Statistics**

In this study, the conventional statistics include mean, maximum, minimum, median, SD, CV, skewness and kurtosis was determined using SPSS of metal elements in soil for both the dry and rainy seasons. This analysis was performed to check the variability of metal elements due to anthropogenic activities as well as from natural parent materials and hence discussed in the following articles.

##### **4.2.2.1 Dry Season**

The descriptive statistics of metal elements in soil for dry season of waste disposal site is provided in Table 4.2. Table 4.2 reveals that for all studied metal elements, the total concentration showed a great degree of variability; indicated by the large CV from 22.11%

of Zn to 59.41% for Cu for dry season. The contours rising of CV values reflected the non-homogeneous distribution of concentrations of anthropogenically emitted metal elements (Li et al., 2012). Result also showed that the greatest and the smallest SD were observed in case of Ti (406.26) and Cd (1.14), respectively.

Table 4.2: Descriptive statistics of metal elements in soil for dry season (n=40)

Metal	Min	Max	Median	Mean	CV (%)	SD	Skewness	Kurtosis
Al	158.35	874.78	458.46	490.25	40.31	197.61	0.303	-0.727
As	1.55	8.77	3.42	4.15	48.79	2.03	0.733	-0.709
Ba	37.88	121.90	60.61	65.25	37.36	24.38	0.667	-0.709
Ca	100.20	318.00	173.19	183.80	33.55	61.67	0.577	-0.793
Cd	2.55	7.03	4.46	4.55	24.99	1.14	0.387	-0.530
Co	3.40	12.02	6.67	7.00	34.50	2.42	0.391	-1.015
Cr	4.19	9.82	5.83	6.03	23.63	1.43	0.841	-0.023
Cu	2.92	16.45	4.82	6.20	59.41	3.68	1.541	1.110
Fe	733.19	1987.7	1386.5	1363.94	25.67	350.15	-0.081	-1.199
Hg	1.98	9.20	4.01	4.63	44.63	2.07	0.797	-0.460
K	104.88	460.33	316.37	292.00	35.98	105.06	-0.416	-0.937
Mn	10.82	30.76	14.16	15.69	29.82	4.68	1.654	2.382
Na	19.50	100.54	77.89	68.30	42.78	29.22	-0.511	-1.326
Ni	2.56	8.06	4.71	4.83	33.02	1.60	0.409	-0.984
Pb	21.29	90.55	33.94	37.61	36.34	13.67	1.840	4.393
Sb	2.92	12.55	5.76	6.07	36.04	2.19	0.787	0.496
Sc	7.30	20.41	11.67	12.16	26.89	3.27	0.451	-0.603
Sr	15.44	54.12	26.62	27.53	30.33	8.35	0.869	0.825
Ti	643.33	1937.3	1223.8	1221.2	33.27	406.26	0.160	-1.198
V	20.83	83.35	40.83	43.44	33.35	14.49	0.736	0.264
Zn	22.79	50.76	34.64	34.57	22.11	7.65	0.612	-0.116

In case of dry season, the skewness of metal elements of Al, Cd, Co, K, Na, Ni, Sc and Ti was found between -0.5 to 0.5, indicated the data were approximately symmetrical. The metal elements of As, Ba, Ca, Cr, Hg, Sb, Sr, V and Zn indicated the data points were moderately skewed as the skewness value varied from -0.1 to -0.5 and 0.5 to 0.1. Moreover, metal elements of Cu, Fe, Mn and Pb were highly skewed exhibited skewness value of <-1 and > 1. In addition, the metal elements of Al, As, Ba, Ca, Cd, Co, Cr, Fe, Hg, K, Na, Ni, Sc, Ti and Zn exhibited platykurtic distribution (Kurtosis<0), whereas the

metal elements of Cu, Mn, Pb, Sb, Sr and V exhibited leptokurtic distribution (Kurtosis >0). These results provide a quite reliable analysis.

#### 4.2.2.2 Rainy Season

The descriptive statistics of metal elements in soil for rainy season of waste disposal site is provided in Table 4.3. For all metal elements, the total concentrations showed a great degree of variability as CV varied from 18.25% of Zn to 77.25% of Mn reflecting contamination by anthropogenic activities.

Table 4.3: Descriptive statistics of metal elements in soil for rainy season (n=20)

Metal	Min	Max	Median	Mean	CV (%)	SD	Skewness	Kurtosis
Al	119.22	412.55	224.02	239.76	38.27	91.77	0.402	-1.118
As	0.87	3.68	1.11	1.56	56.70	0.88	1.335	0.245
Ba	18.20	74.00	39.92	39.58	39.77	15.74	0.304	-0.682
Ca	44.20	153.23	105.32	100.16	32.34	32.39	-0.169	-1.194
Cd	1.20	3.90	1.88	2.03	36.26	0.74	1.029	0.284
Co	1.98	9.02	4.95	5.11	49.65	2.54	0.197	-1.483
Cr	0.77	5.55	1.54	1.94	68.01	1.32	1.511	1.302
Cu	0.73	6.33	2.43	2.55	50.37	1.29	1.120	1.845
Fe	160.71	683.92	287.66	365.79	48.95	179.06	0.423	-1.407
Hg	0.72	5.21	1.12	1.72	75.25	1.29	1.655	1.552
K	55.04	262.08	110.52	126.86	52.53	66.64	0.917	-0.473
Mn	1.02	13.43	3.88	5.41	77.25	4.18	0.545	-1.135
Na	8.23	55.60	21.68	25.74	56.01	14.42	0.688	-0.528
Ni	1.08	5.04	2.70	2.58	43.09	1.11	0.494	-0.327
Pb	10.88	24.29	16.34	17.54	22.12	3.88	0.316	-0.838
Sb	0.98	6.12	3.35	3.41	48.60	1.66	0.127	-1.124
Sc	3.02	12.12	7.87	7.73	34.03	2.63	-0.132	-0.928
Sr	8.88	30.44	16.20	16.92	33.96	5.75	0.668	-0.137
Ti	243.88	1445.56	696.69	707.89	42.08	297.85	0.627	0.295
V	6.92	40.88	20.15	21.43	44.48	9.53	0.345	-0.790
Zn	11.82	24.33	17.23	17.58	18.25	3.21	0.458	-0.064

Moreover, in this study, large SD were found in all the studied metal elements except Cd and As. The greatest and the smallest SD were detected of 297.8519 and 0.736449 for the metal elements of Ti and Cd, respectively (Table 4.3). Moreover, the skewness of metal

elements of Al, Ba, Co, Fe, Ni, Pb, Sb, V and Zn were fairly symmetrical; however the metal elements of Ca, K, Mn, Na, Sc, Sr and Ti indicated the data points were moderately and metal elements of As, Cd, Cr, Cu and Hg were highly skewed. Furthermore, the metal elements of Al, Ba, Ca, Co, Fe, K, Mn, Na, Ni, Pb, Sb, Sc, Sr, V and Zn exhibited platykurtic distribution, whereas the metal elements of As, Cd, Cr, Cu, Hg and Ti exhibited leptokurtic distribution. These results provide a quite reliable analysis.

#### **4.2.2.3 Seasonal Comparison of the Concentration of Metal Elements**

Based on descriptive statistical analysis, large SD was found for all metal elements, especially for Fe, Al, K and Ca in both the dry and rainy seasons indicated wide variation of their concentrations in soil. Table 4.3 showed the highest mean concentration of Fe and Ti both the dry and rainy season. Moreover, based on mean concentration, the level of metal elements in soil can be ordered as Fe> Al> K> Ca> Ba> Na> Pb> V> Ti> Sr> Zn> Mn> Sc> Cu> Sb>Co>Cr>Hg>As>Ni>Cd in dry season and Fe>Al>K>Ca>Ba>Na>V>Ti>Sr>Zn>Pb>Mn>Sc>Co>Cu>Sb>Cr>Hg>As>Ni>Cd in rainy season (Table 4.3). The concentrations of metal elements for rainy season were relatively lower as compared to the dry season and the magnitude of concentrations followed almost same pattern for both the dry and rainy seasons. Result reveals that the CV varied from 22.11% of Zn to 59.41% of Cu in dry season as well as 18.25% of Zn to 77.25% of Mn in rainy season, respectively, which indicated a great degree of variability. The greatest and the smallest SD were detected for metal element of Ti (406.26) and Cd (1.14) in the dry season. Similarly, the greatest and the smallest SD were detected for metal element of Ti (297.85) and Cd (0.74) in the rainy season. This result was also supported by a researcher of Reza et al. (2010) in case of soil for a landfill site.

#### **4.3 Multivariate Statistics**

In this study, multivariate analysis was performed to study how the variables are related to one another, and how they work in combination to distinguish between the cases on which the observations are made. The multivariate statistics including Pearson's correlation, principal component analysis (PCA) and agglomerative hierarchical clustering (AHC) were performed using XLSTAT and discussed in the following articles.



Table 4.4: Correlation analysis and coefficients for the metal elements in dry season

	Fe	Mn	Cr	Cu	Pb	Zn	Ni	Cd	As	Hg	Co	Na	K	Ca	Al	Ti	Sb	Sc	Sr	V	Ba
Fe	1.000	0.554	0.768	0.755	0.739	0.844	0.927	0.888	0.914	0.760	0.948	0.942	0.966	0.947	0.952	0.982	0.951	0.967	0.938	0.929	0.943
Mn	0.554	1.000	0.448	0.761	0.871	0.708	0.631	0.741	0.656	0.726	0.653	0.461	0.570	0.657	0.645	0.598	0.687	0.650	0.708	0.727	0.667
Cr	0.768	0.448	1.000	0.669	0.567	0.710	0.753	0.747	0.739	0.575	0.797	0.738	0.774	0.750	0.776	0.796	0.713	0.765	0.700	0.740	0.752
Cu	0.755	0.761	0.669	1.000	0.804	0.887	0.874	0.815	0.879	0.882	0.855	0.632	0.745	0.883	0.865	0.822	0.852	0.853	0.849	0.878	0.887
Pb	0.739	0.871	0.567	0.804	1.000	0.799	0.761	0.824	0.809	0.788	0.810	0.620	0.738	0.811	0.799	0.759	0.840	0.815	0.866	0.857	0.814
Zn	0.844	0.708	0.710	0.887	0.799	1.000	0.885	0.796	0.867	0.788	0.879	0.752	0.856	0.881	0.888	0.881	0.895	0.893	0.883	0.906	0.887
Ni	<u>0.927</u>	0.631	0.753	0.874	0.761	0.885	1.000	0.880	0.959	0.855	0.941	0.818	0.887	0.958	0.952	0.955	0.943	0.958	0.945	0.940	0.963
Cd	<u>0.888</u>	0.741	0.747	0.815	0.824	0.796	0.880	1.000	0.892	0.807	0.918	0.816	0.866	0.920	0.901	0.914	0.901	0.922	0.912	0.916	0.926
As	0.914	0.656	0.739	0.879	0.809	0.867	0.959	0.892	1.000	0.885	0.955	0.785	0.882	0.974	<u>0.974</u>	0.954	0.952	0.964	0.966	0.958	0.970
Hg	0.760	0.726	<u>0.575</u>	0.882	0.788	0.788	0.855	0.807	0.885	1.000	0.831	0.635	0.743	0.881	0.867	0.815	0.829	0.838	0.862	0.875	0.867
Co	0.948	0.653	0.797	0.855	0.810	0.879	0.941	0.918	0.955	0.831	1.000	0.879	0.926	0.970	0.960	0.973	0.956	0.976	0.956	0.958	0.971
Na	0.942	0.461	0.738	0.632	0.620	0.752	0.818	0.816	0.785	0.635	0.879	1.000	0.950	0.849	0.862	0.915	0.848	0.884	0.832	0.822	0.834
K	0.966	0.570	0.774	0.745	0.738	0.856	0.887	0.866	0.882	0.743	0.926	0.950	1.000	0.918	0.947	0.962	0.925	0.946	0.915	0.921	0.902
Ca	0.947	0.657	0.750	0.883	0.811	0.881	0.958	0.920	0.974	0.881	0.970	0.849	0.918	<u>1.000</u>	<u>0.979</u>	<u>0.978</u>	<u>0.972</u>	<u>0.985</u>	<u>0.974</u>	<u>0.970</u>	<u>0.992</u>
Al	0.952	<u>0.645</u>	0.776	0.865	0.799	0.888	0.952	0.901	0.974	0.867	0.960	0.862	0.947	<u>0.979</u>	<u>1.000</u>	<u>0.976</u>	<u>0.965</u>	<u>0.979</u>	<u>0.967</u>	<u>0.974</u>	<u>0.963</u>
Ti	0.982	0.598	0.796	0.822	0.759	0.881	0.955	0.914	0.954	0.815	0.973	0.915	0.962	<u>0.978</u>	<u>0.976</u>	<u>1.000</u>	<u>0.966</u>	<u>0.987</u>	<u>0.958</u>	<u>0.956</u>	<u>0.975</u>
Sb	0.951	0.687	0.713	0.852	0.840	0.895	0.943	0.901	0.952	0.829	0.956	0.848	0.925	<u>0.972</u>	<u>0.965</u>	<u>0.966</u>	<u>1.000</u>	<u>0.986</u>	<u>0.987</u>	<u>0.984</u>	<u>0.975</u>
Sc	0.967	0.650	0.765	0.853	0.815	0.893	0.958	0.922	0.964	0.838	0.976	0.884	0.946	<u>0.985</u>	<u>0.979</u>	<u>0.987</u>	<u>0.986</u>	<u>1.000</u>	<u>0.985</u>	<u>0.980</u>	<u>0.983</u>
Sr	0.938	0.708	0.700	0.849	0.866	0.883	0.945	0.912	0.966	0.862	0.956	0.832	0.915	<u>0.974</u>	<u>0.967</u>	<u>0.958</u>	<u>0.987</u>	<u>0.985</u>	<u>1.000</u>	<u>0.984</u>	<u>0.973</u>
V	0.929	0.727	0.741	0.878	0.857	0.906	0.940	0.916	0.958	<i>0.876</i>	0.958	0.822	0.921	<u>0.970</u>	<u>0.974</u>	<u>0.956</u>	<u>0.984</u>	<u>0.980</u>	<u>0.984</u>	<u>1.000</u>	<u>0.967</u>
Ba	0.943	0.667	0.753	0.887	0.814	0.887	0.963	0.926	<u>0.970</u>	0.867	0.971	0.834	0.902	<u>0.992</u>	<u>0.963</u>	<u>0.975</u>	<u>0.975</u>	<u>0.983</u>	<u>0.973</u>	<u>0.967</u>	<u>1.000</u>

Table 4.5: Correlation analysis and coefficients for the metal elements in rainy season

	Fe	Mn	Cr	Cu	Pb	Zn	Ni	Cd	As	Hg	Co	Na	K	Ca	Al	Ti	Sb	Sc	Sr	V	Ba
Fe	1.000	0.947	0.818	0.846	0.846	0.800	0.832	0.840	0.809	0.733	0.943	0.938	0.909	0.922	0.971	0.913	0.946	0.927	0.940	0.948	0.928
Mn	<b><u>0.947</u></b>	1.000	0.896	0.803	0.843	0.812	0.874	0.835	0.853	0.718	0.917	0.949	0.918	0.895	0.969	0.891	0.955	0.916	0.925	0.940	0.935
Cr	0.818	0.896	1.000	0.803	0.803	0.653	0.772	0.830	0.879	0.763	0.710	0.859	0.853	0.711	0.847	0.807	0.831	0.779	0.852	0.812	0.843
Cu	0.846	0.803	0.803	1.000	0.774	0.782	0.873	0.888	0.826	0.783	0.792	0.897	0.849	0.844	0.883	0.947	0.879	0.884	0.945	0.901	0.918
Pb	0.846	0.843	0.803	0.774	1.000	0.637	0.774	0.852	0.753	0.697	0.762	0.872	0.861	0.837	0.876	0.865	0.870	0.855	0.870	0.847	0.853
Zn	0.800	0.812	0.653	0.782	0.637	1.000	0.787	0.721	0.733	0.569	0.856	0.845	0.721	0.854	0.850	0.861	0.861	0.870	0.855	0.883	0.850
Ni	0.832	0.874	0.772	0.873	0.774	0.787	1.000	0.872	0.790	0.695	0.844	0.930	0.893	0.890	0.911	0.927	0.922	0.918	0.924	0.913	0.936
Cd	0.840	0.835	0.830	0.888	0.852	0.721	0.872	1.000	0.899	0.873	0.767	0.942	0.930	0.818	0.898	0.927	0.886	0.867	0.933	0.879	0.910
As	0.809	0.853	0.879	0.826	0.753	0.733	0.790	0.899	1.000	0.886	0.726	0.907	0.900	0.691	0.856	0.847	0.814	0.760	0.881	0.830	0.820
Hg	<b><u>0.733</u></b>	0.718	0.763	0.783	0.697	0.569	0.695	0.873	0.886	1.000	0.664	0.827	0.884	0.641	0.772	0.785	0.754	0.708	0.819	0.761	0.745
Co	0.943	0.917	0.710	0.792	0.762	0.856	0.844	0.767	0.726	0.664	1.000	0.910	0.855	0.957	0.949	0.904	0.941	0.935	0.914	0.961	0.927
Na	<b><u>0.938</u></b>	0.949	0.859	0.897	0.872	0.845	0.930	0.942	0.907	0.827	0.910	1.000	0.959	0.917	0.979	0.971	0.969	0.950	0.984	0.973	0.967
K	0.909	0.918	0.853	0.849	0.861	0.721	0.893	0.930	0.900	0.884	0.855	0.959	1.000	0.858	0.949	0.915	0.937	0.901	0.946	0.920	0.909
Ca	0.922	0.895	0.711	0.844	0.837	0.854	0.890	0.818	<b><u>0.691</u></b>	0.641	<b><u>0.957</u></b>	0.917	0.858	<b><u>1.000</u></b>	<b><u>0.956</u></b>	<b><u>0.943</u></b>	<b><u>0.967</u></b>	<b><u>0.985</u></b>	<b><u>0.937</u></b>	<b><u>0.971</u></b>	<b><u>0.961</u></b>
Al	0.971	0.969	0.847	0.883	0.876	0.850	0.911	0.898	0.856	0.772	0.949	0.979	0.949	<b><u>0.956</u></b>	<b><u>1.000</u></b>	<b><u>0.961</u></b>	<b><u>0.986</u></b>	<b><u>0.969</u></b>	<b><u>0.979</u></b>	<b><u>0.986</u></b>	<b><u>0.976</u></b>
Ti	0.913	0.891	0.807	0.947	0.865	<b><u>0.861</u></b>	0.927	0.927	0.847	0.785	0.904	0.971	0.915	<b><u>0.943</u></b>	<b><u>0.961</u></b>	<b><u>1.000</u></b>	<b><u>0.959</u></b>	<b><u>0.967</u></b>	<b><u>0.991</u></b>	<b><u>0.978</u></b>	<b><u>0.977</u></b>
Sb	0.946	0.955	0.831	0.879	0.870	<b><u>0.861</u></b>	0.922	0.886	0.814	0.754	0.941	0.969	0.937	<b><u>0.967</u></b>	<b><u>0.986</u></b>	<b><u>0.959</u></b>	<b><u>1.000</u></b>	<b><u>0.988</u></b>	<b><u>0.973</u></b>	<b><u>0.982</u></b>	<b><u>0.978</u></b>
Sc	0.927	0.916	0.779	0.884	0.855	0.870	<b><u>0.918</u></b>	0.867	0.760	0.708	0.935	0.950	0.901	<b><u>0.985</u></b>	<b><u>0.969</u></b>	<b><u>0.967</u></b>	<b><u>0.988</u></b>	<b><u>1.000</u></b>	<b><u>0.965</u></b>	<b><u>0.981</u></b>	<b><u>0.976</u></b>
Sr	0.940	0.925	0.852	0.945	0.870	0.855	0.924	<b><u>0.933</u></b>	0.881	0.819	0.914	0.984	0.946	<b><u>0.937</u></b>	<b><u>0.979</u></b>	<b><u>0.991</u></b>	<b><u>0.973</u></b>	<b><u>0.965</u></b>	<b><u>1.000</u></b>	<b><u>0.984</u></b>	<b><u>0.979</u></b>
V	<b><u>0.948</u></b>	0.940	0.812	0.901	0.847	0.883	0.913	0.879	0.830	0.761	<b><u>0.961</u></b>	0.973	0.920	<b><u>0.971</u></b>	<b><u>0.986</u></b>	<b><u>0.978</u></b>	<b><u>0.982</u></b>	<b><u>0.981</u></b>	<b><u>0.984</u></b>	<b><u>1.000</u></b>	<b><u>0.982</u></b>
Ba	0.928	0.935	0.843	0.918	0.853	0.850	0.936	0.910	0.820	0.745	0.927	0.967	0.909	<b><u>0.961</u></b>	<b><u>0.976</u></b>	<b><u>0.977</u></b>	<b><u>0.978</u></b>	<b><u>0.976</u></b>	<b><u>0.979</u></b>	<b><u>0.982</u></b>	<b><u>1.000</u></b>

### 4.3.1 Pearson's Correlation

In this study, Pearson's correlation coefficients were calculated for the concentrations of metal elements in soil for both the dry and rainy using XLSTAT. The interrelationship studies between different variables are very helpful tools in promoting research and opening new frontiers of knowledge. The study of correlation reduces the range of uncertainty associated with decision making (Rogers and Nicewander, 1988). The correlation analysis is a preliminary descriptive technique to estimate the degree of association among the variables involved. The purpose of the correlation analysis was to measure the intensity of association between two variables. Such association is likely to lead to reasoning about causal relationship between the variables. The Pearson's correlation matrix of metal elements in soil for both the dry and rainy seasons is provided in Table 4.4 and Table 4.5, respectively. The most significant correlation was observed for Ca and Ba (0.992) in dry season and Ti and Sr (0.991) in rainy season. In dry season, the concentration of Cr showed very weak correlations with Mn, indicated Mn is from different sources than Cr (Li et al., 2012).

However, the concentrations of Hg showed very weak correlations with Zn (0.569) in rainy season. This indicated that Hg was from different sources than Zn (Li et al., 2012). Based on the results of Pearson's correlations matrix on soil of waste disposal site during dry season, it was observed the high positively correlations between Ba and Ca (0.992), Sc and Ti (0.987), Sb and Sc (0.986), Sb and V (0.984), Sc and Ba (0.983), Al and Sc (0.979), Ti and Ba (0.975), As and Al (0.974), V and Ca (0.974) , As and Ba (0.970) Ti and Sb (0.966), Ni and Fe(0.927), Zn and V (0.906), Fe and Cd (0.888), Al and Mn (0.645),Cr and Hg (0.575) as well as Hg and V (0.876) (Table 4.4). In addition, the some parameters in on soil of waste disposal site in rainy season showed also high positively correlated as Sr and Ti (0.991), Al and V (0.986), V and Ba (0.982), Ca and Sc (0.985), Ti and V (0.978) and Co and Ca (0.957), V and Co (0.961), Fe and V (0.948), Fe and Na (0.938), Cd and Sr (0.933), Ni and Sc (0.918), Zn and Ti (0.861), Zn and Sb (0.861), Hg and Fe (0.733), Fe and Mn (0.947) and As and Ca (0.691) (Table 4.5) . Here it should be noted that from Table 4.4, the 8x8 sub matrix at the bottom-right square have similar values indicating strong correlation between metal elements variables. Concentrations of Ca, Al, Ti, Sb, Sc, Sr, V, Ba showed strong correlation with each other in dry season, which indicated same sources of contamination for these metals. Similarly, the 8x8 sub matrix in Table 4.5 also

showed strong correlation between Ca, Al, Ti, Sb, Sc, Sr, V and Ba in rainy season (Olawoyin, 2012) indicating same sources of contamination.

### 4.3.2 Principal Component Analysis

In principal component analysis (PCA), the principal components (PCs) with variables, the high loadings (eigenvalues) depicted greater importance from the contamination sources, whereas, lower loadings (eigenvalues) point to lower importance with regards the sources of these contaminations (Lee et al., 2006; Zou et al., 2015).

Table 4.6: PCA of metal elements in soil for dry and rainy seasons

PCs	Dry season			Rainy season		
	Eigenvalue	Variability (%)	Cumulative (%)	Eigenvalue	Variability (%)	Cumulative (%)
F1	18.267	86.987	86.987	18.533	88.253	88.253
F2	1.07	5.118	92.105	0.906	4.3147	92.567
F3	0.416	1.98	94.085	0.398	1.8962	94.464
F4	0.391	1.861	95.946	0.335	1.5943	96.058
F5	0.23	1.095	97.041	0.263	1.2521	97.310
F6	0.162	0.774	97.814	0.180	0.8575	98.168
F7	0.119	0.566	98.38	0.151	0.7183	98.886
F8	0.076	0.362	98.742	0.080	0.3829	99.269
F9	0.058	0.277	99.019	0.055	0.26	99.529
F10	0.051	0.244	99.263	0.043	0.2068	99.736
F11	0.042	0.202	99.465	0.016	0.077	99.813
F12	0.035	0.167	99.633	0.015	0.0734	99.886
F13	0.026	0.122	99.755	0.009	0.0452	99.931
F14	0.016	0.077	99.832	0.007	0.0318	99.963
F15	0.013	0.063	99.895	0.003	0.0165	99.979
F16	0.008	0.036	99.931	0.002	0.0121	99.991
F17	0.004	0.02	99.952	0.001	0.0047	99.996
F18	0.004	0.018	99.97	0.0005	0.0025	99.999
F19	0.003	0.013	99.983	0.0002	0.001	100
F20	0.002	0.01	99.993			
F21	0.001	0.007	100			

A research conducted by Olawoyin et al. (2012) and stated when the eigenvalues less than unique, the PCs as considered of insignificant. Based on this postulation, in PCA, for cut

off of PCs, eigenvalues of 1 was considered as reference value. In this study, most of the metal element's correlation showed similar trends which indicated same sources of contamination. Moreover, for variability calculation based on eigenvector and factor loadings, PCs of 21 for dry and 19 for rainy were considered for the metal elements and results provided in Table 4.6. The larger eigenvalue obtained for F1 (18.267) indicated large proportion of variability (86.987%) for dry season (Table 4.6). Based on the results of PCA for metal elements of dry season, the eigenvalues upto the second extracted components (F2) were found greater than 1.0. Moreover, as the eigenvalues for the PCs (F3 to F21) as well as (F2 to F19) were found less than 1 for dry and rainy season, respectively, so, these PCs were depicted as very negligible contribution in the analysis. Thus, the variables could be reduced to 2 components model (dry season) with 92.105% variation as well as 1 component model (rainy season) that accounts for 88.253% variation (Table 4.6). In addition, the larger eigenvalue obtained for F1 (18.533) that indicated large proportion of variability (88.253%) for rainy season (Table 4.6). The percentage contribution of the 1<sup>st</sup> to 2<sup>nd</sup> PCs for the metal elements in dry season as represented in Equation 4.1 was 92.105%.

$$\left( \left( \sum_{i=1}^2 \lambda_i \right) / \left( \sum_{i=1}^{21} \lambda_i \right) \right) \dots \dots \dots (4.1)$$

The selection of the first two parameters as the PCs since there was significant evidence of high enough total variance from the percentage contributions. The eigenvalues  $\lambda_i$  ( $\lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8, \lambda_9, \lambda_{10}, \lambda_{11}, \lambda_{12}, \lambda_{13}, \lambda_{14}, \lambda_{15}, \lambda_{16}, \lambda_{17}, \lambda_{18}, \lambda_{19}, \lambda_{20}, \lambda_{21}$ ), had little contributions to the total structure of the data under study. The percentage contribution of the 3th to 21st components (Equation 4.2) is 7.895% (Olawoyin, 2012). This suggested that very little information, which can be considered negligible, will be lost.

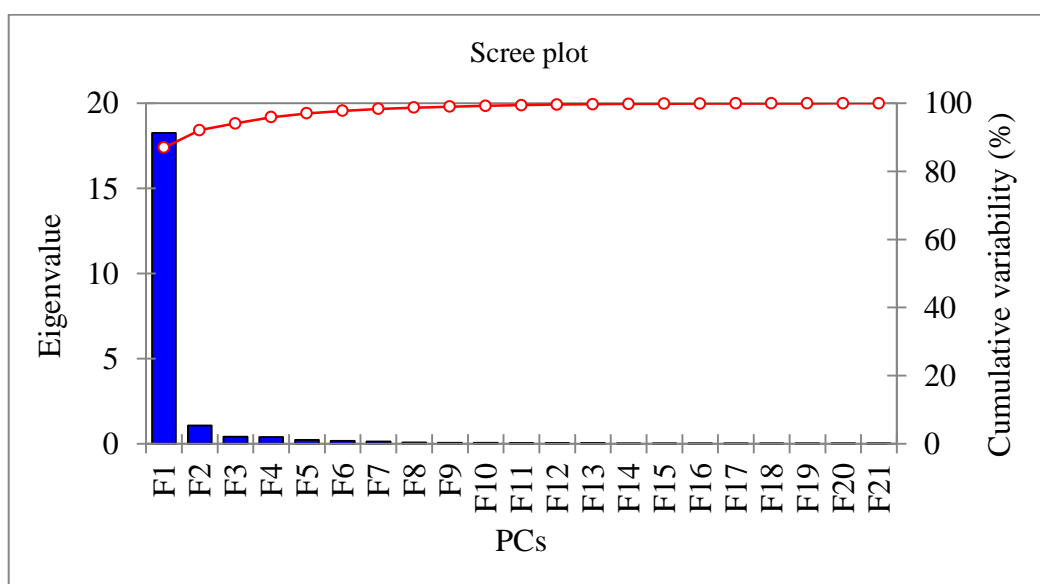
$$\left( \left( \sum_{i=3}^{21} \lambda_i \right) / \left( \sum_{i=1}^{21} \lambda_i \right) \right) \dots \dots \dots (4.2)$$

However, the percentage contribution of the 1st PC for the metal elements in rainy season as represented in Equation 4.3 was 88.25%

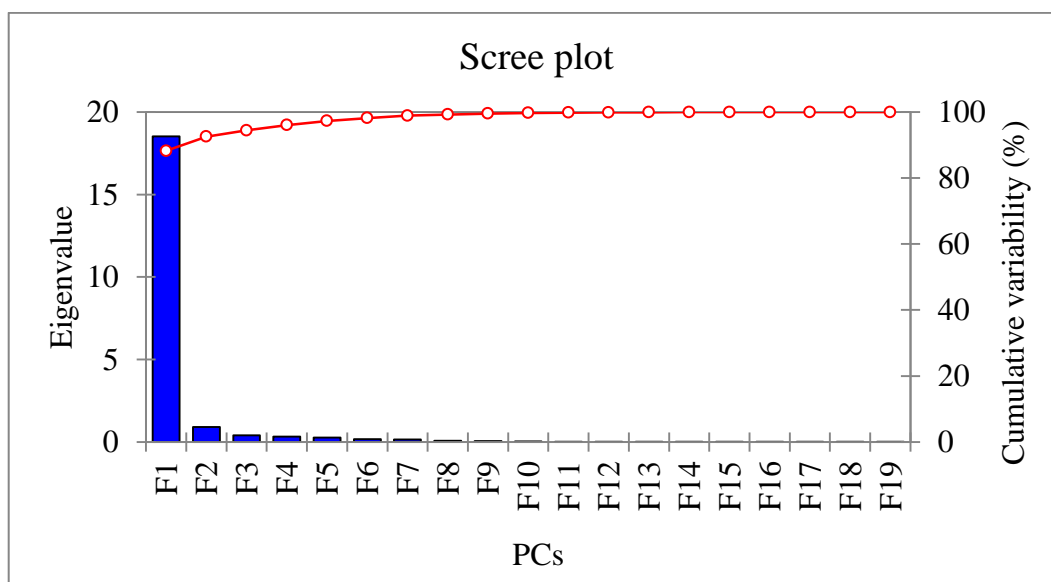
$$\left( \left( \sum_{i=1}^1 \lambda_i \right) / \left( \sum_{i=1}^{18} \lambda_i \right) \right) \dots \dots \dots (4.3)$$

Which verbalized the selection of the first three parameters (eigenvalues, variability and cumulative) as the PCs and the eigenvalues  $\lambda_i$  ( $\lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8, \lambda_9, \lambda_{10}, \lambda_{11}, \lambda_{12}, \lambda_{13}, \lambda_{14}, \lambda_{15}, \lambda_{16}, \lambda_{17}, \lambda_{18}, \lambda_{19}$ ), had minimal contributions to the nature of the general data. The percentage contribution of the 2<sup>nd</sup> to 19<sup>th</sup> PCs as illustrated in Equation 4.4 was 11.75% for rainy season.

$$\left( \left( \sum_{i=2}^{19} \lambda_i \right) / \left( \sum_{i=1}^{19} \lambda_i \right) \right) \dots \dots \dots (4.4)$$



(a)



(b)

Figure 4.3: Scree plot of the PCs (a) dry season; (b) rainy season.

The scree plots for the factor variability of metal elements in both the dry and rainy season are presented in Figure 4.3 (a) and Figure 4.3(b) respectively. The scree plots of PCs were represented in a form of simple line segment plot that shows the fraction of total variance in the data. In this PCA, varimax rotation was applied to simplify the factor interpretation by reducing total number of variables that exhibit high loadings per factor. For all the metal elements analyzed, correlation matrix was built with equal weights. Table 4.7 showed the factor analysis of PCA before and after the varimax rotation for dry season. Natural and anthropogenic sources are one of the root cause of metal element contamination which has caused widespread and variable the hazardous possibilities of environmental and health effect (Tahir et al., 2007). Moreover, some previous investigations indicated first principal component (PC1) and second component (PC2) refers to the contamination of soil due to anthropogenic or human activities and natural parent materials, respectively (Tahir et al., 2007). In this study, factor analysis revealed that metal elements of Al, As, Ba, Ca, Cd, Co, Fe, K, Na, Ni, Sb, Sc, Sr, Ti and V were closely related to PC1 indicated derived from anthropogenic activities and rests of the metal elements of Cu, Hg, Mn, Pb and Zn in soil were related to PC2 indicated derived from natural parent materials. In addition, the metal element of Cr was closed to PC1 and PC2 indicated derived from both the anthropogenic activities and natural parent materials.

In addition, the factor analysis of PCA for rainy season before and after the varimax rotation is provided in Table 4.8. It can be estimated that As, Cr, Hg, K, Mn and Na in soil were related to PC2 indicating derived from natural parent materials. Other metal elements of Al, Ba, Ca, Co, Cu, Fe, Ni, Sb, Sc, Sr, Ti and V were related to PC1 indicating derived from anthropogenic activities. Moreover, as the metal elements of Al, Ba, Ca, Co, Cu, Fe, Ni, Sb, Sc, Sr, Ti and V in soil showed closed to PC1 indicated derived from anthropogenic activities.

Table 4.7: All explained variables and factors derived using the orthogonal varimax rotation method of dry season

Metal	Before rotation																					After rotation	
	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	D1	D2
Fe	0.96	-0.24	0.02	-0.11	0.01	0.01	-0.01	-0.01	0.05	0.01	-0.01	-0.04	-0.05	-0.08	-0.04	0.00	-0.03	0.00	0.00	0.00	0.00	0.91	0.37
Mn	0.72	0.61	0.29	-0.07	0.00	0.10	-0.05	-0.11	0.02	-0.07	0.02	0.02	-0.03	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.91
Cr	0.78	-0.28	0.32	0.44	-0.08	-0.05	0.08	-0.03	0.01	-0.01	0.02	-0.03	-0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.79	0.24
Cu	0.89	0.28	-0.11	0.25	0.09	0.08	-0.10	0.13	-0.04	-0.08	-0.06	-0.03	0.01	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00	0.55	0.75
Pb	0.85	0.40	0.18	-0.14	0.00	-0.16	0.13	0.13	0.06	0.04	-0.04	-0.02	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.45	0.82
Zn	0.92	0.09	0.02	0.10	0.35	0.01	-0.02	-0.02	-0.01	0.12	0.04	0.03	-0.01	0.01	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.68	0.62
Ni	0.96	-0.04	-0.13	0.08	0.01	-0.01	-0.07	-0.09	0.15	0.02	-0.07	-0.04	0.06	0.00	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.80	0.53
Cd	0.94	0.04	0.14	-0.04	-0.23	0.04	-0.16	0.02	-0.07	0.12	-0.03	0.01	0.03	0.00	-0.01	0.00	0.00	0.00	0.00	-0.01	0.00	0.73	0.59
As	0.97	0.01	-0.15	0.05	-0.07	-0.08	0.03	-0.04	0.02	-0.02	-0.05	0.11	-0.03	0.01	-0.02	-0.04	0.00	-0.02	-0.01	0.00	0.00	0.77	0.59
Hg	0.88	0.27	-0.25	0.07	-0.13	0.20	0.17	-0.01	0.02	0.05	0.05	-0.03	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.54	0.74
Co	0.98	-0.08	0.01	0.02	-0.03	-0.02	-0.02	0.08	0.05	-0.03	0.08	0.10	0.06	-0.04	0.02	0.01	0.00	0.01	0.01	-0.01	0.00	0.84	0.52
Na	0.87	-0.38	0.15	-0.18	0.04	0.19	0.00	0.07	0.05	-0.04	0.01	0.00	0.02	0.04	-0.03	0.00	0.01	-0.02	0.00	0.00	0.00	0.92	0.21
K	0.94	-0.23	0.10	-0.12	0.09	0.08	0.09	-0.01	-0.06	0.00	-0.07	0.01	0.01	-0.01	0.07	-0.02	-0.01	0.00	0.00	-0.01	0.00	0.90	0.37
Ca	0.99	-0.03	-0.10	0.00	-0.05	-0.01	-0.02	0.04	-0.02	-0.01	0.02	-0.01	-0.06	0.03	0.02	0.05	-0.01	-0.01	-0.02	-0.01	-0.01	0.81	0.56
Al	0.98	-0.07	-0.07	0.01	0.00	0.01	0.07	-0.03	-0.06	-0.03	-0.08	0.03	-0.01	-0.01	-0.03	0.04	0.02	0.01	0.02	0.00	0.01	0.83	0.53
Ti	0.98	-0.17	-0.02	-0.02	0.00	0.01	-0.02	-0.01	0.00	0.01	0.01	0.02	-0.04	0.00	0.02	-0.01	0.02	0.02	-0.01	0.03	-0.02	0.89	0.44
Sb	0.98	-0.01	-0.04	-0.10	0.05	-0.10	-0.03	-0.03	-0.03	-0.06	0.04	-0.06	0.00	-0.02	-0.01	-0.02	0.03	0.00	0.00	-0.02	0.00	0.80	0.57
Sc	0.99	-0.08	-0.03	-0.05	0.00	-0.05	-0.02	0.00	-0.02	-0.01	0.02	-0.02	0.02	0.04	-0.01	-0.01	-0.01	0.04	-0.02	0.00	0.02	0.85	0.52
Sr	0.98	0.03	-0.06	-0.11	-0.01	-0.10	0.02	-0.04	0.00	-0.02	0.02	-0.01	0.02	0.05	-0.02	-0.01	-0.03	0.00	0.02	0.00	-0.02	0.77	0.61
V	0.99	0.05	-0.02	-0.03	0.01	-0.05	0.04	-0.06	-0.09	-0.02	0.04	-0.03	0.06	-0.03	0.00	0.01	-0.01	-0.02	-0.02	0.02	0.00	0.77	0.62
Ba	0.99	-0.02	-0.09	0.02	-0.04	-0.05	-0.08	0.03	0.01	0.01	0.05	-0.02	-0.05	0.01	0.04	-0.01	0.00	-0.02	0.02	0.01	0.02	0.80	0.57



Table 4.8: All explained variables and factors derived using the orthogonal varimax rotation method of rainy season

Metal	Before rotation																			After rotation
	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	D1
Fe	0.95	-0.08	0.14	0.09	0.11	-0.06	0.12	0.13	0.04	-0.08	-0.04	-0.02	-0.01	0.02	-0.01	0.01	0.00	0.00	0.00	0.95
Mn	0.95	-0.04	0.19	0.21	-0.03	0.08	-0.01	0.00	0.00	0.00	0.04	-0.03	-0.02	-0.02	0.00	-0.02	0.00	0.01	0.00	0.95
Cr	0.87	0.29	0.19	0.20	-0.25	0.01	0.12	-0.10	0.02	0.01	-0.04	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.87
Cu	0.92	0.07	-0.22	-0.14	-0.15	-0.07	0.21	0.06	-0.05	-0.03	0.03	-0.02	0.01	-0.01	0.01	-0.01	0.00	0.00	0.00	0.92
Pb	0.88	0.08	0.33	-0.21	-0.07	-0.19	-0.13	0.01	-0.07	0.03	0.00	-0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.88
Zn	0.85	-0.29	-0.28	0.23	-0.07	-0.15	-0.14	-0.05	-0.01	-0.04	-0.01	-0.01	0.02	0.01	0.00	-0.01	0.00	0.00	0.00	0.85
Ni	0.93	-0.06	-0.09	-0.11	-0.13	0.29	-0.07	0.04	-0.05	0.01	-0.02	-0.03	0.01	0.02	-0.01	0.00	0.00	0.00	0.00	0.93
Cd	0.93	0.25	-0.08	-0.15	-0.02	-0.01	-0.10	0.02	0.17	-0.02	0.00	0.01	0.02	0.00	0.01	0.00	0.01	0.00	0.00	0.93
As	0.88	0.38	-0.10	0.21	-0.02	-0.03	-0.10	0.10	-0.03	0.03	0.03	0.02	-0.02	0.00	-0.02	0.01	0.01	0.00	0.00	0.88
Hg	0.81	0.48	-0.14	-0.03	0.26	-0.02	0.05	-0.11	-0.03	0.02	0.00	-0.04	0.00	0.02	-0.01	0.00	0.00	0.00	0.00	0.81
Co	0.93	-0.28	0.03	0.10	0.19	0.03	0.06	0.03	0.01	0.10	-0.02	0.00	0.04	-0.02	0.00	0.00	0.01	-0.01	0.00	0.93
Na	0.99	0.05	-0.02	0.02	0.02	0.02	-0.07	0.03	0.01	0.02	-0.04	-0.04	-0.04	-0.01	0.03	0.00	-0.01	-0.01	0.00	0.99
K	0.96	0.18	0.07	-0.03	0.13	0.11	-0.04	0.00	-0.05	-0.07	0.00	0.06	0.02	-0.01	0.01	-0.01	0.00	0.00	0.01	0.96
Ca	0.94	-0.30	0.04	-0.11	0.05	-0.01	0.01	-0.04	0.02	0.01	0.02	0.01	-0.01	0.04	0.01	0.00	0.02	0.00	0.00	0.94
Al	0.99	-0.06	0.07	0.04	0.04	0.01	0.00	0.03	0.01	-0.02	0.05	0.02	0.01	0.03	0.00	0.00	-0.02	-0.01	-0.01	0.99
Ti	0.98	-0.04	-0.11	-0.11	-0.04	-0.06	0.00	0.02	-0.01	0.05	-0.04	0.03	-0.01	-0.01	-0.02	-0.02	0.00	0.00	0.00	0.98
Sb	0.99	-0.11	0.05	-0.01	0.02	0.03	-0.01	-0.08	-0.01	-0.06	0.02	-0.02	0.01	-0.03	0.00	0.02	0.00	0.00	0.00	0.99
Sc	0.97	-0.19	0.00	-0.08	0.00	0.00	0.00	-0.09	0.00	-0.06	-0.02	0.02	-0.04	-0.01	-0.02	0.00	0.00	-0.01	-0.01	0.97
Sr	0.99	0.01	-0.06	-0.04	-0.01	-0.03	0.03	0.02	-0.02	0.00	-0.02	0.03	0.01	0.00	0.01	0.02	0.00	0.01	0.00	0.99
V	0.99	-0.12	-0.03	0.01	0.04	-0.02	0.02	-0.02	-0.03	0.06	0.00	0.03	-0.03	0.01	0.02	0.01	0.00	0.00	0.00	0.99
Ba	0.98	-0.09	-0.01	-0.05	-0.07	0.03	0.05	-0.03	0.08	0.06	0.03	0.00	0.00	-0.01	-0.02	0.01	-0.01	0.00	0.01	0.98

In this study, metal elements were distributed in a correlation circle which indicated four distinct quadrants relative to the PCs for dry season shown in Figure 4.4. This circle exhibited a graphical representation of loading vectors for metal elements in order to determine the most influential variables (metal elements) interaction in this model. All the metal elements were positively correlated as the appearance of variables on the same quadrant (Lu et al., 2007). The impact of any individual variable to the entire PCA model was measured by the distance by how far the variable is from the origin (Olawoyin et al., 2012).

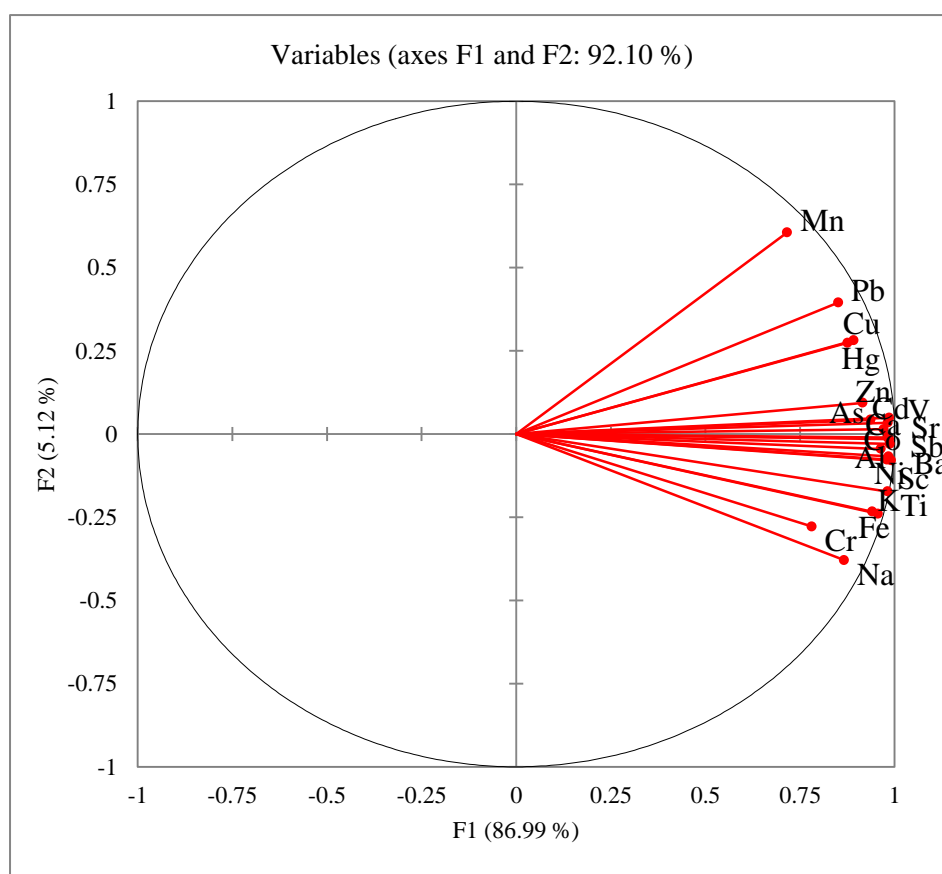


Figure 4.4: Correlation circle for metal elements in soil for dry season.

Variables that showed longer distances from origin of circle have larger impacts on the general architecture of the model than variables with shorter distances (Olawoyin et al., 2012). In the loading plot corresponding to the first two PCs (Figure 4.4); the following eight metal elements of Ca, Al, Ti, Sb, Sc, Sr, V and Ba showed clear positive correlation because they were in the same quadrant. In addition, these metal elements have stronger impact on the PCA model because these metal elements were at a longer distance from

origin. Besides, the metal elements of Na, Pb, Cu, K, Ni, Co, Hg, Fe, As, Zn and Cd, which also positively correlated with less stronger impact because they were comparative in shorter distance from origin than that of Ca, Al, Ti, Sb, Sc, Sr, V and Ba. Moreover, the metal elements of Mn and Cr having the least impact on the PCA model because they were far from each other in the same quadrant. It was noticed that Cu, Hg, Mn, Pb and Zn in soil were located at a distance from origin of circle than that of other metal elements. This indicated the origin of these metal elements was differing from other metal elements.

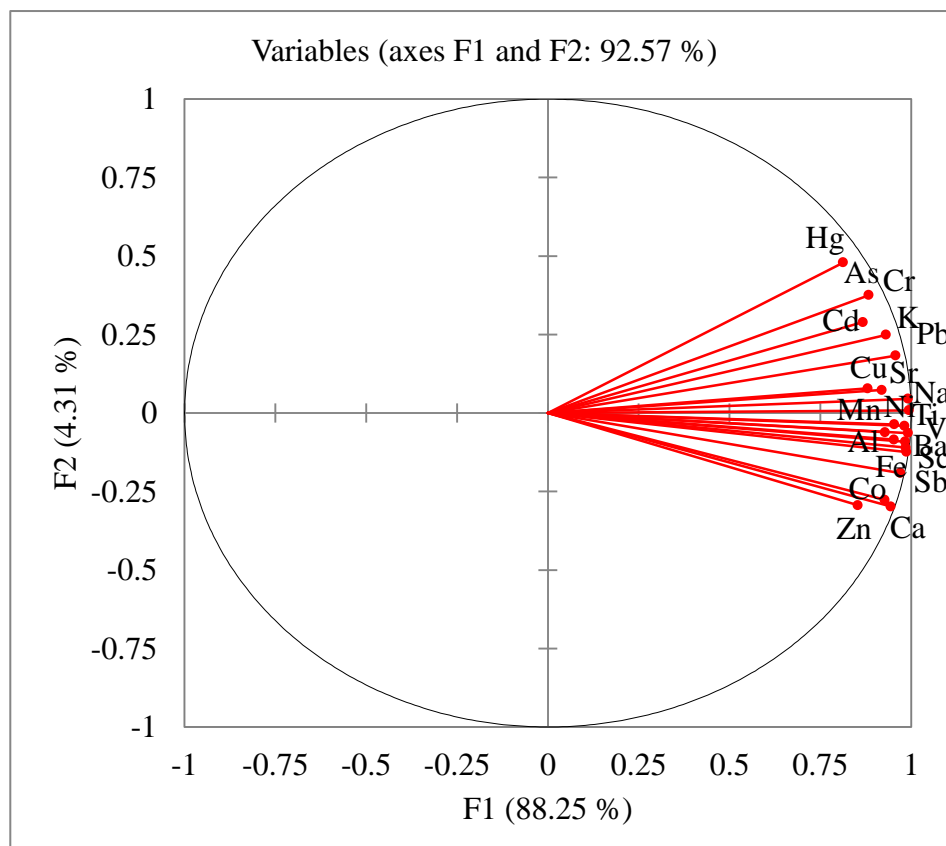


Figure 4.5: Correlation circle for metal elements in soil for rainy season.

Figure 4.5 illustrates the projection of variables and their correlation in the factor space for rainy season. It was clearly illustrated that 8 variables; Ca, Al, Ti, Sb, Sc, Sr, V and Ba in soil showed clear positive correlation but they have stronger impact on the PCA model than that of Na, Pb, Cu, K, Ni, Co, Mn, Fe, As, Cr and Cd, which also correlates positively with 8 variables, whereas Hg and Zn having the least impact on the PCA model (Olawoyin et al., 2012). It was also found from PCA analysis that Mn, Hg, As, Cr, Na and K were located far from other metal elements indicating different origin from other metal elements.

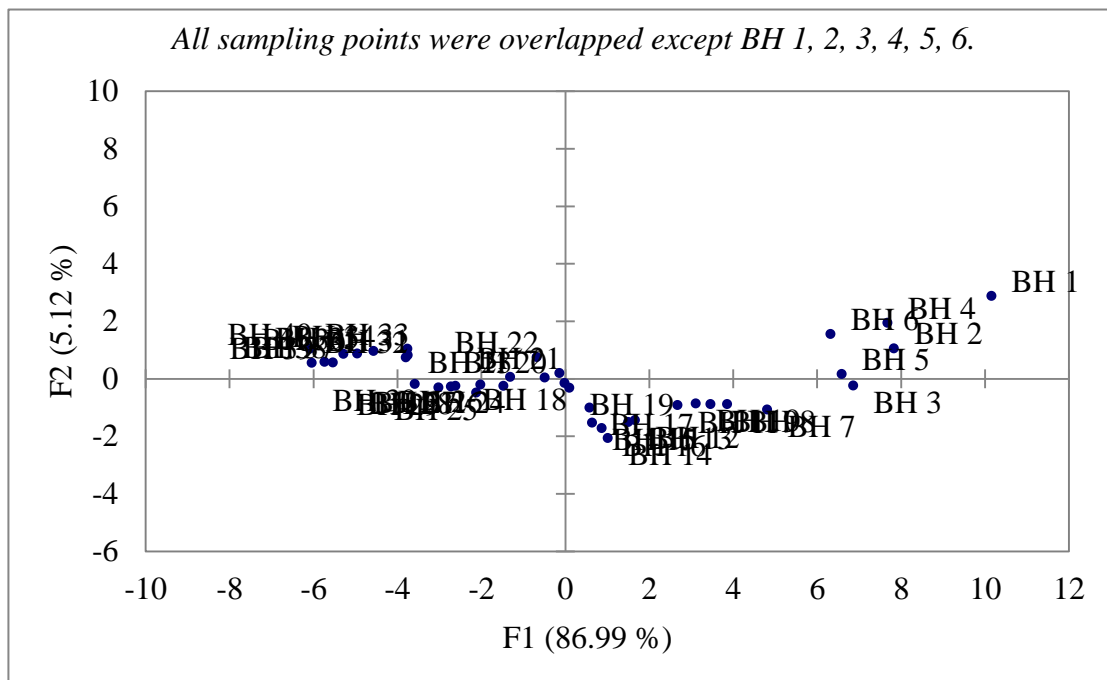


Figure 4.6: Correlation plot of soil sampling locations of dry season.

The correlation soil sampling locations for dry season is shown in Figure 4.6. Figure 4.6 reveals that the same soil sampling location indicated similar pattern of concentration of metal elements in soil. The biplot of soil sampling locations and concentration of metals elements in soil depicts in Figure 4.7.

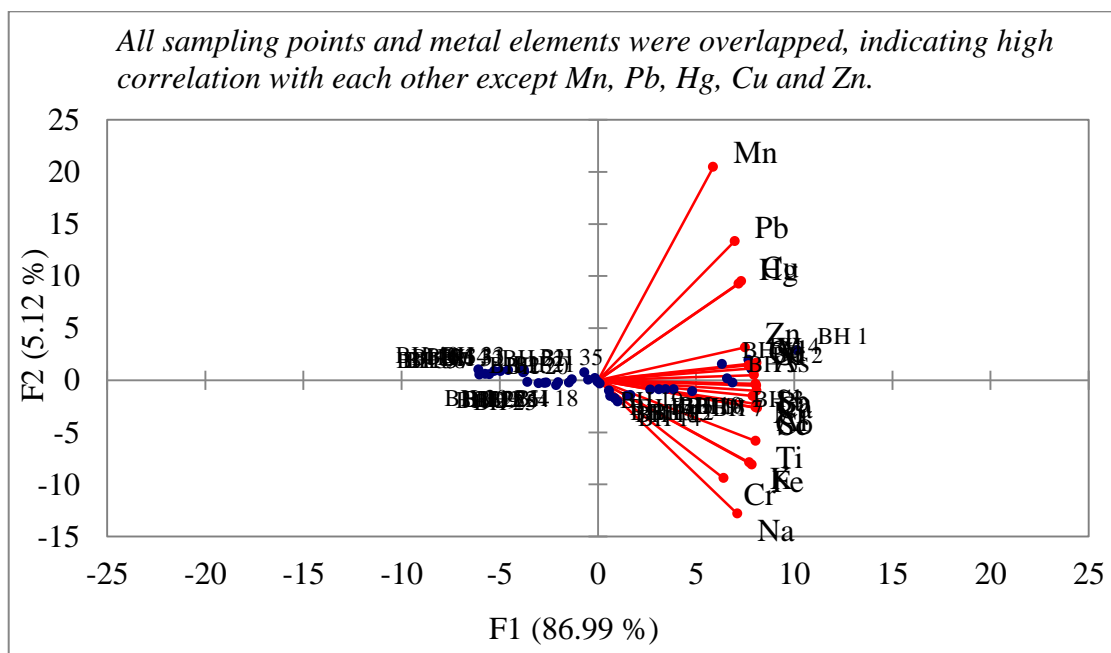


Figure 4.7: Biplot of soil sampling locations and concentrations of metal elements of dry season.

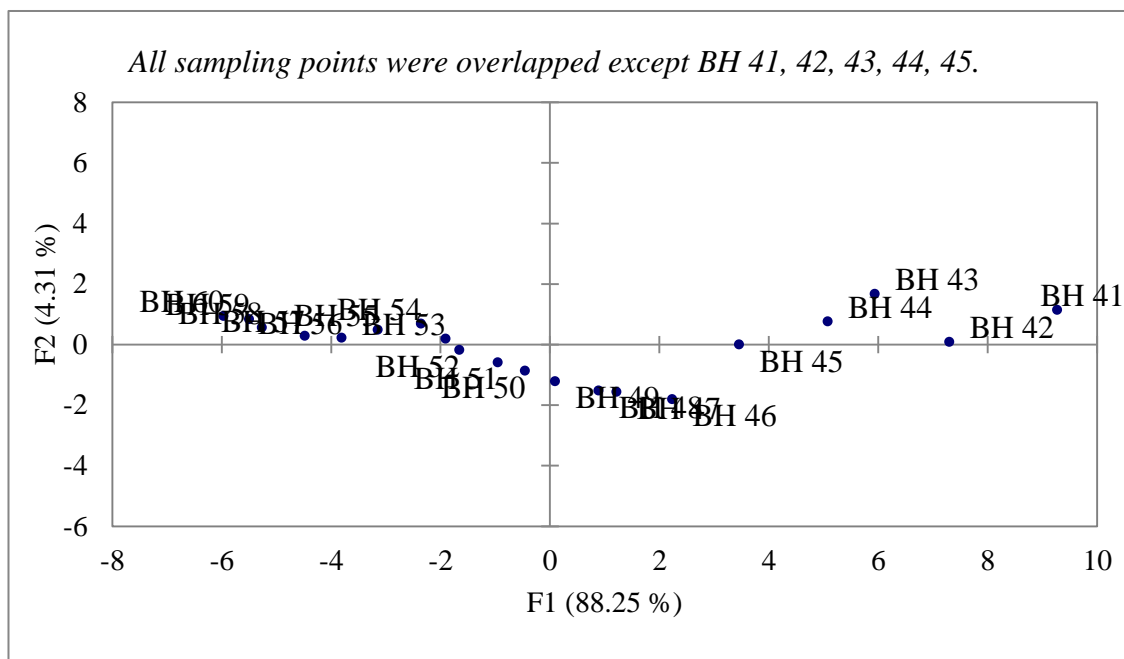


Figure 4.8: Correlation plot of soil sampling locations of rainy season.

Figure 4.7 illustrates the correlation between the observation locations, indicating a nucleated pattern for the most of locations to depict the correlation between the variables and how strongly they impacted the overall model with respect to the observations during dry season (Olawoyin et al., 2012).

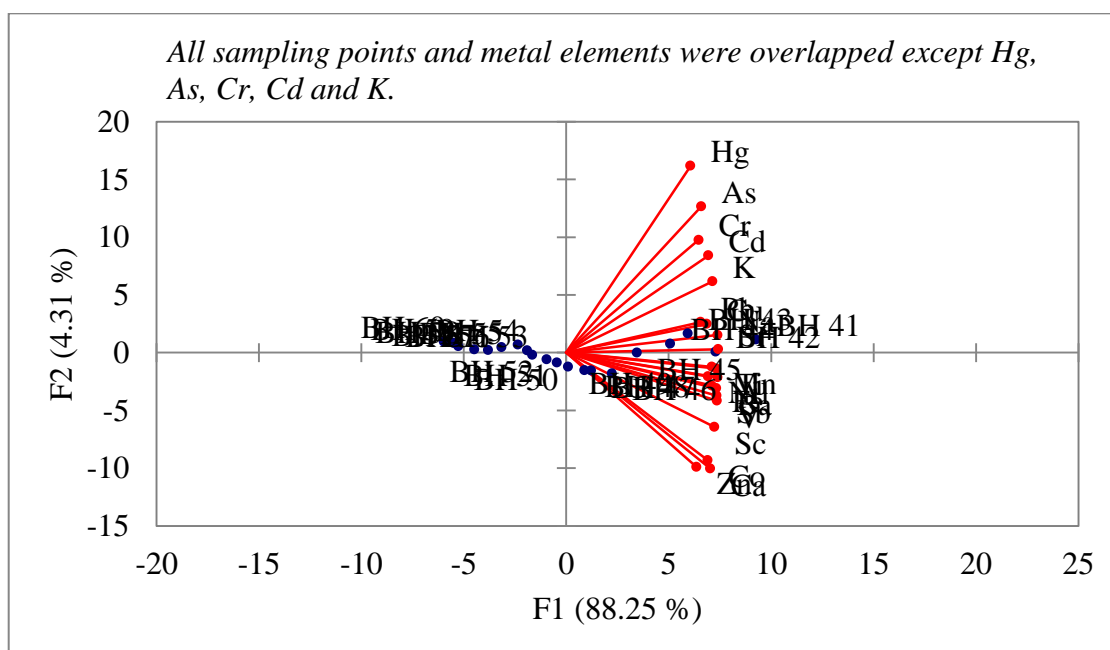


Figure 4.9: Biplot plot of the locations for metal elements (rainy) in soil of waste disposal site.

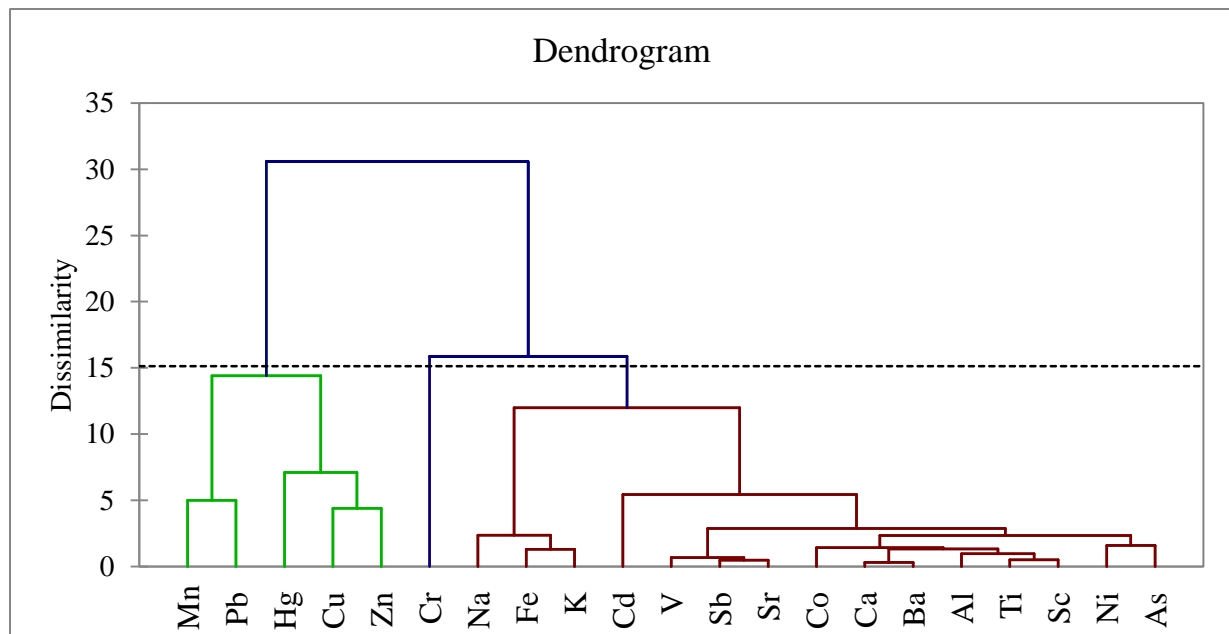
This suggested high similarities between the locations, but due to locational, geological and concentration there were some differences too. There seem to be a consistent trend in the correlation between the metal elements concentrations in soil of waste disposal site. Figure 4.8 and Figure 4.9 illustrated the correlation between the observation locations and the biplots, respectively, for the metal elements variables in rainy season. These plots showed positive correlations between the variables, observations and all areas with positive correlation (Olawoyin et al., 2012).

Based on the results of PCA for metal elements in dry season, it can be decided that the metal elements of Cu, Hg, Mn, Pb, and Zn derived from natural sources as they were closed to PC2, whereas, Al, As, Ba, Ca, Cd, Co, Fe, K, Na, Ni, Sb, Sc, Sr, Ti and V were also derived from anthropogenic activities because they were closed to PC1. Nevertheless, the metal element of Cr also derived from both natural soil parent materials and anthropogenic sources. In context of rainy season, the PCA results demonstrated that metal elements of As, Cr, Hg, K, Mn and Na were derived from different sources as the location of these metal elements were distinct from other metal elements of Al, Ba, Ca, Co, Cu, Fe, Ni, Sb, Sc, Sr, Ti and V. Moreover, it can be estimated as the metal elements of Al, Ba, Ca, Co, Cu, Fe, Ni, Sb, Sc, Sr, Ti and V were derived from anthropogenic activities and As, Cr, Hg, K, Na and Mn were derived from natural parent materials. Other metal elements of Cd, Pb and Zn were derived from both natural parent materials and anthropogenic sources.

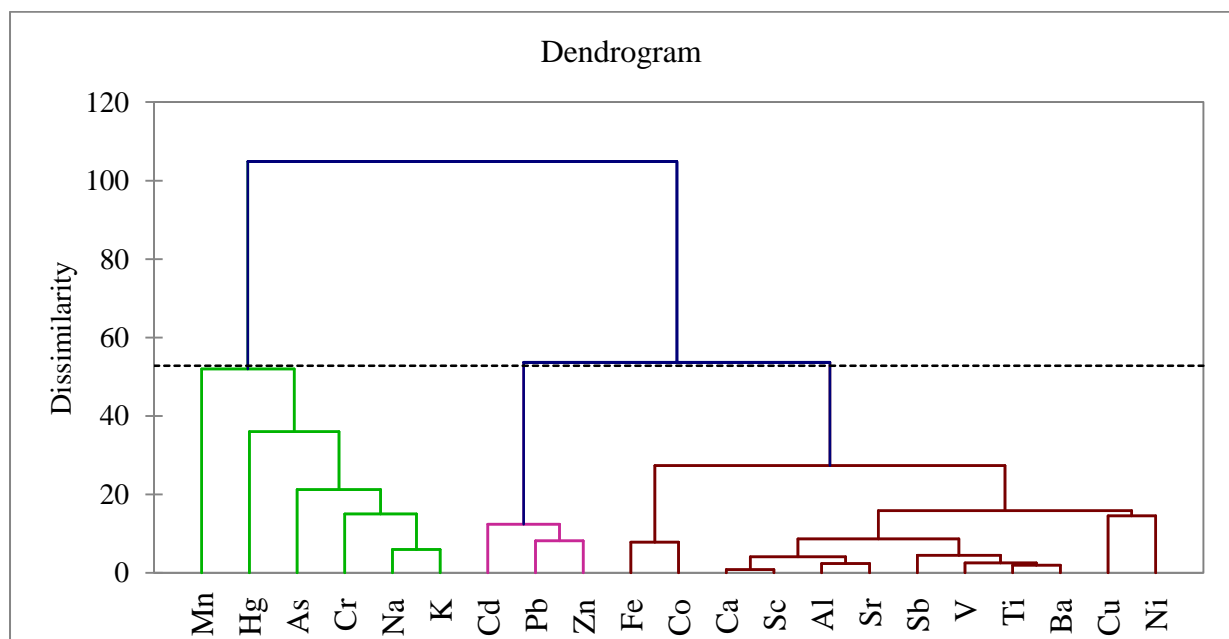
#### **4.3.3 Cluster Analysis: Agglomerative Hierarchical Clustering**

Agglomerative Hierarchical Clustering (AHC) is one of the most popular clustering methods. It is an iterative classification method whose principle is simple. In this study, AHC analysis was performed using XLSTAT software. Figure 4.10(a) and Figure 4.10(b) represents the hierarchical dendrograms for the classification of variables (metal elements) presence in soil for dry and rainy season, respectively. Clustering obtained by cutting the dendrogram at a desired level, each connected metal elements formed a cluster. The cluster formed for both the dry and rainy seasons were found to be very similar results as PCA. A study conducted by Anderberg (1973) reviewed that in case of soil, the minimum distance to centroid referred to close relation between generation sources of metal elements. The

metal elements exhibited smallest distance indicated generation from both the natural and anthropogenic activities; whereas the largest maximum distance to centroid indicated the possible pollutant generation source is natural parent material, while comparative smaller value indicated pollutant due to anthropogenic activities.



(a)



(b)

Figure 4.10: Dendrogram for metal elements in soil during (a) dry and (b) rainy season.

In case of dry season, cluster 1 comprises with metal elements of Al, As, Ba, Ca, Cd, Co, Fe, K, Na, Ni, Sb, Sc, Sr, Ti and V in soil which indicated these metal elements were generated from anthropogenic activities . In addition, cluster 2 comprises with Cu, Hg, Mn, Pb, and Zn, indicating origination from natural sources and Cluster 3 comprises with Cr which derived from both the natural parent materials and anthropogenic sources. Table 4.9 and Table 4.10 showed the results of cluster analysis by class for both the dry and rainy season, respectively.

Table 4.9: Results of cluster analysis for dry season

Class	1	2	3
Objects	15	5	1
Sum of weights	15	5	1
Within-class variance	4511.6859	18600.6792	0.0000
Minimum distance to centroid	24.7282	91.2381	0.0000
Average distance to centroid	58.8082	118.9985	0.0000
Maximum distance to centroid	122.5230	170.2796	0.0000
	Fe, Ni, Cd, As, Co, Na, K, Ca, Al, Ti, Sb, Sc, Sr, V, Ba	Mn, Cu, Pb Zn, Hg	Cr

In this study, maximum distance to centroid was found for cluster 2 of 170.279 between three clusters, indicating generation of metal elements from natural sources (Table 4.9). Moreover, maximum distance to centroid for cluster 1 was found comparatively smaller of 122.5230 than that of cluster 2, indicating generation of metal elements from anthropogenic or human activities. Cluster 3 showed maximum distance to centroid was zero, indicating it was closed to both the clusters, consequently generated from both natural sources and anthropogenic activities. In addition, for rainy season, cluster 1 comprises with metal elements of Al, Ba, Ca, Co, Cu, Fe, Ni, Sb, Sc, Sr, Ti and V which indicated these metal elements were generated from anthropogenic activities (Figure 4.10). Cluster 2 comprises with As, Cr, Hg, Mn, K and Na indicating origination from natural sources and Cluster 3 comprises with Cd, Pb and Zn which derived from both natural parent materials and anthropogenic sources.

Maximum distance to centroid cluster 2 was found as 148.207 between three clusters, indicating generation of metal elements from natural sources. The maximum distance to centroid was found to be smaller for cluster 1 (89.064) than that of cluster 2 (148.2072), indicating generation of metal elements from anthropogenic activities. Cluster 3 showed



maximum distances to centroid was smaller (64.249) than that of cluster 1 and 2, indicating it was closed to both classes, consequently generated from both natural and anthropogenic sources (Table 4.10).

Table 4.10: Results of cluster analysis for rainy season

Class	1	2	3
Objects	12	6	3
Sum of weights	12	6	3
Within-class variance	3042.8452	12931.7509	4559.6888
Minimum distance to centroid	25.5419	66.0451	41.0257
Average distance to centroid	48.7729	99.0971	54.2642
Maximum distance to centroid	89.0645	148.2072	64.2498
	Fe, Cu, Ni, Co, Ca, Al, Ti, Sb, Sc, Sr, V, Ba	Mn, Cr, As, Hg, Na, K	Pb, Zn, Cd

Based on AHC results for dry season, it can be decided that the metal elements of Cu, Hg, Mn, Pb, and Zn in soil derived from natural sources, whereas Al, As, Ba, Ca, Cd, Co, Fe, K, Na, Ni, Sb, Sc, Sr, Ti and V in soil were also derived from anthropogenic activities and Cr also derived from both natural soil parent materials and anthropogenic sources. In context of rainy season, the AHC results demonstrated that metal elements of As, Cr, Hg, K, Mn and Na in soil were derived from different sources as the location of these metal elements were distinct from other metal elements of Al, Ba, Ca, Co, Cu, Fe, Ni, Sb, Sc, Sr, Ti and V. Moreover, it can be estimated as the metal elements of Al, Ba, Ca, Co, Cu, Fe, Ni, Sb, Sc, Sr, Ti and V were derived from anthropogenic activities and As, Cr, Hg, K, Na and Mn were derived from natural soil parent materials. Other metal elements of Cd, Pb and Zn were derived from both natural parent materials and anthropogenic sources.

#### 4.4 Cluster Analysis: Artificial Neural Network

Clustering is the process of training of a neural network on patterns so that the network comes up with its own classifications according to patterns similarity and relative topology (Kohonen, 2001). This analysis is used to identify the possible sources of contamination of metal elements by different color patterns. This analysis is also useful for gaining insight into data, or simplifying it before further processing. Pattern recognition is the process of training a neural network to assign the correct target classes to set of input patterns. In this analysis, self-organizing map (SOM) was used to represent the cooperation, competition

and updating mechanisms in accomplishing unsupervised learning, with the goal of reducing the dimensions found in any data set. In this study, the connection patterns of metal elements were presented in soil of waste disposal site. The weight planes for each element of the input vector that connect each input to each of the neurons were visualized. Darker colors represented highly contamination with high concentration and the correlations between the metal elements were very strong (Kohonen, 2014).

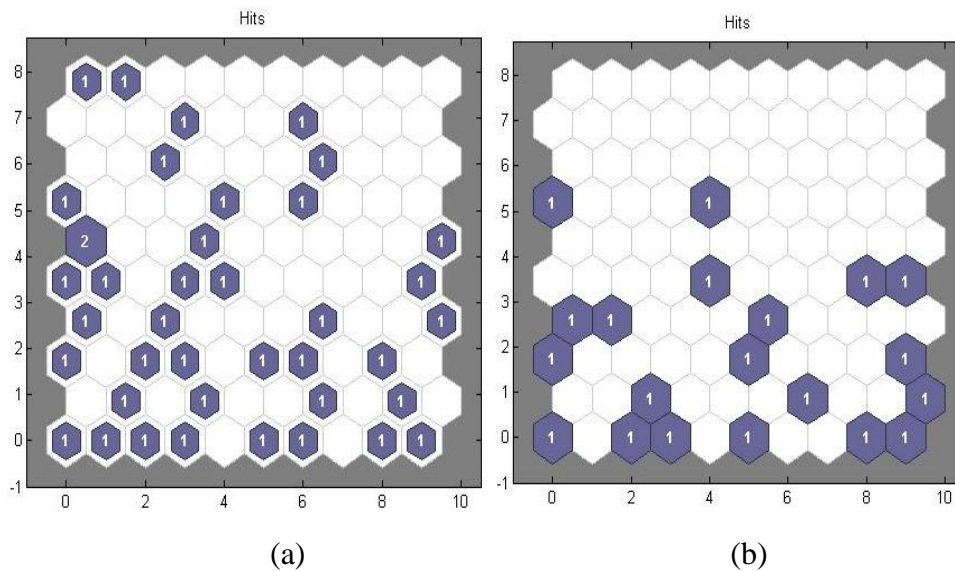


Figure 4.11: Plot of SOM sample hits for metal elements in (a) dry season (b) rainy season.

Figure 4.11 showed the neuron locations in the topology which indicated how many of the training data were associated with each of the neurons (cluster centers). The topology was 10-by-10 grid, so there were 100 neurons by default. The maximum number of hits associated with any neuron was 2 in case of dry season. Thus, there were 2 input vectors in that cluster.

Plot of SOM weight plane for metal elements of Cd, Ni, Pb and Zn in dry season is shown in Figure 4.12 (a) to Figure 4.12(d), respectively. The weight plane for each element of the input vector that connect each input to each of the neurons were visualized in the Annex-B except Cd, Ni, Pb and Zn. In this analysis, darker colors represented higher contamination with high concentrations of metal elements. Connection patterns of the concentration of all studied metal elements were almost similar, which interpreted these metal elements were highly correlated (Figure 4.12). The connection patterns of all metal elements were very similar as these metal elements were highly correlated. In case of dry season, K, Ca, Al,

Ti, Sb, Sc, Sr, V and Ba the connection pattern were very similar as these parameters were highly correlated. This statement was also verified from Pearson's correlation as correlation coefficient for these metal elements in dry season were very high ( $r > 0.90$ ).

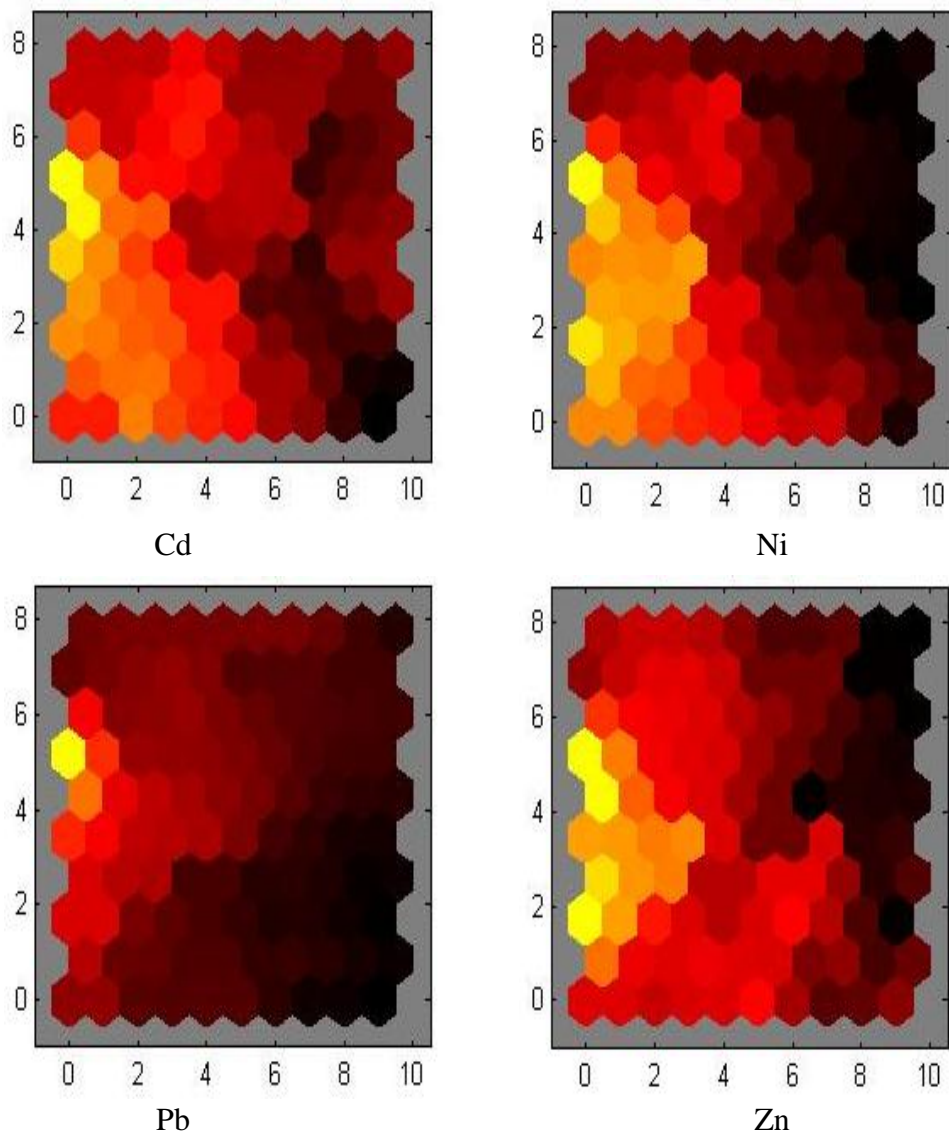


Figure 4.12: Plot of SOM weight plane for metal elements of (a) Cd, (b) Ni, (c) Pb and (d) Zn in dry season.

Figure 4.13 showed the connection pattern for Cd, Ni, Pb and Zn in the rainy season respectively. Rest of the cluster patterns were reported in Figure B.35 and Figure B.36 in the Annex- B for both the dry and rainy season, respectively. It was found from Figure B.35 and Figure B.36, the connection pattern for metal elements such as Fe, Mn, Cr, Cd, As, Hg, Co, Na and K, which were very similar. On the other hand, the cluster pattern for Ca, Al, Ti was almost same, whereas the pattern of Sb, Sc, Sr, V and Ba were similar,

which represent the correlation between metal elements present in soil of waste disposal site.

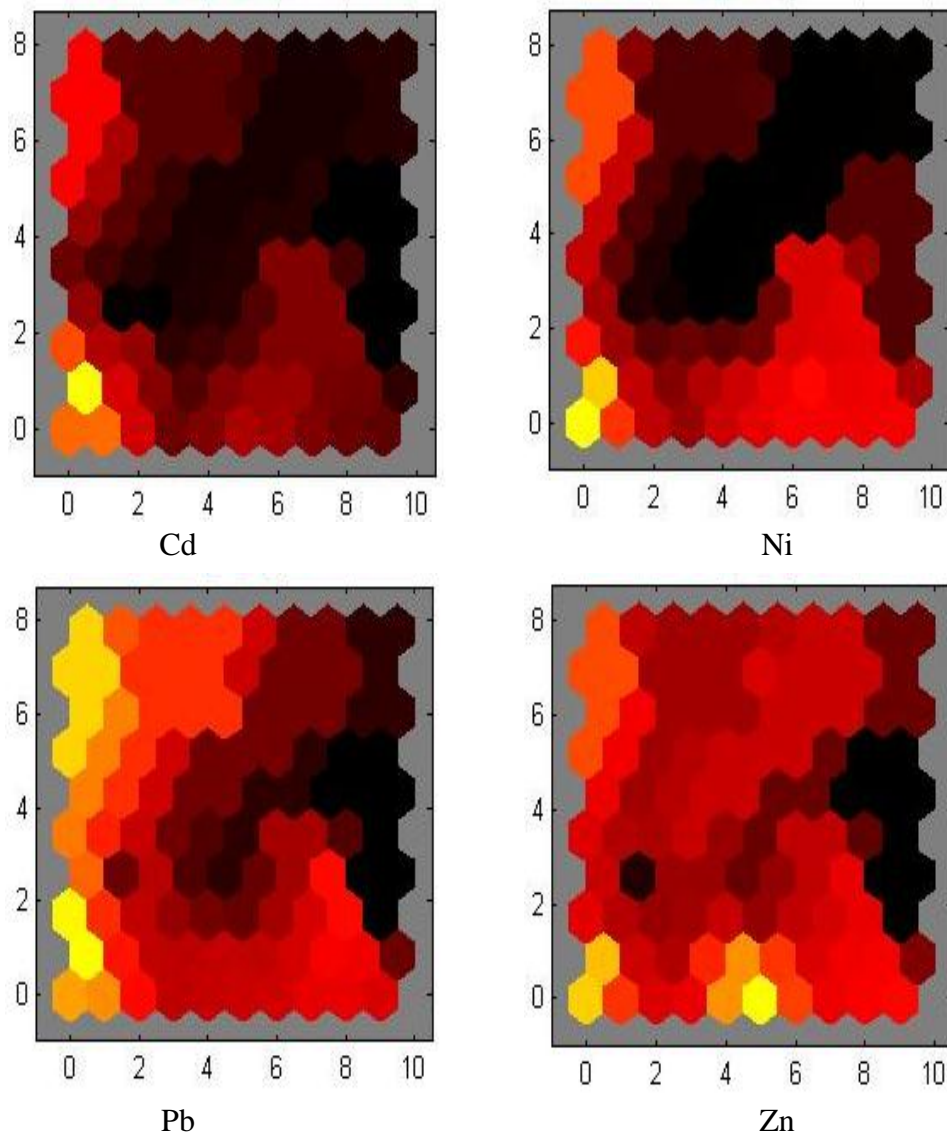


Figure 4.13: Plot of SOM Weight Plane for metal elements of (a) Cd, (b) Ni, (c) Pb and (d) Zn, in rainy season.

#### 4.5 Geostatistical Analysis

In this study, Geostatistical analyses including ArcGIS and Artificial Neural Network (ANN) were considered. Different interpolation methods of ArcGIS showed spatial distribution of metal elements in soil of waste disposal site with prediction errors. Moreover, ANN was also used to predict values considering standard error and hence discussed in the following articles.

### **4.5.1 ArcGIS Approach**

In this study, geostatistical analysis such as ordinary kriging (OK), inverse distance weighting (IDW), local polynomial interpolation (LP) and radial basis function (RBF) were performed using ArcGIS to distribute the metal elements in soil spatially. In addition, semivariogram parameters obtained from ordinary kriging exhibited spatial dependence of metal elements in soil and hence discussed in the following articles.

### **4.5.2 Deterministic Methods**

Deterministic techniques have parameters that control either (1) the extent of similarity (for example, IDW) of the values or (2) the degree of smoothing (for example, RBF). These techniques are not based on a random spatial process model. In this study, deterministic methods included IDW, LPI and RBFs were used for distributing metal elements spatially in soil. In this analysis, various prediction errors such as mean prediction error (MPE) and root mean square prediction error (RMSPE), mean standardized prediction error (MSPE), root mean square standard prediction error (RMSSPE) and average standard prediction error (ASPE) were considered.

#### ***4.5.2.1 Inverse Distance Weighting***

In case of IDW analysis, several powers of 1 to 5 were used to identify predicted surface with least prediction errors. To find out the best interpolation technique of IDW, cross validation of IDW with power 1 to 5 was performed for each metal elements. The result of cross validation for metal elements were reported in Annex-B, except Cd, Ni, Pb and Zn which is explained in details in this section. In the produced predicted surface from IDW interpolation, green with yellow color region represented the level of less contamination whereas orange, orange-red, ocher and red color area represented the level of highly contaminated soil (Yasrebi et al., 2009).

### ***Cross Validation of Cadmium***

To find out the best interpolation technique for IDW, cross validation of IDW with power 1 to 5 was performed for each metal elements and the results of cross validation for Cd is provided in Table 4.11. Figure. 4.14 showed the spatial distribution of Cd for IDW with power 1 to 5. Based on Figure 4.14, it was noticed that IDW with power 1 (IDW1) with highest RMSPE value (1.445) indicated the highest prediction error. Furthermore, in the produced prediction surface, there was a slighter greenish and yellow color region i.e. most of the region covered by red, ocher, orange-red and orange color with Cd concentration ranges approximately from 3.97 to 7.03 mg/kg indicated the soil of the study area was highly contaminated by Cd. In addition, IDW2 showed RMSPE (1.359), IDW3 of RMSPE (1.324) and IDW4 of RMSPE (1.318). The spatial distribution pattern with IDW with power 2 to 5 were almost same, except exhibited comparatively little greenish, little more greenish and more greenish color region, respectively (Figure 4.14).

Table 4.11: Cross validation of IDW for Cd

Metal element	Power	<sup>a</sup> MPE	<sup>b</sup> RMSPE
Cd	1	0.675	1.445
	2	0.594	1.359
	3	0.495	1.324
	4	0.412	1.318
	5	0.349	1.306

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error

In addition, IDW with power 5 (IDW5) exhibited maximum greenish color region and showed the lowest RMSPE value of 1.324 indicated the lowest prediction error in case of Cd. The produced prediction surface, comparatively more greenish color region was observed with Cd concentration ranges approximately from 1.2 to 3.97 mg/kg indicated the soil of the study area was found to be less contaminated in comparison of IDW1 to IDW4. A study conducted by Brovelli et al. (2011) and stated the IDW1 showed comparatively the severe contamination of Cd in soil, whereas for IDW5 was comparatively less contamination. So, the findings of this study are well agreed by the other researcher. From Figure 4.14, it was concluded that IDW5 showed better field condition. From this study, it was found that the contamination hotspots were near the center of the waste disposal site for metal element of Cd.

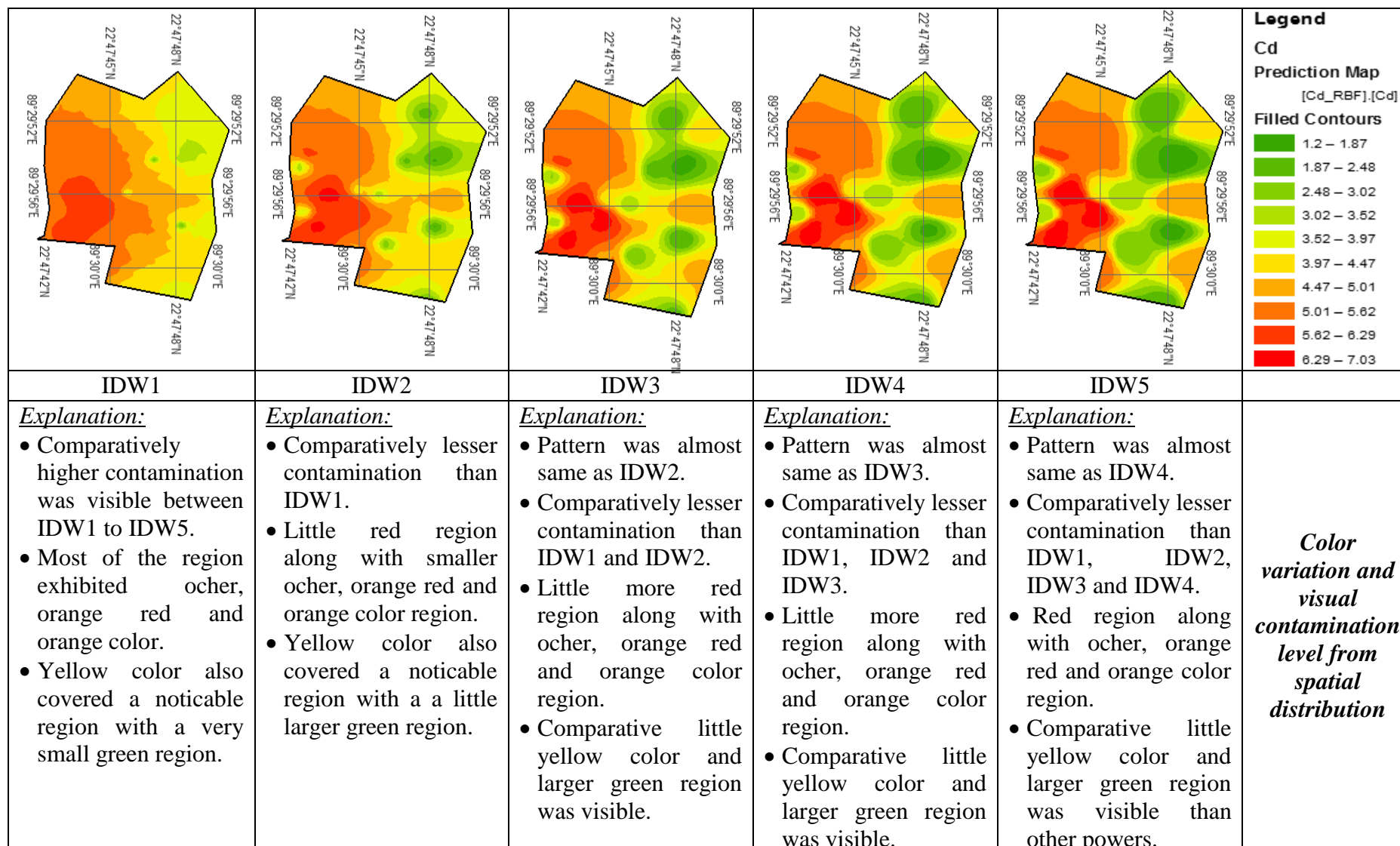


Figure 4.14: Spatial distribution of Cd in soil using IDW with power of 1- 5.

### ***Cross Validation of Nickel, Lead and Zinc***

The result of cross validation for IDW with power 1-5 for Ni, Pb and Zn was listed in Table 4.12. Table 4.12 reveals IDW1 with highest RMSPE value of 1.477 indicated the highest prediction error. In this case, most of the region covered by red and deep orange color which indicated the soil of the study area was highly contaminated for Ni. Moreover, for IDW5, the produced prediction surface showed comparatively more greenish color region which indicated the soil of the study area was found to be less contaminated with respect to IDW1 to IDW4 (Figure 4.15).

Table 4.12: Cross validation of IDW with different power for Ni, Pb and Zn

Metal Elements	Power	<sup>a</sup> MPE	<sup>b</sup> RMSPE
Ni	1	0.766	1.477
	2	0.674	1.389
	3	0.573	1.354
	4	0.491	1.340
	5	0.427	1.339
Pb	1	8.053	15.618
	2	6.667	14.555
	3	5.230	13.967
	4	4.180	13.558
	5	3.458	13.263
Zn	1	3.994	9.577
	2	3.670	9.095
	3	3.120	8.867
	4	2.608	8.742
	5	2.209	8.693

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error

This study (Table 4.12) also reveals the controlling parameters define prediction of variability of spatial distribution of Pb and Zn in soil of study area with IDW of power 1-5. IDW1 with highest RMSPE (15.618) indicated the highest prediction error. Moreover, for IDW5, the RMSPE value was found to be 13.263. Thus, the produced prediction surface showed comparatively more greenish color region indicated the soil of the study area was found to be less contaminated in comparison of IDW of power 1 to 4.



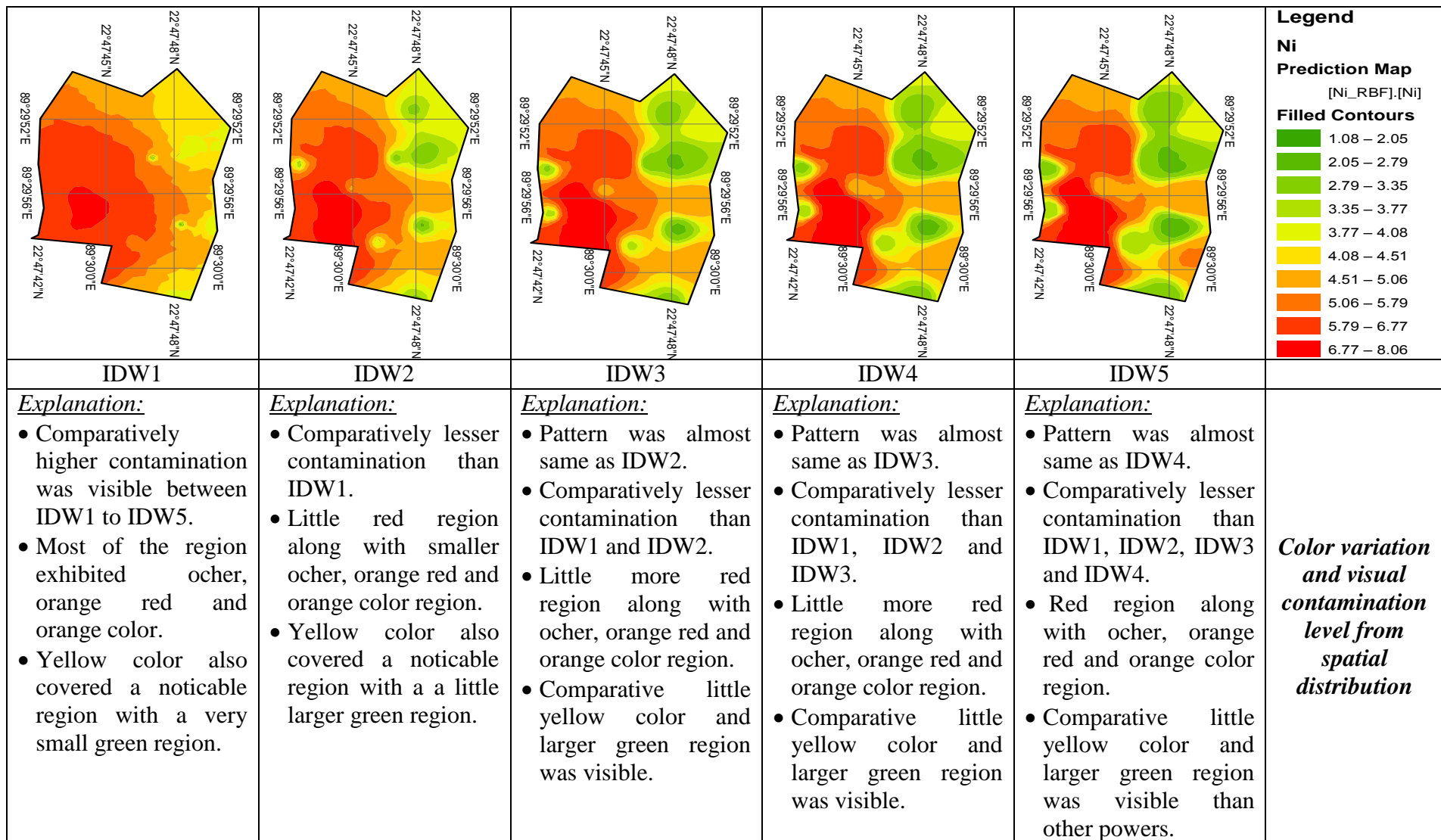


Figure 4.15: Spatial distribution of Ni in soil using IDW with power of 1- 5.

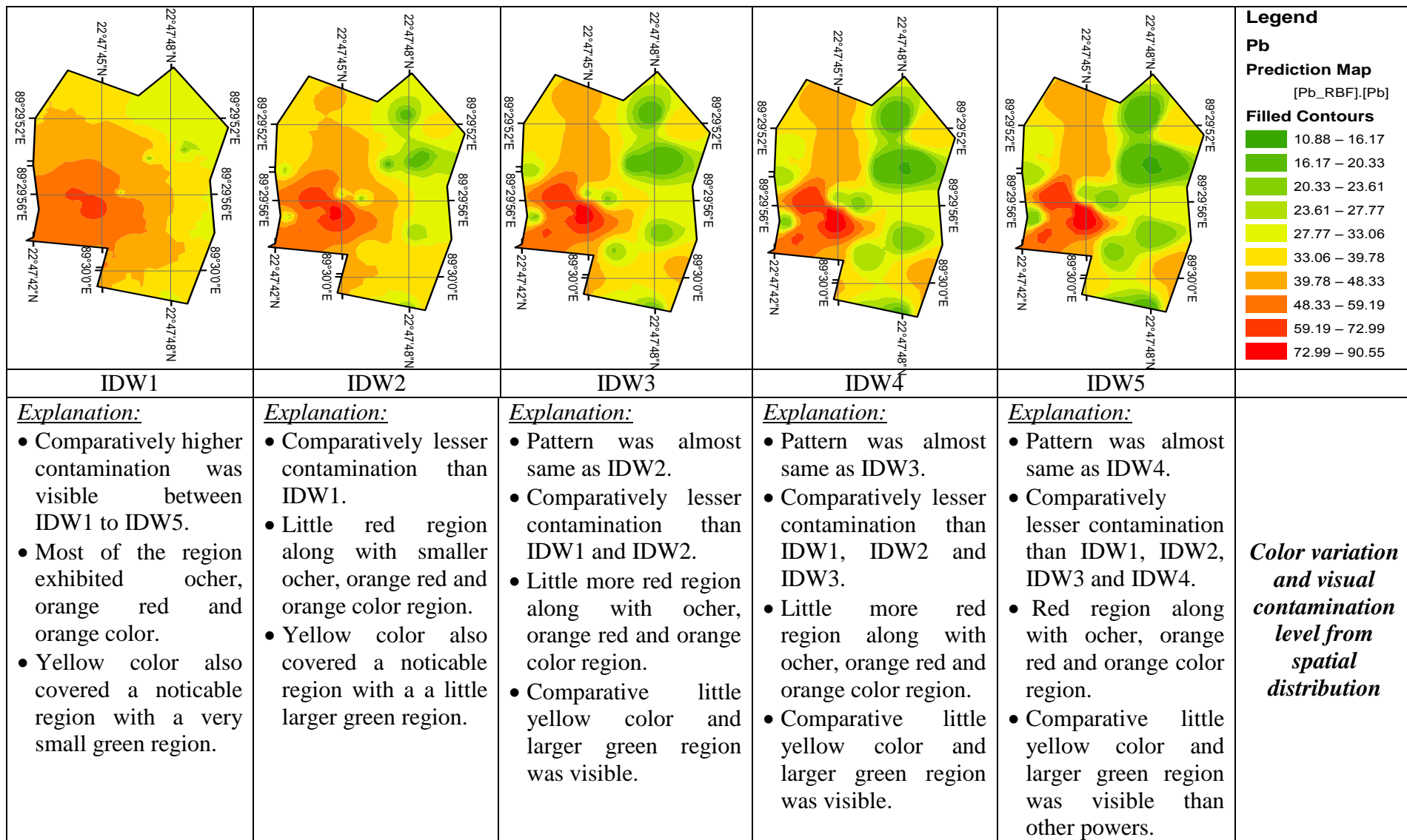


Figure 4.16: Spatial distribution of Pb in soil using IDW with power of 1- 5.

Table 4.12 also exposes IDW with power 1 with highest RMSPE value of 9.577 which indicated the highest prediction error. In this case, most of the region covered by red color with a little light green region which indicated the soil of the study area was highly contaminated for Zn. Moreover, for IDW with power 5, the produced prediction surface showed comparatively more greenish color region which indicated the soil of the study area was found to be less contaminated in comparison of IDW of power 1 to 4 (Figure 4.17).

Figure 4.15, Figure 4.16 and Figure 4.17 displayed IDW5 deliver better field condition for metal elements of Ni, Pb and Zn, respectively in soil of waste disposal site. It was comprehended from the analysis that the contamination hotspots were near the center of the selected disposal site for metal element of Ni, Pb and Zn.

In addition, the cross validation of IDW with power 1 to 5 was performed and the results are provided in Table C.1 to Table C.17 as well as the spatial distribution for metal elements of Al, As, Ba, Ca, Co, Cr, Cu, Fe, Hg, K, Mn, Na, Sb, Sc, Sr, Ti, and V in soil of the study area is depicted in Figure C. 1 to Figure C.17 in Annex-C. Results reveals that in produced prediction surface for all these metal elements,, greenish color region represented the level of less contamination and redish color area represented the level of highly contaminated soil. Almost all the metal elements exhibited severe contamination for IDW1; whereas comparatively less contaminated region was obtained for IDW5 interpolation techniques.

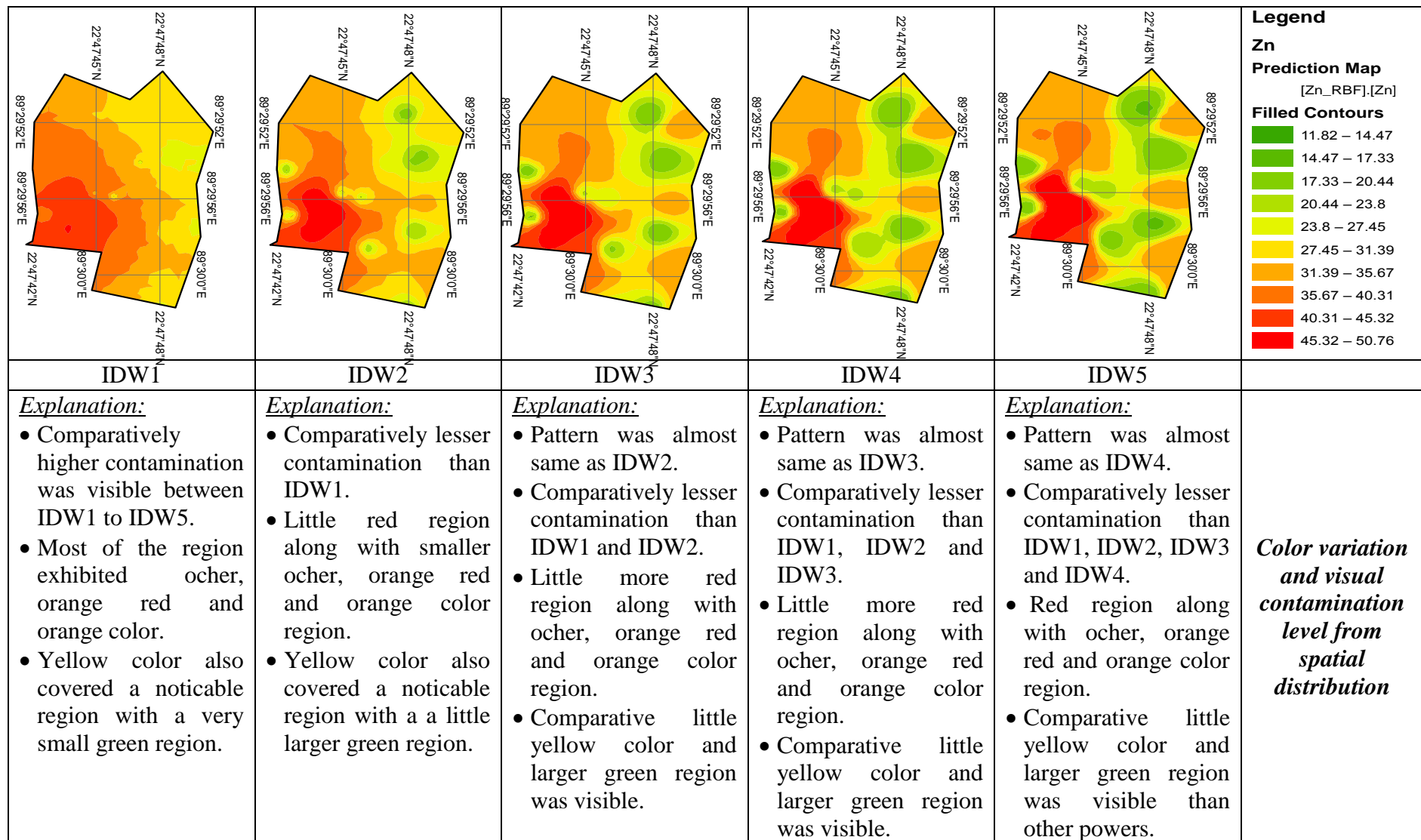


Figure 4.17: Spatial distribution of Zn in soil using IDW with power of 1- 5.

#### ***4.5.2.2 Local Polynomial Interpolation***

Local polynomial (LP) is a moderately quick deterministic interpolator technique. It is more flexible for analysis than that of global polynomial method. In this analysis, cross validation of LP was performed with six kernel function of Exponential, Polynomial, Gaussian, Epanechnikov, Quartic, Constant for order 1 to 3 for each metal elements and hence discussed in the following sections. Interpolation quality may be judged by examining RMSPE of the cross validation. A researcher Brovelli et al. (2011) stated that when RMSPE showed the lowest, the interpolation technique was considered as best fitted model. The cross validation result of metal elements were reported in the Annex-D, except for the metal elements such as Cd, Ni, Pb and Zn which are provided in this section. A study conducted by Li et al. (2013) and stated the interpolation techniques showed red color with high concentration of metal elements indicated the contamination hotspots, while green with yellow color indicated comparatively less contamination of soil by metal elements. In this study, from produced predicted surface by LP interpolation with different order of 1 to 3, it was comprehended that green and yellow color area represented less contamination and orange, orange red, ochre and red color area represented high contamination.

#### ***Cross Validation of Cadmium***

The results of cross validation of LP with order 1 to 3 for different kernel function of Cd are provided in Table 4.13. Based on the interpolation analysis, it was observed LP with order 1 (LP1) for gaussian function showed comparatively the lower value of RMSPE (1.359) than that of other kernel functions. Based on this finding, gaussian function was chosen as the best fitted model for Al. In addition, for second order polynomial (LP2), exponential function showed comparatively the lower of RMSPE with 1.428, and exponential function was chosen as the best fitted model. Finally, LP with order 3 (LP3) for exponential function showed the lowest value of RMSPE (1.436) and this function was chosen as the best fitted model.

Figure 4.18 showed the spatial distribution of Cd for the best fitted model of LP with order 1 to 3. The produced prediction surface area for LP2, there was a small yellow greenish color region i.e. most of the region covered by ochre and orange red color. with Cd concentration approximately ranges from 5.01 to 6.29 mg/kg, indicated the soil of the study area was highly contaminated by Cd (Figure 4.18).

Table 4.13: Cross validation of LP for Cd

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.353	0.282	0.284	0.265	0.282	0.198
<sup>b</sup> RMSPE	1.377	1.371	1.360	1.371	1.369	1.650
<sup>c</sup> MSPE	0.084	0.028	0.041	0.013	0.028	0.049
<sup>d</sup> RMSSPE	0.906	0.955	0.939	0.963	0.955	0.950
<sup>e</sup> ASPE	1.525	1.469	1.472	1.462	1.466	2.395
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.450	0.466	0.455	0.476	0.461	0.584
<sup>b</sup> RMSPE	1.428	1.474	1.461	1.481	1.474	1.526
<sup>c</sup> MSPE	0.174	0.174	0.170	0.180	0.171	0.239
<sup>d</sup> RMSSPE	0.914	0.956	0.947	0.958	0.956	0.963
<sup>e</sup> ASPE	1.542	1.513	1.519	1.521	1.512	1.571
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.342	0.411	0.350	0.392	0.407	0.405
<sup>b</sup> RMSPE	1.436	1.584	1.449	1.450	1.516	1.458
<sup>c</sup> MSPE	0.124	0.104	0.123	0.145	0.140	0.159
<sup>d</sup> RMSSPE	0.933	0.936	0.953	0.956	0.958	0.960
<sup>e</sup> ASPE	1.584	3.092	1.567	1.590	1.627	1.597

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

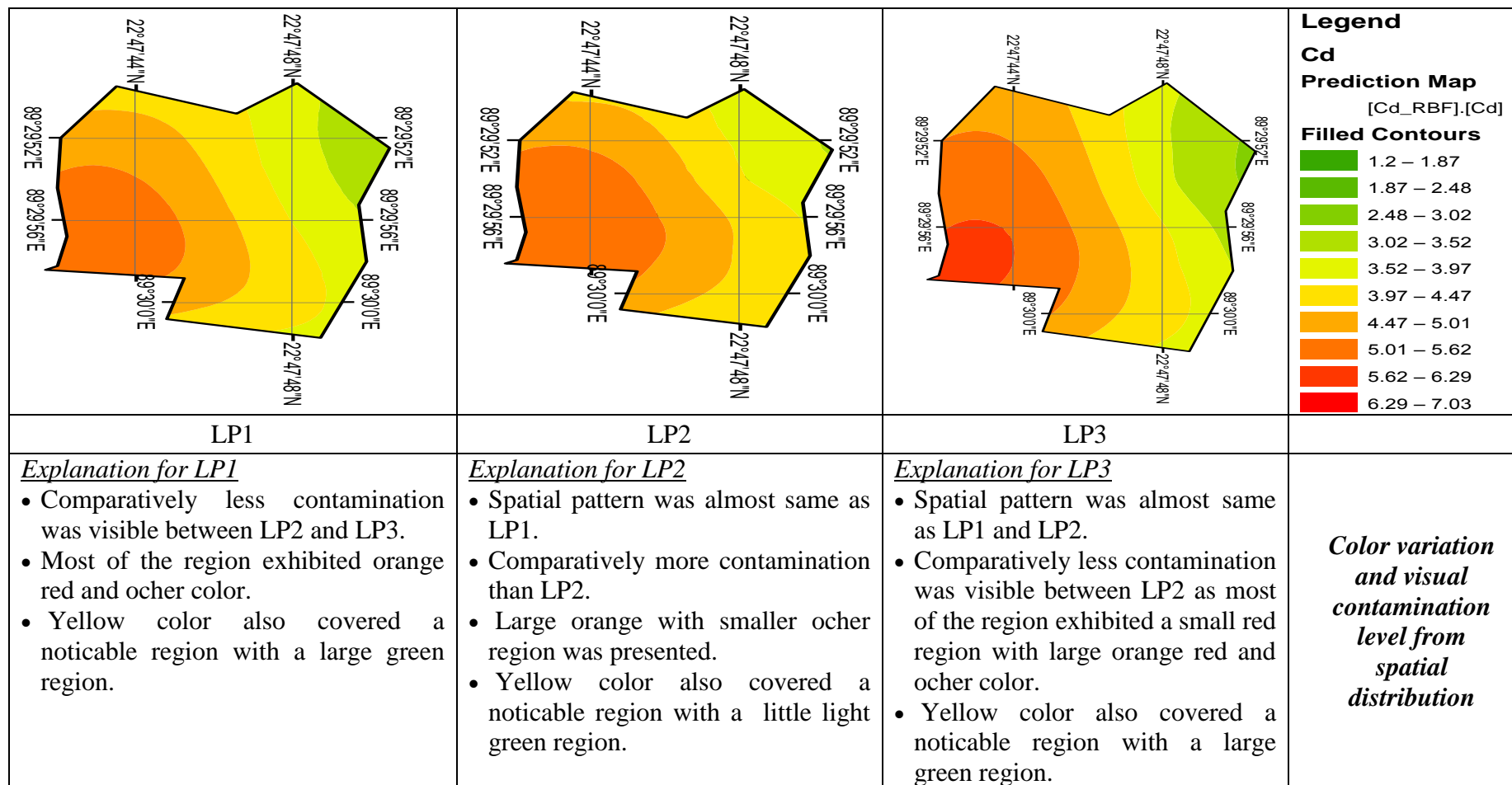


Figure 4.18: Spatial distribution of Cd in soil using LP with order of 1- 3.

