uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Mohammad Hafizur Rahman
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.A.Sc. (Civil Engineering)
GRADE / DEGREE

School of Information Technology and Engineering
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Spatial and Temporal Analysis of Landfill Leachate Characteristics at Trail Road Landfill Site

TITRE DE LA THÈSE / TITLE OF THESIS

M. Warith
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

B. Danesfar
CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

Paul Van Geel

I. Nistor

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

SPATIAL AND TEMPORAL ANALYSIS OF LANDFILL LEACHATE CHARACTERISTICS AT TRAIL ROAD LANDFILL SITE

By

Mohammad Hafizur Rahman

A Thesis
Presented to the University of Ottawa in Partial Fulfillment of the Requirements for Master of Applied Science in Civil Engineering

Department of Civil Engineering
University of Ottawa
Ottawa, Canada
K1N 6N5

November 2009

The M.A.Sc. in Civil Engineering
is a joint program with Carleton University administered by
The Ottawa-Carleton Institute for Civil Engineering

Your file Votre référence
ISBN: 978-0-494-65506-1
Our file Notre référence
ISBN: 978-0-494-65506-1

THE UNIVERSITY OF OTTAWA-FACULTY OF GRADUATE STUDIES

CERTIFICATE OF EXAMINATION

Examining Board

Advisor

The thesis by
Mohammad Hafizur Rahman
entitled
Spatial and Temporal Analysis of Landfill Leachate Characteristics at Trail Road Landfill Site

is accepted in partial fulfillment
of the requirements of the degree of
Master of Applied Science (Water Resources Engineering)
in the
Department of Civil Engineering

Date:

Chairperson of Examining Board

APPROVED BY:

ACKNOWLEDGEMENT

First of all, I am very grateful to my promoter Prof. Dr. Leta Fernandes (late) for suggesting the research and initiating the idea to work in such an interesting field. I feel happy that I was true to her research ideas. Unfortunately, she has passed on, and I wish her heaven life.

I would like to express my deep gratitude to my supervisors Dr. Bahram Daneshfar and Dr. Mostafa Warith for their continuous guidance through my long endeavour for my M.A.Sc. at the University of Ottawa. I was very fortunate to have them as my supervisors. I have learned a lot from their precious suggestions, ideas and techniques. I would like to thank them both for their assistance, guidance, encouragement and support during the two past years. I am also grateful to Dr. Warith who gave me spontaneous support in my research. I have benefited very much not only from my supervisors' support, vast knowledge and experience, but also their very pragmatic advice.

I would like to express my gratitude to the many people who helped me to study at University of Ottawa. All of whom are very friendly and helpful.

I very much enjoyed studying and learning at University of Ottawa.

October 30, 2009

DEDICATION

In the name of Allah, the most gracious,

To

Prof. Dr. Leta Fernandes (late)

Rabeya Begum (late)

Dr. Rabeya Naznin

ABSTRACT

This study is conducted at the Trail Road Landfill, located in Nepean, Ontario, Canada. The objective is to investigate the leachate characteristics of changing spatial-temporal patterns in a landfill groundwater environment by comprehensive analyses of annual spatial data.

Exploratory statistical data analysis identified the association of B (boron) with K, NH_3 and TKN. Raster layers (maps) are created based on the concentrations of required variables in each time interval (year). In this study, it is notable that the raster data layers are used instead of discrete well data.

Several change detection methods are applied to determine the spatial and temporal changes of B and its associated variables and to identify the well locations where the changes occurred. These included post-classification visualization, principal component analysis, standard deviation and unsupervised classification (clustering) methods. The suitability of these methods is also discussed. The results determined that during the 1993-95 time period the concentrations of B and its associates was initially increasing, and then decreased substantially.

In summary, the study analysed characteristics of pollutants in landfill site groundwater environmental monitoring by using raster data in different change detection methods, and discussed the suitability of the applied methods. The same methodology and analysis techniques can be applied to other variables in similar environmental monitoring studies.

TABLE OF CONTENTS

PAGE

PAGE

LIST OF FIGURES

LIST OF TABLES

GLOSSARY/NOTATION

h = hydraulic head

Δh = hydraulic gradient

K = hydraulic conductivity

n = effective porosity

R_c = critical value

R_{Lo} = lower outlier statistics

R_{Hi} = higher outlier statistics

S_x = standard deviation

v = velocity

$\overline{X}$ = mean

X_1 = lowest rank

X_n = highest rank

Abbreviations:

B – Boron

COD – Chemical Oxygen Demand

CV – Coefficient of Variation

EDA – Exploratory Data Analysis

EPA - Environmental Protection Agency

GIS – Geographic Information System

IDW – Inverse Distance Weightage

K – Potassium

KMO – Kaiser-Meyer-Olkin

K-S – Kolmogrov-Smirnov

MSW – Municipal Solid Waste

NH_3 – Ammonia

PCA – Principal Component Analysis

Q-Q – Quantile-Quantile

ROI – Region of Interest

SD – Standard Deviation

TKN – Total Kjeldahl Nitrogen

TRL – Trail Road Landfill

TOC - Total Organic Carbon

XOCs – Xenobiotic Organic Compounds

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Landfill leachate pollution is an increasing problem in most areas, and it is considered one of the most critical aspects of Municipal Solid Waste (MSW) management. Landfilling is the process by which a mixture of solid waste residue is disposed of in a landfill site. Runoff or direct infiltration from landfill due to precipitation often forms leachate, which can cause serious groundwater pollution. Concern about the problem is growing, and MSW management intends to apply new technologies and techniques to solve it. As part of this effort, this study uses the GIS technique to analyze spatial and temporal variations of leachate at the Trail Road Landfill (TRL). TRL is owned and operated by the City of Ottawa, Ontario, Canada, and municipal solid waste has been disposed of there since 1980.

According to previous literature, the TRL has been developed in four stages. Stage 1 (1980-1987) and Stage 2 (1988-1991) do not have a bottom liner. Stage 3 and 4 implemented after 1993, include clay and geomembrane bottom liners and leachate collection systems to prevent leachate contamination through percolation. These structural components of the landfill play an important role in groundwater contamination at the site, and this study uses spatial-temporal analysis to determine the impact.

The study focuses on deep aquifer leachate concentration data. Although the landfill leachate consists of several chemicals, boron (B) and related leachate groundwater pollution compounds are the main subject of the study. Many studies, (Bjerg *et al.*, 2003; Al-Yaqout *et al.*, 2005; and Al-Yaqout *et al.*, 2003), have investigated municipal landfills known to have high contents of organic and inorganic matter. The study focuses on B and associated variables to study their spatial and temporal characteristics. .

Among the many contributing elements in leachate composition, the study focuses on B for specific reasons. It exhibits special behaviour in leachate, and like other conservative elements B is not affected by chemicals from non-landfill sources (i.e. highway salt, agriculture pesticide, septic sludge, etc.). As a conservative element originating from the landfill site, B can represent extents of the contamination plum. Therefore its spatial and temporal variations can indicate such variability of the maximum extent of the contaminated plum. B is an essential nutrient for plants, but in high concentrations it causes the soil to become toxic. B released from landfills requires special attention with regard to both agricultural nutrition and drinking water. Bagchi (1994) determined that the mobility of B in clayey layers is high.

In applying the GIS technique, the availability of local scale GIS data layers was limited, so the local scale spatial data layers have been retrieved from scanned documents through GIS. Microsoft Access database format is used to manage the data, and it is directly linked to GIS. Once a GIS database is developed, it can provide an efficient means of analyzing landfill data. A comprehensive approach of the GIS application combined with statistical data analysis is applied, and is described in Chapter 4.

The example provided in this study is specific to lead in the analysis of B. However, the methodology is applicable for all other leachate compositions identified in the landfill site.

1.2 STUDY AREA

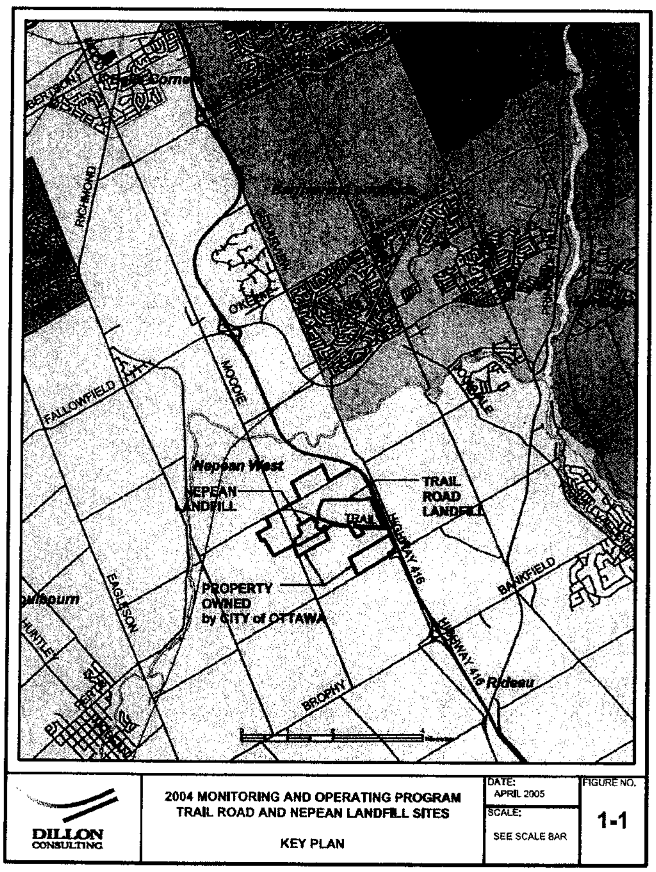
The triangular shaped landfill area is bounded by Cambrian Road to the north, Trail Road to the south and Highway 416 to the east, as shown in Figure 1.1. According to a 2004 Monitoring and Operating Report, the main access road is Moodie Drive west of the landfill, which crosses Fallowfield Road to the north and Bankfield Road to the south. This report indicates that the area north of the Trail Road landfill site shows the most dynamicity of plume due to dewatering pond. Based on the study report of the Trail Optimization/Expansion Project (2001), the landfill consists of the four stages described below.

Stage 1: The landfilling began in 1980, and was capped by geomembrance in 1987. It was located on 26 ha on the east side of the landfill. Leachate maintenance in this stage was by natural attenuation process. The groundwater flow in the shallow aquifer flows to the northeast of the landfill (Figure 1.4) and the deep aquifer of this stage is directed from the north half of the stage to the northwest of the landfill (Figure 1.5). .The grading of the existing base is from 104.0 m at southeast corner of the landfill to 102.0 at northwest corner.

Stage 2: The landfilling of Stage 2 took place from 1988 to 1991. The leachate attenuation of this stage was likely due to the management of Stage 1. The grading of the base of Stage 2 is 98.0 m. The north of the stage meets with underlying clay at Stage 1. The clay plain slopes toward the northeast and is followed by the leachate flows.

Stage 3: The landfilling of Stage 3 was from 1990 to 2000. Like Stage 2, the base elevation is 98.0 m. The stage differs from previous stages due to inclusion of bottom liners which prevent leachate percolating to the deep aquifer.

Stage 4: The landfilling in this stage began in 1999 and still is in operation. The base elevation at the north end was set at 96.6 m. Like Stage 3, Stage 4 consists of bottom liners and a leachate collection system.



(Source: Final Report, 2005 by Dillon Consulting Limited)

Figure 1.1 Key plan of landfill site

1.3 AIMS AND OBJECTIVES

The principal objective of this research is to study spatial and temporal variations of B at the TRL landfill site. The specific objectives of the study are to:

- ◊ evaluate the leachate impacts by assessing its spatial and temporal variation and its link to expanding or receding contamination plume by different change detection methods;
- ◊ evaluate the suitability of applied change detection methods in landfill site environmental monitoring; and,
- ◊ identify the affected wells and their locations in the landfill site.

1.4 OUTLINE OF THESIS

The following chapters are the theoretical framework underlying the spatial and temporal analysis on which the work in this thesis is based.

Chapter 1 describes the background information, site conditions and data availability of the landfills.

Chapter 2 details the methodology and the use of the techniques.

Chapter 3 describes the related literature of the study.

Chapter 4 explains the applied exploratory data analysis and indicates the results of applied multivariate statistical analysis techniques.

Chapter 5 discusses various methods that are applied for change detection of B (and related elements) to determine their spatial-temporal variation.

Chapter 6 summarizes the contents of the thesis and suggests potential directions for future research.

1.5 GEOLOGY AND HYDROGEOLOGICAL SETTINGS

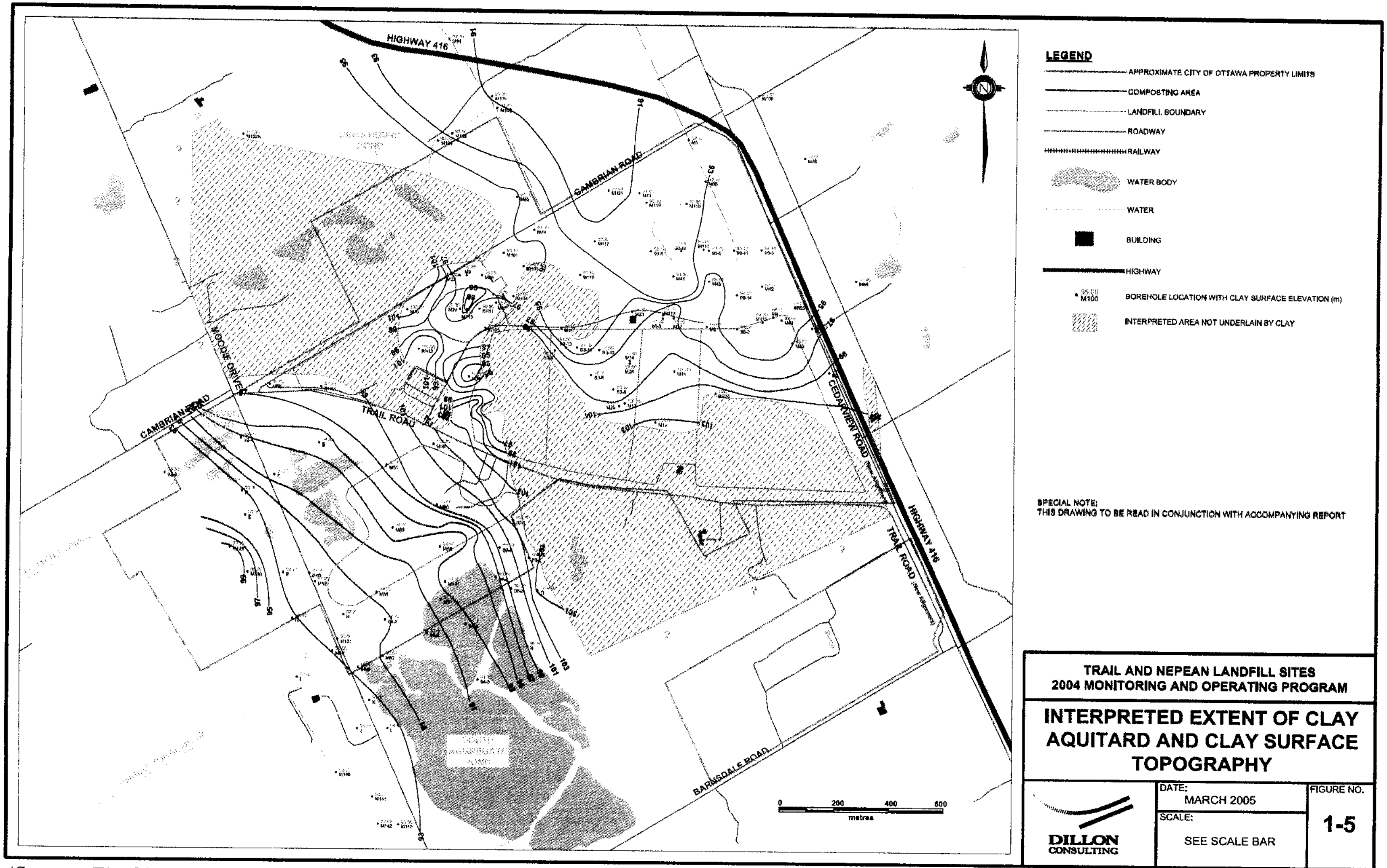
1.5.1 SITE CONDITIONS

According to the Trail Landfill Optimization/Expansion Project (2001), TRL is located on the east side of the trending sand and gravel deposit ridge and is surrounded by clay plains, whereas the Nepean Landfill is located on the west side of the ridge. The ridge varies in silt and clay content, extending from the bedrock surface to the ground surface. The site consists of two main types of aquifers: shallow and deep. The shallow aquifer is comprised of the fine to medium sand, and the deep aquifer of sand and gravel geological units. An extended clay aquitard made up of marine silt and clay geological unit is also found at the site (Figure 1.2). There is a window-like opening that has formed due to discontinuities of the clay aquitard, and this is very important in its hydrogeological regimes, particularly in TRL. The presence of windows in the clay indicates that the shallow aquifer is hydraulically connected or continuous with the water table of the deep aquifer. Where the clay is not present the deep aquifer is the water table aquifer, and where the clay is present the deep aquifer is a confined unit. In the northeastern, western and northern areas of the site, the deep aquifer and shallow aquifer are separated by the clay confining unit. The deep aquifer is found beneath almost the entire areas of the Trail Road and Nepean landfill sites, whereas the shallow aquifer is found only where the discontinuous clay layer is present.

A dewatering pond was excavated at the north side of the TRL to observe landfill performance, particularly that of the deep aquifer. The dewatering pond shifts the direction of groundwater flow beneath the landfill from north-east to north-west, and it plays an important role in the hydrogeological conditions of the site. According to the TRL Asset Management Study (1998) a major change of groundwater flow direction in deep aquifer was identified when the dewatering pond was formed.

The clay layer plays an important role in directing the groundwater flow at both sites. At the north end of the TRL, shallow aquifer flow continues northerly in the direction of the clay layer gradients. In deep aquifer, flow continues to the dewatering pond in the northwest side due to the gradient. It is evident that the groundwater flow in deep aquifer of the southern half of Stage 1 and 2 is northwesterly (Figure 1.4), and the northern half of shallow aquifer flows northeasterly.

The geological setting of deep aquifer comprises a complex layering of sand to sand and gravel beds that follow bedrock underlying the entire study area. In some locations, the layer is found as a sand and gravel bed that follows the underlying discontinuous thick glacial layer.



(Source: Final Report, 2004 by Dillon Consulting Limited)

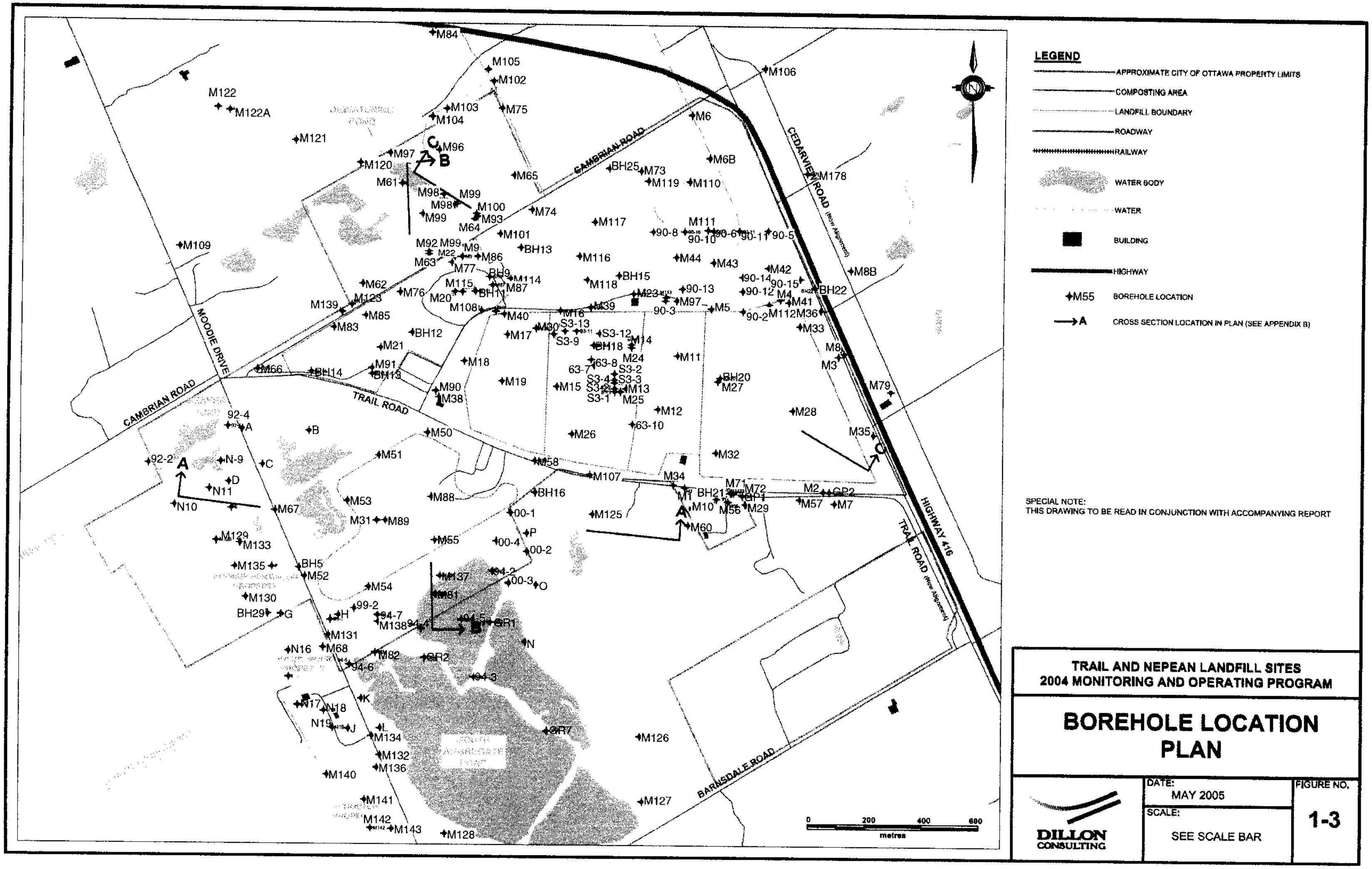
Figure 1.2 Extents of clay aquitard and clay surface topography

1.5.2 DATA AVAILABILITY

Quarterly (4 times in a year) leachate concentrations that are collected from the TRL site have been used in this study. Fourteen years of data (from 1992 to 2005) are collected from approximately 130 monitoring wells. Among them, few wells had less than 14 years monitoring data as some new wells were installed or some were closed during monitoring. The study uses this data, which was compiled by Dillon Consulting Limited, as well as yearly hardcopy monitoring reports of the Trail and Nepean Landfills. In the data series there were very few lower than detection limit values. Such few values were replaced by half of the detection limit values. A Microsoft Access database is created to store the data, which includes all well coordinates and ID's. Different local scale GIS data layers have been input and stored in the database for the study. The local scale spatial data layers, landfill layout, landfill boundary, local roads and well locations have been taken from the Final Report, (2004) by Dillon Consulting Limited. Data regarding highways, water bodies, rivers and greenbelt areas are also available for the study.

