DEVELOPMENT AND EVALUATION OF MODELS FOR ASSESSING GEOCHEMICAL POLLUTION SOURCES WITH MULTIPLE REACTIVE CHEMICAL SPECIES FOR SUSTAINABLE USE OF AQUIFER SYSTEMS: SOURCE CHARACTERIZATION AND MONITORING NETWORK DESIGN

A thesis submitted by

Michael Saah Hayford, B.Sc. Hon, MPhil

This dissertation is submitted in for the degree of

Doctor of Philosophy (PhD), College of Science and Engineering

James Cook University, Australia

2020

STATEMENT OF ACCESS

I, the undersigned, author of this work, understand that James Cook University will make this thesis available for use within the University Library and, via the Australian Digital Theses network, for use elsewhere.

I understand that, as an unpublished work, a thesis has significant protection under the Copyright Act and I wish this work to be embargoed until June 2020, after which I do not wish to place any further restriction on access to this work.

Signature

Date

STATEMENT OF SOURCES

I herewith declare that I have produced this thesis without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This thesis has not previously been presented in identical or similar form to any other Australian or foreign examination board.

Signature

Date

ELECTRONIC COPY

I, the undersigned, the author of this work, declare that the electronic copy of this thesis provided to James Cook University Library is an accurate copy of the print thesis submitted.

Signature

Date

DECLARATION

I declare that this thesis is the result of my own work and includes nothing that is the outcome of work done in collaboration, except where specifically indicated in the text. It has not been previously submitted, in part or whole, to any university or institution for any degree, diploma or other qualification.

Signed:_____

Date:_____

Michael S Hayford

ACKNOWLEDGEMENTS

First off, I want to thank God because that's who I look up to. He has graced my life with opportunities that I know are not of my hand or any other human hand. He has shown me that it's a scientific fact that gratitude reciprocates. (M.M. 2014)

I want to sincerely thank my supervisor, Dr Bithin Datta, who has been a tremendous support to me. His patience, understanding, supervision, encouragement and, ultimately, the PUSH he gave me when I was at my wits' end, has made this thesis possible.

I am also thankful to CRC-CARE, Graduate Research School and College of Science and Engineering of James Cook University for funding my research and providing substantial support for attending several conferences during my study.

I want to sincerely thank Melissa Norton and Jodie Wilson, who made my PhD journey very smooth by resolving the many difficulties I encountered over the years.

I also thank Dr T. T. Akiti, GAEC and the staff of NCERC for their assistance over all these years.

I must also thank the staff of the College of Science and Engineering for their support. To the PhD students of CSE, it was nice knowing you all and you made my stay at JCU a memorable one. Nasif, Korah, Mehdi, Alvin, Raymond, Prince and Peter: thanks for the friendship we shared.

I want to thank Alex Efa, Maabee, Radar and Augustine Afful for the love you showed me when I was unwell.

I thank my wife and friend, Ekua, for her unconditional love and support, for which I am always and forever indebted. Thanks also to Babs, KB, my parents, my siblings and the entire Hayford family for bearing with me after all these years of being away from home. I love you guys.

DEDICATION

This is dedicated to all abstract thinkers. Don't give up on your ideas……

STATEMENT OF CONTRIBUTION OF OTHERS

Dr Bithin Datta provided overall guidance for this thesis and supervised the entire PhD project.

CRC-CARE, Australia provided financial support for this research through Project No. 5.6.0.3.09/10(2.6.03), CRC-CARE-Bithin Datta, JCU. (Dr Bithin Datta's research project).

The Graduate School, James Cook University, provided a Graduate Research School Scholarship (JCUPRS) for this research.

The College of Science and Engineering, JCU supported this research through a college-sponsored scholarship in 2017.

My supervisor, Dr Bithin Datta, provided editorial support and suggestions for writing this thesis.

The professional editor Dr John Gibbens edited the thesis.

Except for the cases stated above, I performed the numerical simulations, analytical work, editorials, optimizations, model design and implementation.

Abstract

Groundwater pollution is a major concern as it is related to the health of humans and the environment. Groundwater pollution is a major concern, as such pollution reduces groundwater quality and increases pollutant migration. In mining areas, groundwater may be contaminated by several substances due to the geochemical reactions and acid mine drainage typical of abandoned mine sites. Hence, identifying the specific sources of specific contaminants is a complex issue.

The most important first step in the management and remediation of contaminated groundwater aquifers is to characterise unknown contaminant sources (source identification). Often, the hydrogeological field data available for use in source identification is very sparse. In addition, hydrogeological and geochemical parameter estimates and field measurements are often uncertain. In complex contaminated sites like abandoned mines, the geochemical processes are very complex. Therefore, characterizing unknown contamination sources in terms of location, magnitude and timing and determining contaminant pathways can be very difficult. The chemically-reactive nature of contaminant species contributes to the complexity of modelling geochemical transport processes. Similarly, source identification inverse problems are often non-unique and ill-posed. Unlike conventional groundwater contaminant transport simulations, which usually provide stable and well-behaved solutions, inverse groundwater contaminant transport problems, such as those involving unknown groundwater source identification problems, may result in non-uniqueness, non-existence and instability. There may also be increased computational challenges due to a paucity of data.

Another related step is the optimal design of monitoring networks based on initial arbitrary well locations to accurately estimate contaminant source characteristics. Since the accuracy and reliability of the source characterization process largely depends on the use of suitable monitoring locations and spatiotemporal groundwater contaminant concentration data, it is essential to implement efficient monitoring networks that provide optimal measurements under uncertain conditions. This is mostly the case when concentration measurements are sparse and taken at arbitrary monitoring locations. When monitoring networks are optimally designed,

accurate and reliable groundwater data can be obtained. This improves the identification of unknown contaminants in the source characterization process.

Firstly, this thesis proposes a three-dimensional groundwater flow and reactive transport model of a complex, heterogeneous, contaminated, mine site aquifer. Equilibrium and kinetic reactions are incorporated into the model to determine the ongoing acid mine drainage and geochemical interactions occurring at the site. The proposed model has the capability to predict the pathways of contamination by multiple species. Comprehensive calibration and validation of this model are performed. The model is representative of the current situation at the mine site. The proposed model can be applied to other complex contaminated aquifers at a regional scale, with transport problems controlled by several reactions, not limited to the chemical processes of aqueous complexation, precipitation-dissolution, adsorption-desorption, ion-exchange, redox and acid-base reactions, and mixed equilibrium and kinetic reactions. The numerical results demonstrate the applicability and limitations of the proposed model.

A new optimization formulation for solving the objective functions of several individual species from distributed sources is described. Five distinct species undergoing geochemical processes are characterised in this thesis. A linked simulation-optimization approach is adapted to characterize unknown groundwater contaminant sources. Numerical flow and reactive transport models are coupled to an optimization model and solved iteratively to obtain optimal source characteristics. The complexity of characterizing multiple species from distributed sources is handled using a robust optimization algorithm (adaptive simulated annealing, ASA) that provides global optimum solutions for the source characterization of individual species simultaneously by collectively solving nonlinear objective functions. The ASA algorithm deals with the fitting of source parameters and retains the nonlinearities inherent in individual species optimization models used to solve complex source identification problems.

Multifractal modelling is employed to examine its scalable ability to model groundwater contamination. It provides potential spatial and temporal dependence in groundwater monitoring networks. Fractal modelling is used to map the contaminated zones in the study area. Fractal mapping provides an adequate description of contamination zones, which can be categorized as impacted, mildly impacted and un-

impacted. Potential locations for monitoring networks suggested by the fractal model are used as input for the implementation of monitoring network design. The potential locations suggested by fractal modelling and arbitrary monitoring locations are optimized. The ASA algorithm is used to solve a multi-objective optimization model based on the arbitrary monitoring network and that of monitoring network by the fractal model. The influence of fractal modelling is determined by error analysis. Concentration data from arbitrary networks and optimal networks are used as inputs to identify unknown sources. A multi-objective optimization formulation is utilized to determine monitoring networks that are optimal for improving the accuracy of contaminant source estimates. Pareto-optimal solutions are obtained, which demonstrate enhanced source identification accuracy and reliability when concentration data from an optimally-designed monitoring network is used as input for unknown groundwater contaminant source identification.

Finally, the proposed approaches are implemented to solve typical problems in contaminant transport. The performance evaluations of the developed methodologies is evaluated for several complex groundwater pollution scenarios involving distributed mineral waste deposits that are sources of reactive chemical species at a former mine site. These methodologies are also applied to a real-life contaminated aquifer to demonstrate their potential applicability. The performance of the linked source characterization model is also evaluated by applying it to a real case study—a complex, abandoned mine site in the Northern Territory, Australia. Extensive performance evaluation results demonstrate the accuracy and applicability of the source characterization and monitoring network design techniques developed in this thesis.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AMD	Acid Mine drainage
AHD	Australian Height Datum
Aq	aqueous
ARD	Acid rock drainage
ASA	Adaptive Simulated Annealing
Cu	Copper
EBFR	East Branch of the Finniss River
EFDC	East Finniss Diversion Channel
FDM	Finite difference method
Fe	Iron
FEM	Finite Element Method
H	Hydrogen
Head	hydraulic head (or water level)
K	Hydraulic conductivity
KH	Horizontal hydraulic conductivity
KV	Vertical hydraulic conductivity
Mn	Manganese
NT	Northern Territory
OH	hydroxide
PAEE	percent average estimation error
RGC	Robertson GeoConsultants Inc.
ASA	Adaptive Simulated Annealing
SO4	Sulphate
UO_2^{2+}	Uranium
WRD	wastes rock dump

List of Notations

θ =	effective moisture content (L^3/L^3)
h =	pressure head (L)
t =	time (T)
z =	potential head (L)
q =	source or sink representing the artificial injection or withdrawal of fluid [($L^3/L^{3)/}$T]
ρ_o =	referenced fluid density at zero biogeochemical concentration (M/L^3)
ρ =	fluid density with dissolved biogeochemical concentrations (M/L^3)
ρ^* =	fluid density of either injection (= ρ^*) or withdraw (= ρ) (M/L^3)
μo =	fluid dynamic viscosity at zero biogeochemical concentration [(M/L)/T]
μ =	fluid dynamic viscosity with dissolved biogeochemical concentrations [(M/L)/T]
α ,=	modified compressibility of the soil matrix (1/L)
β ‘ =	modified compressibility of the liquid (1/L)
ne =	effective porosity (L^3/L^3)
S =	degree of effective saturation of water
g =	gravity (L/T^2)
k =	permeability tensor (L^2)
ks =	saturated permeability tensor (L^2)
Kso =	referenced saturated hydraulic conductivity tensor (L/T)

kr = relative permeability or relative hydraulic conductivity

ν = material volume containing constant amount of media (L^3)

Ci = concentration of the i-th species in mole per unit fluid volume (M/L^3)

Γ = surface enclosing the material volume ν (L^2)

n = outward unit vector normal to the surface Γ

Ji = the surface flux of the i-th species due to dispersion and diffusion with respect to relative fluid velocity [$(M/T)/L^2$]

θri = production rate of the i-th species per unit medium volume due to all biogeochemical reactions [$(M/L^3)/T$]

Mi = external source/sink rate of the i-th species per unit medium volume[$(M/L^3)/T$]

M = number of biogeochemical species

Vi = transporting velocity relative to the solid of the i-th biogeochemical species (L/T)

$\hat{C}_i$ = chemical formula of the ith species

μik = reaction stoichiometry of the ith species in the kth reaction associated with the reactants

νik = reaction stoichiometry of the ith species in the kth reaction associated with the products

{N} = {1, 2, …, N} in which N is the number of reactions

R_k = rate of the *k*th kinetic reaction

A_i = activity of the *i*th species

N_E = number of linearly independent equilibrium reactions.

R_k rate of the *k*th kinetic reaction

R_E rate of the *k*th equilibrium reaction

T_0 = appropriate initial temperature

T_E = a smaller value is selected as the end

$n_{ob:}$ = total number of concentration observation location

nk = total number of concentration observation time periods

nspe= total number of species involved

a = a fraction such that 0 < a < 1

w_{iob}^{k} = weight corresponding to observation location iob, and the time period k.

η = an appropriate constant which is derived as the average of the highest and lowest concentration values of each species or an appropriate constant so that errors at low concentrations do not dominate the solution.

εr = random error term.

$\overline{C}_o$ = actual input concentration of the species observed

$\widetilde{C}_o$ = concentration of the species generated by the optimization model

$pertC_{sim_{iob}}^{k}$ = perturbation numerically simulated concentration;

$C_{sim_{iob}}^{k}$ = the numerically simulated concentration;

ε = the error matrix with normally distributed error terms with a zero mean and one as standard deviation;

$Cobs\,(M_L, t)$ = measured or observed concentration of species at monitoring location M_L at time t;

$Csim\,(M_L, t)$ = simulated concentration at location ML, at time t from the numerical simulation model;

$(f_{i,j})$ =	binary variable represents a yes or no decision on a potential location
K_K^f and K_K^b =	activity-based forward and backward rate constant of the *k*th kinetic reaction
K_K^e =	equilibrium constant of the kth reaction, A_i is the activities of the *i*th species
$C_{sp\ obs_{iob}}^{\ k}$ =	is observed concentration of species at monitoring location iob and kth time period
$C_{sp\ sim_{iob}}^{\ k}$ =	is concentration of species estimated by the source identification model at monitoring location iob and kth time period
Sp1, sp2, sp3, sp4 =	species numbers one, two , three and four respectively involved in the chemical reaction
$f\ (x, y, z, C_{sim})$=	estimated concentration obtained from the transport simulation model at an observation location and source concentrations C_{sim}.
.INTRP (*f* (*i,j*)*C*(*i,j*)) =	the spatially kriged interpolated values of the concentrations at node i.j, interpolated based on the designated monitoring locations given by $f_{i.j}$=1

Statement of Original Authorship

The work contained in this thesis has not been previously submitted to meet the requirements of any award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

Signature: _____

Date: _____

Chapter 1: Introduction

Groundwater is subject to contamination from a variety of sources. Identifying these sources is a critical step in effectively remediating and preserving groundwater. The process of identifying contaminant source characteristics is the predominant procedure required to implement groundwater management strategies. When pollutant concentrations are measured at monitoring locations, the data can be practically and effectively used for estimating unknown pollutant source characteristics, namely, their location, magnitude and duration of activity. To improve the estimation of pollution source characteristics, implementation of a monitoring network that is designed according to optimal monitoring locations is necessary.

The present study discusses two main techniques for determining groundwater contaminant characteristics: 1) the linked simulation-optimization approach and 2) monitoring network design methodologies. In the linked simulation optimization method, an optimization algorithm is used to obtain optimal solutions for source characteristics. In the monitoring network design methodology, multifractal modelling is used to facilitate optimal monitoring network designs that provide improved, efficient, and reliable determination of contaminant source characteristics. Performance evaluation of these developed methodologies is conducted for groundwater contamination occurring in geochemically reactive aquifer systems. The applicability of the developed methodologies is demonstrated using real-life scenarios of contaminated aquifers with several distributed contaminant sources.

Globally, groundwater is the most-extracted natural resource, with a withdrawal rate of 982 km^3 per year. Many countries depend on withdrawn groundwater for domestic supplies. In Australia, groundwater constitutes 30% of overall water usage. In regions where evaporation is greater than rainfall, groundwater is the most common water supply. Annually, approximately 3500 GL of groundwater is used for various purposes. Economic sectors depend heavily on the use of groundwater and have the highest usage percentage, while a smaller usage percentage is attributed to domestic activities. The agricultural sector accounts for nearly 60% usage, while mining and manufacturing industries account for 12% and 17%, respectively. A meagre amount of 5% is used for household water supply and 9% is used as injection into portable

water supply systems. As a country's population gradually expands, the demand for groundwater also increases. The effects of a growing population and increasing anthropogenic activity include increasing demands for groundwater resources. Increasing pressures from various manmade activities have led to reductions in and mismanagement of groundwater resources; hence, there is a need to protect and manage this valuable resource. The map of Australia in Figure 1.1 describes the percentages of the populations of various regions that are dependent on groundwater.

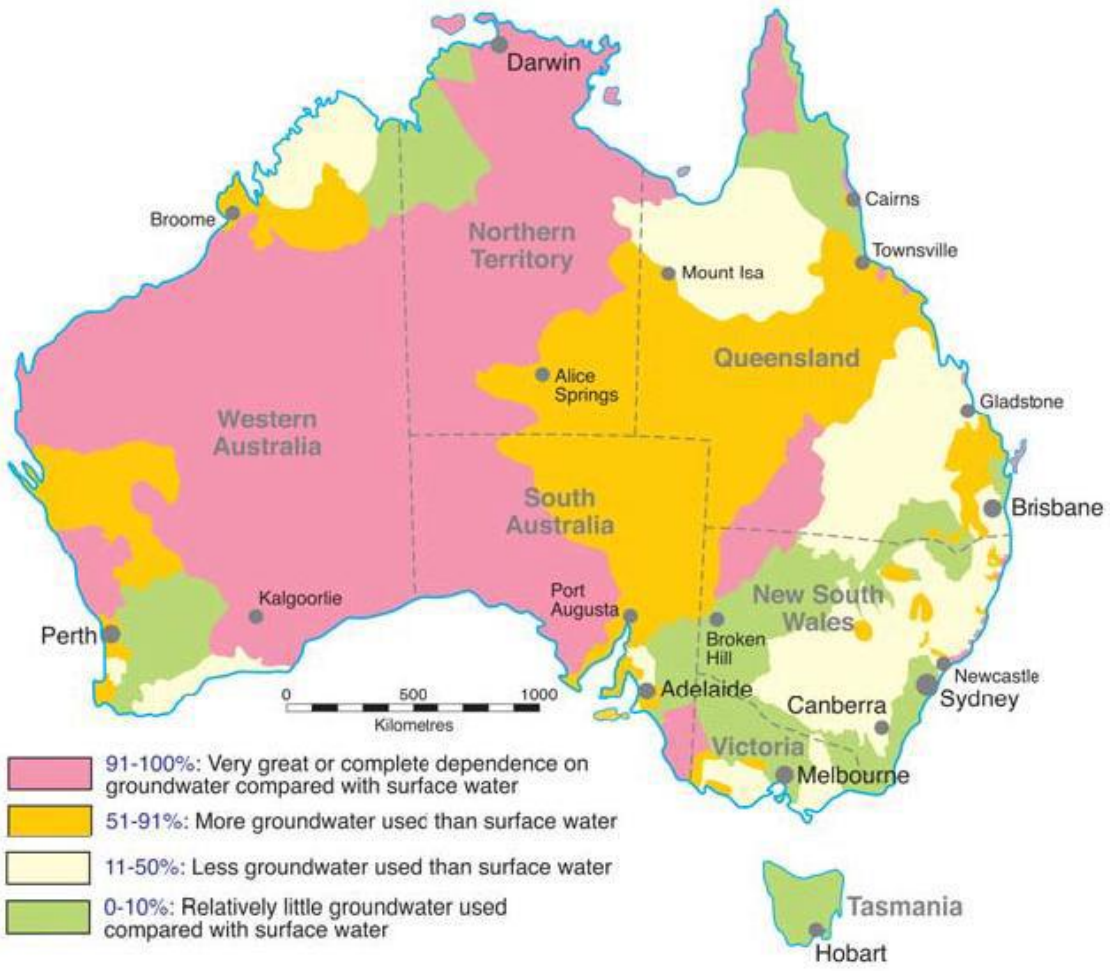


Figure 1.1: Map of Australia indicating the dependence percentages of regional populations that are reliant on groundwater. Credit: NCGRT, 2013.

Unlike surface water, for which obvious measures can be taken to keep people from contaminating it, groundwater has no direct preventive measures. Contaminants

infiltrate aquifer systems after travelling through layers of soil, sediment and rocks. Once groundwater is polluted, the pollution may go unnoticed for years or even decades. There are two main causes of groundwater contamination: naturally-occurring contaminants and anthropogenic contaminants.

Natural contaminants, by nature, exist in soils and geological formations. These include nutrients like phosphate and ammonia, hydrogen sulphide from decomposed organic materials, and radionuclides and radioactive emissions from the geological matrix. The elements iron and manganese also occur naturally in many groundwaters. Anthropogenic contaminants mostly originate from human activities that affect the environment. Such contaminants can be either organic, inorganic or radioactive. Chemicals used in mining, farming, industrial manufacturing and domestic activities have the possibility to contaminate the natural environment by dissolution through the aqueous phase. Anthropogenic contaminants can also enter the groundwater system through leakage from landfills, tailing storage facility, pipes, mine waste damp, and oil pipes.

In Australia, one of the common modes of anthropogenic contamination is heavy metal pollution from waste rock dumps and mining tailings. Australia has a very large and diverse mining sector, which contributes to around 8% of the GDP and about 60% of exports. The mining sector has been a driving force for much of the exploration and extraction of Australia's precious minerals and the development of associated industries. With the high amount of mining activity, there is a need to ensure that it does not leave a lasting impact on groundwater resources, including contamination. This research work focuses on groundwater contamination from mining waste, particularly from abandoned mine sites. This category of groundwater contamination is generally termed *acid mine drainage* (AMD) and is a common source of contamination at many mine sites across the world. Acid mine drainage is a prominent problem with most hard rock and metal sulphide mines. While this research focuses on groundwater contamination from distributed sources such as waste rock dumps, open mine pit ponds/lakes and mine tailings, even rocks in undisturbed environments can cause similar groundwater contamination by acid rock drainage without any anthropogenic influence.

Acid mine drainage is a severe environmental pollution source. It occurs when acids and metals leach from waste rock into the environment and pollute the

groundwater system. This is a result of the oxidation of pyrite, which is mostly found in combination with metal minerals. Since mine tailings and waste rocks have a high surface area due to their small grain size, these mine wastes are more prone to generating AMD, thereby discharging acids and metals into groundwater systems. Metals are dissolved by acid drainage and react with other metals as they move through the groundwater system, producing daughter species that can cause complex contamination. The processes of AMD in the subsurface are complex and involve several geochemical reactive processes. To understand these subsurface processes, the application of geochemical reactive modelling can be an important part of hydrogeological and geochemical investigations at such mine sites.

Stopping the ongoing and increasing rates of acid mine drainage generated by mine waste is critical in protecting groundwater resources. The principal step necessary for protecting and managing contaminated aquifers is identifying the contamination source characteristics and predicting current and future groundwater flows and contaminant transport. When a contaminant is detected, efforts must be made to identify the source and monitor it, then develop a remediation and clean up strategy for the groundwater system. Clean-up is capital-intensive and time-consuming and may span a lengthy period. Considering how much groundwater is extensively polluted, there is an urgent need for efficient procedures for preventing, detecting and remediating groundwater contamination. Groundwater pollutants can stay undetected for several years after their sources become active. Identification of contaminant source characteristics is therefore essential in solving source identification problems in complex conditions. This can provide a platform for effective groundwater pollution prevention and remediation.

1.1 UNKNOWN GROUNDWATER CONTAMINANT SOURCE CHARACTERIZATION

In solving groundwater contaminant source identification problem, understanding the contaminant characteristics is important. Contaminant sources can be characterized in terms of source numbers, strengths, locations and release duration. Often, sources locations and characteristics are unidentified as released information is non-existent and the source may have long ceased to exist. What may be available is information on the past distribution of a contaminant at one or more arbitrary locations within a

contaminated aquifer. This may be critical when determining the source characteristics and plume distribution from limited data. Several challenges impede the source identification process, including uncertainties in the prediction of source flux in the aquifer, insufficient field concentration data and the aquifer's responses to hydraulic and geochemical stresses (Datta et al., 2013). The process of simulating groundwater contaminant transport is done using mathematical models that solve the flow and transport processes that constitute advection, dispersion and chemical reaction. The process is a forward type of modelling used when hydraulic parameters and source flux values are known. Their solutions for the spatial and temporal distributions of contaminant concentrations tend to be uniquely and properly constrained. The solutions for these simulations can be used to predict the flow and transport of contaminants in an aquifer. However, in source identification problems where several unknowns need to be identified, groundwater contaminant transport simulations are no longer a forward modelling process because a straight-forward solution is impossible.

In contrast to stable and definitive models of groundwater contaminant transport processes, models of unknown groundwater contaminant problems are more non-unique, non-existent and unstable (Sun, 1996). The problem of characterizing a contaminant source is an inverse problem in groundwater contaminant transport. Therefore, when inverse modelling is done, because of insufficient field data and the ill-behaved nature of the problem, the process of modelling unknown contaminant sources becomes much more challenging and difficult. Various classifications of the inverse problem exist in the literature, which include identification of contaminant source locations (Gorelick et al., 1983; Wagner, 1992; Mahar et al., 2000; Neupauer et al., 2000; Neupauer & Wilson, 2005), estimation of the number of possible pollution sources (Ayvaz, 2010), estimation of release history (Skaggs & Kabala, 1994, 1995; Alapati & Kabala, 2000; Sidauruk et al., 1998) and reconstruction of the historical distribution of a contaminant (Michalak & Kitanidis, 2004; Bagtzoglou & Atmadja, 2003).

Many researchers have studied the area of groundwater systems optimization over the years (Gorelick, 1983). Most researchers in the field of groundwater optimization studies have addressed the issues of remediation design (Mantoglou & Kourakos 2007; Zheng & Wang 1999), identification of pollution sources (Aral et al., 2001; Bagtzoglou et al., 1991, 1992, 2005, 2014; Singh et al., 2004) and saltwater

intrusion (Abarca et al., 2006; Bhattacharjya & Datta, 2005, 2009; Cheng et al., 2000). The linked simulation-optimization approach used over the years has been practical and efficient in solving contaminant source identification problems (Datta et al., 2009, 2017; Esfahani & Datta, 2015, 2016; Harrouni et al., 1996, 1997; Kitanidis, 2004; Katsifarakis et al., 1999; Lesnic et al., 1998; Michalak & Ayvaz, 2010; Prakash & Datta, 2014; Singh & Datta, 2004; Wagner, 1992).

In the linked simulation-optimization approach, a simulation model representing aquifer processes is coupled with an optimization model to solve the unknown groundwater identification problem. Early linked simulation-optimization approaches solved the inverse problem by utilizing linear programming and response matrixes together with forward simulation. In recent times, advanced algorithms such as genetic algorithms and adaptive simulated annealing have been incorporated in the linked simulation-optimization methodology for solving the inverse problem. These emerging evolutionary algorithms are able to solve source characterization problems by reaching global optimal solutions, making them convenient and reliable for use with the linked simulation-optimization approach. A linked optimization simulation approach is typically feasible and effective for solving unknown source identification problems.

1.2 MONITORING NETWORK DESIGN DEDICATED TO SOURCE CHARACTERIZATION

Collecting groundwater data from monitoring points is essential to contamination assessment and groundwater contamination mitigation. The decision to place monitoring wells at the right locations within the right time frame to collect groundwater data is important yet a challenging task. This is so because groundwater monitoring is an expensive task. The primary objectives of a monitoring network design are: minimizing the contaminated area, minimizing the total cost of the monitoring system and maximizing detection probability. When these objectives are met, the network has reliable capabilities in detecting contaminant concentrations and tracking contaminant movements and plume coverages before they exceed acceptable limits. Designed monitoring networks prevent severe groundwater contamination of the environment at the same time as optimizing the amount of concentration data obtained. When designed monitoring networks are used in source characterization,

these concentration data improve the accuracy and efficiency of identifying unknown groundwater contamination sources.

Groundwater data collection from monitoring wells is important for evaluation of the impacts of resource development. Knowing the time and location to position monitoring wells is important for groundwater administrators, mainly because groundwater monitoring is expensive. Numerical modelling can provide predictive data regarding possible water quality changes stemming from resource development. However, there is often high uncertainty in these predictions, which makes it challenging to undertake investment decisions. In this study, we propose a method for optimizing monitoring network designs in which many plausible groundwater quality outcomes are considered and synthesized to help in the design of the monitoring network.

The most important step in contaminated aquifer remediation and groundwater resource management is effective source characterization. Effective source characterization requires concentration data to be measured across the contaminated site. In real-life situations, several factors, such as incomplete concentration data and limited monitoring wells, makes it challenging to obtain enough data to reflect the true state of the contaminated site. In view of such constraints and limitations, new approaches are needed to optimally design effective monitoring networks that provide reliable contaminant concentration measurements that can improve the accuracy of source identification (Datta & Dhiman, 1996). Prakash and Datta (2015) implemented a sequential monitoring network design that provides better identification of source characteristics. One of the important factors associated with the optimal design of a monitoring network is the initial arrangement of the potential monitoring well locations. In this study, multifractal modelling is utilized to obtain potential well locations based on aerial mapping of contaminated and uncontaminated zones. The maps delineate the scope of the contamination zones by identifying contamination boundaries. With the contamination boundary information obtained from multifractal modelling, identification of potential monitoring wells is more achievable. The locations of these potential monitoring wells, as well as the concentrations measured there, are used as input to an optimal design model. Multi-objective optimization is used to solve the optimal monitoring network design model.

The proposed methodology solves a two-objective optimal monitoring network design model utilizing a C++ based adaptive simulated annealing algorithm. The optimization model is linked to a calibrated groundwater simulation model. Even with constraints on the number of monitoring locations, the designed monitoring network improves the source identification results. The optimal monitoring locations obtained decrease the likelihood of missing the effects of potential contaminant sources on concentration measurements. A designed monitoring network can reduces the extent of non-uniqueness in the measured set of possible aquifer responses to corresponding geochemical stresses, i.e contaminant sources. Sequential designed and implemented optimal monitoring networks can provide relevant feedback for identifying source locations, and therefore, can facilitate efficient characterization of groundwater pollution sources.

1.3 RESEARCH OBJECTIVES

The aim of this research study is to develop methods based on optimization to characterise unknown groundwater contamination sources in a hydrogeologically and geochemically complex heterogeneous anisotropic aquifer with reactive multiple species transport processes.

In previous studies, researchers have typically studied point or ideal-shaped non-point contaminant sources in solving source identification problems, especially when utilizing the optimization approach. Although many groundwater contamination processes can be simulated by considering these kinds of sources, they cannot be used to determine the influences of distributed areal groundwater contamination sources in real life. This is an important problem, particularly in mining sites where groundwater contamination from mining activities affects human health, vegetation and water resource systems. Therefore, optimally identifying distributed areal groundwater contamination sources and their associated individual multiple reactive species is a necessary step before developing a remediation strategy. The objective of this study is to propose a linked simulation-optimization approach to solve groundwater contamination source identification problems involving distributed sources and consideration of individual contaminants.

Contaminant source characterization/identification in complex contaminated sites, like operational or former mine sites, is made even more complex by the distributed nature of the sources, the presence of multiple reactive chemical species, and varying acidity in the environment caused by acid mine drainage. Due to the scale of these studies, the sparsity of available field measurement data and uncertainties related to the nature of chemical reactions and hydrogeological parameter estimates, the source characterization problem is even more complex in highly contaminated mine sites.

When there is little or no information available regarding source locations and pollution magnitudes, characterizing sources becomes very difficult. Having insufficient or limited field data on contaminant concentrations measured at arbitrary monitoring locations increases the difficulty. In the case of real-life source identification problems, where distributed sources are present, there may be multiple sources of distinct contaminant species with different points of entry into the aquifer system and unique species movements. The key aims of this study are to develop methodologies to optimally characterize unknown contaminants from spatially and temporally distributed sources of multiple reactive species, and to improve contaminant source identification using an optimally designed monitoring network in a contaminated mine site aquifer. The performance of the developed methods is evaluated by using a contaminated aquifer site of an abandoned uranium mine site in the Northern Territory of Australia.

1.3.1 Specific objectives

The specific objectives of this work are:

- To calibrate and validate a subsurface, multicomponent, three-dimensional flow and multiple reactive chemical species simulation model of a contaminated aquifer.
- To develop a methodology for characterising unknown reactive species from distributed sources using a linked simulation-optimization approach.
- To design and implement a dedicated monitoring network that enhances the accuracy and efficiency of the characterization of distributed contamination sources.

- To apply the developed methodologies to a real site contaminated by previous mining activities and affected by acid mine drainage .
- To evaluate the performance and practicability of the developed methodology by applying it to a real-life contaminated mine site aquifer controlled by complex geochemical reactions.

1.4 ORGANIZATION OF THE THESIS

The organization of the thesis and an outline of its chapters are presented below:

Chapter 1 provides a brief introduction and background to the problem, and the main objectives of this research.

Chapter 2 presents a detailed review of literature related to source identification problems and monitoring network designs. This chapter distinguishes between the multiple types of source identification problems and presents approaches to solving them. A comprehensive review of the different tools used in such studies is also made. More focused literature reviews relevant to each chapter are also provided at the beginnings of the other chapters.

Chapter 3 presents the development of a three-dimensional multispecies reactive flow and transport model of a geochemically-complex abandoned mine site contaminated by acid mine drainage. The proposed model details all the possible geochemical reactions occurring and how they affect groundwater quality. Details of the calibration and validation procedures, as well as the daughter products of multiple contaminant species, are addressed.

Chapter 4 deals with the design and implementation of efficient source characterisation methodologies that are suitable for use at a complex contaminated mine site. A calibrated flow and transport simulation model that addresses multiple species reactive transport processes for the aquifer site is integrated with a source characterization optimization model. The optimization model consists of an optimization algorithm that seeks to determine a set of source characteristics based on contaminant concentrations simulated (estimated) by the flow and transport simulation model and comparison of these with concentrations measured in the field. In this chapter, a new optimization formulation is used to characterise individual reactive

species undergoing geochemical reactions based on the locations of distributed sources. The performance of the proposed optimization formulation is evaluated using a real contaminated mine site in the Northern Territory, Australia.

Chapter 5 develops a monitoring network design that improves the characterisation of unknown groundwater contaminant sources. This chapter presents a multi-objective optimization method for designing groundwater concentration monitoring networks. The method is useful in designing new optimal monitoring networks, redesigning existing ones, and increasing the extent of existing monitoring networks. Multifractal modelling is used to define potential well locations and contaminant plume boundaries. Combining the results of the multifractal modelling and multi-objective optimization processes yields a Pareto-optimal design for a monitoring network that achieves accurate groundwater contaminant source characterization and cost-effectiveness. The developed methodology is evaluated using contaminant concentration data obtained from the designed network (for performance evaluation purposes only, the concentration data are estimates (simulated) of the calibrated and validated simulation models solved for the source identification optimization model). This procedure is utilized to test the performance of the source identification model and the monitoring network design. It compares source identification results based on measurements from the designed monitoring network, with those based on concentration measurements obtained from an existing arbitrary monitoring network.

Chapter 6 provides a concluding discussion of this study and makes recommendations for further research in this area.

The main literature review for this study is presented in the following chapter.

Chapter 2: Literature Review

This chapter briefly discusses the body of literature relevant to characterising groundwater contaminant sources. The first section of this chapter describes the various approaches that have been used to solve these problems. Additional variations of these problems, involving unknown pollutant source characteristics and source identification, are reviewed. The second section in this chapter recounts the numerous methodologies developed for designing monitoring networks, particularly in the context of pollutant source identification. The final section provides an overview of literature related to the various tools and techniques used for numerical simulation of groundwater flow and transport and for optimization.

2.1 POLLUTANT SOURCE IDENTIFICATION BACKGROUND

In the past two decades, researchers have developed a keen interest in solving the problem of identifying groundwater pollutant sources. Most of these pollutants are noticed at the early stages or even after they cease activity. Pollutant source identification is viewed as an inverse problem. To solve it, questions such as: When was the pollutant released from the source? (release history); Where is the pollutant's source? (source location); and How much pollutant flux was released from the source? (source magnitude) need to be answered. These questions are mostly addressed as estimation functions of the history of contamination. Identification of pollutant source characteristics in terms of magnitude, location and activity duration is described as an inverse problem which is generally ill-posed (Yeh, 1986; Datta, 2002). Several approaches to solving it have been documented (Atmadja & Bagtzoglou, 2001b; Michalak & Kitanidis, 2004a, b; Neupauer et al., 2000; Sun et al., 2006a, b). These methods can be categorized as follows: heat transport inversion, contaminant transport inversion consisting of optimization approaches, probabilistic and geostatistical, analytical solutions, and regression and direct methods (Bagtzoglou &Atmadja, 2005). Comprehensive reviews of these methods are discussed in Atmadja and Bagtzoglou (2001b); Michalak and Kitanidis (2004a, b); Neupauer et al., (2000); and Sun et al. (2006a, b). In general, an inverse problem is considered well-posed if the following

conditions are satisfied: a solution exists, the solution is unique and the solution is stable (Datta, 2002). Additional methods developed over the years can be accessed in Morrison et al. (2000), Chadalavada et al. (2007) and Amirabdollahian and Datta (2013), which contain extensive literature reviews of methods that focus on groundwater contaminant source identification.

