

**COST-EFFECTIVE STRATEGY FOR THE INVESTIGATION AND
REMEDICATION OF POLLUTED SOIL USING GEOSTATISTICS AND
A GENETIC ALGORITHM APPROACH**

YONGQIANG CUI

2016

Acknowledgment

First and foremost, I would like to give my deepest thanks to Professor Yoneda Minoru, Associate professor Shimada Yoko and Lecturer Matsui Yasuto, especially my supervisor Professor Yoneda, who has been always so patiently and kindly giving me guidance and good advices in my study, no matter how busy he was in this five-and-a-half-year academic oversea study in Japan..

I would also like to give my thanks of gratitude to Ms. Hoshihara Hanae who has really helped me a lot to make my study and life so easy since the day first I came to Japan. Because of her kind help, I can really focus on my study and research.

I owe to deepest gratitude for the committee of my thesis, Prof. Yoshihisa Shimizu Sensei, A.Prof. Yoko Fujikawa Sensei, for their review, priceless input and constructive critics.

Then I am very grateful to Thuong San, Raga San, Gomi Kun who are so kind to modify my thesis so seriously. Especially their suggestions both in thesis, published paper and PowerPoint are so helpful to me.

In addition, the research room members, Mori San, Kometani Kun, Shimomura Kun, Li Kun, Wang Kun, Chen Kun who help a lot in the sampling and experiment, will get my deep appreciation.

At last, I will give my special and deep thanks to all members in the lab, when I had questions, I am always looking for help, and every time you would stop your work at hand to help me, which made me so moved. Because of you all, I am living a happy oversea study life in Kyoto, Japan.

Summary

Heavy metals (HMs) or trace metals normally exist in nature, and in small quantities are nutritionally essential to life. However, HMs can also cause serious environmental pollution and greatly threaten the health of animal and human when the concentration of HMs in soil, water or atmosphere exceeds a certain threshold value. In China, because of the recent thirty years' rapid development of industrialization and urbanization partially at price of environment, soil has been seriously polluted by HMs. Based on the latest survey report released on 17th April, 2014, 16.1% of the Chinese soil was polluted and 82.8% of the polluted land was contaminated by toxic inorganic pollutants and the HMs. Compared with China, the soil pollution has been greatly controlled and decreased in Japan; however the accident of Fushishima Daiichi Nuclear Plant (FDNPP) has caused serious soil pollution of radioactive substances like cesium. Remediation or treatment on so large an area of polluted soil makes the people, company, even government impossible to pay for such large money when using the traditional investigation and remediation methods, in addition the cost of human worker also contribute the high cost of the total project. As a result, developing a cost-effective strategy to investigate and remediate the polluted soil is so important. (*Chapter1*)

In this research, there are main three sections to lead to the cost-effective strategy:

1. Development of the cost-effective strategy on the surface soil (*Chapter2-3*). In this section, we use large numbers of assumed actual polluted fields (AAPFs) generated by unconditional simulation (US) to assess a strategy combining traditional systematic grid sampling (SGS) and kriging methods, and then use a genetic algorithm (GA) to find optimal locations

for the additional sampling. Most of the earlier studies have focused on single actual polluted fields and used conditional simulation (CS) to delineate the distribution of pollution. We propose that AAPFs can be used to develop and assess strategies by assuming many polluted fields, in light of the practical difficulties of finding and operating on large numbers of different and actual polluted fields in the real world. To address the issues that most directly concern stakeholders and decision-makers (the total fee (TF) and correct treatment rate (CTR)), we developed a new performance criterion (the adjusted total fee (ATF), which balances the TF and CTR) for the assessment and optimization procedure. Our results indicate that the parameter confidence interval (CI) of kriging has a relationship with the pollution rate (PR) of the field, and that a combination of SGS, kriging, and GA can greatly reduce the ATF compared with that derived solely from SGS. Optimal locations for additional sampling were also suggested that near the boundary between the polluted and unpolluted areas and at places where there are sampling data close to the threshold value; whereas the additional samples are generally dispersed over the full scale of the AAPF rather than concentrated in any one area, at last they are inside the area contoured by the suggested best CI range.

2. Feasibility and development of the cost-effective strategy on both the surface soil and the soil at depth (3D) (*Chapter4*). This section mainly review the interpolation methods (including the disadvantage and advantage of each method), which, when used in the 3D interpolation that only considering the statistical correlation of sampling data, are not enough to simulate the distribution of HMs in soils at depth, but the sampling data at depth may have the strong spatial correlation with the sampling data on the surface. As a result, we suggested the combination

of kriging estimation on surface soil and transport/migration model (functions) at depth to construct the 3D model. And at the end of this chapter, we discussed the existed transport functions and affecting factors to better understand the distribution or migration/transport of HMs in soil. In our future work, an actual polluted site will be selected to construct and validate such 3D model that combines kriging and transport or migration model (functions).

3. A case study in Kawauchi village, Fukushima, Japan (*Chapter5*). In this section, we select a real polluted site to apply the developed 3D model, which uses the kriging to estimate the distribution of ^{134}Cs on the surface soil and the transport/migration functions considering the transport/migration mechanism in the soil at depth. And finally, we improved the 3D radioactivity model by combining the kriging, compartment model, and conditional simulation. Through the results, we successfully constructed the 3D distribution of HMs in the forest soil in Kawauchi village. And based on the threshold value for decontamination, the corresponding remediation ratio that can be used to calculate the polluted soil volume was suggested for the engineer or the government officers.

At last (*Chapter 6*), the main conclusion was given that the cost-effective strategy for the investigation and remediation could well work both on the artificial polluted field and real-world polluted field. The results from the research can give a good reference for researchers, engineers, or government officers who are in charge of the project of the investigation and remediation of the polluted fields.

Table of contents

Acknowledgment	
Summary	i
Table of contents	v
List of Figures	ix
List of Table	xiii
CHAPTER 1 INTRODUCTION	1
1.1 Background	1
1.1.1 Soil pollution by HMs in China	2
1.1.2 Soil pollution by HMs in Japan	4
1.2 Objectives.....	7
1.3 The outline of thesis.....	9
CHAPTER II SELECTION OF THE METHODS FOR THE COST-EFFECTIVE STRATEGY: SAMPLING, REMEDICATION, GEOSTATISTICS AND GENETIC ALGORITHMS (GAS)	11
2.1 Investigation (sampling) method	11
2.1.1 Sampling in the soil at depth	11
2.1.2 Sampling design on the surface soil.....	11
2.1.3 In summary	14
2.2 Remediation methods.....	15
2.1.1 Excavation	15
2.1.2 Solidification/ Stabilization(S/S).....	15
2.1.3 Phytoremediation	16
2.1.4 In summary	16
2.2 Geostatistics.....	18
2.2.1 Kriging	18
2.1.2 Application and Related literature review	23
2.1.3 In summary	23
2.2 Genetic Algorithms (GAs).....	24
2.2.1 Initialization.....	24
2.2.2 Selection.....	25
2.2.3 Genetic operators	26
2.2.4 Termination.....	26

2.2.5 Application and Related literature review	27
2.2.6 In summary	30
2.3 Programming in VBA Excel.....	30
CHAPTER III COST-EFFECTIVE STRATEGY OF THE INVESTIGATION AND REMEDIATION AREA ON THE SURFACE POLLUTED-SOIL (2D).....	31
3.1 Abstract.....	31
3.2 Introduction	32
3.3 Materials and methods	35
3.3.1 Assumed Actual Polluted Fields (AAPFs) Generated by Unconditional Simulation (US).	35
3.3.2 Preliminary Grid Sampling	36
3.3.3 Remediation Method	37
3.3.4 Performance Criteria.....	37
3.3.5 Geostatistical Interpolation (Kriging Estimation).....	39
3.3.6 Optimization of Additional Samples by a Genetic Algorithm (GA) approach	39
3.4 Result and discussion	42
3.4.1 Preliminary Grid Sampling Analysis	42
3.4.2 Geostatistical (Kriging) Analysis.	46
3.4.3 Genetic Algorithm (GA) Optimization Analysis.....	51
3.5 Summary and future works	55
CHAPTER IV OVERVIEW OF THE EXISTED 3D MODEL AND DISCUSSION OF THE DEVELOPMENT FEASIBILITY OF THE COST-EFFECTIVE STRATEGY FOR THE SOIL AT DEPTH (3D).....	57
4.1 Introduction	57
4.2 Review of existed several literatures about 3D distribution of HMs in soil.....	57
4.2.1 Inverse Distance Weighted (IDW)	58
4.2.2 Natural Neighbors (N-N)	59
4.2.3 Kriging-three dimensional (kriging-3D) interpolation.....	60
4.2.4 Discussion and comparison of the three interpolation methods: IDW, N-N, Kriging-3D. ...	61
4.3 Feasibility of the development of 3D model: considering the migration or transport mechanism (function) of HMs at depth.	63
4.3.1 Mechanisms and pathways of the mobility of HMs in soil	63
4.3.2 Factors affecting the mobility and transport of HMs in soil	64
4.3.3 Existed transport modeling (or functions)	66
4.4 In summary and future work	67
CHAPTER V CASE STUDY IN THE POLLUTED FOREST SOIL (3D MODEL) IN KAWAUCHI VILLAGE, IN FUKUSHIMA PREFECTURE, JAPAN	69

5.1 Introduction	69
5.1.1 Fukushima Nuclear Dai-Ichi Power Plant (FNDPP) accident	69
5.1.2 Introduction of cesium	71
5.1.3 Migration of Cesium in the soil	72
5.2 Materials and methods	76
5.2.1 Study area	76
5.2.2 Sample collection and preparation	77
5.2.3 Radioactivity measurements by Gamma spectrometry	79
5.3 Results and discussion	81
5.3.1 Sampling data analysis based on the ^{134}Cs	81
5.3.2 Construction of the 3D ^{134}Cs -distribution model	86
5.3.3 3D-model developed and improved by conditional simulation (CS)	97
5.3.4 Assessment of the soil pollution based on the developed cost-effective strategy in Chapter3.....	102
5.3.5 Assessment of the soil pollution based on conditional simulation (CS)	104
5.4 In summary and future study.....	108
CHAPTER VI GENERAL CONCLUSION.....	103
6.1 Development of the cost-effective strategy on the surface soil (2D).....	109
6.2 Development of the 3D radioactivity-distribution model for the polluted forest soil: applied to a case study.....	110
REFERENCE.....	113

List of Figures

Figure 1.1 2000-2010 the output growth of 10 kinds of nonferrous metals.....	3
Figure 1.2 World urbanization prospects and rapid urbanization in China ¹¹	3
Figure 1.3 Increasing tendency of contaminated sites in Japan. Red color chart- Number of excess cases; Blue color chart-Number of non-excess cases.....	5
Figure 1.4 Number of contaminated sites exceeding the soil environmental standards in 2010	6
Figure 1.5 Number of contaminated sites exceeding the soil environmental standards from 1991-2010 (Reduplication of statistic data exists)	6
Figure 1.6 Zones for planned evacuation and emergency evacuation preparation ⁴⁶	7
Figure 2.1 Regulation of sampling depth of surface soil by Soil Contamination Countermeasures Act.....	11
Figure 2.2 10m*10m sampling method (10-method) ⁴⁸	12
Figure 2.3 5-point mixture sampling method (5-method) ⁴⁸	13
Figure 2.4 5-point mixture sampling method developed by M.Yoneda (5Y-method)	14
Figure 2.5 Traditional treatment (Excavation) of polluted soil by HMs	17
Figure 2.6 Curve of relationship between fee and sample numbers.....	17
Figure 2.7 Variogram	20
Figure 2.8 Flowchart of Ordinary Kriging.....	22
Figure 2.9 Flowchart for Genetic Algorithms	25
Figure 2.10 Some factors to take into account when the initial population is generated randomly ⁹⁵	26
Figure 3.0 Abstract.....	31
Figure 3.1 AAPFs randomly generated by US. Red area: polluted soil ($\geq$ Threshold value); Blue area: unpolluted soil ($<$ Threshold value))	36
Figure 3.2 Grid sampling on an AAPF. n: sample or mesh number of each side of the square field; Δ : sampling point at the center of each mesh	36
Figure 3.3 One assumed remediation scenario on an AAPF. APA: the Actual Polluted Area of the AAPF (red area); ARA: the Actual Remediated Area of the AAPF, (green area); CRA: the Correctly Remediated Area of the APA of the AAPF (area where the red and green areas overlap)	39
Figure 3.4 Flowchart of the GA process used in this study.....	41
Figure 3.5A Relationship between the ATF and sample number determined by summarizing 100 randomly generated AAPFs while changing the values of the sill. Figures (b), (d), and (f) are partial enlargements around the turning points of figures (a), (c), and (e). μ is the average value and σ is the corresponding standard deviation.....	43
Figure 3.5B Relationship between the ATF and sample number determined by summarizing 100 randomly generated AAPFs while changing the values of the correlation scale. Figures (b), (d), and (f) are partial enlargements around the turning points of figures (a), (c), and (e). μ is the average value and σ is the corresponding standard deviation	44

Figure 3.5C Relationship between the ATF and sample number determined by summarizing 100 randomly generated AAPFs while changing the sample price (SP). Figures (b), (d), and (f) are partial enlargements around the turning points of figures (a), (c), and (e). μ is the average value and σ is the corresponding standard deviation.....	45
Figure 3.6 (a) One randomly generated AAPF; (b) Histogram of concentrations on the randomly selected AAPF; (c) Normal plot of concentrations (Log); (d) Sample semivariogram for concentrations (blue line) and the exponential model (red line) fitted to the sample semivariogram; (e) Normal plot of residuals (ϵ s) in kriging; (f) Sample semivariogram for the residuals (ϵ s) in kriging; (g) kriging estimation of pollution distribution (CI=75%); (h) Kriging Standard Deviation (SD) of 121 preliminary grid sampling data	47
Figure 3.7 (a) Relation between PR and the corresponding best CI used in kriging; (b) Relation between PR and SDPR; (c) Relation between SDPR and the corresponding best CI used in kriging.	48
Figure 3.8 The different ATFs (A) and corresponding CTRs (B) for each of the 150 AAPFs. Blue line: ATF (CTR) for the preliminary grid sampling method. Green line: ATF (CTR) after kriging estimation by selecting the best CI value for each AAPF. Red line: ATF (CTR) after kriging estimation by selecting the best average CI value. The AAPF number is ranked in ascending order of PR or SDPR (larger AAPF number, with a larger PR or SDPR value).	50
Figure 3.9 ATFs (a) and CTRs (b) after GA optimization on 5 candidate areas (CI Ranges: 75-99%) of 10 randomly selected AAPFs (SDPR Range: 14-24%). μ is the average value and σ is the corresponding standard deviation.	53
Figure 3.10 Comparison of the ATFs (A) and CTRs (B) calculated by different combinations of preliminary sampling data, kriging estimation, and GA optimization. The 121 and 144 in the legend are the preliminary numbers of samples and 23 is the additional number of samples. The AAPF numbers from 1 to 50 in 10 increments correspond to the SDPR ranges of 4-14%, 14-24%, 24-44%, 44-64%, and 64-84%, respectively.....	54
Figure 3.11 Distribution of 121 preliminary samples (red and blue bubbles) and 23 additional samples (green bubbles) on the AAPF used in Figure S6. The sizes of the red and blue bubbles represent the concentrations of samples above and below the threshold value, respectively, while the green bubbles only represent the additional sample location without considering the size.	54
Figure 4.1 One example to understand the IDW original function. ¹⁷³	59
Figure 4.2 Variograms in horizontal and vertical direction. $\gamma_1(h_1)$ is shown as the red line, while $\gamma_2(h_2)$ is shown as the blue line.	61
Figure 4.3 Natural neighbour interpolation for irregularly spaced data. The crosses represent observation points and the triangle represents a prediction point. The predicted value at 0 depends on the observed value at 1 but not at 2. ¹⁶² Although the advantage of N-N interpolation to deal with the scattered and irregularly spaced data, its implementation is still a problem in practice, especially in three and higher dimensions. ¹⁷⁷ Compared with the IDW and N-N, kriging-3D to give the most	

accurate results and provides the additional benefit of variography that aids in the interpreting the site characteristics. ¹⁶³	62
Figure 4.4 Schematic of the rhizosphere, showing the various exudates and how they can influence abiotic factors and mechanisms in the soil solution interface. Legends: OC=organic carbon; C^+ = cation; A^- = anion; L= ligand; pe = redox potential. ¹⁸⁴	64
Figure 5.1 Distribution map of ^{134}Cs inventory obtained using in situ gamma spectrometry using b of $1.30 \text{ g}\cdot\text{cm}^{-2}$ (b is the relaxation mass depth). Activities were decay-corrected to March 1, 2012. ¹⁸⁸	70
Figure 5.2 Estimated lifetime attributable risk (LAR) for a leukemia, b all solid, c thyroid, and d breast cancers for adult age group. ¹⁸⁹	71
Figure 5.3 Representation of the i th compartment corresponding to the soil layer situated between the x_{i-1} and x_i depths. C is the radionuclide concentration in the i th compartment. The radioactivity input and output in this compartment are denoted k , which represents the intercompartment fractional transfer rate, and λ is the radioactive decay constant. ²⁰³	75
Figure 5.4 Sampling location in Kawauchi village, Fukushima prefecture, Japan	77
Figure 5.5 The investigation of the polluted forest field and the sampling plan	79
Figure 5.6 Core sampling by the auger and the section description of one example soil layer	79
Figure 5.7 The soil sample that was put in the sealed plastic bag	79
Figure 5.8 The soil sample and the plastic bag were put in the plastic box before the detection of the radioactivity	79
Figure 5.9 The equipment used to detect the radioactivity of the soil samples	80
Figure 5.10 The relationship between the counts of ^{134}Cs and ^{137}Cs in the litter layer	82
Figure 5.11 The distribution of ^{134}Cs at different layers based on the radioactivity per unit area (Logarithm data) (Bq/cm^2). The size of the circle represent the radioactive strength, while the red color circle (Log data ≥ 0) is positively proportional to the radioactive strength and the circle without color (Log data < 0) is inversely proportional to the radioactive strength. OL-organic layer; SL-soil layer.	82
Figure 5.12 The distribution of ^{134}Cs at different layers based on the radioactivity per unit cubic (Bq/cm^3). The size of the circle represent the radioactive strength, while the red color circle (Log data ≥ 0) is positively proportional to the radioactive strength and the circle without color (Log data < 0) is inversely proportional to the radioactive strength. OL-organic layer; SL-soil layer.	84
Figure 5.13 The relationship between the average radioactivity (Log, Bq/m^3) of the total 107 sampling data and the soil depth (cm)	85
Figure 5.14 The relationship between the average radioactivity (Log, Bq/m^3) of the 18 sampling data and the soil depth (cm)	86
Figure 5.15 The relationship between the average radioactivity (Log, Bq/m^3) of 12 sampling data and the soil depth (cm)	86
Figure 5.16 Conceptual scheme of the compartment model	88
Figure 5.17 Location and size of the sampling data (Log, unit: Bq)	91

Figure 5.18 Semi-variogram for sampling points. Blue-points is the experimental variogram, red-points is the model variogram.	92
Figure 5.19 Semi-variogram for residues.....	92
Figure 5.20 Kriging estimation of pollution distribution (Log) (CI=50%) on the polluted forest field (50m*50m).	93
Figure 5.21 The standard deviation (SD) for the Kriging estimation of pollution distribution (Log) (CI=50%) on the polluted forest field (50m*50m).	93
Figure 5.22 The relationship between two transport parameters (Po and Ps) of ^{134}Cs	94
Figure 5.23 (a) Histogram of Po, x-axis is distance interval (m), y-axis is the frequency; (b) Sample semivariogram for concentrations (blue points) and the exponential model (red points) fitted to the sample semivariogram; (c) Sample semivariogram for the residuals (ε s) in kriging; (d) kriging estimation of pollution distribution (CI=50%); (e) Kriging Standard Deviation (SD).	95
Figure 5.24 (a) Histogram of Ps, x-axis is distance interval (m), y-axis is the frequency; (b) Sample semivariogram for concentrations (blue points) and the exponential model (red points) fitted to the sample semivariogram; (c) Sample semivariogram for the residuals (ε s) in kriging; (d) kriging estimation of pollution distribution (CI=50%); (e) Kriging Standard Deviation (SD).	96
Figure 5.25 3D distribution of ^{134}Cs by kriging estimation (CI=50%)(Left figures) on different layers and the distribution of polluted soil that needs to be remediated (red area except the area in purple colour in the Figure 5.21, right figures) when threshold value=25Bq.....	98
Figure 5.26 The distribution of polluted soil that needs to be remediated (red area except the area in purple colour in the Figure 5.21) when threshold values are set as 50 (left figures) and 100Bq (right figures) respectively.	99
Figure 5.27 The average radioactivity (blue bar chart) and compartment model error (red bar chart) of total 107 sampling data in different layers.....	100
Figure 5.28 Real, simulated and estimated profiles. Blue line is reality (true profile); Black bold line is kriging estimation; Green dotted line is conditional simulation; Red points are conditioning data (observed data).....	101
Figure 5.29 The remediation area (in red color except the area in purple colour in the Figure 5.21) based on the Table 3 and the cost-effective strategy in Chapter 3. Thr-the threshold value; SDPR-pollution rate reflected by sampling data.	103
Figure 5.30 The distribution of radioactivity strength at different soil layers based on the calculation of CS (1000 realizations and CCDF=50%) and compartment model.....	105
Figure 5.31 The distribution of radioactivity strength at different soil layers based on the calculation of CS (1000 realizations and CCDF=95%) and compartment model.....	106

List of Table

Table 1 Phytoremediation includes the processes and mechanisms of contaminant removal	16
Table 2 The application of GAs introduced by part literatures	27
Table 3 The best CI values of kriging suggested for different SDPR ranges.	49
Table 4 Effects of Soil Factors on Trace Metal Mobility and/or Bioavailability ^{183,185}	65
Table 5 Basic information of Cs. ¹⁹⁴	72
Table 6 Radioactive nuclide half-lives and decay constants. ⁴⁴	74
Table 7 The correlation of radioactive strengths at different layers.	83
Table 8 The correlation coefficient between Po and Ps.....	95
Table 9 The mean and SD of the compartment model errors on different layers	97
Table 10 The remediation ration based on the Table 3 and the cost-effective strategy in Chapter 3.	104
Table 11 The remediation ratio at different soil layers based on the different threshold values (Thr) and different CCDF values.	106
Table 12 The comparison of polluted soil volume (PSV) by different three methods.	108

CHAPTER 1 INTRODUCTION

1.1 Background

Heavy metals (HMs) or trace metals normally exist in nature, and in small quantities are nutritionally essential to life. However, HMs can also cause serious environmental pollution and greatly threaten the health of animal and human when the concentration of HMs in soil, water or atmosphere exceeds a certain threshold value. Due to the non-biodegradable and persistent nature, HMs are accumulated in vital organs of the human body via the food chain and exhibit specific signs of their toxicity¹. For example, chronic exposure to Cd can have adverse effects such as lung cancer, pulmonary adenocarcinomas, prostatic proliferative lesions, bone fractures, kidney dysfunction, and hypertension²⁻⁵. Since there is no good mechanism to eliminate the HMs from the body⁶, the toxicity resulting from the continuous and excessive exposure poses a grave risk factor to human health as pathological changes have been observed in the kidney, liver, gastrointestinal tract, bone, pancreas, testes, blood vessels etc⁷.

The total amount of HMs does not prove anthropogenic pollution since high accumulation of HMs might also result from natural processes⁸. However, their occurrence and accumulation in the environment directly or indirectly result from human activities, such as rapid industrialization, urbanization and anthropogenic sources⁹. The environmental pollution is particularly attributed to industrial activities such as chemical works, service stations, metal fabrication shops, paper mills, tanneries, textile plants, waste disposal sites and intensive agriculture¹⁰. Recently, the rapid urbanization accelerated the industry and economic activities, and caused more resource consumption and chemical emissions¹¹. Anthropogenic sources such as agriculture, manufacturing, mining and the land application of sewage can also contaminated the soil and water¹². Because of the development of economy and technology and the explosion of world population, a large range and amount of HMs are needed to meet our consumption for all aspects of modern daily life⁷. As a result, we are facing with serious pressure of environmental pollution by HMs. For example, e-waste contains a huge number of toxic substances, including plastics and HMs such as lead(Pb), nickel(Ni), chromium(Cr), cadmium(Cd), arsenic(As) and mercury(Hg),

and according to the estimation of Robinson(2009), the global generation of e-waste is 20-25 million tons per year¹³. In most parts of world, large quantities of HMs are directly discharged to nearby land and into surface water¹⁴, and intake of HMs by human via the food chain has been widely reported throughout the world¹. At least 20 metals are classified as toxic substances, half of which have emitted into environment and posed great risks to human health⁹.The most common HMs which can cause environmental pollution are Pb, Cr, As, Cd, Ni, zinc (Zn), copper (Cu), Hg, silver (Ag), and selenium (Se). All these elements are selected for discussion in Environmental Protection Agency(EPA)¹⁵.

Soil plays a principal role in biochemical transformations, the cycling of elements, filtration of water, supporting plants and infrastructure, and many recreational activities¹¹. The quality of soil relates to the water and food safety¹⁰, therefore, soil pollution is getting more worldwide concern. Studies of HMs pollution in soil have included the sources and behavior of HMs, impacts on public health and the environment, site investigation and analysis, management of remedial action and risk assessment, and the techniques for remediation and so on¹⁶. Among the most common HMs, Zn and Cd show the greatest mobility in the soil environment, therefore are of greater concern¹². In the following content, the situation of soil pollution will be discussed through two examples: the developing country-China and the developed country-Japan.

1.1.1 Soil pollution by HMs in China

In recent thirty years' rapid development of industrialization¹⁷ (**Figure 1.1**) and urbanization¹¹ (**Figure 1.1**) partially at price of environment, China has had the serious soil pollution by HMs. In addition, with the largest population of the world, China is still facing with large pressure of soil pollution by HMs in future years. China's 12th Five-Year Plan on environmental protection has urged efforts to carry out the cleansing of polluted soil by HMs¹⁸. Industry emission, wastewater irrigation and waste fertilization are the origin of the polluted soil by HMs¹⁹.

In recent years, the numerous reports of heavy-metal-pollution accidents have garnered headlines of local and federal media and aroused more and more public concerns in China²⁰. In 2011, about 50 million tons of Cr waste was found, and these Cr waste has existed in the six cities of Henan province for decades; besides more than 600 million tons of waste has existed in more than twenty metropolises had reported²¹. At the same year, EPB (Environmental Protection Bureau) of Wuhan city, Hubei province, figured out that the polluted area in Wuhan city was more than 200 acres with

pollutants mainly including Hg, Pb, As, Cd, Cr. In addition, there were other five polluted areas with the polluted area more than 1260 acres²².

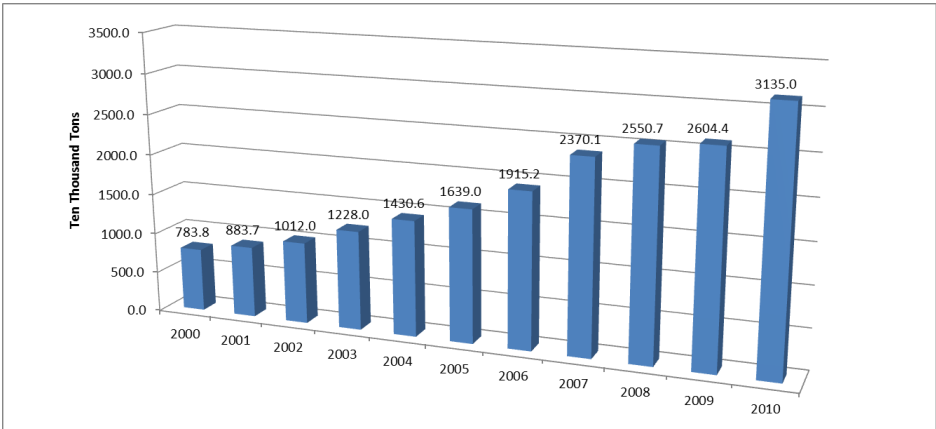


Figure 1.1 2000-2010 the output growth of 10 kinds of nonferrous metals

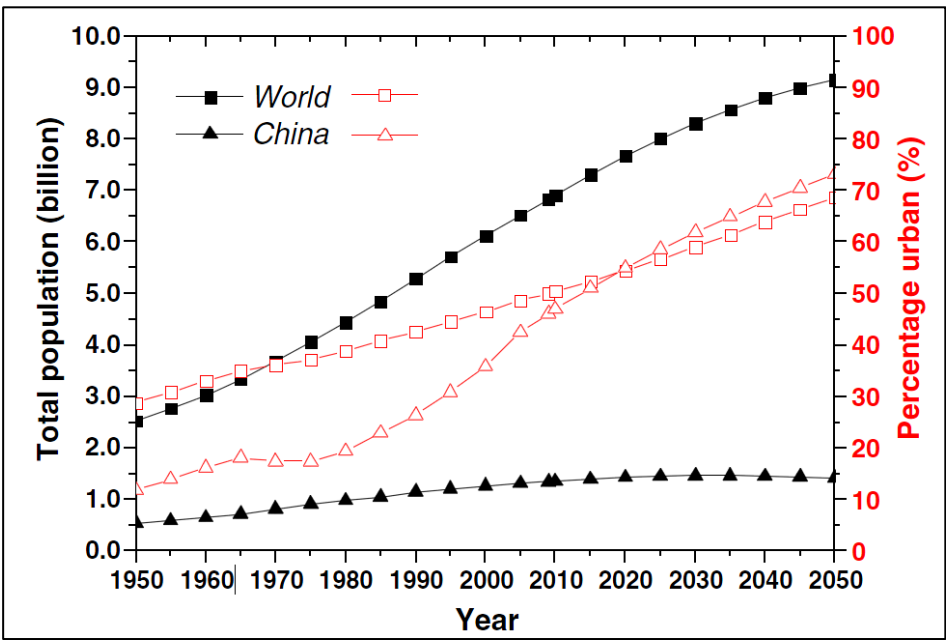


Figure 1.2 World urbanization prospects and rapid urbanization in China¹¹

Many researches of soil pollution by HMs have been done on the agriculture area and urban area in China.^{11,23-31} According to numerous studies, soil heavy metals pollution were mainly cause by anthropogenic sources. The dominant sources of heavy metals

pollution to agricultural soils included sewage irrigation, mining, smelting, waste disposal, urban effluent, vehicle exhausts, sewage sludge, pesticides, fertilizers application and so on, while the urban soil were industrial activities, traffic emissions, fuel combustion, and waste disposal^{2,32,33}. In fact, national investigation of soil heavy metals including approximately 32 urban soil contamination and geochemical surveys or projects has been conducted in China since the 1990s; up to 2012, the survey has already covered an area of 1,650,000 km² and collected high-quality data for the assessment and remediation of trace metal contamination in urban soils in China^{34,35}. Based on the latest survey report released on 17th April, 2014, 16.1% of the Chinese soil was polluted and 82.8% of the polluted land was contaminated by toxic inorganic pollutants and the HMs such as cadmium(7.0%), nickel(4.8%) , arsenic(2.7%), copper(2.1%), mercury(1.6%), lead(1.5%), chromium(1.1%) and zinc(0.9%) contributed to the most common pollutants.³⁶ In comparison with agricultural soils, urban soil are not used for crop products and generally receive lower doses of fertilizers and pesticides, however the HMs in urban soils can be easily transferred into humans through ingestion, inhalation, or dermal routes, etc. ³¹³² In China's major cities, there are more than 5000 brownfield sites left polluted or potentially polluted, and more brownfield sites will be left according to the recently released ambitious urbanization plan that will move more polluting plants from cities. ³⁷ Remediation of polluted soil by HMs not only improve the quality of soil to keep the water and food safe to the human health, but the land after remediation also has the large potential economic or commercial value. In effort to curb soil pollution, the Chinese central government has totally invested over 100 billion RMB (16 billion USD) for the soil environment protection in the 12th Five Year Plan (2011-2015) and the funding for the soil remediation will continuously grow in the long term.³⁸

1.1.2 Soil pollution by HMs in Japan

In Japan, rapid industrialization in 1960s caused dangerous levels of soil pollution by HMs especially by Cd which is known to cause “itai-itai” disease; Because of these issues, the Agricultural Land Soil Pollution Prevention Law was established in 1970 to regulate HMs pollution³⁹. In 1971, the Agricultural Land Soil Pollution Prevention Law came into force, and thereafter, the soil dressing method was implemented to deal with polluted soil of

paddy fields with the Cd concentration 1.0 mg kg⁻¹ or higher, and nowadays there are few new cases of soil pollution by Cd from mines and smelters because of strict regulations and management⁴⁰. In 2003, the Soil Pollution Control Act was enacted in Japan (**Figure 1.3**), and contaminated soil started to become an issue in the development of old factory sites and redevelopment of urban areas⁴¹. However, soil pollution issues did not decrease; conversely, they had the increasing tendency (**Figure 1.3**). As of 2011, 87.8% of the total polluted land (7,575 ha), designated by the Agricultural Land Soil Pollution Prevention, was remedied by applying uncontaminated soil and/or replacing the soil.⁴² From the number of contaminated sites exceeding the soil environmental standards (**Figure 1.4** **Figure 1.5**), we can see that the top three of the second type pollutants(HMs etc.) are Pb, Fluorine, As and their compounds, respectively⁴³. On March 11, 2011, the Tohoku earthquake and the subsequent tsunami resulted in the Fukushima Dai-Ichi Nuclear Power Plant (FDNPP) accident, which was the largest release of nuclear contamination since the Chernobyl disaster of 1986, and cause serious release and pollution of about 73 radionuclides (135 in total including radioactive progeny), such as ¹³⁷Cs, ¹³⁴Cs, ¹³¹I.^{44,45} On April 22, authorities declared the area of “Exclusion zone”, “Emergency evacuation preparation zone” and “Planned evacuation zone”, which are within 20km, 20-30km and 20-about 50km from the FDNPP, respectively (**Figure 1.6**).⁴⁶ The highest activities are concentrated within a 70 km long radioactive plume where initial ¹³⁷Cs contamination exceeded 300 kBq m⁻² covering an area of 3000 km².⁴⁷

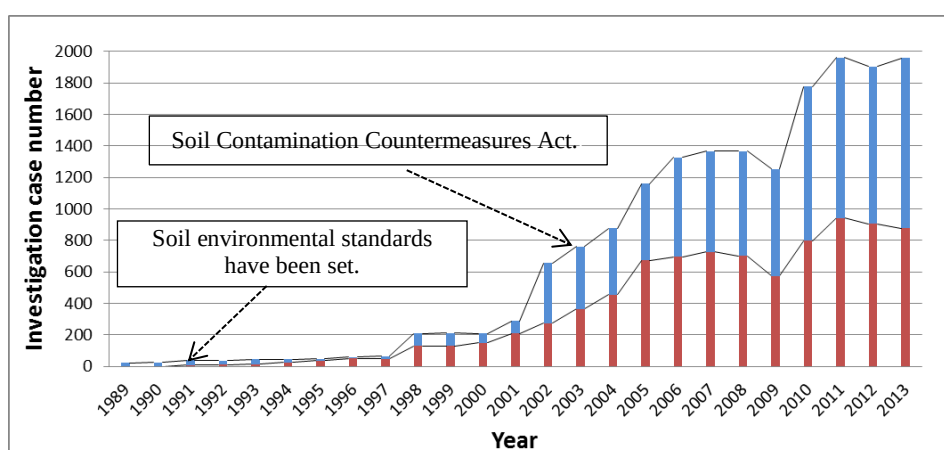


Figure 1.3 Increasing tendency of contaminated sites in Japan. Red color chart- Number of excess cases; Blue color chart-Number of non-excess cases

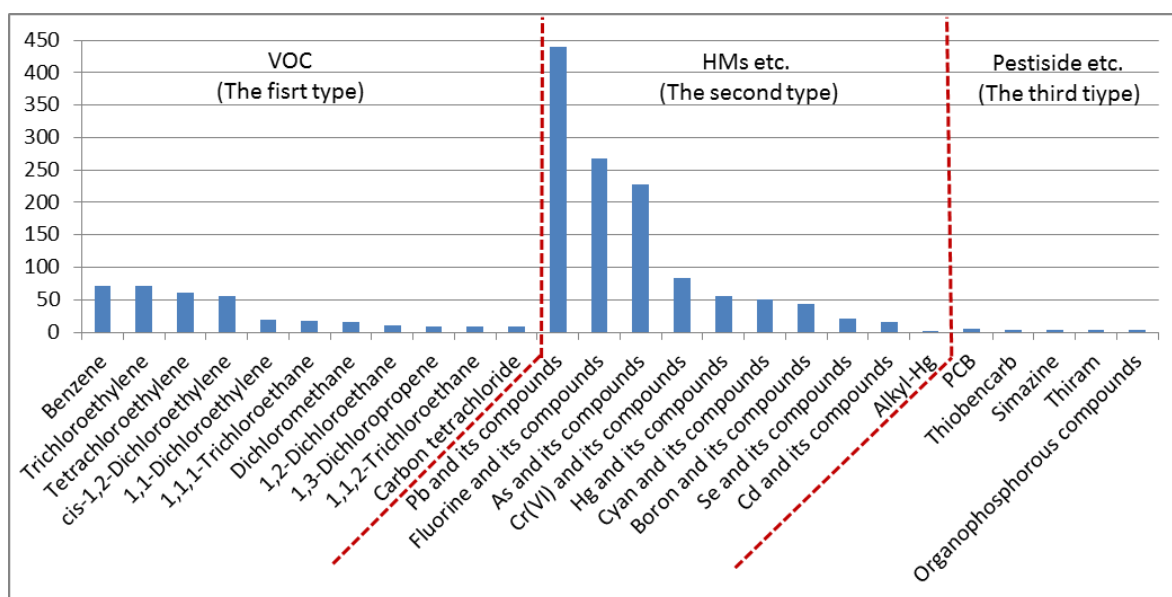


Figure 1.4 Number of contaminated sites exceeding the soil environmental standards in 2010
(Reduplication of statistic data exists)

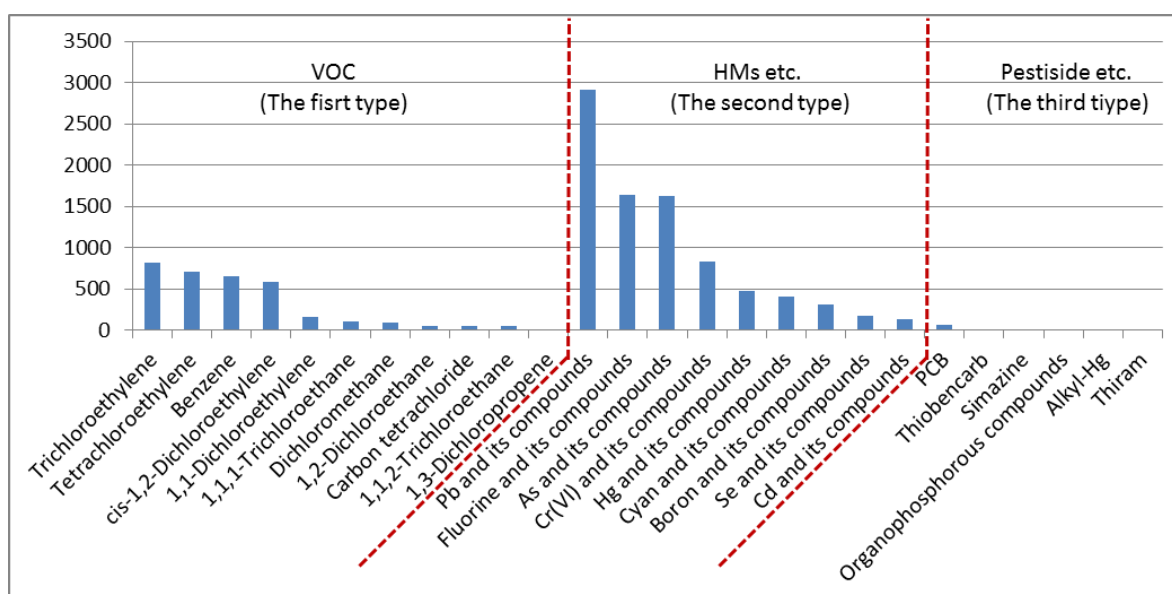


Figure 1.5 Number of contaminated sites exceeding the soil environmental standards from 1991-2010 (Reduplication of statistic data exists)



Figure 1.6 Zones for planned evacuation and emergency evacuation preparation⁴⁶

1.2 Objectives

Based on the background that is previously mentioned, both in China and Japan, even in other developing or developed countries, there existed large sums of polluted area. It will be a serious economic or financial burden, if we investigate and treat the polluted area using traditional methods; as a result, developing a convincing and cost-effective strategy is essential and imperative to solve the real engineering problems confronted in the real-world polluted sites like how many and where the samples should be taken, what kind of remediation method should be used, and where the additional samples should be taken, etc.

To develop such strategy, this study combines the systematic (or uniform) grid sampling (SGS), excavation remediation, geostatistical technique (or kriging) and genetic algorithm (GA) together into the whole strategy of investigation and remediation of polluted field. In addition, to address the issues that most directly concern stakeholders and decision-makers (the total fee (TF) and correct treatment rate (CTR)), we developed a new performance criterion (the adjusted total fee (ATF), which balances the TF and CTR) for the assessment and optimization procedure. In the validation progress of the strategy, we use large numbers of assumed actual polluted fields (AAPFs), instead of real polluted sites, generated by unconditional simulation (US), to assess the ATFs when we change the parameters of strategy, like sampling number and price, confidence interval (CI) of kriging. At last, based on the ATFs, we decide the optimal or cost-effective strategy for the investigation and remediation of the polluted fields.

However, after the development of the strategy, we have to bring and apply this strategy into the real polluted site to check the accuracy and efficiency and to know if it works well. So the pollute forest soil in Kawauchi village in Fukushima Prefecture, which has been affected by the radionuclides from the accident of FDNPP, was selected to assess the developed strategy. In view of the developed strategy only considering two dimensional (2D) or surface soils, the distribution of HMs (mainly ^{134}Cs , ^{137}Cs) in the soil at depth was also studied to construct the three dimensional (3D) model. That means, firstly, to use the SGS and kriging to estimate the distribution of HMs on surface soil, and secondly, to simulate and assess the distribution of HMs at depth through considering the transport or migration functions of HMs in the vertical direction. Based on both distributions on the surface and at depth, the 3D visible distribution of HMs in the soil can be constructed and the corresponding total volume of polluted soil can be also known well. It will be useful for the decision-makers or engineering to operate the remediation, like where and how much the polluted soil should be removed. This developed and validated strategy will demonstrate the advantage of integrating kriging into the practice of soil investigation and remediation for engineers and decision-makers and has a high reference value when used in the other real-world polluted sites.