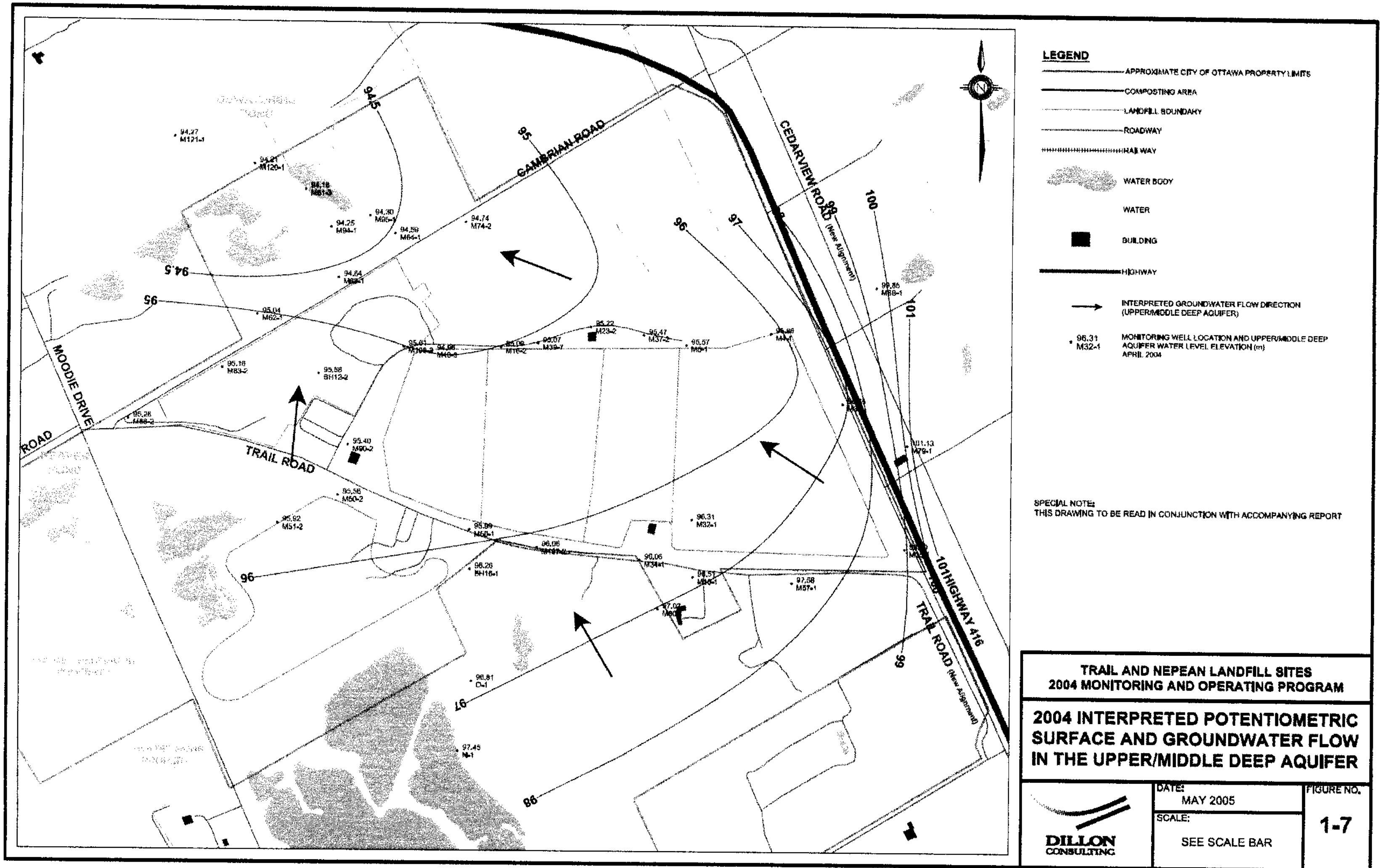
1.5.3 LOCATION OF MONITORING WELLS AND HEAD

Figure 1.3 shows the borehole location plan at the landfill sites. There are no groundwater monitors in the lower deep and upper deep aquifer beneath Stages 3 and 4. The study by Dillon Consulting shows (Figure 1.4 and Figure 1.5) the gradient of the groundwater and the direction of flow at the landfill site. It also indicates that in deep aquifer groundwater flows mostly to the northwest dewatering pond. The dewatering pond plays a significant role in deep aquifer groundwater flow.



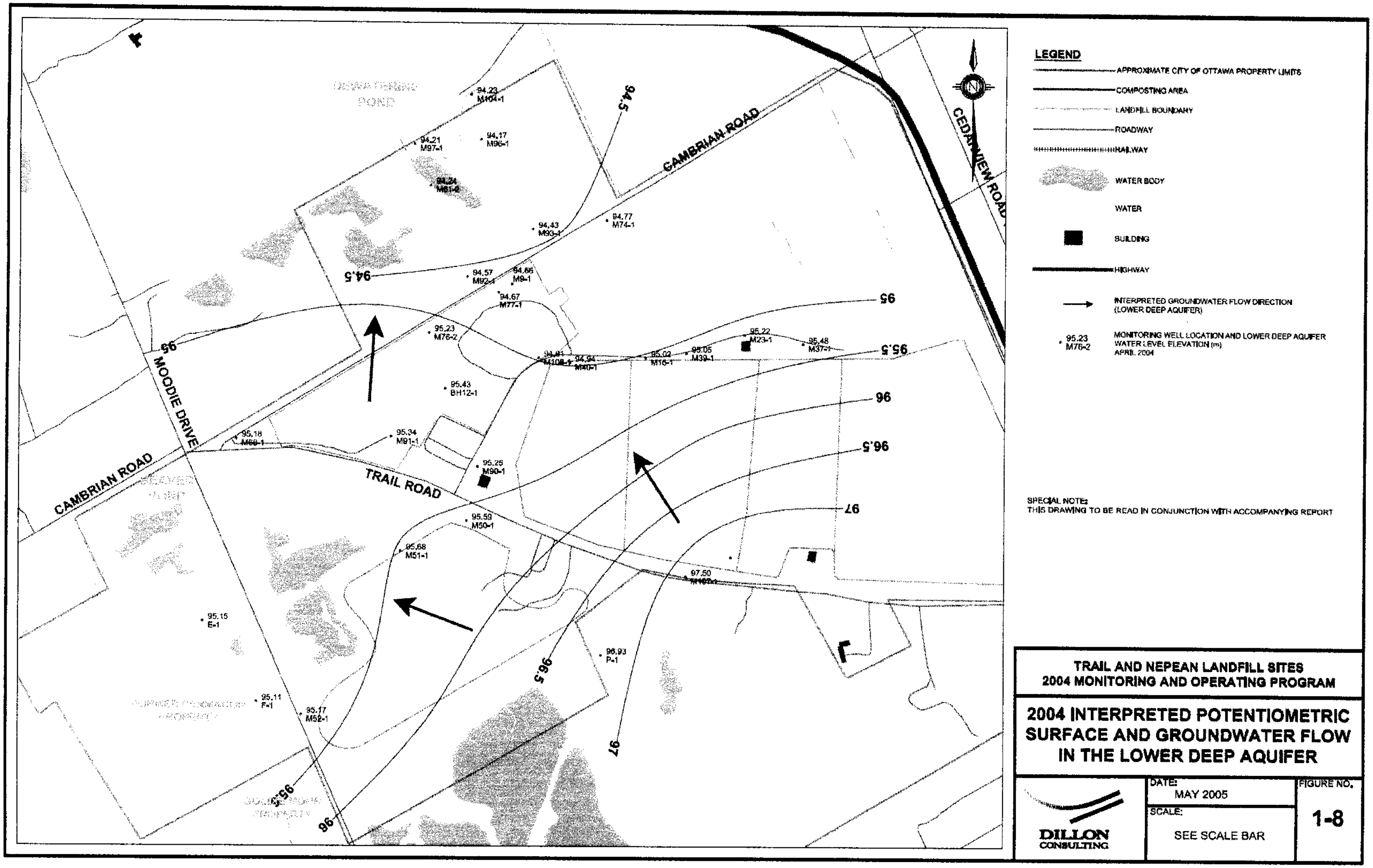
(Source: Final Report, 2004 by Dillon Consulting Limited)

Figure 1.3 Monitoring well locations of the study area



(Source: Final Report, 2004 by Dillon Consulting Limited)

Figure 1.4 Groundwater level and direction of flow in the study area for upper/middle deep aquifer



(Source: Final Report, 2004 by Dillon Consulting Limited)

Figure 1.5 Groundwater level and direction of flow in the study area for lower deep aquifer

CHAPTER 2

BACKGROUND LITERATURE

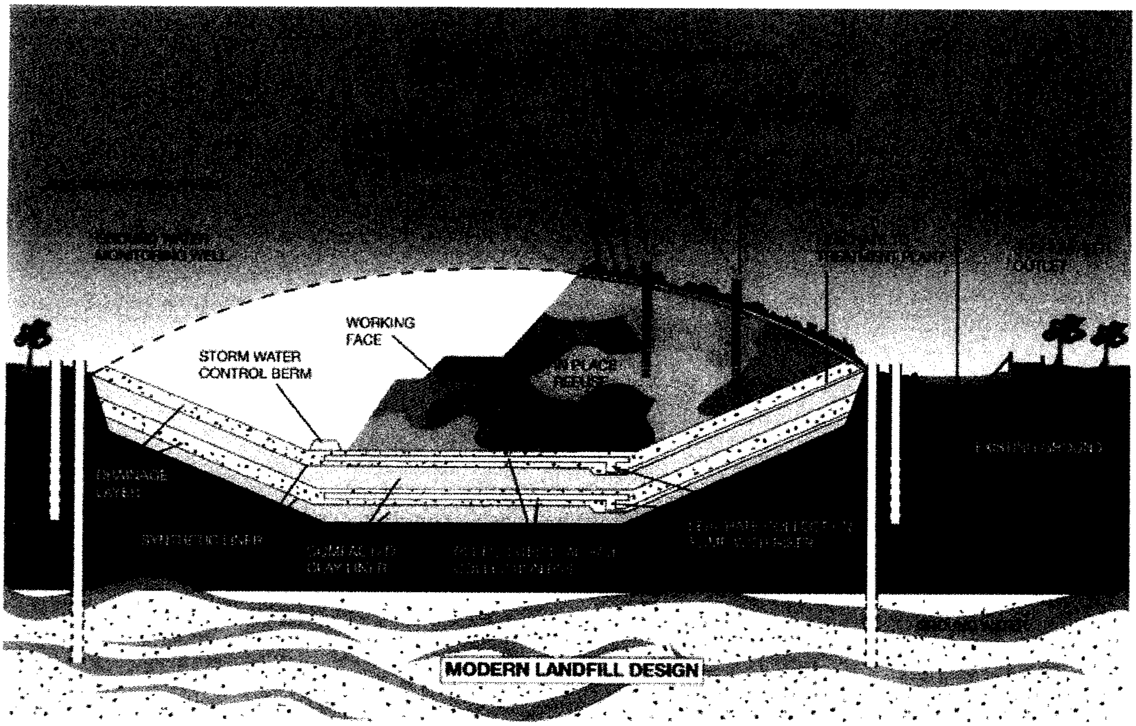
2.1 LEACHATE GENERATION

Leachate generation is an unavoidable problem associated with landfills. Once moisture content exceeds the field capacity, lechate is generated. The total moisture content retained by the waste is subject to the pull of gravity. Precipitation is the main source of moisture in waste, and it determines the degradation or stabilization of landfills. Reinhart and Grosh (1997) explained how degradation and stabilization occur in a landfill through physical, chemical and biological processes occurring in the waste over time. The biological process generates organic and mineral compounds, in addition to the products of the physical and chemical processing of the waste. The products washed out by percolating rainwater form leachate. Wang *et al.* (2004) and Jorstad *et al.* (2004), state the factors affecting leachate generation are:

- ◊ the quality of waste;
- ◊ techniques of landfilling and degree of waste compaction;
- ◊ age of the landfill;
- ◊ biochemical and physical process of waste decomposition;
- ◊ the moisture and absorption capacity of the waste;
- ◊ the topography of the landfill site;
- ◊ the lining system;
- ◊ hydrogeology; and,
- ◊ vegetation.

Precipitation and climate have a significant influence on leachate generation. There are many other factors that influence it as well, including surface runoff, infiltration, evaporation, evapotranspiration, ambient temperature, waste composition, density, initial moisture content and depth of the landfill. Wang *et al.* (2004) state that the absorption capacity of wastes is another factor affecting leachate production. The absorption capacity is governed by the density of the waste and the pathway of liquids infiltrating through it. Generally, an increase of waste density decreases leachate production.

Figure 2.1 displays a cross-section of a modern landfill with its different components. It shows the perforated leachate collection pipe with synthetic and confined clay liners. It plays an important role in the prevention of leachate pollution through percolation.



(Source: RUNCO Environmental, Inc., Birmingham)

Figure 2.1 Cross section of a modern landfill

2.2 LANDFILLS AS GROUNDWATER POLLUTION

Groundwater pollution occurs in many ways when pollutants leach from hazardous waste sources. Among these causes, landfills are an abundant source of groundwater pollution. Bjerg *et al.* (2003) studied leachate and determined it contains a variety of components, including dissolved organic matter, inorganic compounds, heavy metals and XOCs, due to the mixed nature of the waste in landfills. A landfill is seen as a complex source, as it is expected to last for decades or even centuries. Landfills also often cover large areas and exhibit spatial variability in the area due to many different leachate compositions. Without doing any hydrogeological investigation, old landfills can be presumed to be pollution prone if they are located near environmentally sensitive resources. Old landfills are not usually designed with bottom liner and leachate collection systems, which help to prevent serious seeping of leachate into groundwater. Modern landfills typically do have the bottom liner and leachate collection systems, and they have better leachate management techniques as well. However, even though newer landfills are well managed, they are not totally non-polluting. The pollution can be reduced by shredding the waste rather than sporadic dumping. It has been found that if waste is shredded, landfill pollution also becomes less complex. Jones-Lee and Lee (2000) determined that shredded waste significantly reduced landfill gas production.

According to many literature reviews, (i.e. Bjerg *et al.*, 2003), the vulnerability of groundwater pollution at landfill sites depends on a number of factors, such as:

◊ the nature of landfilling;
◊ landfill design;
◊ the nature of aquifers, and whether they are confined or unconfined;
◊ soil porosity; and,
◊ the nature of the vadose zone.

Most household waste is dumped into municipal landfills, and this generates the highest concentration of frequently recorded pollutants: heavy metals, organic compounds and methane. A diverse range of other pollutant materials are also found, some of which are potentially hazardous. In the landfill, all these pollutants form the leachate by mixing with rainwater and percolate into the groundwater.

2.2.1 POLLUTANTS TRANSPORT IN GROUNDWATER

All compounds in leachate entering an aquifer will be subject to physicochemical processes of advection, dispersion (dilution) and sorption in order to transport pollutants in groundwater. Advection transports dissolved contaminant mass due to the bulk flow of groundwater, and according to the seepage velocity in the pore space. It is considered the most dominant contaminant mass transport mechanism. The speed and direction of groundwater flow is characterized by the average linear velocity, which is equal to the rate of contaminant or fluid flowing in porous medium, or seepage velocity. Therefore, advection refers to the transport of contaminants at the same speed as the linear velocity of groundwater flow or the average pore velocity. McBean *et el.* (1995) and Bedient *et al*, (1999) explained this average linear velocity can also be determined from Darcy's law as:

$$v = -\frac{K}{n}\Delta h$$

where, K is the hydraulic conductivity tensor of porous medium $\lfloor LT^{-1} \rfloor$, h is the hydraulic head of the groundwater at a location in space [L], Δh is the hydraulic gradient [-], and n is the porosity of the aquifer material.

Bedient *et al.* (1999) studied how dissolved constituents or solutes introduced to groundwater migrate by diffusion and dispersion processes. Diffusion occurs as the pollutants move from higher concentration to lower concentration. The general process of spreading is termed dispersion. Dispersion is the distribution of mass beyond the region affected by advection. Due to the varying pore velocities and soil heterogeneity, varying flow velocities and directions lead to dispersion.

Dispersion causes mixing of pollutants with uncontaminated groundwater by dilution mechanisms according to the concentration gradients in the fluid. The dispersivity has a longitudinal component (in the flow direction), a vertical component, and a horizontal, transverse component. The longitudinal dispersivity is important only for those concentrations at the front of leachate plumes. The flow of leachate may differ physically from that of groundwater with respect to the following three aspects:

<u>*Local water table gradients*</u> According to Christensen *et al.* (1994), locations below and around landfills are likely different from the general gradients of areas surrounding landfills. As well, local mounding effects are enhanced by the lateral spreading of the leachate plume and the downward directed hydraulic gradients in the groundwater zone beneath the landfill. The latter can cause an unexpected spreading pattern despite homogeneous aquifer conditions. The enhanced lateral spreading of the plume may increase the volume of contaminated groundwater and its spatial extent, but it also increases the dilution of contaminants.

The viscosity of the leachate may differ from that of the groundwater. Christensen *et al.* (2001) showed a higher viscosity leads to lower flow velocities, which can influence the dilution of the leachate plume.

The density of the leachate is a function of the temperature and the concentration of dissolved solids. Density differences may significantly affect the vertical positioning of the plume just below the landfill. As Christensen *et al.* (2001) showed in their study, field observations on the downward movement of the plume are often difficult to separate from the effect of local water table mounds. Density effects are considered the major cause of vertical leachate spreading in aquifers.

2.3 LEACHATE COMPOSITIONS

Leachate from Municipal Solid Waste (MSW) landfills varies over time depending on the physical, chemical and biological activity taking place in the landfill. Types of waste play an important role as leachate constituents. Typically, landfill leachate contains high values of total dissolved solids and chemical oxygen demand (COD), and slightly to moderately low pH levels. As Kulikowska and Klimiuk (2006) investigated in their study, the variation of leachate compositions and the quantity of pollutants from waste are often attributed to the volume of water which infiltrates the landfill, and directly related to the natural processes occurring inside the landfill. A study conducted by Kjeldsen *et al.* (2002) determined that compositions of MSW landfill leachate can be categorized into four groups:

1. Dissolved organic matter, quantified as COD or Total Organic Carbon (TOC), volatile fatty acids that accumulate during the acid phase of the waste stabilization, as explained by Christensen and Kjeldsen, (1989) and more refractory compounds such as fulvic-like and humic-like compounds.
2. Inorganic macrocomponents: Boron (B), calcium (Ca2+), magnesium (Mg2+), sodium (Na+), potassium (K+), ammonium (NH4+), iron (Fe2+), manganese (Mn2+), chloride (Cl-), sulphate ($SO_4$2-) and hydrogen carbonate (HCO3-).
3. Heavy metals: cadmium (Cd2+), chromium (Cr3+), copper (Cu2+), lead (Pb2+), nickel (Ni2+) and zinc (Zn2+).

4. Xenobiotic organic compounds (XOCs) originating from household or industrial chemicals and present in relatively low concentrations (usually less than 1 mg/l of individual compounds). These compounds include a variety of aromatic hydrocarbons, phenols, chlorinated aliphatics, pesticides and plastizers.

The study also found that other compounds may be found in leachate from landfills, including borate, sulfide, arsenate, selenate, barium, lithium, mercury and cobalt in very low concentration.

2.4 LEACHATE QUALITY

Leachate quality varies throughout the operational life of a landfill, and long after it has closed. A study conducted by Kulikowska and Klimiuk (2006) showed biochemical changes and physicochemical processes, including dissolution, precipitation, adsorption, dilution, volatization and others, influence leachate quality. Based on several studies (Reinhart and Grosh (1998) and Kulikowska and Klimiuk (2006)), it can be concluded that leachate quality varies according to the following phases:

Phase 1: Aerobic decomposition starts sooner due to oxygen that was trapped in the waste when it was placed in the landfill. Bacteria (aerobic) consume the oxygen and break down the complex compounds in the waste. This phase continues for days or months until the available oxygen is depleted. Temperatures get higher during the phase, and can become much higher depending on the moisture content. The primary product of this phase is carbon dioxide. Nitrogen is also found in higher concentrations at the beginning of the phase, but decreases over time throughout the phases.

Phase 2: This highly acidic phase of decomposition starts just after the depletion of oxygen in Phase 1. In this phase, anaerobic bacteria convert compounds produced in the previous phase into acetic, lactic and formic acid. Ethanol and methanol alcohol are also produced. The acids mix with the landfill moisture and act to dissolve nutrients. The phase includes nitrogen, phosphorus and gaseous by-products of carbon dioxide and hydrogen.

Phase 3: This phase is known as methanogenesis, and is a neutral environment created in the landfill due to the anaerobic bacteria consuming organic acids and producing acetate. A symbiotic relationship develops between acid-producing bacteria and methanogenic bacteria. The acid-producing bacteria produce compounds beneficial to the methanogenic bacteria, and the methanogenic bacteria consume acetate and carbon dioxide and increase toxicity for the anaerobic bacteria.

Phase 4: The amounts of methane and carbon dioxide are relatively constant in this phase. The gas production comprises 45 to 60 percent methane by volume, 40 to 60 percent carbon dioxide and 2 to 9 percent other gases such as ammonia and sulphides. Many studies showed that this gas production can continue for up to 20 years in ideal conditions.

2.4.1 FACTORS AFFECTING LEACHATE GENERATION

Basically, the leachate generation depends on the rate of decomposition, which is governed by the moisture content, temperature, landfill cover, precipitation, solid waste types and landfill design. Many studies (Reinhart and Grosh (1998) and Wang *et al.* (2004)) found that factors affecting the composition of landfill leachate include the following:

- ◊ landfill material;
- ◊ landfill conditions;
- ◊ water quality (precipitation, pH, temperature, etc.); and,
- ◊ soil characteristics under the landfill (hydrogeological behaviours).

Leachate quality is highly variable. A study conducted by Reinhart and Grosh (1998) determined that factors affecting leachate quality are based on the following key points:

<u>*Waste Composition:*</u> Municipal waste varies greatly in composition and characteristics, which means that the type of waste dictates the quality of the leachate.

Depth of Waste: Under similar conditions of precipitation and percolation, greater concentrations of constituents are found in leachate from deeper landfills.

Age of Landfill: Leachate quality is greatly influenced by the length of time waste has been in the landfill. Variations in leachate quality with respect to age can be expected throughout most of a landfill's life, since organic matter will continue to undergo stabilization.

Ambient Temperature: The temperature at a landfill site has an influence on leachate quality. Temperature affects both bacterial growth and chemical reactions, and mediates many different phenomena that occur in landfill environments.

Moisture Availability: Water is the most significant factor influencing waste stabilization and leachate quality. As soon as waste is dumped in a landfill, it mixes with precipitation and maintains the available moisture in the waste. The addition of moisture also has a stimulating effect on methanogenesis.

Available Oxygen: When waste is put in a landfill it traps oxygen and triggers aerobic decomposition. Therefore, the quantity of free oxygen determines whether the state of landfill decomposition is anaerobic or aerobic. Aerobic decomposition begins as soon as the waste is dumped, when oxygen is most available.

Processed Waste: The shredding or baling of waste can have a great influence on leachate characteristics. Leachate from shredded waste is less contaminated than that from unshredded waste, so landfills of processed waste are less harmful to the environment.

2.5 BORON IN LEACHATE

B is an important element in leachate for analysis, particularly for the TRL site, as it is free from reactions of non-landfill chemicals. It is one of the most troublesome trace elements, and has strong influence on soil management. A study conducted by Gupta (1993) concluded that B is normally present in the soil as a non-ionized form over the pH value suitable for plant growth. B is non-toxic to animals but toxic to plants. Sartaj. and Fernandes. (2003) indicate that B appears in three forms in soils: inside silicate minerals, adsorbed on clay minerals and aluminium and iron hydroxide, and inorganic matter.

According to Sartaj and Fernandes (2003), the presence of organic matter in soil can contribute to B adsorption. Goldberg (1997) studied factors affecting B adsorption in soils, including pH, soil texture, moisture, temperature, organic matter, wetting and drying, soil solution composition and clay mineralogy. A study conducted by Sartaj and Fernandes (2003) found that pH (7-7.5) has a strong effect on B adsorption. They also found that temperature has a negative effect on B adsorption; as the temperature increases B adsorption decreases.

In nature, B behaves as a conservative element. In leachate B does not react or is not influenced greatly by geological materials or leachate chemicals. Compared to other elements, it does not show ionic exchange with geological materials therefore it is more mobile. and in delineating plume extent, B can be traced up to the outer edge of the contamination plume.

2.6 SPATIAL AND TEMPORAL ANALYSIS

Spatial features include location (coordinates). Temporal features are attributes of data over time. Spatial and temporal analysis is when the data being analysed was collected over time as well as space. In spatial analysis a space-based concentration is studied, and in temporal analysis time-based fluctuation is determined.

According to Abdul-Rahman and Pilouk (2008), spatial changes refer to variations across space at a given time, or over a period in which changes are compared at the same sites, or different sites of similar age. They also determined that spatial analysis consists of three basic parameters: attribute, duration, and continuity of the event or process. Temporal information is associated either with the time variation of individual layers, or of spatial objects. Temporal changes occur at different points or periods in time, and they are recognized by changes in spatial properties and/or locations from one time to another.

In spatial-temporal analysis, comparisons are made between two or more sets of spatial and temporal objects (as in this study), to show how the attributes, temporal properties, and spatial characteristics of a process change over time.

CHAPTER 3

MATERIALS AND METHODS

3.1 TECHNICAL BASIS

To perform the spatial-temporal analysis of leachate characteristics, the following tools are employed at various steps:

- ◊ ArcGIS 9.1/9.2 is used to develop spatial and temporal changes of leachate characteristics, geostatistics and mapping requirements;
- ◊ SPSS is used for statistical analysis and graphical presentation of attribute data;
- ◊ MultiSpecW32 coupling with ArcGIS 9.2 is used for change detection in post classification method;
- ◊ GetData is used to convert *.pdf file format data into Excel or Word format;
- ◊ MS Access is used to create a database architectures and for pre-processing of attribute data; and,
- ◊ Excel is used for attribute data and to pre-process the collected data.

3.2 METHODOLOGY

The analysis is done in three steps: a multivariate exploratory data analysis to identify the principal components and variables associated with B, data processing to prepare maps, and application of GIS based change detection techniques to identify spatial and temporal changes. The following processes are performed to complete these steps.