2.1.1 Solving the Source Identification Problem

Two different methodologies exist for solving source identification problems. The first approach involves solving differential equations backwards in time (inverse problem) by using techniques that overcome the problems of non-uniqueness and instability. These techniques include the random walk particle method (Bagtzoglou et al., 1991, 1992), the Tikhonov regularization method (Skaggs & Kabala, 1994), the quasi-reversibility technique (Skaggs & Kabala, 1995), the minimum relative entropy method (Woodbury & Ulrych, 1996), Bayesian theory and geostatistical techniques (Snodgrass & Kitanidis, 1997), the adjoint method (Neupauer & Wilson, 1999, Li et al., 2007), the non-linear least-squares method (Alapati & Kabala, 2000), the marching-jury backward beam equation method (Atmadja & Bagtzoglou, 2001) and genetic algorithms (Aral et al., 2001; Mahinthakumar & Sayeed, 2005).

The second approach to solving source identification problems involves a simulation-optimization approach consisting of a forward-time contaminant transport simulation model with an optimization algorithm. The optimization techniques in this category are linear programming and least-squares regression analysis (Gorelick et al., 1983), statistical pattern recognition (Datta et al., 1989) and non-linear maximum-likelihood estimation (Wagner, 1992). This approach avoids the problems of non-uniqueness and stability associated with formally solving inverse problems, but the iterative nature of the simulation models usually requires high computational effort. Mahar and Datta (1997, 2000) used non-linear programming with an embedding method that eliminates the necessity for external simulation since the governing equations of flow and solute transport are directly incorporated in the optimization model as binding constraints. Artificial neural networks (Singh et al., 2004) offer an alternative method of simulation that is computationally effective. Mirghani et al. (2006) proposed a grid-enabled simulation-optimization approach to solving problems that require a large number of model simulations.

2.1.2 Linked Simulation-Optimization Approach to Source Characterization

The evolution of the linked simulation-optimization approach to source characterization began with Datta et al. (1989), who developed an expert-system embedding pattern-recognition technique for pollution-source characterization. The pattern recognition technique was based on stochastic dynamic programming, which minimized losses due to recognition errors. This pattern recognition model was utilized as a screening model for optimization-based source characterization models.

Wagner (1992) developed an optimization-based methodology for simultaneous estimation of model parameters and source characterization. In his work, Wagner used an inverse model as a non-linear maximum-likelihood estimation problem. Estimates of hydrogeological and source parameters were based on measurements of hydraulic head and pollutant concentration.

Mahar and Datta (1997, 2001) were the first to combine optimal source characterization with the design of a groundwater quality monitoring network to improve the efficiency of the source characterization process. An embedding technique was used in which the simulation model was embedded as a binding constraint in the optimization model. They applied their method to a hypothetical 2-D, homogeneous, isotropic and saturated aquifer with a conservative pollutant plume in a two-step process. In the first step, an optimization model was used to identify an unknown pollution source based on observation data. In the second step, different pollutant plumes were simulated using perturbed sources. These realizations of pollutant plumes were used within an integer programming model to determine the optimal locations of monitoring wells. Pollutant concentration measurements from these monitoring wells were used in a non-linear optimization model to obtain more accurate estimates of sources. Mahar and Datta (2000) were also able to estimate the magnitude, location and duration of pollutant sources using a non-linear optimization technique. Datta et al. (2009a) used an optimal dynamic monitoring network design to identify groundwater pollution sources.

Aral et al. (2001) formulated the pollutant source characterization problem as a nonlinear optimization model, in which pollutant source locations and release histories are defined as explicit unknown variables. The optimization model minimizes the residual errors between observed and simulated pollutant concentrations at observation locations. Simulated concentrations are implicitly embedded in the formulation

through the simulation models. Repeatedly solving these models is computationally intensive but is a necessary feature of the optimization process. A progressive genetic algorithm (PGA) was applied to solve the optimization problem.

Singh et al. (2004) and Singh and Datta (2007) used a trained, multilayer, feed-forward artificial neural network (ANN) for the simultaneous estimation of pollution source characteristics and hydrogeological parameters. The universal function approximation capability of the ANN was utilized to estimate the source characteristics and flow and transport parameters. The ANN was trained on data patterns that consisted of a set of source fluxes and corresponding temporally-varying simulated concentration measurements. The methodology was evaluated with varying degrees of concentration measurement error.

Mahinthakumar and Sayeed (2005) investigated and compared several hybrid optimization approaches that combine genetic algorithms with a number of local search approaches for reconstructing the release histories of pollutant sources. The methodology was evaluated for a three-dimensional, heterogeneous flow field considering multiple sources. The results indicate that hybrid optimization methods, which combine an initial global heuristic approach (for example, a genetic algorithm) with a gradient-based local search approach (for example, conjugate gradients), are very effective in estimating flux release histories.

Singh and Datta (2006) used a simulation-optimization approach for characterising unknown groundwater pollution sources. Its performance was evaluated with combinations of source characteristics, data availability conditions and concentration measurement error levels. The main advantage of this method is that the numerical simulation model can be externally linked to the optimization model. This approach solves source characterization problems involving complex aquifers with multiple pollution sources.

Yeh et al. (2006) proposed an approach called SATS-GWT that combines simulated annealing (SA), Tabu Search (TS) and the three-dimensional groundwater flow and solute transport model (MODFLOW-GWT) to estimate the source location, release concentration and release period. The source location is selected by TS within a suspected source area. Then, SA is used to optimally estimate the release concentration and release period. The search for an optimal estimate of these unknown source characteristics is terminated based on the best objective function value. The

performance of this method was evaluated for homogeneous and heterogeneous aquifers with transient flow conditions.

He et al. (2009) presented a coupled simulation-optimization approach for the optimal design of petroleum-contaminated groundwater remediation strategies involving uncertainty. This approach has several advantages as it (1) addresses stochasticity in the modelling parameters used to simulate the flow and transport of non-aqueous phase liquids (NAPLs) in groundwater, (2) provides a direct and rapid - response bridge between remediation strategies (pumping rates) and remediation performance (pollutant concentrations) via proxy models, (3) alleviates the computational cost of searching for optimal solutions, and (4) gives confidence levels for the optimal remediation strategies obtained.

Datta et al. (2009b) developed a methodology for simultaneous source identification and parameter estimation in groundwater systems in which a simulation model is externally linked to a nonlinear optimization model. The simulator defines the flow and transport processes and serves as a binding equality constraint. The search direction is determined by a Jacobian matrix in the nonlinear optimization model, which links the groundwater flow-transport simulator and the optimization method. This method addresses the limitations in embedding discretised flow and transport equations as equality constraints in the optimization process.

Ayvaz (2010) developed a linked simulation-optimization model in which the locations and release histories of pollution sources are treated as explicit decision variables. The MODFLOW and MT3DMS software packages are used to simulate the flow and transport processes in the groundwater system. These models are then integrated with an optimization model based on the heuristic Harmony Search (HS) algorithm.

Datta et al. (2011) developed a methodology for linking a classical nonlinear optimization model to a flow and transport simulation model. The essential link between the simulator and the optimization method is the derivatives or gradient information required for the optimization algorithm. This methodology lacks some of the computational limitations of some earlier developed methodologies. It uses nonlinear programming with flow and transport process governing equations embedded as equality constraints within the optimization model.

Jha and Datta (2011) presented a linked simulation-optimization method for solving source characterization problems. An adaptive simulated annealing (ASA) optimization algorithm provided superior performance compared with similar methods that use GAs. Jha and Datta (2012b) demonstrated the superior performance of ASA over GA in solving groundwater source characterization problems. An overview of different optimization-based methodologies for characterizing unknown groundwater pollution sources is given in Chadalavada et al. (2011b).

Jha and Datta (2013) designed an ASA algorithm-based solution. This method proved to be computationally efficient in the optimal identification of source characteristics, in terms of execution time and accuracy. Although moderate levels of error were found in the estimated parameters and concentration measurements, the method is computationally efficient.

Prakash and Datta (2013) developed a technique that uses a dedicated monitoring network design and implementation process to screen and identify possible source locations. This procedure combines spatial interpolation of contaminant concentration measurements with a simulated annealing optimization algorithm to design an optimal monitoring network. This method has advantages when used as a screening model for subsequent accurate identification of pollution sources in terms of location, magnitude and activity duration.

Prakash and Datta (2013) developed an approach for estimating groundwater pollution concentrations and identifying their sources. This approach combines the spatial interpolation of concentrations with a simulated annealing optimization algorithm to optimally design a monitoring network. The approach also provides information on plausible source locations that can be used in an optimal source characterization model to accurately estimate the locations, magnitudes and durations of pollution activity. This idea can be applied to various groundwater pollution scenarios where pollutants have been detected. The method performs well with very small amounts of initial pollution concentration data. This capability of identifying pollution source locations with very limited data is significant in improving the accuracy of source characterization using a formal approach.

Datta, Prakash, Campbell and Escalada (2013) presented a methodology for improving the accuracy of groundwater pollution source identification using concentration measurements from a heuristically designed optimal monitoring

network. This type of network is constrained by a maximum number of monitoring locations. This methodology uses genetic programming (GP) and linked simulation-optimization to reconstruct the flux histories of unknown conservative pollutant sources based on limited amounts of spatiotemporal pollution concentration data. The significance of this approach was demonstrated by evaluating its performance in an illustrative study area. The results demonstrated its efficiency in source identification when concentration measurements from a monitoring network were used.

Prakash and Datta (2014) used a simulation-optimization model for efficient source identification. This model incorporated a GP-based impact factor that was used to design an optimal monitoring network. Concentration data from the network is then used in an SA-based model.

Prakash and Datta (2014) proposed an alternative optimization-based source identification model to simultaneously estimate source flux release histories and initiation times. These were used as explicit decision variables that were optimally estimated by a decision model.

Amirabdollahian and Datta (2014) proposed an optimal source identification model that incorporates an ASA optimization algorithm with numerical flow and transport simulation models to identify contaminant source characteristics. The fuzzy logic concept was used to identify the effect of hydrogeological parameter uncertainty on groundwater flow and transport simulation. The fuzzy membership values incorporate parameter reliability into the optimization model. The potential applicability of this method was demonstrated by applying it to an illustrative study area. Incorporation of a fuzzy model within a source identification model increased the applicability of the contaminant source detection models in real-life contaminated aquifers.

Jha and Datta (2015b) utilized ASA as an optimization algorithm for distributed source characterization in complex study areas. The optimal solutions obtained show that the linked simulation-optimization-based methodology is potentially applicable to the characterization of spatially-distributed pollutant sources, such as those typically present at abandoned mine sites.

Esfahani and Datta (2016) used GP models as surrogates for reactive species transport in groundwater to achieve optimal reactive contaminant source

characterization. This design reduces the computational stress of running a numerical simulation several times.

Datta et al. (2017) designed a linked simulation-optimization method comprising a reactive transport simulation model (PHT3D) linked to an SA-based optimum decision model. It was used to address difficulties in identifying the potential sources and pathways of multiple species of chemically-reactive contaminants, which can make the source characterization process more challenging.

2.2 MONITORING NETWORK DESIGN FOR UNKNOWN GROUNDWATER CONTAMINANT SOURCE CHARACTERISATION

Monitoring networks are integral to groundwater management. Monitoring network designs may have different objectives, according to the aquifer properties and budgetary constraints. Monitoring networks are essentially installed for extracting information that achieves the underlying monitoring objectives. A large body of literature exists that deals with the design of monitoring networks for different groundwater quality management objectives.

Designing an optimal monitoring network is dependent on various factors such as aquifer parameters, management objectives, and specific constraints and limitations. Some of the objectives of monitoring networks that have been proposed over the last two decades include detection of contamination (Loaiciga et al., 1992; Loaiciga & Hudak, 1993), reducing the cost of groundwater quality monitoring (Loaiciga, 1989; Fethi et al., 1994), multiple objective groundwater monitoring network design (Reed & Minsker, 2004) and source identification and redundancy reduction using feedback information (Dhar & Datta, 2007; Datta et al., 2009). The methods used include sampling strategy in space and time using a Kalman filter (Kollat et al., 2011; Reed & Kollat, 2013), long-term monitoring using a multi-objective simulation-optimization model with uncertainties (Luoa et al., 2016), and optimal contaminant source characterization by integrating sequential-monitoring-network design and source identification (Prakash & Datta, 2015).

Meyer and Brill Jr (1988) presented a methodology for designing an optimal monitoring network that incorporates contaminant movement simulation under uncertain conditions. The Monte Carlo technique was used to apply uncertainty in

transport parameters in the contaminant concentration distribution. In addition, Meyer et al. (1994) developed a method that incorporates conditions of uncertainty in a monitoring network design. They illustrated three objectives: i) minimizing the number of observation wells in the monitoring network, ii) maximizing the probability of detecting a pollutant leak and iii) minimizing the expected area of pollution at the time of detection.

Chandalavada and Datta (2007) extended the optimal monitoring network design methodology to include uncertainty in both source and aquifer parameters. A genetic algorithm and a geostatistical tool, kriging, are used for solving the optimization model and for computing the variances of estimated concentrations at potential monitoring locations, respectively. Their study applied the following two objective functions to obtain an optimal design of monitoring locations: i) minimization of the summation of unmonitored pollutant concentrations at potential locations where the probable concentrations are high and ii) minimization of estimated concentration variances at potential locations.

Dhar and Datta (2007) also developed a method for the optimal design of well locations using optimization models. Based on a transient pollutant transport process, an optimal design of a monitoring network was obtained using dynamic well locations in different time stages. This time-varying network design is less costly than designing all well locations at a single stage. Dhar and Datta (2010) presented a method for optimally minimizing the number of monitoring wells using an inverse distance weighting method and linear optimization model. Chandalavada et al. (2011b) developed a methodology to track pollutant plumes with minimum integrate monitoring wells. This optimized monitoring network design incorporated uncertainty in pollutant concentration estimates. The developed methodology was applied to a polluted defence site in Australia.

Chandalavada et al. (2011) designed an optimal monitoring network to delineate pollutant plumes using a minimum number of monitoring wells. Monitoring wells were installed at locations with minimum measurement uncertainty. The uncertainty in the study area was quantified by analysing variances in concentration estimates at all potential monitoring locations. Jha and Datta (2012) used a dynamic time warping system to estimate the starting time of source activity. This was used to compare observed and simulated pollutant concentrations over time in a linked simulation-

optimization model. An ASA algorithm was used to solve the optimization problem. Prakash and Datta (2015) proposed a new feedback-based method for the efficient identification of unknown pollutant-source characteristics. It integrates a sequential-monitoring-network design and source identification processes. They evaluated the performance of the method in a polluted aquifer in New South Wales, Australia.

Esfahani and Datta (2018) used fractal singularity-based multi-objective monitoring networks for reactive species contaminant source characterization. The method uses a multi-objective optimization algorithm to solve a two-objective optimal monitoring network design model. The optimization model is linked to a numerical simulation model that simulates the flow and transport processes in the aquifer. While constraining the maximum number of monitoring locations, the optimised monitoring network improves the accuracy of contaminant source characterization. The designed monitoring network can decrease the degree of non-uniqueness in a measured set of possible aquifer responses to geochemical stresses.

2.3 TOOLS AND TECHNIQUES

This section reviews the literature on tools for modelling groundwater flow and transport processes, data interpolation techniques, optimization algorithms and regression modelling. These techniques are used at different stages throughout this study.

2.3.1 Flow and Transport Modelling with Reactive Contaminant Transport

The development of mechanistic reactive chemical transport models has been in the ascendance for the past two decades (e.g. Yeh & Tripathi, 1990; Nienhuis et al., 1991; Engesgaard & Kipp, 1992; Parkhurst & Appelo, 1999; Bacon et al., 2000; Steefel, 2001; Yeh et al., 2001, 2004b; Xu et al., 2003; Yang et al., 2008). These numerical reactive transport models have been diverse in scope. Many models have coupled transport simulations with equilibrium geochemistry (Yeh & Tripathi, 1991; Cheng, 1995; Parkhurst, 1995). Some models couple transport simulations with kinetic geochemistry to describe certain geochemical processes, like precipitation-dissolution (e.g. Lichtner, 1996; Steefel & Yabusaki, 1996; Suarez & Šimunnek, 1996) and redox reactions (Lensing et al., 1994).

More effective general reactive transport models have emerged since the late 1990s and early 2000s. These models are capable of handling a complete suite of geochemical reaction processes (aqueous complexation, adsorption, precipitation-dissolution, acid-base, and reduction-oxidation phenomena). The models also allow individual reactions within any of these geochemical processes to be handled as either equilibrium or kinetic reactions, as appropriate for the system being considered (e.g., Yeh et al., 1996; Bacon et al., 2000; Steefel, 2001; Yeh et al., 2001; Xu et al., 2003, Yeh et. al., 2004). Work in reactive transport modelling has incorporated coupled microbial-mediated processes and some purely geochemical transport processes., Some recent models are able to handle reactive geochemical and microbial types of reactions. Models that can simulate any number of reactions, both geochemical and biological, have gained popularity and have been implemented in a number of codes (e.g., Yeh et al., 2004a, b; Zhang et al., 2007).

HYDROGEOCHEM was among the first comprehensive simulators of hydrologic transport and geochemical reactions in saturated-unsaturated media. This simulator is able to simulate all the hydrogeochemical processes involved in the subsurface. The software code iteratively solves three-dimensional transport and geochemical equilibrium equations (Yeh et al., 1990).The code provides these numerical options for solving reactive geochemical, biochemical transport and heat transfer equations. These are the conventional finite element method (FEM) and hybrid Lagrangian-Eulerian FEM The most recent versions are based on the concept of defining reaction rates, which makes the modelling of mixed fast/equilibrium and slow/kinetic reactions consistent. This thesis utilizes HYDROGEOCHEM 5.0 as a numerical model for simulating the transport of reactive chemical species in a contaminated aquifer.

2.4 DATA INTERPOLATION

Interpolation techniques are used when analysing groundwater flow and transport processes. In order to select an optimal interpolation method to describe the spatial distribution of groundwater parameters, the spatial correlations of parameters such as groundwater level and contaminant species concentrations are interpolated based the inverse distance weighting and kriging interpolation methods. The inverse distance

weighting method is more suitable for data with weak spatial correlation. For parameters that have high spatial correlation, kriging is a reliable method for interpolation. In this work both inverse distance weighting and kriging interpolation methods were used for the various spatial distribution of groundwater data.

2.5 OPTIMIZATION ALGORITHM: ADAPTIVE SIMULATED ANNEALING

Identifying the source of unknown groundwater pollution is a management issue. Groundwater management problems of this kind should be viewed in terms of optimization. Optimization problems include representations of the groundwater system being managed and the management goals and constraints. To solve an unknown groundwater pollution source problem, different optimization algorithms can be applied. The choice of optimization algorithm depends largely on the kind of problem to be solved. In the present study, ADAPTIVE SIMMULATED ANNEALING is used to solve the optimization problem.

Objective functions suitable for identifying pollution sources are complex multi-variate optimization problems. Such formulations are highly non-linear, containing several local and global optima. The ease of implementing complex objective functions and their convergence to a global optimal solution make them suitable for solving ill-posed inverse problems, as is the case with groundwater pollution source characterization.

2.5.1 Adaptive Simulated Annealing (ASA)

Simulated annealing (SA) is a global stochastic optimization algorithm that mimics the metallurgical annealing process (Kirkpatrick, 1983). The objective function is often called the energy *E,* and is assumed to be related to the state, popularly known as temperature T, by a probability distribution. During the course of optimization, new points are sampled and accepted using a probabilistic criterion, such that inferior points also have non-zero probability of being accepted. The temperature is also updated as the algorithm progresses. The search terminates when the temperature has fallen substantially. The original SA algorithm allowed for a very slow rate of decrease in temperature and, hence, a very high cost solution, as compared to ASA. Ingber (1989) developed a version of SA that has complex modifications of the

sampling method, which enables the use of very high cooling rates and, hence, reduced simulation cost. Non-uniform sampling rates for the variables and different cooling rates in the variable and function spaces are used. A concept called *reannealing* updates the cooling rate associated with each parameter by accounting for the sensitivities of the objective function.

Simulated annealing, first introduced by Kirkpatrick et al. (1983), is an extension of the metropolis algorithm (Metropolis et al., 1953). The basic concept of SA is derived from thermodynamics. Each step of the SA algorithm replaces the current solution with a random nearby solution, chosen according to a probability that depends on the difference between the corresponding function values and algorithm control parameters that are gradually decreased during the process. SIMANN is a FORTRAN code for SA developed by Goffe (1996) that has been publicly available since 1996. Ingber (1993) developed an SA computer code in the C-language program called adaptive simulated annealing (ASA). It is a version of the SA algorithm in which the parameters are automatically adjusted as the algorithm progresses. These parameters control the temperature, schedule and random step selection. Thus, the main advantage of the ASA algorithm is that it is more efficient and less sensitive to user-defined parameters.

2.6 FRACTAL / MULTIFRACTAL MODELLING

The evaluation and remediation of contaminated aquifers require accurate delineation of contamination plumes. Ideally, a large number of concentration observations are required to achieve accurate delineation of a contamination plume. However, in practice, due to budgetary constraints, contamination in groundwater resources is detected by a limited number of arbitrarily located or predesigned contamination monitoring wells. Therefore, a technique is required to estimate the boundaries of the plume using the available sparse observation data. The local singularity mapping technique can be used for plume delineation. This technique is based on the multifractal concept. In fractal geometry, a local feature is similar to the whole in terms of shape and structure. Generalized self-similarity is characterized by a power-law relationship. A detailed review of this technique can be found in Zuo and Wang (2016).

Mandelbrot (1967, 1972) introduced the concept of fractals in the last century. Fractal and multifractal analysis has quickly developed into an important branch of non-linear science. It has had significant impacts in many areas of natural science, including the study of earthquakes (Turcotte, 2002), floods (Cheng, 2008), geoscience (Chen et al., 2015) and groundwater studies (Sivakumar et al., 2004).

Fractal behaviour in subsurface solute transport phenomena has been addressed in a number of studies (e.g. Hewett, 1986; Wheatcraft & Tyler, 1988; Benson et al., 2001; Berkowitz & Scher, 2001; Puente et al., 2001a, b). Differences in opinion exist as to the type of fractal behaviour (mono- or multi-fractal) occurring in transport phenomena, the underlying mechanisms involved, and the appropriate predictive methods. For instance, Datta et al. (2016) used local singularity mapping to delineate groundwater contamination plumes. Benson et al. (2001) suggested that a mono-fractional derivative in the advection-dispersion equation may be adequate for solute transport predictions. Puente et al. (2001) investigated the possibility of modelling the dynamics of groundwater contamination plumes using a deterministic fractal-multifractal (FM) approach, via projections of fractal interpolation functions while Esfhani et al. (2017) developed and implemented a fractal singularity based monitoring network design.

Fractal models such as the number-size model (N-S), concentration-area model (C-A; Cheng et al., 1994), spectrum-area model (S-A; Cheng et al., 1999, 2000), simulated size-number (SS–N) model (Sadeghi et al., 2015), singularity index (Chen et al., 2015), and concentration-volume model (C-V; Afzal et al., 2011) have been developed for geochemical data analysis.

In the past two decades, various power law models have been developed based on fractal analysis in mineral exploration, such as the density-area model (Cheng, Agterberg, & Ballantyne, 1994), singularity index (Cheng, 1997), spectrum-area model (Cheng, 1999b), density-distance model (Li, Ma, & Shi, 2003), multifractal singularity decomposition model (Li & Cheng, 2004), and density-volume model (Afzal et al., 2011). The GeoDAS GIS system (Cheng, 2000) for processing non-linear spatial geoscience information, which is based on these multifractal models, has played an important role in identifying geochemical anomalies and modelling the spatial distributions of mineral deposits (Cheng et al., 2001; Ko & Cheng, 2004; Xie and Bao, 2004; Chen et al., 2007; Zuo et al., 2009).

Cheng and his colleagues proposed the singularity theory (Cheng et al., 2010; Cheng, 2012b) based on the C–A fractal model to quantify geo-anomalies according to the invariant properties between fractal measure and scale. A number of case studies have demonstrated that the singularity index is a powerful tool for identifying geochemical anomalies, processing remote sensing information, and analysing the spatial distributions of mineral deposits (Chen et al., 2007; Cheng, 2007; Wang, Zhao, & Cheng, 2011; Zuo, 2011).

The main attraction of fractal/multifractal theory lies in its ability to quantify irregular and complex phenomena or processes that exhibit similarity over a wide range of scales, which is termed self-similarity (Mandelbrot, 1983). Several years of fractal application has proven its ability to effectively characterize spatial distributions in concentrations and relationships between the tonnage and grade of deposits (e.g., Cheng et al., 1994, 2000; Lavallee et al., 1993; Mandelbrot, 1983; Turcotte, 1986, 1997, 2002).

2.6.1 Motivation for this Study

The research has been driven by the substantial challenges associated with the prediction and remediation of contaminant plumes in the environment. Most of these challenges are due to uncertainties associated with contaminant sources (e.g. source locations, strengths, activity times, concentrations) as well as contaminant migration (e.g., velocity, dispersivity,) related to aquifer and contaminant transport properties. Determining the number of contaminant sources, their locations and physical properties is an important task that yields information valuable in predicting the fate and transport of contaminants, making hazard and risk assessments, and remediation. Most often, information about contamination sources and contaminant migration in an aquifer is limited or unavailable, which explains the increasing use of complex numerical inverse models.

Groundwater researchers and modellers are often required to answer questions about unknown groundwater source identification and management of contaminated aquifers. Providing answers to these seemingly straightforward questions requires considerable specific hydrogeochemical information and analyses, as well as general hydrogeologic knowledge, insight and objective functioning. Even relatively simple groundwater problems require knowledge of basic aquifer parameters and hydrologic stresses, such as pumping and recharge rates.

Based on the available methodologies covered in the literature review, there is scope of further developing computationally efficient methodologies for unknown groundwater contaminant characterization in highly geochemically-complex aquifer areas such as mine sites. Reactive transport of chemical species in contaminated groundwater systems, especially with multiple species, is a complex and non-linear process. To increase modelling reliability with real-life field data, uncertainties in hydrogeological parameters and boundary conditions need to be considered as well. Modelling and characterising such complex geochemical processes using efficient numerical models is generally a challenge. In this study, flow and chemically-reactive transport processes are numerically simulated in complex contaminant aquifers by considering the various chemical reactions taking place in the subsurface system.

The source characterization methodology proposed in this study uses groundwater numerical simulation models in a linked simulation-optimization approach. The efficiency and computational feasibility of source characterization are evaluated for complex study areas. A new optimization equation is formulated to solve the characterization of individual reactive species in distributed sources.

The next chapter (Chapter 3) discusses the implementation of a three-dimensional numerical flow and contaminant simulation model of a contaminated aquifer at a former mine site that contains multiple reactive chemical species.

Chapter 3: Calibration and validation of a three-dimensional flow and multicomponent reactive transport model of an abandoned mine site in the Northern Territory, Australia

This chapter describes a three-dimensional finite element-based numerical model of multispecies reactive transport in an aquifer at an abandoned uranium mine in the Northern Territory, Australia. The flow and transport simulation model is based on the geochemical reactions of multiple species in an acidic environment. The numerical groundwater model incorporates kinetic and equilibrium reactions to study the underlying contaminant transport processes in an aquifer at an abandoned uranium mine site. A reasonably accurate calibrated flow and transport simulation model is an essential first step in the process of characterizing unknown sources of contamination. Therefore, the flow and transport model's implementation is described here.

3.1 BACKGROUND TO THE PROBLEM

Predicting the fate and transport of dissolved metals and radionuclides in groundwater systems is important in assessing environmental impacts and developing operational remediation approaches. Contaminants change in terms of concentration and species formation due to physical, chemical and biological processes as they move through an aqueous environment. Hence, the ability to understand and model these processes is fundamental in assessing the efficacy of contaminated aquifer remediation strategies. This chapter describes the set-up, calibration and validation of a three-dimensional multiple species reactive transport model of contaminant transport in an aquifer Calibration and validation of the developed model is performed using very limited field-measured data. The developed simulation model incorporates geochemical reactive processes. The flow and, especially, the multispecies reactive transport model are not based on a simplistic assumption of conservative pollutants, which is an

assumption made in models previously developed by the Northern Territory Department of Mines and Energy (RGC, 2012).

The generation of acidic wastewater at the waste rock dumps and open mining pits of former mine sites in Australia constitutes a matter of great environmental and economic concern. Such acidic waters can transport potentially toxic dissolved metals, which may contaminate groundwater. Over 4000 decommissioned or abandoned mine sites occur in Australia alone (Unger, C., 2020). Many of these abandoned mine sites contain reactive sulfide minerals which cause sulfide to be acidified by oxidation reactions. These reactions can persist for several hundreds of years if not controlled and monitored. At waste rock piles or tailing impoundment locations close to aquifers, it is possible for dissolved metals to move away from mining waste locations and degrade nearby groundwater and surface water bodies. Thus, an understanding of the geochemical interactions between acid mine drainage (AMD) and aquifers is critical in assessing impacts on the quality of water resources. This also provides a basis for the appropriate management and remediation of sites similarly affected by AMD contamination.

The study site used for this investigation is illustrative of mine sites undergoing complications in rehabilitation related to the acid rock drainage (ARD) of metals, including uranium. Such acid rock drainage has serious environmental impacts on groundwater at the mine site and in a nearby river. The acid rock drainage of metals from mining waste is a source of groundwater contamination, including that of copper, iron, manganese, zinc and sulfate. The geochemical processes considered in this study include both kinetic and equilibrium reactions. Uranium and its daughter products have not been emphasized enough in the past and, therefore, contamination due to uranium has not been adequately modelled at this site. This study also developed a simulation model that incorporates the geochemical processes of uranium and its interactions with other chemical species present in the site.

At this site, nonradioactive uranium isotopes are prevalent; therefore, they were incorporated in this study. In addition, the interactions of uranium with other chemical species present at this aquifer site are also modelled.

Substantial modelling of the flow and transport processes was carried out for this site as part of a rehabilitation plan. However, these models consider the chemical species as conservative and only representative of a single species (RGC, 2012). A

well-defined reactive transport model incorporating a wide range of reaction types and rates has not been implemented before. Hence, there was a need to model the geochemical reactive transport process, including that of multiple chemical contaminant species, at this abandoned uranium mining site. This study highlights the geochemical reactions likely to occur based on contaminant concentrations at monitoring points and our understanding of the complex mineralogical and geochemical processes at the site. These reaction types include, but are not limited to, precipitation-dissolution, acid-base, aqueous complexation, ion-exchange, adsorption-desorption, and oxidation-reduction. The objective of the groundwater modelling was to achieve results representative of the current understanding of the interactions between reactive species or metals at waste dumps, open-cut pits and aquifer systems. The design, development and calibration of the model were all adapted to achieve these objectives and provide a framework for its potential use as new information on reactive processes becomes available. The objective was to implement a flow and transport simulation model based on very limited measurement data that describes the complex hydrogeological and geochemical processes associated with acid mine drainage and metalliferous drainage in the contaminated aquifer underlying the former mine.

Acid metal drainage (AMD), also known as acid mine drainage or acid rock drainage (ARD), happens when sulphidic rocks, for example arsenopyrite, chalcopyrite and pyrite, are exposed to oxygen and water. Although AMD occurs naturally, most are humanly induced and originate from reactive sulfide mineral deposits in tailings dams, waste rock dumps (WRDs), mine pits, and leach pads (Claire et al., 2012).

At most mine sites, the main sources of pollutants are open pits, WRDs and tailing dams. The WRDs and tailing dams most often consist of several minerals and geological matrices, which are exposed to weathering conditions. Pyrite oxidation is the fundamental reaction in the leaching of metals and radionuclides into the environment. Leaching of metals and radionuclides from waste rock piles and tailing dams is estimated to last for hundreds to thousands of years (Öhlander et al., 2012). Accurate prediction of the release rates of metals and radionuclides from these sources and their transport into the subsurface environment is critical in the assessment of environmental impacts and the development of effective remediation strategies. To

produce a realistic representation of the system under study and its complex problems, sophisticated models are required.

The development of reactive chemical transport models has progressed in the last three decades (Yeh & Tripathi, 1990; Pruess, 1991; Zyvoloski et al., 1994; Lichtner, 1996; Steefel & Yabusaki, 1996; William Maxwell Aitken et al., 2000; Fang et al., 2003; Yeh et al., 2004; Zhang et al., 2007, Datta et al., 2016, Esfahani & Datta, 2015; Xiao et al., 2018; Druhan & Tournassat, 2019). These numerical reactive transport models have had various scope, and different competencies and accuracies. This chapter describes the development and application of the current advancement in mechanistic based numerical models for simulating fluid flow, thermal transport and reactive transport in variably saturated porous and fractured media. These are amongst the most practical models that can be employed to study the geochemical processes involved in groundwater flow and transport processes under variably saturated conditions.

The HYDROGEOCHEM (Yeh & Tripathi, 1991) flow and transport simulator was the first all-inclusive simulator of hydrological transport and geochemical reactions in saturated-unsaturated media. It solves iteratively two-dimensional transport and geochemical balanced equations. The hybrid Lagrangian-Eulerian finite element model (LEHGC) was developed by Yeh et al . (1995) to simulate transport through saturated-unsaturated medium. The LEHGC model is an improved HYDROGEOCHEM version, which uses only an Eulerian system. This hybrid modification increases the solution's efficiency and allows for greater time steps to be used. HYDROGEOCHEM 2.0 was introduced by Yeh and Salvage (1997) with an extended LEHGC 1.1 model (Yeh et al., 1995) capable of handling varied balance and geochemical kinetic reactions. The updated LEHGC 1.1 model is a more powerful LEHGC version and is more stable as it can use more grid nodes. Version 4.0 (Li, 2003),of the HYDROGEOCHEM is limited to two dimensional simulations. This simulator incorporates heat transfer with reactive geochemical and biochemical transport modelling. HYDROGEOCHEM 5.0 (HGCH; Yeh et al . , 2004; Sun, 2004) is the new version and can be modelled in three dimensions. In addition, this simulator provides three options for numerical calculations of reactive geochemical and biochemical transport and heat transfer equations: 1) conventional finite-element methods (FEM), 2) hybrid Lagrangian-Eulerian FEM, and 3) hybrid Lagrangian-

Eulerian FEM for interior nodes plus FEM for boundary nodes. However, HYDROGEOCHEM 5.0 has certain limits: 1) only one phase of fluid is applicable, and 2) double porous media cannot be used effectively (Yeh et al., 2004). This study simulates the transport of reactive chemical species in a contaminated aquifer using HYDROGEOCHEM 5.0.

The use of a reactive transport model to simulate the transport of reactive multi-species contaminants in a heterogeneous, anisotropic, saturated-unsaturated porous medium at a former uranium mining site in Australia is described here. The mining site contains a number of sources of potential reactive chemical species, such as waste rock dumps and open pits. This chapter outlines the validation of flow modelling, and, to a limited extent, the transport process at the abandoned Rum Jungle uranium mining site. It also discusses the calibration and validation of the flow and transport simulation model.

3.2 METHODOLOGY

In this section, the processes of solving the unsaturated-saturated flow and contaminant reactive transport are described. The governing equations of flow and reactive geochemical transport are solved using numerical approximation based on finite element methods and other numerical schemes. The finite element method is used because of its function-approximation capability and because it produces spatially-continuous solutions. The complexity of the study area also influenced the choice of numerical approximation method. The finite element method has the ability to handle anisotropy and heterogeneity in aquifer systems. Other reasons that influenced the choice of the finite element method are as follows:

- It does not require special formulae to incorporate irregular boundaries,
- the computational effort required may be less than that required by other methods, as fewer nodal points are required to represent the region of interest with similar accuracy,
- unstructured meshes can be incorporated to handle different levels of spatial discretization in different sections of the region of interest

3.2.1 Modelling of groundwater flow and contaminant transport

The 3-D finite element-based reactive transport simulator HYDROGEOCHEM 5.0 was utilized in this study to model the aquifer flow and transport processes. The flow and reactive transport models are described in the following sections.