Table 4.14: Cross validation of LP for Ni

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.207	0.120	0.179	0.098	0.115	0.214
<sup>b</sup> RMSPE	1.385	1.339	1.406	1.353	1.343	1.577
<sup>c</sup> MSPE	0.019	-0.065	-0.011	-0.085	-0.066	-0.047
<sup>d</sup> RMSSPE	0.903	0.945	0.931	0.954	0.944	0.944
<sup>e</sup> ASPE	1.783	1.651	1.787	1.670	1.673	3.894
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.419	0.466	0.406	0.504	0.383	0.394
<sup>b</sup> RMSPE	1.433	1.511	1.484	1.502	1.513	1.519
<sup>c</sup> MSPE	0.163	0.208	0.147	0.228	0.133	0.128
<sup>d</sup> RMSSPE	0.933	0.939	0.961	0.940	0.968	0.977
<sup>e</sup> ASPE	1.561	1.978	1.591	1.955	1.611	1.618
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.328	0.282	0.291	0.244	0.303	0.263
<sup>b</sup> RMSPE	1.466	1.432	1.438	1.445	1.625	1.729
<sup>c</sup> MSPE	0.077	0.0524	0.060	0.029	0.114	0.117
<sup>d</sup> RMSSPE	0.917	0.942	0.933	0.946	0.982	0.999
<sup>e</sup> ASPE	2.351	2.442	2.477	2.676	1.720	1.665

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.



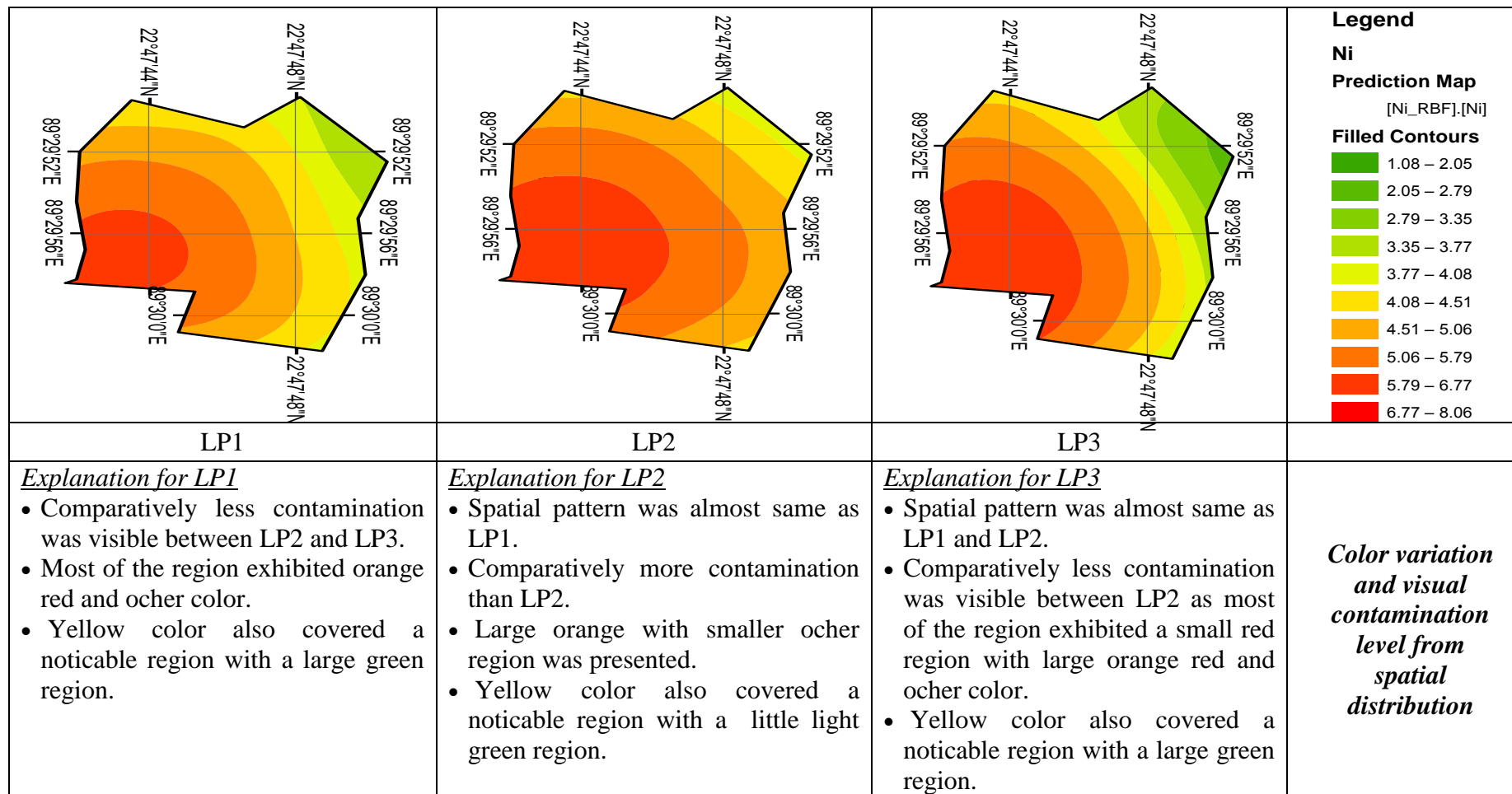


Figure 4.19: Spatial distribution of Ni in soil using LP with order of 1- 3.

In addition, LP1 exhibited maximum yellow with smaller area of greenish color with Cd concentration combindly ranges from 2.48 to 3.97 mg/kg indicated the soil of the study area less contaminated with respect to LP of order 1 and 2. Figure 4.18 displayed LP1 better field condition for metal elements of Cd respectively in soil of waste disposal site. It was comprehended from the analysis that the contamination hotspots were near the center of the selected dispoal site for metal element of Cd. A study conductd by Brovelli et al. (2011) and stated the LP2 showed comparatively the severe contamination of Al in soil, whereas for LP3 was comparatively less contamination. So, the findings of this study are well agreed with the other researchers in the similar cases of studies.

#### ***Cross Validation of Nickel, Lead and Zinc***

The result of cross validation for LP with order 1-3 for Ni, Pb and Zn was listed in Table 4.14, Table 4.15 and Table 4.16 respectively. Table 4.14 reveals polynomial kernal function for LP2 was showed the highest RMSPE value of 1.433 indicated the highest prediction error for Ni. In this case, most of the region covered by red and deep orange color indicated the soil of the study area was highly contaminated for Ni. Moreover, for LP1, the produced prediction surface by showed comparatively more greenish and yellow color region which indicated the soil of the study area was found to be less contaminated in comparison of LP2 and LP3 (Figure 4.19).

Figure 4.20 showed the spatial distribution of Pb for the best fitted model of LP with order 1 to 3. The produced prediction surface area for LP2, there was a small yellow greenish color region i.e. most of the region covered by orange color with Pb concentration approximately ranges from 27.77 to 59.19 mg/kg indicated the soil of the study area was moderately contaminated in case of Pb (Figure 4.20). In addition, LP3 exhibited maximum yellow with smaller area of greenish color with Pb concentration combindly ranges from 6.17 to 23.61 mg/kg which indicated the soil of the study area was found to be less contaminated in comparison of LP of order 1 and 2 (Table 4.15). From Table 4.16 for Zn, exponential kernal function of LP2 showed the highest RMSPE indicated the highest prediction error. In this case, most of the region covered by red and deep orange color indicated the soil of the study area was highly contaminated for Zn (Figure 4.21).

Table 4.15: Results of cross validation for LP with different order for Pb in soil

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	3.025	2.544	2.816	2.485	2.946	-0.938
<sup>b</sup> RMSPE	12.878	12.853	12.913	13.451	13.461	26.938
<sup>c</sup> MSPE	0.057	0.022	0.038	0.012	0.026	-0.148
<sup>d</sup> RMSSPE	0.842	0.908	0.883	0.943	0.917	0.916
<sup>e</sup> ASPE	20.721	19.730	20.180	19.418	21.767	55.672
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	5.287	5.505	5.407	4.571	5.562	4.656
<sup>b</sup> RMSPE	14.878	15.580	15.246	15.530	15.608	15.417
<sup>c</sup> MSPE	0.193	0.191	0.190	0.141	0.195	0.143
<sup>d</sup> RMSSPE	0.916	0.961	0.948	0.979	0.961	0.985
<sup>e</sup> ASPE	19.857	19.694	19.754	18.985	19.724	18.821
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	3.915	3.956	3.973	4.685	4.610	4.749
<sup>b</sup> RMSPE	14.936	15.616	15.459	15.784	16.029	15.960
<sup>c</sup> MSPE	0.125	0.123	0.121	0.151	0.141	0.164
<sup>d</sup> RMSSPE	0.917	0.954	0.947	0.956	0.953	0.963
<sup>e</sup> ASPE	20.987	20.284	20.784	21.137	21.238	21.518

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

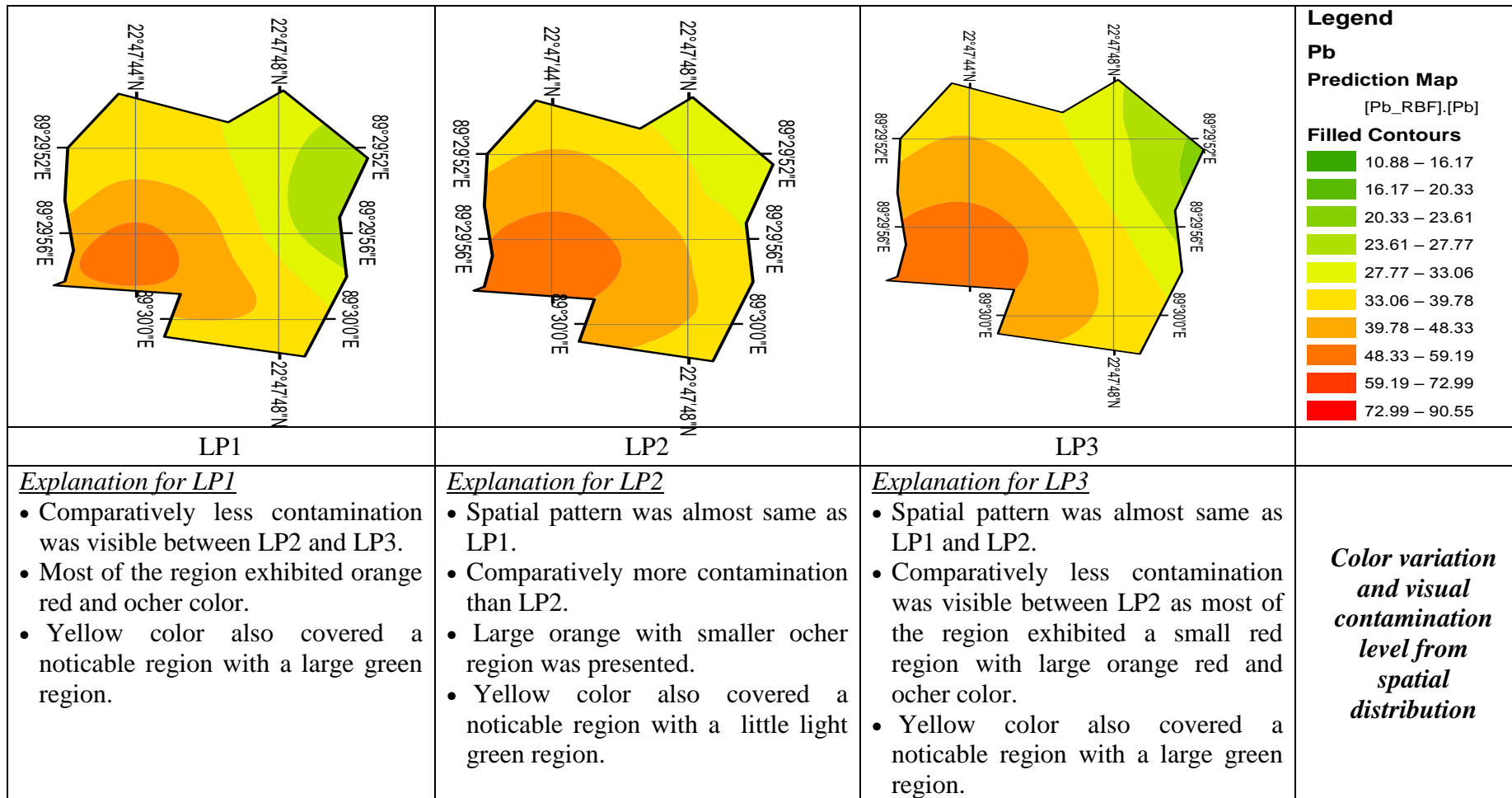


Figure 4.20: Spatial distribution of Pb in soil using LP with order of 1- 3.

Table 4.16: Cross validation of LP for Zn

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	2.421	1.668	2.111	1.631	1.681	1.035
<sup>b</sup> RMSPE	9.149	9.279	9.218	9.405	9.306	11.675
<sup>c</sup> MSPE	0.107	0.049	0.091	0.046	0.053	0.065
<sup>d</sup> RMSSPE	0.858	0.940	0.913	0.949	0.939	0.929
<sup>e</sup> ASPE	12.186	10.886	10.362	11.135	11.020	14.440
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	2.930	2.872	2.854	2.518	2.930	2.475
<sup>b</sup> RMSPE	9.958	10.114	9.980	10.066	9.958	10.095
<sup>c</sup> MSPE	0.192	0.180	0.178	0.127	0.192	0.120
<sup>d</sup> RMSSPE	0.894	0.931	0.921	0.973	0.894	0.978
<sup>e</sup> ASPE	11.338	13.188	13.087	10.427	11.338	10.435
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	2.106	1.924	2.041	1.786	2.170	2.103
<sup>b</sup> RMSPE	9.942	10.126	10.038	10.393	10.291	10.755
<sup>c</sup> MSPE	0.118	0.054	0.067	0.035	0.116	0.137
<sup>d</sup> RMSSPE	0.921	0.936	0.930	0.945	0.958	0.973
<sup>e</sup> ASPE	11.421	16.039	15.537	17.199	11.345	11.175

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

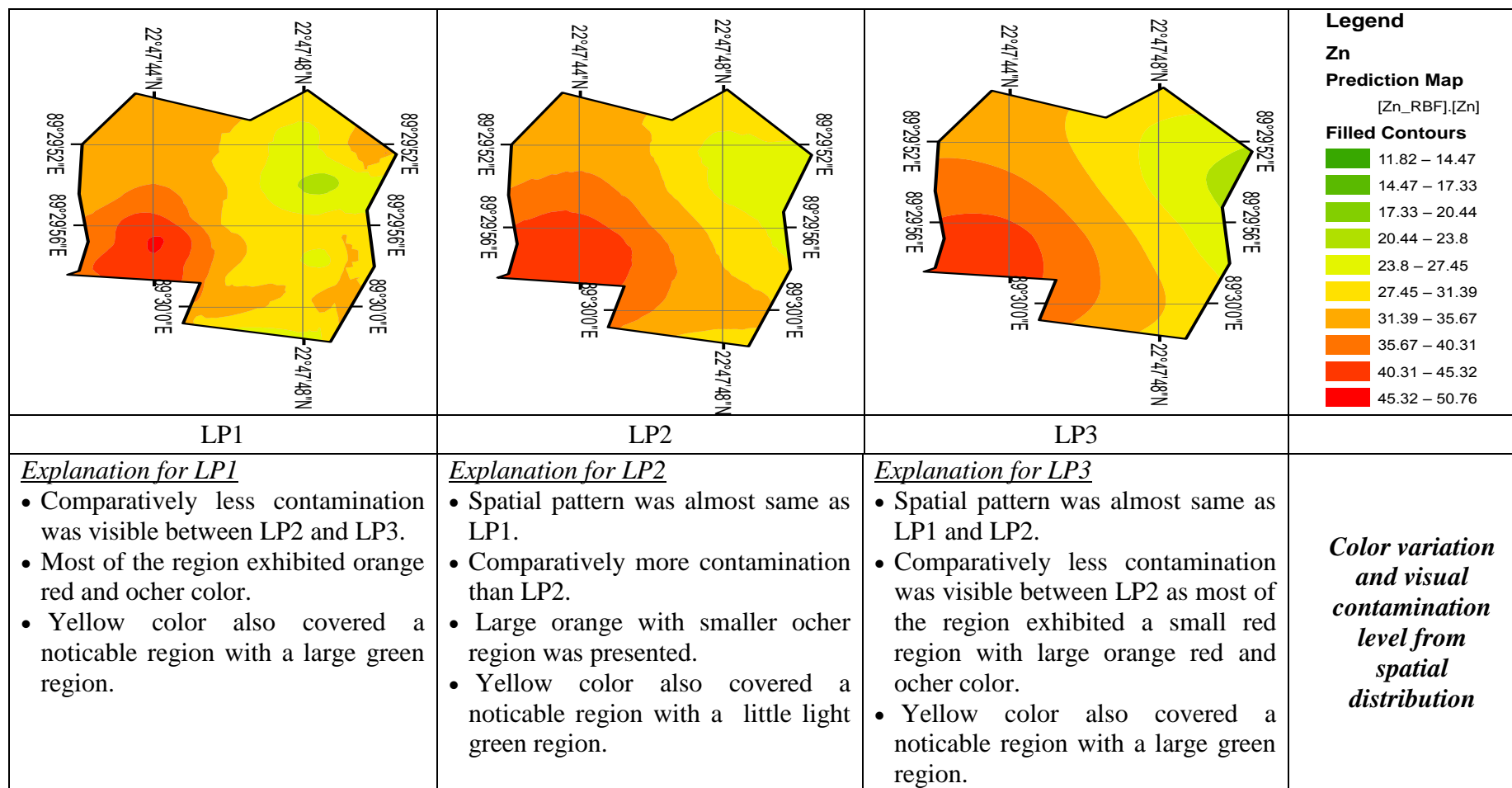


Figure 4.21: Spatial distribution of Zn in soil using LP with order of 1- 3.

Moreover, for LP3, it was found the highest RMSPE value of 14.936. The produced prediction surface by showed comparatively more greenish and yellow color region which indicated the soil of the study area was found to be less contaminated in comparison of LP1 and LP2 (Figure 4.21). Figure 4.19, Figure 4.20 and Figure 4.21 displayed LP1 provided better field condition for metal elements of Ni, Pb and Zn, respectively in soil of waste disposal site. It was comprehended from the analysis that the contamination hotspots were near the center of the selected disposal site for the metal element of Ni, Pb and Zn.

In addition, the cross validation of LP with order 1 to 3 was performed and the results are provided in Table D.1 to Table D.17 as well as the spatial distribution for metal elements of Al, As, Ba, Ca, Co, Cr, Cu, Fe, Hg, K, Mn, Na, Sb, Sc, Sr, Ti, and V in soil of the study area is depicted in Figure D. 1 to Figure D.17 in Annex-D. Results reveals that in produced prediction surface for all these metal elements,, greenish color region represented the level of less contamination and redish color area represented the level of highly contaminated soil. Almost all the metal elements exhibited severe contamination for LP2; whereas comparatively less contaminated region was obtained for LP3 interpolation techniques.

#### **4.5.2.3 Radial Basis Function**

Radial basis functions (RBF) is the name given to a large family of exact interpolators which use a basic equation dependent on the distance between the interpolated point and the sampling points (Brovelli et al., 2011). In this study, RBF analysis, cross validation of RBF with five kernal function of thin-plate spline (TPS), spline with tension (SPT), completely regularized spline (CRS), multi-quadratic function (MQ) and inverse multi-quadratic function (IMQ) was performed for each metal elements and hence discussed in the following sections. The cross validation result of metal elements were reported in the Annex-E, except for the metal elements such as Cd, Ni, Pb and Zn which is provided in this chapter. When kernal function showed higher value of the kernel parameter, it will provide a smoother surface except IMQ. When kernal function showed lower value of the kernel parameter, it will provide a smoother surface in case of IMQ (ESRI, 2001; Zhao et al., 2010; Yao et al., 2013). A researcher Brovelli et al. (2011) stated that when RMSPE showed the lowest value, the interpolation technique was considered as best fitted model.

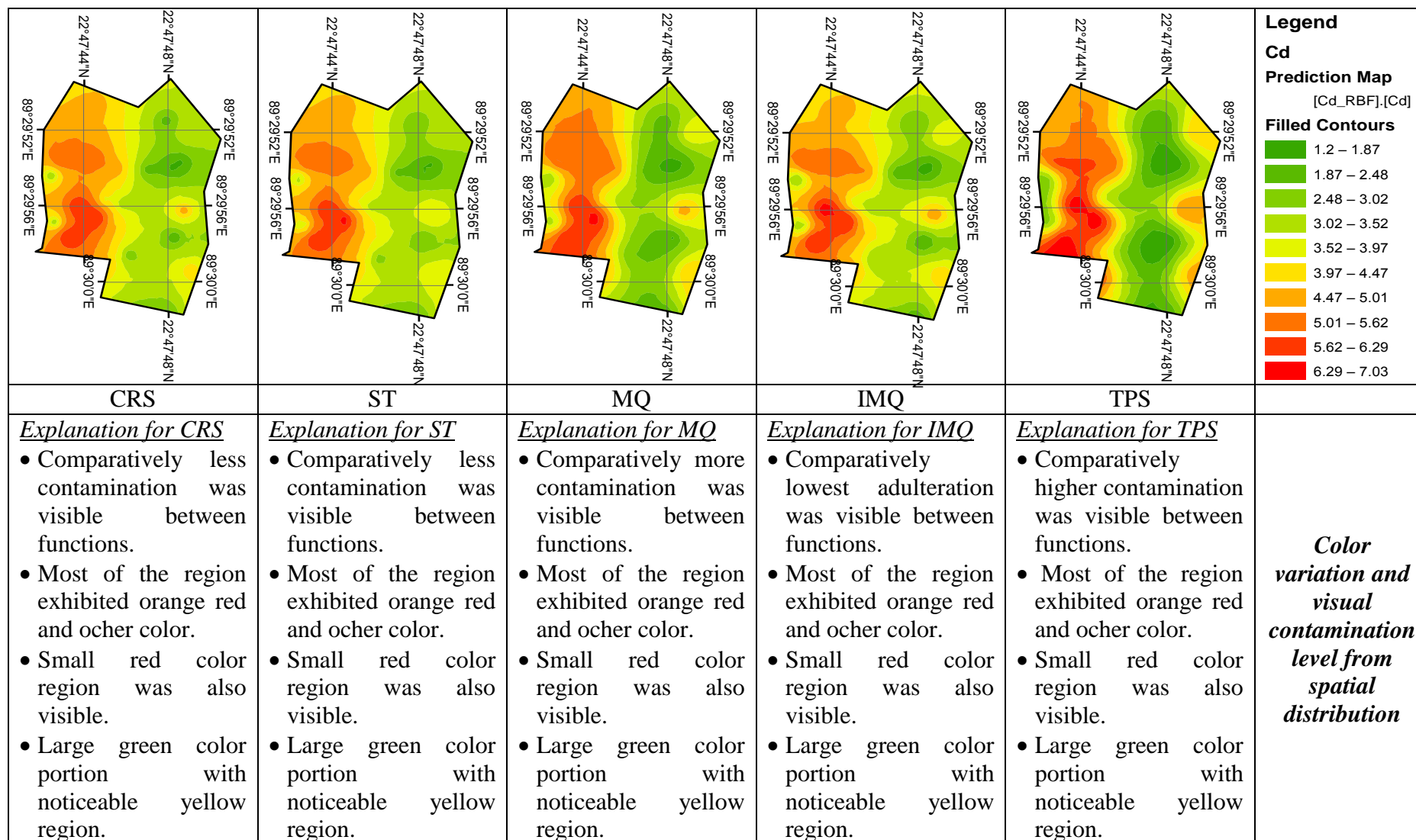


Figure 4.22: Spatial distribution of Cd in soil using RBF's.



### ***Cross Validation of Cadmium***

In RBF analysis, cross validation of RBF with five different kernel functions of CRS, ST, MQ, IMQ and TPS was performed for Cd to find the best fitted model for showing spatial distribution and provided in Table 4.17. Based on the interpolation analysis, it was observed RBF for IMQ showed comparatively the lower value of RMSPE (1.2054) than that of other kernel functions. Based on this finding, IMQ was chosen as the best fitted model for Cd. In case of Cd, the kernel parameter of IMQ showed 0.000121 indicated a very small value tends to zero thus it provided a smoother surface. In addition, the kernel function of TPS showed comparatively the highest value of kernel parameter and provided a smoother surface. Figure 4.18 showed the spatial distribution of Cd for distinct RBF's. The produced prediction surface area for TPS, there exhibited maximum greenish color region with Cd concentration approximately ranges from 1.87 to 3.52 mg/kg. In addition, RBF with TPS exhibited maximum red and orange color with Cd concentration combinedly ranges from 4.47 to 7.03 mg/kg, indicated the soil of the study area was highly contaminated by Cd (Figure 4.22).

Table 4.17: Results of cross validation of RBF with different kernel functions for Cd

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	0.076	0.075	0.068	0.115	0.122
<sup>b</sup> RMSPE	1.2239	1.2241	1.3372	1.2054	1.6159
Kernel Parameter	55181.22	43835.15	0	0.000121	1.00E+20

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error

### ***Cross Validation of Nickel, Lead and Zinc***

In case of Ni, Pb and Zn, the cross validation of RBF with five different kernel functions of CRS, ST, MQ, IMQ and TPS was performed to find the best fitted model and the results are provided in Table 4.18. The kernel function IMQ was chosen as best fitted model for RBF interpolation in case of Ni. It was found from the cross validation result, RMSPE (0.115) was comparatively lower than that of other kernel functions. The kernel parameter of IMQ showed 0.00013 indicated a very small value tends to zero thus it provided a smoother surface. The kernel function of TPS showed comparatively the highest value of kernel parameter and also provided a smoother surface.

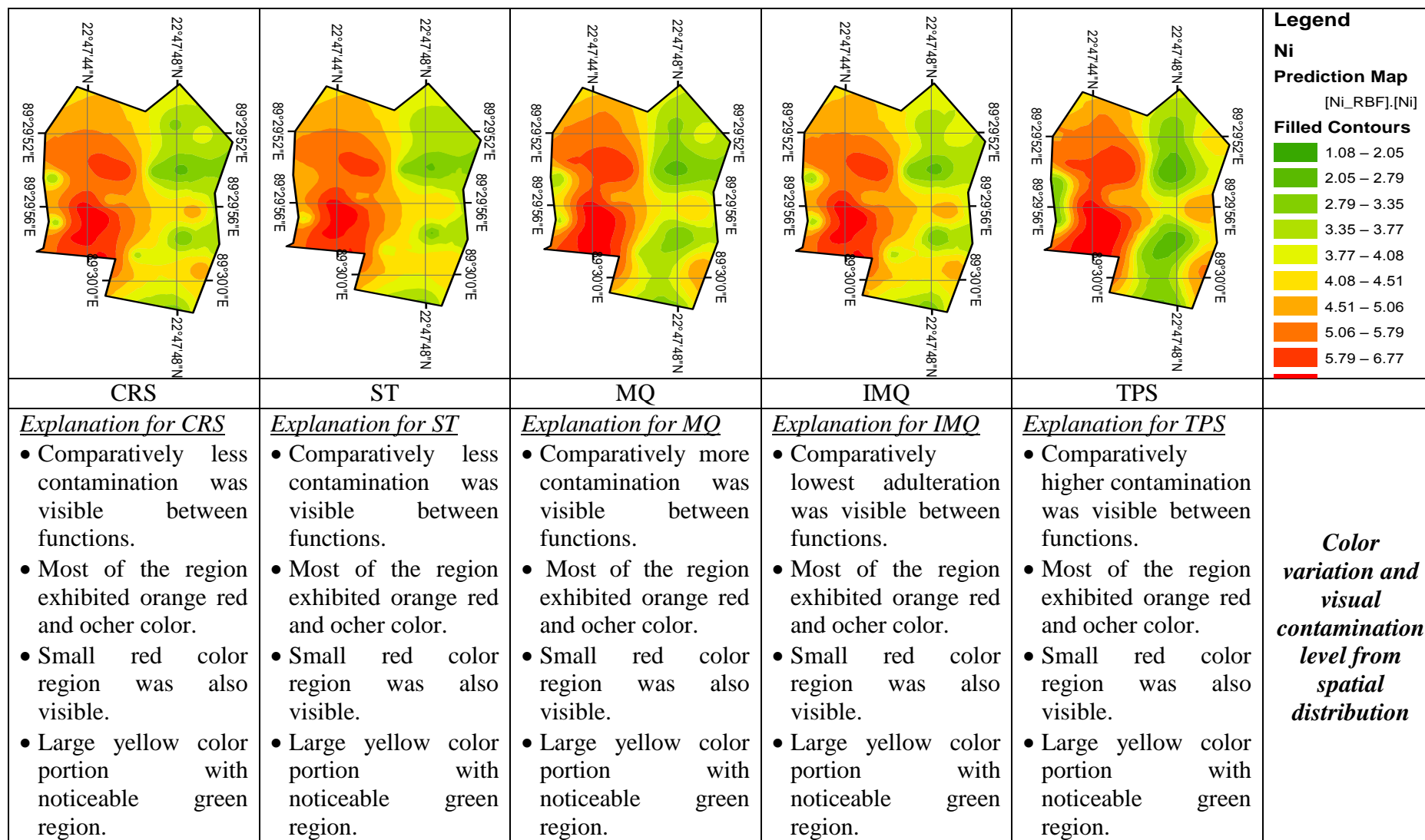


Figure 4.23: Spatial distribution of Ni in soil using RBF's.

Table 4.18: Cross validation of RBF for Ni,Pb and Zn

Metals	Model	<sup>a</sup> MPE	<sup>b</sup> RMSPE	Kernel Parameter
Ni	CRS	0.147	1.204	35188.1
	ST	0.144	1.206	35234.7
	MQ	0.097	1.294	0
	IMQ	0.211	1.197	0.00013
	TPS	0.161	1.586	1.00E+20
Pb	CRS	0.736	11.999	91048.6
	ST	0.738	11.999	71818.2
	MQ	0.477	12.830	0
	IMQ	1.187	12.139	0.00011
	TPS	0.891	15.644	1.00E+20
Zn	CRS	0.613	7.866	20297.9
	ST	0.536	7.915	18199.2
	MQ	0.502	8.336	0
	IMQ	0.713	7.745	0.00016
	TPS	0.849	9.500	1.00E+20

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error

IMQ was chosen as best fitted model for RBF interpolation in case of Ni. It was found from the cross validation result, RMSPE (1.197) was comparatively lower than that of other kernel functions. The kernel parameter of IMQ showed 0.00013 indicated a very small value tends to zero thus it provided a smoother surface. The kernel function of TPS showed comparatively the highest value of kernel parameter and provided a smoother surface. Figure 4.23 showed the spatial distribution of Ni for different RBF's. The produced prediction surface area for TPS, maximum greenish color region with Ni concentration approximately ranges from 1.08 to 4.08 mg/kg. In addition, RBF with TPS exhibited yellow color with concentration of 4.08 to 4.51 mg/kg. Maximum ochre, orange, orange red and red color with Ni concentration combinedly ranges from 4.08 to 8.06 mg/kg, which indicated the soil of the study area was moderately contaminated in case of Ni (Figure 4.23). From Table 4.18 for Pb, it was observed RBF for CRS showed comparatively the lower value of RMSPE (11.999) than that of other kernel functions. Thus, CRS was chosen as the best fitted model for Pb. In case of Pb, the kernel parameter of CRS showed 91048.62 indicated a large value indicated a smoother produced surface. The kernel function of TPS showed comparatively the highest value of kernel parameter and provided more smoother surface. Figure 4.24 showed the spatial distribution of Pb using distinct RBF's. The produced prediction surface area for TPS, there exhibited maximum greenish color region with Pb concentration approximately ranges from 10.88 to 33.06 mg/kg.

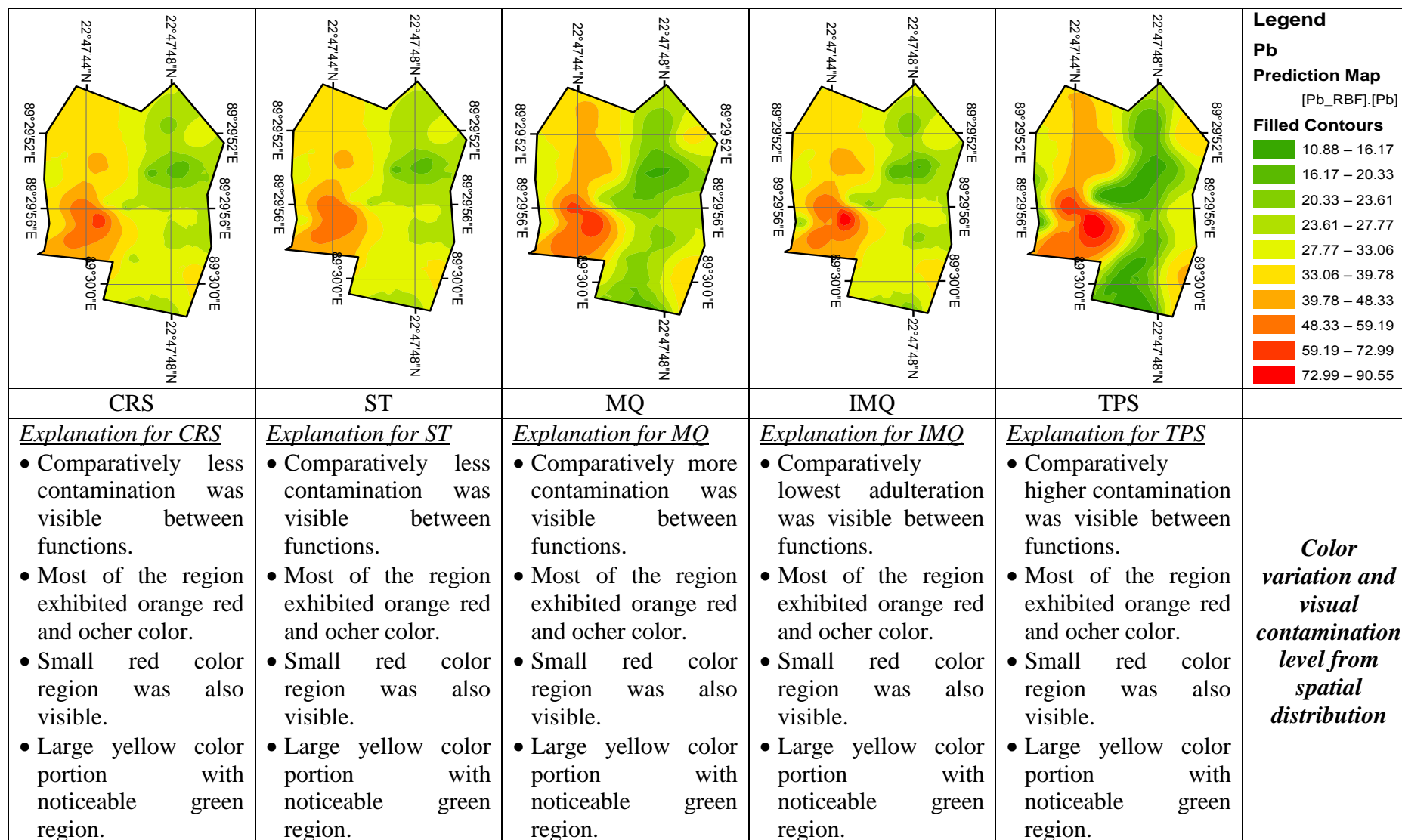


Figure 4.24: Spatial distribution of Pb in soil using RBF's.

In addition, RBF with TPS exhibited maximum ochre, orange, orange-red and red color with Pb concentration combinedly ranges from 39.78 to 90.55 mg/kg, which indicated the soil of the study area was moderately contaminated in case of Pb (Figure 4.24).

In addition, from the cross validation result of RBF, RMSPE (7.745) was comparatively lower for IMQ than that of other kernel functions. The kernel parameter of IMQ showed 0.00016 indicated a very small value tends to zero thus it provided a smoother surface. The TPS showed comparatively the highest value of kernel parameter and provided a smoother surface. The produced prediction surface area for TPS, maximum greenish color region with Zn concentration approximately ranges from 11.82 to 23.8 mg/kg (Figure 4.18). In addition, RBF with TPS also exhibited maximum red and orange color with Zn concentration combinedly ranges from 31.39 to 50.76 mg/kg, which indicated the soil of the study area was moderately contaminated in case of Zn (Figure 4.25). Figure 4.23, Figure 4.24 and Figure 4.25 displayed IMQ provided better field condition for metal elements of Ni, Pb and Zn, respectively in soil of waste disposal site. It was comprehended from the analysis that the contamination hotspots were near the center of the selected disposal site for metal element of Ni, Pb and Zn.

In addition, the cross validation of RBF with five distinct functions was performed and the results are provided in Table E.1 to Table E.17 as well as the spatial distribution for metal elements of Al, As, Ba, Ca, Co, Cr, Cu, Fe, Hg, K, Mn, Na, Sb, Sc, Sr, Ti, and V in soil of the study area is depicted in Figure E. 1 to Figure E.17 in Annex-E. Results reveals that in produced prediction surface for all these metal elements,, greenish color region represented the level of less contamination and redish color area represented the level of highly contaminated soil.

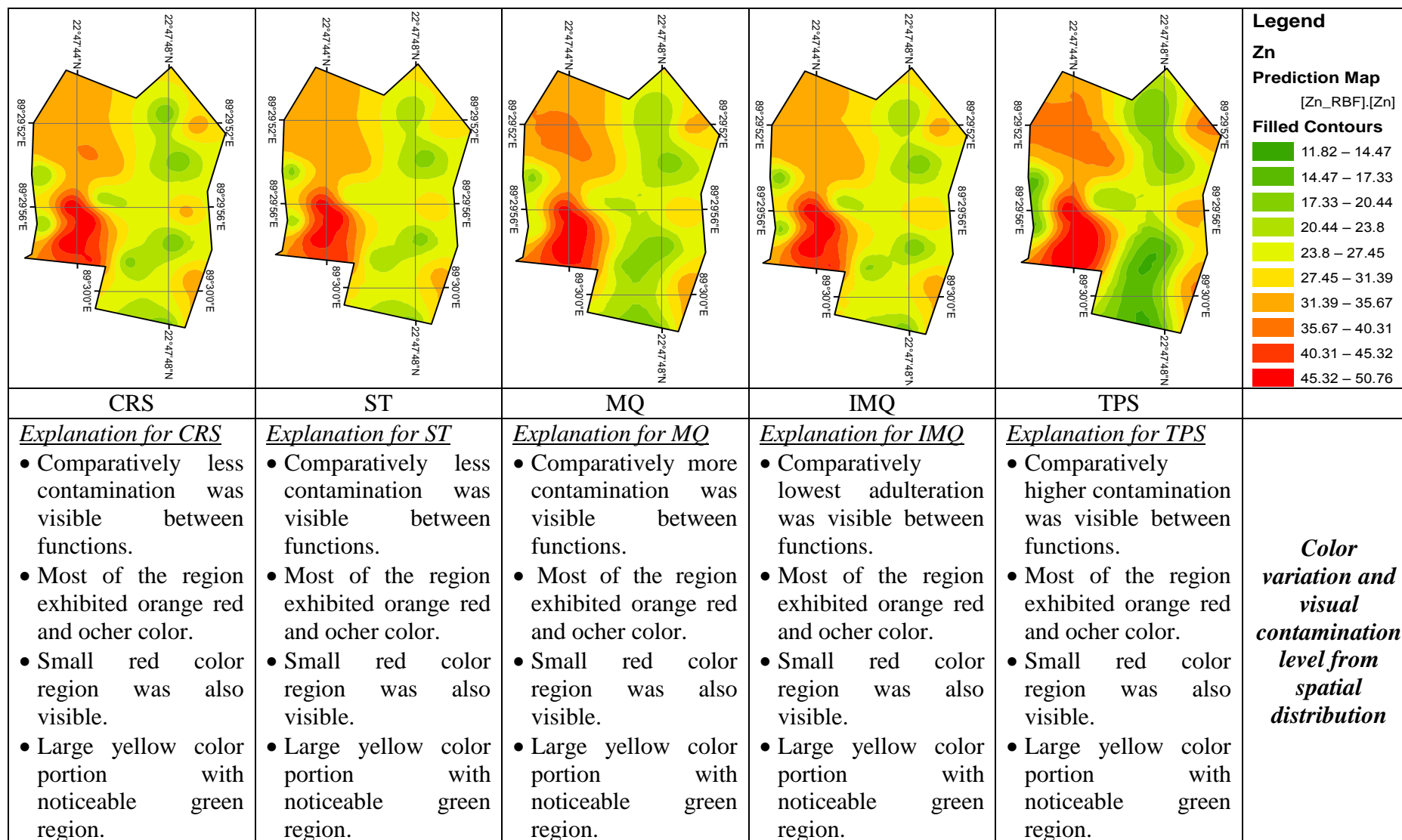


Figure 4.25: Spatial distribution of Zn in soil using RBF's.

### 4.5.3 Geostatistical Methods

Geostatistical techniques assume that at least some of the variation observed in natural phenomena can be modeled by random processes with spatial autocorrelation and require that the spatial autocorrelation be explicitly modeled. Geostatistical techniques can be used to describe and model spatial patterns (variography), predicts values at unmeasured locations (kriging), and assess the uncertainty associated with a predicted value at the unmeasured locations (kriging). The geostatistical wizard offers several types of kriging, which are suitable for different types of data and have different underlying assumption of Ordinary, Simple, Universal, Indicator, Probability, Disjunctive, Areal interpolation etc.. These mean standardized prediction error (MSPE), root mean square standard prediction error (RMSSPE) and average standard prediction error (ASPE) was used to select the best fitted models.

#### 4.5.3.1 Ordinary Kriging

The ordinary kriging (OK) with eleven distinct models such as Circular, Spherical, Tetraspherical, Pentaspherical, Exponential, Gaussian, Rational quadratic, Hole effect, K-Bessel, J-Bessel and Stable was performed to select the best fitted model for each metal elements in soil. This selected model was used for modeling the empirical semivariogram and spatial distribution of the concentration of metal elements in soil. A study conducted by Brovelli et al. (2011) and stated when MSPE closed to 0 and RMSSPE tends to 1 for a particular model, then the model will be choosen as the best fitted model. In addition, when ASPE closed to RMSPE for a specific model, then it can be confidently said that the prediction model was appropriate (Johnston et al., 2001; Zhao et al., 2010).

#### *Cross Validation of Cadmium*

The cross-validation result of eleven distinct models stated earlier using ordinary kriging for Cd is shown in Table 4.19. The values of MSPE ranges from 0.0081 to 0.0358 and RMSSPE from 0.8980 to 1.0139. Result reveals the value of MSPE was closest to zero (0.0358), RMSSPE closest to 1(1.0139) and the ASPE (1.3924) closed to RMSPE (1.2498) for the model of rational quadratic (Table 4.19). So, in this case, rational quadratic was selected as the best fitted model using ordinary kriging interpolation. The spatial distribution of Cd for this best fitted model shown in Figure. 4.26.

Table 4.19: Cross validation results of ordinary kriging for different models of Cd in soil

<b>Models</b>	<b><sup>a</sup>MPE</b>	<b><sup>b</sup>RMSPE</b>	<b><sup>c</sup>ASPE</b>	<b><sup>d</sup>MSPE</b>	<b><sup>e</sup>RMSSPE</b>
Circular	0.0244	1.2724	1.4398	0.0081	0.9097
Spherical	0.0361	1.2590	1.4399	0.0161	0.8994
Tetraspherical	0.0436	1.2520	1.4348	0.0199	0.8986
Pentaspheical	0.0442	1.2475	1.4141	0.0200	0.9148
Exponential	0.0597	1.2521	1.4073	0.0245	0.9244
Gaussian	0.0295	1.2638	1.4375	0.0107	0.9102
<b>Rational quadratic</b>	<b>0.0959</b>	<b>1.2498</b>	<b>1.3924</b>	<b>0.0358</b>	<b>1.0139</b>
Hole Effect	0.0524	1.2634	1.4322	0.0310	0.9219
K-Bessel	0.0550	1.2359	1.4111	0.0231	0.9043
J-Bessel	0.0220	<b>1.2359</b>	1.4284	0.0087	0.8980
Stable	0.0630	1.2322	1.4074	0.0275	0.9036

Table 4.20: Cross validation results of ordinary kriging for different models of Ni in soil

<b>Models</b>	<b><sup>a</sup>MPE</b>	<b><sup>b</sup>RMSPE</b>	<b><sup>c</sup>ASPE</b>	<b><sup>d</sup>MSPE</b>	<b><sup>e</sup>RMSSPE</b>
Circular	0.0730	1.2542	1.4066	0.0295	0.9257
Spherical	0.0883	1.2427	1.4049	0.0392	0.9180
Tetraspherical	0.0967	1.2424	1.4034	0.0442	0.9170
Pentaspheical	0.0978	1.2380	1.4023	0.0445	0.9152
Exponential	0.1151	1.2264	1.4036	0.0519	0.9001
Gaussian	0.0622	1.2659	1.4332	0.0261	0.9234
Rational quadratic	0.1468	1.2260	1.4029	0.0696	0.9516
<b>Hole Effect</b>	<b>0.0938</b>	<b>1.3416</b>	<b>1.4173</b>	<b>0.0484</b>	<b>0.9768</b>
K-Bessel	0.1160	1.2229	1.3948	0.0523	0.9116
J-Bessel	0.0732	<b>1.2664</b>	1.4220	0.0329	0.9255
Stable	0.1186	1.2283	1.3929	0.0541	0.9145

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error



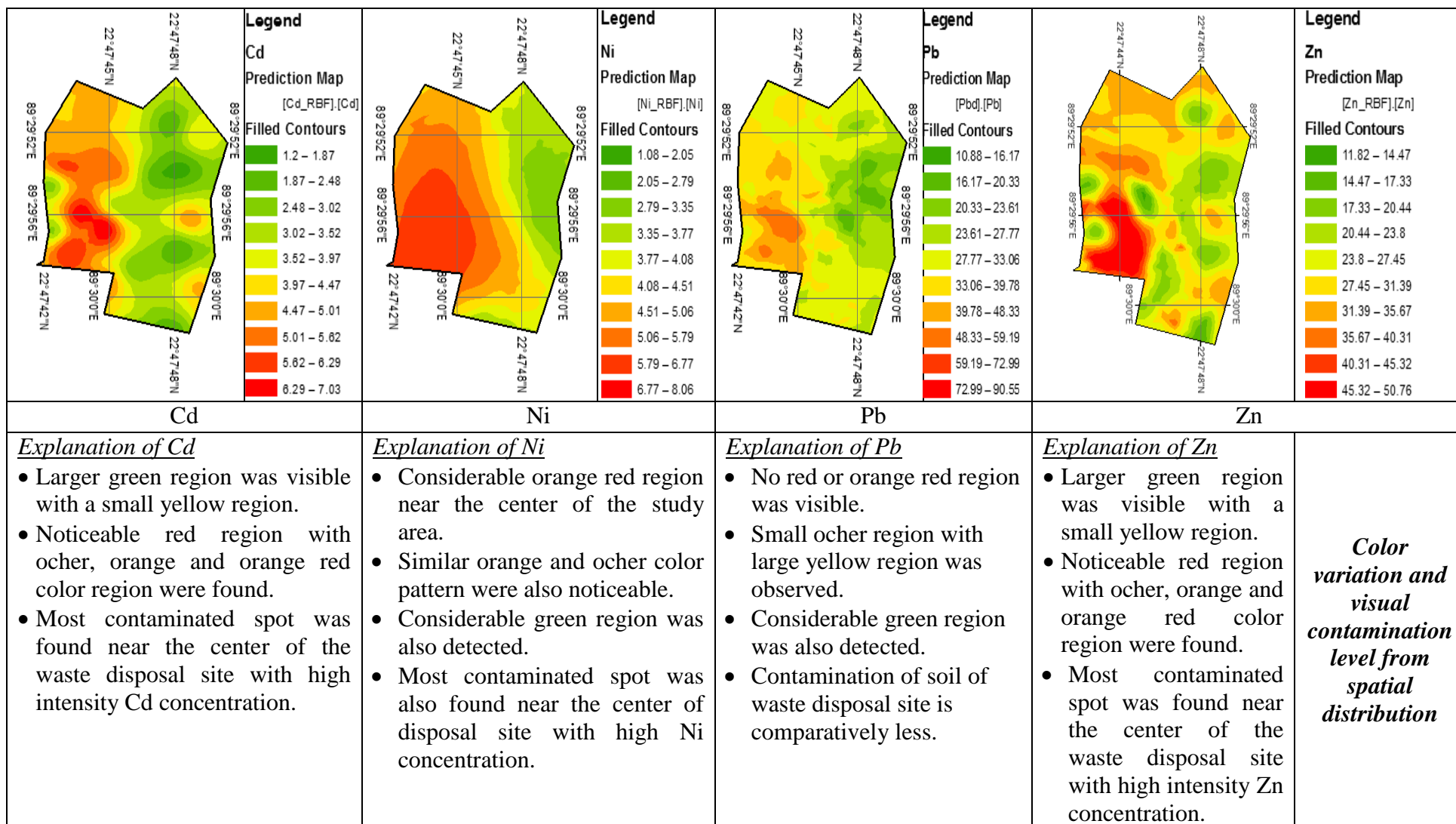


Figure 4.26: Spatial distribution of Cd, Ni, Pb and Zn in soil using ordinary kriging for best fitted model.

The prediction surface produced by rational quadratic model represented larger greenish and yellow region indicated the level of contamination of soil was low. From Figure 4.26, the contamination hotspots were found near the center of the waste disposal site.

#### ***Cross Validation of Nickel, Lead and Zinc***

Among the cross validation results of eleven distinct models for Ni, it was observed that the values of MSPE ranges from 0.0261 to 0.0696 and RMSSPE from 0.9001 to 0.976793. The value of MSPE (0.0484) was found to be closest to zero, RMSSPE (0.9768) closest to 1 as well as ASPE (1.4173) closed to RMSPE with 1.3416 for the model of hole effect (Table 4.20). So, based on this configuration, hole effect model was selected for metal element Ni. The spatial distribution of Ni for this best fitted model shown in Figure. 4.26 and this figure represented larger greenish and yellow region indicated the low level of contamination in soil. Figure 4.26 also showed the contamination hotspots for Pb and Zn near the center of the waste disposal site.

Similarly, the best fitted models were selected through cross validation of eleven distinct models using ordinary kriging interpolation for other metal elements following the same procedure based on the statement of MSPE, RMSSPE, ASPE and RMSPE stated earlier. Cross validation results revealed that the best fitted model for Pb and Zn was same as hole effect (Table 4.21 and Table 4.22). Results reveals in produced prediction surface, greenish color region represented the level of less contamination and redish color area represented the level of highly contaminated soil for both the metal elements of Pb and Zn.

In addition, the cross validation of eleven distinct models using ordinary kriging was performed for the metal elements of Al, As, Ba, Ca, Co, Cr, Cu, Fe, Hg, K, Mn, Na, Sb, Sc, Sr, Ti, and V in soil and the results are provided in Table F.1 to Table F.17 as well as the spatial distribution for these metal elements is depicted in Figure F. 1 to Figure F.17 in the Annex-F. Result reveals that in produced prediction surface for all these metal elements,, greenish color region represented the level of less contamination and redish color area represented the level of highly contaminated soil.

In addition, the best fitted model from cross validation results of eleven distinct models using ordinary kriging interpolation for the studied metal elements of Al, As, Ba, Ca, Cd, Co, Cr, Cu, Fe, Hg, K, Mn, Na, Ni, Pb, Sb, Sc, Sr, Ti, V and Zn in soil is provided in Table 4.23.

Table 4.21: Cross validation results of ordinary kriging for different models of Pb in soil

Models	<sup>a</sup> MPE	<sup>b</sup> RMSPE	<sup>c</sup> ASPE	<sup>d</sup> MSPE	<sup>e</sup> RMSSPE
Circular	0.0998	12.1487	17.6798	-0.0008	0.7172
Spherical	-0.2039	12.1431	17.5225	-0.0183	0.7173
Tetraspherical	-0.2009	12.1397	17.4972	-0.0186	0.7199
Pentaspherical	-0.2148	12.1331	17.4802	-0.0195	0.7202
Exponential	-0.0502	12.2256	17.5881	-0.0101	0.7242
Gaussian	-0.1240	12.1179	17.5628	-0.0141	0.7174
Rational quadratic	0.0842	12.1877	17.5603	-0.0031	0.7257
Hole Effect	<b>-0.5162</b>	<b>12.1974</b>	<b>17.4139</b>	<b>-0.0338</b>	<b>0.7204</b>
K-Bessel	-0.1292	12.1086	17.5570	-0.0145	0.7168
J-Bessel	-0.3339	12.1375	17.3920	-0.0259	0.7232
Stable	-0.1234	12.0964	17.5665	-0.0141	0.7158

Table 4.22: Cross validation results of ordinary kriging for different models of Zn in soil

Models	<sup>a</sup> MPE	<sup>b</sup> RMSPE	<sup>c</sup> ASPE	<sup>d</sup> MSPE	<sup>e</sup> RMSSPE
Circular	0.2522	8.0045	10.3394	0.0017	0.7936
Spherical	0.1441	7.9873	10.2595	-0.0074	0.7829
Tetraspherical	0.1452	8.0009	10.2872	-0.0076	0.7814
Pentaspherical	0.1489	8.0015	10.3103	-0.0072	0.7789
Exponential	0.1497	8.0794	10.5987	-0.0032	0.7682
Gaussian	0.5291	7.9571	10.1408	0.0374	0.9158
Rational quadratic	0.2064	8.0041	10.3895	-0.0009	0.7678
Hole Effect	<b>-0.0197</b>	<b>8.0749</b>	<b>9.1687</b>	<b>-0.0402</b>	<b>0.9258</b>
K-Bessel	0.4592	7.8992	10.1817	0.0262	0.8399
J-Bessel	0.1210	<b>7.6625</b>	9.7297	-0.0117	0.8264
Stable	0.4636	7.9276	10.1875	0.0245	0.8386

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error

Table 4.23: Fitted parameters of the theoretical variogram model for heavy metal  
parameter

Metal elements	Models	Predicted errors				
		<sup>a</sup> MPE	<sup>b</sup> RMSPE	<sup>c</sup> ASPE	<sup>d</sup> MSPE	<sup>e</sup> RMSSPE
Al	Hole Effect	10.844	145.577	170.976	0.05278	0.8876
As	J-Bessel	0.0701	1.4997	1.8929	0.0269	0.8365
Ba	J-Bessel	0.5681	14.5981	17.296	0.02793	0.8887
Ca	Hole Effect	6.5894	48.8408	48.068	0.0529	1.5581
Cd	Rational quadratic	0.0958	1.2498	1.39247	0.03582	1.0139
Co	K-Bessel	0.0469	1.3034	1.2201	0.01969	1.1676
Cr	Circular	-0.0038	2.081	2.217	-0.0045	0.965
Cu	Hole Effect	0.1637	2.5916	3.5444	0.04146	0.7525
Fe	J-Bessel	17.891	496.523	482.812	0.00085	1.0234
Hg	J-Bessel	0.2565	2.0116	1.8865	0.1207	1.7894
K	Exponential	6.7203	80.83	86.4309	0.04847	0.9687
Mn	J-Bessel	0.0597	5.627	6.737	-0.0032	0.88
Na	Rational quadratic	0.6671	24.2788	24.1802	0.0147	1.0495
Ni	Hole Effect	0.0938	1.3416	1.4173	0.0484	0.9767
Pb	Hole Effect	-0.516	12.1973	17.4139	-0.03381	0.7203
Sb	J-Bessel	0.0132	1.54348	1.78	0.00981	0.9004
Sc	J-Bessel	0.0203	2.4072	2.62	0.00956	0.9546
Sr	J-Bessel	0.267	6.185	7.3186	0.033	0.8779
Ti	Rational quadratic	12.739	280.391	296.743	0.02302	1.0204
V	Hole Effect	0.7895	12.1697	14.4548	0.0451	0.8677
Zn	Hole Effect	-0.0196	8.0749	9.1687	-0.0402	0.9258

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error

#### 4.5.3.2 Semivariogram

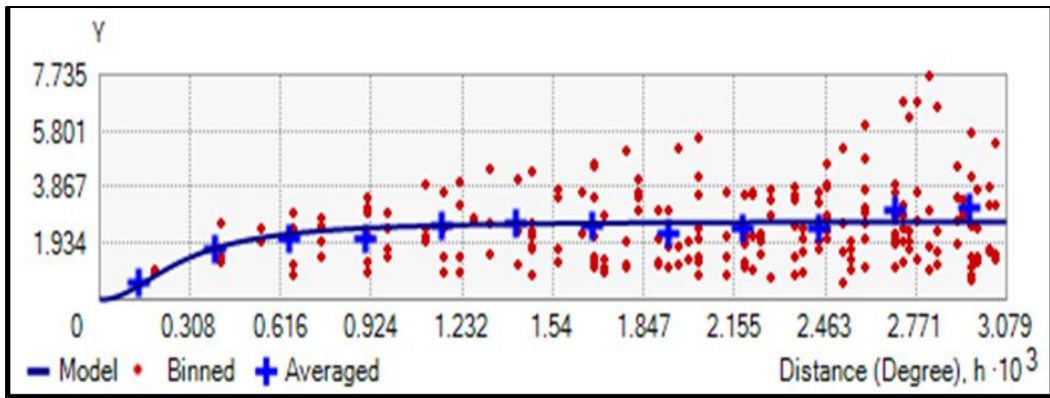
The semivariogram depicts the spatial autocorrelation of the measured sample points. Once each pair of locations is plotted, a model is fit through them. In plot, *ordinate* represents the semivariance and *abscissa* represents lag distance (distance between points). There are certain characteristics or semovariogram parameters such as nugget, sill, range and nugget to sill ratio that are commonly used to describe these models (ESRI, 2001). The full nugget effect should be used at zero distance i.e. the nugget effect refers to the nonzero intercept of the variogram. It is an overall estimate of error caused by measurement inaccuracy and environmental variability (Burgess and Webster, 1980).

Range represents the distance where the model first flattens out. Sample locations separated by distances closer than the range are spatially autocorrelated, whereas locations farther apart than the range are not. In addition, sill indicated the value that semivariogram model attains at the range (the value on the y-axis). The partial sill is the sill minus the nugget (ESRI, 2001).

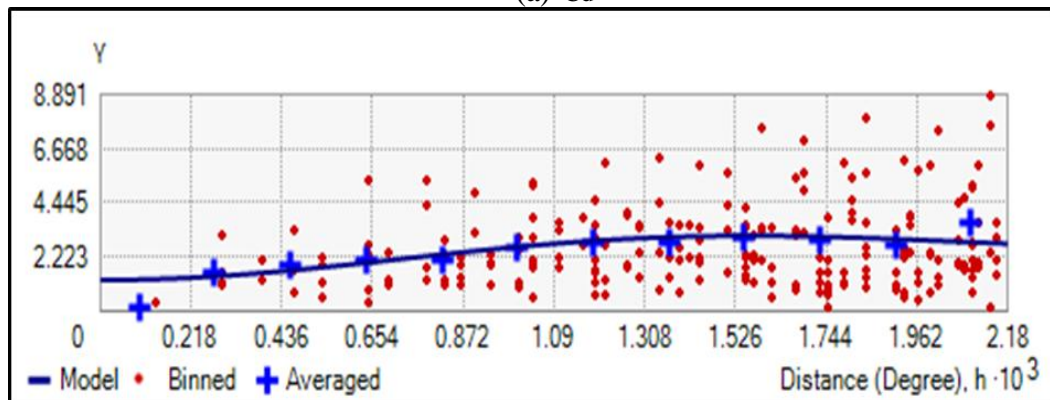
In this study, to find out the best fitted model, cross validation ordinary kriging with eleven distinct models stated earlier was performed for each metal elements in soil. The semivariogram was originated for each metal elements based on this best fitted model. In case of Cd, Rational quadratic model was obtained as the best fitted model and originated semivariogram shown in Figure 4.27(a). Similarly, for Ni, Pb and Zn, the best fitted models were obtained Hole effect model for each metal elements and the originated semivariogram shown in Figure 4.27(b) to Figure 4.27(d), respectively.

Table 4.24: Semivariogram parameters of best fitted models for metal elements

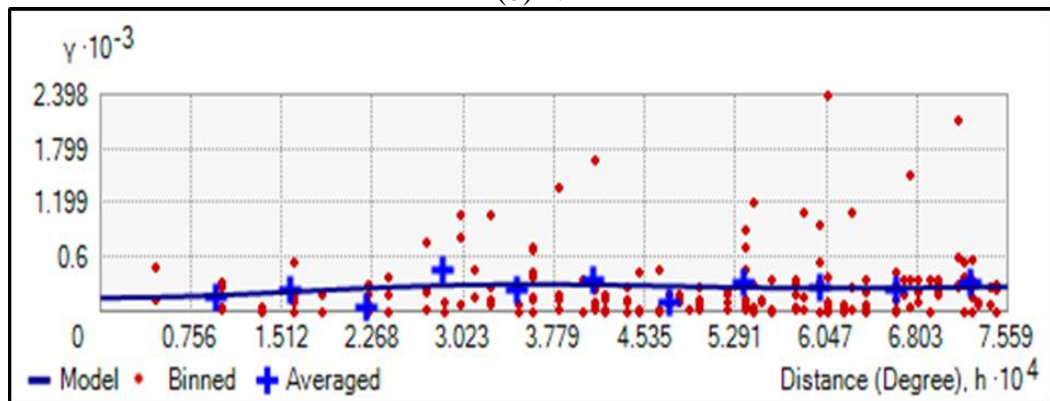
Metals	Models	Range (A)	Nugget (C <sub>0</sub> )	Sill (C+C <sub>0</sub> )	C <sub>0</sub> / (C+C <sub>0</sub> )
Al	Hole Effect	0.0002	21643.0200	39039.8600	0.5544
As	J-Bessel	0.0003	2.5213	5.2603	0.4793
Ba	J-Bessel	0.0041	238.6232	656.6376	0.3634
Ca	Hole Effect	0.0007	0.0000	2701.6600	0.0000
Cd	Rational quadratic	0.0012	0.0027	2.7114	0.0010
Co	K-Bessel	0.0035	0.8716	10.2777	0.0848
Cr	Circular	0.0013	2.2889	6.2986	0.3634
Cu	Hole Effect	0.0036	10.7931	14.5641	0.7411
Fe	J-Bessel	0.0007	0.0000	269384.0000	0.0000
Hg	J-Bessel	0.0010	0.0000	5.1108	0.0000
K	Exponential	0.0028	37.1119	17528.7320	0.0021
Mn	J-Bessel	0.0006	17.6414	47.8002	0.3691
Na	Rational quadratic	0.0058	381.8491	1365.4756	0.2796
Ni	Hole Effect	0.0022	1.2869	2.7768	0.4634
Pb	Hole Effect	0.0005	150.6974	272.1019	0.5538
Sb	J-Bessel	0.0045	2.6217	5.9723	0.4390
Sc	J-Bessel	0.0044	5.6103	14.3584	0.3907
Sr	J-Bessel	0.0042	44.3647	86.1366	0.5151
Ti	Rational quadratic	0.0004	42693.8900	262171.1400	0.1628
V	Hole Effect	0.0026	158.4399	257.0829	0.6163
Zn	Hole Effect	0.0006	0.0000	96.4491	0.0000



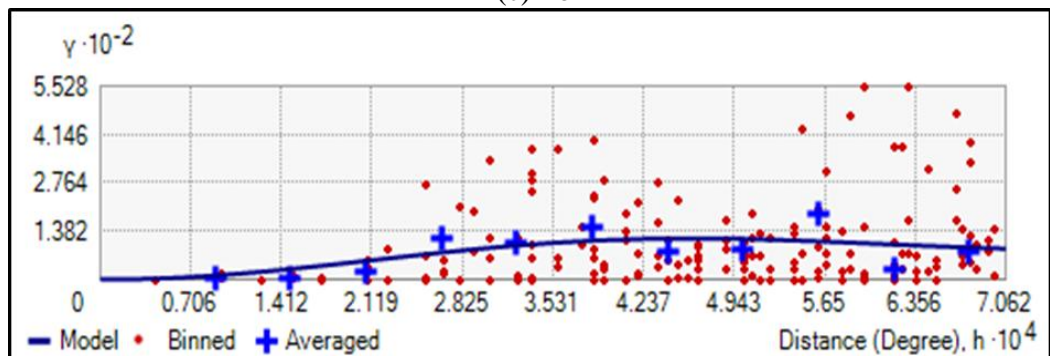
(a) Cd



(b) Ni



(c) Pb



(d) Zn

Figure 4.27: Semi variogram of metal elements.

The semivariogram parameters such as range, nugget, partial sill and nugget to sill ratio from obtained best fitted model for each metal elements are provided in Table 4.24. A study conducted by Yasrebi et al. (2010) and stated that when the nugget to sill ratio,  $C_0 / (C_0 + C)$  was found <25%, 25 to 75% and >75%, it was considered the soil parameters are strongly, moderately and weakly distributed spatially, respectively. In this study, from semivariogram parameters,  $C_0 / (C_0 + C)$  were found to be 0.099, 8.48, 0.21 and 16.28% for the metal elements of Cd, Co, K, and Ti respectively. This statement (less than 25%) that exhibited that the these metal elements were strongly distributed spatially. In addition,  $C_0 / (C_0 + C)$  were found as 55.44, 47.93, 36.34, 36.34, 74.11, 36.91, 27.96, 46.34, 55.38, 43.90, 39.07, 51.50 and 61.63% for Al, As, Ba, Cr, Cu, Mn, Na, Ni, Pb, Sb, Sc, Sr and V, respectively, indicated that these metal elements in soil were moderately correlated spatially. In contrast, the metal elements of Ca, Fe, Hg and Zn were non spatially correlated as the  $C_0 / (C_0 + C)$  was found to be zero.

#### **4.6 Assessment of Method Performance**

The performance of interpolation methods of IDW, LP, RBF and ordinary kriging (OK) were assessed for evaluating the accuracy of these selected models. In this analysis, to check the accuracy of the selected models of IDW, LP, RBF and OK, the mean absolute percentage error (MAPE), Relative improvement (RI) and goodness of prediction (G-value) were considered provided in Table 4.25. A research of Yao et al. (2013) stated that small values (zero) of MAPE represented a model with less errors and more accurate predictions. Furthermore, G-value equal to 1 indicated perfect prediction, a positive value indicated a more reliable model, whereas a negative value specified a less reliable model (Yao et al., 2013). In addition, small values (zero) of RI indicated a model with perfect predictions.

##### **4.6.1 Mean Absolute Percentage Error**

The performance of interpolation methods of IDW, LP, RBF and ordinary kriging (OK) for all studied metal elements were analyzed based on MAPE. On the basis of MAPE, result reveals that IDW with power 1 to 5 and RBF with five functions as CRS, ST, MQ,

IMQ and TPS performed better than that of ordinary kriging and LP with power of 1 to 3 (Table 4.25).

Table 4.25: Performance assessment of interpolation methods based on MAPE

Parameter	MAPE									
	OK	IDW (1-5)	LP			RBF				
			1	2	3	CRS	ST	MQ	IMQ	TPS
Fe	0	0	0.92	1.02	0.987	0	0	0	0	0
Mn	0.28	0	1.41	1.55	1.382	0	0	0	0	0
Cr	0.21	0	0.49	0.92	0.868	0	0	0	0	0
Cu	0.29	0	0.37	0.59	0.417	0	0	0	0	0
Pb	0.06	0	0.33	0.41	0.386	0	0	0	0	0
Zn	0	0	0.14	0.24	0.279	0	0	0	0	0
Ni	0.10	0	0.35	0.38	0.327	0	0	0	0	0
Cd	0	0	0.40	0.41	0.391	0	0	0	0	0
As	0.22	0	0.53	0.55	0.575	0	0	0	0	0
Hg	0	0	0.65	0.98	0.779	0	0	0	0	0
Co	0.05	0	0.08	0.18	0.198	0	0	0	0	0
Na	0.22	0	0.61	0.68	0.692	0	0	0	0	0
K	0.01	0	0.44	0.48	0.460	0	0	0	0	0
Ca	0	0	0.07	0.30	0.224	0	0	0	0	0
Al	0.14	0	0.29	0.38	0.329	0	0	0	0	0
Ti	0.05	0	0.24	0.26	0.013	0	0	0	0	0
Sb	0.13	0	0.30	0.28	0.305	0	0	0	0	0
Sc	0.06	0	0.18	0.20	0.192	0	0	0	0	0
Sr	0.07	0	0.20	0.17	0.201	0	0	0	0	0
V	0.16	0	0.33	0.32	0.023	0	0	0	0	0
Ba	0.07	0	0.21	0.25	0.222	0	0	0	0	0

In this study, IDW with power 1 to 5 exhibited MAPE value of zero for each metal element which means the percentage of relative error for IDW interpolation was zero (Table.4.25). For Cd, MAPE value for LP2 exhibited comparatively poor performance as the MAPE value was found as 0.41. The MAPE value was found to be zero for IDW with power of 1 to 5 and different kernel functions of RBF. The LP2 performed poor for showing spatial distribution of metal elements of Fe, Mn, Cr, Cu, Pb, Ni, Cd, Hg, K, Ca, Al, Ti, Sc and Ba than that of other interpolation methods. In addition, IDW and RBF provided accurate prediction of spatial distribution of metal elements in soil. Furthermore, ordinary kriging also provided better mapping for Fe, Zn, Cd, Hg, Ca. The value of MAPE for Pb, Co, K, Ti, Sc, Sr and Ba was also very insignificant, which indicated better interpolation of metal elements, LP1 showed poor performance for the metal element of V.



In summary, IDW with power 1 to 5 and RBF with different kernel functions provided the most accurate prediction. Moreover, ordinary kriging (OK) and LP3 also made better interpolation for spatial distribution of metal elements. Besides, LP2 showed worst performance between all interpolation techniques.

#### 4.6.2 Goodness of Prediction

The performances of interpolation methods of IDW, LP, RBF and ordinary kriging (OK) for all studied metal elements were analyzed based on the concept of goodness of prediction (G-value). From Table 4.26, it was found that IDW with power 1 to 5 and RBF with five functions of CRS, ST, MQ, IMQ and TPS exhibited the value of 1 for each metal element which means in IDW and RBF, the prediction statistic was perfect.

Table 4.26: Performance assessment of interpolation methods based on G-value

Parameter	G-Value									
	OK	IDW (1-5)	LP			RBF				
			1	2	3	CRS	ST	MQ	IMQ	TPS
Fe	1	1	0.272	0.12	0.169	1	1	1	1	1
Mn	0.86	1	0.196	0.00	0.178	1	1	1	1	1
Cr	0.81	1	0.687	0.13	0.218	1	1	1	1	1
Cu	0.55	1	0.710	0.47	0.618	1	1	1	1	1
Pb	0.82	1	0.367	0.27	0.325	1	1	1	1	1
Zn	1	1	0.259	0.45	0.414	1	1	1	1	1
Ni	0.72	1	0.541	0.47	0.476	1	1	1	1	1
Cd	1	1	0.399	0.37	0.430	1	1	1	1	1
As	0.42	1	0.594	0.55	0.526	1	1	1	1	1
Hg	1	1	0.637	0.21	0.443	1	1	1	1	1
Co	0.87	1	0.954	0.82	0.805	1	1	1	1	1
Na	0.70	1	0.526	0.46	0.452	1	1	1	1	1
K	0.99	1	0.575	0.50	0.536	1	1	1	1	1
Ca	1	1	0.840	0.85	0.888	1	1	1	1	1
Al	0.67	1	0.683	0.52	0.617	1	1	1	1	1
Ti	0.86	1	0.701	0.52	-2.949	1	1	1	1	1
Sb	0.66	1	0.691	0.72	0.682	1	1	1	1	1
Sc	0.66	1	0.623	0.66	0.693	1	1	1	1	1
Sr	0.62	1	0.669	0.59	0.655	1	1	1	1	1
V	0.64	1	0.583	0.54	-2.302	1	1	1	1	1
Ba	0.73	1	0.742	0.65	0.713	1	1	1	1	1

Moreover, LP showed poor prediction, as the G-value was smaller than 1, especially in case of Fe, Mn, Pb, Zn and Cd (G-value < 0.5). For Cr and Hg, LP with order 2 offered very poor performance as the G-values were obtained 0.475 and 0.215, respectively. LP with order 3 provided negative values for Ti and V represented a less reliable model. The metal elements of Co, Ca and Ba contributed better result for all the performed interpolation techniques. On the basis of G-value, the technique of IDW with power 1 and RBF with five functions of CRS, ST, MQ, IMQ and TPS provided the best prediction interpolation for all metal elements in soil.

### 4.6.3 Relative Improvement

Relative improvement (RI) is considered to check the performance of interpolation methods used in this study for all the studied metal elements were analyzed based on The RI of interpolation techniques was also reported in Table 4.27.

Table 4.27: Performance assessment of interpolation methods based on RI

Parameter	RI									
	OK	IDW					LP			RBF
		1	2	3	4	5	1	2	3	Best Fitted
Fe	12.8	0	4.9	7.71	9.39	10.5	5.15	1.005	3.20	23.232
Mn	19.1	0	5.2	7.76	9.55	11.0	5.35	3.253	6.04	26.975
Cr	12.2	0	5.3	8.63	9.90	10.2	20.9	3.211	7.89	23.449
Cu	32.1	0	8.5	14.4	18.5	20.7	33.5	22.78	7.47	24.827
Pb	21.9	0	6.8	10.5	13.1	15.0	52.2	4.678	6.81	23.301
Zn	15.6	0	5.0	7.41	8.72	9.23	21.6	1.536	7.55	18.470
Ni	9.15	0	5.9	8.33	9.27	9.28	15.0	5.700	17.1	24.526
Cd	13.4	0	5.9	8.34	8.75	8.34	17.6	6.372	9.33	25.406
As	23.0	0	10.	17.9	21.2	22.5	7.38	13.13	10.2	22.592
Hg	13.1	0	8.1	13.8	16.9	18.4	27.3	7.083	11.5	25.382
Co	32.3	0	9.5	15.8	20.4	22.9	20.2	8.839	15.9	17.512
Na	16.9	0	7.0	10.5	11.8	12.4	6.26	0.814	5.45	24.512
K	21.5	0	9.5	14.7	17.3	18.7	14.7	3.114	5.91	21.444
Ca	7.46	0	8.1	12.7	15.6	17.4	29.9	7.925	17.1	19.284
Al	12.3	0	9.8	15.3	18.4	20.0	7.63	5.069	12.8	20.196
Ti	14.0	0	9.4	14.1	16.5	17.7	19.3	7.905	28.7	17.546
Sb	12.9	0	10.	16.2	19.9	22.2	19.5	8.476	10.6	19.089
Sc	10.1	0	7.8	12.1	14.5	15.6	32.8	7.579	7.53	17.184
Sr	11.0	0	9.3	14.1	16.8	18.0	26.0	9.735	10.8	19.312
V	10.5	0	8.5	13.2	16.0	17.5	28.0	7.802	21.5	21.812
Ba	18.3	0	11.	17.5	21.6	23.9	20.4	10.31	20.4	16.104

The zero value of this mathematical term exhibited better performance (Yasrebi et al., 2009). In this study, IDW with power 1 contributed zero value for Fe which indicated perfect interpolation, whereas, LP with order 2 has the RI value of 1.005, also indicated better interpolation technique than that of other techniques. The interpolation technique of ordinary kriging and selected best fitted model (IMQ) obtained from cross validation of RBF interpolation offered poor interpolation prediction that contributed poor map quality for metal elements of Fe, Mn, Cr, Pb, Zn, Ni, Hg, Co, Na, K, Al, Ti, Sb, Sc, V and Ba (Table 4.27).

IDW with power 1 followed by LP with order 2 performed better than other interpolation technique for metal elements of Fe, Mn, Cr, Pb, Zn, Ni, Hg, Co, Na, K, Al, Ti, Sb, Sc, V and Ba. Similar results and relation were also obtained by Yao et al. (2013). In this study, result reveals that IDW with power 1 followed by LP with order of 3 performed better than that of other interpolation techniques for the metal elements of Cu and As, as the value of RI was found smaller in compare to others. In addition, for Cd and Sr, IDW with power 1 provided perfect interpolation, whereas, IDW with power 2 also provided better predictions. Moreover, for Ca, IDW with power 1 followed by ordinary kriging performed better than that of other interpolation techniques as the value of RI was found comparatively smaller than that of other techniques.

On the basis of RI, IDW with power 1 provided best prediction as the RI value was found to be 0 for all the studied metal elements. LP with order 2 also contributed better interpolation as the RI value was smaller compared to other interpolation techniques. On the basis of MAPE and G-value, IDW with power 1 to 5 and RBF with different functions of CRS, ST, MQ, IMQ and TPS provided better prediction interpolation as MAPE and G-value was found to be 0 and 1, respectively, for all studied metal elements. Finally, based on the above analysis, it can be concluded that IDW with power of 1 was the most accurate interpolation techniques among all interpolation techniques

#### **4.7 Artificial Neural Network**

In this study, artificial neural network (ANN) was performed to predict the accuracy of the observed data for metal elements obtained from laboratory. Predictions of observed data were evaluated by ANN and a relation between the predicted and observed values was

established for each metal elements presence in soil of waste disposal site. In ANN analysis, the data was divided like this: 70% training data, 20% validation data and 10% test data. The performance of ANN was investigated by plotting mean squared error (MSE) plot for each metal element. The value of MSE is always non-negative, and values are closer to zero which indicated better prediction results (Shahin et al.,2004). In addition, the error histogram plots for training data were to provide additional verification of network performance. The correlation coefficient (R-value) measures the correlation between observed and predicted values. The value of R varies from 1 to 0. The R-value of 1 is meant for a close relationship between observed and predicted data; whereas, The R-value of 0 indicated random relationships between observed and predicted values. The MSE and R-value from ANN for all studied metal elements are provided in Table 4.28. The outputs of ANN for the metal elements of Cd, Ni, Pb and Zn are discussed in details in this section. Moreover, the he outputs for the other metal elements of Al, As, Ba, Ca, Co, Cr, Cu, Fe, Hg, K, Mn, Na, Sb, Sc, Sr, Ti, and V were reported in Figure G.1 to Figure G.51, respective, in the Annex-G.

#### **4.7.1 Mean Standard Error**

MSE measures the average of the squares of the errors or deviations—that is, the difference between the observed and predicted values that occurs because of randomness or because the estimator doesn't account for information that could produce a more accurate estimate (Alkaiem et al., 2016). The performance plot for the metal elements of Cd, Ni, Pb and Zn was presented and hence described in this section.

##### ***Performance Plot of Cadmium***

The performance plot is shown in Figure 4.28. The performance plot for Cd indicated that MSE became small as nu number of epochs (one complete sweep of training, testing and validation) were increased (Figure 4.28). The error of validation set and test set consumed similar characteristics and no significant over fitting had occurred by epoch 1000 (where best validation performance has occurred). From Figure 4.28, the best validation performance for Cd was found as  $5.1068 \times 10^{-6}$ , closed to zero which indicated a close relation between the observed and predicted data for Cd.

### Performance Plot for Nickel, Lead and Zinc

The performance plot for Ni, Pb and Zn were provided in Figure 4.29. It was found that the best performance was obtained at  $1.4143 \times 10^{-5}$  in case of Ni, which indicated the error between observed and predicted value was very negligible.

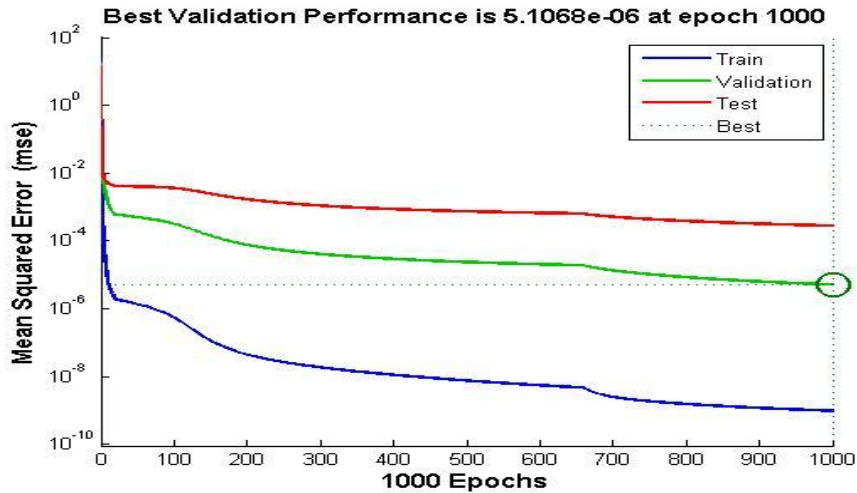


Figure 4.28: Performance plot of Cd.

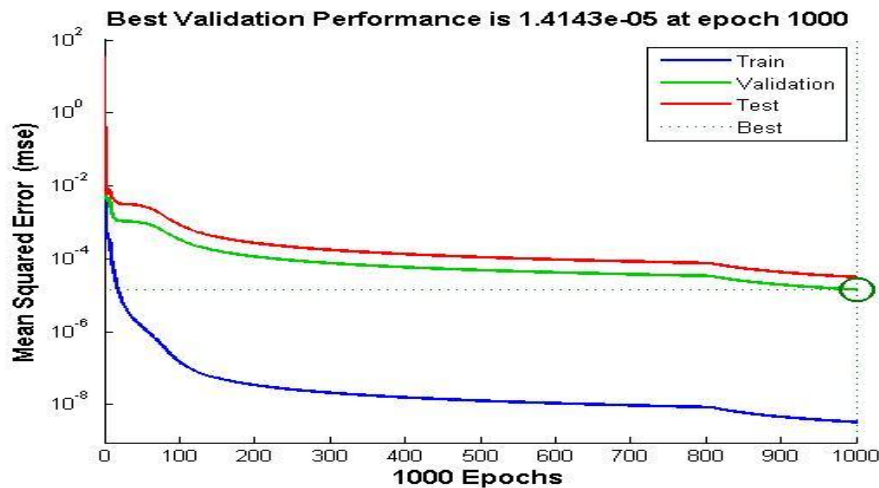


Figure 4.29: Performance plot of Ni.

Thus the prediction was accurate for Ni. Similarly,  $1.1693 \times 10^{-7}$  was reported as the MSE value for Pb, which was closed to zero (Figure 4.30). Thus, the prediction was precise. Furthermore, the best performance was obtained at  $3.1212 \times 10^{-5}$  in case of Zn, which indicated the error between observed and forecasted value was very negligible (Figure 4.31). Thus the prediction was exact. In addition, the performance plot for metal elements of Al, As, Ba, Ca, Co, Cr, Cu, Fe, Hg, K, Mn, Na, Sb, Sc, Sr, Ti, and V were reported in Figure G.1 to Figure G.17, respectively, in Annex-G.

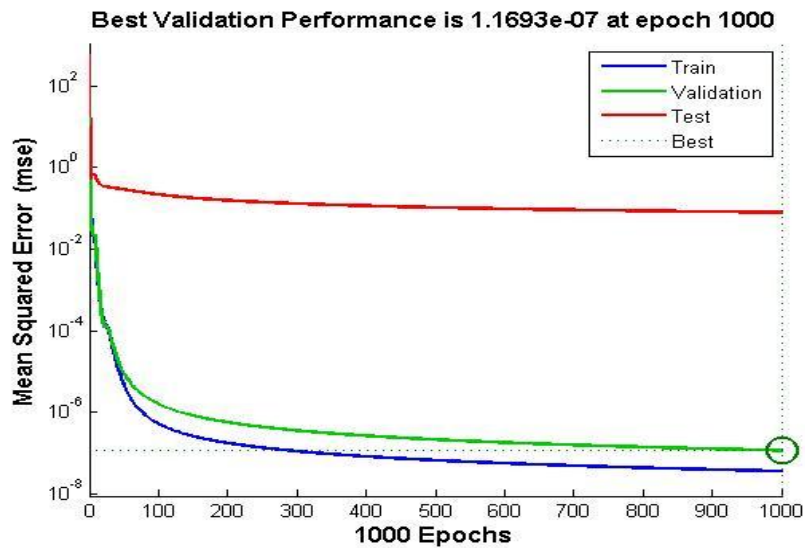


Figure 4.30: Performance plot of Pb.

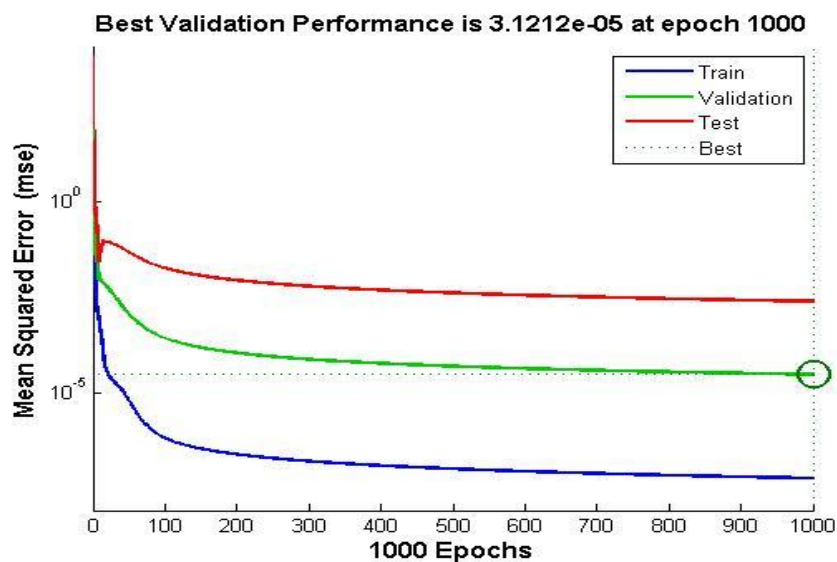


Figure 4.31: Performance plot of Zn.

#### 4.7.2 Regression Coefficient

The correlation coefficient (R-value) measures the correlation between observed and predicted values. The value of R varies from 1 to 0. The R-value of 1 is meant for a close, relationship between observed and predicted data; whereas R-value of 0 indicates random relationships between observed and predicted data. The R-value between the predicted and the actual values of metal elements present in contaminated soil were shown in Figures 4.32-4.35 for training, validation, testing and the whole datasets.

Result reveal that the values of R varies from 0.99922 to 1 obtained for the training, validation, testing and the whole dataset, respectively (Table 4.28). Results showed that the predicted concentration of metal elements were very close to the actual values for all the datasets. The R-value of observed and predicted data for metal elements of Cd, Ni, Pb and Zn was represented in this section.

Table 4.28: Performance assessment of nntool model

Parameter	Mean Standard Error (MSE)	Regression Coefficient (R-value)
Al	$2.19 \times 10^{-7}$	0.99993
As	0.0266	0.99922
Ba	$4.22 \times 10^{-8}$	1
Ca	0.00011	1
Cd	$5.11 \times 10^{-6}$	0.99999
Co	$2.41 \times 10^{-8}$	0.99958
Cr	$7.92 \times 10^{-8}$	0.99942
Cu	$9.61 \times 10^{-6}$	0.99998
Fe	$3.05 \times 10^{-8}$	1
Hg	0.00012	0.99998
K	0.0017	1
Mn	0.095	0.99974
Na	$9.17 \times 10^{-6}$	1
Ni	$1.41 \times 10^{-5}$	1
Pb	$1.17 \times 10^{-7}$	0.99996
Sb	$1.35 \times 10^{-8}$	0.99995
Sc	$9.28 \times 10^{-8}$	0.99991
Sr	$2.99 \times 10^{-6}$	1
Ti	$1.73 \times 10^{-6}$	1
V	$7.27 \times 10^{-5}$	1
Zn	$3.12 \times 10^{-5}$	1

#### ***Regression Coefficient of Cadmium***

In case of Cd, the R-values obtained from training, validation, testing and the whole datasets for metal elements presence in soil of waste disposal site reported in Figure 4.32. The R-value were obtained 0.99999, 0.99969, 0.99995 and 0.99993 for training, validation, testing and the whole datasets, respectively. This R-value was closed to 1, indicated a close relationship between observed and predicted values (Figure 4.32).

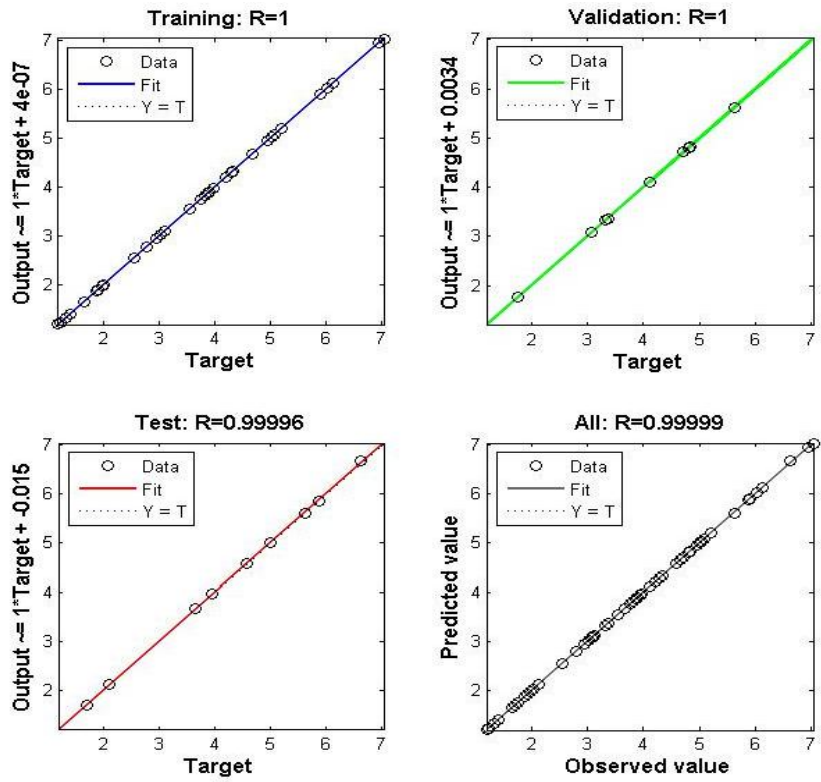


Figure 4.32: Regression coefficient of Cd in soil.

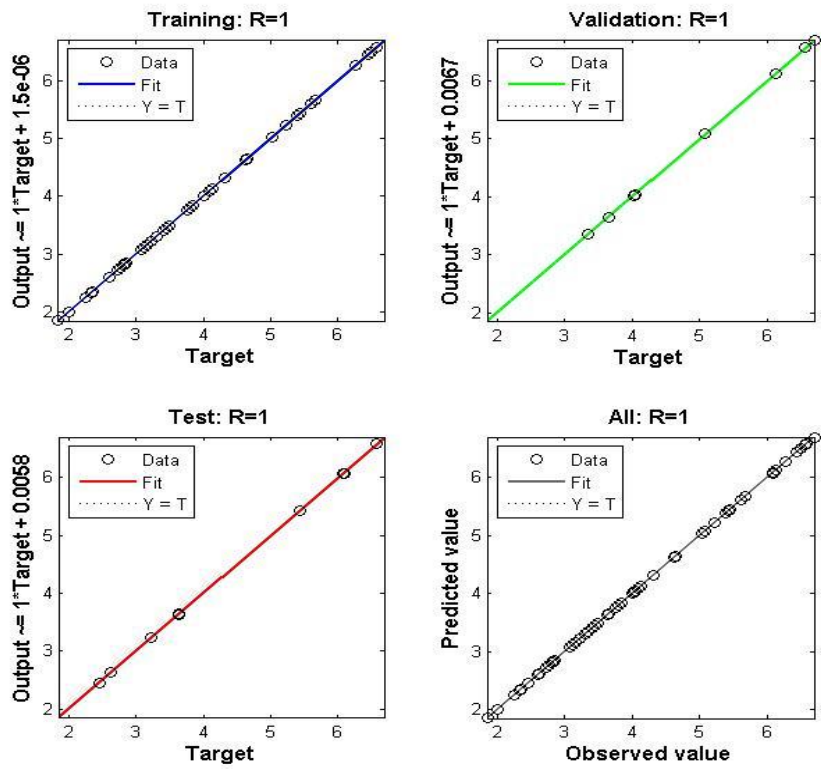


Figure 4.33: Regression coefficient of Ni in soil.



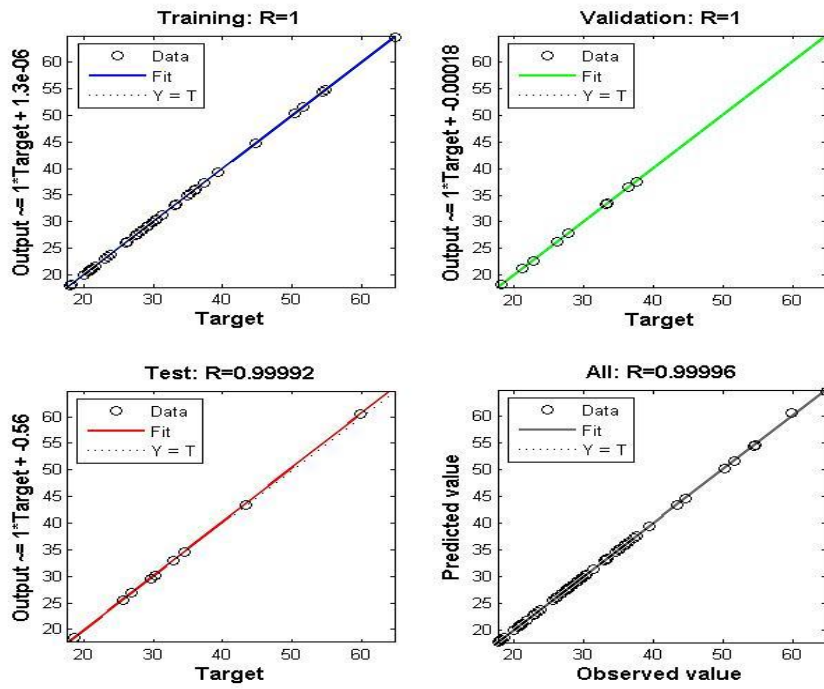


Figure 4.34: Regression Coefficient of Pb.

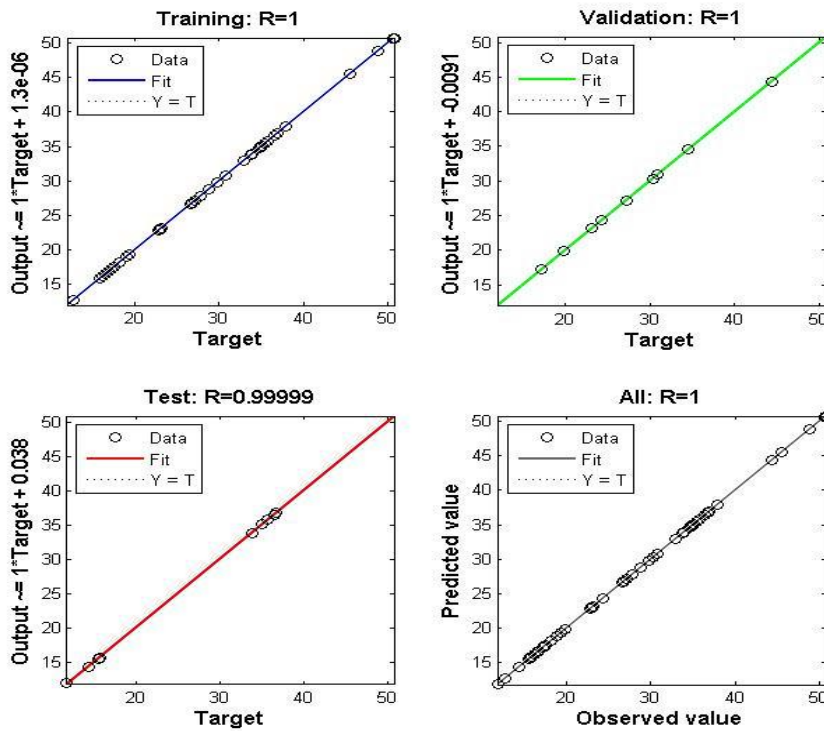


Figure 4.35: Regression Coefficient of Zn.

### ***Regression Coefficient of Nickel, Lead and Zinc***

The R-values of training, validation, testing and the whole datasets were found to be 1, which directly represented the observed values of Ni was predicted accurately (Figure 4.33). From Figure 4.34, the R-values acquired as 1, 1, 0.99992 and 0.99996 from training, validation, testing and the whole datasets, respectively, for the metal element of Pb. As the R-value was closed to 1 (0.99996), the prediction was considered to be precise. Moreover, based on the above statement, it can be decided that there was a close relationship between observed and predicted data in case of Zn (Figure 4.35). In addition, the error histogram plotted by ANN for metal elements of Al, As, Ba, Ca, Co, Cr, Cu, Fe, Hg, K, Mn, Na, Sb, Sc, Sr, Ti, and V provided in Figure G.18 to Figure G. 34 in the Annex-G.

### **4.7.3 Error Histogram**

The error histogram plot for training data provided the additional verification of network performance which showed the distribution of the residuals between targets and network outputs. In this study, the error histogram plot for training data shown in Figure 4.36-4.39. In this plot the yellow color “outliers” indicated the average value of error between target and output. The blue, green and red bars represented training data, validation data and testing data, respectively (Figure 4.36 to Figure 4.39). When the data falls near to zero error line (outliers) indicates low error and perfect prediction of the predicted values (Alkaiem et al., 2016). The error histograms of target and output values were represented for metal elements of Cd, Ni, Pb and Zn were represented in this section.

#### ***Error Histogram of Cadmium***

The error histogram of Cadmium is shown in Figure 4.36. Figure 4.36 reveals that most of the data falls very near to zero error line (outliers ) which provided an idea to check either the data points were similar to the rest of the data set or not. This findings are well agreed with research conducted by Alkaiem et al. (2016) and it was found that most data fall on zero error line included testing and training. In this study, the error was reported as - 0.00026 for Cd, which indicated smaller error, tends to zero with perfect prediction.

#### ***Error Histogram of Nickel, Lead and Zinc***

For metal element of Ni, most data fall on zero error line which provided an idea to check the outliers to indicate those data points were similar to the rest of the data set. From

Figure 4.37, the error was reported as  $-6.3 \times 10^{-5}$ , which indicated smaller error, tends to zero which indicate perfect prediction.

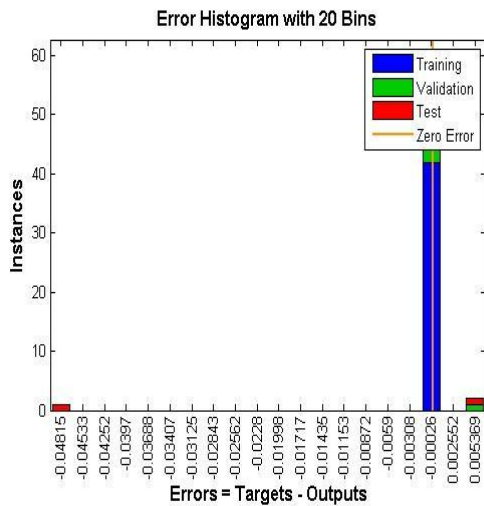


Figure 4.36: Histogram plot of Cd.

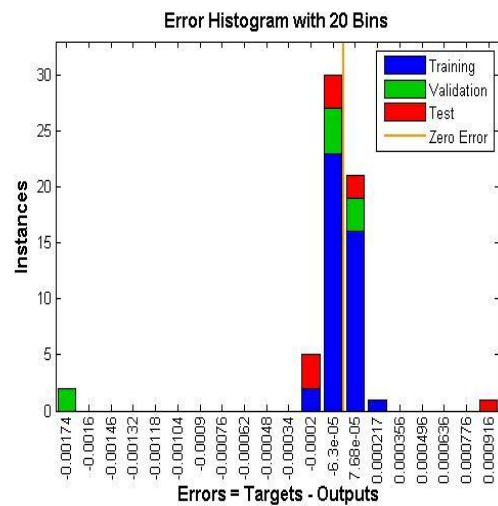


Figure 4.37: Histogram plot of Ni.

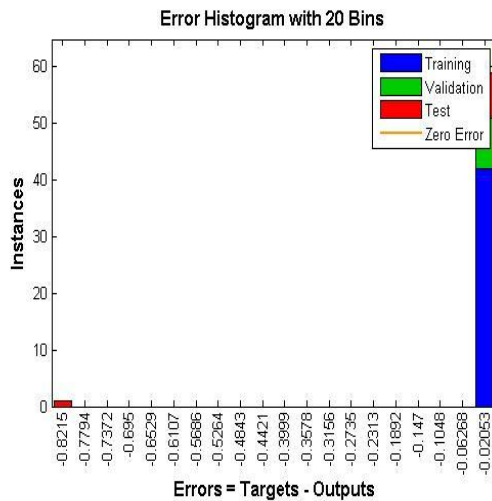


Figure 4.38: Histogram plot of Pb.

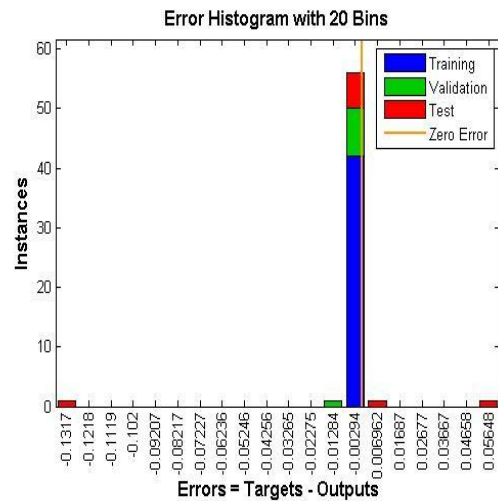


Figure 4.39: Histogram plot of Zn.

Metal element of Pb showed that most data fall on zero error line which provided an idea to check the outliers to indicate those data points were similar to the rest of the data set (Figure 4.38). The error was reported as  $-0.0205$ , which indicated smaller error, tends to zero which indicate perfect prediction. Furthermore, Zn also showed negligible value of error ( $0.0029$ ), which indicate prediction of the concentration of metal element of Zn was faithfully perfect (Figure 4.39). In addition, the error histogram plotted by ANN for metal elements of Al, As, Ba, Ca, Co, Cr, Cu, Fe, Hg, K, Mn, Na, Sb, Sc, Sr, Ti, and V provided in Figure G.35 to Figure G.51 in the Annex-G.

#### **4.8 Concluding Remarks**

The term “landfill” is a unit, designed and operated for the disposal of municipal solid waste (MSW) to protect the environmental components (water, soil, air, etc.) and all the living beings from the contaminants like metal elements presence in MSW stream (Alamgir et al., 2005). MSW can be disposed either in form of open dumping or disposal in sanitary landfill. In developing countries like Bangladesh, open dumping has become a traditional practice to dispose MSW. Due to long retention of MSW on dumping site, MSW decomposes and produces three components of solid, liquid and landfill gas. In addition, leachate and contaminated soil creates vulnerable effects to the environmental components and nearby inhabitants. There has long been concern about the issue of contamination of soil by metal elements because of their toxicity for plant, animal and human beings as well as their lack of biodegradability (Li et al., 2006; Zhuang et al., 2009). Soil is a multi-phase system contaminated with the presence of xenobiotic (human-made) chemicals or other alteration in the natural soil environment due to improper disposal of MSW, industrial activity and agricultural chemicals. Soil contamination occurs when the presence of toxic chemicals, pollutants or contaminants with high concentrations in soil and it has great risk to plants, wildlife, humans and of course for the soil itself (Jia et al., 2010). The concentration of metal elements in soil and their impact on ecosystems can be influenced by many factors such as the parent rock, climate and anthropogenic activities (Jia et al., 2010).

In Khulna city, most of the MSWs are collected from door-to-door without any sorting and dumped in open disposal site at Rajbandh. Due to inadequate management practices of MSW, necessities arise to take steps for the proper disposal of MSW as well as maintenance of disposal site at Rajbandh, Khulna, Bangladesh. Moreover, to date, there is no comprehensive study to examine how the metal elements are correlated to each other as well as their possible sources of contamination such as anthropogenic or human activities and natural parent materials. In addition, there is no ready manual or guidebook related to this research from where one can easily get the information about quantitative distribution of metal elements spatially as well as the level of contamination of soil due to presence of metal elements in soil. Thus unplanned disposal, irregular unloading and improper maintenance of MSW in waste disposal site consequent of contamination of pollutants

such as metal elements in soil and environment of waste disposal site as well as in surrounding areas. This random disposal of MSW also consequences health effects on all living being near the waste disposal site

The main purpose of this study was to find out the possible sources of contamination of metal elements and the distribution of metal elements spatially in soil of waste disposal site. To these endeavors, total sixty soil samples were collected at a depth of 0-30 cm from the existing ground surface from different selected locations within the waste disposal site. These study periods covered both the dry season (March to May, 2016) and rainy season (June to August, 2016). In the laboratory, the concentration of metal elements such as Al, As, Ba, Ca, Fe, Hg, K, Mn, Na, Na, Pb, Sb, Sc, Sr, Ti, V and Zn in soil were measured through standard test methods. To evaluate the nature of measured concentration of all the studied metal elements, the normality test was performed through K-S test and S-W test categorized distinctly for both the dry and rainy season. For more accurate result, normal QQ plot was also plotted to check the distribution of metal elements either normally or not. Conventional statistics, based on mathematical indices such as mean, median, maximum, minimum, CV, SD, skewness and kurtosis were evaluated to show the variability of metal elements in soil for both the dry and rainy season. Moreover, multivariate statistics including pearson correlation was accomplished to measure the linear correlation between metal elements in soil. In addition, The PCA was implemented to explain the major generation sources of metal elements to exploit the soil of waste disposal site. Furthermore, the AHC was also performed to classify the metal elements on the basis of dissimilarity between sets of observations. Perfect prediction of metal elements using different interpolation techniques of IDW, LP, RBF's and ordinary kriging exhibited better field condition of spread out of metal elements in soil of waste disposal site. Moreover, the accuracy of these interpolation techniques were checked based several indices such as MAPE, RI and G-value. In this study, a network model was developed by ANN to predict and depict the validity based MSE and R-value of the observed data of metal elements obtained from laboratory. It was found that the predicted values from ANN were almost same as obtained from laboratory. The main outcome of this study was to know the correlation of metal elements with each other in soil, possible sources of their contamination, distribution of metal elements spatially and the level of contamination of soil due to presence of metal elements in soil.

## **CHAPTER V**

### **COMPARISON AND VALIDATION**

#### **5.1 General**

This chapter deals with the comparison of results obtained from this study with previous studies conducted in different parts of world similar to this study. Previously, conventional statistics were performed in many researches to understand the distribution of metal element concentrations in soil. Based on multivariate statistics such as principal component analysis (PCA), agglomerative hierarchical clustering (AHC), generation sources of heavy metal contamination in soil were identified in previous studies. In addition, geostatistical analyses were performed to identify contamination hotspots of respective study areas. In most of the studies, inverse distance weighting (IDW) and ordinary kriging (OK) were performed to see the actual condition spread of metal elements in study area. In some cases other interpolation techniques such as, radial basis function's (RBF's), local polynomial interpolation (LPI) were also performed. Artificial neural network (ANN) also used to check the accuracy of predicted values in some previous studies. In this study, conventional statistics in terms of mean, median, skewness, kurtosis was evaluated to see the distribution of metal contamination in and around the soil of waste disposal site in Khulna. To identify the possible generation sources of metal elements, whether it comes from anthropogenic activities or natural sources, multivariate statistics such as principal component analysis (PCA), agglomerative hierarchical clustering (AHC) were performed. Moreover, to visualize the contamination pattern of metal elements in and around the soil of waste disposal site, geostatistical analyses such as IDW, LPI, RBF's and ordinary kriging (OK) were performed. A prediction was made for each metal concentration and accuracy was checked using ANN.

#### **5.2 Comparison with Previous Studies**

##### **Case I: Beijing suburbs**

A study was performed in Beijing at 2015 by Zou et al. to analyze spatial variations and sources of heavy metals in farmland soils of Beijing suburbs. In his study, he analyzed for

eight metal elements. The results exhibited Cd and Hg concentrations in the soils showed greater spatial heterogeneity than the concentrations of the other heavy metals, with coefficients of variation of 0.50 and 0.86, respectively (Table 5.1). Because the entire study was located on plains areas without significant differences in the types of soil parent material or land use, the strong spatial heterogeneity in the distribution of these two heavy metals was most likely directly related to human activities. The coefficients of variation for the As, Cr, Ni, Pb and Zn in soil concentrations were small, indicating a relatively uniform external influence on these metals and the likely homogeneity of their concentrations throughout the region.

Table 5.1: Statistics of heavy metal concentrations in soils of the suburbs of Beijing

Heavy metal	Sample count	Minimum (mg/kg)	Maximum (mg/kg)	Mean (mg/kg)	SD (mg/kg)	CV (mg/kg)
Cu	758	11.87	97.22	26.78	9.21	0.34
As	758	3.41	12.82	7.99	1.31	0.16
Cd	758	0.07	1.04	0.20	0.10	0.50
Cr	758	36.38	102.96	58.15	6.74	0.12
Hg	758	0.01	0.83	0.13	0.11	0.86
Ni	758	11.76	31.50	21.22	2.35	0.11
Pb	758	4.81	60.78	22.64	6.51	0.29
Zn	758	37.34	207.47	78.03	17.56	0.23

A significant correlation between the concentrations of Cu and the concentrations of Cd, Pb and Zn were found. The correlation coefficients were large, between 0.53 and 0.89, and highly significant. As and Ni were moderately but highly significantly correlated. In addition, Cr was moderately correlated with Cd and Zn, whereas Hg was only correlated with As and not with the other heavy metals. These results indicate that there were special sources of the Hg pollution.

From PCA, the first principal component (PC1) explained 40.16% of the total variance and was dominated by Cu, Cd, Pb and Zn (Table 5.2). The second principal component (PC2) indicated enrichment for As and Hg, and explained an additional 17% of the total variance; PC3 explained an additional 14% of the total variance and was dominated by Cr, Hg and Ni, with a negative loading for Ni. However, the proportions of soil Hg in F2 and F3 were

similar. There was no significant correlation between the concentration of Cr or Ni with any of the other six elements, indicating a degree of uniqueness.