1.3 The outline of thesis

- *Chapter 1 Introduction*

In this chapter, the background, especially the situations of soil-pollution by HMs in China and Japan, is described to motivate this research to be carried out. The main objectives of this research are also given: 1) to develop a cost-effective strategy to investigate and remediate the polluted site with the respect of the total fee (TF) and correct treatment rate (CTR); 2) to apply the developed strategy into the investigation and remediation at the real-world polluted site, and provide a reference for the engineers, decision-makers or researchers.

- *Chapter 2 Discussion and selection of sampling, remediation and geostatistical methods*

In this chapter, the sampling, remediation and geostatistical methods, which are the main components of the strategy are introduced and discussed in detail. At last, the systematic (or uniform) grid sampling (SGS), excavation remediation, geostatistical technique (Ordinary Kriging: OK) and a genetic algorithm (GA) approach are selected through the comparison of the advantages or disadvantages of each method.

- *Chapter 3 Cost-effective strategy of the investigation and remediation area on the surface polluted-soil (2D)*

This chapter explained, based on the methods selected in the chapter 2, that how the strategy has been developed and validated. Firstly, a large sum of assumed actual polluted fields (AAPFs) are randomly generated by the unconditional simulation (US), and on these AAPFs the systematic grid sampling (SGS) were used to obtain the sampling data. Secondly, through the sampling data and threshold value, we can judge that which part of soil is polluted and should be removed, and then we can further know if the removed part is real polluted soil or the retained part still has some

polluted soil not removed, since the concentration of HMs at each location of AAPF is known beforehand. After calculating the corresponding ATFs for the SGS with different sample numbers, the optimal sample number can be suggested to get lower ATF. Thirdly, after the optimal sample number is fixed, the kriging method can be used to calculate the corresponding ATF, and compare the ATF by kriging and the ATF by only SGS to judge which is better. In this research, the ATF by kriging is lower than by only SGS, and change the confidence interval (CI) of kriging to know which CI has the lowest ATF. At last, based on the best CI decided in last step, use the GA to optimize the additional sampling locations.

- *Chapter 4 Development of the cost-effective strategy for the soil at depth (3D)*

In this chapter, we discussed the difference and the (dis-) advantage of existed 3D model to simulate the distribution of HMs in the polluted soil, and analyze the feasibility of the development of the 3D model in this research. The distribution of HMs on the surface soil takes the method of kriging estimation, and at depth uses transport or migration functions in the vertical direction.

- *Chapter 5 Case study in the polluted forest soil in Kawauchi village, in Fukushima prefecture, Japan*

In this chapter, we tried to construct the 3D model (the distribution of HMs both on the surface soil, but also at depth) on the real-world polluted site to assess the volume of the polluted soil, and give the suggestions of remediation, for example the volume the polluted soil, the total fee, to decision-makers.

- *Chapter 6 Discussion and general conclusion*

In this chapter, all the important findings are summarized and emphasized. Finally, the future plan and the opportunity to improve the model developed in this research are described.

CHAPTER II SELECTION OF THE METHODS FOR THE COST-EFFECTIVE STRATEGY: SAMPLING, REMEDIATION, GEOSTATISTICS AND GENETIC ALGORITHMS (GAS)

2.1 Investigation (sampling) method

2.1.1 Sampling in the soil at depth

In Japan, the depth of sampling is regulated by the law of Soil Contamination Countermeasures Act ⁴⁸. In detail, the surface soil is divided into two parts by depth: topsoil (0~5cm) and subsoil (5~45cm). Subsoil can further be uniformly divided into several smaller parts. Take a certain volume soil from each smaller part and mix them together to one. Then the mixed soil from these smaller parts mixes again with the same amount of soil from topsoil, and at last get the final sample of surface soil (**Figure 2.1**).

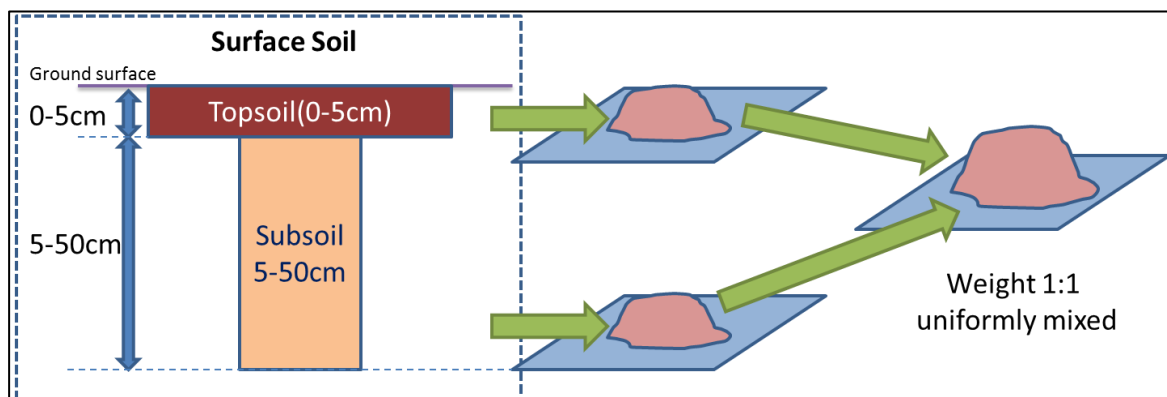


Figure 2.1 Regulation of sampling depth of surface soil by Soil Contamination Countermeasures Act

2.1.2 Sampling design on the surface soil

There are mainly three kinds of sampling designs on the surface soil:

- 10*10m method (in this research, the method is called “10-method” in short)

When the pollution probability is relatively high, usually use the 10 method to divide the area into 10*10m square meshes, and then take one sample for every one mesh. At last, based on the sampling data and the guideline, make a decision whether or not to remove the soil in the mesh **Figure 2.2**.

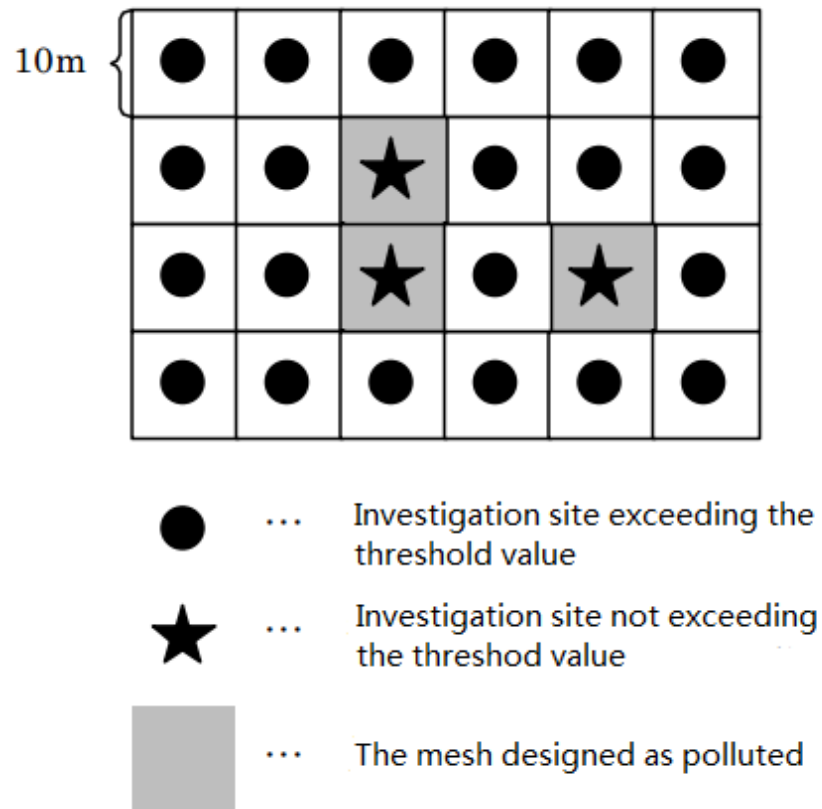


Figure 2.2 10m*10m sampling method (10-method)⁴⁸

➤ 5-point mixture method(in this research, the method is called “5-method” in short)

When the pollution probability is relatively low, usually use the 5-method to decide the sampling positions on topsoil. Firstly, divide the square area (30m*30m) into 9 uniform meshes, and take 5 samples and mix them together to get the final mixed sample; then based on the data of mixed sample, we can judge removing the soil of the whole square area (30*30m) or not (**Figure 2.3**).

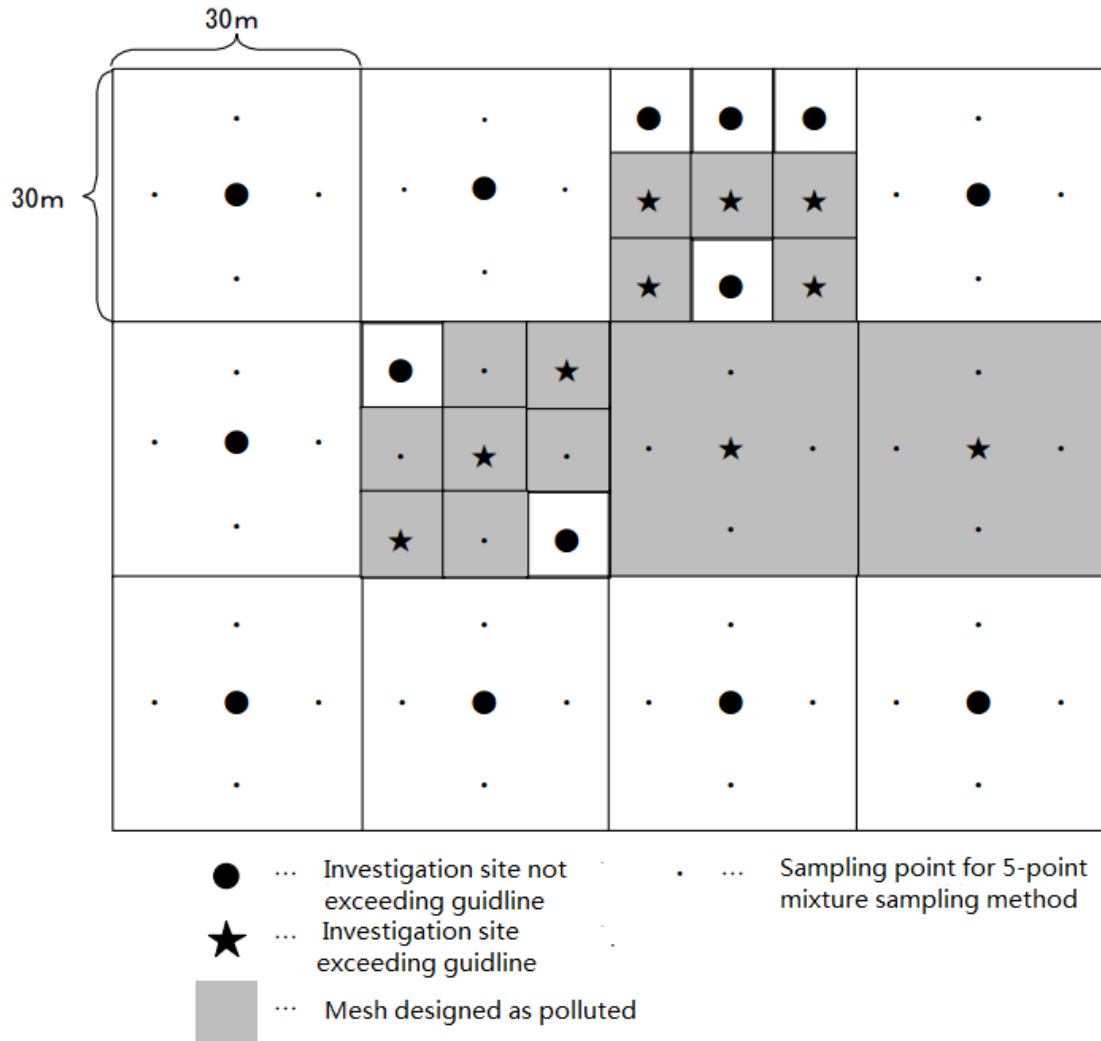


Figure 2.3 5-point mixture sampling method (5-method) ⁴⁸

When doing the further samplings, two choices can be selected.

- A) Directly using 10-method to do further sampling on area to be needed like **(Figure 2.3)** (the square of row 1, column 3).
 - B) Do further samplings in the 4 meshes in which no sample was taken as shown in **Figure 2.3** (the square of row 2, column 2).
- 5-point mixture method developed by M.Yoneda **(Figure 2.3)** (in this research, the method is called “5Y-method” in short)

M.Yoneda assumed a field with stochastic characters of weak stationarity, isotropy and three types of covariance functions, and developed a new sampling design of 5-point mixture sampling method through analyzing the stochastic characteristics and minimizing the estimation variance by Genetic Algorithm (GA). The results show that the sampling points should be on the center and about 13m apart from it on the diagonal line. If the square area is not 1000m^2 , the function $R=13*\beta/31.6$ (β is the length of square side, in this research $\beta=30\text{m}$) can be used to calculate the distance R ⁴⁹.

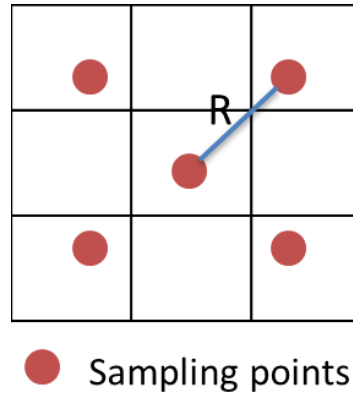


Figure 2.4 5-point mixture sampling method developed by M.Yoneda (5Y-method)

2.1.3 In summary

Except the three kinds of sampling designs, there also exist other sampling methods. Among these three sampling designs, the $10*10\text{m}$ method is extensively used, since in the real-world polluted site, it is hard to judge where is heavily polluted and where is lightly polluted to use the 5-method. Of course, if one polluted field has the background documents or has the distribution data of pollution, taking the 5-method is acceptable. The 5Y-method, though it may have been proved better than 5-method, is not used popularly, maybe because the complicated operation in practical project issues. In this research, the $10*10\text{m}$ -method is used, and changing the mesh size can result in different sample numbers.

2.2 Remediation methods

Since the contaminated soil by HMs is getting more worldwide concern, the remediation of contaminated soils by HMs is still recognized nowadays to be one of the most difficult problems to be solved even if few, albeit costly, technologies may be employed⁵⁰. Various techniques of in-situ and ex-situ have been used to reduce the impact of metals in the soil, including excavation, solidification/stabilization, soil washing, electro-remediation and phytoremediation⁵¹. In-situ methods mean remediation without excavation of contaminated soil, while ex-situ techniques require removal of the contaminated soil for treatment on- or off-site⁵². Some remediation methods mentioned above are detailed as follows, others please refer to⁵³.

2.1.1 Excavation

Excavation of soil contaminated by HMs and disposal to a secure landfill are very common technology. Excavation, which physically moves away the contaminated soil, can completely remove the contaminants and rapidly clean up the contaminated site, and is highly effective at low risk. However, the method just simply moves the contaminants to a different place, so it needs further monitoring and treatment; otherwise, other attention should be paid on the risk of spreading of contaminated soil and dust during removing and transporting. In addition, the method is always related to high cost.⁵⁴

2.1.2 Solidification/ Stabilization(S/S)

Solidification physically encapsulates the contaminants in a solid matrix, while stabilization uses chemical reactions to form minerals containing HMs in a form that is less soluble, immobile, toxic, and not easily absorbed by plants, animals and human being^{53–55}. Wide ranges of fixing additives including organic, inorganic and combined organic-inorganic materials are summarized in detail in literature⁵⁶. However, some metals such as As, Cr (VI), and Hg are not suitable for using the remediation method, because they do not form hydroxides of poor solubility⁵⁵. Also, the remediation methods

based on S/S technology are no longer considered as a permanent environmental solution because of some certain disadvantages⁵⁷.

2.1.3 Phytoremediation

Definition of Phytoremediation derived from (“phyto” meaning plant, and the Latin suffix “remedium” meaning to clean or restore), in which growing plants are used to remove, transfer, stabilize and/or degrade contaminants in soil, sediment and water^{58,59}. Phytoremediation has five main progresses or subgroups (**Table 1**), which has been introduced and discussed in detail in several literatures^{60–62}. Many examples of hyper-accumulators and their bioaccumulation potential or accumulation characteristics are given in literature⁵⁸. The method is considered as relatively low cost (less than a quarter of the cost of excavation or in situ fixation) and get wide public acceptance, though it is time-consuming⁵⁴. Please refer to⁶² for detailed advantages and disadvantages/limitation.

Table 1 Phytoremediation includes the processes and mechanisms of contaminant removal

Process	Mechanism	Contaminant
Rhizofiltration	Rhizosphere accumulation	Organics/Inorganics
Phytostabilisation	Complexation	Inorganics
Phytoextraction	Hyper-accumulation	Inorganics
Phytovolatilization	Volatilisation by leaves	Organics/Inorganics
Phytotransformation	Degradation in plant	Organics

2.1.4 In summary

In Japan, moving all possibly polluted soil (Excavation method) is always used because this can almost remove all of the total contaminants without spending too much time. The method also doesn’t introduce other chemicals and has no need for doing further monitoring and treatment on site. As a result, the method reduces the risk of the contaminants in the soil spreading to other soil, water or the plants, though the cost is high. If we can use certain methods or techniques to reduce the cost, the method can become an ideal one in actual polluted soil remediation.

In fact, no matter which techniques or methods are used in the remediation, if we know better about the concentration and spatial distribution of HMs in soil, it will be of great help for risk assessment and remediation decision making. Regarding excavation, for example, if the distribution of contaminants in soil is not identified clearly, all of the possibly contaminated soils have to be moved including much unpolluted soil. As a result, the fee of transport and treatment will greatly increase (Figure 2.5).

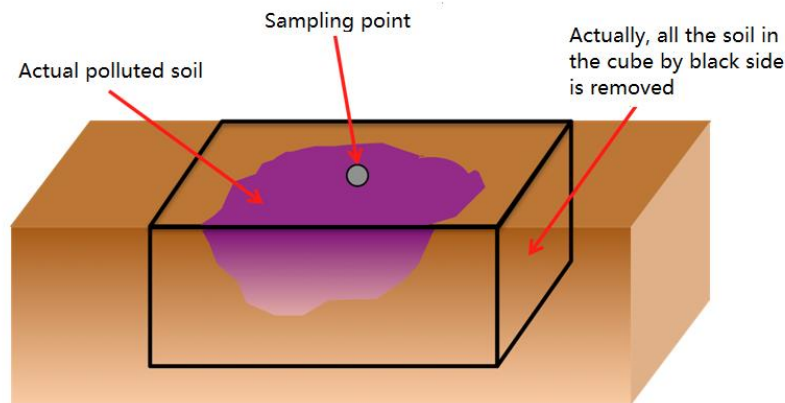


Figure 2.5 Traditional treatment (Excavation) of polluted soil by HMs

Developing better sampling design (including sampling number and position) is a good method to solve the problem of high cost. In practice, taking more samples can make the estimation more accurate and decrease the cost by reducing the risk of removing the unpolluted soil. However, this way certainly increases investigation (sampling) fee (as shown in Figure 2.6). As a result, optimal strategy of sampling and remediation area of excavation method is considered to make the total fee the lowest by balancing the investigation fee and treatment fee.

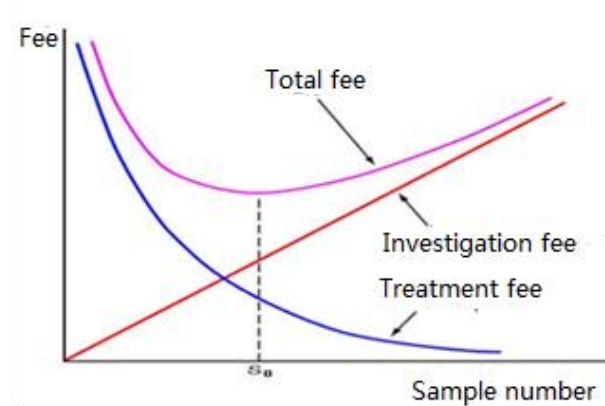


Figure 2.6 Curve of relationship between fee and sample numbers

2.2 Geostatistics

Geostatistics is a branch of statistics which aims at providing quantitative description of natural variables (e.g. concentrations of pollutants in a contaminated site) distributed in space or in time and space⁶³. Geostatistics was proposed by D.G. Krige when dealing with estimating the distribution of minerals of the whole mine zone through few sampling data in 1951 and founded by G.Matheron in 1962 who was the first to use this term extensively^{64,65}.

2.2.1 Kriging

Kriging, as the central method of geostatistics, is always regarded as having the same meaning as geostatistics⁶⁵. Kriging is a BLUE (Best Linear Unbiased Estimator) interpolator, which predicts values at unknown locations from data observed at known locations, which include Simple Kriging, Ordinary Kriging, Indicator Kriging, Co-Kriging, and so on.

2.2.1.1 Simple Kriging (SK)

SK assumes that the random function (RF) $\{Z(x): x \in R^d\}$ was second order stationary meeting the following two conditions:

1. The mean of RF $\mu = E[Z(x)]$ is constant.
2. The covariance is a function of the distance $h = x_1 - x_2$ between the reference points x_1 and x_2 .

$$\text{Cov}(x_1, x_2) = E[(Z(x_1) - \mu)(Z(x_2) - \mu)] = C(h)$$

The value at unobserved points x_0 is estimated by the function $\hat{Z}(x_0) = \sum_{i=1}^n w_i Z(x_i) + \mu$, and the weights w_i is solution of the function of $AX = Y$.

$$A = \begin{bmatrix} C(0) & C(x_1 - x_2) & \cdots & C(x_1 - x_n) \\ \vdots & \vdots & \ddots & \vdots \\ C(x_n - x_1) & C(x_n - x_2) & \cdots & C(0) \end{bmatrix}$$

$$X = (w_1, w_2, \cdots w_n)^T$$

$$Y = (C(x_1 - x_0), C(x_2 - x_0), \cdots C(x_n - x_0))^T$$

And the SK estimation variance is given by:

$$\sigma_{SK}^2 = C(0) - \sum_{i=1}^n w_i C(x_i - x_0)$$

However, in most applications neither the expectation nor the covariance is known beforehand, this makes the model very robust but more restrictive⁶⁶.

2.1.1.2 Ordinary Kriging

Ordinary Kriging uses variogram (mainly means semivariogram in this research) to express spatial variation and utilizes BLUE to keep the error of estimation to a minimum. The variogram includes two parts: experimental variogram and model variogram: An experimental variogram $\gamma(h)$ for an interval lag distance class h , is represented by

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

Where h denotes the lag distance that separates pairs of points; $N(h)$ represents the number of pairs separated by the lag distance h ; $Z(x), Z(x + h)$ denotes the actual values at location $x, x + h$ respectively⁶⁷. After the experimental variogram is computed, the next step is to define a model variogram. A model variogram is a simple mathematical function that models the trend in the experimental variogram showed in [Figure 2.7](#), and mainly includes three parameters: range, sill, and nugget⁶⁸.

- 1) Range is a distance at which the data no longer have autocorrelation and the model first flattens out.
- 2) Sill is the corresponding variogram value at the distance of range mentioned above, and means the total variance of data set; the partial sill is the sill minus the nugget.
- 3) Nugget is the variogram value that exhibits some value more than zero at the distance h where the theoretical value should be zero; the nugget effect is attributed to spatial variability at distance smaller than sampling intervals, including measurement error inherent in measuring devices.

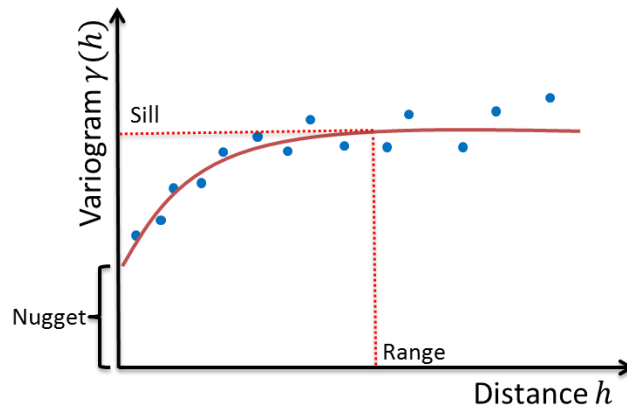


Figure 2.7 Variogram

The five most frequently used models are:

Nugget:
$$g(h) = \begin{cases} 0 & \text{if } h = 0 \\ c & \text{otherwise} \end{cases}$$

Spherical:
$$g(h) = \begin{cases} c \cdot \left(1.5 \left(\frac{h}{a} \right) - 0.5 \left(\frac{h}{a} \right)^3 \right) & \text{if } h \leq a \\ c & \text{otherwise} \end{cases}$$

Exponential:
$$g(h) = c \cdot \left(1 - \exp \left(\frac{-3h}{a} \right) \right)$$

Gaussian:
$$g(h) = c \cdot \left(1 - \exp \left(\frac{-3h^2}{a^2} \right) \right)$$

Power:
$$g(h) = c \cdot h^\omega \quad \text{with } 0 < \omega < 2$$

Where c represents sill and h lag distance, a (practical) range⁶⁹. In this research, the model of Exponential (nugget=0) was used. Once the model variogram was constructed, it was used to compute the weights used in kriging. The criterion used to determine the weights is whether they minimize the estimation variance which can be represented by:

$$\sigma^2_{OK} = \sum_{i=1}^N w_i \gamma(x_i - x_o) + \lambda$$

Where λ is the Lagrange multiplier. Based on the weights, the value of the random variable at an un-sampled location x_o can be estimated:

$$Z^*(x_o) = \sum_{i=1}^N w_i Z(x_i)$$

$Z^*(x_o)$ is the estimated value at location x_o ; x_i is the location of the sampling point for variable Z ; w_i is the kriging estimation weight of $Z(x_i)$; N is the number of the variable Z involved in the estimation. To ensure a non-skewed estimator, the condition that the sum of all weights is 1 is mandatory⁶⁶. In detail, the weights w_i is the solution of the function of $AX = Y$, where

$$A = \begin{bmatrix} 0 & \gamma(x_1 - x_2) & \cdots & \gamma(x_1 - x_n) & -1 \\ \gamma(x_2 - x_1) & 0 & \cdots & \gamma(x_2 - x_n) & -1 \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ \gamma(x_n - x_1) & \gamma(x_n - x_2) & \cdots & 0 & -1 \\ -1 & -1 & -1 & -1 & 0 \end{bmatrix},$$

$$X = (w_1, w_2, \cdots w_n, \lambda)^T,$$

$$Y = (\gamma(x_1 - x_0), \gamma(x_2 - x_0), \cdots \gamma(x_n - x_0), -1)^T$$

The flowchart of the method is shown in **(Figure 2.8)**, and the method is used in this research.

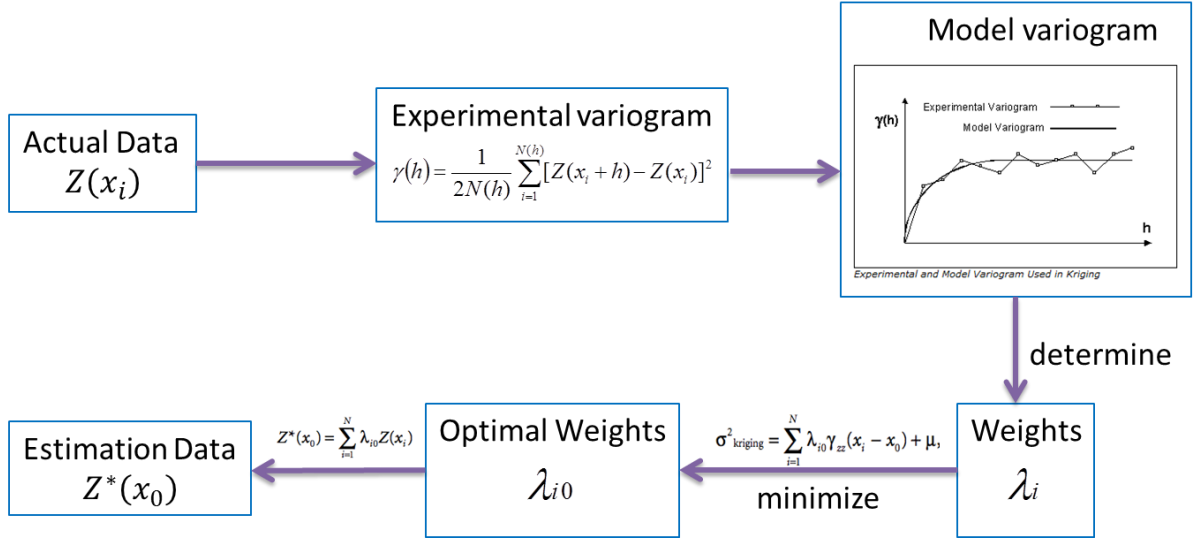


Figure 2.8 Flowchart of Ordinary Kriging

2.1.1.3 Co-Kriging(CK)

Co-Kriging not only uses the main variable of Z_1 , but also introduces additional variable(s) $Z_2, \dots, Z_k$ which exhibit some correlation with the primary variable Z_1 , which means Co-Kriging uses both autocorrelation for Z_1 and cross-correlations between Z_1 and all other variable(s) to make better predictions; the estimator based on the total variables $Z_1, \dots, Z_k$ is given by

$$\hat{Z}(x_0) = \sum_{i=1}^n \sum_{j=1}^k w_{ij} z_j(x_i)$$

To minimize the variance of the Co-Kriging variance, the weights meet the following conditions:

$$\sum_{i=1}^n w_{i0} = 1 \text{ And } \sum_{i=1}^n w_{ij} = 0, 1 \leq j \leq k$$

The detailed calculation of the weights referring to^{65,70} is relatively complicated.

There are other Kriging methods like Universal Kriging(UK), Indicator Kriging(IK); Making the choice which kriging should be used depending on the data characteristics and the type of spatial model desired ⁷¹.

2.1.2 Application and Related literature review

Kriging, as an interpolation method that is widely applied in soil research, can analyze the variability of soil properties such as soil salinity, texture, exchangeable sodium percentage (ESP), electrical conductivity (ECe), pH and cation exchange capacity (CEC)^{72,73}. This method can also analyze the spatial distribution of HMs in soil and does good for the delineation of contaminated area, the decision of additional sampling, the choice for remediation method or the risk assessment^{64,74–80}. There are also other literatures using the Kriging in other fields like spatial patterns of bird diversity⁶⁷ and groundwater monitoring⁸¹. Kriging has also been widely used combined with GIS to map the distribution of soil pollution^{82–87}.

2.1.3 In summary

Kriging can estimate the concentration of contaminants at un-sampled location by spatial interpolation using statistics theory of spatial correlation, and is used as an important interpolation method at different fields such as groundwater, radiological, rainfall and medical application,⁸⁸ especially in studies on soil pollution^{89,90}. Compared with many classical statistical procedures, kriging has the advantages that it incorporates the spatial correlation of the data and allow the incorporation of the data support; while compared with other contouring techniques, kriging has the advantage of its ability to quantify the estimation variance.⁸⁸ kriging will work regardless of the existence of spatial correlation between the data; if the observations are independent, kriging estimates will have the same meaning with estimates determined by using least-squares regression analysis.⁸⁸ In addition, though all interpolation algorithms (inverse distance squared, splines, radial basis functions, triangulation, etc.) estimate the value at a given location as a weighted sum of data values at surrounding locations, kriging assigns weights according to a (moderately)

data-driven weighting function, rather than an arbitrary function, but it is still just an interpolation algorithm and will give very similar results to others in many cases.⁹¹ Kriging has its disadvantages that it always has smooth effect and is based on the assumption of the normality of the distribution. In this research, the ordinary kriging (OK) is mainly used, and cooperate with lognormal kriging and indicator kriging to keep the normal distribution of the data and decrease the effect of smoothing, respectively.

2.2 Genetic Algorithms (GAs)

Genetic algorithms (GAs) are a subset of heuristic evolutionary algorithms that generate globally optimal or near-optimal solutions for diverse, complex, and globally distributed problems through mimicking the process of natural evolution. Though heuristic optimization methods don't ensure global optimal solution, they have the advantage of searching a discrete solution-space globally which is not possible by gradient-based search methods that require continuous solution-space and have a possibility to get stuck in a local optimal solution.⁹² The three fundamentals used for running GAs programming are the representation of the parameters to be optimized or genes, the genetic operator, and the objective function. In this study, additional samplings are represented as genes and the total fee of investigation and treatment of contaminated soil by HMs is the objective function⁹³. The genes should be represented by binary encoding (an array of bits), value encoding (real numbers, characters, or objects), permutation encoding (a string of sequence numbers) or tree encoding to encode the potential solutions to the objective function in a form that the computer can process⁹⁴. After the genetic representation and the fitness function are defined, the process shown in **Figure 2.9** can be run to search population with better solution to objective function. In general, the process can be divided into four main parts: initialization, selection, genetic operators, and termination, which will be introduced in detail in following subsections.

2.2.1 Initialization

The initialization process randomly generates a population of n chromosomes (candidate solutions to the problem) with an array of genes (variables). The population size depends

on what kind of problem to be solved; “small” population size could guide the algorithm to poor solutions, and a “large” population size could make the algorithm expend more computation time in finding a solution; **Figure 2.10** shows some factors may influence the initial population size⁹⁵.

2.2.2 Selection

Selection is based on the fitness value evaluated by fitness function (better the fitness, bigger the chance to be selected); the methods most commonly used are Roulette wheel selection, Rank selection, Boltzmann selection, Steady state selection, and Tournament selection⁹⁴; however in this research the elitist selection was chosen to guarantee the most fit members of each generation will be selected into next generation to avoid no better members turning up in next generation by other selection methods.

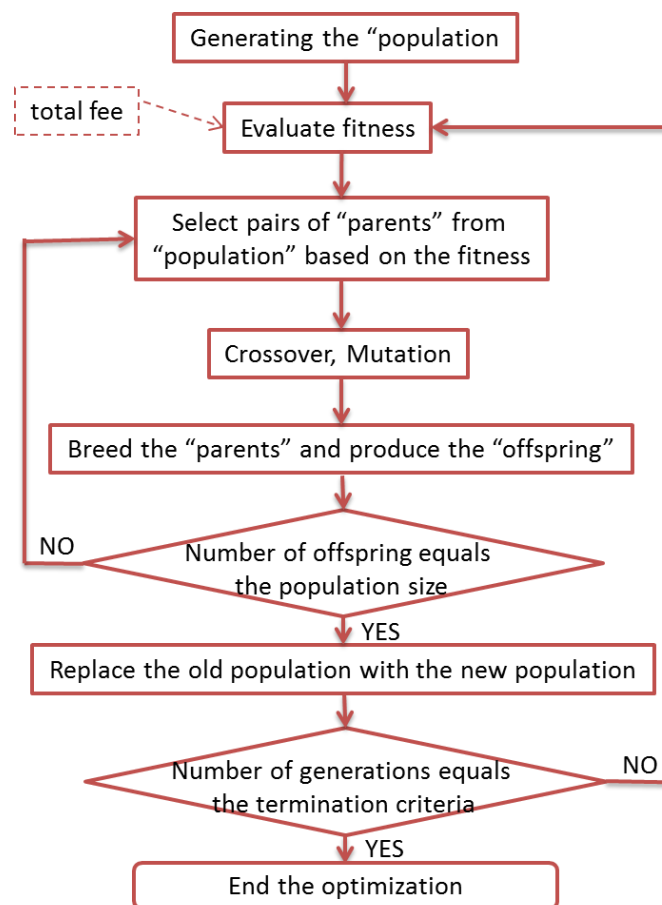


Figure 2.9 Flowchart for Genetic Algorithms

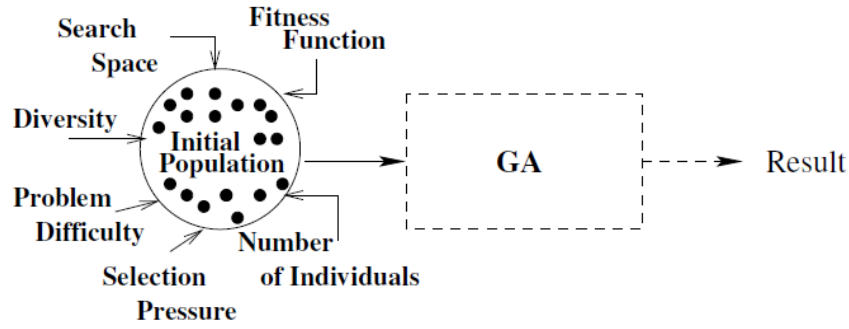


Figure 2.10 Some factors to take into account when the initial population is generated randomly⁹⁵

2.2.3 Genetic operators

After selecting the initial population, the next step is generating the second generation of solutions through genetic operators mainly crossover and/or mutation.

2.2.3.1 Crossover

Crossover combines the features of two parent chromosomes at a randomly chosen point to form two offspring. The crossover operators are of many types, and some commonly used are one-point crossover, two-point crossover, cycle crossover, and uniform crossover; however, the most common is one-point crossover and is used in this research^{96,97}.

2.2.3.2 Mutation

Mutation alters one or more genes of a selected chromosome according to a specified probability P_m . Mutation is the last and final genetic operator, but it is an important part of the genetic search. It can maintain genetic diversity from one generation to next one, and helps to prevent the population from stagnating at any local optima⁹⁴.

2.2.4 Termination

Termination is the condition or criteria used in GAs to stop the generational process when it is reached. The GAs process could be terminated through fixing a number of generations,

locating a string with a certain high fitness value, or attaining a certain degree of homogeneity among the strings in the population (a large number of strings have identical bits at most positions)⁹⁸.

2.2.5 Application and Related literature review

The GAs can also be divided two kinds: single- and multi-objective. Single-objective GA is used to maximize or minimize only one objective function, while multi-objective GA generate the solutions which makes trade-off or balance between or among objective functions. GAs are always used combining with other methods like Kriging. The application of GA is shown in **Table 2**). In this research, the methods of Ordinary Kriging and GAs are used together or/and separately to assess different sampling strategies by total fee estimation.

Table 2 The application of GAs introduced by part literatures

Refer ence	Method	Application	Remarks
⁹⁹	IDW, OK or IDW- OK + GA	Long-term groundwater monitoring design	Sampling costs could be reduced by as much as 60% without significant loss in accuracy of the global mass estimates
¹⁰⁰	NLSI+ Elitist NSGA	Cost effective long- term groundwater monitoring	Identified 34 of 36 members of the Pareto optimal set and represents substantial improvement over standard NSGA which found at most 21 of 36 members
¹⁰¹	QK+NSG A-II	Long-term groundwater monitoring design	Reduce the decision space from 500 million designs to 1,156
¹⁰²	SGA+ FK	Groundwater monitoring network design	Use 20, 25 and 30 monitoring wells in both short scale (28.5 km) and long scale (40 km)
¹⁰³	FOK+NSG A-II	Groundwater monitoring network	demonstrate the potential applicability of the proposed method for the design under epistemic

		design	uncertainty
¹⁰⁴	Two-simulation Models + GA	Groundwater development strategies	Compared with the trial-and-error approach of previous study, the methods are both effective and efficient.
¹⁰⁵	MCSGA and NGA	Cost-effective sampling network design	Compared with MCSGA, NGA achieves a potential cost saving of 45% and reduces the optimization runtime by a factor of 6.5, though moment estimation errors are higher
¹⁰⁶	Log-Normal Kriging + NSGA-II	Monitoring network design for phytoremediation systems	Increasing the number of secondary data sampled resulted in a significant decrease in global uncertainty with a minimal increase in cost
¹⁰⁷	OK and IK +GA	Monitoring and estimating the flow conditions and fish presence probability	the method yields a current velocity 0.32 (m/s) and water depth 0.29 (m), and classifies the flow conditions in the study area as pool, run, riffle and slack
¹⁰⁸	Elitist NSGA-II	Design of falling-film evaporator system	Give well-distributed Pareto-optimal solutions
¹⁰⁹	Kriging + GA	Monitoring network design for transient transport of pollutants in Groundwater aquifers	Established the potential applicability of the proposed methodology for the monitoring network design
¹¹⁰	NSGA-II	Water distribution system	Confirmed the methodology ability to supply a multidimensional tool for sensor placement decision making
¹¹¹	MOGA	Water distribution system	The optimal meter placement problem is solved through trade-off between model precision and

			accuracy
¹¹²	NSGA-II	Water distribution system	A reasonable level of contaminant protection can be achieved using a small number of strategically located sensors
¹¹³	Kriging + GA + AHP	River water quality monitoring	35 stations are selected and using Entropy theory to calculate sampling frequency
¹¹⁴	SGA	Wind turbine micrositeing	The minimum for the cost per unit power is 1.3292×10^{-3} with a maximum power production 27,169.52KW of 54 turbines
⁶⁷	Kriging + GA CK + GA	Additional sampling schemes for spatial patterns and future sampling of bird diversity	GA efficiently obtained additional sampling sites for measuring the amount of bird diversity, and improved the Kriging and CK estimation
¹¹⁵	GA+SIK →GSIK	Model two sources of uncertainty of Indicator Kriging	Results showed a decrease in MAE(6.5925) compared to SIK (8.4364) and IK (8.4366) and the coefficient of determination GSIK(0.8683) compared to SIK(0.8423) and IK (0.8421)
¹¹⁶	SWAT+NS GA-II	The selection and placement of BMP for NPS pollution control	The optimized BMP plans resulted in potential reductions of 33%, 32%, and 13% in sediment, phosphorus, and nitrogen loads, respectively, from the watershed.
⁹³	SWAT+G A	Optimize structural BMP to improve water quality goals	The method is able to optimize the cost of BMPs to meet the defined level of sediment and nutrient load reduction (e.g. 20%, 40%, and 60%)

IDW-Inverse Distance Weighting; **NLSI**-Nonlinear Spatial Interpolation; **QK**-Quantile Kriging; **FK**-Factorial Kriging; **FOK**- Fuzzy Ordinary Kriging; **MCSGA**-Monte Carlo simple genetic algorithm; **NGA**-Noisy genetic Algorithm; **MOGA** -multi-objective genetic algorithm; **AHP**-Analytical Hierarchy Process; **SIK**-Soft Indicator Kriging; **BMP**- Best Management Practices; **NPS**- Nonpoint Source; **NSGA- II**- Non-dominated Sorting Genetic Algorithms- II.