3.2.2 Governing flow equations

The general equations for flow through saturated-unsaturated media are based on: 1) fluid continuity, 2) solid continuity, 3) fluid movement (Darcy's law), 4) stabilization of media, and 5) water compressibility (Chen et al., 1998):

$$\frac{\rho}{\rho_o} F \frac{\partial h}{\partial t} = -\nabla \cdot \left(\frac{\rho}{\rho_o} \mathrm{V} \right) + \frac{\rho^*}{\rho_o} q \qquad (3.1)$$

Where F = a generalized storage coefficient (1 /L), defined as:

$$F = \alpha' \frac{\theta}{n_e} + \beta' \theta + n_e \frac{dS}{dh} \qquad (3.2)$$

K = the hydraulic conductivity tensor (L/T), defined as:

$$K = \frac{\rho g}{\mu} \quad K = \frac{\rho / \rho_O}{\mu / \mu_O} \frac{\rho_o g}{\mu_O} \quad K_s K_r = \frac{(\rho / \rho_o)}{(\mu / \mu_o)} \mathrm{K}_{\mathrm{so}} k_r \qquad (3.3)$$

and V = Darcy's velocity (L/T), described as:

$$V = -K \cdot \left(\frac{\rho_o}{\rho} \nabla h + \nabla z \right) \qquad (3.4)$$

where ρ is the fluid density [M/L^3], ρ_o is the reference fluid density at zero chemical concentration and at the reference temperature [M/L^3], F is the generalized storage coefficient [1/L], h is the pressure head [L], t is time [T], ρ^* is the fluid density of either injection ($\rho^* = \rho$) or withdrawal ($= \rho$) [M/L^3], q is a source or sink representing artificial injection or withdrawal of fluid [(L^3/L^3)/T], V is the specific discharge or Darcy's velocity [L/T], K is the hydraulic conductivity tensor [L/T], z is the potential head [L], α' is the modified compressibility of the media [1/L], θ is the effective moisture content [L^3/L^3], n_e is the effective porosity [L^3/L^3], β' is the modified compressibility of the liquid [1/L], S is the degree of saturation of water, g is acceleration due to gravity (L/T^2), μ_o is the fluid dynamic viscosity at zero chemical

concentration and at the reference temperature M/(L/T), μ is the fluid dynamic viscosity M/(L/T), K_{so} is the reference saturated hydraulic conductivity tensor [L/T], and k_r is the relative permeability or relative hydraulic conductivity (dimensionless).

The finite element method is used to solve Equations (3.1), (3.2), (3.3) and (3.4), and the constitutive relationships between pressure heads, hydraulic conductivity tensor, and degree of saturation along with the appropriate initial and boundary conditions. The initial conditions for this study were obtained from field measurements. In the case of transient simulation, the initial conditions must be realistic and consistent. Initial conditions that are not appropriate are likely to introduce either non-convergence or non-realistic solutions, and therefore these conditions need to be specified as close to the actual situation. In a number of situations, it may be impossible to measure the initial pressure field across an entire study domain. In such situations, an alternative way of setting the initial conditions is to assume that, in general, steady state flow conditions may have existed. Therefore, the simulation results from of a steady-state simulation with steady-state specified boundary conditions are used.

For the modelling purposes of this study, three main boundary conditions were assigned. These are the Dirichlet, Cauchy and variable boundary conditions (Yeh et al., 2004). In the Dirichlet boundary conditions, the pressure head is set as a function of time. In the Cauchy boundary conditions, the volumetric flux is set as a function of time. For variable boundary conditions, either the flux or Dirichlet boundary conditions can be given, depending on the infiltration capacity and the rainfall intensity during precipitation periods, or based on the evaporative capacity of the media and the evaporation potential of the atmosphere during non-precipitation periods. Using the HYDROGEOCHEM 5.0 code, the temporal-spatial distributions of the hydrological variables were simulated, including total head, Darcy's velocity, pressure head, and moisture content.

3.2.3 Governing Reactive Transport equations

The governing equations for the reactive transport of the reactive biogeochemical system are discussed below. The governing equations for transport were derived based on the continuity of mass and Fick's flux laws (Yeh et al., 2000). The main transport processes are advection, dispersion & diffusion, source & sink and

biogeochemical reactions (including radioactive decay). The general transport equation governing the temporal-spatial distribution of any biogeochemical species in a reactive system is described below. Let C_i be the concentration of the i^{th} species; then, the governing equation for C_i is obtained by applying the principle of mass balance in integral form, as follows (Yeh et al., 2000):

$$\frac{\partial \theta C_i}{\partial t} + \theta \alpha' \frac{\partial h}{\partial t} C_i = L(C_i) + \theta r_i + M_i, i \in \{M\} \quad (3.5)$$

Where L is the transport operator denoting:

$$L(C_i) = -\nabla \bullet (VC_i) + \nabla \bullet \left[\theta D \bullet \nabla C_i\right] \quad (3.6)$$

$$\frac{D}{Dt}\int_v \theta\, C_i\, dv = -\int_\Gamma n \bullet (\theta\, C_i)\; V_i d\,\Gamma - \int_\Gamma J_i d\,\Gamma + \int_v \theta\, r_i dv + \int_v M_i dv,\ \mathrm{i} \in \mathrm{M} \quad (3.7)$$

where C_i is the concentration of the i^{th} species in moles per unit volume [M/L^3]; v is the material volume containing a constant amount of media (L^3); Γ is the surface enclosing the material volume v (L^2); n is the outward unit vector normal to the surface Γ; r_i is the production rate of the i^{th} species due to biogeochemical reactions, in chemical mass per water volume per unit time [M/L^3/T]; {M} = {1,2,…,M}, in which M is the number of biogeochemical species; D is the dispersion coefficient tensor [L^2/T]; and M_i is the source/sink of the i^{th} species in chemical mass per unit volume of media [M/L^3/T]; M is the number of biogeochemical species; v_i is the transporting velocity relative to the solid of the i^{th} biogeochemical species (L/T); θr_i is the production rate of the i^{th} species per unit medium volume due to all biogeochemical reactions [(M/L^3)/T], Ji is the surface flux of the i^{th} species due to dispersion and diffusion with respect to the relative fluid velocity [(M/T)/L^2] and Vi is the transporting velocity relative to the solid of the i^{th} biogeochemical species (L/T).

As in the flow model , in order to simulate reactive transport across a wide range of problems, appropriate transport boundary conditions were applied in the model. The physical definitions and mathematical descriptions of these boundary conditions are comparable to those of the flow model (Yeh et al., 2004).

Equations for geochemical processes

The production of a species along its transport path is governed by a number of biogeochemical processes. One of the difficult aspects of geochemical modelling is the formulation of a governing rate equation to represent the chemical processes governing the rate of production of any species (r_i in equation 3.5) and its associated parameters. The formulation of rate equations related to all N reactions is a critical issue in the modelling of mixed equilibrium and geochemical kinetic reactions. A rate equation is essential for the quantitative description of a general geochemical reaction that is written as follows (Fang et al., 2003):

$$\sum_{i\in\{M\}} \mu_{ik} \hat{C}_i \leftrightarrow \sum_{i\in\{M\}} \nu_{ik} \hat{C}_i, k \in \{N\} \tag{3.8}$$

where $\hat{C}_i$ is the chemical formula of the i^{th} species; μ_{ik} is the reaction stoichiometry of the i^{th} species in the k^{th} reaction associated with the reactants; ν_{ik} is the reaction stoichiometry of the i^{th} species in the k^{th} reaction associated with the products; and {N} = {1, 2, …, N}, in which N is the number of reactions.

For all geochemical reactions, two categories of reactions exist: kinetic and equilibrium. Assuming that there are NE fast/equilibrium reactions (all of which must be independent) and NK slow/kinetic reactions (Yeh et al., 2010), then the number of reactions will be: N = NE + NK.

Kinetic Reactions

For a fundamental kinetic reaction, the rate law is given by collision theory as:

$$R_K = K_K^f \prod_{i=1}^{M} \left(A_i\right)^{\mu_{ik}} - K_K^b \prod_{i=1}^{M} \left(A_i\right)^{\nu_{ik}}, K \in N_K \tag{3.9}$$

where R_k is the rate of the k^{th} kinetic reaction, A_i is the activity of the i^{th} species, K_K^f and K_K^b are the activity-based forward and backward rate constants of the k^{th} kinetic reaction, respectively, and N_K is the number of kinetic reactions. The forward and backward rate constants cannot be determined sequentially from the concentration-time curves of all species because the N_K equations in Equation (3.8) are coupled regarding the forward and backward rate constants.

Equilibrium reactions

If the reaction is an equilibrium reaction, the reaction rate is infinity, which results in the law of mass action as:

$$R = \infty \exists K_K^e = \left(\coprod_{i=1}^{M} (A_i)^{\mu_{ik}} \right) / \left(\prod_{i=1}^{M} ({}_i A_i)^{\nu_{ik}} \right), K \in N_E \tag{3.10}$$

Where R_E is the rate of the k^{th} equilibrium reaction, K_K^e is the equilibrium constant of the k^{th} reaction, A_i is the activity of the i^{th} species, and N_E is the number of linearly-independent equilibrium reactions. The equilibrium constants may be determined sequentially by the measurement of the activities of all species.

Due to the nature of the study area and its complexities, a reaction-based formulation was chosen to represent all the geochemical processes in the aquifer site. In a reaction based formulation, all biogeochemical processes are conceptualized and transformed into a reaction network (Fang et al., 2003). This is to consider the contributions of the individual process reaction interplays in the aquifer system and avoid the possibility of the production rate being represented as a lumped rate of all reactions of a particular process which, in this specific instance, will not identify individual reaction rates.

The difficulty in applying the reactive transport model to real-world problems is in transforming the understanding of biogeochemical processes into reaction networks with a rate equation for each reaction. The transformation is a complex task without which our understanding of the aquifer system will be incomplete or inadequate.

3.2.4 Solution technique

Equations 3.1 to 3.10 are a set of partial differential equations that require solving and coupling through flow and transport solutions. The linearized matrix equations can be solved by using the finite element method by applying a number of numerical schemes. A two-step method was used to solve the chemical transport equations and chemical equilibrium equations. Once the solutions for a specific time-step converges, the calculation continues to the next time step.

Finite element methods were used for temporal discretization of the governing partial differential equations in the flow model and reactive transport model. The

Galerkin finite element method was used for spatial discretization of the modified Richards equation. Meanwhile, scalar reactive transport equations were solved using both conventional finite element methods and hybrid Lagrangian-Eulerian finite element methods for spatial discretization. The solutions to the chemical equilibrium equations were obtained using the Newton-Raphson or Picard methods. Three numerical schemes (iteration approach, operator splitting approach, and predictor-corrector approach) were used to couple the hydrological transport and geochemical reactions. Since this study involves a complex three-dimensional model containing more than 6000 elements, 4000 nodes and 33 species, it was more efficient to solve it using the operator splitting approach. Hence, coupling of the transport and geochemical reactions was achieved using the operator splitting approach.

3.3 STUDY SITE

3.3.1 Overview

The study area is the former site of one of Australia's first major uranium mines, Rum Jungle (Figure 2.1). The site is located approximately 105 km south of Darwin in the headwaters of the East Branch of the Finniss River, near Batchelor in the Northern Territory, Australia. The site was proclaimed a Restricted Use Area in 1989 under the *Soil Conservation and Land Utilisation Act* of the Northern Territory and is closed to public access. The mine was in operation from 1954 to 1971. Uranium (U) and Copper (Cu) were the major minerals mined. Although the last uranium ore was extracted in 1963, the processing of U and Cu continued from the ore stock until April 1971. A total of 863,000 tonnes of uranium ore were processed. Its average grade was 0.28–0.41% and 3,520 tonnes of U_3O_8 were produced from different deposits in the Rum Jungle. Between 1954 to 1958, uranium was mined from the White and Dyson open pits from while copper was extracted from the intermediate pit. Within the intermediate waste rock dumps (WRDs), quartz is the dominant constituent and pyrite is a minor component. Twenty percent of Dyson's WRDs consist of pyritic black shale with an average content of pyrite. The WRD of the White Pit is a combination of minerals from carbonaceous slates, visual schists, phosphate, sulphate, and sulfide. At Rum Jungle all three WRDs were coated in clay and gravel. The clay caps have cracked, and water can now enter during the wet season, which has led to acidic mine

drainage (Mudd & Patterson, 2010). Acid mine drainage (AMD) and heavy metal mobilization at the site have resulted in significant environmental impacts on local groundwater and the East Branch of the Finniss River, with radioactive tailings remaining in some areas (Kraatz, 2004). The site was rehabilitated from 1983 to 1986 at a total cost of A$ 18.6 million. Several rehabilitation works have been carried out since 1986 and rehabilitation is still ongoing.

Figure 3.1 (a) shows the mine site's layout during its operational years, while Figures 3.1 (b) provides a closer look at the open pits and waste rock damps (WRDs) afters years of ongoing contamination.

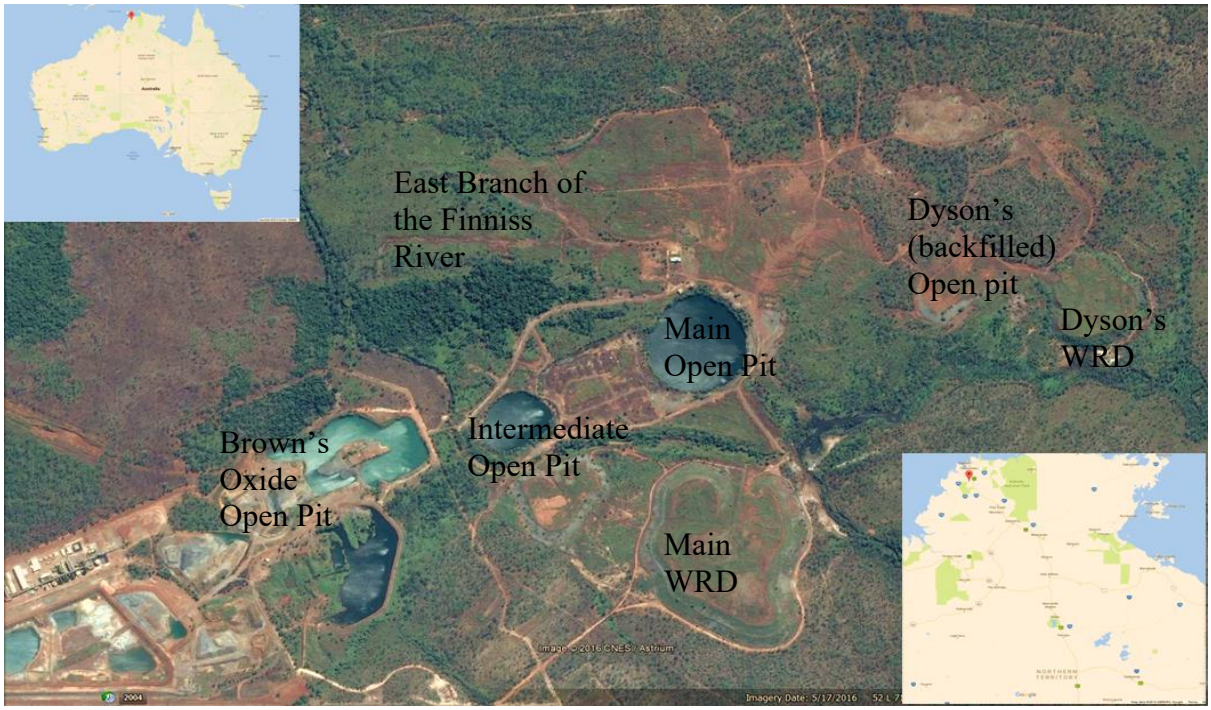


Figure 3.1 (a): Satellite image of the Rum Jungle mine site. Google Earth Imagery © 2019. Insets show its location in the Northern Territory (lower right) and Australia (upper left).

Figure 3.1 (b): Aerial view of the study site, showing the Intermediate Pit, Overburden Heap, White's Pit, and Dyson's Overburden Heap. Photo: NT Government Department of Primary Industry and Resources.

Geological and hydrogeological characterization

The geology and hydrology of the Rum Jungle area are complex. Detailed descriptions are given in Mudd and Petterson (2010), McKay and Miezitis (2001) and CR (2005), and are summarised herein.

3.3.2 Hydrology

The Rum Jungle field consists of the Finniss River Eastern Branch (EBFR), which is situated approximately 8.5 km upstream of its confluence with the Western Branch. From the east, surface water reaches the mine site via the Finniss River's upper east branch and via Fitch Creek, from the southeast. Both creeks met near the northeastern corner of the Main WRD before mining, and eventually flowed eastward into the natural course of the river. The river was redirected during mining to the Eastern Finnish Diversion Channel (EFDC) to allow access to the main and intermediate ores (Figure 3.1A). Through a canal near the former acid dam, the upper east branch of the Finniss River and Fitch Creek flows directly into the east Finniss drainage channel and, during heavy flows, to the Main Pit. Water then flows through

a channel from the Main Pit to the Intermediate Pit, which roughly follows the course of the pre-mining river. Outflow from the Intermediate Pit to the east Finnish diversion channel occurs near the mine site 's western boundary and combined flows from the Main and Intermediate Pits and east Finnish diversion channel continue north into the Finnish River's Eastern Branch natural course. River flows vary predictably in response to intra-annual rainfall variability and typically vary by several orders of magnitude over the course of a year. Early wet season flows in the river are usually observed at the beginning of December or January in response to high rainfall.

3.3.3 Geology

The mine area mineral field contains numerous polymetallic ore deposits, such as the Ranger and Woodcutters ore deposits and the ore deposits associated with the Rum Jungle Mine (i.e. the Main, Intermediate, Dyson’s, and Brown’s Oxide ore deposits). The Rum Jungle Mine Site is located in a triangular area of the Rum Jungle mineral field bounded to the south by the Giant's Reef Fault, and a series of northward east-trending ridges. This triangular area is known as "The Embayment" and is situated on the shallow-dipping limb of a northeast-trending, south-westerly plunging asymmetric syncline formed by northerly dipping faults. The primary lithological units in The Embayment are the Mount Partridge Group's Rum Jungle Formation, and meta-sedimentary and subordinate meta-volcanic rocks. The Rum Jungle Complex is made up primarily of granites and occurs mostly on the southeastern side of the Giant's Reef Fault, while the Mount Partridge Group occurs north of the fault and consists of the following sedimentary units (from younger to older): the Geolsec Formation, Whites Formation, the Coomalie Dolostone and the Crater Formation. The Crater Formation comprises coarse- and medium-grained siliciclastic, whereas the Coomalie Formations comprise magnesite and dolomite with minor chert lenses (McCready et al., 2001). In contrast, Whites Formation (which hosts uranium and polymetallic mineralisation) comprises graphitic, calcareous, slate-phyllite-schist, sericitic, and chloritic. Hence, Whites Formation marks a distinct change in the sedimentary and environmental conditions that occurred in the early Proterozoic.

3.3.4 Aquifer Characterization

The Rum Jungle mine site features a shallow porous aquifer unit and a deeper fractured aquifer unit. These units are hydraulically connected and, hence, it appears that the aquifer may reasonably be regarded as a single aquifer with two units rather than separate shallow and deep aquifers. The shallow unit consists of mixed deposits of in-situ weathered bedrock and soil material with a colluvium-alluvium mix. There are zones of permeable clayey-sand that are interspersed with mottled zones of ferruginous sandy clays. In proximity to the East Finnis River and its tributaries, the shallow soils are predominantly comprised of riverine sands. The deeper unit consists of several lithologies, including granite, dolostone, shale and schist. Investigation by Water Studies (2000) showed that karstic zones with high groundwater inflows may be found in the dolostone, and some aquifer testing results could indicate the presence of permeable karstic features within the fresh to slightly weathered Coomalie Dolostone. In general, groundwater flow in the shallow aquifer unit is controlled by the primary permeability of unconsolidated overburden soils or highly weathered bedrock, whereas groundwater flow in the deeper aquifer unit is controlled by secondary permeability (faults, fractures and/or karstic features). The major fault intersecting the Browns Oxide project is believed to be very transmissive (Water Studies, 2002). In contrast, the Giant's Reef Fault has been inferred to be a hydraulic barrier due to the presence of granitic material on its southeastern flank (Water Studies, 2002; Coffey, 2006).

3.4 CONCEPTUAL AND NUMERICAL DEVELOPMENT

3.4.1 Conceptual Approach Overview

The primary purpose of a conceptual model is to formulate a simplification of the groundwater field problem and collate corresponding field data to visualize and analyse the aquifer system. The conceptualization includes synthesis and framing-up of data relating to, hydrology, geology, and geochemistry. A numerical groundwater flow model was constructed to simulate variations in the groundwater flow system at the Rum Jungle Mine site from 2010 to 2012. This numerical flow model is a mathematical representation of a conceptual model that enables a quantitative representation of real field features. The numerical representation is based on the

assumption that the aquifer system at the mine site is subdivided into hydro-stratigraphic units that represent the waste rock dumps and naturally occurring bedrock aquifers. Each hydro-stratigraphic unit is characterized as a single model layer with representative hydraulic properties. Recharge is estimated as a percentage of incident rainfall assigned to infiltration sections at the site based on elevation. The waste rock dumps and open pits represent a single top layer of variable thickness. The other geological aquifer units are represented as model layers with constant thickness across the model domain. The flooded open pit is represented by specified head boundary conditions that are equivalent to the water levels observed in the pits during the simulation period. Groundwater movement in the hydro-stratigraphic units follows Darcy's law.

3.5 NUMERICAL MODEL SETUP

3.5.1 Flow Model

The finite element method was used for temporal discretization of the underlying partial differential equations in the flow model. The finite element mesh generated for the numerical model consists of 6,587 nodes and 10,704 elements. The Galerkin finite element method was used for spatial discretization of the modified Richards equation governing the pressure fields. The numerical model starts at year 2010 because this is when data became available and ends in 2012. Therefore, the numerical model covers a period of about 730 days. For time discretization, time steps of 30 days were set. This time discretization criterion resulted in a total of 24 time steps.

The numerical model domain was spatially discretised into a 3-dimensional mesh with a triangular wedge mesh. In planar view, each element is represented triangularly, whereas the thickness of the elements depends on the number of layers used to vertically discretise the model domain. For this model, the elements in layers were assigned a set of hydraulic properties based on different material types to represent the complex heterogeneity of the aquifer. The thickness of the elements varies according to lithology. The model is made up of six layers and covers a maximum elevation of approximately 110 m. Surface topography elevation values obtained from a digital elevation model (DEM) were used to define the top of Layer 1. Figure 3.2(a) shows an aerial view of the current ground elevations across most of

the model domain. The remaining discretization and resulting finite-element mesh are shown in Figure 3.2(b), which shows details of the spatial discretization process performed for the numerical model. Layer thicknesses are vertical in depth, with an overall model thickness of 150 m. Layer thicknesses were assigned as follows. Layer 1, which mainly consists of waste rock dumps and tailings, was assigned variable thickness. Layer 2 thickness = 0–7.5 m, Layer 3 = 7.5–15 m, Layer 4 = 15–45 m, Layer 5 = 45–105 m, and Layer 6 = 105–150 m. The tops and bottoms of Layers 3 to 6 were set to the thickness values listed above and were fixed throughout the modelling process.

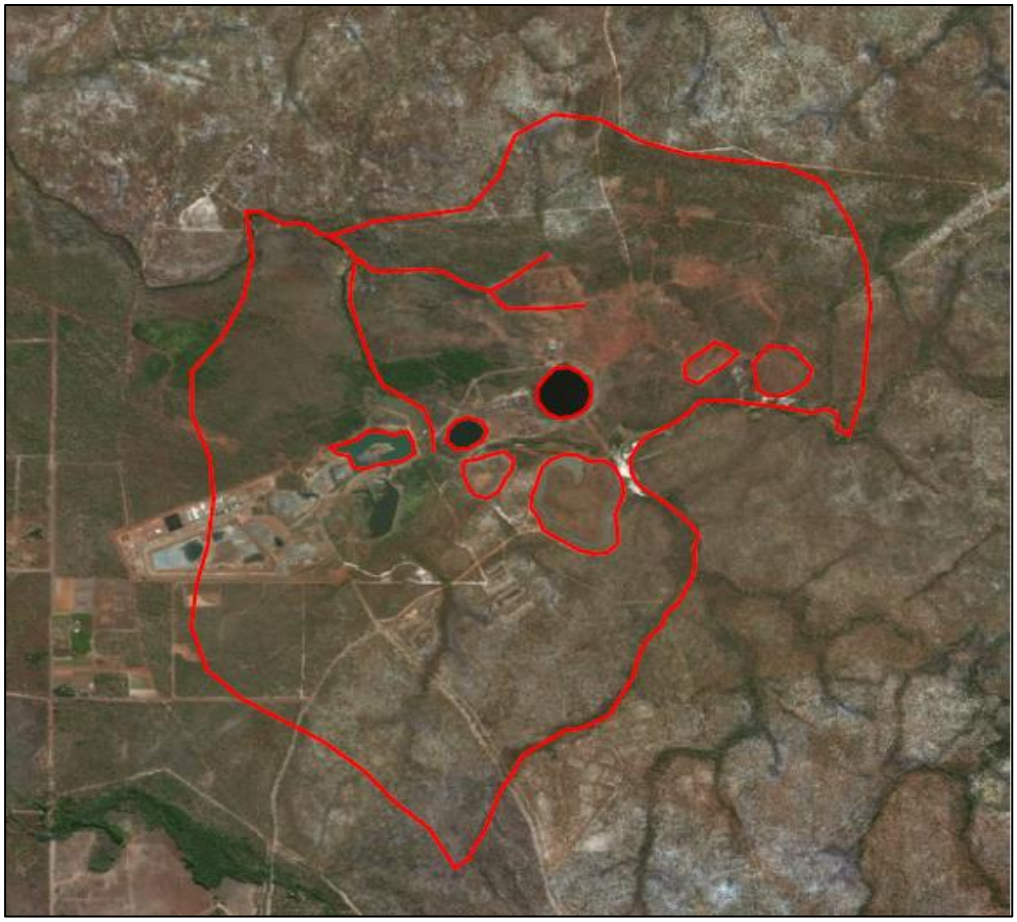


Figure 3.2 (a): Aerial view of the study site with the model domain highlighted in red

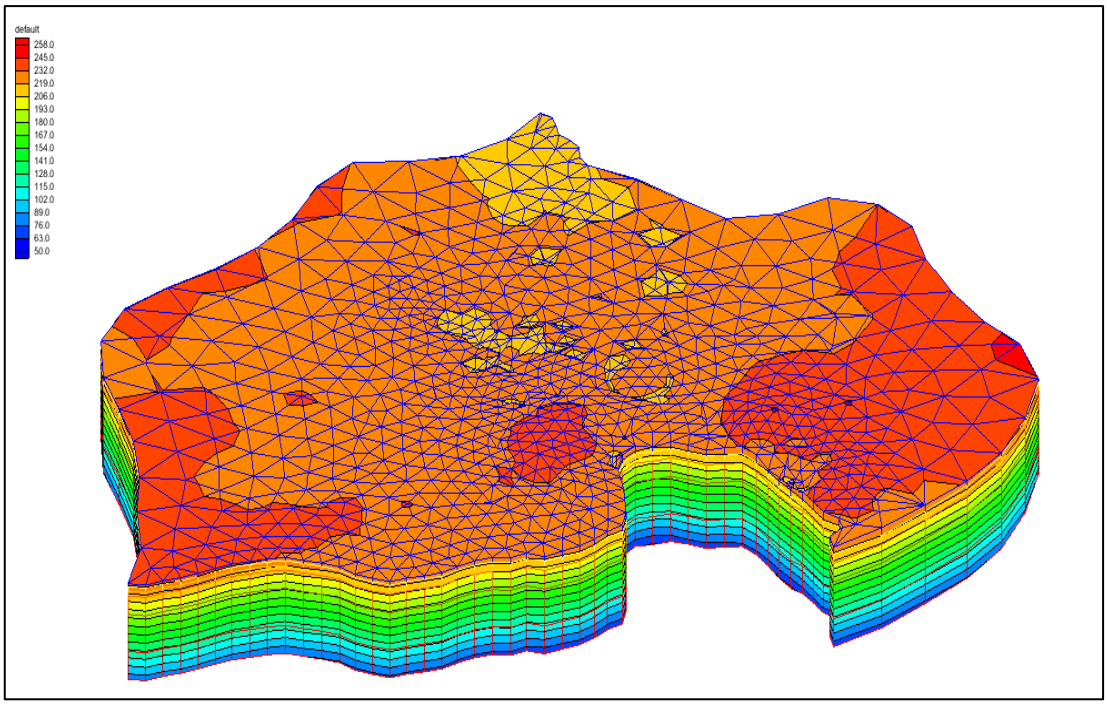


Figure 3.2 (b): Three-dimensional mesh representation of the model domain and study site

Boundary conditions

Every model needs an appropriate set of boundary conditions that represent the system's relation with the surrounding region. Several boundary conditions were set for modelling the flow and transport processes of this study area. Specified heads were set to element nodes that interconnect the perimeters of the flooded pits in Layers 2 and 3. Elements and nodes surrounding the pits representing the bedrock aquifer in contact with water inside the pits were set to head values equal to the measured groundwater level in the pits. The model does not simulate flows within the flooded open pits themselves, so elements within the head boundary are not active. Pit water levels and groundwater levels at monitoring points nearby are adapted to represent the open pit as a head boundary condition derived from the groundwater levels measured in monitoring points situated nearer to this boundary. The northern boundary where the Finnish River is located was assigned a constant head boundary condition, as were the creeks alongside the Finnish River's east side and the creeks at the southern boundaries of the model. For these boundaries, river-bed elevations, and temporally varying river head values, obtained through field observations, were assigned. A transient constant head boundary was set that simulates groundwater water level changes in the main, intermediate and brown pits, and the river. The finite element

method provided ease of using variable meshes, ease of incorporating all the boundary conditions and accurate geometric representations of the aquifer system. The discretization of the study area and model boundary conditions is shown in Figure 3.3.

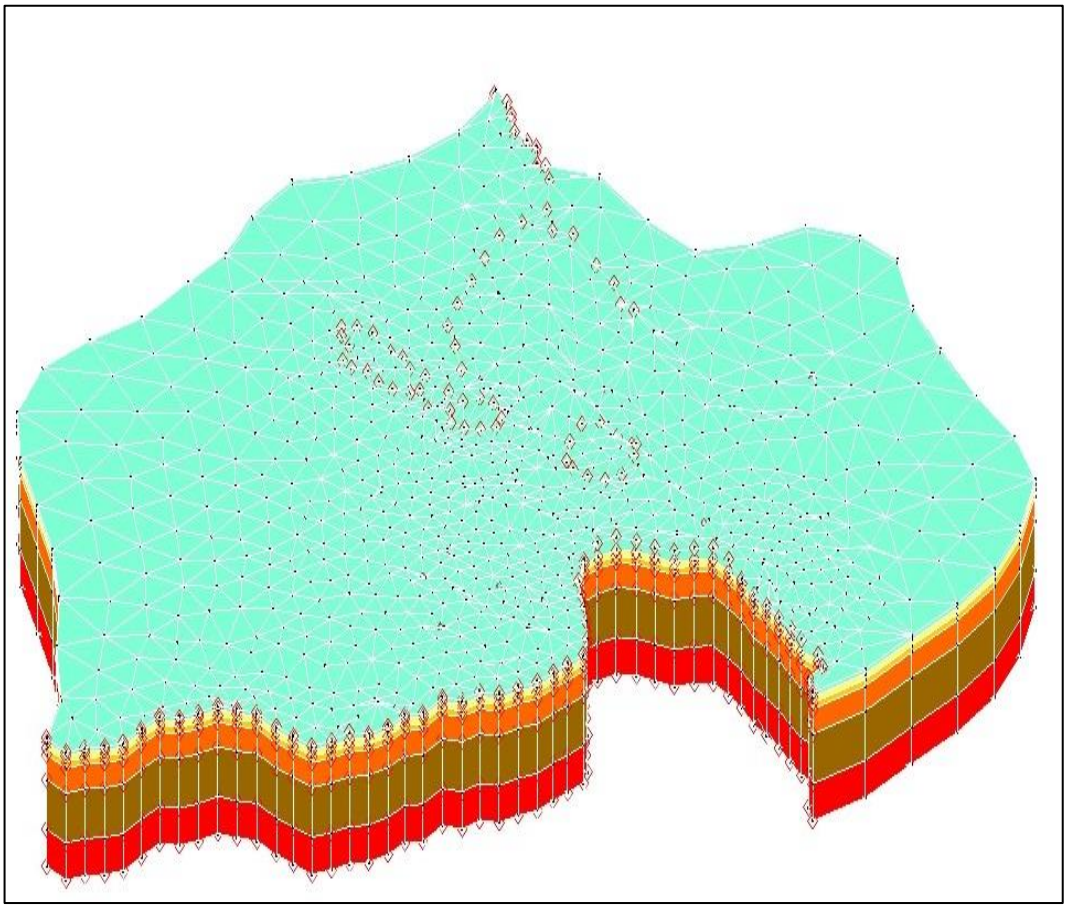


Figure 3.3: Three-dimensional representation of the model's boundary conditions

Model input parameters

Several input parameters for aquifer properties and boundary conditions were used to broadly describe characteristics of the aquifer. Aquifer properties defines the geological medium through which groundwater flows in terms of porosity, hydraulic conductivity, bulk density, and moisture content, while groundwater boundary conditions describe the water flux between aquifer layers and surface features such as groundwater recharge rate and well pumping schedule. The hydraulic conductivity values and specific yield/specific storage values used for the model's hydro-stratigraphic units were estimated from pumping tests described in previous studies (RGC, 2010).

The initial hydraulic head for the model was based on groundwater level data measured at 22 monitoring wells in August 2010. The initial hydraulic head ranged from 50 to 71 m Australia height datum (mAHD). Groundwater recharge was estimated from annual rainfall data recorded at the site. It was assumed that net recharge by rainfall and flows from the flooded open pits were the only sources of water input to the groundwater system within the model domain. The aquifer is recharged largely by rainfall and sources of surface water bodies. The rainfall distribution acrosss the study area was estimated from annual rainfall records for the study area, which was 2372 mm per year. Based on calibration and validation of the flow models, the vertical annual recharge was assumed to be 25% of the average gross rainfall over the study area. Groundwater flow processes were simulated using the hydrogeological parameters listed in Table 3.1. The vertical and horizontal hydraulic conductivities (K_x and K_z, respectively) are shown in Table 3.1 (a) for the different geological layers of the study area. The hydraulic conductivity in the other horizontal direction, K_y, was assumed to be same as the horizontal hydraulic conductivity K_x.

Table 3.1(a): Hydraulic conductivity and layer thickness values of Aquifer

	Hydraulic conductivity		
Model layer	K_x (m/day)	K_z (m/day)	Thickness (m)
Layer 1	0.13	0.12	3
Layer 2	1.21	1.21	7.5
Layer 3	0.44	0.44	7.5
Layer 4	0.65	0.65	30
Layer 5	0.11	0.11	60
Layer 6	0.04	0.04	45

Table 3.1(b): The Study Aquifer's Physical and Hydrogeological Properties

Parameter	Value
Number of nodes	6587
Number of elements	10,704
Effective porosity, θ	0.28
Longitudinal dispersivity, αL	10 m/d
Transverse dispersivity, αT	0.1 m/d
Vertical dispersivity, αV	0.01
Average rainfall	2372 mm/year

3.5.2 Reactive Transport Model

To represent the reactive geochemical processes that occur at the study site, a set of reaction networks consisting of equilibrium and kinetic reactions that describe the aquifer's geochemical processes was formulated, based on contamination and groundwater quality data. HYDROGEOCHEM 5.0 was used to simulate reactive chemical transport and the long-term behaviour of contaminant movement in the aquifer. This study considered the hydrogeochemical transport of six components—OH^-, Cu^{2+}, Fe^{2+}, Fe^{3+}, Mn^{2+} and UO_2^{2+}—and an aqueous complexation of 17 species and three minerals (pyrite, uranite and chalcopyrite). Geochemical reactions of precipitation and dissolution aqueous complexation, mineral dissolution were also incorporated into the transport model.

Since the actual fluxes of the six assumed sources of contamination (four waste rock dumps and two pits) are not known, the groundwater contaminant concentrations measured at several monitoring locations were used as initial concentration data. The first step of contaminant transport modelling is to observe how the contaminant spreads over the field with varying groundwater heads. The groundwater contaminant concentrations measured at some monitoring locations were used in the reactive transport modelling. These concentrations were constant at selected monitoring points on specific days. Also, some of these contaminant concentrations were quite low and tended to decrease very quickly or even almost disappear after a few days, probably due to nominal immediate dilution, which is not realistic. To properly model and achieve realistic modelling, it was necessary to interpolate the known concentration data for the whole study area before using them in transport modelling. Hence, an interpolation of the concentrations was performed for the model domain area before conducting the contaminant transport modelling. Only a few points, which were unevenly scattered across the site, were available for the interpolation, but this is quite typical in hydrogeological studies.

The reason for interpolating contaminant concentrations is to provide a more realistic contaminant distribution as compared to a concentration values scattered even in a non-contaminated area whereas addressing the problem of concentration dilution. Furthermore, there was only one concentration measurement available for each location and each contaminant. Even if this value was entered as a transient contaminant concentration, it would be interpreted as a constant concentration and

cause the modelling of the contaminant process to be insignificant. Thus, even in a large time-scale simulation, such as that used in this study, contaminant transport would not be detectable. The contaminant concentration data obtained from interpolation were used only in the contaminant transport model as starting concentrations. Thus, the interpolated concentration data were used as initial concentrations. The reactive transport model was simulated (run) for a period of two years (from 2012 to 2014) due to the availability of data. Assumed initial conditions for contaminant concentrations were specified in the simulation model based on groundwater quality data from 2011. Copper (Cu^{2+}), sulfate (SO_4^{2-}), manganese (Mn^{2+}), uranium (UO_2^{2+}) and iron (Fe2+) were introduced as initial contaminants in this study.