Table 5.2: All explained variables and factors derived using the varimax rotation method.

Element	Before rotation			After rotation		
	F1	F2	F3	F1	F2	F3
Cu	0.924	0.004	-0.040	0.916	0.125	0.021
As	0.307	0.798	0.026	0.199	0.770	0.315
Cd	0.879	-0.239	0.079	0.893	-0.047	-0.191
Cr	0.286	-0.231	0.562	0.266	0.075	-0.612
Hg	0.060	0.666	0.597	-0.074	0.857	-0.252
Ni	0.183	0.382	-0.689	0.188	0.076	0.782
Pb	0.711	0.020	-0.190	0.716	0.044	0.166
Zn	0.931	-0.103	0.026	0.932	0.059	-0.085

Semivariogram parameters showed that the  $C_0 / (C_0 + C)$  values of the six heavy metals other than Ni and Cr were less than 25% and thus exhibited strong spatial correlation. The  $C_0 / (C_0 + C)$  value for Ni and Cr were 50% and 30.4%, respectively, indicating moderate spatial correlation. The interpolated spatial distribution of the eight heavy metal concentrations plotted with the optimal interpolation model showed that the peak distribution was in the northwest for Cu, Cd, Pb and Zn; in the southeast for As; mainly on the urban fringe for Hg; and mainly in the southwest for Cr and Cd (Figure 5.1).

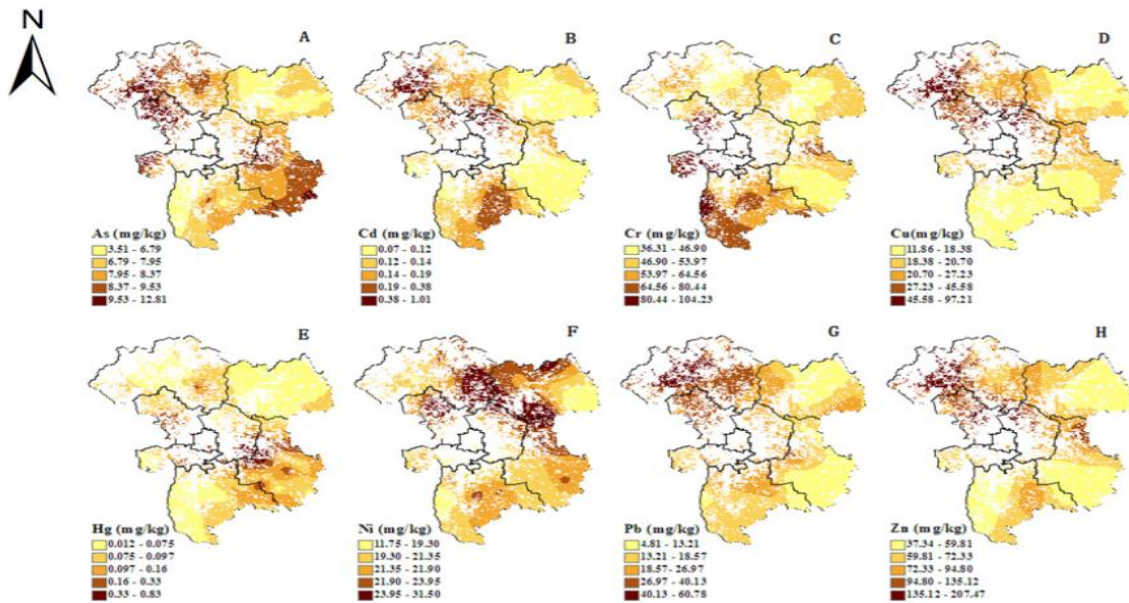


Figure 5.1: Spatial distribution maps of soil heavy metal concentrations



## Case –II: Northeast China

Another study of Li et al. (2012) made analysis on heavy metal contamination of urban soil in an old industrial city (shenyang) in northeast china. For all metals, the total concentrations showed a great degree of variability, indicated by the large CV from 21.99% of Mn to 139.14% of Cd. The elevated coefficients of variation reflected the non-homogeneous distribution of concentrations of anthropogenically emitted heavy metals. Large standard deviations were found in all heavy metals except Hg. This also indicated the wide variation of concentrations in urban soils. The extent of this skew was shown by the differences between the mean and the median values. The difference was greatest for Cd (105%) and Pb (66.6%) and the least for Cr (4.3%). This clearly demonstrated the anthropogenic contribution and revealed the significant pollution levels in the area. The distribution of concentrations was skewed by a small number of large values (contamination hotspots). The concentration of Pb showed a high significant positive relationship with Cu (0.793), Cr (0.363), Zn (0.730), Cd (0.835), As (0.650), and Hg (0.525). Additionally, the correlations between Cu and Pb, Cr, Zn, Mn, Cd, As and Hg were significantly positive. However, the concentrations of Cr and Mn showed very weak correlations with Cd, As and Hg. This indicates that Cr and Mn were from different sources than Cd, As and Hg.

The initial component matrix (Table 5.3) indicated that Pb, Cu, Zn, Cd, As and Hg are closely associated with the first principal component (PC1), while Cr and Mn are mainly distributed with the second component (PC2). The metals in the PC1 mainly come from anthropogenic sources, such as industrial production and traffic activities. Cr and Mn in PC2 are strongly correlated and clearly separate from the other heavy metals regarding their correlation coefficient analysis and PCA. This separation between them and other heavy metals may suggest that Cr and Mn mainly came from non-anthropogenic sources, indicating that they originated from natural sources.

Table 5.3: Matrix of the principal component analysis loadings of heavy metals.

	Pb	Cu	Cr	Zn	Mn	Cd	As	Hg
PC1	0.936	0.844	0.49	0.807	0.362	0.821	0.705	0.638
PC2	-0.134	0.142	0.701	0.006	0.751	-0.222	-0.368	-0.269

The spatial distribution of metal concentrations is a useful aid to assess the possible sources of enrichment and to identify hotspots with high metal concentrations. The estimated maps of Pb, Cu, Cr, Zn, Mn, Cd, As and Hg are presented in Figure 5.2; several hotspots of high metal concentration were identified by the geochemical maps. In these metals, Pb and Cu showed a very similar spatial pattern, with contamination hotspots located simultaneously in the north and east of the study area, indicating that they were from the same sources. This provided a refinement and reconfirmation of the results of the statistical analysis, in which strong associations were found between these two metals.

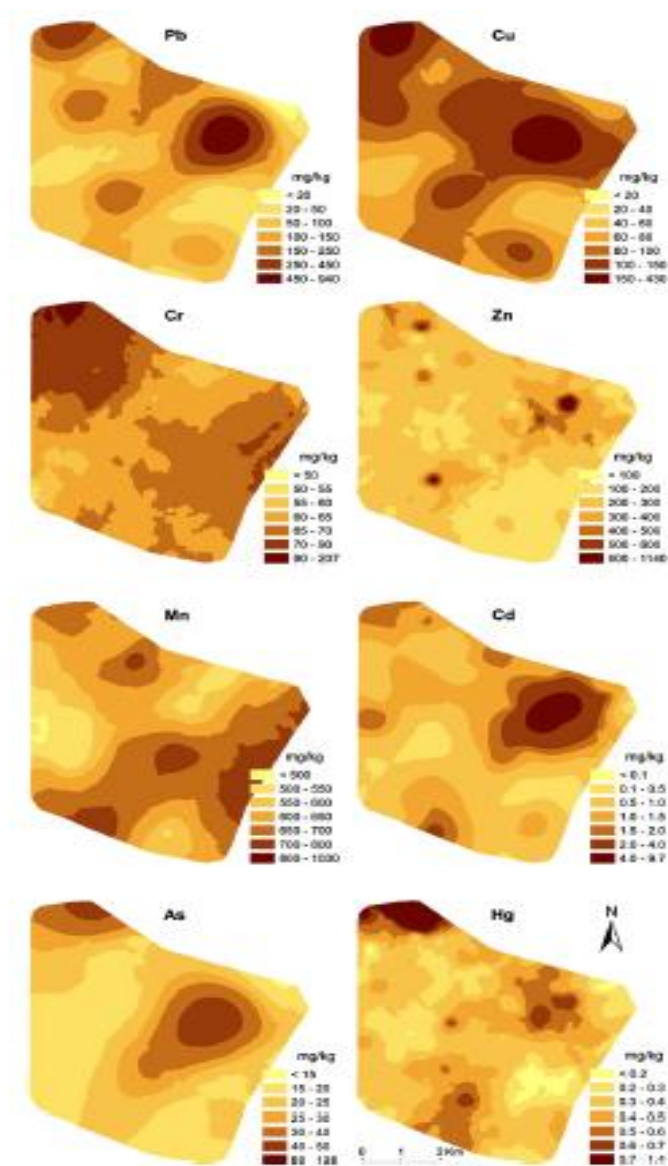


Figure 5.2. Estimated Ordinary Kriging concentration maps for Pb, Cu, Cr, Zn, Mn, Cd, As and Hg (mg/kg).

A study of Yasrebi et al. (2009) on evaluation and comparison of ordinary kriging and IDW methods for prediction of spatial variability of some soil chemical parameters aimed at determining degree of spatial variability of soil chemical properties with Ordinary Kriging (OK) and Inverse Distance Weighting (IDW) methods for 6 soil chemical properties were examined in a fallow land in Bajgah, Iran. The range of spatial dependence was found to vary within soil parameters. Overall, the results obtained from the comparison of the two applied interpolation methods indicated that kriging was the most suitable methods for prediction and mapping the spatial distribution of soil chemical properties in this area. The results showed, IDW with powers of 4 and had almost the same precisions, but IDW with power of 4 was better than IDW with power of 5. Results also revealed that although the IDW is relatively simple and easy to use, but is less accurate than OK. In comparison to OK, IDW increase the error more than 22% for pH, 15% for EC, TN, K, OM and 20% for available P. The results of present study were in agreement to other studies.

### **Context of this Study**

In this study, Results from conventional statistics depicted that the concentration of all studied metal elements showed a great degree of variability due to the generation of metal elements from anthropogenic activities. The greatest SD of Ti indicated that the concentration of Ti were spread out from the mean value of Ti, while, and the smallest SD of Cd indicated Cd were very closed to the mean value of Cd. Results of skewness for Al, Cd, Co, K, Na, Ni, Sc and Ti in soil indicated the data were fairly symmetrical; As, Ba, Ca, Cr, Hg, Sb, Sr, V and Zn were moderately skewed as well as Cu, Fe, Mn and Pb were highly skewed for dry season. Results of Kurtosis revealed that the metal elements of Al, As, Ba, Ca, Cd, Co, Cr, Fe, Hg, K, Na, Ni, Sc, Ti and Zn in soil exhibited platykurtic distribution, whereas, Cu, Mn, Pb, Sb, Sr and V exhibited leptokurtic distribution for dry season. Based on descriptive statistical analysis, the level of metal elements can be ordered as Fe> Al> K> Ca> Ba> Na> Pb> V> Ti > Sr> Zn> Mn> Sc> Cu > Sb> Co> Cr > Hg> As> Ni> Cd in dry season; and Fe> Al>K>Ca>Ba> Na>V>Ti> Sr >Zn> Pb > Mn> Sc >Co> Cu> Sb>Cr>Hg>As>Ni>Cd in rainy season.

Results of Pearson's correlation revealed that most of the significant correlations were observed for Ca and Ba (0.992) in dry season, while, Ti and Sr (0.991) in rainy season.

Almost all the metal elements were strongly correlated with each other indicating these metal elements were derived from the same generation sources.

Results of PCA revealed that Cu, Hg, Mn, Pb, and Zn in soil derived from natural parent materials; Al, As, Ba, Ca, Cd, Co, Fe, K, Na, Ni, Sb, Sc, Sr, Ti and V from anthropogenic activities and Cr from both the natural and anthropogenic sources in dry season. PCA results also demonstrated that As, Cr, Hg, K, Na and Mn in soil derived from natural parent materials; Al, Ba, Ca, Co, Cu, Fe, Ni, Sb, Sc, Sr, Ti and V from anthropogenic activities as well as Cd, Pb and Zn from both the natural parent and anthropogenic sources in rainy season. Results from AHC also proved the same generation sources of metal elements as similar of PCA in soil for both the dry and rainy seasons. Cluster analysis showed similar contamination pattern for almost all the metal elements indicated the same sources of contamination for both the dry and rainy season. This statement was also confirmed from the results of Pearson's correlation analysis.

Geostatistical analysis demonstrated IMQ was the best fitted model almost all the metal elements for least RMSPE; Hole effect model for Al, Ca, Cu, Ni, Pb, V; as well as J-Bessel for As, Ba, Fe, Hg, Mn, Sb, Sc and Sr in soil. In addition, rational quadric, circular and exponential model was the best fitted model for Cd, Na and Ti; Cr as well as K, respectively. Produced prediction surface for all the interpolation techniques showed most of the contaminated hotspots was found near the central point of the disposal site for all studied metal elements. Semivariogram obtained from ordinary kriging interpolation showed strong spatial correlation between Cd, Co, K and Ti in soil; moderate correlation for Al, As, Ba, Cr, Cu, Mn, Na, Ni, Pb, Sb, Sc, Sr and V as well as no spatial correlation between Ca, Fe, Hg and Zn. Based on least value of MAPE, IDW with power 1 to 5 and RBF with different kernel functions showed comparatively more accurate prediction than that of other interpolation techniques. Based on RI, IDW1 showed best performance followed by ordinary kriging. Based on all indices, IDW1 showed the best technique for metal elements in soil considered in this study. Finally, the obtained results in this study were agreed well with the results published by various researchers for similar cases of studies.

## CHAPTER VI

### CONCLUSION AND RECOMMENDATIONS

#### 6.1 Conclusion

Despite the terrible effects of metal contamination in soil of the selected waste disposal site, no proper step is taken to control or reduce the spread of metal elements in and around the soil of disposal site. Therefore, it is necessary to carry out an intensive study and monitor the nature and extent of such metal elements in and around of the disposal site. Based on the findings of the study, the following conclusions have been drawn:

1. From normal QQ plot, it can be concluded that almost all the metal elements in soil for both the dry and rainy seasons were distributed normally, except As.
2. Results from conventional statistics depicted that the concentration of metal elements considered in this study showed a great degree of variability due to the generation of metal elements from anthropogenic activities.
3. Based on descriptive statistical analysis, the level of metal elements can be ordered as Fe> Al> K> Ca> Ba> Na> Pb> V> Ti > Sr> Zn> Mn> Sc> Cu > Sb> Co> Cr > Hg> As> Ni> Cd in dry season; and Fe> Al>K>Ca>Ba> Na>V>Ti> Sr >Zn> Pb > Mn> Sc >Co> Cu> Sb>Cr>Hg>As>Ni>Cd in rainy season.
4. Results of Pearson's correlation revealed that all the metal elements were strongly correlated with each other indicating these metal elements were derived from the same generation sources.
5. The generation sources of all studied metal elements in soil obtained from PCA were completely in agreement with AHC for both the dry and rainy seasons.
6. Cluster analysis showed similar contamination pattern for almost all the metal elements indicated the same sources of contamination for both the dry and rainy season. This statement was also confirmed from the results of Pearson's correlation analysis.
7. Produced prediction surface for all the interpolation techniques showed most of the contaminated hotspots was found near the central point of the disposal site for all studied metal elements.

8. Semivariogram obtained from ordinary kriging interpolation showed strong spatial correlation between Cd, Co, K and Ti in soil; moderate correlation for Al, As, Ba, Cr, Cu, Mn, Na, Ni, Pb, Sb, Sc, Sr and V as well as no spatial correlation between Ca, Fe, Hg and Zn.
9. Based on least value of MAPE, RI and G-value, IDW1 showed best performance followed by ordinary kriging.
10. The predicted values of metal elements in soil from ANN were almost same as obtained from laboratory.

Finally, it can be concluded that the results from the study will also be used to find out the area needed to take under immediate remedial action to remove the metal elements where the levels were too high.

## **6.2 Recommendation for Further Study**

The soil samples should be collected more carefully to obtain better field condition. As the S-W test gives better results for smaller dataset (< 2000). If the number of observation is greater than 2000, the S-W test can be avoided as it gives less reliable result. The updated version of softwares can be used to comprehend better results from further study.

## REFERENCES

- Abdi, H. and Williams, L.J. (2010). PCA, WIREs Comp Stat, Vol. 2, 433–459.
- Aborjona and Paribesh, (<http://www.wasteconcern.org/newsletters/issue5/issue5.html>).
- Agusa, T., Kunito, T., Nakashima, E., Minh, T. B., Tanabe, S., Subramanian, A. and Viet, P. H. (2003). Preliminary on trace element contamination in dumping sites of municipal wastes in India and Vietnam, *J. Phys. IV*, Vol. 107, 21–24.
- Ahmed, F., Fakhruddin, A.N.M. and Imam, M.D.T. (2016). Spatial distribution and source identification of heavy metal pollution in roadside surface soil: a study of Dhaka. *Ecol Process*, Vol. 5:2, 85-96.
- Alamgir, M., Ahsan, A., McDonald, C.P., Upreti, B.N. and Rafizul, I. M. (2005). Present status of municipal solid wastes management in Bangladesh, *Waste-The Social Context*, Canada, Vol.16:1-10.
- Alkaiem, L., Keller, F. and Sternberg H. (2016). Analysis of inclination measurement by means of artificial neural networks – A comparison of static and dynamic networks. D-20457 Hamburg, Germany, Peer JISDM review paper Vienna.
- Anderberg, M.R. (1973). *Cluster Analysis for Applications*. Academic Press, New York.
- Anderson, T.W. and Darling, D.A. (1954). A Test of Goodness of Fit, *J. of the American Statistical Association*, 49: 765–769.
- Asante-Duah, K. (1998). Public Health Risk Assessment for Human Exposure to Chemicals, *Environ. Pollu.* Vol. 27, pp-55-62.
- Ashraf, M.Y., Rafique, N., Ashraf, M., Azhar, N., and Marchand, M. (2013). Effect of supplemental potassium (K<sup>+</sup>) on growth, physiological and biochemical attributes of wheat grown under saline conditions. *J. of Plant Nutrition*, 36: 443-458.
- Bai, E., Li, S., Xu, W., Li, W., Dai, W. and Jiang, P.A. (2013). Meta-analysis of experimental warming effects on terrestrial nitrogen pools and dynamics, *New Phytol.*, Vol.199, pp.441-451.
- Bhuiyan, M.A.H., Doza, M.B. and Islam, A.R.M.T. (2016). Assessment of groundwater quality of Lakshimpur district of Bangladesh using water quality indices, geostatistical methods and multivariate analysis, *Environ. Earth Sci.*, 2016, Vol 75(12), pp. 1
- Brovelli, M.A., Carcano, L. and Minghini, M. (2011). ArcGIS exercise 4- Spatial Interpolation GIS Exercise, Laboratorio di Geomatica, Politecnico di Milano-Polo Regionale di Como via valleggio.
- Burgess, T.M. and Webster, R. (1980). Optimal interpolation and arithmetic mapping of soil properties. The semivariogram and punctual kriging, *J. Soil. Sci.*, 31: 315-331.
- Cambardella, C.A., Moorman, T.B., Novak, J.M., Parkin, T.B., Karlen, D.L., Turco R.F. and Konopka, A.E. (1994). Field-scale variability of soil properties in central Iowa soils. *Soil Sci. Soc. Am. J.*, Vol. 58 (2),: pp. 150-1511.
- Candeias, C., Silva, E.F.D., Salgueiro, A.R., Pereira, H.G., Reis, A.P., Patinha, C., Matos, J. X. and Vila, P.H.A. (2011). The use of multivariate statistical analysis of geochemical data for assessing the spatial distribution of soil contamination by potentially toxic elements in the Aljustrel mining area), *Environ Earth Sci*, 62:1461–1479.
- Chary, N.S., Kamala, C.T. and Raj, D.S.S. (2008). Assessing risk of heavy metals from consuming food grown on sewage irrigated soils and food chain transfer, *Ecotoxicol Environ Saf.*, 69: 513–524.
- Chen, C., Ma, Y., Wan, J. and Huang, M. (2012). Effects of C/N ratio on physiological biochemical characteristics of anaerobic granular sludge, *J. of Environ Sci*, 32: 478–484.

- Conover, W.J. (1999). *Practical Nonparametric Statistical*, 3rd Ed., John Wiley & Sons, Inc. New York. pp.428-433.
- Creutin, J. D. and Obled, C. (1982). Objective analysis and mapping techniques for rainfall fields: An objective comparison, *Water Resour. Res.*, 18 (2): 413-431.
- Daniel, D.E. and Koerner, R.M. (1995). *Waste Containment Facilities, Guidance for Construction Quality Assurance and Quality Control of Liner and Cover System*, ASCE, Vol.13, pp.25-27.
- Debnath, B. (2015). The Present Scenario of Solid Waste Disposal and Management Practice in Chittagong City Corporation, ICMER, CUET, pp. 35-42.
- Ehab, A.I. and Ahmed, E.A. (2015). Effect of Soil Amendments on Growth, Seed Yield and NPK Content of Bottle Gourd Grown in Clayey Soil, *Int. J. of Soil Sci.*, 10: 186-194.
- Einax, J. W. and Soldt, U. (1995). Geostatistical investigations of polluted soils, Fresenius, *Journal of Analytical Chemistry*, Vol. 351, pp. 48–53.
- Environment and Climate Change : Canada,  
(<https://www.ec.gc.ca/eauwater/default.asp?lang=En&n=6A7FB7>).
- EPA (1984). Environmental Protection Agency, Rules and Regulations, Federal Register 49: No. 209.
- ESRI (2001). *ArcGIS Geostatistical Analyst: Statistical Tools for Data Exploration, Modeling and Advanced Surface Generation*. ESRI: Redlands, California, pp. 19.
- Facchinelli, A., Sacchi, E. and Mallen, L.(2001). Multivariate statistical and GIS-based approach to identify heavy metal sources in soils, *Enviro. Pollu.*, 114(3): 313-324.
- Heckman, J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models, *Annals of Econo.and Social Measurement*, 5, pp. 475-92.
- Engineering Statistics Handbook  
(<http://www.itl.nist.gov/div898/handbook/eda/section3/qqplot.html>).
- Heavy metal treatment project,  
(<http://www.sagasikikankyo.co.jp/en/jukinzoku/what/index.html>).
- <https://jessicamudditt.com/2011/02/04/dhakas-largest-waste-site-photographs-of-matuail-landfill/>.
- [http://earthsci.org/basicgeol/solid\\_waste/solid\\_waste.html](http://earthsci.org/basicgeol/solid_waste/solid_waste.html).
- Hu, H.W., Zhang L.M., Dai, Y., Di, H.J. and He, J. Z. (2013). pH-dependent distribution of soil ammonia oxidizers across a large geographical scale as revealed by high-throughput pyrosequencing, *J. Soils Sediments* Vol.13, pp. 1439–1449.
- Isaake, E.H. and Srivastava, R.M. (1989). *An Introduction to Applied Geostatistics*, Oxford University Press, New York.
- Islam, M.R., Alamgir, M., Bidlingmaier, W., Stepniewski, W. and Rafizul, I.M. (2011). Field Performance of Compacted Clay Liner used in a Pilot Scale Sanitary Landfill in Bangladesh, *Proceedings of the WasteSafe 2009*, Khulna, Bangladesh, pp. 431-440.
- Jia, L., W. Wang, Y. Li and Yang, L. (2010). Heavy metals in soil and crops of an intensively farmed area: A case study in Yucheng City, Shandong Province, China, *International Journal of Environmental Research and Public Health*, Vol. 7 (2), pp. 395-412.
- Johnston, K., Hoef, J.M.V., Krivoruchko K. and Lucas, N. (2001). *Using ArcGIS Geostatistical Analyst*. ESRI, 380 NY, USA.
- Kadar, I. (1995). Effect of heavy metal load on soil and crop, *Act. Agro. Hung.*, 43: pp. 3-9.
- Keesstra, S.D., Geissen, V., van Schaik, L., Mosse, K. and Piirainen, S. (2012). Soil as a filter for groundwater quality, *Current Opinion in Environmental Sustainability*, 4: 507–516.



- Kerry, R. and Oliver, M.A. (2004). Average variograms to guide soil sampling, *Int. J. Applied Earth Observation and Geoinformation*, Vol.5 (3): pp. 307-325.
- Kohonen, T. (2001). *Self-Organizing Maps*. Berlin, Springer.
- Kohonen, T. (2014). *MATLAB Implementations and Applications of the Self-Organizing Map* Unigrafia Oy, Helsinki, Finland, pp.222-223.
- Lee, C.S.L., Li, X., Shi, W., Cheung, S.C.N. and Thornton, I. (2006). Metal contamination in urban, suburban, and country park soils of Hong Kong: a study based on GIS and multivariate statistics. *Sci. Total. Environ.* 356, pp. 45–61.
- Li, J., Xie, Z.M., Xu, J.M. and Sun, Y.F. (2006). Risk assessment for safety of soils and vegetables around a lead/zinc mine, *Environ Geochem Health*, Vol. 28, pp. 37–44.
- Liu, W.X., Li, X.D., Shen, Z.G., Wang, D.C., Wai, O.W.H. and Li, Y.S. (2003). Multivariate statistical study of heavy metal enrichment in sediments of the Pearl River Estuary, *Environ Pollut*, 121:377–88.
- Lu, A., Wang, J., Qin, X., Wang, K., Ha, P. and Zhang S. (2012). Multivariate and geostatistical analyses of the spatial distribution and origin of heavy metals in the agricultural soils in Shunyi, Beijing, China, *Sci. of the Total Environ*, Vol. 425, pp. 66 – 74.
- Mahmoudabadi, E., Sarmadian, F., Savaghebi, G.H. and Alijani, Z. (2012). Accuracy Assessment of Geostatistical Methods for Zoning of heavy metals in soils of urban-industrial areas. *Inter. Research J. of Applied and Basic Sci.*, Vol., 3 (5), pp. 991-999.
- Marina, I., Margherita, L., Angela, N. and Alfredo, P. (2003). Toxicity identification evaluation of leachates from MSW landfills: a multispecies approach, *Chemo.*, Vol. 52, pp. 85–94.
- McLean, J. E. and Bledsoe, B. E.(1992). Behavior of metals in soil. U.S. EPA/540/S-92/018.
- Mehidi, S.H., Chakrabarty, N., Mohiuddin, H.M. (2014). An Application of ANN Process to Assess Risk in Cement Industries in Bangladesh, *Indus. Engg. & Manag.*, 3: 4.
- Mireles, A., Solís, C., Andrade, E., Lagunas-Solar M., Pinã, C. and Flocchini, RG. (2004). Heavy metal accumulation in plants and soil irrigated with wastewater from Mexico City, *NIM Physics Response B Beam Interact Mater Atoms*, pp. 219– 220.
- Misra, S. and Mani, D. (2009) *Soil Pollution*, APH Publishing Corp., New Delhi, pp. 29-59.
- Murtaza, M.G. (2012). *Solid Waste Management in Khulna City*, Plan Plus,1(1): 6-15.
- Nabulo, G., Oryem Origa, H., Nasinyama, G.W. and Cole, D. (2008). Assessment of Zn, Cu, Pb and Ni contamination in wetland soils and plants in the Lake Victoria basin, *Int. J. Environ. Sci. Tech.*, Vol 5 (1), pp. 65-74.
- Neuner, H. (2012). Model selection for system identification by means of artificial neural networks, *Int. J. of Applied Geodesy*. Vol. 6, pages 117–124.
- Nielsen, D.R. and Wendroth, O. (2003). Spatial and temporal statistics: sampling field soils and their vegetation, pp. 398.
- Nriagu, J.O. and Pacyna, J.M. (1988). Quantitative assessment of worldwide contamination of air, water and soils by trace metals, *Nature* Vol. 333, pp.134–139.
- Olawoyin, R., Oyewole, S.A. and Grayson, R. L.(2012). Potential risk effect from elevated levels of soil heavy metals on human health in the Niger delta, *Ecotoxicol. Environ. Saf.*, Vol.85(1), pp. 120–130.
- Omran, E.S. and El Razek, A. (2012). Mapping and screening risk assessment of heavy metal concentrations in soils of the Bahr El-Baker Region, Egypt, *J. of Soil Sci. and Environ. Manag.*, Vol. 6(7), pp. 182–195.
- Oyeku, O.T. and Eludoyin, A.O. (2010). Heavy metal contamination of groundwater resource in a Nigerian urban settlement, *Afr. J. Environ. Sci. Technol.*, 4(4): 201–214.

- Pebesma, E.J., Jong, K. and Briggs D.J. (2007). Visualizing uncertain spatio-temporal data under different scenarios: air quality example. *International J. of GIS*, 21 (5): 515-527.
- Pugh, M. (1999). The Path to Affordable Landfills. *Waste Management*, 58-59.
- Qasim, S.R. and Chiang, W. (1994). *Sanitary Landfill Leachate: Generation, Control and treatment*, Technomic Publishing, Lancaster, pp. 1153-1158.
- Rafizul, I.M., Alamgir, M., Islam and M.M. (2011). Evaluation of Contamination Potential of Sanitary lysimeter Using Leachate Pollution Index. *Proceedings Sardinia 2011, 13th International Waste Management and Landfill Symposium S. Italy*.
- Rafizul, I.M., Howlader, M.K. and Alamgir, M. (2012). Construction and evaluation of simulated pilot scale landfill lysimeter in Bangladesh, *Waste Management, ScienceDirect, ELSEVIER*, Vol. 32 (11), pp. 2068-2079.
- Razali, N. and Wah, Y.B. (2011). Power comparisons of Shapiro – Wilk, Kolmogorov–Smirnov, Lilliefors and Anderson–Darling tests, *J. of Stat. Model. and Anal.*, 2 (1): 21–33.
- Reza, S.K., Sarkari, D., Daruah, U. and Das, T.H. (2010). Evaluation and comparison of ordinary kriging and inverse distance weighting methods for prediction of spatial variability of some chemical parameters of Dhalai district, Tripura. *Agropedology*, Vol. 20(1), pp. 38-48.
- Rogers, G. and Nicewander, D. (1988). Thirteen Ways to Look at the Correlation Coefficient, *The American Statistician*, 42 (1): 59–66.
- Rooki, R., Doulati Ardejani, F., Aryafar, A. and Bani Asadi, A. (2011). Prediction of heavy metals in acid mine drainage using artificial neural network from the Shur River of the Sarcheshmeh porphyry copper mine, Southeast Iran. *Environ Earth Sci*, Vol.55:pp. 39-47.
- Ross, S. (1994). Toxic metals in Soil-Plant Systems, *Environ. Pollu.*, Vol.155 , pp. 254-261.
- Royston, P. (1991). Estimating departure from normality, *Stat Med.*, 10(8):1283–93.
- Samiul, F.A. (2016). Solid waste management system in Dhaka city of Bangladesh, *J. of Modern Science and Technology* Vol. 4. No. 1. September 2016 Issue. pp. 192 – 209.
- Shahin, M., Maierm and H., Jaksa, M. (2004). Data Division for Developing Neural Networks Applied to Geotechnical Engineering, *J. Comput. Civ. Eng.*, ASCE, Vol. 105, pp. 105-114.
- Shapiro, S.S. and Wilk, M.B. (1965). An analysis of variance test for normality, *Biometrika*, 52 : 591-611.
- Silipo, R., Berthold, M. and Hand, D. J. (1999). *Neural Networks. Intelligent Data Analysis: An Introduction*. Springer-Verlag New York, Inc.: 432.
- Smith, M. J. d., Goodchild, M.F. and Longley, P.A. (2007). *Geospatial Analysis: A Comprehensive Guide to Principles. Techniques and Software Tools*, 2nd ed., Troubador Publishing Ltd.
- Stuart, A. and Kendall's J.K. Ord. (1987). *Advanced Theory of Statistics: Vol. 1, Distribution Theory*. Oxford University Press, New York.
- St. Mary's Country: Maryland, <http://www.co.saint-marys.md.us/dpw/recycleoverview.asp>.
- Tahir, N.M., Chee, P.S. and Maisarah, J. (2007). Determination of heavy metals content in soils and indoor dusts from nurseries in dungun, *The Malay. J. Analy. Sci.*, Vol. 11, pp. 280-286.
- Terzano, R., Spagnuolo, M., Vekemans, B., De Nolf, W., Janssens, K., Falkenberg, G., (2007). Assessing the origin and fate of Cr, Ni, Cu, Zn, Pb, and V in industrial polluted soil by combined microspectroscopic techniques and bulk extraction methods. *Environ. Sci. & Techn.*, 41(19): 6762–6769.

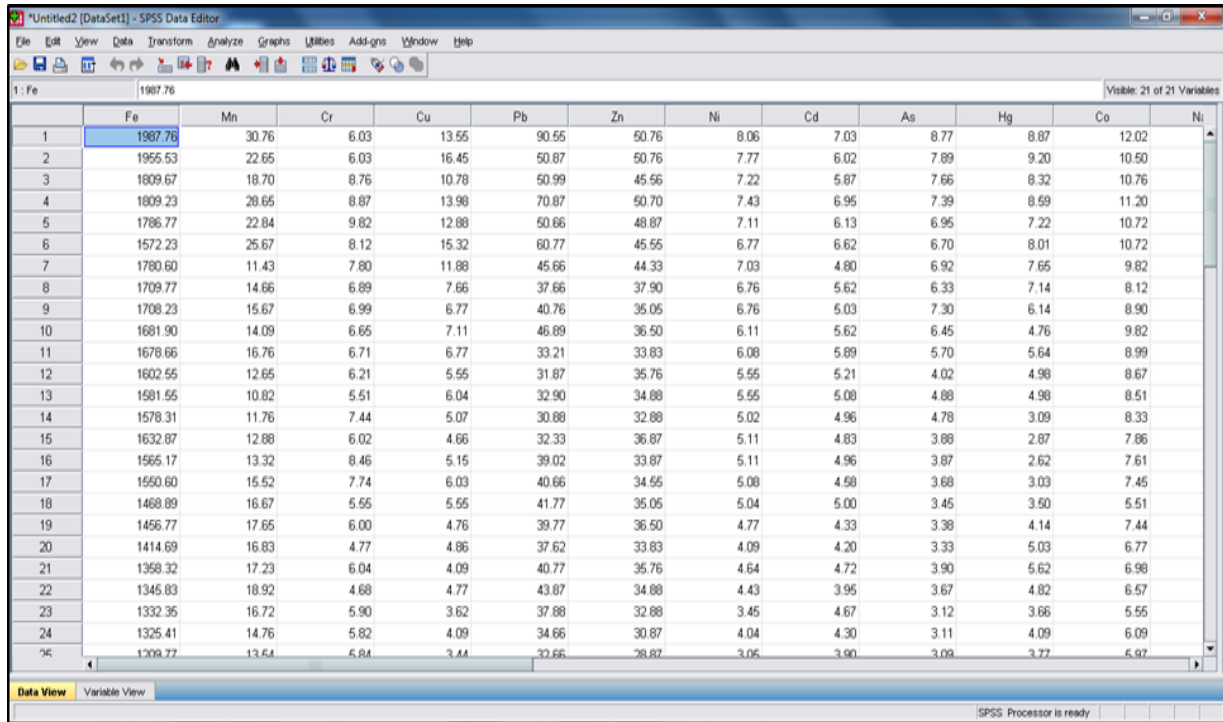
- UN Habitat (2010). Solid Waste Management in the World's Cities Water and Sanitation in the World's Cities, UN Human Settlements Programme, London, Washington, DC.
- Vaalgamaa, S. and Conley, D.J. (2008). Detecting Environmental Change in Estuaries; Nutrient and Heavy Metal Distributions in Sediment Cores in Estuaries from Gulf of Finland, Baltic Sea. *Estuarine, Coastal and Shelf Science*, 76 (1): 45-56.
- Wang, H.Q., Zhao, Q., Zeng, D.H., Hu, Y.L., and Yu, Z.Y. (2015). Remediation of a magnesium-contaminated soil by chemical amendments and leaching, *Land Degrad. Dev.*, 26: 613–619.
- Xiaoyu, L., Lijuan, L., Yugang, W., Geping, L., Xi, C., Xiaoliang, Y., Myrna, H.P.H., Ruichao, G., Houjun, W., Jiehua, C. and Xingyuan H. (2013). Heavy metal contamination of urban soil in an old industrial city (Shenyang) in Northeast China, *Geoderma*, Vol. 192, pp. 50–58.
- Xie, Y., Chen, T., Lei, M., Yang, J., Guo, Q., Song, B. and Zhou X. (2011). Spatial distribution of soil heavy metal pollution estimated by different interpolation methods: Accuracy and uncertainty analysis, *Chemosphere*, 82: 468–476.
- Yan, W., Mahmood, Q., Peng, Da., fu, W., Chen, T., Wang, Y., Li, S., Chen, J. and Liu, D. (2015). The spatial distribution pattern of heavy metals and risk assessment of moso bamboo forest soil around Lead-Zinc mine in Southeastern China. *Soil and Tillage Research*, pp. 153.
- Yang, J., Meng, X.Z., Duan, Y.P., Liu, L.Z., Chen, L. and Cheng, H.F. (2013). Spatial distributions and sources of heavy metals in sediment from public park in Shanghai, the Yangtze River Delta. *Appl. Geochem.* 44: 54–60.
- Yao, X., Fu, B., Lu, Y., Sun, F. and Wang, S. (2013) Comparison of Four Spatial Interpolation Methods for Estimating Soil Moisture in a Complex Terrain Catchment. *PLoS ONE* 8(1): e54660. doi:10.1371/journal.pone.0054660.
- Yasrebi, J., Saffari M., Fathi, H., Karimian, N., Moazallahi M. and Gazni R. (2009). Evaluation and Comparison of Ordinary Kriging and Inverse Distance Weighting Methods for Prediction of Spatial Variability of Some Soil Chemical Parameters, *Research Journal of Biological Sciences*, Vol. 4 (1): 93-102.
- Yun, L. and Uchimura, K., (2007). Using self-organizing map for road network extraction from IKONOS Imagery, *International Journal of Innovative Computing, Information and Control*, 3(3): 641–656.
- Zamani, A.A., Yaftian, M.R. and Parizanganeh, A.H. (2012). Multivariate statistical assessment of heavy metal pollution sources of ground water around a lead and zinc plant, *Journal of Environmental Health Science and Engineering*, Vol. 9: 1-10.
- Zhang, J. (2012). Graph degree linkage: Agglomerative clustering on a directed graph, 12th European Conference on Computer Vision, Florence, Italy.
- Zhang, Y. (2011). Introduction to Geostatistics, Course Notes, Dept. of Geology & Geophysics, University of Wyoming.
- Zhuang, P., Mcbride, B.B., Xia, H.P., Li, N.Y. and Li, Z.A. (2009). Health risk from heavy metals via consumption of food crops in the vicinity of Dabaoshan mine, China, *Sci. Total Env.*, Vol. 407:1551–61.
- Zou, J., Dai, W., Gong, S. and Ma, Z. (2015). Analysis of Spatial Variations and Sources of Heavy Metals in Farmland Soils of Beijing Suburbs, Vol. 10(2): e0118082.

# Annex-A

## Screenshots of steps of all performed analysis

### Steps of S-W and K-S test

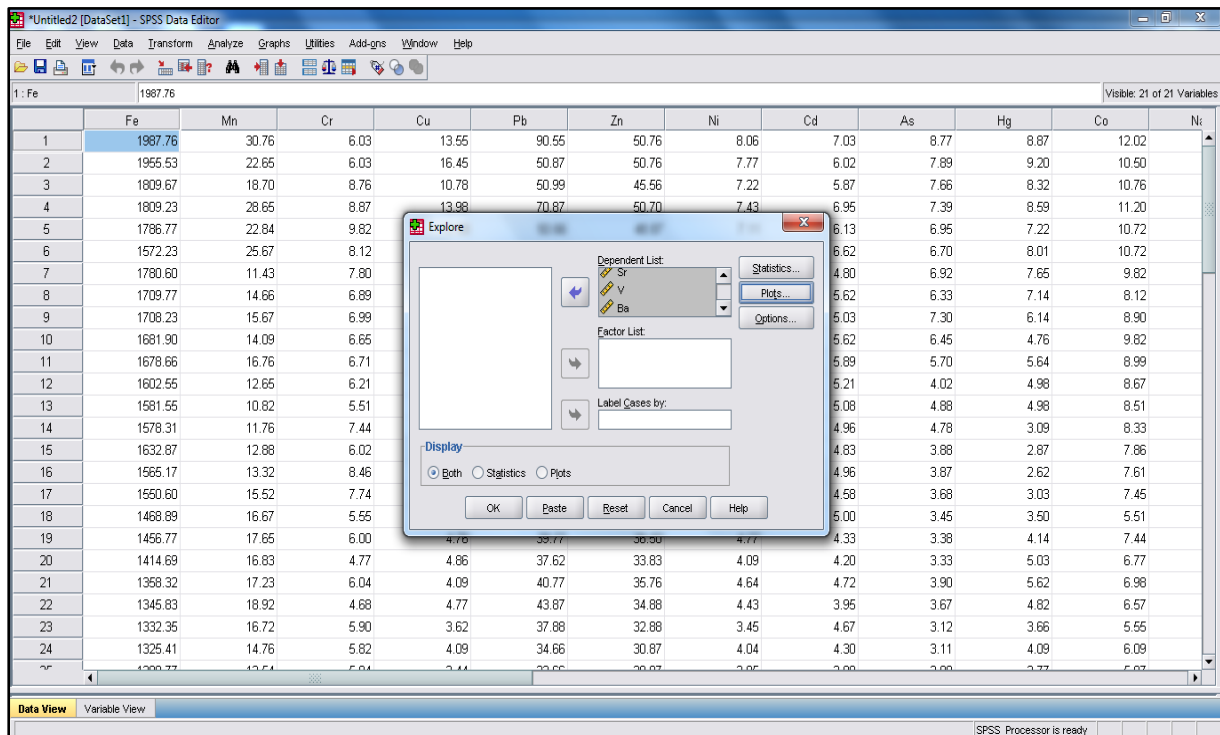
Step 1: The “Data View” of SPSS window represented the data required for the analysis.



The screenshot shows the SPSS Data Editor window with a data table. The table has 24 rows and 14 columns. The first column is labeled '1: Fe' and the second column is labeled '1987.76'. The other columns are labeled with metal symbols: Mn, Cr, Cu, Pb, Zn, Ni, Cd, As, Hg, Co, and Ni. The data values are numerical, representing concentrations of these metals in various samples.

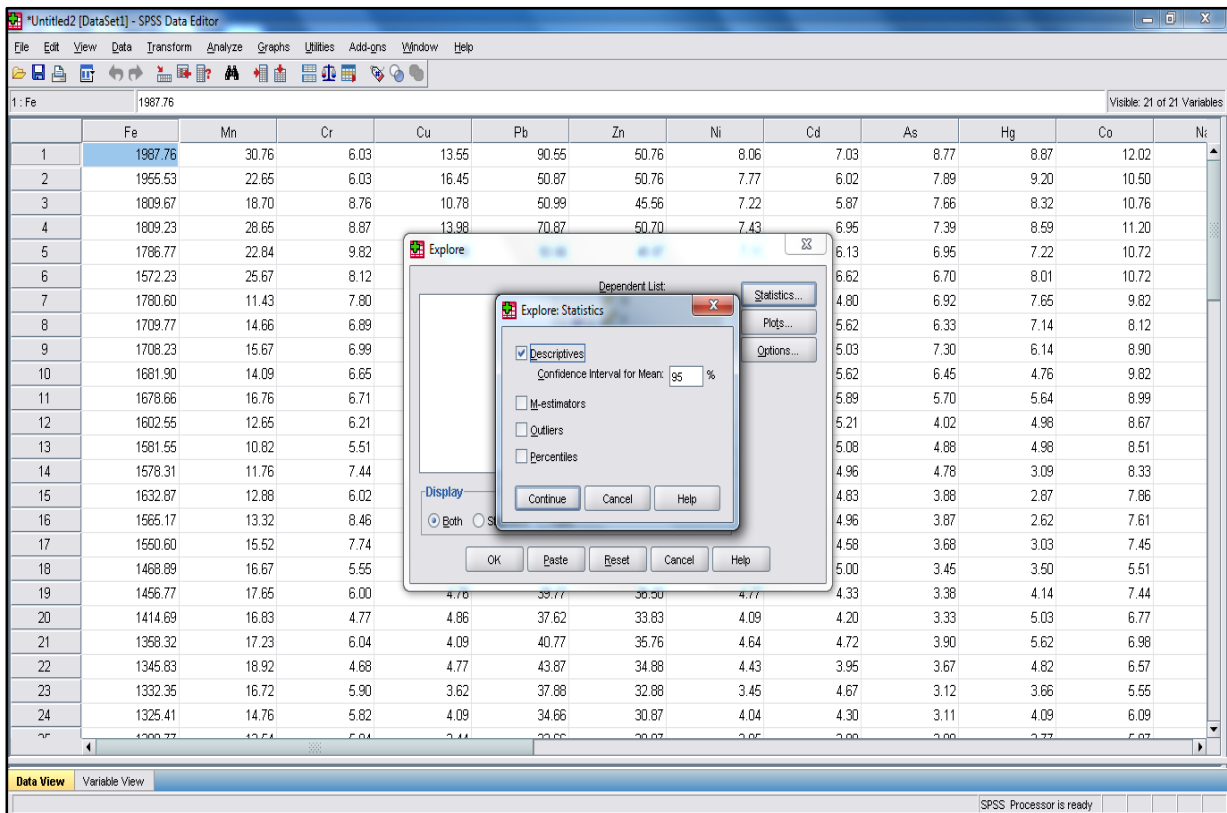
	Fe	Mn	Cr	Cu	Pb	Zn	Ni	Cd	As	Hg	Co	Ni
1	1987.76	30.76	6.03	13.55	90.55	50.76	8.06	7.03	8.77	8.87	12.02	
2	1955.53	22.65	6.03	16.45	50.87	50.76	7.77	6.02	7.89	9.20	10.50	
3	1809.67	18.70	8.76	10.78	50.99	45.56	7.22	5.87	7.66	8.32	10.76	
4	1809.23	28.65	8.87	13.98	70.87	50.70	7.43	6.95	7.39	8.59	11.20	
5	1786.77	22.84	9.82	12.88	50.66	48.87	7.11	6.13	6.95	7.22	10.72	
6	1572.23	25.67	8.12	15.32	60.77	45.55	6.77	6.62	6.70	8.01	10.72	
7	1780.60	11.43	7.80	11.88	45.66	44.33	7.03	4.80	6.92	7.65	9.82	
8	1709.77	14.66	6.89	7.66	37.66	37.90	6.76	5.62	6.33	7.14	8.12	
9	1708.23	15.67	6.99	6.99	40.76	35.05	6.76	5.03	7.30	6.14	8.90	
10	1681.90	14.09	6.65	7.11	46.89	36.50	6.11	5.62	6.45	4.76	9.82	
11	1678.66	16.76	6.71	6.77	33.21	33.83	6.08	5.89	5.70	5.64	8.99	
12	1602.55	12.65	6.21	5.55	31.87	35.76	5.55	5.21	4.02	4.98	8.67	
13	1581.55	10.82	5.51	6.04	32.90	34.88	5.55	5.08	4.88	4.98	8.51	
14	1578.31	11.76	7.44	5.07	30.88	32.88	5.02	4.96	4.78	3.09	8.33	
15	1632.87	12.88	6.02	4.66	32.33	36.87	5.11	4.83	3.88	2.87	7.86	
16	1565.17	13.32	8.46	5.15	39.02	33.87	5.11	4.96	3.87	2.62	7.61	
17	1550.60	15.52	7.74	6.03	40.66	34.55	5.08	4.58	3.68	3.03	7.45	
18	1468.89	16.67	5.55	5.55	41.77	35.05	5.04	5.00	3.45	3.50	5.51	
19	1456.77	17.65	6.00	4.76	39.77	36.50	4.77	4.33	3.38	4.14	7.44	
20	1414.69	16.83	4.77	4.86	37.62	33.83	4.09	4.20	3.33	5.03	6.77	
21	1358.32	17.23	6.04	4.09	40.77	35.76	4.64	4.72	3.90	5.62	6.98	
22	1345.83	18.92	4.68	4.77	43.87	34.88	4.43	3.95	3.67	4.82	6.57	
23	1332.35	16.72	5.90	3.62	37.88	32.88	3.45	4.67	3.12	3.66	5.55	
24	1325.41	14.76	5.82	4.09	34.66	30.87	4.04	4.30	3.11	4.09	6.09	
25	1209.77	13.54	5.84	3.44	32.66	28.87	3.05	3.90	3.09	3.77	5.97	

Step 2: “Descriptive Statistics” under “Analysis” dialog box was selected. Then the dialog box of “Explore” was selected. All metal elements were selected to analyze normality test.

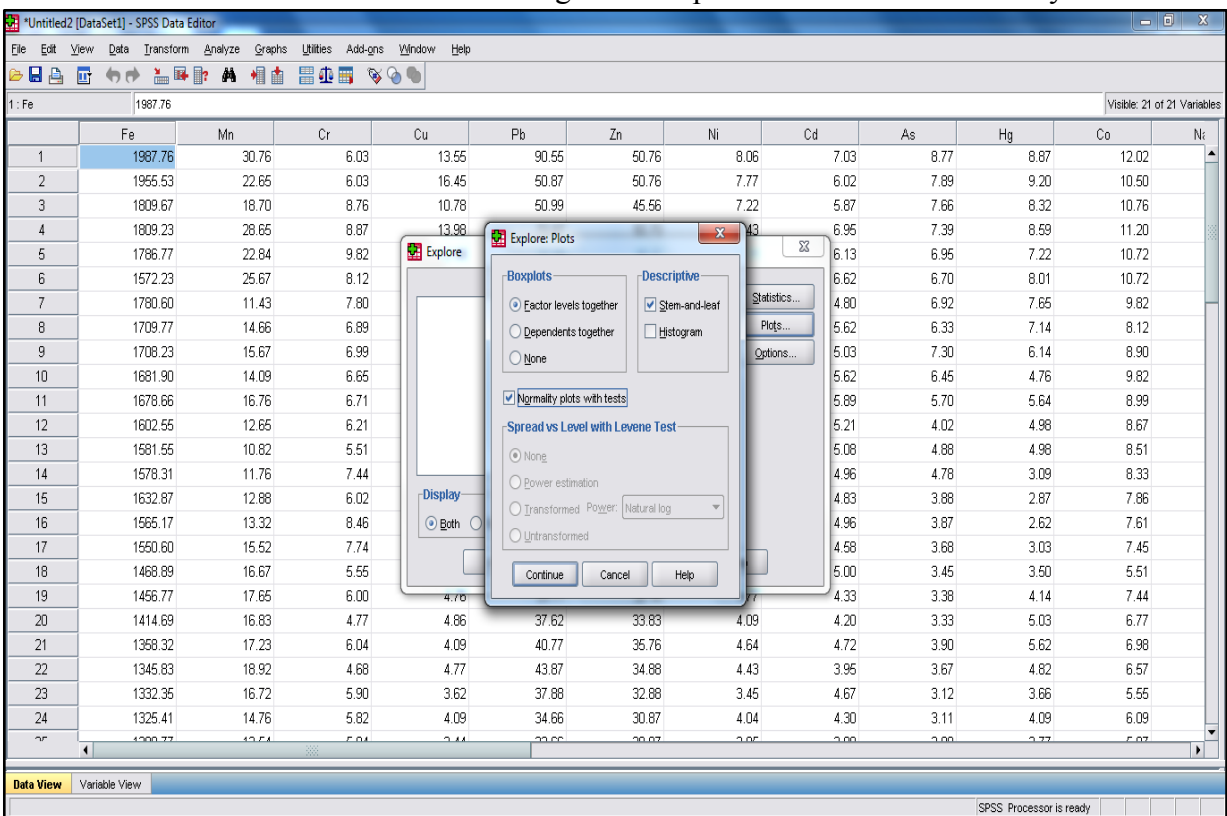


The screenshot shows the same SPSS Data Editor window as in Step 1, but with the 'Explore' dialog box open in the foreground. The dialog box is titled 'Explore' and has several sections: 'Dependent List' (containing 'Sr'), 'Factor List' (empty), and 'Label Cases by:' (empty). There are also 'Statistics...', 'Plots...', and 'Options...' buttons. At the bottom, there are radio buttons for 'Display' with options 'Both', 'Statistics', and 'Plots'. The 'Both' option is selected. The background data table is partially visible behind the dialog box.

Step 3: The dialog box “Descriptive Statistics” was selected to analyze descriptive data for 95% confidence level for mean. “Continue” in dialog box was pressed to continue analysis.



Step 4: “Explore Plot” under the “Analysis” dialog box was selected. “Normality plots with tests” was chosen. “Continue” in the dialog box was pressed to continue the analysis.



Step 5: SPSS output file represented the result originated from SPSS analysis.

<b>Tests of Normality</b>						
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Fe	.105	60	.095	.935	60	.003
Mn	.147	60	.002	.943	60	.008
Cr	.130	60	.013	.944	60	.008
Cu	.215	60	.000	.792	60	.000
Pb	.105	60	.096	.892	60	.000
Zn	.094	60	.200 <sup>*</sup>	.950	60	.015
Ni	.105	60	.099	.961	60	.051
Cd	.095	60	.200 <sup>*</sup>	.967	60	.108
As	.153	60	.001	.883	60	.000
Hg	.100	60	.200 <sup>*</sup>	.925	60	.001
Co	.108	60	.081	.973	60	.198
Na	.154	60	.001	.880	60	.000
K	.147	60	.002	.931	60	.002
Ca	.141	60	.005	.944	60	.008
Al	.086	60	.200 <sup>*</sup>	.933	60	.003
Ti	.106	60	.093	.959	60	.044
Sb	.073	60	.200 <sup>*</sup>	.975	60	.248
Sc	.084	60	.200 <sup>*</sup>	.985	60	.696
Sr	.079	60	.200 <sup>*</sup>	.961	60	.054
V	.103	60	.184	.972	60	.176
Ba	.148	60	.002	.936	60	.003

a. Lilliefors Significance Correction

\*. This is a lower bound of the true significance.

Figure A.1: Stepwise analysis procedure for K-S and S-W test.

## Steps of Normal QQ Plot

Step 1: The data needed to analyze was displayed in the excel sheet.

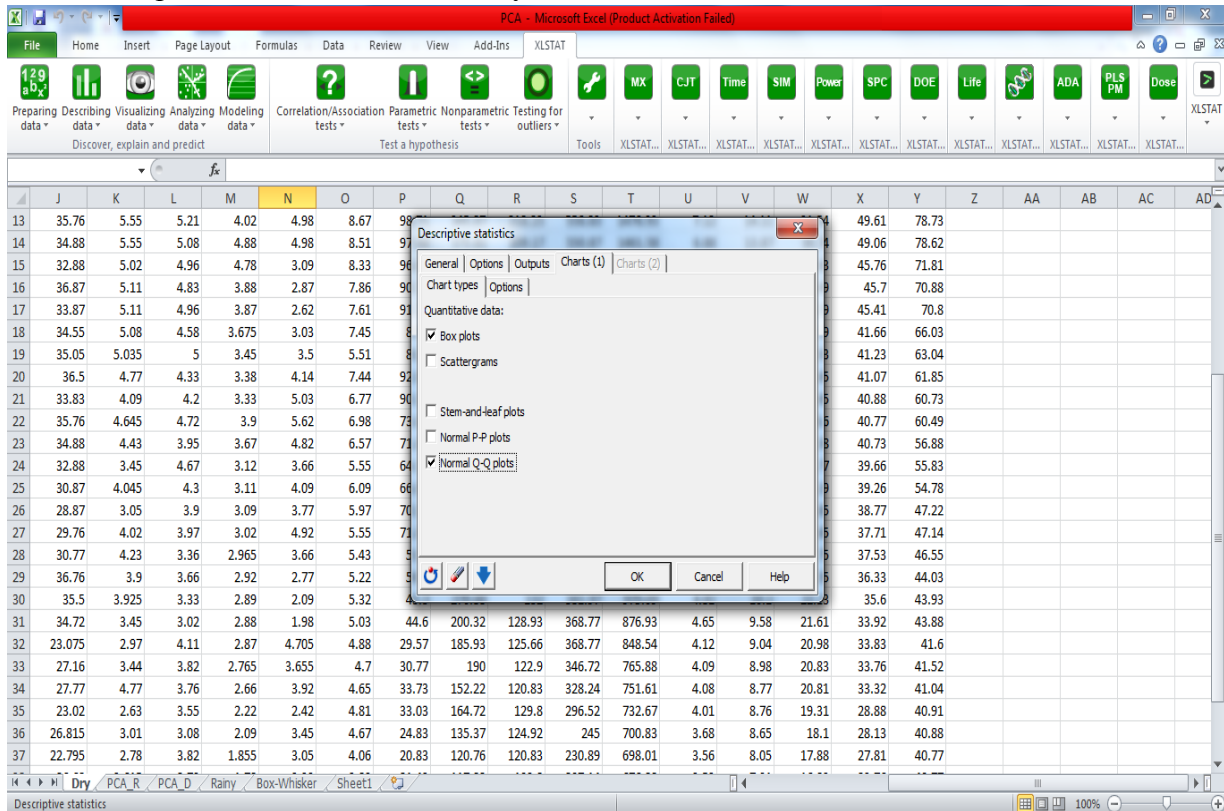
Borehole	Fe	Mn	Cr	Cu	Pb	Zn	Ni	Cd	As	Hg	Co	Na	K	Ca	Al	Ti	Sb	Sc	Sr	V
BH 1	1987.76	30.76	6.03	13.55	90.55	50.76	8.06	7.03	8.77	8.87	12.02	100.54	460.33	318.00	874.78	1937.36	12.55	20.41	54.12	83.3
BH 2	1955.53	22.65	6.03	16.45	50.87	50.76	7.77	6.02	7.89	9.2	10.5	99.20	453.55	302.82	874.50	1915.90	10.77	17.64	40.77	71.0
BH 3	1809.67	18.7	8.76	10.78	50.99	45.56	7.225	5.87	7.66	8.32	10.76	88.76	431.73	299.46	800.87	1899.29	9.69	17.04	40.74	70.7
BH 4	1809.23	28.65	8.87	13.98	70.87	50.7	7.43	6.95	7.39	8.59	11.2	92.82	419.85	272.93	777.74	1809.77	8.87	16.87	38.88	67.9
BH 5	1786.77	22.84	9.82	12.88	50.66	48.87	7.11	6.13	6.95	7.22	10.72	98.76	405.34	262.34	767.43	1765.91	8.87	16.83	37.75	65.8
BH 6	1572.23	25.67	8.12	15.32	60.77	45.55	6.77	6.62	6.7	8.01	10.72	98.71	413.51	270.77	765.77	1679.03	7.99	15.93	36.70	60.2
BH 7	1780.60	11.43	7.8	11.88	45.66	44.33	7.03	4.80	6.92	7.65	9.82	97.66	402.77	266.50	761.43	1675.65	7.95	15.88	36.16	56.3
BH 8	1709.77	14.66	6.89	7.66	37.66	37.9	6.765	5.62	6.33	7.14	8.12	88.00	387.29	256.76	740.00	1673.67	7.85	15.88	35.83	55.2
BH 9	1708.23	15.67	6.99	6.77	40.76	35.05	6.76	5.03	7.3	6.14	8.9	88.87	365.88	240.77	707.98	1667.74	7.77	14.77	35.77	50.8
BH 10	1681.90	14.09	6.65	7.11	46.89	36.5	6.11	5.62	6.45	4.76	9.82	87.34	355.38	238.65	657.88	1566.17	7.58	14.76	33.66	50.5
BH 11	1678.66	16.76	6.71	6.77	33.21	33.83	6.08	5.89	5.7	5.64	8.99	98.76	350.25	227.28	569.67	1546.83	7.14	14.60	32.65	50.5
BH 12	1602.55	12.65	6.21	5.55	31.87	35.76	5.55	5.21	4.02	4.98	8.67	98.71	345.87	218.23	556.83	1476.93	7.12	14.11	31.54	49.6
BH 13	1581.55	10.82	5.51	6.04	32.9	34.88	5.55	5.08	4.88	4.98	8.51	97.66	375.81	209.17	550.87	1461.58	6.88	13.87	30.74	49.0
BH 14	1578.31	11.76	7.44	5.07	30.88	32.88	5.02	4.96	4.78	3.09	8.33	96.65	348.50	197.84	548.61	1436.82	6.87	13.84	28.63	45.7
BH 15	1632.87	12.88	6.02	4.66	32.33	36.87	5.11	4.83	3.88	2.87	7.86	90.76	345.87	188.03	510.22	1425.68	6.83	13.20	28.49	45.7
BH 16	1565.17	13.32	8.46	5.15	39.02	33.87	5.11	4.96	3.87	2.62	7.61	91.76	350.52	183.83	506.50	1374.60	6.72	12.93	28.19	45.4
BH 17	1550.60	15.52	7.74	6.03	40.66	34.55	5.08	4.58	3.675	3.03	7.45	85.90	335.94	182.88	490.67	1331.17	6.66	12.57	27.19	41.6
BH 18	1468.89	16.67	5.55	5.55	41.77	35.05	5.035	5.00	3.45	3.5	5.51	81.90	330.83	185.83	479.09	1257.27	6.09	12.24	27.03	41.2
BH 19	1456.77	17.65	6	4.76	39.77	36.5	4.77	4.33	3.38	4.14	7.44	92.82	330.83	178.87	466.90	1247.83	6.06	12.03	26.65	41.0
BH 20	1414.69	16.83	4.77	4.86	37.62	33.83	4.09	4.20	3.33	5.03	6.77	90.11	290.81	177.83	461.98	1234.98	5.91	11.77	26.65	40.8
BH 21	1358.32	17.23	6.04	4.09	40.77	35.76	4.645	4.72	3.9	5.62	6.98	73.87	312.73	168.54	454.93	1212.68	5.61	11.57	26.60	40.7
BH 22	1345.83	18.92	4.68	4.77	43.87	34.88	4.43	3.95	3.67	4.82	6.57	71.87	320.00	167.64	447.90	1130.11	5.49	10.99	26.58	40.7
BH 23	1332.35	16.72	5.9	3.62	37.88	32.88	3.45	4.67	3.12	3.66	5.55	64.37	300.83	162.83	445.03	1069.97	5.35	10.96	25.77	39.6
BH 24	1325.41	14.76	5.82	4.09	34.66	30.87	4.045	4.30	3.11	4.09	6.09	66.87	295.67	160.64	440.83	1050.77	5.21	10.88	24.19	39.2

Step 2: The “Descriptive Statistics” dialog box was opened. The “Quantitative Data” was selected to continue the analysis. The chemical formula of metal elements and their concentrations for all boreholes cells were selected.

Descriptive statistics dialog box settings:

- General tab selected
- Quantitative data:  (Range: Sheet4!\$B\$1:\$V\$6)
- Qualitative data:
- Subsamples:
- Variable-Category labels:
- Sample labels:

Step 3: The “Chart” dialog box was selected and Normal Q-Q plot was selected for analysis. “OK” was pressed to continue the analysis.



Step 4: The Normal Q-Q plot was originated for each metal element under this study.

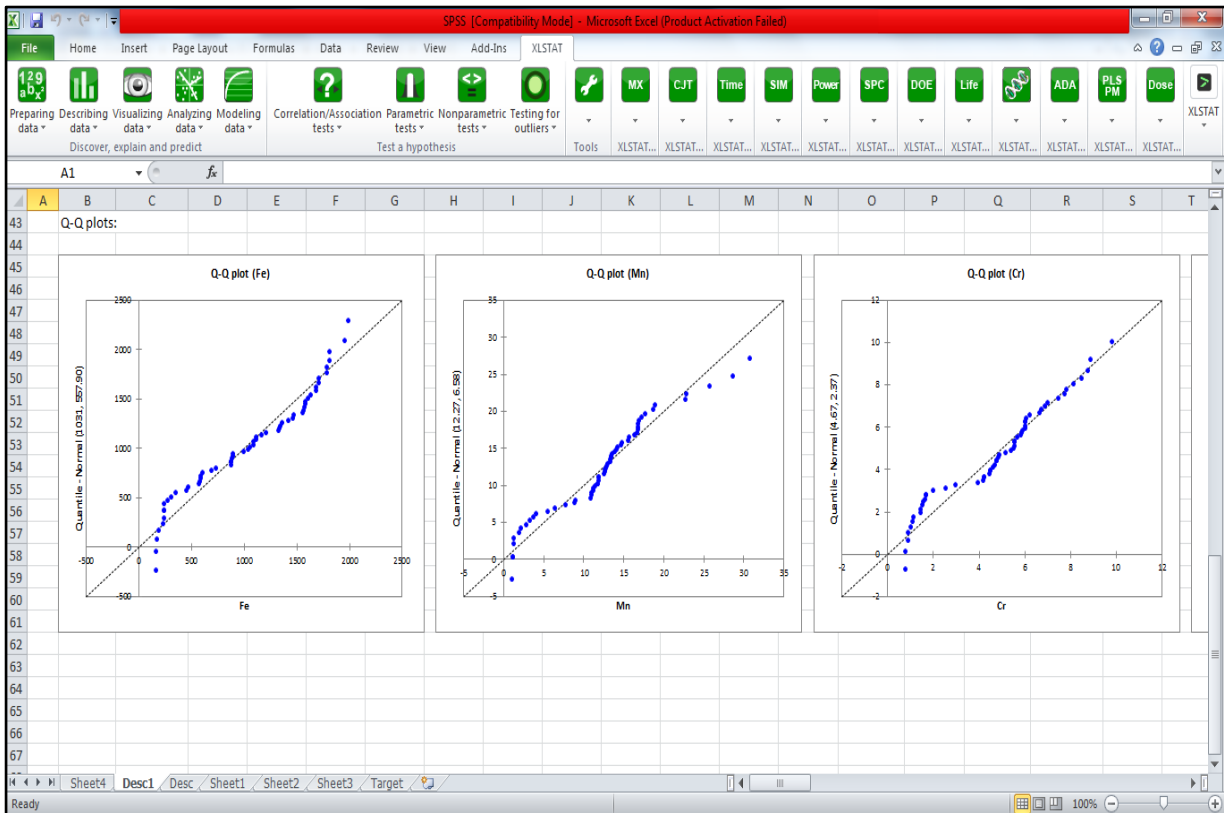


Figure A.2: Stepwise plotting procedure for Normal QQ Plot.



## Steps of Conventional Statistics

Step 1: “Descriptive Statistics” under the “Analysis” dialog box was selected. All the metal elements were selected to analyze for computing conventional statistics.

The screenshot shows the SPSS Data Editor interface with a data table containing 25 rows and 14 columns of metal concentration data. A 'Descriptives' dialog box is overlaid on the table, showing a list of selected variables: K, Ca, Al, Ti, Sb, Sc, Sr, V, and Ba. The 'Save standardized values as variables' checkbox is unchecked. The dialog box has buttons for OK, Paste, Reset, Cancel, and Help.

Step 2: “Conventional Statistics Option” was selected. All the required parameters were selected to compute. “Continue” was pressed to run the analysis.

The screenshot shows the same SPSS Data Editor interface with the 'Descriptives: Options' dialog box open. This dialog box allows for selecting specific statistical parameters to compute. Under 'Mean', the 'Mean' checkbox is checked. Under 'Dispersion', 'Stdeviation', 'Variance', 'Range', 'Minimum', 'Maximum', and 'S.E. mean' are all checked. Under 'Distribution', 'Kurtosis' and 'Skewness' are checked. Under 'Display Order', 'Variable list' is selected. The 'Continue' button is highlighted.

Step 3: SPSS output file represented the result of descriptive statistics originated from SPSS analysis.

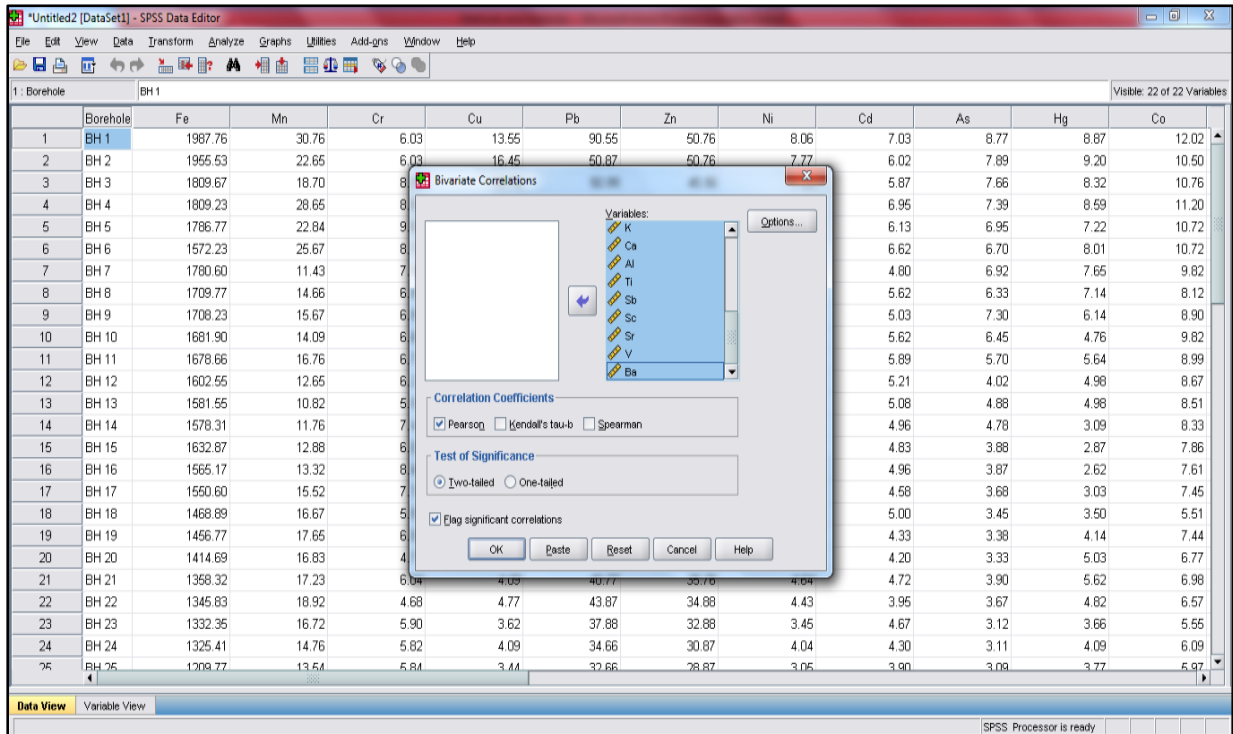
**Descriptive Statistics**

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Fe	60	1827.05	160.71	1987.76	1.0312E3	72.63182	562.60369	3.165E5	-.094	.309	-1.292	.608
Mn	60	29.74	1.02	30.76	12.2656	.85624	6.63242	43.989	.264	.309	.536	.608
Cr	60	9.05	.77	9.82	4.6704	.30792	2.38510	5.689	-.120	.309	-.768	.608
Cu	60	15.72	.73	16.45	4.9861	.45676	3.53804	12.518	1.781	.309	2.789	.608
Pb	60	79.67	10.88	90.55	30.9187	1.91207	14.81082	219.360	1.511	.309	3.714	.608
Zn	60	38.94	11.82	50.76	28.9058	1.33664	10.35360	107.197	.333	.309	-.602	.608
Ni	60	6.98	1.08	8.06	4.0846	.23195	1.79668	3.228	.434	.309	-.537	.608
Cd	60	5.83	1.20	7.03	3.7115	.20271	1.57020	2.466	.152	.309	-.798	.608
As	60	7.90	.87	8.77	3.2877	.27333	2.11721	4.483	.946	.309	-.004	.608
Hg	60	8.48	.72	9.20	3.6581	.29636	2.29560	5.270	.763	.309	-.122	.608
Co	60	10.04	1.98	12.02	6.3730	.33525	2.59683	6.744	.210	.309	-.801	.608
Na	60	92.31	8.23	100.54	54.1128	4.16460	32.25887	1.041E3	.149	.309	-1.599	.608
K	60	405.29	55.04	460.33	2.3695E2	15.75259	122.01903	1.489E4	.093	.309	-1.346	.608
Ca	60	274	44	318	155.92	8.595	66.578	4.433E3	.688	.309	-.164	.608
Al	60	7.5558E2	1.1922E2	8.74783E2	4.0675E2	2.6678E1	2.06647220E2	4.270E4	.650	.309	-.397	.608
Ti	60	1.6934E3	2.4388E2	1.93738E3	1.0501E3	5.7331E1	4.44090700E2	1.972E5	.339	.309	-.833	.608
Sb	60	1.1568E1	...	1.25487E1	5.1869E0	...	2.37853375E0	5.657	.606	.309	.629	.608
Sc	60	1.7390E1	3.0200E0	2.04105E1	1.0686E1	...	3.70672517E0	13.740	.296	.309	-.136	.608
Sr	60	4.5239E1	8.8800E0	5.41196E1	2.3992E1	1.1699E0	9.06243687E0	82.128	.755	.309	.792	.608
V	60	7.6431E1	6.9200E0	8.33511E1	3.6105E1	2.1504E0	1.66575548E1	277.474	.563	.309	.233	.608
Ba	60	1.0370E2	1.8200E1	1.21901E2	5.6693E1	3.2187E0	2.49320202E1	621.606	.761	.309	-.004	.608
Valid N (listwise)	60											

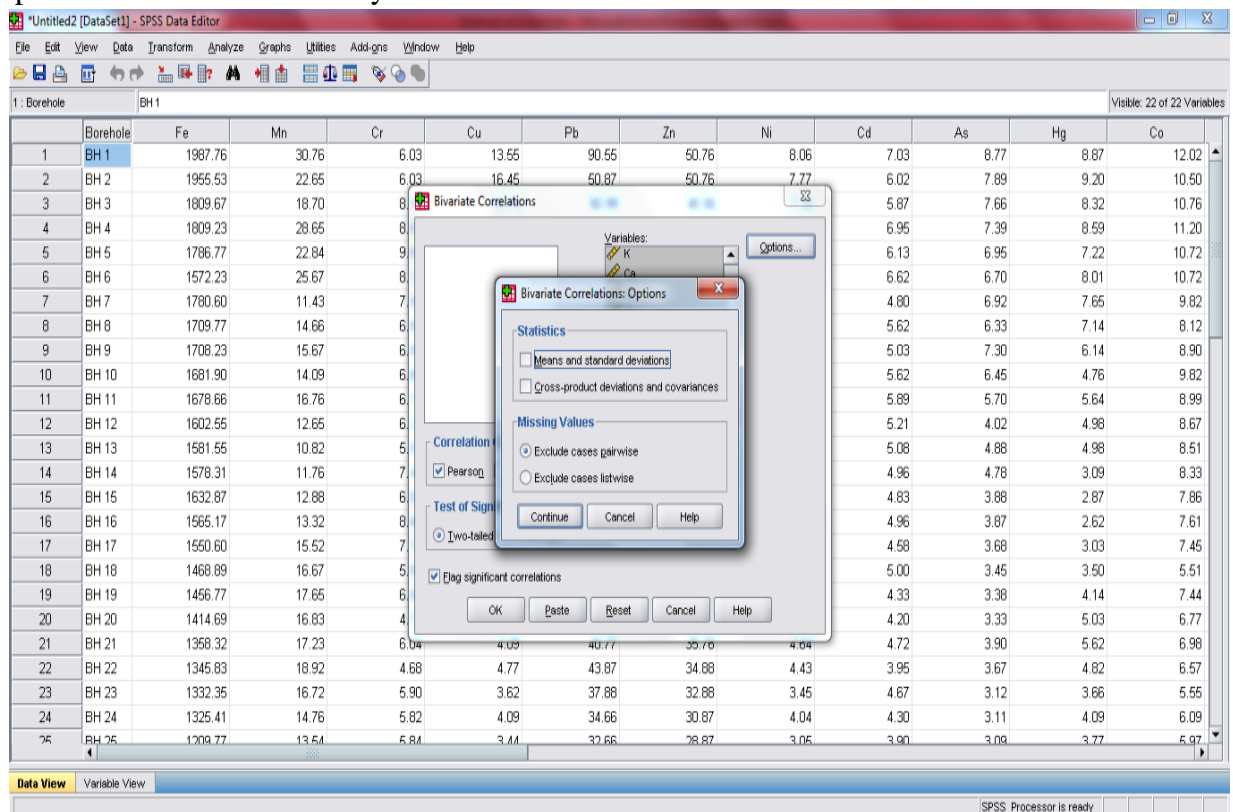
Figure A.3: Stepwise computation procedure for conventional statistics.

## Steps of Pearson's Correlation

Step 1: "Correlate" under the "Analysis" dialog box was selected. Then the dialog box of "Bivariate Correlation" was selected. Correlation coefficient for Pearson was specified, with two tailed significance value. Flag significance correlations were also selected.



Step 2: The "Bivariate Correlations Option" was selected. "Continue" in the dialog box was pressed to continue the analysis.



Step 3: The “Correlation Coefficient” was originated for metal elements was recorded.

The screenshot shows the SPSS Viewer interface with a 'Correlations' table. The table is titled 'Correlations' and is for '[DataSet1]'. It displays the Pearson Correlation, Sig. (2-tailed), and N for each element. The elements are Fe, Mn, Cr, Cu, Pb, Zn, Ni, Cd, As, Hg, Co, Na, K, Ca, and A. The table is a lower triangular matrix, with the diagonal elements all having a Pearson Correlation of 1.000 and a Sig. (2-tailed) of .000. The N for all elements is 40.

		Fe	Mn	Cr	Cu	Pb	Zn	Ni	Cd	As	Hg	Co	Na	K	Ca	A
Fe	Pearson Correlation	1	.554	.768	.756	.739	.845	.928	.889	.914	.760	.949	.942	.967	.948	.9
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40
Mn	Pearson Correlation	.554	1	.449	.762	.872	.708	.632	.741	.656	.727	.653	.462	.570	.658	.6
	Sig. (2-tailed)	.000		.004	.000	.000	.000	.000	.000	.000	.000	.000	.003	.000	.000	.000
	N	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40
Cr	Pearson Correlation	.768	.449	1	.669	.567	.711	.753	.747	.740	.575	.797	.738	.774	.750	.7
	Sig. (2-tailed)	.000	.004		.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40
Cu	Pearson Correlation	.756	.762	.669	1	.805	.887	.875	.815	.880	.882	.856	.833	.745	.884	.8
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40
Pb	Pearson Correlation	.739	.872	.567	.805	1	.800	.761	.825	.810	.789	.811	.821	.738	.812	.7
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40
Zn	Pearson Correlation	.845	.708	.711	.887	.800	1	.886	.797	.867	.789	.879	.753	.857	.881	.8
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40
Ni	Pearson Correlation	.928	.632	.753	.875	.761	.886	1	.880	.959	.855	.942	.818	.887	.958	.9
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000		.000	.000	.000	.000	.000	.000	.000	.000
	N	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40
Cd	Pearson Correlation	.889	.741	.747	.815	.825	.797	.880	1	.892	.807	.918	.817	.867	.920	.9
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000		.000	.000	.000	.000	.000	.000	.000
	N	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40

Figure A.4: Stepwise computation procedure for Pearson’s correlation.

## Steps of Principal Component Analysis

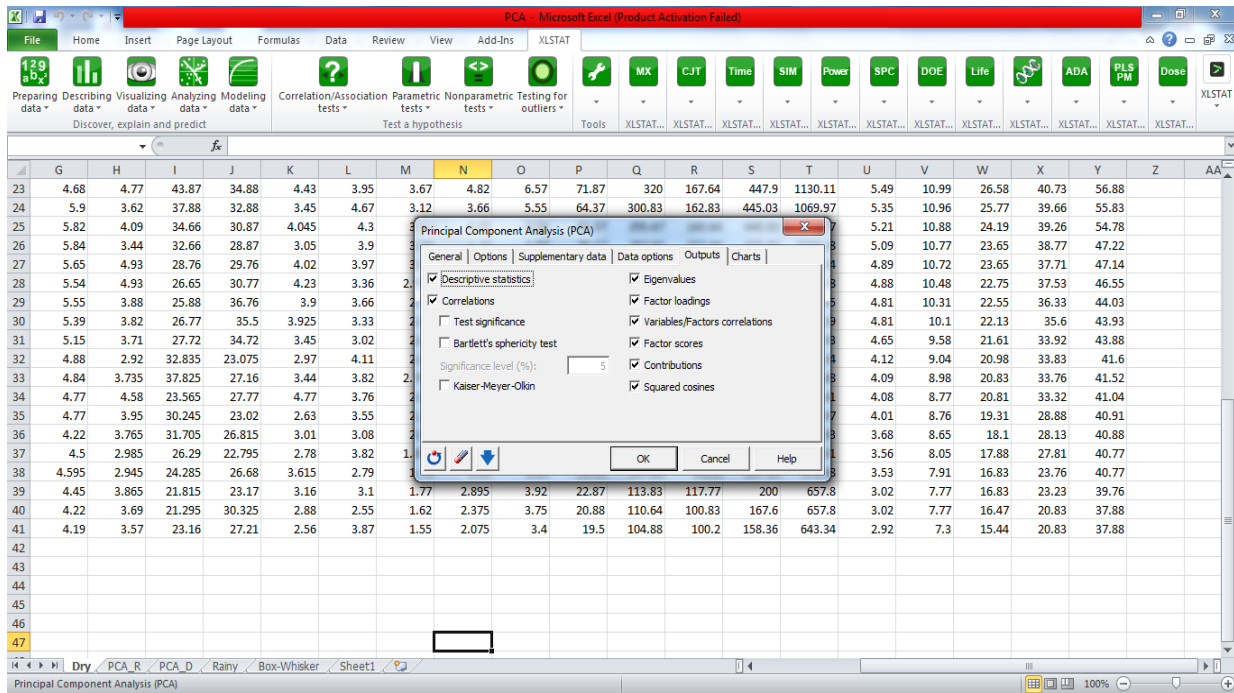
Step 1: The data needed to analyze was displayed in the excel sheet.

Borehole	Fe	Mn	Cr	Cu	Pb	Zn	Ni	Cd	As	Hg	Co	Na	K	Ca	Al	Ti	Sb	Sr	V	
BH 1	1987.76	30.76	6.03	13.55	90.55	50.76	8.06	7.03	8.77	8.87	12.02	100.54	460.33	318.00	874.78	1937.36	12.55	20.41	54.12	83.3
BH 2	1955.53	22.65	6.03	16.45	50.87	50.76	7.77	6.02	7.89	9.2	10.5	99.20	453.55	302.82	874.50	1915.90	10.77	17.64	40.77	71.0
BH 3	1809.67	18.7	8.76	10.78	50.99	45.56	7.225	5.87	7.66	8.32	10.76	88.76	431.73	299.46	800.87	1899.29	9.69	17.04	40.74	70.7
BH 4	1809.23	28.65	8.87	13.98	70.87	50.7	7.43	6.95	7.39	8.59	11.2	92.82	419.85	272.93	777.74	1809.77	8.87	16.87	38.88	67.9
BH 5	1786.77	22.84	9.82	12.88	50.66	48.87	7.11	6.13	6.95	7.22	10.72	98.76	405.34	262.34	767.43	1765.91	8.87	16.83	37.75	65.8
BH 6	1572.23	25.67	8.12	15.32	60.77	45.55	6.77	6.62	6.7	8.01	10.72	98.71	413.51	270.77	765.77	1679.03	7.99	15.93	36.70	60.2
BH 7	1780.60	11.43	7.8	11.88	45.66	44.33	7.03	4.80	6.92	7.65	9.82	97.66	402.77	266.50	761.43	1675.65	7.95	15.88	36.16	56.3
BH 8	1709.77	14.66	6.89	7.66	37.66	37.9	6.765	5.62	6.33	7.14	8.12	88.00	387.29	256.76	740.00	1673.67	7.85	15.88	35.83	55.2
BH 9	1708.23	15.67	6.99	6.77	40.76	35.05	6.76	5.03	7.3	6.14	8.9	88.87	365.88	240.77	707.98	1667.74	7.77	14.77	35.77	50.8
BH 10	1681.90	14.09	6.65	7.11	46.89	36.5	6.11	5.62	6.45	4.76	9.82	87.34	355.38	238.65	657.88	1566.17	7.58	14.76	33.66	50.5
BH 11	1678.66	16.76	6.71	6.77	33.21	33.83	6.08	5.89	5.7	5.64	8.99	98.76	350.25	227.28	569.67	1546.83	7.14	14.80	32.65	50.5
BH 12	1602.55	12.65	6.21	5.55	31.87	35.76	5.55	5.21	4.02	4.98	8.67	98.71	345.87	218.23	556.83	1476.93	7.12	14.11	31.54	49.6
BH 13	1581.55	10.82	5.51	6.04	32.9	34.88	5.55	5.08	4.88	4.98	8.51	97.66	375.81	209.17	550.87	1461.58	6.88	13.87	30.74	49.0
BH 14	1578.31	11.76	7.44	5.07	30.88	32.88	5.02	4.96	4.78	3.09	8.33	96.65	348.50	197.84	548.61	1436.82	6.87	13.84	28.63	45.7
BH 15	1632.87	12.88	6.02	4.66	32.33	36.87	5.11	4.83	3.88	2.87	7.86	90.76	345.87	188.03	510.22	1425.68	6.83	13.20	28.49	45.7
BH 16	1565.17	13.32	8.46	5.15	39.02	33.87	5.11	4.96	3.87	2.62	7.61	91.76	350.52	183.83	506.50	1374.60	6.72	12.93	28.19	45.4
BH 17	1550.60	15.52	7.74	6.03	40.66	34.55	5.08	4.58	3.675	3.03	7.45	85.90	335.94	182.88	490.67	1331.17	6.66	12.57	27.19	41.6
BH 18	1468.89	16.67	5.55	5.55	41.77	35.05	5.035	5.00	3.45	3.5	5.51	81.90	330.83	185.83	479.09	1257.27	6.09	12.24	27.03	41.2
BH 19	1456.77	17.65	6	4.76	39.77	36.5	4.77	4.33	3.38	4.14	7.44	92.82	330.83	178.87	466.90	1247.83	6.06	12.03	26.65	41.0
BH 20	1414.69	16.83	4.77	4.86	37.62	33.83	4.09	4.20	3.33	5.03	7.07	90.11	290.81	177.83	461.98	1234.98	5.91	11.77	26.65	40.8
BH 21	1358.32	17.23	6.04	4.09	40.77	35.76	4.645	4.72	3.9	5.62	6.98	73.87	312.73	168.54	454.93	1212.68	5.61	11.57	26.60	40.7
BH 22	1345.83	18.92	4.68	4.77	43.87	34.88	4.43	3.95	3.67	4.82	6.57	71.87	320.00	167.64	447.90	1130.11	5.49	10.99	26.58	40.7
BH 23	1332.35	16.72	5.9	3.62	37.88	32.88	3.45	4.67	3.12	3.66	5.55	64.37	300.83	162.83	445.03	1069.97	5.35	10.96	25.77	39.6
BH 24	1325.41	14.76	5.82	4.09	34.66	30.87	4.045	4.30	3.11	4.09	6.09	66.87	295.67	160.64	440.83	1050.77	5.21	10.88	24.19	39.2

Step 2: "Principal Component Analysis (PCA)" under "Analyzing data" was selected. The data for "Observation/variables tables" was selected. "Observation labels" was selected. "OK" was clicked to continue. Options" wizard of PCA was selected. "Rotation" was selected.

G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA
20	6	4.76	39.77	36.5	4.77	4.33	3.38	4.14	7.44	92.82	330.83	178.87	466.9	1247.83	6.06	12.03	26.65	41.07	61.85	
21	4.77	4.86	37.62	33.83	4.09	4.2	3.33	5.03	6.77	90.11	290.81	177.83	461.98	1234.98	5.91	11.77	26.65	40.88	60.73	
22	6.04	4.09	40.77	35.76	4.645	4.72									5.61	11.57	26.6	40.77	60.49	
23	4.68	4.77	43.87	34.88	4.43	3.95									5.49	10.99	26.58	40.73	56.88	
24	5.9	3.62	37.88	32.88	3.45	4.67									5.35	10.96	25.77	39.66	55.83	
25	5.82	4.09	34.66	30.87	4.045	4.3									5.21	10.88	24.19	39.26	54.78	
26	5.84	3.44	32.66	28.87	3.05	3.9									5.09	10.77	23.65	38.77	47.22	
27	5.65	4.93	28.76	29.76	4.02	3.97									4.89	10.72	23.65	37.71	47.14	
28	5.54	4.93	26.65	30.77	4.23	3.36	2.09	3.45	4.67	24.83	135.37	124.92	245	700.83	4.88	10.48	22.75	37.53	46.55	
29	5.55	3.88	25.88	36.76	3.9	3.66	1.855	3.05	4.06	20.83	120.76	120.83	230.89	698.01	4.81	10.31	22.55	36.33	44.03	
30	5.39	3.82	26.77	35.5	3.925	3.33									4.81	10.1	22.13	35.6	43.93	
31	5.15	3.71	27.72	34.72	3.45	3.02									4.65	9.58	21.61	33.92	43.88	
32	4.88	2.92	32.835	23.075	2.97	4.11									4.12	9.04	20.98	33.83	41.6	
33	4.84	3.735	37.825	27.16	3.44	3.82									4.09	8.98	20.83	33.76	41.52	
34	4.77	4.58	23.565	27.77	4.77	3.76									4.08	8.77	20.81	33.32	41.04	
35	4.77	3.95	30.245	23.02	2.63	3.55									4.01	8.76	19.31	28.88	40.91	
36	4.22	3.765	31.705	26.815	3.01	3.08	2.09	3.45	4.67	24.83	135.37	124.92	245	700.83	3.68	8.65	18.1	28.13	40.88	
37	4.5	2.985	26.29	22.795	2.78	3.82	1.855	3.05	4.06	20.83	120.76	120.83	230.89	698.01	3.56	8.05	17.88	27.81	40.77	
38	4.595	2.945	24.285	26.68	3.615	2.79	1.78	3.39	3.89	21.43	117.82	109.8	207.14	676.88	3.53	7.91	16.83	23.76	40.77	
39	4.45	3.865	21.815	23.17	3.1	1.77	2.895	3.92	22.87	113.83	117.77		200	657.8	3.02	7.77	16.83	23.23	39.76	
40	4.22	3.69	21.295	30.325	2.88	2.55	1.62	2.375	3.75	20.88	110.64	100.83	167.6	657.8	3.02	7.77	16.47	20.83	37.88	
41	4.19	3.57	23.16	27.21	2.56	3.87	1.55	2.075	3.4	19.5	104.88	100.2	158.36	643.34	2.92	7.3	15.44	20.83	37.88	

Step 3: "Chart" wizard of PCA was selected. All parameters were chosen to analyze.



Step 4: SPSS output file represented the result of PCA originated from XLSTAT analysis.

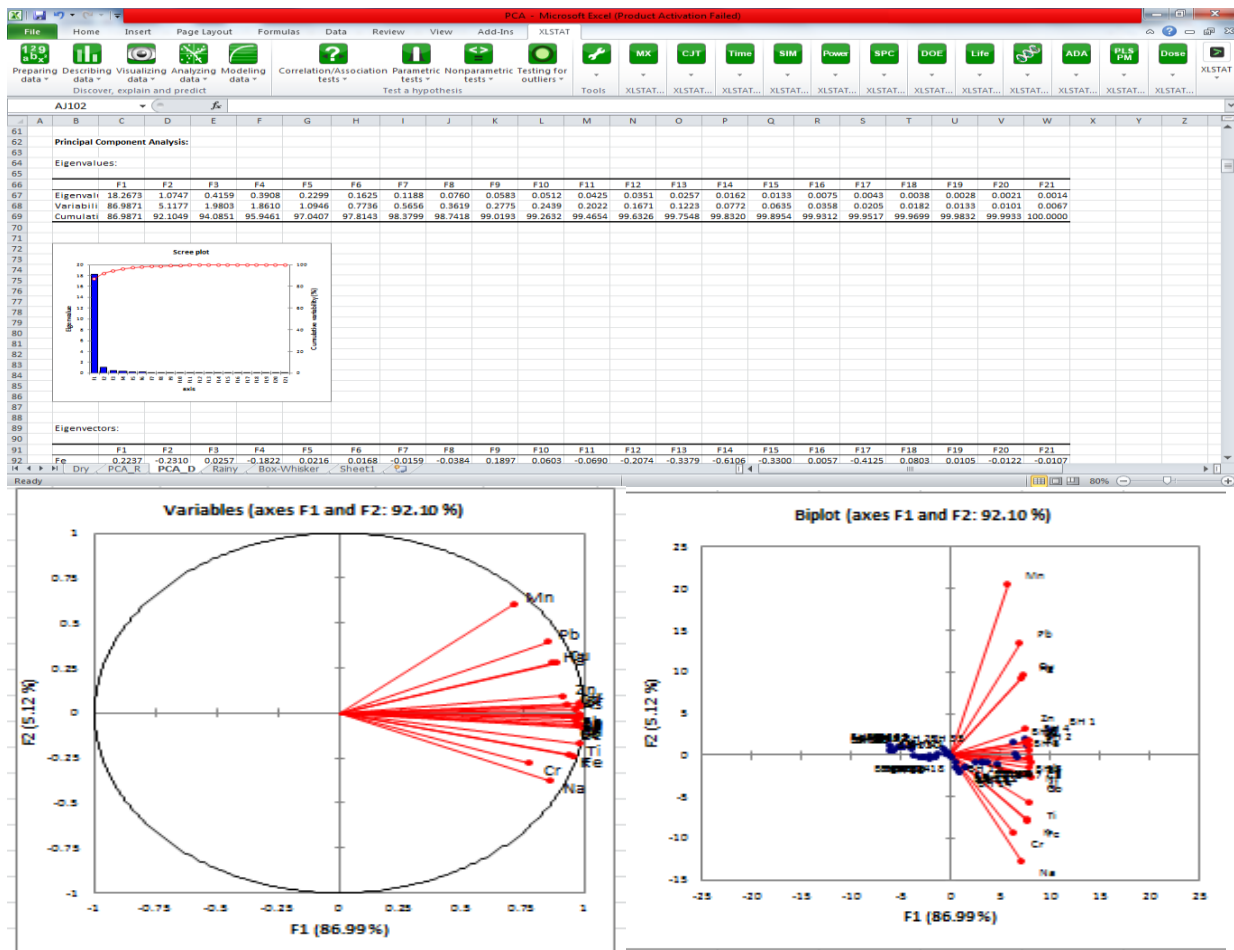
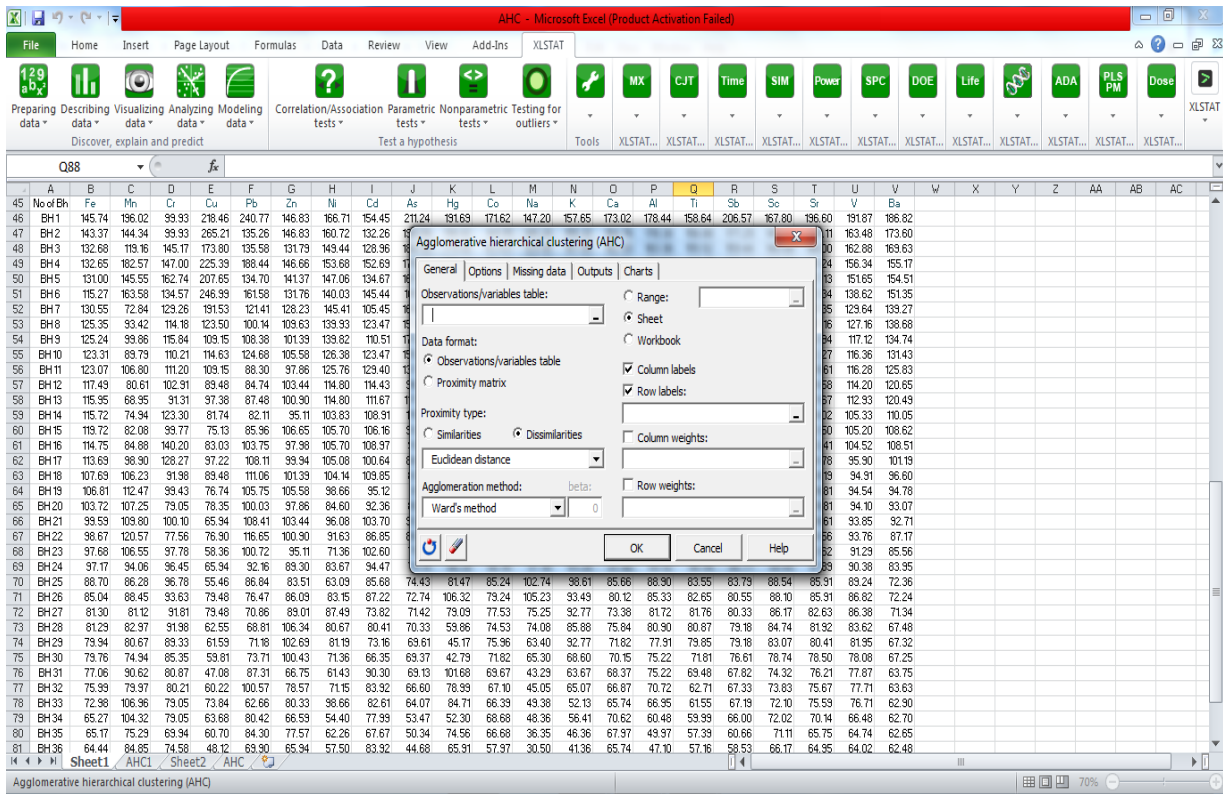


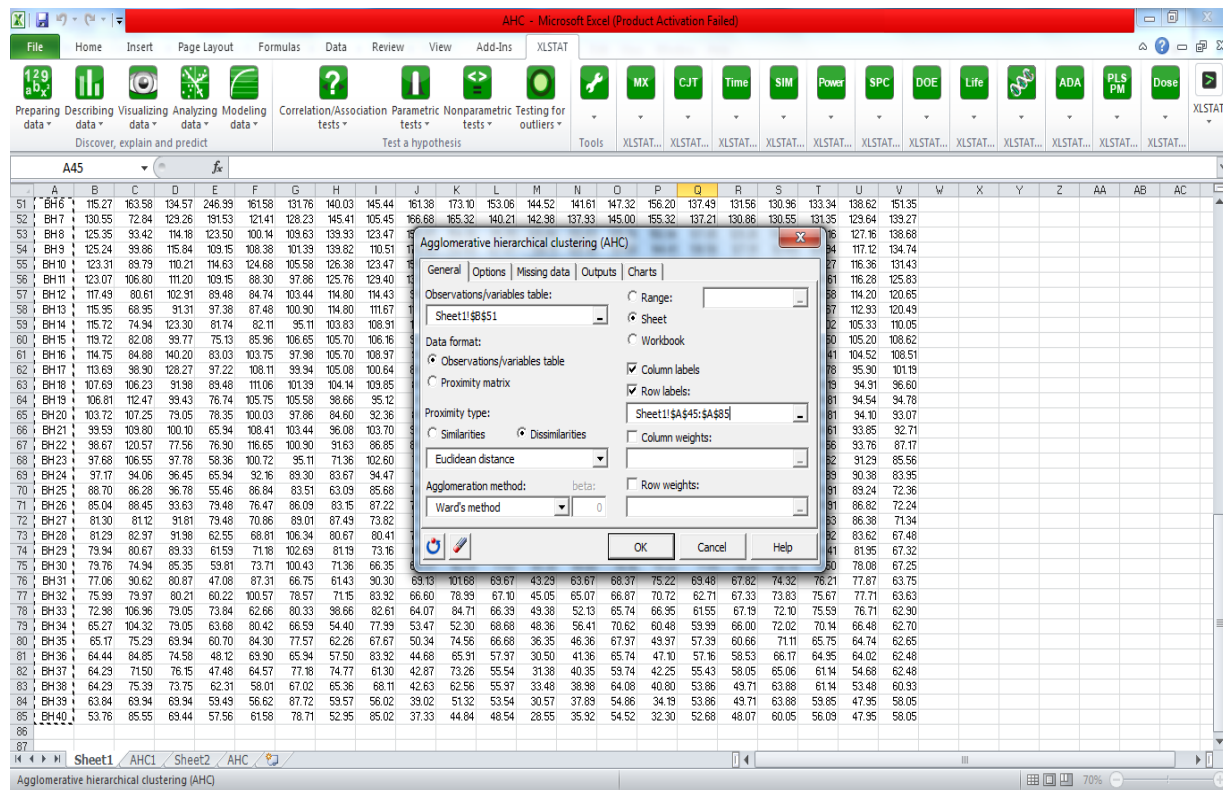
Figure A.5: Stepwise Analysis of Principal Component Analysis.

# Steps of Agglomerative Hierarchical Clustering

Step 1: "Agglomerative Hierarchical Analysis (AHC)" under "Analyzing data" was selected.



Step 2: The data for "Observation/variables tables" was selected. "Row labels" was selected. "OK" was clicked to continue.



Step 3: SPSS output file represented the result of AHC originated from XLSTAT analysis.

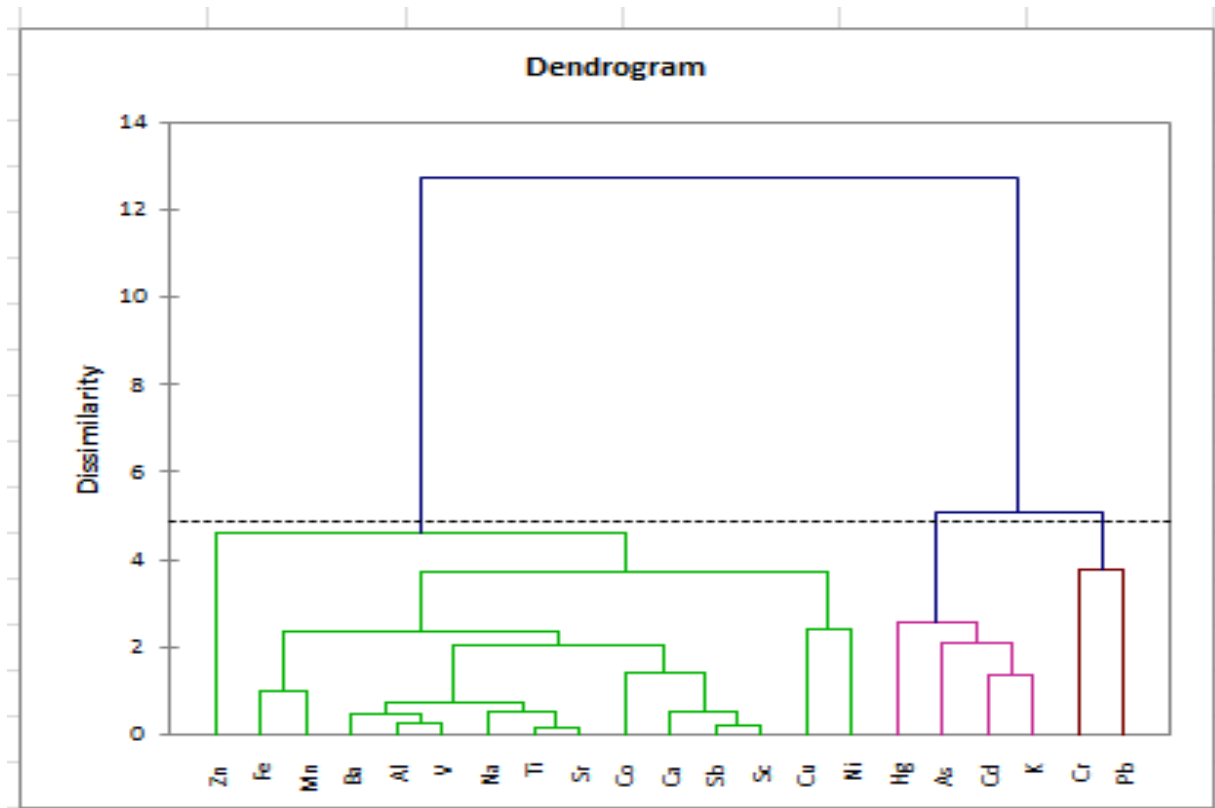
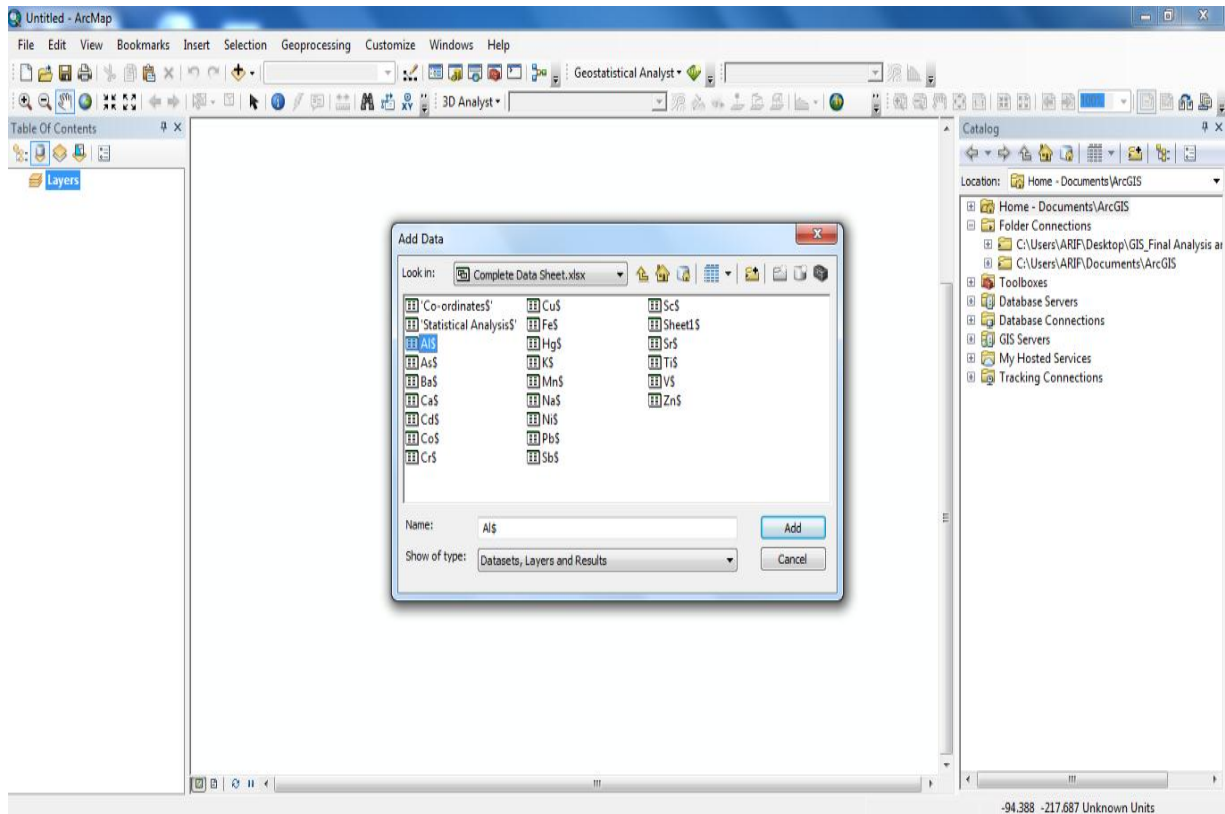


Figure A.6 : Stepwise analysis procedure of AHC.

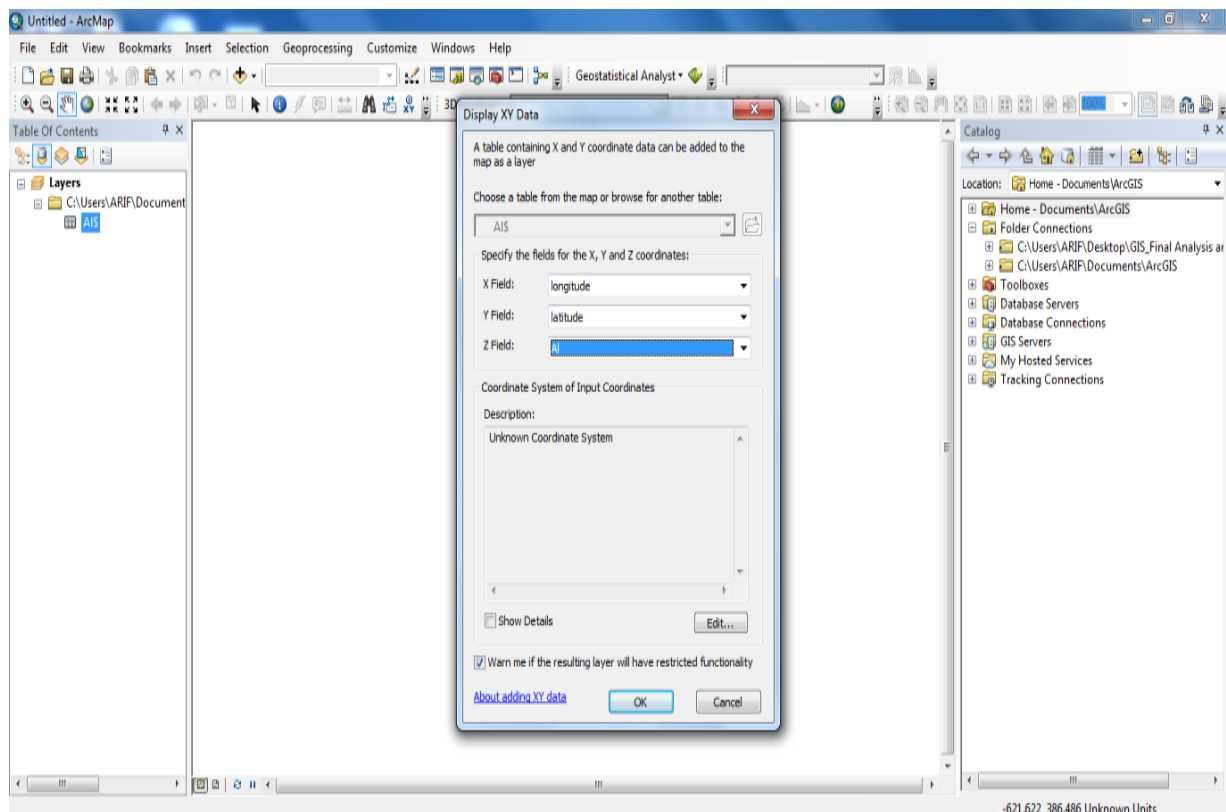


## Common Steps of Geostatistical Analysis

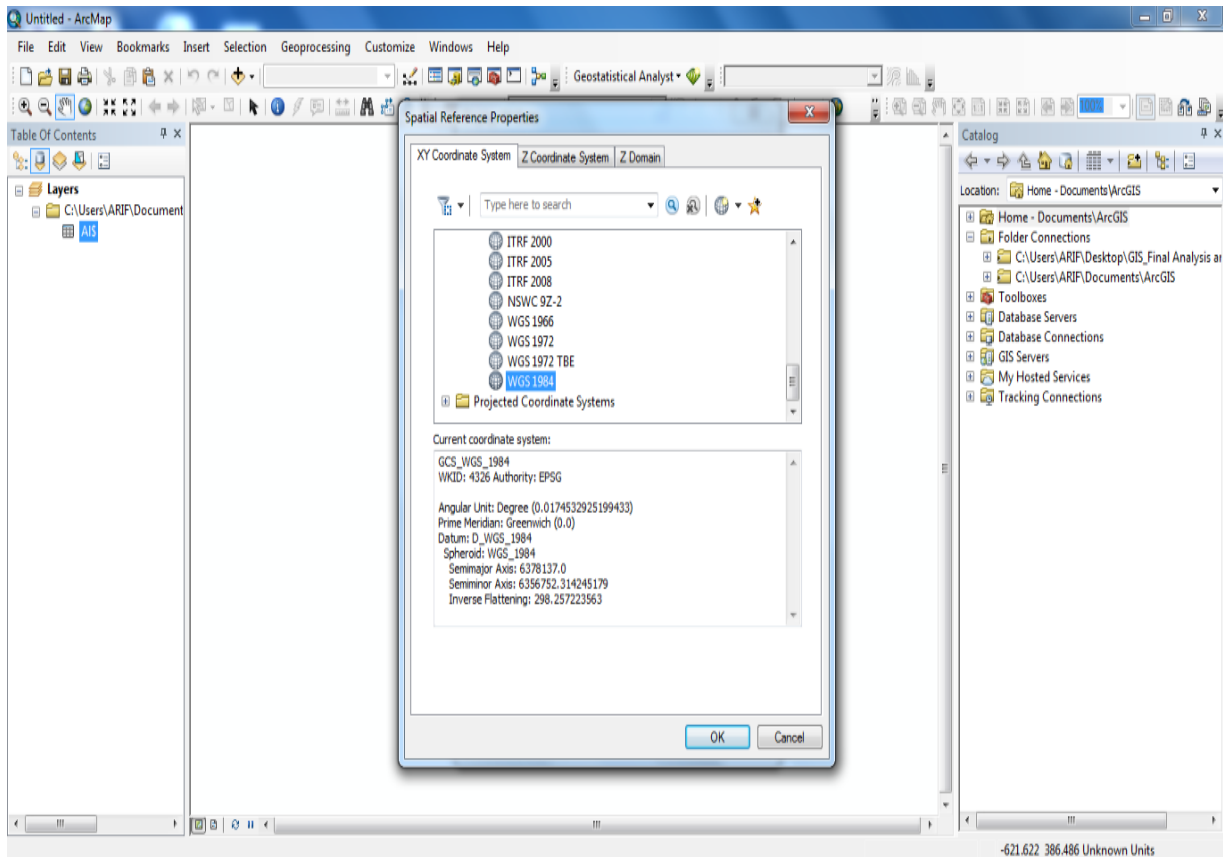
Step 1: The excel sheet was added to ArcGIS map.