2.2.6 In summary

S.N. Sivanandam and S. N. Deepa has introduced in detail about the comparison with other (conventional) optimization techniques, including the 4 kinds of differences with these techniques, the 17 advantages, 14 limitations and 13 applications of GAs; finally they summarized the GAs are a powerful and robust optimization technique through the ability of the algorithm to explore and exploit simultaneously, a growing amount of theoretical justification, and successful application to real-world problems.¹¹⁷

GAs have been used in many fields, in this research, it is mainly used to optimize the additional sampling, which can result in the relative lower ATFs. The parameters of GA used in this research will also be discussed and decided in detail in Chapter 3.

2.3 Programming in VBA Excel

The programming languages such as C, C++, FORTRAN and VBA Excel can be used to express the Kriging or GA. Nowadays, Excel is the most common software, since computers are so popular in the world. As a result, once the program is created in Excel, and it is easy to extend and share the program with other researchers. On the other hand, many engineers are familiar with MS Excel and always use it to analyze or calculate data in academia or industry¹⁰⁸. So in this research, the chosen program was created in Excel. All the statistic and optimization analyses in this study are performed with the programs created and developed in the EXCEL VBA.

This chapter mainly introduced two methods: Kriging and GAs. Kriging makes it feasible to estimate distribution of HMs in soil through existed samples; GAs are used to optimize the additional sampling design to decrease the ATFs furthermore. The processes of Kriging and GA are both transferred into VBA programs in Excel, and run the programs to get the final results of optimization.

CHAPTER III COST-EFFECTIVE STRATEGY OF THE INVESTIGATION AND REMEDIATION AREA ON THE SURFACE POLLUTED-SOIL (2D)

3.1 Abstract

The geostatistical technique of kriging has extensively been used for the investigation and delineation of soil heavy metal pollution. Kriging is rarely used in practical circumstances, however, because the parameter values are difficult to decide and relatively optimal locations for further sampling are difficult to find. In this study we used large numbers of assumed actual polluted fields (AAPFs) randomly generated by unconditional simulation (US) to assess the adjusted total fee (ATF), an assessment standard developed for balancing the correct treatment rate (CTR) and total fee (TF), based on a traditional strategy of systematic (or uniform) grid sampling (SGS) and kriging. We found that a strategy using both SGS and kriging is more cost-effective than a strategy using only SGS. Next, we used a genetic algorithm (GA) approach to find optimal locations for the additional sampling. We found that the optimized locations for the additional sampling were at the joint districts of polluted areas and unpolluted areas, where abundant SGS data appeared near the threshold value. This strategy was less helpful, however, when the pollution of polluted fields showed no spatial correlation.

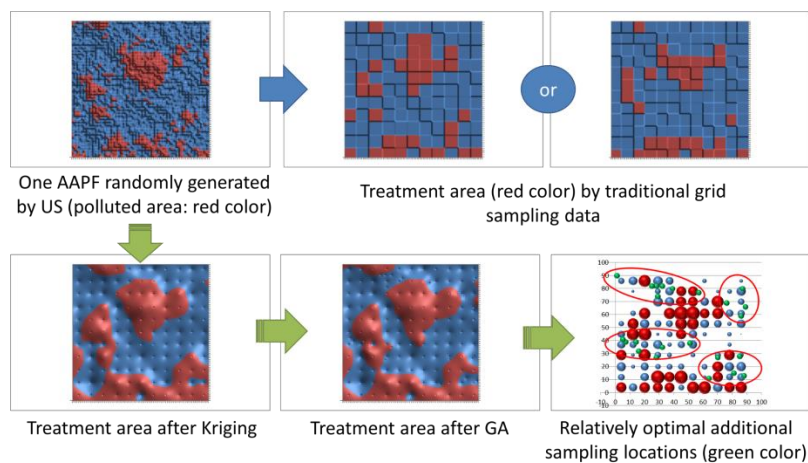


Figure 3.0 Abstract

3.2 Introduction

Soil heavy metal pollution is becoming an increasingly severe global environmental problem, especially in developing countries such as China, due to the toxic, non-biodegradable, and persistent nature of heavy metals (HMs).^{2,36,37,118,119} The pollution situations in China and Japan have been introduced in detail in chapter 1. Remediation of polluted soil by HMs improves not only the quality of soil to keep food and water safe for human health, but also the economic and commercial value of the land after remediation. Yet when we deal with a real-world polluted site, we invariably struggle to solve complex issues such as the number of samples to be taken, the actual site of sampling, the remediation method to use, and where the additional samples should be located in order to trade off the total fee (TF: fees for investigation, analysis, transportation, remediation, etc.) and correct treatment rate (CTR: the percentage of the area or volume of treated polluted soil and the total polluted soil before treatment). We therefore seek to develop a cost-effective strategy for trading off the TF and CTR in this study. Somebody must always take responsibility for obtaining representative samples before analysis. Practitioners of chemometrics often fail to heed that the error for representative sampling is typically far larger (10-100 times, even as much as 100-1000 times) than that incurred in pre-analysis treatment and analysis.¹²⁰⁻¹²² Sampling design involves a sampling approach and a sample density. Two principal sampling patterns are always recommended to collect metal data in soils: random sampling and systematic grid sampling (SGS).^{14,123} Yet the former, random sampling, is inadequate and inefficient in reflecting site conditions and spatial differences in contamination levels.^{78,123} On one hand, the site soil is heterogeneous and the sampling method tends to produce clustering.^{124,125} On the other hand, the large numbers of samples required for random sampling can incur high costs in time and money.¹⁰ For this study we therefore select the SGS approach, which gives an optimum between minimum variance and cost, as an optimal strategy.^{126,127} Sample density by SGS depends on the sampling area and budget. In the soil sampling guidelines used in Europe (ESSG), the number of samples ranges from 2/ha to 144/ha, depending on the country, and the sampling density can reach levels as high as 20/400 m² (Finland) and 16/100 m² (Switzerland).¹²⁸ In Japan, two grid sampling methods have been regulated by Law of Soil Contamination Countermeasures Act (2003): 1) taking one sample at the center of

10m*10m mesh where the pollution probability is relatively high; 2) a composite sample mixed from five sub-samples on the area of 30m*30m, where the pollution probability is relative low.⁴⁸

Even when sparsely distributed across sites, soil sampling data cannot provide all of the suitable information for the effective delineation of contaminated areas (information on the distribution of heavy metals at unsampled locations can be lacking, for example).

Kriging interpolation has been used to model and map spatial distributions of soil heavy metal contamination for more than three decades.^{75,129,130} Kriging is more advanced than other interpolation methods such as the inverse distance weighted method, since kriging considers two sets of distances (the distance between a location of interest and a sample location, and the distance between two sample locations).¹³¹ Yet kriging-based delineation can be confounded by the variation of the kriging estimation (e.g., the smoothing effect). This confounding effect can lead to both false positive ($z^*(x_0) > z_c$ given $z(x_0) < z_c$, where $z^*(x_0)$, z_c and $z(x_0)$ are kriged value, threshold value, and true value at location x_0 separately) and false negative ($z^*(x_0) < z_c$ given $z(x_0) > z_c$) delineation, which in turn can lead to financial risks and environmental risks due to over- or under-estimation at unsampled locations.^{74,75,132,133} The variogram, which is essential for kriging, requires adequate sampling data to generate more precise estimation. Webster and Oliver have proposed that variograms require at least 100 data for computation, while those computed on fewer than 50 data are of little value.¹³⁴ When fewer than 50 or even 100 observations are used, the kriging estimation will lack any clear advantage and lead to a large proportion of false delineations.^{135,136} The data, moreover, must be normally distributed to produce the best possible estimate.⁶⁶ The observed data sets from real polluted sites always diverge from the normal distribution (e.g., unduly high skewness and outliers), hence the standard geostatistical methods are inappropriate for datasets. A logarithmic or Box-Cox data transform is often applied before analysis in this situation, and a Q-Q normal plot is also used to test the normality of the transformed data.¹³⁷⁻¹³⁹ Though, sometimes, data transformation does not always lead to a normal distribution, estimators associated with the normal distribution are usually quite robust, and they do not get seriously affected by the fact that the variable real distribution is not absolutely normal.⁶⁶

The sampling at the beginning of a contaminated site investigation is mainly designed not for geostatistical interpolations, but for risk assessment based on legal regulations or intuition (e.g., the deviating color of soil).¹⁴⁰ Additional sampling therefore becomes necessary at a later stage. A number of sampling approaches to improve the reliability of kriging-based delineation have been proposed. An adaptive cluster sampling (ACS) approach (based on the regulation threshold and kriging variance and an adaptive sampling approach using the cumulative distribution function of order statistics (CDFOS)), for example, was proposed for additional sampling to reduce kriging errors and misclassification rates.^{74,75} The remediation cost uncertainty or expected total loss from misclassifications of the contamination status of the soil has also been used as performance criteria for additional or later-phase sampling.^{140,141} Many studies have focused on conditional simulation (CS) or sequential Gaussian simulation to generate probability maps for the assessment of soil pollution.^{129,132,133,140–145} Even though CS is not aimed at minimizing local variance error, the simulation can better reproduce the spatial variability modeled by the same sample variogram.¹⁴² Additional sampling locations are commonly suggested at sites where the conditional probability of exceedance is around 50%, while the number of additional samples taken depends on the decision-maker's stance on risk (a risk-prone investigator prefers taking fewer additional samples and a risk-averse decision-maker prefers taking more).^{140,141,146} However, these approaches or performance criteria do not lead to an optimal trade-off between TF and CTR simultaneously.

The commonly used validation methods such as cross-validation tend to waste some of the existing sampling data or require more independent samples to evaluate the effectiveness of the variogram model. Cross-validation only validates the prediction accuracy at sampling sites and cannot reflect the accuracy at unsampled sites.¹⁴⁷ Contaminant concentrations are usually impossible to determine at every single location of a real-world site unless the surveyor is willing to pay the prohibitive costs in time and money required to take exhaustive samples. The costs would clearly impede a whole field-scale validation. A complete validation becomes feasible, however, using something like stochastic unconditional simulation (US), a technique capable of generating a perfectly known “synthetic” reality. This type of validation has been widely applied within different fields

of application (e.g., nuclear waste, petroleum engineering, the environment).¹³² Geostatistical techniques have been repeatedly proven to be among the best available methods in the delineation of polluted soil, yet most of the world's soil remediation projects seem to avoid the use of geostatistics in the field. According to a survey in Belgium, for example, 84% of the 32 Dutch soil remediation firms rarely or never use geostatistics in practical circumstances.¹⁴⁴ This clearly underscores the need to report convincing results from research to demonstrate the advantages of integrating spatial statistics into the practice of soil remediation.

In this study we randomly generated a series of assumed actual polluted fields (AAPFs) by US and then used them to validate a cost-effective strategy based on the main two performance criteria: the CTR and TF. We then developed a cost-effective strategy using SGS and kriging and determined relatively optimal locations for additional sampling through GA optimization.

3.3 Materials and methods

3.3.1 Assumed Actual Polluted Fields (AAPFs) Generated by Unconditional Simulation (US).

Methods based on conditional simulation (CS) generate a series of realizations (equally probable solutions) of pollution distribution through original sampling data from one polluted field. In the present research we developed another approach based on US. First, we randomly generated a series of AAPFs. Next, we used the sampling data (logarithm) on each AAPF to generate the realization by kriging. US honors the overall mean, variance, and spatial correlation while disregarding observations at specific sampling locations.¹⁴⁸ As a consequence, an AAPF generated by US keeps a certain spatial correlation and can show the concentration (logarithm data) at every single location. The parameters of a semivariogram based on the exponential model to generate AAPFs are set as follows: sill=1; correlation scale=30 m; nugget=0; mean=1. The total area of each AAPF is 90 m×90 m=8100 m². The threshold value is changeable, and different threshold values can generate AAPFs with different pollution rates (PRs), as shown in [Figure 3.1](#).

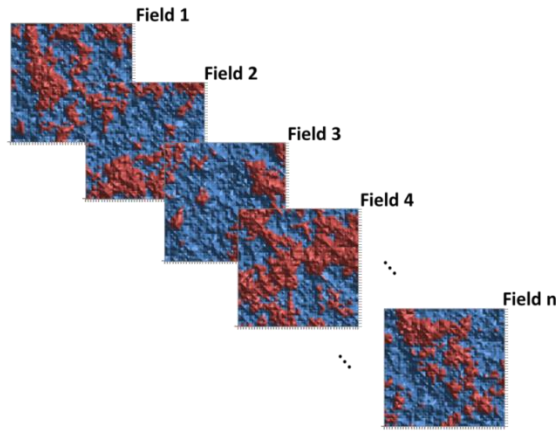


Figure 3.1 AAPFs randomly generated by US. Red area: polluted soil ($\geq$ Threshold value); Blue area: unpolluted soil ($<$ Threshold value))

3.3.2 Preliminary Grid Sampling

Each AAPF is regularly or uniformly divided into $N = n \times n$ square meshes, and a sample is taken at the center of each mesh. The total number of samples is equal to the total number of the square meshes (N), depending on the number of samples, or number of meshes (n) of each side of the AAPF (**Figure 3.2**). In this study we set n in a range from 1 to 45, so the total number of samples ranges from 1 to 2025.

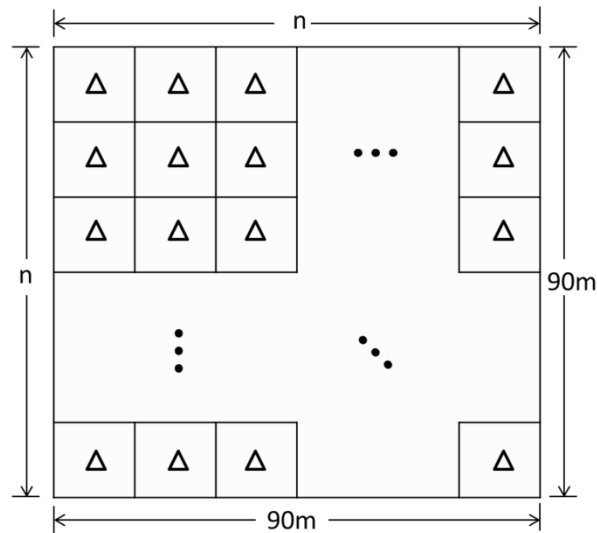


Figure 3.2 Grid sampling on an AAPF. n : sample or mesh number of each side of the square field; Δ : sampling point at the center of each mesh

3.3.3 Remediation Method

Excavation, solidification/stabilization, soil washing, electro-remediation, and phytoremediation, mentioned in chapter 2, have all been used as in-situ and ex-situ techniques to reduce the impact of metals in soil. Notwithstanding, the remediation of contaminated soils by HMs is still recognized as one of the most difficult problems to solve. A few technologies are available for this purpose, but all are costly.^{50,51,149} The policy in Japan is usually to move all of the “possibly polluted” soil (the excavation method): in total, about 86% of all remediation projects in Japan for HM-polluted soil adopt this method.¹⁵⁰ The excavation method is costly, but it removes almost all of the total contaminants, as well as the risk of contamination to neighboring soils, waters, and plants. It also can be performed in relatively little time without introducing other chemicals or requiring further monitoring or treatment on site. If we can use certain methods or techniques to reduce the cost, excavation can become an ideal method for the actual remediation of polluted soil. In this study we treat the excavation method as the optimal strategy.

3.3.4 Performance Criteria

To apply the optimal strategy to real-world remediation projects, the parameters used in the strategy should be as close as possible to real-world conditions and reflect the issues that most directly concern stakeholders or decision-makers: TF and CTR. TF consists of a sampling fee (SF: includes the preliminary sampling fee and additional sampling fee) and remediation fee (RF: includes the analysis fee, transport fee, treatment fee, etc.). One remediation scenario is assumed before the calculation of TF and CTR shown in **Figure 3.3**. The calculation formula is as follows:

$$TF = SF + RF = \text{number of samples} \times 17.5 + ARA \times 4.5 \times 0.5;$$

$$CTR = \frac{CRA}{APA} \times 100\%$$

Where 17.5 is the average price of a boring sampling given by the Japanese companies Jiorizo-mu and Pasture (unit: 10,000 Yen); 4.5 is the average price of polluted soil remediated by the Japanese company Chugai Technos (unit: 10,000Yen); and 0.5 is the depth of the excavated soil (unit: m). The unit prices of sampling and remediation (17.5 and 4.5) are selected with actual market unit prices, respectively. We had two reasons for selecting a soil depth of 0.5. First, HMs generally accumulate more in soil between 0 and 5 cm than at other depths, and most of the studies so far performed have focused on the sampling of surface soil (≤ 20 cm).^{123,151–154} Besides, HMs tend to accumulate in the soil layers rich in organic matter and clay at depths from 0 to 50 cm, because HMs have a high affinity with organic matter and clay that prevents them from rapidly leaching.^{8,155} These layers were also a target of focus to quantify the effect of spatial variability of soil contaminants in cost calculations for soil remediation.¹⁵⁶ Second, more layers (>50 cm) result in more variation, because remediation costs increase quickly with increasing depth and uncertainty increases as more layers are used due to increasing scarcity of available data at deeper depths.¹⁵⁶

To reduce the number of performance criteria in the GA optimization procedure for additional samples and optimally operate the trade-off between TF and CTR, the adjusted total fee (ATF) is introduced with the following function:

$$ATF = \frac{TF}{CTR^\partial} (\partial \geq 0)$$

Where the value of ∂ decides the weight between TF and CTR, that is, where ATF focuses more on TF when the value of ∂ gets smaller (or TF has more influence on the final ATF value than on CTR) and focuses more on CTR when the value of ∂ gets bigger. When the value of ∂ is fixed, the strategy with a lower ATF value is considered better. In a real-world project, the optimal strategy is always expected to reach a high level of accuracy at a given stage of investigation. In this research, the value of ∂ is subjectively set as 2. The performance criterion ATF, however, is imperfect, insufficient, and subject to its own constraint as an assessment standard, since the strategies may have the same ATF values

but different TFs and CTRs (e.g., two strategies have the same ATFs (such as $ATF_1=ATF_2$), whereas the TFs and CTRs may differ (such as $TF_1<TF_2$, $CTR_1<CTR_2$). In this case, the ATF (e.g., ATF_2) with a high CTR (CTR_2) is selected as the relatively better strategy in this study.

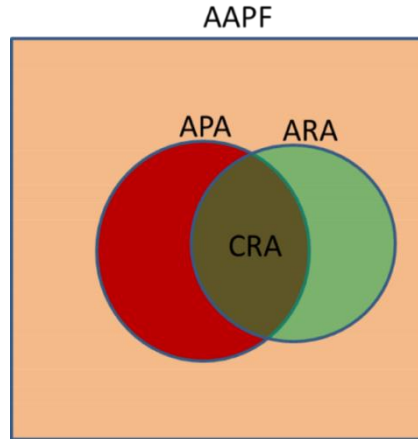


Figure 3.3 One assumed remediation scenario on an AAPF. APA: the Actual Polluted Area of the AAPF (red area); ARA: the Actual Remediated Area of the AAPF, (green area); CRA: the Correctly Remediated Area of the APA of the AAPF (area where the red and green areas overlap)

3.3.5 Geostatistical Interpolation (Kriging Estimation)

We performed this study using ordinary kriging (OK), a method illustrated in detail in earlier literature.^{67–69} To implement kriging here, the exponential semivariogram model is adopted for the spatial interpolation at unsampled locations using log-transformed sampling data, and a semivariogram is calculated for log-transformed data to minimize the effect of extreme outliers.¹²³ Log-normal kriging has theoretical limits, primarily due to the sensitivity of the back transformation from logarithmic data space, but the method yields the best results when data are highly skewed or relatively sparse.¹⁴⁵

3.3.6 Optimization of Additional Samples by a Genetic Algorithm (GA) approach

Genetic algorithms (GAs) are a subset of heuristic evolutionary algorithms that generate globally optimal or near-optimal solutions for diverse, complex, and globally distributed

problems by mimicking the process of natural evolution. While heuristic optimization methods cannot ensure global optimal solutions, they have the advantage of searching discrete solution-spaces globally. Gradient-based search methods lack this capability, as they require continuous solution-spaces and run the risk of getting stuck in local optimal solutions.¹⁵⁷ The three fundamentals used for running GA programming are the representation of the parameters to be optimized (or genes), the genetic operator, and the objective function.⁹³ In this study, additional samplings are represented as genes, and the ATF is the objective function. A detailed description of the GA process is shown in **Figure 3.4**, which can be considered to consist of the following steps:

1. Initialization. The initialization process randomly generates a population of 100 parent chromosomes (called parents in short, which here mean candidate solutions to additional samples design), which each parent chromosome consists of an array of genes (additional samples).
2. Selection. Selection is based on the fitness value evaluated by fitness function (better the fitness, bigger the chance to be selected). In this research, the value of ATF was used as evaluate fitness, and the elitist selection was chosen to guarantee the most fit members of each generation will be selected into next generation to avoid no better members turning up in next generation by other selection methods.
3. Genetic operators: crossover and mutation. 1) Crossover combines the features of two parent chromosomes at a randomly chosen point to form two offspring chromosomes (called children in short).The crossover operators are of many types, and some commonly used are one-point crossover, two-point crossover, cycle crossover, and uniform crossover; however, the most common is one-point crossover and is used in this research.^{96,97} 2) Mutation alters one or more genes of a selected chromosome according to a specified probability P_m (in this research, $P_m=0.05$). Mutation is the last and final genetic operator, but it is an important part of the genetic search. It can maintain genetic diversity from one

generation to next one, and helps to prevent the population from stagnating at any local optima.⁹⁴

4. Termination. A single GA cycle, known as a “generation”, includes process of selection and genetic operators mentioned at step2 and step3. Termination is the condition or criteria used in GAs to stop the generational process when it is reached. The GAs process could be terminated through locating a string with a certain high fitness value, attaining a certain degree of homogeneity among the strings in the population (a large number of strings have identical bits at most positions) or fixing a number of generations.⁹⁸ This research was terminated arbitrarily using the number of generations was set as 10.

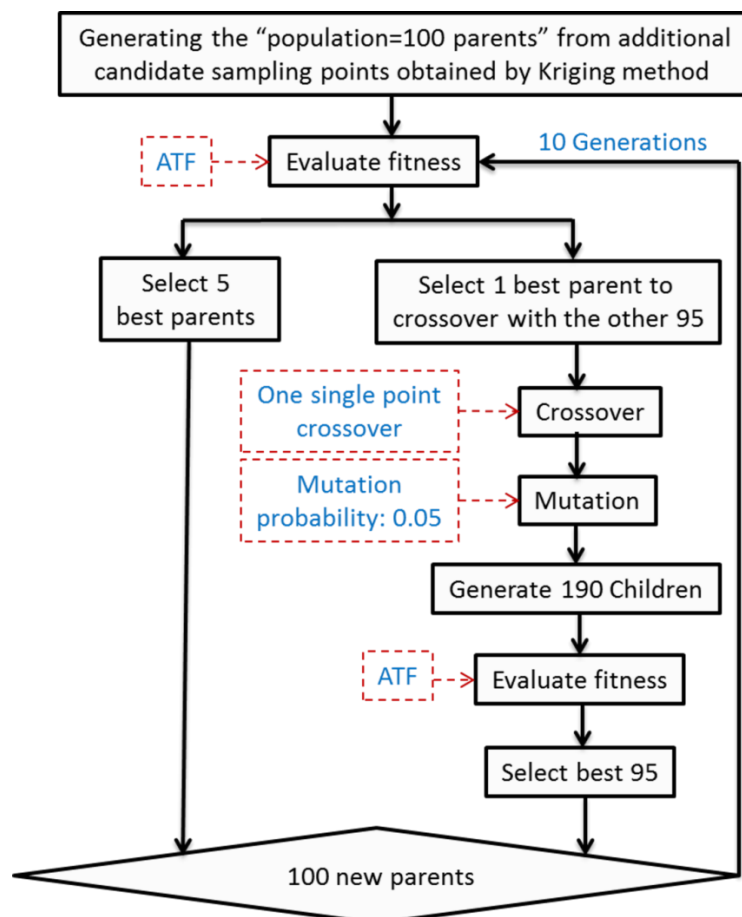


Figure 3.4 Flowchart of the GA process used in this study

3.4 Result and discussion

3.4.1 Preliminary Grid Sampling Analysis

Figure 3.5A(c) shows the relationship between the ATF and number of samples by summarizing 100 randomly generated AAPFs. The general trend of the average ATF (μ) first decreases and then increases as the number of samples increases. The standard deviations (SDs: σ s) decrease sharply as the number of sample increases while the number of samples is still small, but change little afterwards. When the number of samples is quite small, the ATF and SD fluctuate largely. Hence, taking a small number of samples from a real-world polluted field always introduces a high risk of uncertainty for both the TF and CTR. From **Figure 3.5A(d)**, a partial enlarged drawing around the turning point of **Figure 3.5A(c)**, we can see a somewhat lower ATF when the number of samples ranges from 81 to 169, and is lowest when the number of samples is 121. Based on the performance criterion of ATF in this research, SGS with 121 samples, that is, more samples than required by Japanese law (=81), was performed on a 90 m×90 m square polluted field. Other parameters used for the generation of AAPFs by US such as the sill, correlation scale, and sample price (SP) are also discussed to elucidate the change between the ATF and number of samples.

When sill values vary in the manner shown in **Figure 3.5A**, the optimal numbers of samples for keeping a relatively lower ATF does not widely differ (from 81-169), and the relative lowest ATF appears when number of samples is near 121. Hence, the optimal number of samples for keeping a lower ATF is not closely related to the sill. The SD (σ_{ATF}) of the ATF, however, decreases as the sill increases. This behavior can be attributed to that, the data distribution of the AAPF concentration (generated by US with higher sill values) has a relatively higher SD (σ_{data}) and higher average value (μ_{data}), and the area requiring remediation (concentration data $\geq$ threshold value) is larger (or the PR of the AAPF will be larger), which enables a reduced fluctuation of ATFs when using the same grid sampling on different AAPFs. Meanwhile, the average ATF (μ_{ATF}) may increase as the sill increases. The higher σ_{data} means the larger fluctuation of the AAPF concentration, and the grid sampling data cannot reflect sufficiently the real situation of pollution distribution leading to more false positive and negative delineation, hence the μ_{ATF} (when

the number of samples is the same) may increase (red line or red-point line shown in (Figure 3.5A) when the sill is higher.

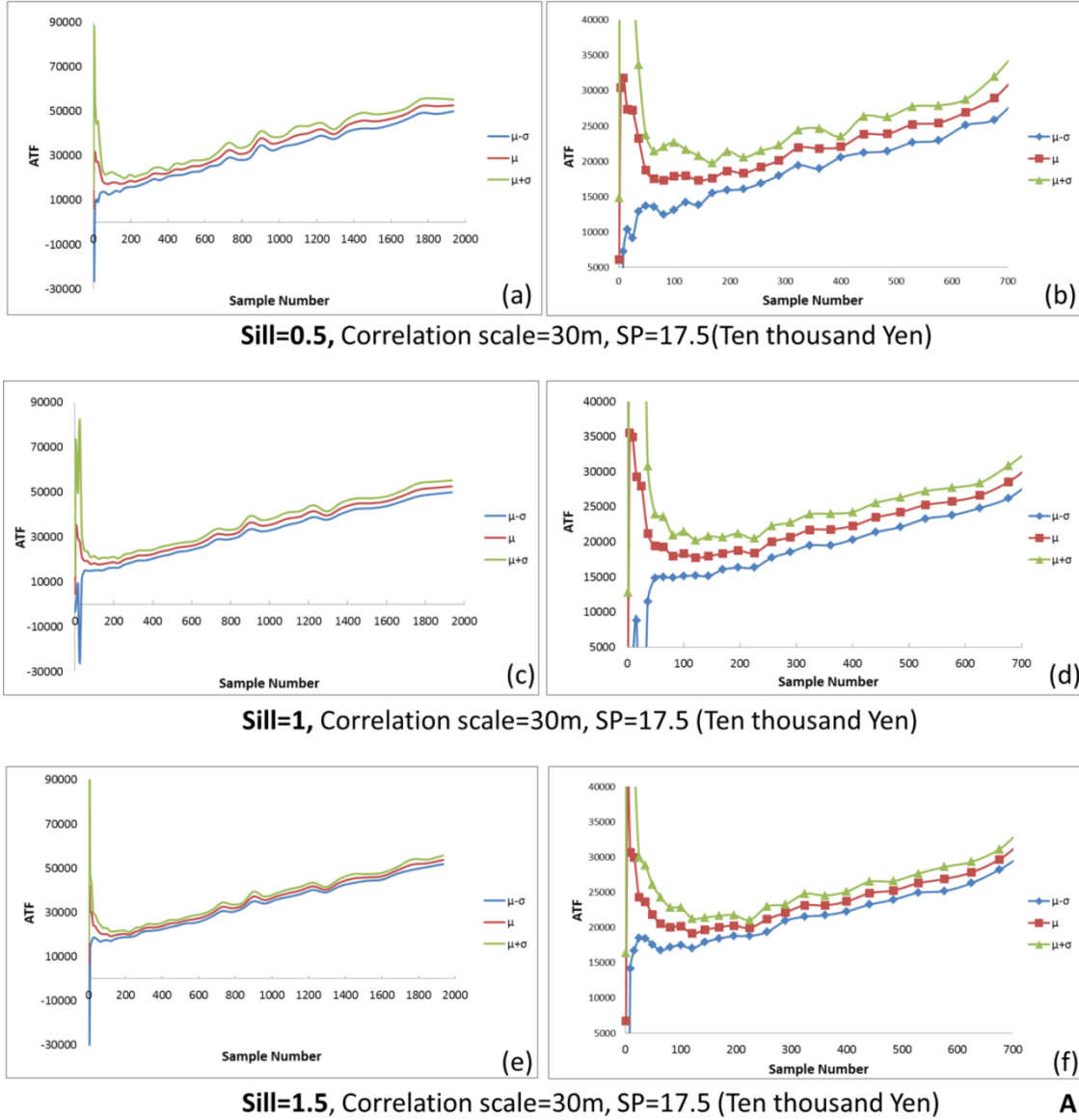


Figure 3.5A Relationship between the ATF and sample number determined by summarizing 100 randomly generated AAPFs while changing the values of the sill. Figures (b), (d), and (f) are partial enlargements around the turning points of figures (a), (c), and (e). μ is the average value and σ is the corresponding standard deviation.

When the values of correlation scale differ in the manner shown in **Figure 3.5B**, the SDs of ATF do not show much obvious change. The optimal number of samples for a relatively lower ATF decreased as the correlation scale increased, because a lower correlation scale value when the sill is fixed signifies that the AAPF only has spatial correlation on a smaller scale, which requires denser sampling data to give a more accurate estimation.

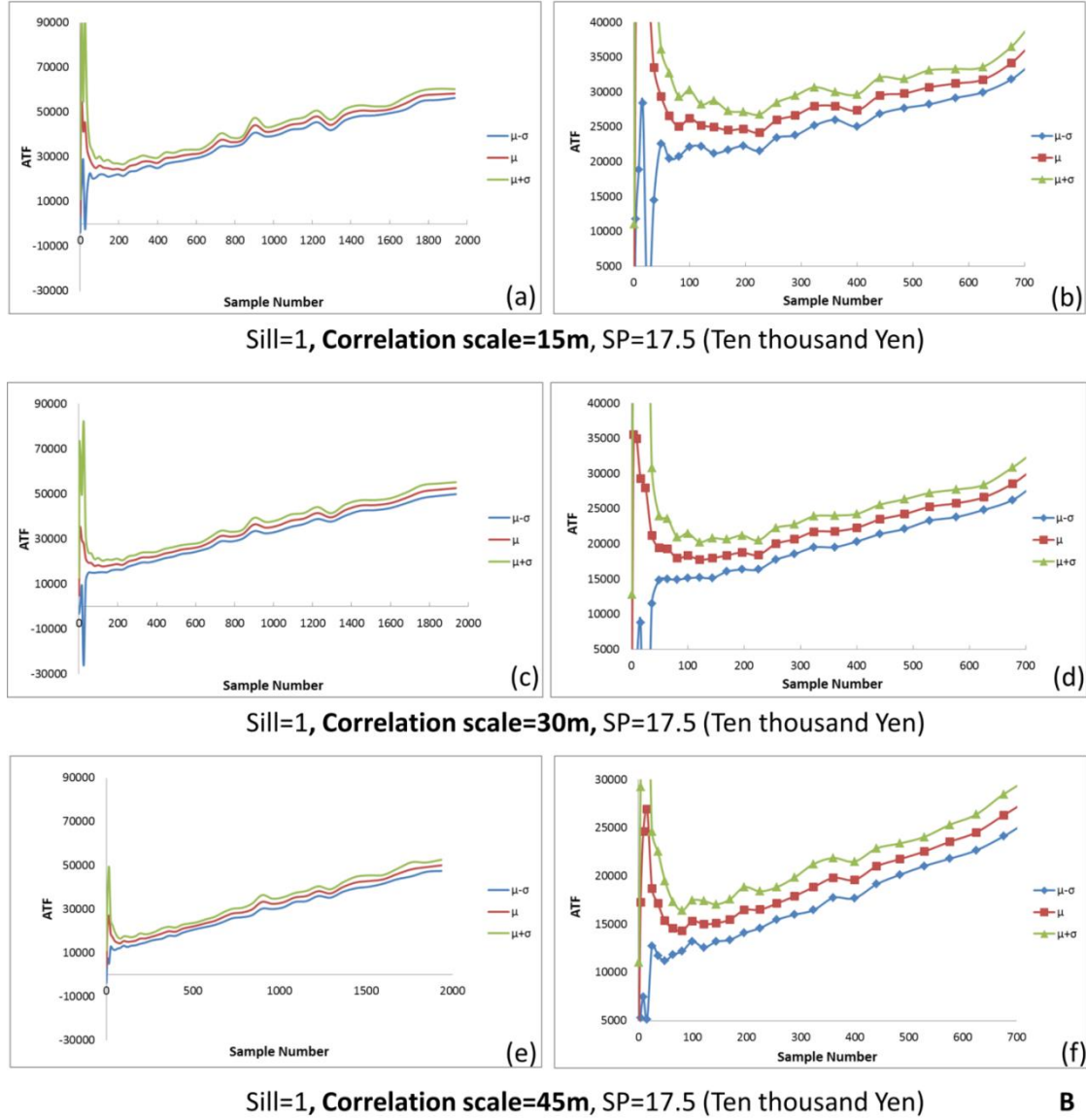


Figure 3.5B Relationship between the ATF and sample number determined by summarizing 100 randomly generated AAPFs while changing the values of the correlation scale. Figures (b), (d), and (f) are partial enlargements around the turning points of figures (a), (c), and (e). μ is the average value and σ is the corresponding standard deviation

While SP changes, the optimal number of samples with the relative lowest ATF varies. As **Figure 3.5C** illustrates, the optimal number of samples with the relative lowest ATF appears around 225 when the SP decreases from 17.5 to 10, and approaches 81 when the sample price increases from 17.5 to 20.

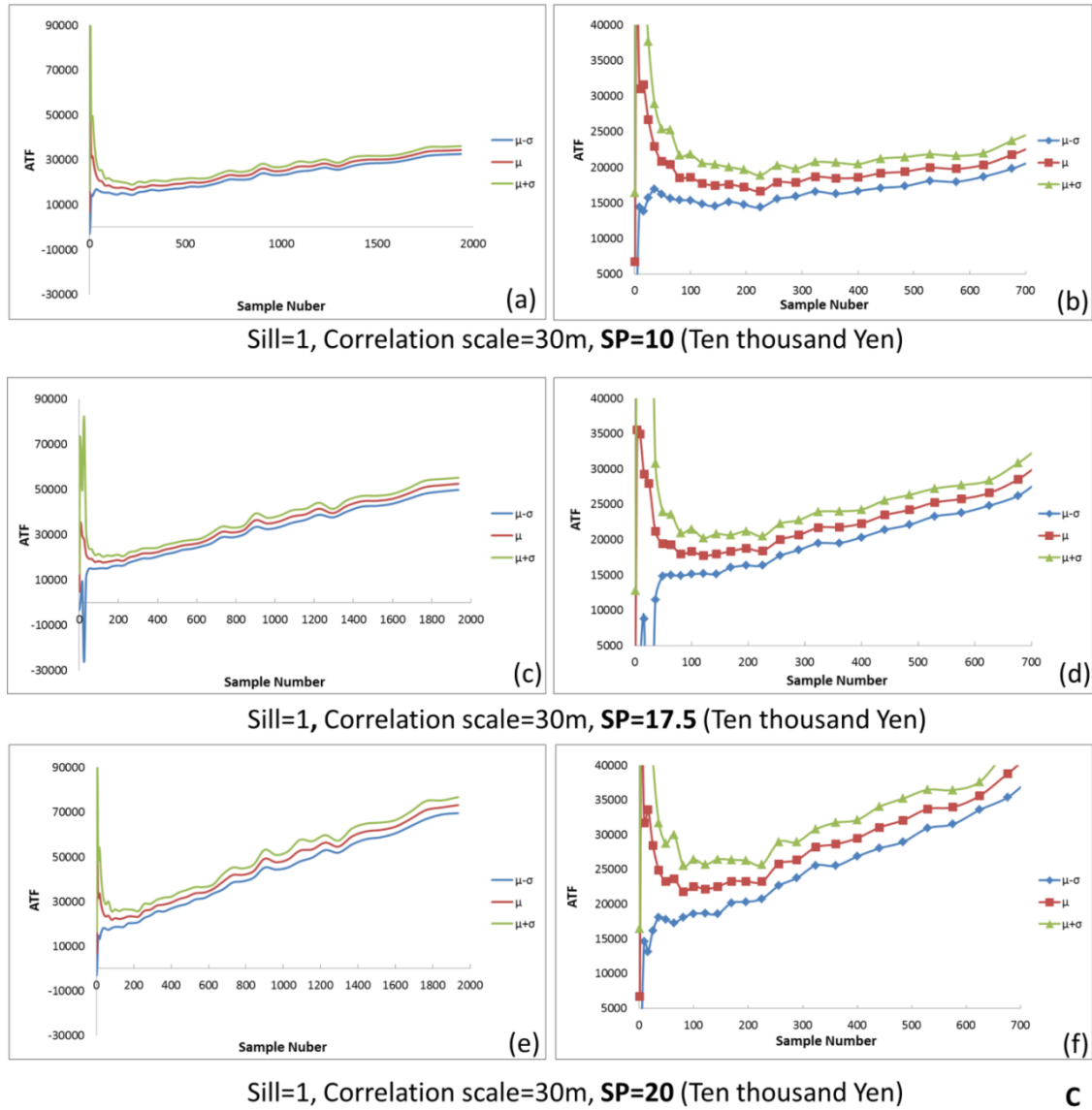


Figure 3.5C Relationship between the ATF and sample number determined by summarizing 100 randomly generated AAPFs while changing the sample price (SP). Figures (b), (d), and (f) are partial enlargements around the turning points of figures (a), (c), and (e). μ is the average value and σ is the corresponding standard deviation.

Additionally, the SD of the ATF generally increases as the SP increases, because the same uncertainty with higher SP will result in a larger change in the ATF value. We therefore propose that more samples can be taken in the real-world treatment of a polluted field when the sample price decreases, and *vice versa*.

3.4.2 Geostatistical (Kriging) Analysis.

Based on the result of preliminary grid sampling analysis (optimal number of samples, 121; assumed threshold value, 1.5), one AAPF (shown in **Figure 3.6(a)**) from a total of 150 randomly generated AAPFs was selected as one example to continually perform the geostatistical analysis. The analysis basically consisted of three steps: (1) exploratory data analysis; (2) structural analysis (variogram model selection, parameter determination, and model validation); and (3) prediction (kriging estimation error and simulations).¹⁴³

A histogram of the sampling data on this randomly selected AAPF is illustrated in **Figure 3.6(b)**. Although OK requires no data to distribute normally, predictions are usually performed better if the sample histogram is displayed without strong skewness. The use of normally distributed data, is essential, however, if one plans to use the kriging SD to build confidence intervals (CIs) for the estimation (e.g., in this study).^{136,143} From both **Figure 3.6(b)** and **Figure 3.6(c)**, we see that the sampling data generally conform well with the symmetrical shape of the normal distribution.

The sample semivariogram calculated to account for the spatial structure of concentrations (see **Figure 3.6(d)**) is well fitted to the exponential model (sill=1, correlation scale=28.5m). The goodness of fit was evaluated by the residuals (ϵ s) in kriging, which is introduced in detail in Kitanidis' book on the application of geostatistics to hydrogeology.¹⁵⁸ When we follow the same calculation with **Figure 3.6(c)**, we learn from **Figure 3.6(e)** that the ϵ s generally display the normal distribution well. Since the ϵ s are orthonormal (i.e., uncorrelated with each other and normalized to have unit variance), the ϵ s should only have the sample semivariogram with a nugget-effect value of about 1 at the short-correlation-scale distance. At a long-correlation-scale distance, on the other hand, they might show heavily fluctuated scatter points from 1. Clearly, as shown in **Figure**

3.6(f), the sample semivariogram values of ε_s are very close to 1 at short-correlation-scale distance, which illustrates again that the exponential model fits the sample semivariogram better. **Figure 3.6(g)** and **Figure 3.6(h)** respectively show the estimation values of kriging (CI=75%) and the corresponding SD distributions. As **Figure 3.6(h)** reveals, the SDs around the existing sampling locations are small and will increase as the distance (between the unsampled locations to be estimated by kriging and the existing sampling locations) increases.