Conceptual reactive transport modelling

The reactive transport model for the study area was built upon the flow model by implementing the necessary transport conditions. These include the transport boundary conditions, the total number of components and the species to be simulated. The assumption is that the transport of contaminants is based solely on the simulated flow fields, which may involve waste rock interactions at the four waste rock dumps (Figure 3.1a) or mineral interactions with the aquifer rock bed and formation of daughter products or additional metal species.

The conceptual model for the reactive geochemical system is based on equilibrium and kinetic reactions. The reactive system is entirely described by identifying chemical reactions and the total number of chemical species involved in them. The standard equilibrium reactions with appropriate equilibrium constant is used to represent all the fast reactions, like aqueous complexation reactions and the precipitation of secondary phases. Slow reactions are represented by kinetic reactions and associated rate constants. This addresses the dissolution reactions involving the major minerals occurring in the waste rock dumps and pits. The reaction network describing the geochemical system of the study area, and the associated rate constants obtained by modifying Yeh et al.'s (2004) equations, are shown in Table 3.2.

Table 3.2: Proposed Chemical Reactions During Contaminant Transport (Yeh et al., 2004)

Chemical reaction	Constant rate (log K)
$H_2O(aq) \rightarrow H^+ + OH^-$	-13.99
$H^+ + SO_4 \rightarrow HSO_4^-$	1.99
$Cu^{2+} + H_2O \leftrightarrow Cu(OH)^+ + H^+$	-9.19
$Cu^{2+} + SO_4^{2-} \rightarrow CuSO_4$	2.36
$Cu^{2+} + 2H_2O \rightarrow Cu(OH)_2 + 2H^+$	-16.19
$Cu^{2+} + 3H_2O \rightarrow Cu(OH)_3^- + 3H^+$	-26.9
$Fe^{2+} + H_2O \rightarrow H^+ + FeOH^+$	-9.50
$Fe^{2+} + SO_4^{2-} \rightarrow FeSO_4$	2.20
$Fe^{2+} + 2H_2O \rightarrow 2H^+ + Fe(OH)_2$ (aq)	-20.57
$Fe^{2+} + 3H_2O \rightarrow 3H^+ + Fe(OH)_3^-$	-31.00
$Fe^{2+} + 4H_2O \rightarrow 4H^+ + Fe(OH)_4^{2-}$	-46.00
$Mn^{2+} + SO_4^{2-} \rightarrow MnSO_4$	2.26
$Mn^{2+} + H_2O \rightarrow MnOH^+ + H$	-10.59
$Mn^{2+} + 3H_2O \rightarrow Mn(OH)_3^- + 3H$	-34.08
$UO_2^{2+} + SO_4^{2-} \rightarrow UO_2SO_4$	3.15
$UO_2^{2+} + SO_4^{2-} \rightarrow UO_2(SO4)_2^{2-}$	4.14
$UO_2^{2+} + 2H_2O \leftrightarrow UO_2(OH)_2$ aq $+ 2H^+$	12.15
$Fe(OH)_{3(s)} + 3H^+ \leftrightarrow Fe^{3+} + 3H_2O$	Kf = 0.05
$FeOOH_{(S)} + 3H^+ \leftrightarrow Fe^{3+} + 3H_2O$	Kf = 0.07
$FeOOH_{(S)} \leftrightarrow FeOH$	Kf = 0.05

All equations specified in Table 3.2 were derived by modifying equations from Yeh et al. (2004).

3.6 MODEL CALIBRATION

The groundwater model was calibrated by iterative adjustment of aquifer parameters and stresses to achieve the best match between the observed and simulated water

levels. A well-calibrated model accurately replicates hydrogeological conditions of real-world, which is the first goal of modelling. Therefore, the calibrated model can provide confidence in the predicted changes to the groundwater regime due to mining.

3.7 VALIDATION OF FLOW MODEL

To verify the performance of the groundwater flow and transport models, the transient groundwater model was run for an extended two-year simulation period to replicate groundwater levels from 2012 to 2014/2015. The developed flow model's outputs (hydraulic heads) were verified against field data from 2014/2015. The model was validated by comparing the solutions of the calibration simulation with a separate set of field measurements not utilized for calibration. It was not possible to calibrate the contaminant transport model as the contaminant source magnitudes, timings and locations could not be specified accurately. The results obtained for this validation exercise are presented in Figure 3.7; they match satisfactorily with field measurements and thereby demonstrate good predictive ability.

3.8 RESULTS AND DISCUSSION

3.8.1 Calibration of the Flow Model

For a model to be able to adequately and accurately simulate field parameters, it needs proper calibration. The aim of calibrating a model is to tune the hydrogeological parameters until the model approximates field measurements such as hydraulic heads and concentrations. The idea is to simulate the physical processes in the aquifer accurately. The flow model was calibrated for hydrogeological parameters and boundary conditions by running the forward simulation repeatedly and manually adjusting the input parameters selected for calibration, including boundary conditions, within their allowable ranges in each run until a satisfactory match between the modelled and field results was achieved. In this study, a trial-and-error procedure was used. It is worthwhile noting that the numerical simulation codes utilized in this study do not follow an automated calibration procedure such as PEST (Doherty, 2015). Manual trial-and-error calibration runs are conceptually straightforward and require intuitive judgment of the results obtained from multiple forward simulation runs. This

process is flexible, allowing logical adjustments of parameter values and structures, including changes in mesh designs and the representation of the geological framework. For this study, the purpose of calibration was to obtain hydraulic conductivity and recharge estimates for the modelled aquifer based on limited field measurements. Calibration was attempted without changing the hydrogeological zones defined by the distribution of hydrogeological units present at the site. Calibration was attained through refinement of the model parameters and features, including hydraulic conductivity in the horizontal and vertical dimensions, and recharge and hydraulic conductivity assigned to the sections used to replicate the influence of the fault.

Since the flow model simulation started in December 2010, groundwater head data measured from 20 monitoring locations in December 2010 were used to calibrate it. Throughout the calibration process, the hydraulic conductivity values of the different soil materials and rock types were varied within the acceptable range of field-measured hydraulic conductivity values in Table 3.1. Simulation runs were repeated until a reasonable match between the observed and estimated hydraulic head values was reached under the transient-state simulation conditions. Different values of hydraulic conductivity were obtained after calibration. The hydraulic conductivity values differed from layer to layer based on the varying material properties of each layer. Each layer had specific hydraulic conductivity values. The hydraulic conductivity values ranged from 0.1–5 m/day with the exception of the fault zone that cuts through the east side of the model, which had values as high as 75 m/day.

Hydraulic head measurements from 22 observation locations were used to calibrate the simulation model. Data from 2010 to 2012 were used to calibrate the flow model, while data from 2012 to 2014 were used to validate it. For the purposes of calibration, a percentage of annual rainfall was defined as a recharge value in the model. The calibration targets for the developed model were set to be within 2 m of the hydraulic head values observed at the observation locations.

Determining the exact boundary conditions of a model domain is a difficult task when measurement data are limited. Without exact or characteristic boundary conditions, a model may not accurately represent a field process. It is therefore necessary to implement realistic boundary conditions in a model that reflects the conditions of the site. One of the most difficult tasks in the calibration process is to properly assign the correct boundary conditions. The boundary conditions must

therefore also be determined on the basis of the preliminary results of the calibration. The boundary conditions of the model are manually adjusted to achieve the calibration targets. In this study, simulated hydraulic heads were compared with field-measured hydraulic heads at monitoring points. Groundwater generally flows from the higher elevations of the study area towards the central area, towards the East Branch of the Finniss River northwards.

The hydraulic head calibration results are illustrated by the bar graphs in Figure 3.4 (a), which compare the simulated and measured values. The same comparison is presented in Figure 3.4 (b) but as a line graph. A graph of the simulated and observed groundwater levels at selected monitoring locations after calibration is given in Figure 3.4. This figure shows the monitoring location of wells on the *x*-axis and the head values in metres on the *y*-axis. Figures 3.4a and 3.4b compare the observed and simulated heads for the two-year period of 2010–2012. Figures 3.4a and 3b show that there is an acceptable match between the field-measured and simulated heads. There is a small difference between the measured and simulated heads, which may be due to various errors and uncertainties in measurement, measurement/estimation of parameters, and boundary conditions. Variations within the model elements due to minor deviations in the hydraulic parameters, as is commonly found in groundwater modelling investigations, may be an additional cause. The deviation between the simulated and measured hydraulic heads does not exceed 5% of the field-measured values, so the calibration process can be said to have made the model reasonably approximate the observed groundwater head values. The calibration results, as a comparison between observed and simulated groundwater levels, are shown in Figure 3.5.

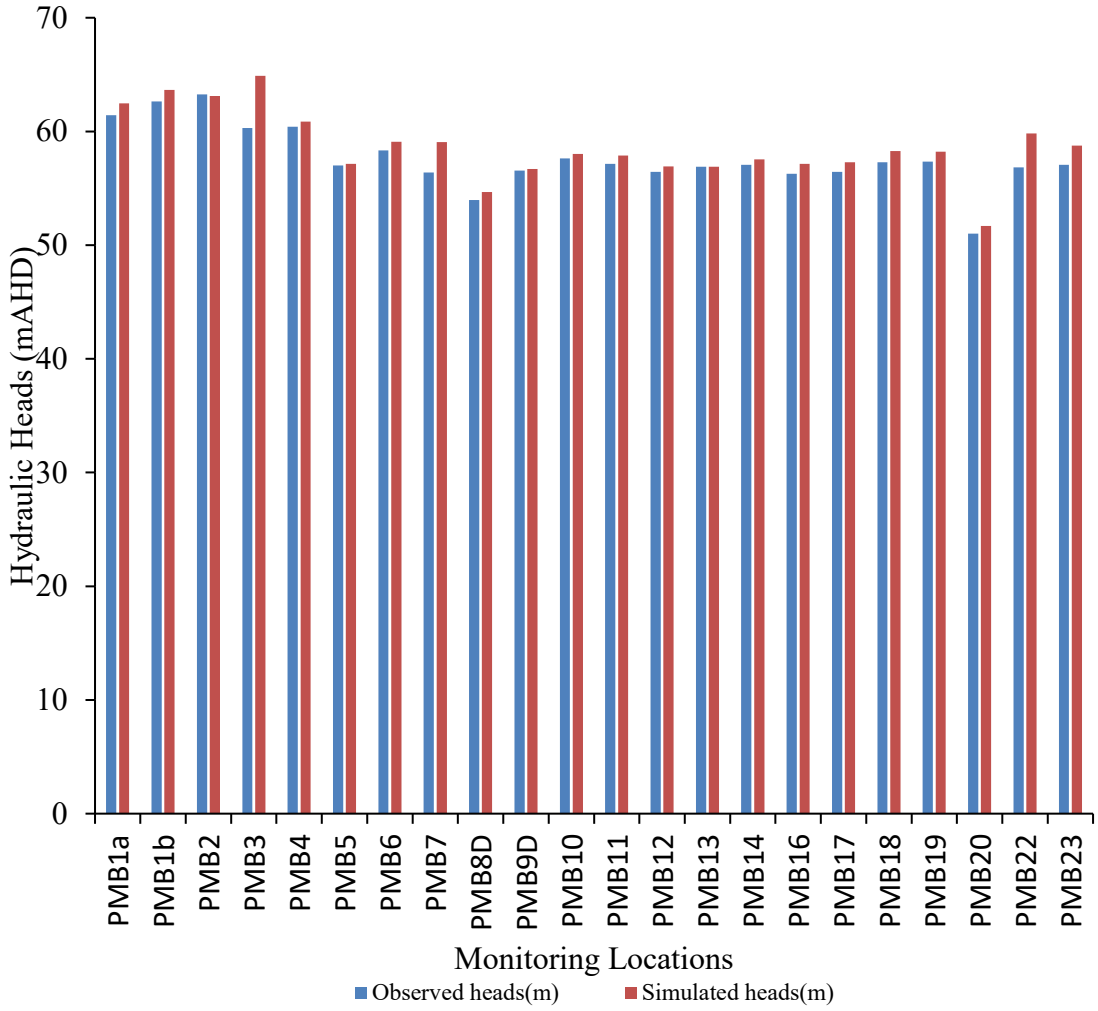


Figure 3.4 (a): Comparison of hydraulic heads observed in 2012 with those estimated by the calibrated model

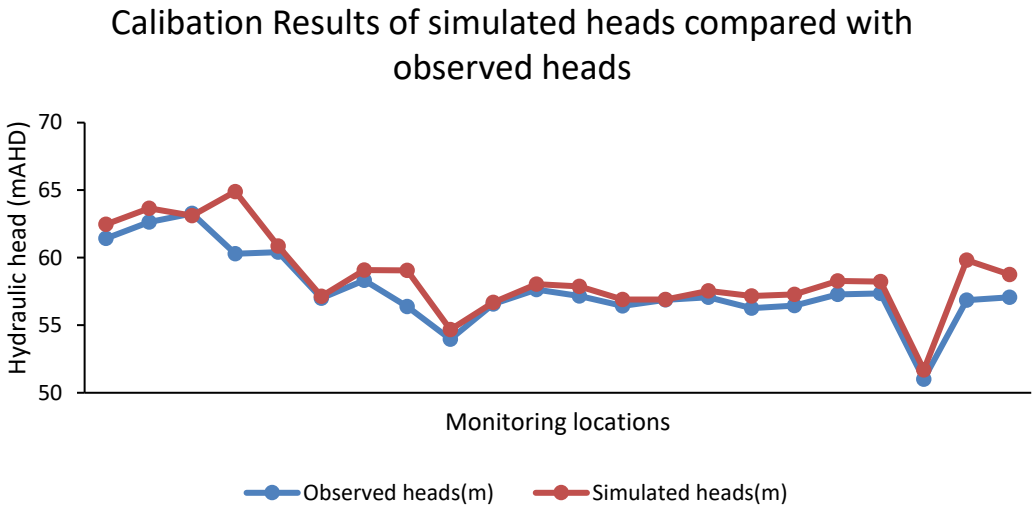


Figure 3.4 (b): Calibrated results of hydraulic heads for 2012

A scatterplot of simulated vs. observed values can be considered as a calibration graph. A plot for the calibration period (2010-2012) is shown in Figure 3.5. Statistical analysis of the calibrated model results shows that the residual mean (RM) groundwater levels at the monitoring locations during the calibration period ranged from 0.01 m to 2.97 m. The mean absolute error (MAE) was calculated as 0.82 m, and the standard deviation is 0.77 m. The normalized root mean squared (NRMS) was 0.09 and the root mean square error is 1.10. The correlation coefficient (R) is 0.90. From the scatterplot in Figure 3.5 and the comparison of observed and simulated hydraulic heads, we can see that the simulated head levels were within an acceptable range of the measured heads. Figure 3.6 shows the correlation between the measured and simulated heads with 95% confidence intervals on the mean observed and measured values.

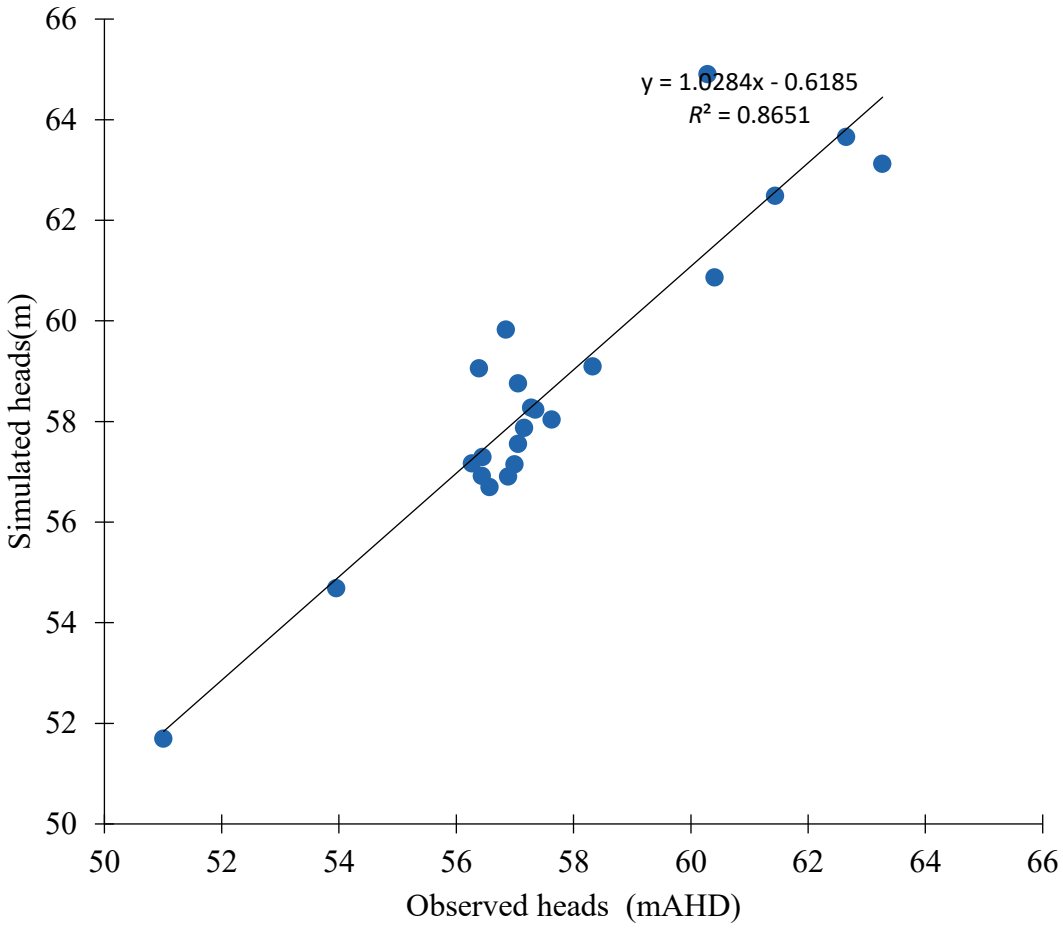


Figure 3.5: Correlation between simulated and observed groundwater heads

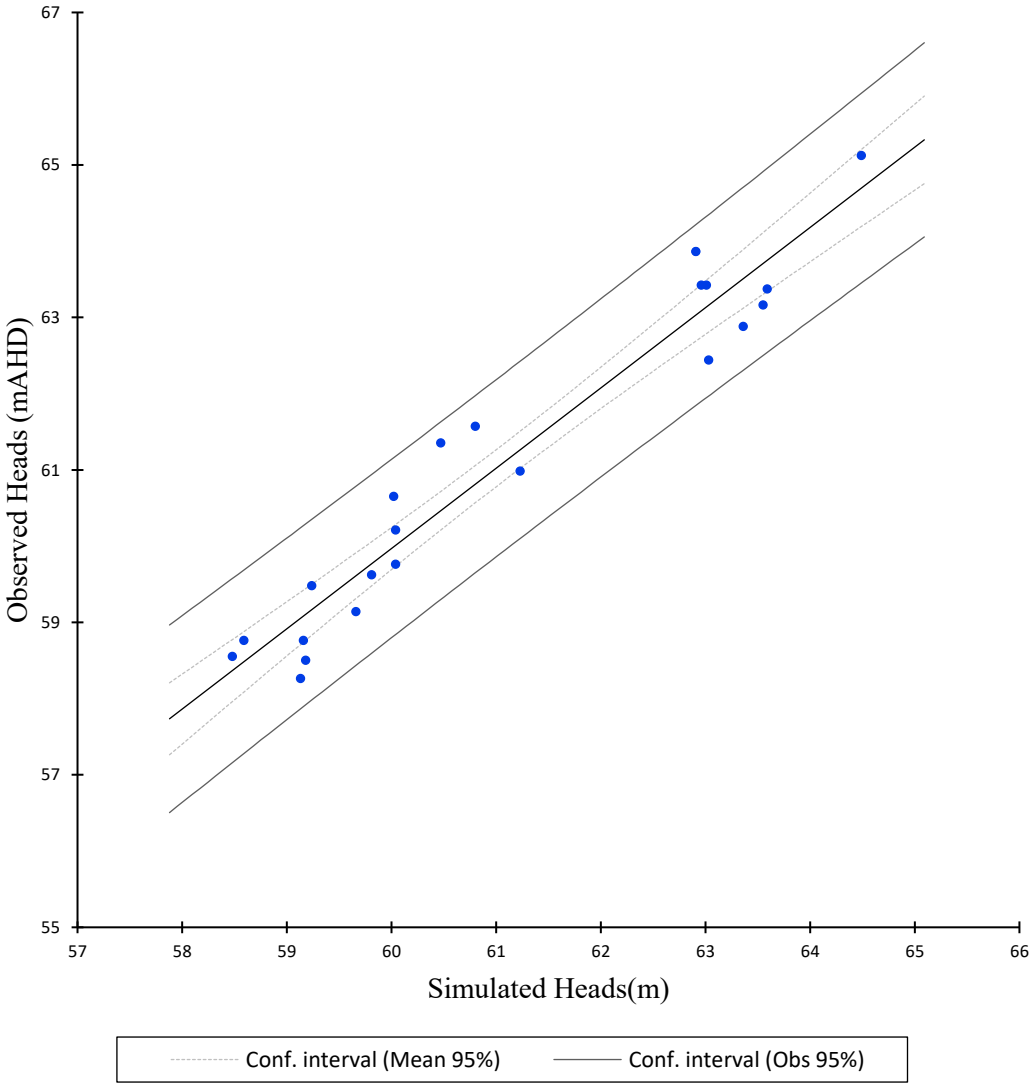


Figure 3.6. Linear relationship between measured and simulated hydraulic heads with 95% confidence intervals for the mean of the observed values and the observed values

The results of calibrating the simulated heads is presented in Table 3.3. The calibration of a numerical model is typically considered good if the NRMS error is < 5%. The computed NRMS values for the simulated heads are well below the target of 5%, suggesting good calibration to head targets and simulated hydraulic parameters. The residual average error for the total head data sampled at the 22 monitoring wells in 2010 is 1.0168. The heads range in value from 50 m to 70 m. The simulated heads at the monitoring points were then compared with the observed heads. Figure 3.4 shows bar charts that indicate a close relationship in the hydraulic heads. A Pearson correlation coefficient test was applied to the calibration results of the total heads. The

correlation coefficient (r) of 0.918 shows a close linear relationship between the observed and simulated heads.

Field data approximately corresponding to the first 738 days were used to calibrate the numerical model at all monitoring points. Field data from after this period were reserved for validating the numerical model. Figure 3.7 shows the evaluation of the numerical model's performance in the calibration period. The simulated groundwater heads predict the field-measured data reasonably well for the validation stage. Residuals between simulated and measured groundwater heads were also calculated by means of mean absolute error (MAE), NRMSE and RMSE. Table 3.5 shows the statistics for the calibration and validation periods. Both Figures 3.7 and 3.8 show that the numerical model maintains its calibrated accuracy throughout the validation period.

Table 3.3: Calibration Results for the Two-year Simulation Period (m AHD)

Monitoring well	Observed head (m)	Simulated head (m)	Residual (m)
PMB1a	61.44	62.48	-1.04
PMB1b	62.65	63.65	-1
PMB2	63.27	63.12	0.15
PMB3	60.29	64.9	-0.11
PMB4	60.41	60.86	-0.45
PMB5	57	57.14	-0.14
PMB6	58.33	59.09	-0.76
PMB7	56.39	59.05	-2.66
PMB8D	53.96	54.68	-0.72
PMB9D	56.57	56.69	-0.12
PMB10	57.63	58.03	-0.4
PMB11	57.16	57.87	-0.71
PMB12	56.44	56.91	-0.47
PMB13	56.89	56.9	-0.01
PMB14	57.06	57.55	-0.49
PMB16	56.27	57.16	-0.89
PMB17	56.45	57.29	-0.84
PMB18	57.28	58.27	-0.99
PMB19	57.35	58.23	-0.88
PMB20	51.01	51.69	-0.68
PMB22	56.85	59.82	-2.97
PMB23	57.06	58.75	-1.69

3.8.2 Validation of the Flow Model

The validation results for the validation period of 2012–2014 are shown in Figure 3.7. Table 3.4 presents the validation results in terms of hydraulic heads.

The availability of hydraulic head field measurements from 2012 to 2014 allowed the model to be validated over this period, using the 2012 measured head distribution as the initial condition. Simulation was carried out until 2014 using a time step period of 30 days. Figure 3.7 compares the hydraulic heads observed and simulated for this period, illustrating a wide correspondence. Thus, after transient-state

validation, the model is shown to simulate groundwater levels with a reasonable level of accuracy. Total groundwater head contours and direction of flow are shown in Figure 3.9.

Table 3.4: Validation of Hydraulic Head Values for 2015 (values in m AHD)

Monitoring well	Observed head (m)	Simulated head (m)	Residual (m)
PMB1a	63.42	63.01	0.41
PMB1b	63.86	62.91	0.95
PMB4	62.44	63.03	-0.59
PMB5	58.26	59.13	-0.87
PMB6	58.76	58.59	0.17
PMB7	58.55	58.48	0.07
PMB8D	59.48	59.24	0.24
PMB10	59.62	59.81	-0.19
PMB11	59.14	59.66	-0.52
PMB12	59.76	60.04	-0.28
PMB13	60.21	60.04	0.17
PMB14	61.35	60.47	0.88
PMB17	60.65	60.02	0.63
PMB23	58.76	59.16	-0.4
RN22039	61.57	60.8	0.77
RN22081	60.98	61.23	-0.25
RN22085	65.12	64.49	0.63
RN22543	58.5	59.18	-0.68
RN23051	63.16	63.55	-0.39
RN23413	63.42	62.96	0.46
RN23419	63.37	63.59	-0.22
RN29993	62.88	63.36	-0.48

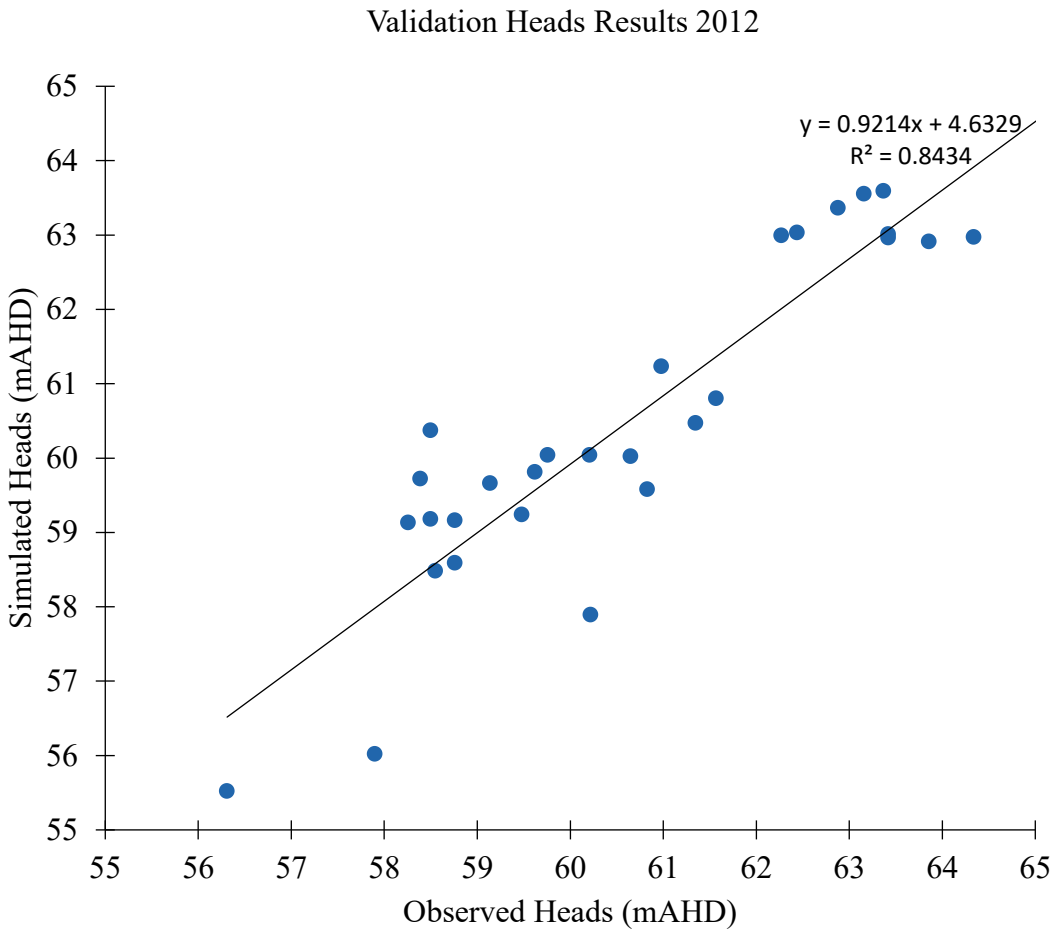


Figure 3.7: Correlation between observed and simulated groundwater levels during the validation period

The validation residual results between the observed and simulated groundwater levels are given in Figure 3.8. This graph shows that the residuals were typically less than 1 m (with the average depth of the aquifer being approximately 150 m) throughout the majority of the model domain, except in four monitoring well locations: PMB2, PMB9D, PMB16 and PMB 18. These had larger head residuals and tended to be in areas with a higher topographic relief and/or deeper water table, factors that tend to cause high seasonal fluctuations in groundwater levels. In general, these areas were more difficult to calibrate. Nevertheless, the spatial bias in head residuals was considered acceptable for the purposes of this study.

A summary and comparison of the calibration and validation results are given in Table 3.5. The maximum deviations in the predicted and measured hydraulic heads are 0.07 m and 0.95 m, respectively. Table 3.5 also shows that the deviations between the measured and predicted heads are more or less of the same order for both the

calibration and validation results. Therefore, given the situation of having limited measured data and a large and hydrogeologically complex study area with a multilayer aquifer, these validation results show that the calibrated model can be utilized for the prediction of hydraulic head under different input scenarios.

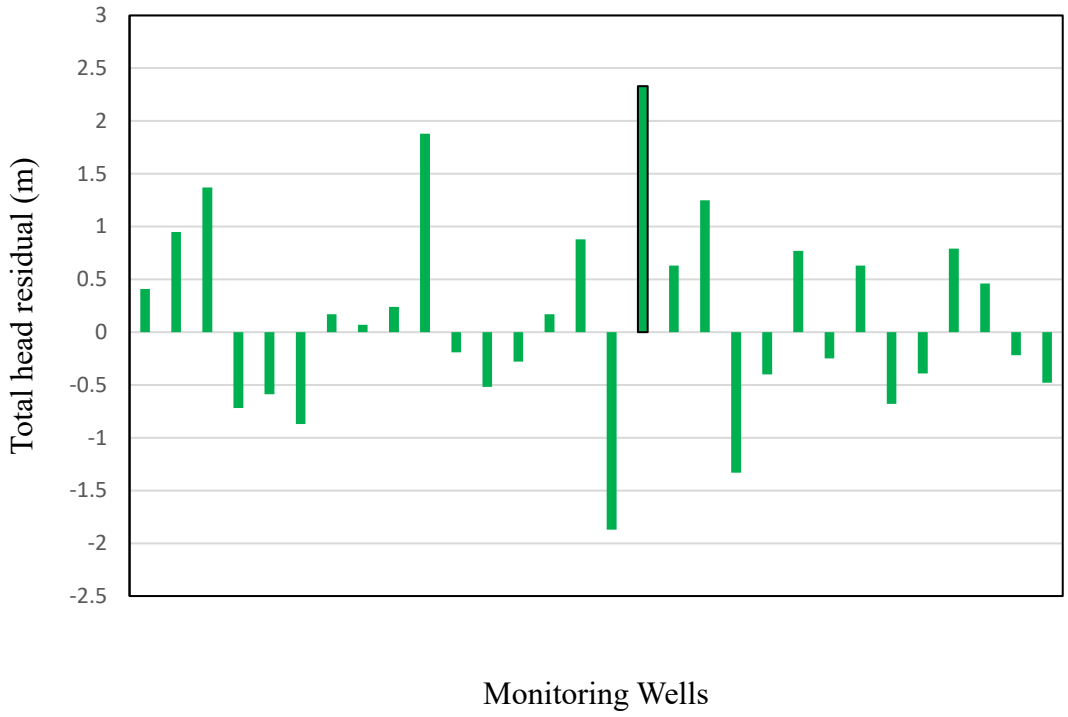


Figure 3.8: Errors between observed and simulated hydraulic heads during the validation period (2014)

Table 3.5: Statistics of the residuals for the calibration and validation periods

Statistic	Calibrated period	Validation period
Simulation period (days)	730	730
MAE (m)	0.82	0.72
Standard deviation (m)	0.77	0.91
NRMSE (m)	0.09	0.10
RMSE (m)	1.10	0.91

The numerical flow model for the Rum Jungle Mine Site was calibrated using annual groundwater level data measured at a selection of bores since August 2010.

Moreover, the validation process suggests that the current calibration provides a reasonable approximation of current flow conditions at the Rum Jungle Mine Site and can be used to predict the response of the groundwater system for rehabilitation planning.

Concerning the calibration process, it is worth noting that the calibrated values of hydraulic conductivity lie within the range of uncertainty in the values obtained by means of hydrogeological characterization. Average hydraulic conductivity values were used due to the complexity of the aquifer system. The nature of the aquifer requires that a layer has about five different conductivity values, which causes convergence issues. The model encountered challenges in converging when 14 distinct hydraulic conductivity values were applied, as average values were used to represent the heterogenous and anisotropic nature of the site as much as possible and to achieve model convergence.

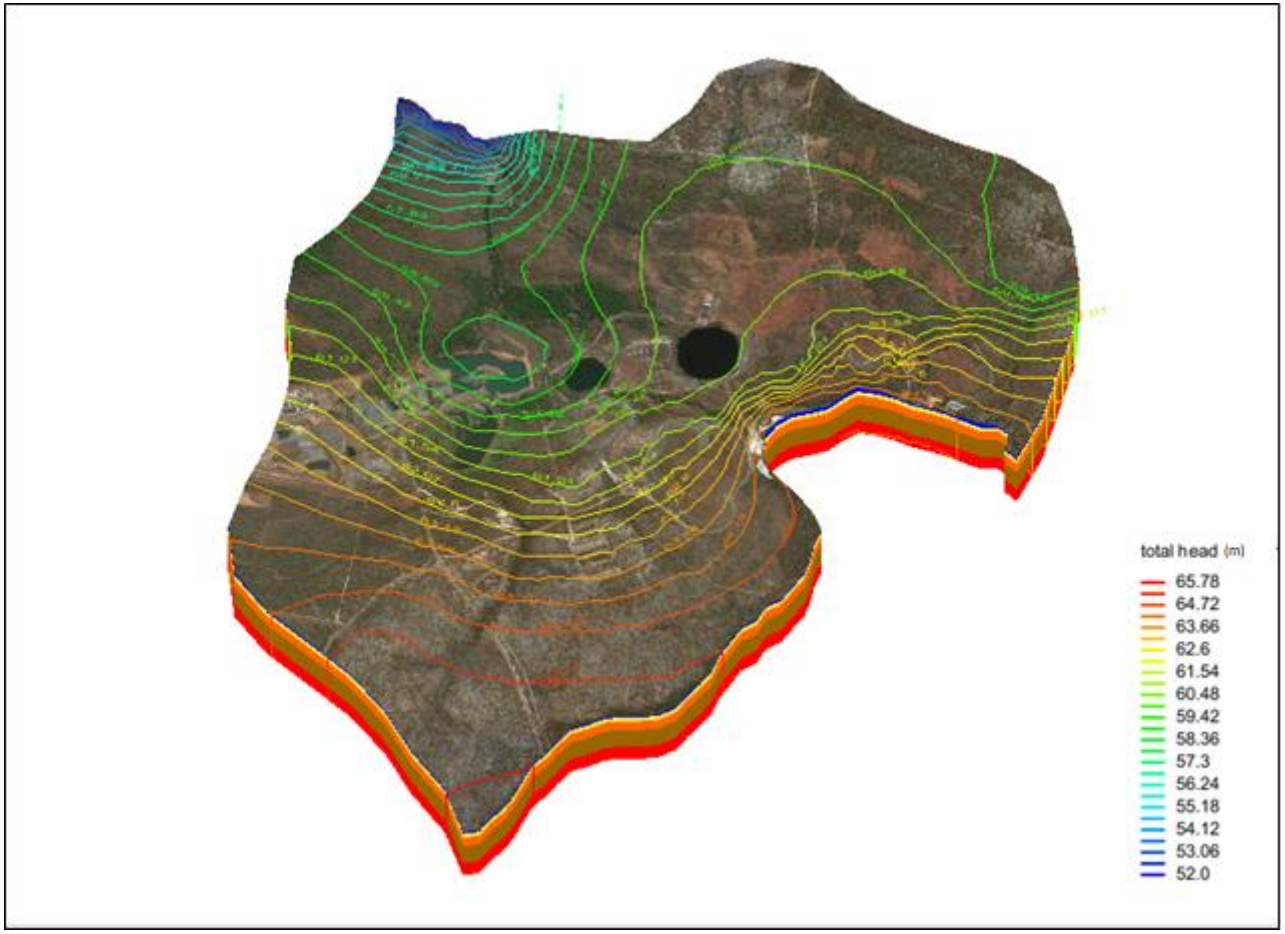


Figure 3.9: Three-dimensional model showing total head contours and the various material layers of the study area

Reactive Transport Modelling

Contrary to the groundwater flow numerical model, almost no calibration was possible for the reactive transport model as the contaminant sources were unknown. Monitoring locations were selected to compare measured and numerically-computed concentrations. The selected monitoring points correspond to all the available boreholes, with the concentration values being of the same magnitude as the groundwater contaminant concentration measurements made in the study area.