3.2.1 FLOWCHART OF METHODOLOGY

Two main types of data are used in this study, as shown in Figure 3.1. GIS layers and well coordinates are transformed to a georeferenced coordinate system. Attribute data as tabular values of leachate concentration are processed by statistical analysis. An exploratory data analysis is performed to identify the data quality before using it in the analysis. Factor analysis, including Principal Component Analysis (PCA), is applied to extract the variables associated with B. The PCA technique identifies the fewest number of variables among the many variables which are associated with B. These extracted variables and GIS layers are then processed in ArcGIS. Different spatial data layers are created for different change detection methods. The following flowchart shows the overall methodology of the study:

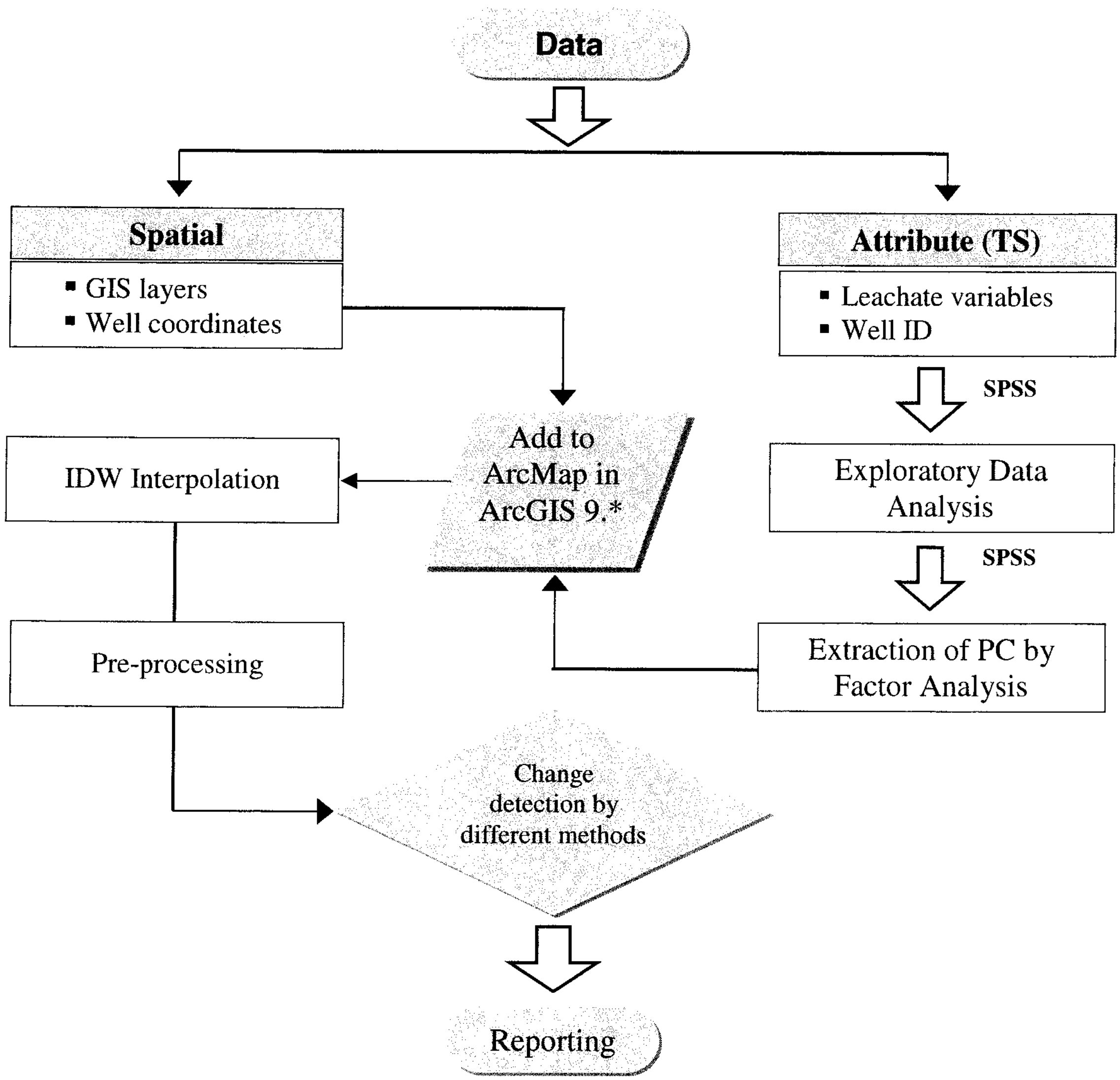


Figure 3.1 Flowchart of study methodology

3.2.2 *DATA COLLECTION*

Quarterly data are collected from Dillon Consulting Limited, Golder Associates, University of Ottawa archives and Public Library paper reports. The primary data are collected from approximately 130 wells in the landfill area by the sole source, the City of

Ottawa. Golder Associates and Dillon Consulting Limited had been using the data in annual monitoring and operation of the Nepean and TRL for many years, and most of the available information comes from their yearly monitoring and operating reports. The time series data from 1992 to 2005 and geospatial (well location) data are collected from the reports, which are archived in the Central Public Library. The collected data are in several formats, and have been converted into the required format of the study. GIS layers of the City of Ottawa and local scale GIS layers are collected from University of Ottawa Archives, and Monitoring and Operating Reports (2004) prepared by Dillon Consulting Limited. Shapefiles of different layers have been produced from paper reports, and then transformed into known coordinate systems which have congruity with other layers. A comprehensive statistical analysis is applied to observe the data quality before it is used in the study.

3.2.3 DATA PROCESSING AND CREATING DATABASE

The collected raw data has been cross-checked at the start of the study. The GIS data layers are transformed into a common coordinate system. The local scale GIS data (e.g. landfill boundary, local roads, etc.) have been retrieved from paper reports by on-screen digitizing and conversion into a common coordinate system. Finally, they are stored as layers in shapefile format.

The attribute data are exported into ArcGIS software to create shapefile for data points. Ultimately, all the attribute and geospatial data are transformed into shapefile formats similar to those which are predominantly being used for spatial-temporal analysis.

A database structure is created with Access database tools containing all the prepared data used for the analysis. A personal geodatabase has also been created in ArcGIS to store and analyse the data.

3.2.4 EXPLORATORY DATA ANALYSIS

3.2.4.1 Basic Statistics and Outlier Detection

Primary parameters, the mean (or average) and the standard deviation (SD) are determined in order to detect the most obvious outliers. As outliers can cause misleading statistical results, outlier detection is important before post processing data for analysis. In this study statistical analyses are applied to determine the presence of outliers. To identify these outliers, normality tests, quantile-quantile (Q-Q) plot and Dixon-Thompson tests are applied, as detailed in Chapter 4.

3.2.4.2 Correlation

Correlation commonly produces useful statistics which show how strongly variables in a group are related. The type of correlation test used depends on the nature of the data. If data are normally distributed the Pearson Correlation test is conducted, otherwise Spearman Correlation is applied. The KMO statistics vary between 0 and 1. A value of 0 means the sum of partial correlation is large relative to the sum of correlations, indicating diffusion in the pattern of correlation, and demonstrating that factor analysis is likely to be inappropriate. A value close to 1 indicates that patterns of the correlations are relatively compact, and factor analysis should yield distinct and reliable results. Kaiser (1974) recommends regarding values greater than 0.5 as acceptable (values below this suggest either collecting more data or re-evaluating which variables to include). Furthermore, values from 0.5 to 0.7 are mediocre, values from 0.7 to 0.8 are good, values from 0.8 to 0.9 are excellent and values above 0.9 are exceptional.

3.2.4.3 Factor Analysis

The objective of the correlation-matrix with seventeen variables is to find the correlation of variables with B concentration, as described in Chapter 4. Correlation determines the association between variables as bi-relational, while Factor Analysis is a multivariate data analysis which determines the correlation as multi-relational. Inspection of the correlation matrix shows that the correlations are substantial, indicating a considerable presence of general factors. Thus, factor analysis may provide optimal results in this case.

One type of output from factor analysis is a matrix of factor loadings. Factor loading is the degree to which every variable correlates with a factor. The meanings of the rotated factors are inferred from the variables significantly loaded on their factors. A decision must be made regarding what constitutes significant loading. Lawley and Maxwell (1971) investigated a frequently used rule that an absolute value of factor loading greater than 0.3 is significant, greater than 0.4 is more important and greater than 0.5 is very important. Thus, the rotated factor loadings greater than 0.50 are highlighted to help interpret the meanings of the factors.

A study conducted by Johnson & Wichern (1982) suggests that principal-component analysis provides a method of processing large numbers of samples that have been analyzed for several chemical constituents. The study determines the total variance of a data set among several influential variables or the principal components. Principal-component scores are calculated for each sample and constituent within a variance of 1. The constituents and samples are related to the principal components by the value of their scores. The principal component scores are analogous to the correlation between the constituent or sample and the principal component.

3.2.4.4 Scatter Plot

A Scatter Plot displays a set of collection of points, each having the value of one variable determine its position on the horizontal axis, and the value of the other variable determine its position on the vertical axis. It is a visual indicator of whether correlation exists among data points, and it suggests how the variation of a specific parameter relates to other parameters. One of the most powerful aspects of a scatter plot is that it has ability to show nonlinear relationships between the variables.

3.2.5 SPATIAL-TEMPORAL ANALYSIS

3.2.5.1 Create Wells Shapefile

The readable dBASE file of xy coordinates is added to an ArcMap window to create a new shapefile for well locations. This dBASE file includes leachate compositions, as well as X and Y coordinates which have been used in further approaches of spatial analysis.

3.2.5.2 Create Buffering

A buffer represents the area within a specified distance from an object in space. For such a study, Ontario guidelines indicate a minimum 100 m radius of buffering distance—which has been followed (see Appendix A)—for data processing or creating raster layers.

Based on well distribution, a 350 meter radius of buffering from well locations is applied as an input feature to observe the movement tendency within the area.

3.2.5.3 Interpolation

Maps (rasters) are created by the interpolation method. A raster represents spatial data that defines space as an array of cells of equal size that are arranged in rows and columns. Each cell contains an attribute value.

The maps are created for concentrations of variables. The IDW method is applied to create the interpolated maps. In the interpolation process, buffering files are used as boundary files to limit the extent of the interpolated map. It has several input parameters, which are shown in Appendix A.

3.2.5.4 Reclassification

All the interpolated surfaces have been reclassified to the same scale to visualize the spatial-temporal changes of variables. A common scale reclassification is required for a visual change analysis.

3.2.5.5 Quantile-Quantile (Q-Q) Plot

A Q-Q plot is performed to show the quantiles of two distributions against each other. For an identical distribution the plot shows a straight line. The plot allows comparing the quantiles of the B concentration with the quantiles of a standard normal distribution. It can also show statistical non-normality of distribution. In a Q-Q plot, any extreme event in data series displays further away from the continuous data positions. Q-Q plots are applied to check the statistical distribution and number of existing statistical populations of the variables.

3.2.5.6 Change Detection

Suitable change detection techniques are required for this study in order to understand the variability of successive maps (rasters) of each time interval. Four common and useful methods: post-classification visualization (reclassification to a common scale), standard deviation, principal component analysis, and unsupervised classification (clustering) methods are applied.

3.2.6 RESULT SYNTHESIS

The study discusses different change detection techniques. A synthesis of change detection approaches (described in Chapter 6) is applied to determine the suitability of the applied methods for this study, and for other similar studies.

CHAPTER 4

EXPLORATORY DATA ANALYSIS

4.1 INTRODUCTION

Exploratory Data Analysis (EDA) is a way to analyse data, and it is used for visualizing and studying data to uncover statistical irregularities that might not be apparent. Applying EDA, summarization, and displays of data makes it more comprehensible to users. According to Haining and Wise (1997), EDA identifies data properties for the purposes of pattern detection in data, and the applicability of the data in modeling or analysis.

The data which is to be applied to spatial-temporal analysis in this section is first studied thoroughly. Four aspects of the data are examined in turn: the basic statistics, the variability, the correlation structure and the multivariate correlations. After log-transformation of data, the data shows better characteristics in statistical analysis. Factor analysis is a multivariate data analysis technique. According to Bakac and Kumru (2000), multivariate statistical methods are used to quantify relationships between more than two variables under simultaneous consideration of their interactions. It has been applied to find the multivariate relation among leachate variables closely correlated to B. Drton *et al.* (2006) indicate that factor analysis reduces the dimensionality of data, and many possible observed variables are summarized by fewer factors.

Detailed cross-checking of the data was done at the outset. The optimal visualization and error detection technique is applied to identify the extreme events, and the decision to apply data correction is taken after studying the physical characteristics of the landfill site. In addition, data showing zero and missing values in the database are corrected.

4.2 EXPLORATORY STATISTICS

4.2.1 SAMPLE MOMENTS

Moments are useful descriptors of data. For example, the mean, which is a moment, is an important characteristic of a set of observations on a random variable, such as the concentration of leachate pollutants. According to Davis and McCuen (2005), in practice the following moments are generally effective for statistical analysis.

- ◊ The mean: the first moment of values.
- ◊ The standard deviation: the square root of second moment of values measured by variance.
- ◊ The variance: the second moment of values measured about the mean.

In descriptive statistics of leachate composition, variance, mean and standard deviation are calculated. The following table shows the sample moment statistics of the leachate data used in this study.

Table 4.1 Descriptive statistics of leachate composition (log-transformed)

Descriptive Statistics

	Minimum	Maximum	Mean	Std. Deviation	Variance	CV
ALK	4.01	9.80	5.5171	.74868	.561	.003
B	-5.30	1.90	-3.2605	1.41322	1.997	-.012
Br	-4.61	4.99	-1.8724	1.29492	1.677	-.017
Ca	1.14	8.77	4.4063	.75038	.563	.004
Cl	-.51	7.98	3.3855	1.52080	2.313	.010
COND	2.84	10.55	6.0196	1.13126	1.280	.006
DOC	-.69	9.83	1.1348	1.62617	2.644	.033
Fe	-5.12	6.31	-1.5477	2.31582	5.363	-.037
CaCO3	1.95	9.95	5.7425	.74565	.556	.003
K	-2.30	5.68	.9112	.83084	.690	.020
Mg	.83	7.16	3.2394	.75735	.574	.005
Na	-2.41	7.75	2.7905	1.25933	1.586	.010
NH3	-4.61	6.08	-2.7522	1.65798	2.749	-.012
PH	1.77	2.16	2.0400	.05533	.003	.001
SO4	-.56	6.60	3.1313	1.08396	1.175	.008
TSS	-.06	13.05	4.9563	2.43643	5.936	.014
TKN	-3.00	6.38	-1.2705	1.48593	2.208	-.025

Data characteristics can be found by comparing the mean and standard deviation values. A high standard deviation indicates scattered, less interrelated data spread out over a large range of values. The CV values represent the dispersion variation of the data. In Table 4.1, the mean and standard deviations of data are checked for possible outliers.

Table 4.1 displays the descriptive statistics of leachate compositions. The values are measured to check for any extreme events in the data series. The variables showing a poor measure of central tendency are prone to have extreme events which are eventually found. The dubious mean values indicate it is necessary to check the extreme events. The unexpected shift of mean values is considered to explain the possible reasons. Standard deviation is a useful descriptor of the dispersion or spread of either a sample of data or a distribution function. Chlorine, DOC, Fe, NH_3 and TSS in leachate are found to have higher variability compared to other variables.

Coefficient of Variation (CV)

CV is a statistical measure of the dispersion of data in a data series around the mean. It is calculated as the ratio of standard deviation to the expected return or mean. It is a useful statistic for comparing the degree of variation from one data series to another, even if the means are very different. It is also useful because the standard deviation of data may not always be understood in the context of the mean. The CV is a dimensionless number, so when comparing between data sets with different units or highly different means, it can be used for comparison instead of the standard deviation. Generally, $CV<1$ is considered as a low variance and $CV>1$ is considered as a high variance of data.

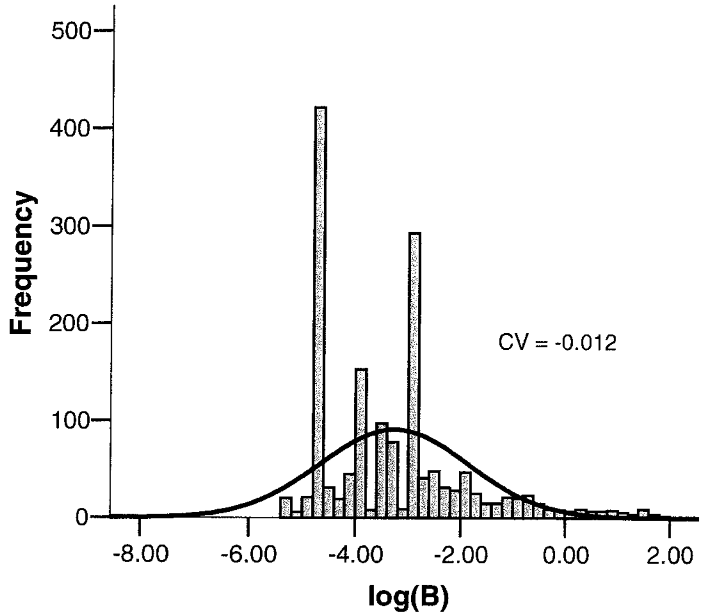


Figure 4.1 Histogram of log(B) showing CV and normal curve

Figure 4.1 shows a histogram with normal curve of log(B). It also shows the CV value of -0.012.

4.2.2 OUTLIERS DETECTION

Outliers can be considered as any extreme event that is much larger or much smaller than the rest of the sample values. Some data samples contain extreme events, which can create problems in the data analysis. For example, an extremely large value can cause the sample mean and standard deviation to be much larger than the actual values. An extreme value can adversely influence the sample value in many ways (e.g. correlation, regression and normal distribution) so that it may not reflect the true relationship among the variables. Therefore, it is essential to check for outliers.

Comprehensive crosschecking is done in this study to identify any visual extreme event present in the data. In data series, zero (0) values recorded in quarterly reported values are corrected by the mean of the other reported quarterly values after studying Operation and Monitoring report (2004) on Trail and Nepean Landfills. The dubious values are corrected by the mean value of the nearest values.

A Dixon-Thompson test is conducted to detect outliers in B concentration of the data set, after cross-checking and preliminary (zero value) correction. To perform the

Dixon-Thompson test, the data are ranked from smallest to largest values, with the smallest value denoted as X_i (i=1) and the largest value denoted as X_n. The subscripts indicate the rank of the value from smallest to largest. The test statistic R (low/high) and critical value R_c depend on the sample size. Table 4.2 is derived from the outcome of following equations.

For low outliers test statistics:

$$R_{Lo} = \frac{\overline{X} - X_1}{S_x}$$

For high outliers test statistics:

$$R_{Hi} = \frac{X_n - \overline{X}}{S_x}$$

where, $\overline{X}$ indicates the mean of the data series, X_1 indicates data at lowest rank and X_n indicates data at highest rank, R_{Lo} indicates lower outlier statistics, R_{Hi} indicates higher outlier statistics and S_x indicates the standard deviation of series. In this test, the critical value (R_c) is calculated as:

$$R_c = 2.2795 + 0.025012n - 0.00018427n^2 + 4.61106 \times 10^{-7} n^3$$

Table 4.2 Statistics of outlier detection for B

Year	NoObs	Mean	Std	X1	Xn	RLo	RHi	Rc	Outliers
2005	91	0.149	0.435	0.01	2.91	0.319	6.345	3.377	Yes
2004	124	0.189	0.671	0.01	4.93	0.267	7.067	3.427	Yes
2003	107	0.186	0.620	0.01	4.64	0.284	7.188	3.411	Yes
2002	131	0.242	0.876	0.02	6.06	0.254	6.643	3.430	Yes
2001	107	0.203	0.627	0.01	4.91	0.308	7.509	3.411	Yes
2000	102	0.194	0.573	0.01	4.4	0.322	7.340	3.403	Yes
1999	126	0.169	0.470	0.01	2.63	0.339	5.231	3.428	Yes
1998	121	0.063	0.191	0.01	1.74	0.276	8.789	3.425	Yes
1997	87	0.129	0.602	0.01	5.27	0.198	8.534	3.364	Yes
1996	112	0.101	0.295	0.01	2.52	0.309	8.201	3.417	Yes
1995	121	0.309	1.036	0.01	6.71	0.288	6.181	3.425	Yes
1994	166	0.198	0.670	0.005	4	0.289	5.676	3.463	Yes
1993	166	0.151	0.404	0.005	2.8	0.362	6.561	3.463	Yes
1992	91	0.132	0.248	0.02	1.73	0.452	6.450	3.377	Yes

Table 4.3 Numbers of detected outliers

Year	NoObs	Mean	Std	No. of Outlers	Discarded Outliers
2005	148	0.270	0.499	2	No
2004	172	0.381	0.734	3	No
2003	137	0.194	0.579	2	No
2002	168	0.261	0.816	4	No
2001	139	0.240	0.623	2	No
2000	150	0.268	0.745	1	No
1999	190	0.329	0.792	2	No
1998	170	0.342	0.900	1	No
1997	142	0.540	1.205	1	No
1996	169	0.301	1.041	1	No
1995	268	0.695	1.403	1	No
1994	288	0.398	1.020	2	No
1993	247	0.311	1.252	1	No
1992	140	0.305	0.780	1	No

Statistical Significance

The calculated statistics indicate that the critical values range within the lower and the higher outlier values. Therefore extreme events or outliers are present in the data series.

A comprehensive review of the data is performed throughout the data series after determining the number of outliers (Table 4.3) present for each year. It seems that even though annual data series contain outliers (Table 4.2), high values due to extreme events or any natural phenomena at the site were not discarded as outliers after comprehensive review of the Monitoring and Operating Report (2004).

The data collected from locations M27-1 and M28-1 beneath Stage 1 of the landfill was deleted, as these wells cannot be sampled due to blockage (Monitoring and Operating Report, 2004). An extreme TSS value is found in well BH14, a deep (upper/mid) aquifer. The eastern limit of the aquifer is defined by the presence of unsaturated conditions above the aquitard. These unsaturated conditions are defined by a line running approximately from monitoring well BH14, south through M50 and extending close to N-2 (Monitoring and Operating Report, 2004). Multiple gas measurement ports are located at M50 and M51, but they are damaged. Data from the M51 deep lower well show extreme values, so it is discarded from the analysis. The measured concentrations of inorganic parameters at well M51-3 continue to decrease, indicating diminishing leachate effects. Leachate influence at this location is attributed to a poor seal at a former adjacent well that was abandoned several years ago (Monitoring and Operating Report, 2004).

4.2.3 NORMAL Q-Q PLOT

The Q-Q or quantile-to-quantile plot is applied to observe the statistical distribution of the data shown in Figures 4.2 and 4.3. The Q-Q graph tests the conformity between the empirical distribution and the given theoretical distribution. This method is applied to verify the normality of the data and the number of statistical populations. If the points are aligned along the line, the data are normally distributed. The extreme points have greater variability than those in the center of the distribution. Thus, a U-shaped graph indicates that one distribution is skewed relative to another, while an S-shaped graph indicates that the distributions represent a greater influence of extreme values on another distribution (long tail).

As an example for interpretation, the Q-Q plot of B in Figure 4.2 shows that two populations of data are present in the data series. The majority of values correspond to low-medium-high values of B (population-1). Another smaller B population can be observed in the data that represents very high values of B in the data (population-2).