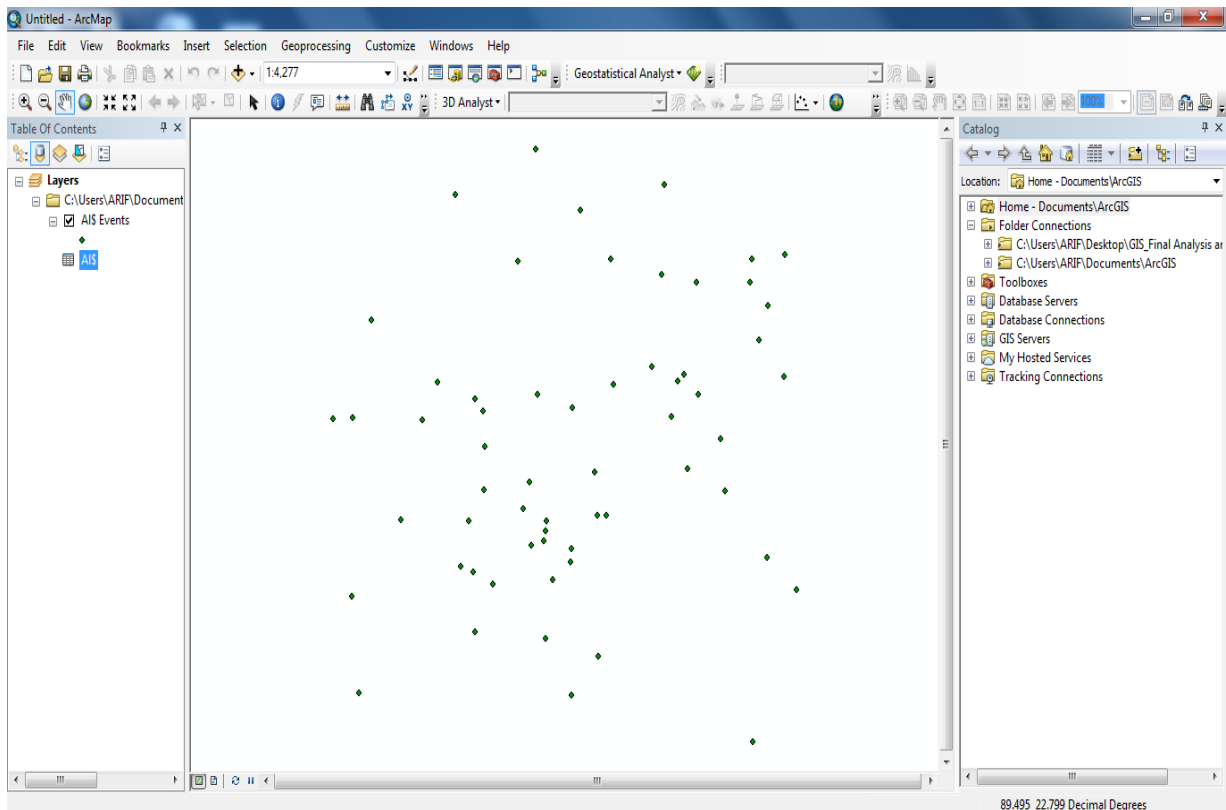
Step 2: The “Display XY Data” dialog box was opened and X, Y and Z field was specified.



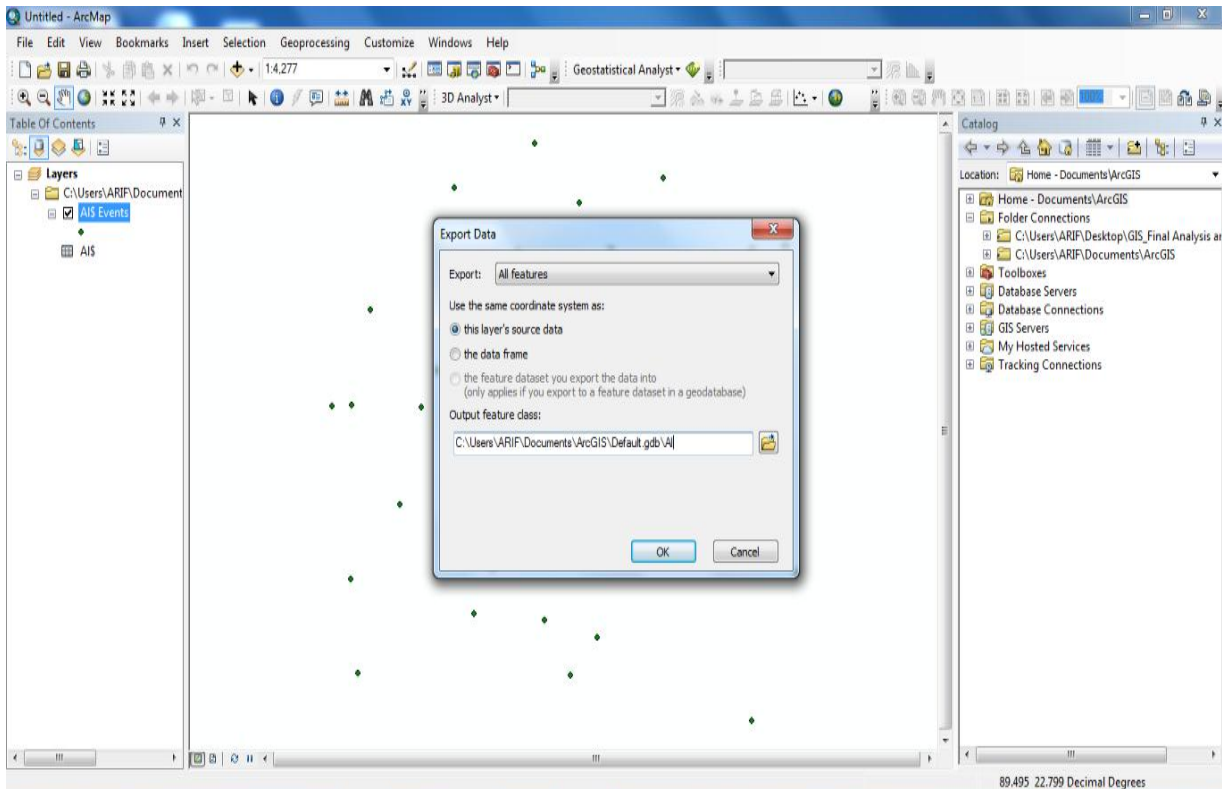
Step 3: “World Coordinate System (WGS 1984)” under “Geographic Coordinate System” was selected.



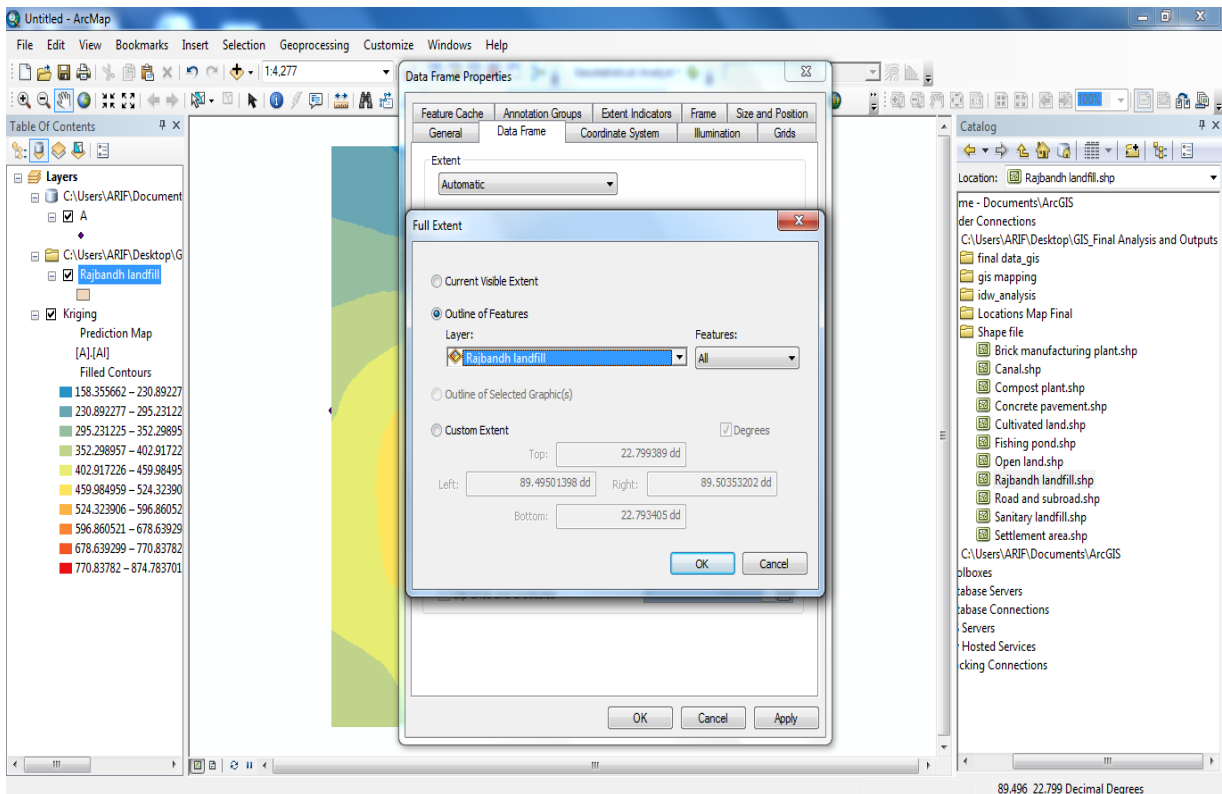
Step 4: The sampling locations of the study were displayed on ArcMap wizard.



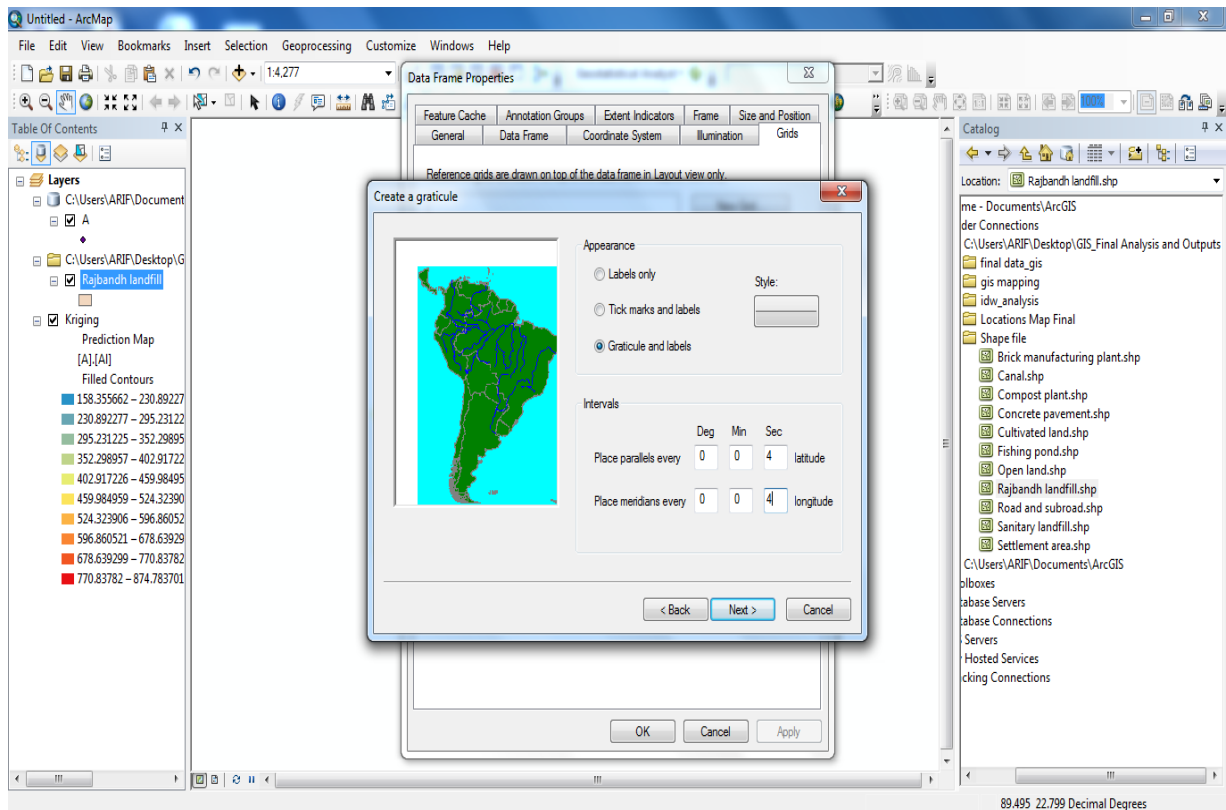
Step 5: Right click was done to “Layer” that lead to “Export”, followed by “Export Data”. “OK” was pressed to continue the analysis.



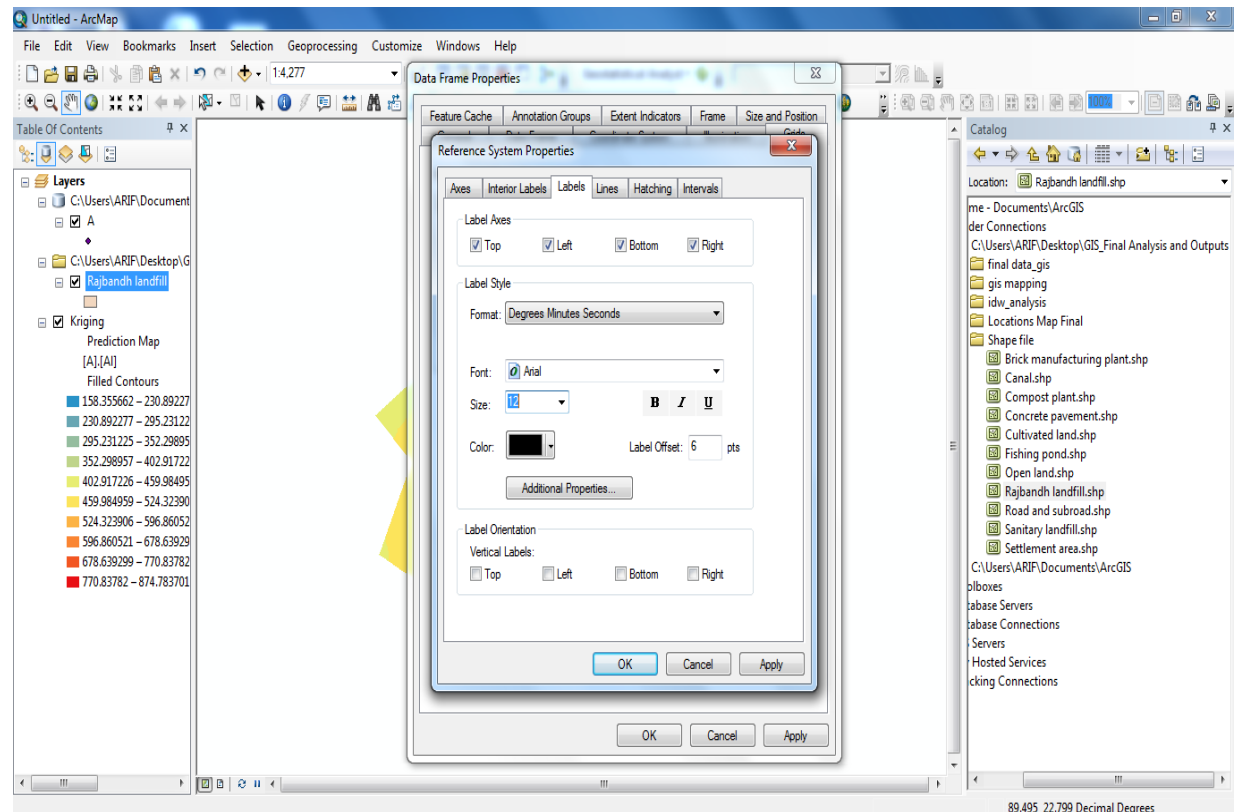
Step 6: “Data Frame Properties” was selected and the extent was specified as “Rajbandh landfill”. “Clip Options” was selected to clip to shape of Rajbandh landfill. “OK” was clicked to continue.



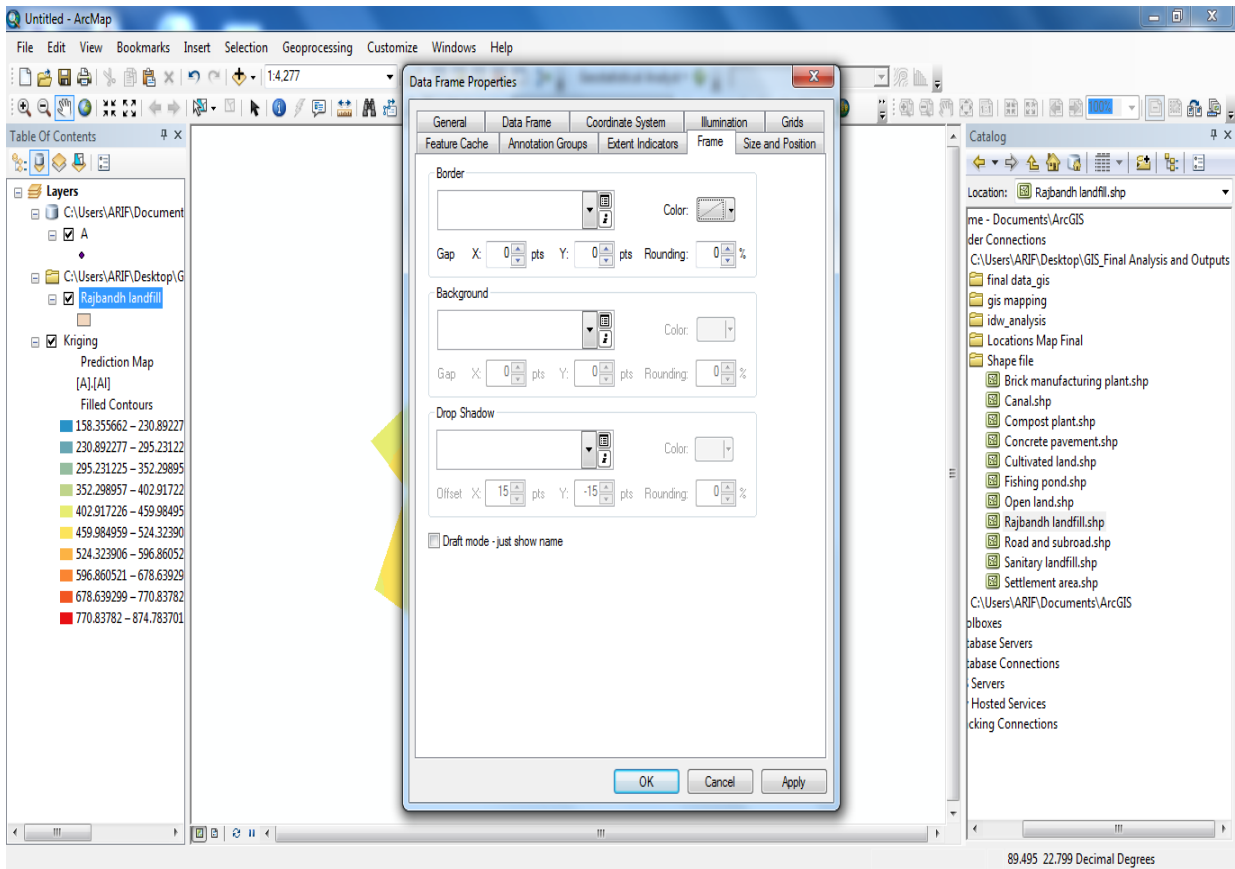
Step 7: “Graticule” was selected to display the grids in the ArcMap. The specifications for the wizard “New Grid” were identified.



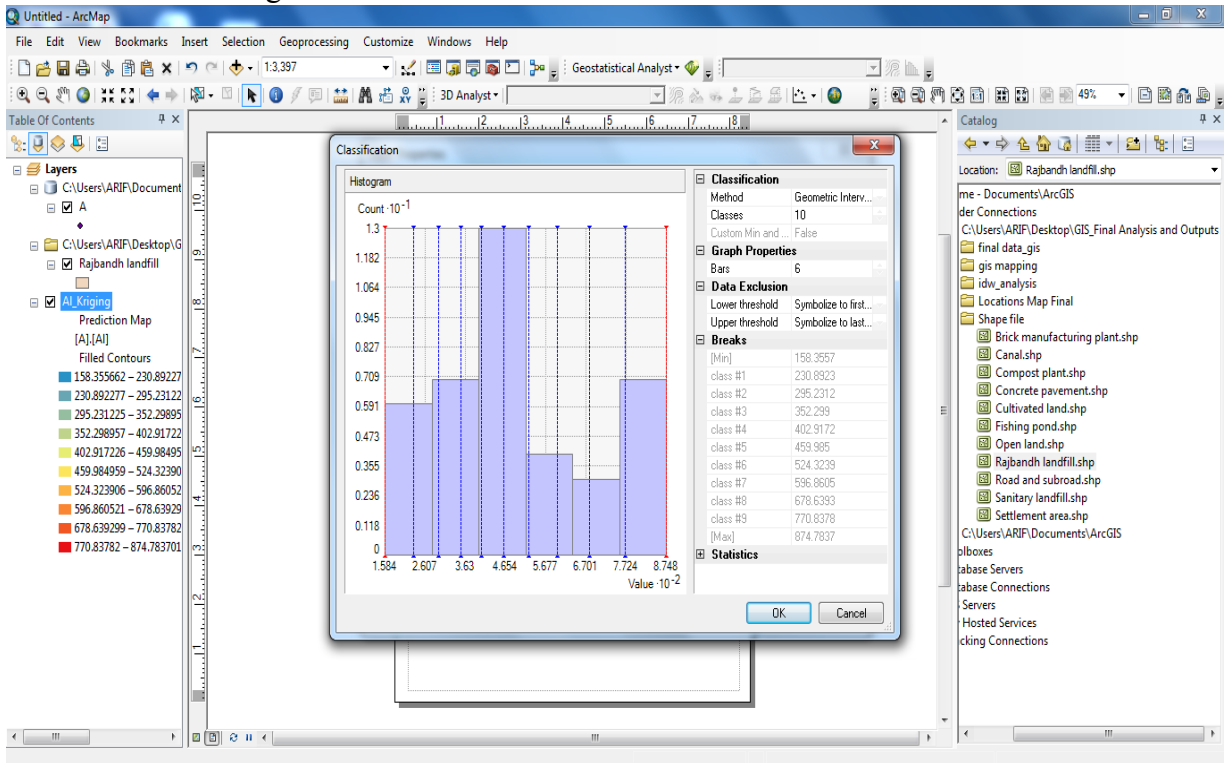
Step 8:”Font Size” under the “Properties” wizard was also specified,



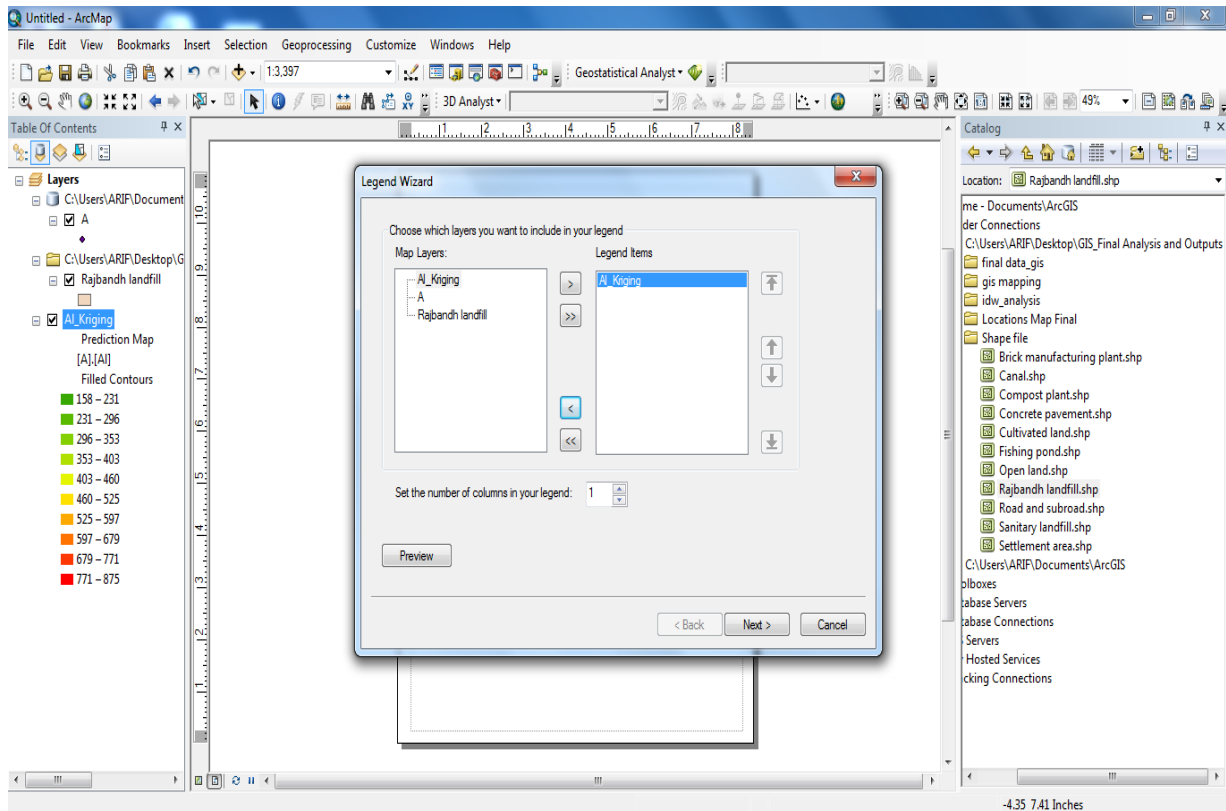
Step 9: “Frame” was chosen to change the border color of ArcMap. “No color” was selected. “OK” was pressed to continue.



Step 10: The “Symbology” wizard under the “Properties” of generated surface was selected. Modification in “Legend” was done.”OK” was clicked.



Step 11: “Insert” option was chosen and the legend of the shape file needed was chosen. “Next” was clicked to continue.



Step 12: The generated output surface map with the” Legend” in the “Layout view “.

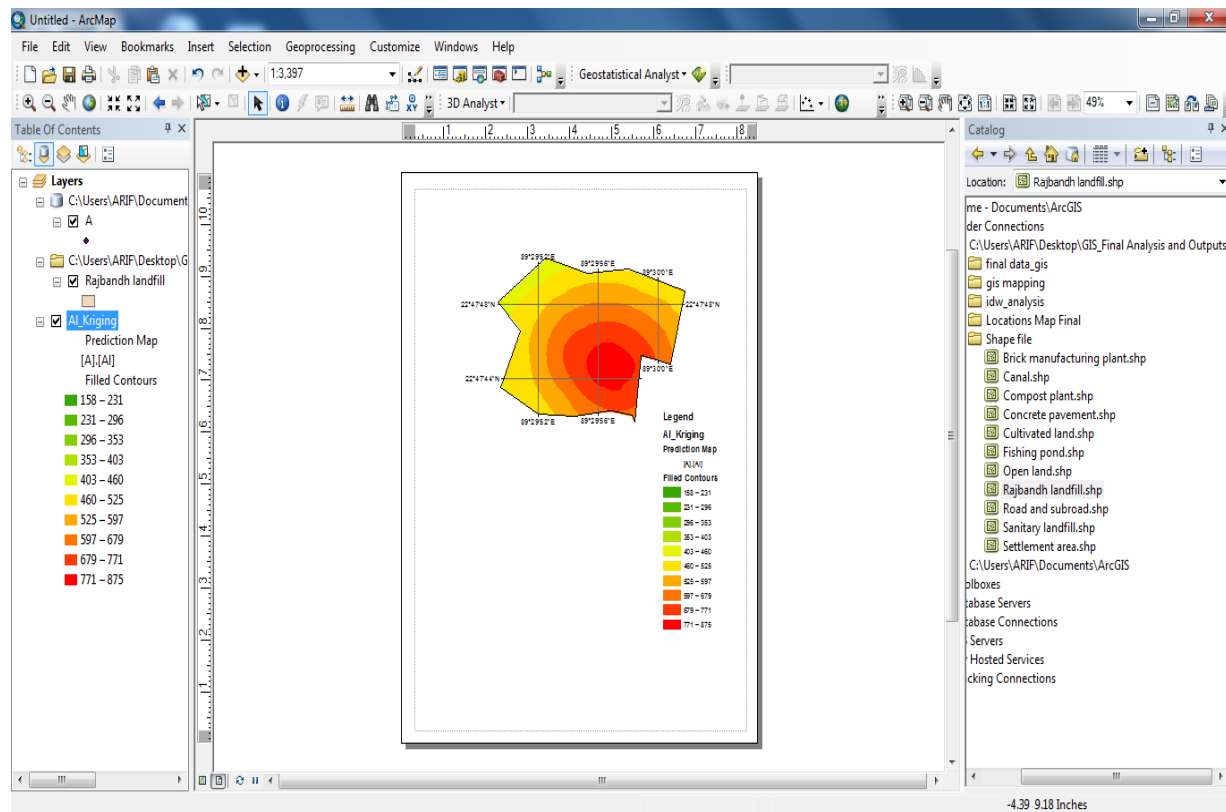
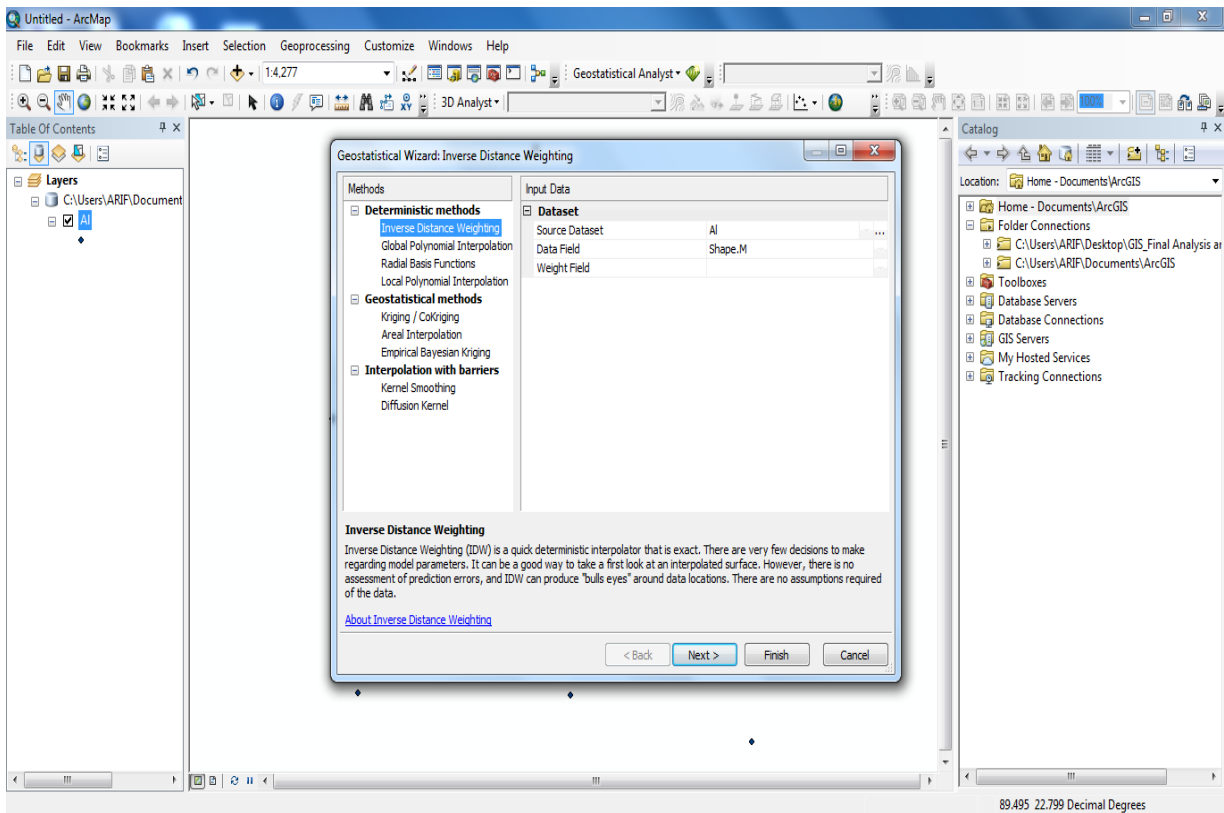


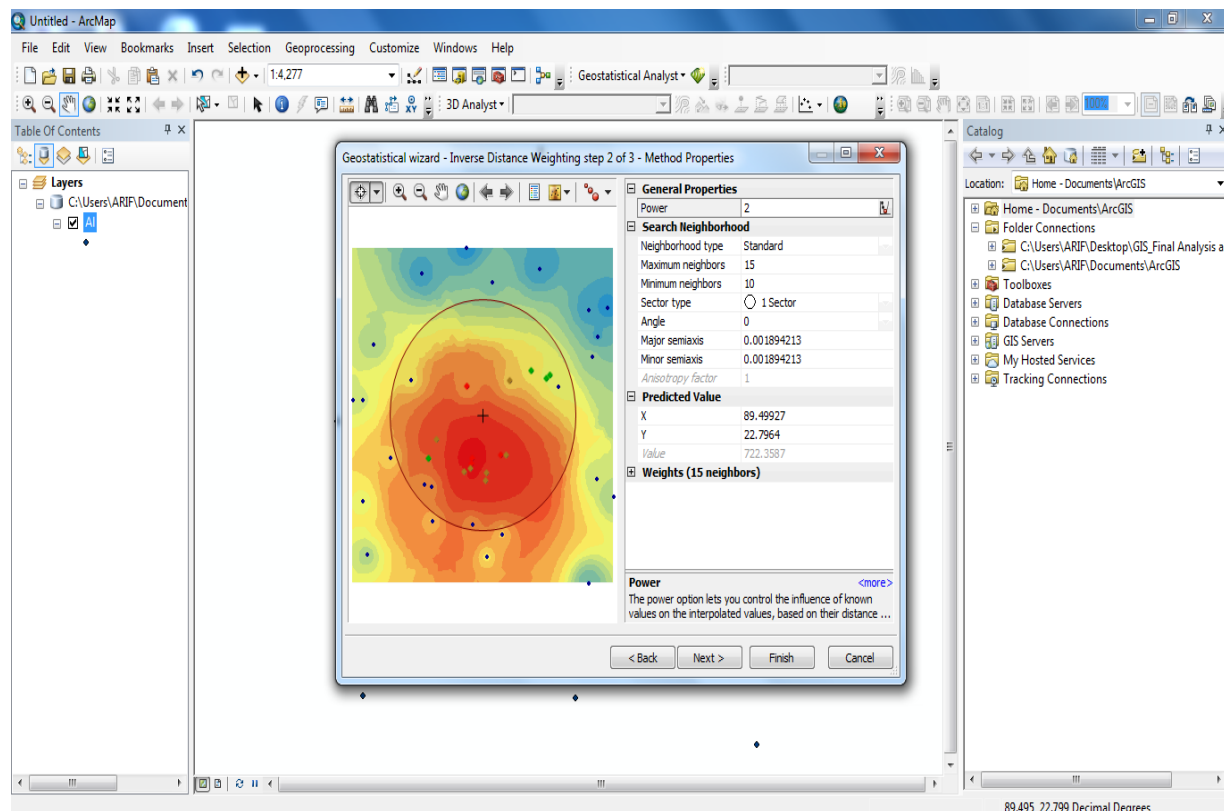
Figure A.7 : Common steps of geostatistical analysis by ArcGIS.

## Steps of Inverse Distance Weighting

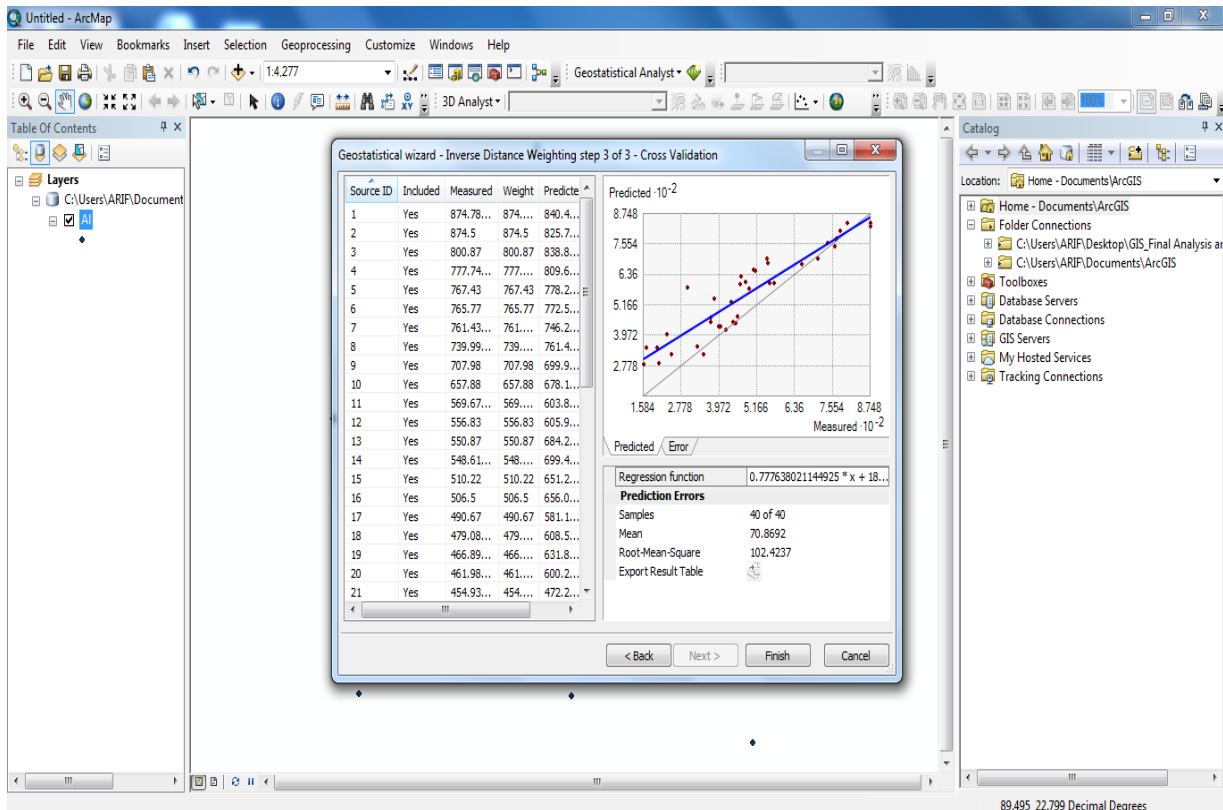
Step 1: “Inverse Distance Weighting” under “Geostatistical Analyst” was chosen. “Data Field” and “Weight Field” was chosen.



Step 2: “Power of the Method” was chosen. ”Next” was clicked to continue.



Step 3: This wizard represented the prediction errors of interpolation techniques in terms of MPE and RMSPE. "Finish" was clicked to continue.



Step 4: "OK" was clicked to continue.

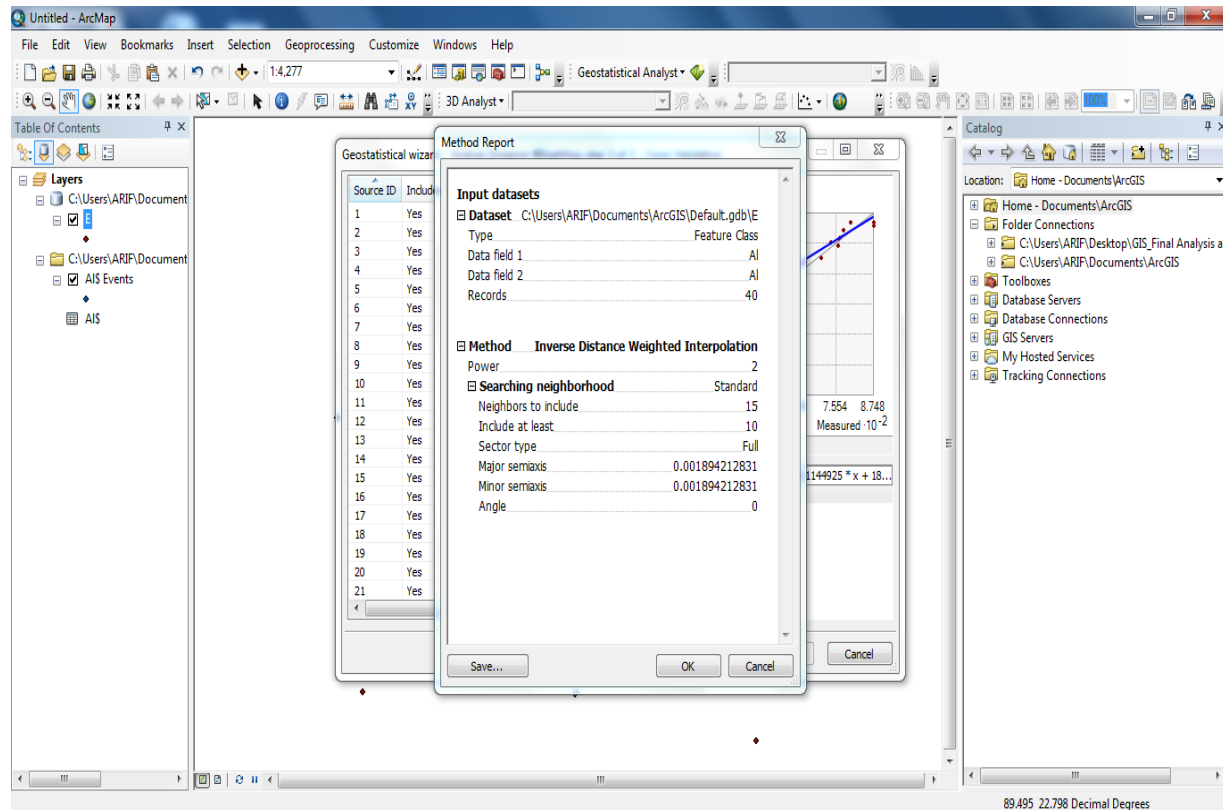
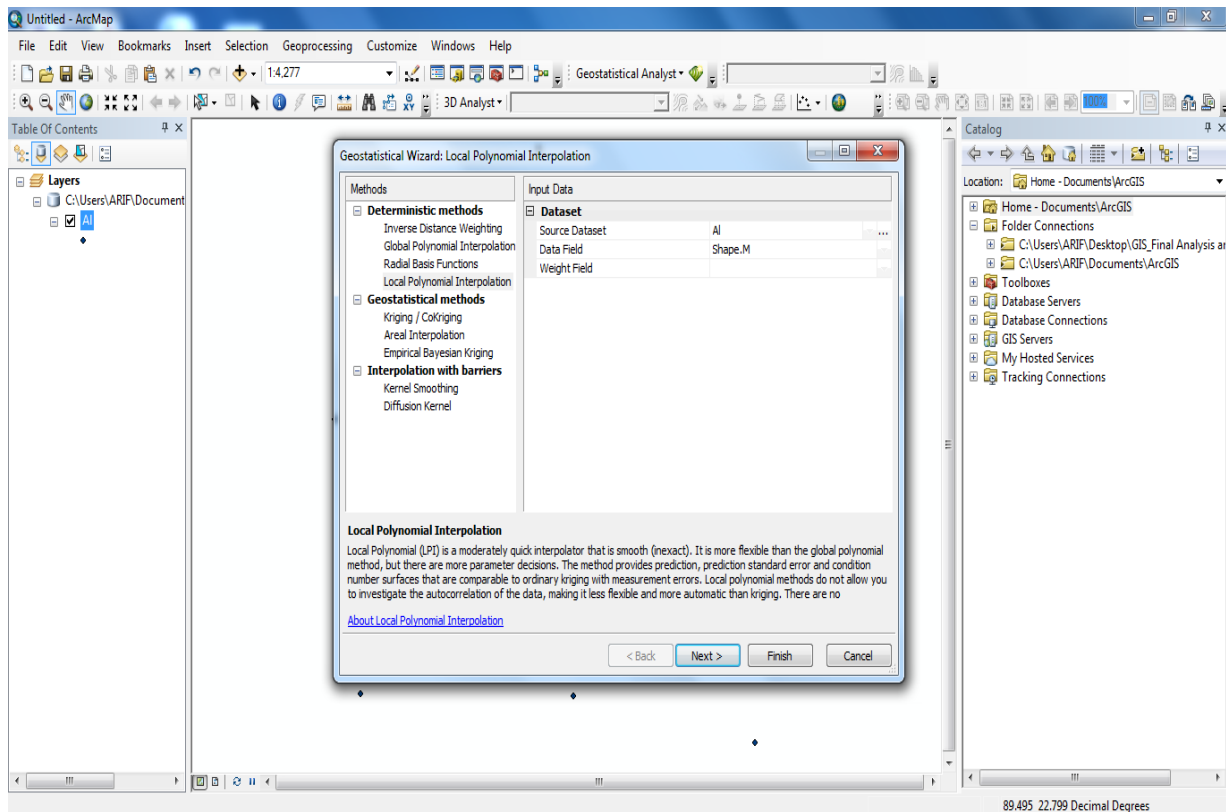


Figure A.8: Stepwise interpolation techniques of IDW.

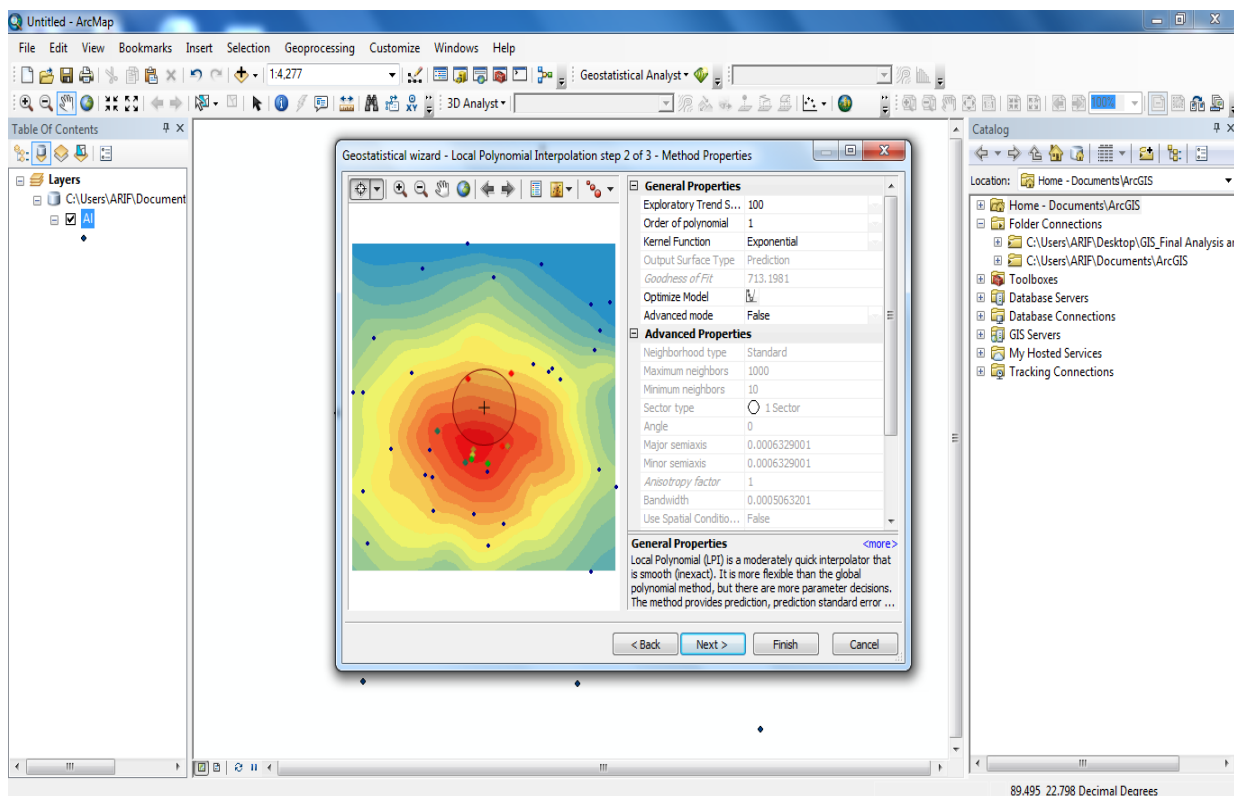


## Steps of Local polynomial Interpolation

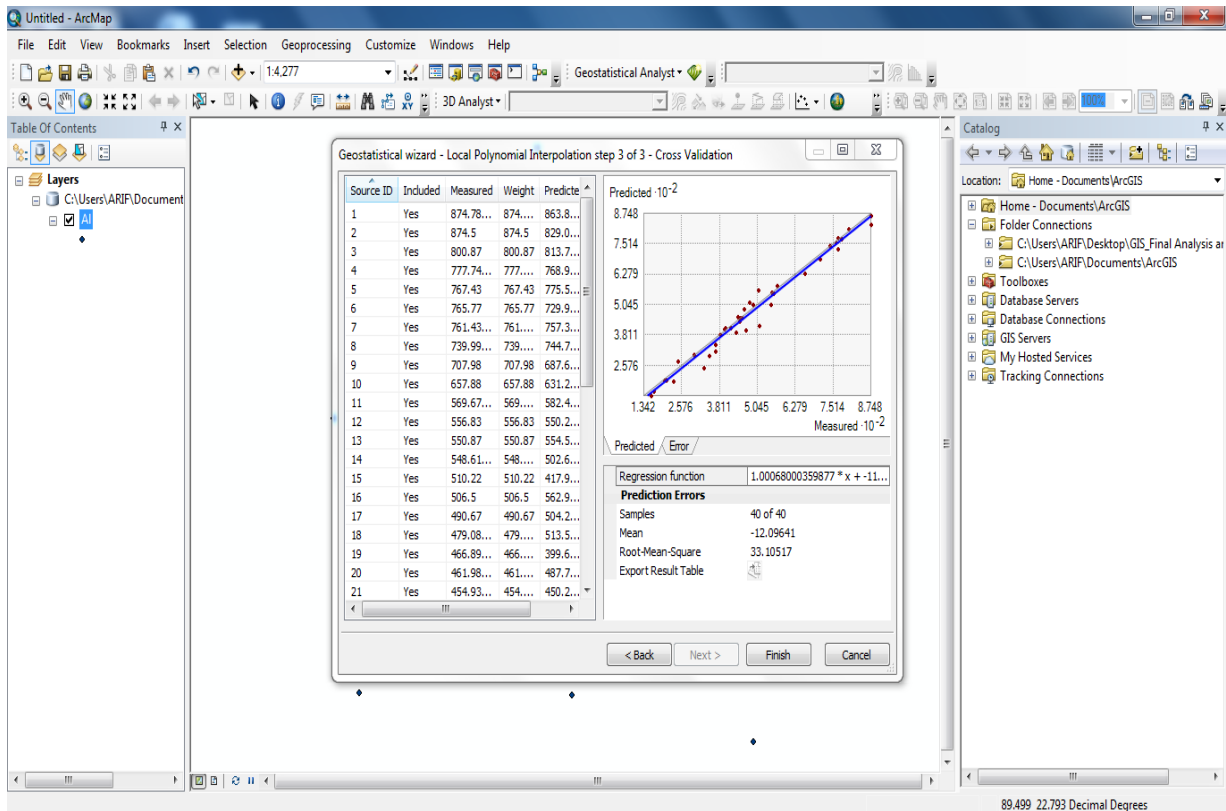
Step 1: “Local Polynomial Interpolation” under “Geostatistical Analyst” was chosen. “Data Field” and “Weight Field” was chosen.



Step 2: “Kernel Function” was chosen. ”Next” was clicked to continue



Step 3: This wizard represented the prediction errors of interpolation techniques in terms of MPE and RMSPE. "Finish" was clicked to continue.



Step 4: "OK" was clicked to continue.

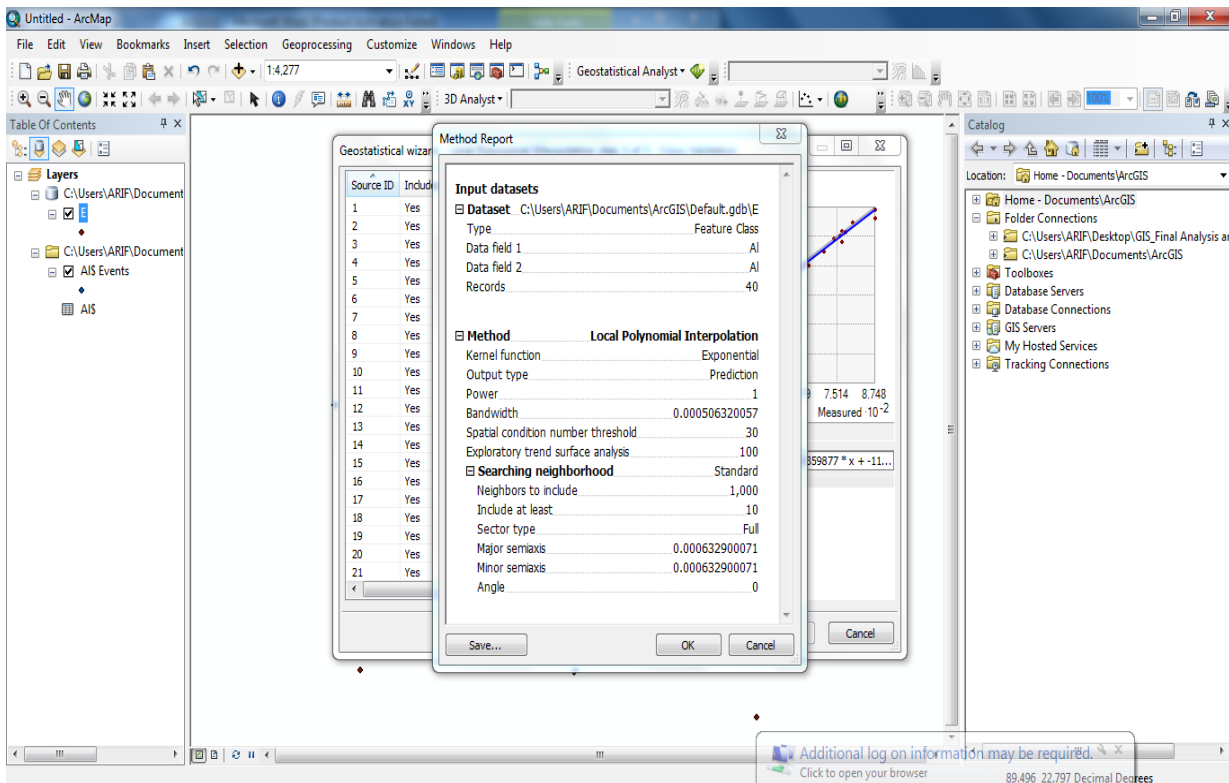
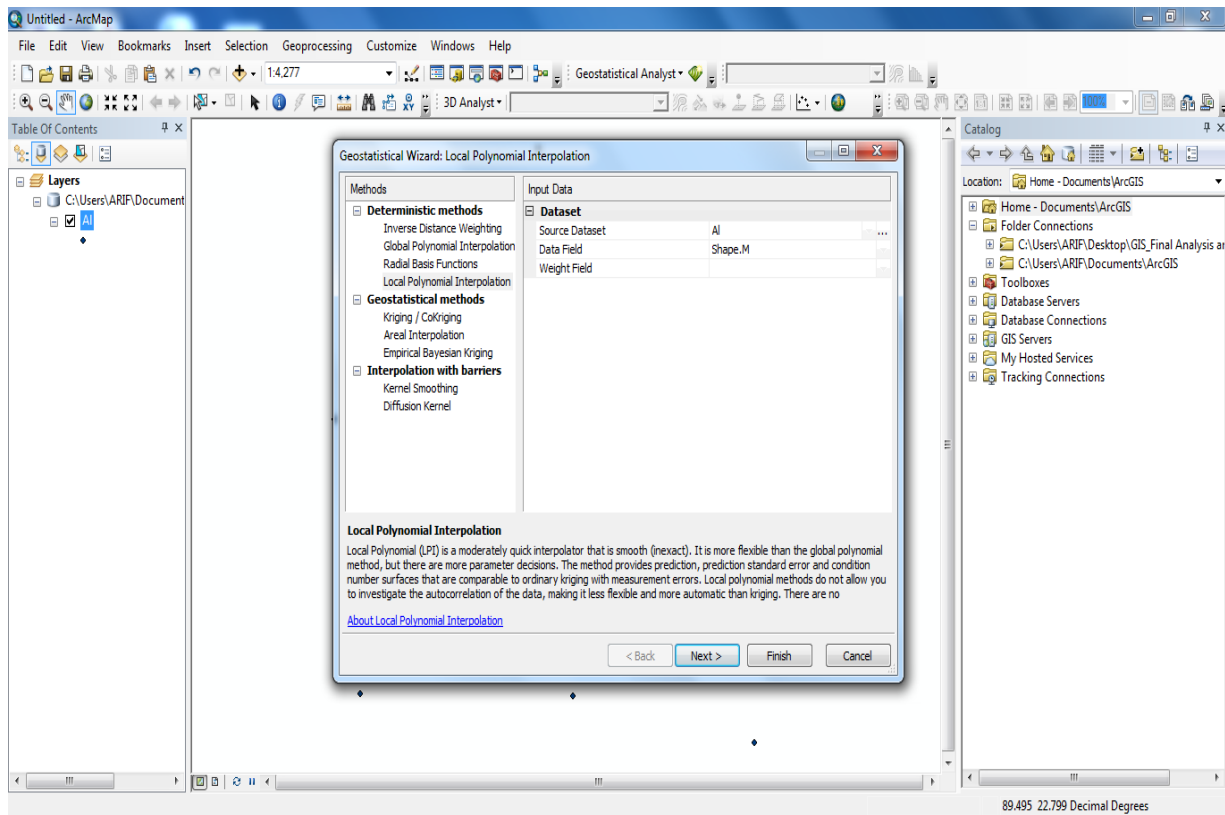


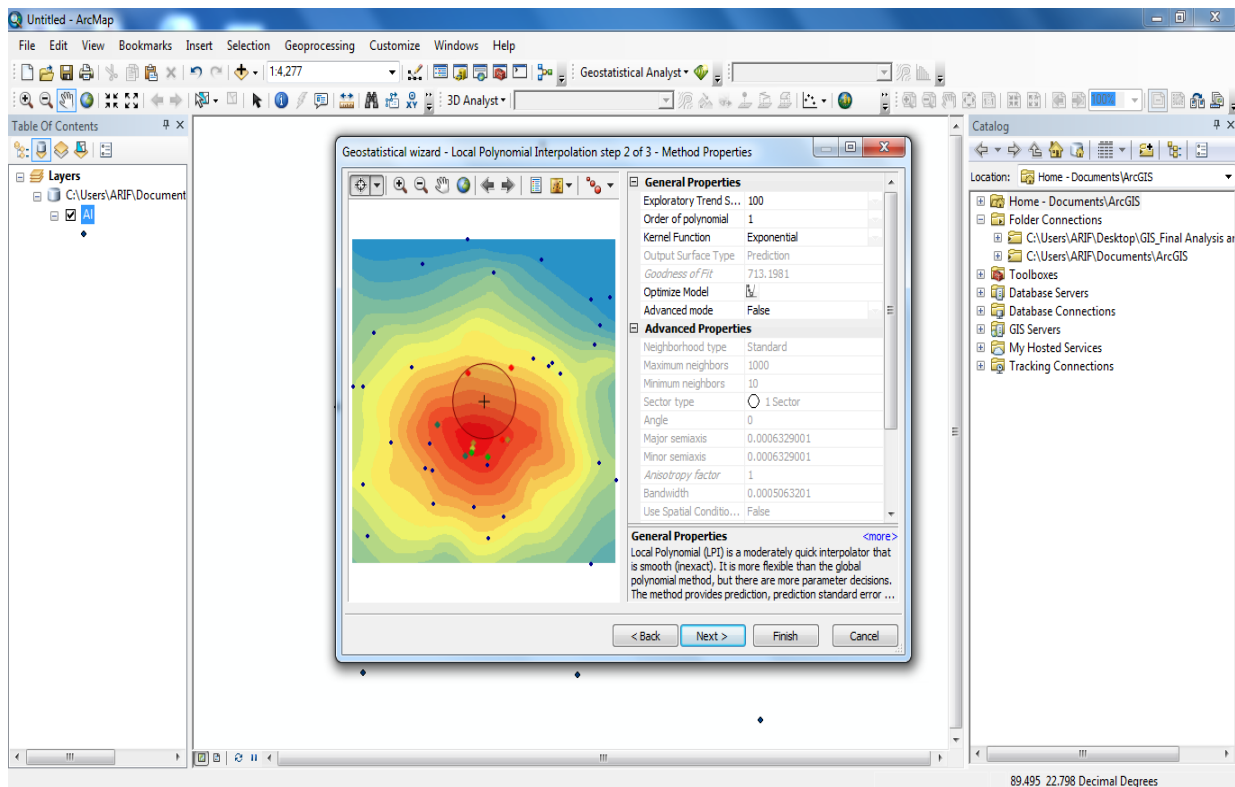
Figure A.9: Stepwise interpolation techniques of LPI.

## Steps of Radial Basis Functions

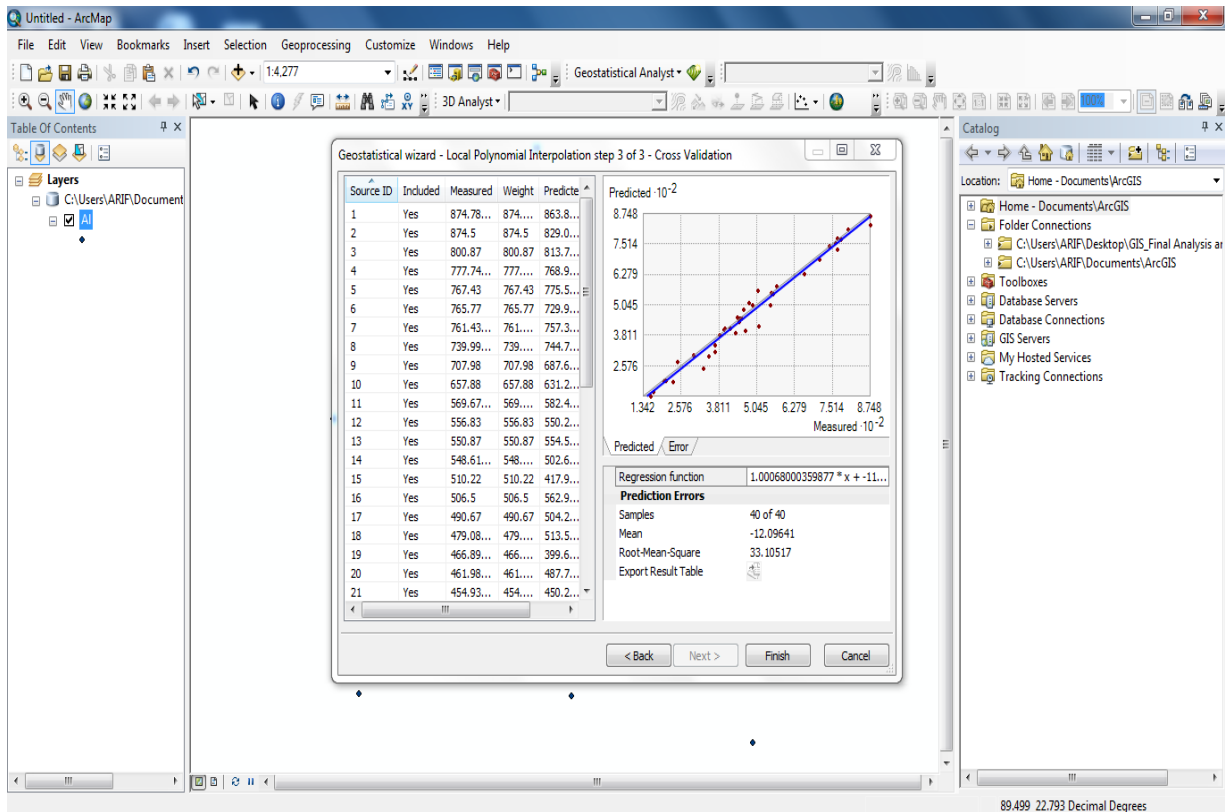
Step 1: “Radial Basis Function’s” under “Geostatistical Analyst” was chosen. “Data Field” and “Weight Field” was chosen.



Step 2: “Kernel Function” was chosen. ”Next” was clicked to continue



Step 3: This wizard represented the prediction errors of interpolation techniques in terms of MPE and RMSPE. "Finish" was clicked to continue.



Step 4: "OK" was clicked to continue.

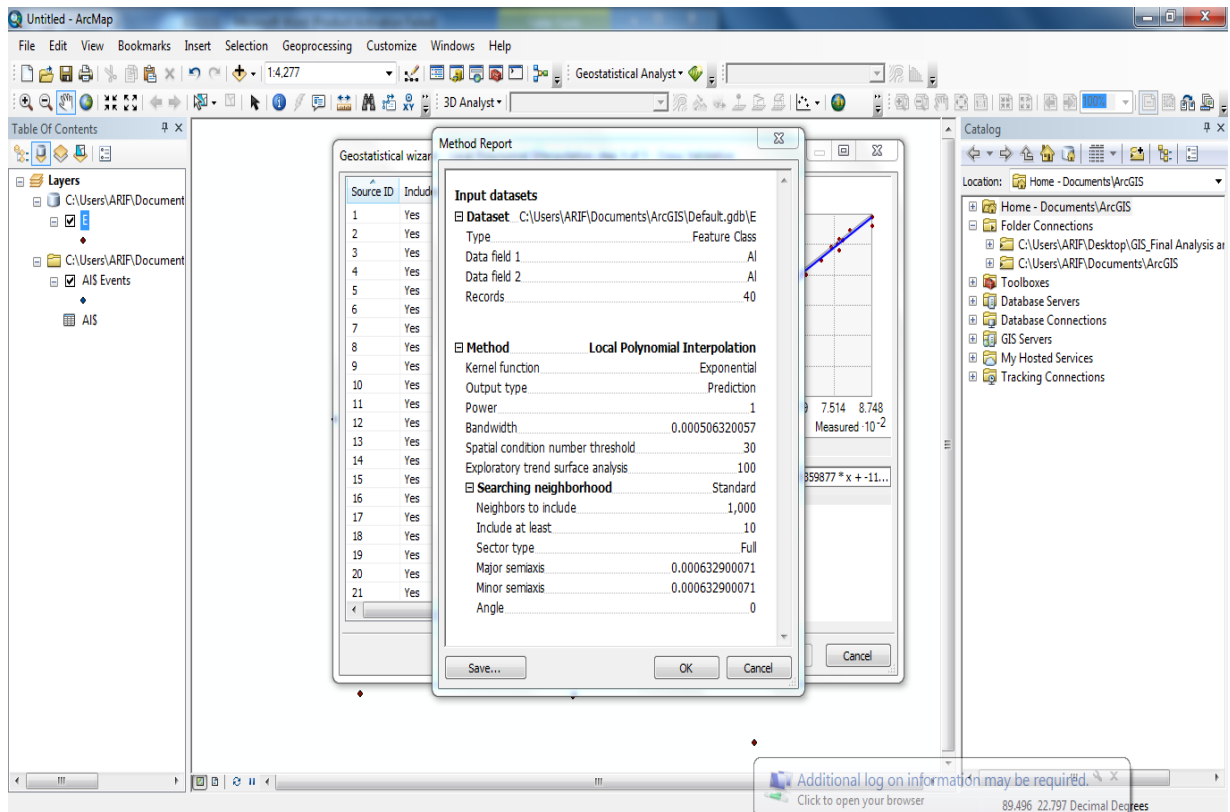
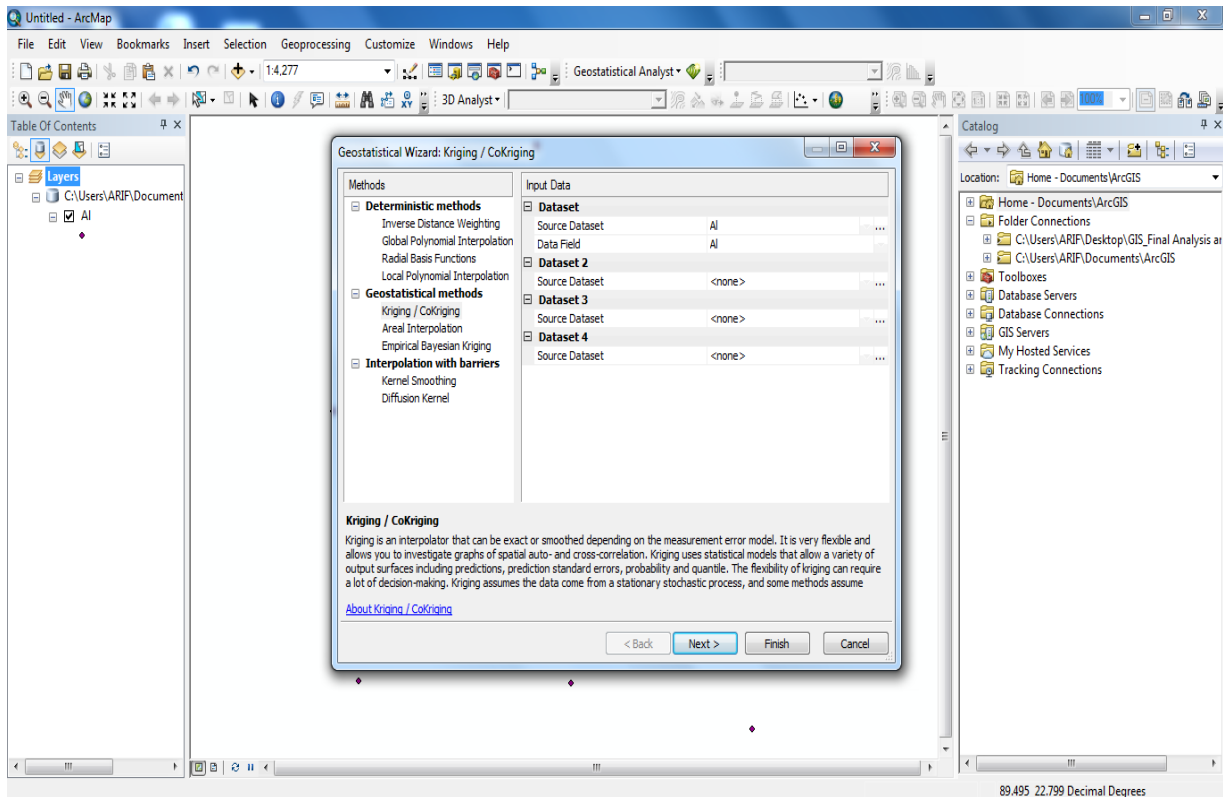


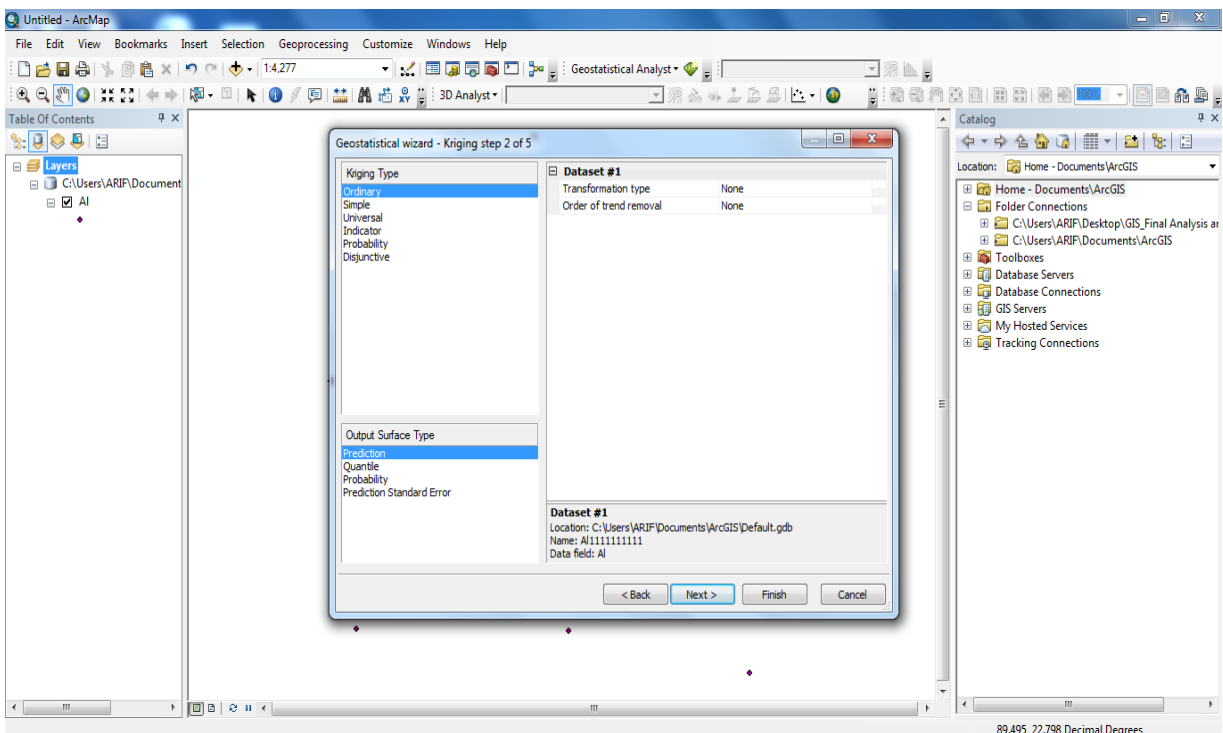
Figure A.10: Stepwise interpolation techniques of RBF's Interpolation.

## Steps of Ordinary Kriging

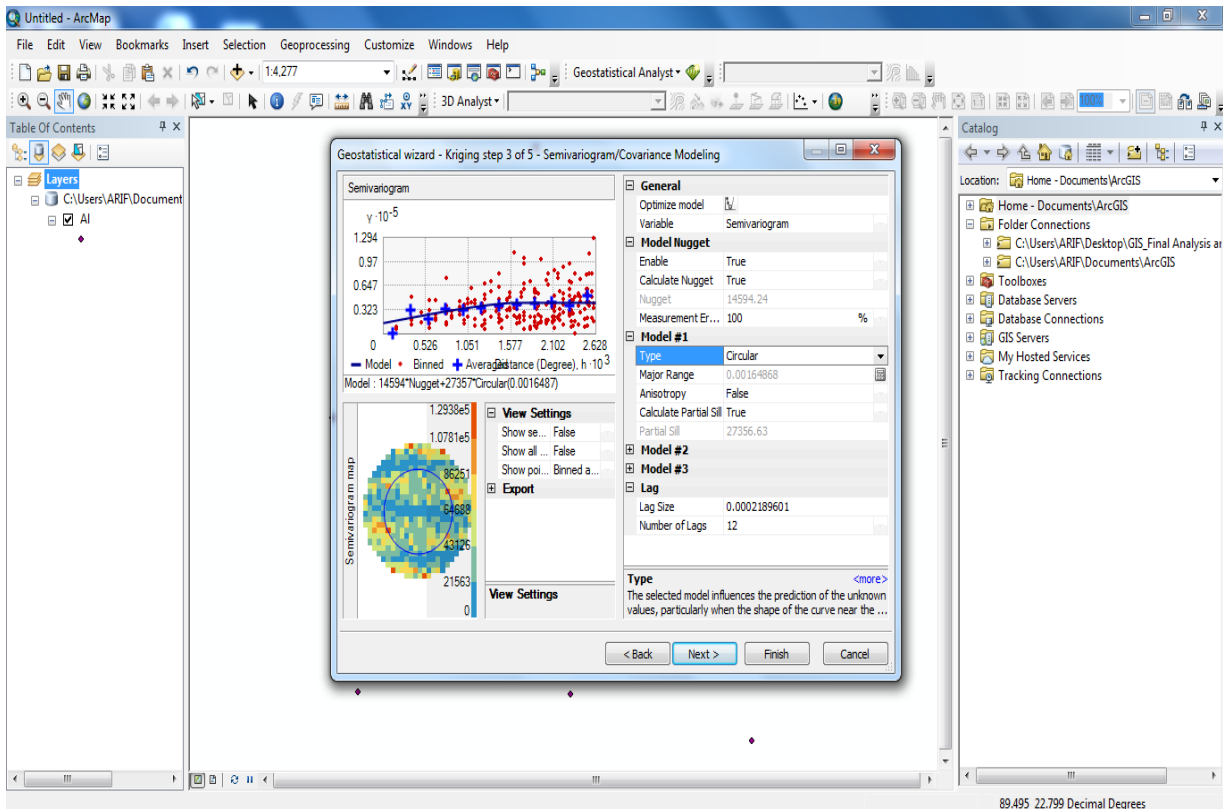
Step 1: “Geostatistical Analyst” wizard was opened and “Kriging/Cokriging” was selected. The inputs were inserted and “Next” were pressed to continue.



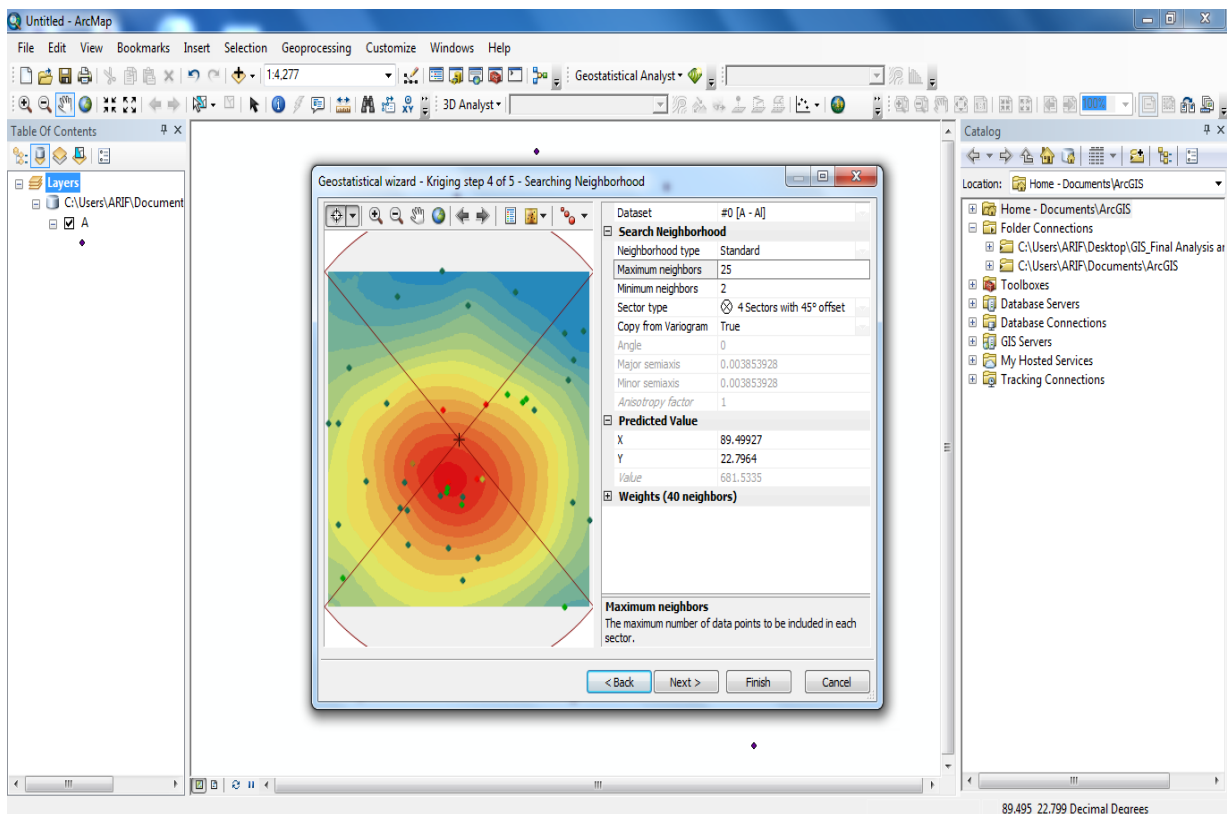
Step 2: “Ordinary Kriging Type” and “Prediction type of output surface” was selected. “Transformation Type” and “Order of trend removal” was selected as “None”. “Next” was pressed to continue.



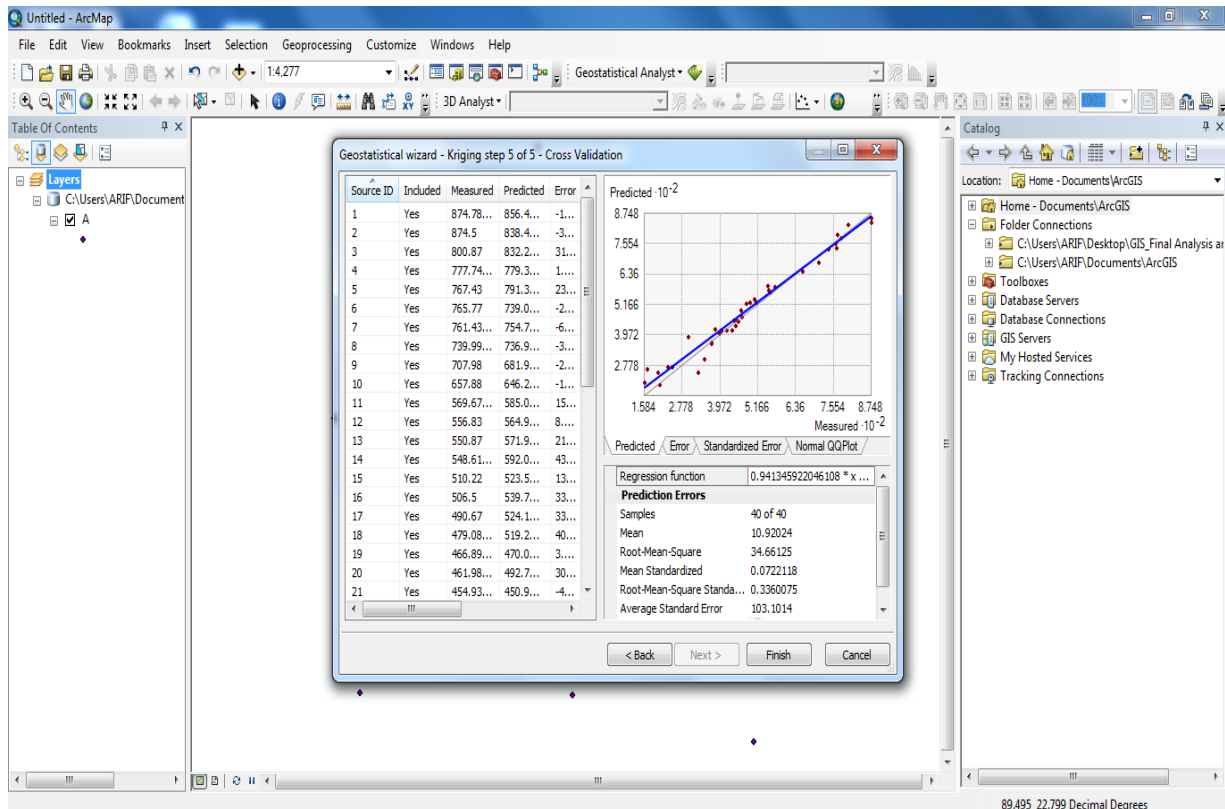
Step 3: “The variable type” was selected as ‘Semivariogram’, The “Model#1’ was chosen. The nugget and partial sill were obtained from this wizard.



Step 4: “Maximum neighbors’ and ‘Minimum neighbors’ were given as input. “Next” was clicked to continue the analysis.



Step 5: This wizard represented the prediction errors of interpolation techniques in terms of MPE, RMSPE, MSPE, RMSSPE and ASPE. "Finish" was clicked to continue.



Step 6: "OK" was clicked to continue.

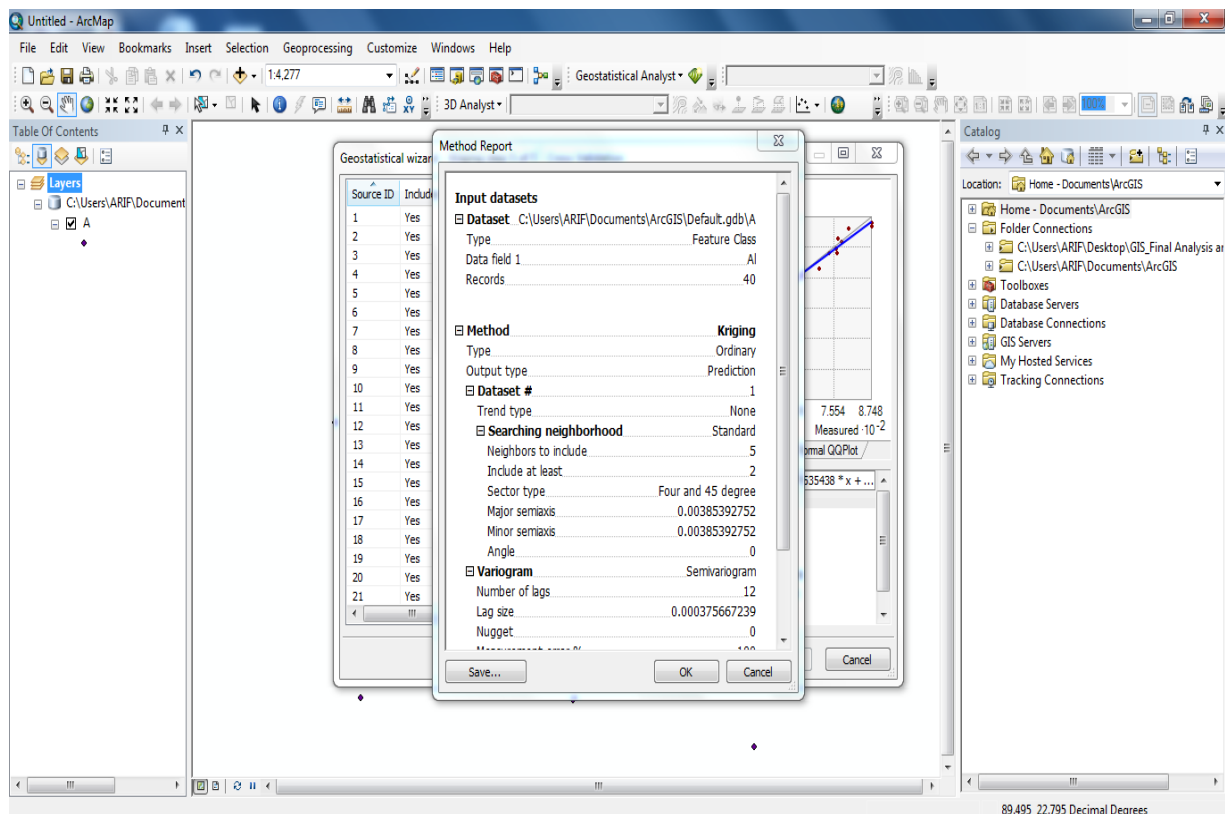
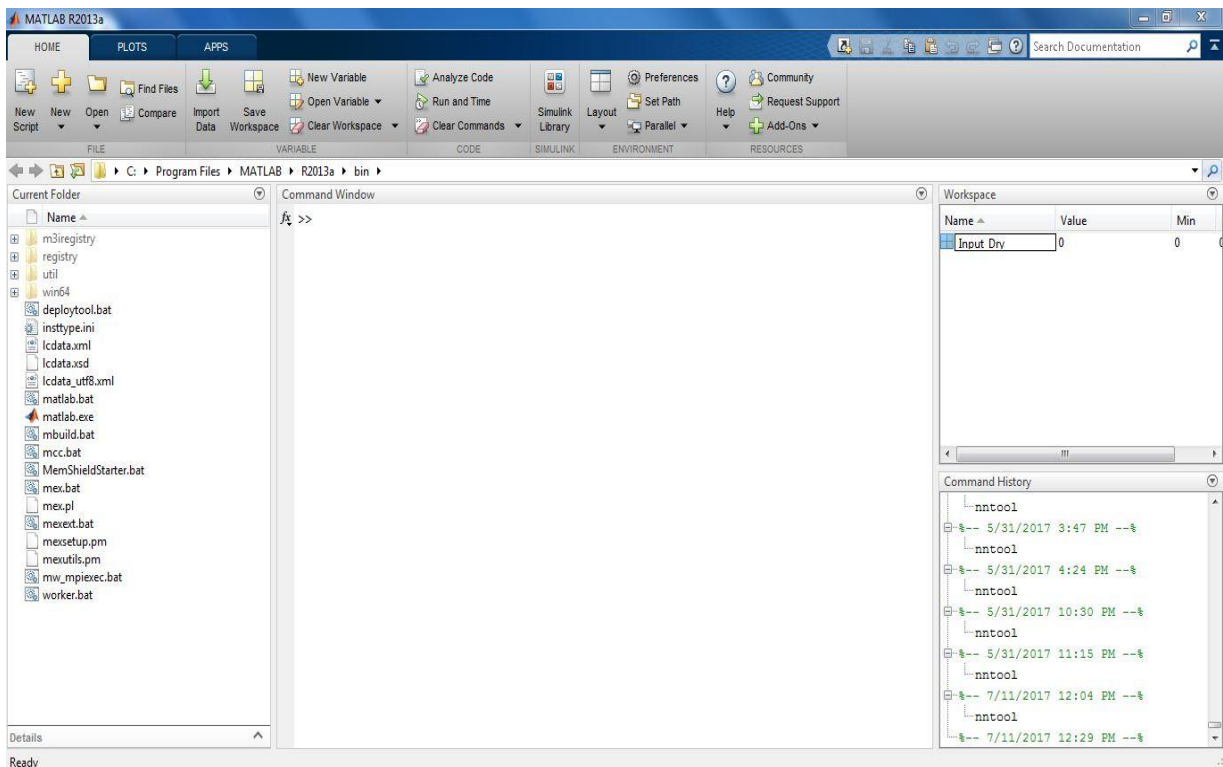


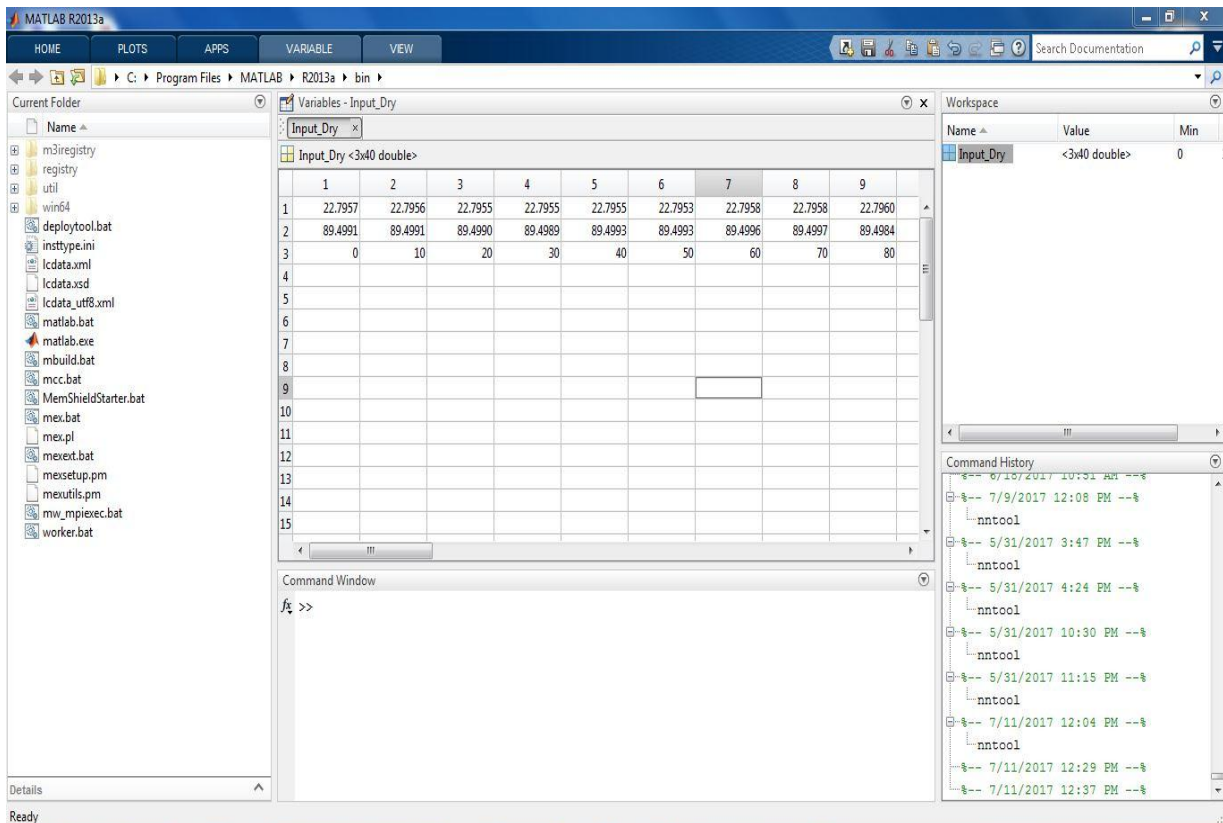
Figure A.11: Stepwise interpolation techniques of ordinary kriging Interpolation.

## Steps of Artificial Neural Network

Step 1: A new file was added in the workspace, renamed as “Input\_Dry”.

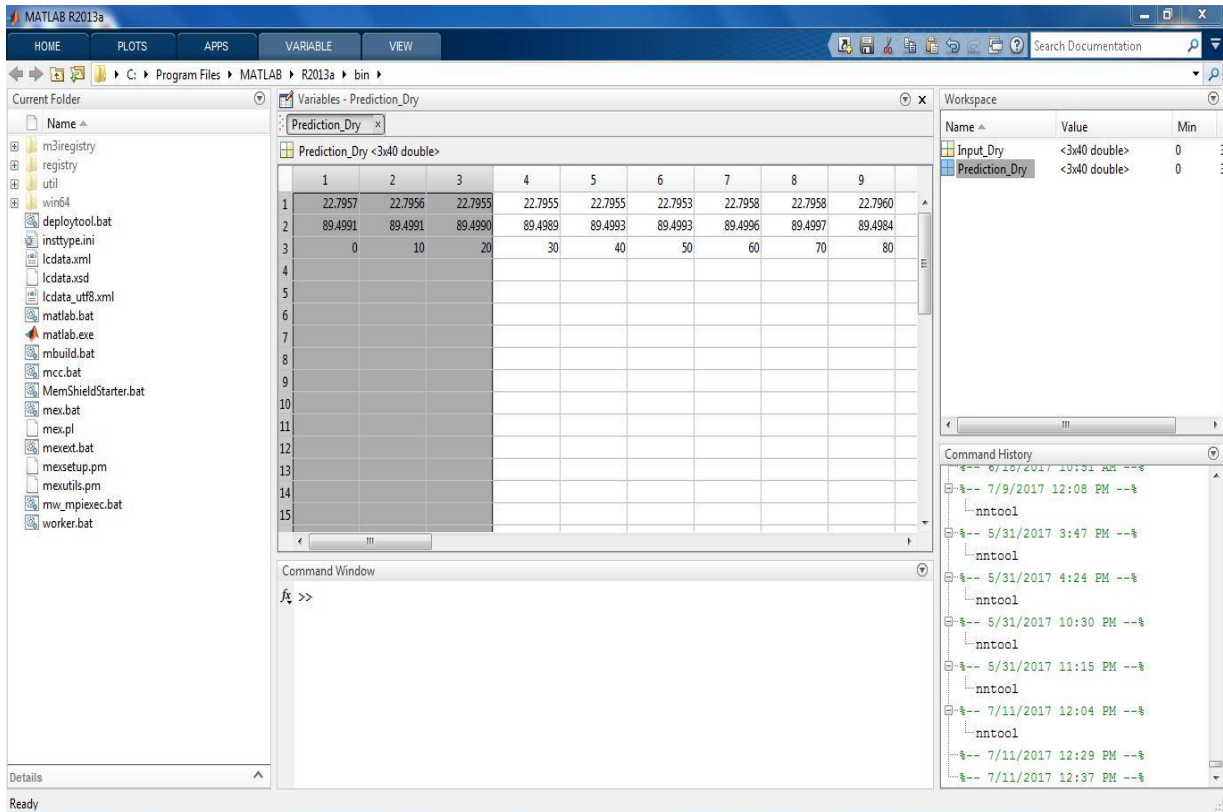


Step 2: The copied data was pasted in the “Input\_Dry” file. The data was transposed in form.

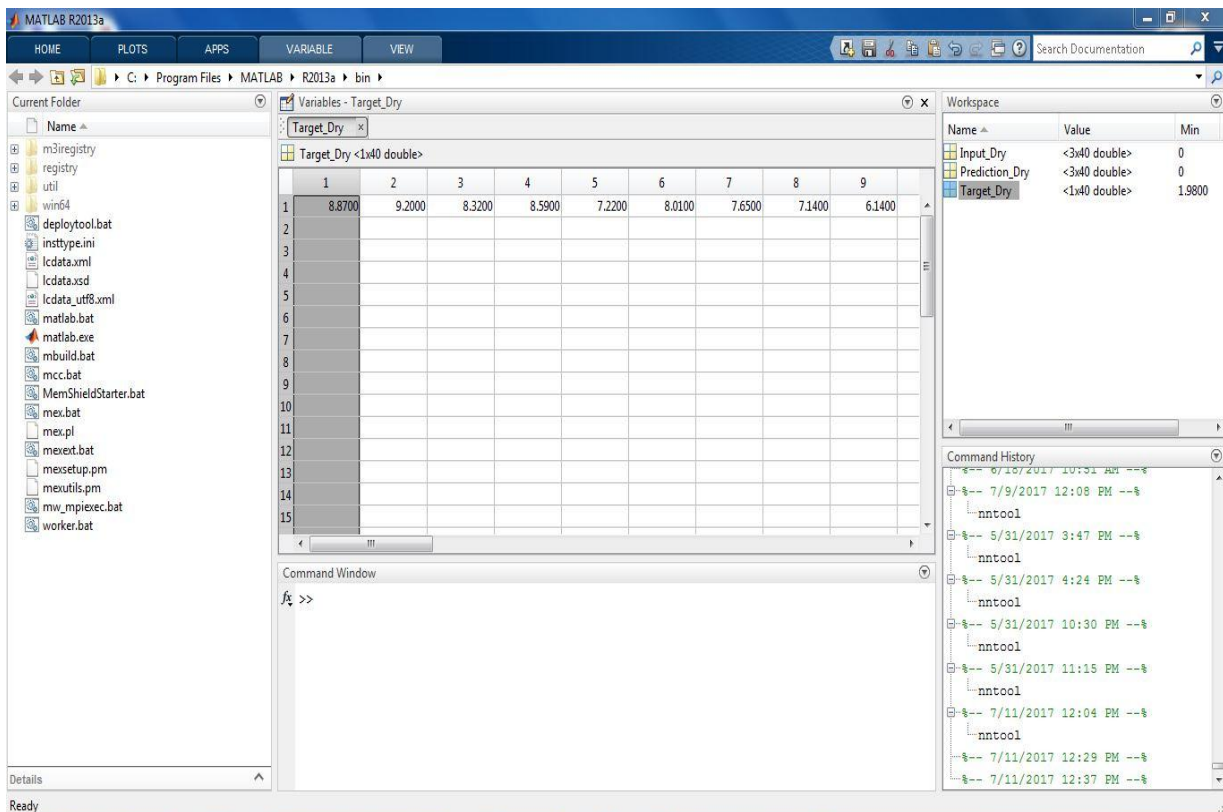




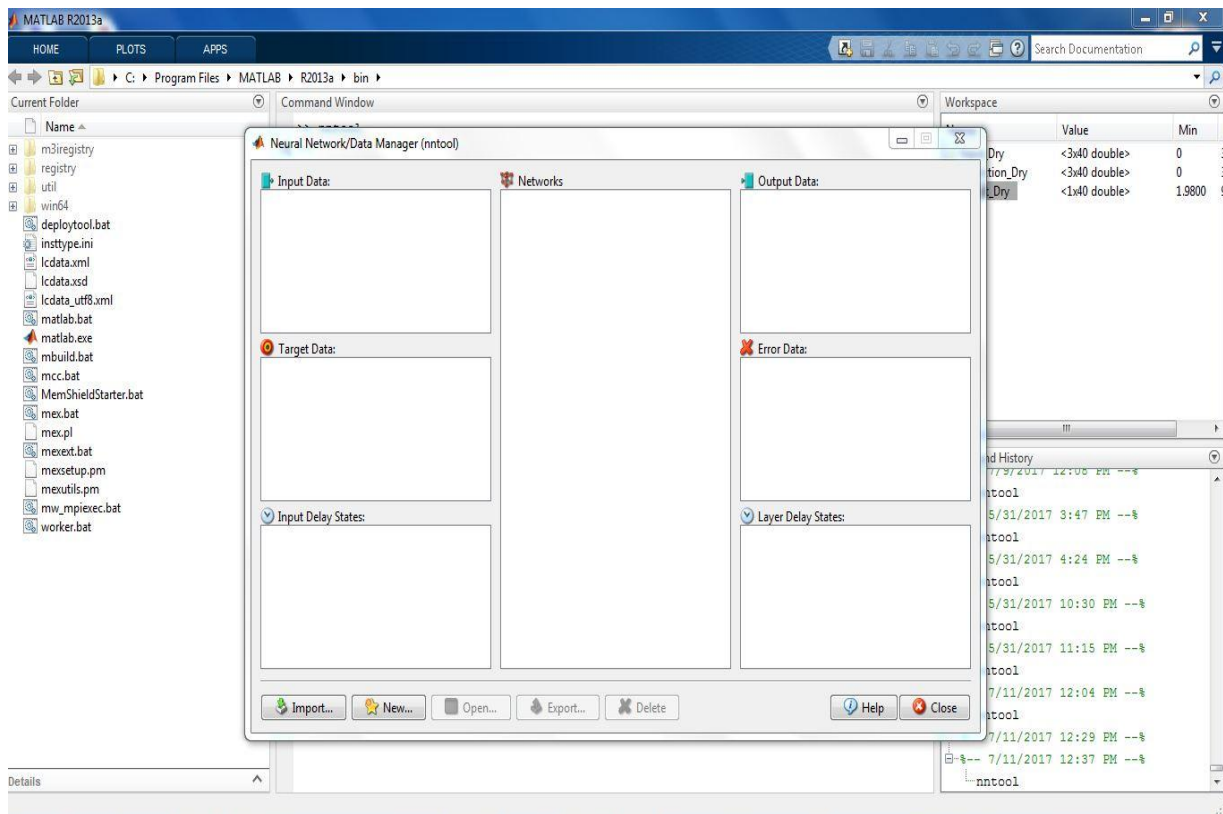
Step 3: A new file was opened as “Prediction\_Dry” and same transpose data was pasted here too.



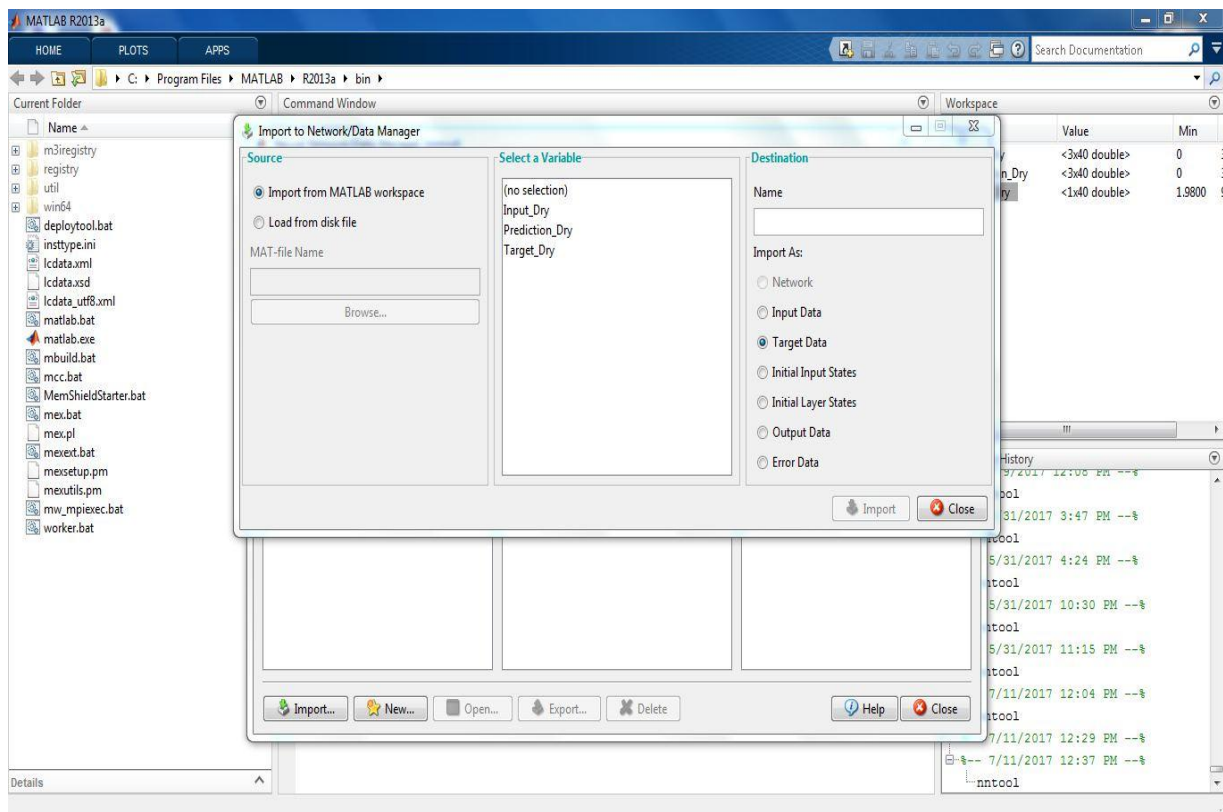
Step 4: A new file was opened as “Target\_Dry” and the concentration of each metal element was pasted here in transpose form.



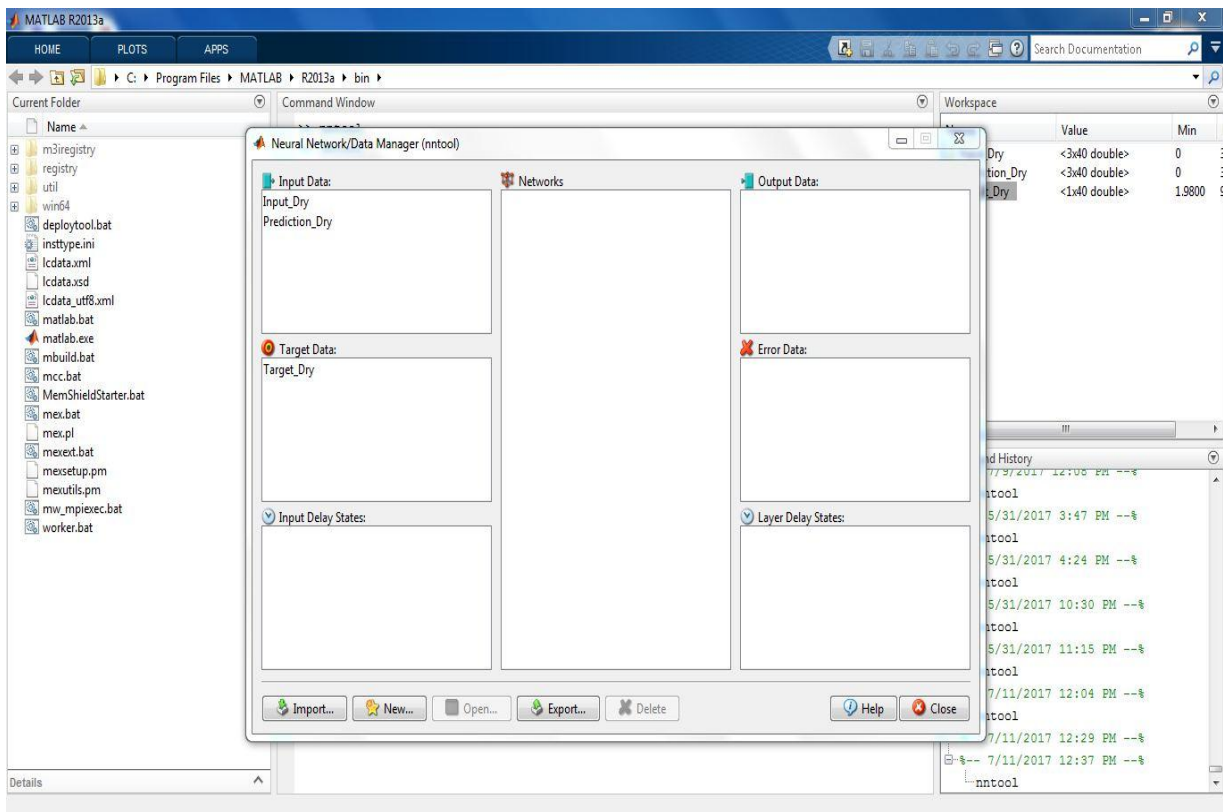
Step 5: “nntool” model was opened to start the analysis.