In most cases of reactive transport in contaminated aquifers at abandoned mine sites, transport of dissolved contaminants are transported in the liquid phase. These contaminants in the nature of chemical species can undergo chemical interactions with other existing dissolved chemical species. The solid species mostly include mineral phases (precipitates), exchanged species (from the change complex) and absorbed species. These species undergo transportation through the porous medium, and also undergo geochemical processes. A set of equilibrium and kinetic reactions are used to represent the geochemical processes listed in Table 3. The species of focus in this abandoned uranium mining site are Cu, SO_4, Fe, Mn and U, which are considered reactive species in the transport process. Essentially, the contaminant transport model cannot be calibrated, as that would require accurate information on contaminant sources (magnitude, location and duration). Due to the complex hydrogeology of the site, reasonable assumptions regarding recharges, boundary conditions, and initial head and concentration values were used iteratively to obtain an acceptable calibrated flow and transport model. The results obtained from reactive transport simulation of a two-year period reveal the presence of geochemical reactions in this aquifer.

It is worth to note that the numerical model relied on previous work by RGC (2012, 2013, 2014, 2016), which was used to create the model domain and define the initial conditions. The dimensions and geometry of the different hydrogeological zones used in this study are different from those of previous works, which considered a grid type of discretization; finite difference method while the present study triangular wedge discretization; employed finite element method. The recharge rates of the model were defined according to the work of Ferguson et al. (2011) and the groundwater chemical composition and hydraulic parameters were obtained from RGC (2016). The opinions expressed in this chapter have been based on information contained in the

RGC reports made by the Northern Territory Department of Resources (DOR). The opinions in this work are provided in response to the objectives of this study. The author has exercised all due care in reviewing the information in the RGC reports. Whilst this research has compared key supplied data with estimated values, the accuracy of the results and conclusions from this work are entirely reliant on the accuracy and completeness of the data used.

3.9 CONCLUSIONS

This study outlined the development and calibration of a transient groundwater flow model for a contaminated aquifer underlying an abandoned uranium mine in the Northern Territory of Australia. This work was initiated as a fundamental step to understand the flow of the aquifer system and to establish a multiple species reactive transport model for a hydrogeologically-complex contaminated aquifer. The outcome of the model is the basis for a transient flow model and for reactive transport and predictive modelling. The implementation of this model was based on subjective judgment of the selection of appropriate data, due to its sparseness and reliability. The numerical flow model was calibrated a two-year period of 2010 to 2012. The calibrated model was also validated according to selected hydrogeological parameters with data from 2012 to 2014. The calibrated flow model and transport model were used to simulate the heads and concentrations at various points in the study area. Overall, the calibrated model provided a reasonable match with field observations, demonstrating strong hydraulic connection to the materials at each layer. This validation process suggests that the current calibrated model is a reasonable approximation of the current flow conditions at the Rum Jungle Mine Site and can be used to predict the responses of the groundwater system.

The challenge in using reactive transport models for modelling real-life scenarios is in selecting the geochemical reactions that best describe the geochemical processes occurring at the study site and transforming them into transport equations of 1) kinetic variables, 2) components and 3) equilibrium variables. These can then be used for the computation of equilibrium and kinetic rates to facilitate numerical contaminant transport simulations.

There was an acceptable level of agreement between the observed and simulated hydraulic heads. Transient simulation of the groundwater hydraulic heads and movement of reactive contaminants was accomplished with the developed simulation model. The concentrations predicted for the species in the reaction network are of the same order of magnitude as those of available measurements. Quantification of the contaminant sources in terms of magnitude, location and duration of activity was not possible. Therefore, the contaminant transport simulation calibration is, at best, subjective in this scenario. Hence, the simulation models, once reasonably calibrated to site conditions, are potentially good approximations, and similar approaches can be used for similar sites with similar complex challenges.

The next chapter presents the design and implementation of a contaminant source identification methodology. This has the potential to increase the accuracy of characterizing distributed sources in a complex contaminated aquifer.

Chapter 4: Application of Contaminant Source Identification Methodology to a Mine Site

This chapter discusses the application of a source identification method for characterizing contaminant sources at a contaminated mine site aquifer containing complex multiple reactive species. The performance evaluation of the method is assessed using synthetic data and limited field data.

4.1 BACKGROUND TO THE PROBLEM

The remediation of contaminated aquifers is a challenge task in groundwater resource management. Effective and reliable management of groundwater resources first requires the identification of the contaminant sources (Datta & Kourakos, 2015). Numerous problem-solving approaches have been proposed in recent decades to address contaminant source problems (Atmadja & Bagtzoglou, 2001; Amirabdollahian & Datta, 2013). Among these approaches, linked simulation-optimization models have been progressively applied to identify groundwater contaminant source characteristics in contaminated aquifers. However, over the years, most researchers studied point sources or ideally shaped non-point pollution sources. More so, such studies considered the contaminants in the transport model as non-reactive in a homogeneous geological media whiles solving the unknown source identification optimization problem. Even though such models can be simulated by considering ideally-shaped or point sources, they cannot be used to determine the characteristics of distributed groundwater contaminant sources with reactive contaminants. Furthermore, in highly heterogeneous geological media involving geochemical reactions (both kinetic and equilibrium reactions) of reactive contaminant species, the simulation of the transport process becomes more complex and difficult.

The optimal characterization of the contaminant sources requires accurate simulation of the flow and transport processes occurring in the contaminated aquifer. The contamination scenario at an operational or abandoned mine site is generally very

complex due to geological heterogeneity, distributed sources and the presence of multiple reactive species in the mined mineral ores. In addition, at abandoned mine sites, monitoring of spatially- and temporarily-varying hydraulic heads and concentrations is usually very sparse and inadequate. Optimal characterization of contaminant sources also requires the use of an accurate flow and transport simulation model. Accurate and reliable identification of contaminant sources at an abandoned mine site is, therefore, especially complex and difficult. Source characterization is the first step towards reliable and sustainable contamination remediation. The issue of accurately characterizing contaminant sources at poorly monitored abandoned mining sites is crucial since the soil matrix and groundwater contamination have important influences on human health, vegetation and ecological systems. Thus, solving contaminant source problems in complex aquifers characterized by many aquifer parameter dissimilarities requires methods that are robust, efficient and able to handle data uncertainty.

To optimally identify spatially-distributed groundwater contamination sources, the source flux, activity, duration and time of initiation need to be determined. However, when multiple reactive chemical species are present as contaminants, identification of the sources in terms of the individual species involved is required before developing a remediation strategy.

Previous distributed contamination source characterization models for typical mine sites were reported in Jha and Datta (2015a). They developed a linked simulation-optimization-based methodology for estimating the release histories of spatially-distributed fixed pollution sources at an illustrative abandoned mine site, but the pollution sources were considered as conservative. In their method, adaptive simulated annealing (ASA) was used as an optimization algorithm to determine source concentrations at the pit. Similarly, Ayvaz (2015) developed a genetic algorithm-based simulation-optimization model to determine the spatial distributions and source fluxes of areal groundwater pollution sources. This model was evaluated on a simple hypothetical aquifer model under ideal conditions. Eshafani and Datta (2016) developed genetic programming models as surrogate models for the characterization of distributed contaminant sources at a contaminated mine site. Genetic programming-based trained surrogate models were used to approximate complex transport processes

involving reactive species. However, this study did not individually characterize the constituents of multiple contaminant species.

The objective of the present research was to develop a linked simulation-optimization approach to solving distributed groundwater pollution source identification problems in a complex groundwater system with emphasis on multiple reactive species. To the best of our knowledge, the potential of this approach for addressing multiple reactive species contaminant sources is still unexplored.

For this work, it was assumed that the observed concentrations are measured at several monitoring locations, dispersed in space, and monitoring contaminant transients over a period. If there are multiple contamination sources in an aquifer, each monitoring location detects a mixture of contamination fields (plumes) originating from different locations. It is assumed that each contaminant source releases a different geochemical constituent that is mixed in the aquifer, and that the resultant mixture is detected at the observation locations. Also, the geochemical constituents are reactive, and their transport is impacted by geochemical reactions or other fluid/solid interactions in the porous media where the flow occurs. The aim is to identify the number, locations and activity times of contamination sources.

The scenario of having very limited flow and concentration measurement data was incorporated to represent the typical situations of contaminated sites. The effective integration of a source identification optimization modelling technique with an accurate simulation of contaminant transport from distributed sources with complex pollutant geochemistry is addressed. The proposed approach is then evaluated for efficiency, accuracy and applicability. In the proposed approach, the groundwater flow and reactive transport processes are simulated by modelling the aquifer system of an abandoned uranium mine located in the Northern Territory of Australia. The three-dimensional coupled physical and chemical transport process simulator HYDROGEOCHEM 5.0 (HGCH; Yeh et al., 2004) is used to numerically simulate the flow and reactive chemical transport processes.

An initial attempt to optimally identify groundwater contamination sources was made by Gorelick (1983) using a combination of linear response matrixes and linear programming. This was followed by an innovative approach of using statistical pattern recognition (Datta et al., 1989) to characterize sources in space and time. These source identification methods were extended to incorporate different scenarios and to estimate

flow and transport model parameters. Statistical methods were also proposed to solve the unknown source characterization problem; i.e., use of nonlinear maximum likelihood estimation to solve the inverse problem of estimating contaminant sources in groundwater aquifers. (Wagner, 1992). Source identification using embedded nonlinear optimization techniques was proposed by Mahar and Datta (1997, 2000, 2001). Other researchers proposed genetic algorithm (GA)-based optimization algorithms (Aral et al., 2001). Singh and Datta (2006) proposed use of an artificial neural network (ANN) approach. Use of classical optimization for solving optimal source identification problems was proposed by Datta et al. (2009). Source identification based on heuristic harmony was proposed by Ayvaz (2010). Also, simulated annealing (SA)-based linked optimization algorithms for source identification were proposed by Jha and Datta (2011), Chandalavada et al. (2012) and Prakash and Datta (2014a, 2015); and a differential evolution algorithm for groundwater source identification was proposed by Ayvaz (2014) and Gurarslan and Karahan (2015). A linked simulated annealing-based optimization model with a PHT3D simulation model for chemically reactive transport processes was proposed by Datta et al., (2017). Self-organizing map-based surrogate models was proposed by Hazrati & Datta, (2017); GP surrogate models for source identification in chemically reactive medium was proposedby Esfahani & Datta (2017). Reviews of optimization techniques for solving source identification problems are presented in Chandalavada et al. (2011), Amirabdollahian and Datta (2013), Ketabchi and Ataie-Ashtiani (2015) and Sreekanth and Datta (2015).

In this chapter, the objectives are three-fold. The first objective is to evaluate the performance of an inverse source identification problem formulation to identify contaminant source characteristics based on a synthetic case study with highly complex hydrogeological and aquifer properties. This is important because, despite the advantages of using a simulation-optimization formulation to solve an inverse unknown groundwater problem, this approach has largely been applied to statistical type heterogeneities, knowing well how the type of heterogeneities largely influence mass transport.

The second objective is to test the efficiency and advantages of using an adaptive simulated annealing optimization algorithm to demonstrate the feasibility of characterising multiple species concentrations from distributed sources. ASA is a

global optimization algorithm that depends on randomly sampling important parameter space. In contrast to the deterministic approaches, the exponential annealing schedules in the ASA allow resources to be used adaptively on re-annealing and convergence in all dimensions, guaranteeing extensive global search in the first phases of the search and quick convergence in the final phases of the projected stopping criteria in the formulated problem of source identification. While ASA optimization has been applied to a variety of optimization problems, we believe that this is the first time it has been applied to the characterisation of multiple species concentrations in a contaminant source identification problem.

The third objective is to provide an optimization benchmark case that allows optimization strategies on objective functions defined over a discrete domain and inspired by real applications, which are not commonly available in the optimization community. The performance of the developed method is evaluated using a real complex aquifer system. It initially uses synthetic data (to be able to evaluate the performance for different variations of the scenarios in the field) on the multiple species concentration identification at the distributed sources. The performance evaluation is carried out using limited concentration data obtained at the study site to characterise the contaminant sources. The developed method is applied to an abandoned uranium mining site in the Northern Territory, Australia, and an associated contaminated groundwater system.

4.2 METHODOLOGY

In this section, a general mathematical formulation of the source identification problem and its notation are presented. The proposed linked simulation-optimization method of source characterization has a two-phase structure of numerical simulation and optimization.

4.2.1 Groundwater Flow and Transport Simulation

The governing equations for groundwater flow and reactive transport are extensively detailed in Chapter 3, Section 3.2.

The first phase consists of a numerical simulation of the physical processes of flow and reactive transport in the groundwater system. To solve the source

identification problem, the governing equations of groundwater flow and transport are solved to accurately represent the flow and transport processes occurring in the contaminated aquifer system. The simulation of these processes requires data on the hvydraulic head fields and contaminant concentrations so that the governing groundwater flow and transport equations, respectively, can be solved. Multiple reactive transport and mixtures of contaminant plumes in an aquifer present an intricate problem likely influenced by, but not limited to, complexation, precipitation-dissolution, adsorption-desorption, advection, dispersion and diffusion, and the chemical processes of aqueous, ion-exchange, redox and acid-base reactions.

This study used contaminant concentration measurements for the mine site that were obtained by a previous study (RGC, 2012, 2016). These data are very limited in number and only cover two years of monitoring. The concentration calibration results of the present study were compared with these limited measurement data. No other concentration data was available. This study illustrates the limitations in modelling flow and transport processes at such a site with very limited field measurements. The limitations in concentration data, and the very limited knowledge on aquifer parameter values, were a challenge to the development of a well-calibrated simulation model. Such a challenge is common in this area of research. This is also one of the reasons why a large number of modelling iterations with different assumptions of recharge, boundary conditions, and initial head and concentration values was needed to achieve an acceptably calibrated model. The implementation of flow and transport simulation models should be considered in light of such limitations and the challenges common to contaminated aquifer sites such as the current one.

4.2.2 Adaptive Simulated Annealing Optimisation Algorithm (ASA)

The second step involves using an optimization algorithm to find optimal candidate solutions. The present study uses an ASA optimization algorithm in the optimal source characterization model. This algorithm is preferred for its comparative efficiency in reaching a global optimal solution. The optimization algorithm is used hereafter to minimize the objective function formulation.

The ASA global optimization algorithm relies on random importance-sampling of parameter space (Ingber, 1993a). It was created with the objective of speeding up

the convergence of standard SA methods (Ingber, 1989, 1993, 2017; Chen & Luk, 1999). The basic structure of the ASA algorithm is the same as that of classical SA. There are, nevertheless, some key differences. It has new distributions for the acceptance and state functions and a new annealing schedule. It uses independent temperature scales for each fitted parameter and for the acceptance function. It also performs reannealing at specific intervals. The ASA algorithm maintains the advantages of SA but converges faster. Hence, the ASA algorithm is a powerful global optimization tool for solving complex parameter estimation problems.

The major advantage of ASA is that the algorithm parameters are modified in an adaptive way and that the solutions do not differ greatly if the parameters are updated within acceptable limits. This contrasts with other optimization algorithms, where only small variations in parameters, such as mutation probability, crossover probability, or population size, cause major differences in solutions. Additional benefit of ASA over SA is that it overcomes the speed issue of traditional SA approaches and ensures fast convergence towards a global minimum solution. A comprehensive discussion of the ASA formulation is provided in the literature review chapter (Chapter 2).

Based on the principles of ASA, the following set of instructions were set and performed repeatedly until the stopping criterion was met.

Set iterative criteria: Criteria are set to decide whether to accept the new state when a new current is selected. The objective function value of the original state and the new state are determined and the new state is judged according to the above criteria.

Initial temperature T_0: The appropriate initial temperature is selected by attempt.

Termination temperature T_E: A smaller value is selected as the end temperature.

Cooling rate: The temperature from a certain state T_k to the next state T_{k+1} should satisfy: $T_{k+1} = \alpha T_k$, where $\alpha \in (0,1)$. Cooling is slow when α is large.

Selection of neighbour state: A new solution is generated by random variation within a certain neighbourhood.

The whole simulated annealing process outlined above is summarised in the flow diagram in Figure 4.1.

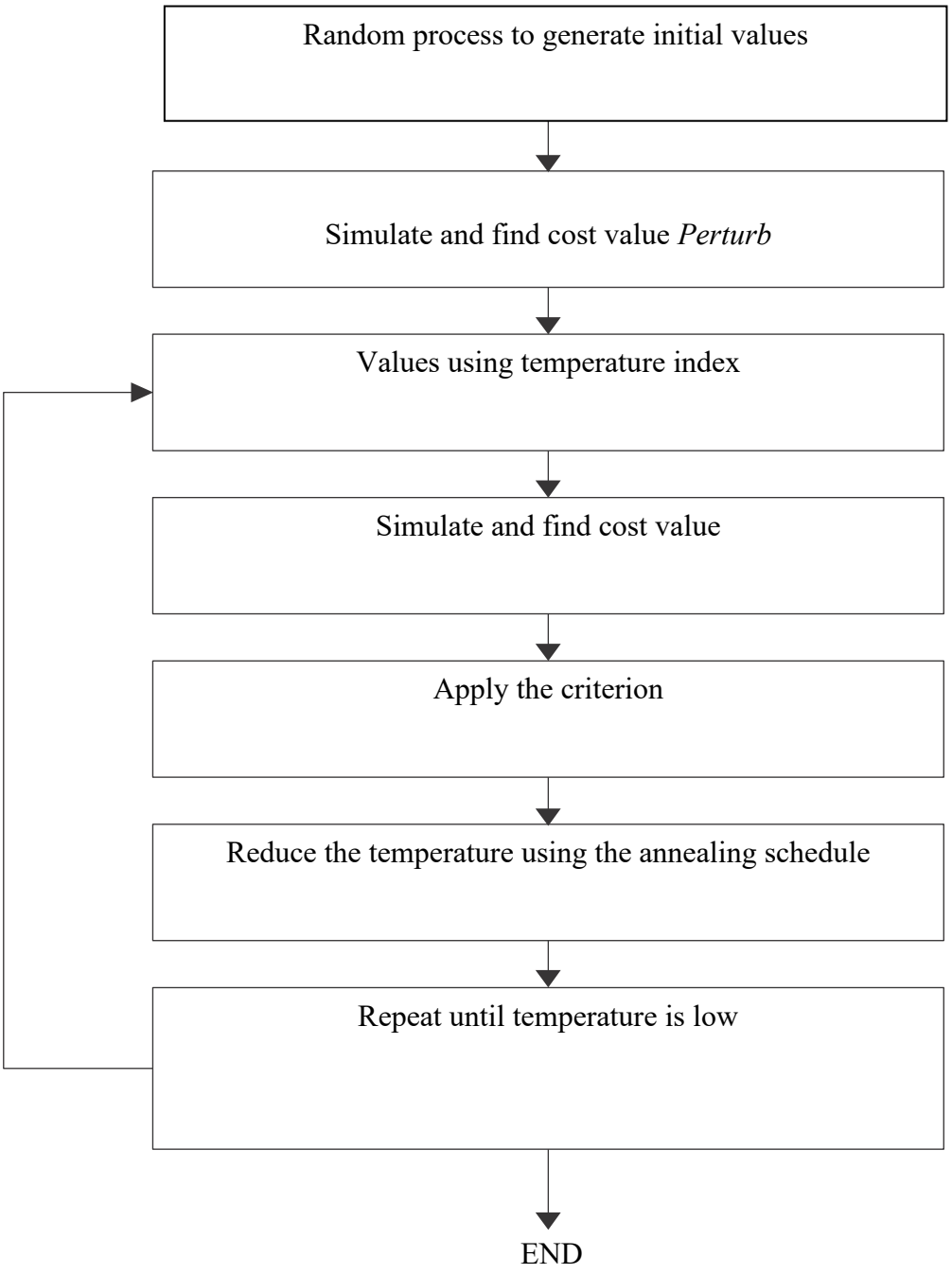


Figure 4.1: Flow diagram illustrating the operations involved in the simulated annealing (SA) process

4.2.3 Source Identification Using Simulation-Optimization

The simulation-optimization model simulates the physical processes of flow and reactive transport within an optimization model. The flow and reactive transport simulation models are considered as important binding constraints for the optimization model. Therefore, any feasible solution of the optimization model needs to satisfy the flow and transport simulation models. The advantage of this approach is that it is possible to connect any complex numerical model to the optimization model. In this identification model, the flow and transport simulation models are linked to the optimization model using the ASA algorithm to find the solution.

In simulation-optimization models, the groundwater contaminant source identification problem is formulated as a forward-time simulation in combination with an optimization model. The simulation-optimization model simulates the physical processes of flow and reactive transport within the optimization model. The flow and reactive transport simulation models are treated as important binding constraints for the optimization model to ensure that the simulated source responses fare properly simulated. Therefore, any feasible solution of the optimization model is expected to satisfy and based on the implemented flow and transport simulation models. In this approach, several forward-time simulations of the groundwater reactive transport equation are solved with different sets of potential sources and their respective strengths. The solutions of these forward runs are compared to measured spatial and temporal concentrations of the contaminant species. The optimization model solution selects a set of sources(strength, location etc.), which results in the simulated concentrations closely matching the species concentration measurements. The advantage of this approach is that it becomes possible to incorporate any complex numerical simulation model to the optimization model. In the source identification model, the flow and transport simulation models are linked to the optimization model using the ASA algorithm to find a solution. Figure 4.2 shows a flowchart of the simulation optimization and program execution sequence for groundwater contaminant source identification.

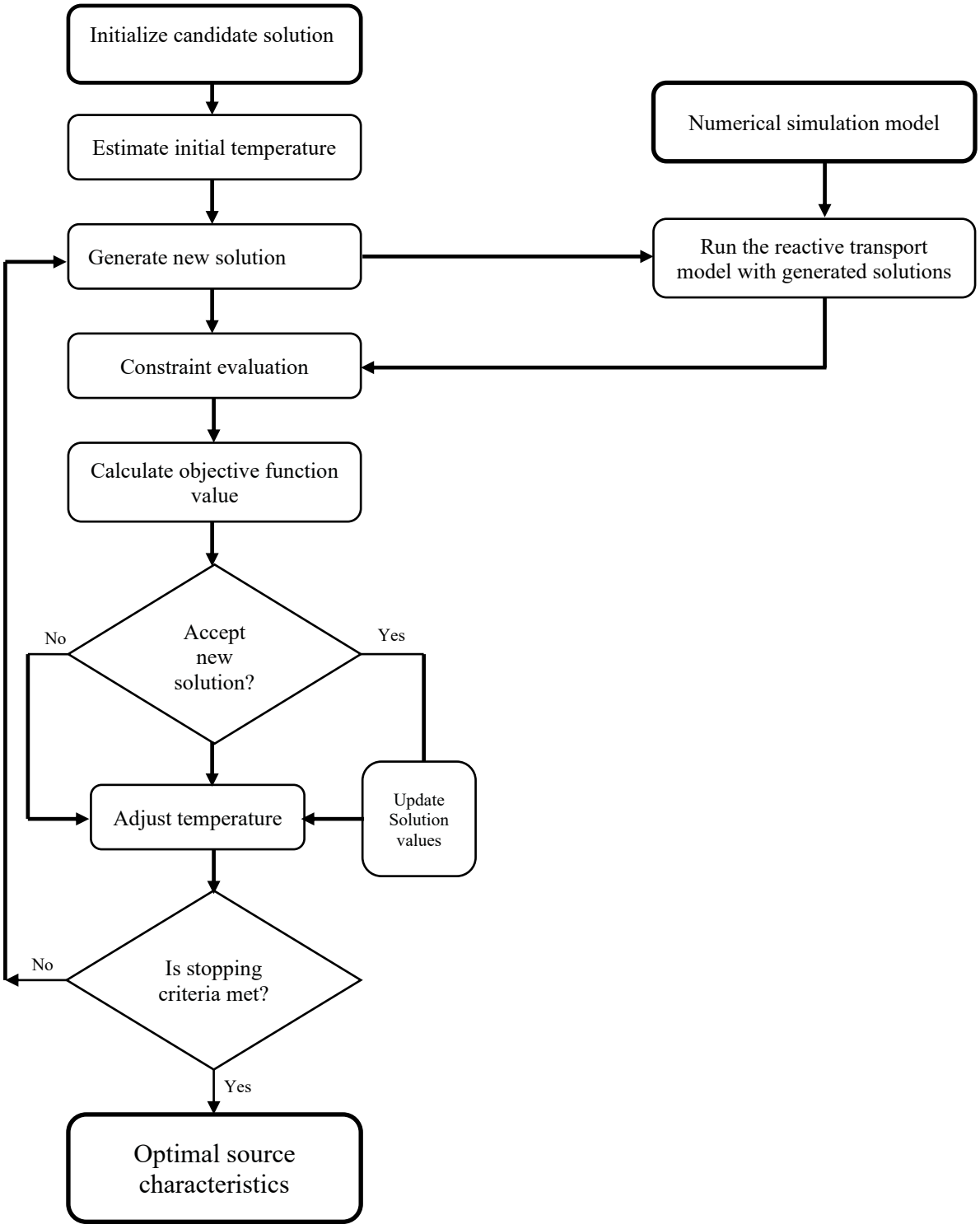


Figure 4.2: Schematic illustration of the linked simulation-optimization model using adaptive simulated annealing

4.2.4 Mathematical Formulation of Simulation-Optimization Models

The main goal of the contaminant source identification model is to characterise each source according to the geochemically-reactive species that are reacting in the aquifer. The source characteristics of interest include location and release duration and magnitude. The source identification model uses the optimization approach to provide candidate solutions for a set of source characteristics. It minimizes a weighted objective function of the differences between the simulated and observed species concentrations at monitoring locations within the model domain.

The optimization model generates candidate concentrations of species associated with each potential distributed source location. In this case, five candidate solutions of species concentrations are generated by the optimization algorithm at six separate potential distributed source locations. These candidate concentration solutions generated by the optimization algorithm are utilized to estimate spatial and temporal contaminant species concentrations in different time periods for monitoring locations at which field concentrations have been measured. Appropriate constraint conditions can be imposed on the parameters of the model. The optimization algorithm then evaluates the objective function. The objective function value is defined as a function of the differences between the observed and simulated concentrations of different reactive species at monitoring locations in different time periods. Optimal source characterization is obtained by solving the optimization model to minimize the objective function. The objective function of the simulation-optimization model used for source characterization was formulated as follows:

$$\text{Mininimze F} = \sum_{k=1}^{n_k} \sum_{iob=1}^{n_{ob}} \left(C_{sp1_{obs_{iob}}}^{k} - C_{sp1_{sim_{iob}}}^{k} \right)^2 \bullet w_{iob}^{k} + \sum_{k=1}^{n_k} \sum_{iob=1}^{n_{ob}} \left(C_{sp2_{obs_{iob}}}^{k} - C_{sp2_{sim_{iob}}}^{k} \right)^2 \bullet w_{iob}^{k} + \sum_{k=1}^{n_k} \sum_{iob=1}^{n_{ob}} \left(C_{sp3_{obs_{iob}}}^{k} - C_{sp3_{sim_{iob}}}^{k} \right)^2 \bullet w_{iob}^{k} + \sum_{k=1}^{n_k} \sum_{iob=1}^{n_{ob}} \left(C_{sp4_{obs_{iob}}}^{k} - C_{sp4_{sim_{iob}}}^{k} \right)^2 \bullet w_{iob}^{k} \quad (4.1)$$

$$\text{Subject to: } C_{sim_{iob}}^{k} = f\left(x, y, z, C_{sim}\right) \quad (4.2)$$

$$\text{The weight } w_{iob}^{k} \text{ can be described as: } w_{iob}^{k} = \frac{1}{\left(C_{obs_{iob}}^{k} + \eta\right)^2} \quad (4.3)$$

Where $C^{k}_{sp\,obs_{iob}}$ is the observed concentration of a species at monitoring location *iob* in the k^{th} time period;

$C^{k}_{sp\,sim_{iob}}$ is the concentration of a species estimated by the source identification model at monitoring location *iob* in the k^{th} time period;

sp1, sp2, sp3, sp4 are species numbers one, two, three and four, respectively, involved in the chemical reaction;

n_{ob} is the total number of concentration-observation locations;

nk is the total number of concentration observation time periods;

nspe is the total number of species involved;

$f\,(x, y, z, C_{sim})$ represents the concentration simulation results at location coordinates defined by x, y, z obtained from the transport simulation model.

w^{k}_{iob} is an assigned weight corresponding to observation location iob and time period k; and

η is an appropriate constant that is the average of the highest and lowest concentrations of each species. This ensures that errors at low concentrations do not dominate the solution.

4.2.5 Development of a Linked Simulation-Optimization Model for Source Characterization based on Multiple Species of Contaminants

A finite element-based three-dimensional numerical simulator (HYDROGEOCHEM 5.0) was used to simulate the flow and transport processes in the study area aquifer. Hydrological variables, including Darcy's velocity and moisture content, are necessary in determining the transport of contaminants through saturated-unsaturated subsurface systems. These variables need to be specified to solve the basic governing equations in a simulation model. These variables can be iteratively estimated by calibration of a simulation model.

A linked simulation-optimization approach linking the groundwater numerical simulation model with an optimization model that incorporates an ASA algorithm was implemented. Integrating a numerical simulation model with an optimization

algorithm results in significant performance improvement over source identification results obtained through conventional standalone simulation or optimization methods. In the linked simulation-optimization methodology, the first step is to develop a groundwater simulation model that computes head and species concentration values at different monitoring locations in different time steps. In the second step, the simulation model is externally linked with the optimization model. Whenever the optimization procedure requires the objective function and/or constraint evaluation, it calls the simulation model while passing the candidate solutions to the simulator. Then the simulation model executes and returns the resulting concentrations. An ASA algorithm acts as a driver model that calls the simulation model by passing variables and gets back the corresponding objective function value. The ASA then adjusts the variables to compute a new objective function and continues for several iterations until there is no further improvement or the stopping criteria are satisfied. The ASA used as the optimization routine calls the calibrated simulation model during each iteration.

A computer code was written in C++ language to interface the ASA and calibrated groundwater simulation model, thus facilitating communication between the simulation and optimization models. The C++ based code acts as a subroutine that has a set of instructions designed to communicate between the FORTRAN-based numerical simulation program and the C-language-based optimization program of ASA to perform frequently used operations within the linked simulation-optimization methodology.

For the simulation and optimization models to work efficiently together to optimally solve the source identification problem, they must be interfaced. This requires the design and implementation of a linking script that facilitates communication (recursive calls) between each module, hence the C++ code. Each time an optimization model requires a function evaluation or constraint evaluation, it calls the simulation model. Figure 4.1 shows the program's execution sequence. The ASA algorithm starts from an initial guessed solution given by the user. Candidate source concentrations are generated randomly by the ASA algorithm as possible solutions. For each species' set of source concentration values, the numerical simulation is executed once to update the concentration in response to the source concentrations. The output of the simulation consists of concentration values for all nodal points in the simulation domain. The concentration values for the selected observation locations are

forwarded to the optimization module. The optimization algorithm evaluates the constraints and checks for termination. If the constraints are not satisfied, then it computes new source concentrations based on the ASA model and passes them to the numerical simulation model. Based on the HYDROGEOCHEM results, a new set of source concentrations is formed and HYDROGEOCHEM is called again to compute the concentrations. This process is continued until an optimal solution is reached based on the objective function and the constraints. The time period for optimization is given by the user. The numerical model will start running from the initial time period, irrespective of the optimization time period and, through this, the numerical model will take care of the time relationship. HYDROGEOCHEM 5.0 uses many input files, but during optimization, only the candidate source concentrations of individual contaminants change. All other parameters do not change; hence, only the source concentration files are modified using the linking C++ code during each iteration. The idea here is to introduce new candidate solutions through the source concentration files and then run the whole optimization model. The file containing the concentrations of individual contaminant species at different time steps and different observation locations is checked by solving the objective function, whether the constraints are satisfied or not.

4.3 PERFORMANCE EVALUATION

To evaluate the performance of the proposed optimal contaminant source characterization methodology, concentration measurements taken at monitoring locations in different time steps were utilized as a part of the real-life illustrative scenario. To test the reliability and robustness of the proposed methodology in real scenarios, concentration measurement errors were incorporated by introducing various amounts of synthetically-generated, normally-distributed error in the simulated concentration values (Singh & Datta, 2006). The perturbed simulated concentrations represent erroneous measurements, and are defined as follows:

$$pertC^{k}_{sim_{iob}} = C^{k}_{sim_{iob}} + \varepsilon \bullet a \bullet C^{k}_{sim_{iob}} \tag{4.4}$$

Where:

$pertC^{k}_{sim_{iob}}$ is a perturbed simulated concentration;

$C^{k}_{sim_{iob}}$ is a simulated concentration;

ε is a normally-distributed error term with a zero mean and standard deviation of one;

a is a fraction, such that 0 < a < 1. For example; a is varied from 0.05 to 0.2. a < 0.10 corresponds to a low noise level, 0.10 < a < 0.15 corresponds to a moderate noise level, and a. 0.15 corresponds to high noise level (Singh & Datta, 2006).

Model for Erroneous Concentration Measurement Values

For evaluation purposes, simulated concentration values were perturbed to represent measurement errors. Adding randomly-generated random errors to the simulated concentrations perturbed these simulated values. The normally-distributed random error terms are used to simulate the errors that generally occur in field measurements. The perturbed concentration values were computed as follows (Prakash & Datta, 2013):

$$Cobs\ (M_L,\ t) = \mathrm{Csim}\ (M_L,\ t) + \varepsilon\ r \qquad (4.5)$$

Where:

Cobs (M_L, t) is the measured or observed concentration of a species at monitoring location M_L at time t;

Csim (M_L, t) is the simulated concentration at location M_L and time t from the numerical simulation model; and

$\varepsilon\ r\ =$ is a random error term.

Here, the random variable $\varepsilon\ r$ is assumed to follow a normal distribution with mean $= 0$ and standard deviation = $a.Csim\ (M_L,\ t)$. Furthermore, the error term is defined as:

$$\varepsilon\ r = e.\ a.\ Csim\ (M_L,\ t) \qquad (4.6)$$

where a = a fraction ($0 \le a \le 1.0$) and e = normal deviates.

For this study, MATLAB (R2017a) software was used to generate standard normal deviates (e). The value of a was varied from 0.05 to 0.2, where higher values indicate a higher level of noise in the data. It was assumed that values of $a < 0.10$ correspond to a low noise level, $0.1 \leq a \leq 0.15$ corresponds to a moderate noise level, and $a > 0.15$ corresponds to a high noise level (Singh et al., 2004; Prakash & Datta 2013). Also, for performance evaluation purposes a normal distribution of errors is assumed. Any other suitable distribution function may be incorporated. The value of $Cobs\ (M_L,\ t)$ can be negative if e is negative, a is large and $Csim\ (M_L,\ t)$ is small. Generally, such a situation is less probable if a is small and e is also small. Otherwise, a truncated normal distribution may be used.

4.4 OPTIMAL SOURCE CHARACTERIZATION

The numerical simulation model used to solve the three-dimensional flow and reactive biogeochemical transport processes was utilized in the linked simulation-optimization model for optimal source characterization of different distributed sources in the contaminated mine site area. For evaluation purposes only, concentration measurements at specified monitoring locations were simulated (synthetic measurements) by the calibrated numerical flow and transport simulation model. These concentration data were generated for specified source characteristics. Then, they were used with the linked simulation-optimization models to evaluate the potential applicability, accuracy and feasibility of the developed methodology. After evaluating the performance utilizing synthetic simulated measurement data, the developed methodology was applied to characterize sources based on contamination measurements obtained in the study area.

However, as is generally the case for such large-scale, complex, contaminated aquifer sites, the contaminant sources are unknown. Hence, the source concentration magnitudes or source activity starting times cannot be validated. Hence, synthetic concentration values were utilized with the calibrated flow model to evaluate the performance of the source characterization methodology. Concentrations of contaminants measured in the field in 2011 and 2012 (Ferguson et al., 2012) were used in the source characterisation process.

4.5 RESULTS AND DISCUSSION

The aim of the performance evaluation process was to evaluate whether the source characterization model can recover the actual source characteristics based on synthetic concentration values. Hence, in the linked simulation-optimization model for source characterization, the calibrated numerical simulation models were used to generate synthetic (simulated) concentration data. The reason for using synthetic data for performance evaluation was to ensure that unknown errors in the actual measurements did not distort the performance evaluation results. Also, by using synthetic concentration data, it is possible to test the performance with different scenarios of measurement error. Furthermore, the actual sources may not be known, so the performance is evaluated with synthetic concentrations generated for specified contaminant sources. This is to ensure that the estimation results can be verified. Therefore, the source characteristics recovered can be verified from the specified sources used to generate the synthetic data. If the performance is satisfactory, the methodology should be useful for recovering the actual source characteristics.