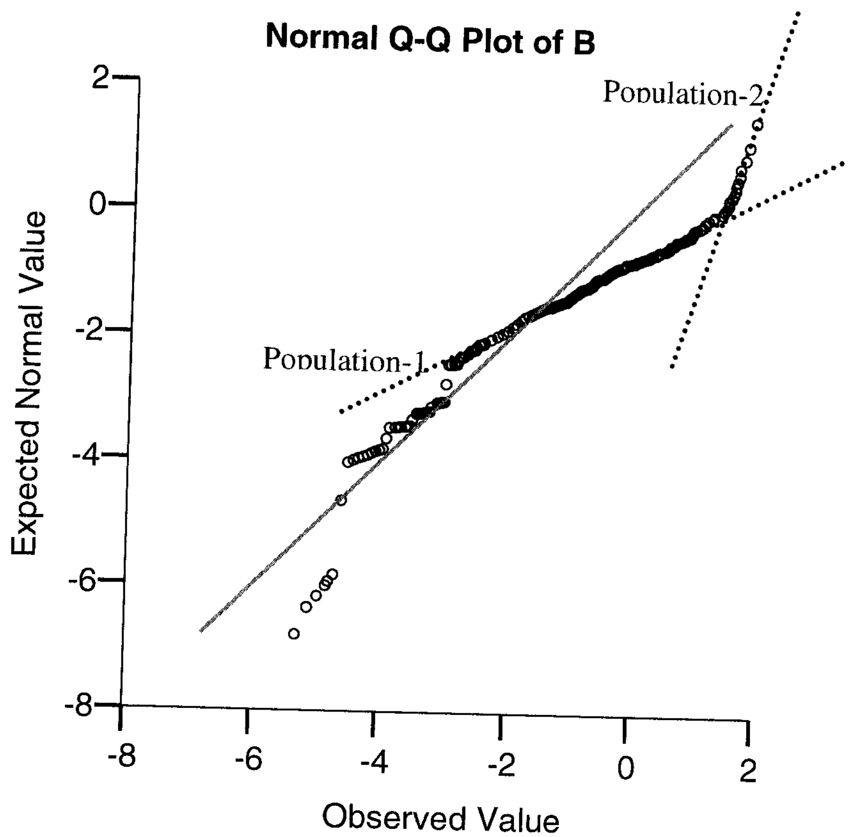
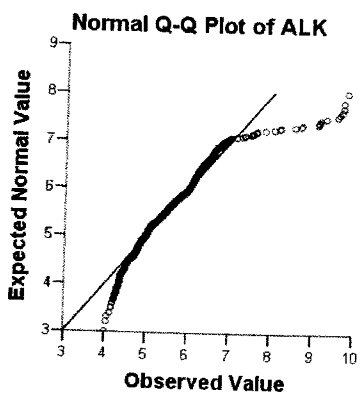
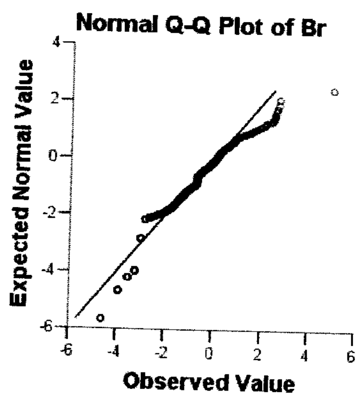
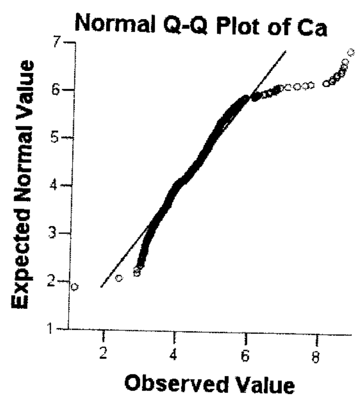


Figure 4.2 Normal Q-Q plot of B







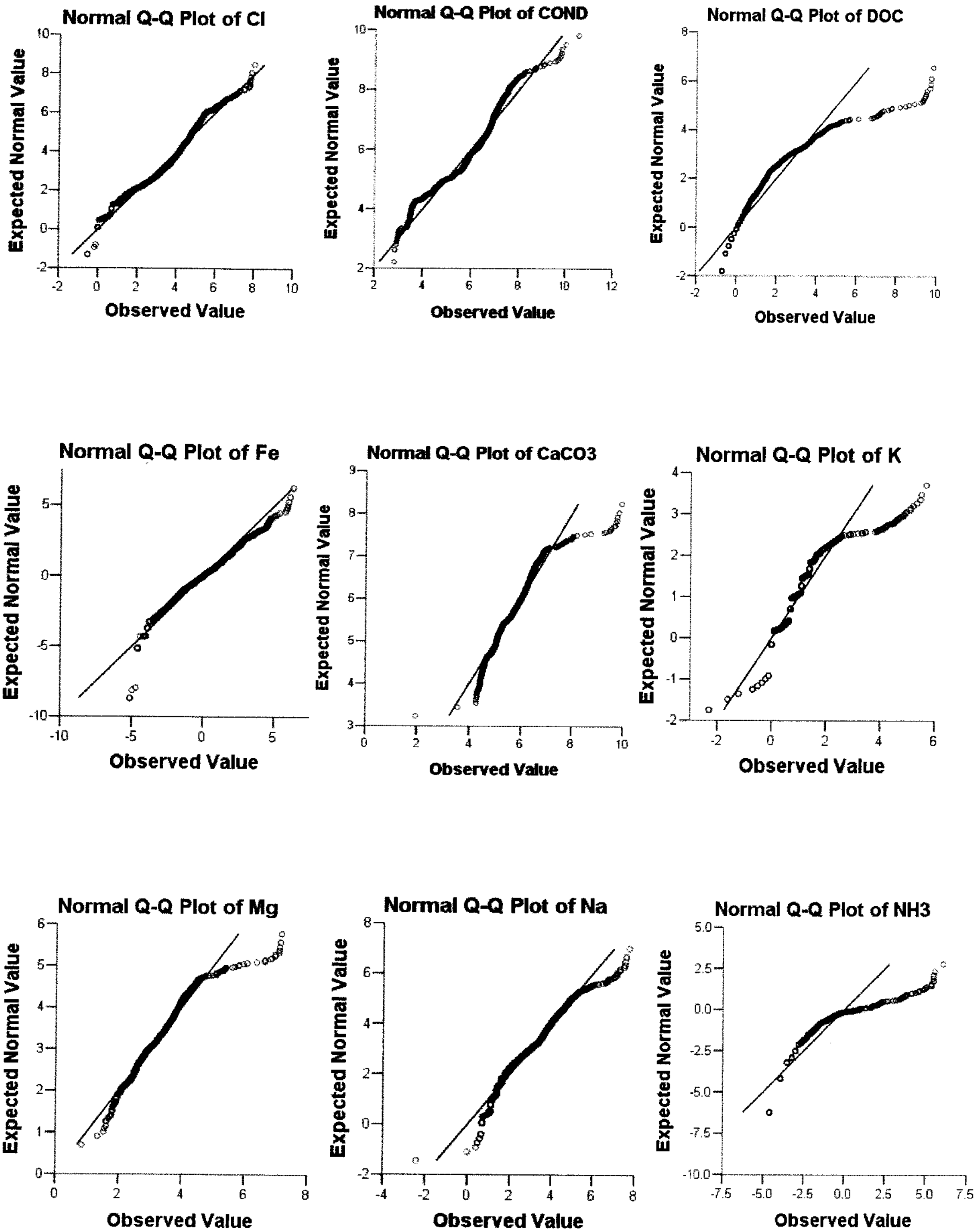


Figure 4.3 Normal Q-Q plots of variables

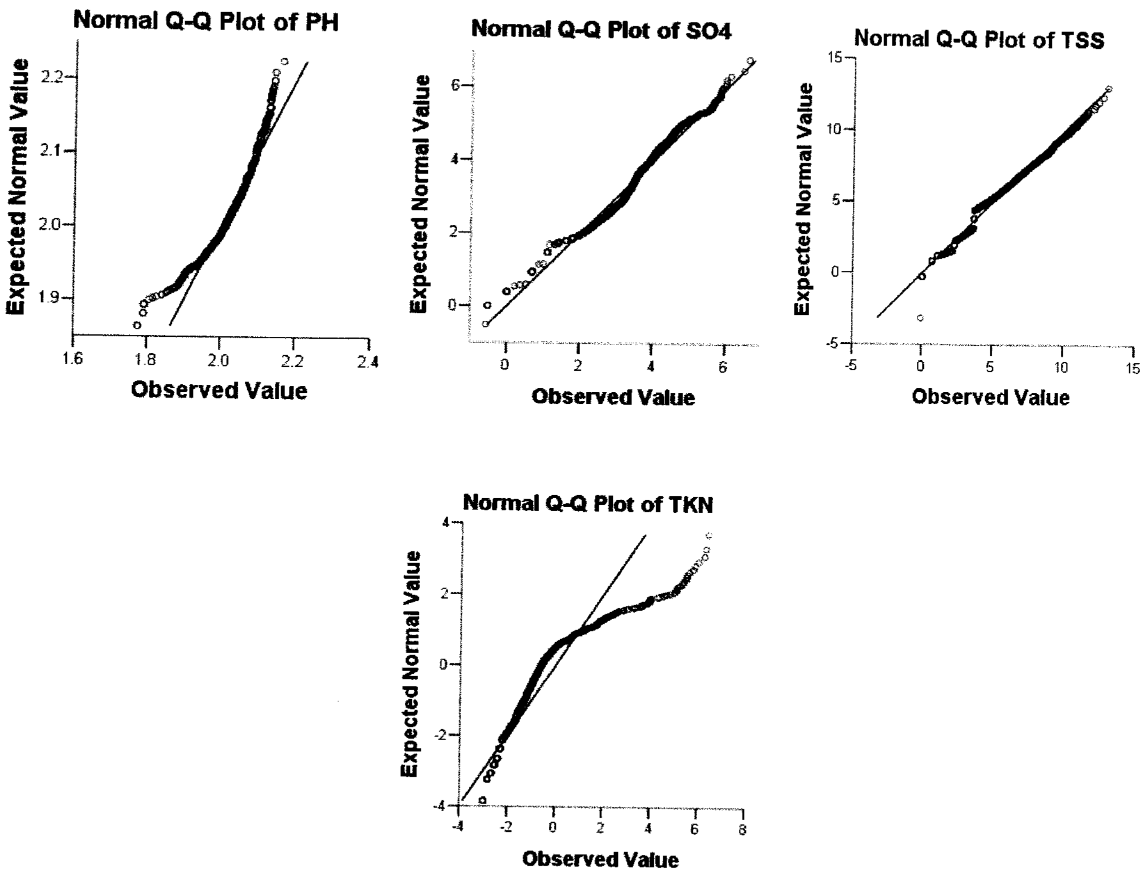


Figure 4.3 Normal Q-Q plots of variables (cont'd)

In Figure 4.3, Br, Fe, TSS, SO_4 and Cl show near to one major population. NH_3 possibly indicates one population which is not normally distributed.

4.2.4 NORMALITY TEST

The Kolmogrov-Smirnov (K-S) normality test is applied to test the normality of leachate variables, and to determine whether these variables have a normal distribution after log-transformation.

The test results in Table 4.4 indicate that the K-S values are significantly greater than the CV values in both normal and lognormal conditions, which is not consistent with normal distribution. As well, the significance level in all cases is less than 0.05 which is very low, indicating that the data are not normally distributed.

Table 4.4 Kolmogorov-Smirnov test for normality

	CV	Normal Parameters		Kolmogorov-Smirnov Z	Significance
		Mean	Std. Deviation		
ALK	.003	5.541	.742	1.800	.003
B	-.012	-3.313	1.413	5.024	.000
Br	-.017	-1.892	1.250	7.695	.000
Ca	.004	4.421	.741	2.340	.000
Cl	.010	3.421	1.511	2.726	.000
COND	.006	6.129	1.074	3.552	.000
DOC	.033	1.158	1.616	4.743	.000
Fe	-.037	-1.517	2.327	3.360	.000
CaCO3	.003	5.760	.732	2.160	.000
K	.020	.899	.836	5.787	.000
Mg	.005	3.253	.753	1.875	.002
Na	.010	2.824	1.246	2.727	.000
NH3	-.012	-2.776	1.708	7.835	.000
PH	.001	2.036	.054	2.892	.000
SO4	.008	3.096	1.110	3.781	.000
TSS	.014	5.058	2.434	3.774	.000
TKN	-.025	-1.271	1.497	4.833	.000

Another approach can be applied to determine data usability. According to Kirkman, (1996), in a normality test, 15% of the data should lay more than 1 standard deviation below the mean (i.e. below *mean – standard deviation*). Similarly only about 2% of the data should be more than 2 standard deviations above the mean (i.e., above *mean + 2*standard deviation*).

Table 4.5 Descriptive statistics of KS test for normality test after log-transformation

Variables	Mean	SD	1 SD	2 SD	below 15%	above 2%	Normal Distribution?
ALK	5.54	0.74	4.80	7.02	15.58	1.86	No
B	-3.31	1.41	-4.73	-0.49	2.98	5.46	Yes
Br	-1.89	1.25	-3.14	0.61	9.93	3.29	Yes
Ca	4.42	0.74	3.68	5.90	14.53	2.17	Yes
Cl	3.42	1.51	1.91	6.44	16.51	1.61	No
COND	6.13	1.07	5.05	8.28	23.71	0.99	No
DOC	1.16	1.62	-0.46	4.39	8.63	3.41	Yes
Fe	-1.52	2.33	-3.84	3.14	22.84	3.04	No
CaCO3	5.76	0.73	5.03	7.22	13.53	2.36	Yes
K	0.90	0.84	0.06	2.57	1.43	2.61	Yes
Mg	3.25	0.75	2.50	4.76	9.37	2.11	Yes
Na	2.82	1.25	1.58	5.32	12.66	1.92	No
NH3	-2.78	1.71	-4.48	0.64	3.35	4.90	Yes
pH	2.04	0.05	1.98	2.14	9.12	0.06	No
SO4	3.10	1.11	1.99	5.32	12.04	1.80	No
TSS	5.06	2.43	2.62	9.93	8.63	2.05	Yes
TKN	-1.27	1.50	-2.77	1.72	4.03	4.10	Yes

Table 4.5 shows that some variables show normal distribution with log-transformation of original data, which means that this data could be used in spatial and temporal analysis.

4.2.5 CORRELATION

Correlation analysis is a preliminary descriptive technique used to estimate the degree of association among associated variables. The purpose of the correlation analysis is to measure the level of association between B and other variables. Such associations are likely to lead to consideration of the causal relationship of the variables in bi-relation. Table 4.6 shows that B has the highest association with K, Na, NH_3, DOC and TKN.

The correlation output matrix of the parameters in Table 4.6 is calculated based on log-transformed of variables, and provides information regarding association of the variables.

Table 4.6 Pearson's correlation coefficients

Pearson's Correlations

	ALK	B	Br	Ca	Cl	COND	DOC	Fe	CaCO3	K	Mg	Na	NH3	PH	SO4	TSS	TKN
ALK	1																
B	.416	1															
Br	.544	.380	1														
Ca	.910	.357	.521	1													
Cl	.628	.219	.517	.725	1												
COND	.520	.168	.358	.558	.481	1											
DOC	.740	.461	.491	.703	.543	.544	1										
Fe	.601	.372	.399	.504	.378	.355	.617	1									
CaCO3	.913	.349	.536	.974	.728	.559	.716	.534	1								
K	.560	.452	.309	.545	.339	.306	.567	.478	.531	1							
Mg	.906	.348	.563	.920	.716	.559	.707	.562	.950	.523	1						
Na	.737	.418	.524	.742	.835	.497	.645	.479	.764	.426	.734	1					
NH3	.468	.433	.254	.425	.224	.268	.576	.521	.423	.709	.378	.351	1				
PH	-.714	-.328	-.395	-.703	-.462	-.379	-.558	-.515	-.685	-.466	-.622	-.540	-.430	1			
SO4	-.155	-.137	-.157	.054	.003	-.072	-.332	-.403	.018	-.096	-.044	-.098	-.203	.093	1		
TSS	.113	.032	.007	.067	.033	.009	.002	.142	.071	.120	.058	.116	.120	-.132	.002	1	
TKN	.618	.485	.396	.587	.430	.415	.756	.562	.584	.685	.544	.548	.791	-.509	-.254	.071	1

4.2.6 FACTOR ANALYSIS

Factor analysis is a statistical procedure that finds the most inter-related variables and arranges them into groups or factors, and a method of data reduction. Hudec, (1984) defines factor analysis as the most useful multivariate statistical technique. Factor analysis attempts to bring inter-correlated variables together under more general, underlying variables. More specifically, the goal of factor analysis is to reduce the dimensionality of the data, and to find the multivariate structure and association of all variables. There are many different methods that can be used to conduct factor analysis, including principal axis factor, maximum likelihood and least squares. In factor analysis, two types of rotation techniques, varimax and equimax can be applied to ensure that the factors can not be correlated with one another. A number of factors are extracted using the principal component analysis technique, which is the main objective of factor analysis.

Factor analysis typically requires a large sample size, an important factor for analysis. Factor analysis is based on the correlation matrix of the variables involved, and correlations usually need a large sample size before they can stabilize. Tabachnick and Fidell (2001) advised the following regarding sample size: 50 is very poor, 100 is poor, 200 is fair, 300 is good, 500 is very good, and 1000 or more is excellent.

The matrix of inter-correlated variables in this study is shown in Table 4.6. In correlation analysis, correlations among variables are determined by bi-relation, which is a univariate process. However, this data requires multivariate data analysis, for which factor analysis may give a better reading of data association.

In factor analysis, factors are extracted based on the principal component analysis and eigenvalues over 1. One of the outputs of factor analysis is a matrix of factor loadings. Factor loading is the degree to which every variable correlates with a factor, and factor loadings are the basis for imputing a label to different factors. The meanings of the rotated factors are inferred from the variables significantly loaded on their factors. A decision needs to be made regarding what constitutes a significant loading. Lawley and Maxwell (1971) showed in their study that the absolute value of a factor loading greater than 0.3 is usually considered significant, greater than 0.4 more important and greater than 0.50 very important. Thus, rotated factor loadings greater than 0.50, are highlighted to assist in the interpretation of the meanings of the factors. In the initial factor solution, the first factor accounts for the highest variance, the second accounts for the next highest amount of variance, and so on.

The factor loadings for the varimax orthogonal rotation show how the variables are associated with each factor. It represents the correlation between the variables and the factor. A varimax rotation attempts to maximize the squared loadings of the columns.

4.2.6.1 KMO and Bartlett's Test

The KMO statistics vary between 0 and 1. A value of 0 indicates that the sum of partial correlation is large relative to the sum of correlations, indicating diffusion in the pattern of correlation; hence, factor analysis is likely to be inappropriate. A value close to 1 indicates that patterns of the correlations are relatively compact, and so factor analysis should yield distinct and reliable factors. Kaiser (1974) recommended deeming values greater than 0.5 as acceptable (values below this can lead to either collecting more data or rethinking which variables to include). Values from 0.5 to 0.7 are mediocre, from 0.7 to 0.8 good, from 0.8 to 0.9 excellent and above 0.9 exceptional.

KMO statistics show a value of 0.881 in Table 4.7, which is an excellent value for factor analysis. As well, the Bratlett's test shows highly significant as $P < 0.001$. Therefore factor analysis can be useful in this study.

Table 4.7 KMO and Bartlett's test

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.881
Bartlett's Test of Sphericity	Approx. Chi-Square	26002.903
	df	136
	Sig.	.000

4.2.6.2 Communality Test

According to Field (2000), as communalities become lower the importance of sample size increases. Rietveld & Van Hout, (1993) determined that communalities can be seen as a continuation of factor loadings. The communality of a variable is the sum of the loadings of the variable on all extracted factors. As such, the communality of a variable represents the proportion of the variance of that variable that can be accounted for by all extracted factors. Thus, if the communality of a variable is high, the extracted factors account for a large proportion of the variable's variance. This means that the variable is reflected well via the extracted factors, and so the factor analysis is reliable. When the communalities are not very high, however, the sample size has to compensate for this.

Table 4.8 Communality test

Communalities

	ALK	B	Br	Ca	Cl	COND	DOC	Fe	CaCO3	K	Mg	Na	NH3	PH	SO4	TSS	TKN
Initial	1.00	1.00	1.000	1.000	1.000	1.000	1.000	1.00	1.000	1.0	1.00	1.00	1.000	1.00	1.000	1.000	1.000
Extraction	.863	.416	.499	.925	.735	.464	.780	.635	.945	.736	.887	.738	.823	.544	.790	.279	.792

Extraction Method: Principal Component Analysis.

The extraction of communalities in this study shows the average value of extraction is 0.70. In Table 4.8, TSS shows a much lower value (which is equal to 0.70) if there are less than 30 variables. Therefore, the statistical analysis signifies that KMO and Bratlett's tests are appropriate for extraction of factors, and factor analysis can be meaningful. As well, a lower value of extraction would require more data in the analysis.

4.2.6.3 Variance of Factors Extraction

The output of factor analysis is shown in Table 4.9. This helps to determine the number of components/factors to be retained for further analysis. The factors with eigenvalues over 1 are extracted to see which factor is more associated with B and other related variables. The first three factors are considered dominant factors, and account for 69.71% of the total variability of the data (Table 4.9).

Extraction is done in the analysis process by selecting both correlation and covariance options in factor analysis, in order to have comparative scenarios. Generally, covariance measures how much the deviation of two or more variables or processes match, and it is analogous to the variance of a single variable. Correlation measures how much one random variable depends upon the other, and how accurately the value of one variable can be predicted from the value of the other. Therefore, the process of correlation method is selected for parameters extraction.

Table 4.9 Test for total variance explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.940	52.590	52.590	8.940	52.590	52.590	6.541	38.475	38.475
2	1.770	10.412	63.002	1.770	10.412	63.002	4.019	23.639	62.113
3	1.140	6.708	69.710	1.140	6.708	69.710	1.291	7.597	69.710
4	.991	5.831	75.541						
5	.784	4.609	80.150						
6	.674	3.963	84.113						
7	.542	3.187	87.300						
8	.476	2.800	90.100						
9	.400	2.353	92.453						
10	.377	2.219	94.673						
11	.305	1.793	96.465						
12	.212	1.245	97.710						
13	.136	.800	98.511						
14	.129	.758	99.269						
15	.064	.374	99.643						
16	.059	.347	99.990						
17	.002	.010	100.000						

Extraction Method: Principal Component Analysis.

The "Total variance explained" output of the factors extraction table (Table 4.9) shows the following components in the table:

Component - The initial number of components or factors is the same as the number of variables used in the factor analysis. However, not all 17 factors are retained; in this example only the first three factors are retained (as selected).

Initial Eigenvalues - Eigenvalues are the variances of the factors. Because the factor analysis is conducted on the correlation matrix, the variables are standardized, which means that the each variable has a variance of 1, and the total variance is equal to the number of variables used in the analysis; in this case, 17.

Total - This column shows the eigenvalues. The first factor accounts for the most variance (and hence has the highest eigenvalue), and the next factor accounts for as much of the remaining variance as it can, and so on. Hence, each successive factor accounts for less variance.

% of Variance - This column shows the percentage of total variance accounted for by each factor.

Cumulative % - This column shows the cumulative percentage of variance accounted for by the current and all preceding factors. For example, the third row shows a value of 69.71, which means that the first three factors together account for 69.71% of the total variance.

Extraction Sums of Squared Loadings - The number of rows in this panel of the table correspond to the number of factors retained. In this example, three factors are selected to be retained, so there are three rows, one for each. The values in this panel are calculated in the same way as the values in the left panel, except that here the values are based on the common variance. Since common variance is always smaller than total variance, the values in this panel are always lower than the values in the left panel.

Rotation Sums of Squared Loadings - The values in this panel of the table represent the distribution of the variance after varimax rotation. Varimax rotation tries to maximize the variance of each of the factors, so the total amount of variance accounted for is redistributed over the three extracted factors.

4.2.6.4 Scree Plots

A Scree Plot is a simple line segment plot that shows the fraction of total variance in the data. It graphs the eigenvalue against the factor number. The plot may consider the side of a mountain, and 'scree' refers to the debris fallen from mountain which is lying at its base. The scree plot proposes to stop analysis at the point the mountain ends and the debris begins. In this instance that point coincides with the eigenvalue criterion (the first three factors).

The scree plot in Figure 4.4 shows some of the information in Table 4.9. From the third factor on, it is seen that the line is almost flat, meaning the each successive factor is accounting for smaller and smaller amounts of the total variance.

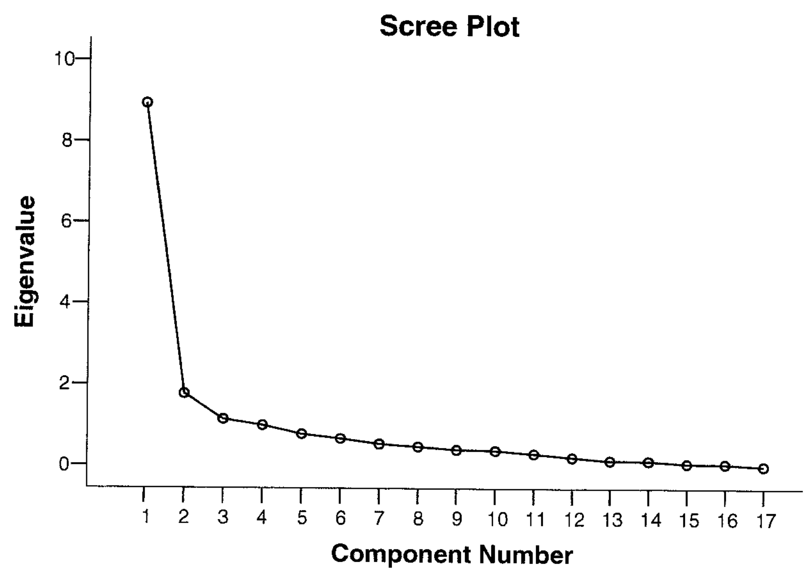


Figure 4.4 Scree plot of leachate factors

4.2.6.5 Rotated Component Matrix

Factor Matrix - This table showed the unrotated factor loadings, which are the correlations between the variable and the factor. Because these are correlations, possible values range from -1 to +1. Coefficient values of 0.3 or less are not shown in Table 4.9. This made the output easier to read by removing the lower coefficients, which have little influence. Table 4.9 is one of the most important outputs of this factor analysis. It indicates that B is dominantly associated with principal component 2. It also shows that in a mathematical multivariate space, B is associated mainly with NH_3, TKN and K. This component can be considered the "B indicator." This result is considered in the analyses in the next chapter. Component 1 most likely represents alkalinity, since it has the highest association with carbonate environment elements and compounds.