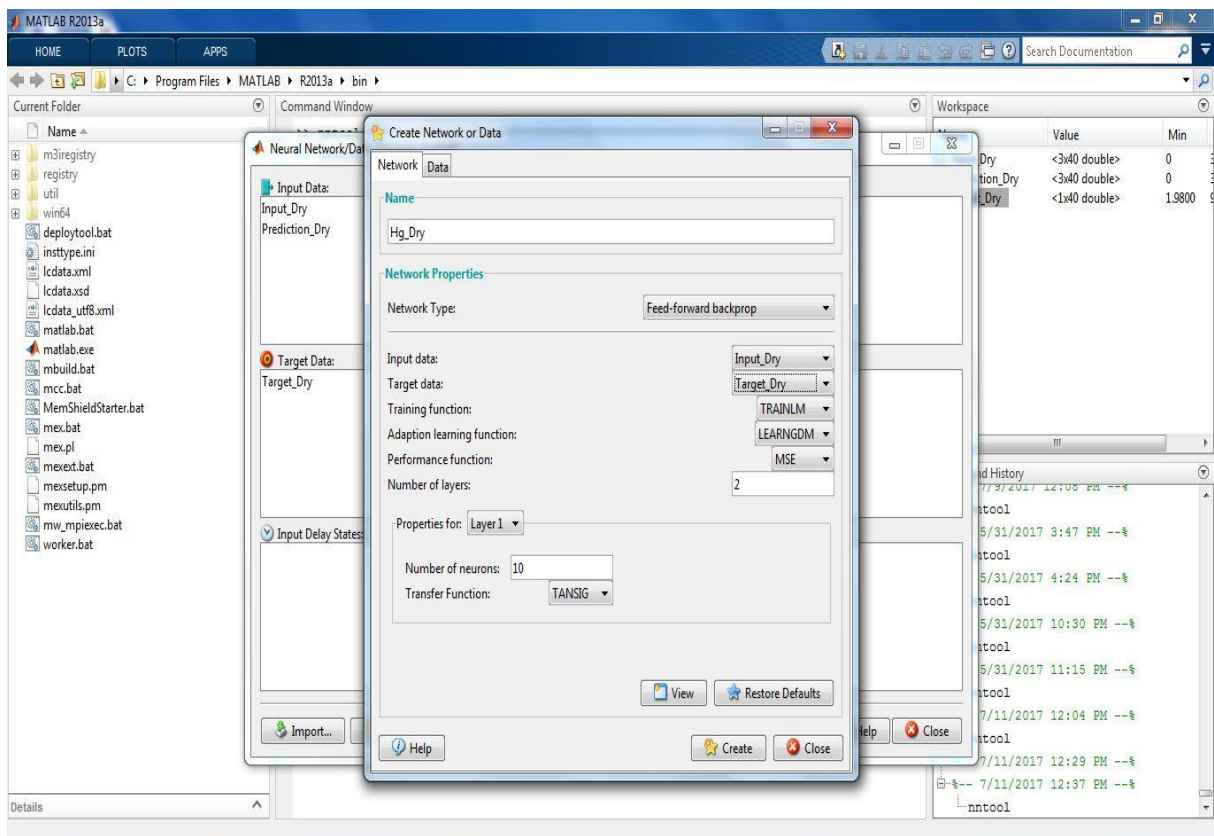
Step 6: The “Input\_Dry, Prediction\_Dry, Target\_Dry” file was selected as input data.



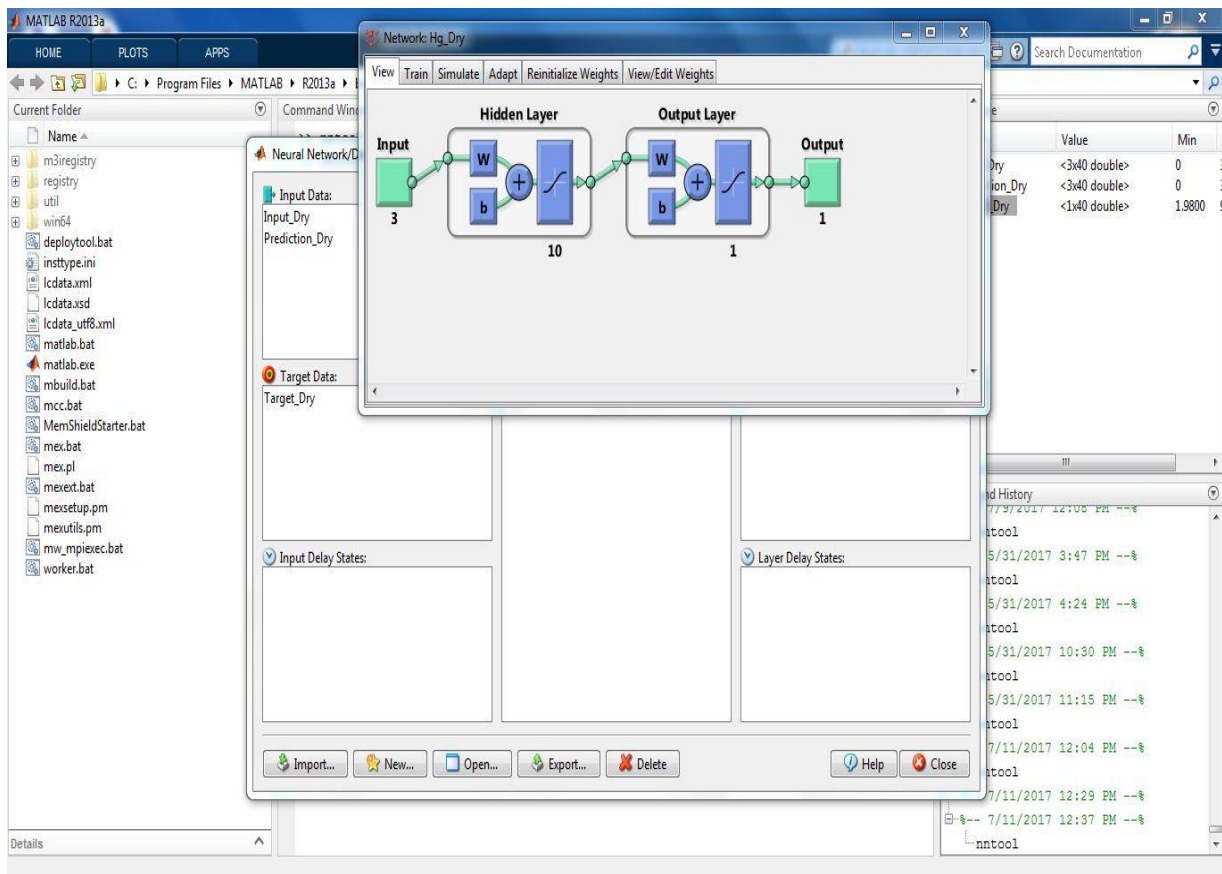
Step 7: The selected files were displayed in the nntool wizard.



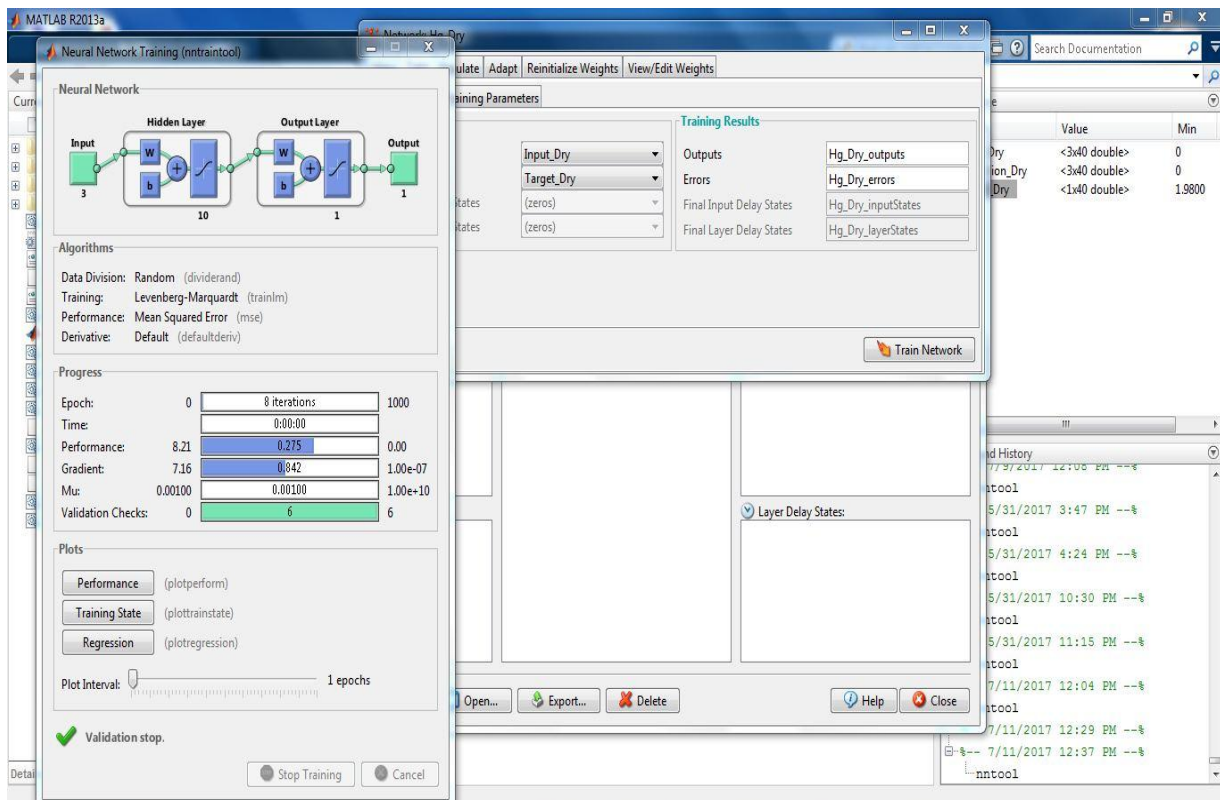
Step 8: A new network “Hg\_Dry” was created, where input and target data was inserted. MSE was selected as performance function.



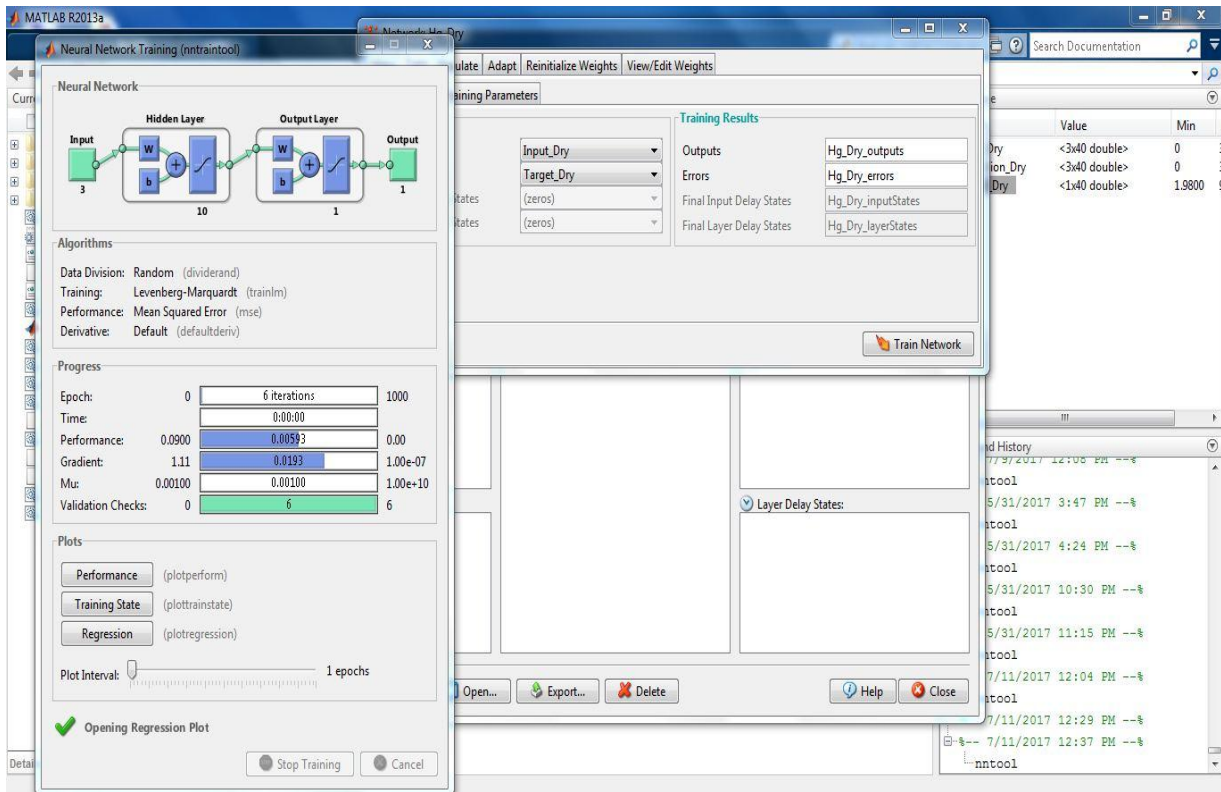
Step 9: A dynamic network was created showing the structure of the analysis.



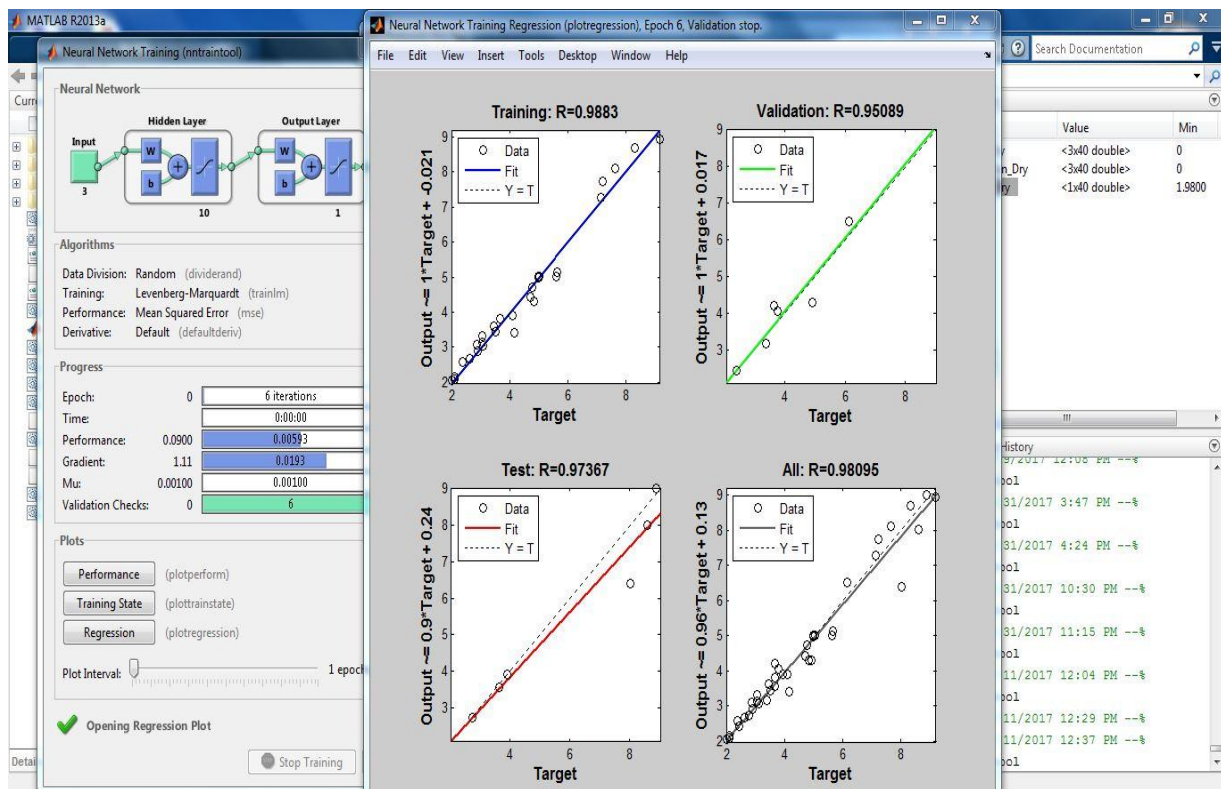
Step 10: Neural Network Training was started where “Input\_Dry” was treated as “Input data: and “Target\_Dry” as target data.



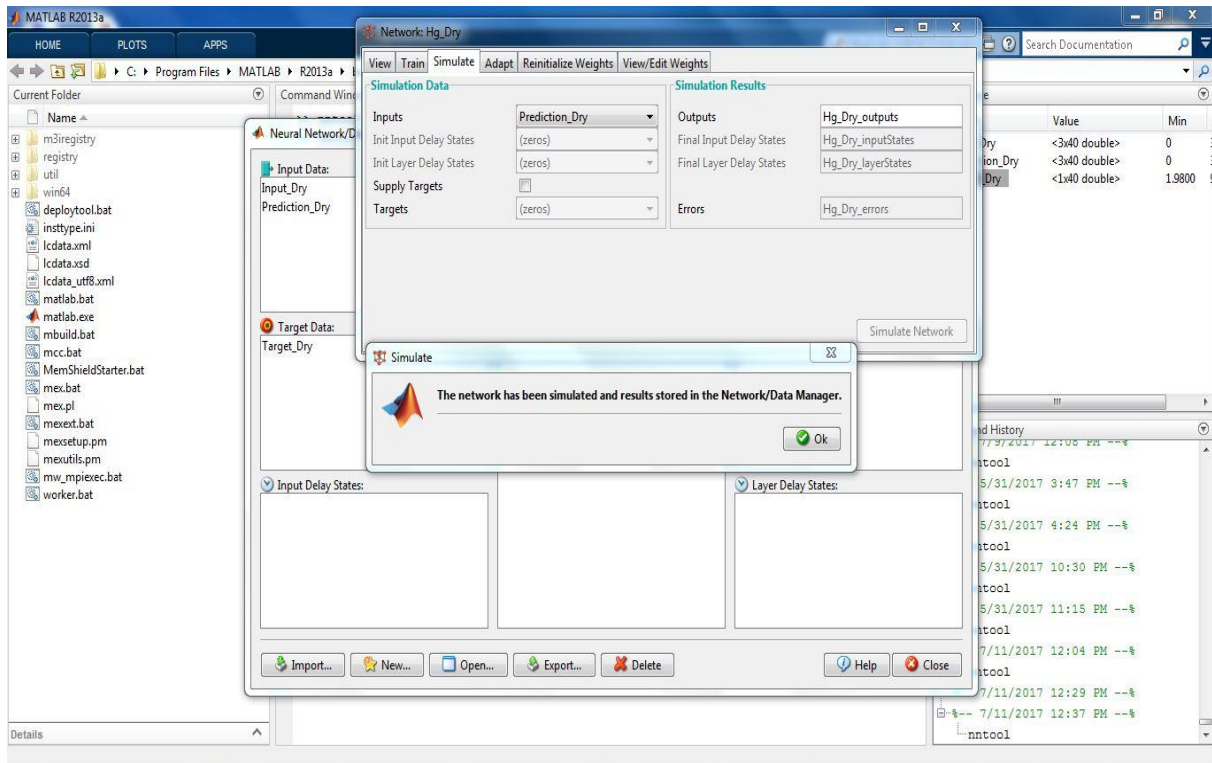
Step 11: The network was trained several times to get the best result. It was found that the MSE value was 0.0900, close to 0, indicate a better prediction.



Step 12: Regression coefficient was obtained for training, validation, test and overall method, respectively. The R-value was found 0.98095, close to 1. This also indicated a better prediction.



Step 13: Prediction of future data was estimated by simulating the network.



Step 14: The output was exported to the workspace in matrix form.

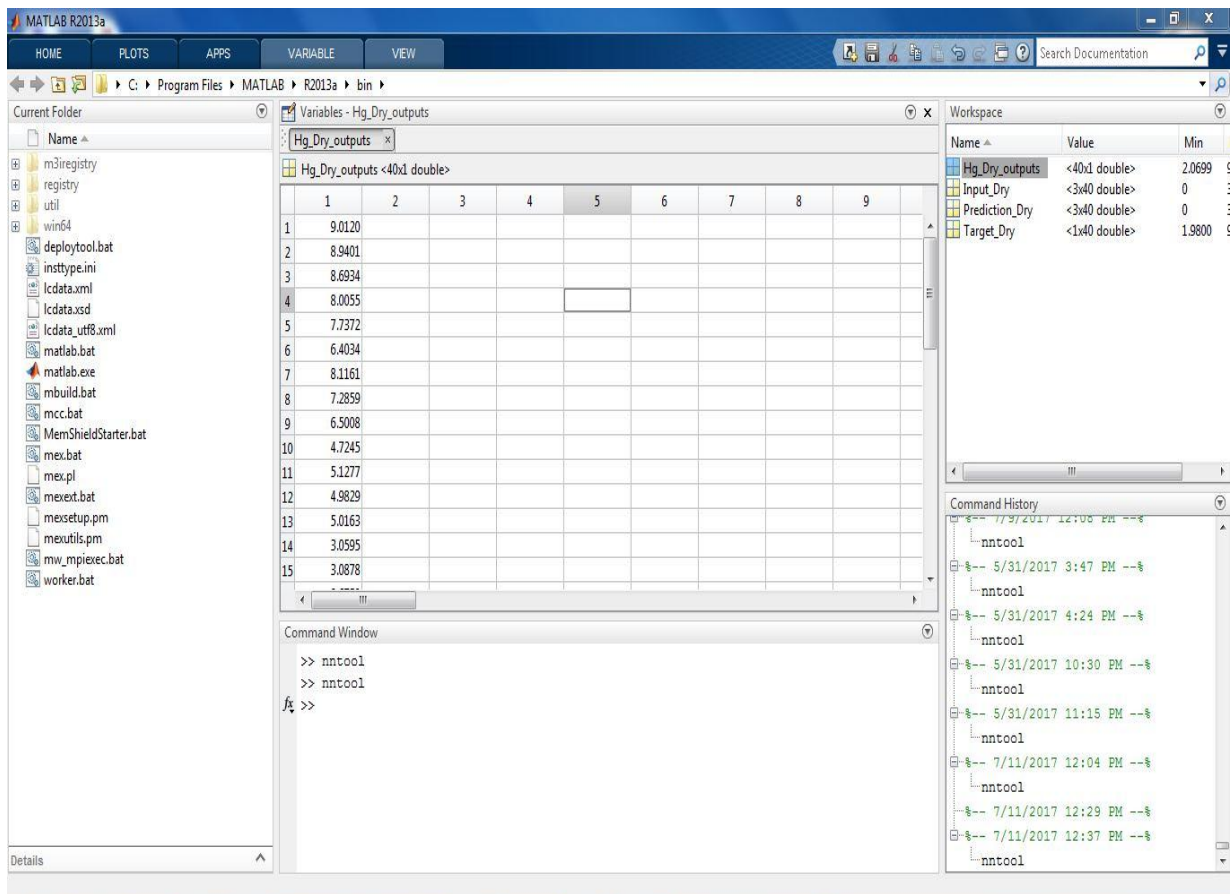


Figure A.12: Stepwise analysis of nntool model using ANN.

## Annex- B

### Representation of Normal Q-Q Plot and Self Organizing Map (SOM) of Metal Elements using XLSTAT

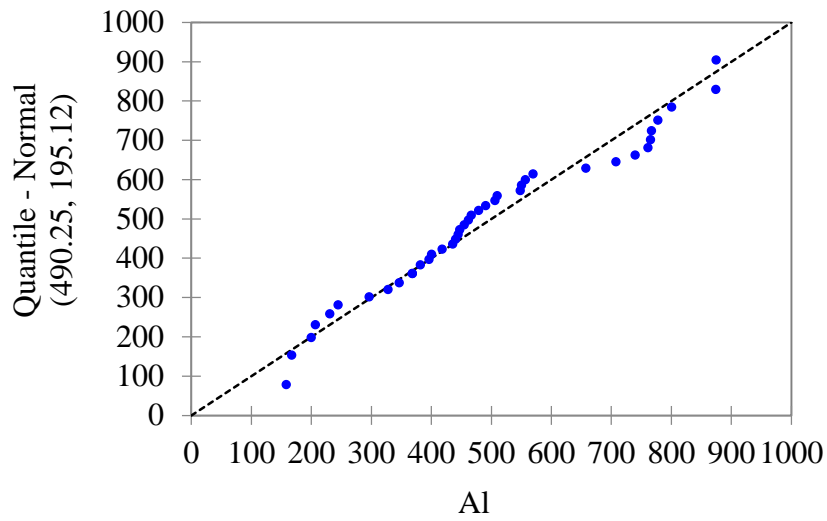


Figure B.1: Normal QQ plot of Al for the dry season (n=40).

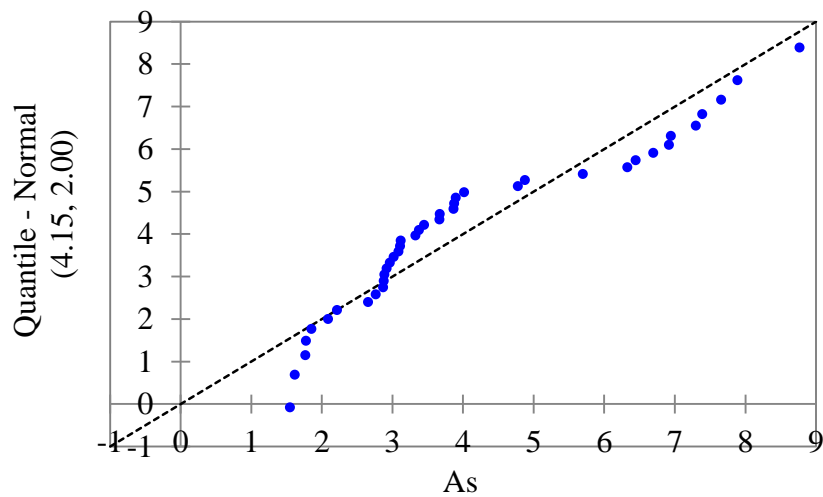


Figure B.2: Normal QQ plot of As for the dry season (n=40).

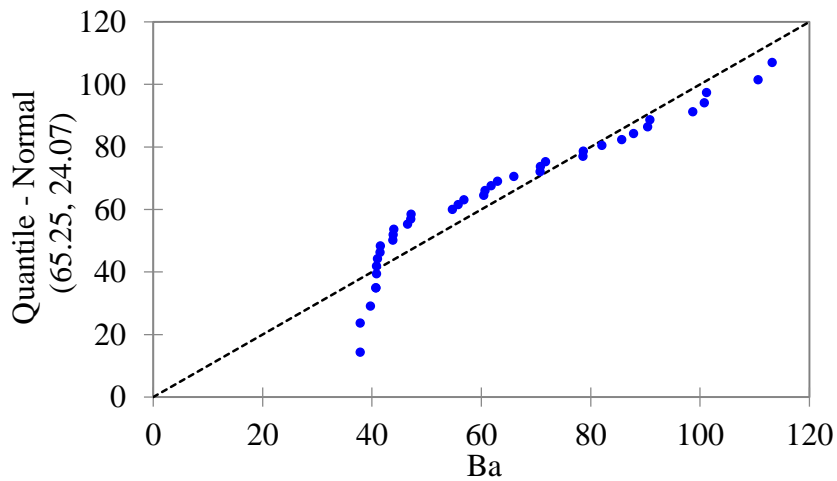


Figure B.3: Normal QQ plot of Ba for the dry season (n=40).

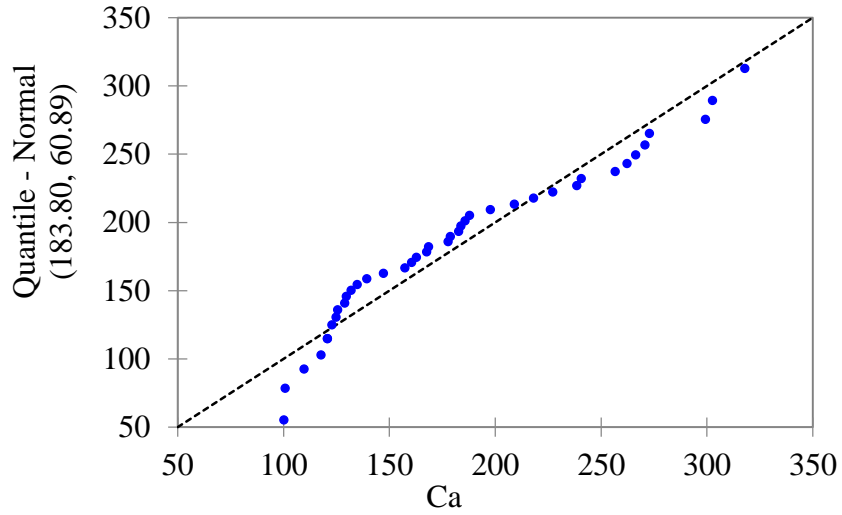


Figure B.4: Normal QQ plot of Ca for the dry season (n=40).

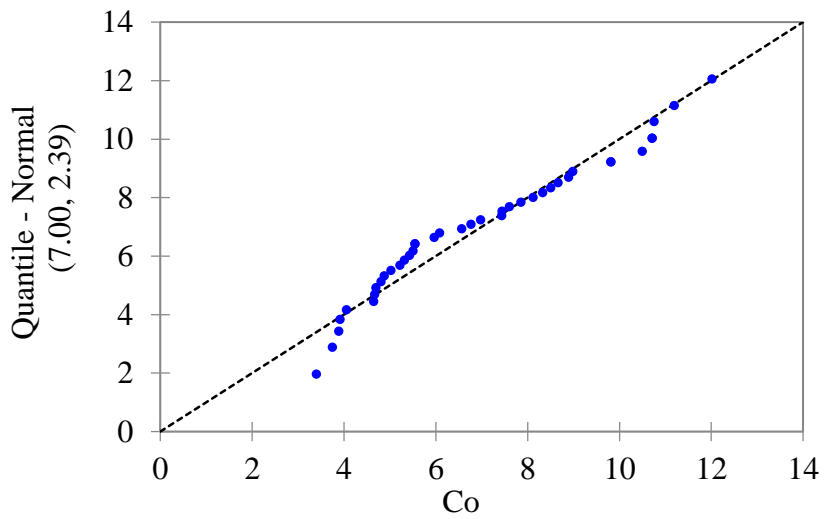


Figure B.5: Normal QQ plot of Co for the dry season (n=40).

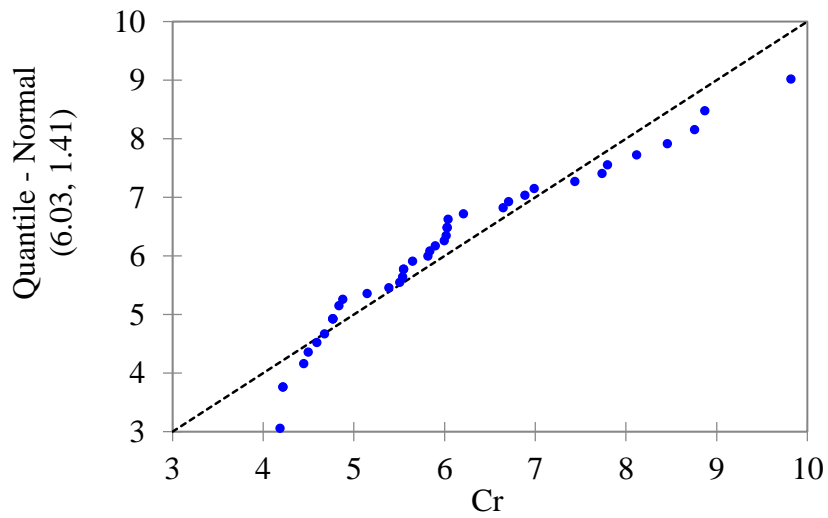


Figure B.6: Normal QQ plot of Cr for the dry season (n=40).



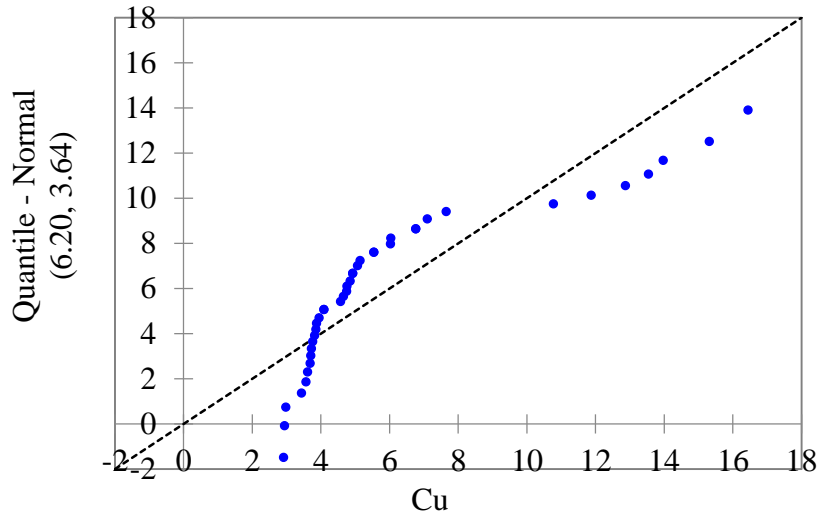


Figure B.7: Normal QQ plot of Cu for the dry season (n=40).

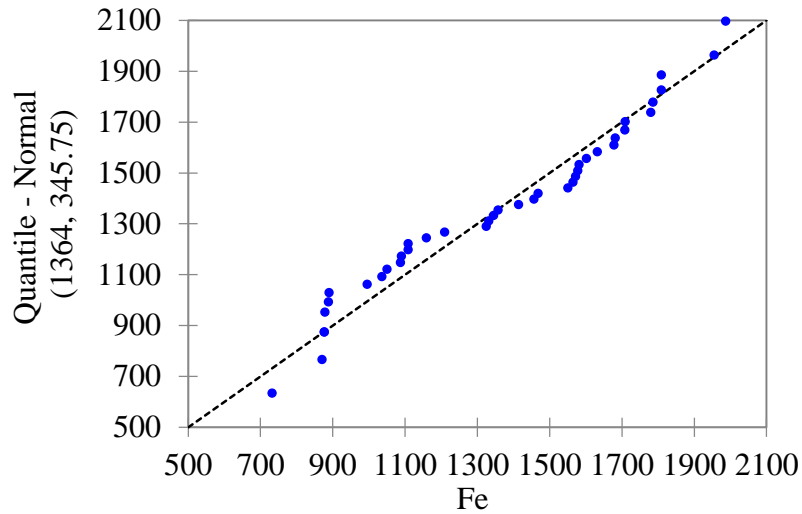


Figure B.8: Normal QQ plot of Fe for the dry season (n=40).

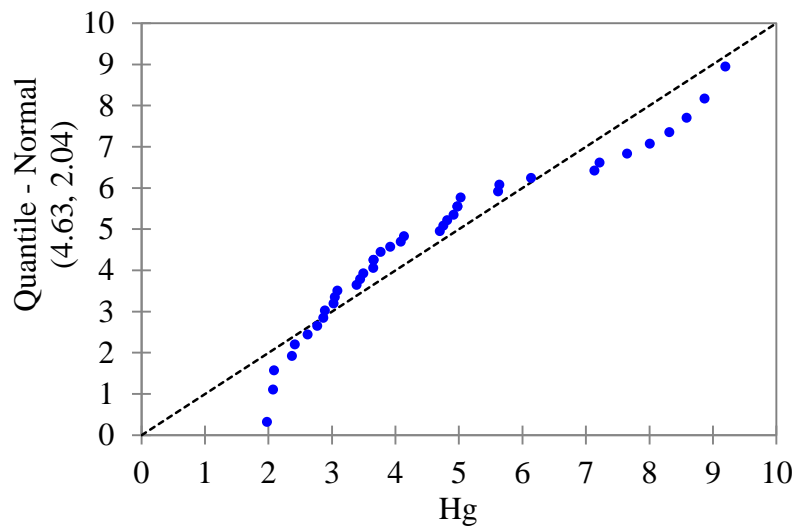


Figure B.9: Normal QQ plot of Hg for the dry season (n=40).

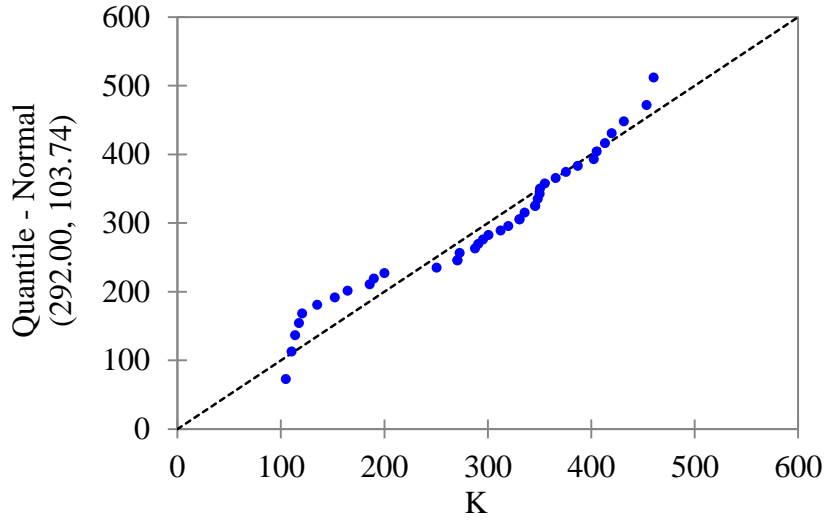


Figure B.10: Normal QQ plot of K for the dry season (n=40).

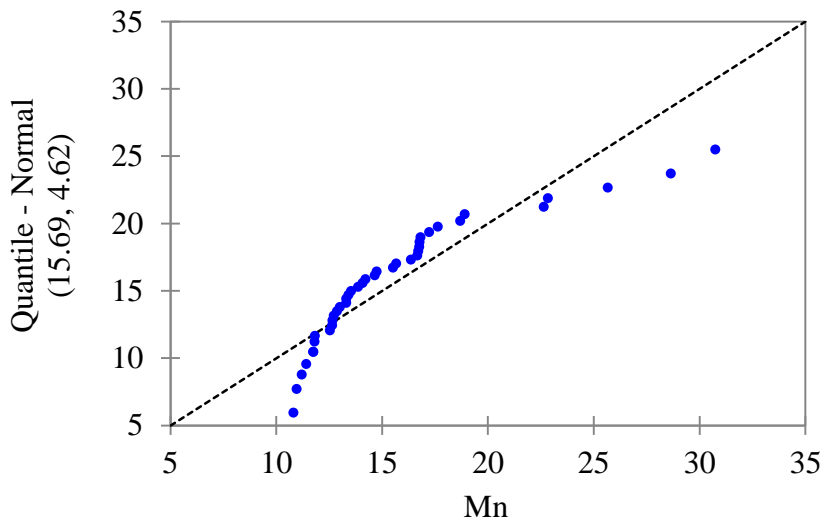


Figure B.11: Normal QQ plot of Mn for the dry season (n=40).

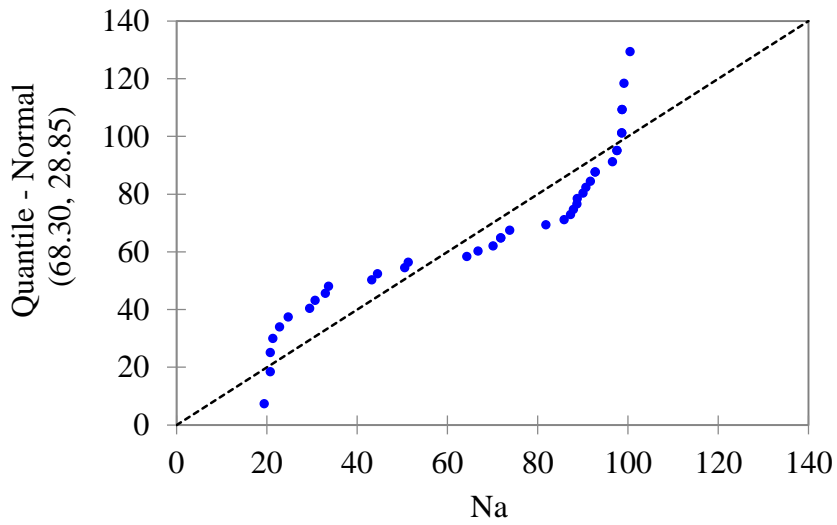


Figure B.12: Normal QQ plot of Na for the dry season (n=40).

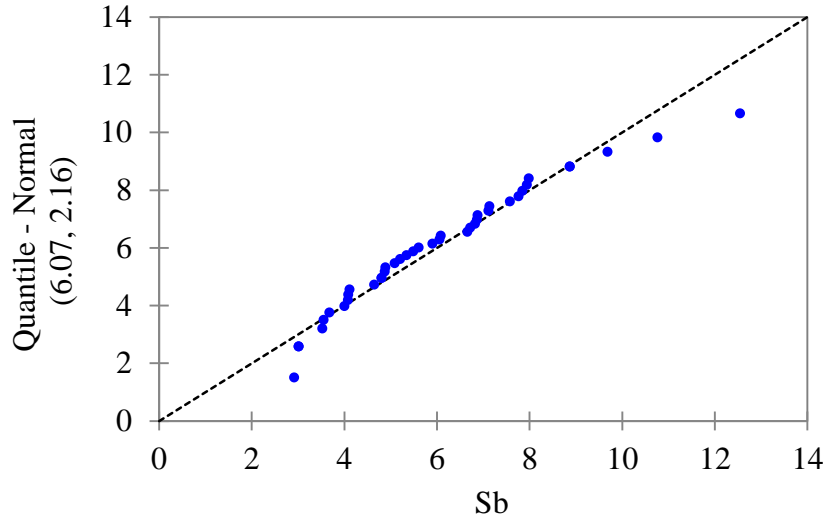


Figure B.13: Normal QQ plot of Sb for the dry season (n=40).

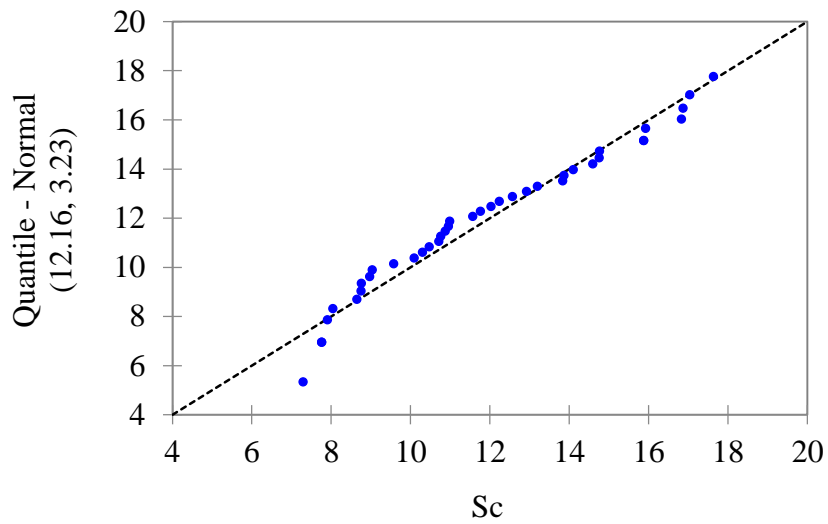


Figure B.14: Normal QQ plot of Sc for the dry season (n=40).

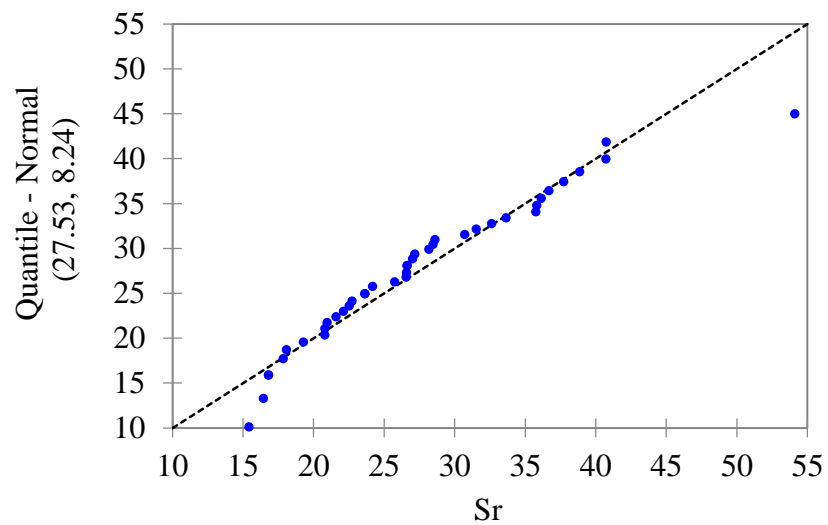


Figure B.15: Normal QQ plot of Sr for the dry season (n=40).

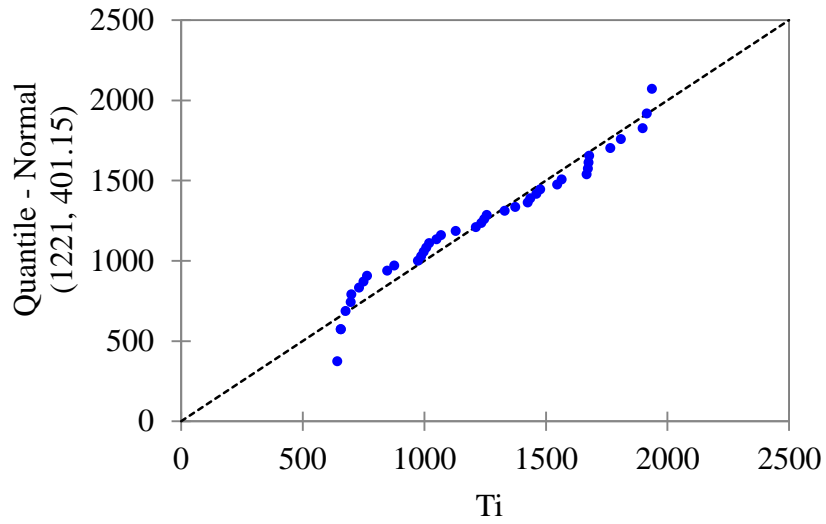


Figure B.16: Normal QQ plot of  $T_i$  for the dry season ( $n=40$ ).

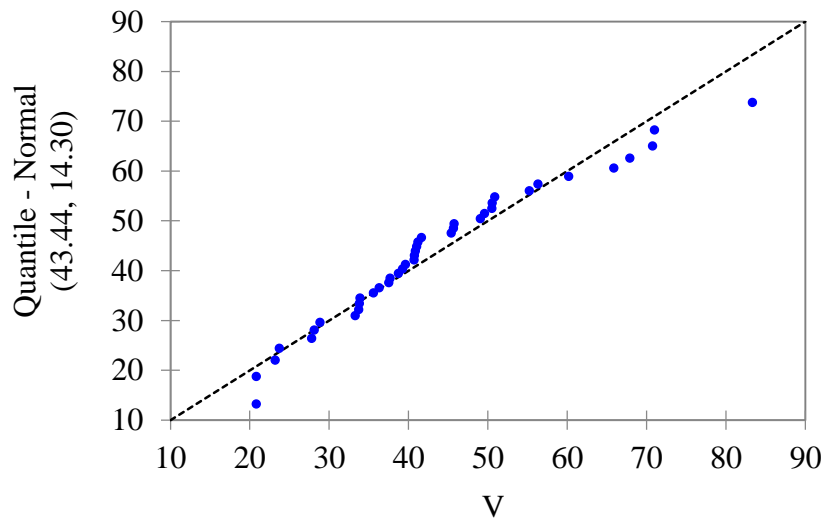


Figure B.17: Normal QQ plot of  $V$  for the dry season ( $n=40$ ).

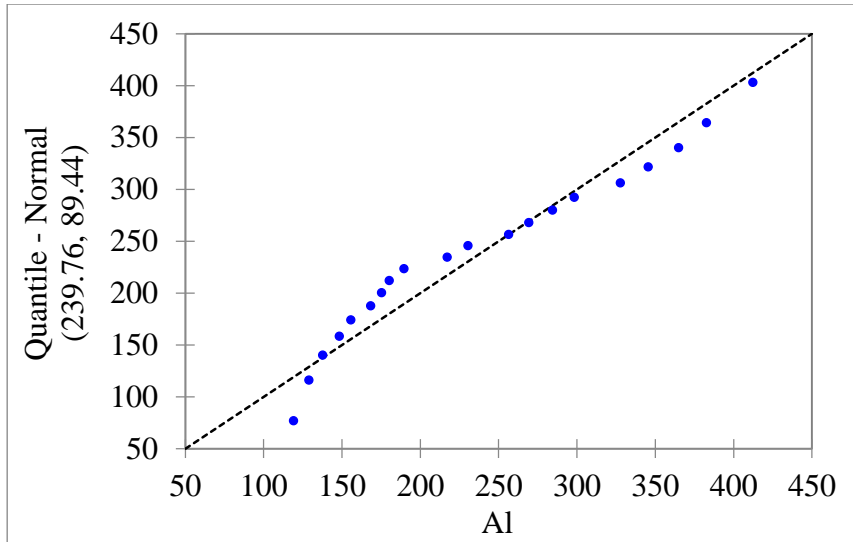


Figure B.18: Normal QQ plot of Al for the rainy season (n=20).

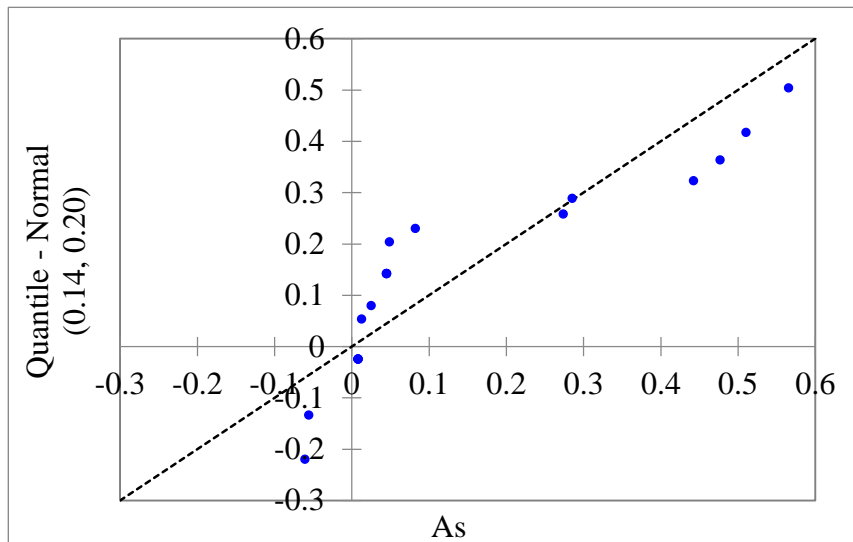


Figure B.19: Normal QQ plot of As for the rainy season (n=20).

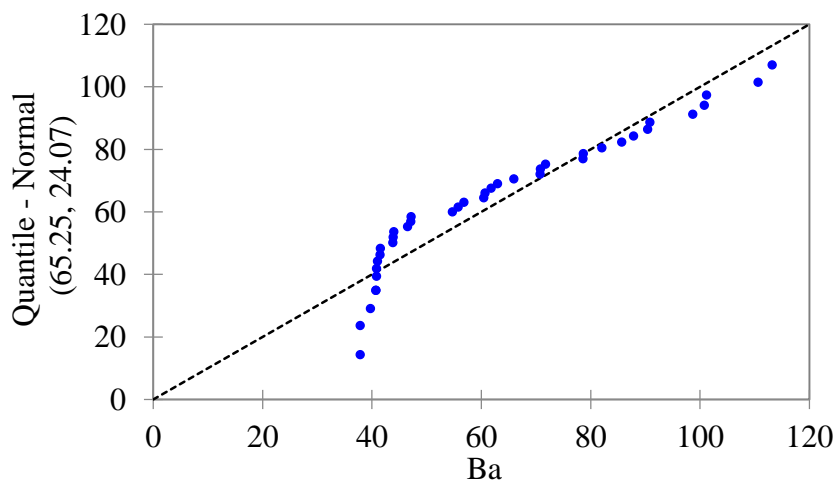


Figure B.20: Normal QQ plot of Ba for the rainy season (n=20).

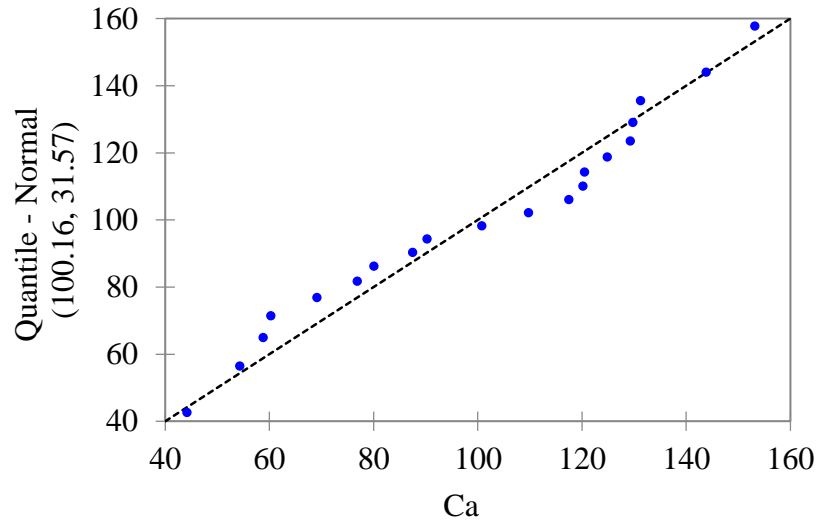


Figure B.21: Normal QQ plot of Ca for the rainy season (n=20).

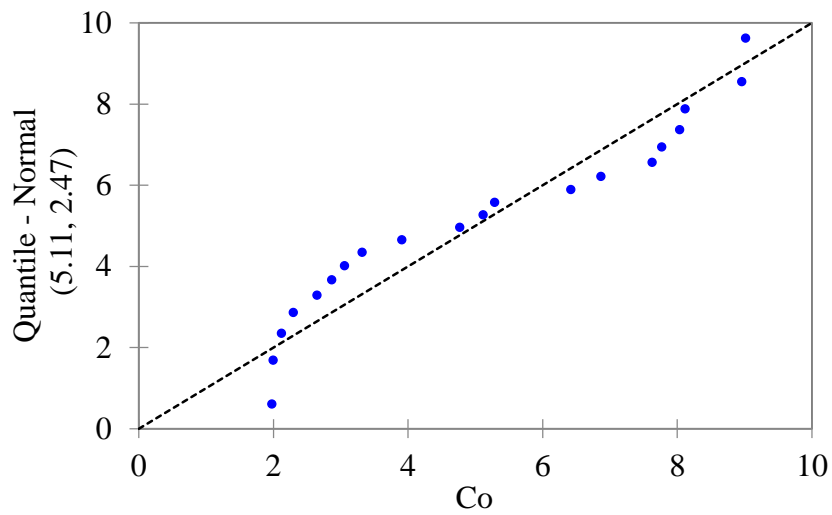


Figure B.22: Normal QQ plot of Co for the rainy season (n=20).

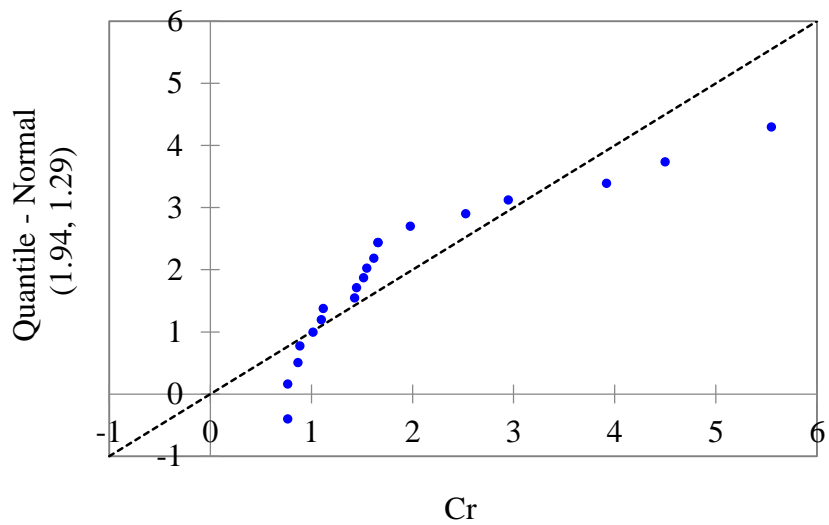


Figure B.23: Normal QQ plot of Cr for the rainy season (n=20).

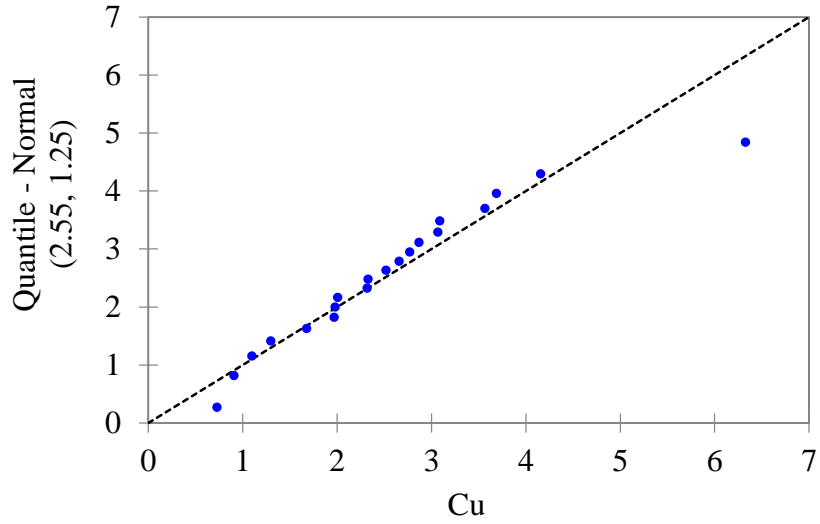


Figure B.24: Normal QQ plot of Cu for the rainy season (n=20).

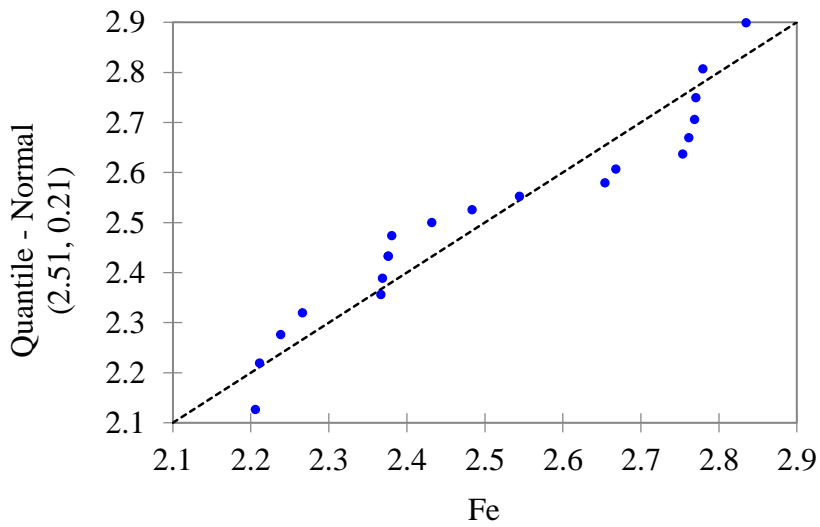


Figure B.25: Normal QQ plot of Fe for the rainy season (n=20).

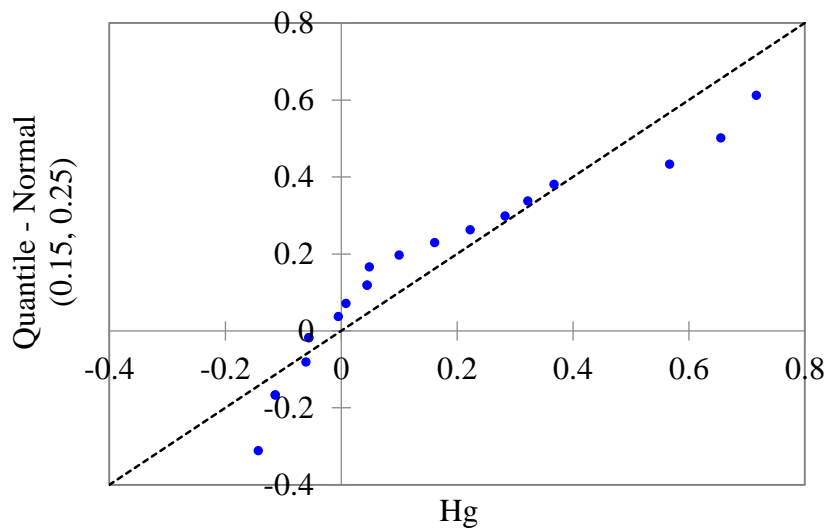


Figure B.26: Normal QQ plot of Hg for the rainy season (n=20).

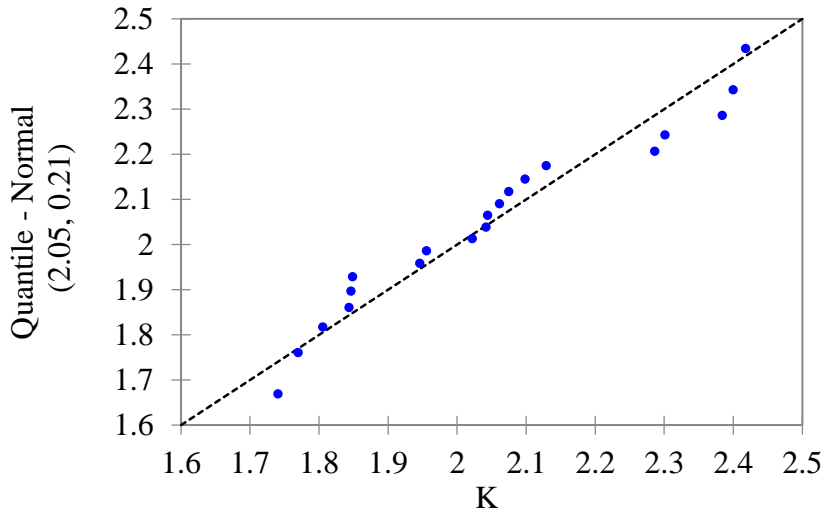


Figure B.27: Normal QQ plot of K for the rainy season (n=20).

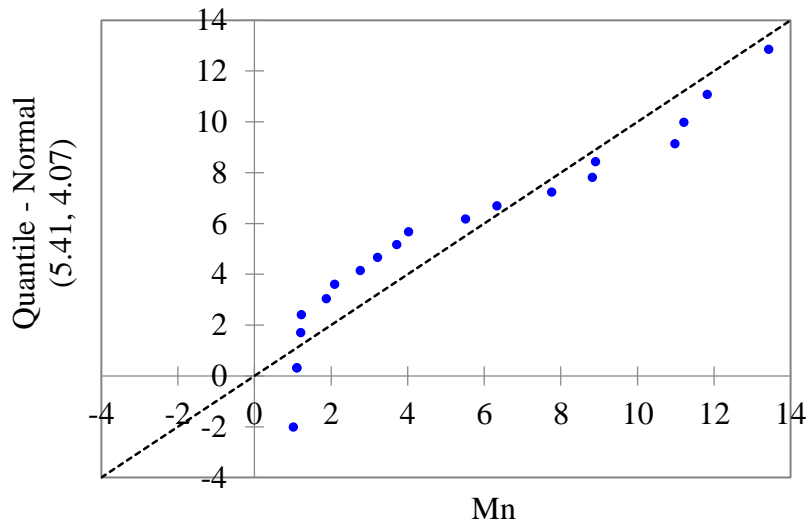


Figure B.28: Normal QQ plot of Mn for the rainy season (n=20).

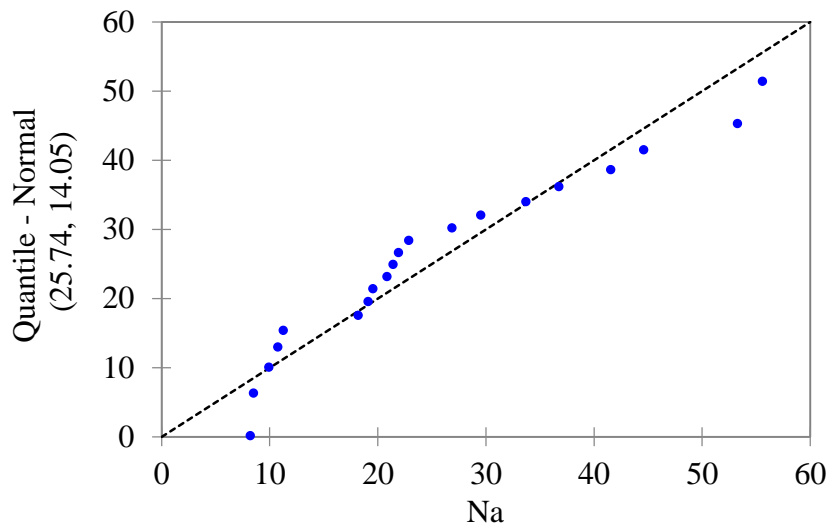


Figure B.29: Normal QQ plot of Na for the rainy season (n=20).



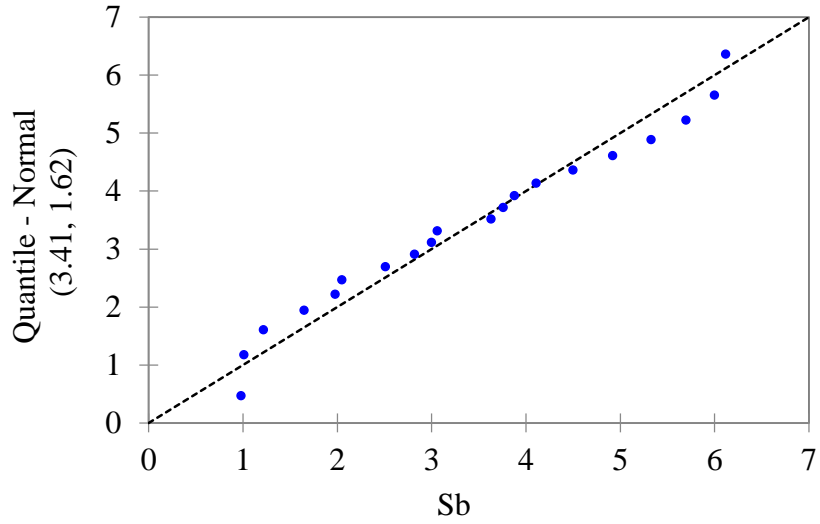


Figure B.30: Normal QQ plot of Sb for the rainy season (n=20).

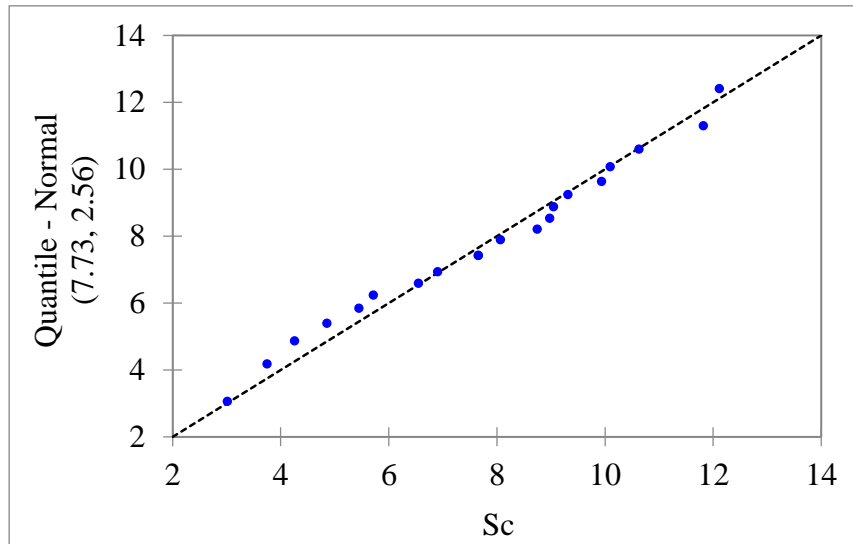


Figure B.31: Normal QQ plot of Sc for the rainy season (n=20).

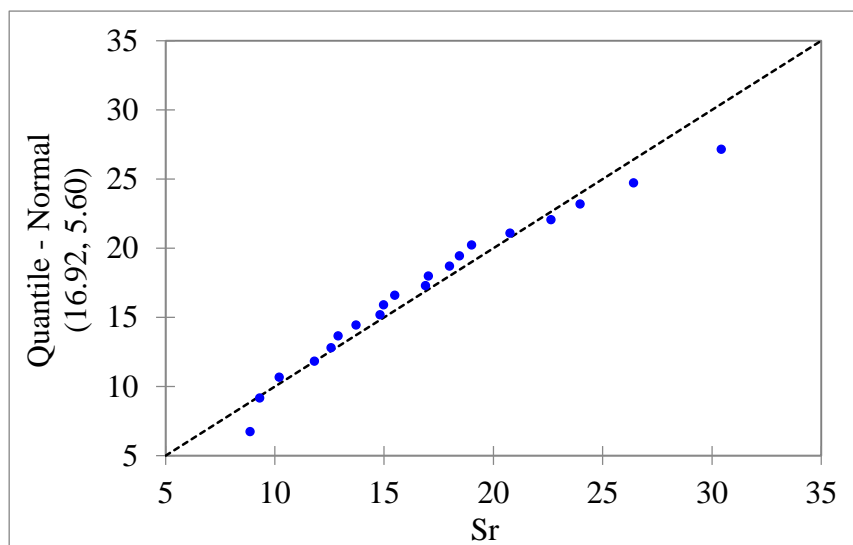


Figure B.32: Normal QQ plot of Sr for the rainy season (n=20).

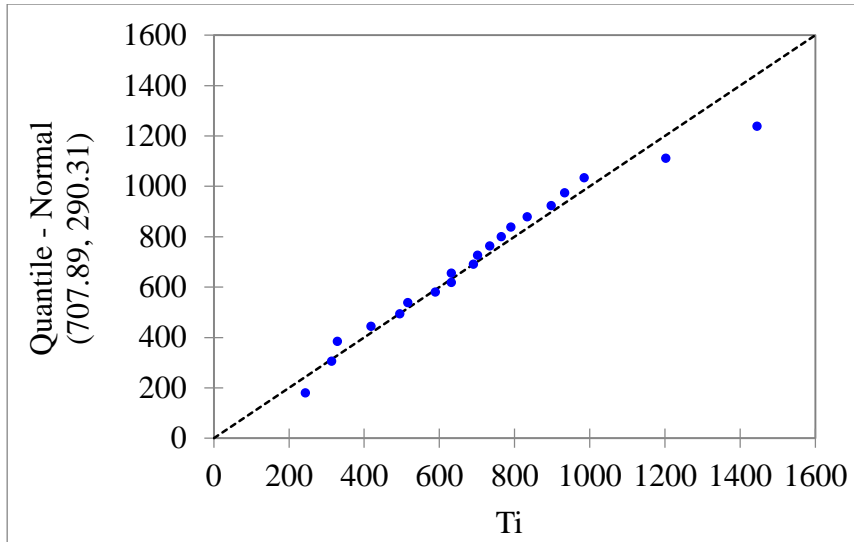


Figure B.33: Normal QQ plot of  $T_i$  for the rainy season ( $n=20$ ).

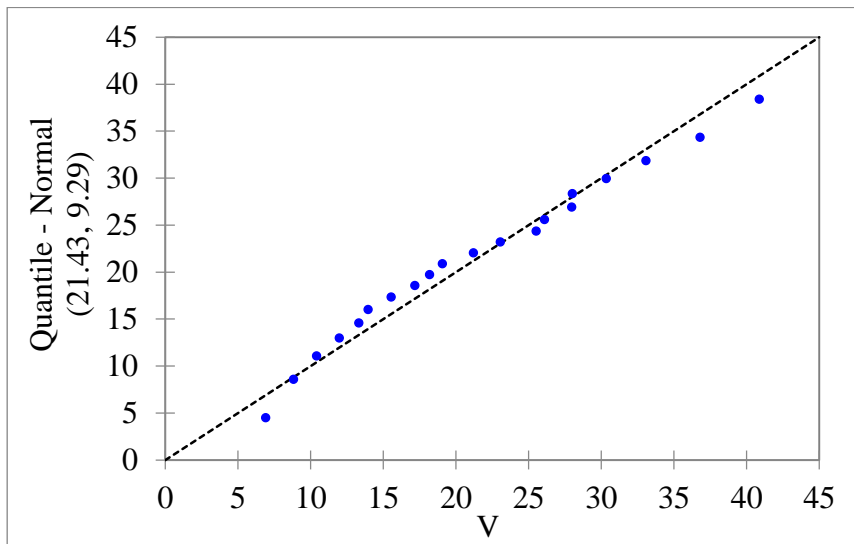
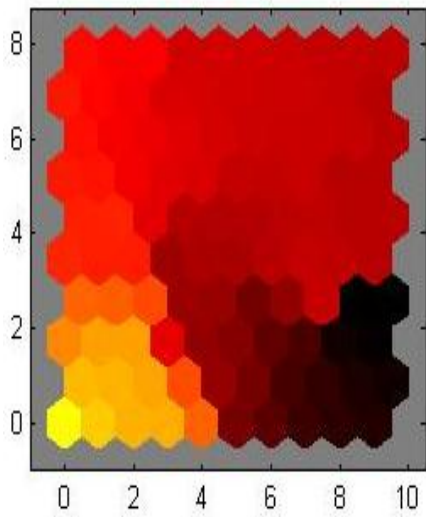
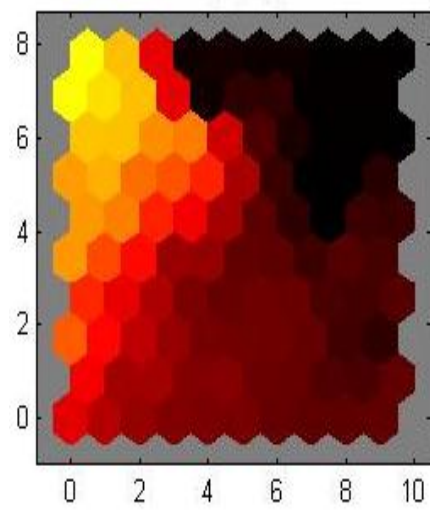


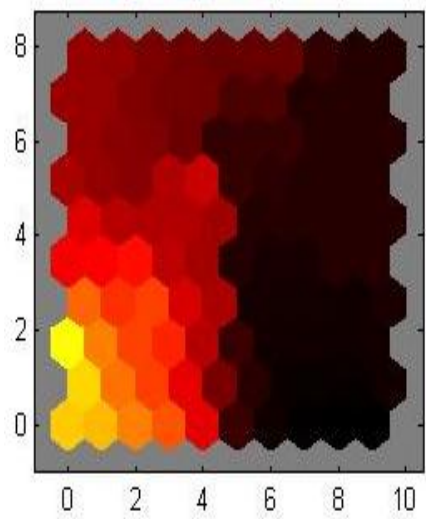
Figure B.34: Normal QQ plot of  $V$  for the rainy season ( $n=20$ ).



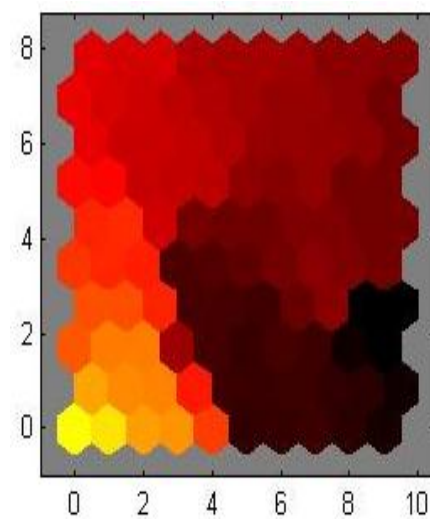
Al



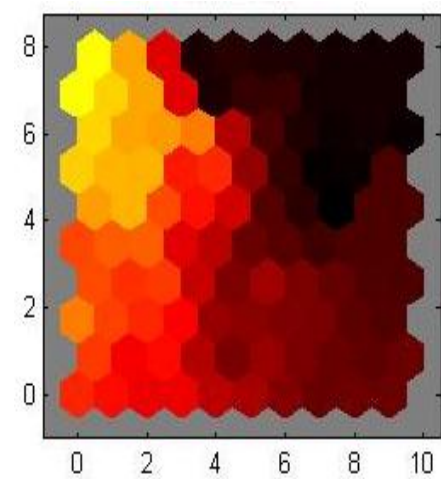
As



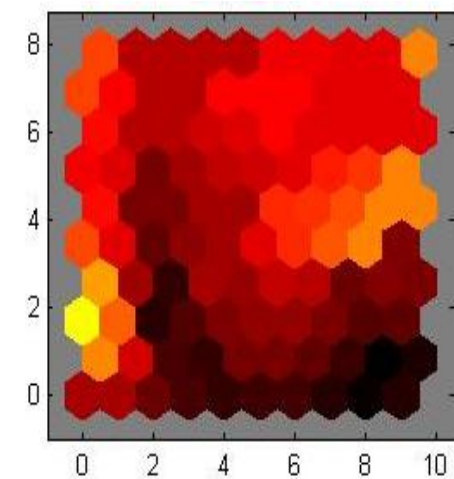
Ba



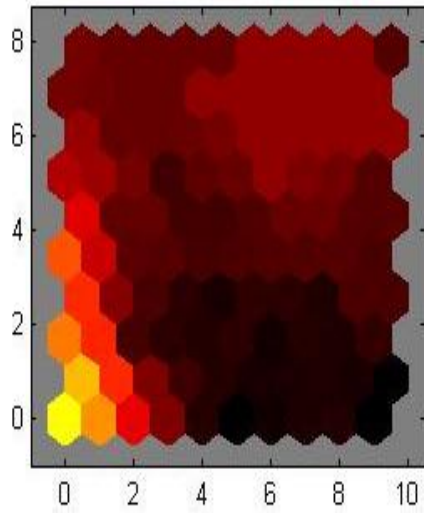
Ca



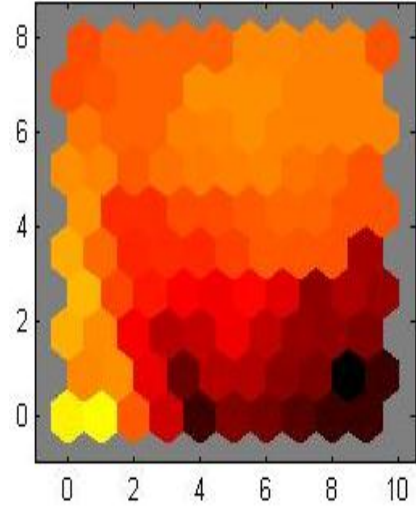
Co



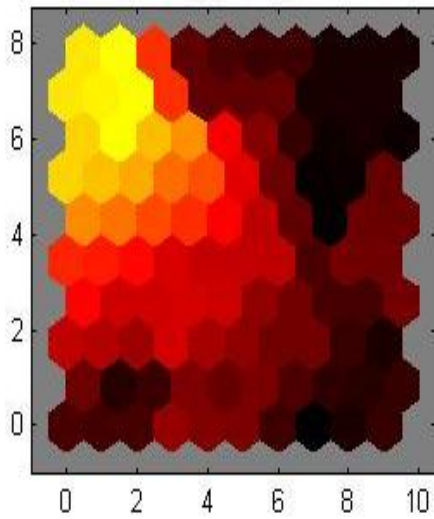
Cr



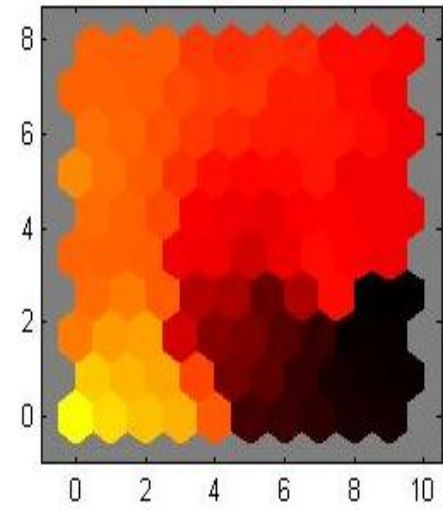
Cu



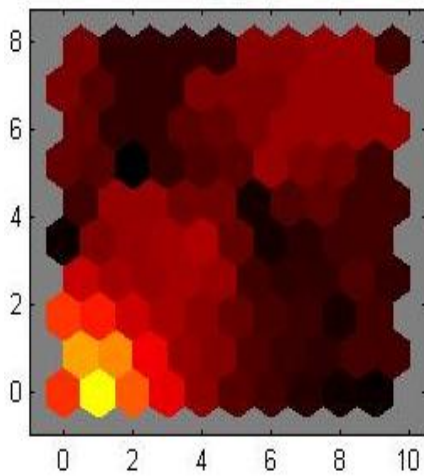
Fe



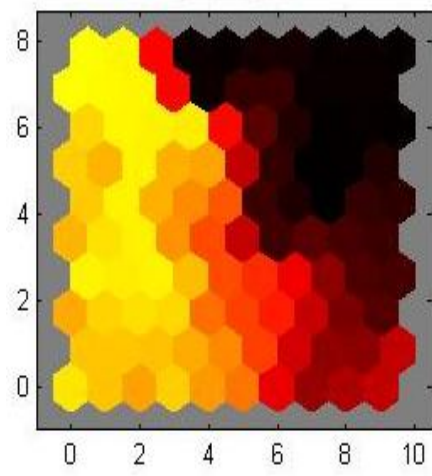
Hg



K



Mn



Na

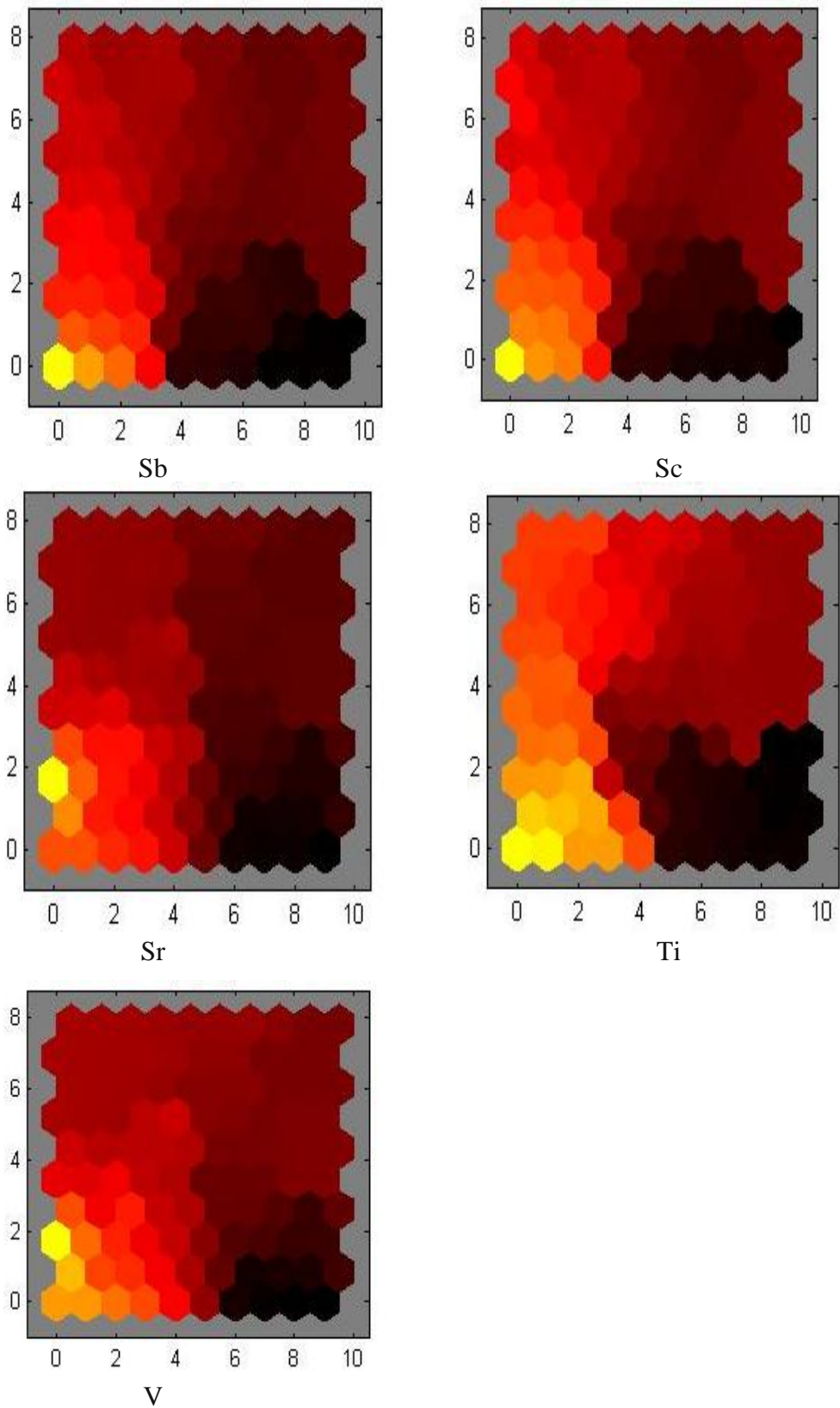
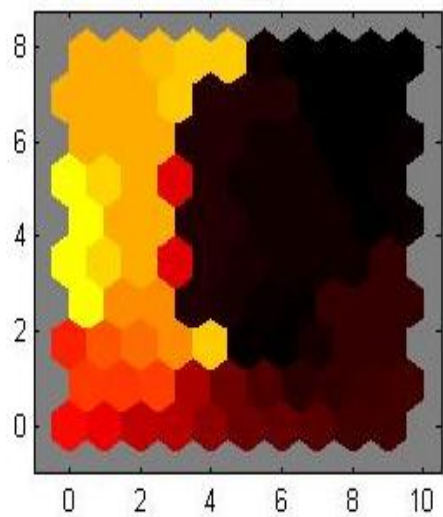
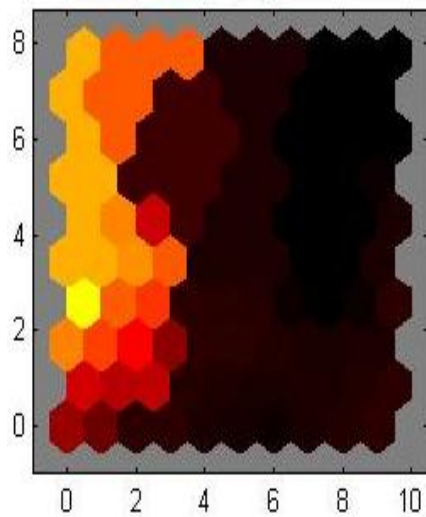


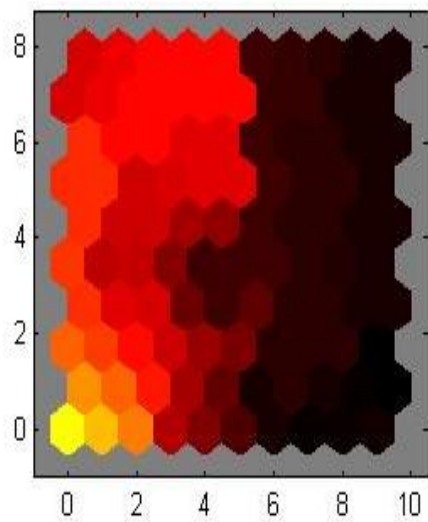
Figure B.35: SOM obtained using ANN for metal elements in dry season.



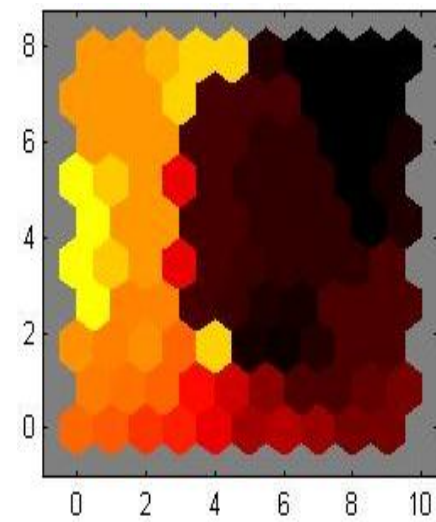
Al



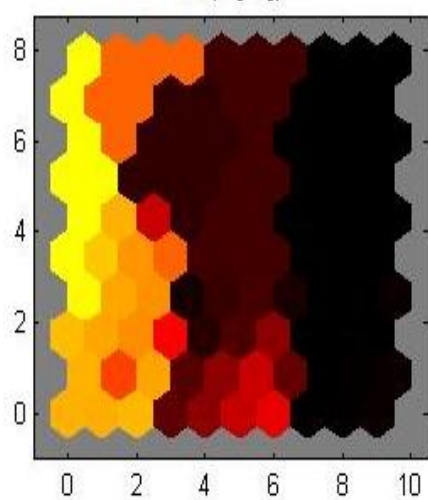
As



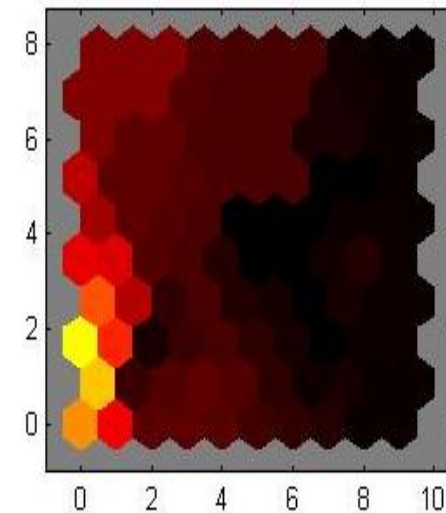
Ba



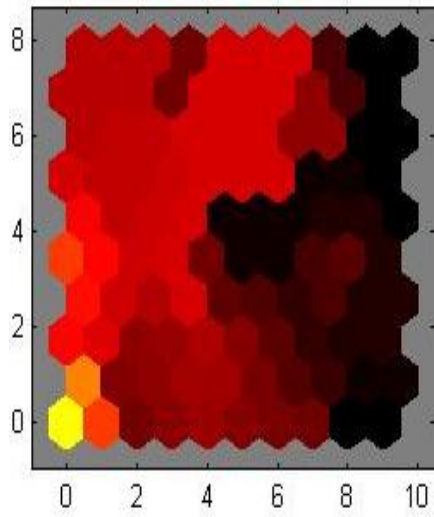
Ca



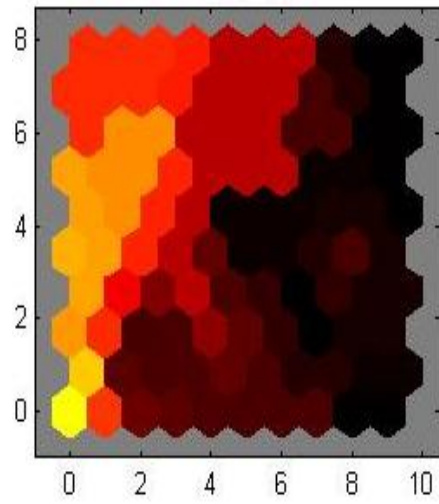
Co



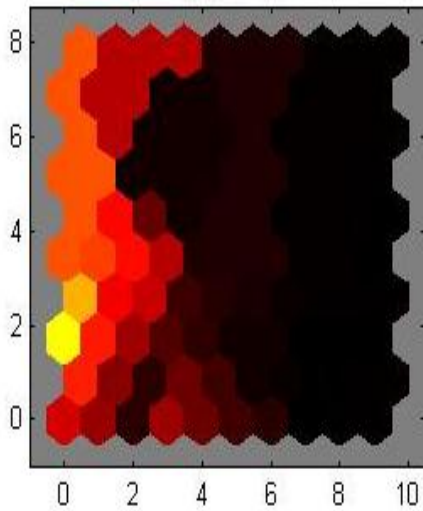
Cr



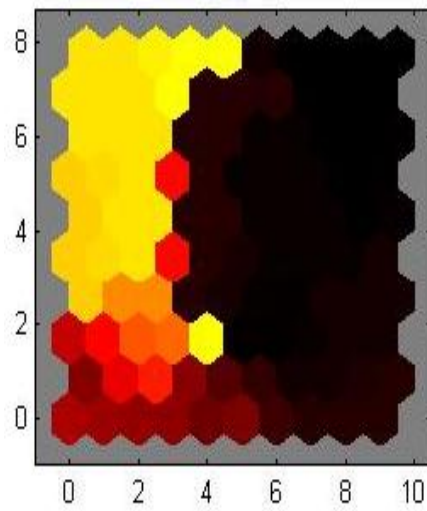
Cu



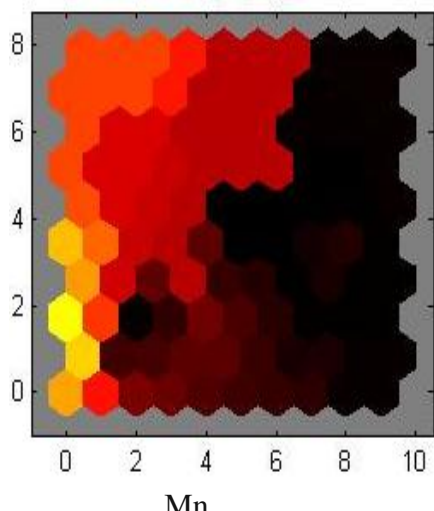
Fe



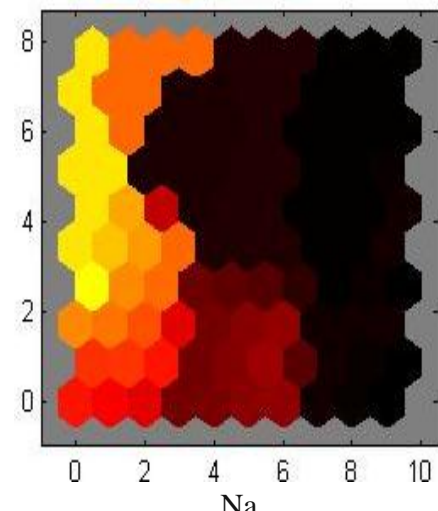
Hg



K



Mn



Na

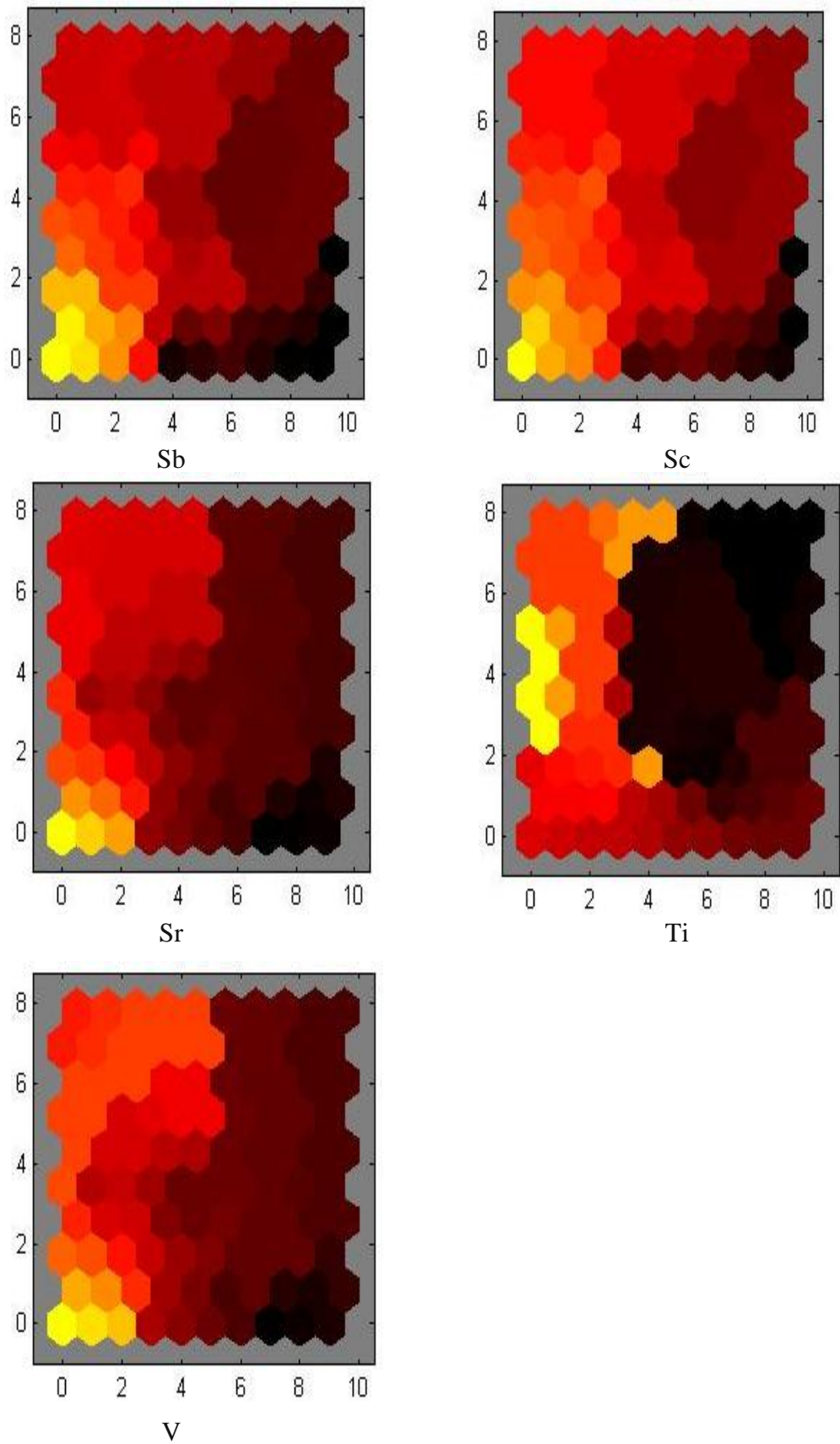


Figure B.36: SOM obtained using ANN for metal elements in rainy season.



## Annex – C

### Cross Validation Results and Spatial Distribution of Metal Elements using IDW

**Table C.1: Cross validation of IDW for Al**

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	96.217	166.117
IDW 2	82.885	149.770
IDW 3	68.693	140.647
IDW 4	57.664	135.536
IDW 5	49.862	132.793

**Table C.2: Cross validation of IDW for As**

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	1.195	1.949
IDW 2	1.014	1.739
IDW 3	0.820	1.600
IDW 4	0.674	1.536
IDW 5	0.574	1.510

**Table C.3: Cross validation of IDW for Ba**

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	10.354	17.869
IDW 2	8.689	15.878
IDW 3	7.016	14.738
IDW 4	5.714	14.012
IDW 5	4.767	13.586

**Table C.4: Cross validation of IDW for Ca**

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	28.391	52.770
IDW 2	24.831	48.494
IDW 3	20.899	46.068
IDW 4	17.679	44.492
IDW 5	15.290	43.565

**Table C.5: Cross validation of IDW for Co**

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	0.963	1.639
IDW 2	0.785	1.482
IDW 3	0.619	1.379
IDW 4	0.493	1.304
IDW 5	0.403	1.262

**Table C.6: Cross validation of IDW for Cr**

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	1.117	2.371
IDW 2	1.001	2.244
IDW 3	0.845	2.166
IDW 4	0.706	2.136
IDW 5	0.593	2.129

**Table C.7: Cross validation of IDW for Cu**

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	1.981	3.425
IDW 2	1.683	3.131
IDW 3	1.373	2.930
IDW 4	1.127	2.790
IDW 5	0.947	2.713

**Table C.8: Cross validation of IDW for Fe**

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	305.811	569.813
IDW 2	277.247	541.882
IDW 3	241.160	525.837
IDW 4	208.707	516.264
IDW 5	181.154	509.430

**Table C.9: Cross validation of IDW for Hg**

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	1.395	2.316
IDW 2	1.252	2.127
IDW 3	1.070	1.996
IDW 4	0.919	1.923
IDW 5	0.806	1.889

**Table C.10: Cross validation of IDW for K**

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	57.983	103.040
IDW 2	49.877	93.217
IDW 3	40.762	87.876
IDW 4	33.486	85.119
IDW 5	28.142	83.754

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error.

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	3.805	6.959
IDW 2	3.392	6.596
IDW 3	2.893	6.419
IDW 4	2.474	6.294
IDW 5	2.123	6.191

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	16.341	29.218
IDW 2	14.165	27.152
IDW 3	11.751	26.134
IDW 4	9.826	25.763
IDW 5	8.373	25.573

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	1.028	1.773
IDW 2	0.867	1.587
IDW 3	0.699	1.485
IDW 4	0.565	1.420
IDW 5	0.466	1.378

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	1.330	2.680
IDW 2	1.157	2.469
IDW 3	0.963	2.357
IDW 4	0.798	2.292
IDW 5	0.675	2.261

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	3.638	6.952
IDW 2	3.115	6.304
IDW 3	2.550	5.968
IDW 4	2.106	5.785
IDW 5	1.790	5.694

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	176.868	326.413
IDW 2	149.913	295.712
IDW 3	122.780	280.265
IDW 4	101.694	272.254
IDW 5	0.675	2.261

Method	<sup>a</sup> MPE	<sup>b</sup> RMSPE
IDW 1	7.457	13.604
IDW 2	6.532	12.440
IDW 3	5.462	11.806
IDW 4	4.574	11.426
IDW 5	1.790	5.694

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error

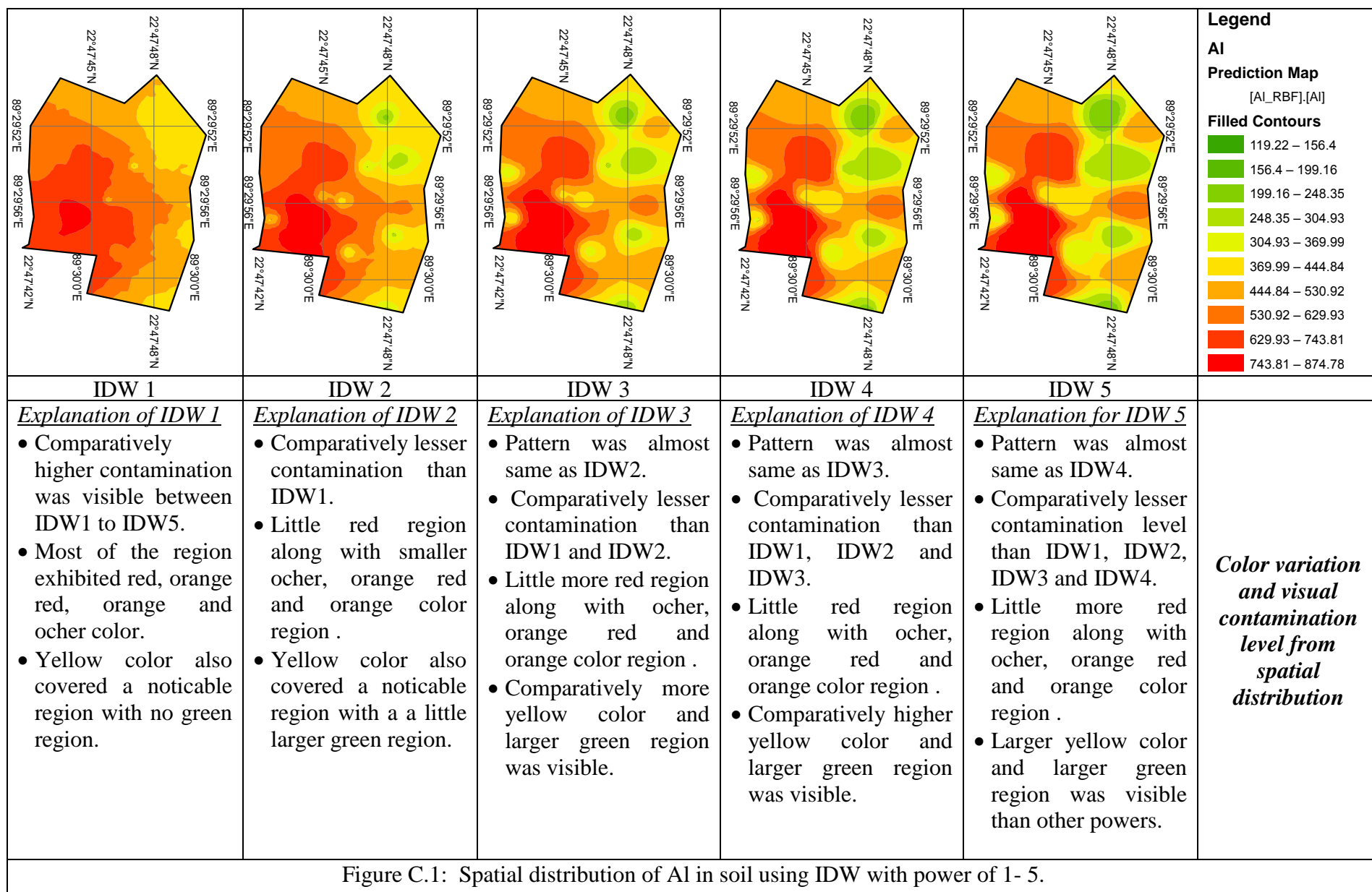


Figure C.1: Spatial distribution of AI in soil using IDW with power of 1- 5.

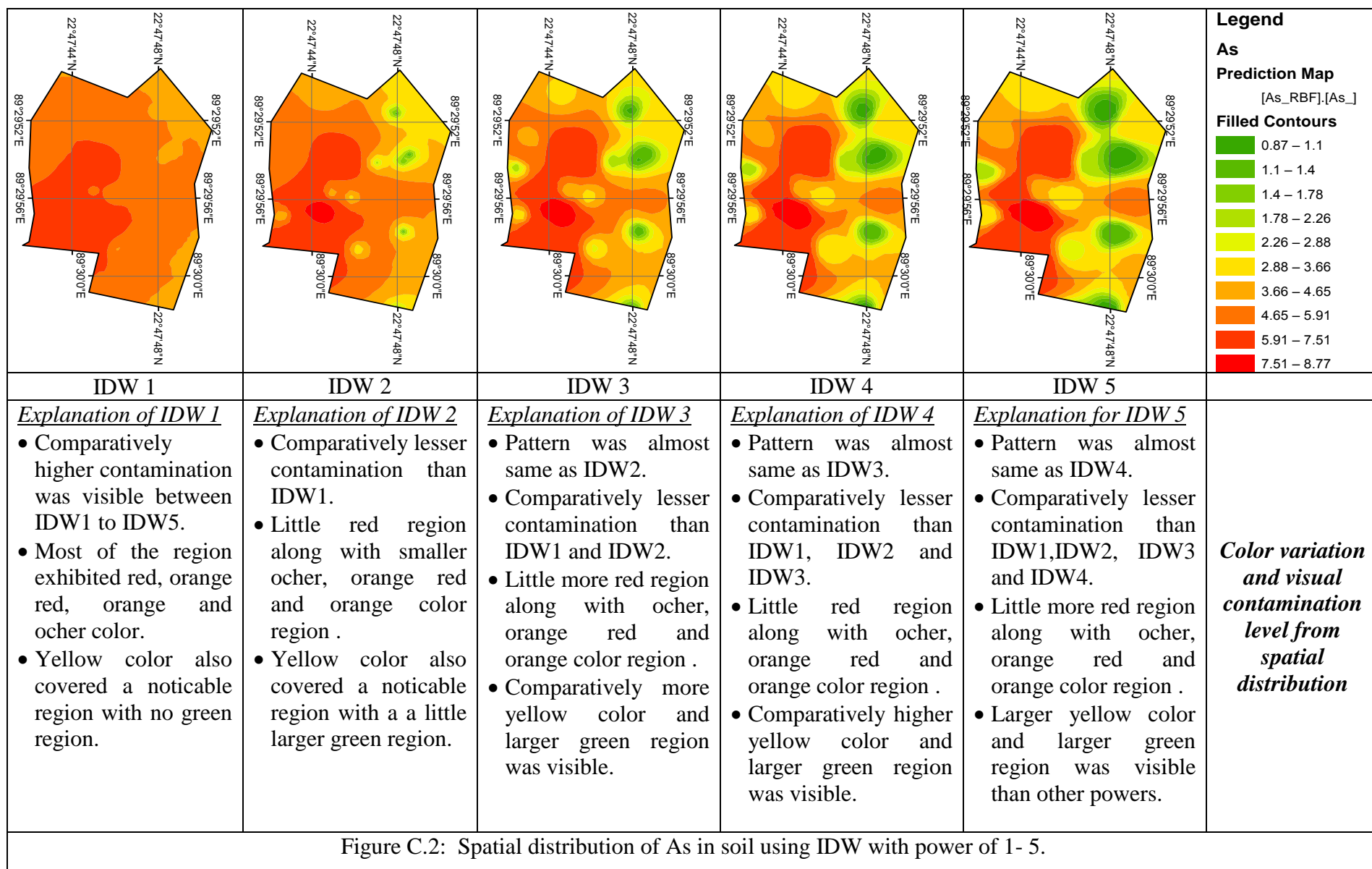


Figure C.2: Spatial distribution of As in soil using IDW with power of 1- 5.

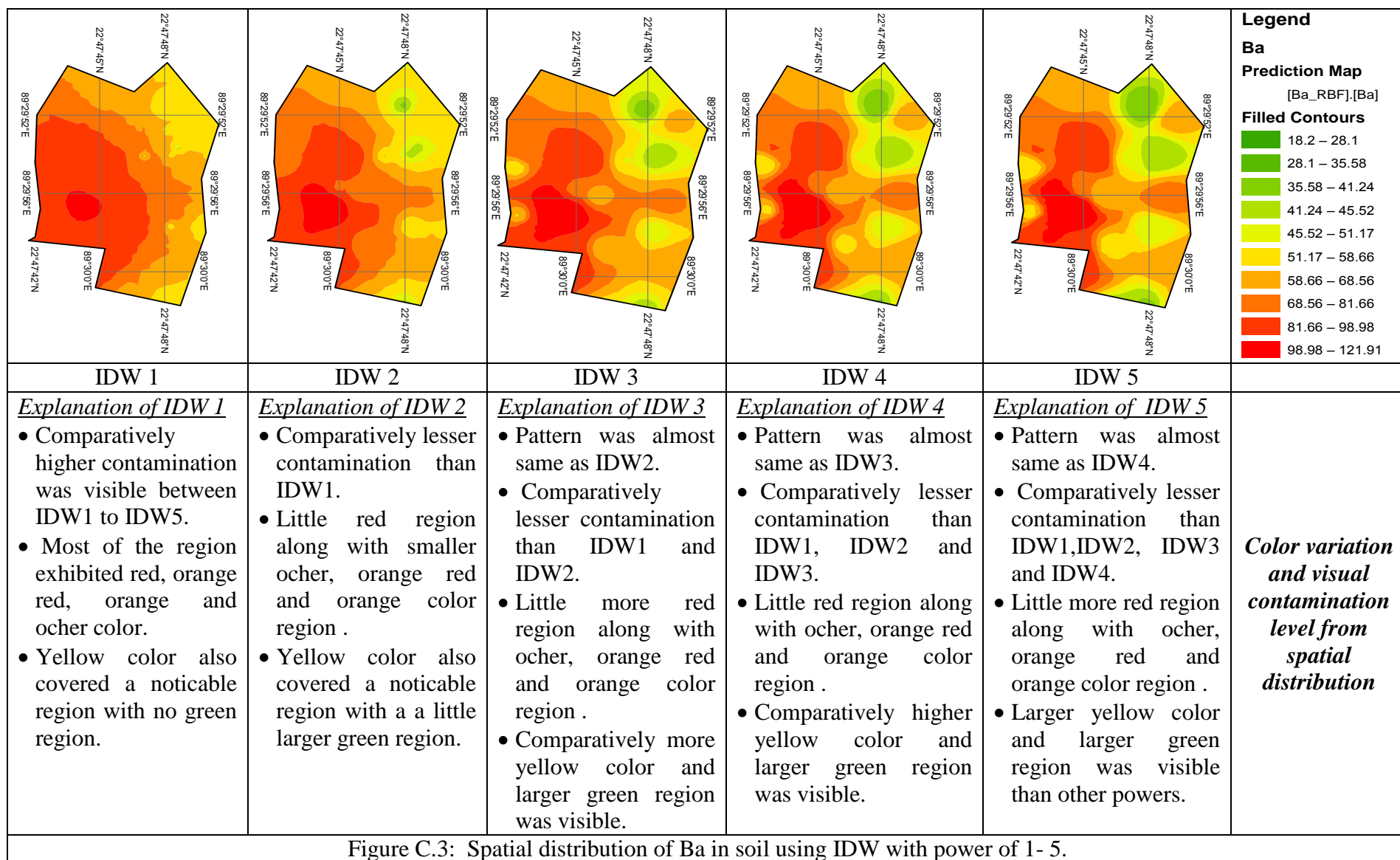


Figure C.3: Spatial distribution of Ba in soil using IDW with power of 1- 5.