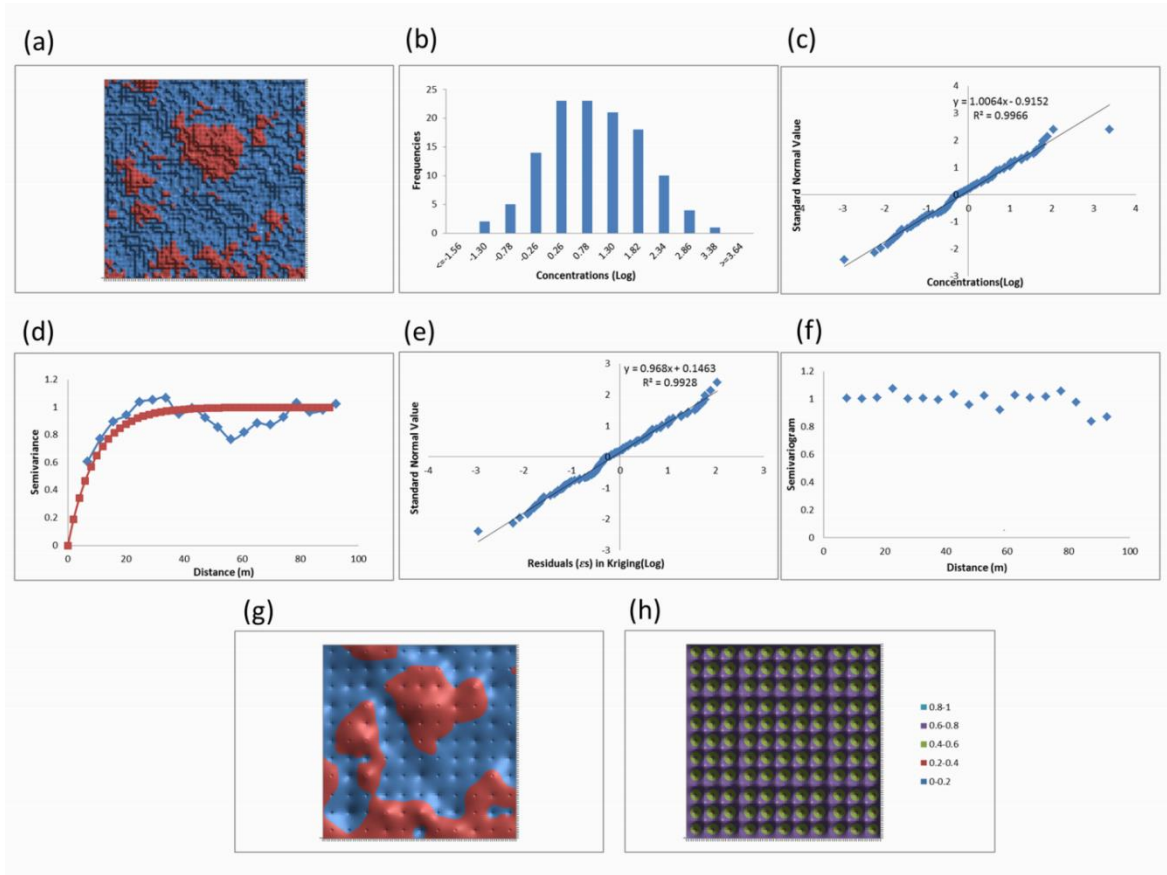


Figure 3.6 (a) One randomly generated AAPF; (b) Histogram of concentrations on the randomly selected AAPF; (c) Normal plot of concentrations (Log); (d) Sample semivariogram for concentrations (blue line) and the exponential model (red line) fitted to the sample semivariogram; (e) Normal plot of residuals (ε_s) in kriging; (f) Sample semivariogram for the residuals (ε_s) in kriging; (g) kriging estimation of pollution distribution (CI=75%); (h) Kriging Standard Deviation (SD) of 121 preliminary grid sampling data

By continuously using different Confidence Intervals (CIs from 50% to 95% in 5% increments, and 99%) on this AAPF after the structural analysis, the kriging generates solutions (realizations) for the treatment areas with different corresponding ATFs until the best solution with the lowest ATF can be obtained. The best CI values, however, differ for different AAPFs. By analyzing the best CI values for different AAPFs, we find some relation between the best CI values and pollution rates (PR: APA/ total area of AAPF*100%, as shown in **Figure 3.3**). An exploration of this relation with the 150 AAPFs (PR: 4%-84%) reveals a stepwise reduction of CIs as the PR of the AAPFs increase (**Figure 3.7(a)**). In real-world polluted sites, meanwhile, there is no way of knowing the PR beforehand (as with the AAPF), because the concentration of HM is difficult to determine at each single location. This obstacle makes it quite difficult to use the PR to suggest the corresponding best CI value used in the kriging. We do, however, have the option of using the PR reflected by the sampling data (SDPR) as a proxy for the PR, since the SDPR and PR have a linear relation (see **Figure 3.7(b)**). After using the SDPR instead of PR, we can obtain the relation between the SDPR and corresponding best CI. When we use this approach to obtain the relation shown **Figure 3.7(c)**, we still find a stepwise reduction of the CI value as the SDPR increases.

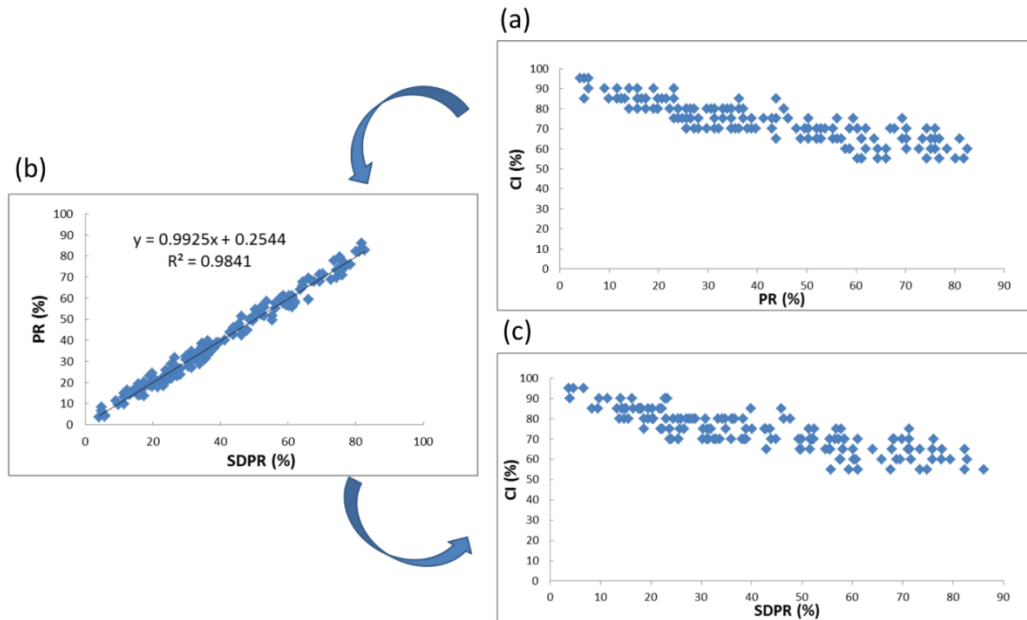


Figure 3.7 (a) Relation between PR and the corresponding best CI used in kriging; (b) Relation between PR and SDPR; (c) Relation between SDPR and the corresponding best CI used in kriging.

The best CI range and best average CI value (calculated from the best CI range) can be suggested for AAPFs with different SDPR ranges, which are generally summarized in **Table 3**. We also find suggestions for the best CI range and best average CI value when the SDPRs are at 0-4% and 84-100%, but these are subjectively based on the trend of the best average CI values suggested in the other SDPR ranges in **Table 3**. **Figure 3.8A** compares the results between the ATF of the preliminary grid sampling analysis and the ATF of the kriging analysis. The ATFs that use the corresponding best CI value for each AAPF in the kriging estimation (green line) can be substituted for the ATFs that use the corresponding best average CI value (red line), since they fluctuate at almost the same level and have similar values. Both of the ATFs (red and green line) may be far smaller than the ATFs based only on the preliminary sampling data (blue line), although the differences among them become smaller as the PR rate of the AAPF (AAPF number) increases. If the CTR is considered apart from the performance criterion ATF, the CTRs that use the best CI value (green line) and best average CI value (red line) fluctuate at the same level and have similar values, as shown in **Figure 3.8B**. Both of the CTRs are greatly improved compared with the CTR using only the preliminary sampling data (blue line), although the differences among them become smaller as the PR (or SDPR) rate of the AAPF (or AAPF number) increases.

Table 3 The best CI values of kriging suggested for different SDPR ranges.

SDPR Ranges (%)	0-4 ^a	4-14	14-24	24-44	44-64	64-84	84-100 ^a
Best CI range for kriging (%)	95-99 ^a	85-95	80-90	70-80	65-75	55-70	50-60 ^a
Best average CI values for kriging (%)	95 ^a	90	85	75	70	60	55 ^a
Best CI range for additional samples (%)	95-99 ^a	90-95	90-95	80-85	75-80	65-70	55-60 ^a

^aThese values are extrapolated based on the trend of best CI values suggested in the other SDPR ranges.

Based on the analysis above, the combination of preliminary grid sampling and the kriging method can result in more accurate estimations of pollution distribution with relatively higher CTRs, which in turn can result in more cost-effective strategies with relatively lower ATFs for decision-makers for remediation. As a result, [Table 3](#) can serve as a reference when using kriging to estimate the pollution distribution on a single polluted field before making a cost-effective plan for remediation.

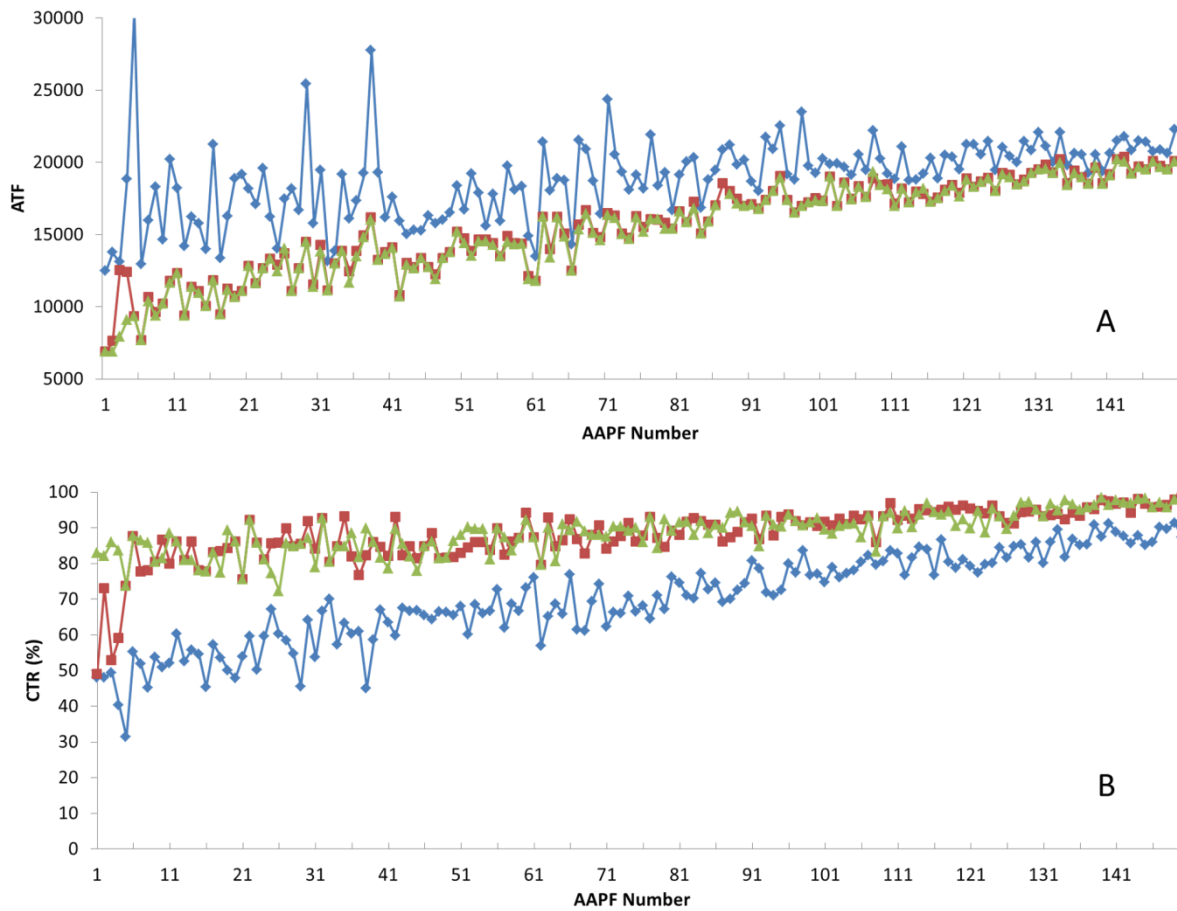


Figure 3.8 The different ATFs (A) and corresponding CTRs (B) for each of the 150 AAPFs. Blue line: ATF (CTR) for the preliminary grid sampling method. Green line: ATF (CTR) after kriging estimation by selecting the best CI value for each AAPF. Red line: ATF (CTR) after kriging estimation by selecting the best average CI value. The AAPF number is ranked in ascending order of PR or SDPR (larger AAPF number, with a larger PR or SDPR value).

3.4.3 Genetic Algorithm (GA) Optimization Analysis.

Ten AAPFs were randomly selected in each SDPR range (except the first range (1-4%) and last range (84-100%)) based on **Table 3**. Hence, a total of 50 AAPFs were used to conduct the optimization work. Five candidate areas contoured by two neighboring kriging CI values (from 50% to 95% in 5% increments, and 99%) around the best average CI value (**Table 3**) were selected to conduct the GA optimization on each randomly selected AAPF. Assume, for example, that the SDPR of one randomly selected AAPF is 8%, which belongs to the SDPR range 4-14%, so that the corresponding best average CI value suggested is 90% based on **Table 3**. The 5 section candidate areas with CI ranges (90-95%, 85-90%, 95-99%, 80-85%, and 75-80%) around this best average CI value (90%) can thus be selected to operate the GA optimization and then search the relatively optimal additional samples, which keep the ATF of the whole strategy lowest. The additional number of samples should preferably be kept small to run the optimization procedure at a fine resolution.¹⁴⁴ In this research we selected 23 additional samples because the ATF calculated by the 121 preliminary grid sampling data together with the 23 additional sampling data can be compared with the ATF calculated by 121 or 144 preliminary grid sampling data either in combination with the kriging estimation or without it combined with kriging estimation or not. Besides, the additional sampling number 23 has no obvious effect on the semivariogram structure when using the kriging estimation.

Here we use additional samples taken on 5 candidate areas (contoured by CI ranges: 75-80%, 80-85%, 85-90%, 90-95%, and 95-99%) of 10 randomly selected AAPFs (SDPR Range: 4-14%) as one example to show the changes of the ATFs after GA optimization. **Figure 3.9(a)** illustrates that the relative lowest ATF appeared in the CI range of 90-95%, whereas none of the ATFs after GA optimization on the 5 candidate areas showed much difference. The corresponding CTRs, however, have the general decreasing trend shown in **Figure 3.9(b)**. Compared with both **Figure 3.9(a)** and **Figure 3.9(b)**, we can always obtain a relatively lower ATF with a higher CTR in the CI range of 90-95%. We can also obtain a relatively lower ATF with a higher CTR in the CI range of 95-99%, but this CI range always contributes to a larger candidate area and increases the difficulty in selecting additional sampling locations. Hence, we selected 90-95% as the best CI range for the

candidate area for additional sampling. **Table 3** shows the best CI range suggested for the additional sampling based on the same evaluation of the ATFs on other randomly selected AAPFs with different SDPR ranges (14-24%, 24-44%, 44-64%, 64-84%). When the SDPR ranges are at 0-4% and 84-100%, the best CI ranges are subjectively suggested to be 95-99% and 55-60%, respectively, based on the trend of the best CI ranges values suggested in other SDPR ranges. As a result, the corresponding ATFs and CTRs for the 50 AAPFs were compared with the ATFs and CTRs calculated by the other three approaches, after GA optimization using the best CI range suggested for additional samples in **Table 3**: 144 preliminary grid sampling data, 144 preliminary sampling data combined with kriging estimation, and 121 preliminary sampling data combined with kriging estimation. As shown in **Figure 3.10A**, the ATFs after kriging estimation (blue line and green line) are considerably lower than the ATF (red line) calculated only from the preliminary sampling data. We find, however, that the ATFs (purple line) after GA optimization and kriging estimation are always the lowest, even though this may be difficult to tell from the visual plot in **Figure 3.10A**. As the SDPR of the AAPFs increases (or the AAPF number increases), the ATFs generated by these four approaches become inconspicuous and bunched more closely together, since an AAPF with a higher SDPR always offers relative less space for optimization. Even though the ATFs (blue, green, and purple lines) are closer to each other, the structures of the ATFs (the ratios or weights of the TFs and CTRs) vary. From **Figure 3.10B** we can see that the corresponding CTR (purple line) is always the highest. We also find, in both **Figure 3.10A** and **Figure 3.10B**, that the ATFs and CTRs (blue lines and green lines) show relatively little difference. From this we can infer that a traditional increase of grid sampling data will not necessarily generate a more cost-effective effort. Through the additional sampling data optimized by GA, we can lower the ATF and increase the CTR simultaneously. We also need to narrow the candidate area for additional samples and consider their detailed locations, which is invariably difficult when dealing with real polluted sites. Limitations in the GA optimization procedure in this study may preclude identification of the optimal additional sampling locations, since the termination was set at 10 generations. Even so, the optimized additional samples by GA in this study suggested the relative optimal locations.

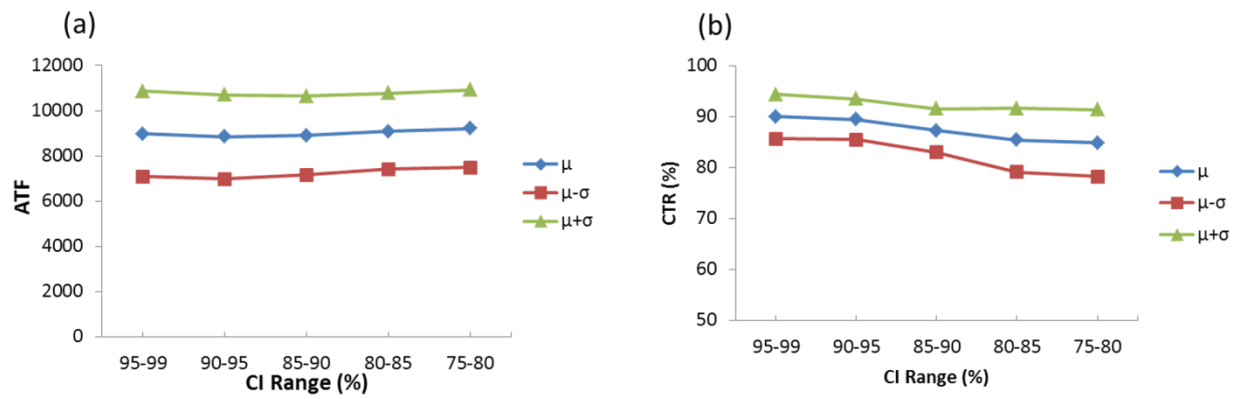


Figure 3.9 ATFs (a) and CTRs (b) after GA optimization on 5 candidate areas (CI Ranges: 75-99%) of 10 randomly selected AAPFs (SDPR Range: 14-24%). μ is the average value and σ is the corresponding standard deviation.

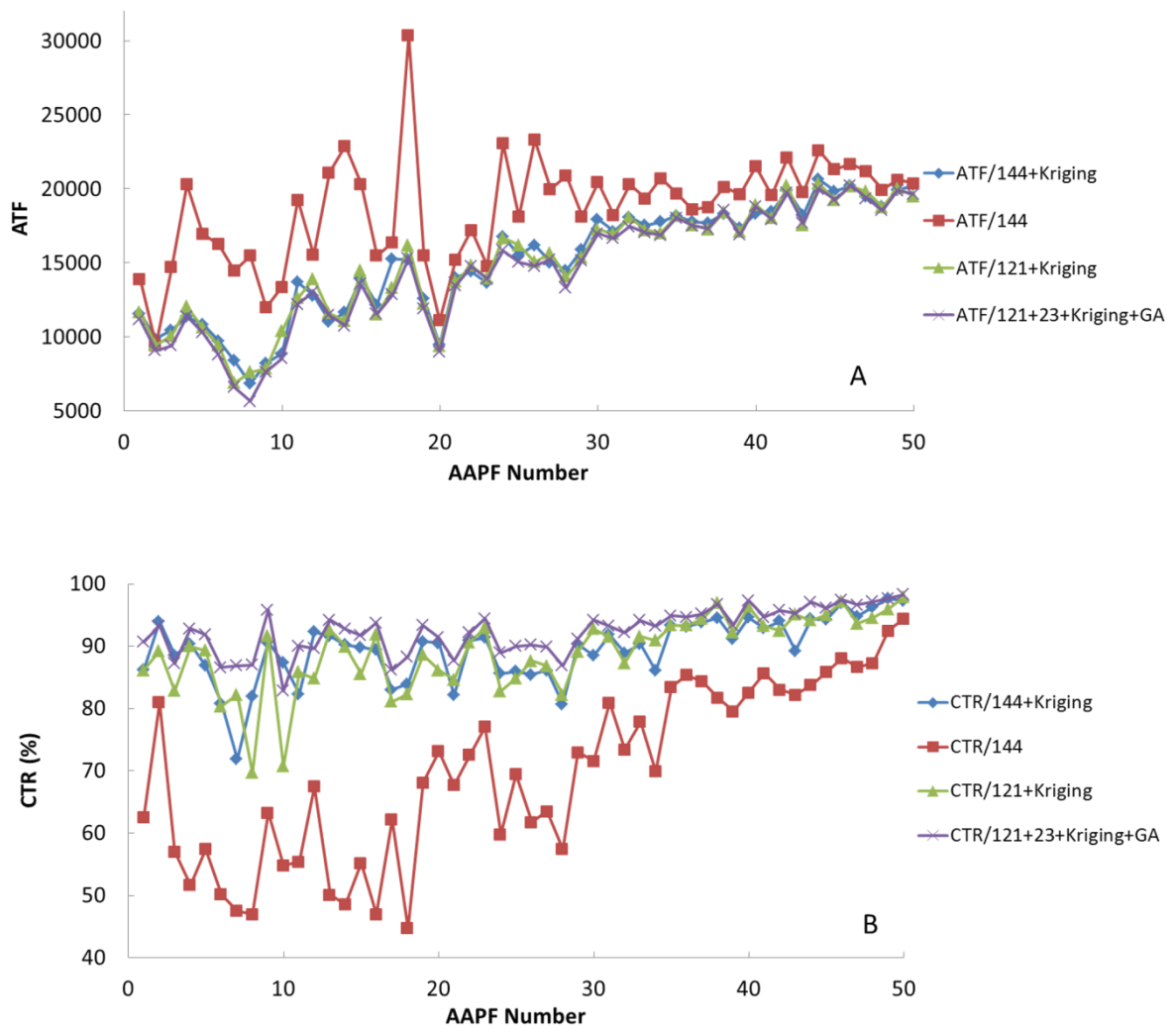


Figure 3.10 Comparison of the ATFs (A) and CTRs (B) calculated by different combinations of preliminary sampling data, kriging estimation, and GA optimization. The 121 and 144 in the legend are the preliminary numbers of samples and 23 is the additional number of samples. The AAPF numbers from 1 to 50 in 10 increments correspond to the SDPR ranges of 4-14%, 14-24%, 24-44%, 44-64%, and 64-84%, respectively.

In a study on multi-phase sampling for soil remediation surveys by Marchant et al., the additional samples were generally dispersed on the boundaries between areas which, according to the first phase, either required or did not require remediation.¹⁴¹ Our result shown in **Figure 3.11** is similar. The additional samples (green bubbles inside the red circles) are always selected at the joint locations of polluted and unpolluted areas, where there are abundant sampling data near the threshold value, whereas the additional samples are generally dispersed over the full scale of the AAPF rather than concentrated in any one area.

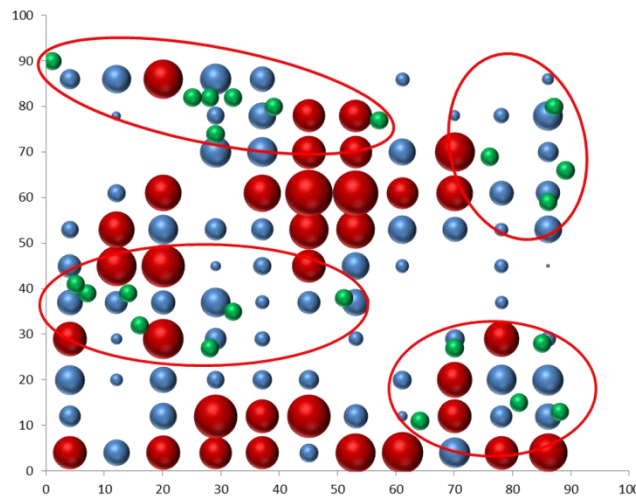


Figure 3.11 Distribution of 121 preliminary samples (red and blue bubbles) and 23 additional samples (green bubbles) on the AAPF used in Figure S6. The sizes of the red and blue bubbles represent the concentrations of samples above and below the threshold value, respectively, while the green bubbles only represent the additional sample location without considering the size.

3.5 Summary and future works

In conclusion, our study combined the grid sampling method, kriging estimation, and GA optimization for additional samples into an overall strategy for investigating and remediating polluted fields. Many different AAPFs generated by US are used in place of real-world polluted sites to give relative optimal suggestions to engineers or decision-makers, such as parameters to use in kriging estimation and suggested locations for additional samples, to develop more cost-effective (balanced TF and CTR) strategies before dealing with the real-world polluted sites. The study thus exemplifies the advantages of integrating geostatistical interpolation and GA optimization into the practice of soil investigation and remediation, even though the model relies on certain assumptions and the strategy is unhelpful when the pollution of a polluted field shows no spatial correlation. We are now considering ways to use and validate this strategy in practical circumstances in the future.

CHAPTER IV OVERVIEW OF THE EXISTED 3D MODEL AND DISCUSSION OF THE DEVELOPMENT FEASIBILITY OF THE COST-EFFECTIVE STRATEGY FOR THE SOIL AT DEPTH (3D)

4.1 Introduction

The concentration of heavy metals and the position information of sampling are not enough to make more accurate estimation of the amount and spatial distribution of heavy metals at un-sampled sites without consideration of the behavior of heavy metals in the soil. The behavior including adsorption, precipitation, transport, migration, etc. not only depends on the physiochemical properties of the metals but mostly on the physical and chemical properties of the soil, like, for example, soil organic matter content, clay fraction content, mineralogical composition, pH, temperature and so on.¹⁵⁹ A lot of research has focused on the two-dimensional (2D) investigation using Kriging method.^{74,81,83–85,87,160} However, three-dimensional (3D) model will be more accurate and visual to simulate the amount and distribution of heavy metals in the soil, and do better for the selection of remediation method and the amount of contaminated soil needed to be treated. There have been literatures using Kriging to estimate the 3D soil textures and soil strength in a durum wheat field,^{161,162} but few literatures dealing with the amount and spatial distribution of heavy metals in soil.

In chapter 2, the distribution of HMs on surface soil (2D) is introduced, however the soil considered in the research at depth is assumed as constant ($\leq 0.5\text{m}$). In real world, the distribution at depth is quite different, since the mobility of different kinds of HMs is different. The distribution is also affected by other factors, like the soil properties, precipitation, etc. As a result, the correct simulation of the distribution of HMs at depth exist much difficulty.

4.2 Review of existed several literatures about 3D distribution of HMs in soil

Several two-dimensional spatial prediction methods have been classified as global or local, interpolating or non-interpolating, and smooth or non-smooth predictors, and the techniques include like global means and medians, moving averages, inverse distance weighted (IDW) interpolation, Akima's interpolation, natural neighbor (N-N) interpolation, quadratic trend surface, Laplacian smoothing splines, local polynomial, radial basis

functions and ordinary kriging.^{147,163} Even though the difficulty in correctly simulating the 3D distribution of HMs in soil, three interpolation methods have been commonly used: Kriging-3D, Inverse Distance Weighted (IDW), natural neighbor(N-N).¹⁶⁴ Interpolation methods have been used in many researches to construct the 3D distribution model, like the HMs in the sediments or soils,^{165–168} soil strength,¹⁶² soil magnetic susceptibility,¹⁶⁹ soil organic carbon,¹⁷⁰ soil nitrate-nitrogen,¹⁷¹ hydraulic conductivity,¹⁷² clay content,¹⁷³ etc. The three interpolation methods will be introduced in detail as follows:

4.2.1 Inverse Distance Weighted (IDW)

Inverse distance weighted (IDW) is a deterministic, nonlinear interpolation technique that uses a weighted average of the attribute (i.e., phenomenon) values from nearby sample points to estimate the magnitude of that attribute at non-sampled locations.¹⁷⁴ The assigned values to unknown points are calculated with a weighted average of the values available at the known points. The basic equation used in IDW interpolation is

$$Z(x) = \sum_{i=1}^n w_i c_i$$

Where n is the number of samples, w_i are the weights, and c_i are the concentration values at each sample. Inside the function, the weights (w_i) is inversely proportional to distance, as the following function shows:^{130,164}

$$w_i = \frac{1/h_i^p}{\sum_{j=1}^n 1/h_j^p}$$

where h_i is the distance from the sampling location to the interpolation point; p is the power parameter, which determines how fast the weights decay with distance. In most cases, $p = 2$. The alternate form of the weight function has also been used:

$$w_i = \frac{[(R - h_i)/Rh_i]^2}{\sum_{j=1}^n [(R - h_j)/Rh_j]^2}$$

where R is the distance from the interpolation point to the most distant sample.

One simple example to understand the original function is shown in following **Figure 4.1**. Here, the green points represent known sample points and the red point is an unknown location for which we would like to estimate a value (Z(x)). The distances between sample points and the unknown point are shown in black text, while the attribute values of the sample points are shown in blue. You can see how the distance affects the influence of a point on the estimated value in the equation at the right: nearer points has a significantly greater effect on the estimated value than more distant points.¹⁷⁴

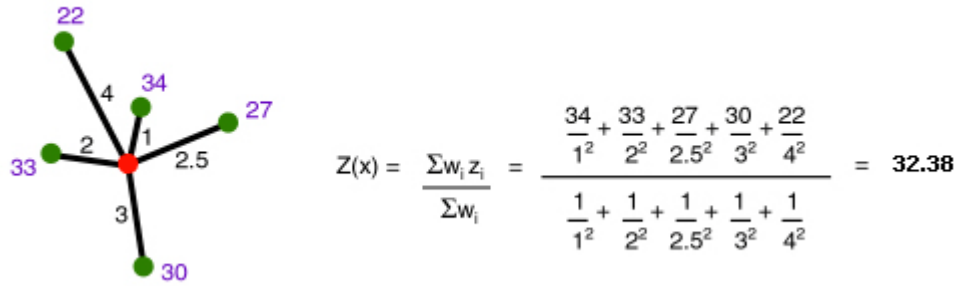


Figure 4.1 One example to understand the IDW original function.¹⁷⁴

4.2.2 Natural Neighbors (N-N)

Natural neighbor (N-N) interpolation was first introduced by Sibson (1981). N-N interpolation share the same function with the IDW, however the method to compute the weight function is different. Before going into further understanding the N-N interpolation method, the definitions of the Voronoi diagram and Delaunay triangulation as well as the relationship between them, which build the blocks of N-N interpolation, should be considered.¹⁷⁵ Here we do not introduce in detail, and the definition, function and features have been introduced in detail in the literatures or websites.^{163,164,175–179}

4.2.3 Kriging-three dimensional (kriging-3D) interpolation

Kriging 2D interpolation has been introduced in detail in chapter 2. Kriging-3D interpolation is based on the kriging-2D interpolation, and simultaneously considers the distribution of HMs at depth in the soil. To construct the 3D model, we should build (or compute) the experimental variogram $\gamma_1(h_1)$ as a function only of the separation distance h_1 in the horizontal direction and then we build the experimental variogram $\gamma_2(h_2)$ as a function only of the separation distance h_2 in the vertical direction.¹⁵⁸ Assume two variograms with the same sill value 0.06, as shown in **Figure 4.2**. The experimental variogram $\gamma_2(h_2)$ (red line) reach the sill value at the shorter range (about 10m) compared with that (about 30m) of the $\gamma_1(h_1)$ (blue line), because there is much more correlation in the horizontal direction than there is in the vertical direction. At last, we can combine the two experimental variograms into one to get the experimental variograms for the kriging-3D interpolation. The function of the krigng-3D can be described as follows:

$$\gamma(\Delta x, \Delta y, \Delta z) = C_0 + C \left[1 - \exp \left\{ - \sqrt{\left(\frac{\Delta x}{a_x} \right)^2 + \left(\frac{\Delta y}{a_y} \right)^2 + \left(\frac{\Delta z}{a_z} \right)^2} \right\} \right]$$

where Δx , Δy , Δz are the distance between the interpolated location with the existed sampling locations in the x , y (x , y horizontal direction) and z axis (vertical direction) respectively; $\gamma()$ is the experimental variogram; C_0 is the nugget value, the sum of C and C_0 equals sill value; a_x , a_y , and a_z are the correlation scale in the x , y , and z direction separately.

If the spatial correlations structure of the data set depends only on the separation distance, or no anisotropy was detected in the horizontal direction, the omni-directional variogram should be selected. For example, Chong Chen et. al used the omni-directional variogram in the 3D mapping of soil organic carbon,¹⁸⁰ and Ying Ouyang et.al selected the omni-directional variogram in the 3D kriging analysis of sediment mercury distribution;¹⁶⁶ in

addition, H. Y. Li et. al didn't observe the directional effect in the horizontal plane when mapping the 3D variability of soil electrical conductivity, and an omni-directional horizontal variogram was calculated.¹⁸¹ However, in the case of the anisotropic case, the variogram depends on the separation vector $\mathbf{h} = [h_1, h_2, h_3]$ instead of the scalar separation distance $h = \sqrt{h_1^2 + h_2^2 + h_3^2}$. That means, in the anisotropic case, the separation between two points not only depends on the scalar distance but also on the orientation.¹⁵⁸ As a result, the directional experimental variogram can deal with the degree of anisotropy of the data. Some researches use different number of the directions, for example, Katsuaki Koike et.al constructed the experimental variogram for six directions when doing the three-dimensional distribution analysis of phosphorus content of limestone through a combination of geostatistics and artificial neural network.¹⁸²

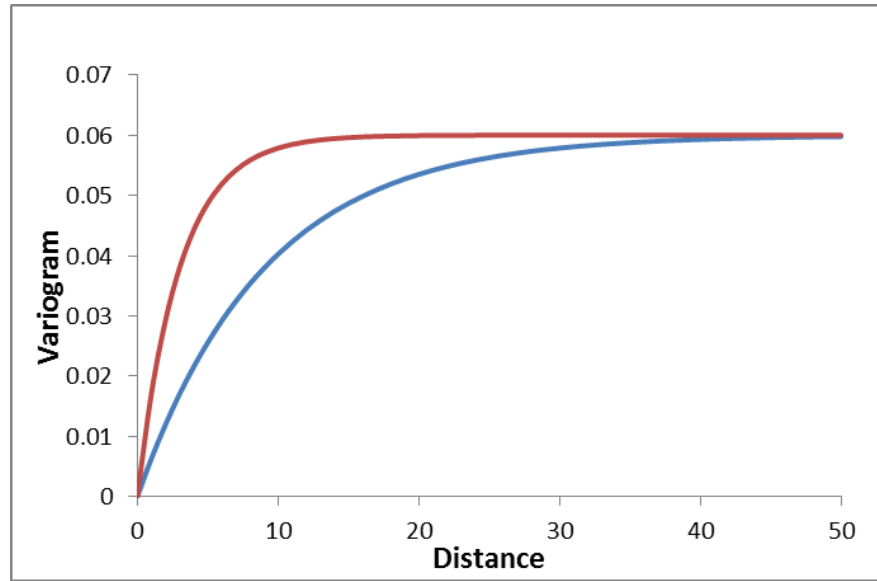


Figure 4.2 Variograms in horizontal and vertical direction. $\gamma_1(h_1)$ is shown as the red line, while $\gamma_2(h_2)$ is shown as the blue line.

4.2.4 Discussion and comparison of the three interpolation methods: IDW, N-N, Kriging-3D.

The method of the IDW, which is sensitive to clustering of sites and cannot reproduce the planar surface, is popular in the mining industry. Because the weights are positive, it cannot track the peaks and valleys.¹⁶³ For the sampling data, which don't show the existence spatial structure, IDW is a better choice than OK with a nugget model. The IDW has the advantage that it is intuitive and efficient hence recommended for the nominal data, and compared with kriging, it has more simple procedure and fewer steps.¹⁸³ The accuracy of both IDW and kriging is described almost the same, in most literatures, though the kriging was considered to superior "on average", IDW was superior under certain conditions.¹⁶⁴

The weights of N-N interpolation is based on both distances and topological relationships, N-N interpolation can be more accurate than IDW interpolation when the sampling data are clustered.¹⁶⁴ However finding the appropriate distance for N-N interpolation to select neighbors is difficult, and it requires a priori knowledge of a dataset.¹⁷⁸ In addition, N-N interpolation has its own disadvantage that it can lead to strange results. Like the **Figure 4.3** shows interpolation at point 0 depends on the value at point 1 but not point 2.¹⁶³

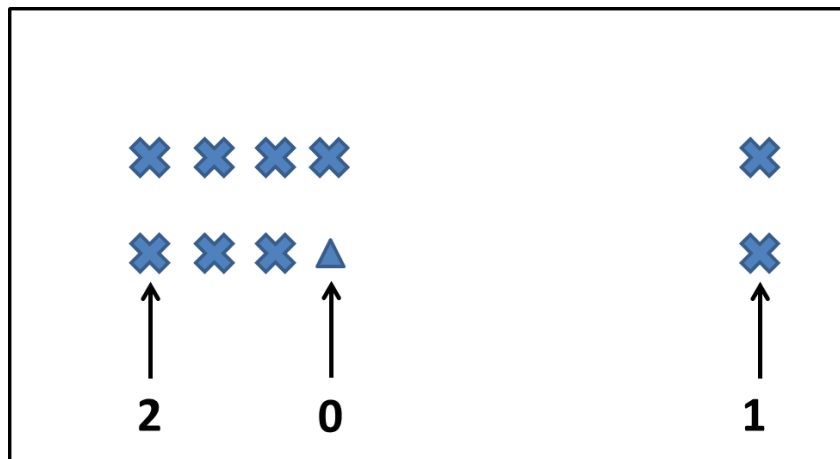


Figure 4.3 Natural neighbour interpolation for irregularly spaced data. The crosses represent observation points and the triangle represents a prediction point. The predicted value at 0 depends on the observed value at 1 but not at 2.¹⁶³ Although the advantage of N-N interpolation to deal with the scattered and irregularly spaced data, its implementation is still a problem in practice, especially in three and higher dimensions.¹⁷⁸ Compared with

the IDW and N-N, kriging-3D to give the most accurate results and provides the additional benefit of variography that aids in the interpreting the site characteristics.¹⁶⁴

Even though kriging-3D has its own advantage compared with IDW and N-N interpolation, it is the statistical methods, which only deals with the relationship of among sampling data. In some way, it cannot correctly describe the distribution of HMs at depth, since it lose the respect on distribution of HMs at depth from the view of the transport or migration mechanism of HMs in soil. Besides, the sampling data always don't show strong spatial correlation even no spatial correlation at depth, but the concentration at depth always have the strong relationship with concentration on surface soil or up-layer soil, as a result, it may be a good idea to combine the kriging (on surface soil) and migration or transport mechanism of HMs (at depth) to construct the 3D distribution model.

4.3 Feasibility of the development of 3D model: considering the migration or transport mechanism (function) of HMs at depth.

Soil is a so complicated and heterogeneous system with so many different components. Soil may quite differ from place to place. As a result, it is difficult to model the migration or transport of HMs in the vertical direction in soil; however there are some literatures that well describe and conclude the general mechanisms and pathways of the mobility of HMs in soil.^{159,184} In the following sections, the content is generally introduced not in detail.

4.3.1 Mechanisms and pathways of the mobility of HMs in soil

There are three main basic processes that control the mobility of HMs in soils: physical, chemical, and biological. For the physicochemical process, most HMs distributed in soil is bonded to the soils' solid surface, while small portion of HMs is dissolved in the soil solution. Inside the soil solution, most HMs is bonded to the organic acid, while the remained part is distributed in the form of inorganic ion pairs and free ions. The partition of HMs between solid-liquid phases in soils is always reflected by K_d coefficient, and the function is as follows:

$$K_d = \frac{\text{Sorbed metal}}{(\text{dissoved metal})^n}$$

n is freundlich-exponent, usually $n = 1$.