4.6 CASE STUDIES FOR MODEL DEMONSTRATION

In the following sections, the model for identification of groundwater contaminant sources is demonstrated using two case studies involving a mining area; (i) with simulated (synthetic) measurements for testing (where the sources are known); and (ii) same study area with observed concentrations to demonstrate practical applicability.

4.6.1 Calibrated model testing with simulated data

This section focuses on a contaminated aquifer site (Figure 4.3) with six distributed sources of contamination consisting of four waste rock dumps and two open mine pits filled with water. The boundary conditions for flow are represented by the red border lines shown in Figure 4.3. This case study involves a calibrated model of a real mine site in the Northern Territory, Australia. It is adapted in this study to test the performance of the multiple species source identification model, where individual contaminant species from distributed sources are examined. There were six possible sources: four waste rock dump sites, D1 (Main WRD), D2 (Intermediate WRD), D3

(Dyson Backfilled Pits), D4 (Dyson WRD) and two open pits, P1 (Main Pit) and P2 (Intermediate Pit; Figure 4.2). There were nine observation wells (W1, W2, W3, W4, W5, W6, W7, W8 and W9) at the study site (Figure 4.2), which recorded concentrations of six contaminant species over a two-year period. The thickness of the aquifer is 150 m. The aquifer is presumed to be anisotropic. The effective porosity is taken to be 0.3. Table 4.1 shows the schedule of release of contaminant species, with locations and magnitudes (contaminant concentrations). The releases were assumed to occur over a two-year period and to be continuous thereafter.

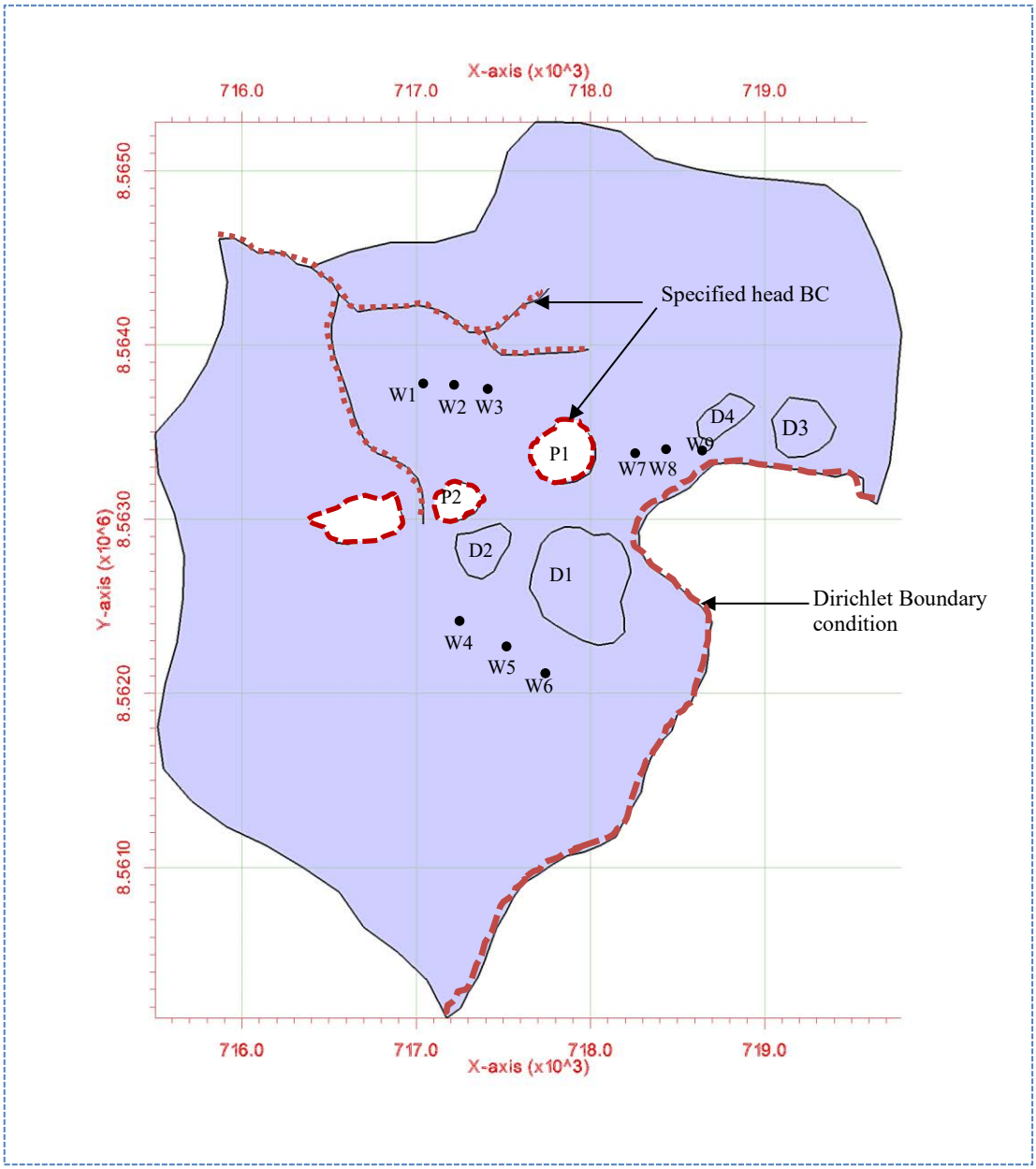


Figure 4.3 Mine site aquifer with four contaminated waste rock disposal sites and nine observation wells

Table 4:1 Contamination Release Schedule from Distributed Sources –Simulated Measurements for Testing (As Sources are Known)

Site	Contaminant concentration (mg/L)			
	Cu^{2+}	SO_4^{2-}	UO_2^{2+}	Fe^{2+}
Main WRD	4.19	3430	0.568	0
Intermediate WRD	34.9	13800	1.840	349
Dyson Backfilled Pits	30.0	2500	1.590	4.865
Dyson WRD	4.63	579	0.155	2.74
Main Pit	55.0	8200	0.000	430
Intermediate Pit	60.0	3100	0.000	2.00

A groundwater flow and transport model was run in HYDROGEOCHEM 5.0 to simulate the flow and transport processes and generate synthetic concentrations at specified observation wells. The simulation was carried out for a two-year period with a time-step of 30 days. The concentrations at the selected observation wells estimated by the simulation are shown in Figure 4.3, which compares actual and predicted concentrations for species at specific observation wells. These concentrations estimated by the simulation model were treated as input observed concentration data for the source identification model. The observed concentration data was fed to the source identification model for characterization of the sources. The optimization algorithm used for this problem was ASA, which tries to find a set of candidate source concentrations of individual contaminants by minimizing the objective function defined in Equation 4.1

4.7 RESULTS AND DISCUSSION

This section discusses the evaluation results obtained from the source identification model using a calibrated model with simulated measurements for testing (as sources are known) the performance. The source identification model was able to identify separate contaminants concentrations. Six sources were used in the characterization. It is important to note that, for this study, the contaminant sources (waste rock dumps) were characterised in terms of concentrations rather than fluxes. This is so because species concentrations are expressed in terms of moles per litre in the software used; hence, concentration multiplied by volume flux gives the mass flux. Also, in the case of pits, it was assigned as a boundary condition and hence expressed in terms of concentration.

The results of the source characterisation evaluation are presented in Figures 4.4 to 4.7. Table 4.2 summarizes the error statistics related to the source characterization using error-free concentration measurement data. Figures 4.4 to 4.7 compare the actual concentrations with those estimated by the optimization model of specific contaminant species at the six sources.

Figure 4.4 is a graph showing the concentration of the contaminant copper, where the source concentrations are compared with the concentration solutions obtained from the optimization model. The concentrations compared very well at all source locations when error-free data was utilized.

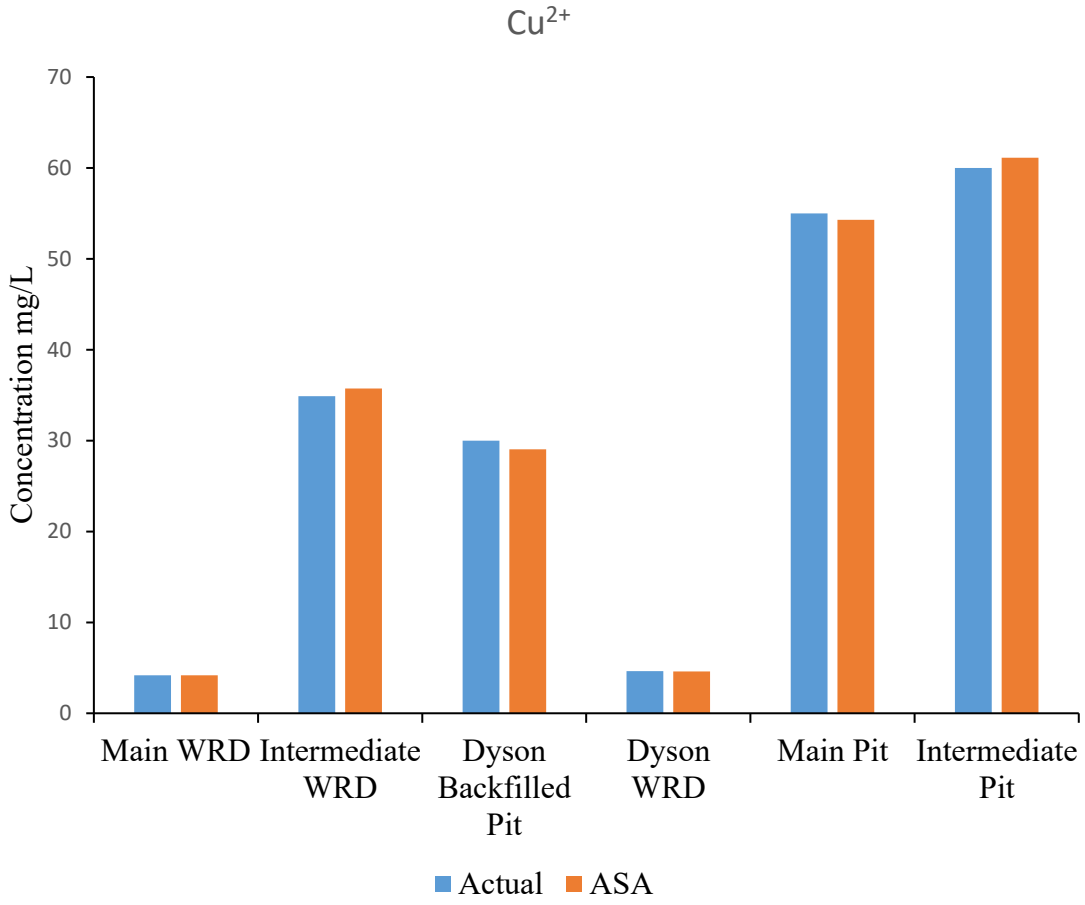


Figure 4.4: Comparison of Cu^{2+} species concentrations (mg/L) estimated by the ASA-linked optimization model and actual data

Figure 4.5 shows sulfate concentrations, where the source concentrations are compared with the estimates of the optimization model. The concentrations compared very well at all source locations when error-free data was utilized as synthetic concentration measurements.

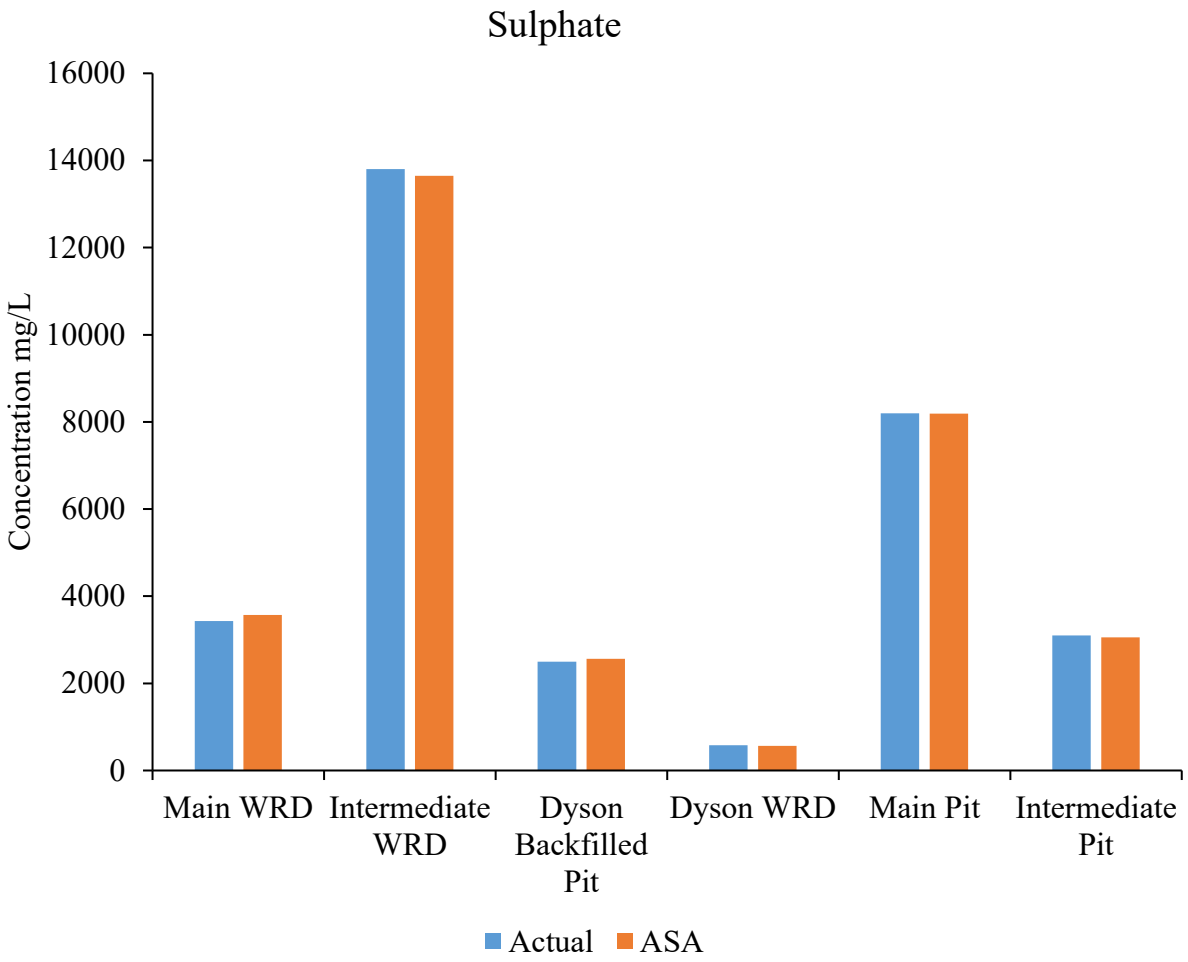


Figure 4.5: Comparison of SO_4^{2+} species concentrations (mg/L) from the ASA-linked optimization model and synthetic data

Figure 4.6 shows the concentrations of the contaminant uranium (in the form of UO_2^{2+}), where the source concentrations are compared with the concentration solutions obtained from the optimization model. The actual and estimated concentrations compared very well at all source locations when error-free data was utilized. The Main Pit and Intermediate Pit sources did not show any bars since the source concentration of uranium was near zero at these points. Hence, the optimization model was able to provide same value.

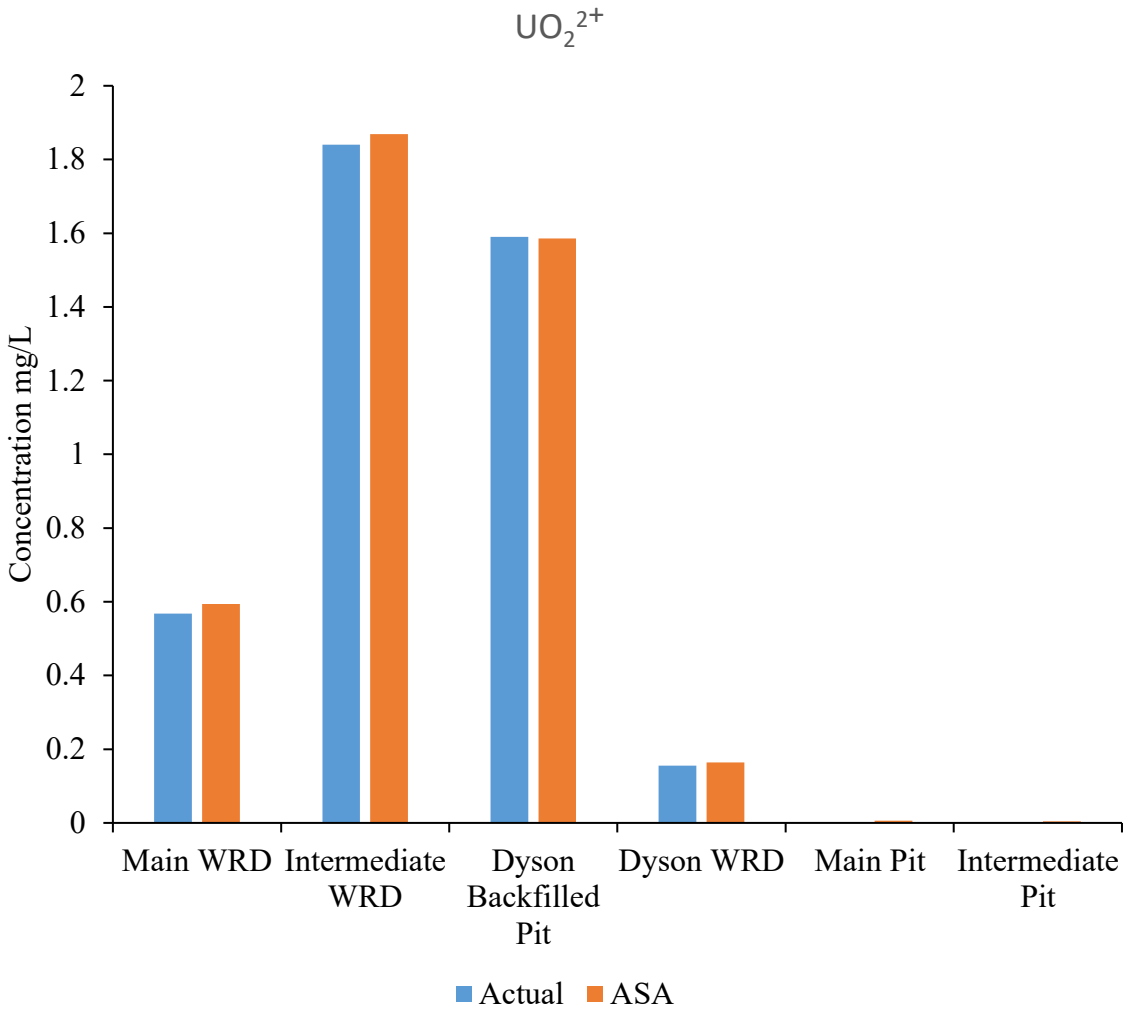


Figure 4.6: Comparison of UO_2^{2+} species concentrations (mg/L) from the ASA-linked optimization model and actual data

Figure 4.7 shows the concentrations of the contaminant iron (in the form of Fe^2), where the source concentrations are compared with the concentration solutions obtained from the optimization model. The concentrations compared very well at all locations when error-free synthetic data were used. Source locations at the Main WRD did not peak up, which reflects the same results for the optimization model. At the Intermediate Pit, the concentration of iron was low, so does not show on the figure.

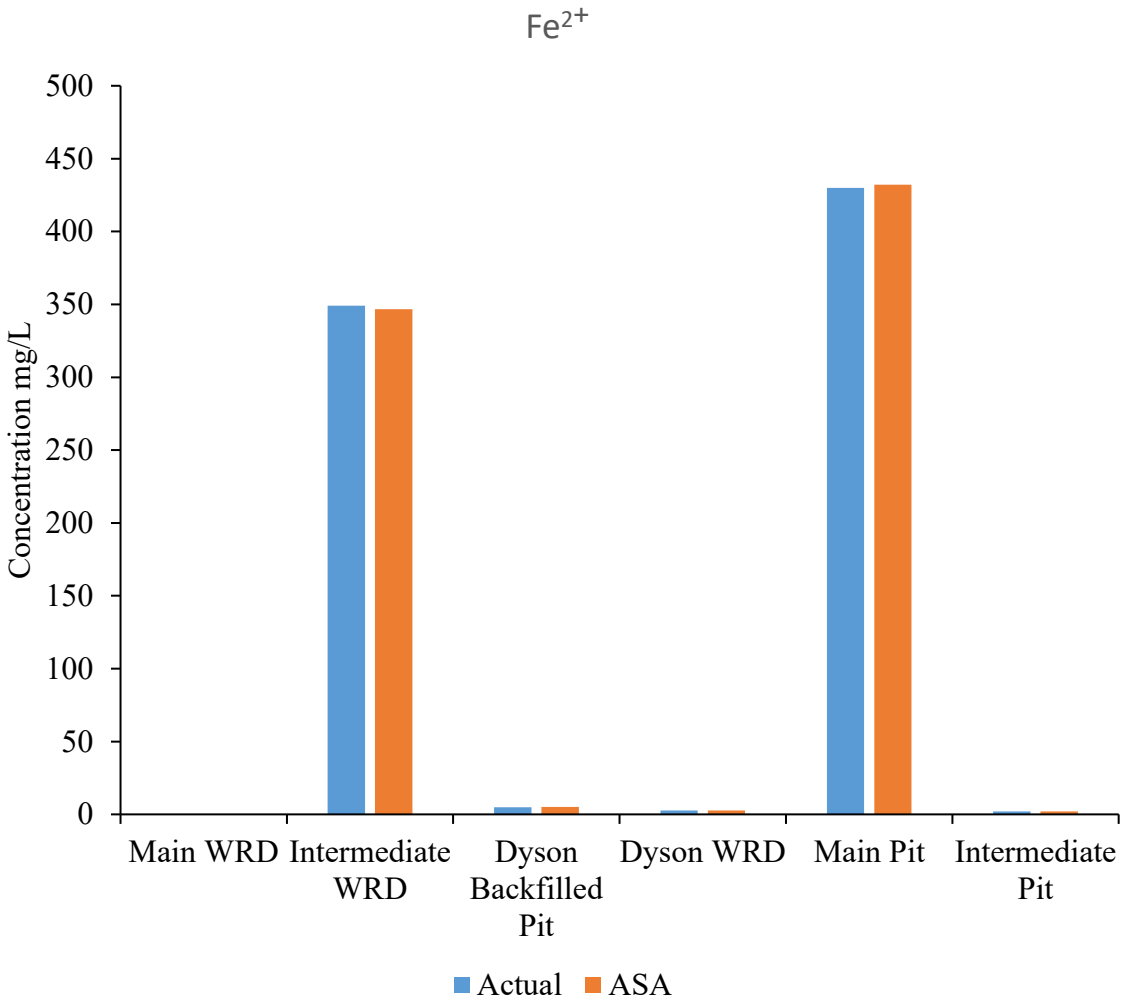


Figure 4.7: Comparison of Fe^{2+} species concentrations (mg/L) from the ASA-linked optimization model and the actual data

The performance of the linked simulation-optimization source characterization model using an ASA optimization algorithm was evaluated in terms of percent average contaminant source estimation error (PAEE). PAEE was used to compare the input concentrations of the generated and actual source locations. Mahar and Datta (2001) performed such a comparison in terms of PAEE; the formula is:

$$\text{PAEE}\ (\%) = \frac{\left|\overline{C}_o - \widetilde{C}_o\right|}{\widetilde{C}_o} \times 100 \qquad (4.5)$$

Where $\overline{C}_o$ is the actual input observed species concentration and $\widetilde{C}_o$ is the species concentration generated by the optimization model.

Using measured concentration data in the optimal source identification model, the actual and optimization-generated species concentrations were accurately estimated within a 10% average estimation error value. This is except for iron, which had a PAEE value of 6.5% at the Intermediate WRD. It can be seen that the PAEE value tends toward zero when the generated concentrations approach the actual ones. Table 4.2 compares the calculated PAEE values for the species at different source locations.

Table 4.2: Comparison of percent average estimation error (PAEE, %) obtained for species at distributed source locations using error-free data

Source	PAEE (%)			
	Cu	SO_4^{2-}	U	Fe
Main WRD	0.1	3.8	4.3	0.0
Intermediate WRD	2.3	1.10	1.55	6.5
Dyson Backfilled Pit	3.32	2.35	0.26	5.35
Dyson WRD	0.33	1.99	5.5	5.22
Main Pit	1.30	0.11	0	0.50
Intermediate Pit	1.82	1.35	0	4.0

4.8 SOURCE IDENTIFICATION FOR A FIELD PROBLEM

This case study considers the transport of reactive contaminant species through an aquifer system. The contamination came from acid mine drainage from waste rock dumps and open mine pits that occurred continuously over a long time. The aquifer shown below was described in Chapter 3. It is associated with an abandoned uranium mine in Australia. The modelled area is approximately 12 km^2. It is bounded by the upper east branch of the Finniss River on the left side and a specified transient head boundary condition to the borders of the Main and Intermediate Open Pits. The water levels in the open pits were used as the constant head values at the boundaries of the pit. Groundwater contaminant concentrations measured at several monitoring locations were used as specified concentrations in the reactive transport model. These concentrations were estimated as constant values for species at selected monitoring points at specific time intervals. There were six potential sources, consisting of two

open pits and four waste rock dumps: OP1, OP2, MWRD, IWRD, DBP and DWRD. These released reactive species contaminants associated with AMD via a reactive transport process over a period of two years. It is assumed that after two years of release, the contaminants stopped being released at the study site. Nine observation wells, labelled OB1, OB2, OB3, OB4, OB5, OB6, OB7, OB8 and OB9, recorded contaminant concentrations over a two-year simulation period.

4.9 RESULTS AND DISCUSSION

Based on the concentrations measured at the monitoring wells, source identification was attempted using the optimization formulation in Equation (4.1). This equation provided optimal source concentrations that matched the field measurements. Corresponding measured concentrations were used to formulate the optimization model used in this study. The concentrations of the species considered were used for optimal source characterisation of the distributed sources.

The optimal source characteristics of the species concentrations obtained from the optimization model were then utilized as input source concentration parameters to model and simulate the concentration plumes from the sources in this study area.

The use of concentrations measured at monitoring locations to identify sources is discussed. The contours of simulated plumes of different species at different source locations are shown in Figures 4.8 to 4.12.

Figure 4.8 shows a simulated contour map of Cu concentration plumes based on field-measured concentrations. Figures 4.9 and 4.10 are simulated contour maps of SO_4^{2+} concentration plumes represented in Layers 1 and 2, respectively. Comparing Figures 4.9 and 4.10, it can be seen that as the SO_4^{2+} moves from one layer to the other, the concentrations at the sources also change. These contour maps show variations in concentrations as the plumes move through different layers of the model. Figures 4.11 and 4.12 show simulated contour maps of UO_2^{2+} concentration plumes as characterized in Layers 1 and 2, respectively. These contour maps show some variation in the concentrations as the plumes move through different layers of the model.

Figures 4.8 to 4.11 represent the concentration plumes of species at different layer resulting from characterized sources.

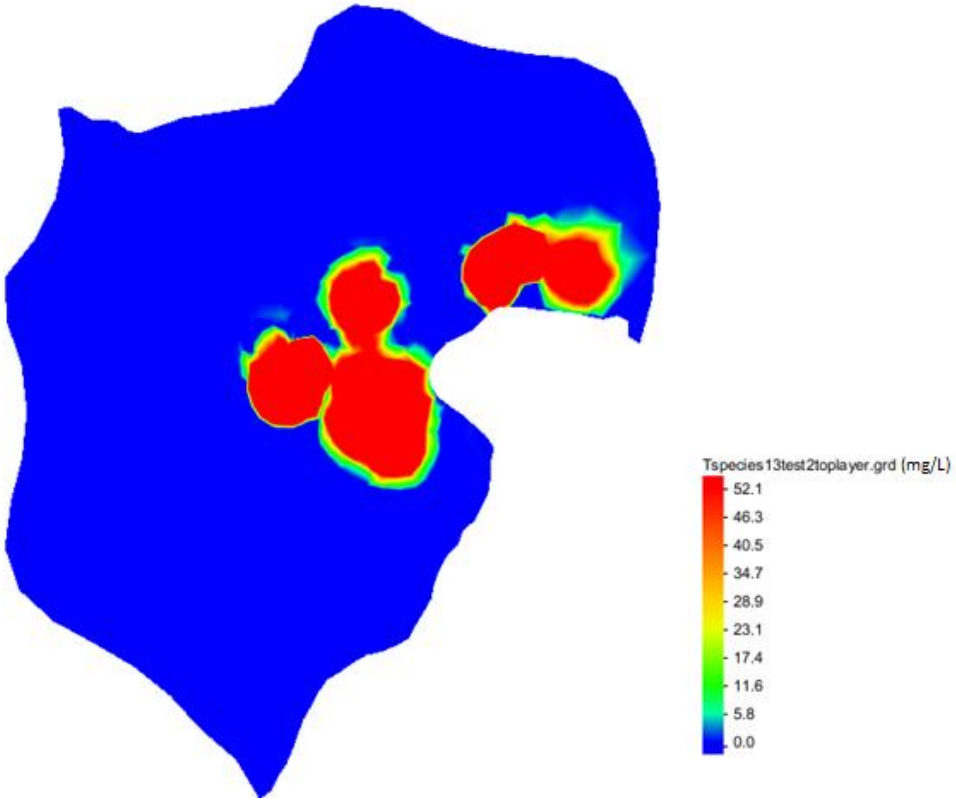


Figure 4.8: Concentration plumes resulting from characterized sources for Species 1: Cu Layer 1

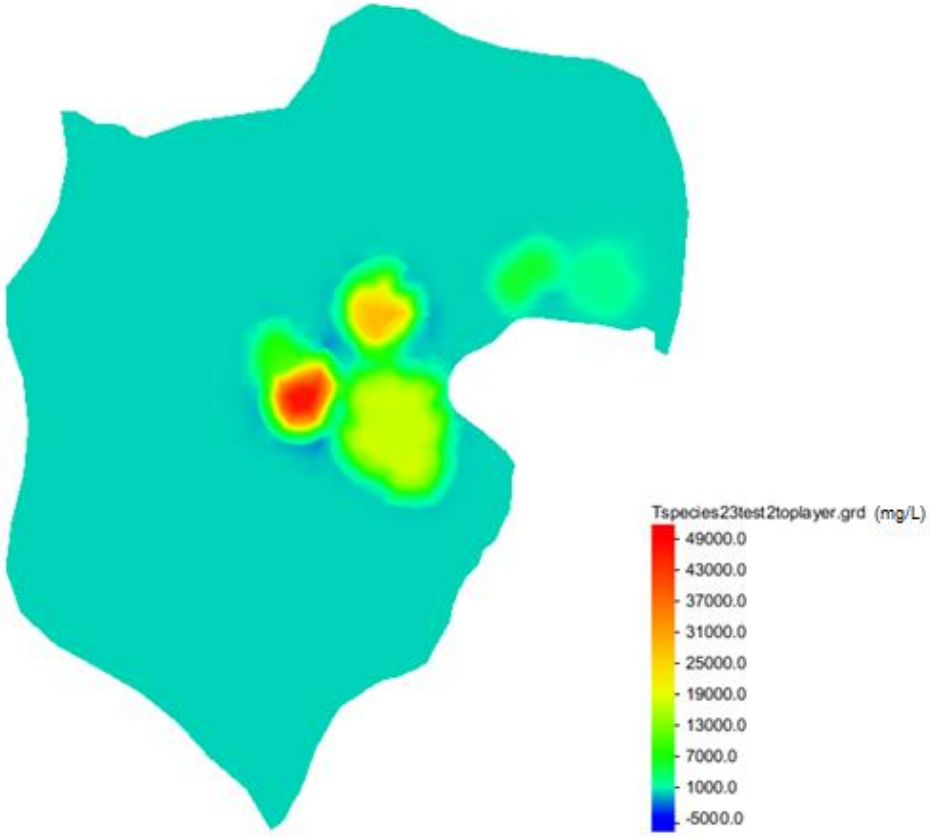


Figure 4.9: Concentration plumes resulting from characterized sources for Species 2: SO_4 Layer 1

Figure 4.10: Concentration plumes resulting from characterized sources for Species 2: SO_4 Layer 2

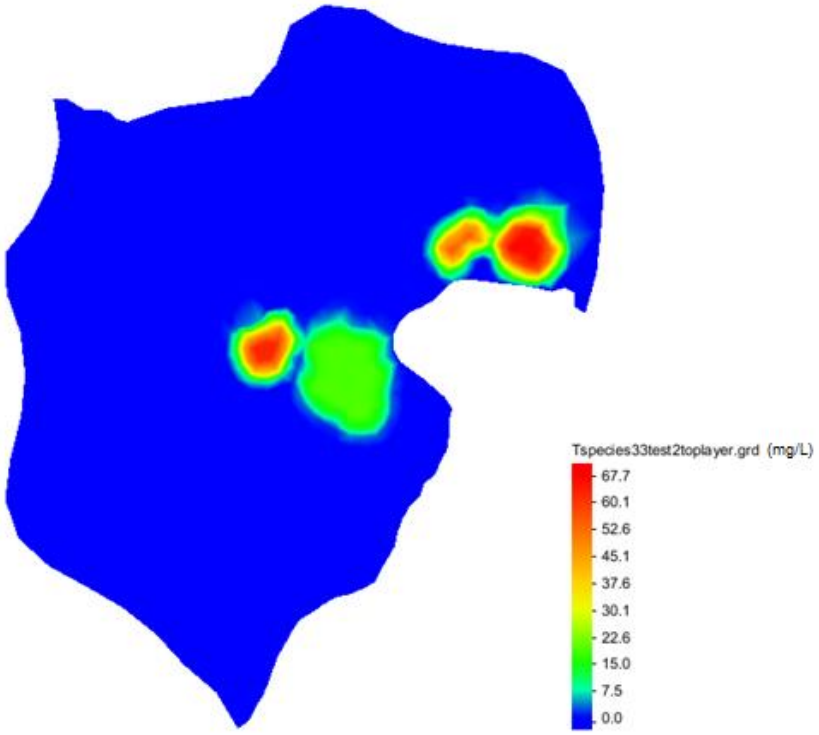


Figure 4.11: Concentration plumes resulting from characterized sources for Species 3: UO_2 Layer 1

Figure 4.12: Concentration plumes resulting from characterized sources for Species 3: UO_2 Layer 2

These contours can be used to compare point measurements of concentrations with concentration estimates based on the contours. However, such comparisons are possible only when the actual sources in the field are the same as those identified by the source characterization inverse model, which is not possible in this case. However, few pointwise concentration values seem to be of the same order as those in the field. This may help in an intuitive validation of the source characterization process, as the physical conditions assumed for the performance evaluations resemble the field conditions to a certain extent. Selected points in specific areas in the point concentration comparison and the source identification intuitive validation are discussed below.

The following monitoring bores were used for point concentration comparisons. Two monitoring bores, MB10-10 and MB10-11, were positioned below the former copper extraction pad. Monitoring bores MB 10-3 and MB 10-4 were located near the East Finniss Diversion Channel. Monitoring bores PMB 10-5 and PMB 10-6 were situated near the intermediate waste rock dump (IWRD). Monitoring bores MB 10-7, MB 10-12, MB 10-13 and MB 10-16 were located in the central mining area. Monitoring bores PMB 10-9S and PMB 10-9D were sited near the east branch of the Finnish River and the intermediate pit. Monitoring bores PMB 10-20 and PMB 10-21 were also positioned downstream of the mine site. The simulated (sim) and field-observed (obs) concentrations of sulfate, copper, uranium and iron at the monitoring bores are compared in Table 4.3.