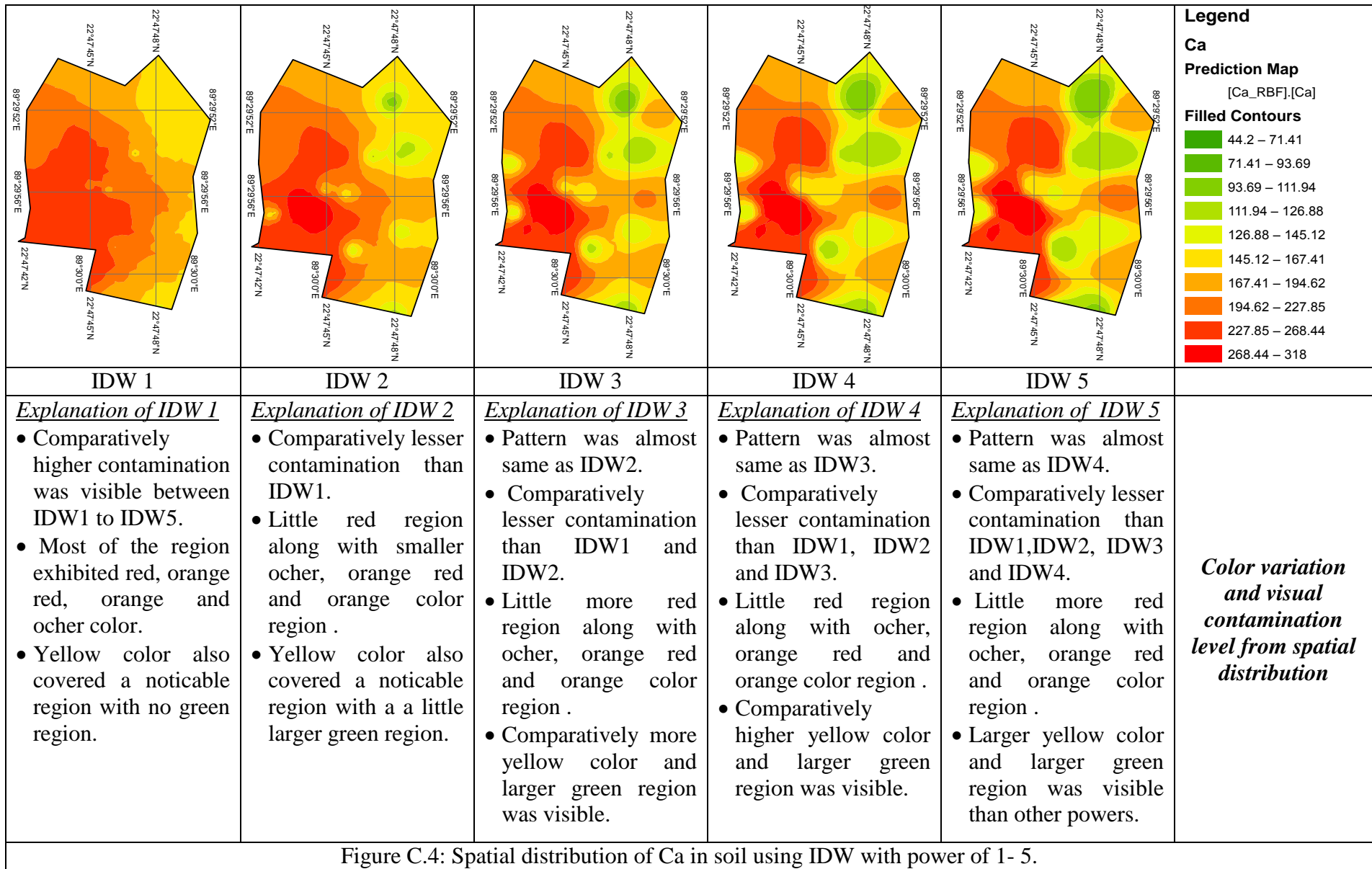


Figure C.4: Spatial distribution of Ca in soil using IDW with power of 1- 5.

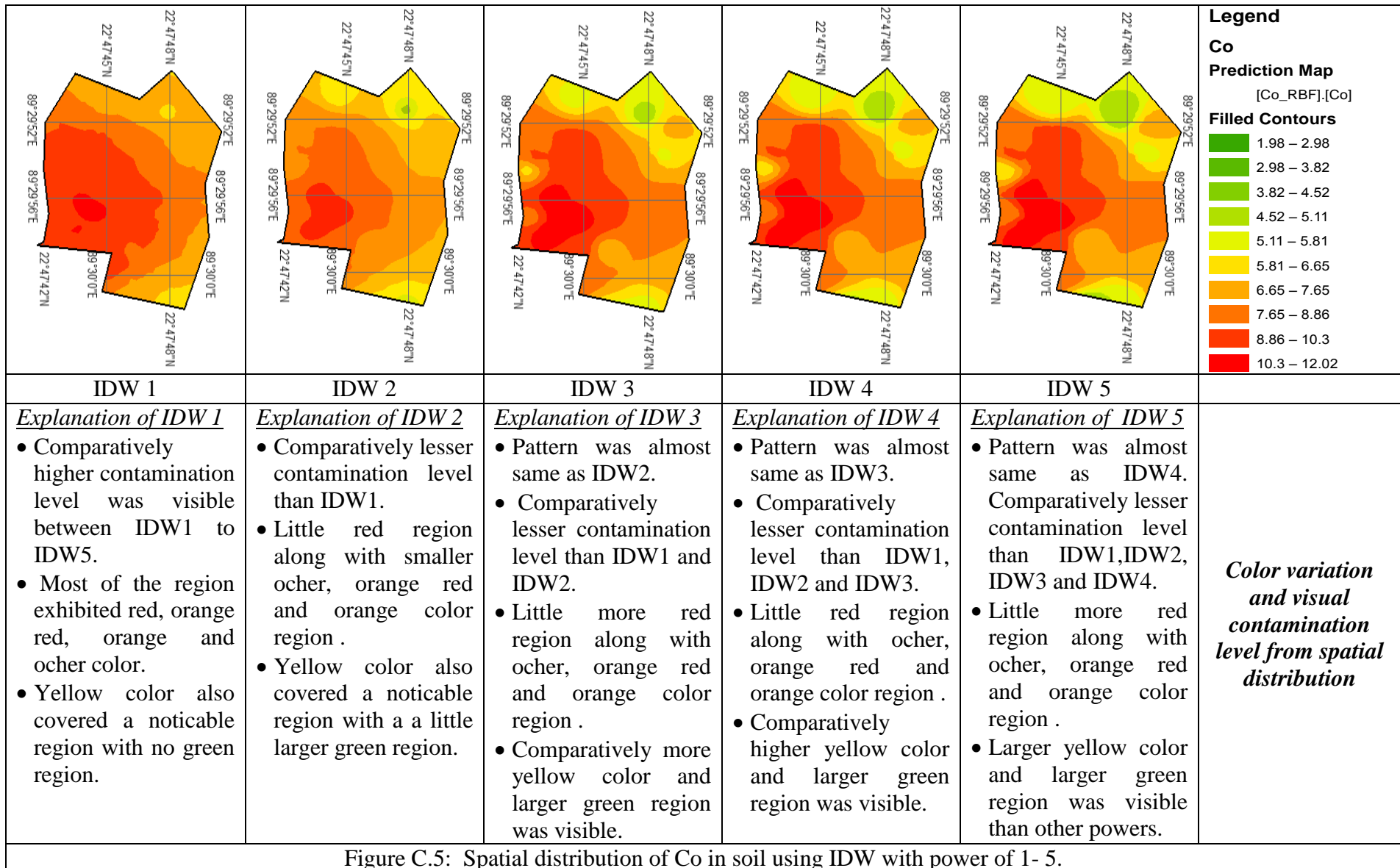
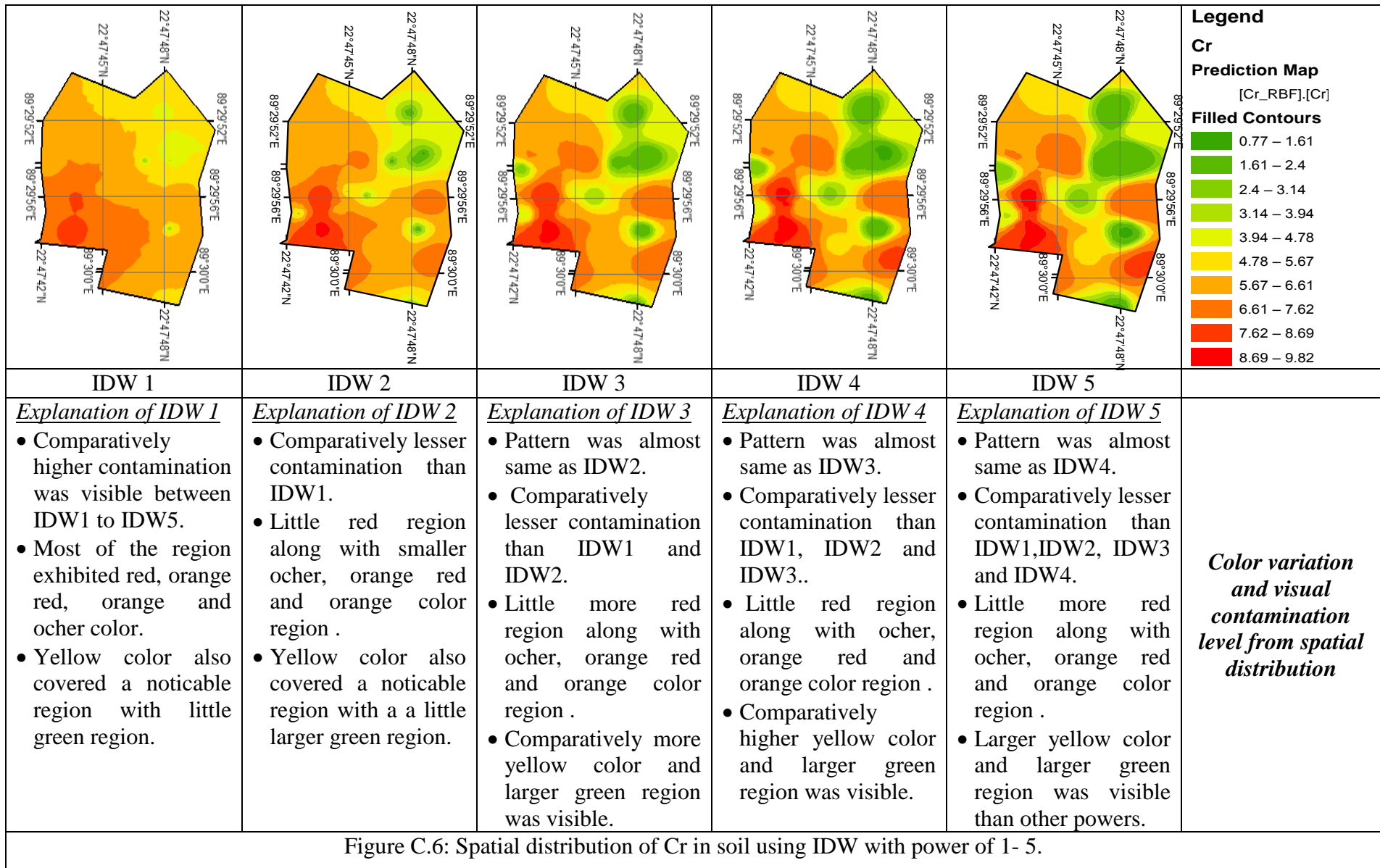


Figure C.5: Spatial distribution of Co in soil using IDW with power of 1- 5.





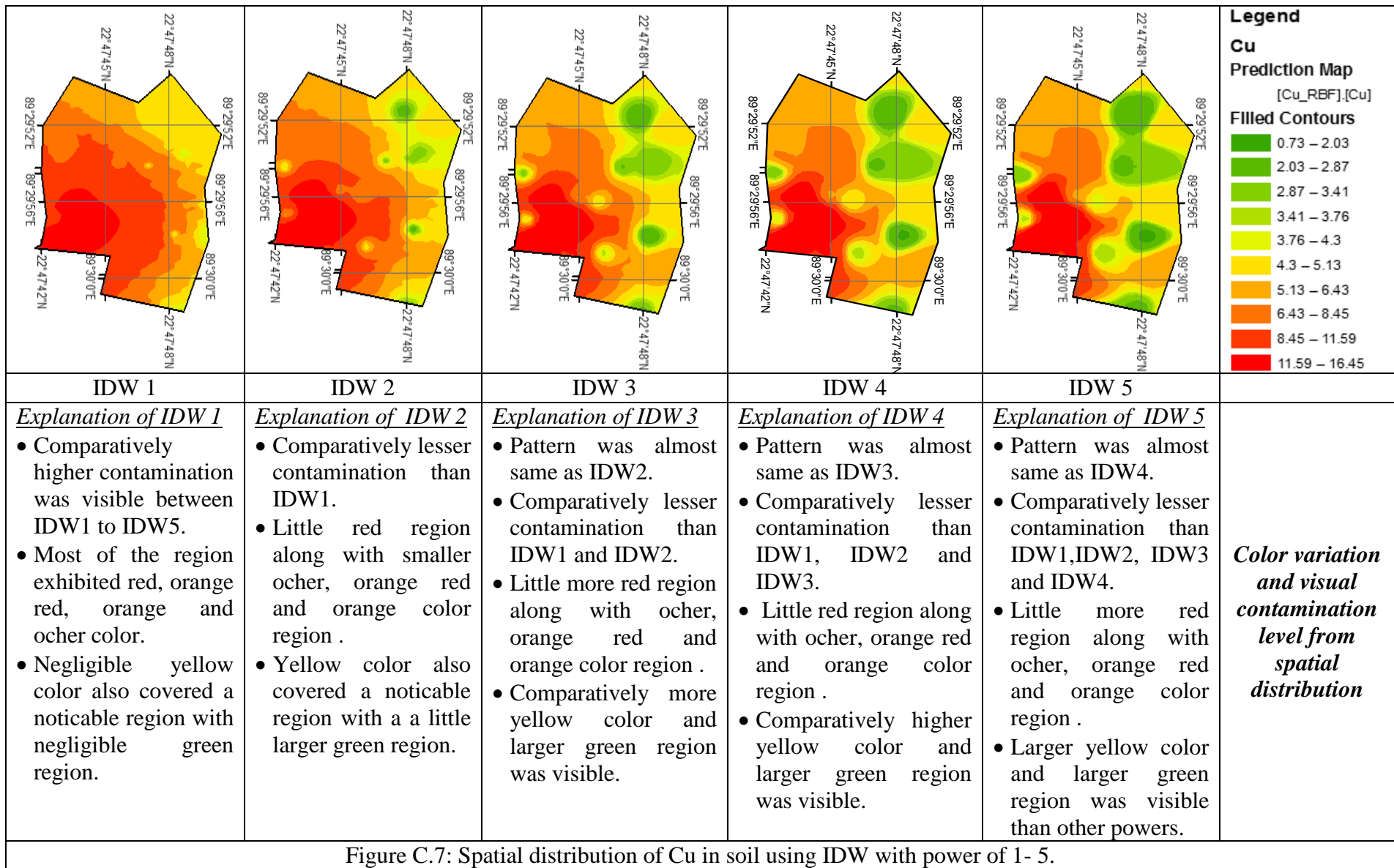


Figure C.7: Spatial distribution of Cu in soil using IDW with power of 1- 5.

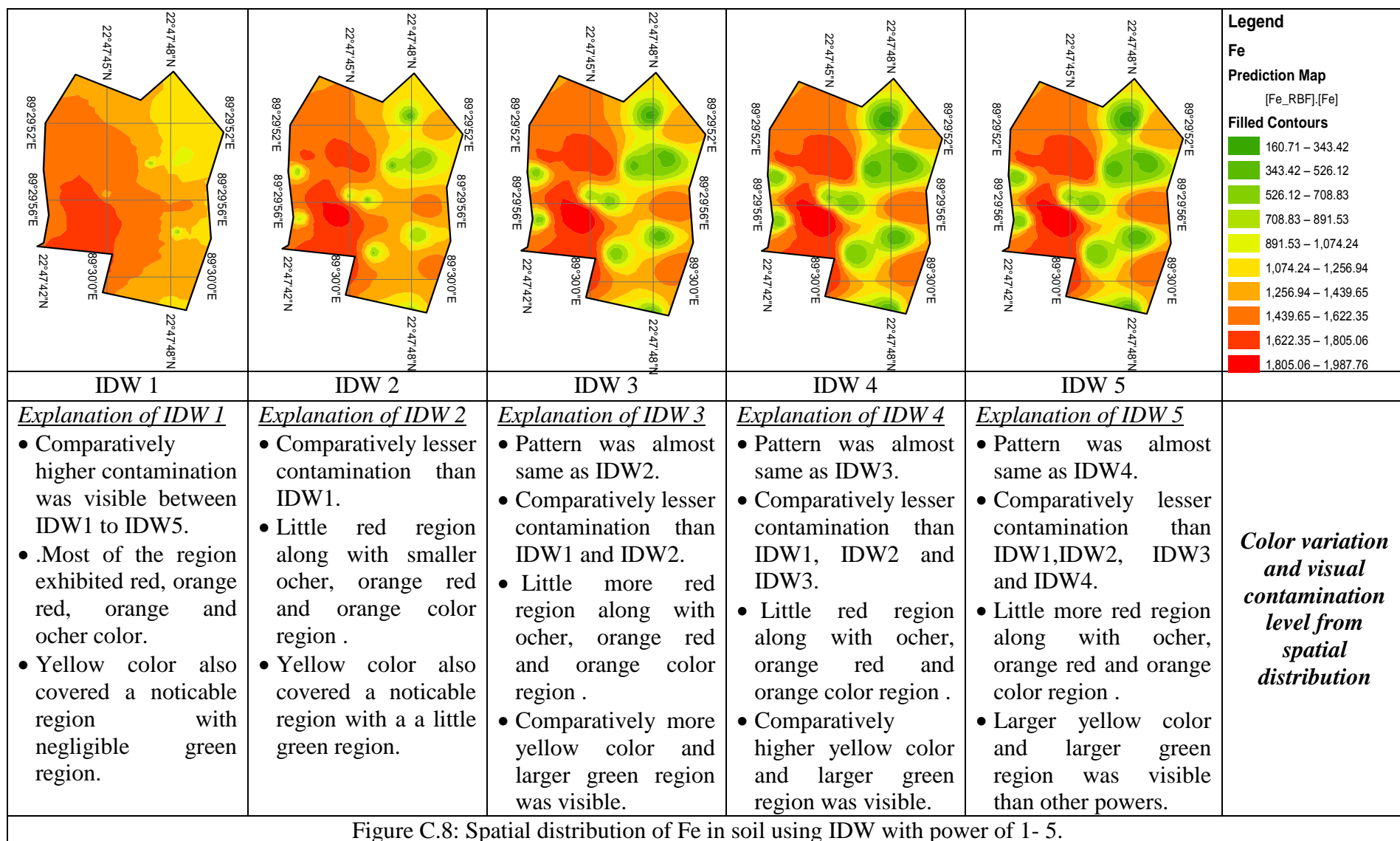


Figure C.8: Spatial distribution of Fe in soil using IDW with power of 1- 5.

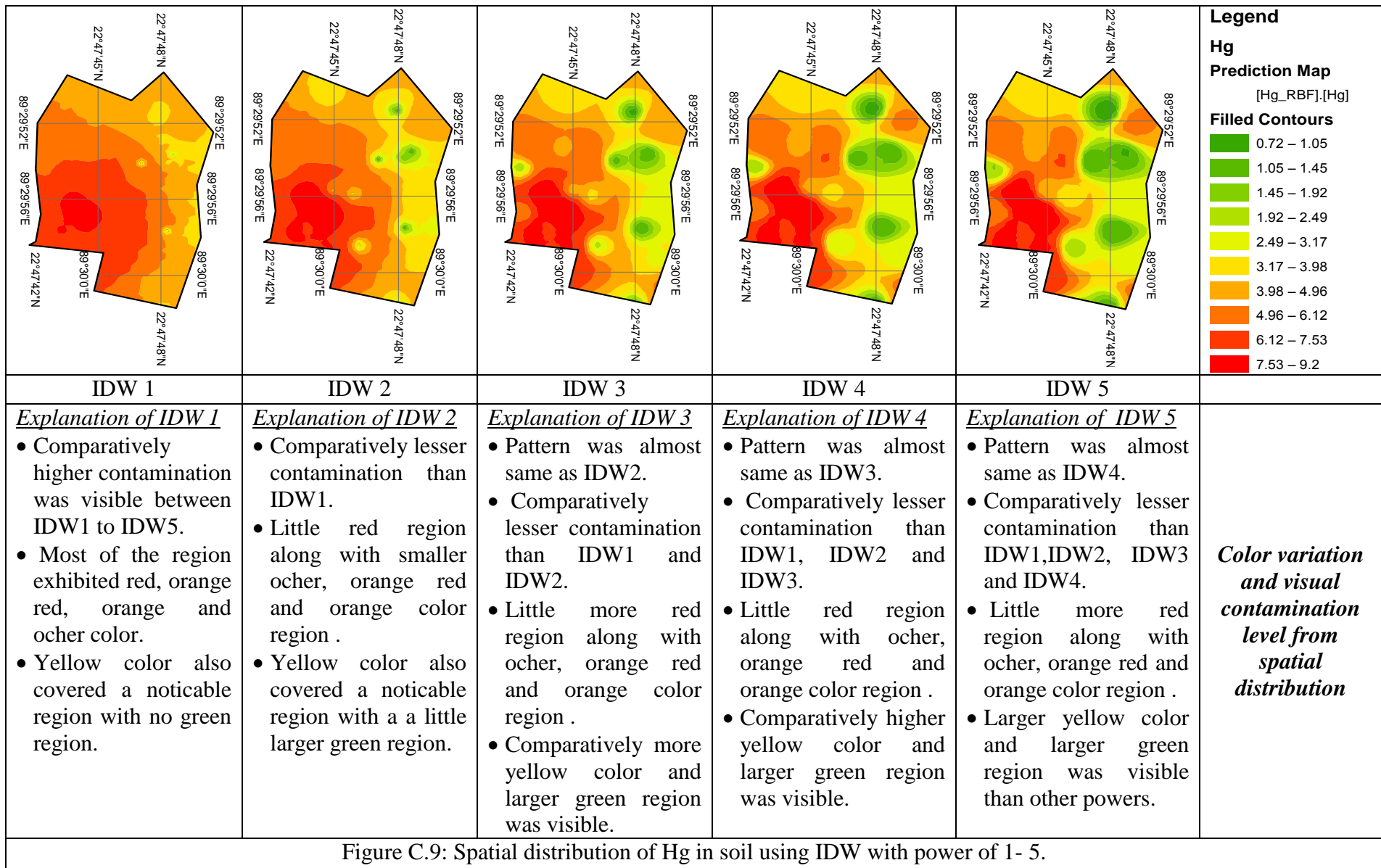


Figure C.9: Spatial distribution of Hg in soil using IDW with power of 1- 5.

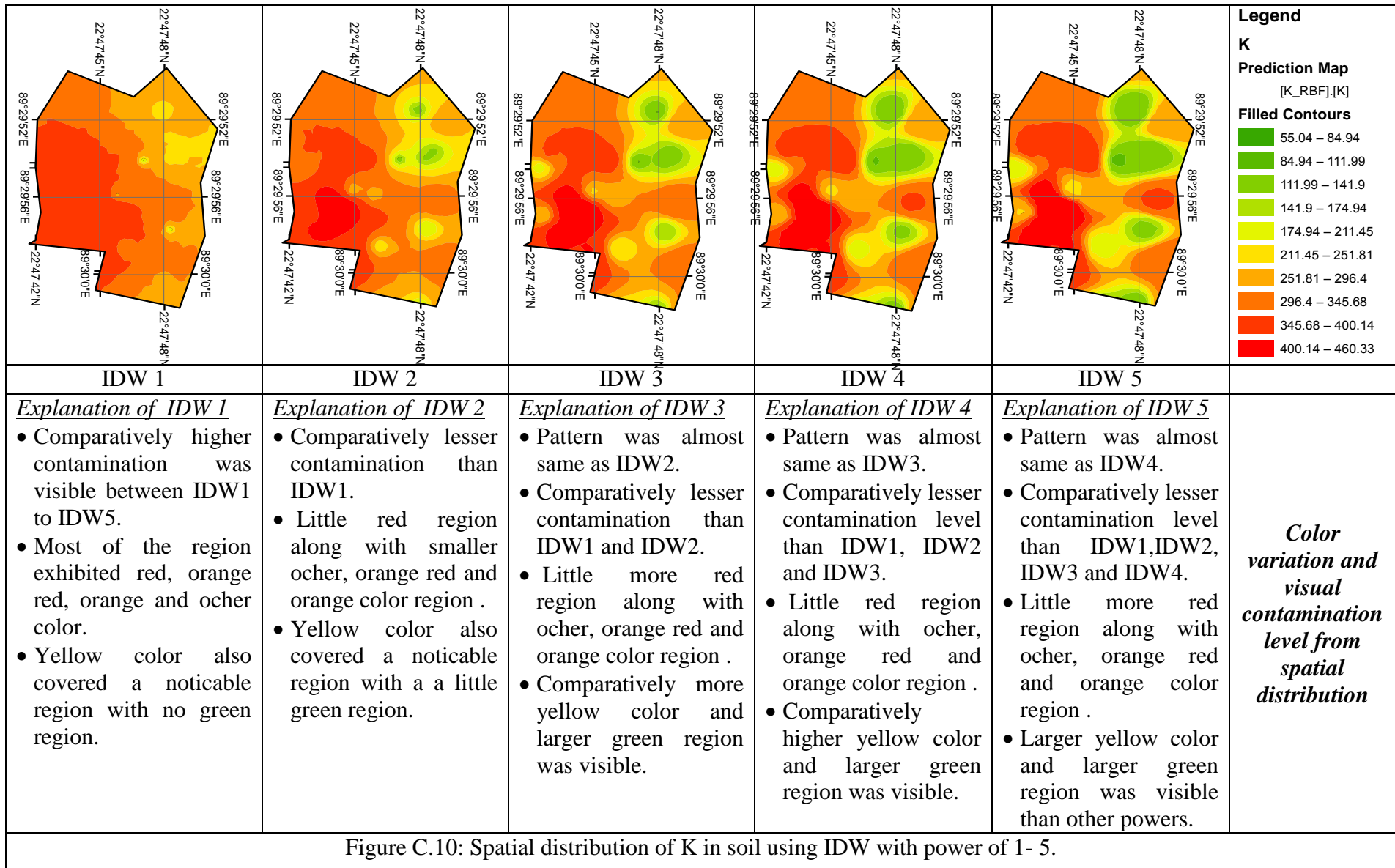


Figure C.10: Spatial distribution of K in soil using IDW with power of 1- 5.

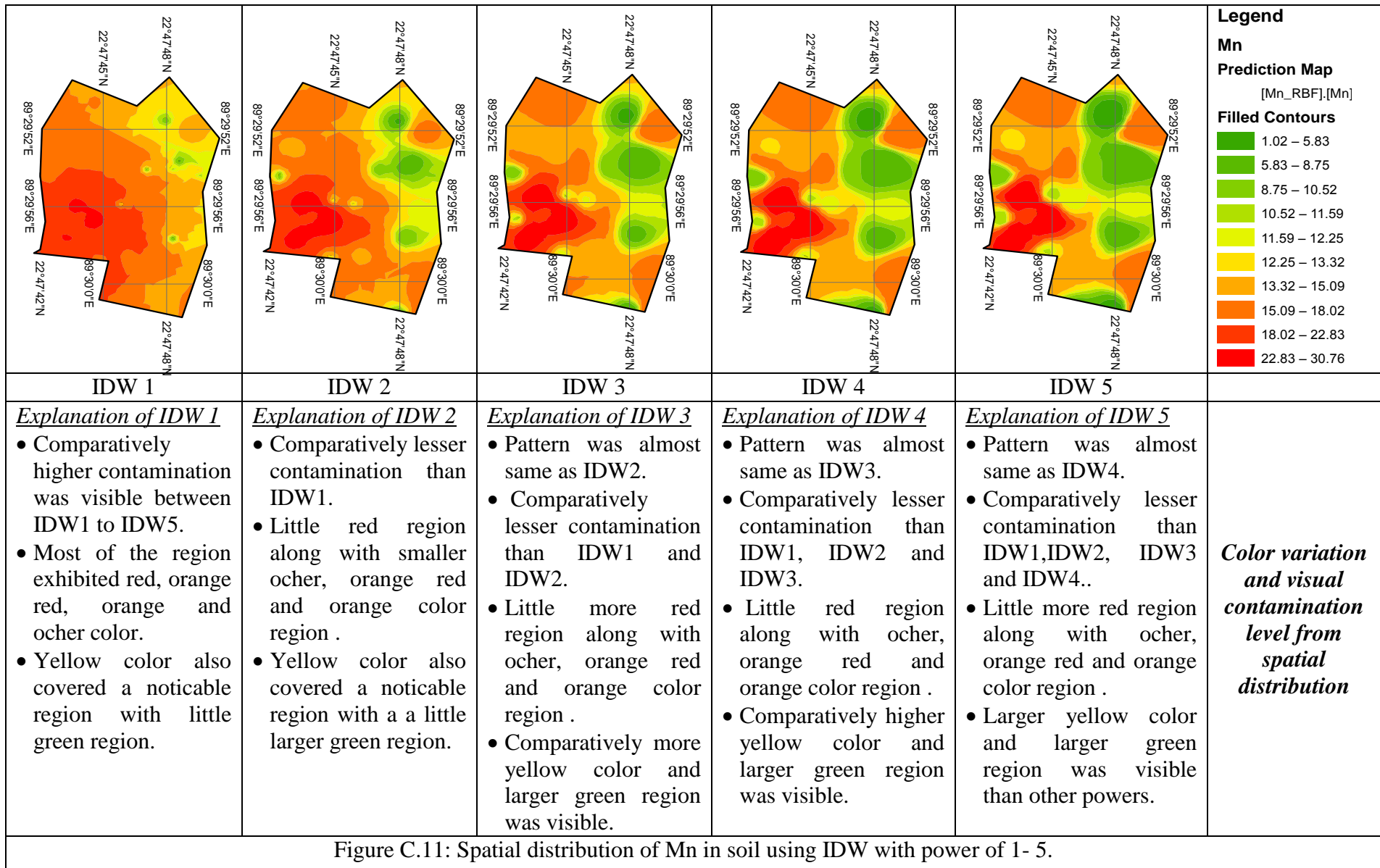


Figure C.11: Spatial distribution of Mn in soil using IDW with power of 1- 5.

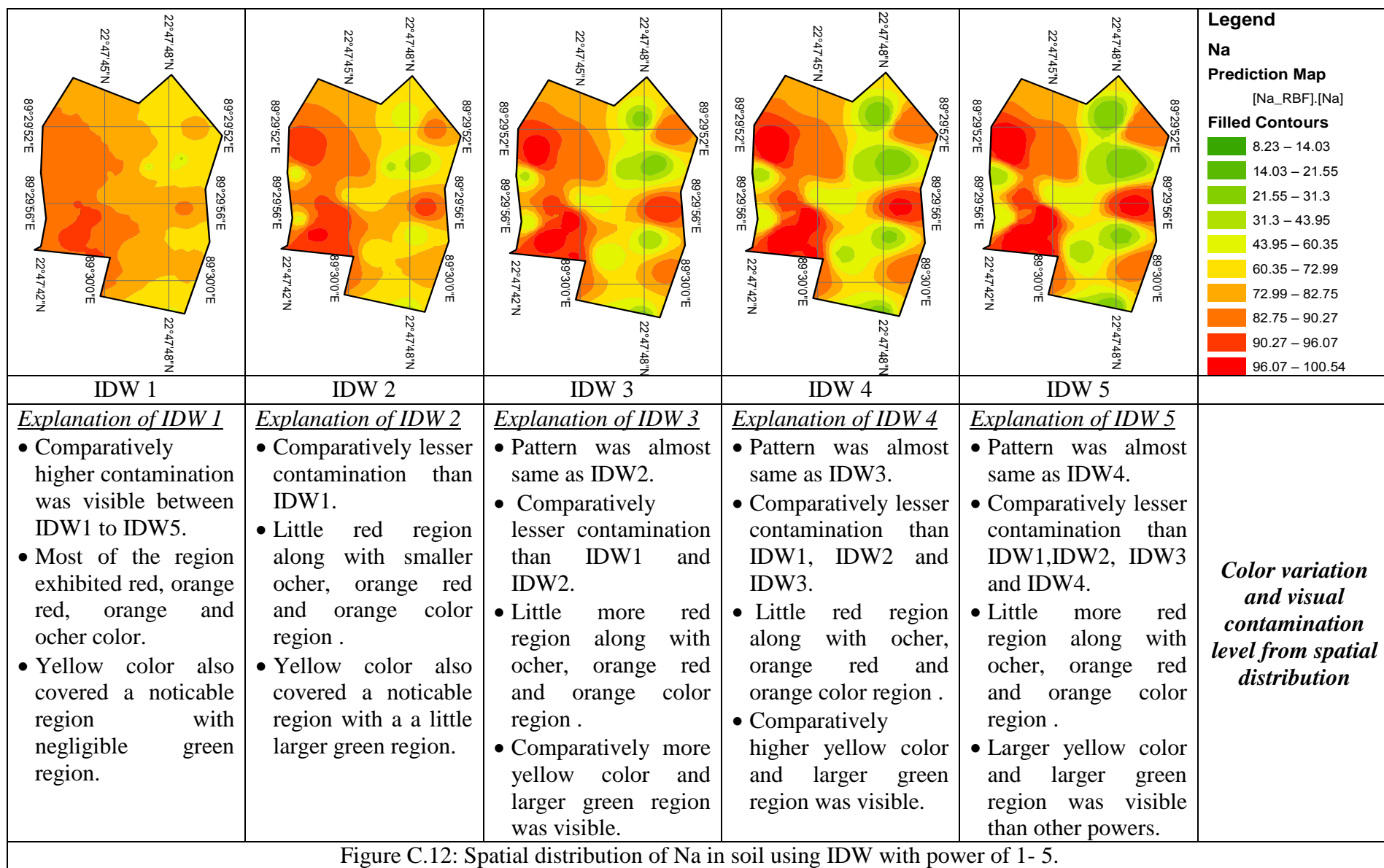


Figure C.12: Spatial distribution of Na in soil using IDW with power of 1- 5.

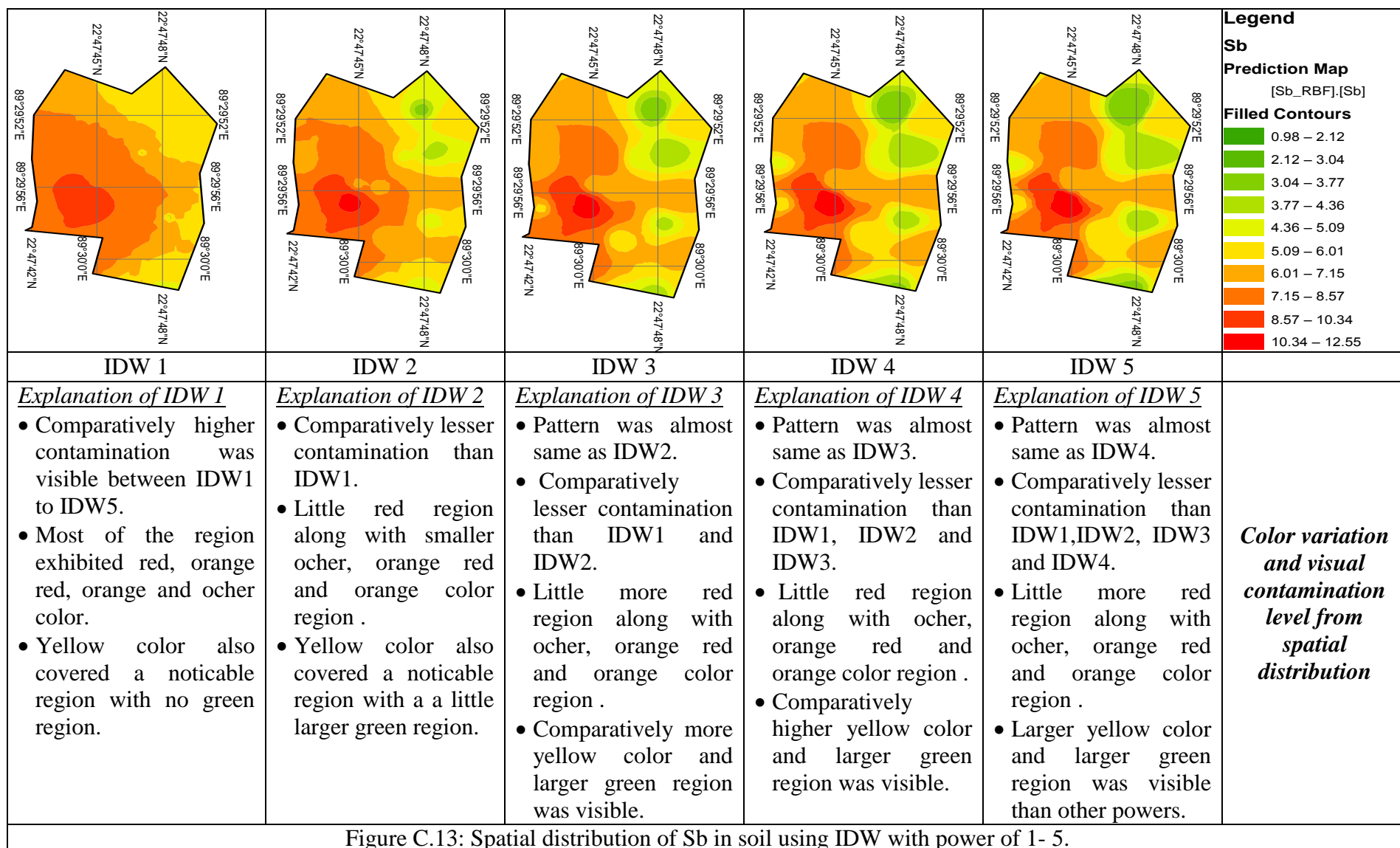


Figure C.13: Spatial distribution of Sb in soil using IDW with power of 1- 5.

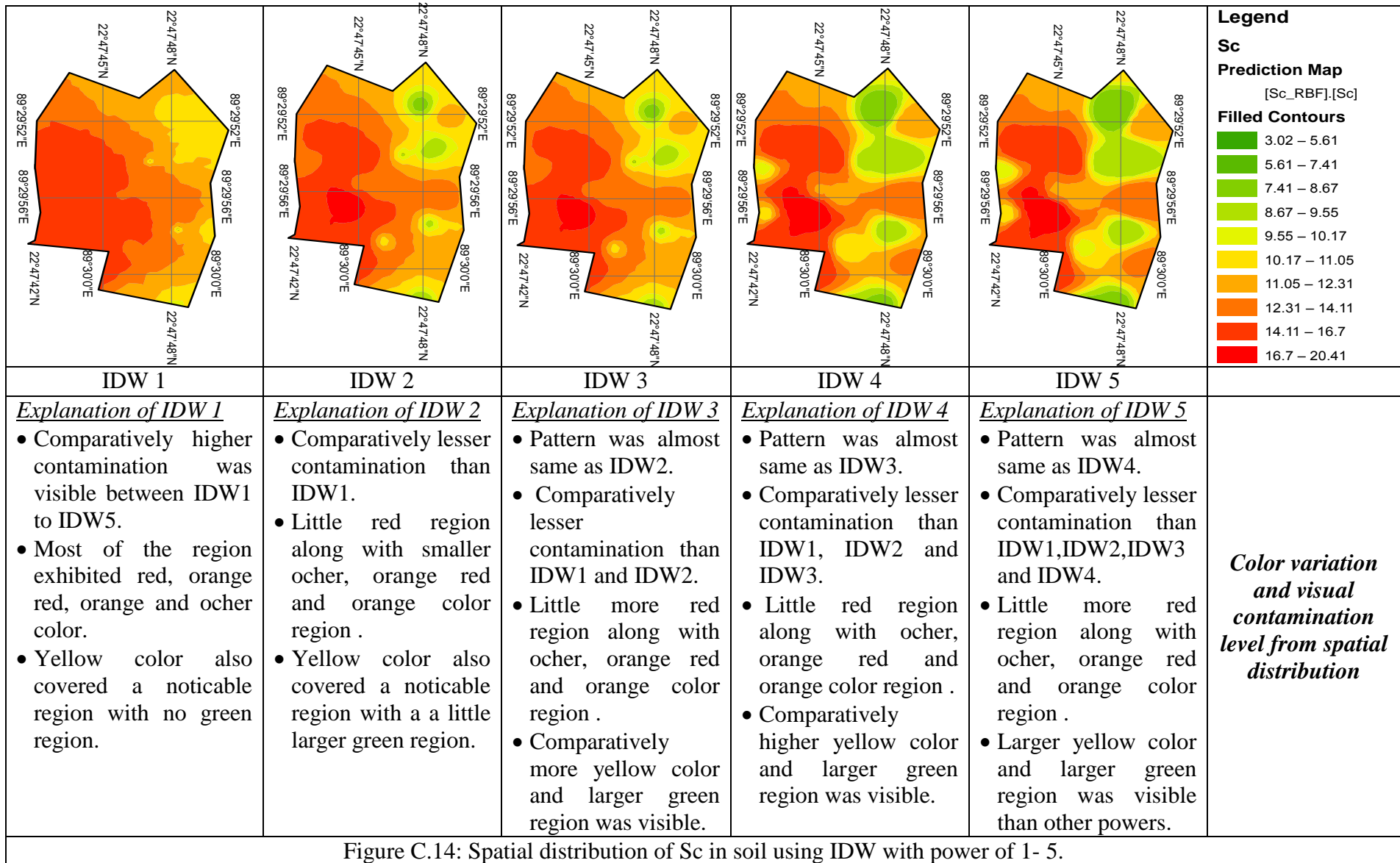


Figure C.14: Spatial distribution of Sc in soil using IDW with power of 1- 5.



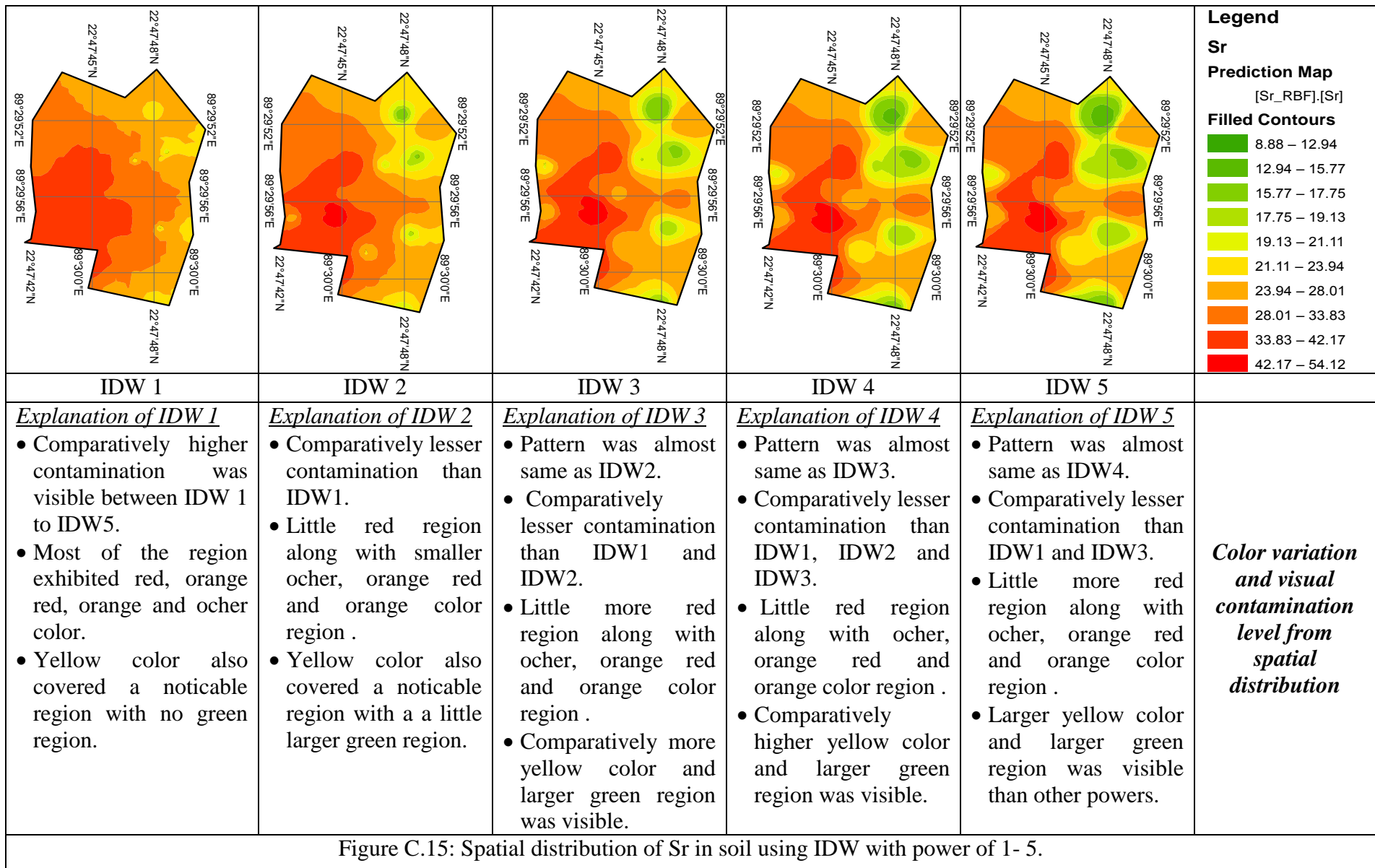
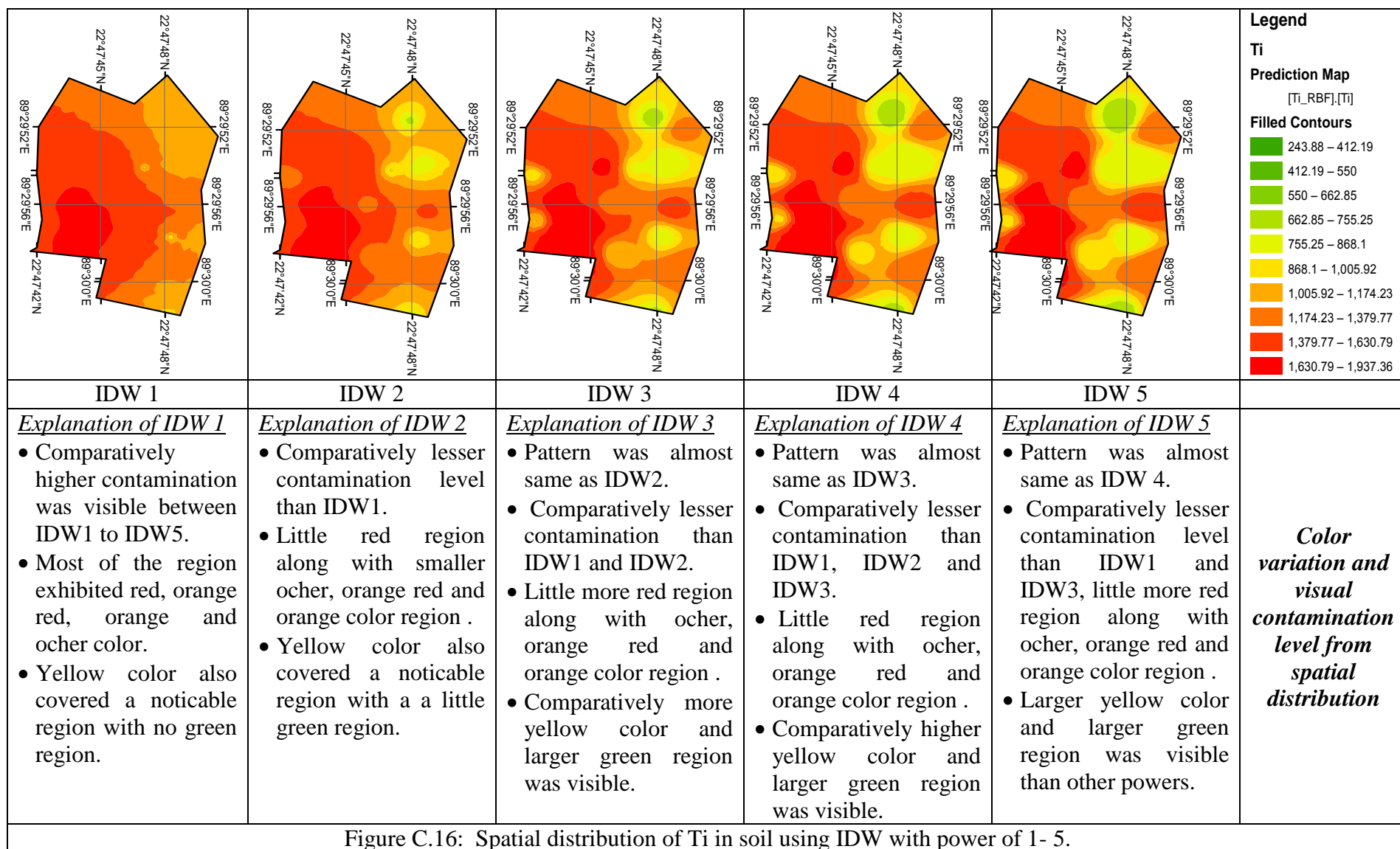


Figure C.15: Spatial distribution of Sr in soil using IDW with power of 1- 5.



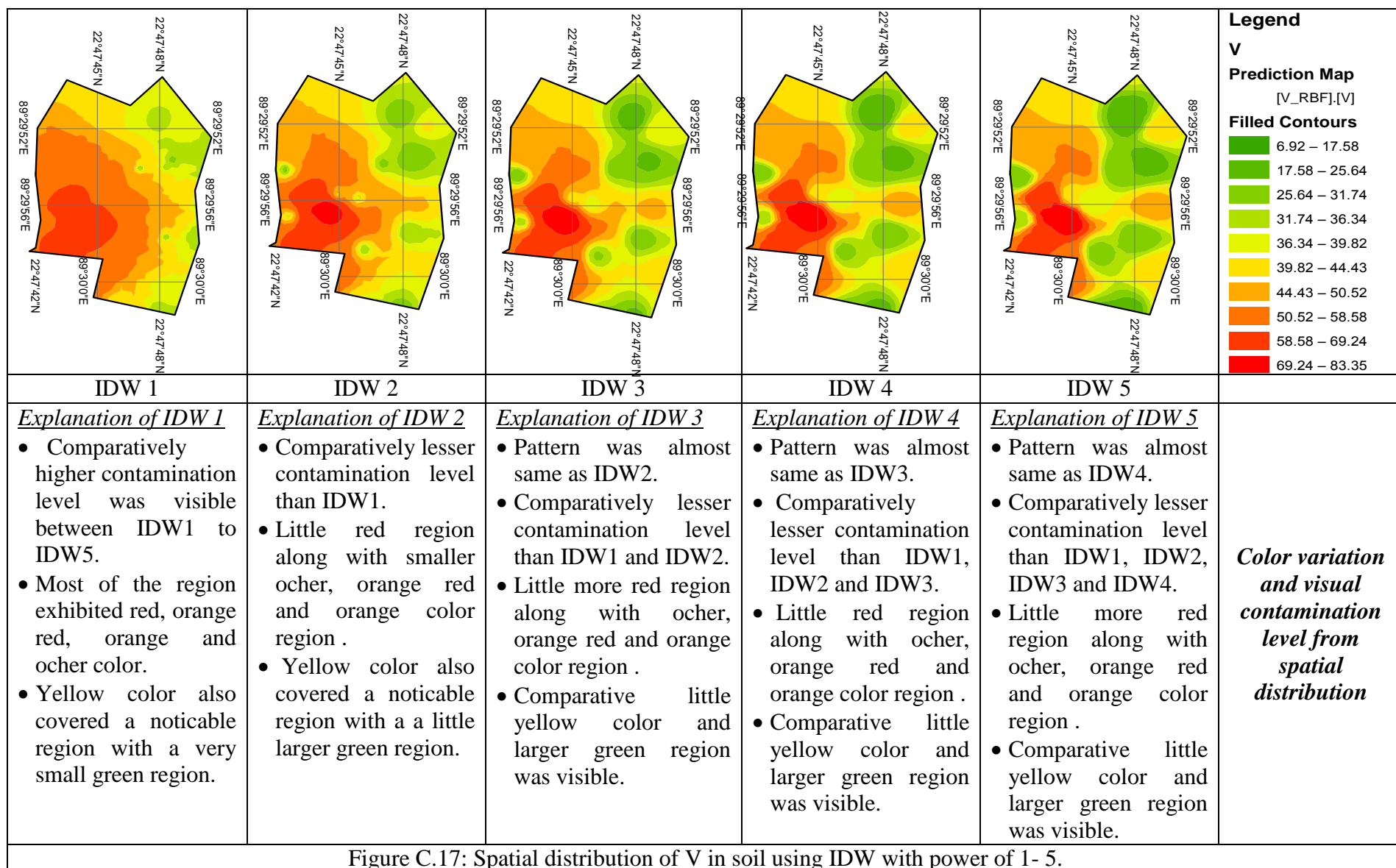


Figure C.17: Spatial distribution of V in soil using IDW with power of 1- 5.

**Annex –D**

**Cross Validation Results and Spatial Distribution of Metal Elements using LPI**

Table D.1: Cross validation of LP for Al

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	38.3474	22.5334	22.0472	17.9884	20.9309	0.0743
<sup>b</sup> RMSPE	149.4701	141.5211	138.0518	144.7918	141.7094	141.2395
<sup>c</sup> MSPE	0.0715	-0.0192	-0.0044	-0.0528	-0.0235	-0.1044
<sup>d</sup> RMSSPE	0.8683	0.8996	0.8834	0.9273	0.9055	0.9745
<sup>e</sup> ASPE	226.7868	215.0220	198.8298	215.6663	191.9776	167.4523
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	44.7554	59.7076	55.2460	61.3524	59.8381	44.7554
<sup>b</sup> RMSPE	158.2909	159.4017	160.3474	159.3847	166.7433	158.2909
<sup>c</sup> MSPE	0.1645	0.2122	0.2062	0.2267	0.2216	0.1645
<sup>d</sup> RMSSPE	0.9185	0.9067	0.9262	0.9151	0.9434	0.9185
<sup>e</sup> ASPE	187.8413	242.2169	196.3436	239.9590	198.4051	187.8413
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	36.8085	41.2147	37.8243	41.0332	39.7708	40.9897
<sup>b</sup> RMSPE	158.8542	170.1771	163.9462	170.9420	168.0804	182.2876
<sup>c</sup> MSPE	0.1198	0.1188	0.1122	0.1398	0.1258	0.1549
<sup>d</sup> RMSSPE	0.9066	0.9462	0.9381	0.9552	0.9490	0.9692
<sup>e</sup> ASPE	207.2221	209.9345	206.7214	206.1396	206.5775	206.9261

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table D.2: Cross validation of LP for As

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.4076	0.3524	0.3260	0.3260	0.3457	0.2890
<sup>b</sup> RMSPE	1.5350	1.5432	1.5227	1.5791	1.5494	1.6441
<sup>c</sup> MSPE	0.0505	-0.0029	0.0062	-0.0283	-0.0083	-0.0416
<sup>d</sup> RMSSPE	0.8508	0.8960	0.8862	0.9258	0.9047	0.9286
<sup>e</sup> ASPE	2.2839	2.4583	2.2414	2.4327	2.4275	5.3608
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.6729	0.8191	0.7792	0.8520	0.8324	0.6259
<sup>b</sup> RMSPE	1.7519	1.7917	1.7587	1.8259	1.7993	2.0168
<sup>c</sup> MSPE	0.2352	0.2687	0.2499	0.2892	0.2753	0.1651
<sup>d</sup> RMSSPE	0.8364	0.8988	0.8810	0.9106	0.8993	0.9783
<sup>e</sup> ASPE	3.0450	2.7237	2.7399	2.7058	2.7113	2.2364
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.6013	0.5266	0.5574	0.4138	0.4970	0.5908
<sup>b</sup> RMSPE	1.8618	1.7753	1.7784	1.6731	1.7741	1.8633
<sup>c</sup> MSPE	0.1139	0.0823	0.0997	0.0816	0.0721	0.1196
<sup>d</sup> RMSSPE	0.8635	0.9053	0.8984	0.9264	0.9023	0.9395
<sup>e</sup> ASPE	3.1453	3.1887	2.9450	2.3149	3.4999	2.6828

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table D.3: Cross validation result of LP for Ba

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.8571	1.1494	2.0730	0.3634	0.8361	-0.8900
<sup>b</sup> RMSPE	18.3029	14.5529	14.7503	15.0827	14.6713	14.6843
<sup>c</sup> MSPE	-0.0219	-0.0669	-0.0171	-0.1115	-0.0838	-0.1405
<sup>d</sup> RMSSPE	0.9230	0.8937	0.8797	0.9280	0.9038	0.9706
<sup>e</sup> ASPE	24.5952	23.7291	20.9875	20.8050	23.6350	17.5082
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	4.6885	5.8079	5.0301	6.0707	5.8498	7.3310
<sup>b</sup> RMSPE	16.4951	16.3065	16.2908	16.3134	16.2937	18.1627
<sup>c</sup> MSPE	0.2039	0.2161	0.2000	0.2345	0.2217	0.2624
<sup>d</sup> RMSSPE	0.8475	0.9062	0.8890	0.9174	0.9064	0.9404
<sup>e</sup> ASPE	35.3169	27.1548	28.9570	26.0910	27.1646	25.5920
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	2.6264	2.8849	2.2781	2.2226	2.4224	2.8833
<sup>b</sup> RMSPE	16.6680	17.8794	15.8664	16.2989	16.4151	19.9388
<sup>c</sup> MSPE	0.0930	0.0823	0.0175	-0.0061	0.0070	0.1237
<sup>d</sup> RMSSPE	0.9087	0.9497	0.8946	0.9120	0.9058	0.9753
<sup>e</sup> ASPE	21.9288	22.4668	34.1811	40.5156	36.6516	22.7962

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table D.4: Cross validation of LP for Ca

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	6.1596	6.2387	8.7111	4.2527	5.6160	5.2003
<sup>b</sup> RMSPE	45.2333	45.9318	46.6204	47.3228	46.5236	64.5781
<sup>c</sup> MSPE	0.0351	-0.0190	0.0254	-0.0614	-0.0312	-0.0861
<sup>d</sup> RMSSPE	0.8372	0.8955	0.8785	0.9275	0.9099	0.8868
<sup>e</sup> ASPE	65.7952	69.3335	62.6609	69.5774	60.6907	102.8070
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	13.5223	17.3282	15.5816	17.9585	17.0742	12.4510
<sup>b</sup> RMSPE	54.3850	51.7428	50.8292	51.1013	52.3593	55.2043
<sup>c</sup> MSPE	0.1886	0.1979	0.1831	0.2131	0.2007	0.1140
<sup>d</sup> RMSSPE	0.8472	0.9079	0.9239	0.9170	0.9400	0.9779
<sup>e</sup> ASPE	105.7380	80.4923	61.8614	76.9287	62.3092	61.1293
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	9.8409	11.0696	8.7111	10.8182	10.4356	10.7018
<sup>b</sup> RMSPE	50.4415	53.6921	46.6204	53.2313	52.5837	56.2887
<sup>c</sup> MSPE	0.1047	0.0998	0.0254	0.1211	0.1077	0.1350
<sup>d</sup> RMSSPE	0.9024	0.9426	0.8785	0.9508	0.9457	0.9624
<sup>e</sup> ASPE	65.2710	66.0995	62.6609	64.4359	64.5886	66.0564

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table D.5: Cross validation of LP for Co

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.1109	0.0525	0.1382	-0.0572	0.0109	-0.1794
<sup>b</sup> RMSPE	1.1691	1.3074	1.3196	1.3760	1.3256	1.4657
<sup>c</sup> MSPE	-0.0164	-0.1131	-0.0533	-0.1860	-0.1396	-0.2596
<sup>d</sup> RMSSPE	0.8548	0.9118	0.9148	0.9360	0.9195	0.9574
<sup>e</sup> ASPE	1.5581	1.8736	1.6192	1.9511	1.8862	2.1057
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.4978	0.5489	0.5105	0.5945	0.5715	0.5835
<sup>b</sup> RMSPE	1.4495	1.4725	1.4532	1.4981	1.4830	1.5901
<sup>c</sup> MSPE	0.2296	0.2542	0.2342	0.2790	0.2641	0.2859
<sup>d</sup> RMSSPE	0.8947	0.9256	0.9175	0.9305	0.9258	0.9605
<sup>e</sup> ASPE	1.8667	1.9714	1.9223	2.0181	2.0127	1.7400
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.2088	0.1808	0.2237	0.1032	0.2847	0.3318
<sup>b</sup> RMSPE	1.4155	1.4034	1.4171	1.6688	1.5285	1.6011
<sup>c</sup> MSPE	0.0189	-0.0131	0.0121	-0.0628	0.0959	0.1369
<sup>d</sup> RMSSPE	0.9164	0.9301	0.9251	0.9608	0.9519	0.9651
<sup>e</sup> ASPE	2.2783	2.2836	2.2128	1.8818	1.7702	1.8511

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.



Table D.6: Cross validation of LP for Cr

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.7301	0.7566	0.8245	0.7085	0.7464	0.4587
<sup>b</sup> RMSPE	2.1203	2.2708	2.2941	2.2816	2.2745	2.6814
<sup>c</sup> MSPE	0.1398	0.1099	0.1463	0.0762	0.1049	0.0744
<sup>d</sup> RMSSPE	0.8502	0.9390	0.9255	0.9585	0.9403	0.9887
<sup>e</sup> ASPE	2.5769	2.5150	2.6326	2.3734	2.5122	3.3372
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.9772	0.9799	0.9869	0.9620	0.9753	1.0834
<sup>b</sup> RMSPE	2.3716	2.4278	2.4104	2.4279	2.4254	2.4502
<sup>c</sup> MSPE	0.2186	0.2156	0.2213	0.2076	0.2140	0.2634
<sup>d</sup> RMSSPE	0.9115	0.9538	0.9420	0.9559	0.9524	0.9535
<sup>e</sup> ASPE	2.4790	2.4048	2.4303	2.4066	2.4083	2.5069
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.8541	0.7667	0.8744	0.9303	0.8677	0.8764
<sup>b</sup> RMSPE	2.3274	2.5268	2.3600	2.3843	2.3529	2.3714
<sup>c</sup> MSPE	0.1764	0.1145	0.1813	0.2011	0.1847	0.1953
<sup>d</sup> RMSSPE	0.9203	0.9317	0.9411	0.9475	0.9450	0.9543
<sup>e</sup> ASPE	2.5380	2.9201	2.5049	2.5250	2.4655	2.4854

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table D.7: Cross validation of LP for Cu

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.6032	0.3770	0.3925	0.5378	0.4081	0.0885
<sup>b</sup> RMSPE	2.4525	2.5627	2.5000	2.5361	2.5796	3.6902
<sup>c</sup> MSPE	0.0447	-0.0133	0.0005	-0.0042	-0.0066	-0.1533
<sup>d</sup> RMSSPE	0.8303	0.9034	0.8838	0.9382	0.9137	0.9585
<sup>e</sup> ASPE	4.1827	4.5969	4.2428	4.1410	4.5913	7.1241
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.8536	1.1235	1.0635	1.2643	1.1693	1.5250
<sup>b</sup> RMSPE	3.3442	3.2290	3.1319	3.2889	3.2317	4.0559
<sup>c</sup> MSPE	0.2099	0.2417	0.2282	0.2700	0.2512	0.2730
<sup>d</sup> RMSSPE	0.8654	0.9279	0.9184	0.9308	0.9263	0.9144
<sup>e</sup> ASPE	5.7158	5.1581	5.0615	5.5094	5.2507	137.2448
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.4153	0.3793	0.3616	0.5075	0.3789	0.6472
<sup>b</sup> RMSPE	3.0396	3.1774	3.2668	3.0417	3.1040	3.2852
<sup>c</sup> MSPE	0.0602	0.0444	0.0445	0.0357	0.0355	0.0559
<sup>d</sup> RMSSPE	0.8985	0.9365	0.9282	0.9390	0.9347	0.9502
<sup>e</sup> ASPE	6.1200	5.9179	5.9814	6.1621	6.1004	6.1687

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table D.8 Cross validation of LP for Fe

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	183.2705	195.3167	174.5920	173.9638	168.1433	87.7411
<sup>b</sup> RMSPE	523.4628	538.8153	526.6695	526.6449	525.5781	551.9420
<sup>c</sup> MSPE	0.1167	0.1229	0.0934	0.0843	0.0750	0.0246
<sup>d</sup> RMSSPE	0.8779	0.9200	0.9226	0.9289	0.9406	0.9311
<sup>e</sup> ASPE	566.9098	574.5741	547.0827	584.9735	538.1291	708.1670
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	228.0063	231.3812	230.8207	230.0194	230.6624	225.5643
<sup>b</sup> RMSPE	554.0003	559.6246	558.8425	559.0286	559.2687	557.7059
<sup>c</sup> MSPE	0.2164	0.2267	0.2243	0.2247	0.2255	0.2164
<sup>d</sup> RMSSPE	0.8862	0.9386	0.9284	0.9402	0.9368	0.9528
<sup>e</sup> ASPE	568.2313	537.9670	544.1347	538.3904	539.1265	530.6620
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	212.8792	214.6612	216.0453	222.7809	230.7473	223.6761
<sup>b</sup> RMSPE	548.4714	556.2259	550.8707	551.7820	566.6045	550.8948
<sup>c</sup> MSPE	0.1938	0.1985	0.2004	0.2134	0.2120	0.2213
<sup>d</sup> RMSSPE	0.8931	0.9264	0.9199	0.9284	0.9253	0.9355
<sup>e</sup> ASPE	589.2222	570.1125	573.9597	574.0499	591.3344	571.6585

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table D.9: Cross validation of LP for Hg

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.5558	0.6096	0.5914	0.5152	0.5196	0.5416
<sup>b</sup> RMSPE	1.8675	2.0502	1.9208	1.9389	1.9184	2.5691
<sup>c</sup> MSPE	0.0889	0.0288	0.0656	0.0190	0.0293	-0.0756
<sup>d</sup> RMSSPE	0.9108	0.9359	0.9410	0.9523	0.9438	0.9441
<sup>e</sup> ASPE	2.2292	21798214.1900	2.3130	2.3863	2.3514	3.4805
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.9273	0.9373	0.9832	0.9141	0.9522	0.9534
<sup>b</sup> RMSPE	2.2317	2.2495	2.2692	2.2422	2.2368	2.4018
<sup>c</sup> MSPE	0.2404	0.2443	0.2560	0.2522	0.2485	0.2068
<sup>d</sup> RMSSPE	0.9378	0.9417	0.9470	0.9341	0.9380	0.9811
<sup>e</sup> ASPE	2.4466	3.2593	2.5495	3.5903	3.2315	2.5943
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.6885	0.6325	0.6739	0.7349	0.7297	0.6999
<sup>b</sup> RMSPE	2.1904	2.2222	2.2775	2.3546	2.3135	2.4772
<sup>c</sup> MSPE	0.1721	0.0955	0.1565	0.1848	0.1743	0.1828
<sup>d</sup> RMSSPE	0.9611	0.9802	0.9742	0.9756	0.9734	0.9851
<sup>e</sup> ASPE	2.5826	2.4313	2.5988	2.6761	2.6722	2.6227

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table D.10: Cross validation of LP for K

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	24.3834	12.4496	14.7802	10.7630	12.2381	14.2105
<sup>b</sup> RMSPE	90.0177	88.3450	88.0716	88.6391	88.2734	103.2877
<sup>c</sup> MSPE	0.0880	-0.0353	-0.0077	-0.0603	-0.0403	0.0597
<sup>d</sup> RMSSPE	0.8951	0.9366	0.9288	0.9457	0.9376	0.9279
<sup>e</sup> ASPE	111.4180	108.9467	103.6384	108.5254	108.5496	168.0926
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	29.6641	29.6782	29.9749	28.6604	29.3338	28.8146
<sup>b</sup> RMSPE	93.4976	96.0027	95.5725	96.2202	95.9345	96.5028
<sup>c</sup> MSPE	0.1697	0.1625	0.1668	0.1533	0.1599	0.1544
<sup>d</sup> RMSSPE	0.9008	0.9497	0.9391	0.9564	0.9492	0.9669
<sup>e</sup> ASPE	105.5451	101.7685	102.9100	100.6093	101.6367	99.4799
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	26.1237	25.7761	26.1237	28.5106	30.0185	26.9004
<sup>b</sup> RMSPE	93.4137	97.8116	93.4137	95.2750	99.2911	94.8895
<sup>c</sup> MSPE	0.1367	0.1079	0.1367	0.1522	0.1512	0.1515
<sup>d</sup> RMSSPE	0.9106	0.9238	0.9106	0.9404	0.9387	0.9474
<sup>e</sup> ASPE	110.5204	152.0699	110.5204	109.4085	112.7709	106.7050

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table D.11: Cross validation of LP for Mn

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	2.4752	2.3410	2.4000	2.4749	2.3849	0.8529
<sup>b</sup> RMSPE	6.9628	6.6564	6.5899	6.7015	6.6676	6.8799
<sup>c</sup> MSPE	0.1077	0.0945	0.1054	0.1035	0.1000	-0.0051
<sup>d</sup> RMSSPE	0.9315	0.9511	0.9285	0.9498	0.9518	0.9452
<sup>e</sup> ASPE	9.0623	8.8016	9.1781	9.3061	8.7896	10.1622
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	3.0179	3.1444	2.9997	2.9539	2.9458	2.9788
<sup>b</sup> RMSPE	7.0494	7.2864	7.1897	7.2223	7.2674	7.1887
<sup>c</sup> MSPE	0.2102	0.2215	0.1949	0.1869	0.1896	0.1887
<sup>d</sup> RMSSPE	0.9405	0.9567	0.9653	0.9724	0.9697	0.9780
<sup>e</sup> ASPE	8.7966	9.2559	8.8728	8.8574	8.9303	8.7733
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	2.5631	2.8303	2.5517	2.6989	2.7061	3.0281
<sup>b</sup> RMSPE	7.1084	7.5660	7.1830	7.3362	7.2895	7.2708
<sup>c</sup> MSPE	0.1683	0.1703	0.1590	0.1814	0.1737	0.2080
<sup>d</sup> RMSSPE	0.9438	0.9606	0.9569	0.9627	0.9599	0.9702
<sup>e</sup> ASPE	9.5066	9.8590	9.4611	9.6051	9.6461	9.1412

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table D.12: Cross validation of LP for Na

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	6.9290	4.8300	5.4521	8.7793	5.0860	3.0048
<sup>b</sup> RMSPE	24.4825	24.2281	24.2850	25.8485	24.2722	24.7396
<sup>c</sup> MSPE	0.0736	0.0088	0.0305	0.1082	0.0130	-0.0523
<sup>d</sup> RMSSPE	0.9535	0.9505	0.9446	0.9961	0.9557	0.9703
<sup>e</sup> ASPE	28.9314	30.9012	30.7415	28.9047	30.8758	32.8656
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	8.9819	9.0801	9.0741	9.4030	9.2094	9.7725
<sup>b</sup> RMSPE	25.5843	25.6581	25.6355	25.6903	25.6561	25.7944
<sup>c</sup> MSPE	0.1752	0.1783	0.1797	0.1911	0.1834	0.2036
<sup>d</sup> RMSSPE	0.9430	0.9605	0.9620	0.9685	0.9624	0.9744
<sup>e</sup> ASPE	28.7541	28.3623	28.1163	27.8549	28.1733	27.7148
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	9.2009	9.2009	9.2898	9.1895	9.5348	8.9639
<sup>b</sup> RMSPE	26.5412	26.5412	26.8312	26.8870	27.7154	26.2035
<sup>c</sup> MSPE	0.1695	0.1695	0.1722	0.1692	0.1709	0.1702
<sup>d</sup> RMSSPE	0.9280	0.9280	0.9459	0.9524	0.9467	0.9595
<sup>e</sup> ASPE	31.7319	31.7319	31.1203	31.0220	31.9916	30.2065

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table D.13 Cross validation of LP for Sb

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.1919	0.1995	0.2643	0.1316	0.1710	-0.4740
<sup>b</sup> RMSPE	1.9285	1.5516	1.5714	1.5831	1.5530	1.5536
<sup>c</sup> MSPE	0.0622	-0.0363	-0.0057	-0.0741	-0.0494	-0.3464
<sup>d</sup> RMSSPE	0.8923	0.9011	0.8911	0.9298	0.9098	0.8598
<sup>e</sup> ASPE	2.8898	2.5364	2.3020	2.5331	2.5155	4.4231
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.5380	0.6577	0.6088	0.6768	0.6044	0.4238
<sup>b</sup> RMSPE	1.6881	1.7230	1.7046	1.7311	1.8037	1.8445
<sup>c</sup> MSPE	0.1964	0.2145	0.1994	0.2310	0.1946	0.1076
<sup>d</sup> RMSSPE	0.8598	0.9135	0.8968	0.9216	0.9512	0.9798
<sup>e</sup> ASPE	2.9985	2.8055	2.8494	2.8278	2.2761	2.2168
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.3838	0.4462	0.3509	0.4278	0.4149	0.4198
<sup>b</sup> RMSPE	1.6775	1.8657	1.7419	1.7929	1.7849	1.8781
<sup>c</sup> MSPE	0.1027	0.1050	0.0900	0.1239	0.1109	0.1362
<sup>d</sup> RMSSPE	0.9056	0.9457	0.9425	0.9527	0.9482	0.9643
<sup>e</sup> ASPE	2.4041	2.4559	2.3687	2.3985	2.4021	2.4529

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.



Table D.14: Cross validation of LP for Sc

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.5435	0.2709	0.3087	0.2562	0.2562	0.3518
<sup>b</sup> RMSPE	2.4791	2.4476	2.4372	2.4599	2.4599	3.6293
<sup>c</sup> MSPE	0.0669	-0.0202	-0.0078	-0.0336	-0.0336	-0.0122
<sup>d</sup> RMSSPE	0.8928	0.9086	0.9073	0.9202	0.9202	0.9271
<sup>e</sup> ASPE	3.4692	3.4543	3.0242	3.0372	3.0372	4.6562
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.7424	0.8149	0.7538	0.8640	0.7363	0.8724
<sup>b</sup> RMSPE	2.5662	2.6553	2.6362	2.6468	2.6899	2.7767
<sup>c</sup> MSPE	0.1765	0.1906	0.1752	0.2071	0.1703	0.2204
<sup>d</sup> RMSSPE	0.9002	0.9210	0.9309	0.9266	0.9481	0.9549
<sup>e</sup> ASPE	3.0715	3.7565	3.0844	3.7022	3.0596	3.3540
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	0.4627	0.5310	0.2960	0.5071	0.4935	0.5234
<sup>b</sup> RMSPE	2.5894	2.8004	2.7359	2.6662	2.6793	2.7353
<sup>c</sup> MSPE	0.0898	0.0859	0.0139	0.1047	0.0926	0.1225
<sup>d</sup> RMSSPE	0.9139	0.9471	0.9144	0.9517	0.9484	0.9608
<sup>e</sup> ASPE	3.2365	3.2928	5.1505	3.2041	3.2155	3.2181

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table D.15: Cross validation of LP for Sr

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	1.2621	0.6048	0.8960	0.3570	0.4901	0.6000
<sup>b</sup> RMSPE	6.2697	6.0837	6.1866	6.2596	6.1202	8.2229
<sup>c</sup> MSPE	0.0394	-0.0367	-0.0015	-0.0698	-0.0500	0.0153
<sup>d</sup> RMSSPE	0.8549	0.8997	0.8896	0.9295	0.9090	0.9154
<sup>e</sup> ASPE	9.9260	9.5698	8.5826	8.4181	9.5048	13.5659
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	2.0750	2.2732	1.9352	2.3195	2.0420	1.4485
<sup>b</sup> RMSPE	6.5960	6.8420	7.0409	6.8327	7.1743	7.3073
<sup>c</sup> MSPE	0.1837	0.1959	0.1843	0.2067	0.1745	0.0968
<sup>d</sup> RMSSPE	0.8677	0.9151	0.8989	0.9231	0.9530	0.9794
<sup>e</sup> ASPE	10.5401	11.0016	11.9573	10.8568	8.5890	8.3549
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	1.2219	1.4605	1.0412	1.3558	1.3267	1.3732
<sup>b</sup> RMSPE	6.8222	7.5907	7.1267	7.3086	7.2749	7.6491
<sup>c</sup> MSPE	0.0905	0.0906	0.0737	0.1070	0.0956	0.1210
<sup>d</sup> RMSSPE	0.9108	0.9503	0.9490	0.9572	0.9528	0.9686
<sup>e</sup> ASPE	9.0857	9.2670	8.8339	9.0345	9.0571	9.2764

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table D.16: Cross validation of LP for Ti

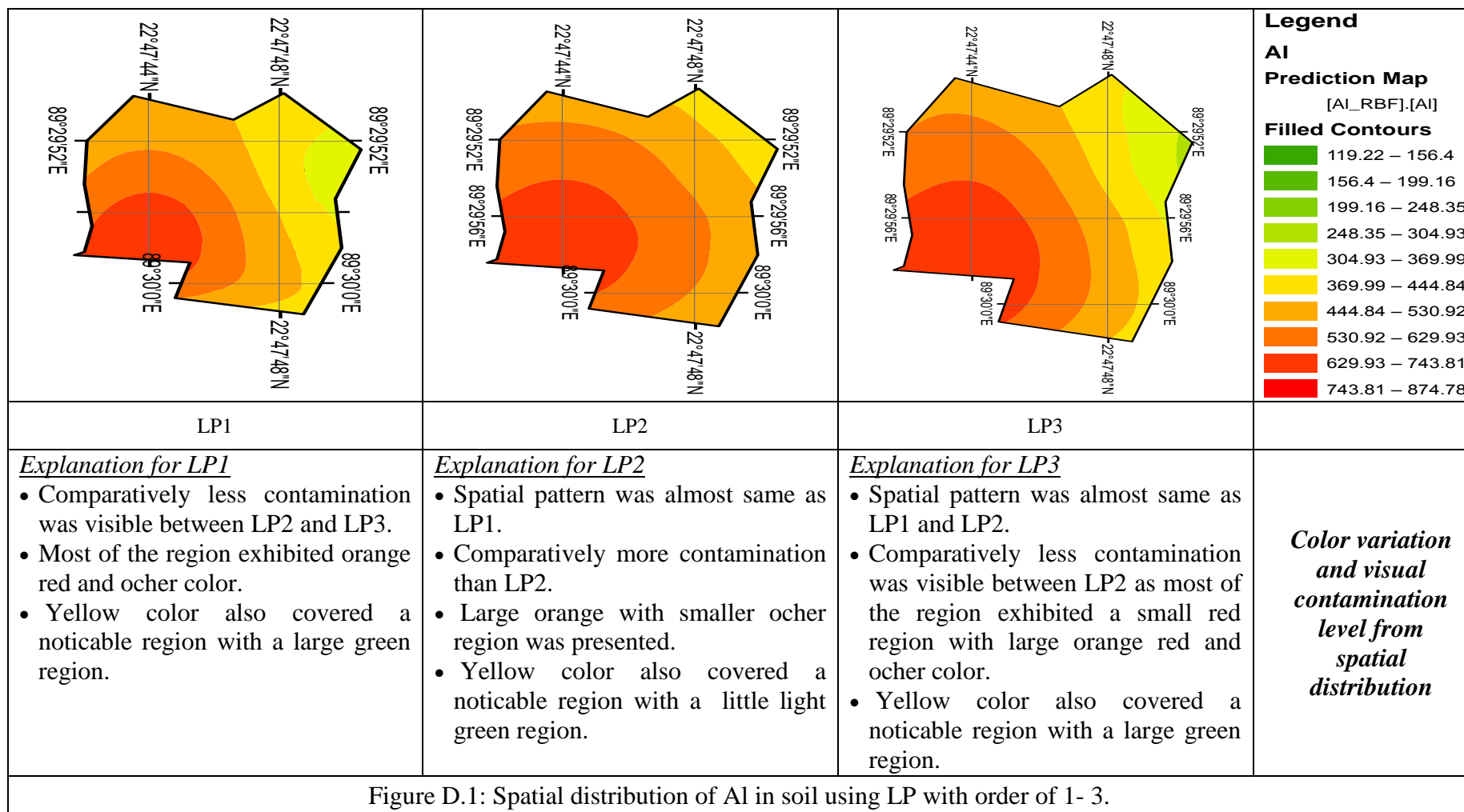
Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	27.8406	34.4402	35.9088	18.0595	27.8406	-18.7581
<sup>b</sup> RMSPE	283.2086	282.1172	282.6346	288.3609	283.2086	349.6378
<sup>c</sup> MSPE	-0.0505	-0.0307	-0.0194	-0.0853	-0.0505	-0.2601
<sup>d</sup> RMSSPE	0.9138	0.9069	0.9072	0.9303	0.9138	0.9593
<sup>e</sup> ASPE	408.0986	410.6313	360.2554	410.8174	408.0986	557.3885
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	81.1144	81.1144	86.8835	107.4198	88.3055	68.9752
<sup>b</sup> RMSPE	300.5327	300.5327	311.0297	309.6579	318.3270	326.3302
<sup>c</sup> MSPE	0.1594	0.1594	0.1671	0.2155	0.1677	0.1069
<sup>d</sup> RMSSPE	0.8998	0.8998	0.9346	0.9175	0.9485	0.9749
<sup>e</sup> ASPE	355.4000	355.4000	355.8500	452.8700	356.8215	349.4247
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	41.2382	71.8272	56.7519	64.2719	57.1410	65.7470
<sup>b</sup> RMSPE	242.6174	340.5921	312.0925	316.1629	294.6907	320.5818
<sup>c</sup> MSPE	0.1104	0.1013	0.0883	0.1137	0.0668	0.1303
<sup>d</sup> RMSSPE	0.9765	0.9478	0.9430	0.9536	0.9336	0.9591
<sup>e</sup> ASPE	288.8967	387.8213	370.9719	372.6687	358.9454	373.6147

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table D.17: Cross validation of LP for V

Order	1					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	3.7059	2.2955	2.4294	2.0153	2.1748	-0.2915
<sup>b</sup> RMSPE	13.1669	12.3488	12.3179	12.5643	12.3519	17.1316
<sup>c</sup> MSPE	0.0804	0.0007	0.0181	-0.0231	-0.0067	-0.0841
<sup>d</sup> RMSSPE	0.8684	0.9038	0.8990	0.9307	0.9112	0.8919
<sup>e</sup> ASPE	18.9594	17.9464	15.9653	16.1406	17.8291	27.9832
Order	2					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	4.5060	4.9092	4.8307	5.0050	5.0052	3.8837
<sup>b</sup> RMSPE	13.1147	13.5578	13.4810	13.6689	14.0946	14.2245
<sup>c</sup> MSPE	0.2015	0.2080	0.2016	0.2226	0.2093	0.1327
<sup>d</sup> RMSSPE	0.8651	0.9152	0.9023	0.9202	0.9480	0.9768
<sup>e</sup> ASPE	19.7136	19.8385	19.7048	20.4761	16.6833	16.1783
Order	3					
Kernal function	Exponential	Polynomial	Gaussian	Epanechnikov	Quartic	Constant
<sup>a</sup> MPE	2.4563	3.6357	3.4081	2.6776	3.4860	3.4046
<sup>b</sup> RMSPE	11.9549	14.2389	14.1614	13.7793	14.0206	15.2368
<sup>c</sup> MSPE	0.1224	0.1199	0.1359	0.0271	0.1257	0.1473
<sup>d</sup> RMSSPE	0.9043	0.9455	0.9525	0.9270	0.9481	0.9689
<sup>e</sup> ASPE	16.1363	17.8743	17.2930	27.2407	17.5692	17.9151

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.



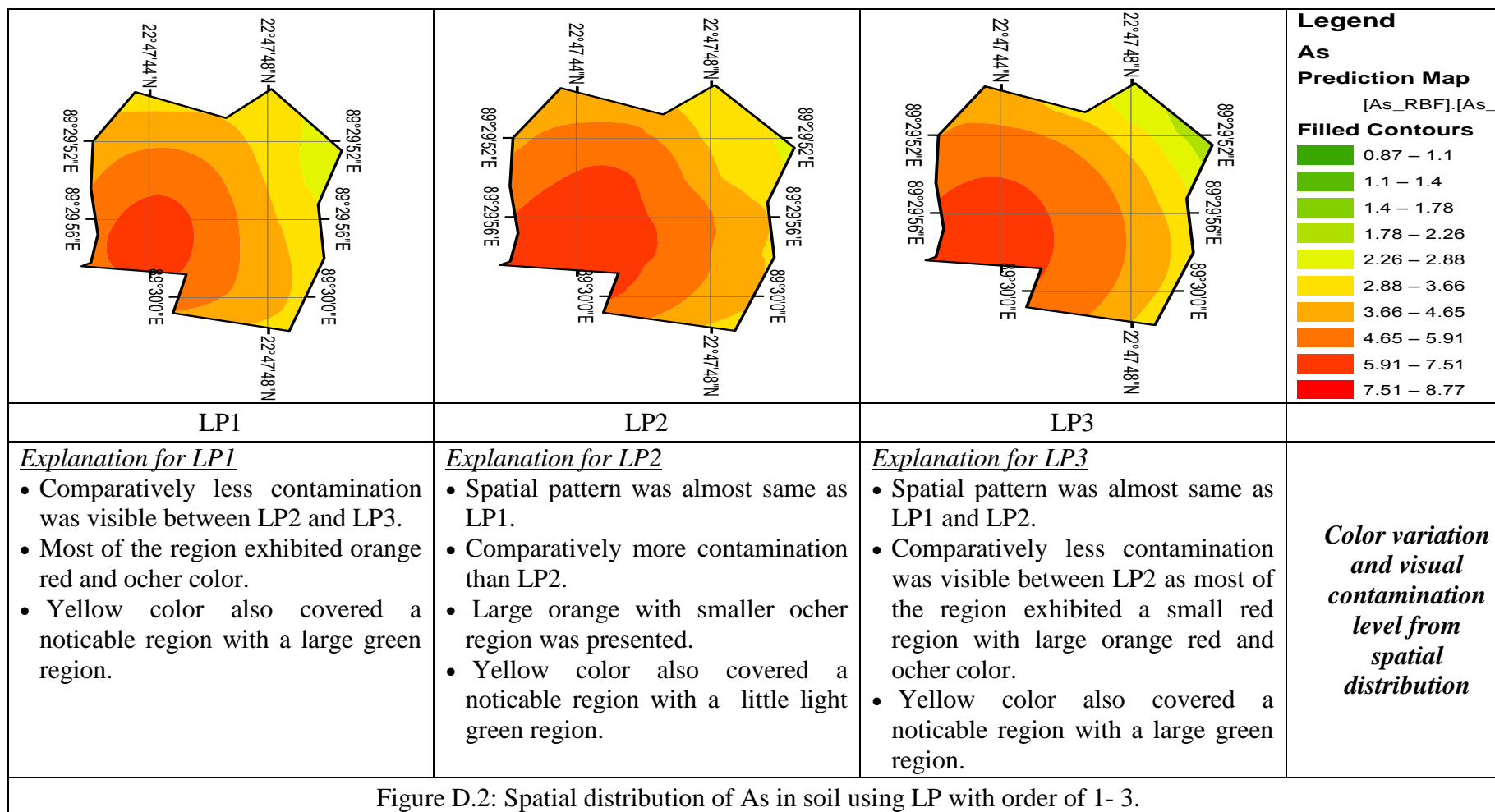
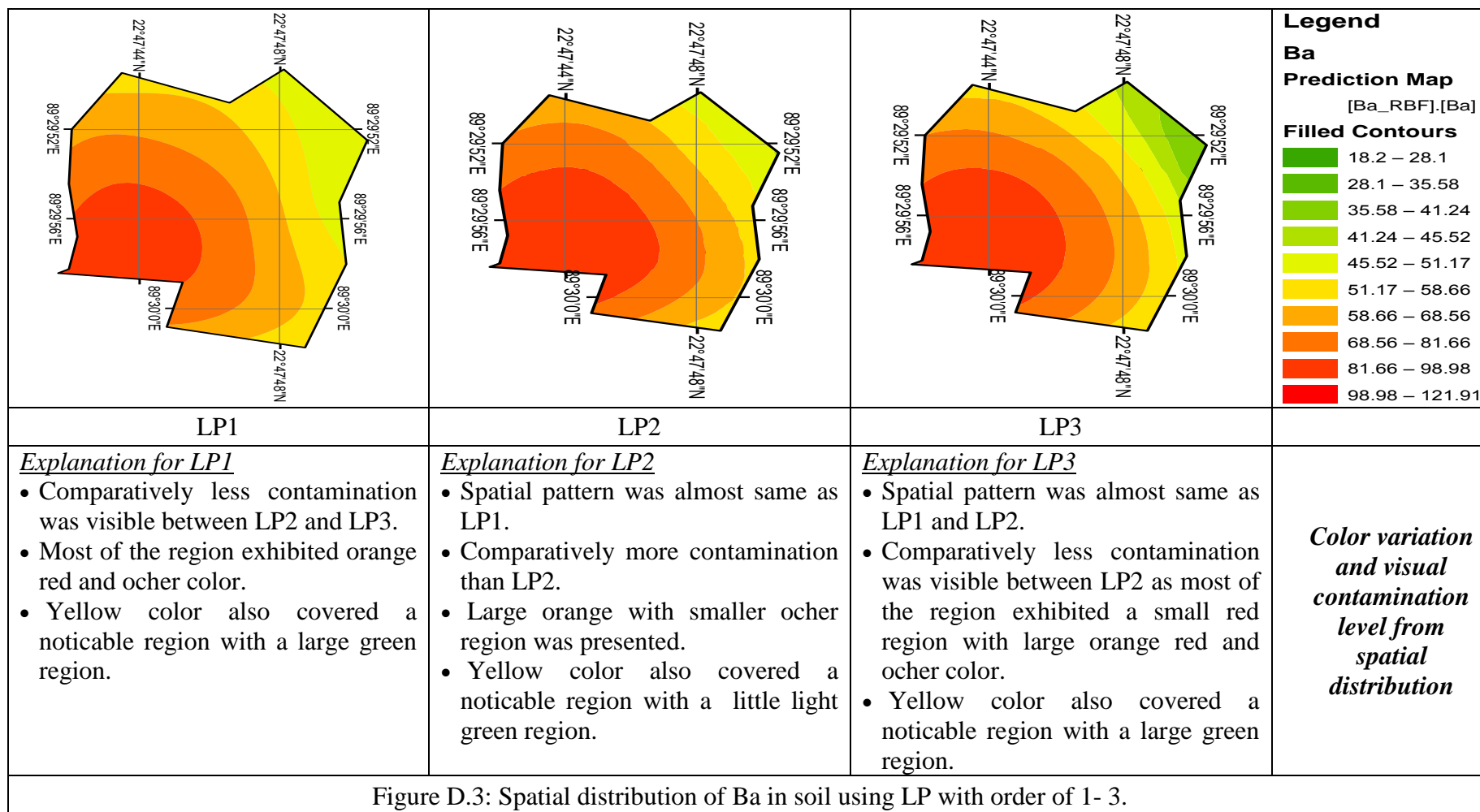


Figure D.2: Spatial distribution of As in soil using LP with order of 1- 3.



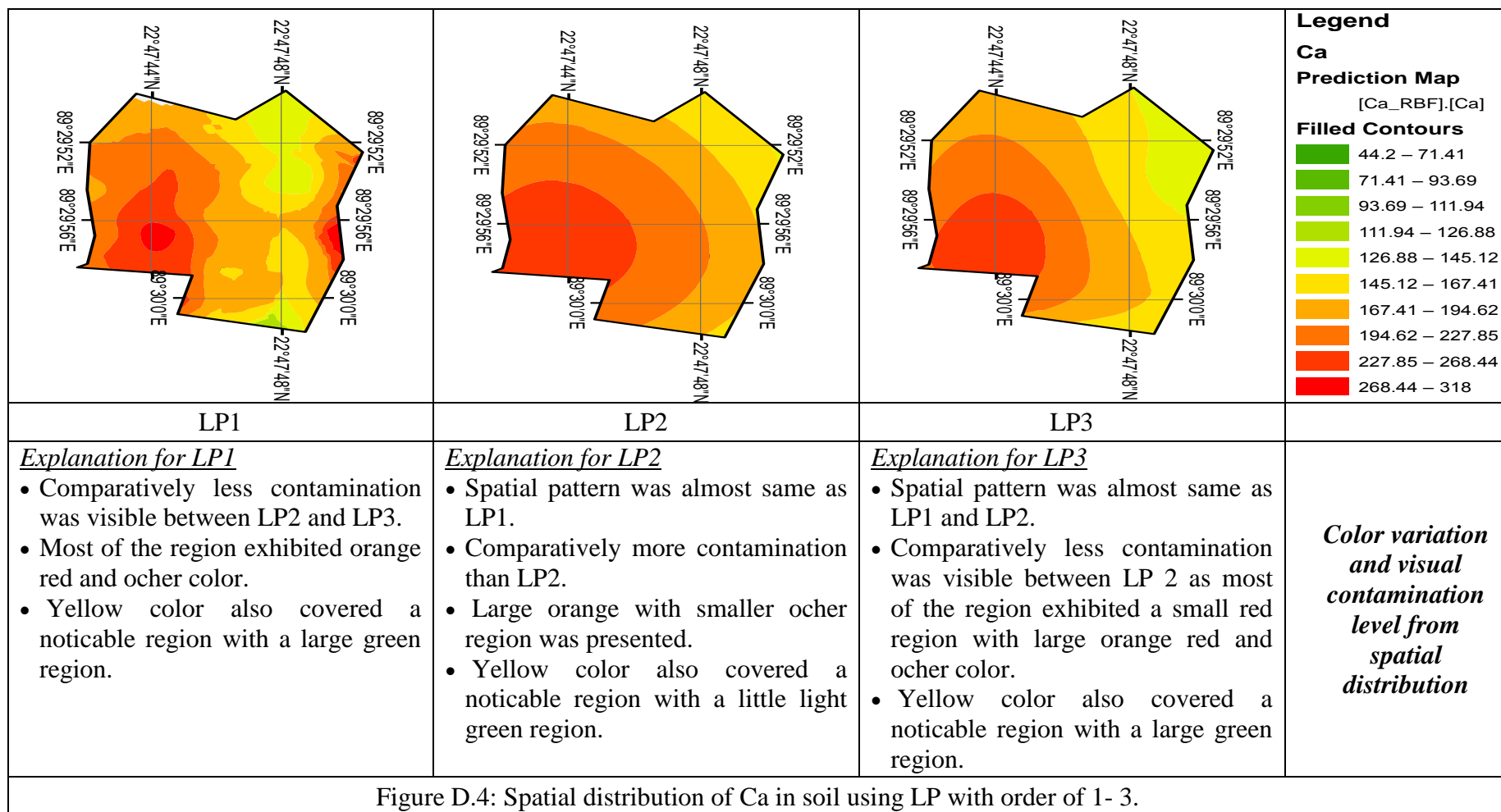
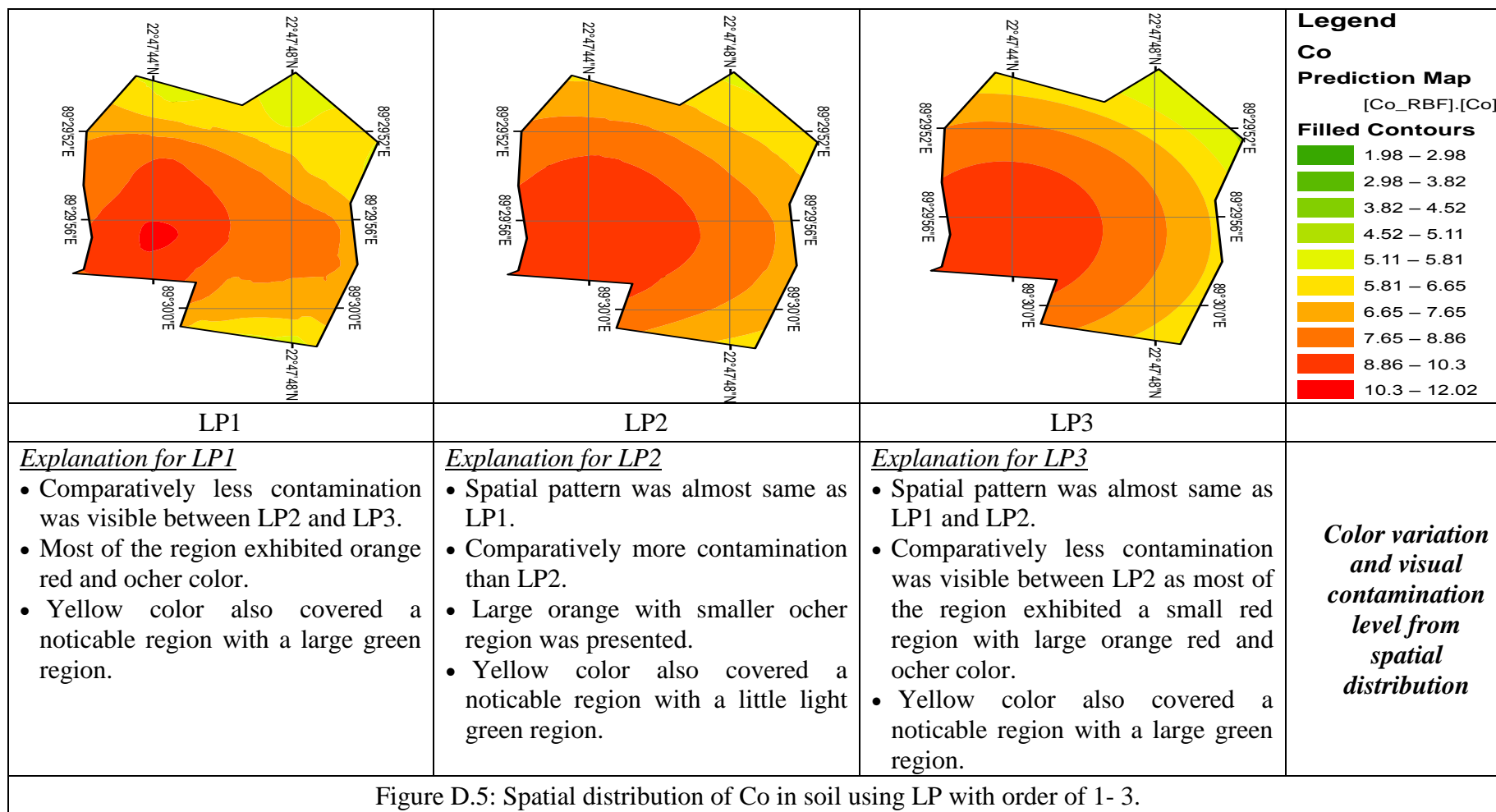
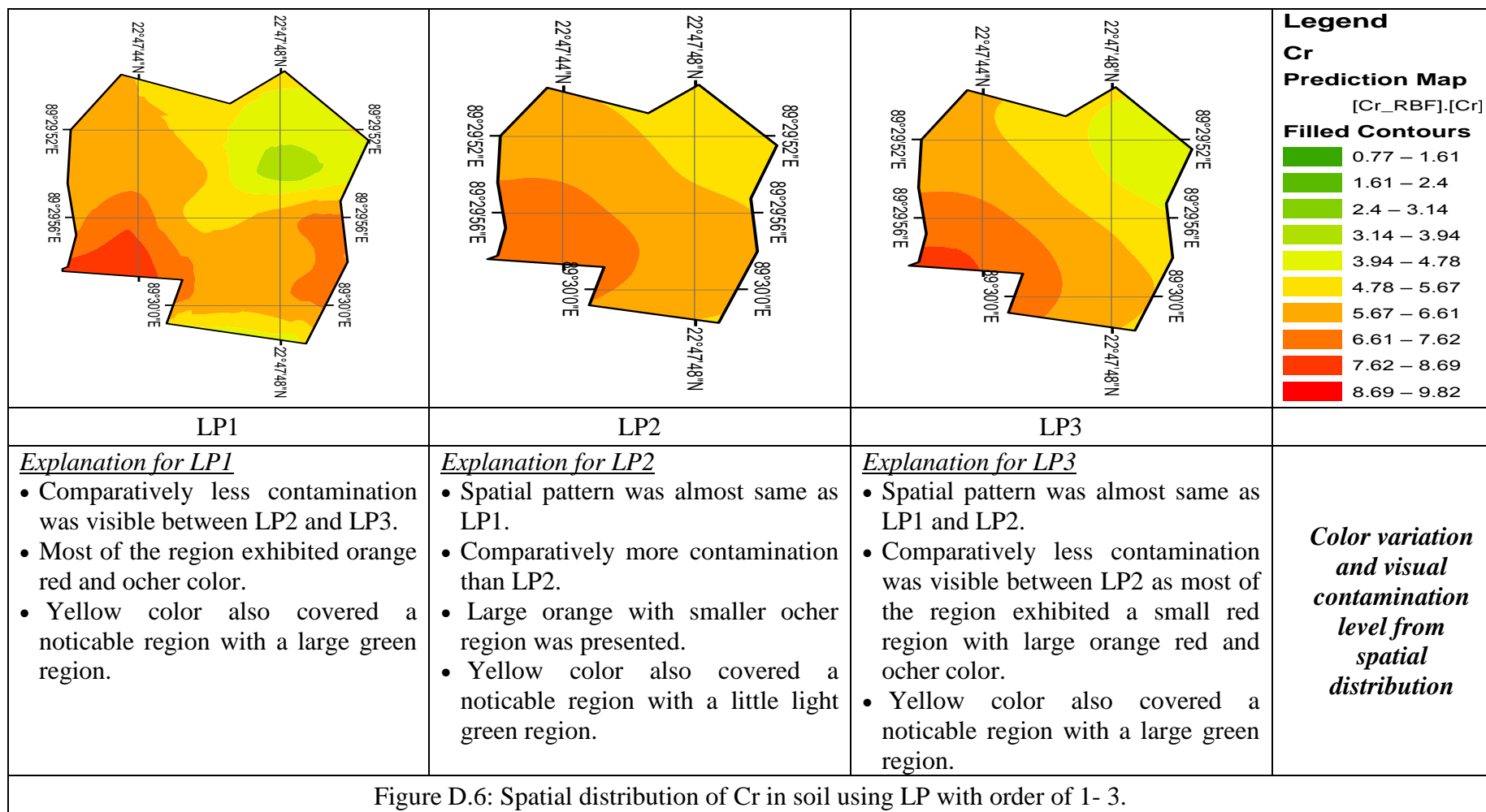
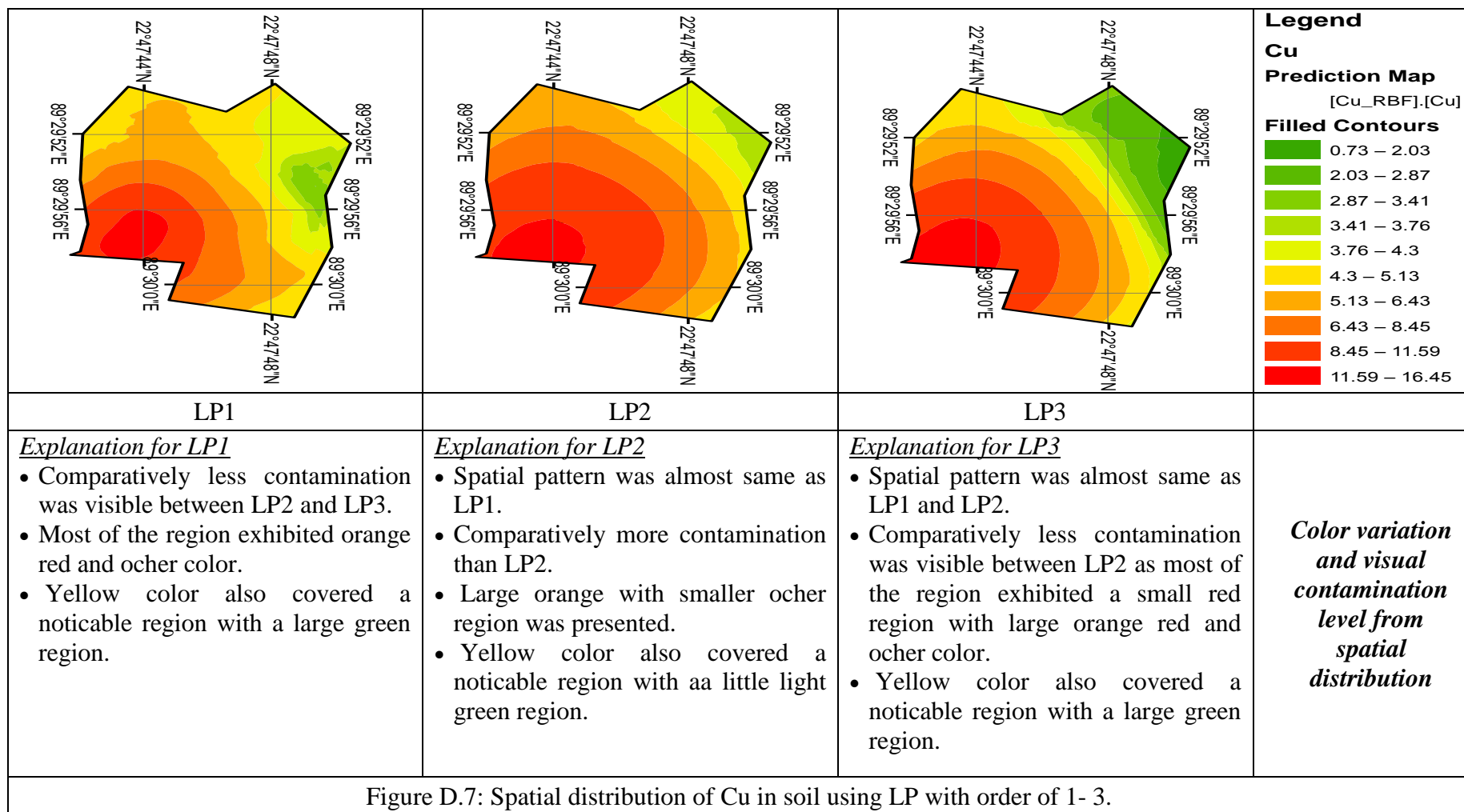


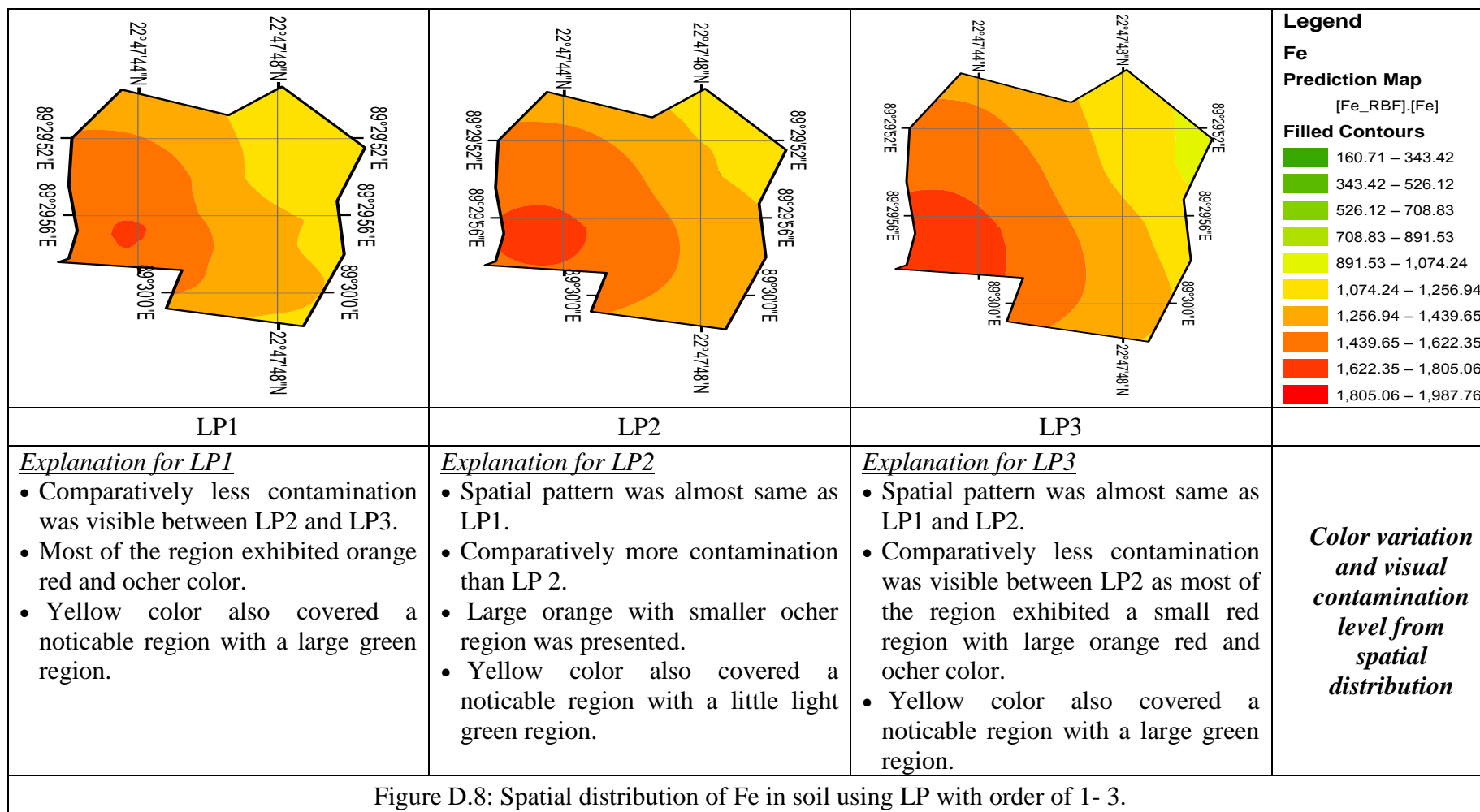
Figure D.4: Spatial distribution of Ca in soil using LP with order of 1- 3.