HMs also happen the complexation with dissolved natural organic matter (DOM). Some plants (roots) can modify their rhizosphere by uptaking the HMs or just releasing the compounds to complex with HMs in the rhizosphere (as shown in [Figure 4.4](#)).¹⁸⁵ Except for the plants, the microbes and invertebrates can also affect the migration or transport of HMs in soil

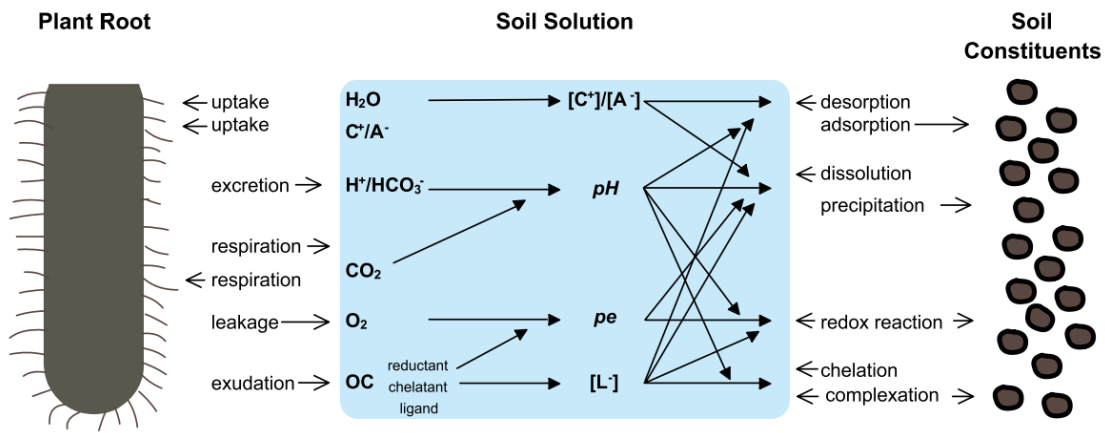


Figure 4.4 Schematic of the rhizosphere, showing the various exudates and how they can influence abiotic factors and mechanisms in the soil solution interface. Legends: OC=organic carbon; C⁺ = cation; A⁻ = anion; L= ligand; pe = redox potential.¹⁸⁵

4.3.2 Factors affecting the mobility and transport of HMs in soil

There are many factors affecting the mobility and transport of HMs in soil, such as soil pH, chemical speciation, soil organic matter, fertilizers and soil amendments, redox potential, clay content and soil structure, as shown in [Table 4](#).¹⁸⁴ A useful empirical regressions can be used to estimate the concentration of HMs dissolved in soil solution; the model can be given as the following function:

$$\begin{aligned} \text{Log}_{10}(\text{dissolved metal}) = & a + bpH + c\text{Log}_{10}(\text{total soil metal}) \\ & + d(\text{soil organic matter}) \end{aligned}$$

Table 4 Effects of Soil Factors on Trace Metal Mobility and/or Bioavailability^{184,186}

Soil factor	Affected process	Effect on process
Low pH	Decreasing sorption of cations onto oxides of Fe and Mn	Increase
	Increasing sorption of anions onto oxides of Fe and Mn	Decrease
High pH	Increasing precipitation of cations as carbonates and hydroxides	Decrease
	Increasing sorption of cations onto oxides of Fe and Mn	Decrease
	Increasing complexation of certain cations by dissolved ligands	Increase
	Increasing sorption of cations onto (solid) humus material	Decrease
	Decreasing sorption of anions	Increase
High-clay content	Increasing ion exchange for trace cations (at all pH)	Decrease
High-swelling clays	Forming structured soils, which allow bypass flow	Increase
High OM (Solid)	Increasing sorption of cations onto humus material	Increase
High-(soluble) humus content	Increasing complexation for most trace cations	Decrease/increase
Competing ions	Increasing competition for sorption sites	Increase
Dissolved inorganic ligands	Increasing trace metal solubility	Increase
Fe and Mn oxides	Increasing sorption of trace cations with increasing pH	Decrease
	Increasing sorption of trace anions with decreasing pH	Decrease
Low redox	Decreasing solubility at low Eh as metal sulfides	Decrease
	Decreasing solution complexation with lower Eh	Increase/decrease

4.3.3 Existed transport modeling (or functions)

A lot of transport modeling (or functions) have been introduced in the literatures, like Richards equation for the variably saturated water flow, dual-porosity or dual-permeability or multi-porosity or multi-permeability model for preferential flow, convection-dispersion equation, Freundlich or Langmuir isotherms for sorption, cation exchange function, precipitation-dissolution process, water flow equation, solute transport equation, mass balance equation, complexation reaction equations, charge balance equation on the exchange site, etc. ^{184,187} While convection-dispersion equation was always used in the simulation of transport of HMs in the soil. One of the most well-known models was based on dispersion-convection-sorption equalizations as following functions: ¹⁵⁹

$$\frac{\partial C}{\partial t} = \frac{D_d \partial^2 C}{\partial x^2} - \frac{v \partial C}{\partial x} - \frac{(\frac{\rho}{\theta}) \partial S}{\partial t}$$
$$\left(1 + \frac{(\frac{\rho}{\theta}) \partial S}{\partial C} \right) * \frac{\partial C}{\partial t} = \frac{D_d \partial^2 C}{\partial x^2} - \frac{v \partial C}{\partial x}$$
$$S = \frac{m_a}{m_s}$$

where: ρ - concentration of pollutant in soil solution; C - total concentration of pollutant in soil; θ - volumetric moisture content; D_d - dispersion coefficient; m_a - mass of pollutant in unitary volume of soil; m_s - mass of sorbent in unitary volume of soil.

These models can be divided into two groups: single-component model, which only describe the transport of single solutes, and multi-components model that simulate the transport and interaction of multiple solutes, simultaneously. These models or functions are complicated to calculate, while there existed some software that have been widely used in the researches, such as MACRO, PHREEQC, HYDRUS, POLLUTE, etc. ^{184,188}

4.4 In summary and future work

This chapter mainly review the interpolation methods, which, when used in the 3D interpolation that considering the statistical correlation of sampling data, are not enough to simulate the distribution of HMs in soils at depth, since the sampling data at depth may has only strong spatial correlation with the sampling data on the surface. As a result, we suggested the combination of kriging estimation on surface soil and transport/migration model (functions) at depth to construct the 3D model. And at the end of this chapter, we discussed the existed transport functions and affecting factors to better understand the distribution or migration/transport of HMs in soil. The future work, a real polluted site will be selected to construct and validate such 3D model that combines kriging and transport/migration model (functions).

CHAPTER V CASE STUDY IN THE POLLUTED FOREST SOIL (3D MODEL) IN KAWAUCHI VILLAGE, IN FUKUSHIMA PREFECTURE, JAPAN

5.1 Introduction

5.1.1 Fukushima Nuclear Dai-Ichi Power Plant (FNDPP) accident

As the description partly mentioned in the section 1.1.2, chapter 1, on 11th, March happened very serious earthquake and tsunami, which caused the accident of FNDPP. It was estimated that ^{131}I (153-160 PBq), ^{129}I (5.71 GBq), $^{137,134}\text{Cs}$ (13-15 PBq), $^{239,240}\text{Pu}$ (1.0-2.4 TBq), ^{241}Pu (1.1-2.6 TBq), ^{238}Pu (2.9-6.9 TBq) have been released to the atmosphere, while less than 20% deposited on the landscape. The affected area is nearly 10,000km² (within 80km), where about 80% is forest, about 15% is agricultural land and about 5% is inhabited area. About 30% of the forest plants or trees are evergreen coniferous species, while 50% are the deciduous broadleaf species, and the cultivated lands are dominated by paddy fields.

To obtain fundamental information and take corresponding effective and appropriate countermeasures against the impacts of radionuclides, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) commissioned the Japan Atomic Energy Agency (JAEA) in June 2011 to investigate the level of radioactive contamination as a leading role, including mapping the radionuclide distribution in the ground soil around the FDNPP.¹⁸⁹

After two months of the preparation including a pilot study, three campaigns had been completed, and at last finished the distribution-mapping project that started in June, 2011 by the end of 2012, take the ^{134}Cs -distribution as the example (Figure 5.1, other radioactive like ^{137}Cs , $^{110\text{m}}\text{Ag}$, and the distribution map of air dose rates at 1 m above ground, please refer to the original reference):¹⁸⁹ C. V. Srinivas et.al. also assessed the radiation doses and lifetime attributable health risk to the members of public in a 40-km impact zone around the FDNPP during 11-31March 2011, and radiation risk models were

adopted to estimate the health risk for leukemia, breast, thyroid, and all solid cancers using simulated distributed lifetime organ doses, as shown in [Figure 5.2](#).¹⁹⁰

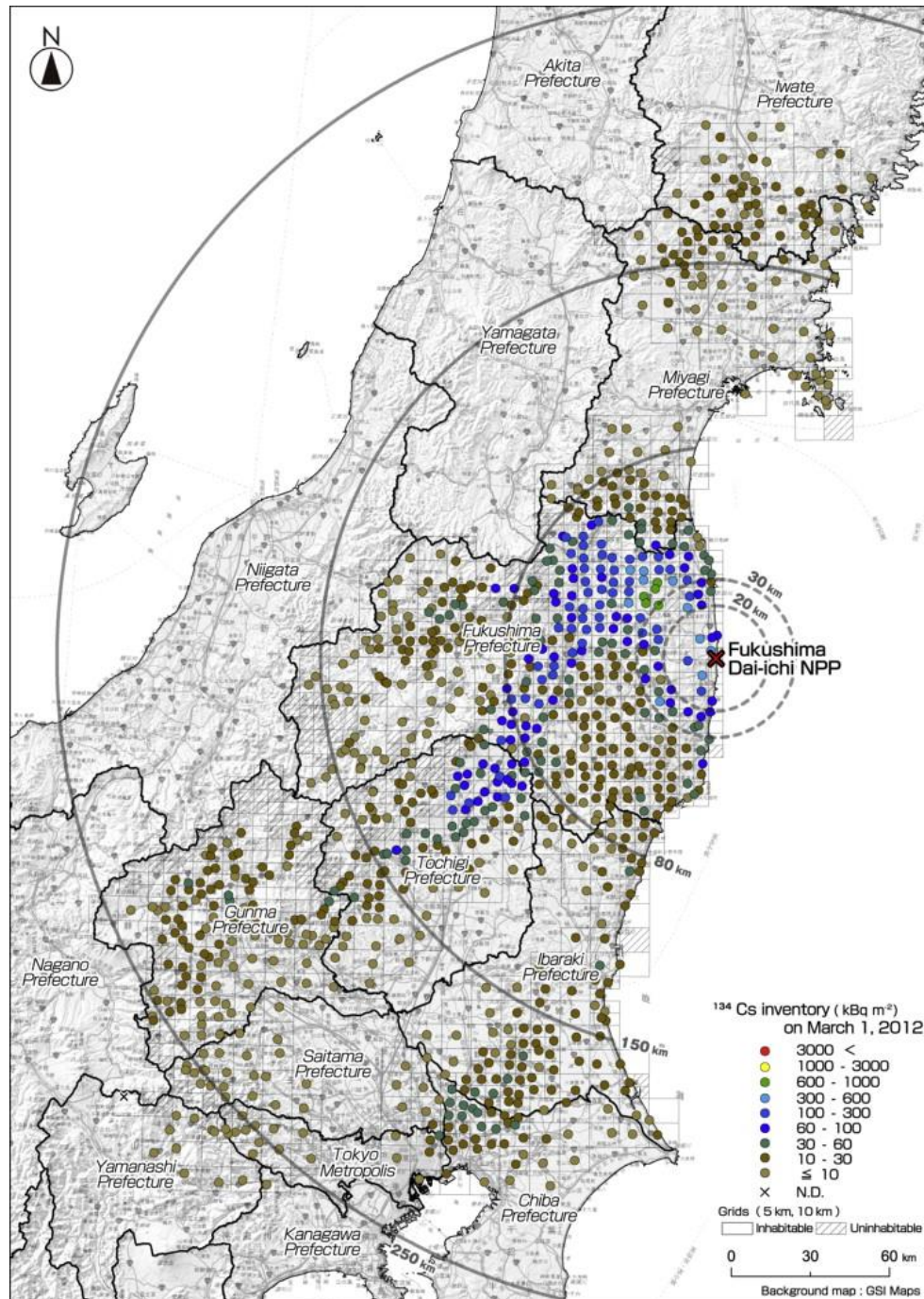


Figure 5.1 Distribution map of ^{134}Cs inventory obtained using in situ gamma spectrometry using b of $1.30 \text{ g}\cdot\text{cm}^{-2}$ (b is the relaxation mass depth). Activities were decay-corrected to March 1, 2012.¹⁸⁹

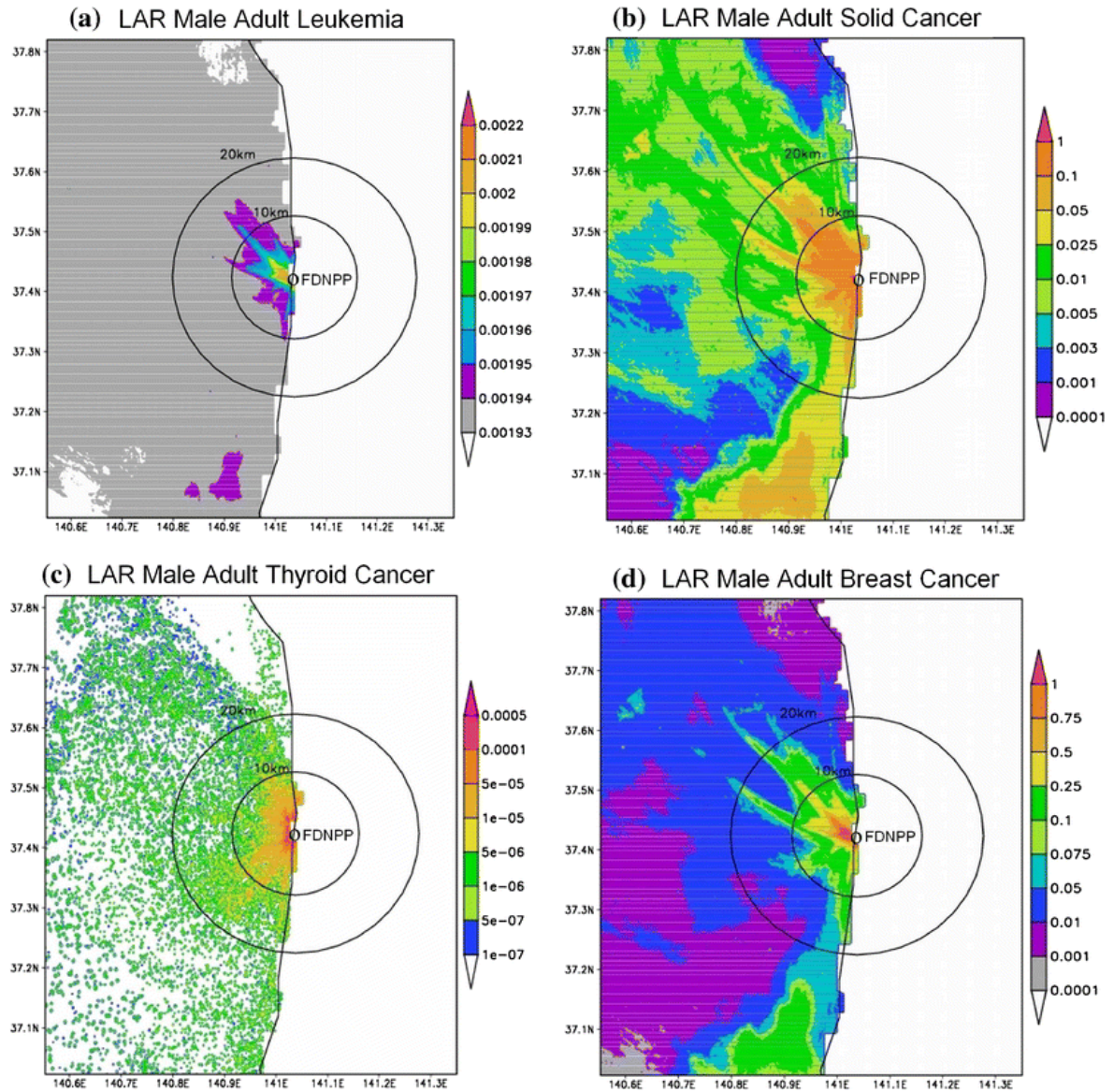


Figure 5.2 Estimated lifetime attributable risk (LAR) for **a** leukemia, **b** all solid, **c** thyroid, and **d** breast cancers for adult age group.¹⁹⁰

5.1.2 Introduction of cesium

Cesium is a member of the alkali family, which includes lithium, sodium, potassium, rubidium, and francium. The basic information for Cesium is shown in [Table 5](#). Cesium is considered the most active metal, though in theory it is after francium, which is too rare to have any commercial uses.¹⁹¹ Cesium has been found in rocks, soil, and dust at low concentration as a naturally-occurring element. The average concentration in Granites and

sedimentary rocks are ppm, 4ppm respectively. Natural cesium is stably present in the environment in a formation of isotope ^{133}Cs . Cesium is a mineral known as pollucite, usually about 5-32% of which is cesium oxide (Cs_2O). Radioactive forms of cesium are generated by the fission of uranium from like the normal operation of nuclear power plants, explosion of nuclear weapons, and accidents involving nuclear power plants or nuclear powered satellites or submarines.¹⁹² ^{134}Cs and ^{137}Cs are the two most important radioactive isotopes of cesium, which will decay into stable isotopes: xenon 134 (^{134}Xe) or barium 134 (^{134}Ba) and barium 137 (^{137}Ba) respectively.¹⁹² In the process of decaying, beta particles and gamma radiation are given off, and the rate of decay can be described by the curie (Ci) and the Becquerel (Bq) (1 Bq = 27 pCi.) The quantities and units (conversions) has been well described in the references.^{193,194} Nonradioactive (stable) cesium can be neither created nor destroyed under typical environmental conditions, but can change into different cesium compounds by reacting with other compounds found in the environment. Cesium compounds can travel a long distance in the air before falling on the earth by rainfall and gravitational settling. In water and moist soils, most cesium compounds are very soluble. Cesium can binds strongly to the soil, and hinder its migration far below the soil's surface. Therefore it is hard for cesium to enter the plant through uptake by roots, except that if falls onto the leaves of plants.¹⁹²

Table 5 Basic information of Cs.¹⁹⁵

Atomic Number	55	Number of Protons/Electrons	78
Atomic Mass	132.90546 amu	Crystal Structure	Cubic
Melting Point	28.5 °C	Density (293 K)	1.873 g/cm ³
Boiling Point	678.4 °C	Color	Silver white
Isotopes (Half-life)	Cs-126 (1.6 minutes); Cs-129 (1.3 days); Cs-131 (9.7 days); Cs-132 (6.4 days); Cs-133 (Stable); Cs-134 (2.1 years); Cs-134m (2.9 hours); Cs-135 (2300000.0 years); Cs-136 (13.2 days); Cs-137 (30.2 years); Cs-138 (32.2 minutes); Cs-139 (9.3 minutes)		

5.1.3 Migration of Cesium in the soil

The transport or migration of cesium in the soil has been reported very low in many literatures, and the most of the cesium retained within the top 5 cm of the soil surface, especially in the forest soil that is rich in organic matters on top soil layers.^{47,196–198} While based on the research on Mediterranean hillslopes, the cesium will reach even deeper to 47cm in cultivated soils compared with the uncultivated soil that in the upper 12cm.¹⁹⁹ According to previous studies in section 4.3.2, the initial distribution of radioactive cesium in surface soils and their post-depositional redistribution within the soils are influenced by various physical, physicochemical, and biological processes. However, as the radioactive cesium especially ¹³⁴Cs (residence half time is 2.06 year), the decay of the cesium should be taken into the account when consider the distribution or mobility of cesium in soil at depth. To approximate the distribution of the cesium in soil at depth, some fitting function, such as the Gaussian, double exponential, and polynomial functions, have been proposed or developed.¹⁹⁶ The exponential function or model has successfully described the young fallout in the first years after deposition, even though it has the difficulty in describing the depth distribution after longer time periods.²⁰⁰ There are three main transport models (dispersion-convection model, residence time model, and back-flow model) to obtain the frequency distributions of the characteristic transport parameters.²⁰¹

5.1.3.1 Convection-dispersion model

Different from the dispersion-convection-sorption equalizations mentioned in the 4.3.3, the transport of cesium, especially the ¹³⁴Cs that has the low residence half-time, should consider the decay constant as shown in **Table 6**. While in most cases, the convection–dispersion equation (CDE) model produces good descriptions for near-surface soil layers.²⁰² The CDE function is often described as following function:

$$\frac{\partial A}{\partial t} = \frac{\partial}{\partial t} \left(D \frac{\partial A}{\partial z} - VA \right) - \lambda A$$

With A —the total ¹³⁷Cs activity per volume of soil (Bq/cm³), z —the soil depth (cm), t —the time (year), D —the diffusion coefficients (cm²/year), V —the convection velocity (cm/year) and λ —the radioactive decay constant (year⁻¹)

After considering the initial boundary condition with a pulse-like input of radionuclides, one approximate solution is like the following function:

$$a(z, t) = a_0 \exp(-\lambda t) \frac{1}{2[\pi Dt]^{\frac{1}{2}}} \times \exp \left\{ -\frac{[z - Vt]^2}{4Dt} \right\}$$

With a_0 is the initial surface concentration, $a(z, t)$ is the concentration at depth z and at time t after the contamination. In addition, Marianna Jagercikova et.al and S. Almgren et.al also give the relative complicated solution for the function.^{200,201,203} The distribution of the fallout in undisturbed soils can also be approximated by the following exponential functions based on the Beck's research, which the parameters of the model have been greatly decreased:^{204,205}

$$C(x) = C_0 \exp(-\alpha x)$$

Where $C(x)$ (Bq/kg) is the radionuclide concentration at depth x , C_0 is the radionuclide concentration as x approaches zero, and α -factor (cm^{-1}) is the characteristic coefficient of the distribution.

Table 6 Radioactive nuclide half-lives and decay constants.⁴⁴

Nuclide	Half-life ($T_{1/2}$)	Radioactive decay constant λ (d^{-1})
Cs-134	2.06y	9.22×10^{-4}
Cs-134	30.0y	6.33×10^{-5}
I-131	8.04d	8.62×10^{-2}
Sr-89	50.5d	1.37×10^{-2}
Sr-90	29.1y	6.53×10^{-5}

However the model has the limitation that it fails to describe “young” profiles shortly after fallout.²⁰²

5.1.3.2 Compartment model

Compared with the CDE, the compartment model is a relative simple model. In compartment analysis, the soil is divided into N layers of arbitrary thickness, usually 1–2 cm.²⁰¹ The distribution of radionuclides is considered uniform in each compartment. The movement of the radionuclides is represented by a series of transfers among compartments. The transfer rate of the neighbor tow layer (the layer i and the next layer $i + 1$) is assumed to be proportional to the amount of radioactivity in the layer i , but independent of the radioactivity in the layer $i + 1$. The transfer rate, which is defined by the coefficient k as shown in **Figure 5.3**, also means the removal of radioactivity from the compartment. The function can be written as follows:

$$\frac{dC_i}{dt} = -\lambda C_i + k_{i-1,i} C_{i-1} - k_{i,i+1} C_i,$$

$$i = 1, 2, \dots, n.$$

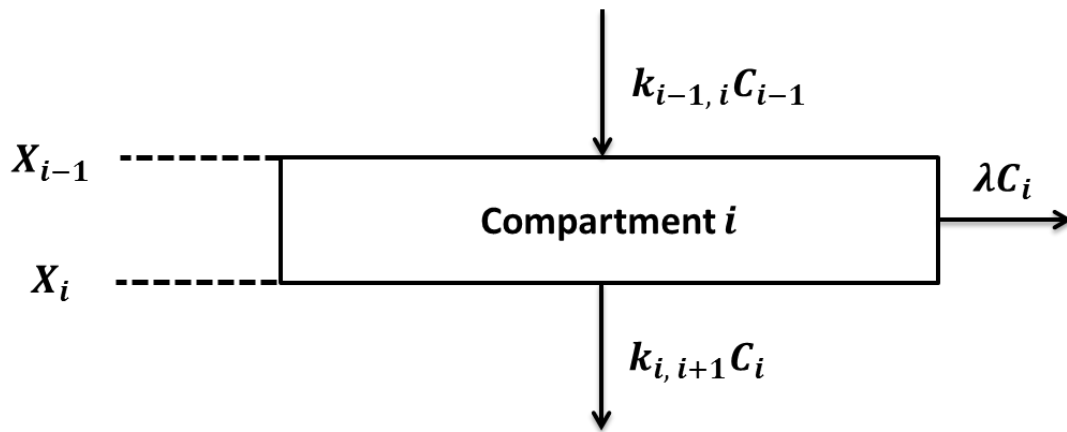


Figure 5.3 Representation of the i th compartment corresponding to the soil layer situated between the x_{i-1} and x_i depths. C is the radionuclide concentration in the i th compartment. The radioactivity input and output in this compartment are denoted k , which represents the intercompartment fractional transfer rate, and λ is the radioactive decay constant.²⁰⁴

The advantage of this compartment model is that it does not need any information on the actual transport processes of the radionuclide in the soil (e.g. sorption properties, speciation, water infiltration, etc.); however the disadvantage of this model is that one can only evaluate the residence times of a radionuclide in the various soil layers, but cannot obtain any information about the physicochemical processes process that controls the migration progress.^{206,207}

5.1.3.3 Back-flow model

Back-flow model is similar with the convection-dispersion model, but considering additionally the upward transport caused by the diffusion of radionuclides. In the back-flow model, the convective transport is small as compared to diffusion, and it is more appropriate to use a linear compartment model including both downward and upward transfer between adjacent soil layers. The model is like the following function based on the function at section 5.1.3.2:²⁰¹

$$\frac{dC_i}{dt} = -\lambda C_i + k_{i-1,i}C_{i-1} - k_{i,i+1}C_i + k_{i+1,i}C_{i+1}$$

$$i = 1, 2, \dots n.$$

Thus, downward transport can be due to diffusive and convective processes, whereas the upward transport is caused by diffusion only. All $k_{i+1,i}$ are assumed to be equal and are denoted as the *upward transfer rate*, since all $k_{i,i+1}$ are considered as *downward transfer rate*.

5.2 Materials and methods

5.2.1 Study area

The study was conducted in Kawauchi village, Fukushima prefecture, Japan, located in the area (about 50m×50m) about (37.378499 N, 140.856818 E). The sampling area is about 16.25km from the FNDPP in the south-west direction and 6.14km from Kawauchi village

in the direction of north-east, where is a forest with a small not-paved road inside with few people passing through there, as shown in [Figure 5.4](#).

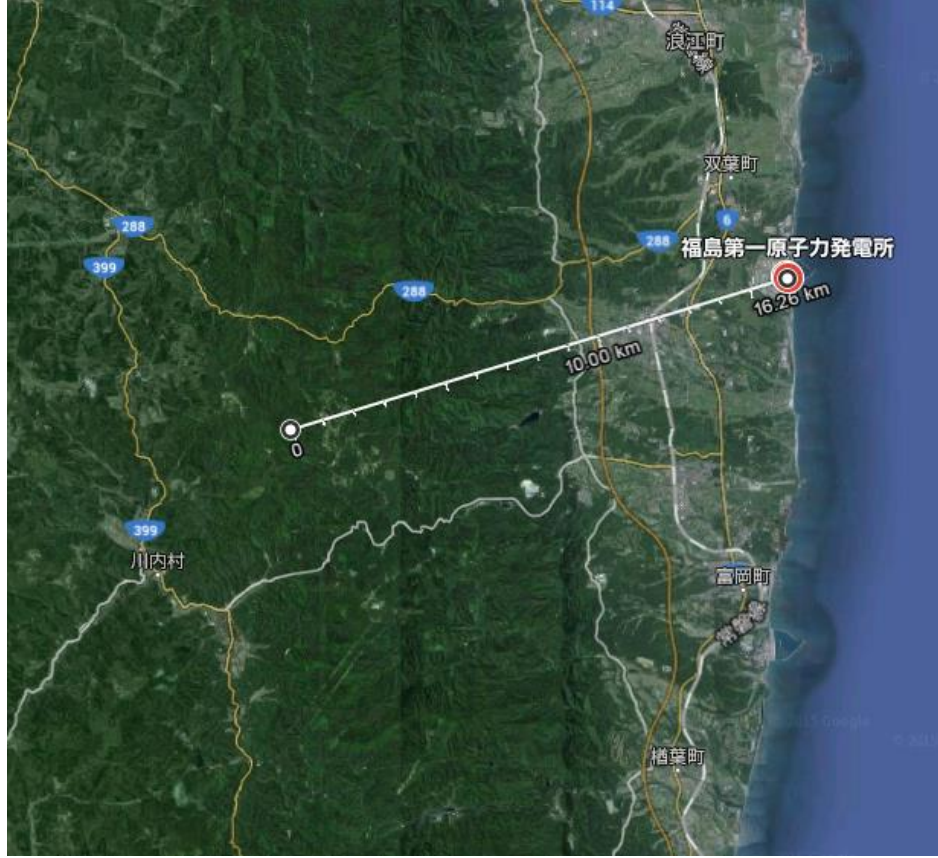


Figure 5.4 Sampling location in Kawauchi village, Fukushima prefecture, Japan

5.2.2 Sample collection and preparation

Because of the irregular distribution of the trees, the intervals of the samples keep a distance from 1.5m to 5m, even though 5m was planned before doing the samplings as shown in [Figure 5.5](#). We selected 107 locations for the sampling in this forest area (about 50m×50m), and at each location we took 6 samples from the surface to the 7cm at depth. That means we totally took $107 \times 6 = 642$ samples. The total 6 samples at each location were



Figure 5.5 The investigation of the polluted forest field and the sampling plan



Figure 5.6 Core sampling by the auger and the section description of one example soil layer



Figure 5.7 The soil sample that was put in the sealed plastic bag



Figure 5.8 The soil sample and the plastic bag were put in the plastic box before the detection of the radioactivity

taken from each layer divided beforehand (the surface: organic layer; soil: 0-1cm, 1-2cm, 2-3cm, 3-5cm, and 5-7cm) using a soil auger(section area: 20cm², length: 20cm) in 3rd-5th, November, 2015. At the each layer of 3-5cm and 5-7cm, one sample were taken (**Figure 5.6**), while at the other each layer, two samples were taken and mixed together to keep the volume of the sample the same. The sample was sealed into the plastic bag (**Figure 5.7**) separately and then put into a plastic box (**Figure 5.8**) waiting for the radioactivity measurement by Gamma spectrometry.

5.2.3 Radioactivity measurements by Gamma spectrometry

The radioactivity of the prepared sample was analysed by a high resolution gamma spectrometry. The gamma spectrometry system consist of a p-type coaxial hyper pure Ge-Detector (HPGe, EG&G ORTEC) connected to a multi-channel analyser, a detector shielding, amplifier, high voltage power supply. The background of the radioactivity was minimized by using the 10 cm thick shield of the lead. The equipment is shown in the **Figure 5.9**.

Counts per second (CPS) corresponding to the energy of the emitted photons of 606.43 keV emitted from ¹³⁴Cs (97.6%) and 661.65 keV emitted from ¹³⁷Cs (85.0%). The standard

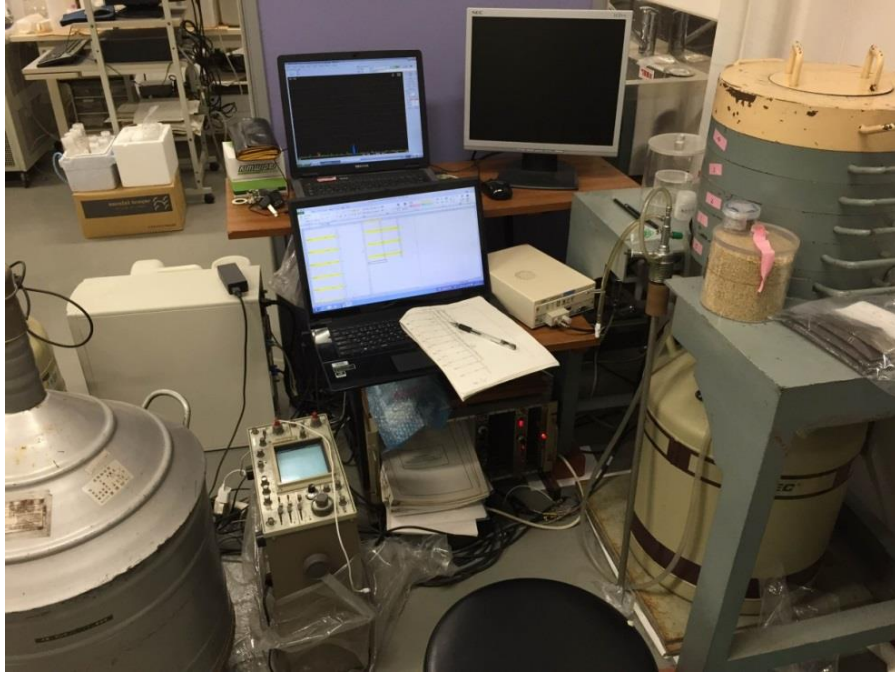


Figure 5.9 The equipment used to detect the radioactivity of the soil samples

geometry was used in accordance to the Japan Society for Analytical Chemistry (JSAC). The detector was properly calibrated using certified reference standard samples in the same geometry as the measured samples. All the measured activities were corrected for radioactive natural decay between the sample collection time and the radioactivity measurement time. The concentration of radiocesium was measured in Bq/cm² (for organic matter on the soil's surface) or Bq/cm³ (for the soil sample at depth from 0-7cm and organic matter) by a calculation using the following equation.

$$R = \frac{C}{\varepsilon * (A \text{ or } A * D) * P_{\gamma} * t}$$

Where:

R is the radioactivity concentration (Bq, Bq/cm² or Bq/cm³)

C is the number of counts per second (CPS)

ε represents the counting efficiency, here is 1.5697%

A is the section area of the auger (cm²)

D is the depth of the sample (cm)

P_γ is the emission probability ($^{137}\text{Cs} = 0.85$ and $^{134}\text{Cs} = 0.976$)

t is the measurement time for the sample, here is 6minutes=36000 (seconds/100)

5.3 Results and discussion

5.3.1 Sampling data analysis based on the ^{134}Cs

In this study, the sampling data analysis of ^{134}Cs was used instead of considering ^{137}Cs or both ^{134}Cs and ^{137}Cs , due to the following reasons:

- 1) Because of the relatively short half-life time as shown in [Table 5](#) or [Table 6](#), ^{134}Cs can better reflect the pollution of radioactive nuclides from FDNPP, since the resource of ^{137}Cs that has a long half-life time was affected both by the accident of FDNPP and the Nuclear Weapon Test (NWT).
- 2) There existed some linear relationship of the radioactivity strength (counts per second: CPS) between ^{134}Cs and ^{137}Cs based on the sampling data at different layers, e.g. in the litter layers as shown in [Figure 5.10](#). The data analysis of ^{134}Cs can reflect that of ^{137}Cs .

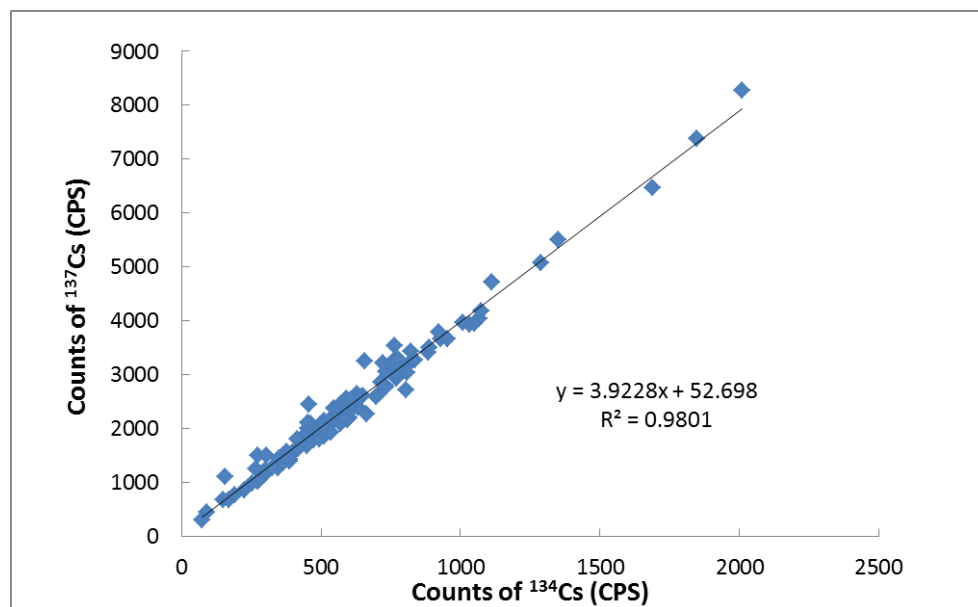


Figure 5.10 The relationship between the counts of ^{134}Cs and ^{137}Cs in the litter layer.

5.3.1.1 Sampling data analysis based on the radioactivity per unit area (Bq/cm^2)

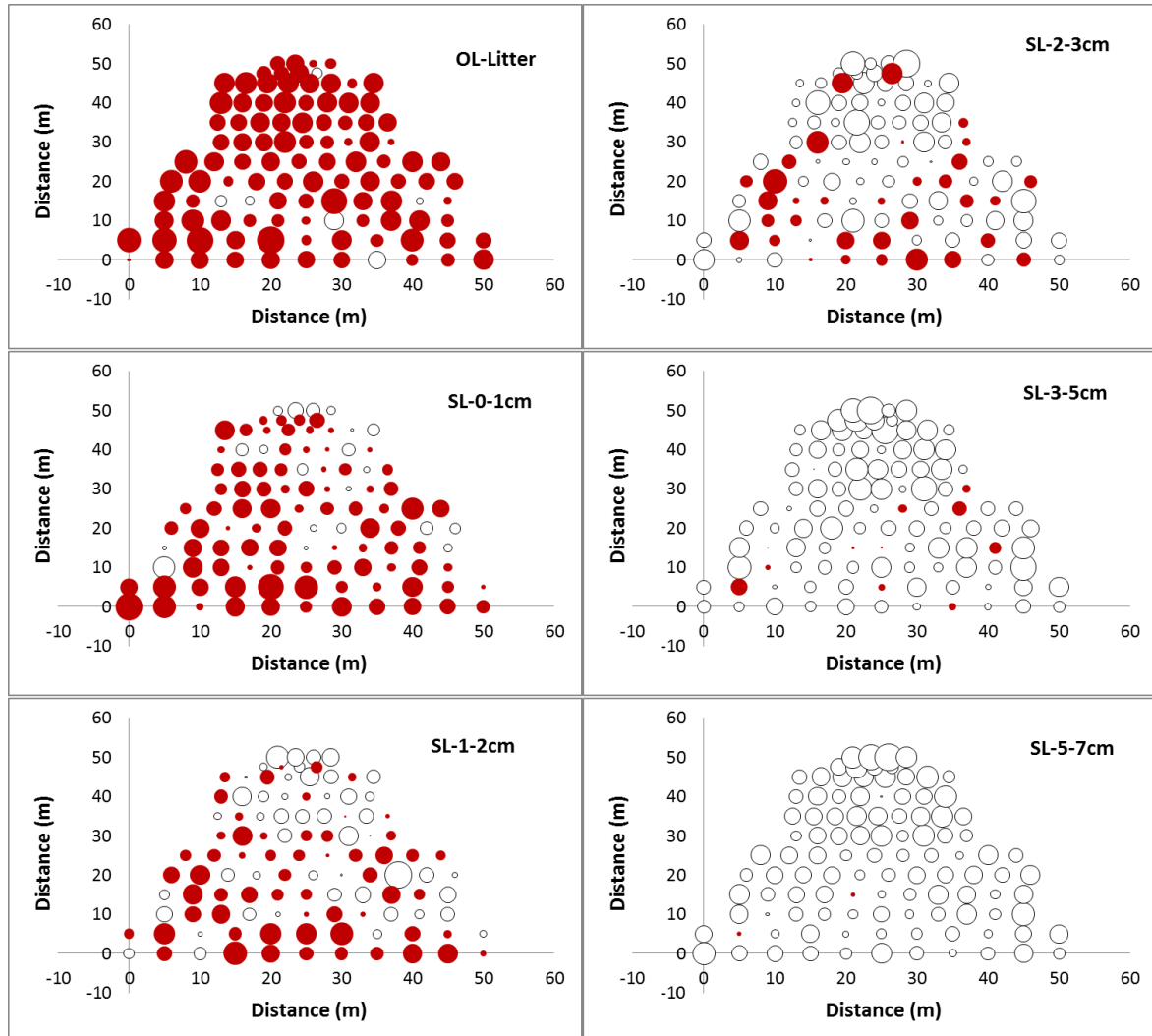


Figure 5.11 The distribution of ^{134}Cs at different layers based on the radioactivity per unit area (Logarithm data) (Bq/cm^2). The size of the circle represent the radioactive strength, while the red color circle (Log data ≥ 0) is positively proportional to the radioactive strength and the circle without color (Log data < 0) is inversely proportional to the radioactive strength. OL-organic layer; SL-soil layer.

Sampling data analysis based on the radioactivity per unit area (Bq/cm^2) includes the radioactivity analysis both on the surface soil (including the organic matter) and the soil at depth (from 0-7cm). The concentration of the radioactivity at each location is shown in **Figure 5.11**. From the **Figure 5.11** we can see that the radioactive strength of the ^{134}Cs on the surface forest soil (Organic layer that is mainly litter) was the largest, while the radioactive strength decreased as the soil depth increased. **Table 7** shows the correlation of radioactive strengths at different layers. As shown in **Table 7**, the neighbor soil layers always showed the strong correlation, and the OL has very weak correlation with SLs even the neighbor layer of SL-0-1cm. In addition, the upper soil layers had a decreasing trend of correlation as the soil depth increased.

Table 7 The correlation of radioactive strengths at different layers.

	Litter	0-1	1-2	2-3	3-5	5-7
Litter	1					
0-1	0.169452	1				
1-2	0.124789	0.629219	1			
2-3	0.09733	0.492177	0.73147	1		
3-5	0.048062	0.485004	0.544088	0.525121	1	
5-7	0.176373	0.363094	0.477508	0.406465	0.70435	1

5.3.1.2 Sampling data analysis based on the radioactivity per unit cubic (Bq/cm^3)

Sampling data analysis based on the radioactivity per unit cubic (Bq/cm^3) includes the radioactivity analysis both on the surface soil (including the organic matter) and the soil at depth (from 0-7cm). The results of the analysis are the similar as that based on the radioactivity per unit area (Bq/cm^2). The concentration of the radioactivity at each location is shown in **Figure 5.12**. From the **Figure 5.12** we can see that the radioactive strength of the ^{134}Cs on the surface forest soil (Organic layer that is mainly litter) was the largest, while the radioactive strength decreased as the soil depth increased. The correlation of radioactive strengths at different layers is the same with **Table 7**, since sampling data using the unit of Bq/m^3 is half the value of that using the unit of Bq/m^2 and don't affect the structure of the sampling data. As shown in **Table 7**, the neighbor soil

layers always showed the strong correlation, and the OL has very weak correlation with SLs even the neighbor layer of SL-0-1cm. In addition, the upper soil layers had a decreasing trend of correlation as the soil depth increased.