Component - The columns under this heading are the unrotated factors that are extracted. The rotated component matrix shows that the B concentration is associated with K, NH_3, and TKN, represented by component 2 of the Principal Component Analysis shown in Table 4.10. The plot in Figure 4.5 shows the variables in the rotated factor space. It demonstrates how the variables are organized in the factor space represented by first three factors.

Table 4.10 Rotated component matrix

Rotated Component Matrix[a]

	Component		
	1	2	3
CaCO3	.912		
Ca	.897		
Mg	.895		
Cl	.854		
ALK	.818	.435	
Na	.815		
COND	.659		
Br	.616		
DOC	.612	.560	
PH	-.589	-.440	
NH3		**.890**	
TKN		**.787**	
K		**.786**	
Fe		.588	
B		**.575**	
SO4			-.834
TSS			-.415

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 5 iterations.

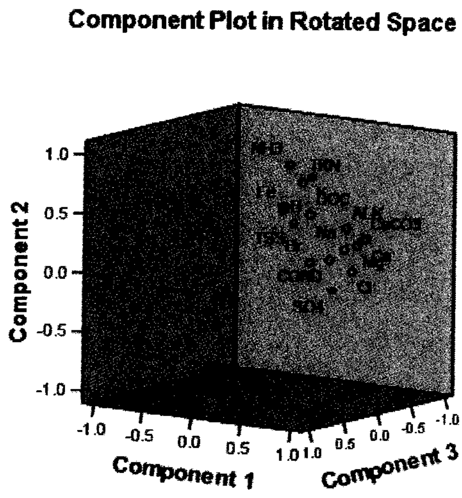
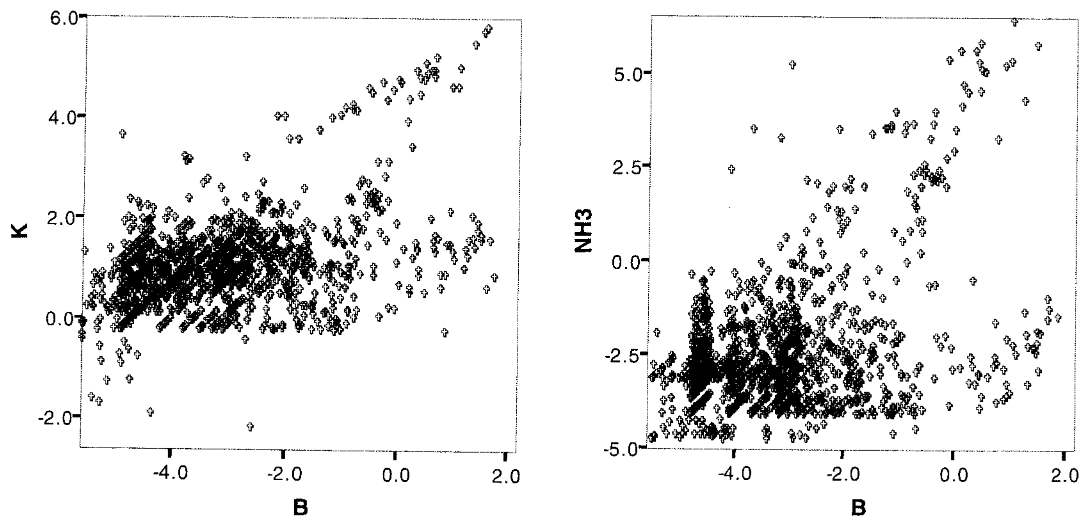


Figure 4.5 Component plot in rotated space displaying B with associated variables

4.2.7 SCATTER PLOT

Scatter plots are often an effective way to display data. Considering the extracted results in Table 4.10, there are seven variables (K, NH_3, TKN, Fe, DOC, pH and ALK) which have a correlation of more than 0.3 with principal component 2. B is associated with these variables. Variation of B concentration versus K, NH_3 and TKN is displayed in Figure 4.6. This figure also displays the positive association between B and K, NH_3 and TKN, as shown in Table 4.10.



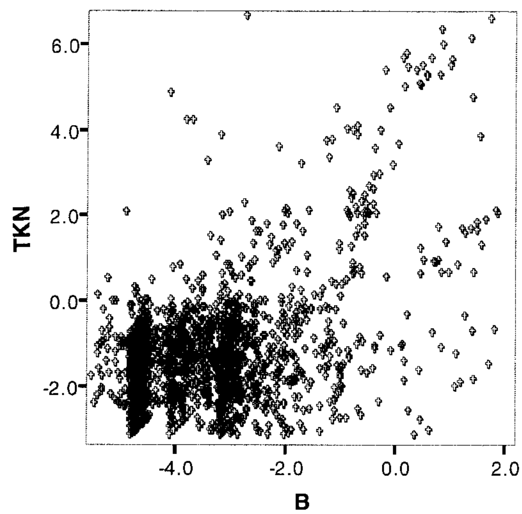


Figure 4.6 Scatter plots of extracted components associated with B

4.3 CONCLUSION

Statistical analysis often pre-supposes that the variable being analyzed is normally distributed. However, if it is not, a significant deviation could arise and might be interpreted as a misspecification of the analysis. Therefore, in such analyses it is a good idea to determine whether the normality assumption is true. In this study, the sample moments test is applied to reveal the basic characteristics of the parameters. Mean, standard deviation and variance are calculated to determine the data characteristics. The Dixon-Thompson test for outlier detection is applied to identify the extreme outliers that could represent erroneous data. For this reason, the outlier detection is calculated by the yearly data frequency. The analysis shows that few extreme events appear in each year. Then, a comprehensive review of physical or landfill site characteristics is conducted and evaluated, to come to a decision whether the extreme values should be discarded or used in the analysis. Normal Q-Q plots are also applied to review data normality. The normality of the data is not completely met, even after applying log-transformations. Eventually, after transformation the data are not perfectly normally distributed, but some indicated variables are close to it.

A correlation is calculated between all variables; Pearson's correlation method is applied. The next step of the data analysis is to extract principal components to determine which variables B is most associated with, and in a multivariate space which factor represents B and associated variables in the study. A factor analysis is applied to extract the parameters based on standardized variance or eigenvalues over 1 (one).

The output of the exploratory data analysis shows the association of B with K, NH_3, and TKN, which are all represented by principal component (factor) 2. These results are applied in both spatial and temporal analysis.

CHAPTER 5

SPATIAL AND TEMPORAL ANALYSIS

5.1 INTRODUCTION

In this chapter, the spatial and temporal variations of B and associated variables of the leachate in the groundwater at the Trail Road Landfill site are studied. This analysis applies an integrated GIS/statistical approach. Spatial and temporal changes are detected and measured using several of the common change detection techniques. These techniques are mainly applied for image processing. Images in the study are rasters created by the interpolation of numerical variables (e.g. B and associated variables). The chapter shows that these change detection methods can also be used in an environmental example related to the landfill site. Timely and accurate change detection in groundwater environments is extremely important for understanding interactions among physical phenomena, periodical changes and leachate compositions, and to support improved decision making.

Landfill leachate chemicals in groundwater are dynamic, and some change rapidly over time due to their interactions and natural events. Conversely, some of the chemicals are not as dynamic and change more slowly. Spatial distribution of these variables can be mapped by interpolation techniques for a specific time interval. The output raster (map) can be considered a snap-shot, representing how a variable is distributed in 2-dimensional space in the study area. By repeating similar interpolations for the same variable and other time intervals, new maps are created. A unique aspect of this thesis is that, instead of quantitative analysis being performed on the original measured values in tables related to the original datasets, quantitative methods of change detection are applied directly on the 2-dimentional maps representing spatial distribution of B and other related variables. The maps are the input for the applied change detection techniques.

The dimensions of the change detection region of interest (ROI) are carefully determined in the area of the monitoring wells. The geographic extent of the ROI covers the locations of all wells having records of the studied variables during the timeframe of the study (1992-2005). Results support the applicability of the change detection methods to study spatial-temporal variations in this landfill site. In addition, they define regions with differing levels of change for B and other related variables, to determine the high-risk areas and detect the effect of time on the distribution of B. The methods are also used for K, NH_3 and TKN to determine changes related to them. Factor analysis is applied to determine the variables associated with B (in Chapter 4). The annual raster data layers are created by interpolation of all these variables (Appendix A). These maps required some processing before using them in quantitative change detection methods. Post classification visualization, principal component analysis, standard deviation and unsupervised classification or clustering are applied to the maps to detect the spatial-temporal changes. These methods can be considered as univariate and multivariate statistical analysis techniques in a spatial context. A synthesis of output from change detection is examined to determine the suitability of these methods. The study requires the knowledge and technology of spatial data processing in applications of change detection. Relationships between the accuracy of change detection, and recommendations to achieve more accurate results are discussed.

5.2 CHANGE DETECTION PROCESS

This section is organized into three subsections. The first provides a general overview of the methodology, the second illustrates the process of change detection, and the third detects the changes by the synthesis of outputs.

Singh (1989) described change detection as the process of identifying differences in the state of an object or phenomenon by observing it at different times. Detecting and presenting change provides valuable information regarding the possible trend of transformations of a given scenario over time. Timely and accurate change detection of features on the earth's surface provides better understanding of the relationships and interactions between humans and natural phenomena, and can improve the management and use of resources. In general, change detection involves the application of multi-temporal datasets for quantitative analyzes, and visual assessment of the temporal effects by reviewing the maps of a study area.

The basic steps to conduct change detection by using maps (raster data layers) are: 1) specify the nature of the change detection problem, 2) conduct data analysis and create maps representing the spatial distribution of variables, 3) extract change information by applying appropriate quantitative change detection techniques and 4) accurately evaluate the change detection results. By following these steps, users can determine whether their change detection results are of value. More specifically, the quantitative and non-quantitative change detection methods in this study provide information about changed and unchanged areas, and spatial relationships within the changed areas.

Many change detection techniques available in quantitative image processing have been developed based on their use. Maps require pre-processing to prepare them for application in different change detection methods. Finally, synthesis of the change detection results provides information of temporal changes that occurred at this landfill site.

5.2.1 CONSIDERATIONS BEFORE IMPLEMENTING CHANGE DETECTION

Change detection techniques are mainly applied to image processing of remote sensing data or maps (raster data layer) derived from satellite images or in a GIS environment. The study applies these methods for this environmental application. MacLeod and Congalton (1998) described four important aspects of change detection for monitoring natural resources: detecting if changes occurred, identifying the nature of the changes, measuring the areal extent of the changes, and assessing the spatial pattern of the changes. Successfully implementing change detection analysis using raster layers requires careful consideration of the data layers, environmental characteristics and image processing methods. The temporal and spatial resolutions of maps have a significant impact on the success of a change detection technique. For this study, annual intervals of measurement seemed adequate for temporal resolution, although there is no continuity of available data for the variables of particular well locations for every year. The other positive aspect of the TRL site data is the spatial distribution pattern of samples, which are almost uniformly distributed within the study area. This enabled the use of the applied change detection methods.

Interpolation of required variables, followed by reclassifying the result maps into common scales, are among the most important pre-processing steps for this type of change detection in environmental applications. The IDW method is applied for interpolation of B and other related variables to create continuous raster layers for change detection. The interpolated rasters (maps) are then reclassified to equalize the spatial variation of concentration to a common scale, to allow comparative temporal changes to be detected by visual assessment.

In this study, the following conditions are satisfied before implementing change detection analysis: 1) the common interpolation method is applied to B and associated variables in each time interval (1992-2005), 2) a common colour symbology is applied to find comparative differences, and 3) a common scale for reclassification was applied to the maps.

Various change detection techniques can provide different types and suites of maps and outputs. Some techniques applied in this study, such as certain quantitative GIS techniques, provided changed/unchanged information. Other techniques, such as post-classification or reclassification comparison, provided a complete matrix of changed locations and their trend directions. For this study, several change detection methods are applied to determine the spatial and temporal changes over time. Finally, the results of the applied change detection methods are compared and discussed.

5.3 METHODOLOGY

The steps to conduct change detection by using maps (raster data) as inputs are detailed in the following flowchart (Figure 5.1). The steps are based on the recommendations of Suh (2008). By following these steps, users are able to decide whether their change detection results are suitable. The process of selecting appropriate raster data layers and change detection techniques according to the nature of the change detection problem is still under investigation, and this is critical for change detection studies. As indicated by the last step of Figure 5.1, in ideal conditions an accuracy assessment is required, and can complete the process. Information accuracy assessment is not done in this study due to the unavailability of new information. This step can be considered for future approaches and continuing studies at the TRL site, or for similar environmental projects.

State the Problem

- Specify change detection geographic region of interest (ROI)
- Define the classes of interest in a classification system
- Select the change detection techniques

Consideration for the Significance of Spatial and temporal Changes

Process Interpolated Raster to Extract Change Information

✓ Create raster layers by suitable methods
✓ Pre-process the interpolated raster
✓ Apply change detection using GIS algorithms
  - Post classification visual assessment
  - Principal Component Analysis method
  - Standard deviation method
  - Unsupervised classification (clustering) methods

Perform Accuracy Assessment

- Result synthesis
- Evaluating the change detection results

Figure 5.1 Flowchart of methodology

5.4 REVIEW OF SOME CHANGE DETECTION TECHNIQUES

The objective of change detection is to compare spatial and temporal representations of two points in time by controlling variability caused by differences in variables. For the sake of convenience, the change detection methods are grouped into four categories: mathematical transformation, classification, GIS approaches, and visual analysis.

Mathematical Transformation

PCA is a multivariate statistical analysis method that reduces the number of variables by applying a linear mathematical transformation, and merging the initial variables into fewer new variables. If maps of the same variable (e.g. B) related to various time intervals are entered by the PCA method, instead of various variables entered as maps, regions that behave similarly in time can be detected. With this approach, change information can be highlighted in the principal components. The method assumes that the inputs (maps of the same variable in time) are correlated with each other which is a valid assumption in this study.

Classification:

Classification for change detection can be applied in two main ways: common scaling (Post-classification comparison) and unsupervised classification.

Common scaling (Post-classification comparison)

In common scaling methods maps are created separately, then a common scale (class) is assigned for each map related to the studied time interval. The common scaling enables pixel-by-pixel comparison of the maps to record change. The major advantage of this method is the capability to provide a matrix of change information, which can indicate which classes have changed relative to each other, or have remained constant for a pair of maps.

Unsupervised classification in change detection

Maps of the same variable in time can be input with this method. This is a multivariate statistical method, in which groups or clusters of pixels with similar characteristics (in time) can be identified. Various change levels across the study area are provided as results in the form of maps.

GIS approach

GIS can be applied to create binary maps, overlay them, and visually and quantitatively integrate them by any logical or mathematical function or group of functions related to a method. Within a GIS, it is also possible to consider masks as maps. These capabilities are specific to GIS, and are required and useful for quantitatively revealing the change dynamics in each categorical or numerical variable (e.g. B and related variables).

Visual analysis

Visual analysis includes visual interpretation of multi-temporal raster layer composites, identification of changed or unchanged areas, and on-screen digitizing of the identified areas. This method can often require all a user's experience and knowledge. The shape, size and pattern of classes of maps representing a variable in time are key elements for identifying spatial and temporal change through visual interpretation. These elements are not often used in digital change detection analysis due to the difficulty of extracting the elements and taking them into account.

5.5 APPLICATION OF CHANGE DETECTION TECHNIQUES

5.5.1 POST-CLASSIFICATION VISUAL METHOD

According to Lu *et al.* (2003), post classification is one of the most common methods of change detection, and it is considered the most straightforward. It involves each map being classified independently, followed by a comparison of the corresponding pixel labels to identify areas where change occurred. Changed areas are simply those which are different from one date to a different date.

The visual analysis category includes visual interpretation of the maps featuring spatial-temporal changes of the raster layers. After creating a map representing continuous distribution of the variable, a common scale (classification) is applied for all maps representing the same variable but for different time intervals. In these reclassified maps, lower than detection limit values are not considered as the lowest class (dark blue).

The resulting shapes and patterns are key elements for the identification of spatial change through visualization, by comparing these reclassified maps. This procedure is applied in this study for B, K, NH_3 and TKN as indicated in Figures 5.2 to 5.5. A 10-class scale (classification) has been used as the common scale for maps of each of these variables in time, to enable a visual comparison of maps in each variable (Figures 5.2 to 5.5). Figure 5.2 shows that B is increasing from 1992 to 1995, decreasing in 1996, again increasing from 1997 and then eventually it is decreasing by 2004 and 2005. The associated variables of B (K, NH_3 and TKN) show almost the similar patterns. Some dissimilarities are also found in the visual comparison of the variables (i.e. NH_3,1995). It is assumed that the main reasons for existing noise in the applied dataset and results are due to errors added in sampling, chemical analysis, changing various analytical labs, data entry and also variations in leachate release and groundwater movements.

In the B maps, the plume shows two parts with high concentrations (M32 and B16 wells). The reason for this observed pattern can be two different gradients of flow towards the dewatering pond (north) and beaver pond (west).

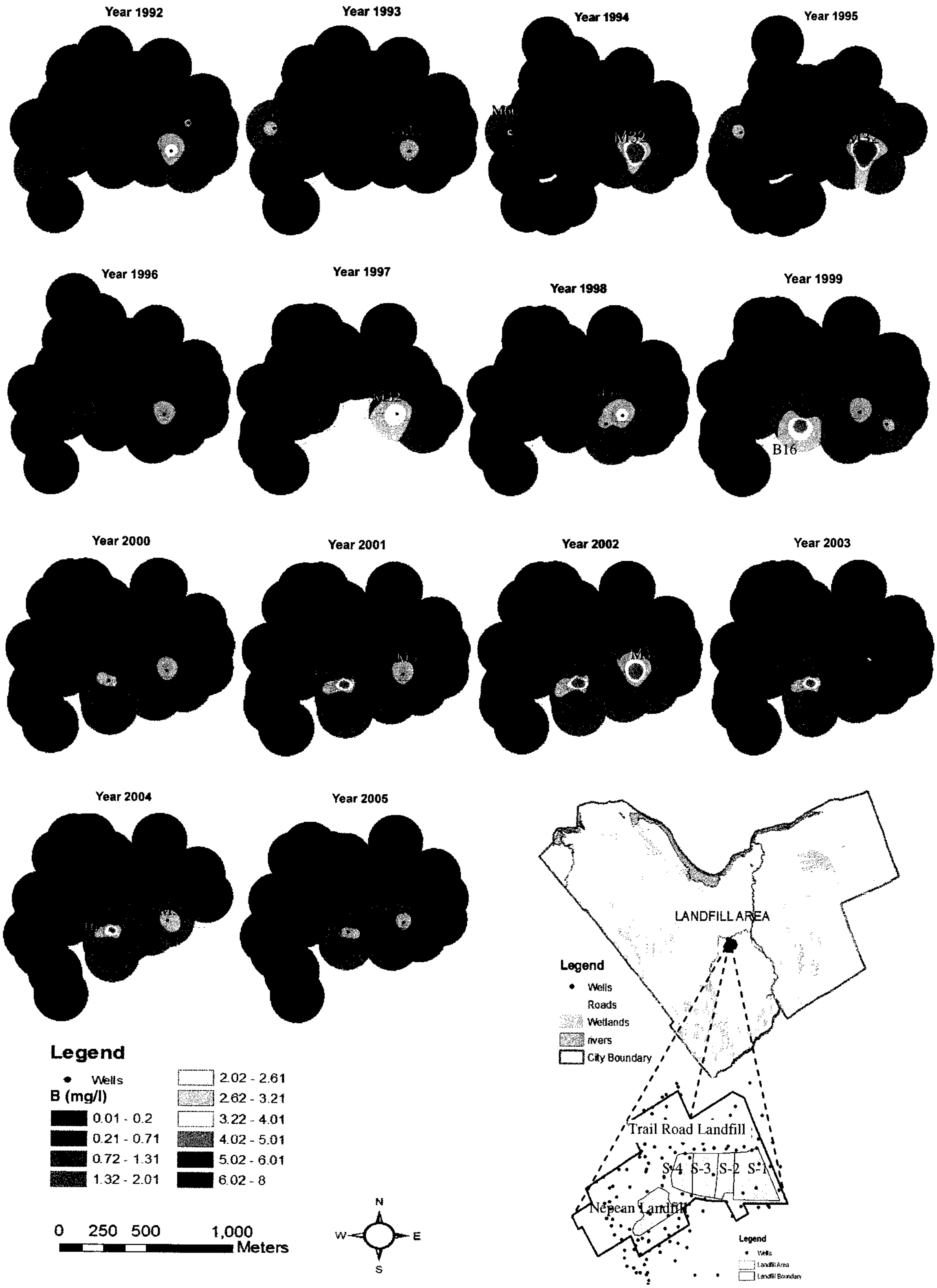


Figure 5.2 Reclassification maps showing spatial variation of B concentration in deep aquifer (for 1992-2005) in a 10-class common scale.

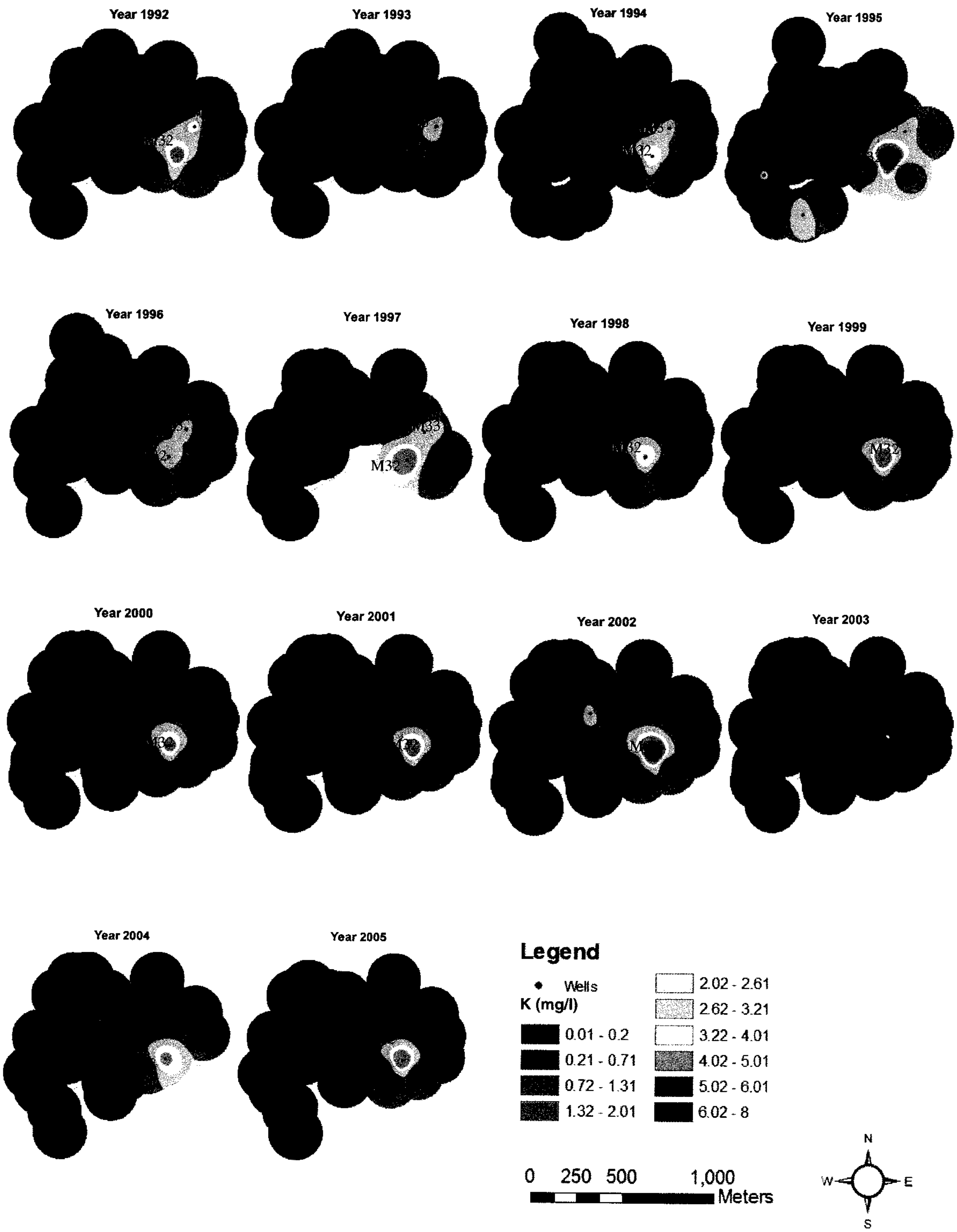


Figure 5.3 Reclassification maps showing spatial variation of K concentration in deep aquifer (for 1992-2005) in a 10-class common scale.