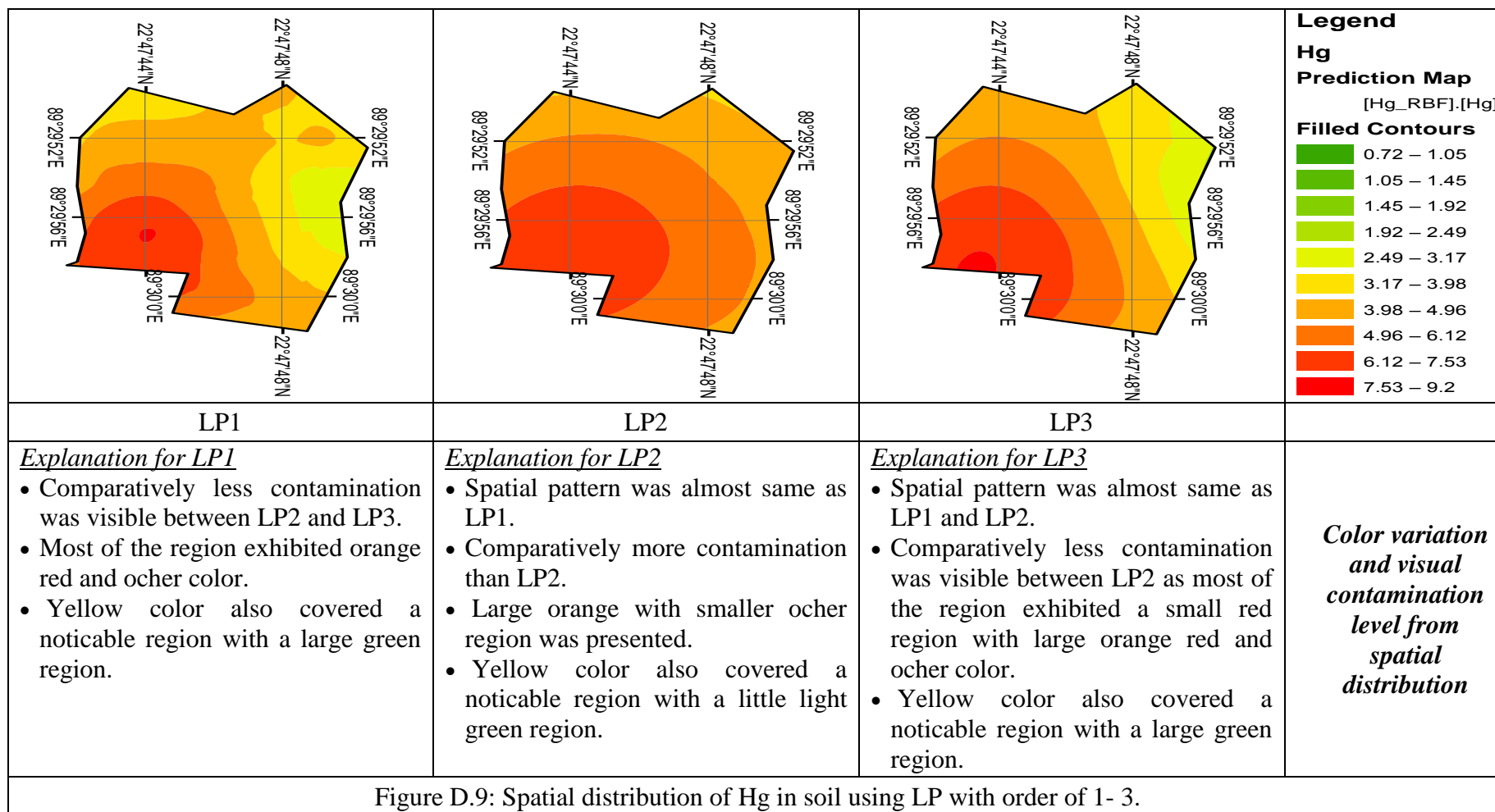


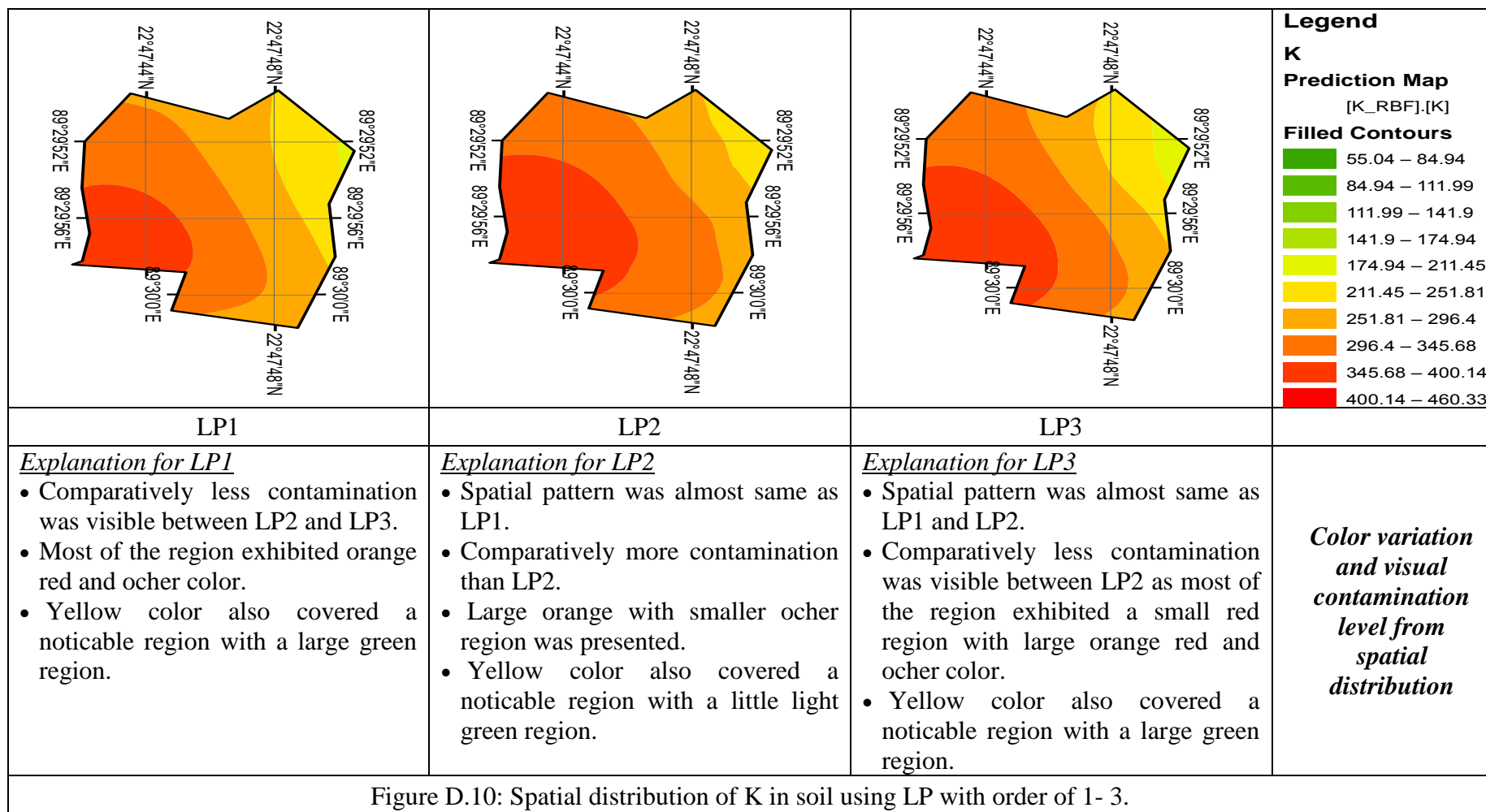












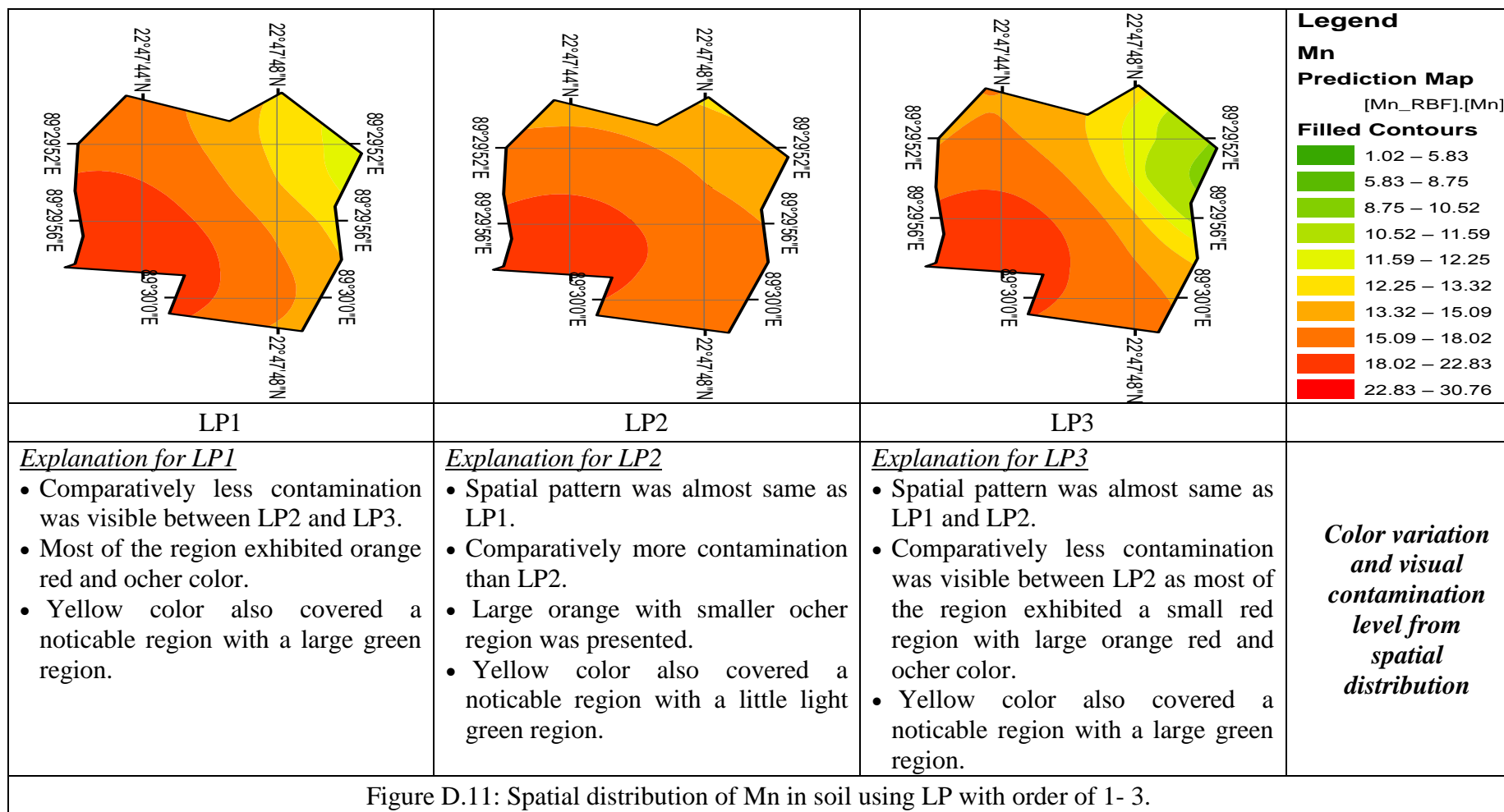
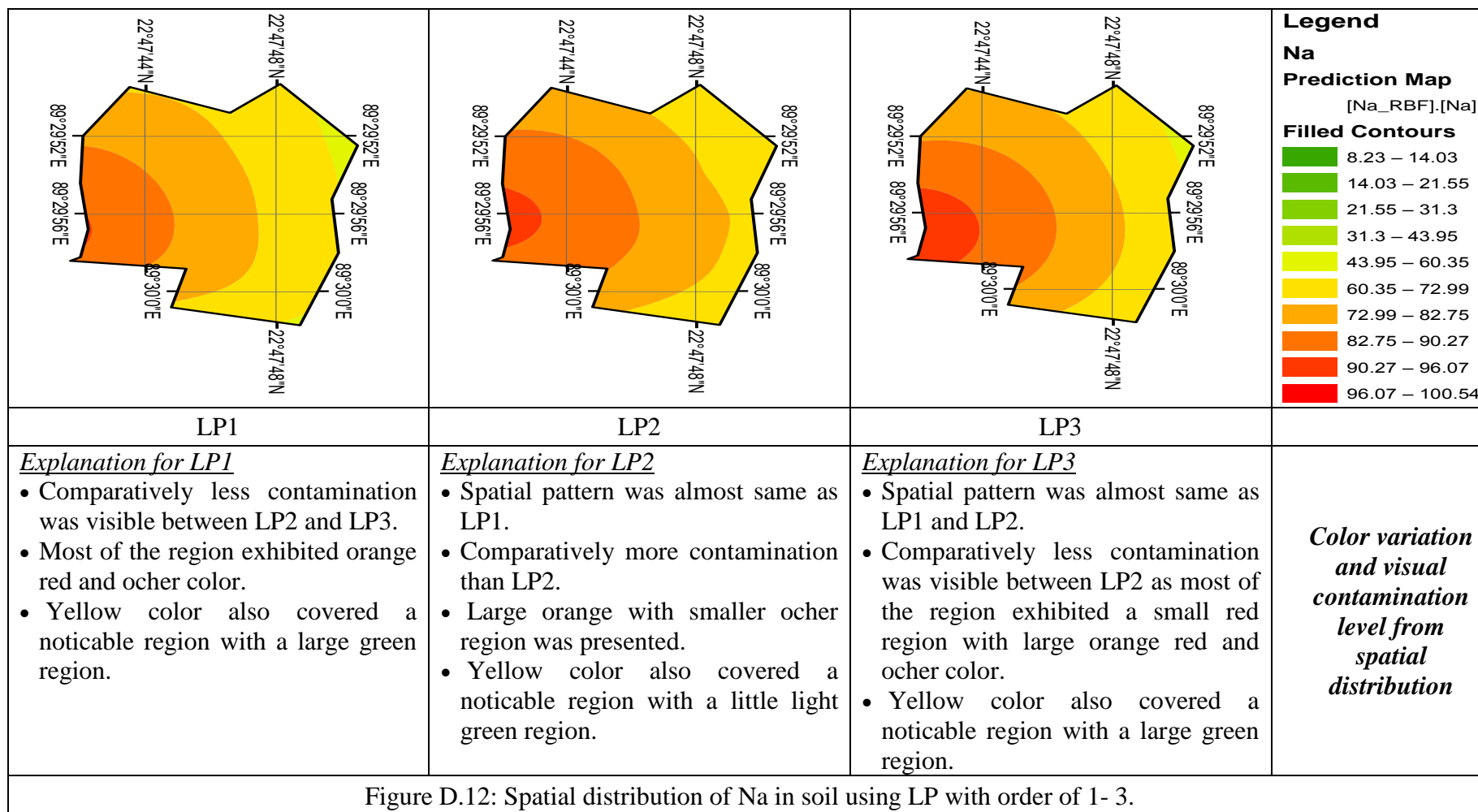


Figure D.11: Spatial distribution of Mn in soil using LP with order of 1- 3.





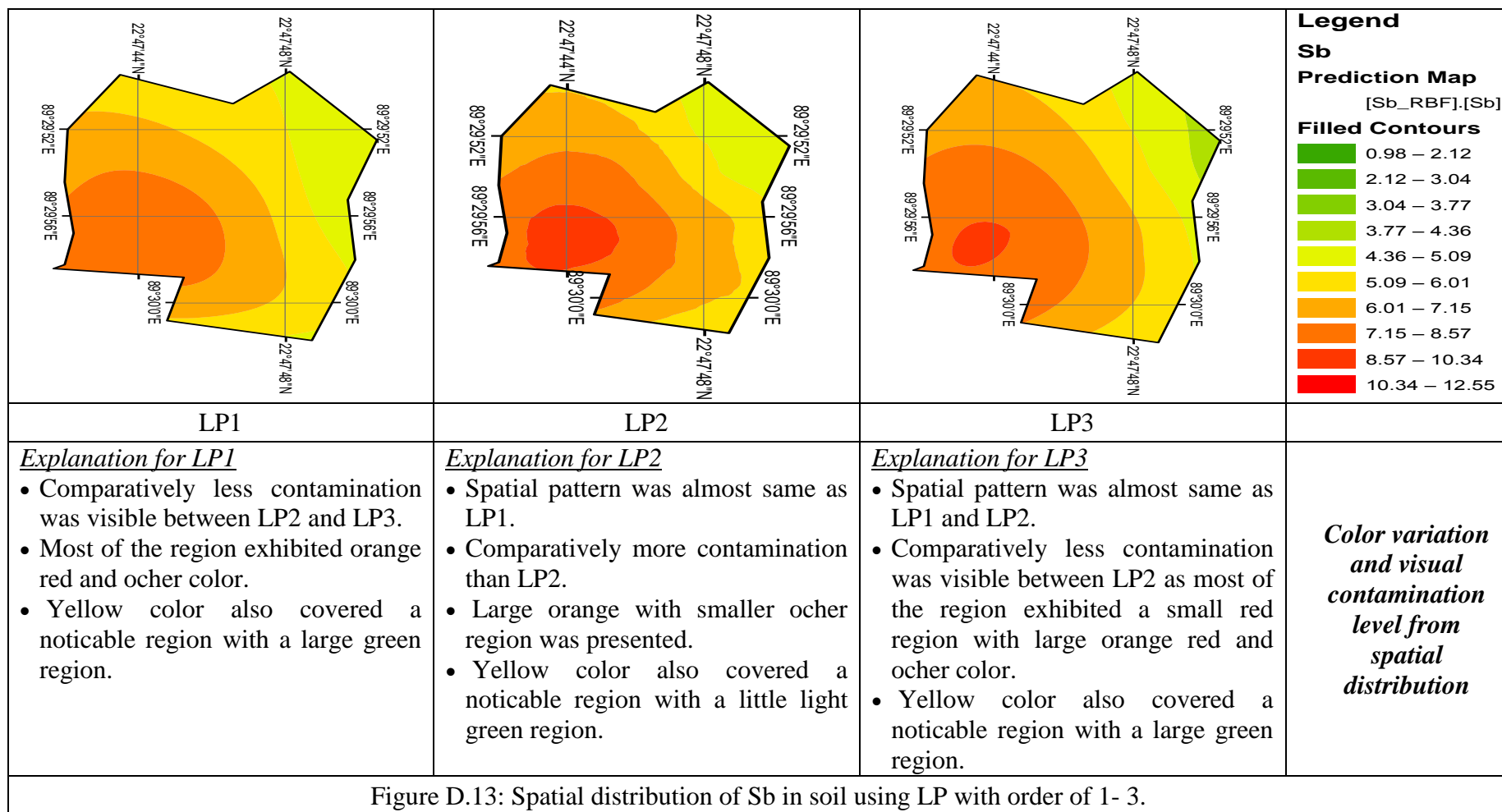
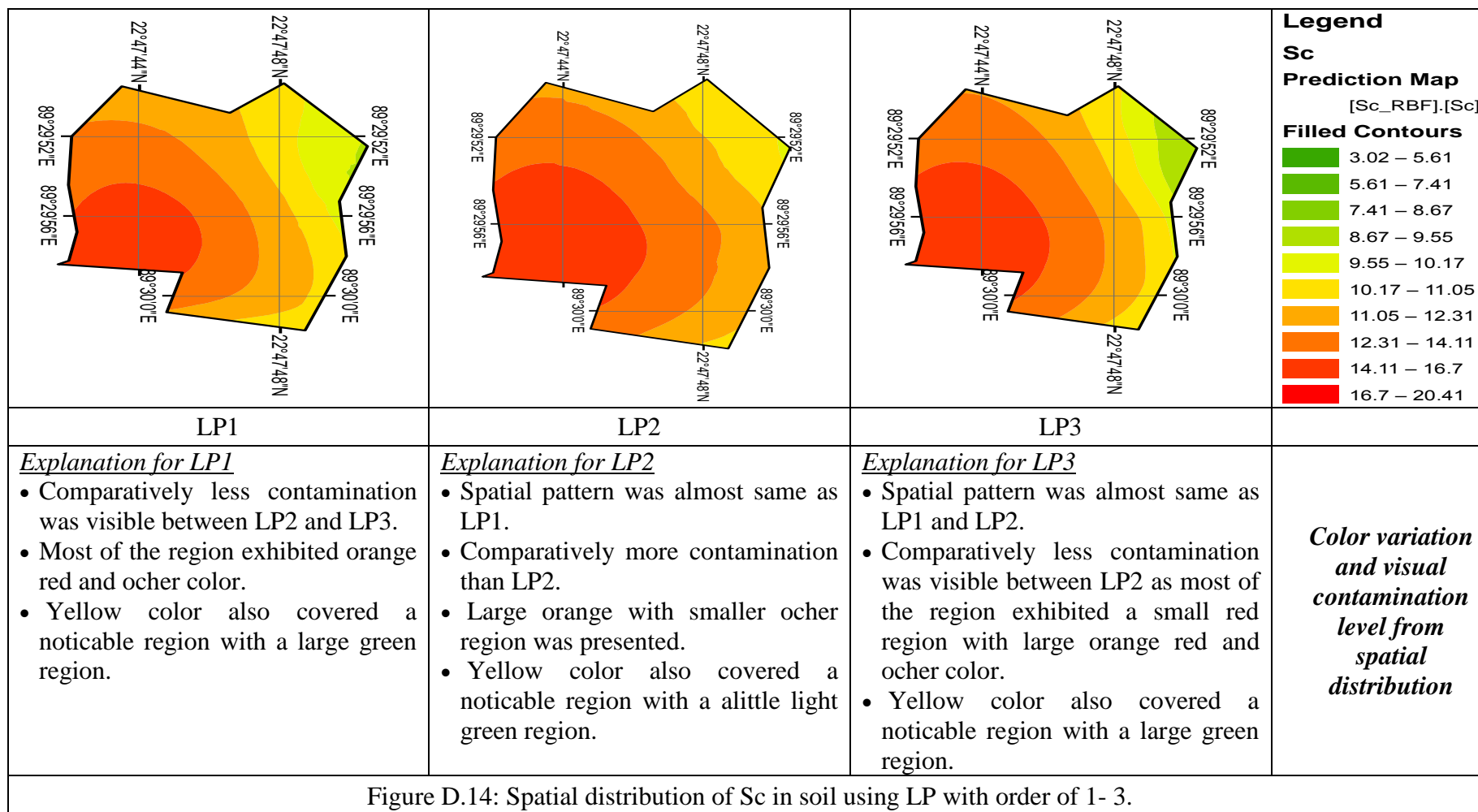
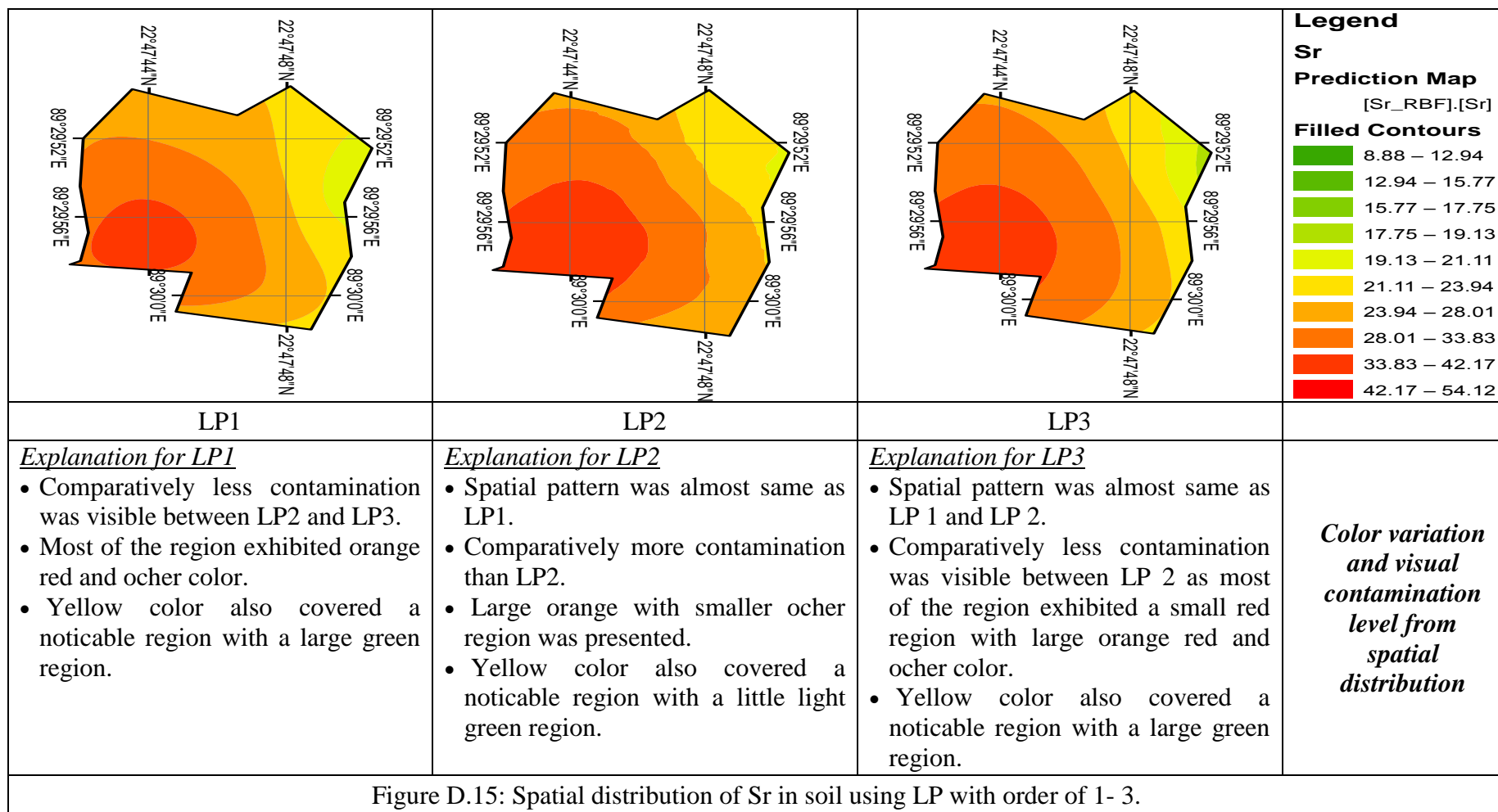
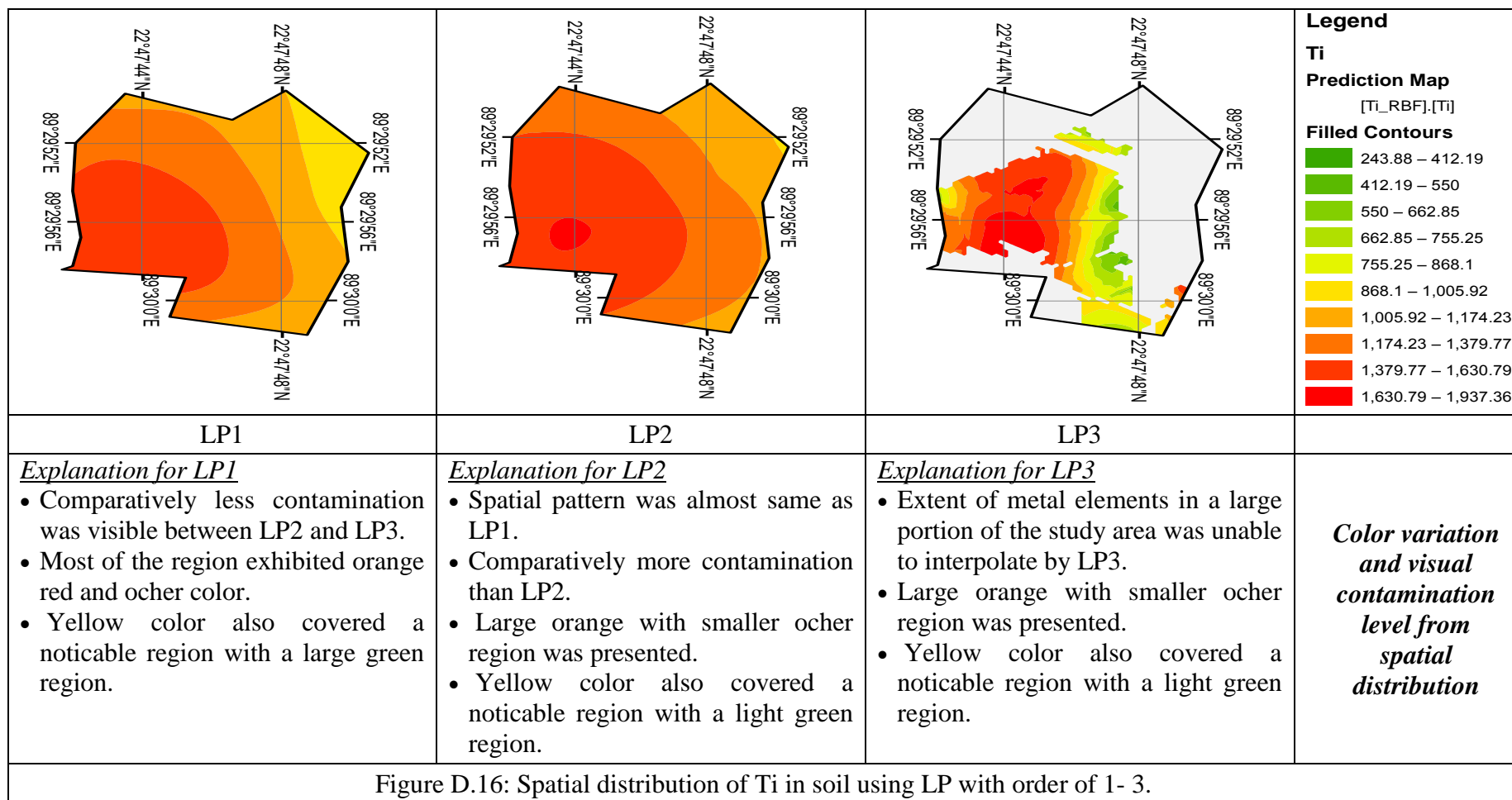
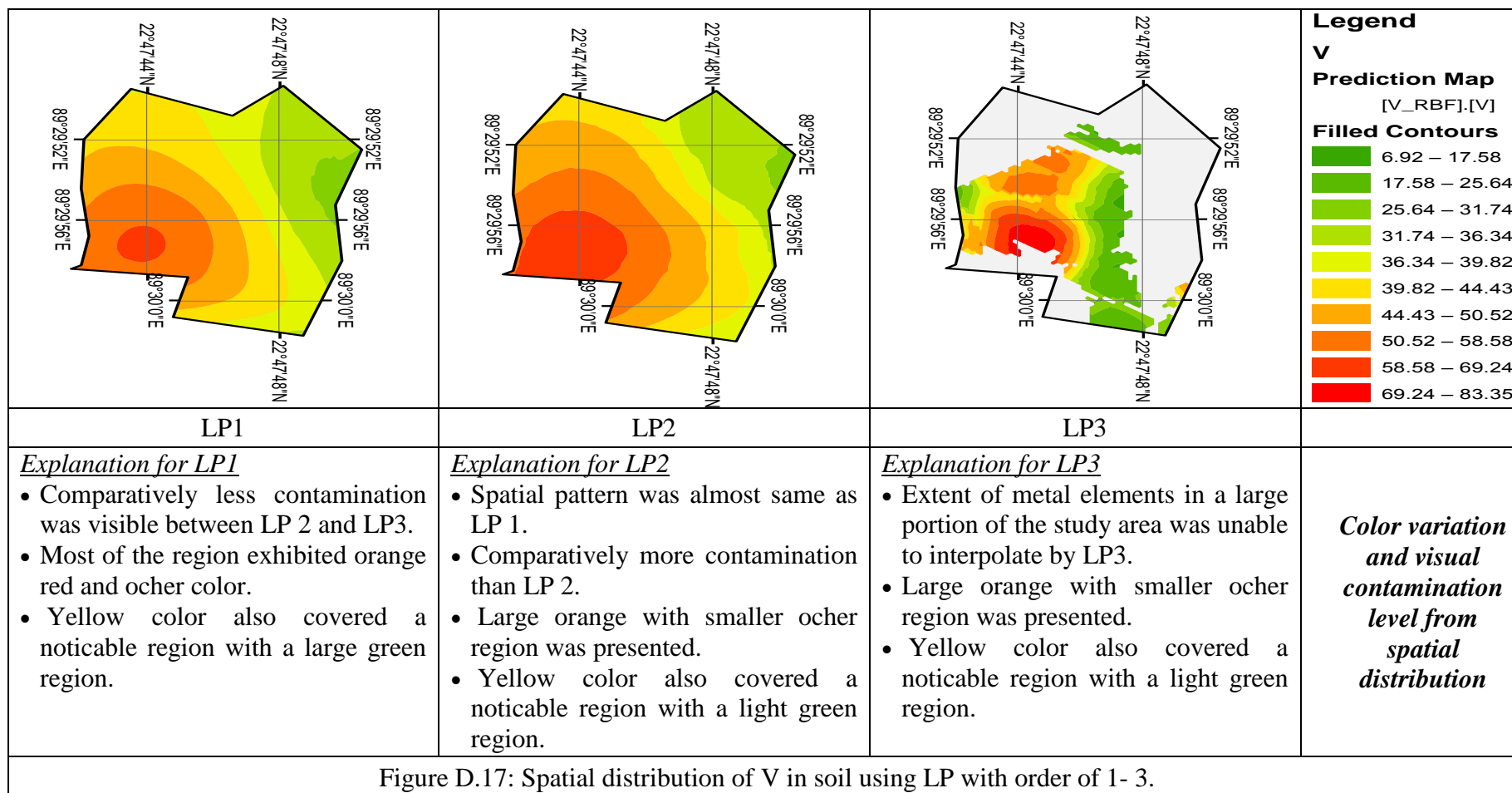


Figure D.13: Spatial distribution of Sb in soil using LP with order of 1- 3.









## Annex – E

### Cross Validation Results and Spatial Distribution of Metal Elements using RBF's

Table E.1: Cross validation of RBF's for Al

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	18.380	17.126	10.958	24.137	15.759
<sup>b</sup> RMSPE	118.158	119.072	126.136	117.522	147.141
Kernel Parameter	20297.85	16393.93	0	0.00017	1.00E+20

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error.

Table E.2: Cross validation of RBF for As

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	0.1627	0.1579	0.0860	0.2396	0.1155
<sup>b</sup> RMSPE	1.3397	1.3418	1.4295	1.3410	1.7307
Kernel Parameter	31702.47	21858.47	0	0.000137	1.00E+20

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error.

Table E.3: Cross validation of RBF for Ba

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	2.2042	2.0635	1.0405	3.0503	1.4402
<sup>b</sup> RMSPE	12.6961	12.7171	13.0694	12.8418	15.1331
Kernel Parameter	21771.43	14833.09	0	0.000177	1.00E+20

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error.

Table E.4: Cross validation of RBF for Ca

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	6.0454	5.7355	3.7103	8.2862	5.6899
<sup>b</sup> RMSPE	39.9360	40.0504	42.0299	39.8239	49.3384
Kernel Parameter	25261.25	18582.56	0	0.000151	1.00E+20

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error.

Table E.5: Cross validation of RBF for Co

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	0.1982	0.1903	0.0867	0.2867	0.1544
<sup>b</sup> RMSPE	1.1999	1.1974	1.2015	1.2438	1.4516
Kernel Parameter	25261.25	16393.93	0	0.000192	1.00E+20

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error.

Table E.6: Cross validation of RBF for Cr

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	0.036001	0.03529	0.070301	0.051787	0.176951
<sup>b</sup> RMSPE	1.998	1.9999	2.1436	1.9691	2.5723
Kernel Parameter	30970.43	21858.47	0	0.000142	1.00E+20

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error.

Table E.7: Cross validation of RBF for Cu

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	0.2648	0.2637	0.1527	0.4132	0.1691
<sup>b</sup> RMSPE	2.1957	2.1957	2.3459	2.2583	2.9209
Kernel Parameter	56077.92	42160.04	0	9.35E-05	1.00E+20

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error.

Table E.8: Cross validation of RBF for Fe

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	30.8837	29.3903	34.0156	38.2507	65.0609
<sup>b</sup> RMSPE	458.8827	459.9148	492.0969	451.4522	588.072
Kernel Parameter	29800.49	21858.47	0	0.000136	1.00E+20

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error.

Table E.9: Cross validation of RBF for Hg

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	0.1916	0.1860	0.1748	0.2588	0.2988
<sup>b</sup> RMSPE	1.5867	1.5893	1.7244	1.5819	2.1201
Kernel Parameter	35188.06	28268.34	0	1.34E-04	1.00E+20

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error.

Table E.10: Cross validation of RBF for K

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	8.5122	8.3415	6.3503	11.9169	9.9865
<sup>b</sup> RMSPE	76.8219	76.9904	83.1852	75.8925	96.6094
Kernel Parameter	33069.1	21858.47	0	0.000142	1.00E+20

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error.

Table E.11: Cross validation of RBF for Mn

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	0.1460	0.1015	0.1974	0.2045	0.6123
<sup>b</sup> RMSPE	5.4496	5.4696	5.8785	5.4164	7.4172
Kernel Parameter	25261.25	36183.49	0	0.000129	1.00E+20

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error.

Table E.12: Cross validation of RBF for Na

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	2.3547	2.3467	1.9191	3.3361	2.9474
<sup>b</sup> RMSPE	22.6961	22.7055	24.8567	22.2998	29.5410
Kernel Parameter	49963.41	36183.49	0	0.000117	1.00E+20

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error.

Table E.13: Cross validation of RBF for Sb

Model	CRS	ST	MO	IMO	TPS
<sup>a</sup> MPE	0.1809	0.1687	0.1003	0.2457	0.1617
<sup>b</sup> RMSPE	1.2749	1.2842	1.3587	1.2667	1.5655
Kernel Parameter	20297.85	16393.93	0	0.000174	1.00E+20

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error

Table E.14: Cross validation of RBF for Sc

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	0.2941	0.2820	0.1771	0.4112	0.2699
<sup>b</sup> RMSPE	2.1859	2.1905	2.2959	2.1771	2.6288
Kernel Parameter	25261.25	21858.47	0	0.000155	1.00E+20

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error.

Table E.15: Cross validation of RBF for Sr

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	0.6956	0.6852	0.4197	1.0424	0.6577
<sup>b</sup> RMSPE	5.4249	5.4180	5.6799	5.4907	6.7148
Kernel Parameter	42129.16	21858.47	0	0.000134	1.00E+20

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error.

Table E.16: Cross validation of RBF for Ti

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	37.1608	35.4811	20.9896	51.0683	30.705
<sup>b</sup> RMSPE	254.1218	254.8953	267.9789	253.0138	306.8545
Kernel Parameter	22775.35	18514.25	0	0.000169	1.00E+20

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error

Table E.17: Cross validation of RBF for V

Model	CRS	ST	MQ	IMQ	TPS
<sup>a</sup> MPE	1.2233	1.1507	0.8884	1.6596	1.4865
<sup>b</sup> RMSPE	10.3253	10.3339	10.9972	10.2142	13.0635
Kernel Parameter	25261.25	21858.47	0	0.000153	1.00E+20

<sup>a</sup>MPE=Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error



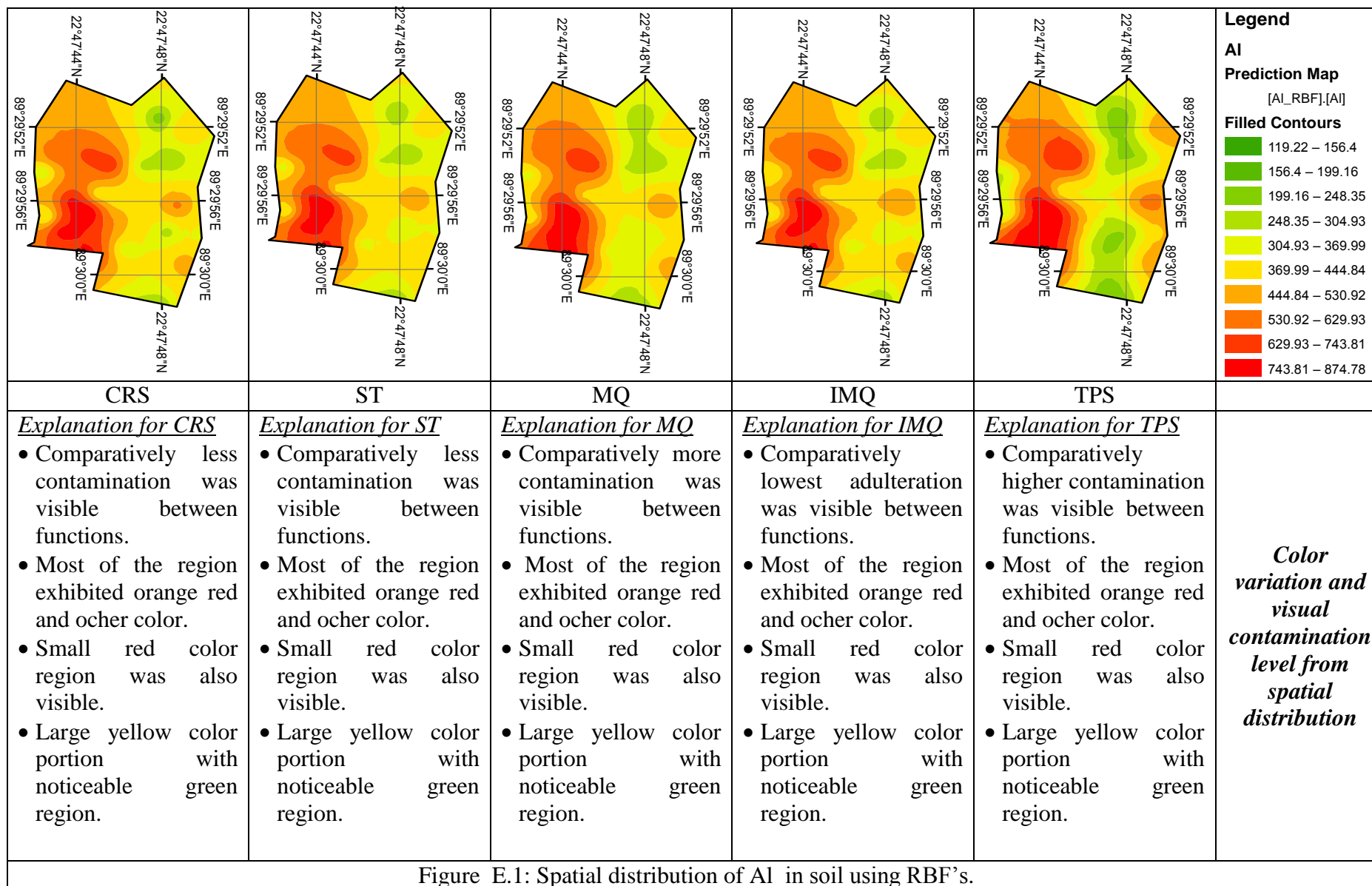


Figure E.1: Spatial distribution of AI in soil using RBF's.

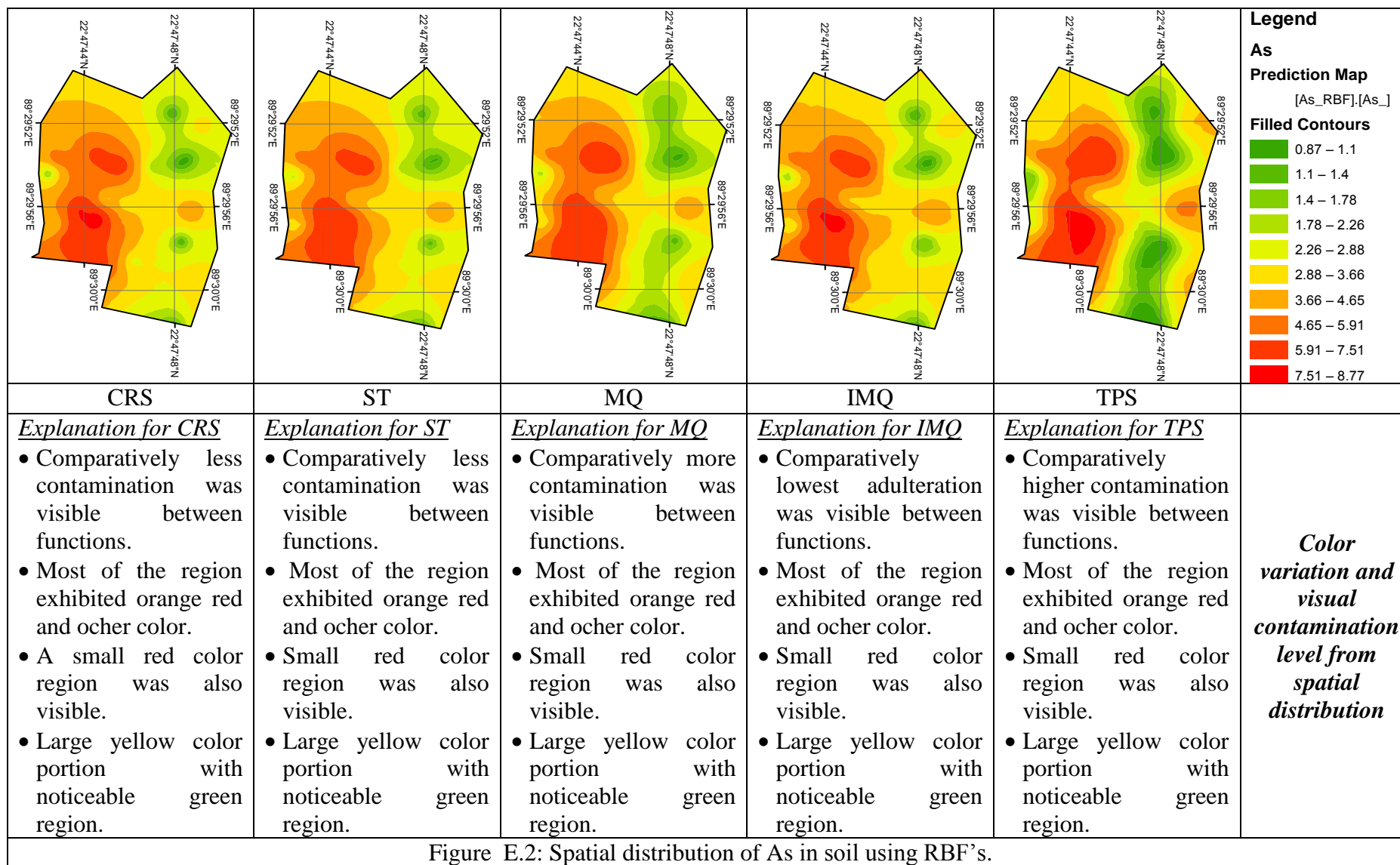


Figure E.2: Spatial distribution of As in soil using RBF's.

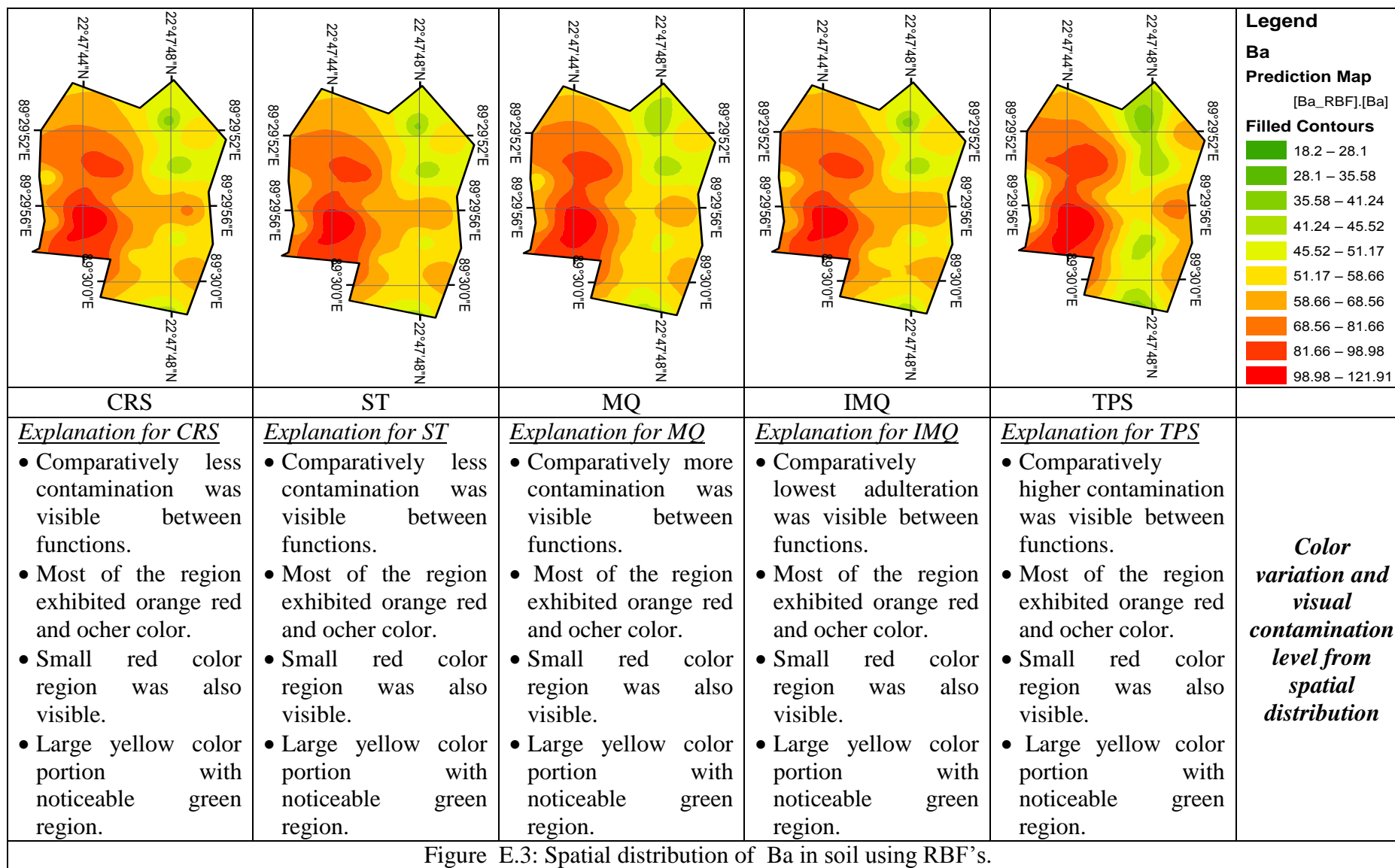


Figure E.3: Spatial distribution of Ba in soil using RBF's.

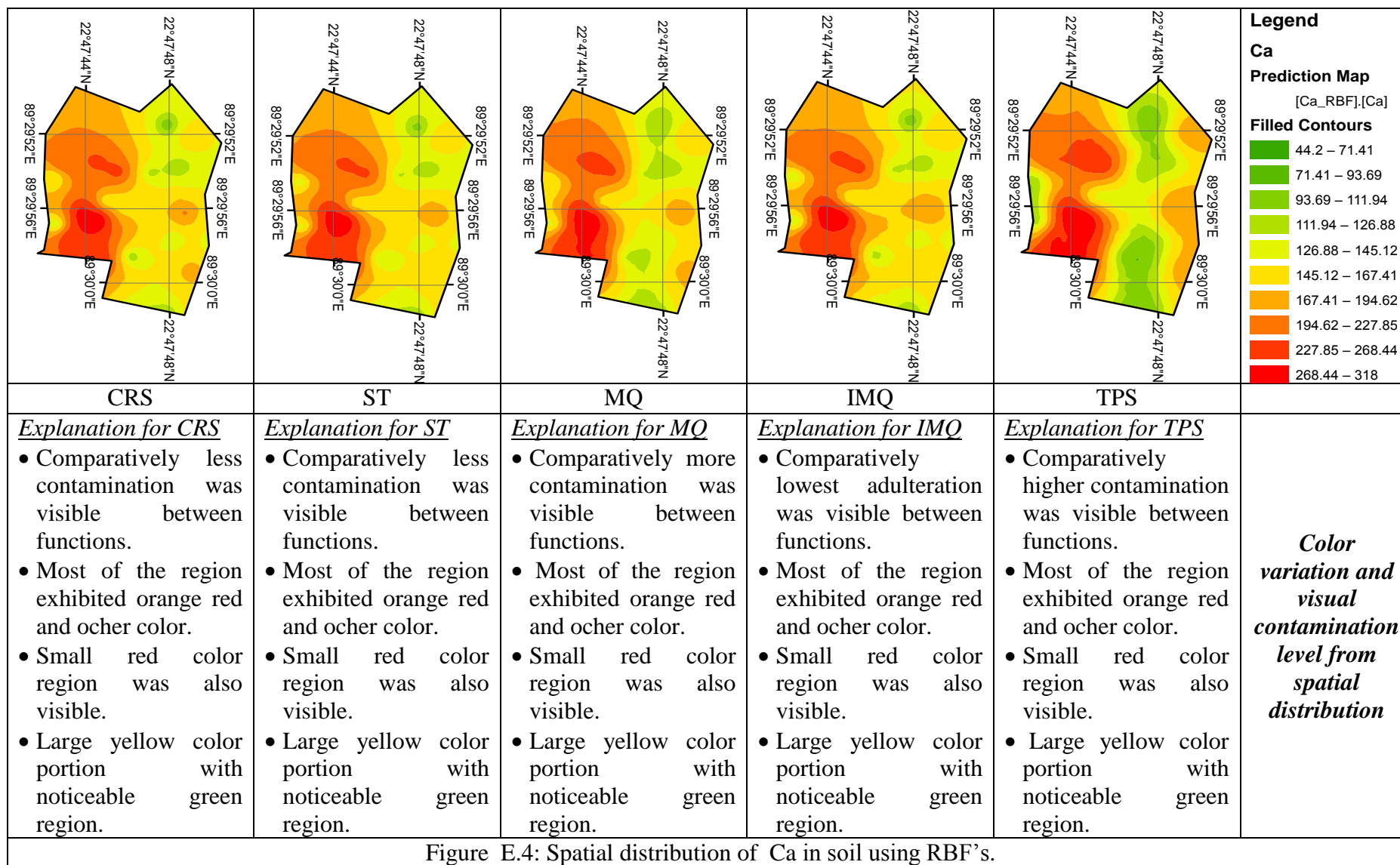


Figure E.4: Spatial distribution of Ca in soil using RBF's.

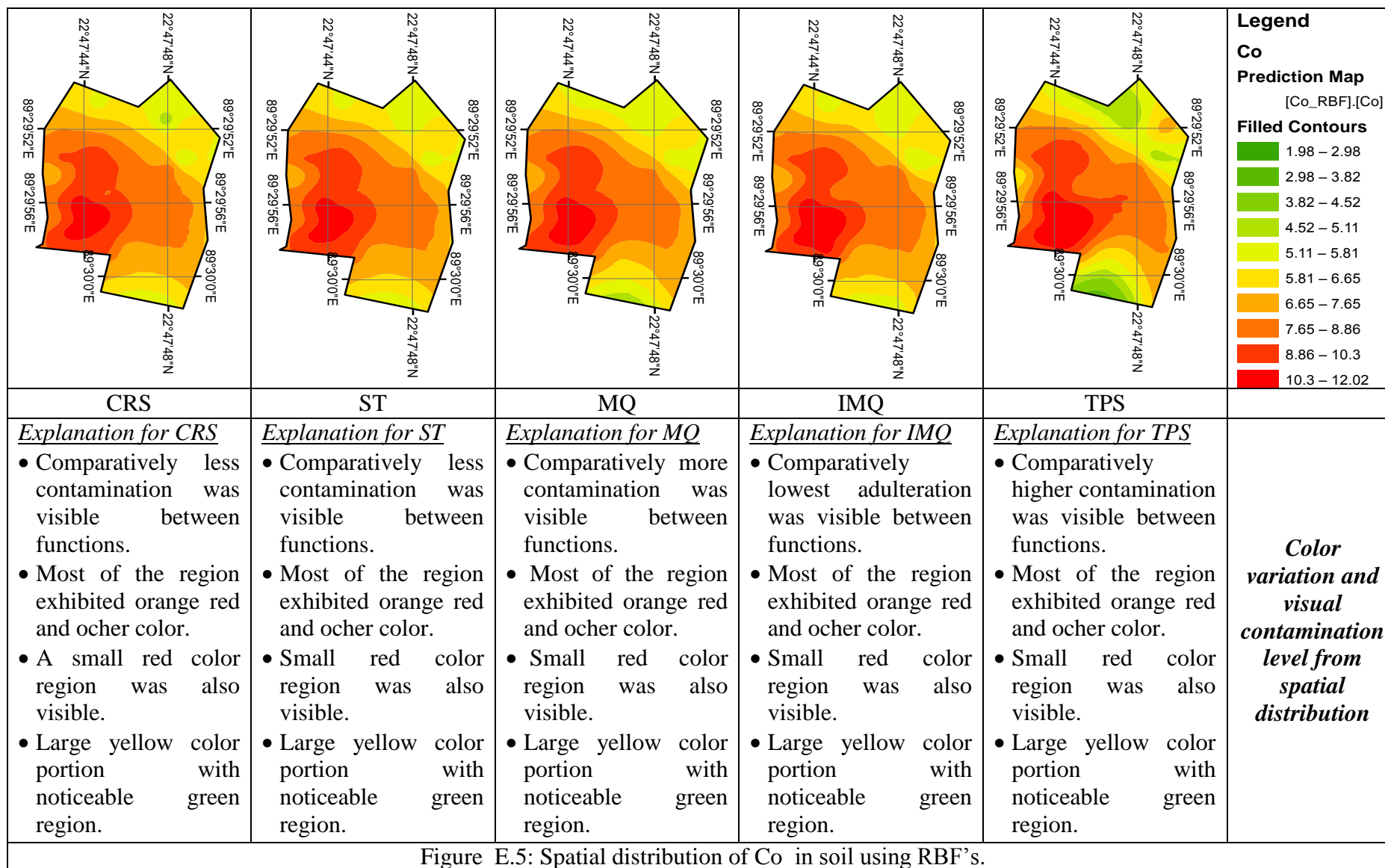


Figure E.5: Spatial distribution of Co in soil using RBF's.

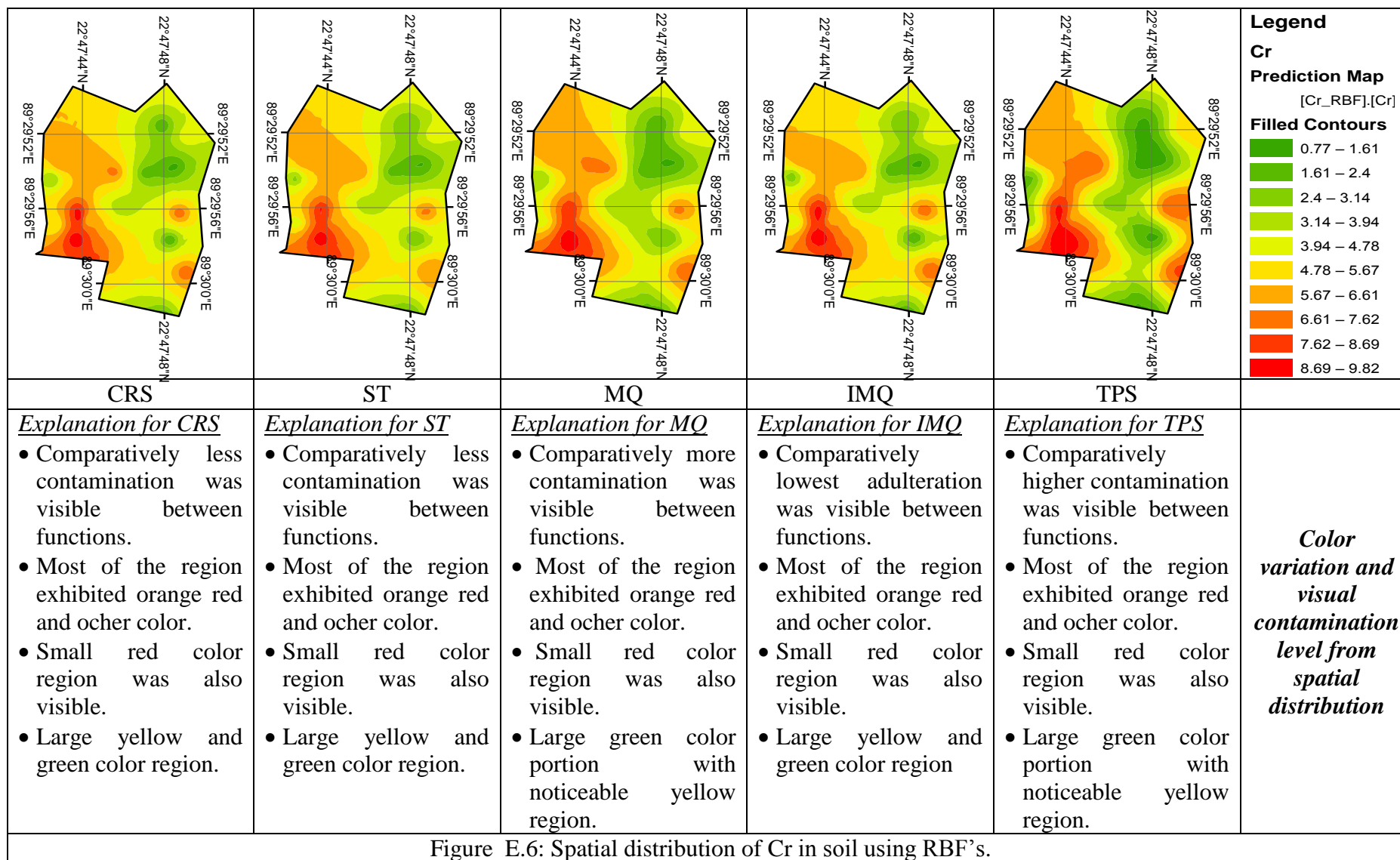


Figure E.6: Spatial distribution of Cr in soil using RBF's.

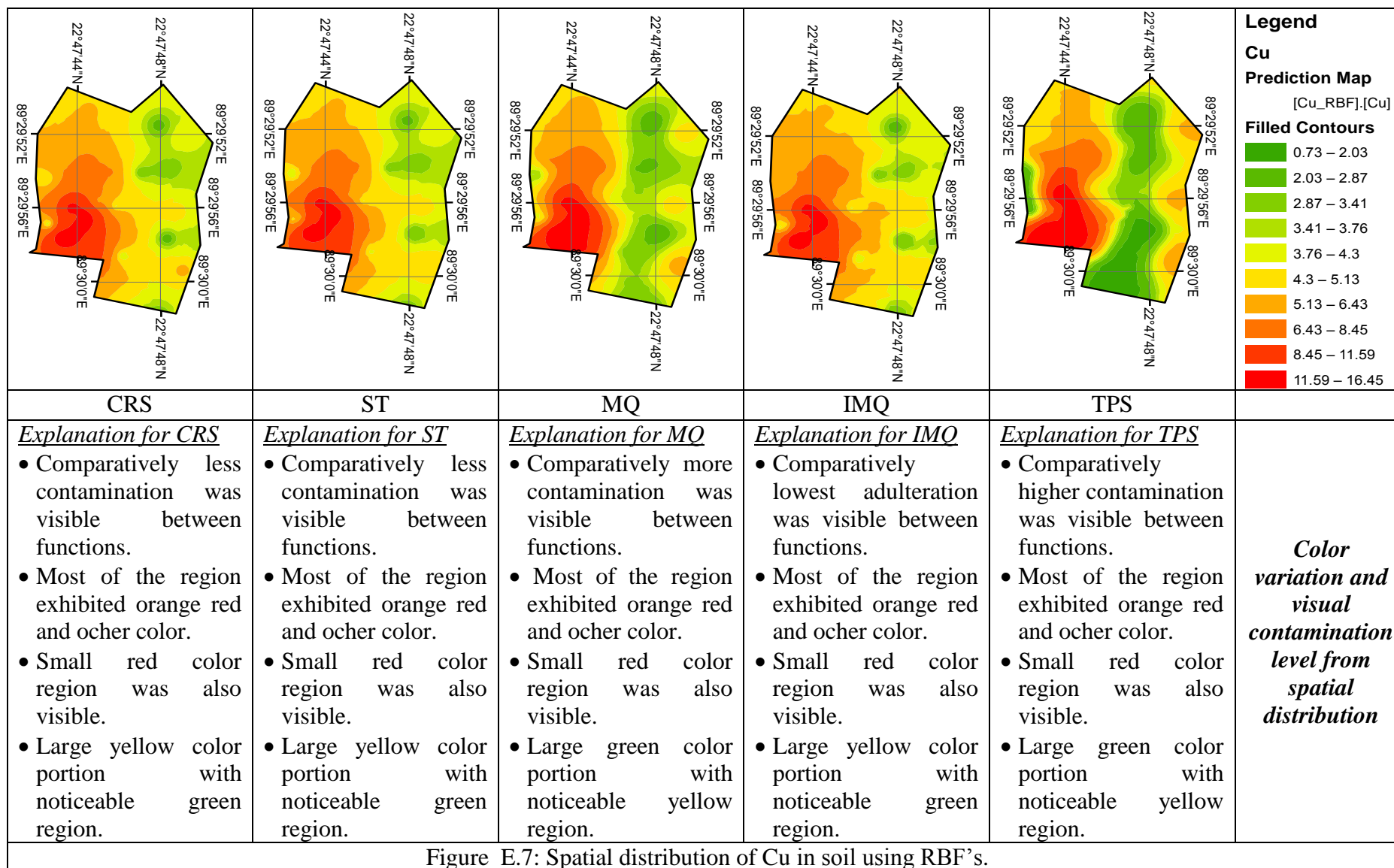


Figure E.7: Spatial distribution of Cu in soil using RBF<sup>2</sup>s.

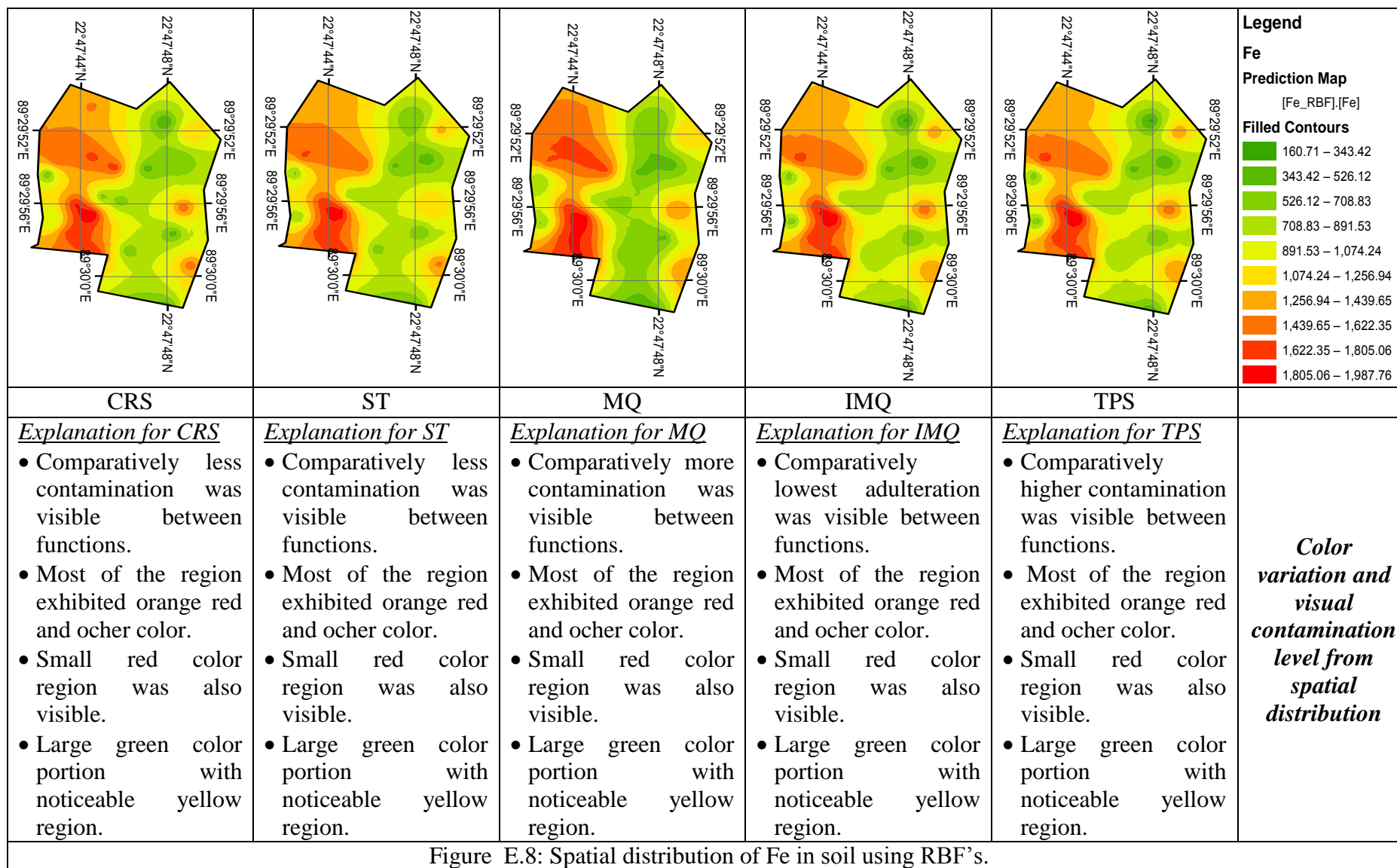


Figure E.8: Spatial distribution of Fe in soil using RBF's.



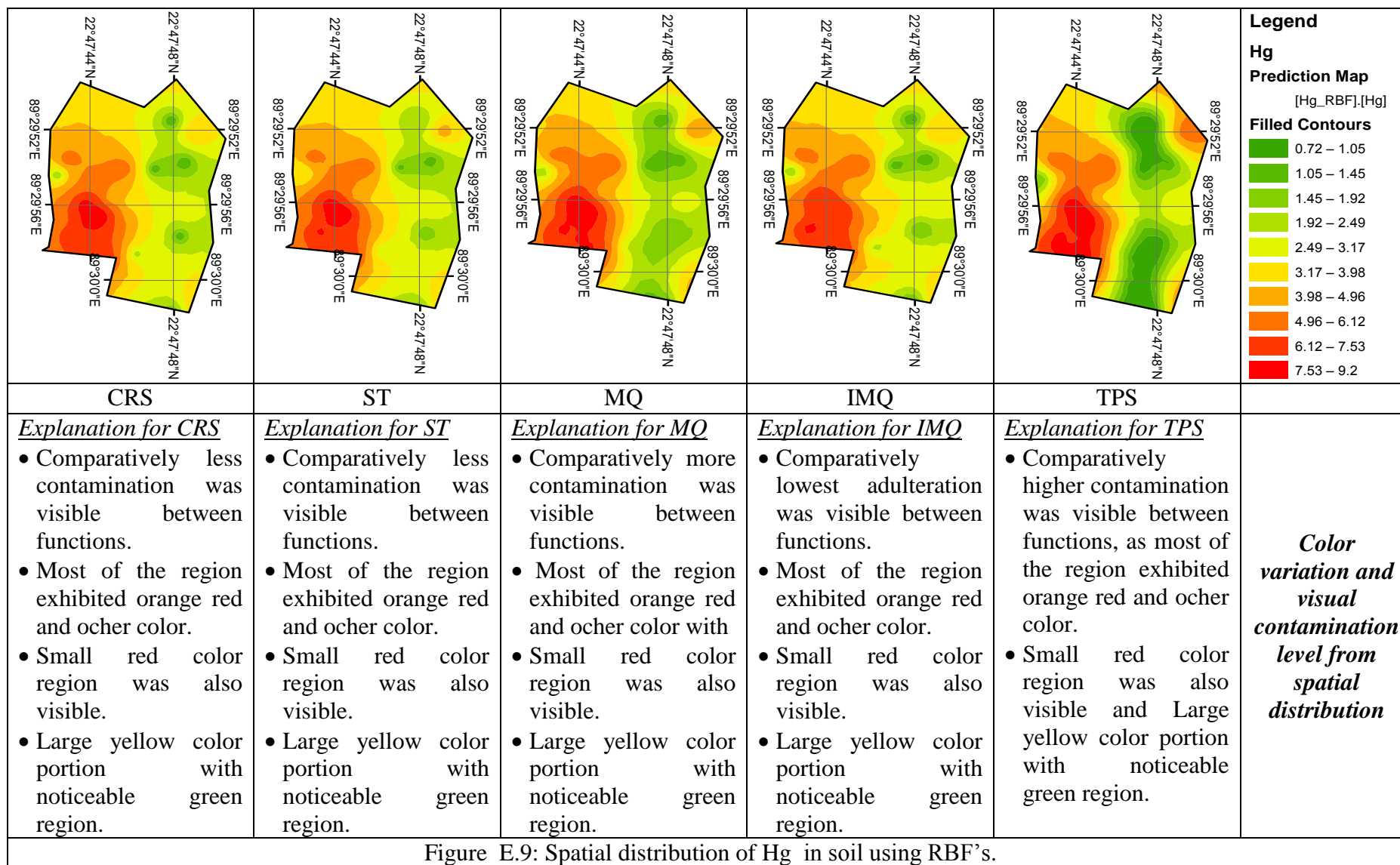


Figure E.9: Spatial distribution of Hg in soil using RBF's.

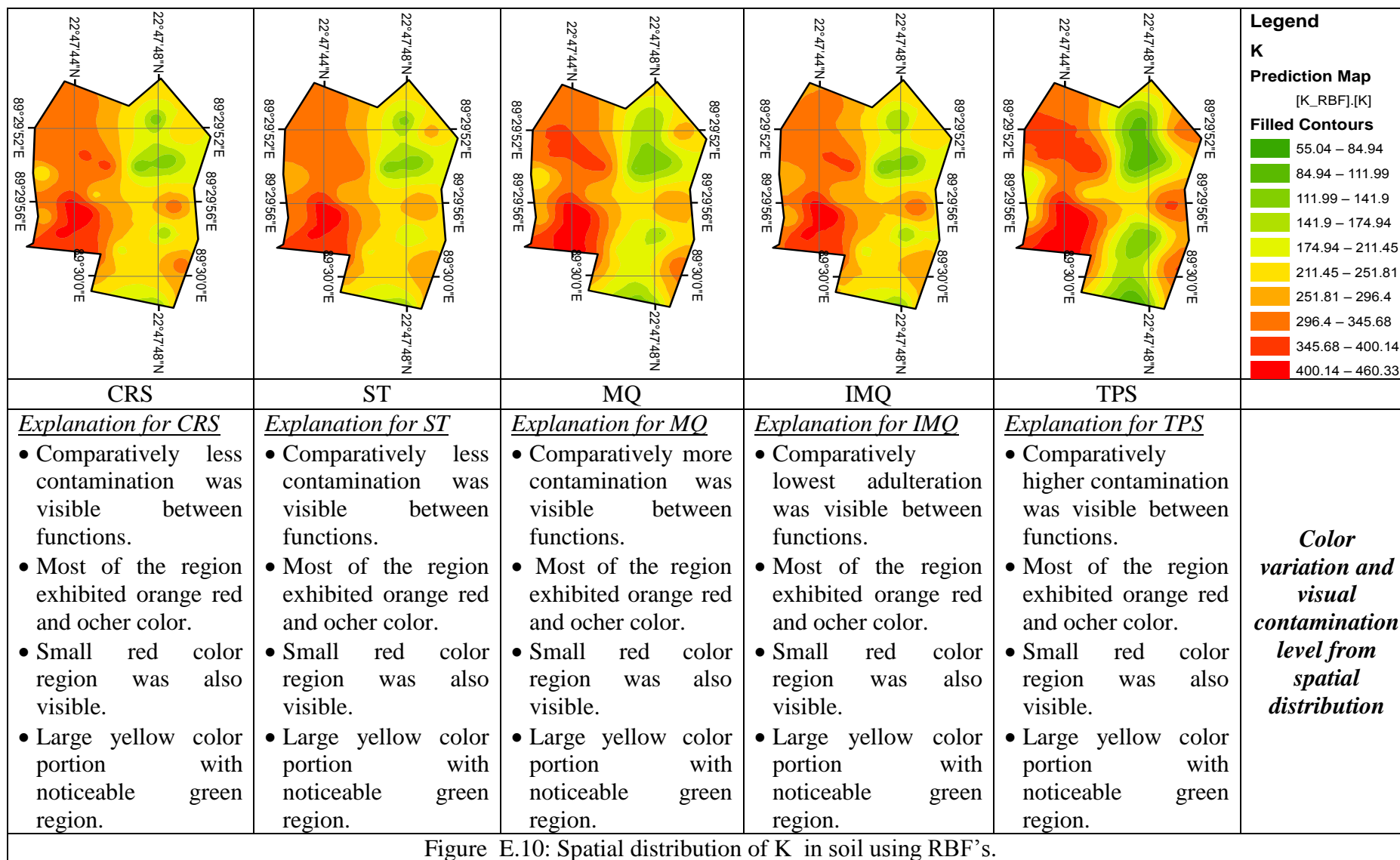


Figure E.10: Spatial distribution of K in soil using RBF's.

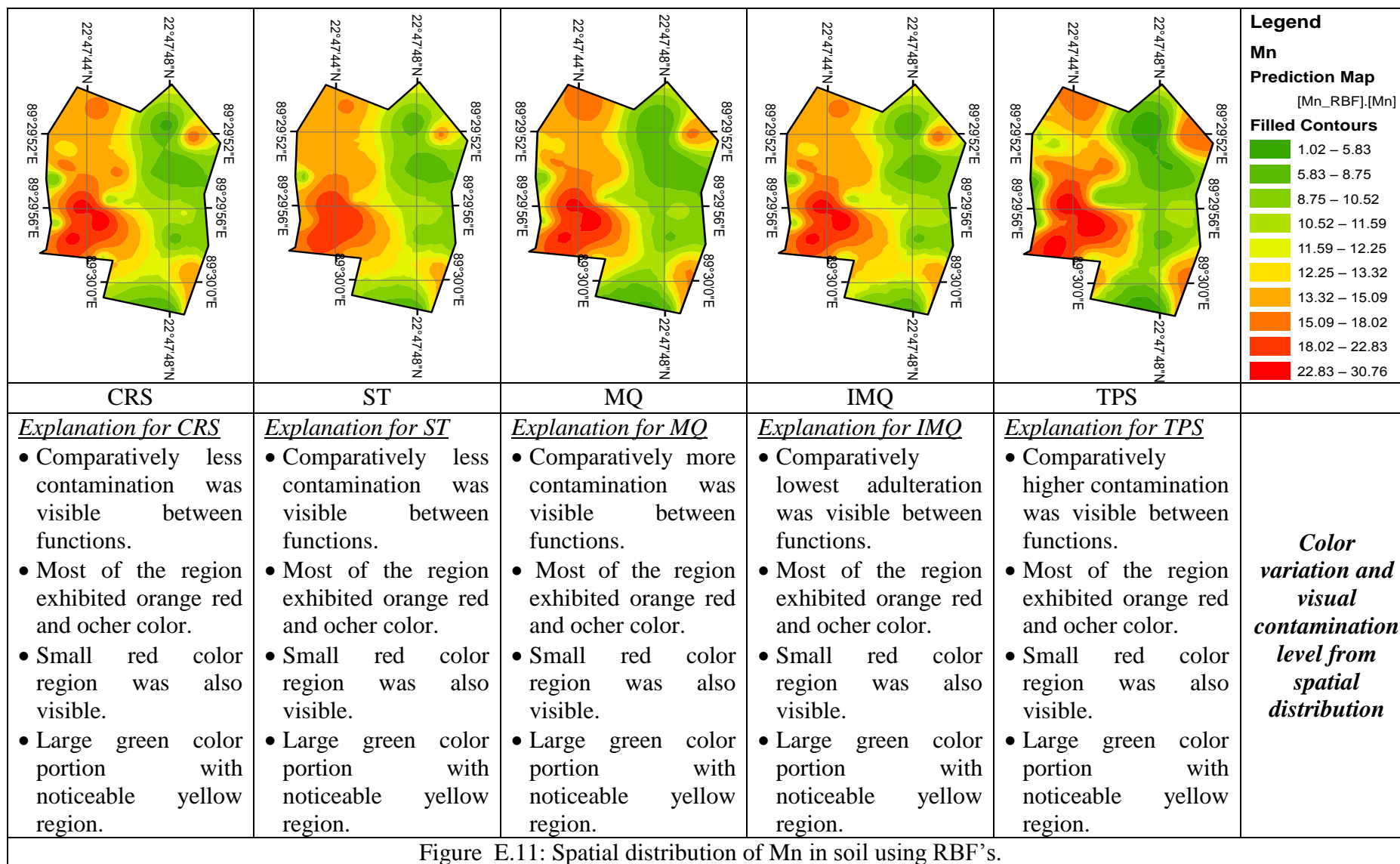


Figure E.11: Spatial distribution of Mn in soil using RBF's.

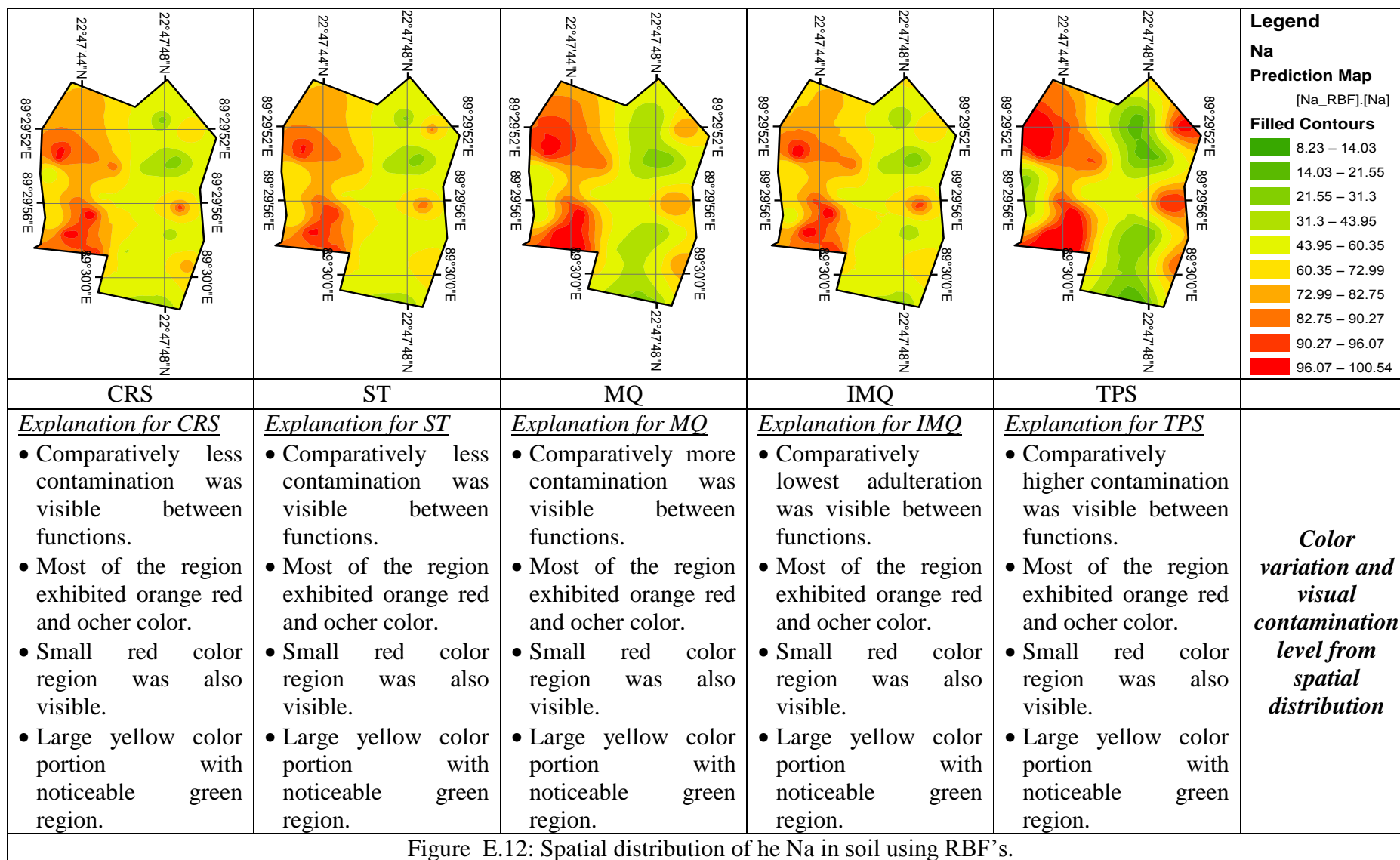


Figure E.12: Spatial distribution of he Na in soil using RBF's.

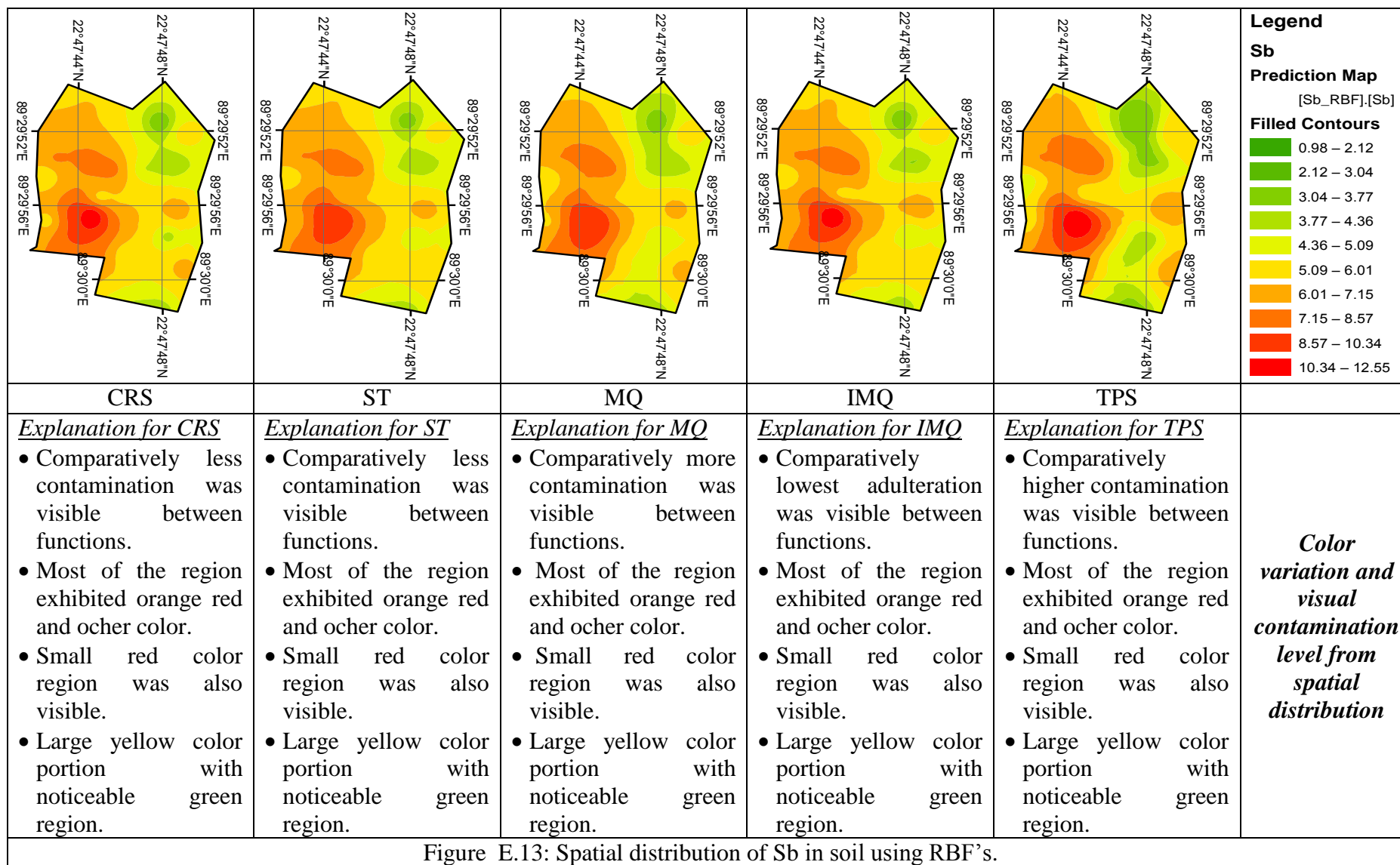


Figure E.13: Spatial distribution of Sb in soil using RBF's.

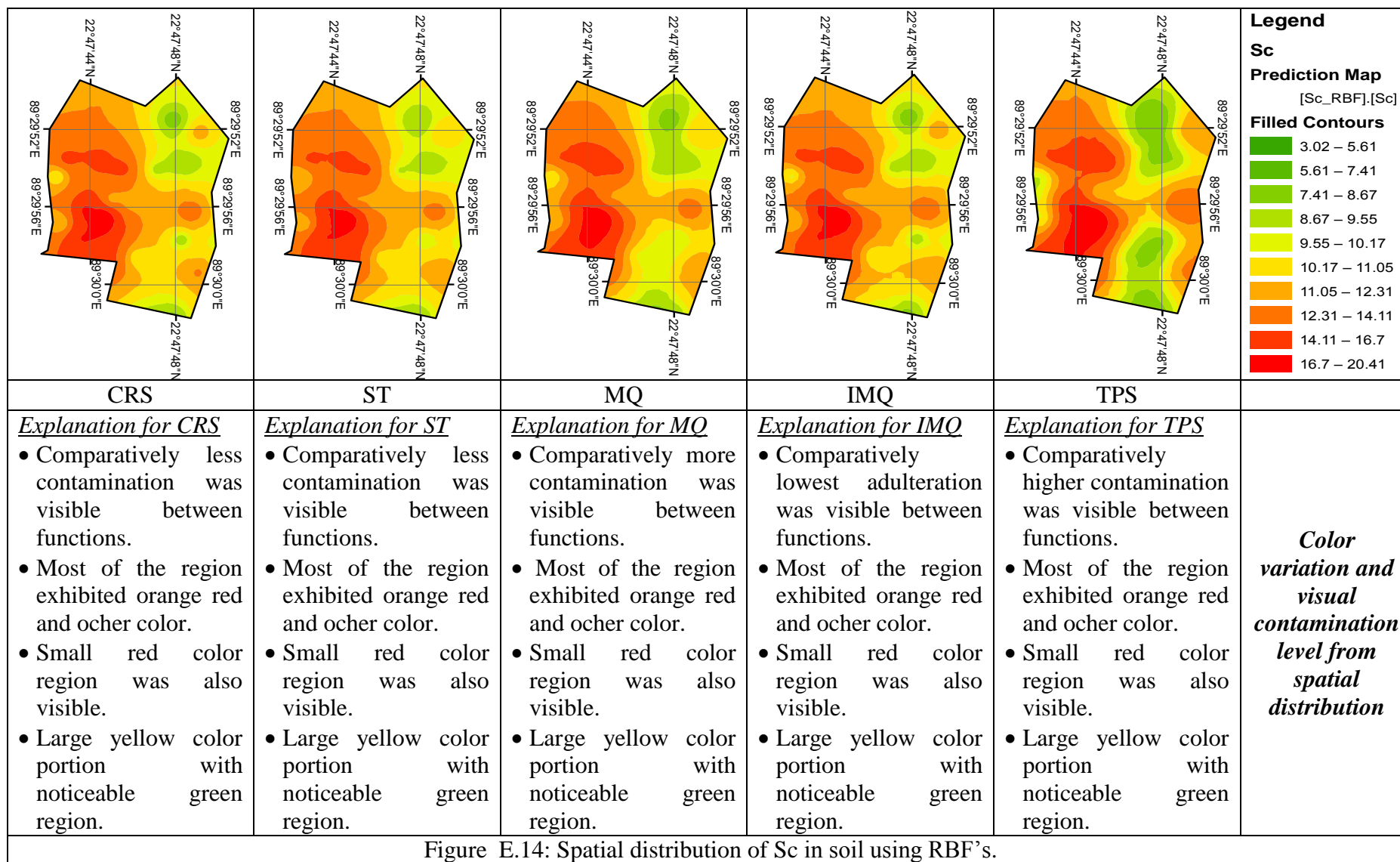


Figure E.14: Spatial distribution of Sc in soil using RBF's.

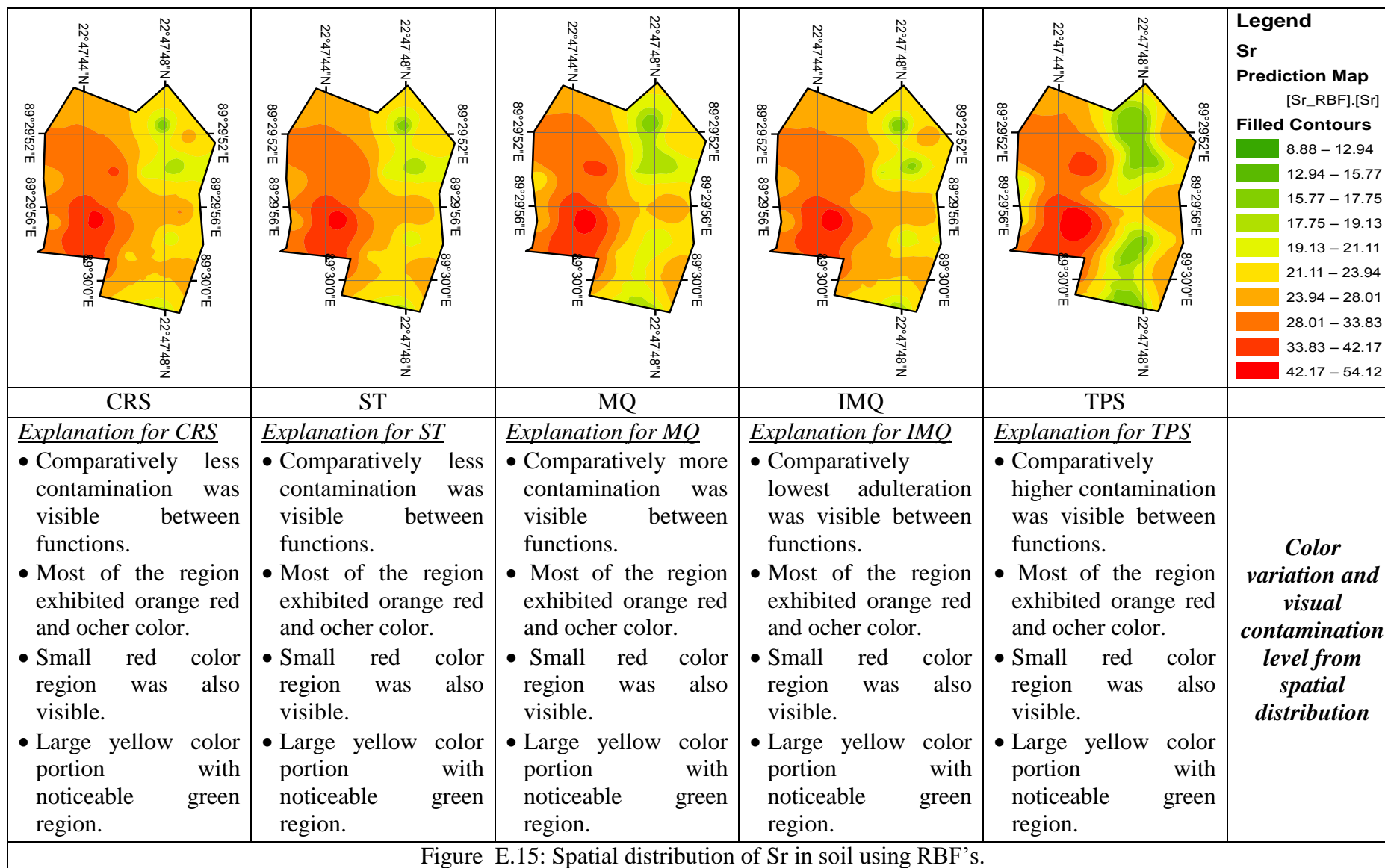


Figure E.15: Spatial distribution of Sr in soil using RBF's.

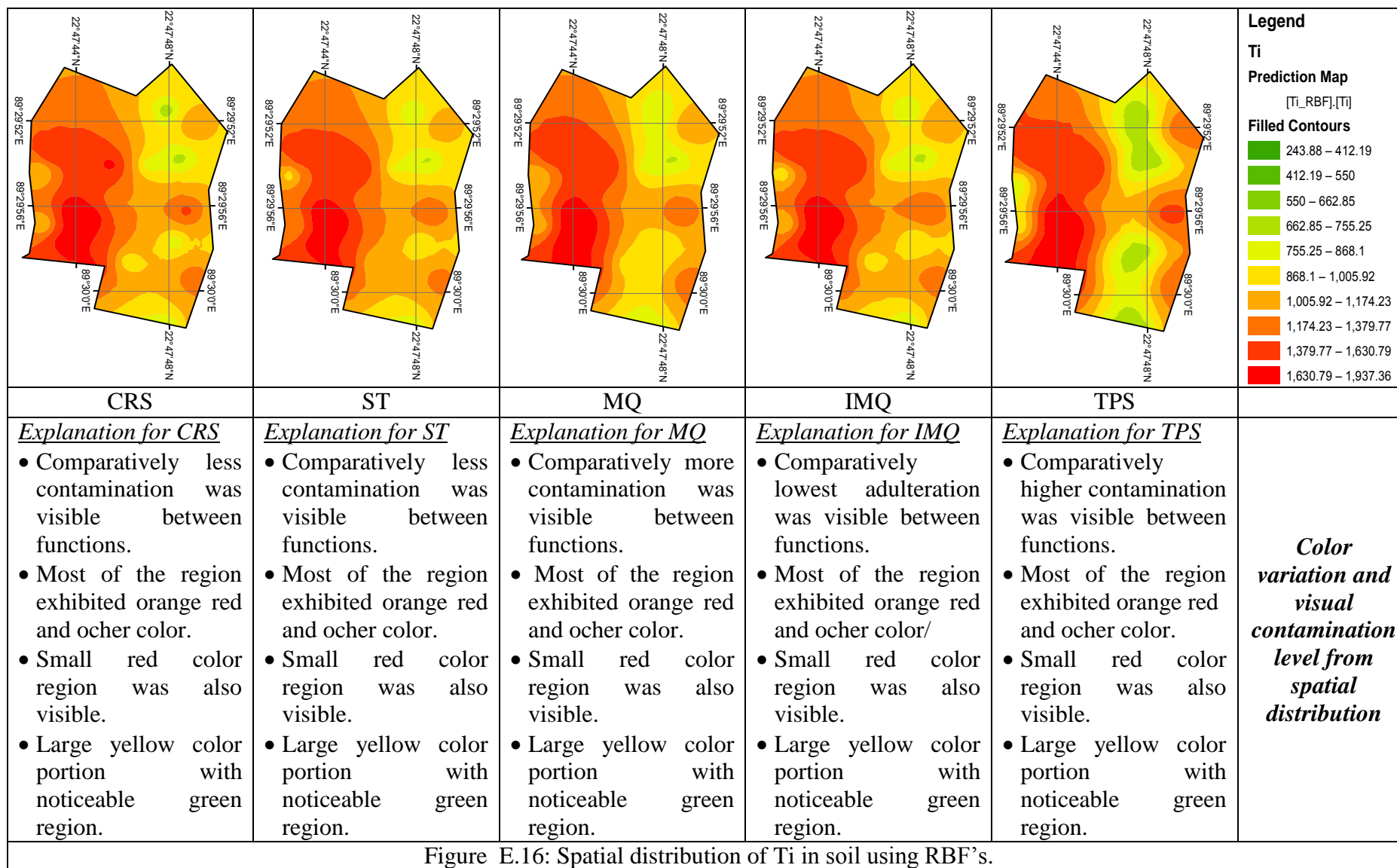


Figure E.16: Spatial distribution of Ti in soil using RBF's.



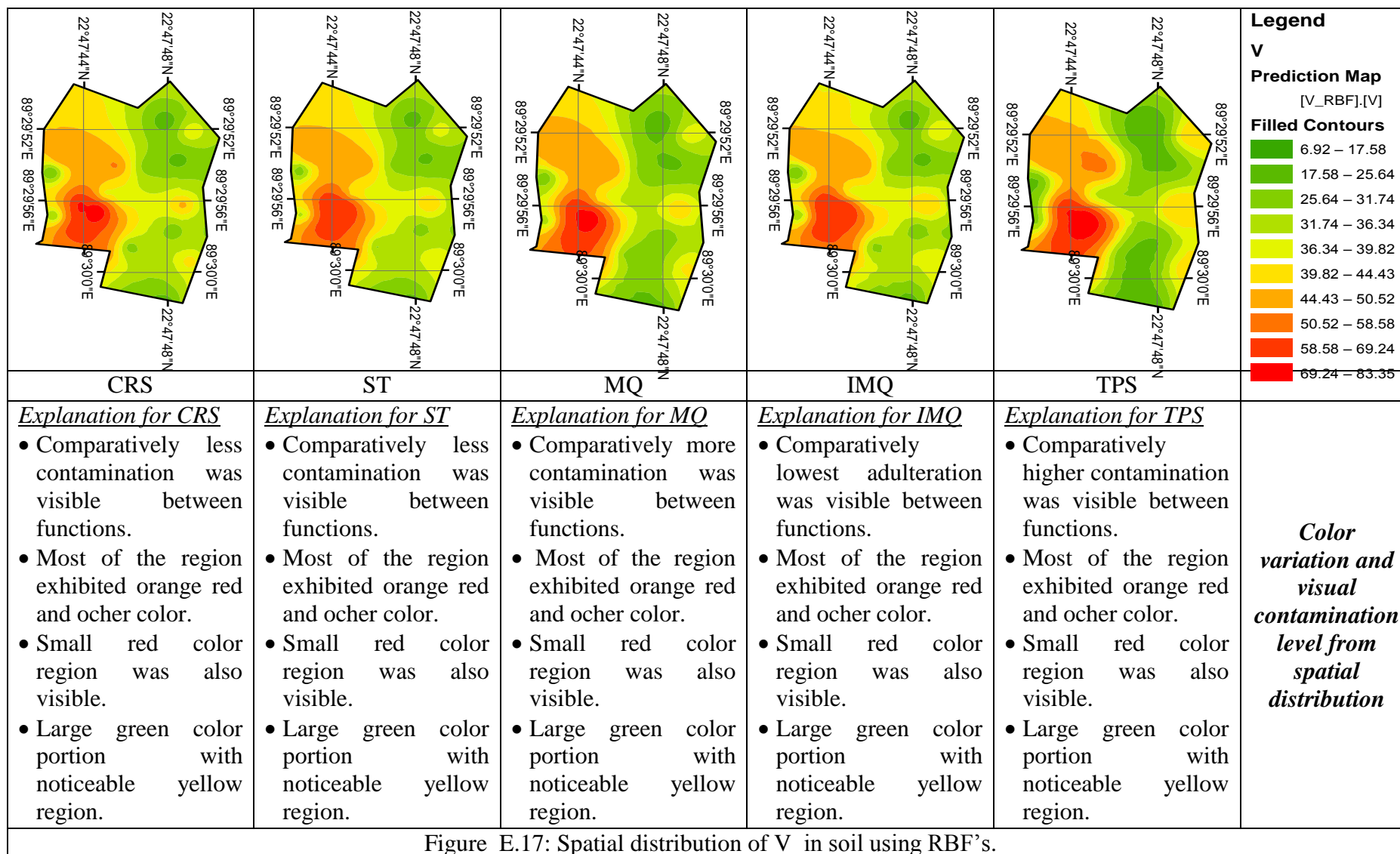


Figure E.17: Spatial distribution of V in soil using RBF's.

**Annex – F**

**Cross Validation Data and Spatial Distribution of Metal Elements using OK**

Table F.1: Cross validation of OK for Al

<b>Models</b>	<b><sup>a</sup>MPE</b>	<b><sup>b</sup>RMSPE</b>	<b><sup>c</sup>ASPE</b>	<b><sup>d</sup>MSPE</b>	<b><sup>e</sup>RMSSPE</b>
Circular	7.9811	131.1189	170.1680	0.0298	0.8138
Spherical	8.3799	129.1836	169.5963	0.0310	0.8031
Tetraspherical	8.2673	127.7511	169.7848	0.0308	0.7930
Pentaspherical	9.0704	127.7458	169.1689	0.0340	0.7937
Exponential	13.8886	118.8524	166.0145	0.0539	0.7440
Gaussian	7.0387	135.3897	172.3418	0.0273	0.8328
Rational quadratic	18.1793	116.1146	166.6314	0.0762	0.7342
Hole Effect	10.8442	145.5775	170.9768	0.0528	0.8876
K-Bessel	14.0068	119.4834	165.4350	0.0547	0.7502
J-Bessel	8.2231	137.1476	170.6431	0.0345	0.8455
Stable	13.9179	119.6027	164.9862	0.0544	0.7525

Table F.2: Cross validation of OK for As

<b>Models</b>	<b><sup>a</sup>MPE</b>	<b><sup>b</sup>RMSPE</b>	<b><sup>c</sup>ASPE</b>	<b><sup>d</sup>MSPE</b>	<b><sup>e</sup>RMSSPE</b>
Circular	0.0634	1.4289	1.8967	0.0218	0.8012
Spherical	0.0603	1.4190	1.8919	0.0201	0.7983
Tetraspherical	0.0686	1.4178	1.8882	0.0235	0.7976
Pentaspherical	0.0836	1.4187	1.8816	0.0307	0.7995
Exponential	0.0925	1.3498	1.8387	0.0308	0.7847
Gaussian	0.0496	1.4739	1.9091	0.0172	0.8220
Rational	0.1067	1.3443	1.8399	0.0364	0.7981
Hole Effect	0.0841	1.5398	1.9331	0.0378	0.8265
K-Bessel	0.1068	1.3456	1.8353	0.0361	0.7792
J-Bessel	0.0702	1.4997	1.8930	0.0270	0.8366
Stable	0.1198	1.3435	1.8251	0.0420	0.7827

<sup>a</sup>Mean Percentage Error, <sup>b</sup>Root Mean Square Percentage Error, <sup>c</sup>Average Standard Percentage Error, <sup>d</sup>Mean Standardized Percentage Error, <sup>e</sup>Root Mean Square Standard Percentage Error.

Table F.3: Cross validation of OK for Ba

<b>Models</b>	<b><sup>a</sup>MPE</b>	<b><sup>b</sup>RMSPE</b>	<b><sup>c</sup>ASPE</b>	<b><sup>d</sup>MSPE</b>	<b><sup>e</sup>RMSSPE</b>
Circular	0.7639	13.0556	17.7070	0.0266	0.7866
Spherical	1.0014	13.2106	17.5297	0.0366	0.8056
Tetraspherical	0.9895	13.1541	17.3851	0.0357	0.8127
Pentaspherical	0.9871	13.1994	17.2951	0.0351	0.8200
Exponential	1.3768	13.0460	16.7660	0.0508	0.8494
Gaussian	0.5709	14.9903	18.1179	0.0218	0.8812
Rational quadratic	0.8295	14.1263	17.2743	0.0275	0.9025
Hole Effect	0.5545	14.6336	17.5492	0.0257	0.8823
K-Bessel	0.5670	14.9074	18.3094	0.0203	0.8646
J-Bessel	0.5682	14.5982	17.2958	0.0279	0.8887
Stable	0.5709	14.9903	18.1179	0.0218	0.8812

Table F.4: Cross validation of OK for Ca

<b>Models</b>	<b><sup>a</sup>MPE</b>	<b><sup>b</sup>RMSPE</b>	<b><sup>c</sup>ASPE</b>	<b><sup>d</sup>MSPE</b>	<b><sup>e</sup>RMSSPE</b>
Circular	3.7648	41.0653	55.0245	0.0347	0.7653
Spherical	3.1377	40.9446	54.5189	0.0264	0.7542
Tetraspherical	3.2083	41.0547	54.6894	0.0272	0.7542
Pentaspherical	3.2002	41.0241	54.8289	0.0273	0.7506
Exponential	3.0570	41.0198	56.4323	0.0283	0.7324
Gaussian	6.0369	40.9597	54.0746	0.0683	0.9453
Rational quadratic	3.4196	40.8019	55.2144	0.0314	0.7425
Hole Effect	6.5894	48.8409	48.0683	0.0529	1.5581
K-Bessel	5.7057	40.7235	54.0170	0.0625	0.8854
J-Bessel	5.8123	43.2724	51.4492	0.0538	1.0899
Stable	5.8001	40.9089	53.9264	0.0650	0.8999

<sup>a</sup>Mean Percentage Error, <sup>b</sup>Root Mean Square Percentage Error, <sup>c</sup>Average Standard Percentage Error, <sup>d</sup>Mean Standardized Percentage Error, <sup>e</sup>Root Mean Square Standard Percentage Error.