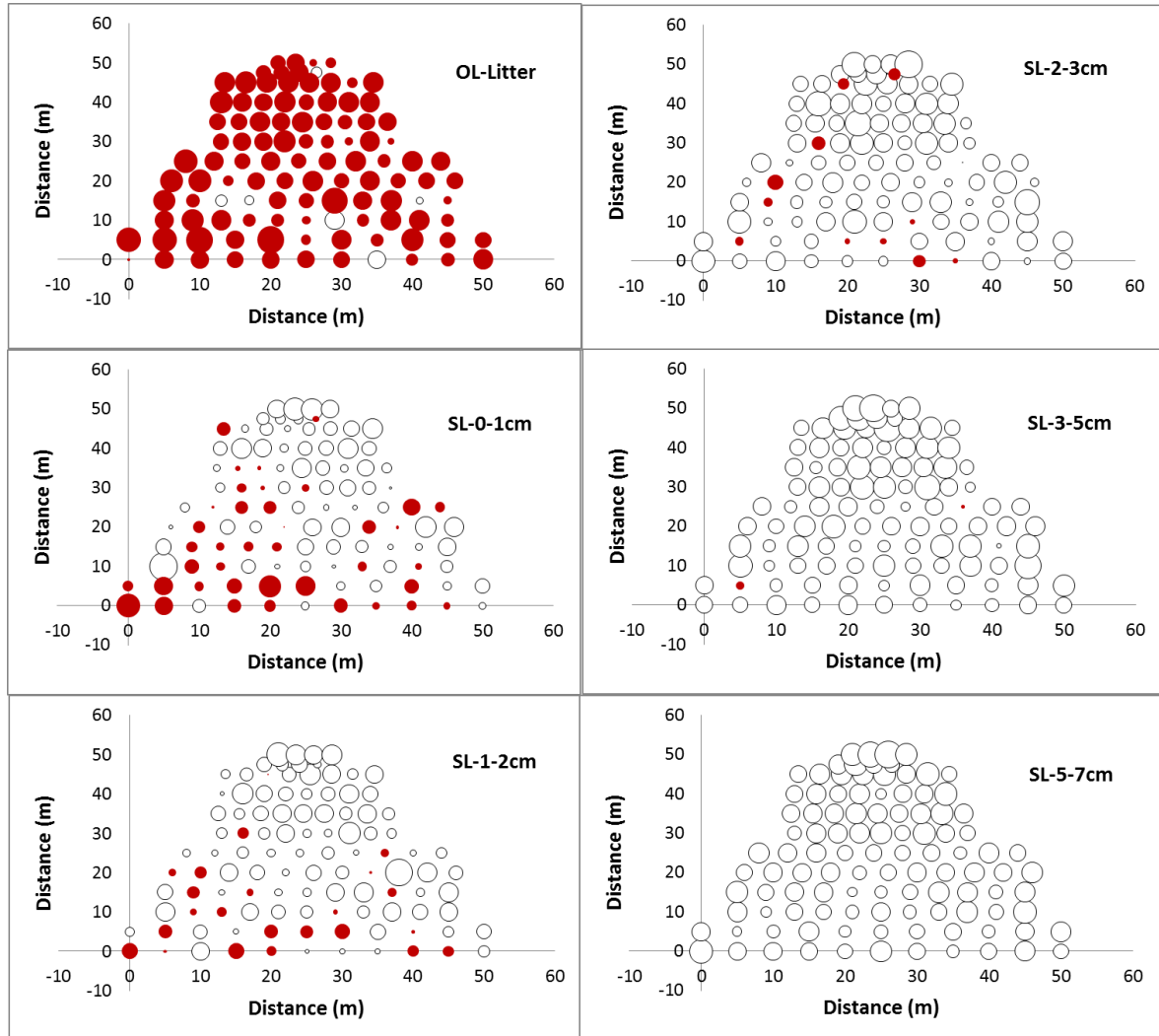


Figure 5.12 The distribution of ^{134}Cs at different layers based on the radioactivity per unit cubic (Bq/cm^3). The size of the circle represent the radioactive strength, while the red color circle (Log data ≥ 0) is positively proportional to the radioactive strength and the circle without color (Log data < 0) is inversely proportional to the radioactive strength. OL-organic layer; SL-soil layer.

5.3.1.3 Descriptive analysis of sampling data at depth.

The section of 5.3.1.1 and 5.3.1.2 is the descriptive analysis of sampling data in the horizontal direction, while the descriptive analysis of sampling data at depth is shown in **Figure 5.13**. The logarithm-transformed average value of total 107 sampling data showed a linear relationship with the soil layers (0-7cm), which also means the radioactivity (not logarithm-transformed data) is exponentially distributed at the soil depth. However, within the total 107 sampling data, there are 18 sampling data not showing linear relationship as shown in **Figure 5.14**, 12 sampling data as shown in **Figure 5.15**, and 2 sampling data showing the peak radioactivity at soil depth 3-5cm, which covers up $(18+12+2)/107*100\%=29.9\%$ of the total 107 sampling data.

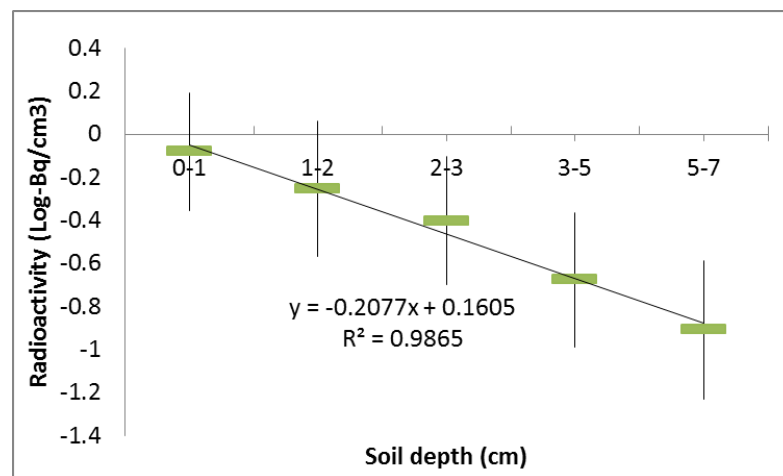


Figure 5.13 The relationship between the average radioactivity (Log, Bq/m³) of the total 107 sampling data and the soil depth (cm).

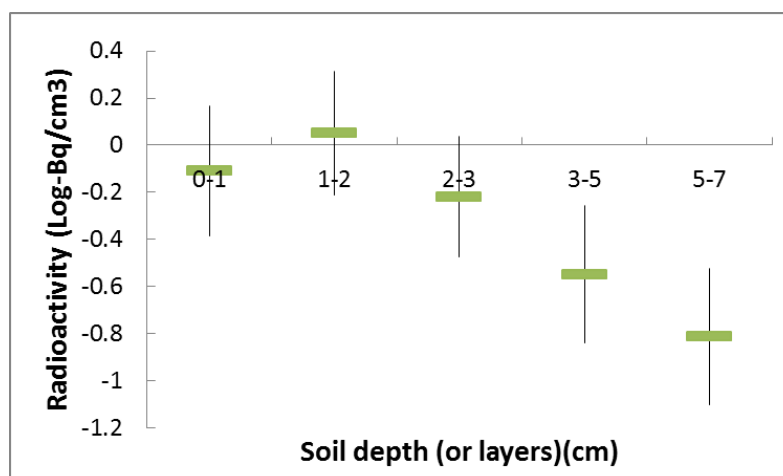


Figure 5.14 The relationship between the average radioactivity (Log, Bq/m³) of the 18 sampling data and the soil depth (cm).

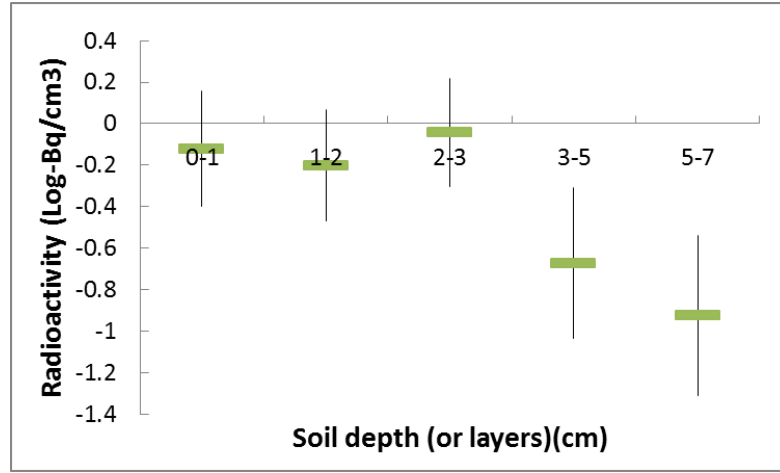


Figure 5.15 The relationship between the average radioactivity (Log, Bq/m³) of 12 sampling data and the soil depth (cm).

5.3.2 Construction of the 3D ¹³⁴Cs-distribution model

5.3.2.1 Selection and application of compartment model

Based on the Chapter4, we should construct the 3D model considering the migration of ¹³⁴Cs at soil depth. While considering the migration of ¹³⁴Cs, there are three types of model introduced in detail in section 5.1.3, chapter5. In this study, the compartment model is selected due to the following reasons:

- 1) The complex structure of the forest soil to use convection-dispersion model. Even though the average distribution of Cs at soil depth is usually approximate to the exponential function (Figure 5.13), not all sampling data follows the exponential distribution as shown in Figure 5.14-15. If arbitrarily use the exponential function to fit the sampling data, it will lead to inaccurate estimation. In addition, the forest soil contains very thick organic layer, which contains most the Cs and make the Cs hard to transport to soil layers (that means has a quite different migration rate of Cs between the organic layers and soil layers with that among the soil layers). As a

result, it is difficult to calculate the parameters, e.g. D —the diffusion coefficients (cm^2/year), V —the convection velocity (cm/year), if using the convection-dispersion model.

- 2) The back-flow model will introduce more parameters (e.g. upward transport rate between neighbour layers), which will complex the calculation of model. In addition, the upward transport caused by diffusion only can be neglected compared with the downward transport.

5.3.2.2 Assumptions for the compartment model

Though the compartment model is suitable for this study compared with other models, it is still based on the following assumptions:

- 1) The distribution of radioactive caesiums (R-Cs) is expressed by quantity of R-Cs in each layer. The layers are an organic layer (OL), 0-1cm soil layer (SL01), 1-2cm soil layer (SL12), 2-3cm soil layer (SL23), 3-5cm soil layer (SL35), and 5-7cm soil layer (SL57) (in this chapter, the unit of radioactivity used in the compartment model is Bq/compartment unit ($20\text{cm}^2 \times 2\text{cm}$)).
- 2) Only adsorption and vertical downward transport phenomena are considered. That is, the other phenomena, for examples, recirculation, absorption by plants, and upward movement, are neglected.
- 3) With the assumptions above, the distribution of R-Cs in forest soil is decided by three parameters: The first one is the total quantity of R-Cs existed at the organic layer for the first (15th March, 2011) expressed as M_0 . The second one is the transport coefficient from OL to SL01, the transport ratio of R-Cs in OL to SL01 in one month expressed as P_0 which is supposed to depend on characteristics of organic layer and climate condition. The third one is the transport coefficient between soil layers expressed as P_s , which is supposed to depend on characteristics of soil and location condition, is assumed same in all soil layers for the unit depth.

Therefore, the quantity of R-Cs in each layer at n-th month from the initial condition is expressed by the following equation and the flowchart is as shown in **Figure 5.16**:

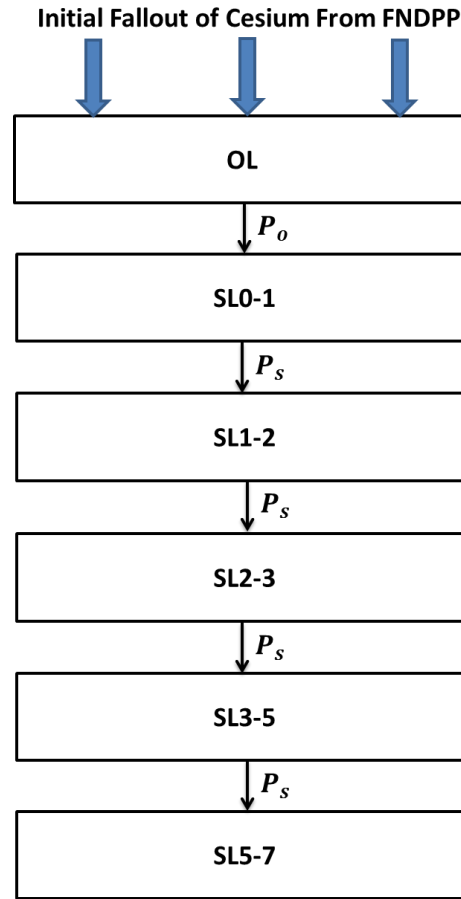


Figure 5.16 Conceptual scheme of the compartment model.

Mog(n) (The amount of R-Cs in the organic layer at n-th month):

$$\text{Mog}(n) = \text{Mog}(n-1) - \text{Mog}(n-1) * P_o = \text{Mog}(n-1) * (1 - P_o) = \text{Mog}(0) * (1 - P_o)^n$$

Ms01(n) (The amount of R-Cs in the 0-1cm soil layer at n-th month):

$$\text{Ms01}(n) = \text{Ms01}(n-1) + \text{Mog}(n-1) * P_o - \text{Ms01}(n-1) * P_s$$

Ms12(n) (The amount of R-Cs in the 1-2cm soil layer at n-th month):

$$\text{Ms12}(n) = \text{Ms12}(n-1) + \text{Ms01}(n-1) * P_s - \text{Ms12}(n-1) * P_s$$

Ms23(n) (The amount of R-Cs in the 2-3cm soil layer at n-th month):

$$Ms23(n)=Ms23(n-1)+M12(n-1)*Ps-Ms23(n-1)*Ps$$

Ms35(n) (The amount of R-Cs in the 3-5cm soil layer at n-th month):

$$Ms35(n)=Ms35(n-1)+M23(n-1)*Ps-Ms35(n-1)*Ps/2$$

Ms57(n) (The amount of R-Cs in the 5-7cm soil layer at n-th month):

$$Ms57(n)=Ms57(n-1)+M35(n-1)*Ps/2-Ms57(n-1)*Ps/2$$

With initial condition:

$$Mog(0)=M0$$

$$Ms01(0)= Ms12(0)= Ms23(0)= Ms35(0)= Ms57(0) =0$$

In this study, the radioactive decay constant λ introduced in section 5.1.3.3 ([Figure 5.3](#)) is neglected and the reason can be explained as follows:

In case of radioactivity, radioactive decay in one month is expressed by the constant reduction rate, that is $\exp^{(-\text{"decay constant"}*\text{"one month"})}$. It is expressed by Q. The constant Q is multiplied to quantity of every compartment at every month.

Mog(n) (The amount of R-Cs in the organic layer at n-th month):

$$Mog(n)=(Mog(n-1)-Mog(n-1)*Po)*Q=Mog(n-1)* (1-Po)*Q=Mog(0)*(1-Po)^n*Q$$

Ms01(n) (The amount of R-Cs in the 0-1cm soil layer at n-th month):

$$Ms01(n)=(Ms01(n-1)+Mog(n-1)*Po-Ms01(n-1)*Ps)*Q$$

Ms12(n) (The amount of R-Cs in the 1-2cm soil layer at n-th month):

$$Ms12(n)=(Ms12(n-1)+M01(n-1)*Ps-Ms12(n-1)*Ps)*Q$$

Ms23(n) (The amount of R-Cs in the 2-3cm soil layer at n-th month):

$$Ms23(n)=(Ms23(n-1)+M12(n-1)*Ps-Ms23(n-1)*Ps)*Q$$

Ms35(n) (The amount of R-Cs in the 3-5cm soil layer at n-th month):

$$Ms35(n)=(Ms35(n-1)+M23(n-1)*Ps-Ms35(n-1)*Ps/2)*Q$$

Ms57(n) (The amount of R-Cs in the 5-7cm soil layer at n-th month):

$$Ms57(n)=(Ms57(n-1)+M35(n-1)*Ps/2-Ms57(n-1)*Ps/2)*Q$$

With initial condition:

$$Mog(0)=M0$$

$$Ms01(0)= Ms12(0)= Ms23(0)= Ms35(0)= Ms57(0) =0$$

These equations show that if we calculate the quantity of each compartment at n^{th} month from that of $(n-1)^{th}$ month neglecting the decay reduction, each value is $1/Q$ times bigger than the real value. This means that if we neglect the decay reduction at the calculation of every month, the quantity of each compartment is $(1/Q)^n$ times bigger than the real value. This means that if we use Q^n times smaller value as initial condition and neglecting decay reduction, the result of the calculation is same as that with real value of initial condition and considering decay reduction. In the program, the initial quantity of $Mog(0)=M0$ is calculated as just the sum of the quantity of all compartment at n -th month, which means that $M0$ is Q^n times smaller than the real value. So, the result of the calculation is same if we consider decay reduction at each month and in the calculation of initial quantity. This means that the calculation for n -th month can neglect Q , if we calculate the initial quantity using the total mass at n -th month neglecting decay reduction. However the attention should be paid on that this is the phenomenon of radioactive decay, in which the same decay reduction occurs at every compartment and at every mass flow. In the case of chemical reduction, in which the chemical reduction occurs at only some part of compartments, the simple calculation above cannot be applied. If there is a need to calculate at another month, the decay reduction should be considered.

So, in the real calculation, the decay reduction can be neglected, if we calculate the value of n^{th} month using the initial value estimated by the existing quantity of n^{th} month neglecting decay reduction,.

5.3.2.3 Calculation of initial radioactivity of ^{134}Cs (Co)

Since the calculation results of radioactivity in different layers after n^{th} month are the same when considering or without considering the decay constant, the following calculation of Co will not include decay constant (the calculated Co is Q^n times smaller than the real initial radioactivity that considers decay constant). Here, Co is the sum of the radioactivity of organic and all soil layers.

The distribution of sampling data (Log-transformed, unit: Bq) is shown in [Figure 5.17](#). The data didn't show so much difference in sampling data (Log) and no obvious direction trend of distribution.

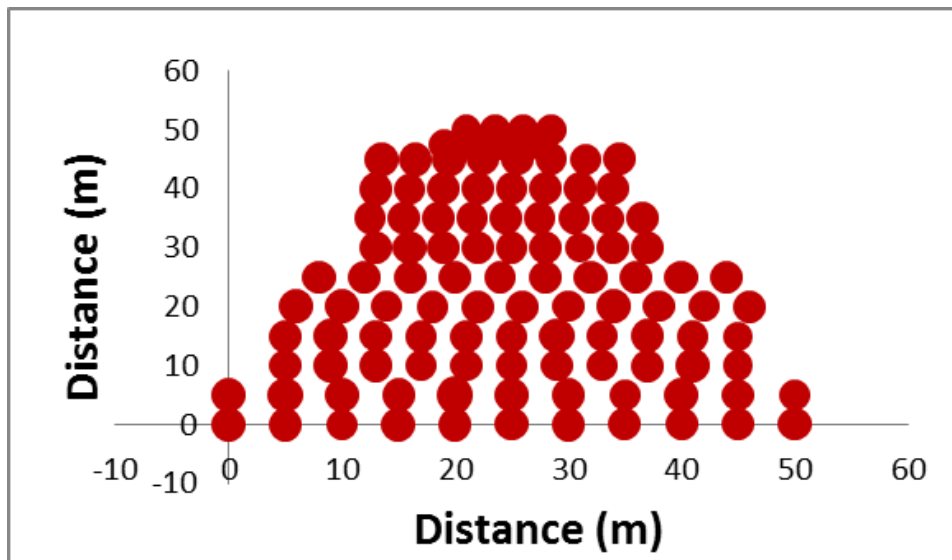


Figure 5.17 Location and size of the sampling data (Log, unit: Bq)

While the histogram shows that the sampling data follows approximately the normal distribution as shown in [Figure 5.18](#). In this study, the spherical model variogram is more suitable to be used compared with other models like exponential model, since the range for the spatial correlation is short (the experimental variogram value reach the sill value in short distance as shown in [Figure 5.18](#)). The parameters for the spherical model are as follows: sill=0.024m, correlation scale = 6.5m, interval distance=4m, nugget is assumed as 0. From the [Figure 5.19](#), we can obtain that the semi-variogram values are fluctuated

around 1, so the parameters like the sill, correlation scale are proper in this spherical model. **Figure 5.20** shows the kriging estimation of radioactivity distribution and the **Figure 5.21** shows the standard deviation (SD) of the kriging estimation.

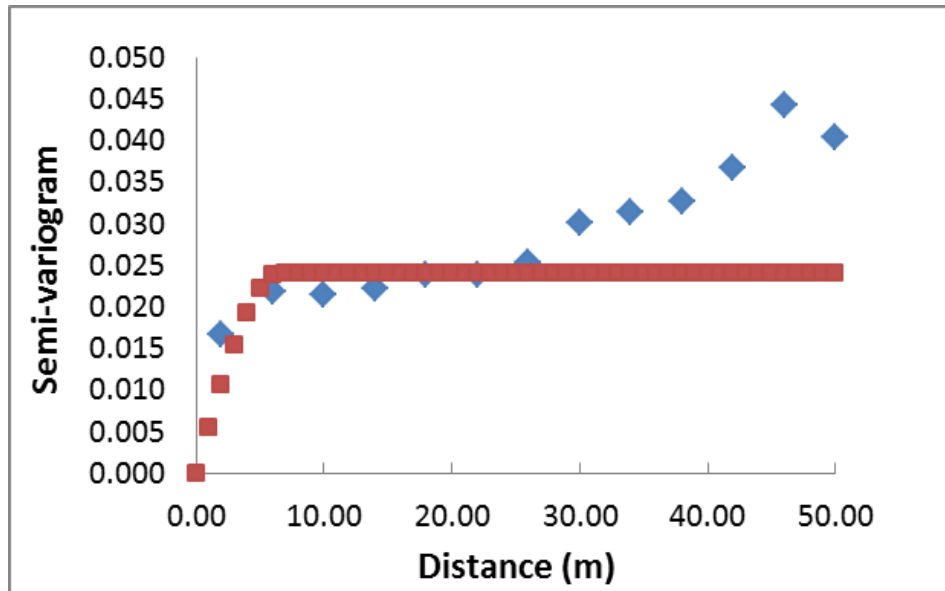


Figure 5.18 Semi-variogram for sampling points. Blue-points is the experimental variogram, red-points is the model variogram.

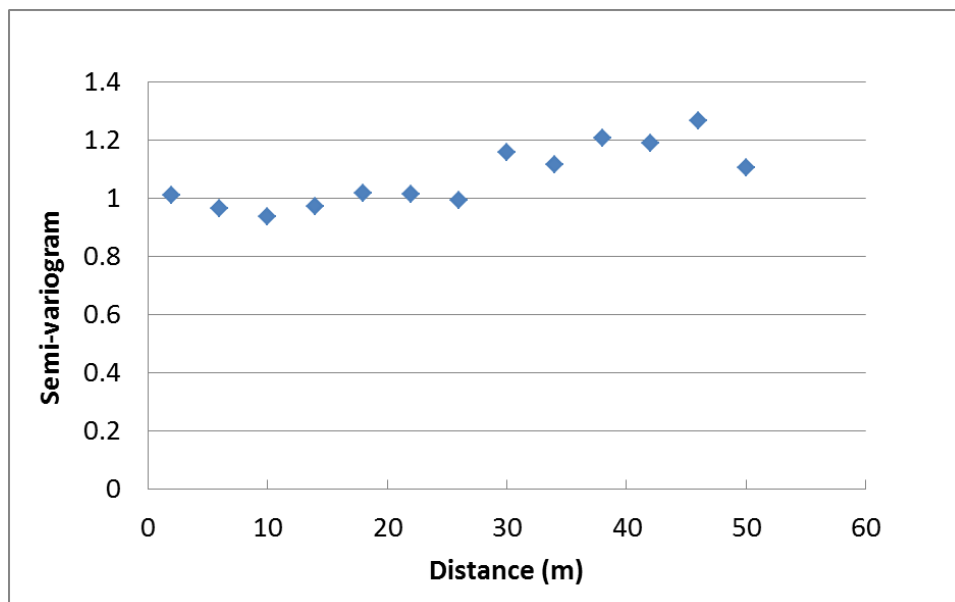


Figure 5.19 Semi-variogram for residues.

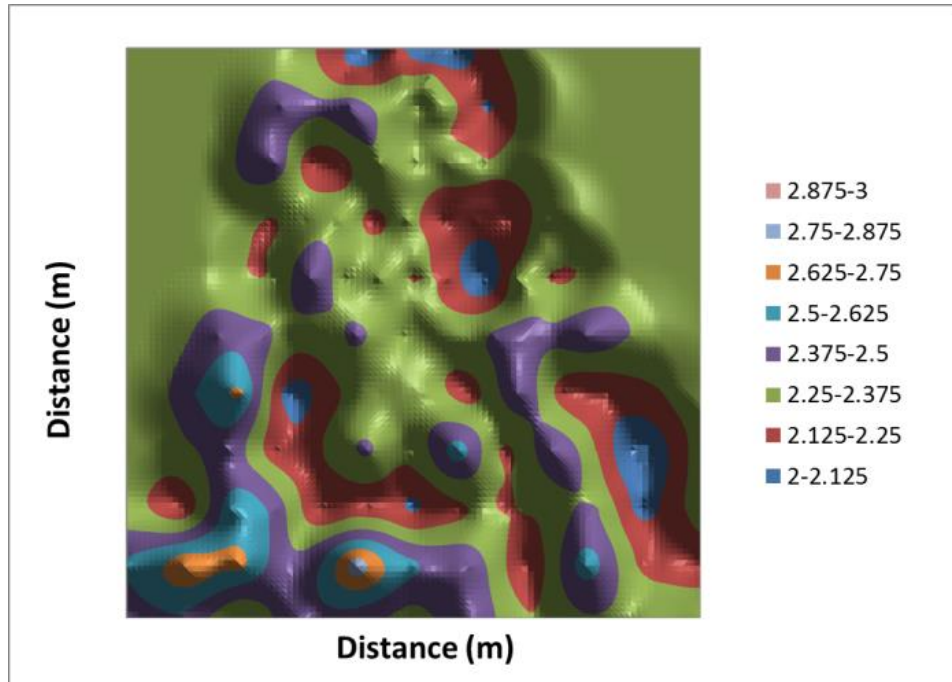


Figure 5.20 Kriging estimation of pollution distribution (Log) (CI=50%) on the polluted forest field (50m*50m).

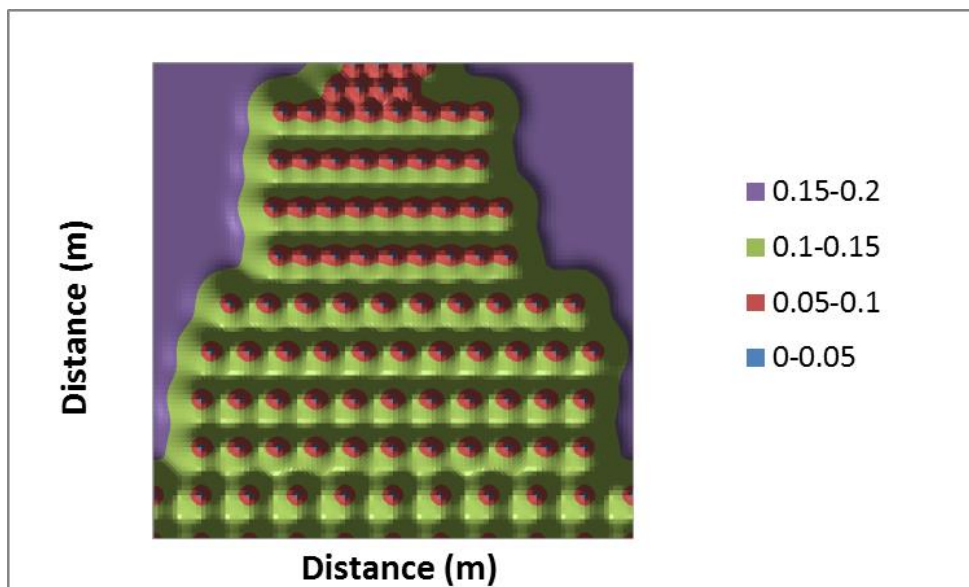


Figure 5.21 The standard deviation (SD) for the Kriging estimation of pollution distribution (Log) (CI=50%) on the polluted forest field (50m*50m).

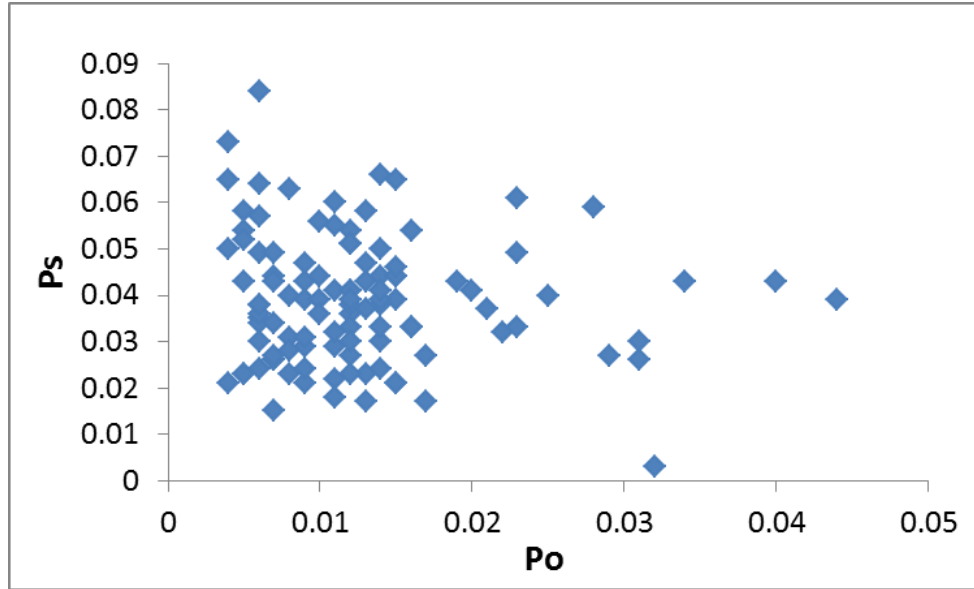


Figure 5.22 The relationship between two transport parameters (Po and Ps) of ^{134}Cs

5.3.2.4 Calculation of transport rate of ^{134}Cs (Po and Ps)

From the original sampling data at different layers, we can get the distribution of radioactivity of ^{134}Cs at each sampling location. Then we use the Po (range from 0.001 to 0.050 in the increment of 0.001) and Ps (range from 0.001 to 0.100 in the increment of 0.001) to fit the distribution of radioactivity at each location. The relatively best Po and Ps for each sampling location will be decided through minimizing the sum of the least-square-error of different layers at this location. After getting the best Po and Ps, we check the relation of the two parameters to judge if they are independent with each other. As shown in **Figure 5.22**, the parameter Po and Ps don't show much relationship with each other; besides the correlation coefficient is -0.1082 as shown in **Table 8**. Therefore they are considered as independent parameters, and we can estimate the Po and Ps by kriging for different layers at each location in the same way with the calculation of Co as shown in **Figure 5.23** and **Figure 5.24**, respectively.

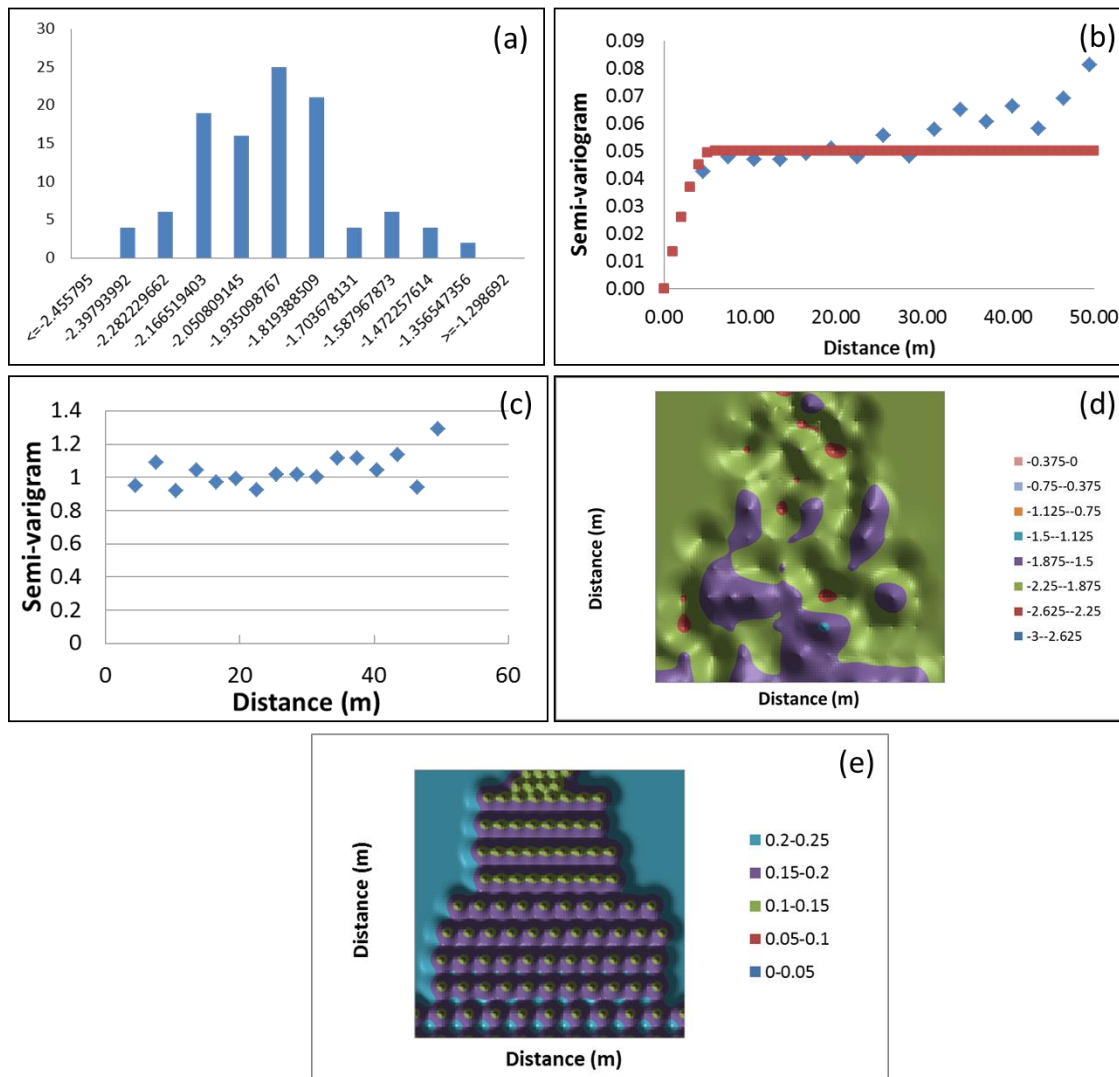


Figure 5.23 (a) Histogram of Po, x-axis is distance interval (m), y-axis is the frequency; (b) Sample semivariogram for concentrations (blue points) and the exponential model (red points) fitted to the sample semivariogram; (c) Sample semivariogram for the residuals (ε_s) in kriging; (d) kriging estimation of pollution distribution (CI=50%); (e) Kriging Standard Deviation (SD).

Table 8 The correlation coefficient between Po and Ps

	Po	Ps
Po	1	
Ps	-0.1082	1

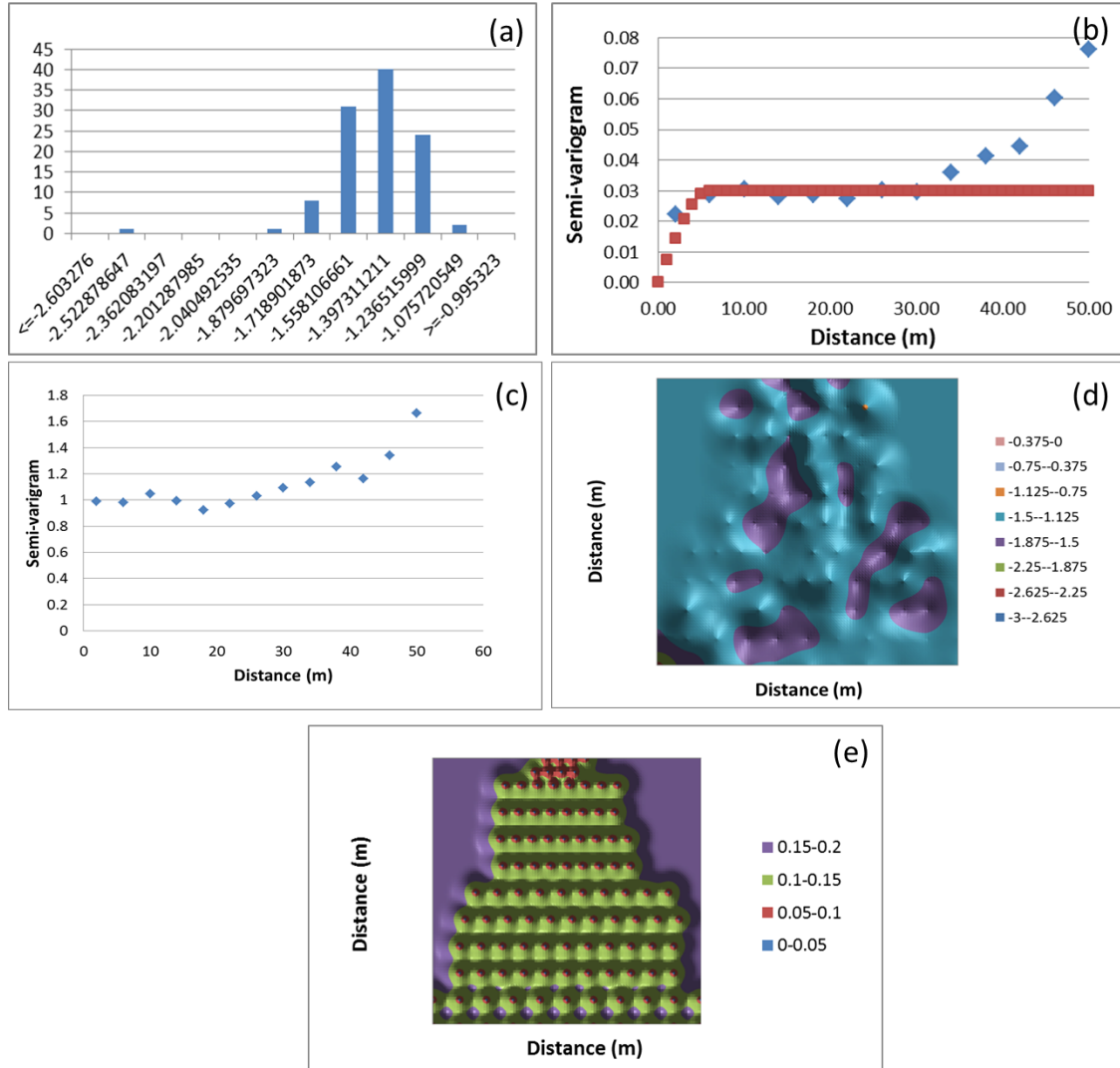


Figure 5.24 (a) Histogram of Ps, x-axis is distance interval (m), y-axis is the frequency; (b) Sample semivariogram for concentrations (blue points) and the exponential model (red points) fitted to the sample semivariogram; (c) Sample semivariogram for the residuals (ε_s) in kriging; (d) kriging estimation of pollution distribution (CI=50%); (e) Kriging Standard Deviation (SD).

5.3.2.5 The 3D model based on the Co, Po, and Ps

After the calculation of the Co, Po and Ps, we can construct the 3D radioactivity-distribution model for ^{134}Cs . Here we take the kriged value (CI=50%) for the Co, Po and Ps as an example to explain the distribution. As shown in **Figure 5.25 (left figures)**, the

kriging (CI=50%), which keeps the least-square-errors at the observed locations, can give the average estimation of ^{134}Cs -radioactivity at each location on different layers. When we tried to set the threshold value as the 25, 50 and 100Bq in this research, the corresponding remediation area (red area in **Figure 5.25 (left figures)** and **Figure 5.26**) can be calculated. The remediation area decrease as the soil depth increase, and the number of layer need to be remediated decrease as the threshold value set up increases. However, except the existed kriging uncertainty, the constructed 3D model should consider the compartment model errors in real radioactivity distribution of ^{134}Cs as shown in **Table 9**. Based on the **Figure 5.26**, compared with the average radioactivity of 107 total sampling in different layers, the compartment model errors is relatively small, and the compartment model errors can be neglected in the calculation of radioactivity.