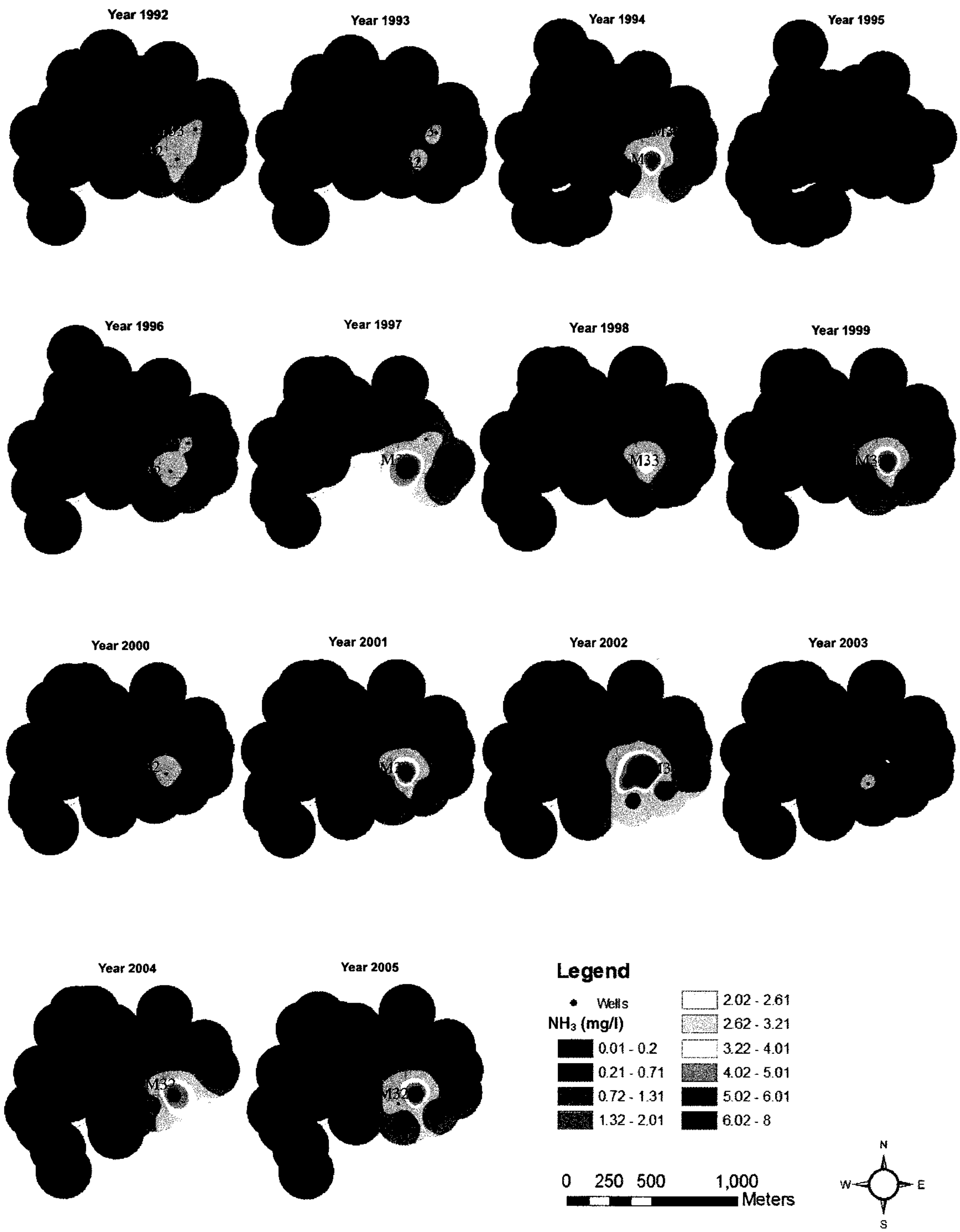


Figure 5.4 Reclassification maps showing spatial variation of NH_3 concentration in deep aquifer (for 1992-2005) in a 10-class common scale.

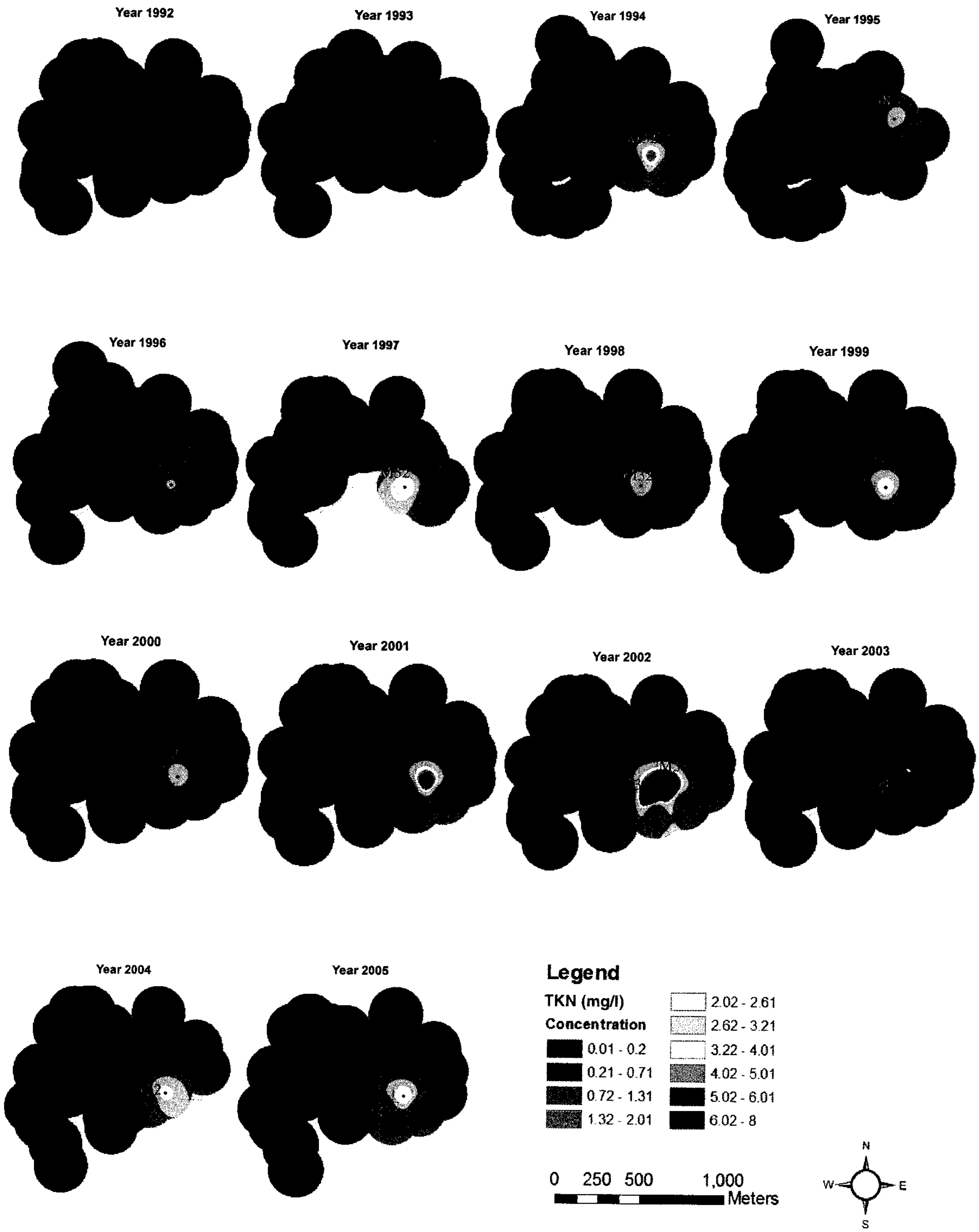


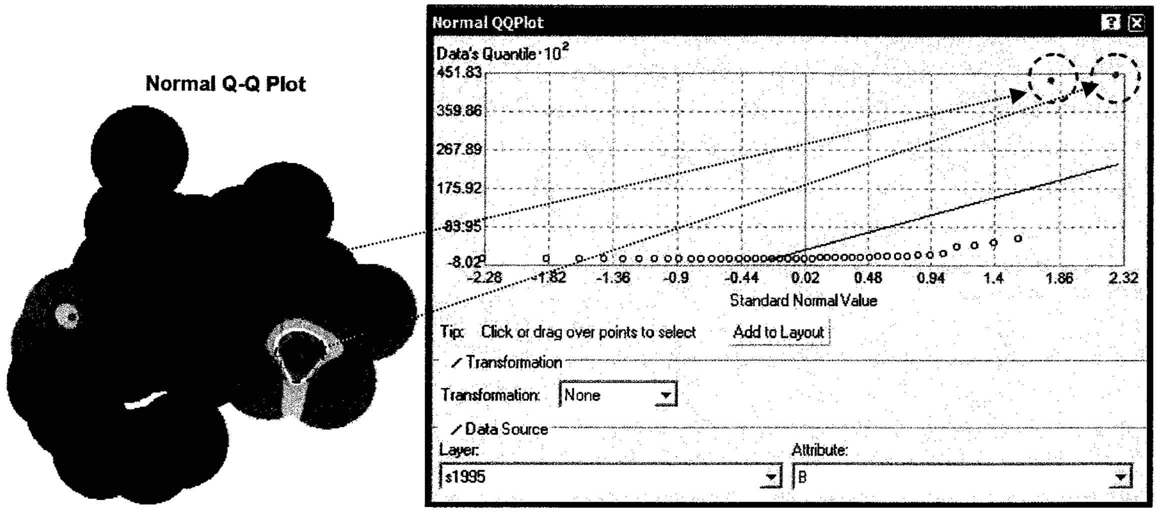
Figure 5.5 Reclassification maps showing spatial variation of TKN concentration in deep aquifer (for 1992-2005) in a 10-class common scale.

The IDW interpolation method is applied to convert the primary point data related to monitoring wells in the deep aquifer into maps (continuous surfaces). The resulting maps are on a continuous scale and are not reclassified (Appendix A). Reclassification is done to create a common scale for all maps related to each of B, K, NH_3 and TKN, as displayed in Figures 5.2 to 5.5. These reclassifications have been applied for each variable separately, which reduces the spatial pattern complexity of the spatial variability of the maps. This simplifies visual comparison and identification of the spatial and temporal changes of each variable. Finally, a relationship of spatial and temporal changes for B with its associates is determined.

Figure 5.2 shows changes of B concentration in the leachate is decreasing over time, which could be due to the lowering rate of leachate generation from Stage 1 and Stage 2. On the other hand, Stage 3 and Stage 4 comprise bottom liners which retard leachate percolation in groundwater. Therefore Stage 1 and Stage 2 could be the source to generate the detected plume. As displayed in Figure 5.2, a slight reducing trend can be identified for B during recent years. A common location of high concentration is apparent in Figure 5.2, 5.3, 5.4 and 5.5, which corresponds to well M32, located in Stage-1. The background study of the TRL site records showed that Stage-1 had no bottom liner, which is still considered as source of pollutants.

Outlier Identification by Normal Quantile-Quantile (Q-Q) Plot

Data exploration by Normal Q-Q plot is applied to identify the wells with high concentrations of B, K, NH_3 and TKN and distinguish them as outliers. Wells with extreme values for B are identified on Q-Q graphs, and dynamically linked with the maps to determine their locations. In Figure 5.6, the light blue dots on the Q-Q plot represent extreme values for B, and the same symbol is applied on the map for corresponding points. As indicated in Figure 5.6 (a) and (b), two wells had extreme values for B in 1995, and can be considered outliers. Corresponding locations on the map indicate where these outliers are located (Figure 5.6(a)). The similar process has been applied for all layers of the post classification method, and the output is shown in the Table 5.1.



(b) (a)

Figure 5.6 (a) Normal Q-Q plot of B in 1995 to identify wells with extreme values (b) Interpolated map of B in 1995. Corresponding well locations with extreme B values encircled by blue dots.

Table 5.1 shows the results of the identification of wells with extreme values for B, K, NH_3 and TKN in different years. As shown in the table, wells M32, M33, M34, BH16 and M8 are common locations in which extreme concentrations of these variables can be observed.

Table 5.1 Wells with extreme values for B, K, NH_3 and TKN (1992-2005)

Variables	B	K	NH_3	TKN
1992	M34	M32 M33	M32 M33	M34
1993	M66 M32	M33	M32	M32
1994	M66 M32	M32 M33	M32	M32
1995	M32	M32	M32 M33	M32
1996	M32	M32 M33	M32 M33	M32
1997	M32	M32	M32	M32
1998	M32	M32	M32	M32
1999	M32	M32	M32	M32
2000	M32	M32	M32	M32
2001	M32	M32	M32	M32
2002	M32 BH16	M32	M32 M34	M32 M34
2003	BH16		M34	M34
2004		M32 M40	M34	M32
2005	M32 BH16	M32	M32	M32

5.5.2 SIMPLIFIED POST CLASSIFICATION

Interpolated maps are then simplified by reclassification to reveal three levels of variability of B, K, NH_3 and TKN for 1992-2005 maps. MultiSpecW32 (a GIS tool) has been applied to identify and quantify the three levels of change. Figure 5.7 shows them as high, medium, and low-or-no change variability of the area of maps created by MultiSpecW32. The tool can also determine temporal trends of area changes in different categories as presented in Figure 5.8.

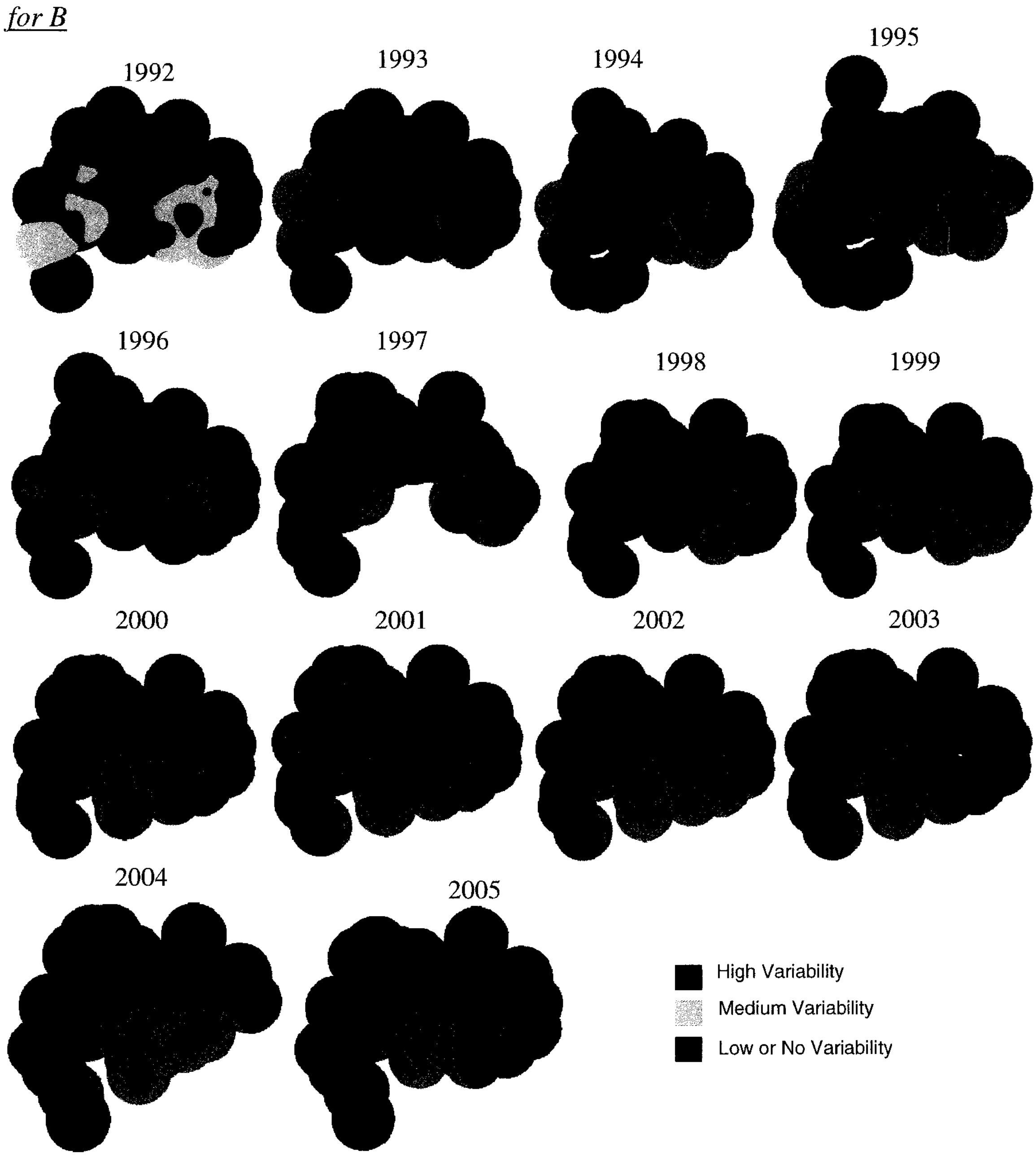


Figure 5.7 Identification of three levels of variability of B based on its corresponding interpolated maps for 1992-2005

The maps of the simplified post-classification method in Figure 5.7 show increasing B variability in the beginning followed by a gradual decrease. Though the overall changes are calculated to be 1.78% of the area, the high variability area in the site accounts for this percentage. Figure 5.8 shows there is also a little declining trend that may indicate an improvement in leachate pollution. This weak trend can be also identified from the reclassified maps of Figure 5.7.

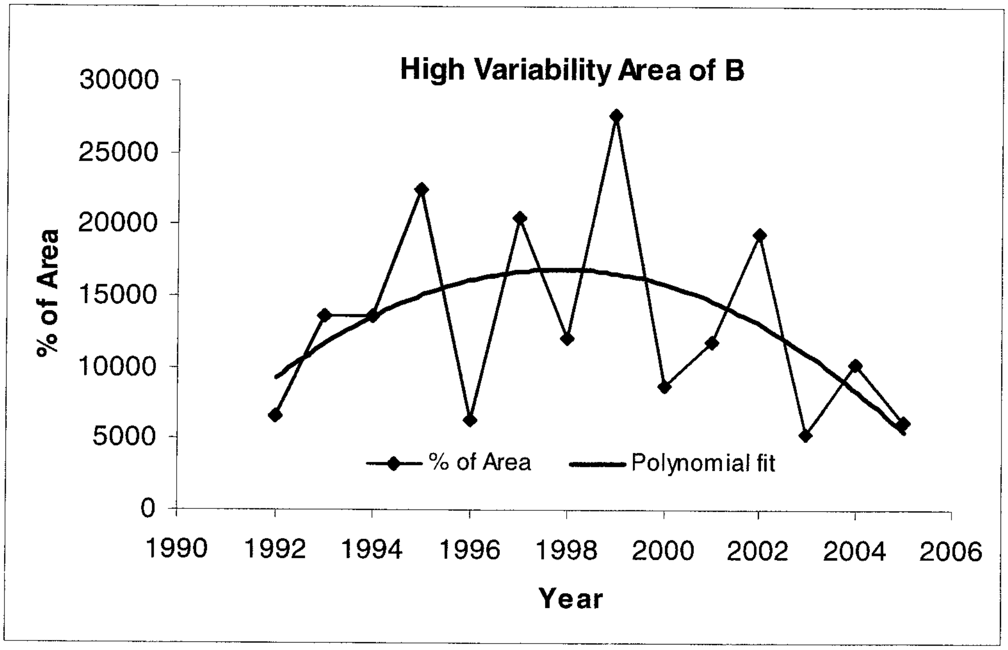


Figure **5.8** Temporal trends of area of high variability class for B for 1992-2005

5.5.3 CHANGE DETECTION BY PRINCIPAL COMPONENT ANALYSIS

According to Chang and Yoon (2003) and McAdams and Demirci (2006), Principal Component Analysis (PCA) is a multivariate statistical analysis method that can identify major components that explain most of the observed variability of variables. PCA is a quantitative method of spatial-temporal change detection. It involves a mathematical procedure that transforms a number of possibly correlated variables into fewer uncorrelated variables, called principal components. The transformation is done from an *n* dimensional space into an *m* dimensional space, so that $m<n$ (number of components is smaller than the number of original variables). The first principal component (PC1) accounts for as much of the variability in the data as possible, and each successive component accounts for as much of the remaining variability as possible. Each component is created by a linear combination of the original variables. These smaller numbers of new variables (or components) preserve the relationships present in the original data.

PCA can be applied on:

1) many variables related to samples taken for one time interval

or

2) one variable related to samples but for many time intervals (temporal application)

In this part of this study, PCA analysis is applied in its temporal mode (by method 2) to identify groups of pixels on maps that act and are temporally similar to each other. Results from 1992-2005 reveal zones of no change and zones of change for each of the studied variables (B, K, NH_3 and TKN). Temporal application of PCA can be a useful method to detect change over time in maps, for the control and monitoring of environmental pollution. The output of temporal PCA is presented for B, K, NH_3 and TKN in a common scale (Figure 5.9).

PCA is not applied to the tabular data of wells in the study, but directly on the interpolated maps of each variable separately in different time intervals (1992 to 2005). Results of PCA indicate changes in B, K, NH_3 and TKN over time in different locations of the TRL site (Figure 5.9).

The inputs to this multivariate analysis for each variable (B, K, NH_3 and TKN) are 14 interpolated maps, each related to a year in 1992-2005. Six principal components are extracted from these variables. For B, K, NH_3 and TKN, the second principal component (PC2) represented the regions with high variability. Figure 5.9 shows the distribution of PC2 scores for each variable. In this analysis, the remaining PCs show the remaining variability and are all smaller than PC2, while PC2 shows more obvious patterns of change compared to the other PCs. The dark blue regions in all maps in Figure 5.9 indicate minor or almost no change. Most of the areas in all the maps are pale blue, indicating that variability in time is low in the majority of regions. The small, pale brown parts of the areas represent regions with high change for B, K, NH_3 and TKN. These could be considered regions with a high concentration of B, and where more B variability in time was recorded. K and NH_3 show very similar patterns as B, and the variability pattern for TKN was generally in the same region. Such similarities with B are expected, as the initial principal component analysis in the exploratory data analysis indicated association of B with K, NH_3 and TKN.

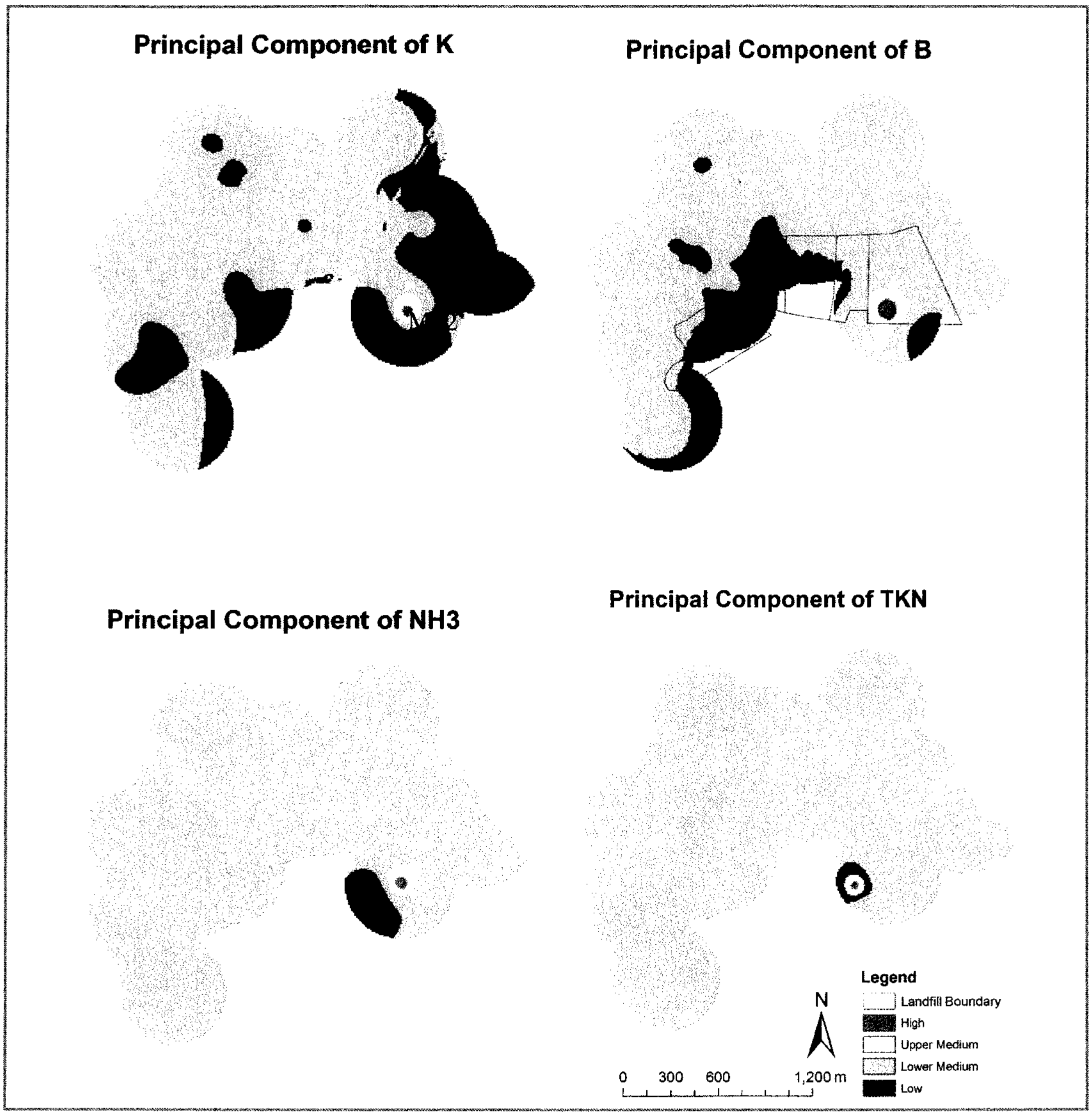


Figure 5.9 Distribution of PC2 scores for B, K, NH_3 and TKN representing regions with similar change characteristics for these variables from 1992-2005

5.5.4 CHANGE DETECTION BY STANDARD DEVIATION

Change detection by standard deviation (SD) is an important method in map analysis, and is applied in many studies. It is traditionally one of the simplest and most popular approaches for change detection among all the developed methods. In this approach, the maps created by IDW interpolation are the inputs to calculate standard deviation maps (Appendix A).