Table F.5: Cross validation of OK for Co

<b>Models</b>	<b><sup>a</sup>MPE</b>	<b><sup>b</sup>RMSPE</b>	<b><sup>c</sup>ASPE</b>	<b><sup>d</sup>MSPE</b>	<b><sup>e</sup>RMSSPE</b>
Circular	0.1180	1.2117	1.3446	0.0481	1.0523
Spherical	0.1217	1.2202	1.3511	0.0503	1.0607
Tetraspherical	0.0999	1.2066	1.0727	0.0409	1.0727
Pentaspherical	0.1048	1.2117	1.3764	0.0422	1.0616
Exponential	0.1174	1.1985	1.6388	0.0397	0.8508
Gaussian	0.0551	1.3343	1.2166	0.0237	1.1859
Rational quadratic	0.1263	1.2228	1.4419	0.0461	0.9982
Hole Effect	-0.0022	1.3244	1.1496	0.0013	1.2298
K-Bessel	0.0469	1.3035	1.2202	0.0197	1.1677
J-Bessel	-0.0317	1.3411	1.1090	-0.0085	1.2689
Stable	0.0554	1.3331	1.2169	0.0238	1.1846

Table F.6: Cross validation of OK for Cr

<b>Models</b>	<b><sup>a</sup>MPE</b>	<b><sup>b</sup>RMSPE</b>	<b><sup>c</sup>ASPE</b>	<b><sup>d</sup>MSPE</b>	<b><sup>e</sup>RMSSPE</b>
Circular	-0.0039	2.0813	2.2178	-0.0046	0.9649
Spherical	0.0003	2.0645	2.2123	-0.0031	0.9591
Tetraspherical	0.0091	2.0602	2.2091	0.0009	0.9577
Pentaspherical	0.0045	2.0434	2.2049	-0.0017	0.9517
Exponential	0.0227	2.0161	2.1966	0.0006	0.9345
Gaussian	-0.0045	2.0800	2.2363	-0.0069	0.9569
Rational quadratic	0.0423	1.9899	2.1944	0.0058	0.9355
Hole Effect	-0.0475	2.0300	2.2114	-0.0190	0.9475
K-Bessel	0.0164	2.0155	2.1949	-0.0020	0.9356
J-Bessel	-0.0001	2.0554	2.2199	-0.0023	0.9542
Stable	0.0270	2.0154	2.1945	0.0033	0.9347

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table F.7: Cross validation of OK for Cu

<b>Models</b>	<sup>a</sup> <b>MPE</b>	<sup>b</sup> <b>RMSPE</b>	<sup>c</sup> <b>ASPE</b>	<sup>d</sup> <b>MSPE</b>	<sup>e</sup> <b>RMSSPE</b>
Circular	0.1095	2.2359	3.5069	0.0210	0.6839
Spherical	0.1262	2.2081	3.5006	0.0247	0.6795
Tetraspherical	0.1418	2.1966	3.4932	0.0284	0.6780
Pentaspherical	0.1403	2.1826	3.4895	0.0281	0.6742
Exponential	0.1505	2.1650	3.4627	0.0281	0.6764
Gaussian	0.0857	2.2258	3.4985	0.0148	0.6860
Rational quadratic	0.1453	2.1377	3.4507	0.0263	0.6767
Hole Effect	0.1638	2.5917	3.5444	0.0415	0.7525
K-Bessel	0.1449	2.1495	3.4613	0.0257	0.6735
J-Bessel	0.1013	2.1794	3.4742	0.0198	0.6766
Stable	0.1617	2.1254	3.4478	0.0289	0.6710

Table F.8: Cross validation of OK for Fe

<b>Models</b>	<sup>a</sup> <b>MPE</b>	<sup>b</sup> <b>RMSPE</b>	<sup>c</sup> <b>ASPE</b>	<sup>d</sup> <b>MSPE</b>	<sup>e</sup> <b>RMSSPE</b>
Circular	15.5978	473.5868	526.5564	0.0064	0.8837
Spherical	14.0698	473.7971	528.6544	0.0040	0.8770
Tetraspherical	13.4335	474.4378	530.6899	0.0030	0.8741
Pentaspherical	13.4994	474.8952	532.4967	0.0036	0.8718
Exponential	9.9977	473.7029	545.1669	0.0006	0.8579
Gaussian	30.1262	473.3400	522.1413	0.0364	0.9126
Rational quadratic	12.3342	472.1590	535.1266	0.0021	0.8576
Hole Effect	21.3869	515.5247	459.5201	0.0103	1.1554
K-Bessel	28.1910	471.0704	521.7186	0.0316	0.8911
J-Bessel	17.8915	496.5230	482.8127	0.0009	1.0235
Stable	30.1262	473.3400	521.1676	0.0364	0.9143

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table F.9: Cross validation of OK for Hg

<b>Models</b>	<sup>a</sup> <b>MPE</b>	<sup>b</sup> <b>RMSPE</b>	<sup>c</sup> <b>ASPE</b>	<sup>d</sup> <b>MSPE</b>	<sup>e</sup> <b>RMSSPE</b>
Circular	0.1721	1.6429	1.9178	0.0528	0.9540
Spherical	0.1728	1.6347	1.9276	0.0544	0.9353
Tetraspherical	0.1855	1.6332	1.9546	0.0581	0.9074
Pentaspherical	0.1924	1.6314	1.9779	0.0600	0.8831
Exponential	0.1636	1.5863	2.0577	0.0481	0.7874
Gaussian	0.2177	1.7181	1.9206	0.0845	1.2411
Rational quadratic	0.2095	1.5941	2.0216	0.0679	0.8200
Hole Effect	0.2569	1.9193	1.7835	0.0984	1.4018
K-Bessel	0.2391	1.7386	1.9103	0.0922	1.1471
J-Bessel	0.2566	2.0116	1.8865	0.1208	1.7895
Stable	0.2354	1.7193	1.9143	0.0823	1.0969

Table F.10: Cross validation of OK for K

<b>Models</b>	<sup>a</sup> <b>MPE</b>	<sup>b</sup> <b>RMSPE</b>	<sup>c</sup> <b>ASPE</b>	<sup>d</sup> <b>MSPE</b>	<sup>e</sup> <b>RMSSPE</b>
Circular	3.3108	81.3544	95.4607	0.0235	0.8825
Spherical	3.9613	81.3065	94.2829	0.0282	0.8938
Tetraspherical	3.8030	81.1803	93.5258	0.0263	0.9009
Pentaspherical	3.8236	81.3200	93.0550	0.0261	0.9067
Exponential	6.7203	80.8301	86.4310	0.0485	0.9688
Gaussian	2.5606	87.8723	95.8443	0.0182	0.9549
Rational quadratic	3.3852	83.1154	91.0340	0.0216	0.9633
Hole Effect	0.7908	87.7489	96.0632	0.0084	0.9449
K-Bessel	5.8927	79.1427	91.0463	0.0394	0.8883
J-Bessel	0.7759	88.6311	95.5048	0.0097	0.9583
Stable	6.1178	79.3421	90.3025	0.0414	0.8969

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error

Table F.11: Cross validation of OK for Mn

<b>Models</b>	<sup>a</sup> <b>MPE</b>	<sup>b</sup> <b>RMSPE</b>	<sup>c</sup> <b>ASPE</b>	<sup>d</sup> <b>MSPE</b>	<sup>e</sup> <b>RMSSPE</b>
Circular	-0.0424	5.4551	6.7255	-0.0122	0.8431
Spherical	-0.0201	5.4858	6.7195	-0.0105	0.8515
Tetraspherical	0.0101	5.5161	6.7161	-0.0073	0.8593
Pentaspherical	0.0086	5.5081	6.7123	-0.0083	0.8606
Exponential	0.0001	5.4503	6.6959	-0.0111	0.8579
Gaussian	-0.0674	5.6086	6.7535	-0.0207	0.8738
Rational quadratic	0.0148	5.3970	6.6924	-0.0122	0.8537
Hole Effect	0.0795	5.5538	6.6673	0.0045	0.8585
K-Bessel	0.0004	5.4504	6.6937	-0.0111	0.8584
J-Bessel	0.0598	5.6270	6.7378	-0.0032	0.8802
Stable	0.0183	5.4579	6.6912	-0.0078	0.8585

Table F.12: Cross validation of OK for Na

<b>Models</b>	<sup>a</sup> <b>MPE</b>	<sup>b</sup> <b>RMSPE</b>	<sup>c</sup> <b>ASPE</b>	<sup>d</sup> <b>MSPE</b>	<sup>e</sup> <b>RMSSPE</b>
Circular	1.1712	23.6778	24.7096	0.0323	0.9807
Spherical	1.1046	23.7292	24.5166	0.0288	0.9911
Tetraspherical	0.8044	23.4226	24.4146	0.0195	0.9895
Pentaspherical	0.7201	23.3740	24.3630	0.0169	0.9913
Exponential	1.0030	23.5503	23.6622	0.0228	1.0269
Gaussian	0.8527	24.2030	25.1575	0.0242	0.9857
Rational quadratic	0.6671	24.2789	24.1803	0.0147	1.0496
Hole Effect	0.2724	23.4469	24.8908	0.0093	0.9728
K-Bessel	0.8188	23.5122	24.4669	0.0178	0.9905
J-Bessel	0.1945	23.5170	24.7507	0.0076	0.9812
Stable	0.7350	23.8103	24.9044	0.0177	0.9854

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table F.13: Cross validation of OK for Sb

<b>Models</b>	<b><sup>a</sup>MPE</b>	<b><sup>b</sup>RMSPE</b>	<b><sup>c</sup>ASPE</b>	<b><sup>d</sup>MSPE</b>	<b><sup>e</sup>RMSSPE</b>
Circular	0.0788	1.3858	1.8104	0.0301	0.8054
Spherical	0.0800	1.3840	1.8023	0.0296	0.8100
Tetraspherical	0.0778	1.3851	1.7987	0.0282	0.8121
Pentaspherical	0.0773	1.3863	1.7965	0.0276	0.8138
Exponential	0.0842	1.3430	1.7608	0.0280	0.8098
Gaussian	0.0551	1.5386	1.8383	0.0229	0.8763
Rational quadratic	0.0682	1.4500	1.7911	0.0225	0.8704
Hole Effect	0.0150	1.5334	1.7967	0.0093	0.8907
K-Bessel	0.0444	1.5266	1.8370	0.0179	0.8725
J-Bessel	0.0132	1.5435	1.7800	0.0098	0.9005
Stable	0.0551	1.5386	1.8383	0.0229	0.8763

Table F.14: Cross validation of OK for Sc

<b>Models</b>	<b><sup>a</sup>MPE</b>	<b><sup>b</sup>RMSPE</b>	<b><sup>c</sup>ASPE</b>	<b><sup>d</sup>MSPE</b>	<b><sup>e</sup>RMSSPE</b>
Circular	0.1225	2.2521	2.6607	0.0292	0.8858
Spherical	0.1375	2.2557	2.6429	0.0330	0.8957
Tetraspherical	0.1363	2.2678	2.6301	0.0319	0.9043
Pentaspherical	0.1473	2.2796	2.6210	0.0343	0.9105
Exponential	0.0716	2.4712	2.7192	0.0180	0.9507
Gaussian	0.0716	2.4712	2.7192	0.0180	0.9507
Rational quadratic	0.1131	2.3652	2.5998	0.0242	0.9750
Hole Effect	0.0274	2.4032	2.6521	0.0099	0.9459
K-Bessel	0.0884	2.4574	2.7136	0.0225	0.9480
J-Bessel	0.0203	2.4072	2.6201	0.0096	0.9546
Stable	0.0894	2.4383	2.7191	0.0227	0.9388

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.



Table F.15: Cross validation of OK for Sr

<b>Models</b>	<b><sup>a</sup>MPE</b>	<b><sup>b</sup>RMSPE</b>	<b><sup>c</sup>ASPE</b>	<b><sup>d</sup>MSPE</b>	<b><sup>e</sup>RMSSPE</b>
Circular	0.2620	5.6219	7.4673	0.0241	0.7900
Spherical	0.3324	5.6434	7.4303	0.0314	0.7971
Tetraspherical	0.3020	5.6138	7.4012	0.0277	0.7989
Pentaspherical	0.3049	5.6152	7.3871	0.0278	0.8010
Exponential	0.3732	5.4821	7.2911	0.0315	0.7998
Gaussian	0.2278	6.1425	7.5146	0.0227	0.8552
Rational quadratic	0.2848	5.7021	7.3102	0.0241	0.8319
Hole Effect	0.2283	6.1783	7.3757	0.0268	0.8719
K-Bessel	0.3013	5.6363	7.4151	0.0262	0.8027
J-Bessel	0.2679	6.1855	7.3187	0.0330	0.8779
Stable	0.2514	5.8078	7.4630	0.0221	0.8182

Table F.16: Cross validation of OK for Ti

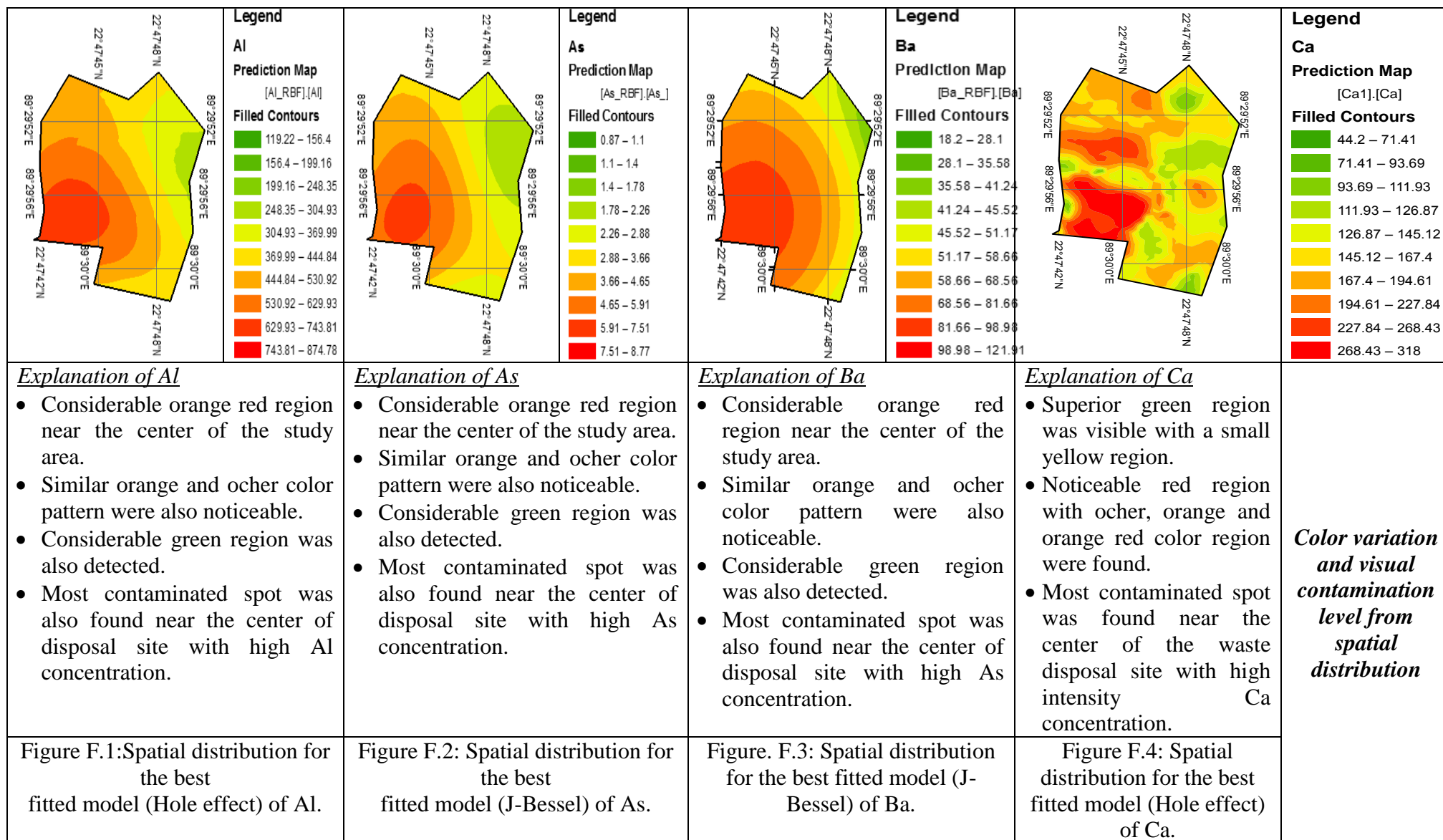
<b>Models</b>	<b><sup>a</sup>MPE</b>	<b><sup>b</sup>RMSPE</b>	<b><sup>c</sup>ASPE</b>	<b><sup>d</sup>MSPE</b>	<b><sup>e</sup>RMSSPE</b>
Circular	12.5442	264.7115	308.0773	0.0230	0.8985
Spherical	17.9196	265.8008	303.4312	0.0363	0.9171
Tetraspherical	17.2483	265.9167	300.3178	0.0340	0.9290
Pentaspherical	18.7797	267.7512	298.6754	0.0375	0.9382
Exponential	25.0203	265.8430	289.5669	0.0522	0.9575
Gaussian	9.3971	294.3402	315.9850	0.0191	0.9806
Rational quadratic	12.7394	280.3912	296.7434	0.0230	1.0204
Hole Effect	6.2457	283.8775	308.6859	0.0169	0.9642
K-Bessel	18.1275	262.7769	305.4450	0.0340	0.8953
J-Bessel	3.7947	283.5454	304.7957	0.0130	0.9712
Stable	15.3638	273.4276	311.5708	0.0294	0.9238

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

Table F.17: Cross validation of OK for V

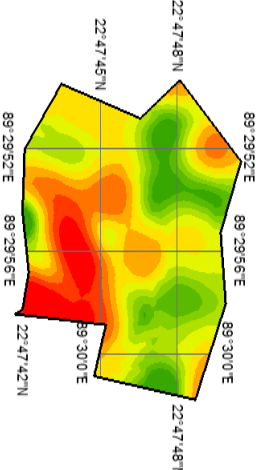
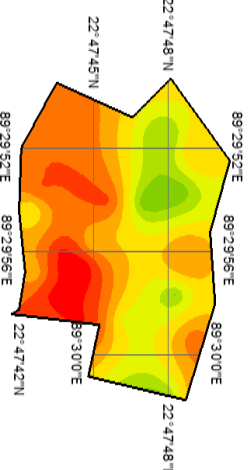
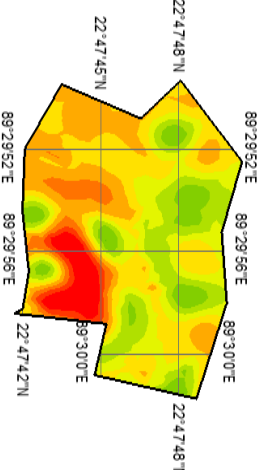
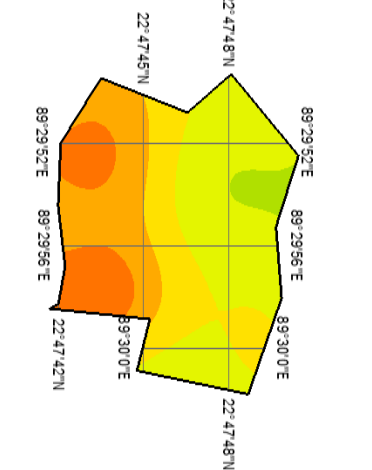
<b>Models</b>	<b><sup>a</sup>MPE</b>	<b><sup>b</sup>RMSPE</b>	<b><sup>c</sup>ASPE</b>	<b><sup>d</sup>MSPE</b>	<b><sup>e</sup>RMSSPE</b>
Circular	0.4901	11.2247	14.4912	0.0218	0.8042
Spherical	0.5535	10.9931	14.4034	0.0247	0.7937
Tetraspherical	0.5700	10.8401	14.3249	0.0255	0.7879
Pentaspherical	0.6037	10.6842	14.2530	0.0267	0.7809
Exponential	0.9192	10.3221	14.0155	0.0395	0.7722
Gaussian	0.5173	11.1754	14.4656	0.0232	0.8112
Rational quadratic	1.3221	10.3888	13.9480	0.0516	0.9463
Hole Effect	0.7895	12.1698	14.4549	0.0451	0.8677
K-Bessel	0.9419	10.3017	13.9821	0.0414	0.7715
J-Bessel	0.5437	11.1034	14.3451	0.0247	0.8095
Stable	0.9095	10.2412	13.9460	0.0399	0.7692

<sup>a</sup>MPE= Mean Prediction Error, <sup>b</sup>RMSPE= Root Mean Square Prediction Error, <sup>c</sup>MSPE= Mean Standardized Prediction Error, <sup>d</sup>RMSSPE= Root Mean Square Standard Prediction Error, <sup>e</sup>ASPE= Average Standard Prediction Error.

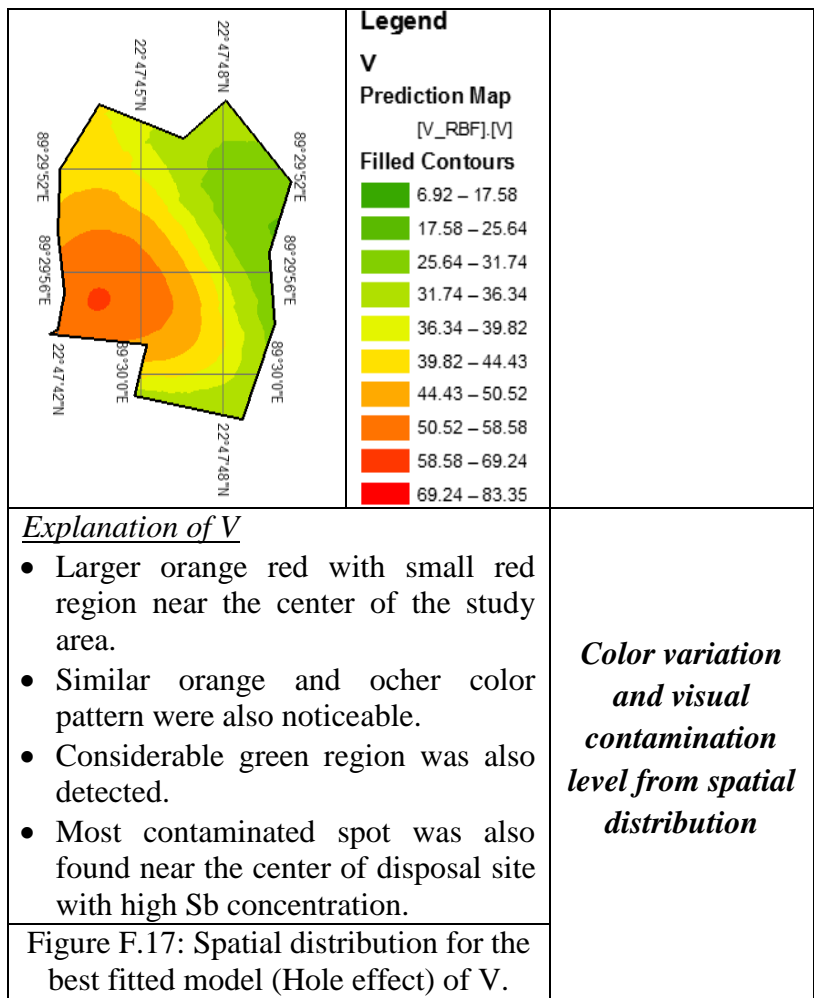


<p><b>Legend</b>  <b>Co</b>  <b>Prediction Map</b>  [Co_RBF].[Co]  <b>Filled Contours</b></p> <ul style="list-style-type: none"> <li>1.98 – 2.98</li> <li>2.98 – 3.82</li> <li>3.82 – 4.52</li> <li>4.52 – 5.11</li> <li>5.11 – 5.81</li> <li>5.81 – 6.65</li> <li>6.65 – 7.65</li> <li>7.65 – 8.86</li> <li>8.86 – 10.3</li> <li>10.3 – 12.02</li> </ul>	<p><b>Legend</b>  <b>Cr</b>  <b>Prediction Map</b>  [Cr_RBF].[Cr]  <b>Filled Contours</b></p> <ul style="list-style-type: none"> <li>0.77 – 1.61</li> <li>1.61 – 2.4</li> <li>2.4 – 3.14</li> <li>3.14 – 3.94</li> <li>3.94 – 4.78</li> <li>4.78 – 5.67</li> <li>5.67 – 6.61</li> <li>6.61 – 7.62</li> <li>7.62 – 8.69</li> <li>8.69 – 9.82</li> </ul>	<p><b>Legend</b>  <b>Cu</b>  <b>Prediction Map</b>  [Cu_RBF].[Cu]  <b>Filled Contours</b></p> <ul style="list-style-type: none"> <li>0.73 – 2.03</li> <li>2.03 – 2.87</li> <li>2.87 – 3.41</li> <li>3.41 – 3.76</li> <li>3.76 – 4.3</li> <li>4.3 – 5.13</li> <li>5.13 – 6.43</li> <li>6.43 – 8.45</li> <li>8.45 – 11.59</li> <li>11.59 – 16.45</li> </ul>	<p><b>Legend</b>  <b>Fe</b>  <b>Prediction Map</b>  [Fe_RBF].[Fe]  <b>Filled Contours</b></p> <ul style="list-style-type: none"> <li>160.71 – 343.42</li> <li>343.42 – 526.12</li> <li>526.12 – 708.83</li> <li>708.83 – 891.53</li> <li>891.53 – 1,074.24</li> <li>1,074.24 – 1,256.94</li> <li>1,256.94 – 1,439.65</li> <li>1,439.65 – 1,622.35</li> <li>1,622.35 – 1,805.06</li> <li>1,805.06 – 1,987.76</li> </ul>
<p><u>Explanation of Co</u></p> <ul style="list-style-type: none"> <li>• Larger orange red region near the center of the study area.</li> <li>• Similar orange and ocher color pattern were also noticeable.</li> <li>• Considerable green region was also detected.</li> <li>• Most contaminated spot was also found near the center of disposal site with high Co concentration.</li> </ul>	<p><u>Explanation of Cr</u></p> <ul style="list-style-type: none"> <li>• No red or orange red region was visible.</li> <li>• Small ocher region with large yellow region was observed.</li> <li>• Considerable green region was also detected.</li> <li>• Contamination of soil of waste disposal site is comparatively less</li> </ul>	<p><u>Explanation of Cu</u></p> <ul style="list-style-type: none"> <li>• Larger orange red region near the center of the study area.</li> <li>• Similar orange and ocher color pattern were also noticeable.</li> <li>• Considerable green region was also detected.</li> <li>• Most contaminated spot was also found near the center of disposal site with high Co concentration</li> </ul>	<p><u>Explanation of Fe</u></p> <ul style="list-style-type: none"> <li>• Larger green region was visible with a small yellow region.</li> <li>• Noticeable red region with ocher, orange and orange red color region were found.</li> <li>• Most contaminated spot was found near the center with high intensity Fe concentration.</li> </ul>
<p>Figure F.5: Spatial distribution for the best fitted model (K-Bessel) of Co.</p>	<p>Figure F.6: Spatial distribution for the best fitted model (Circular) of Cr.</p>	<p>Figure F.7: Spatial distribution for the best fitted model (Hole effect) of Cu.</p>	<p>Figure F.8: Spatial distribution for the best fitted model (J-Bessel) of Fe.</p>

*Color variation and visual contamination level from spatial distribution*

	<p><b>Legend</b></p> <p><b>Hg Prediction Map</b> [Hg_RBF],[Hg]</p> <p><b>Filled Contours</b></p> <ul style="list-style-type: none"> <li>0.72 – 1.05</li> <li>1.05 – 1.45</li> <li>1.45 – 1.92</li> <li>1.92 – 2.49</li> <li>2.49 – 3.17</li> <li>3.17 – 3.98</li> <li>3.98 – 4.96</li> <li>4.96 – 6.12</li> <li>6.12 – 7.53</li> <li>7.53 – 9.2</li> </ul>		<p><b>Legend</b></p> <p><b>K Prediction Map</b> [K_RBF],[K]</p> <p><b>Filled Contours</b></p> <ul style="list-style-type: none"> <li>55.04 – 84.94</li> <li>84.94 – 111.99</li> <li>111.99 – 141.9</li> <li>141.9 – 174.94</li> <li>174.94 – 211.45</li> <li>211.45 – 251.81</li> <li>251.81 – 296.4</li> <li>296.4 – 345.68</li> <li>345.68 – 400.14</li> <li>400.14 – 460.33</li> </ul>		<p><b>Legend</b></p> <p><b>Mn Prediction Map</b> [Mn_RBF],[Mn]</p> <p><b>Filled Contours</b></p> <ul style="list-style-type: none"> <li>1.02 – 5.83</li> <li>5.83 – 8.75</li> <li>8.75 – 10.52</li> <li>10.52 – 11.59</li> <li>11.59 – 12.25</li> <li>12.25 – 13.32</li> <li>13.32 – 15.09</li> <li>15.09 – 18.02</li> <li>18.02 – 22.83</li> <li>22.83 – 30.76</li> </ul>		<p><b>Legend</b></p> <p><b>Na Prediction Map</b> [Na_RBF],[Na]</p> <p><b>Filled Contours</b></p> <ul style="list-style-type: none"> <li>8.23 – 14.03</li> <li>14.03 – 21.55</li> <li>21.55 – 31.3</li> <li>31.3 – 43.95</li> <li>43.95 – 60.35</li> <li>60.35 – 72.99</li> <li>72.99 – 82.75</li> <li>82.75 – 90.27</li> <li>90.27 – 96.07</li> <li>96.07 – 100.54</li> </ul>
<p><u>Explanation of Hg</u></p> <ul style="list-style-type: none"> <li>• Larger green region was visible with a small yellow region.</li> <li>• Noticeable red region with ocher, orange and orange red color region were found.</li> <li>• Most contaminated spot was found near the center of the waste disposal site with high intensity Hg concentration.</li> </ul>	<p><u>Explanation of K</u></p> <ul style="list-style-type: none"> <li>• Noticeable green region was visible with a large yellow region.</li> <li>• Clear red region with ocher, orange and orange red color region were found.</li> <li>• Most contaminated spot was found near the center of the waste disposal site with high intensity K concentration.</li> </ul>	<p><u>Explanation of Mn</u></p> <ul style="list-style-type: none"> <li>• Green region was visible with large yellow region.</li> <li>• Red region with ocher, orange and orange red color region were found.</li> <li>• Most contaminated spot was found near the center of the waste disposal site with high intensity Mn concentration.</li> </ul>	<p><u>Explanation of Na</u></p> <ul style="list-style-type: none"> <li>• No orange red region near the center of the study area.</li> <li>• Noticeable orange and ocher color pattern were also noticeable.</li> <li>• Considerable green region was also detected.</li> <li>• Most contaminated spot was also found near the center of disposal site with high Na concentration.</li> </ul>	<p style="text-align: center;"><i>Color variation and visual contamination level from spatial distribution</i></p>			
<p>Figure F.9: Spatial distribution for the best fitted model (J-Bessel) of Hg.</p>	<p>Figure F.10: Spatial distribution for the best fitted model (Exponential) of K.</p>	<p>Figure F.11: Spatial distribution for the best fitted model (J-Bessel) of Mn</p>	<p>Figure F.12: Spatial distribution for the best fitted model (rational quadratic) of Na.</p>				

	<p><b>Legend</b></p> <p><b>Sb</b> Prediction Map [Sb_RBF],[Sb]</p> <p><b>Filled Contours</b></p> <ul style="list-style-type: none"> <li>0.98 – 2.12</li> <li>2.12 – 3.04</li> <li>3.04 – 3.77</li> <li>3.77 – 4.36</li> <li>4.36 – 5.09</li> <li>5.09 – 6.01</li> <li>6.01 – 7.15</li> <li>7.15 – 8.57</li> <li>8.57 – 10.34</li> <li>10.34 – 12.55</li> </ul>		<p><b>Legend</b></p> <p><b>Sc</b> Prediction Map [Sc_RBF],[Sc]</p> <p><b>Filled Contours</b></p> <ul style="list-style-type: none"> <li>3.02 – 5.61</li> <li>5.61 – 7.41</li> <li>7.41 – 8.67</li> <li>8.67 – 9.55</li> <li>9.55 – 10.17</li> <li>10.17 – 11.05</li> <li>11.05 – 12.31</li> <li>12.31 – 14.11</li> <li>14.11 – 16.7</li> <li>16.7 – 20.41</li> </ul>		<p><b>Legend</b></p> <p><b>Sr</b> Prediction Map [Sr_RBF],[Sr]</p> <p><b>Filled Contours</b></p> <ul style="list-style-type: none"> <li>8.88 – 12.94</li> <li>12.94 – 15.77</li> <li>15.77 – 17.75</li> <li>17.75 – 19.13</li> <li>19.13 – 21.11</li> <li>21.11 – 23.94</li> <li>23.94 – 28.01</li> <li>28.01 – 33.83</li> <li>33.83 – 42.17</li> <li>42.17 – 54.12</li> </ul>		<p><b>Legend</b></p> <p><b>Ti</b> Prediction Map [Ti12],[Ti]</p> <p><b>Filled Contours</b></p> <ul style="list-style-type: none"> <li>243.88 – 412.19</li> <li>412.19 – 550</li> <li>550 – 662.85</li> <li>662.85 – 755.25</li> <li>755.25 – 868.1</li> <li>868.1 – 1,005.91986</li> <li>1,005.91986 – 1,174.23</li> <li>1,174.23 – 1,379.78</li> <li>1,379.78 – 1,630.8</li> <li>1,630.8 – 1,937.36</li> </ul>
<p><u>Explanation of Sb</u></p> <ul style="list-style-type: none"> <li>• Larger orange red region near the center of the study area.</li> <li>• Similar orange and ocher color pattern were also noticeable.</li> <li>• Considerable green region was also detected.</li> <li>• Most contaminated spot was also found near the center of disposal site with high Sb concentration.</li> </ul>	<p><u>Explanation of Sc</u></p> <ul style="list-style-type: none"> <li>• Larger orange red region near the center of the study area.</li> <li>• Similar orange and ocher color pattern were also noticeable.</li> <li>• Considerable green region was also detected.</li> <li>• Most contaminated spot was also found near the center of disposal site with high Sc concentration.</li> </ul>	<p><u>Explanation of Sr</u></p> <ul style="list-style-type: none"> <li>• Considerable orange red region near the center of the study area.</li> <li>• Similar orange and ocher color pattern were also noticeable.</li> <li>• Considerable green region was also detected.</li> <li>• Most contaminated spot was also found near the center of disposal site with high Sr concentration.</li> </ul>	<p><u>Explanation of Ti</u></p> <ul style="list-style-type: none"> <li>• Larger red and orange red region near the center of the study area.</li> <li>• Similar orange and ocher color pattern were also noticeable.</li> <li>• Considerable green region was also detected.</li> <li>• Most contaminated spot was also found near the center of disposal site with high Ti concentration.</li> </ul>	<p><b>Color variation and visual contamination level from spatial distribution</b></p>			
<p>Figure F.13: Spatial distribution for the best fitted model (J-Bessel) of Sb.</p>	<p>Figure F.14: Spatial distribution for the best fitted model (J-Bessel) of Sc.</p>	<p>Figure F.15: Spatial distribution for the best fitted model (J-Bessel) of Sr.</p>	<p>Figure F.16: Spatial distribution for the best fitted model (Rational quadratic) of Ti.</p>				



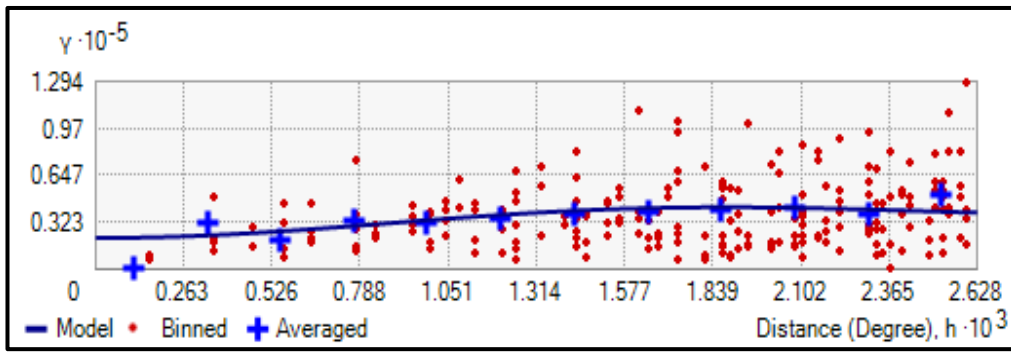


Figure F.18: Semivariogram of Al.

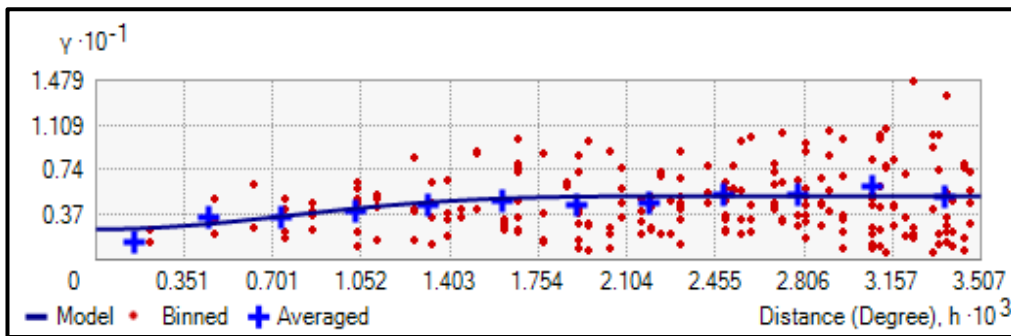


Figure F.19: Semivariogram of As.

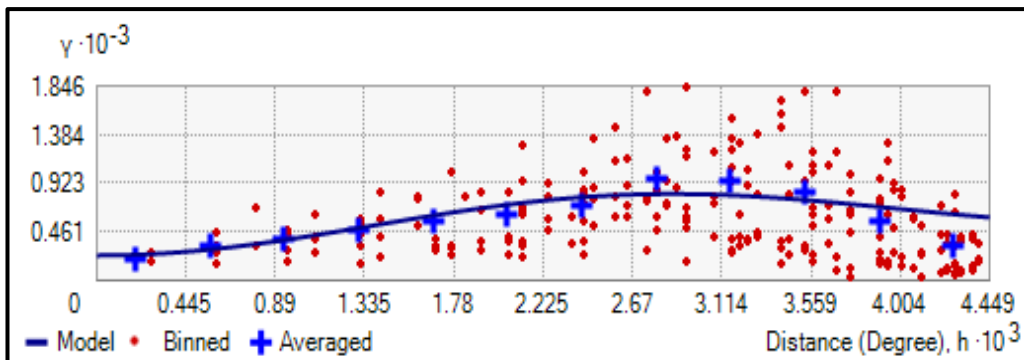


Figure F.20: Semivariogram of Ba.

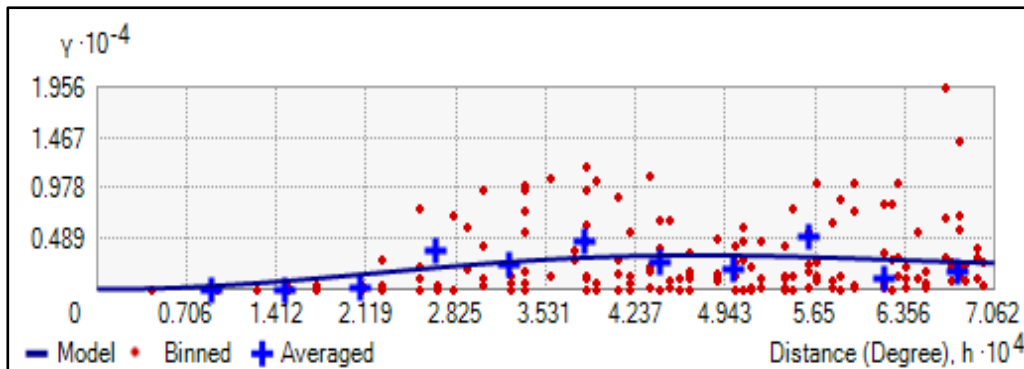


Figure F.21: Semivariogram of Ca.



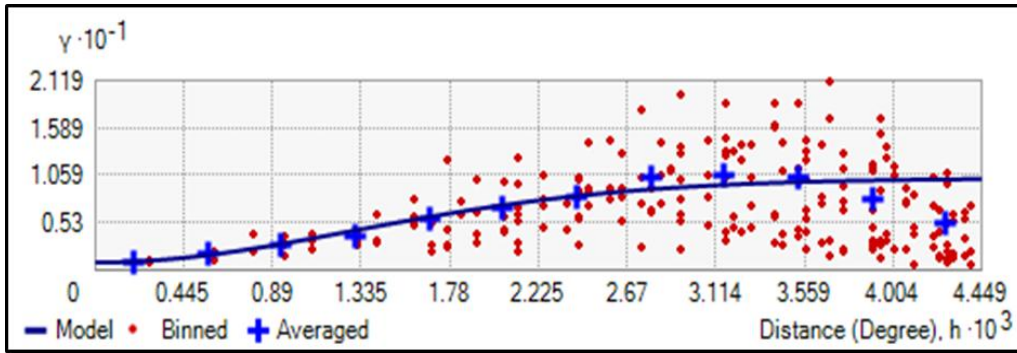


Figure F.22: Semivariogram of Co.

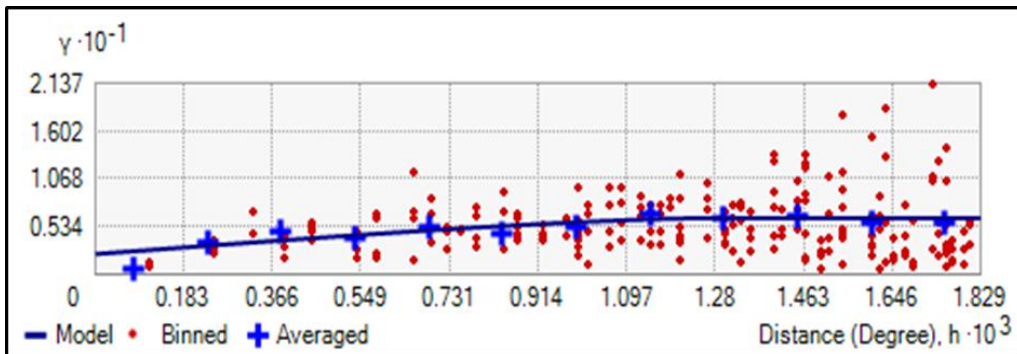


Figure F.23: Semivariogram of Cr.

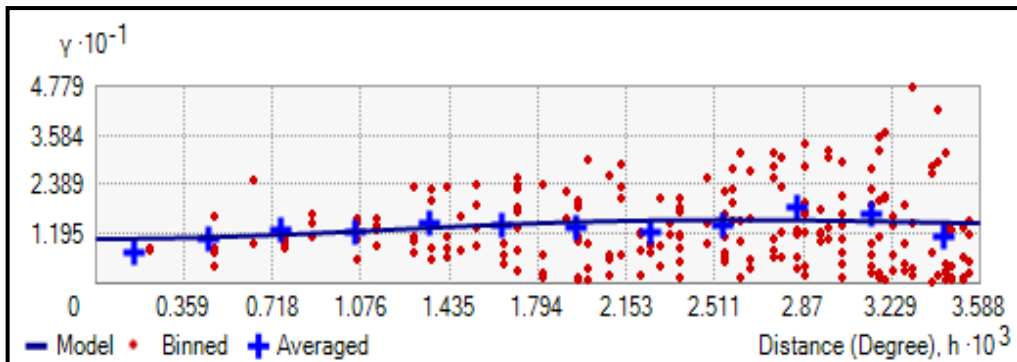


Figure F.24: Semivariogram of Cu.

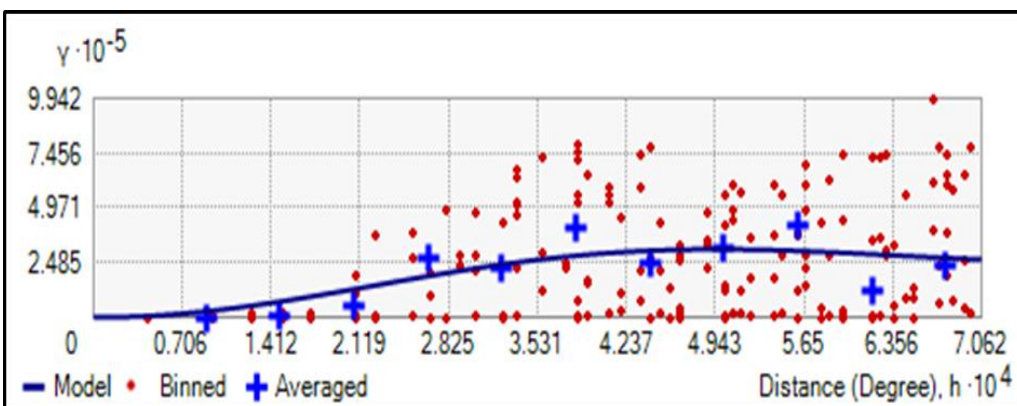


Figure F.25: Semivariogram of Fe.

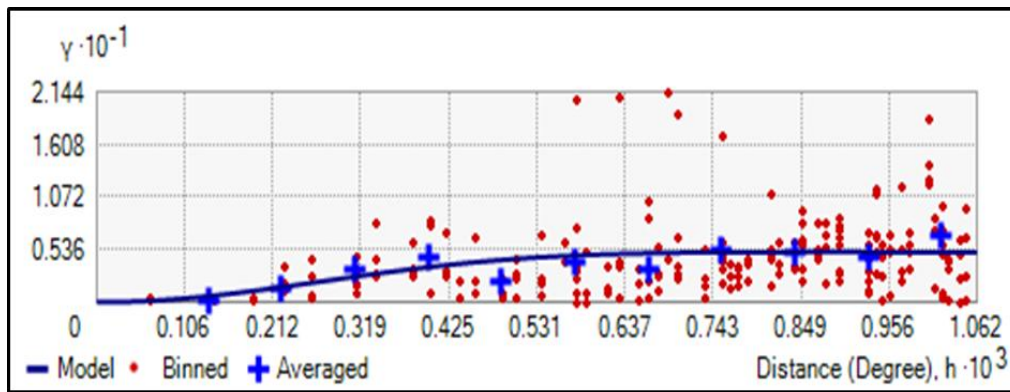


Figure F.26: Semivariogram of Hg.

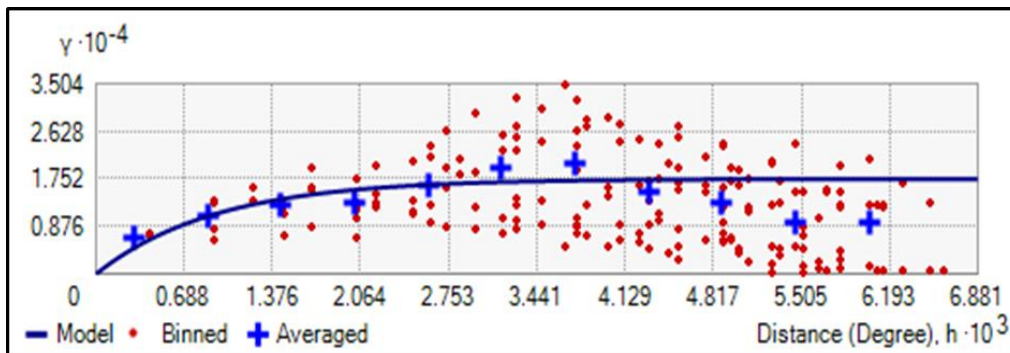


Figure F.27: Semivariogram of K.

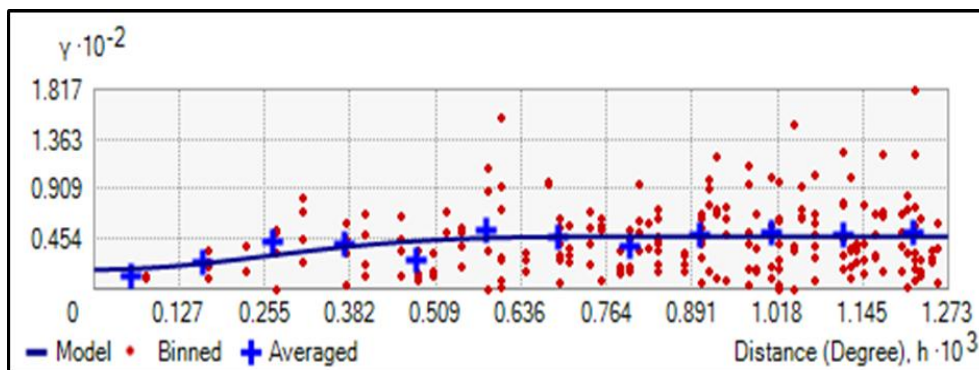


Figure F.28: Semivariogram of Mn.

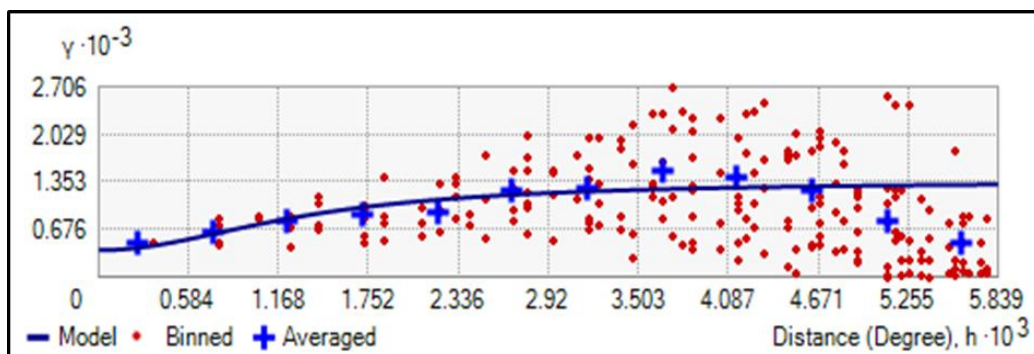


Figure F.29: Semivariogram of Na.

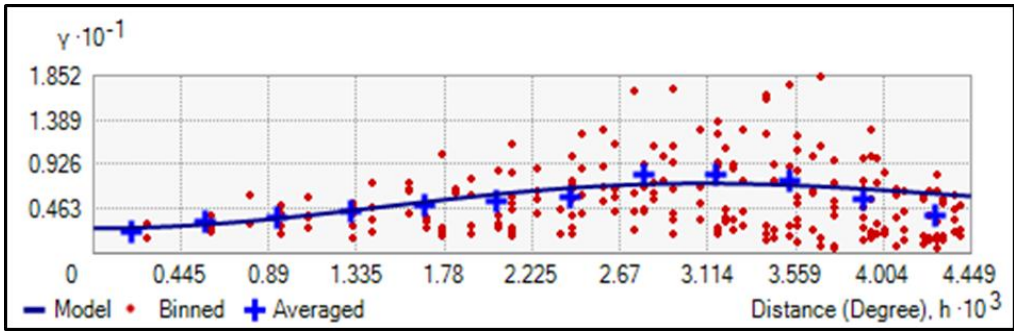


Figure F.30 Semivariogram of Sb

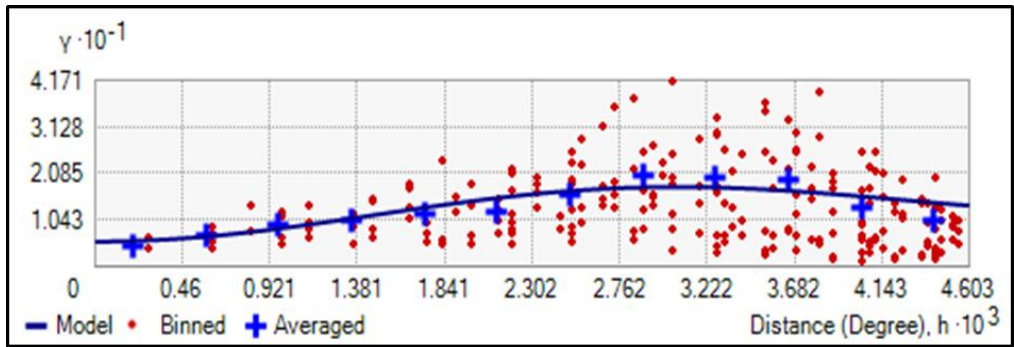


Figure F.31: Semivariogram of Sc.

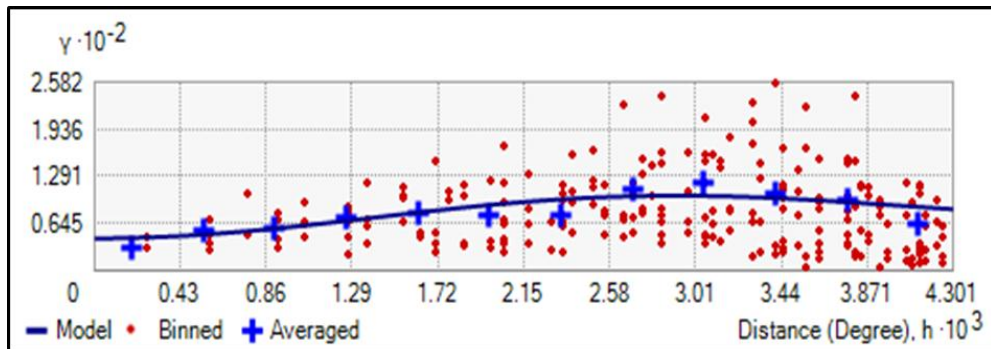


Figure F.32 Semivariogram of Sr.

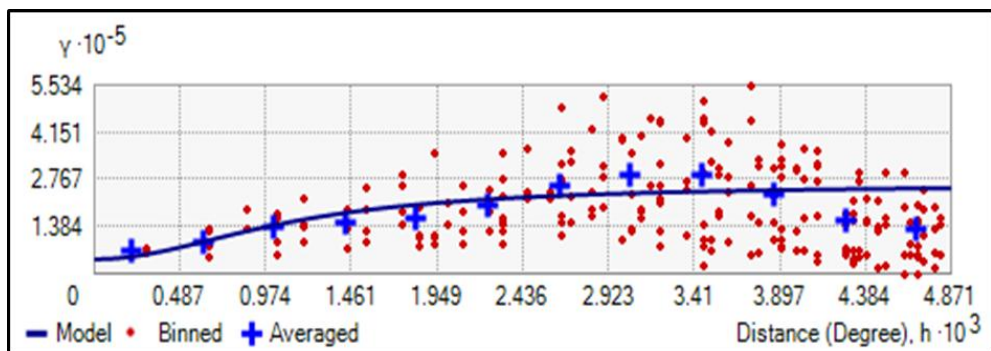


Figure F.33: Semivariogram of Ti.

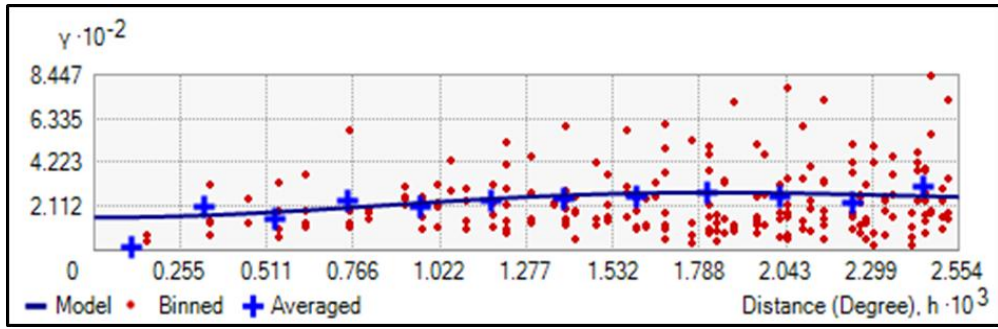


Figure F.34: Semivariogram of V.

## Annex –G

### Performance Assessment of Predicted Data obtained using ANN

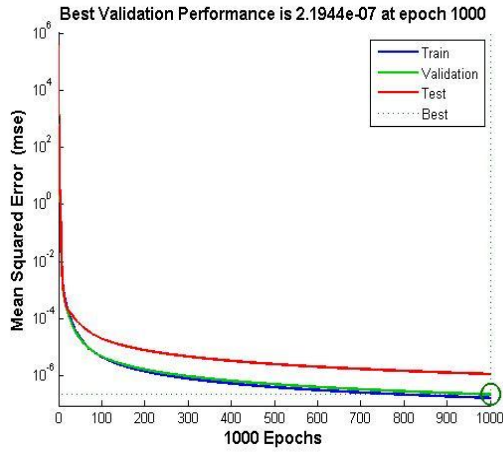


Figure G.1: Performance plot of A1.

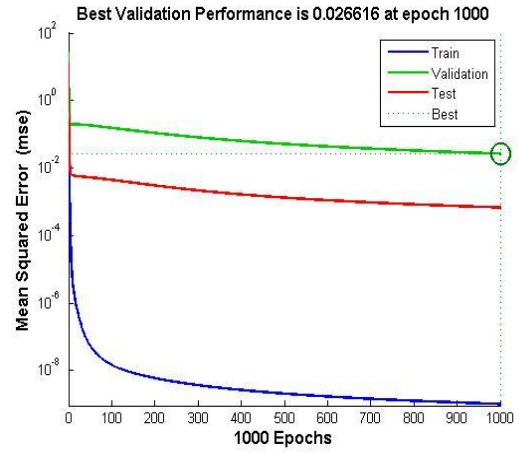


Figure G.2: Performance plot of As.

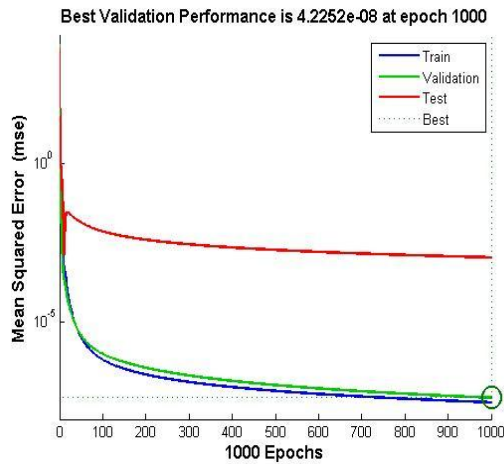


Figure G.3: Performance plot of Ba.

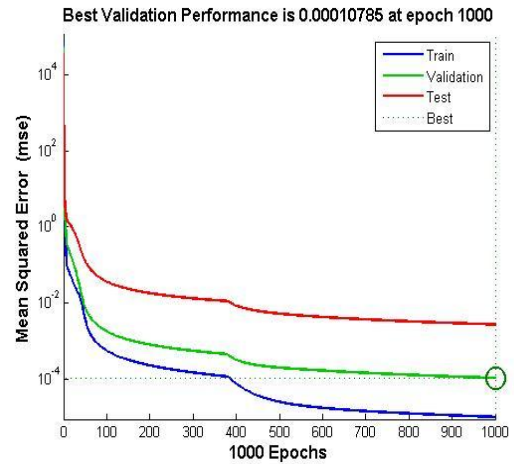


Figure G.4: Performance plot of Ca.

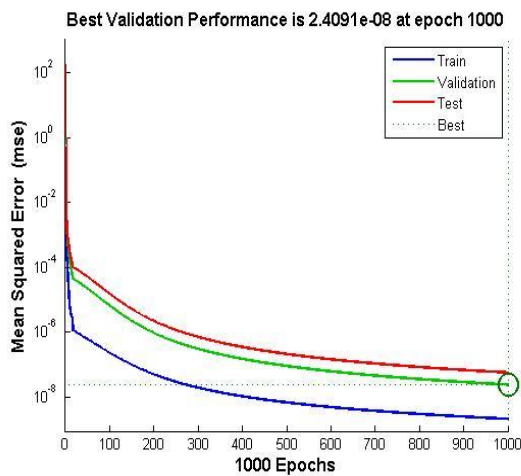


Figure G.5: Performance plot of Co.

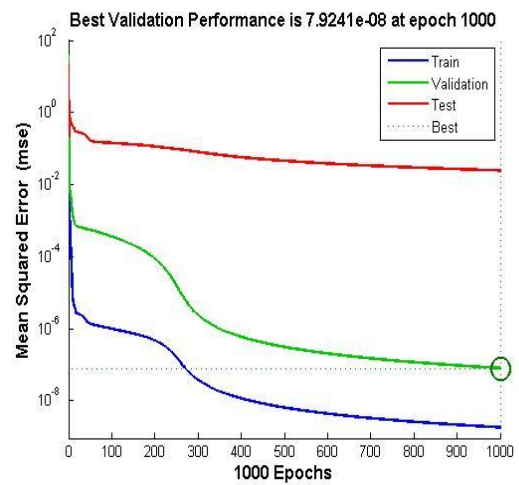


Figure G.6: Performance plot of Cr.

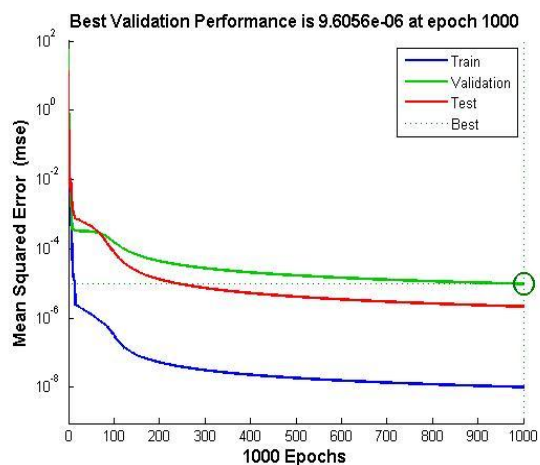


Figure G.7: Performance plot of Cu.

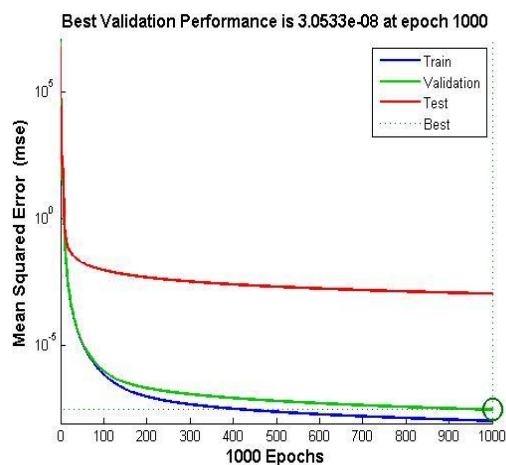


Figure G.8: Performance plot of Fe.

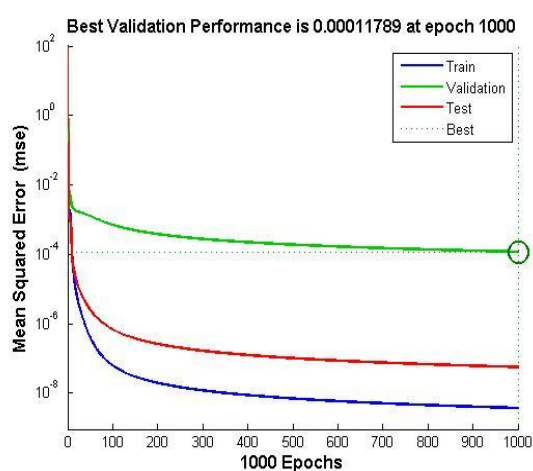


Figure G.9: Performance plot of Hg.

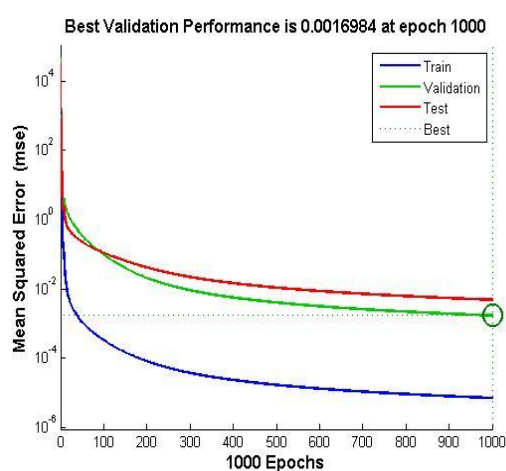


Figure G.10: Performance plot of K.

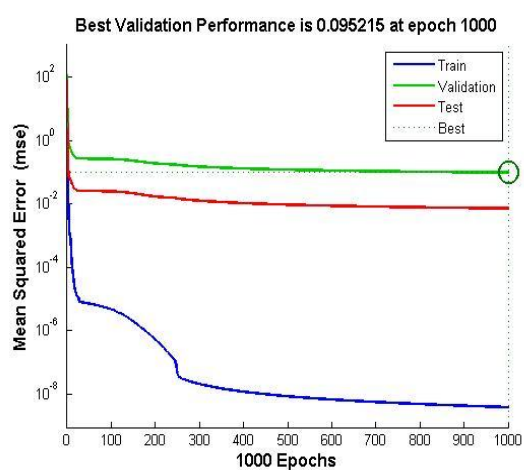


Figure G.11: Performance plot of Mn.

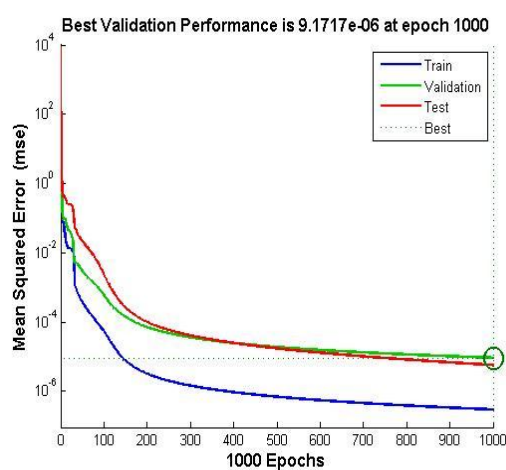


Figure G.12: Performance plot of Na.

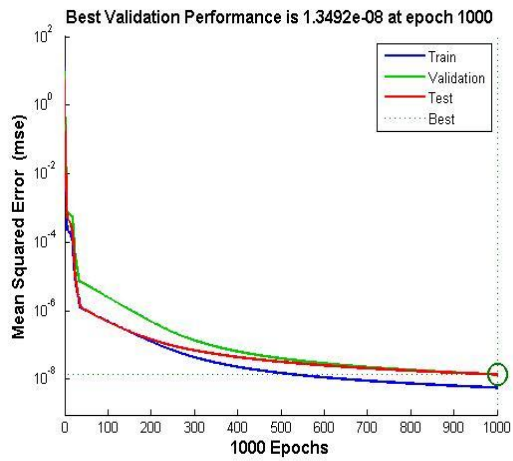


Figure G.13: Performance plot of Sb.

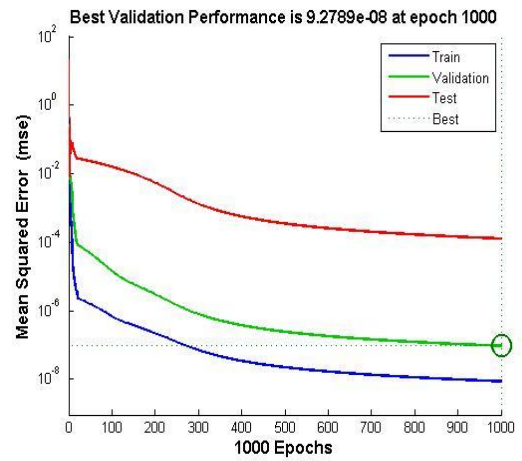


Figure G.14: Performance plot of Sc.

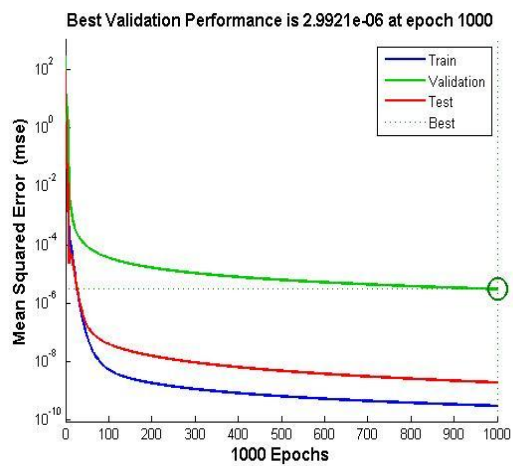


Figure G.15: Performance plot of Sr.

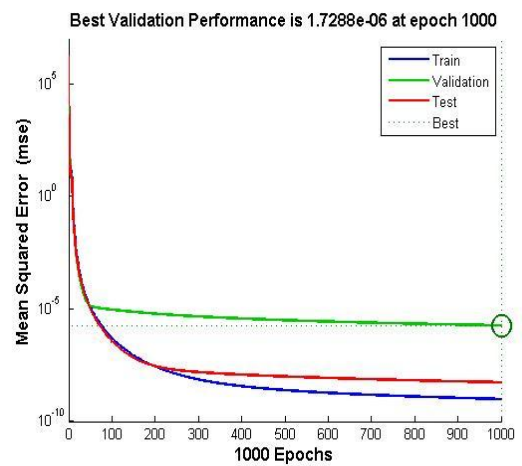


Figure G.16: Performance plot of Ti.

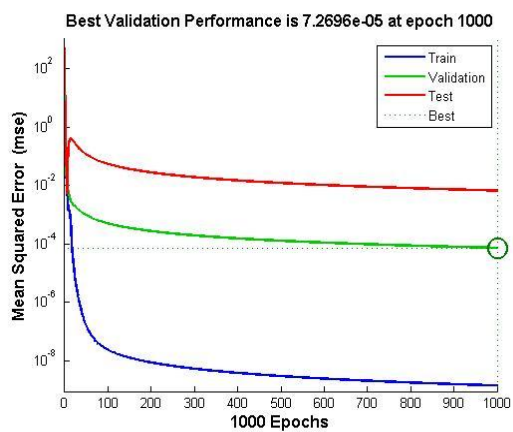


Figure G.17: Performance plot of V

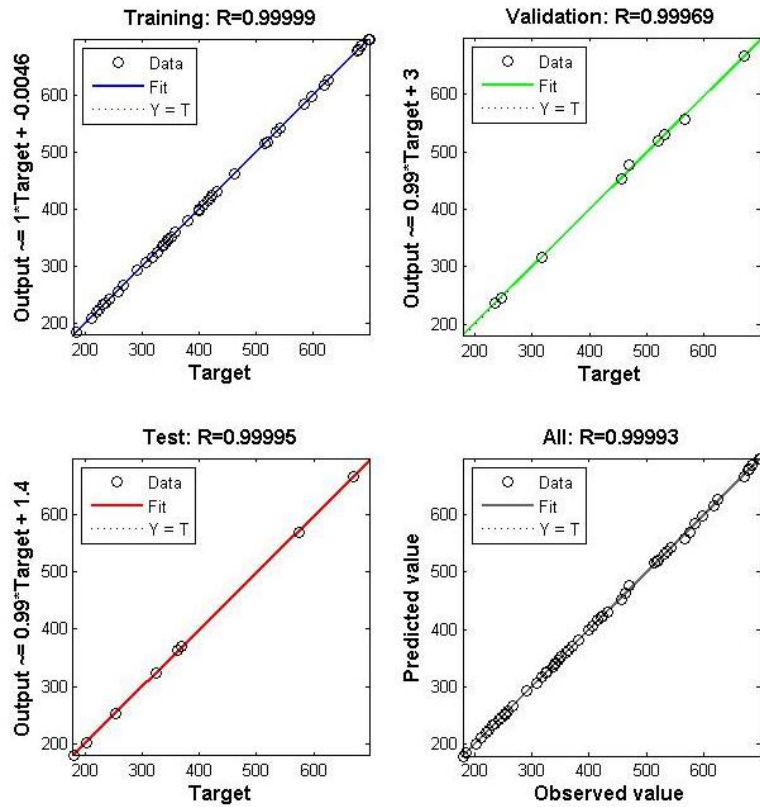


Figure G.18: Regression Coefficient of AI.

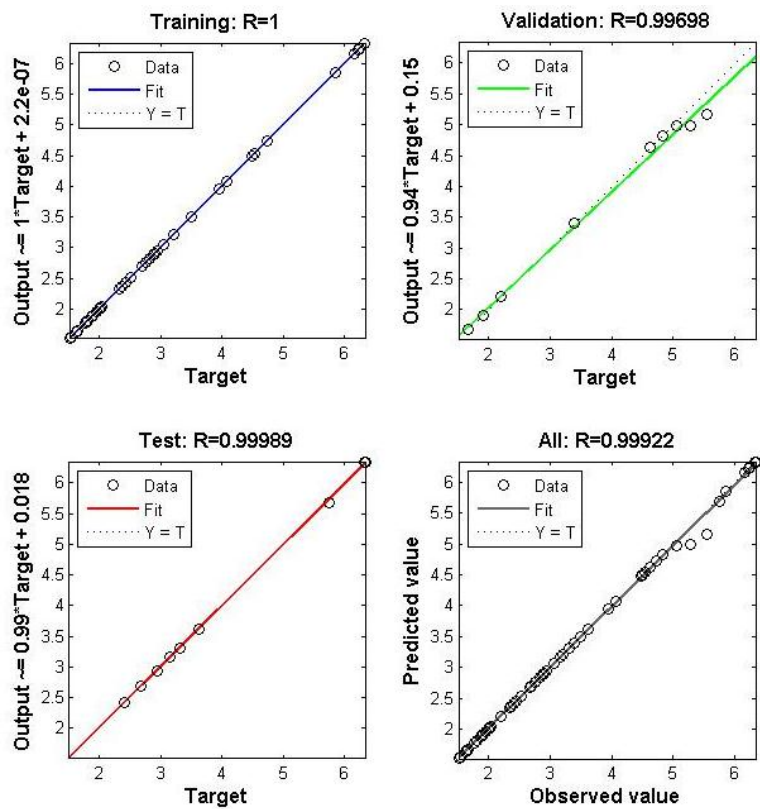


Figure G.19: Regression Coefficient of As.



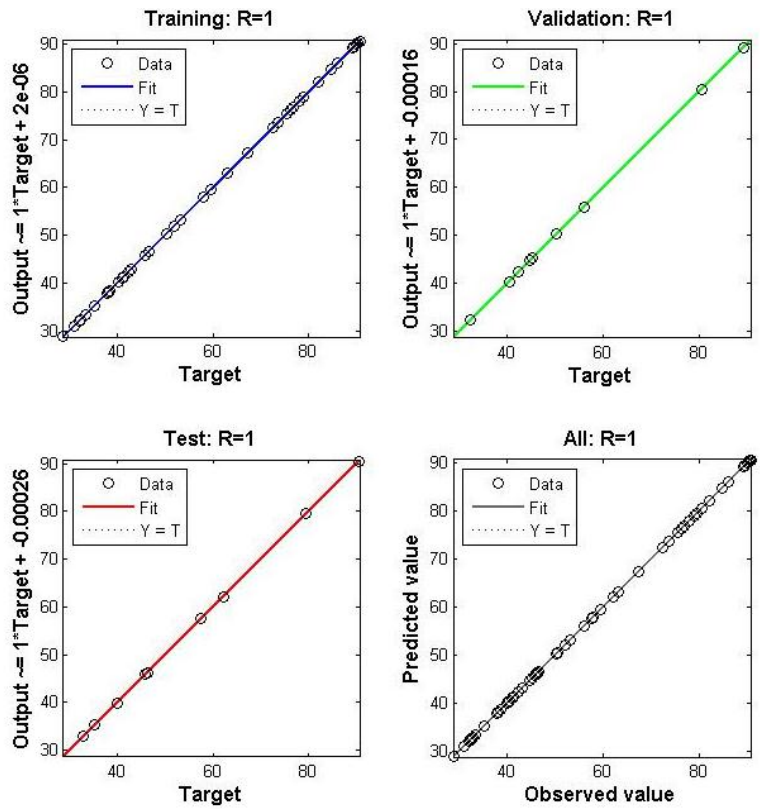


Figure G.20: Regression Coefficient of Ba.

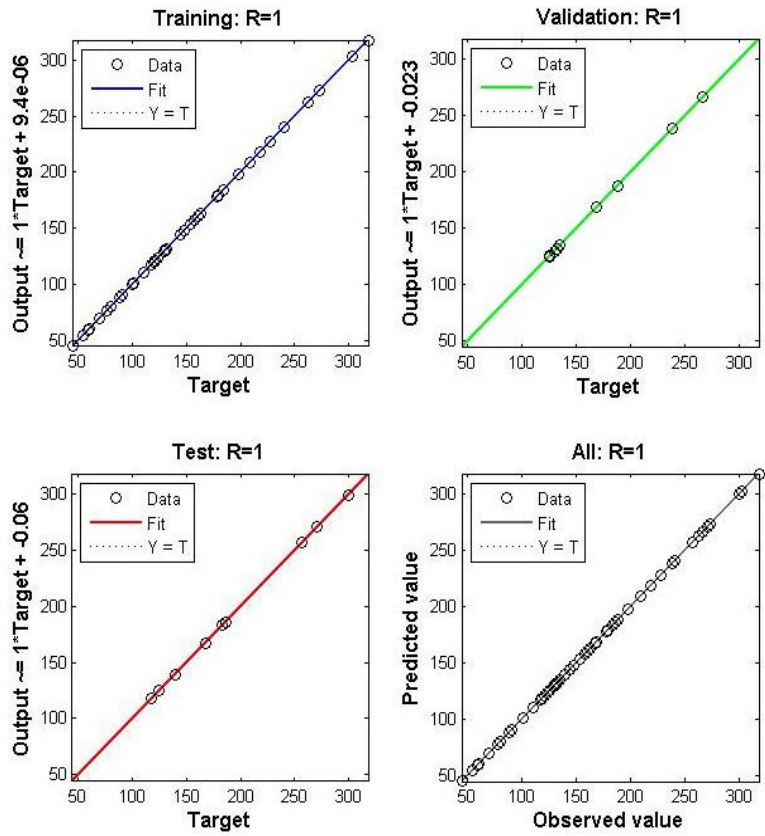


Figure G.21 Regression Coefficient of Ca

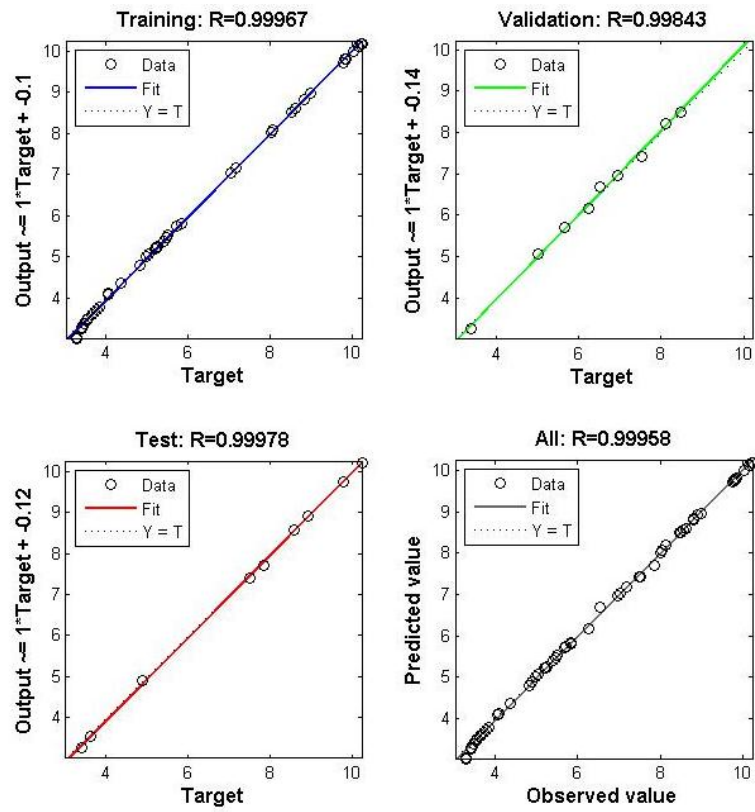


Figure G.22: Regression Coefficient of  $C_o$ .

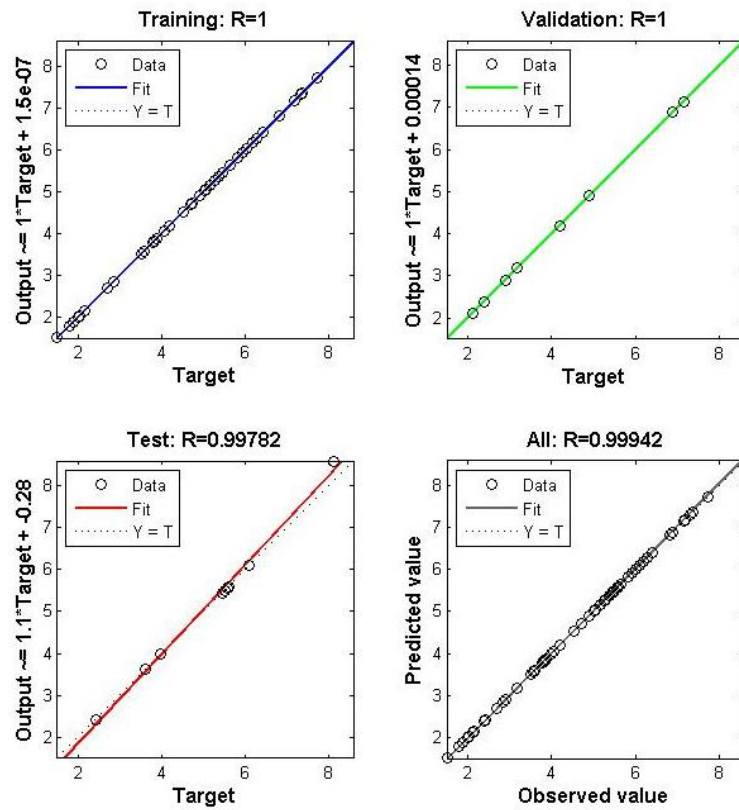


Figure G.23: Regression Coefficient of  $C_r$ .

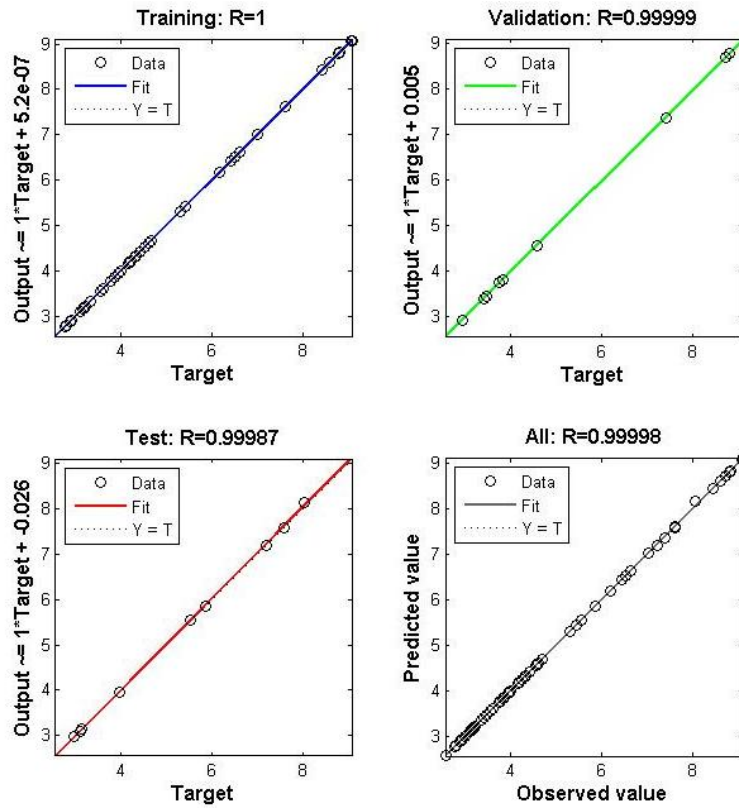


Figure G.24: Regression Coefficient of Cu.

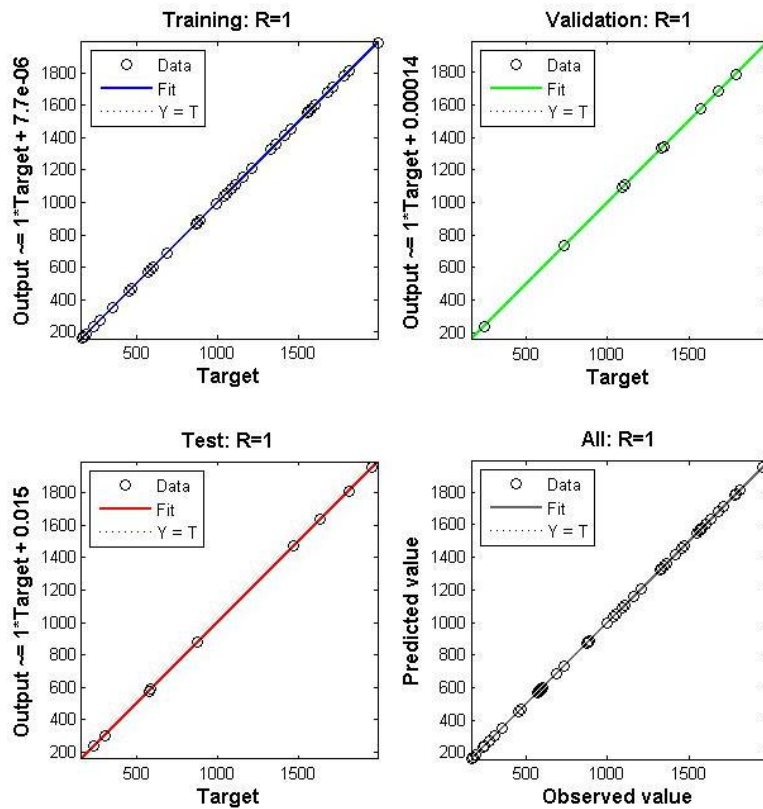


Figure G.25: Regression Coefficient of Fe.

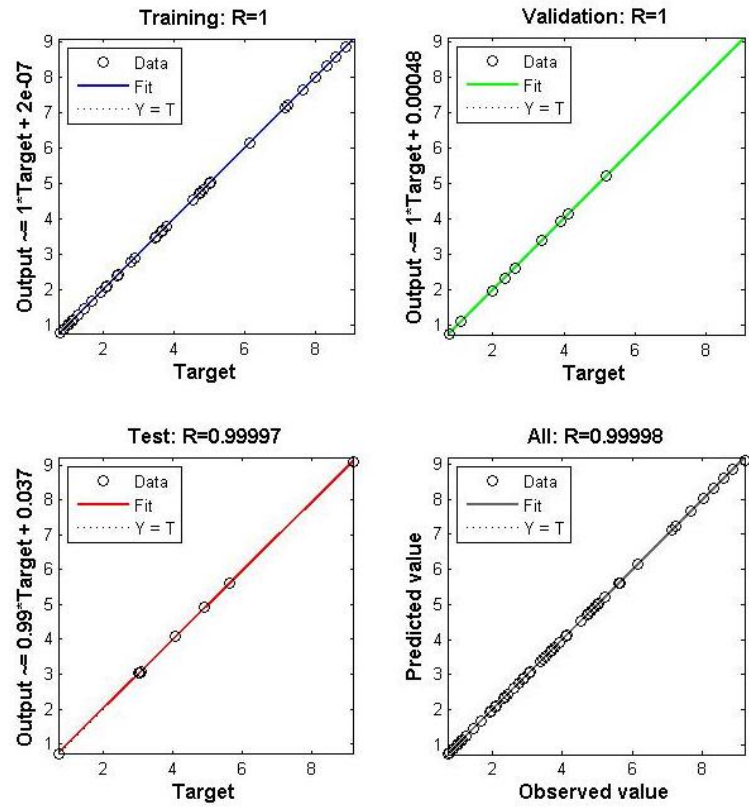


Figure G.26: Regression Coefficient of Hg.

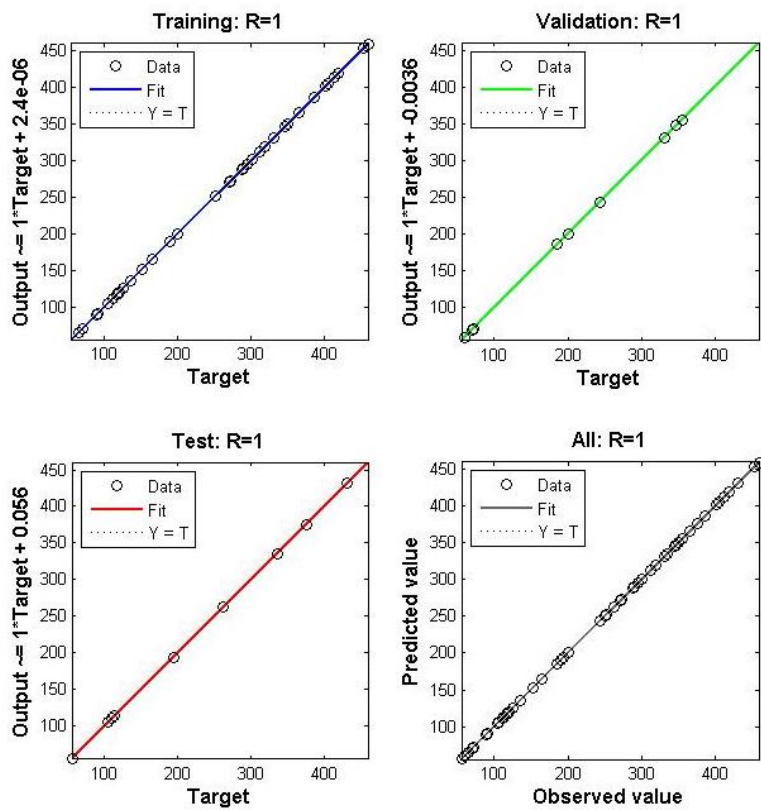


Figure G.27: Regression Coefficient of K.

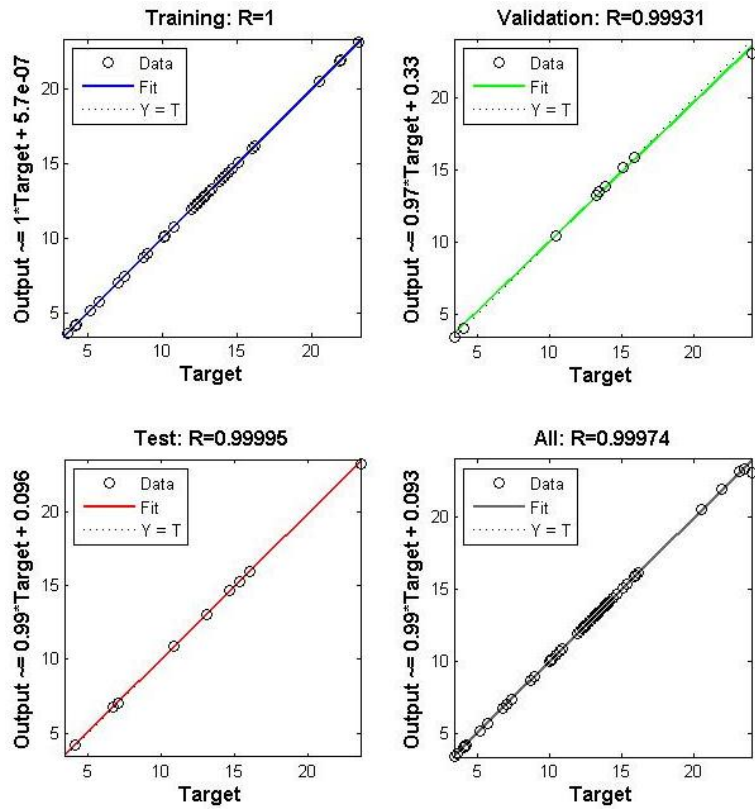


Figure G.28: Regression Coefficient of Mn.

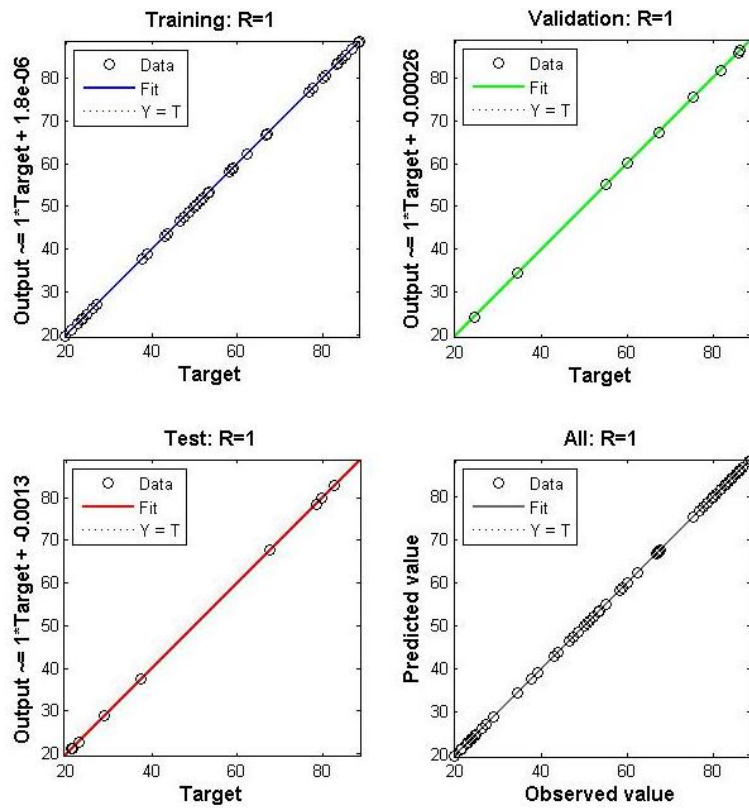


Figure G.29: Regression Coefficient of Na.

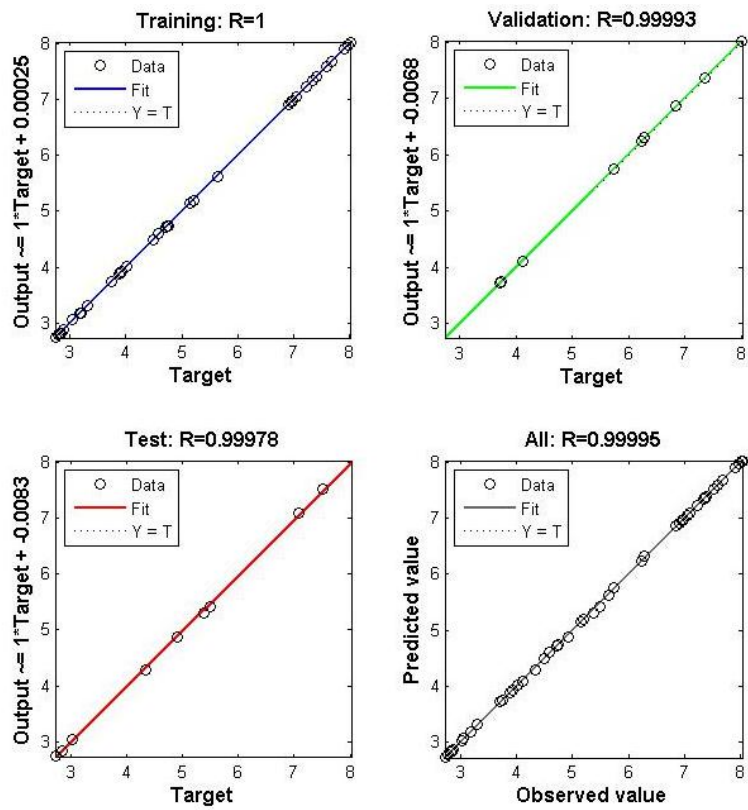


Figure G.30: Regression Coefficient of  $S_b$ .

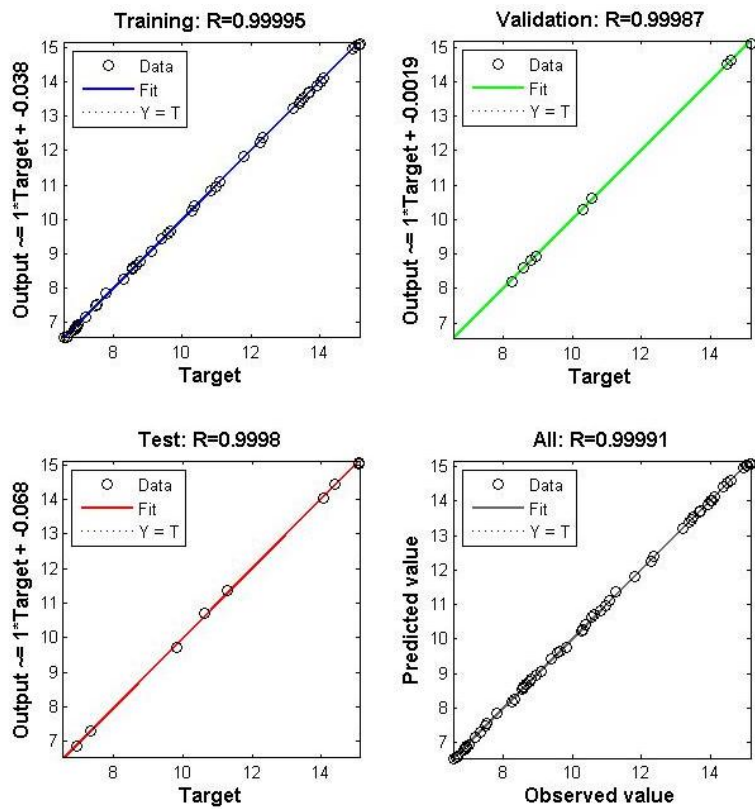


Figure G.31: Regression Coefficient of  $S_c$ .

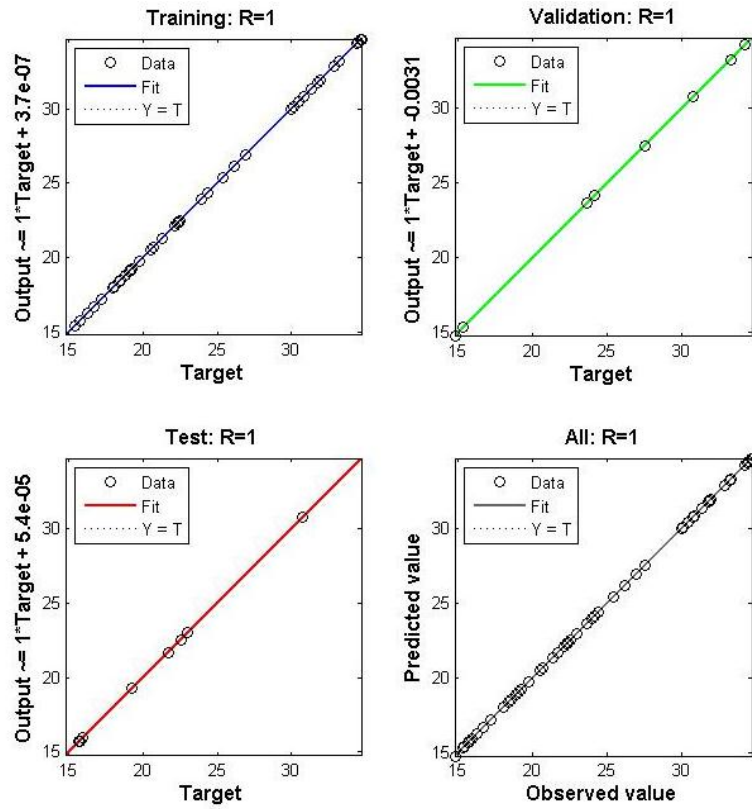


Figure G.32: Regression Coefficient of  $S_r$ .

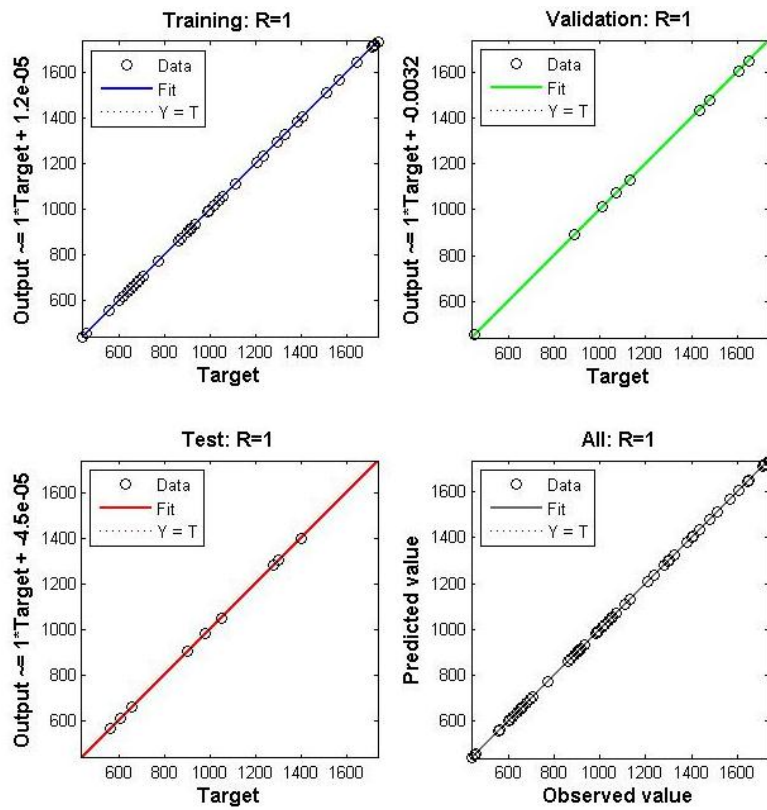


Figure G.33: Regression Coefficient of  $T_i$ .

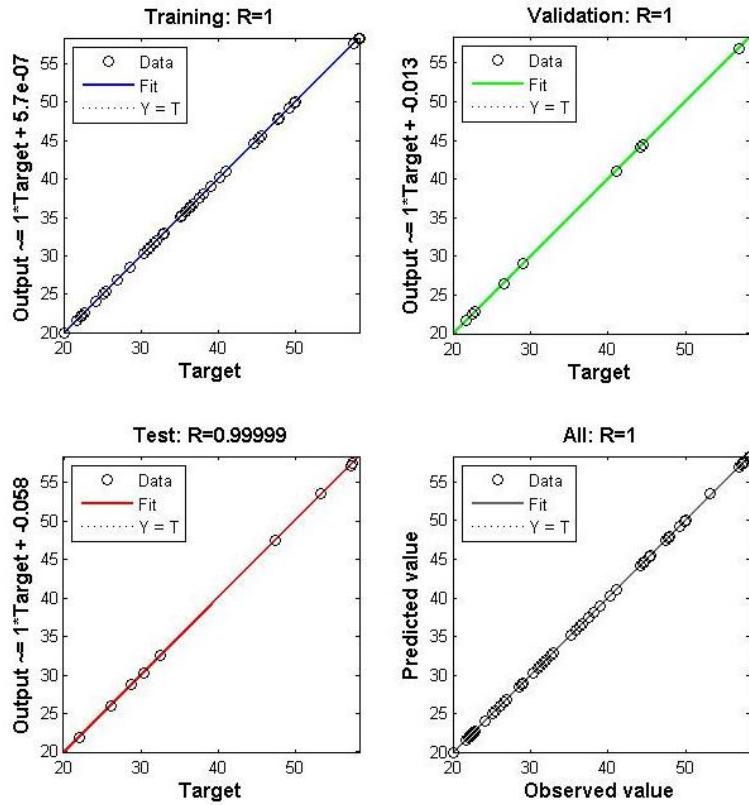


Figure G.34: Regression Coefficient of V.

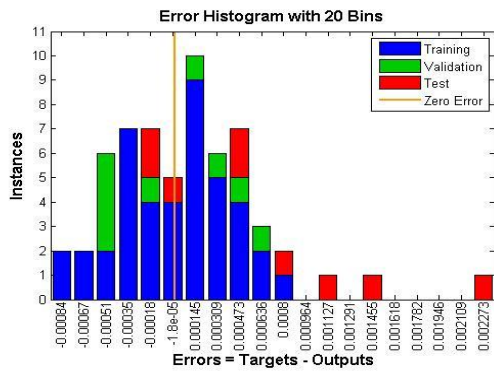


Figure G.35: Histogram plot of A1.

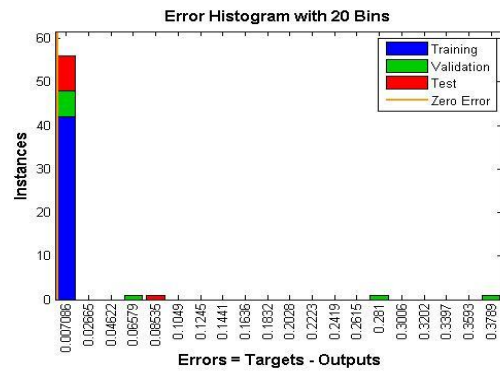


Figure G.36: Histogram plot of A5.

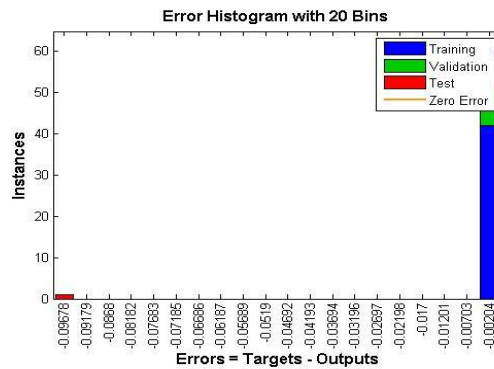


Figure G.37: Histogram plot of B4.

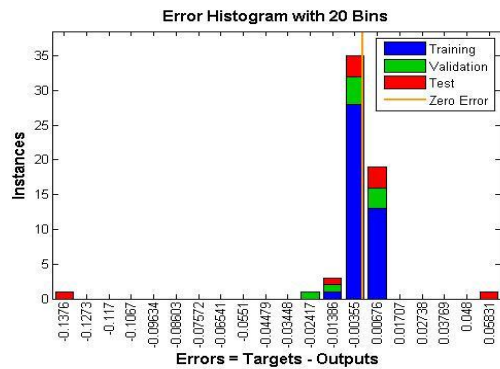


Figure G.38: Histogram plot of C4.



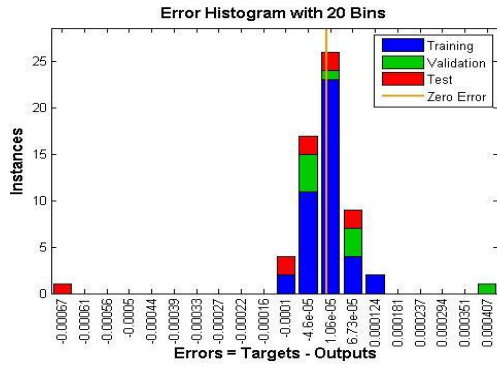


Figure G.39: Histogram plot of Co.

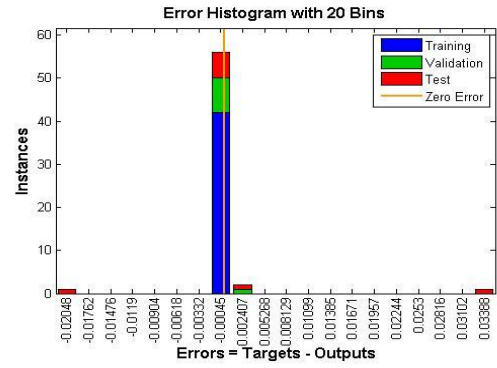


Figure G.40: Histogram plot of Cr.

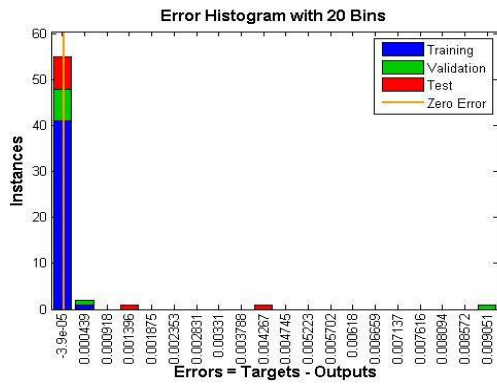


Figure G.41: Histogram plot of Cu.

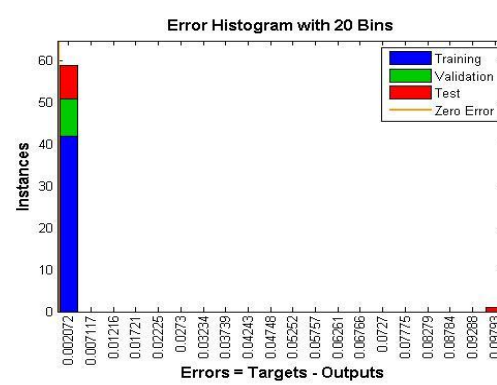


Figure G.42: Histogram plot of Fe.

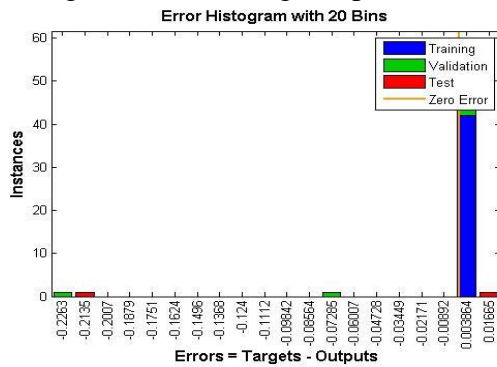


Figure G.43: Performance plot of Hg.

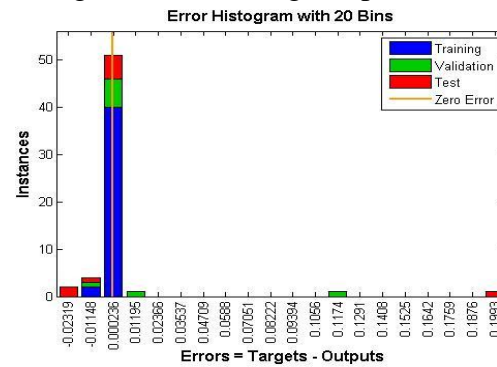


Figure G.44: Performance plot of K.

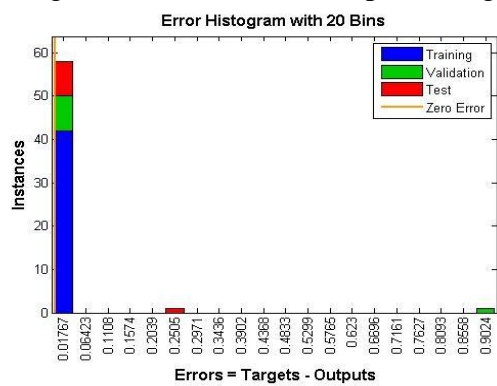


Figure G.45: Performance plot of Mn.

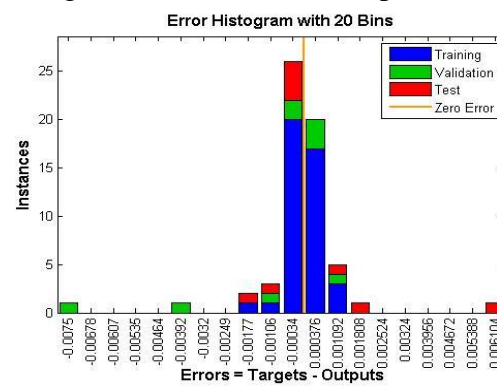


Figure G.46: Performance plot of Na.

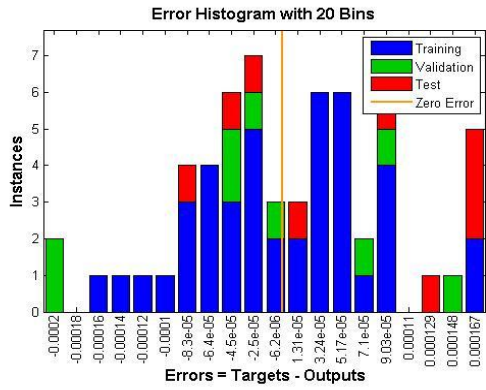


Figure G.47: Performance plot of Sb.

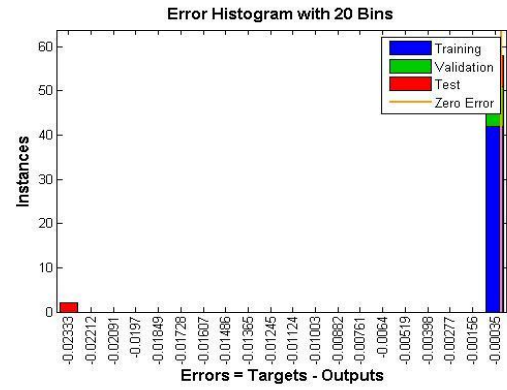


Figure G.48: Performance plot of Sc.

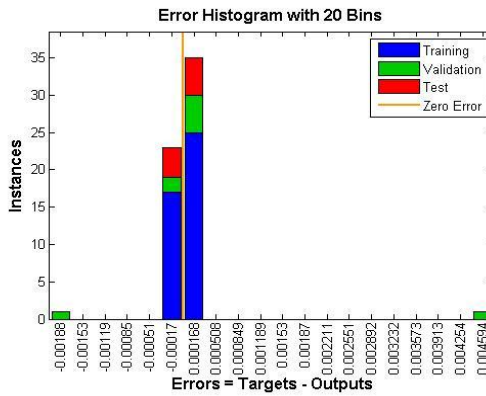


Figure G.49: Performance plot of Sr.

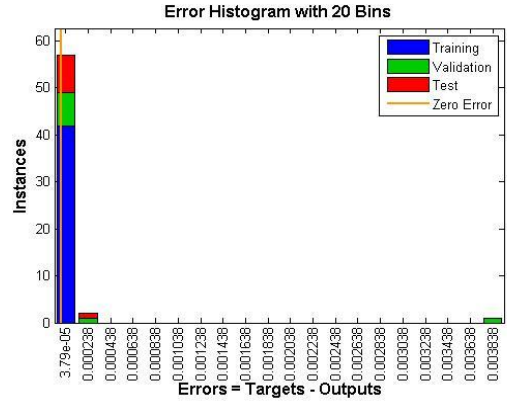


Figure G.50: Performance plot of Ti.

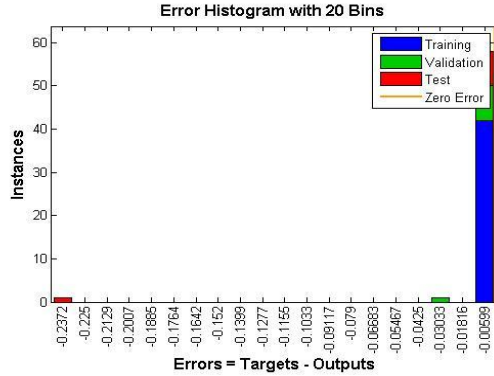


Figure G.51: Performance plot of V.