Table 9 The mean and SD of the compartment model errors on different layers

	OL-Litter	SL-0-1	SL-1-2	SL-2-3	SL-3-5	SL-5-7
Average	2.081	3.413	6.746	5.099	4.093	3.730
SD	1.681	4.335	7.377	7.615	4.656	3.257

5.3.3 3D-model developed and improved by conditional simulation (CS)

Even though the 3D distribution model for ^{134}Cs was constructed, kriging provides one solution to achieve adequate estimates with minimised variance on observed locations. However, in solving for minimised variance, kriging produces estimates associated with the smoothing process.²⁰⁸ The average of many such conditional simulations at a point is equivalent to a kriging estimate.²⁰⁹ The interpolation such as kriging doesn't take into account the uncertainty associated with the estimates. For example, an estimate by kriging honours the known points, but does not reflect the variability between points, because of necessary smoothing, while CS can be used to solve the problem of measuring the uncertainty and variability associated with an estimate.¹⁴²

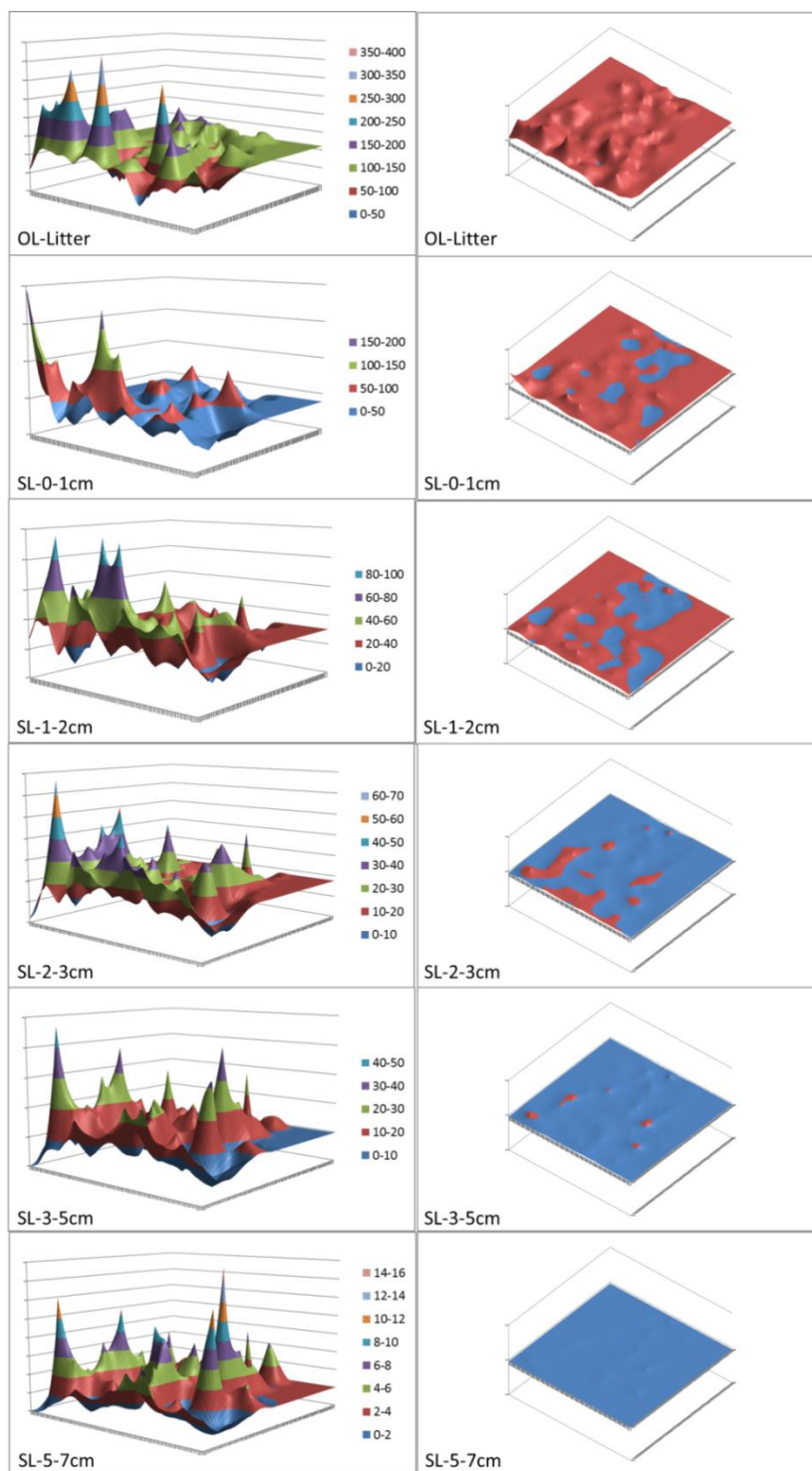


Figure 5.25 3D distribution of ^{134}Cs by kriging estimation (CI=50%)(Left figures) on different layers and the distribution of polluted soil that needs to be remediated (red area except the area in purple colour in the [Figure 5.21](#), right figures) when threshold value=25Bq.

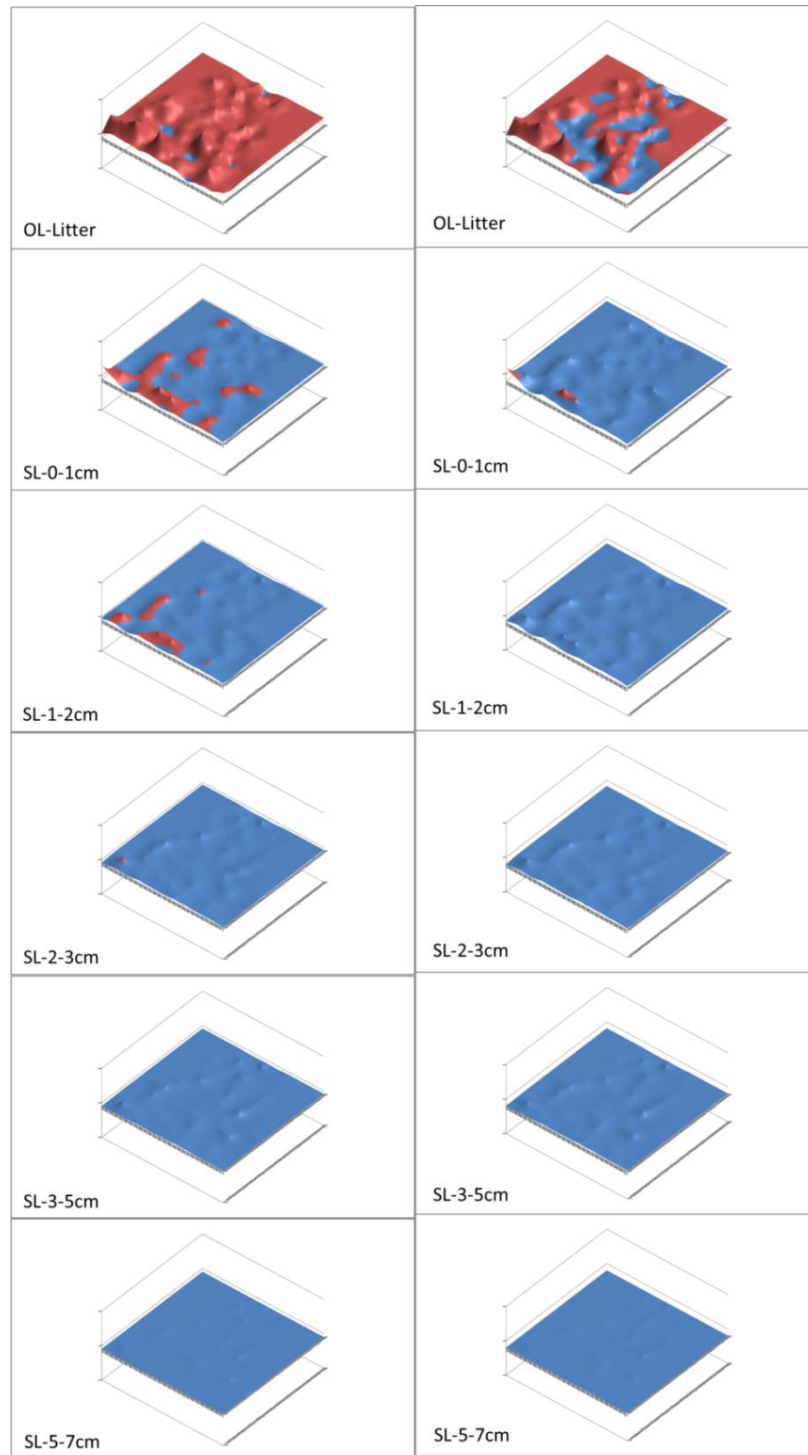


Figure 5.26 The distribution of polluted soil that needs to be remediated (red area except the area in purple colour in the [Figure 5.21](#)) when threshold values are set as 50 (left figures) and 100Bq (right figures) respectively.

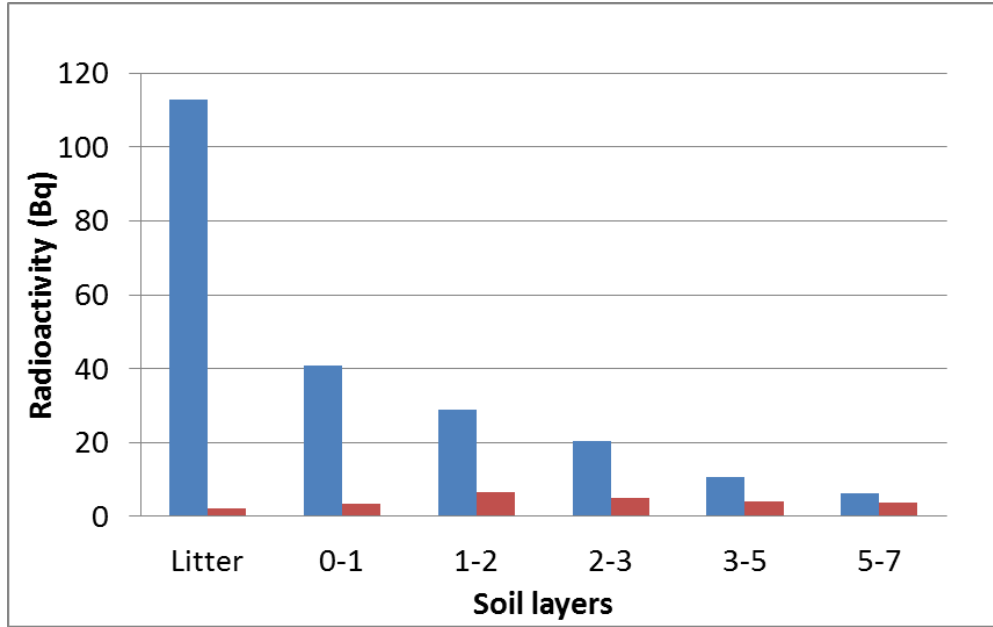


Figure 5.27 The average radioactivity (blue bar chart) and compartment model error (red bar chart) of total 107 sampling data in different layers

5.3.3.1 Comparison between “estimation” and “simulation”

Before the detailed introduction of CS, we should know better about difference of two words: simulation and estimation. Their difference lies in their objectives:²¹⁰

- 1) The objective of estimation is to give, at each point x , an estimator $z^*(x)$ which is as close as possible to the true unknown value $z_o(x)$. The validations of the estimation are unbiasedness and minimal mean quadratic error, or estimation variance $E\{[Z(x) - Z^*(x)]^2\}$. However, these estimators don't reproduce the spatial variability of the true value $\{z_o(x)\}$. In the case of kriging, minimizing the estimation variance leads to a smoothing of the true spatial variance.
- 2) On the other hand, the simulation $\{z_s(x)\}$, or better the CS $\{z_{sc}(x)\}$, has the same first two experimentally found moments (mean and covariance or variogram, as well as the histogram) as the real values $\{z_o(x)\}$. It identifies the main spatial

variation characteristics of these true values. While, at each point x , the simulated $z_s(x)$ or $z_{sc}(x)$ is not the best possible estimator of $z_o(x)$.

In general, the objectives of simulation and estimation are not compatible. As shown in **Figure 5.28**, even though the estimation curve $z^*(x)$ is, on average, closer to the real curve $z_o(x)$, the simulation curve $z_{sc}^*(x)$ is a better reproduction of the fluctuations of the real curve.

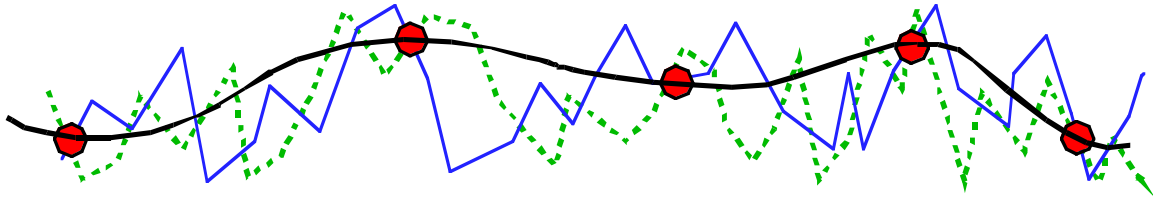


Figure 5.28 Real, simulated and estimated profiles. Blue line is reality (true profile); Black bold line is kriging estimation; Green dotted line is conditional simulation; Red points are conditioning data (observed data).

5.3.3.2 The theory of conditional simulation (CS)

The conditional simulation can be conceptualized by the following decomposition:^{209–211}

$$Z(x) = Z^*(x) + [Z(x) - Z^*(x)]$$

True value = kriging estimator + kriging error

The kriging error is of course unknown since $Z(x)$ is unknown. A characteristic property of kriging is that the kriging error $[Z(x) - Z^*(x)]$ is orthogonal to the kriged value. Now consider the same equality for $Z_s(x)$, which can be generated by the unconditional simulation (US) and has the same covariance (and variogram) as $Z(x)$, and $Z_{sk}^*(x)$ is the kriging estimator as if the simulation were known only at the same observed data.

$$Z_s(x) = Z_{sk}^*(x) + [Z_s(x) - Z_{sk}^*(x)]$$

This time the true value $Z_s(x)$ is known and so is $Z_s(x) - Z_{sk}^*(x)$.

To obtain the desired conditional simulation, it is thus enough to replace the unknown kriging error $[Z(x) - Z^*(x)]$ by an isomorphic and independent kriging error $[Z_s(x) - Z_{sk}^*(x)]$. This gives the $Z_{sc}(x)$ defined by:

$$Z_{sc}(x) = Z^*(x) + [Z_s(x) - Z_{sk}^*(x)]$$

Conditional simulation = kriging estimator + simulation of kriging error

For the conditional simulation, typically 100 simulations may be needed, but this number depends largely on the distribution of the parameters of interest: if it is skewed and if we care even about extreme situations, a large number of simulation is needed, but the extreme value predictions usually lack robustness.²¹¹ In this study, 100 realizations was used also considering the calculation time by computer.

5.3.4 Assessment of the soil pollution based on the developed cost-effective strategy in Chapter3

When the threshold value changes (**Table 3**), the corresponding best average CI value used in the kriging method are different. In this case, we consider the forest soil (including the OL and all SLs) as the whole compartment (that means OL and all SLs are regarded as surface soil). In addition, the radioactivity at each location is the sum of radioactivity of OL and all SLs (only consider Co). If the $Co \geq$ threshold value, the forest soil at this location is considered polluted. As a result, the remediation ratio (that means the ratio of remediation area (kriged value $\geq$ threshold value) and the total field area multiplied by 0.8682) is shown in **Table 10**, where the parameter of 0.8682 means the sampling area coves up 86.82% of the total field, and the rest of the field is not considered in the study since no sampling data existed in the area and has a large estimation error by kriging as shown the area in purple colour in the **Figure 5.21**. From the **Figure 5.29**, we can see that the corresponding best remediation area based on the the **Table 3** and the cost-effective strategy in Chapter 3. Through the soil depth, sample and remediation unit price, the total fee can be calculated.

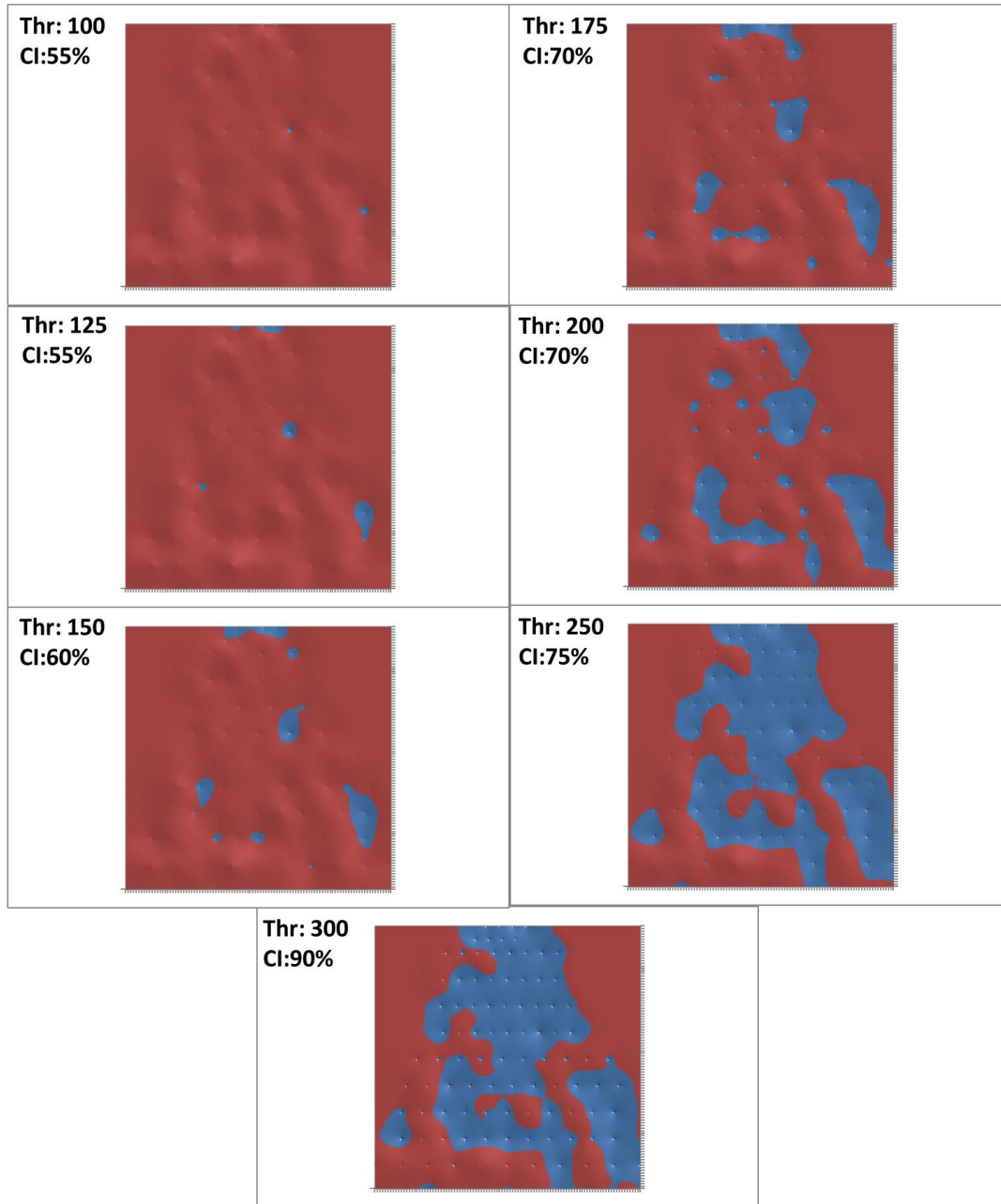


Figure 5.29 The remediation area (in red color except the area in purple colour in the **Figure 5.21**) based on the **Table 3** and the cost-effective strategy in Chapter 3. Thr-the threshold value; SDPR-pollution rate reflected by sampling data.

However, in consideration of the distribution of ^{134}Cs -radioactivity is different at different layers as shown in **Figure 5.25** and **Figure 5.26**, the calculation of remediation area will not be so correct if we assumed the forest soil as one whole compartment. So much unpolluted soils like the SL-3-5cm and SL-5-7cm, which are not polluted so heavily, will

be removed and treated, and then leads to high cost and less cost-effective. That is also the practical reason we want to use condition simulation assessment on 3D distribution of ^{134}Cs -radioactivity.

Table 10 The remediation ration based on the **Table 3** and the cost-effective strategy in Chapter 3.

Threshold Value	100	125	150	175	200	250	300
SDPR (%)	96.26	91.59	82.24	62.62	47.66	26.17	12.15
Best average CI value	55	55	60	70	70	75	90
Remediation ratio (%)	99.83	98.60	95.42	91.34	82.13	58.69	56.33

5.3.5 Assessment of the soil pollution based on conditional simulation (CS)

In this study, due to the limitation of computer calculation memory (or CPU), the simulation size is 51×51 , interval is set up as 1m. The number of the realization is set up as 1,000. The conditional cumulative distribution function (CCDF: %) is set up as from 50% to 95% in 5% increment, and 99%. The result of corresponding remediation ratios at each layer (OL and all SLs) are discussed based on the threshold values 25Bq, 50Bq, and 100Bq, separately (as shown in **Table 11**). Take the CCDF 50% and 95% for examples, the distribution of radioactivity strength (or concentration) are shown in **Figure 5.30** and **Figure 5.31**. In general, after the conditional simulations and the compartment model calculation, the distribution of radioactivity at each layer (OL and all SLs) can be known based on the different CCDF values (e.g. 50%, 95% as shown in **Figure 5.30** and **Figure 5.31**). Since the distribution of radioactivity is known, it is easy to know the remediation ratio, and then the remediation area or volume can be also calculated after setting the threshold value (part of remediation ratios as shown in **Table 11**). The calculation function is as follows:

$$PSV = FA \times (RR_{OL} + RR_{SL01} + RR_{SL12} + RR_{SL23} + RR_{SL35} \times 2 + RR_{SL57} \times 2) \times 10^{-2}$$

Where Polluted (or Remediation) Soil Volume (PSV): m³;

Field Area (FA): $50 \times 50 = 2500\text{m}^2$; Remediation Ratio (RR): %

It is hard to tell which CCDF is best to be selected to decide the remediation ratio, but depends on the budget, remediation time, or other issues considered by decision-makers, government officers, or stake-holders etc. If we assume the organic layer as one compartment of soil layer such as soil layer0-1cm, the [Table 12](#) gives the PSV by three methods or strategies. We can see that the PSV by 3D strategy using CS (CCDF=95%) is quite different from that by only kriging (CI=50%). And the PSV by 3D strategy using CS is regarded as more close to the reality, since PSV by only kriging has the smoothing effect. In addition, the PSV by 2D strategy always overestimate or underestimate that of reality, since it does not consider the distribution at depth.

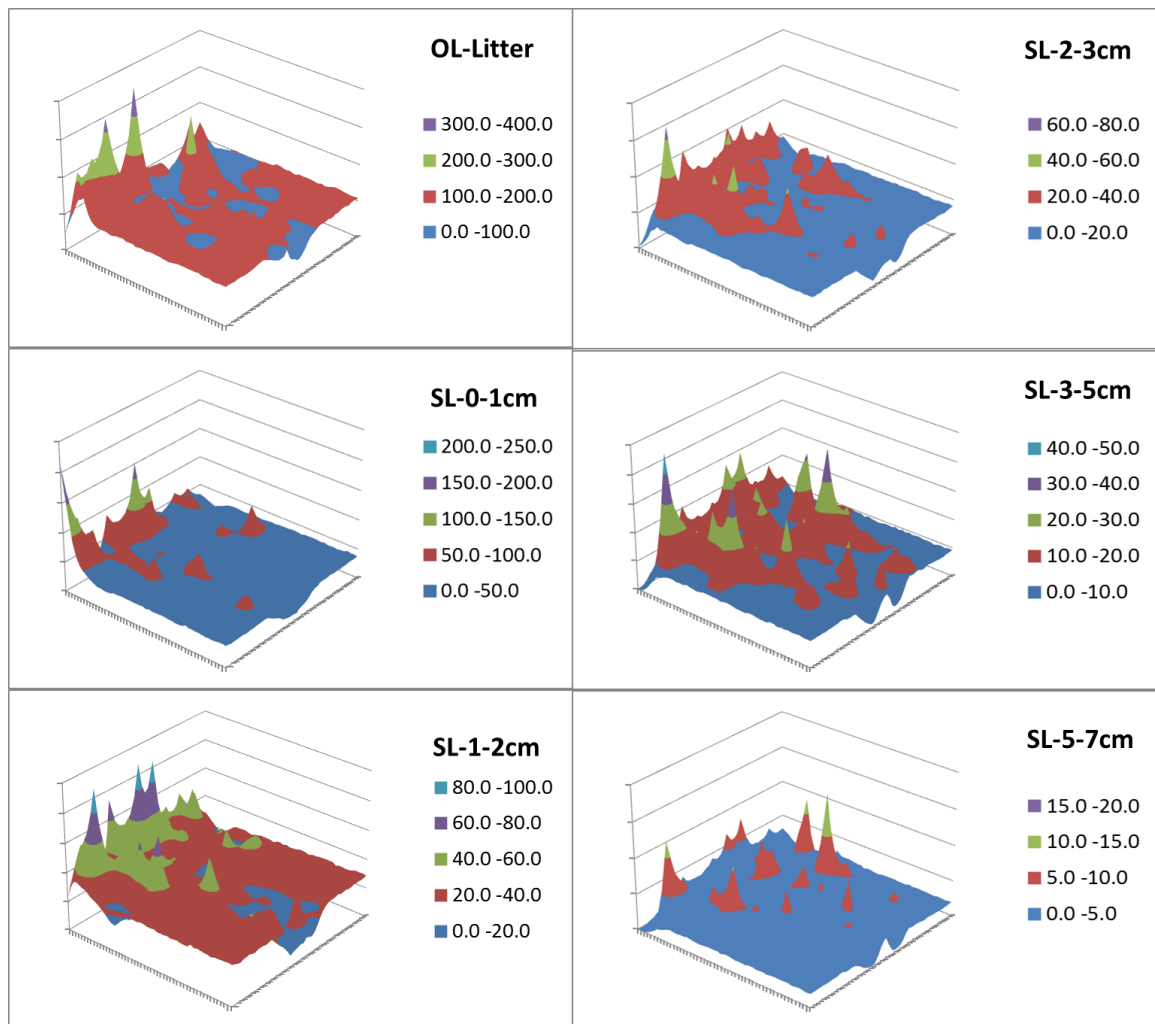


Figure 5.30 The distribution of radioactivity strength at different soil layers based on the calculation of CS (1000 realizations and CCDF=50%) and compartment model.

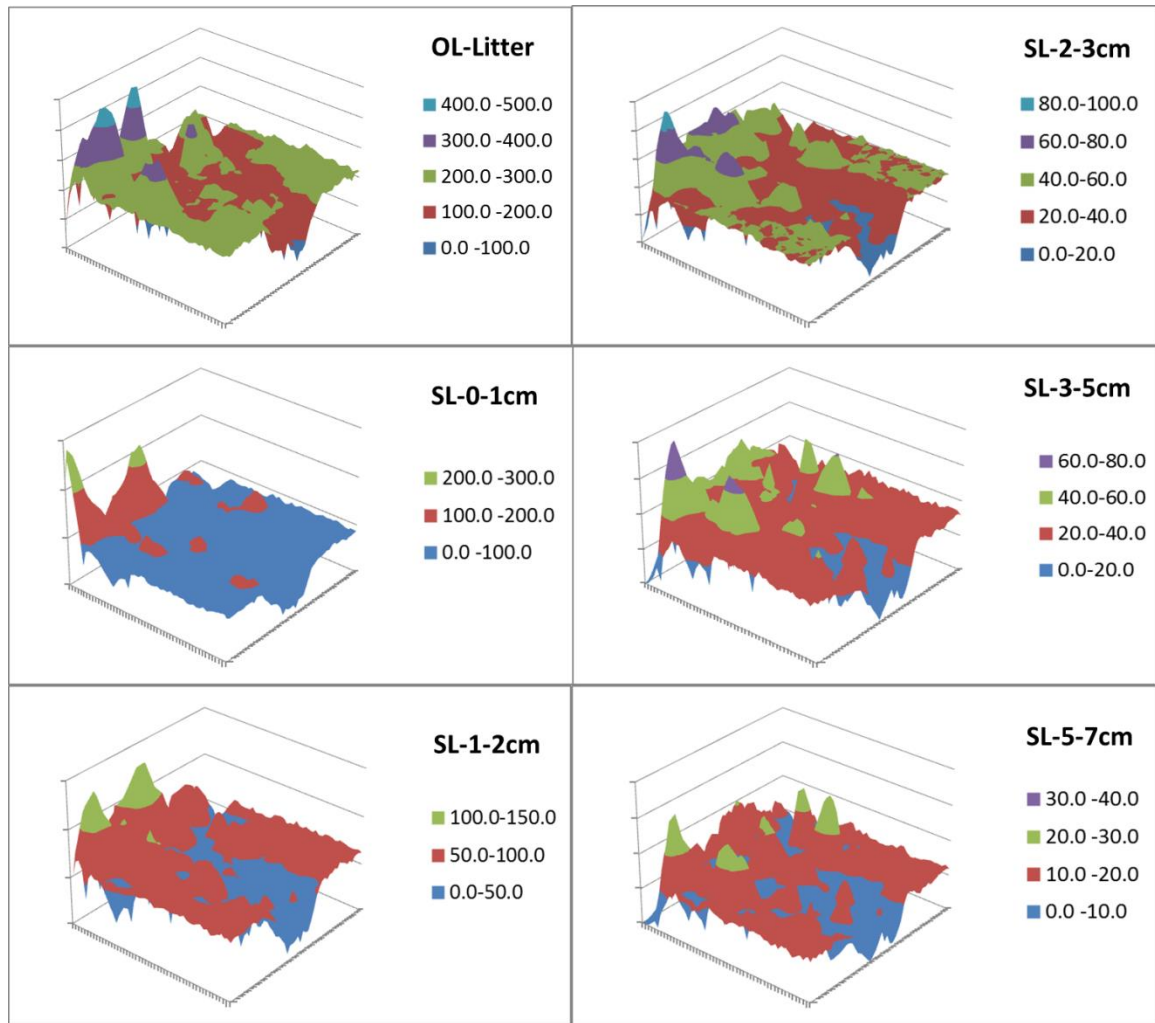


Figure 5.31 The distribution of radioactivity strength at different soil layers based on the calculation of CS (1000 realizations and CCDF=95%) and compartment model.

Table 11 The remediation ratio at different soil layers based on the different threshold values (Thr) and different CCDF values.

Thr (Bq)	CCDF (%)	OL- litter	SL 0-1cm	SL 1-2cm	SL 2-3cm	SL 3-5cm	SL 5-7cm
25	50	0.9992	0.8428	0.6778	0.1234	0.0131	0.0000
	55	0.9992	0.8681	0.7220	0.1449	0.0185	0.0000
	60	0.9992	0.8931	0.7520	0.1684	0.0250	0.0000
	65	0.9992	0.9139	0.7889	0.1942	0.0381	0.0000

	70	0.9992	0.9319	0.8247	0.2334	0.0577	0.0000
	75	0.9992	0.9512	0.8558	0.2930	0.0888	0.0000
	80	0.9992	0.9612	0.8820	0.5521	0.1311	0.0000
	85	0.9992	0.9700	0.9073	0.6444	0.2015	0.0000
	90	0.9992	0.9792	0.9419	0.7170	0.4648	0.0000
	95	0.9992	0.9850	0.9646	0.8051	0.6521	0.0069
	99	0.9992	0.9885	0.9804	0.9004	0.8128	0.1027
50	50	0.9800	0.1496	0.0504	0.0015	0.0000	0.0000
	55	0.9831	0.1757	0.0638	0.0015	0.0000	0.0000
	60	0.9869	0.1995	0.0804	0.0019	0.0000	0.0000
	65	0.9900	0.2334	0.0965	0.0023	0.0000	0.0000
	70	0.9915	0.3014	0.1165	0.0042	0.0000	0.0000
	75	0.9942	0.5275	0.1488	0.0065	0.0000	0.0000
	80	0.9958	0.5978	0.1815	0.0123	0.0008	0.0000
	85	0.9965	0.6686	0.2391	0.0238	0.0031	0.0000
	90	0.9969	0.7401	0.5313	0.0507	0.0104	0.0000
	95	0.9969	0.8220	0.6424	0.1057	0.0338	0.0000
	99	0.9969	0.9204	0.7982	0.4967	0.3160	0.0000
100	50	0.6859	0.0127	0.0000	0.0000	0.0000	0.0000
	55	0.7332	0.0161	0.0000	0.0000	0.0000	0.0000
	60	0.7751	0.0192	0.0000	0.0000	0.0000	0.0000
	65	0.8116	0.0211	0.0000	0.0000	0.0000	0.0000
	70	0.8470	0.0242	0.0000	0.0000	0.0000	0.0000
	75	0.8743	0.0308	0.0012	0.0000	0.0000	0.0000
	80	0.8989	0.0365	0.0031	0.0000	0.0000	0.0000
	85	0.9208	0.0481	0.0069	0.0000	0.0000	0.0000
	90	0.9443	0.0719	0.0123	0.0000	0.0000	0.0000
	95	0.9616	0.1330	0.0273	0.0000	0.0000	0.0000
	99	0.9777	0.5329	0.0965	0.0031	0.0004	0.0000

Table 12 The comparison of polluted soil volume (PSV) by different three methods.

Threshold value	PSV (3D-strategy ^a)	PSV (Kriging ^b)	PSV (2D-strategy ^c)
25	126.80	68.93	190.84
50	65.87	30.55	112.66
100	28.05	18.30	4.08

^a-using the conditional simulation when CCDF=95%; ^b-kriging average estimation when CI=50%; ^c-kriging estimation based on the suggested best average CI value in [Table 3](#).

5.4 In summary and future study

In this chapter, one real case (polluted forest soil by Cs in Fukushima prefecture, Japan) was studied to build the 3D radioactivity-distribution model. From chapter 4, we considered to build the 3D model using both kriging and transport functions. However, the sampling data is not enough to calculate the parameters, i.e. D —the diffusion coefficients (cm^2/year), V —the convection velocity (cm/year), of the convection-dispersion model. Due to the complex situation for forest soil layers, the compartment model is better to be selected to construct the 3D model. Through the calculation of the parameters C_o , P_o and P_s for compartment model, the quasi-3D model has been constructed. From the 3D model, the distribution of ^{134}Cs can be clearly visible for the decision makers or engineers to operate the remediation on the field; however, the area or volume needed to be remediated depends on the CI value used in the kriging method and the threshold value set up in the real project by law or government, et.al. Because of the limitation in the kriging, the conditional simulation (CS) was used to develop the 3D model. One hundred of realizations of CS were generated to assess the probability of simulated values that exceed the threshold value at each location. From the cumulative percent, 95% was selected to assess the area or volume of polluted soil that should be remediated. Finally, the 3D model combining the kriging method, compartment model and conditional simulation (CS) was constructed and was efficient for the assessment of the forest soil pollution.

In future study, the additional sampling method based on the 3D model is focused on to make simulation or estimation more accurate to reflect the reality and finally make the corresponding strategy more cost-effective.

CHAPTER VI GENERAL CONCLUSION

Heavy metals (HMs) or trace metals normally exist in nature, and in small quantities are nutritionally essential to life. However, HMs can also cause serious environmental pollution and greatly threaten the health of animal and human when the concentration of HMs in soil, water or atmosphere exceeds a certain threshold value. In China, because of the recent thirty years' rapid development of industrialization and urbanization partially at price of environment, soil has been seriously polluted by HMs. Based on the latest survey report released on 17th April, 2014, 16.1% of the Chinese soil was polluted and 82.8% of the polluted land was contaminated by toxic inorganic pollutants and the HMs. Compared with China, the soil pollution has greatly controlled and decreased in Japan; however the accident of Fushishima Daiichi Nuclear Plant (FDNPP) has caused serious soil pollution of radioactive HMs like cesium. Remediation or treatment of so large an area of polluted soil makes the people, company, even government impossible to burden such large of the money to carry out using the traditional investigation and remediation methods, in addition the cost of human worker also contribute the high cost of the total project. As a result, developing a cost-effective strategy to investigate and remediate the polluted soil is so important.

In this research, there are two main conclusions are as follows:

6.1 Development of the cost-effective strategy on the surface soil (2D).

In this study, we use large numbers of assumed actual polluted fields (AAPFs) generated by unconditional simulation (US) to assess a strategy combining traditional systematic grid sampling (SGS) and kriging methods,

and then use a genetic algorithm (GA) to find optimal locations for the additional sampling. Most of the earlier studies have focused on single actual polluted fields and used conditional simulation (CS) to delineate the distribution of pollution. We propose that AAPFs can be used to develop and assess strategies by assuming many polluted fields, in light of the practical difficulties of finding and operating on large numbers of different and actual polluted fields in the real world. To address the issues that most directly concern stakeholders and decision-makers (the total fee (TF) and correct treatment rate (CTR)), we developed a new performance criterion (the adjusted total fee (ATF), which balances the TF and CTR for the assessment and optimization procedure. Our results indicate that the parameter confidence interval (CI) of kriging has a relationship with the pollution rate (PR) of the field, and that a combination of SGS, kriging, and GA can greatly reduce the ATF compared with that derived solely from SGS. Optimal locations for additional sampling were also suggested that the additional samples are always selected at the joint locations of polluted and unpolluted areas, where there are abundant sampling data near the threshold value, whereas the additional samples are generally dispersed over the full scale of the AAPF rather than concentrated in any one area.

6.2 Development of the 3D radioactivity-distribution model for the polluted forest soil: applied to a case study.

To develop the 3D radioactivity-distribution model for the polluted forest soil, one real case (polluted forest soil by Cs in Kawauchi village, Fukushima prefecture, Japan) was studied to build the 3D radioactivity-distribution model. We have considered to build the 3D model using both kriging and transport functions. However, the sampling data is not enough to calculate the

parameters, i.e. D —the diffusion coefficients (cm^2/year), V —the convection velocity (cm/year), of the convection-dispersion model. Due to the complex situation for forest soil layers, the compartment model is better to be selected to construct the 3D model. Through the calculation of the parameters C_o , P_o and P_s for compartment model, the quasi-3D model has been constructed. From the 3D model, the distribution of ^{134}Cs can be clearly visible for the decision makers or engineers to operate the remediation on the field; however, the area or volume needed to be remediated depends on the CI value used in the kriging method and the threshold value set up in the real project by law or government, et.al. Because of the limitation in the kriging, the conditional simulation (CS) was used to develop the 3D model. 1000 realizations of CS were generated to assess the probability of simulated values that exceed the threshold value at each location. From the cumulative percent, 50% and 95% was selected as examples to give the distribution map of radioactivity at each layer and then to assess the area or volume of polluted soil that should be remediated. The remediation ratios are also suggested based on different threshold value and CCDF values. Finally, the 3D model combining the kriging method, compartment model and conditional simulation (CS) was constructed and was efficient for the assessment of the forest soil pollution.

We also should pay more attention on the limitations on that in the 2D strategy, the efficiency of 2D strategy compared with traditional methods also depends on the field size, the pollution ration (PR) of the field, and unit price of sampling and remediation of the local market. In addition, we also should notice that the 3D strategy developed in Chapter 5 is one case study, and this 3D strategy maybe can't be used in the whole area (or each area) in the polluted soil in Fukushima prefecture. Before using the 3D strategy, we should judge if the polluted field has the similar phenomena with the case

using in this study, if yes, the 3D strategy can be used well. Of course, the 3D strategy developed in this study mainly focus on the methods, and these methods will be a good reference for the investigation and remediation for the polluted soil in the real sites.

At last, the main conclusion was given that the cost-effective strategy for the investigation and remediation on surface soil pollution and the 3D radioactivity-distribution model for the polluted forest soil have been developed and suggested in this study. The results from the research can give a good reference for researchers, engineers, or government officers who are in charge of the project of the investigation and remediation of the polluted fields.

REFERENCE

1. Singh A, Sharma RK, Agrawal M, Marshall FM. Health risk assessment of heavy metals via dietary intake of foodstuffs from the wastewater irrigated site of a dry tropical area of India. *Food Chem Toxicol*. 2010;48(2):611–619. doi:10.1016/j.fct.2009.11.041.
2. Li Z, Ma Z, Kuijp TJ Vander. A review of soil heavy metal pollution from mines in China: Pollution and health risk assessment. *Sci Total Environ*. 2014;468-469:843–853.
3. Tan M-Z, Xu F-M, Chen J, Zhang X-L, Chen J-Z. Spatial Prediction of Heavy Metal Pollution for Soils in Peri-Urban Beijing, China Based on Fuzzy Set Theory. *Pedosphere*. 2006;16(5):545–554. doi:10.1016/S1002-0160(06)60087-8.
4. Li W, XU B, Song Q. The identification of “hotspots” of heavy metal pollution in soil–rice systems at a regional scale in eastern China. *Sci Total Environ*. 2014;472:07–420.
5. Cheng S. Heavy Metal Pollution in China: Origin, Pattern and Control. *Env Sci Pollut Res*. 2003;10(3):192–198.
6. Suruchi KP. Assessment of Heavy Metal contamination in different vegetables grown in and around Urban areas. 2011:162–179.
7. B.L.Kawatra BP. Consumption of Heavy Metal and Minerals by Adult Women through Food in Sewage and Tube -Well Irrigated Area around Ludhiana City (Punjab , India). *JHumEcol*. 2008;23(4):351–354.
8. Hernandez L, Probst A, Probst JL, Ulrich E. Heavy metal distribution in some French forest soils: evidence for atmospheric contamination. *Sci Total Environ*. 2003;312(1-3):195–219. doi:10.1016/S0048-9697(03)00223-7.
9. Akpor OB, Muchie M. Remediation of heavy metals in drinking water and wastewater treatment systems : Processes and applications. 2010;5(12):1807–1817.
10. Bhargava A, Carmona FF, Bhargava M, Srivastava S. Approaches for enhanced phytoextraction of heavy metals. *J Environ Manage*. 2012;105:103–120. doi:10.1016/j.jenvman.2012.04.002.
11. Luo X, Yu S, Zhu Y, Li X. Trace metal contamination in urban soils of China. *Sci Total Environ*. 2012;421-422:17–30. doi:10.1016/j.scitotenv.2011.04.020.
12. Satyawali Y, Seuntjens P, Van Roy S, et al. The addition of organic carbon and nitrate affects reactive transport of heavy metals in sandy aquifers. *J Contam Hydrol*. 2011;123(3-4):83–93. doi:10.1016/j.jconhyd.2010.12.009.
13. Herat S, Agamuthu P. E-waste: a problem or an opportunity? Review of issues, challenges and solutions in Asian countries. *Waste Manag Res*. 2012;30(11):1113–1129. doi:10.1177/0734242X12453378.
14. Kabir E, Ray S, Kim K-H, et al. Current status of trace metal pollution in soils affected by industrial activities. *Sci World J*. 2012;2012:1–18. doi:10.1100/2012/916705.
15. E.M. Joan, M.R B. Environmental Protection Agency. EPA/540/S92/018 October. 1992.
16. Chen HM, Zheng CR, Tu C, Shen ZG. Chemical methods and phytoremediation of soil contaminated with heavy metals. *Chemosphere*. 2000;41(1-2):229–34. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22722978>.
17. China Industry Security Guide. 入世10周年有色金属工业发展状况（2000-2010年）.