The standard deviation maps in Figure 5.10 are displayed in four classes of standard deviation. Various levels of calculated standard deviation represent the magnitude of temporal variability of B, K, NH_3 and TKN from 1992-2005 in different regions.

Areas showing higher standard deviation indicate higher variability of B, K, NH_3 and TKN during 1992-2005, and areas with lower standard deviation represent lower variability.

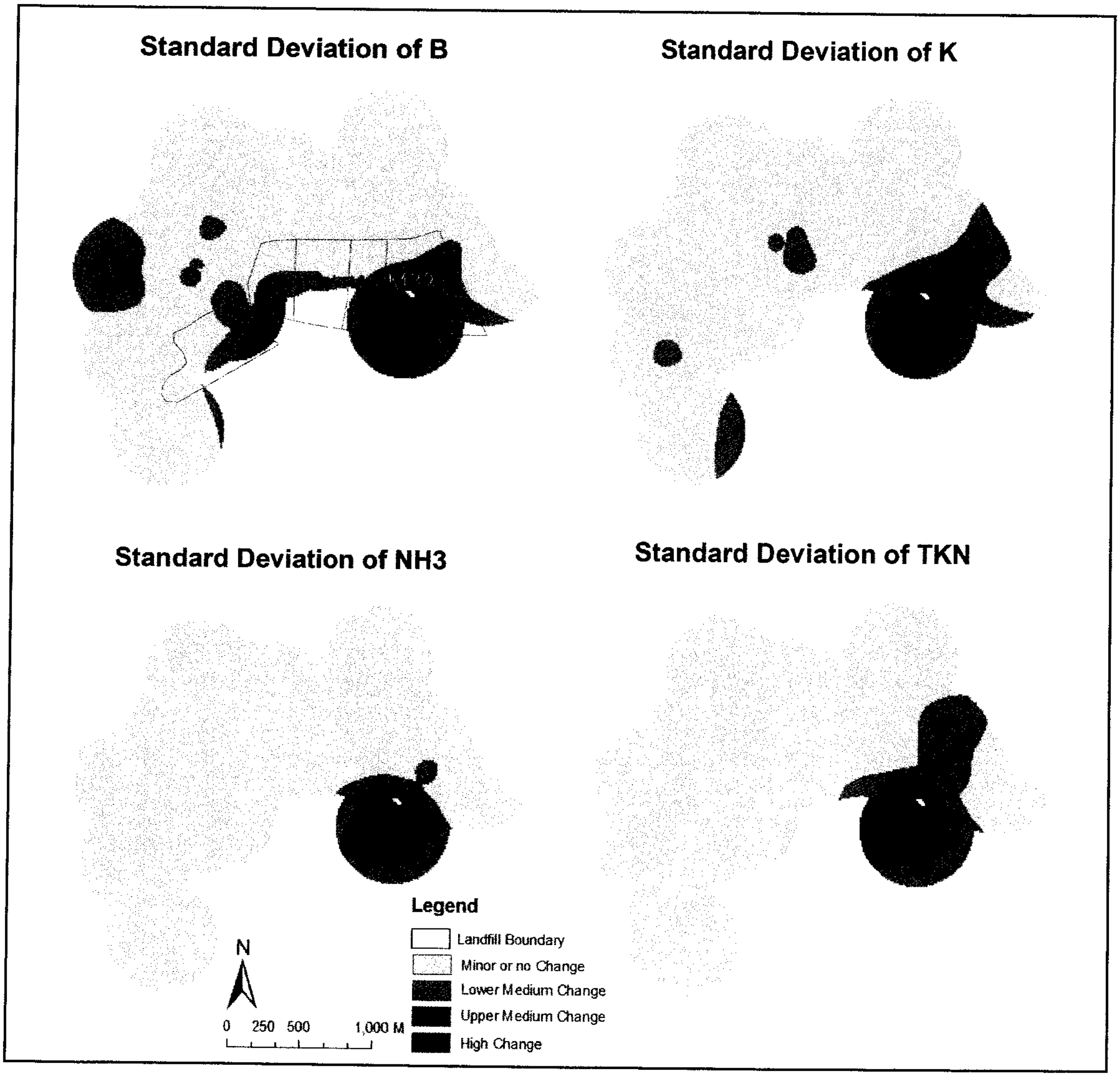


Figure 5.10 Change detection maps based on SD of pixels of maps related to 1992-2005

In Figure 5.10, the darker blue areas indicate higher standard deviation and represent the higher variability of the area. Most of the regions with higher standard deviation can be found in Stage 1 of the landfill, where well M32 is located. Based on statistical analysis, post-classification visualization and other studies, the standard deviation method provides results that may be better matched with background of the TRL (Monitoring Report/2004 by Dillon Consulting Limited).

5.5.4.1 Coefficient of Variation (CV)

The application of CV for spatial-temporal change detection is also tested in this study. The CV map in Figure 5.11 is calculated by dividing standard deviation maps by mean maps derived from interpolated maps related to 1992-2005 for B. As Figure 5.11 indicates, the normalizing of standard deviation by mean by calculating the coefficient of variation does not reveal the spatial-temporal changed pattern for B, as identified in the output maps related to the previous methods. In Figure 5.11, low CV (blue areas) represents low variability regions, and high CV (red areas) high variability regions.

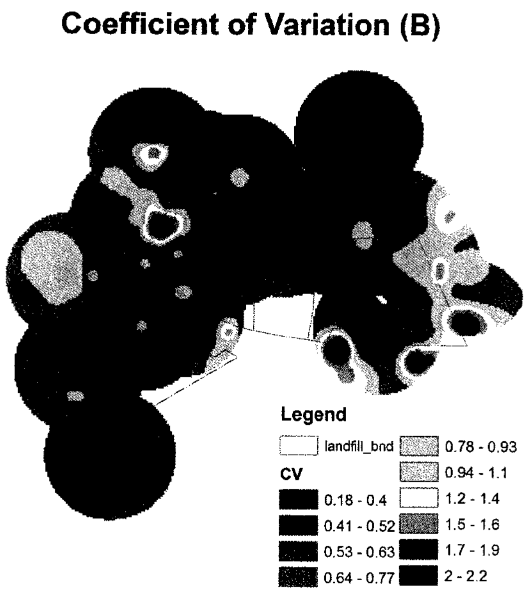


Figure 5.11 Distribution of coefficient of variation of B

5.5.4.2 Comparing Change Detection Results of PCA and SD for B

The regions with high B variation in time that are examined in the previous sections can be compared to each other to find regions where high levels of change have been detected by both the SD and PCA methods.

To find regions of high temporal variability via standard deviation, a boolean binary map is created to display the landfill site area in classes of high and low variability. High variability regions can be considered areas of "change," and the remaining areas can generally be considered "unchanged" (Figure 5.12). A similar map of changed and unchanged regions is also created using the temporal PCA analysis scores for B.

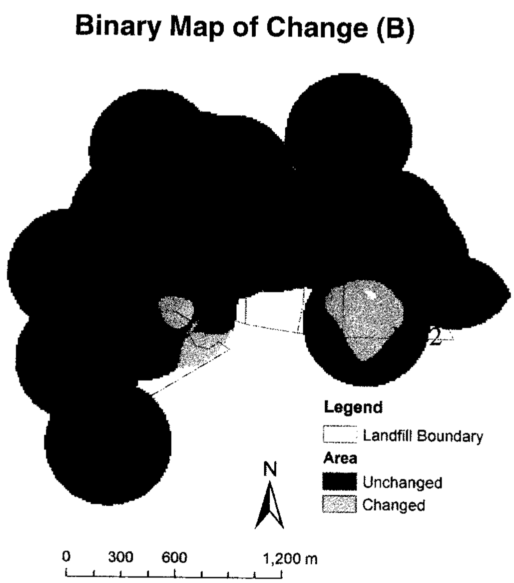


Figure 5.12 Binary map of "changed" and "unchanged" areas based on binary classification of SD of interpolated maps for B from 1992-2005.

These binary maps are then integrated by Boolean AND to find areas which are considered "changed" in both binary maps. The method is useful for visual comparison of changed and unchanged classes on the maps (Figure 5.13).

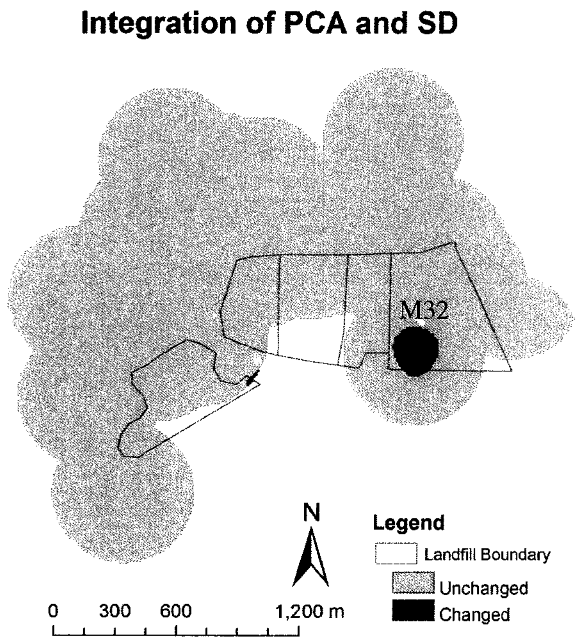


Figure 5.13 Identifying regions considered as "changed" created by boolean AND operator applied on the binary maps of variation from PCA and SD of B

5.5.5 CHANGE DETECTION BY UNSUPERVISED CLASSIFICATION (CLUSTER ANALYSIS)

K-mean clustering is a multivariate statistical method that can determine the statistical structure of the data by classifying observations in clusters or groups. Measurements of objects within each group are similar to each other, and clusters are selected so there are maximum differences between them. This method is also called unsupervised classification, since it does not require training data, and only needs the number of expected clusters as input. K-mean is one of the simplest methods of unsupervised classification, and an easy way to classify a given data set through a certain number of clusters (assume k clusters).

Unsupervised classification based on the k-mean clustering method is applied on the data extracted from interpolated maps related to each year in 1992-2005 for each of the B, K, NH_3 and TKN variables. The results of the k-mean clustering for each variable indicate regions with similar variability in time from 1992 to 2005.

To apply this method, the pixel value of each interpolated map related to each year is extracted at the location of a dense grid designed for this area (Figure 5.14). To design the grid, a software called FIELDS—provided by the Environmental Protection Agency (EPA, 2009)—is applied as an extension of ArcGIS (Figure 5.14). Figure 5.14 displays the grid of points produced by FIELDS. Extraction of map pixel values at the grid points is done for the B, K, NH_3 and TKN variables for all years (1992-2005).

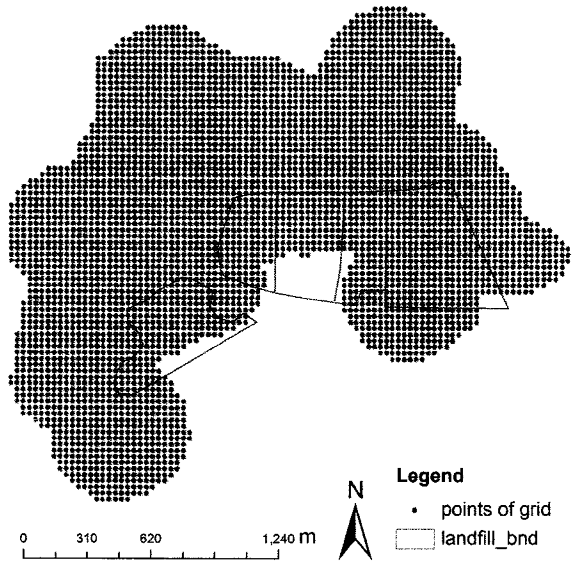


Figure 5.14 Spatial distributions of samples (grid points) created to extract raster values for cluster analysis

For each variable (e.g. B), the values at grid points extracted for all the years (1992-2005) are entered into SPSS (statistical software) for k-mean clustering, in order to calculate which grid points belong to which clusters for B across the area of the landfill site. In k-mean clustering, each sample (record) membership of every extracted cluster can be calculated and saved in the outputs. These calculated cluster membership values can be mapped in GIS to visualize spatial distribution of the cluster that has the maximum membership value for each grid point across the landfill site (Figure 5.15). As displayed in Figure 5.15, four clusters are extracted for each variable, based on the above. Each of the identified clusters (e.g. B) in Figure 5.15 represents a level of variability in time for B. In Figure 5.15, the zones of clusters 1 and 2 for B are displayed in red and blue, and represent maximum change and variability in time.

The spatial-temporal changes are detected separately for B, K, NH_3 and TKN, based on the k-mean clustering as explained. As shown in Figure 5.15, cluster 1, which indicates a zone of high variability in time, appears in almost the same area for all the variables. The cluster 2 distribution, representing high variability areas for B, K, NH_3 and TKN, is slightly less than the cluster 1 pattern. Cluster 1 and 2 together cover almost same general area in all the maps, and highlight where the maximum spatial-temporal changes occurred in 1992-2005.

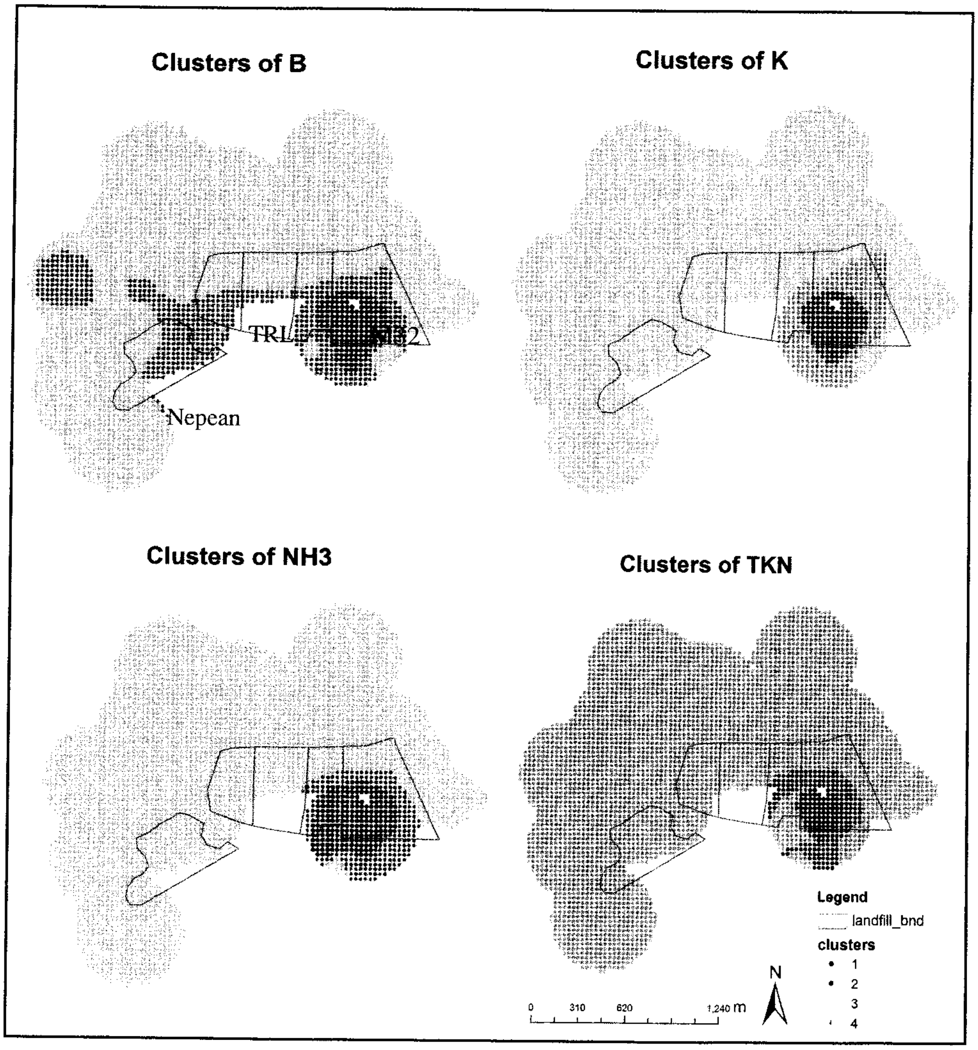


Figure 5.15 Spatial distributions of 4 extracted clusters

Cluster 1 (red dots) represents high variability, and its distribution corresponds to the Stage 1 site where the M32 well is located. These observations are consistent with the results of previous change detection methods. It is assumed that the reason for the occurrence of cluster 1 is the lack of bottom liner at Stage 1 during the lifetime of the TRL site.

5.6 OUTPUT COMPARISON

The above results prove the applicability of the methods employed for such environmental applications. Four types of change detection techniques are applied to detect regions with different levels of variation for B, K, NH_3 and TKN in time during 1992-2005: post-classification comparison, principal component analysis, standard deviation and unsupervised classification. Interpolation has been used to create input maps for change detection techniques (except k-mean clustering) for leachate variables (B, K, NH_3 and TKN).

These four types of change detection are applied in a GIS (ArcGIS) environment, with the statistical assessment done by using SPSS. GIS is a set of tools that are useful in many change detection studies and projects, particularly when multivariate data are involved. Visual assessment or post-classification for change detection has been applied to classify interpolated maps. Change detection based on the post-classification method requires much effort to implement common scale classification. However, it is a widely used method of monitoring change detection for groundwater environmental pollution control, as it determines the different classes of pollution intensity at the affected well locations. Compared to other methods, post-classification produces a complete matrix of change information that shows the intensity of spatial variability, and it provides a good understanding of the nature of the change based on non-quantitative results. The key is to create representative, reclassified maps. MultiSpecW32 has been applied on the classified raster layers to help with the quantitative analysis in the post-classification method. MultiSpecW32 simplified B maps and quantified the trend of temporal changes by calculating the changed area. This combination of post-classification methods provides comprehensible outputs of both quantitative and non-quantitative measurements. When a quantitative approach is needed, the above combination of procedures would likely be effective. However, there are limitations to using this combination in cases which involve long data series, as the method can be tedious and time consuming.

PCA involves more steps and needs more interpretation, but it can provide useful information for change detection representing the location and intensity of change (Figure 5.16). For example, PCA requires interpretation of extracted PCs and the spatial distribution of the derived pattern for their scores, to determine which PC represents change in time. Compared to using PCA for change detection in time, applying standard deviation on interpolated maps is easier, and requires fewer steps and intermediate interpretations. The results of these two quantitative methods are comparable, and they are similar for the regions with high variation in time as expected (Figures 5.16 & Figure 5.17).

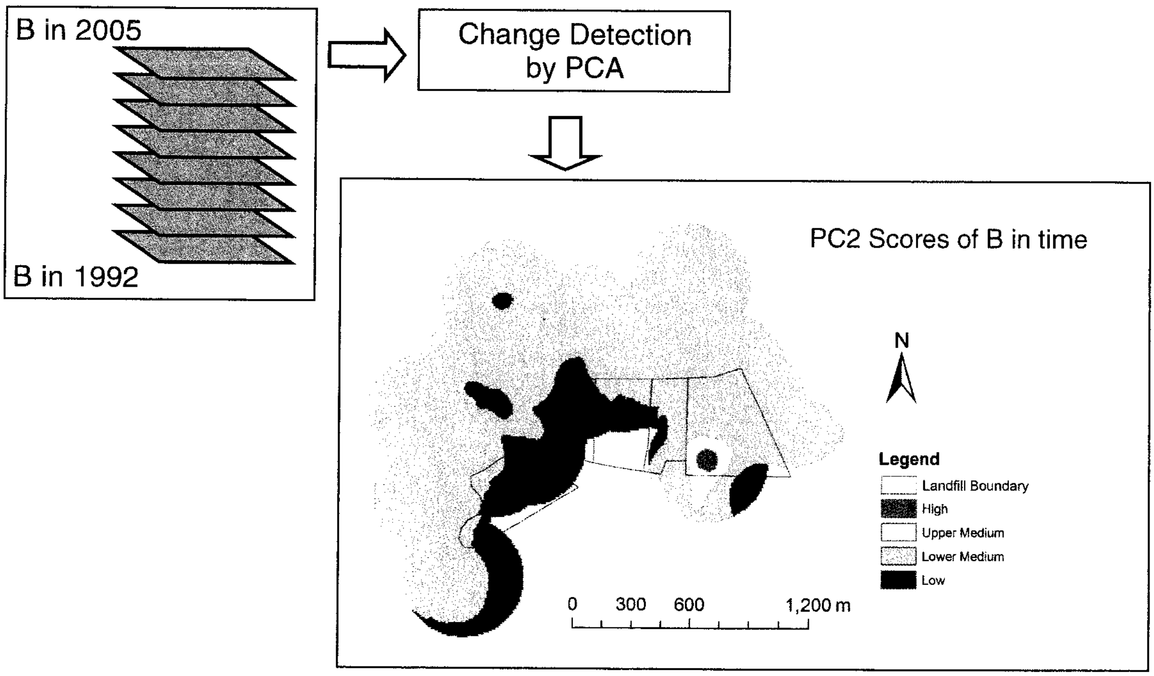


Figure 5.16 Application of Principal Component Analysis (PCA) for change detection of B in time (1992-2005)

One of the negative aspects of using standard deviation on temporal maps (as in this study) rather than PCA, is SD's higher sensitivity to extreme values (outliers). The presence of outliers, due to measurement errors, can have great influence on the final standard deviation map, even with one outlier value related to only one year, and not the previous or following years. In order to avoid this, outliers should be investigated before the interpolation of variables to create maps. Those outliers present for only one year should be tested and, if not valid, eliminated from the data before interpolating maps. Extreme values and outliers those are present more than one year should not be removed, since they can be an indication of "real" change in an area. The application of PCA method can reduce the effect of existing noise in the data and can reveal the overall classes of change.. However, coefficient of variation maps is not as efficient as using standard deviation or PCA results to identify changed regions for the studied variables.

The standard deviation map in this study reveals slightly more regions with high variability of studied variables in time, compared to other quantitative methods applied for change detection.

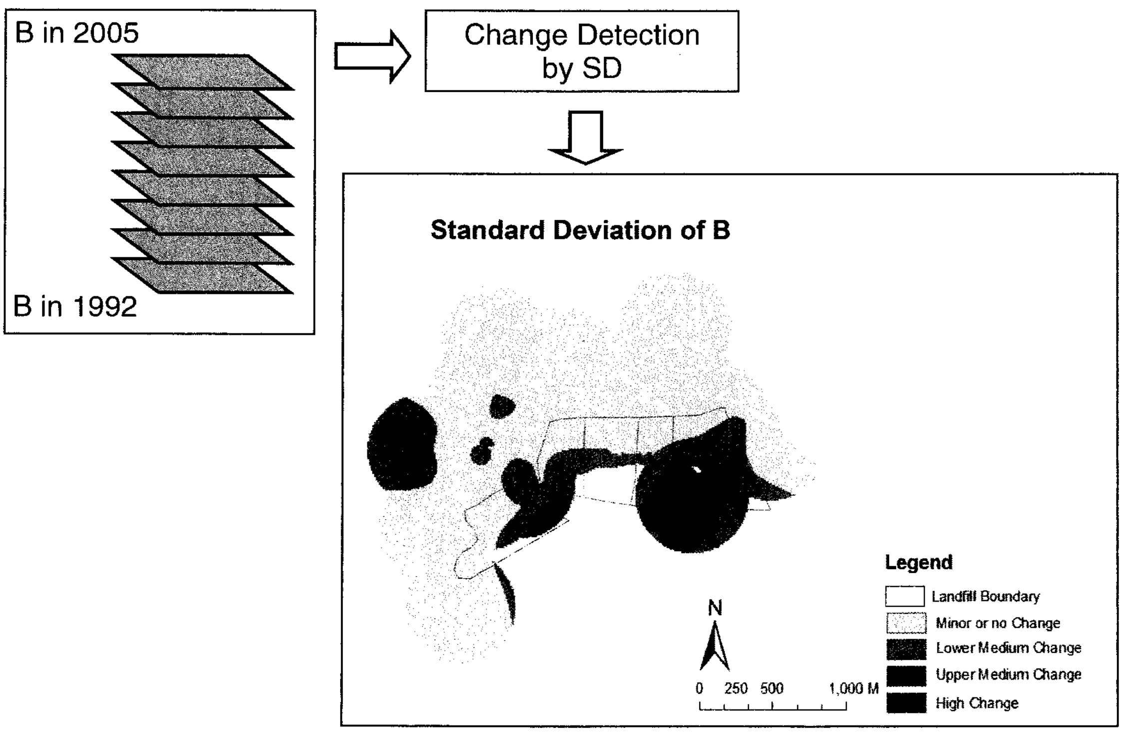


Figure 5.17 Application of SD for change detection of B in time (1992-2005)

Application of the unsupervised classification by k-mean clustering method is also an efficient means of identifying changed areas. Similar to PCA, this is a multivariate statistical method, and is also more robust than applying standard deviation.