Table 4.3: Comparison of observed (Obs) and simulated (Sim) contaminant concentrations (mg/L) at monitoring locations based on final optimal sources from the source characterisation model

	Sulphate (SO_4^{2+})		Iron (Fe^{2+})		Copper (Cu^{2+})		Magnesium (Mn^{2+})		Uranium (UO_2^{2+})	
Bore	Obs	Sim	Obs	Sim	Obs	Sim	Obs	Sim	Obs	Sim
PMB 10-3	493	500	0.542	0.55	2.41	3	1.81	1.79	0.073	0.07134
PMB 10-4	1250	1511	0.336	0.4	0.015	0.019	0.117	0.115	0.012	0.009117
PMB 10-5	212	200	0.054	0.06	0.001	0	0.084	0.09	0.003	0
PMB 10-6	1090	1050	0.03	0.04	0	0	0.649	0.7127	0.002	0
PMB 10-10	756	650	1.8	1.6	0.006	0.001112	0.22	0.21	0.085	0.0964
PMB 10-11	5180	5170	1.49	1.55	77.2	80	144	150	0.024	0.035
PMB10-22	3810	3800	37.4	37.1	561	590	124	140	0.109	0.1763
PMB10-24	1050	1000	0.56	0.61	52.6	55	18.3	20	0.162	0.113
RN022543	1340	1200	0	0	0.008	0.009	0.081	0.089	0.003	0.00354
RN022543	1140	1020	0.014	0.017	0.018	0.015	0.069	0.075	0.003	0.003
PMB10-7	1450	1400	0.002	0.002113	0.003	0.002	0.002	0.01	0.006	0.005126
PMB10-9S	350	285	0.22	0.33	0.003	0.002	0.474	0.5	0.009	0.008
PMB10-9D	3270	3300	2.35	2.4	0.040	0.03	5.42	5.66	0.316	0.611

4.9.1 Results for Erroneous Concentration Measurement Values

Figures 4.12 to 4.27 compare the ASA-generated species concentrations at the distributed sources with error perturbation values of 0.05, 0.10, 0.15 and 0.2. Each of the unknown species' source concentrations is marked on the *y*-axes. The *x*-axes have three bars for each potential source location corresponding to concentration of a species at the source. The first bar is the actual value, the second represents estimated values based on error-free concentration measurements, and the third bar represents estimated values based on concentration measurements with perturbed erroneous measurement data.

The results of source concentration identification using error-free data closely match those using the actual source concentration values for all species and source areas, as displayed in Figures 4.12 to 4.27.

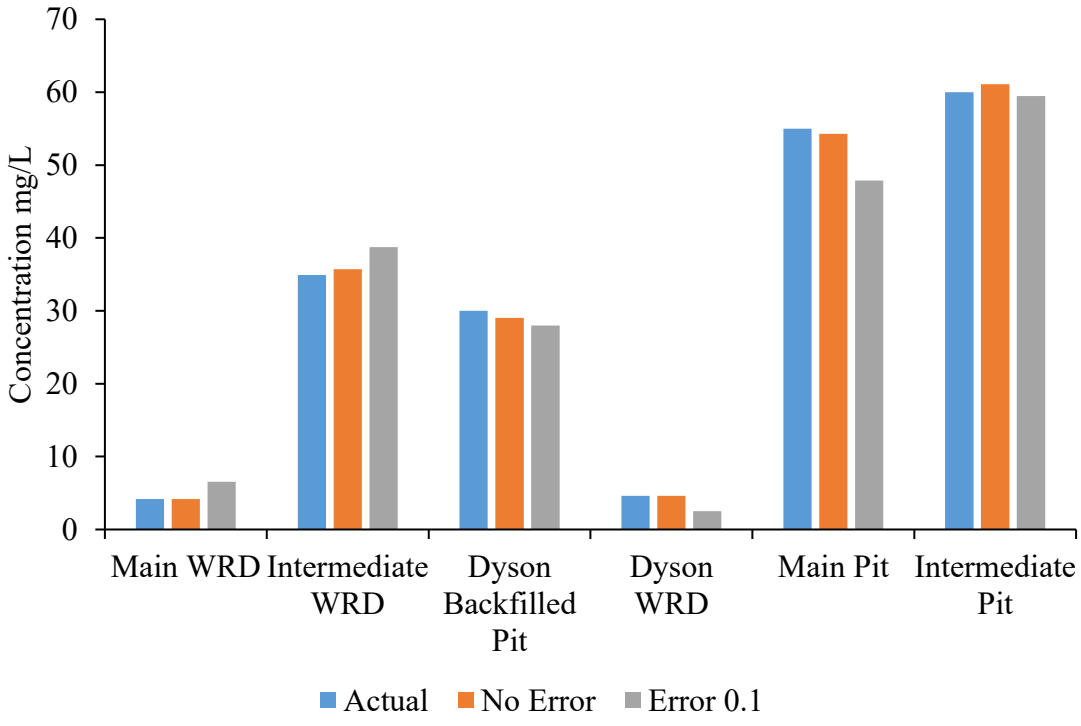


Figure 4.12: Comparison of Cu^{2+} species concentrations at sources with a perturbed error of 0.1

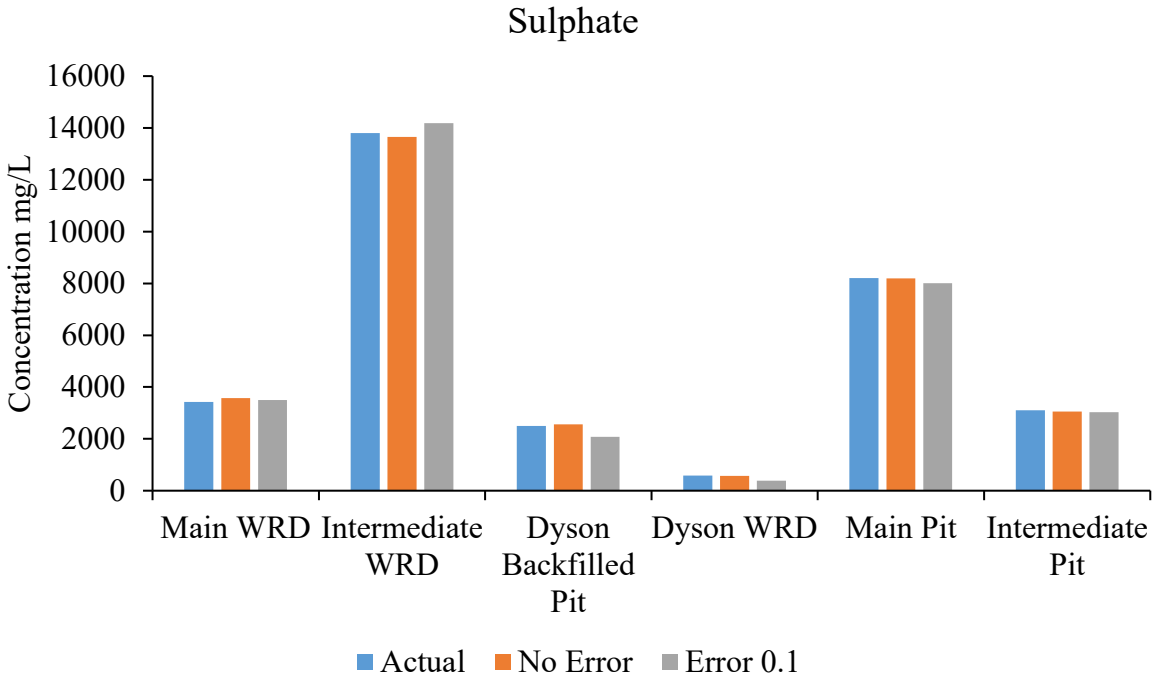


Figure 4.13: Comparison of SO_4^{2+} species concentrations at sources with a perturbed error of 0.1

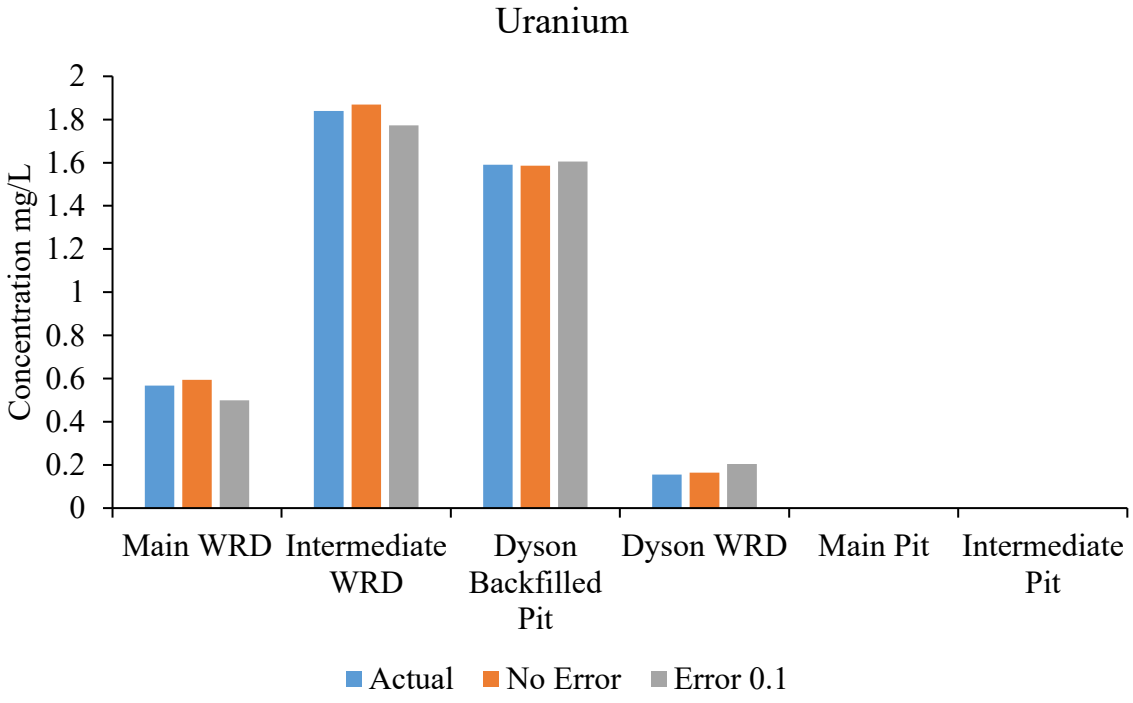


Figure 4.14: Comparison of UO_2^{2-} species concentrations at sources with a perturbed error of 0.1

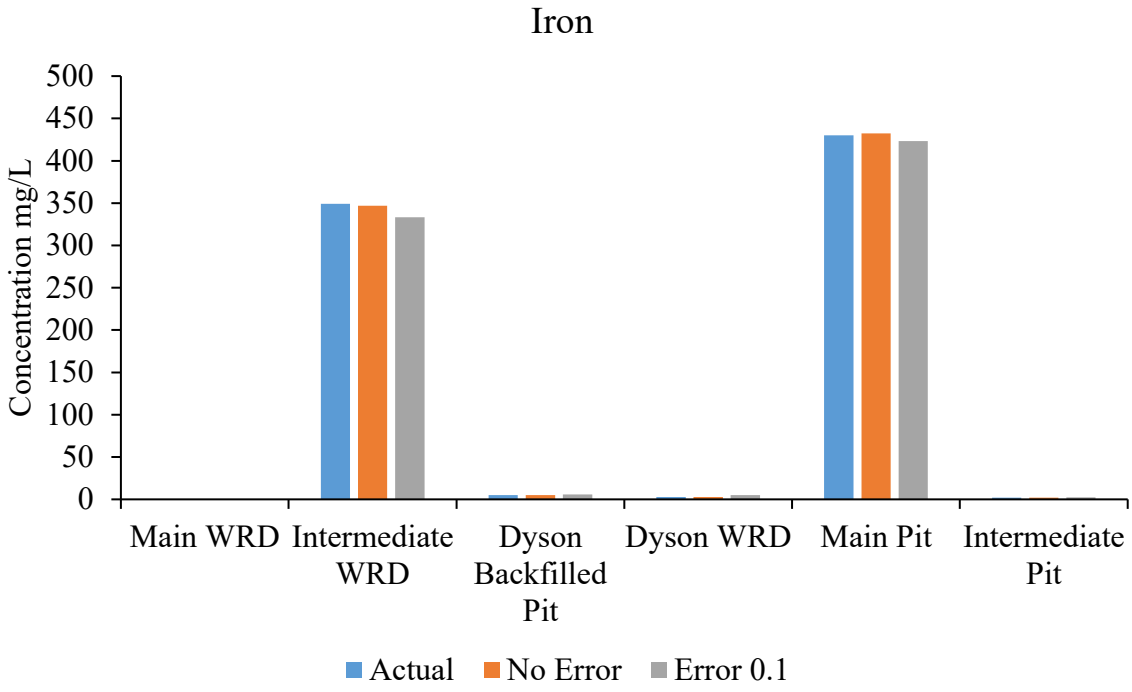


Figure 4.15: Comparison of Fe^{2+} species concentrations at sources with a perturbed error of 0.1

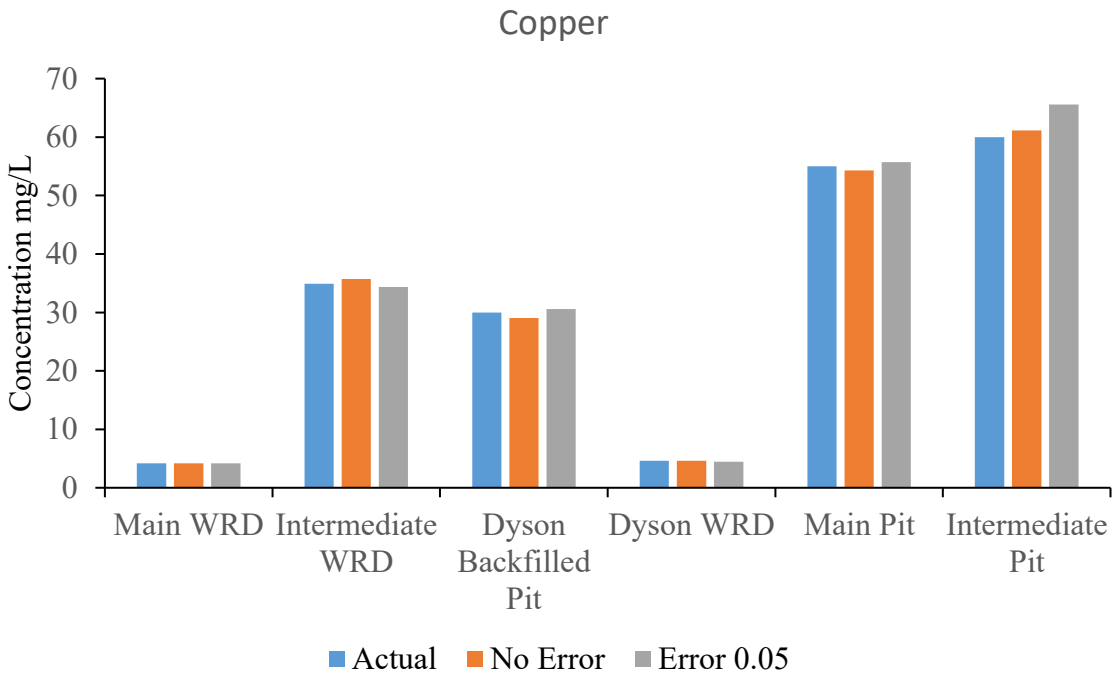


Figure 4.16: Comparison of Cu^{2+} species concentrations at sources with a perturbed error of 0.05

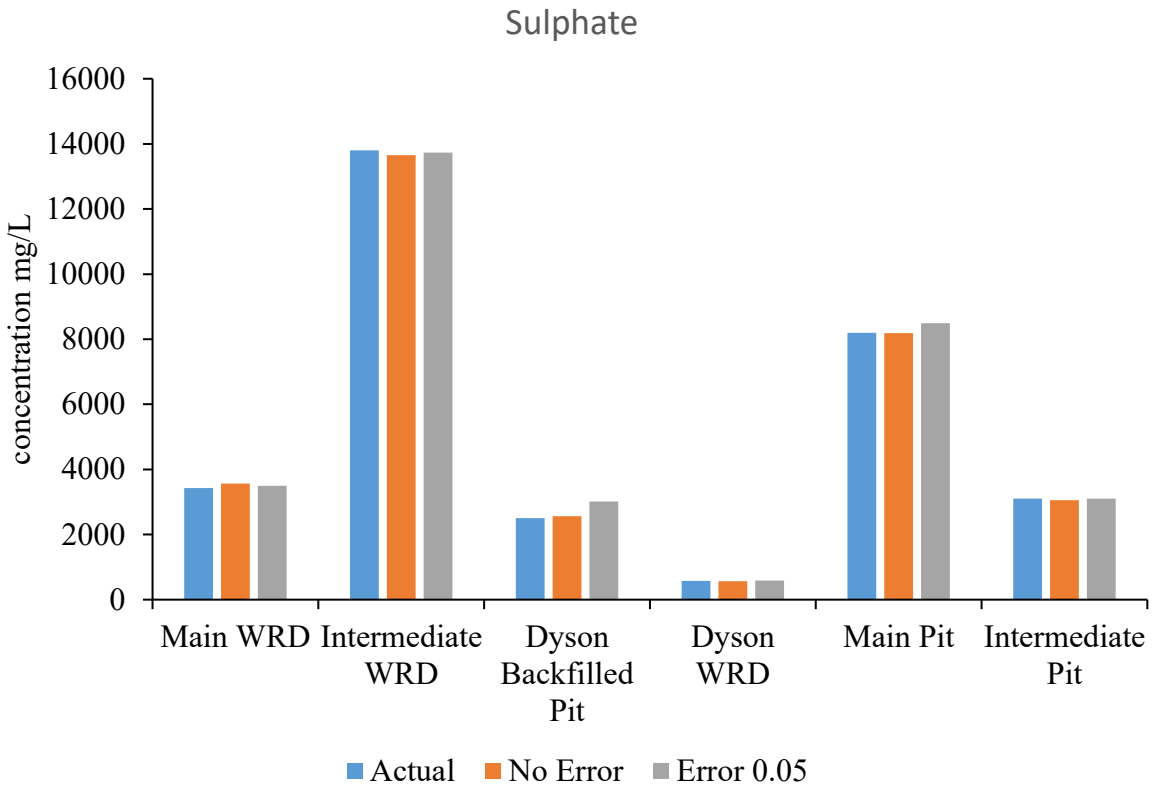


Figure 4.17: Comparison of SO_4^{2+} species concentrations at sources with a perturbation error of 0.05

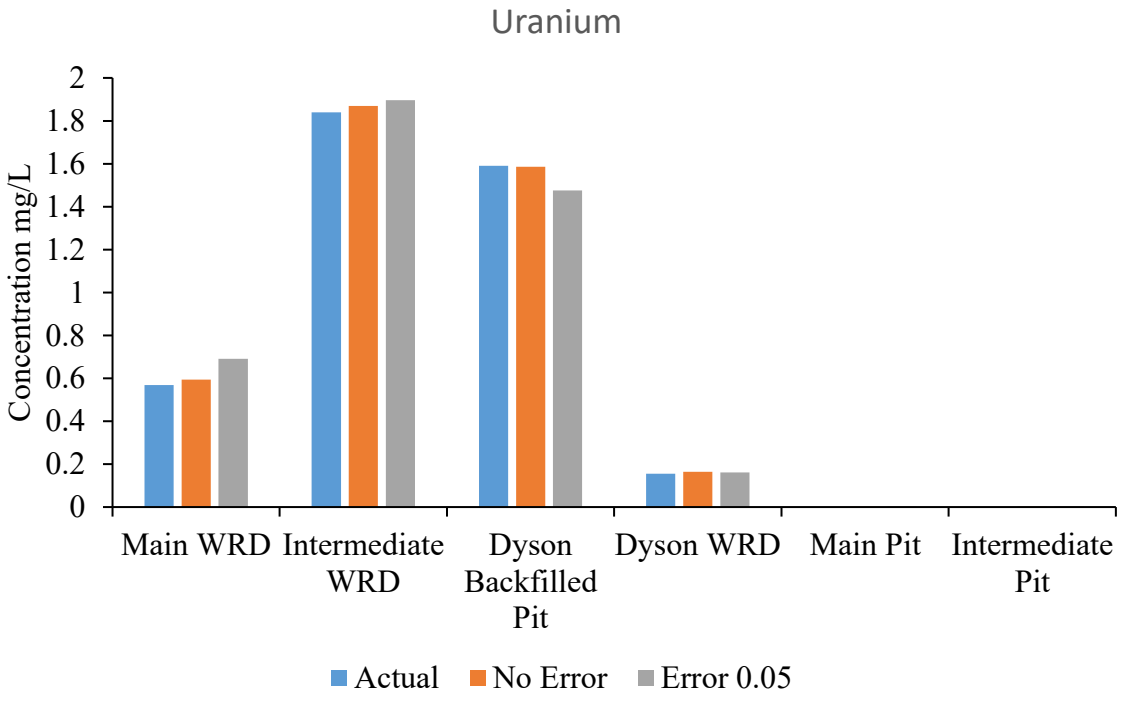


Figure 4.18: Comparison of UO_2^{2-} species concentrations at sources with a perturbation error of 0.05

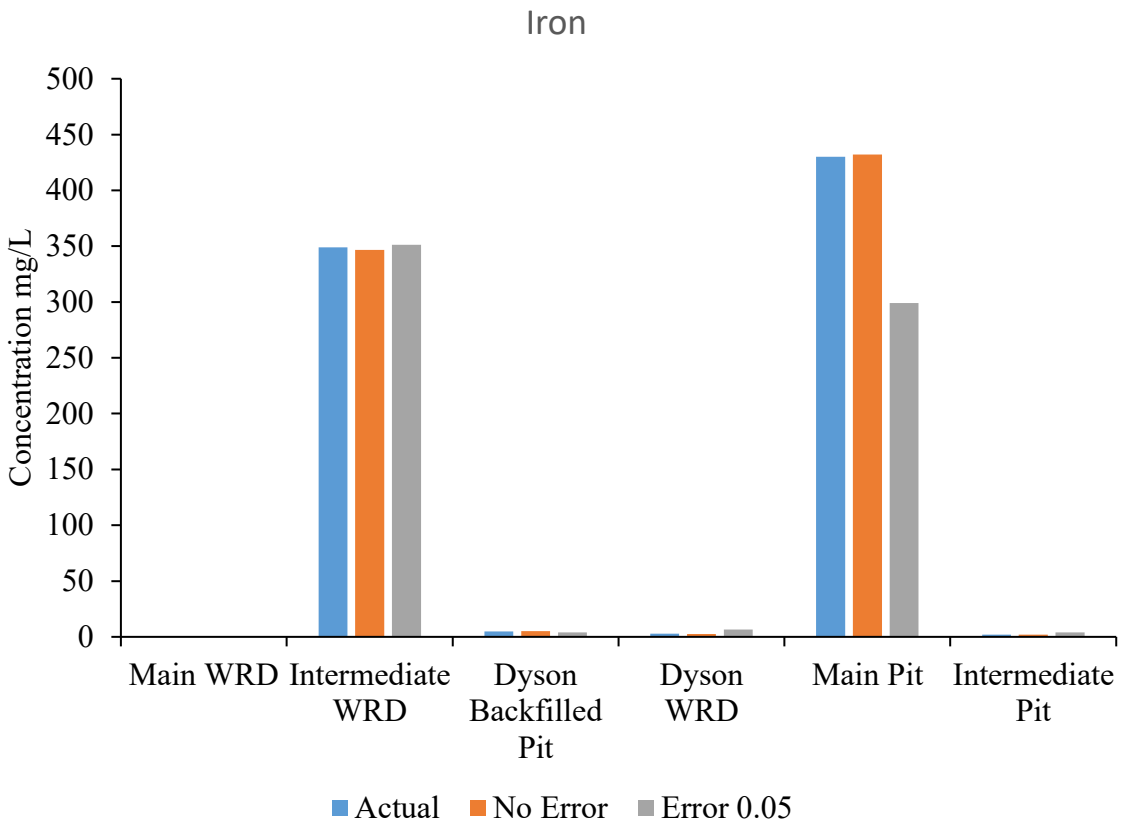


Figure 4.19: Comparison of Fe^{2+} species concentrations at sources with a perturbation error of 0.05

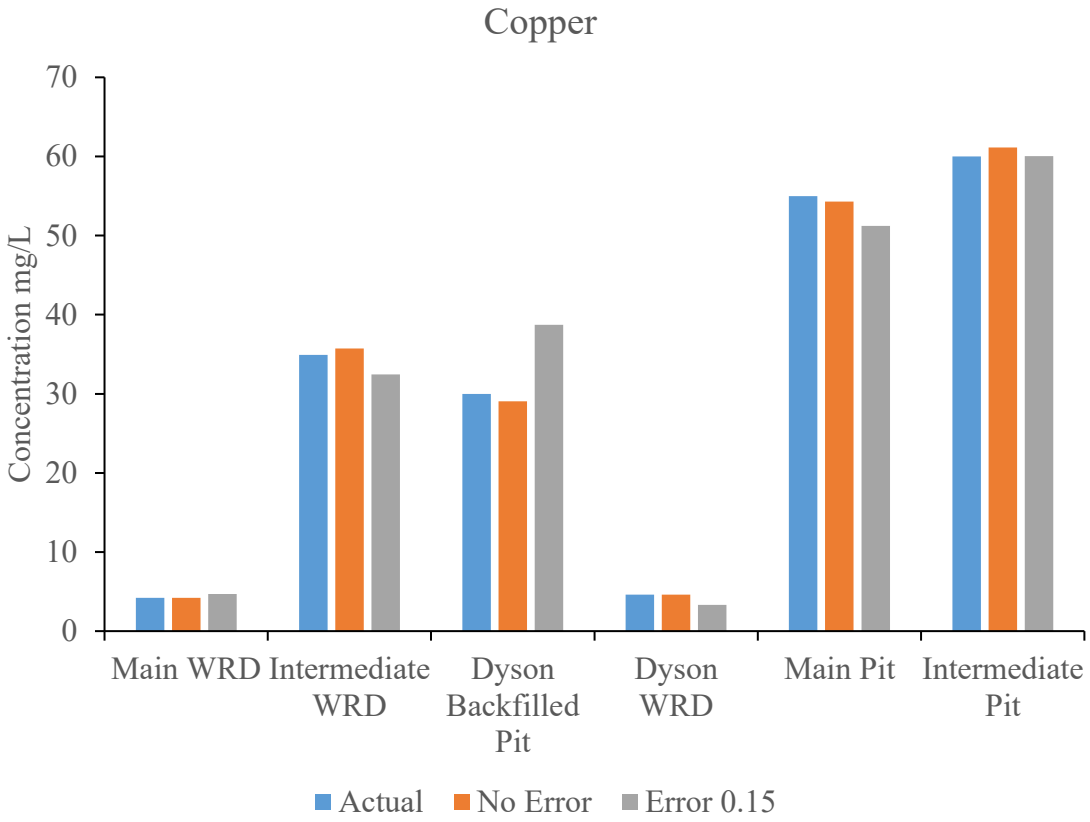


Figure 4.20: Comparison of Cu^{2+} species concentrations at sources with a perturbation error of 0.15

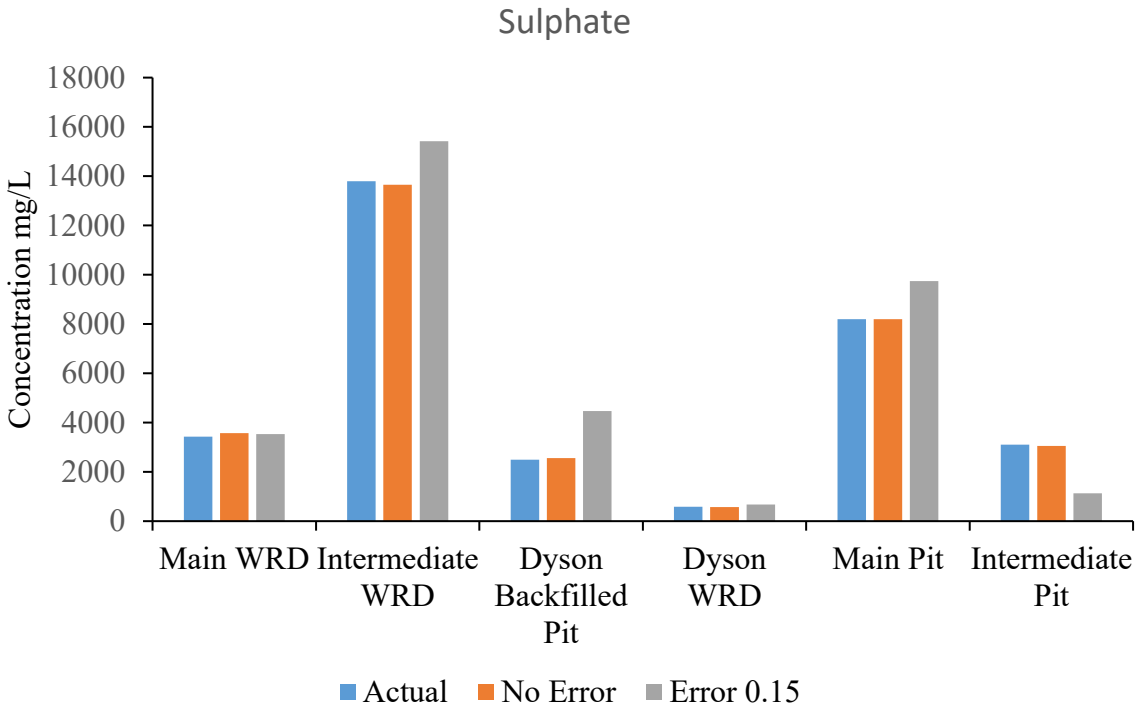


Figure 4.21: Comparison of SO_4^{2-} species concentrations at sources with a perturbation error of 0.15

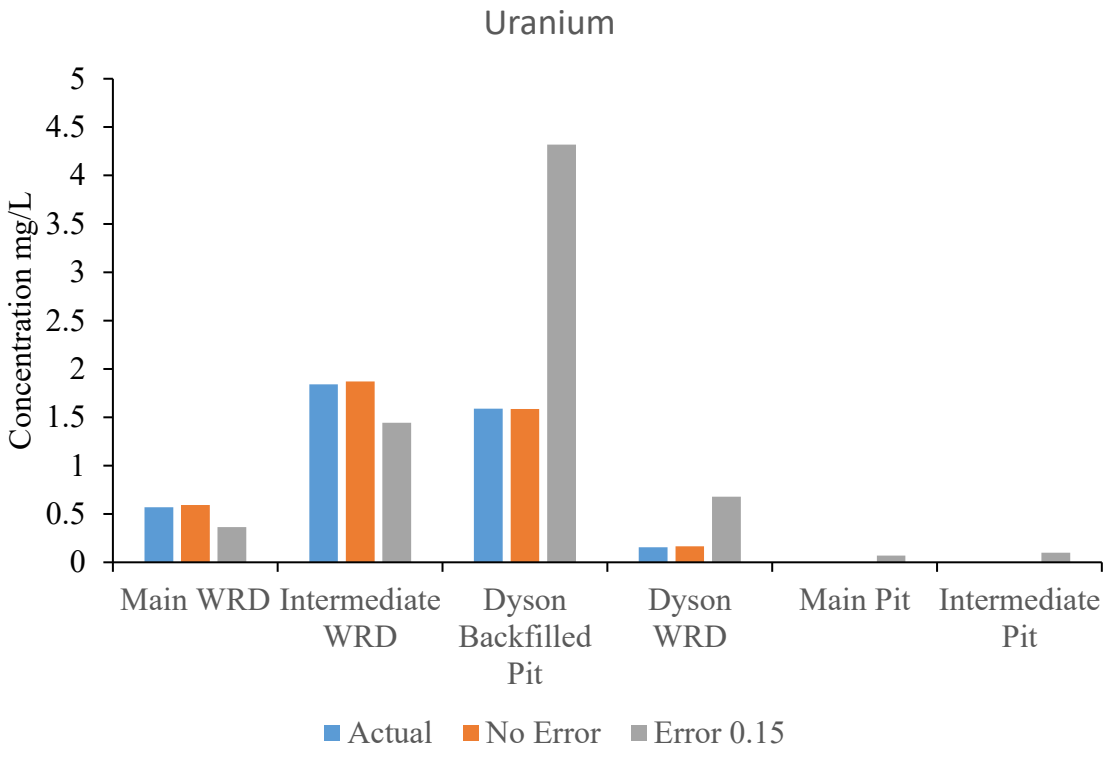


Figure 4.22: Comparison of UO22- species concentrations at sources with a perturbation error of 0.15

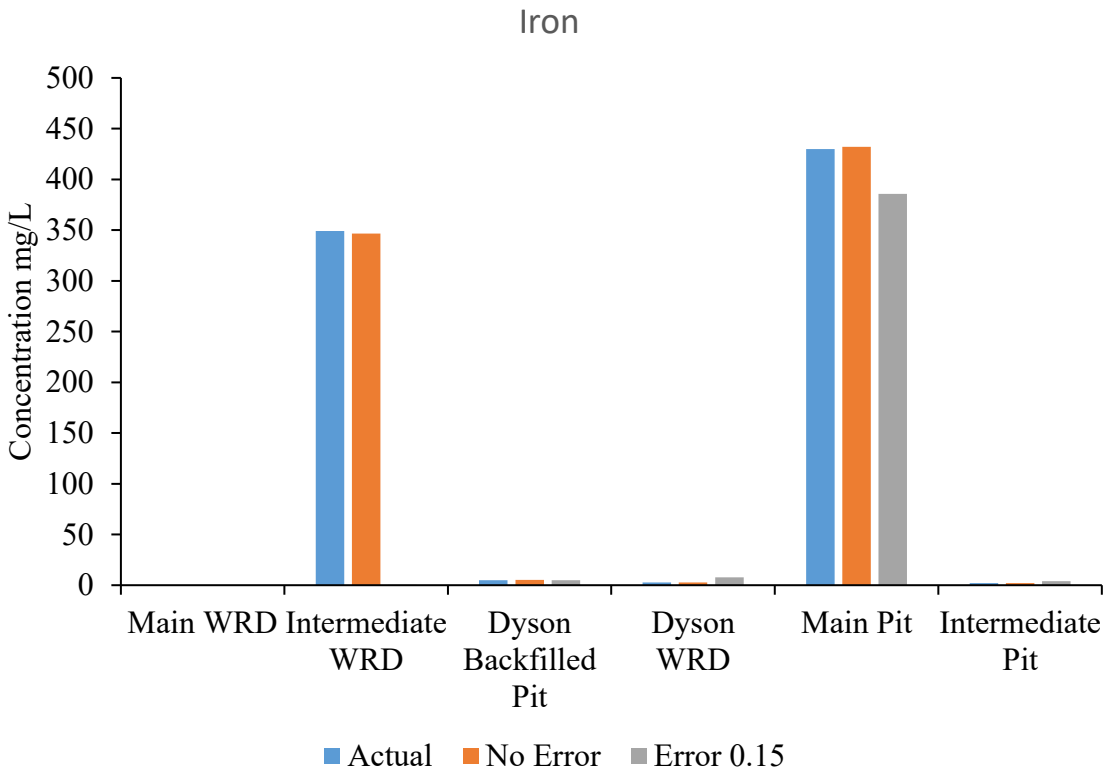


Figure 4.23: Comparison of Fe^{2+} species concentrations at sources with a perturbation error of 0.15

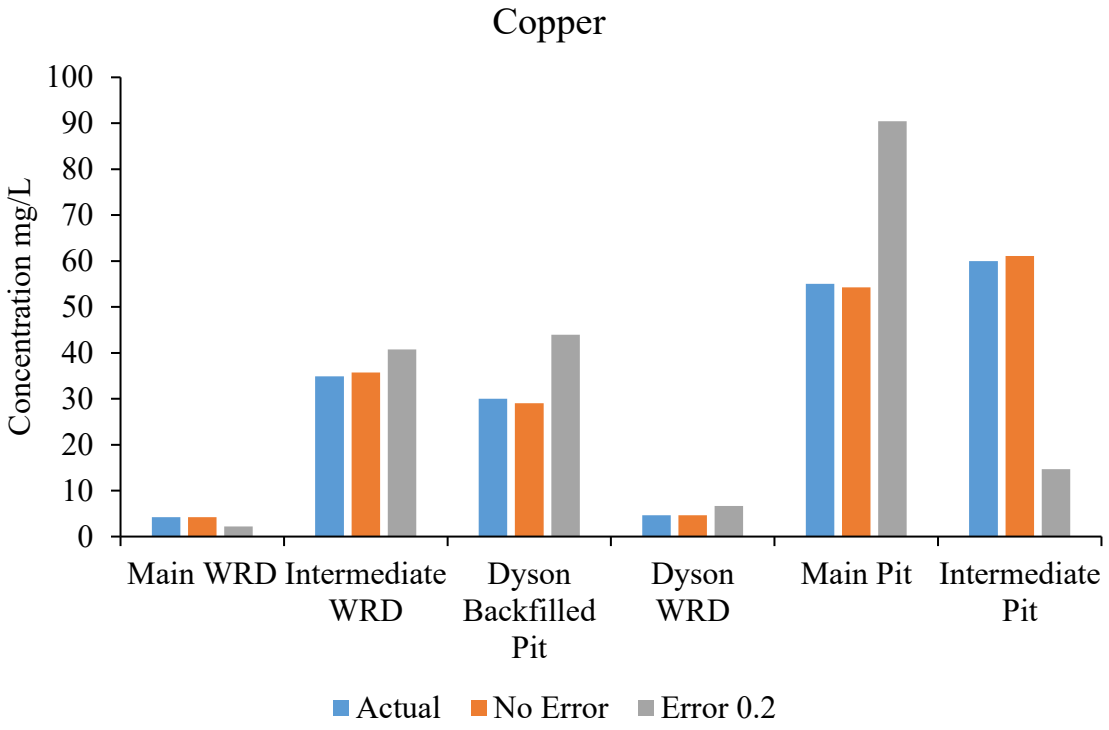


Figure 4.24: Comparison of Cu^{2+-} species concentrations at sources with a perturbation error of 0.2