2011. Available at: <http://acs.mofcom.gov.cn/sites/aqzn/jrkd.jsp?contentId=2646067625534>.
18. Forum on HM Soil Remediation and Ecological Restoration. Beijing. 2012.
19. Cheng S. Heavy metal pollution in China: origin, pattern and control. *Environ Sci Pollut Res Int*. 2003;10(3):192–198.
20. Niu L, Yang F, Xu C, Yang H, Liu W. Status of metal accumulation in farmland soils across China: From distribution to risk assessment. *Environ Pollut*. 2013;176:55–62.
21. 20 城市 600 万吨 铬 渣 污 染 数 十 年 将 影 响 几 代 人 . 2010. Available at: <http://www.yw.zj.cn/xw3898.html>.
22. 任世磊. 武汉开发陷两难 耗资数亿清理千亩毒地污染. 每日经济新闻. 2012. Available at: <http://news.dichan.sina.com.cn/2012/09/14/564863.html>.
23. Chen W, Lu S, Peng C, Jiao W, Wang M. Accumulation of Cd in agricultural soil under long-term reclaimed water irrigation. *Environ Pollut*. 2013;178:294–9. doi:10.1016/j.envpol.2013.03.036.
24. Xiaoni H, Hong L, Danfeng S, Liandi Z, Baoguo L. Multi-scale spatial structure of heavy metals in agricultural soils in Beijing. *Environ Monit Assess*. 2010;164(1-4):605–16. doi:10.1007/s10661-009-0916-7.
25. Liu G, Liu X, Hou J. Heavy metal speciation and pollution of agricultural soils along Jishui River in non-ferrous metal mine area in Jiangxi Province, China. *J Geochemical Explor*. 2013;132:156–163.
26. Shan Y, Tysklind M, Hao F. Identification of sources of heavy metals in agricultural soils using multivariate analysis and GIS. *J Soils Sediments*. 2013;13:720–729.
27. Cai L, Xu Z, Ren M. Source identification of eight hazardous heavy metals in agricultural soils of Huizhou, Guangdong Province, China. *Ecotoxicol Environ Saf*. 2012;78:2–8.
28. Lu A, Wang J, Qin X. Multivariate and geostatistical analyses of the spatial distribution and origin of heavy metals in the agricultural soils in Shunyi, Beijing, China. *Sci Total Environ J*. 2012;425:66–74.
29. Huang L, Deng C, Huang N. Multivariate statistical approach to identify heavy metal sources in agricultural soil around an abandoned Pb–Zn mine in Guangxi Zhuang Autonomous Region, China. *Env Earth Sci*. 2013;68:1331–1348.
30. Li X, Feng L. Spatial distribution of hazardous elements in urban topsoils surrounding Xi'an industrial areas, (NW, China): Controlling factors and contamination assessments. *J Hazard Mater*. 2010;174:662–669.
31. Xia X, Chen X, Liu R, Liu H. Heavy metals in urban soils with various types of land use in Beijing, China. *J Hazard Mater*. 2011;186(2-3):2043–50. doi:10.1016/j.jhazmat.2010.12.104.
32. Luo X, Yu S, Zhu Y, Li X. Trace metal contamination in urban soils of China. *Sci Total Environ*. 2012;421-422:17–30. doi:10.1016/j.scitotenv.2011.04.020.
33. Wei B, Yang L. A review of heavy metal contaminations in urban soils, urban road dusts and agricultural soils from China. *Microchem J*. 2010;94:99–107.
34. Cheng H, Li M, Zhao C. Overview of tracemetals in the urban soil of 31 metropolises in China. *J Geochemical Explor*. 2013. Available at: <http://dx.doi.org/10.1016/j.gexplo.2013.08.012>.

35. Zheng Y, Chen T, He J. Relationship Between Heavy Metals and Soil Minerals (Subject Editor : Stefan Norra) Multivariate Geostatistical Analysis of Heavy Metals in Topsoils from. 2008;8(1):51–58.
36. Chen R, Sherbinin ADE, Ye C, Shi G. China ' s Soil Pollution : Farms on the Frontline. *Science(Letters)*. 2014;344(6185):691.
37. Yang, H., Huang, X., Thompson, J.R., Flower RJ. China ' s Soil Pollution: Urban Brownfields. *Science(Letters)*. 2014;344(6185):691–692.
38. Fact Sheet---China Site Soil Remediation Opportunities. In: *The 13th China International Environmental Protection Exhibition & Conference*. The 13th China International Environmental Protection Exhibition & Conference; 2011. Available at: <http://www.brisea.com/CIEPEC/Fact Sheets/2. Fact Sheet-soil.pdf>.
39. Akino TM, Uo YL, Longhua WU, et al. Heavy Metal Pollution of Soil and Risk Alleviation Methods Based on Soil Chemistry. *Pedologist*. 2010;(April):38–49.
40. Arao T, Ishikawa S, Murakami M, Abe K, Maejima Y, Makino T. Heavy metal contamination of agricultural soil and countermeasures in Japan. *Paddy Water Environ*. 2010;8(3):247–257. doi:10.1007/s10333-010-0205-7.
41. Maki Y, Shibata J, Murayama N. Remediation of Heavy-Metal-Contaminated Soil with Chelating Agent and Treatment of Waste Solution. *J Chem Eng Japan*. 2011;44(10):701–707.
42. 環境省 編. 環境白書/循環型社会白書/生物多様性白書. 2015;248. Available at: <http://www.env.go.jp/policy/hakusyo/h27/pdf/full.pdf>.
43. Ministry of the Environment. Survey of contaminated soil and countermeasure cases in Japan. 2010. Available at: <http://www.env.go.jp/water/report/h23-02/05.pdf>.
44. Krisanangkura P, Itthipoonthanakorn T, Udomsomporn S. Environmental dose assessment using Ecolego: case study of soil from Japan. *J Radioanal Nucl Chem*. 2013;297(3):443–450. doi:DOI 10.1007/s10967-012-2412-8.
45. Aliyu AS, Evangeliou N, Mousseau TA, Wu J, Ramli AT. An overview of current knowledge concerning the health and environmental consequences of the Fukushima Daiichi Nuclear Power Plant (FDNPP) accident. *Environ Int*. 2015;85:213–228. doi:10.1016/j.envint.2015.09.020.
46. Yotaro Hatamura, Seiji Abe, Masao Fuchigami, Naoto Kasahara, Kenji Iino. *The 2011 Fukushima Nuclear Power Plant Accident: How and Why It Happened*. Woodhead Publishing; 2014.
47. Lepage H, Evrard O, Onda Y, Lefèvre I, Laceby JP, Ayrault S. Depth distribution of cesium-137 in paddy fields across the Fukushima pollution plume in 2013. *J Environ Radioact*. 2015;147:157–164. doi:10.1016/j.jenvrad.2015.05.003.
48. Ministry of the Environment. law of Soil Contamination Countermeasures Act. 2003. Available at: <http://www.env.go.jp/council/toshin/t10-h1407/05-2.pdf>.
49. 米田 稔, 森澤 真輔, 留美 西村. 土壤汚染概況の5地点混合方式における試料採取地点最適配置. 環境衛生工学研究. 1999;13(2):5–18.
50. Montinaro S, Concas A, Pisu M, Cao G. Remediation of heavy metals contaminated soils by ball milling. *Chem Eng Trans*. 2012;28(4):187–192. doi:10.1016/j.chemosphere.2006.11.009.

51. Scanferla P, Ferrari G, Pellay R, Volpi Ghirardini A, Zanetto G, Libralato G. An innovative stabilization/solidification treatment For contaminated soil remediation: demonstration project results. *J Soils Sediments*. 2009;9(3):229–236. doi:10.1007/s11368-009-0067-z.
52. Khan a G, Kuek C, Chaudhry TM, Khoo CS, Hayes WJ. Role of plants, mycorrhizae and phytochelators in heavy metal contaminated land remediation. *Chemosphere*. 2000;41(1-2):197–207. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10819202>.
53. MARQUES APGC, RANGEL ANOS. Remediation of HM Contaminated Soils: Phytoremediation as a Potentially Promising Clean-Up Technology. *Crit Rev Environ Sci Technol*. 2009;39:622–654.
54. Lambert M, Leven BA. New Methods of Cleaning Up Heavy Metal in Soils and Water Innovative solutions to an environmental problem. *Environ Sci Technol briefs citizens*.:1–3.
55. Mulligan CN, Yong RN, Gibbs BF. Remediation technologies for metal-contaminated soils and groundwater: an evaluation. *Eng Geol*. 2001;60(1-4):193–207. doi:10.1016/S0013-7952(00)00101-0.
56. Guo G, Zhou Q, Ma LQ. Availability and assessment of fixing additives for the in situ remediation of heavy metal contaminated soils: a review. *Environ Monit Assess*. 2006;116(1-3):513–528. doi:10.1007/s10661-006-7668-4.
57. Dermont G, Bergeron M, Mercier G, Richer-Laflèche M. Soil washing for metal removal: a review of physical/chemical technologies and field applications. *J Hazard Mater*. 2008;152(1):1–31. doi:10.1016/j.jhazmat.2007.10.043.
58. Padmavathiamma PK, Li LY. Phytoremediation Technology: Hyper-accumulation Metals in Plants. *Water Air Soil Pollut*. 2007;184(1-4):105–126. doi:10.1007/s11270-007-9401-5.
59. Hughes JB, Shanks J, Vanderford M, Lauritzen J, Bhadra R. Transformation of TNT by aquatic plants and plant tissue cultures. *Environ Sci Technol*. 1997;31:266–271.
60. Pulford ID, Watson C. Phytoremediation of heavy metal-contaminated land by trees-a review. *Environ Int*. 2003;29(4):529–40. doi:10.1016/S0160-4120(02)00152-6.
61. Carlos G, Itziar A. Carlos Garbisu, Itziar Alkorta. Phytoextraction: a cost-effective plant-based technology for the removal of metals from the environment. *Bioresour Technol*. 2001;77:229–236.
62. Ghosh M, Singh SP. A Review on Phytoremediation of Heavy Metals and Utilization of It's by Products. *Asian J energy Environ*. 2005;6(04):214–231.
63. Jean-Paul C, Pierre D. *Geostatistics: Modeling Spatial Uncertainty*.; 1999.
64. Dong J, Yu M, Bian Z, Wang Y, Di C. Geostatistical analyses of heavy metal distribution in reclaimed mine land in Xuzhou, China. *Environ Earth Sci*. 2011;62(1):127–137. doi:10.1007/s12665-010-0507-5.
65. 間瀬 茂. 地球統計学とクリギング法: R と geoR によるデータ解析.; 2010.
66. Caro a, Legarda F, Romero L, et al. Map on predicted deposition of Cs-137 in Spanish soils from geostatistical analyses. *J Environ Radioact*. 2013;115:53–59. doi:10.1016/j.jenvrad.2012.06.007.
67. Lin Y-P, Yeh M-S, Deng D-P, Wang Y-C. Geostatistical approaches and optimal additional sampling schemes for spatial patterns and future sampling of bird diversity. *Glob Ecol Biogeogr*. 2008;17(2):175–188. doi:10.1111/j.1466-8238.2007.00352.x.
68. ArcGIS Resource Center. Understanding a semivariogram: The range, sill, and nugget.

- Available at:
http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/Understanding_a_semivariogram_The_range_sill_and_nugget/0031000000mq000000/.
69. Geoff B. Introduction to geostatistics and variogram analysis. 2005. Available at: <http://people.ku.edu/~gbohling/cpe940/Variograms.pdf>.
 70. Graler B. cokriging and indicator kriging. 2011. Available at: http://ifgi.uni-muenster.de/~b_grae02/copulaIntro/02_Co-Kriging_Indicator-Kriging.pdf.
 71. United States Department of Agriculture:Forest Service. Why Use Ordinary Kriging. 2007. Available at: <http://webcam.srs.fs.fed.us/pollutants/ozone/spatial/kriging.shtml>.
 72. Ashraf S, Abbaspour H. Study Of Soil Properties, Using Kriging and In Inverse Distance Weighting Methods. *Adv Environ Biol*. 2011;5(9):2870–2880.
 73. Some'e BS, Hassanpour F, Ezani a., Miremadi SR, Tabari H. Investigation of spatial variability and pattern analysis of soil properties in the northwest of Iran. *Environ Earth Sci*. 2011;64(7):1849–1864. doi:10.1007/s12665-011-0993-0.
 74. Juang K-W, Lee D-Y, Teng Y-L. Adaptive sampling based on the cumulative distribution function of order statistics to delineate heavy-metal contaminated soils using kriging. *Environ Pollut*. 2005;138(2):268–277. doi:10.1016/j.envpol.2005.04.003.
 75. Juang K-W, Liao W-J, Liu T-L, Tsui L, Lee D-Y. Additional sampling based on regulation threshold and kriging variance to reduce the probability of false delineation in a contaminated site. *Sci Total Environ*. 2008;389(1):20–28. doi:10.1016/j.scitotenv.2007.08.025.
 76. Reis a P, Sousa a J, Da Silva EF, Fonseca EC. Application of geostatistical methods to arsenic data from soil samples of the Cova dos Mouros mine (Vila Verde-Portugal). *Environ Geochem Health*. 2005;27(3):259–270. doi:10.1007/s10653-004-5554-y.
 77. Hofer C, Papritz A. Predicting Threshold Exceedance by Local Block Means in Soil Pollution Surveys. *Math Geosci*. 2010;42(6):631–656. doi:10.1007/s11004-010-9287-4.
 78. Chu H-J, Lin Y-P, Jang C-S, Chang T-K. Delineating the hazard zone of multiple soil pollutants by multivariate indicator kriging and conditioned Latin hypercube sampling. *Geoderma*. 2010;158(3-4):242–251. doi:10.1016/j.geoderma.2010.05.003.
 79. Reis a P, da Silva EF, Sousa a J, Patinha C, Fonseca EC. Spatial patterns of dispersion and pollution sources for arsenic at Lousal mine, Portugal. *Int J Environ Health Res*. 2007;17(5):335–349. doi:10.1080/09603120701628412.
 80. Lin C. MAPPING SOIL LEAD AND REMEDIATION NEEDS IN CONTAMINATED SOILS. *Environ Geochem Health*. 2002;24:23–33.
 81. Cameron K, Hunter P. Using spatial models and kriging techniques to optimize long-term ground-water monitoring networks: a case study. *Environmetrics*. 2002;13(5-6):629–656. doi:10.1002/env.582.
 82. Sollitto D, Romic M, Castrignanò A, Romic D, Bakic H. Assessing heavy metal contamination in soils of the Zagreb region (Northwest Croatia) using multivariate geostatistics. *Catena*. 2010;80(3):182–194. doi:10.1016/j.catena.2009.11.005.
 83. Tavares MT, Sousa a. J, Abreu MM. Ordinary kriging and indicator kriging in the cartography of trace elements contamination in São Domingos mining site (Alentejo, Portugal). *J Geochemical Explor*. 2008;98(1-2):43–56. doi:10.1016/j.gexplo.2007.10.002.
 84. Hani A. Spatial Distribution and Risk Assessment of As , Hg , Co and Cr in Kaveh

- Industrial City, using Geostatistic and GIS. *Int J Environ earth Sci*. 2010;38–43.
85. Yang CL, Wu ZF, Zhang HH, Guo RP, Wu YQ. Risk assessment and distribution of soil Pb in Guangdong, China. *Environ Monit Assess*. 2009;159(1-4):381–391. doi:10.1007/s10661-008-0636-4.
 86. Wang L, Wu J. Spatial variability of heavy metals in soils across a valley plain in Southeastern China. *Environ Geol*. 2008;55(6):1207–1217. doi:10.1007/s00254-007-1066-2.
 87. Liu X, Wu J, Xu J. Characterizing the risk assessment of heavy metals and sampling uncertainty analysis in paddy field by geostatistics and GIS. *Environ Pollut*. 2006;141(2):257–264. doi:10.1016/j.envpol.2005.08.048.
 88. Journal A, Sciences A, Publications S. Estimating Soil Contamination with Kriging Interpolation Method. *Am J Appl Sci*. 2006;3(6):1894–1898. doi:10.3844/ajassp.2006.1894.1898.
 89. Guagliardi I, Cicchella D, Rosa R. A Geostatistical Approach to Assess Concentration and Spatial Distribution of Heavy Metals in Urban Soils. *Water, Air, Soil Pollut*. 2012;223(9):5983–5998. doi:10.1007/s11270-012-1333-z.
 90. Hani A, Pazira E. Heavy metals assessment and identification of their sources in agricultural soils of Southern Tehran, Iran. *Environ Monit Assess*. 2011;176(1-4):677–691. doi:10.1007/s10661-010-1612-3.
 91. Bohling G. KRIGING. 2005:-3. Available at: <http://people.ku.edu/~gbohling/cpe940/Kriging.pdf>.
 92. Maringanti C, Chaubey I, Arabi M, Engel B. Application of a multi-objective optimization method to provide least cost alternatives for NPS pollution control. *Environ Manage*. 2011;48(3):448–461. doi:10.1007/s00267-011-9696-2.
 93. Kaini P, Artita K, Nicklow JW. Optimizing Structural Best Management Practices Using SWAT and Genetic Algorithm to Improve Water Quality Goals. *Water Resour Manag*. 2012;26(7):1827–1845. doi:10.1007/s11269-012-9989-0.
 94. Chakraborty R. Fundamentals of Genetic Algorithms. 2010. Available at: http://www.myreaders.info/09_Genetic_Algorithms.pdf.
 95. Diaz-Gomez PA, Hougen. DF. Initial Population for Genetic Algorithms: A Metric Approach. Available at: <http://www.cameron.edu/~pdiaz-go/GAsPopMetric.pdf>.
 96. Wu J, Zheng C, Chien CC. Cost-effective sampling network design for contaminant plume monitoring under general hydrogeological conditions. *J Contam Hydrol*. 2005;77:41–65. doi:10.1016/j.jconhyd.2004.11.006.
 97. Marler RT, Arora JS. Survey of multi-objective optimization methods for engineering. *Struct Multidiscip Optim*. 2004;26(6):369–395. doi:10.1007/s00158-003-0368-6.
 98. Srinivas M, Patnaik LM. Gentic Algorithms: A Survey. 1994. Available at: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=294849&tag=1>.
 99. Reed P, Minsker B, Valocchi AJ. Cost-effective long-term groundwater monitoring design using a genetic algorithm and global mass interpolation. *Water Resour Res*. 2000;36(12):3731–3741.
 100. Reed P, Minsker BS, Goldberg DE. A multiobjective approach to cost effective long-term groundwater monitoring using an elitist nondominated sorted genetic algorithm with historical data. *J hydroinformatics*. 2001:71–89.

101. Reed PM, Minsker BS. Striking the Balance : Long-Term Groundwater Monitoring Design for Conflicting Objectives. *J water Resour Plan Manag.* 2004;130:140–149.
102. Yeh M-S, Lin Y-P, Chang L-C. Designing an optimal multivariate geostatistical groundwater quality monitoring network using factorial kriging and genetic algorithms. *Environ Geol.* 2006;50(1):101–121. doi:10.1007/s00254-006-0190-8.
103. Dhar A, Patil RS. Multiobjective Design of Groundwater Monitoring Network Under Epistemic Uncertainty. *Water Resour Manag.* 2012;26(7):1809–1825. doi:10.1007/s11269-012-9988-1.
104. Wu J, Zheng L, Liu D. Optimizing groundwater development strategies by genetic algorithm: a case study for balancing the needs for agricultural irrigation and environmental protection in northern China. *Hydrogeol J.* 2007;15(7):1265–1278. doi:10.1007/s10040-007-0200-2.
105. Wu J, Zheng C, Chien CC, Zheng L. A comparative study of Monte Carlo simple genetic algorithm and noisy genetic algorithm for cost-effective sampling network design under uncertainty. *Adv Water Resour.* 2006;29(6):899–911. doi:10.1016/j.advwatres.2005.08.005.
106. Gopalakrishnan G, Minsker BS, Valocchi AJ. Monitoring network design for phytoremediation systems using primary and secondary data sources. *Environ Sci Technol.* 2011;45(11):4846–4853. doi:10.1021/es1042657.
107. Lin Y-P, Wang C-L, Yu H-H, et al. Monitoring and estimating the flow conditions and fish presence probability under various flow conditions at reach scale using genetic algorithms and kriging methods. *Ecol Modell.* 2011;222:762–775.
108. Sharma S, Rangaiah GP, Cheah KS. Multi-objective optimization using MS Excel with an application to design of a falling-film evaporator system. *Food Bioprod Process.* 2012;90(2):123–134. doi:10.1016/j.fbp.2011.02.005.
109. Chadalavada S, Datta B. Dynamic Optimal Monitoring Network Design for Transient Transport of Pollutants in Groundwater Aquifers. *Water Resour Manag.* 2008;22(6):651–670. doi:10.1007/s11269-007-9184-x.
110. Preis A, Ostfeld A, Asce M. Multiobjective Contaminant Sensor Network Design for Water Distribution Systems. *J water Resour Plan Manag.* 2008;134:366–377.
111. Kang D, Lansey K. Optimal Meter Placement for Water Distribution System State Estimation. *J water Resour Plan Manag.* 2010:337–348.
112. Weickgenannt M, Kapelan Z, Blokker M, Savic DA. Risk-Based Sensor Placement for Contaminant Detection in Water Distribution Systems. *J water Resour Plan Manag.* 2010:629–636.
113. Karamouz M, Kerachian R, Akhbari M, Hafez B. Design of River Water Quality Monitoring Networks: A Case Study. *Environ Model Assess.* 2009;14(6):705–714. doi:10.1007/s10666-008-9172-4.
114. Rajper S, Amin IJ. Optimization of wind turbine micro-siting: A comparative study. *Renew Sustain Energy Rev.* 2012;16(8):5485–5492. doi:10.1016/j.rser.2012.06.014.
115. Zoraghein H, Alesheikh A a., Alimohammadi A, Vahidnia MH. The utilization of soft transformation and genetic algorithm to model two sources of uncertainty of Indicator Kriging. *Comput Environ Urban Syst.* 2012;36(6):592–598. doi:10.1016/j.compenvurbsys.2012.04.001.
116. Maringanti C, Chaubey I, Popp J. Development of a multiobjective optimization tool for the

- selection and placement of best management practices for nonpoint source pollution control. *Water Resour Res.* 2009;45:1–15. doi:10.1029/2008WR007094.
117. Sivanandam SN, Deepa SN. *Introduction to Genetic Algorithms*. Berlin, He. Springer; 2008.
 118. Qu C-S, Ma Z-W, Yang J, Liu Y, Bi J, Huang L. Human exposure pathways of heavy metals in a lead-zinc mining area, Jiangsu Province, China. *PLoS One.* 2012;7(11):e46793. doi:10.1371/journal.pone.0046793.
 119. Zhao F, Ma Y, Zhu Y, Tang Z, Mcgrath SP. Soil Contamination in China: Current Status and Mitigation Strategies. *Environ Sci Technol.* 2015;49:750–759. doi:10.1021/es5047099.
 120. Petersen L, Minkinen P, Esbensen KH. Representative sampling for reliable data analysis: Theory of Sampling. *Chemom Intell Lab Syst.* 2005;77(1-2):261–277. doi:10.1016/j.chemolab.2004.09.013.
 121. Fortunati G, Pasturenzi M. Quality in soil sampling. *Química Analítica.* 1994;13(Suppl 1):S5–S20.
 122. Markert B. Quality assurance of plant sampling and storage. In: Quevauviller P, ed. *Quality assurance in environmental monitoring sampling and sample pre-treatment*. New York: VCH Weinheim; 1995:215–254.
 123. Andronikov S V, Davidson DA, Spiers RB. Variability in contamination by heavy metals: sampling implications. *Water Air Soil Pollut.* 2000;120:29–45.
 124. Hooda PS, Glavinandp RJ. A Practical Examination of the Use of Geostatistics in the Remediation of a Site with a Complex Metal Contamination History. *Soil Sediment Contam.* 2005;14(2):155–169. doi:10.1080/15320380590911814.
 125. Lord DW. Appropriate site investigation. In: Caimey T, ed. *Reclaiming Contaminated Land*. Blackie. London; 1987:62–113.
 126. Flatman GT, Yfantis AA. Geostatistical strategy for soil sampling: the survey and the census. *Environ Monit Assess.* 1984;4(1):335–349.
 127. Royal AG. *A Practical Introduction to Geostatistics*. Unpublished Course Notes, Leeds University; 1981.
 128. Theocharopoulos S., Wagner G, Sprengart J, et al. European soil sampling guidelines for soil pollution studies. *Sci Total Environ.* 2001;264(1-2):51–62. doi:10.1016/S0048-9697(00)00611-2.
 129. Lin Y-P, Chu H-J, Huang Y-L, Cheng B-Y, Chang T-K. Modeling spatial uncertainty of heavy metal content in soil by conditional Latin hypercube sampling and geostatistical simulation. *Environ Earth Sci.* 2010;62(2):299–311. doi:10.1007/s12665-010-0523-5.
 130. Milillo TM, Sinha G, Gardella J a. Use of geostatistics for remediation planning to transcend urban political boundaries. *Environ Pollut.* 2012;170:52–62. doi:10.1016/j.envpol.2012.06.006.
 131. Ha H, Olson JR, Bian L, Rogerson P a. Analysis of heavy metal sources in soil using kriging interpolation on principal components. *Environ Sci Technol.* 2014;48(9):4999–5007. doi:10.1021/es405083f.
 132. Demougeot-Renard H, De Fouquet C. Geostatistical approach for assessing soil volumes requiring remediation: validation using lead-polluted soils underlying a former smelting works. *Environ Sci Technol.* 2004;38(19):5120–5126.
 133. Qu M, Li W, Zhang C. Assessing the risk costs in delineating soil nickel contamination

- using sequential Gaussian simulation and transfer functions. *Ecol Inform.* 2013;13:99–105. doi:10.1016/j.ecoinf.2012.06.005.
134. Webster R, Oliver MA. Sample adequately to estimate variograms of soil. *J Soil Sci.* 1992;43:177–192.
 135. Lettenmaier DP, HUGHES J. Data Requirements for Kriging: Estimation and Network Design. 1981;17(6):1641–1650.
 136. Saito H, Goovaerts P. Geostatistical Interpolation of Positively Skewed and Censored Data in a Dioxin-Contaminated Site. *Environ Sci Technol.* 2000;34(19):4228–4235. doi:10.1021/es991450y.
 137. McGrath D, Zhang C, Carton OT. Geostatistical analyses and hazard assessment on soil lead in Silvermines area, Ireland. *Environ Pollut.* 2004;127(2):239–248. doi:10.1016/j.envpol.2003.07.002.
 138. Saby NP a., Marchant BP, Lark RM, Jolivet CC, Arrouays D. Robust geostatistical prediction of trace elements across France. *Geoderma.* 2011;162(3-4):303–311. doi:10.1016/j.geoderma.2011.03.001.
 139. Aelion CM, Davis HT, Liu Y, Lawson a B, McDermott S. Validation of Bayesian kriging of arsenic, chromium, lead, and mercury surface soil concentrations based on internode sampling. *Environ Sci Technol.* 2009;43(12):4432–8.
 140. Demougeot-Renard H, De Fouquet C, Renard P. Forecasting the number of soil samples required to reduce remediation cost uncertainty. *J Environ Qual.* 2004;33(5):1694–702.
 141. Marchant BP, McBratney a. B, Lark RM, Minasny B. Optimized multi-phase sampling for soil remediation surveys. *Spat Stat.* 2013;4:1–13. doi:10.1016/j.spasta.2012.11.001.
 142. Ersoy A, Yunsel TY, Faculty E. Geostatistical Conditional Simulation for the Assessment of Contaminated Land by Abandoned Heavy Metal Mining. *Environ Toxicol.* 2008;96–109. doi:10.1002/tox.
 143. Ersoy A, Yunsel TY, Cetin M. Characterization of Land Contaminated by Past Heavy Metal Mining Using Geostatistical Methods. *Environ Contam Toxicol.* 2004;46:162–175. doi:10.1007/s00244-003-2269-5.
 144. Verstraete S, Van Meirvenne M. A multi-stage sampling strategy for the delineation of soil pollution in a contaminated brownfield. *Environ Pollut.* 2008;154(2):184–91. doi:10.1016/j.envpol.2007.10.014.
 145. Gopalakrishnan G, Minsker BS, Valocchi AJ. Monitoring network design for phytoremediation systems using primary and secondary data sources. *Environ Sci Technol.* 2011;45(11):4846–53. doi:10.1021/es1042657.
 146. Meerschman E, Cockx L, Van Meirvenne M. A geostatistical two-phase sampling strategy to map soil heavy metal concentrations in a former war zone. *Eur J Soil Sci.* 2011;62(3):408–416. doi:10.1111/j.1365-2389.2011.01366.x.
 147. Xie Y, Chen T, Lei M, et al. Spatial distribution of soil heavy metal pollution estimated by different interpolation methods: accuracy and uncertainty analysis. *Chemosphere.* 2011;82(3):468–76. doi:10.1016/j.chemosphere.2010.09.053.
 148. Hicks PJ. Unconditional sequential Gaussian simulation for 3-D flow in a heterogeneous core. *J Pet Sci Eng.* 1996;16:209–219. doi:10.1016/S0920-4105(96)00041-1.
 149. Tica D, Udovic M, Lestan D. Immobilization of potentially toxic metals using different soil amendments. *Chemosphere.* 2011;85(4):577–583. doi:10.1016/j.chemosphere.2011.06.085.

150. Wada S. Dynamics of heavy metals in soils[In Japanese]. *Glob Environ Res.* 2010;15(1):15–21.
151. Ağca N. Spatial distribution of heavy metal content in soils around an industrial area in Southern Turkey. *Arab J Geosci.* 2014. doi:10.1007/s12517-013-1240-7.
152. Chen T-B, Zheng Y-M, Lei M, et al. Assessment of heavy metal pollution in surface soils of urban parks in Beijing, China. *Chemosphere.* 2005;60(4):542–51. doi:10.1016/j.chemosphere.2004.12.072.
153. Dayani M, Mohammadi J. Geostatistical Assessment of Pb, Zn and Cd Contamination in Near-Surface Soils of the Urban-Mining Transitional Region of Isfahan, Iran. *Pedosphere.* 2010;20(5):568–577. doi:10.1016/S1002-0160(10)60046-X.
154. Ettler V, Mihaljevič M, Kříbek B, Majer V, Šebek O. Tracing the spatial distribution and mobility of metal/metalloid contaminants in Oxisols in the vicinity of the Nkana copper smelter, Copperbelt province, Zambia. *Geoderma.* 2011;164(1-2):73–84. doi:10.1016/j.geoderma.2011.05.014.
155. Nguyen Ngoc M, Dultz S, Kasbohm J. Simulation of retention and transport of copper, lead and zinc in a paddy soil of the Red River Delta, Vietnam. *Agric Ecosyst Environ.* 2009;129(1-3):8–16. doi:10.1016/j.agee.2008.06.008.
156. Broos MJ, Aarts L, Tooren CF Van, Stein A. Quantification of the effects of spatially varying environmental contaminants into a cost model for soil remediation. *J Environ Manage.* 1999;56:133–145.
157. Maringanti C, Chaubey I, Arabi M, Engel B. Application of a multi-objective optimization method to provide least cost alternatives for NPS pollution control. *Environ Manage.* 2011;48(3):448–61. doi:10.1007/s00267-011-9696-2.
158. Kitanidis PK. *Introdction to geostatistics:application to hydrogeology.* Cambridge University Press; 1997.
159. Dube A, Kowalkowski RZT, Cukrowska E, Buszewski B. Adsorption and Migration of Heavy Metals in Soil. 2001;10(1):1–10.
160. Lin Y-P, Chang T-K, Shih C-W, Tseng C-H. Factorial and indicator kriging methods using a geographic information system to delineate spatial variation and pollution sources of soil heavy metals. *Environ Geol.* 2002;42(8):900–909. doi:10.1007/s00254-002-0600-5.
161. He Y, Chen D, Li BG, et al. Sequential indicator simulation and indicator kriging estimation of 3-dimensional soil textures. *Aust J Soil Res.* 2009;47(6):622–631.
162. Maiorana M, Fornaro F, Lopez N, Castrignano a. 3D spatial variability of soil strength and its change over time in a durum wheat field in Southern Italy. *Soil Tillage Res.* 2002;65(1):95–108. doi:10.1016/S0167-1987(01)00288-4.
163. Laslett GM, McBratney AB, Pahl PJ, Hutchinson MF. Comparison of several spatial prediction methods for soil pH. *Eur J Soil Sci Soil Sci.* 1987;38(2):325–341. doi:10.1111/j.1365-2389.1987.tb02148.x.
164. Jones NL, Davis RJ, Sabbah W. A Comparison of Three-Dimensional Interpolation Techniques for Plume Characterization. *GROUND WA T E R.* 2003; 4 1 (4):411–419.
165. Perroy RL, Belby CS, Mertens CJ. Mapping and modeling three dimensional lead contamination in the wetland sediments of a former trap-shooting range. *Sci Total Environ.* 2014;487:72–81. doi:10.1016/j.scitotenv.2014.03.102.
166. Ouyang Y, Higman J, Campbell D, Davis J. Three-dimensional kriging analysis of sediment

- mercury distribution: A case study. *J Am Water Resour Assoc.* 2003;39(3):689–702. doi:10.1111/j.1752-1688.2003.tb03685.x.
167. Ouyang Y, Higman J, Thompson J, O'Toole T, Campbell D. Characterization and spatial distribution of heavy metals in sediment from Cedar and Ortega rivers subbasin. *J Contam Hydrol.* 2002;54(1-2):19–35. doi:10.1016/S0169-7722(01)00162-0.
 168. Piedade TC, Melo VF, Souza LCP, Dieckow J. Three-dimensional data interpolation for environmental purpose: lead in contaminated soils in southern Brazil. *Environ Monit Assess.* 2014;186(9):5625–5638. doi:10.1007/s10661-014-3808-4.
 169. Zawadzki J, Magiera T, Fabijańczyk P, Kusza G. Geostatistical 3-dimensional integration of measurements of soil magnetic susceptibility. *Environ Monit Assess.* 2012;184:3267–3278. doi:10.1007/s10661-011-2187-3.
 170. Mulder VL, Lacoste M, Richer-de-Forges AC, Martin MP, Arrouays D. National versus global modelling the 3D distribution of soil organic carbon in mainland France. *Geoderma.* 2016;263:16–34. doi:10.1016/j.geoderma.2015.08.035.
 171. Van Meirvenne M, Maes K, Hofman G, Meirvenne M Van, Maes K. Three-dimensional variability of soil nitrate-nitrogen in an agricultural field. *Biol Fertil Soils.* 2003;37(3):147–153. doi:10.1007/s00374-003-0581-x.
 172. Liu C-W, Jang C-S, Chen S-C. Three-dimensional spatial variability of hydraulic conductivity in the Choushui River alluvial fan, Taiwan. *Environ Geol.* 2002;43(1-2):48–56. doi:10.1007/s00254-002-0648-2.
 173. Chen C, Hu K, Li W, Li Z, Li B. Three-dimensional mapping of clay content in alluvial soils using hygroscopic water content. *Environ Earth Sci.* 2015;73(8):4339–4346. doi:10.1007/s12665-014-3720-9.
 174. Gruver A, A, Dutton J. Inverse Distance Weighting. Available at: <https://www.e-education.psu.edu/geog486/node/1877>.
 175. Liang L, Hale D, Liang L, Hale D. A stable and fast implementation of natural neighbor interpolation. 2010:1–14. Available at: <http://cwp.mines.edu/Documents/cwpreports/cwp-657.pdf>.
 176. Sibson R. *A brief description of natural neighbor interpolation (Chapter 2).*; 1981.
 177. Boissonnat J-D, Cazals F. Smooth surface reconstruction via natural neighbour interpolation of distance functions. *Comput Geom.* 2002;22:185–203. doi:10.1016/S0925-7721(01)00048-7.
 178. Ledoux H, Gold C. An efficient natural neighbour interpolation algorithm for geoscientific modelling. *Dev Spat Data Handl.* 2004:97–108. doi:10.1007/3-540-26772-7_8.
 179. How Natural Neighbor works. Available at: http://resources.esri.com/help/9.3/arcgisengine/java/gp_toolref/spatial_analyst_tools/how_natural_neighbor_works.htm.
 180. Chen C, Hu K, Li H, Yun A, Li B. Three-Dimensional Mapping of Soil Organic Carbon by Combining Kriging Method with Profile Depth Function. *PLoS One.* 2015;10(6):e0129038. doi:10.1371/journal.pone.0129038.
 181. Li HY, Li FH, Shi Z, Huang MX. Three Dimensional Variability of Soil Electrical Conductivity Based on Electromagnetic Induction Approach. *Artif Intell Comput Intell Int Conf.* 2010;3(December 2006):219–223. doi:<http://doi.ieeecomputersociety.org/10.1109/AICI.2010.284>.

182. Koike K, Gu B, Ohmi M. Three-dimensional distribution analysis of phosphorus content of limestone through a combination of geostatistics and artificial neural network. *Nonrenewable Resour.* 1998;7(3):197–210. doi:10.1007/BF02767670.
183. Setianto A, Triandini T. Comparison of Kriging and Inverse Distance Weighted (IDW) interpolation methods in lineament extraction and analysis. *J Southeast Asian Appl Geol.* 2013;5(1):21–29.
184. Carrillo-González R, Šimůnek J, Sauvé S, Adriano D. Mechanisms and Pathways of Trace Element Mobility in Soils. *Adv Agron.* 2006;91(06):111–178. doi:10.1016/S0065-2113(06)91003-7.
185. Adriano DC, Wenzel WW, Vangronsveld J, Bolan NS. Role of assisted natural remediation in environmental cleanup. *Geoderma.* 2004;122(2-4):121–142. doi:10.1016/j.geoderma.2004.01.003.
186. Adriano DC. *Trace Elements in Terrestrial Environments: Biogeochemistry, Bioavailability and Risks of Metals.* New York: Springer; 2001.
187. Wu G, Li LY. Modeling of heavy metal migration in sand/bentonite and the leachate pH effect. *J Contam Hydrol.* 1998;33(3-4):313–336. doi:10.1016/S0169-7722(98)00075-8.
188. Lo IMC, Asce M, Luk AFT, Yang X. Migration of Heavy Metals in Saturated Sand and Bentonite Õ Soil Admixture. 2004;(August):906–910.
189. Mikami S, Maeyama T, Hoshide Y, et al. Spatial distributions of radionuclides deposited onto ground soil around the Fukushima Dai-ichi Nuclear Power Plant and their temporal change until December 2012. *J Environ Radioact.* 2015;139:320–43. doi:10.1016/j.jenvrad.2014.09.010.
190. Srinivas C, Rakesh P. Assessment of atmospheric dispersion and radiological impact from the Fukushima accident in a 40-km range using a simulation approach. *Air Qual Atmos Heal.* 2014;7(2):209–227. doi:10.1007/s11869-014-0241-3.
191. CESIUM. *Gale Group, Inc.* 2006. Available at: <http://www.encyclopedia.com/topic/cesium.aspx>.
192. ATSDR. Publication Healthy Statment: Cesium. 2004. Available at: <http://www.atsdr.cdc.gov/ToxProfiles/tp157-c1-b.pdf>.
193. Radiation - Quantities and Units of Ionizing Radiation. Available at: http://www.ccohs.ca/oshanswers/phys_agents/ionizing.html.
194. Guidance for Radiation Accident Management. Available at: <https://orise.ornl.gov/reacts/guide/measure.htm>.
195. Bentor Y. Chemical Element.com - Cesium. 2015. Available at: <http://www.chemicalelements.com/elements/cs.html>.
196. Matsuda N, Mikami S, Shimoura S, et al. Depth profiles of radioactive cesium in soil using a scraper plate over a wide area surrounding the Fukushima Dai-ichi Nuclear Power Plant, Japan. *J Environ Radioact.* 2015;139:427–34. doi:10.1016/j.jenvrad.2014.10.001.
197. Ivanov Y a., Lewyckij N, Levchuk SE, et al. Migration of ¹³⁷Cs and ⁹⁰Sr from Chernobyl fallout in Ukrainian, Belarussian and Russian soils. *J Environ Radioact.* 1997;35(1):1–21. doi:10.1016/S0265-931X(96)00036-7.
198. Szerbin P, Koblinger-Bokori E, Koblinger L, Végvári I, Ugron a. Caesium-137 migration in Hungarian soils. *Sci Total Environ.* 1999;227(2-3):215–27. doi:10.1016/S0048-9697(99)00017-0.

199. Gaspar L, Navas A. Vertical and lateral distributions of ^{137}Cs in cultivated and uncultivated soils on Mediterranean hillslopes. *Geoderma*. 2013;207-208:131–143. doi:10.1016/j.geoderma.2013.04.034.
200. Almgren S, Isaksson M. Vertical migration studies of ^{137}Cs from nuclear weapons fallout and the Chernobyl accident. *J Environ Radioact*. 2006;91:90–102. doi:10.1016/j.jenvrad.2006.08.008.
201. Bunzl K, Schimmack W, Zelles L, Albers BP. Spatial variability of the vertical migration of fallout ^{137}Cs in the soil of a pasture, and consequences for long-term predictions. *Radiat Environ Biophys*. 2000;39:197–205. doi:10.1007/s004110000062.
202. Bossew P, Kirchner G. Modelling the vertical distribution of radionuclides in soil. Part 1: The convection-dispersion equation revisited. *J Environ Radioact*. 2004;73(2):127–150. doi:10.1016/j.jenvrad.2003.08.006.
203. Jagercikova M, Cornu S, Le Bas C, Evrard O. Vertical distributions of ^{137}Cs in soils: a meta-analysis. *J Soils Sediments*. 2015;15(1):81–95. doi:10.1007/s11368-014-0982-5.
204. Velasco RH, Menegont S. Vertical transport of radiocesium in surface soils: model implementation and dose-rate computation. 1993;(3).
205. Chamard P, Velasco R, Belli M, Disilvestro G, Ingrao G, Sansone U. Caesium-137 and strontium-90 distribution in a soil profile. *Sci Total Environ*. 1993;136(3):251–258. doi:10.1016/0048-9697(93)90313-U.
206. Bunzl K, Kracke W, Schimmack W, Auerswald K. Migration of fallout $^{239} + ^{240}\text{Pu}$, ^{241}Am and ^{137}Cs in the various horizons of a forest soil under pine. *J Environ Radioact*. 1995;28(1):17–34. doi:10.1016/0265-931X(94)00066-6.
207. Bunzl K, Schimmack W, Krouglov SV, Alexakhin RM. Changes with time in the migration of radiocesium in the soil, as observed near Chernobyl and in Germany, 1986–1994. *Sci Total Environ*. 1995;175(1):49–56. doi:10.1016/0048-9697(95)04842-1.
208. Saikia K, Sarkar BC, Sinha PM. Application of kriging and simulated annealing for spatial variability modelling of a coal seam. *Appl Earth Sci IMM Trans Sect B*. 2007;116(1):37–48. doi:10.1179/174327507X167000.
209. Rogowski AS, Wolf JK, Simmons DE. Conditional simulation of flow and transport. *J Contam Hydrol*. 1991;7(1–2):95–121. doi:10.1016/0169-7722(91)90040-8.
210. Journel AG, Huijbregts CJ. *Mining Geostatistics*. London: ACADEMIC PRESS; 1997.
211. Jean-Paul C, Pierre D. *Geostatistics: Modeling Spatial Uncertainty*. New York: A Wiley-Interscience Publication; 1999.