The existence of individual outliers using PCA can have less influence on the final results than with the standard deviation method. PCA is a reliable technique to understand the temporal-spatial behaviour of environmental data, like the data in this study. Different tools or software can be applied to perform the cluster analysis. A grid approach to extract the input data for the clustering k-mean method has been adopted in this study. With this grid data extraction approach, the output depends on the density of grid points to extract the pixel values of maps. The output of cluster analysis has the same congruity of output as other change detection methods, and the results for B, K, NH_3 and TKN have shown less noise. It is also easier to compare the results with those of other change detection techniques by reducing the complexity of the output maps. Hence, cluster analysis is strongly recommended as an efficient method to monitor this or similar landfill sites.

Based on above, results of all three applied quantitative methods (PCA, standard deviation and unsupervised clustering) seem to be consistent and useful for delineating plume regions and corresponding wells that had variability or considerable change relative to the other wells. Using these change detection methods on rasters (maps), rather than directly using on the points (tabular) data of wells (points), is one of the unique aspects of this thesis. Since the input maps (rasters) to these methods are created by IDW interpolation method which is based on a spatial moving average technique, less noise exist in each map compared to the initial well tabular data. This approach reduces noise in the inputs and outputs, and increases robustness of the results and derived change patterns.

Change detection methods are active research topics, and new methodologies and applications continue to be developed. Users often select several methods to implement change detection in their studies, compare them, then identify the optimal results through accuracy assessment. As discussed earlier, the important factors for successfully implementing change detection are authentic data collection, pre-processing of the data, determining the change categories, and using the appropriate change detection technique.

Among these, the accuracy assessment for change detection is the most challenging, due to availability of new and independent data. An accuracy assessment step is required to quantitatively differentiate the efficiency of the applied change detection methods. Once data is available for the TRL site an accuracy assessment could be a part of future studies. This step should also be considered for similar environmental projects.

CHAPTER 6

CONCLUSION AND RECOMMENDATION

6.1 CONCLUSION

This research contributes to the development of a framework and solution techniques for delineating landfill leachate plume characteristics in space and time in groundwater environments. It also enhances and optimizes the application of change detection methods in similar environmental projects. This chapter presents the principal conclusions drawn from this research, and a set of recommendations for future research and development in such fields of environmental study.

The first conclusion is related to the approaches of data pre-processing. The second conclusion addresses the optimization of change detection, in particular the similarities and differences of various change detection techniques. The recommendations emphasize the need for both spatial and temporal data collection and preparation, and effective methodologies for quality assessment of landfill waste management. They also advise that the validity of outputs from current landfill approaches should be viewed as complementary.

In the exploratory data analysis detailed in Chapter 4, three major tasks have been performed to assess the data quality, to detect outliers, and to extract the principal variables that are associated with B (boron). Correlation and descriptive statistics are calculated to assess the data quality and the strength of the relationships between variables. Outlier detection is performed using the Dixon-Thompson test. A comprehensive Factor Analysis is performed to extract the variables highly associated with B, and the analysis extracted them in a multivariate data context. Due to the large number of variables analyzed in leachate, and the results of data monitoring over many years, the study would become too complicated without the application of multivariate methods. And it could be that some of the variables measure different aspects of the same underlying phenomenon. For situations such as these, (exploratory) factor analysis can be a very efficient statistical multivariate method. Most of this exploratory data analysis has been done by carrying out several statistical analyses on the data using SPSS (statistical software). The results of this exploratory data analysis are considered in subsequent steps of the study. Thus, Chapter 4 not only offers the possibility of achieving a clear view of the data, but also the potential for using similar approaches in the pre-processing steps, which are often required in spatial and temporal analysis, or in other similar environmental projects.

In Chapter 5, it is determined that change detection can be a complicated process influenced by multiple factors. Knowledge related to the landfill site history is important for better interpretation of the results of the change detection process. Pre-processing of the data is required in order to apply the change detection techniques. These procedures are explained in Appendix A.

Change detection for the variability of B, K, NH_3 and TKN in samples are investigated and illustrated individually. Two approaches are unique to this study: 1) the application of various change detection analysis methods that are typically used in image processing in an environmental analysis, and 2) applying the change detection methods on the maps created for the studied variables (B, K, NH_3 and TKN), which reduces spatial noise and enables more robust interpretations of results, rather than directly using point or tabular well data.

Classification to a common scale is applied first, which enables a visual identification of change and trends in time for the studied variables. This is followed by the application of three quantitative change detection techniques on temporal inputs related to interpolated maps of each variable during 1992-2005:

1) Principal Component Analysis (PCA)
2) Standard deviation (SD)
3) Unsupervised classification (k-mean clustering)

The study shows that the overall variability of B is decreasing until 2005. Stage 1 and Stage 2 is closed and capped. Therefore, it is assumed that leachate producing rate is decreasing in order of time. In addition, Stage 3 and Stage 4 comprise bottom liners which retard leachate infiltration in groundwater. Therefore, Stage 1 and Stage 2 will be only active source term for delineating the pollution plume at the site and as both of the stages are closed, leachate generation rate will be reduced in the long term. It is expected that future studies on this landfill site indicate reduction of the leachate plume extent traceable by B.

Due to the lack of recent, independent well data for the TRL site, no accuracy assessment was performed to determine the quantitative efficiency of the outputs of the three methods. The results of the methods are similar and comparable, and all identify generally the same regions as areas with high levels of change for B and associated variables during 1992-2005. These findings are supported by the background and history records of the landfill site. Application of each of the three methods is recommended, and can be useful in similar environmental research to discover and differentiate regions of "change" and "no-change" in time. The advantages and disadvantages of applying each are discussed in chapter 5.

6.2 RECOMMENDATION FOR SIMILAR STUDIES

As an integrated approach, the following categorized recommendations can strengthen the research. For application of the applied procedures and methodologies in similar studies, the following are recommended:

- ◊ Apply various exploratory data analysis before analyzing monitoring data;
- ◊ To reduce the existing noise in data and to have a continuous representation of variables use interpolated maps rather than tabular data of individual wells;
- ◊ Apply PCA (a useful multivariate statistical method for spatial-temporal change detection in its temporal mode, as explained in the text), followed by standard deviation and, if possible, the unsupervised classification (clustering) method;
- ◊ Review the background and design history of the landfill site before interpreting any change detection results;
- ◊ Assessment of results is also important, if independent data is available; and,
- ◊ In the case of a lack of data for accuracy assessment, patterns calculated from the application of other suitable and applicable change detection methods can be compared.

REFERENCES

Abdul-Rahman, A., Pilouk. M., (2008), "Spatial data modelling for 3D GIS" Springer Berlin, Heidelberg, New York.

Al-Yaqout, A. F, Hamoda, M. F., and Zafar, M., (2005), "Characteristics of Wastes, Leachate, and Gas at Landfills Operated in Arid Climate" Journal of ASCE, Vol. 9, no. 2.

Al-Yaqout, A. F, Hamoda, M. F., (2003), "Evaluation of landfill leachate in arid climate - a case study" Journal of Science Direct, Vol. 29, no. 5, pp. 593-600.

Bagchi, A., (1994), "Design, Construction and Monitoring of Landfills" Sanitary Landfills, John Wiley & Sons, Inc, New York.

Bakac, M., Kumru, M. N., (2000), "Factor analysis in the geochemical studies along the Gediz river, Turkey" Journal of Radioanalytical and Nuclear Chemistry, Vol. 249, no. 3, pp. 617–624.

Bedient, P. B., Rifai, H. S., Newell, C. J., (1999), "Ground Water Contamination-Transport and Remediation" Second Edition, Prentice Hall PTR, Upper Saddle River, New Jersey.

Bjerg, P. L., Albrechtsen, H. J., Kjeldsen, P., Christensen, T. H., and Cozzarelli, I., (2003), "Treatise on Geochemistry" Editor: Barbara Sherwood Lollar, Executive Editors: Heinrich D. Holland and Karl K. Turekian, Elsevier, Vol. 9, p.579-612.

Chang, H., Yoon, W. S., (2003), "Improving the Classification of LANDSAT Data using Standardized Principal Components Analysis" KSCE Journal of Civil Engineering, Vol. 7, no. 4, pp. 469-474.

Christensen, T. H., Kjeldsen, P., Albrechtsen, H. J., Heron, G., Nielsen, P. H., Bjerg P. L., Holm, P. E., (1994), "Attenuation of landfill leachate pollutants in aquifers" Critical Reviews in Environmental Science and Technology, Vol. 24, no. 2, pp.119-202.

Christensen, T. H., Kjeldsen, P., Bjerg P. L., Jensen, D. L., Christensen, J. B., Baun, A., Albrechtsen' H. J., Heron, G., Nielsen, (2001), "Biogeochemistry of Landfill Leachate Plume" Journal of Applied Geochemistry, Vol. 16, pp. 659-718.

Davis, A. P., McCuen, R. H., (2005), "Statistical Methods for Data Analysis" Stormwater Management for Smart Growth, Hydraulic Engineering, pp. 37-62.

Drton M., Sturmfels, B., Sullivant, S., (2006), "Algebraic factor analysis: tetrads, pentads and beyond" Probability Theory and Related Fields, Vol. 138, pp. 463–493.

Final Report (2005), Trail Road and Nepean Landfill Sites, 2004 Monitoring and Operating Report, Dillon Consulting Limited, Ottawa, ON, Canada.

Final Report (2004), Trail Road and Nepean Landfill Sites, 2003 Monitoring and Operating Report, Dillon Consulting Limited, Ottawa, ON, Canada.

Field, A., (2000), "Discovering statistics: Using SPSS for Windows" Sage Publications, London, pp. 486.

Goldberg, S., (1997), "Reactions of Boron with Soils" Plant and Soil, Kluwer Academic, The Netherlands, Vol. 193, pp. 35-38.

Gupta, U.C., (1993), "Boron and its role in crop production" CRC Press, Cleveland, Ohio, USA.

Haining R., Wise, S., (1997), "Exploratory Spatial Data Analysis" Retrieved August 20, 2009 from http://www.ncgia.ucsb.edu/giscc/units/u128/u128f.html.

Hudec, P. P., (1984), "Statistical Methods of Aggregate Durability Evaluation" Bulletin of the International Association of Engineering Geology.

Johnson, R., Wichern, D., (1982), "Applied Multivariate Statistical Analysis" Prentice-Hall, Inc.: Englewood Cliffs, NJ, USA.

Jones-Lee, A., Fred Lee G., (2000), "Appropriate Use of MSW Leachate Recycling in Municipal Solid Waste Landfilling" G. Fred Lee & Associates, El Macero, CA, USA.

Kaiser, H., (1974), "An index of factorial simplicity" Psychometrika Vol. 39, no. 1, pp.31-36.

Kirkman, T. W., (1996), "Statistics to Use, Retrieved" August 25 from http://www.physics.csbsju.edu/stats, Saint John's University, Minnesota, USA.

Kjeldsen, P., Barlaz, A, M., Rooker, P, A., Baun, A., Ledin, A., and Christensen, H, T., (2002) "Present and Long-Term Composition of MSW Landfill Leachate" Critical Reviews in Environmental Science and Technology, Vol. 32, Issue 4, pp 297-336.

Kulikowska, D., Klimiuk, E., (2006), "The effect of landfill age on municipal leachate composition" Journal of Bioresource Technology, Vol. 99, pp. 5981-5985.

Jorstad, L. B., J. Jankowski, R. I. Acworth, (2004), "Analysis of the distribution of inorganic constituents in a landfill leachate-contaminated aquifer Astrolabe Park, Sydney, Australia" Journal of Environmental Geology, Vol. 46, No. 2, pp. 263-272.

Lawley, D. N., Maxwell, A. E., (1971), "Factor Analysis as a Statistical Method", Second ed, London, Butterworths, pp. 154, 228-235.

Lu, D., Mausel, P., Brondizio, E., Moran, E., (2004), "Change detection techniques" International Journal of Remote Sensing, Vol. 25, no. 12, pp. 2365-2407.

Macleod R. S., Congalton R. G., (1998), "A Quantitative Comparison of Change Detection Algorithms for Monitoring Eelegrass from Remotely Sensed Data" Photogrammetric and Remote Sensing Environment., Vol. 64, pp. 207-216.

McAdams, M., Demirci, A., (2006), "The use of Principle Component Analysis in data reduction for GIS Analysis of water quality data" Geography Department, Fatih University, Istanbul, Turkey, Retrieved March 10, 2009 from http://www.fatih.edu.tr/~mcadams/pub/PCA_GIS_9.pdf.

McBean, E. A., Rovers, F. A., Farquhar G. J., (1995), "Solid Waste Landfill Engineering and Design" Prentice Hall PTR, Enlewood Cliffs, New Jersey.

Reinhart, D. R., Grosh, C. J., (1997), "Analysis of Florida MSW Landfill Leachate Quality" Civil and Environmental Engineering Department, University of Central Florida, Report no. 97-3.

Rietveld, T., Van Hout, R., (1993), "Statistical techniques for the study of language and language behaviour" Linguistics-Statistical Method, Berlin (New York): Mouton de Gruyter.

RUNCO Environmental, Inc., (2009), Landfill Cross Section. Retrieved April 16, 2009 from http://runcoenv.com/landfill.htm.

Sartaj, M., Fernandes, L., (2005), "Adsorption of boron from landfill leachate by peat and the effect of environmental factors" Journal of Environmental Engineering Science, NRC Canada, Issue 4, pp. 19-28.

Singh, A., (1989), "Digital Change Detection Techniques Using Remotely Sensed-Data" International Journal of Remote Sensing, Vol. 10, no. 6, pp. 989-1003.

Suh, Y. C., (2008), "Analysis of Environmental Change Using GIS" Pukyong National University, Pusan, Korea. Retrieved July 20, 2009 from http://www.pul2008.ul.pt/documentos_pul/Programa_links/abstract_yong_cheol_suh.pdf.

Tabachnick, B. G., Fidell, L. S., (2001), "Using Multivariate Statistics," Prentice Hall: Allyn & Bacon.

Trail Landfill Optimization/Expansion Project, (2001), Draft EA/EPA Document, City of Ottawa, Ottawa, ON, Canada.

Trail Road Landfill Asset Management Study: Landfill Optimization Report, (1998), Final draft for the Regional Municipality of Ottawa-Carleton, Ottawa, J.L. Richards & Associates Limited.

University of Ottawa, (2003), Map Library, Retrieved August 15, 2008 from www.biblio.uottawa.ca/map/index-e.html.

Wang, K. L., Hung Y. T., Lo H. H., (2004), "Handbook of Hazardous Industrial Waste Treatment" Second Edition, pp.1156-1157.

APPENDIX A

DATA PROCESSING

Data processing is an important task prior to spatial and temporal data analysis. Different sources of data are used in data preparation to create the data required for spatial-temporal analysis. Such data processing include, changing data formats, applying map projections, georeferencing the layers, and/or retrieving data from scanned maps.

In this study, GIS is mainly used for preprocessing (data preparation) and post processing for analysis, display and visualization. This appendix indicates generally the applied types of data processing (Figure A.1). Similar data processing may be required for similar studies related to environmental pollution control in groundwater environment.

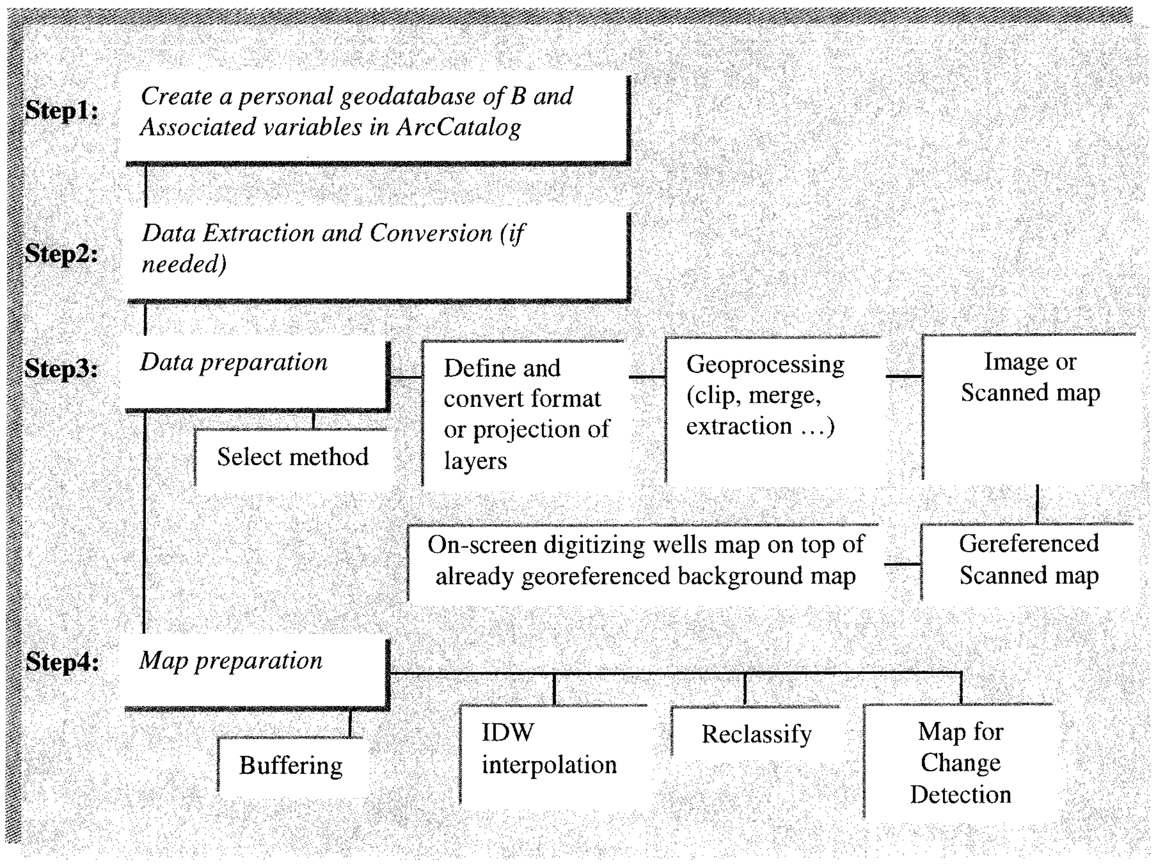


Figure A.1 Flowchart of applied data processing

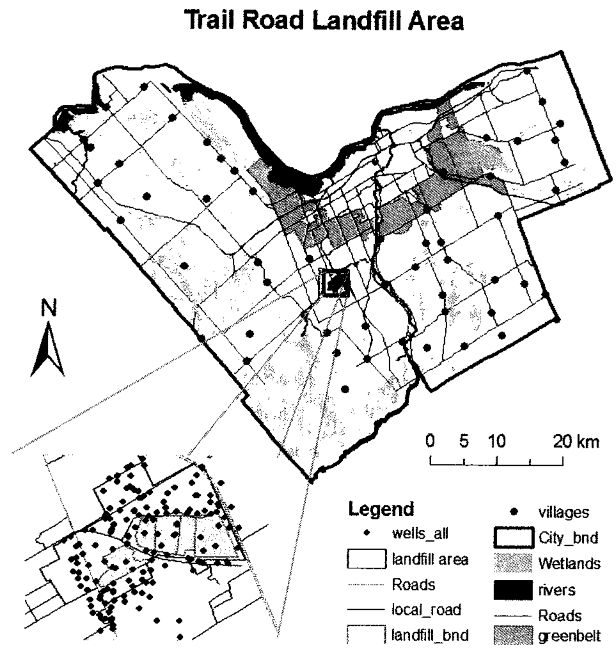


Figure A.2 Projections of shapefiles in ArcMap

For all data layers a common projection and coordinate system is adopted (UTM zone 18N with NAD83 as the datum) (Figure A.2).

In this study local scale spatial data layers of landfill boundary, landfill area boundary and local roads are unavailable. Due to unavailability of the data layers some data have been created by first scanning from available printed reports. Georeferencing technique has been applied on scanned maps that transform the unknown coordinate system with a known coordinate system so that all the layers have similar coordinate system. The georeferenced background maps are then the basis for on-screen digitizing to create the proper GIS layers for the area of the landfill site.

Since the extent of the collected data layers exceed the boundary of the study area (Figure A.3), data clipping is required.

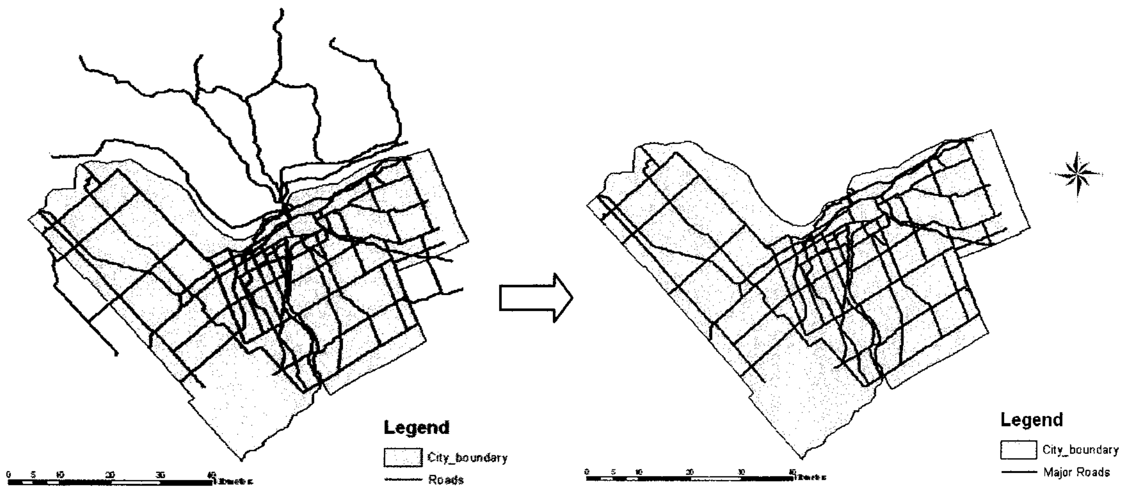


Figure A.3 Clipping "Roads" layer (before and after)

Also for some other adjacent datasets merging is required and applied (e.g. "wetlands" and "water bodies" shapefile layers) to convert them into a single layer.

IDW method is applied for interpolating numerical variables discussed in this study. IDW is a method of interpolation that estimated cell values by averaging the values of sample data points in the neighbourhood of each processing cell. It estimates values for location with no data based on known values which provides understanding of spatial behaviour of phenomena. The closer a point is to the center of the cell being estimated, the more influence, or weight is counted. This method assumes that the variable being mapped decreases in influence with distance from its sampled location. The output value for a cell using IDW is limited to the range of the values used to interpolate. Because IDW is a weighted distance average, the average cannot be greater than the highest or less than the lowest input. The influence of an input point on an interpolated value is isotropic.

Output rasters has been reclassified to convert the continuous range of interpolated values into intervals or classes. Natural Breaks reclassification method has been applied to specify the boundaries between interval classes of numerical variables applied in this study (Figure A.4).

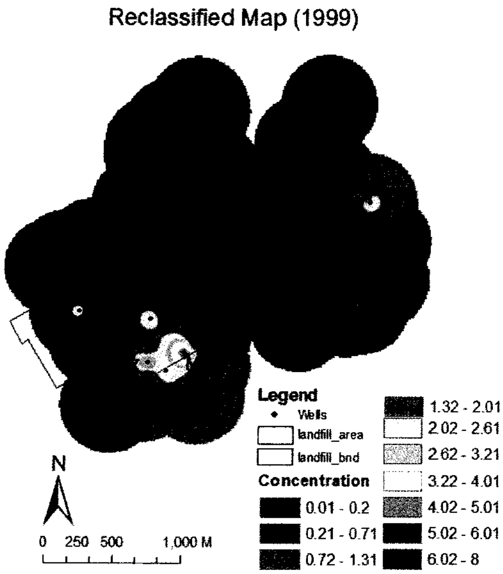


Figure A.4 Reclassification of the interpolated map of B