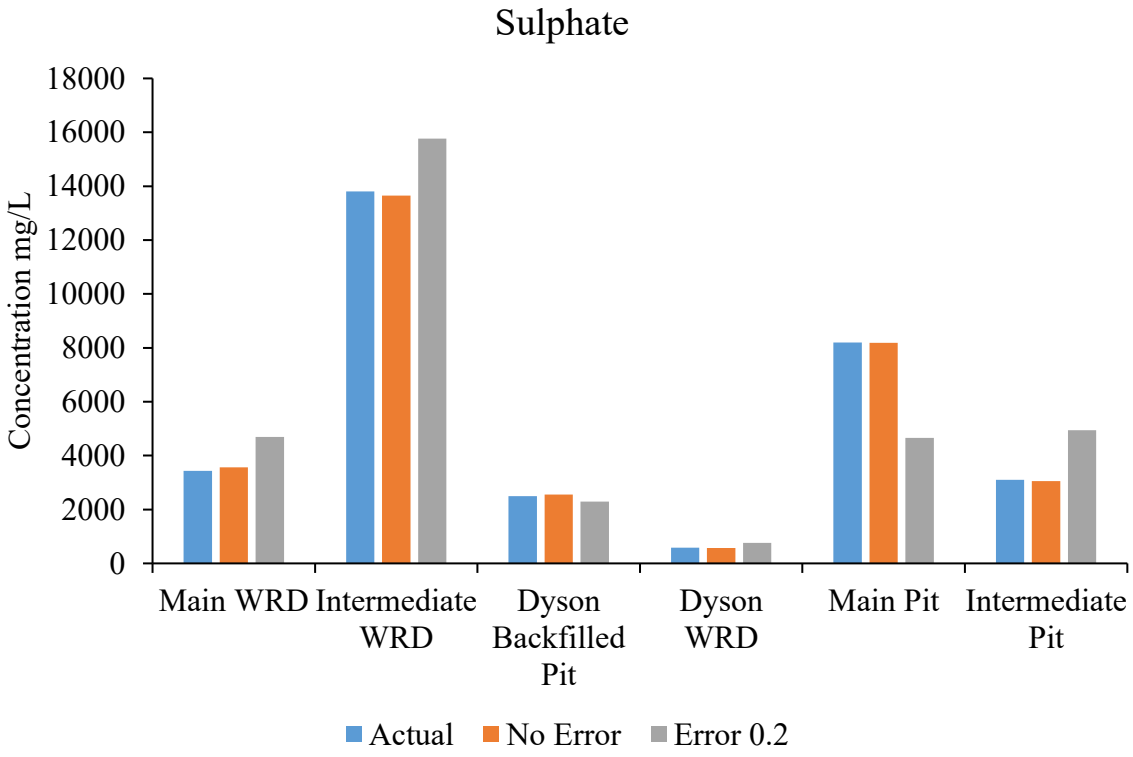


Figure 4.25: Comparison of SO_4^{2-} species concentrations at sources with a perturbation error of 0.2

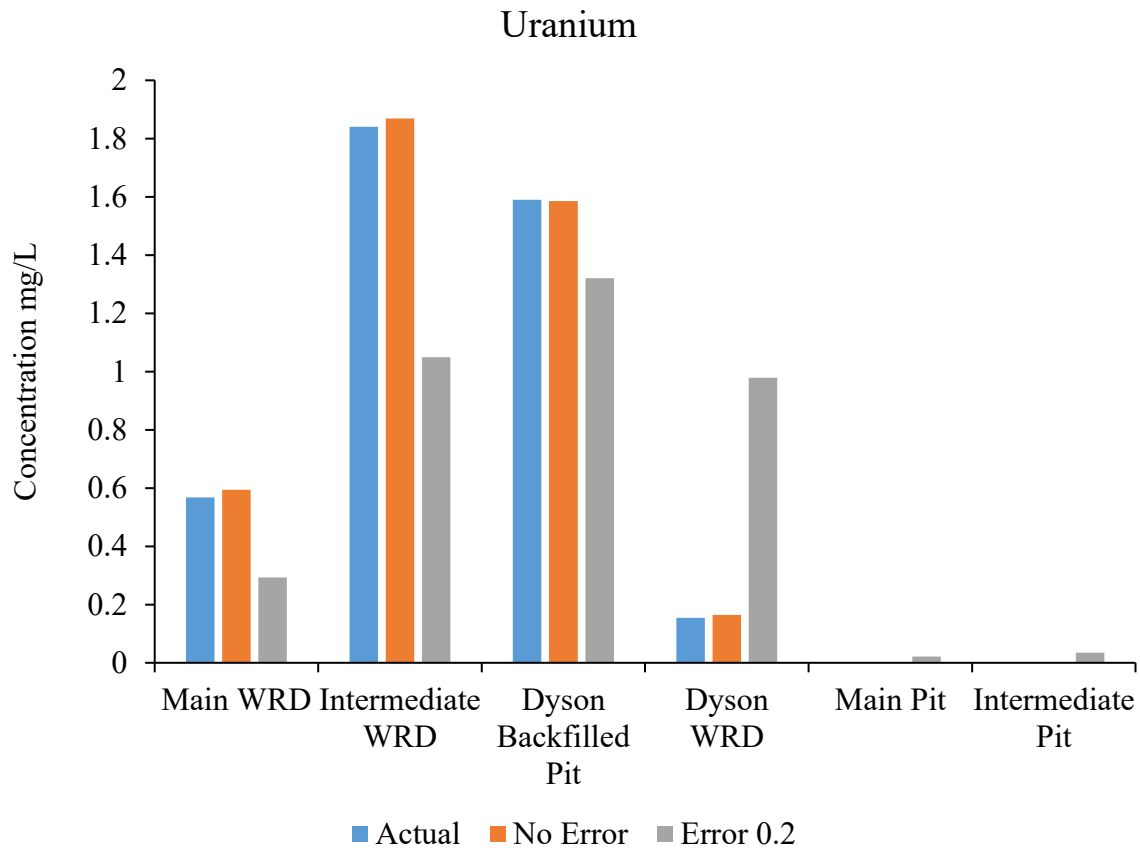


Figure 4.26: Comparison of UO_2^2species concentrations at sources with a perturbation error of 0.2

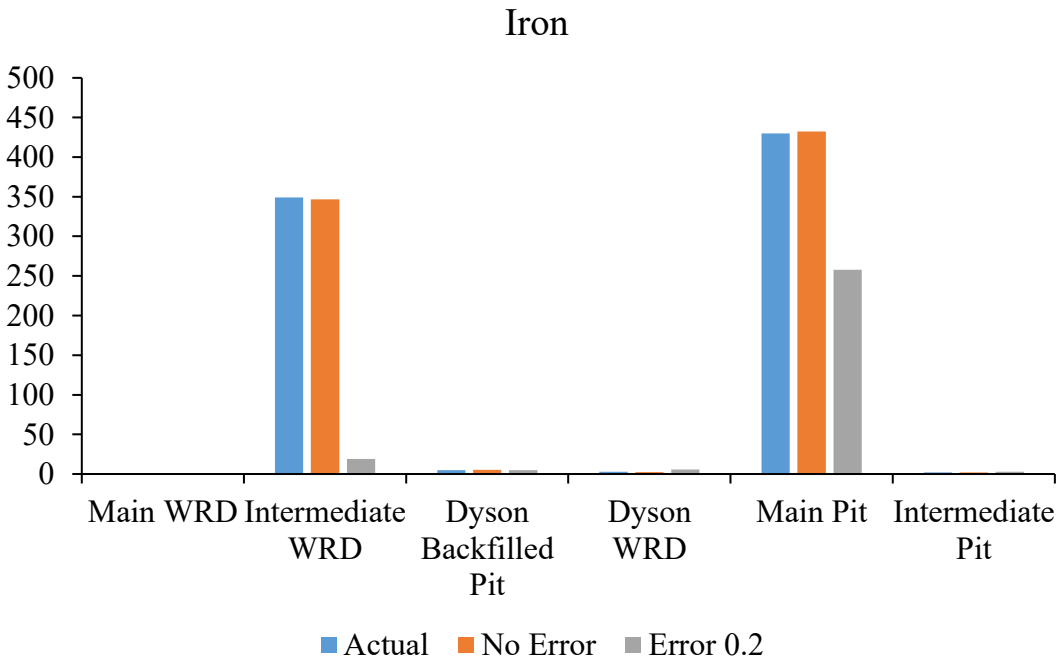


Figure 4.27: Comparison of Fe^{2+}species concentrations at sources with a perturbation error of 0.2

Figures 4.28 to 4.31 show the errors in source estimation for different scenarios. The figures compare the errors in a linear format with four error perturbation values of 0.1, 0.2, 0.15 and 0.05. The error graph justifies the fact that when concentrations are perturbed, as in the case of field measurements where uncertainties in measurement or recording are likely, the optimization model shows that it is robust and capable of handling any form of data and, at the same time, providing an optimal solution of source characteristics with a realistic margin of error. Figures 4.30 to 4.31 show how minimal errors are observed when erroneous data (randomly perturbed) are used as input for source characterization. In Figures 4.30 to 4.31, the *y*-axis represents source estimation errors. The *x*-axis depicts the different source locations, where 1 = Main WRD, 2 = Intermediate WRD, 3 = Dyson Backfilled, 4 = Dyson WRD, 5 = Main Pit, 6 = Intermediate Pit.

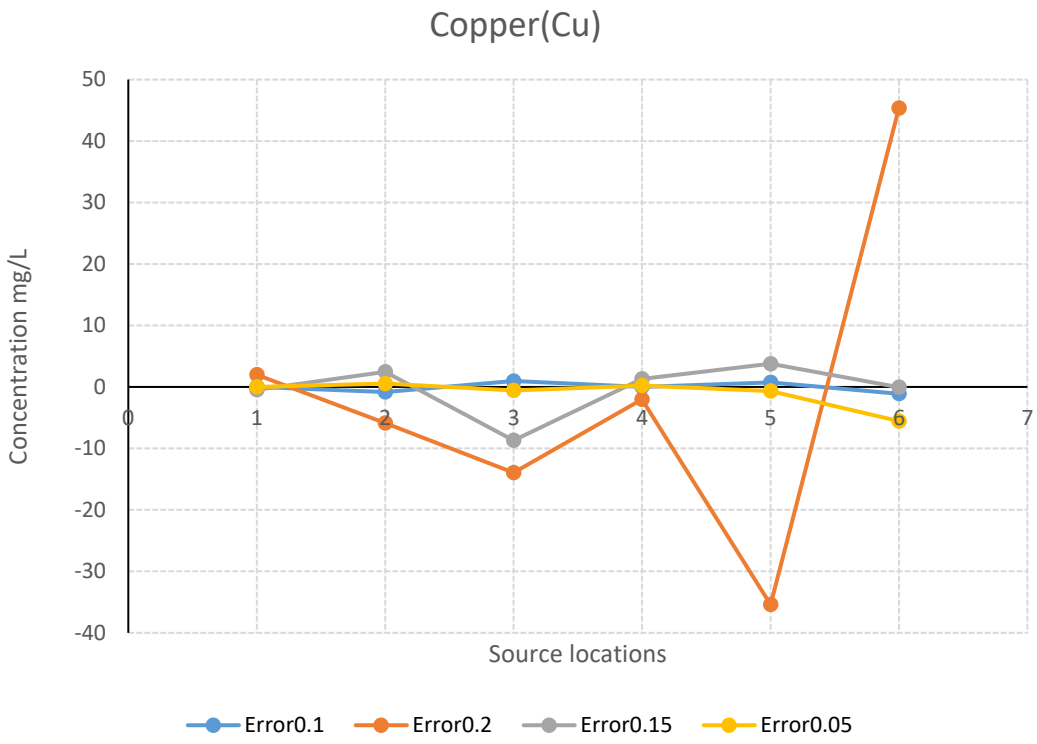


Figure 4.28: Comparison of the errors obtained using different amounts of error

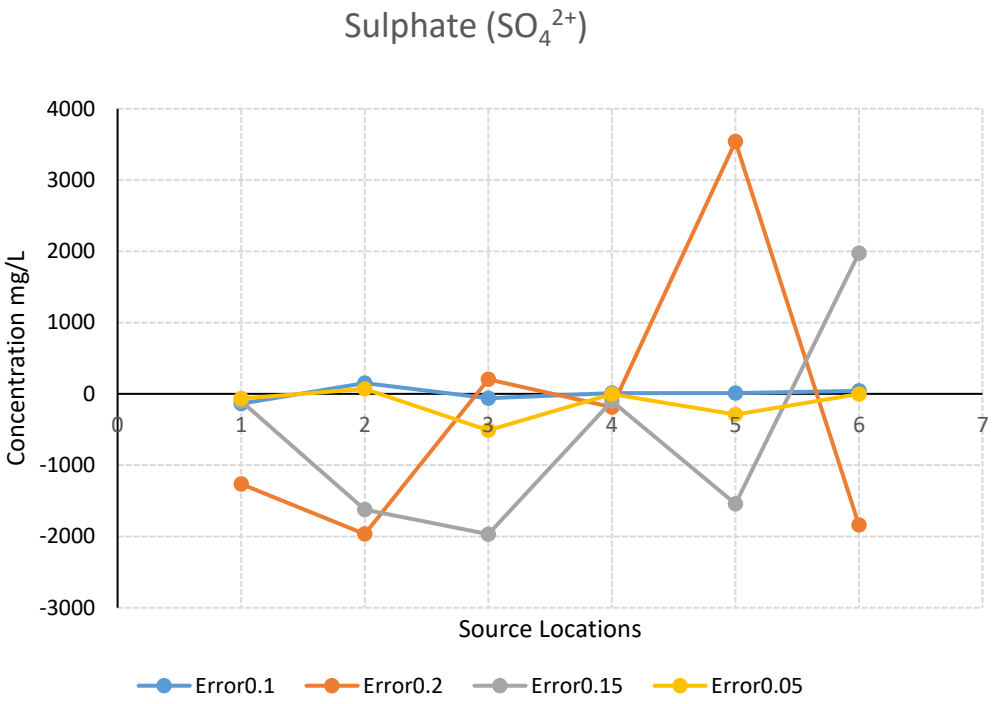


Figure 4.29: Comparison of the source identification of sulfate errors obtained using different erroneous measurement data

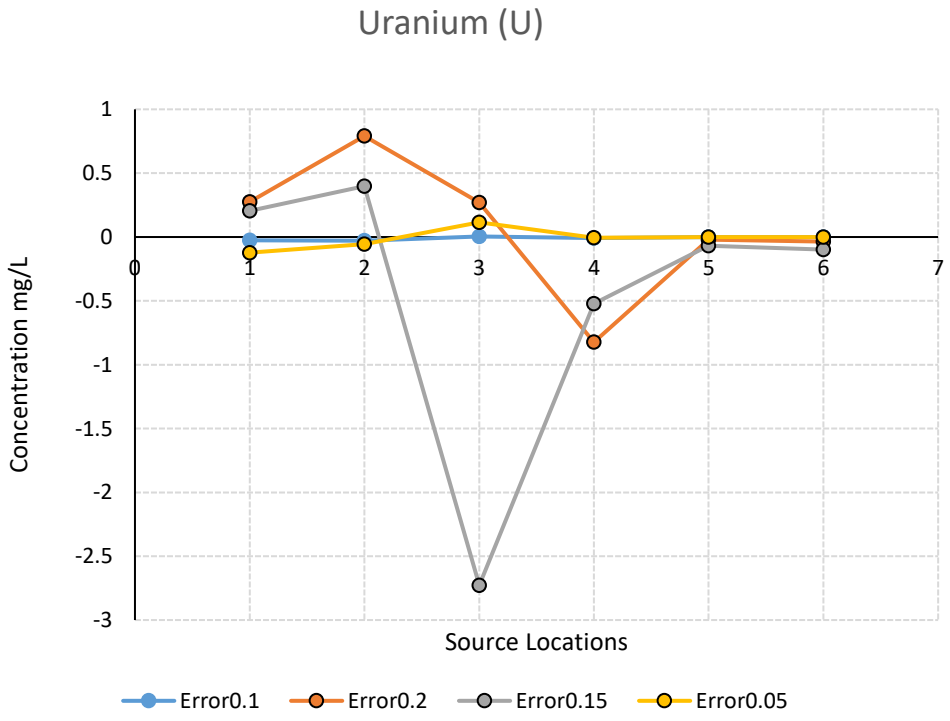


Figure 4.30: Comparison of the errors obtained using different amounts of error Comparison of the source identification errors obtained using different erroneous measurement data

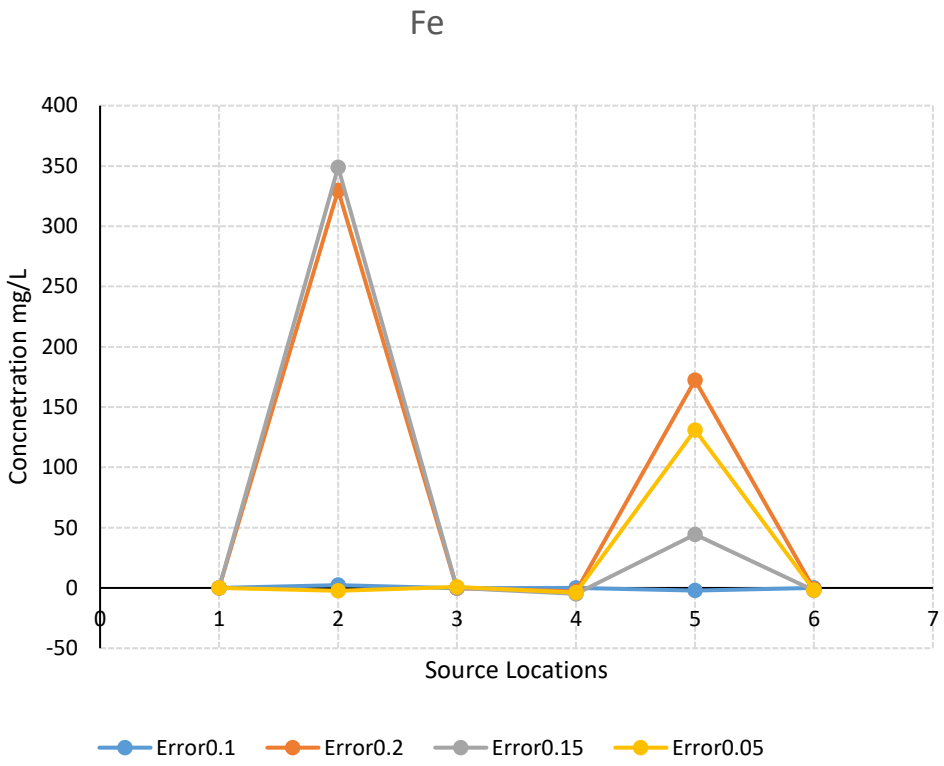


Figure 4.31: Comparison of the errors obtained using different amounts of error

4.10 CONCLUSIONS

This chapter described the application of a source characterization method. The method uses a linked ASA optimisation algorithm to model the transport of multiple chemically-reactive species in a contaminated aquifer beneath a former mine site in Australia. A numerical simulation model of the flow and transport processes was developed that incorporates heterogeneous, anisotropic, hydrogeological parameters. A performance evaluation demonstrated the potential applicability of this method to simultaneously identifying the spatial distributions and input concentrations of unknown areal groundwater contaminant sources. The method uses a limited amount of contaminant species' concentration data at different time intervals and monitoring well locations.

The preliminary evaluation results are encouraging and point towards the feasibility of using the proposed method to optimally characterize the sources and pathways of contamination in complex aquifers. The range of errors obtained for error

and erroneous cases. For example between 4% and 16% etc. Also mention may be erroneous measurements are actually randomly perturbed concentration measurements

This proposed methodology can overcome some of the shortcomings of some of the currently available methods applied to optimal characterization of unknown contaminant sources, particularly at very complex contaminated aquifer sites like abandoned mines that contain multiple species of reactive chemical contaminants. Such characterization is an essential initial step for solving critical environmental problem and to design effective contamination remediation strategies. This study also highlights limitations in utilizing and implementing calibrated flow and transport simulation models calibrated with very limited available site measurements, particularly in hydrogeologically- and geochemically-complex contaminated aquifer sites.

The next chapter (Chapter 5) presents the development and application of a fractal/multifractal modelling approach to the design of optimal monitoring networks. The performance of such a network at a contaminated aquifer site is evaluated in terms of contaminant source characterization accuracy and efficiency.

Chapter 5: Integrated Monitoring Network Design for Improved Contamination Source Characterization at a Mine Site

In this chapter, multifractal modelling is used to determine monitoring well locations that are optimal for assessing the spread of contaminants. The fractal mapping process provides input data for the design of a network of optimal monitoring locations. Multi-objective optimization is performed to target potential relevant locations in the study area and to eliminate locations that will not aid source identification, with consideration of uncertainty in hydrogeological conditions. A performance evaluation shows that source identification efficiency is improved when concentration measurements from an optimally-designed monitoring network are used.

5.1 BACKGROUND TO THE PROBLEM

Characterizing unknown contamination sources is the first step in groundwater remediation and management strategies. At abandoned mine sites, long-term pollution can cause long-term environmental impacts due to contamination from waste rock dumps, tailing dams and open pits exposed to oxygen and rainfall, which leads to acid mine drainage (AMD). To be able to efficiently monitor and detect contamination in groundwater due to persistent and undetected AMD, more groundwater quality measurements are needed. However, data availability is often unsatisfactory for efficiently solving groundwater problems. Unrelated, inadequate or inefficient data collected at less relevant monitoring locations or at the wrong time can constrain the quality of groundwater contaminant data. Hence, it is necessary to implement an optimally-designed monitoring network that can collect data relevant and adequate to source characterization. Also, such water quality monitoring data are essential for assessing the spatial and temporal distributions of contaminants in a groundwater system. Furthermore, contaminant concentration monitoring is essential for

contaminant plume detection and the effective management of contaminated aquifers. The uncertainties associated with the mapping of contaminant plumes, and the financial limitations related to installation of monitoring wells, require the design of an optimal monitoring network, as defined by the objectives.

Monitoring networks are integral part of efficient, effective, and economical management of the groundwater. Since the remediation of contaminated groundwater is prolonged and cost-intensive, detailed information concerning the contamination source characteristics is critical. Hence, an optimal monitoring network that can gather reliable and efficient data is essential for the identification of groundwater contaminant sources and to inform models that accurately describe groundwater flow and contamination transport.

One of the fundamental aspects of appropriate monitoring network design is to identify potential bore/well locations. These must be carefully selected according to the contamination situation. One effective way to determine potential bore/well locations is to map the boundaries of contaminant plumes so that all locations are marked within the boundaries. In this study, fractal modelling (Wang & Zuo, 2015; Datta et al., 2016; Esfahani & Datta, 2018) is applied to map contaminant plume boundaries in contaminated aquifers. This is used as initial input data for the monitoring network's design.

Groundwater contaminant concentrations obtained at multiple locations are useful for estimating source characteristics such as magnitude, spatial distribution, and activity duration. However, it is challenging to identify the optimal locations for monitoring unknown contaminant sources, particularly when the sources are distributed and data are scarce. Moreover, uncertainties related to source flux distribution predictions, hydrogeochemical parameters and field-measured concentrations make accurate source characterization even more challenging. The aim of this study was to develop a method for designing a monitoring network that is optimal for detecting groundwater contamination and assessing its spread at an abandoned mine site.

In real-world groundwater remediation problems, competing objectives must be considered simultaneously. When multiple objectives compete, it can lead to a sequence of solutions recognized as *Pareto-optimal* or *non-dominated solutions* (Qiankun et al., 2016). This study develops a multiple objective optimization model

coupled with a flow and transport simulation model to search for Pareto-optimal solutions to a multi-objective optimization problem associated with groundwater monitoring network design (Esfahani & Datta, 2018). The multi-objective design is based on Pareto-optimal solutions with two objectives 1) minimize the maximum normalized error between contaminant concentrations estimated via interpolation of measurements from designed monitoring network locations, and actual measurements from potential monitoring well locations, and 2) maximize the summation of the product of estimated concentration gradients and the simulated (or measured) concentrations at the selected monitoring locations.

There is a considerable volume of literature relating to the design of optimal monitoring networks. These designs have used heuristic optimization tools like genetic algorithms (GA; Cieniawski et al., 1995; Wu et al., 2005; Yeh et al., 2006; Chadalavada et al., 2011), simulated annealing (SA; Prakash & Datta, 2012, 2015) and genetic programming (GP; Prakash & Datta, 2013; Datta et al., 2013, 2014). Datta and Dhiman (1996) designed a monitoring network using an optimization algorithm known as MIP. Chance constraints were used to define the reliability of the proposed network. Data interpolation techniques like geostatistical kriging have been used in groundwater monitoring network design (Yeh et al., 2006; Feng-guang et al., 2008; Chadalavada et al., 2011; Prakash & Datta, 2012). Chadalavada and Datta (2007) proposed models for determining optimal and efficient sampling locations for contamination plume detection. Dhar and Datta (2010) developed an optimization-based solution for reducing redundancy in a groundwater quality monitoring network. Bashi-Azghadi and Kerachian (2010) developed a new methodology for optimally locating monitoring wells using a Monte Carlo technique. The main aims of all these methods have been to reduce the cost of monitoring by reducing redundancy, and to improve pollutant detection in uncertain dynamic scenarios.

Initial attempts to design optimal sampling networks to improve the accuracy of pollutant source identification were made by Mahar and Datta (1997), Datta et al. (2009a, b) and Prakash and Datta (2013, 2014). Datta et al. (2013) used GP-based monitoring factors in the design of an optimal monitoring network to improve the accuracy of pollutant source identification. The potential applicability of GP to groundwater management problems was first demonstrated by Sreekanth and Datta (2012). These studies used trained GP models to calculate the impact factor of the

sources on the candidate monitoring locations. The issue of sequentially designing monitoring networks to gather feedback information regarding compliance with saltwater intrusion management strategies in coastal aquifers was discussed in Sreekanth and Datta (2014, 2015). Datta and Singh (2014) presented a method that uses a kriging-linked optimization model incorporating uncertainty to design contaminant monitoring networks. Esfahani and Datta (2018) proposed the use of fractal singularity and GP surrogate models linked to a multi-objective optimization algorithm to solve a two-objective optimal monitoring network design model.

This chapter presents an optimal monitoring network design methodology that uses a linked, adaptive simulated annealing, ASA-based, multi-objective optimization model. Fractal/multifractal modelling (Datta et al., 2016) is employed to delineate and estimate contaminant plumes and to determine optimal locations for an efficient monitoring network. The objective is to determine the best Pareto-optimal solutions based on the ASA algorithm. The ASA search algorithm's candidate population size is exploited to construct an estimated Pareto front by gathering non-dominated solutions while exploring the feasible domain. Then, different monitoring network patterns are evaluated to determine the best compromise between cost and source identification accuracy.

5.2 METHODOLOGY

In this section, a multi-objective simulation-optimization model for optimal networking monitoring design based on a two-way approach is discussed.

The first step in the approach consists of two parts. The first part is to quantify the spatial distributions of contaminants concentrations using the multifractal spectrum. The second part identifies the contaminant concentration boundaries using fractal/multifractal models. Fractal/multifractal modelling is used to calculate the singularity indices of the area of study, making use of measured concentration data. Generally, concentrations are interpolated at each node, based on initially available concentrations measured at arbitrary locations. Interpolated concentrations are obtained by ordinary kriging and generating a gridded content map representative of the entire study area. Concentration values, measured or interpolated, are needed at every node to calculate the singularity indices. The calculated indices representing

mapped areas are categorized into two subsets, namely, contaminated and uncontaminated (clean) areas. The regions connecting these two categories are identified as the contamination plume border. The singularity indices in multifractal modelling are regarded as a structural characterization property (Chen et al., 2007) of the contaminant concentration spread in an aquifer. In terms of monitoring network design, if the monitoring design model is used only in evaluation mode for unknown contaminant source characterization, and the contaminant source characteristics are known, these concentration values can be obtained from the groundwater flow and transport simulation model using source characterization values as inputs.

The second step of the method has three main parts: i) groundwater flow and transport simulation, ii) interpolation using kriging and iii) optimization using ASA. A multi-objective optimization formulation is applied to design a set of monitoring networks based on the optimization algorithm (ASA) with a maximum number of monitoring locations/wells set as a constraint. The optimization model is a two-objective formulation (usually conflicting ones) that focuses on minimizing the maximum normalized errors between actual (observed) concentrations and those estimated (simulated) by the interpolation model based on data from monitoring locations, and maximizing the weighted concentration gradients (Prakash & Datta, 2014; Esfahani & Datta, 2018) at the selected monitoring locations. Pareto-optimum monitoring networks are implemented based on the solutions of the two-way objective formulation that are considered Pareto-optimal.

5.2.1 Fractal/Multifractal Modelling

In this section, fractal/multifractal modelling is described that uses a program coded in MATLAB software for processing geochemical data. This MATLAB program, known as the Anomaly Identification System (AIS; Zuo et al., 2013; Datta et al., 2016) consists of two key procedures. The first function quantifies the spatial characteristics of geochemical patterns (in this study, the spatial distribution of contaminant concentrations) using the multifractal spectrum. The second function classifies geochemical anomalies using singularity index analysis (in this study, concentration differences/delineation). The singularity index is used to create a singularity distribution map that estimates the expected contamination plume boundary and delineates potential areas for monitoring wells from which optimal locations are chosen. The main implementation processes of the fractal/multifractal

modelling are outlined in the steps below (Zuo et al., 2013). To identify geochemical anomalies, the steps below are used.

1. Grid map generation by interpolation of the original contaminant concentration data.
2. Conversion of the interpolated map from the spatial domain to the Fourier domain using fast Fourier transform.
3. Calculation of the power spectrum of the converted map to form a dataset consisting of the power spectrum density (E) and the area with power spectrum density values ≥ E. Plotting of these data on a log-log scale.
4. Find breakpoints to divide the data pairs into several segments with different scaling properties, and use these to build up filters.
5. Apply the filters to the map in the frequency domain and transform the filtered map back to the spatial domain using inverse Fourier transform.

The next stage is to divide the background and anomaly maps, then estimate the singularity index from the concentration map. This forms the original singularity analysis. The following steps are used (Zuo et al., 2013).

1. Define a sequence of sliding rectangular windows $A(r_i)$ with variable window sizes $r_{min} = r_1 < r_2 < \ldots < r_n = r_{max}$ for a location on the concentration map of the study area.
2. Calculate the average element concentration $C[A(r_i)]$ for each window size (this is equal to the sum of all the cells' concentrations divided by the total number of cells within the window).
3. Plot $C[A(r_i)]$ against r_i on a log-log graph to obtain a linear relationship as per the equation:

   $$\log C[A(r_i)] = C + (\alpha - 2) \log r_i \qquad (5.1)$$

4. Estimate the singularity index (α) at locations of interest as the slope plus 2 using the linear relationship of Equation 5.1.
5. Create a singularity distribution map and delineate areas of concentration anomalies based on the singularity index (α), where $\alpha \approx 2$ indicates a boundary, $\alpha < 2$ indicates the inside and $\alpha > 2$ indicates the outside of the boundary.

The schematic diagram in Figure 5.1 describes the structural and operational processes involved in the fractal/multifractal modelling in the MATLAB-based AIS program.

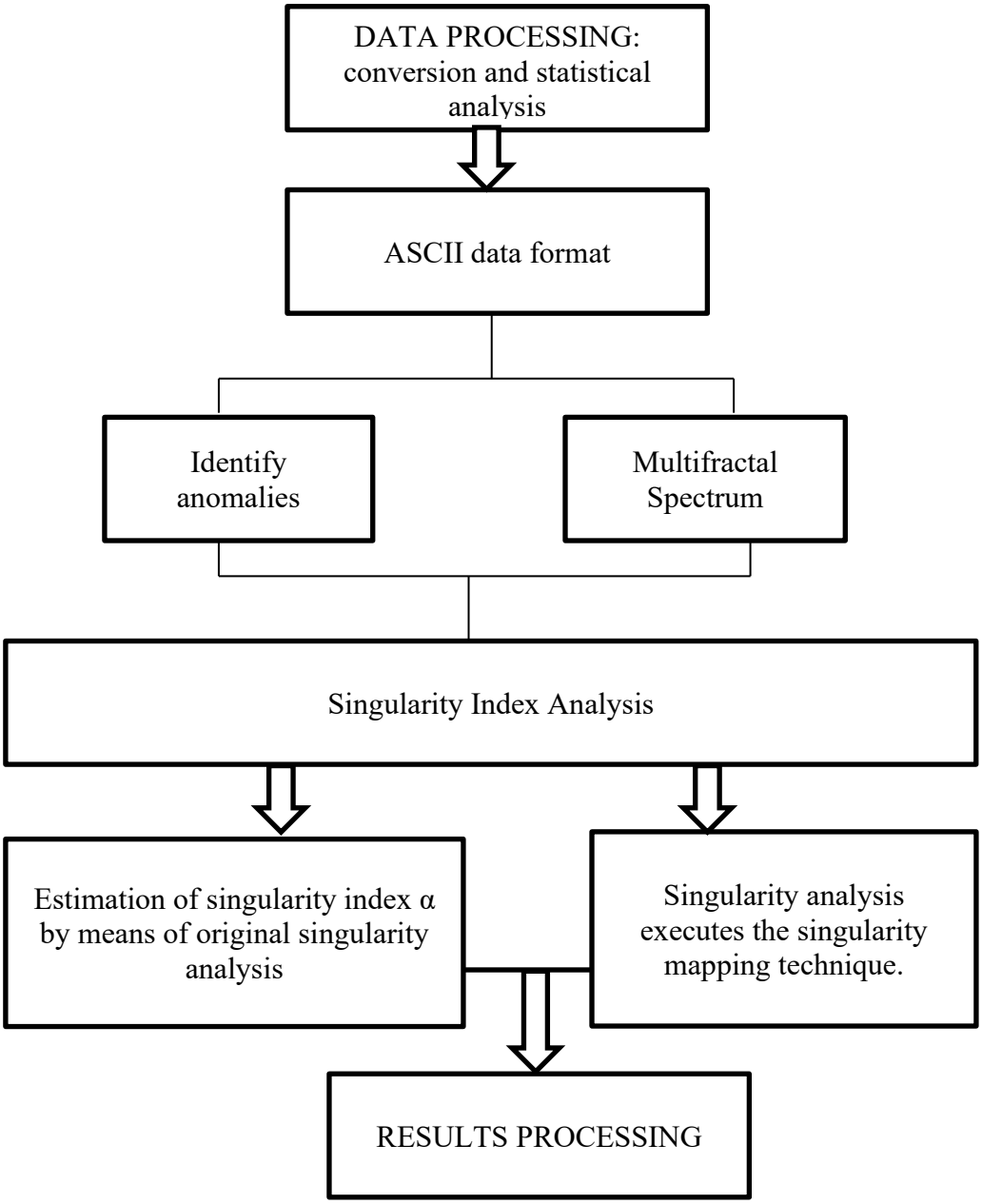


Figure 5.1 Flowchart of the structure and operational procedure of AIS (Wang & Zuo, 2015)

For the purposes of monitoring network design, the approach of applying a singularity index to map the boundary of the contaminant contamination plume and determine potential well locations and numbers can be viewed as a search strategy objective for identifying source characteristics. In this study, the singularity mapping technique of fractal/multifractal modelling in two-dimensional raster data is

characterized as a power-law relationship in subareas in a sampled zone (Datta et al., 2016a). The singularity mapping technique in fractal/multifractal modelling can be useful for detecting the edges and boundaries of different frames. The boundary of a groundwater contaminant plume is mapped on the basis of a practical horizontal gradient existing close to the plume boundary, which relates to the inflection point of an anomaly. The plume boundary also corresponds to the inflection point of the anomaly. Therefore, singularity index $\alpha < 2$ indicates high-density and positive singularity, $\alpha > 2$ indicates low-density and negative singularity, and $\alpha = 2$ indicates non-singularity of linear behaviour and a boundary.

5.2.2 Formulation of a Multi-objective Optimization Model for Monitoring Network Design

Two objectives were considered to design an optimal monitoring network. Two different objectives were defined. These two are: to maximize the estimated weighted concentration gradients, and to improve the accuracy of contaminant plume mass estimate. Therefore, these two objectives to be optimized (Prakash & Datta, 2014; Esfahani & Datta, 2018) are:

i. Minimization of the normalized error between contaminant concentrations estimated by interpolating concentrations measured at designed monitoring network locations, and actual contaminant concentration at potential monitoring well locations, and

ii. Minimization of the summation of the product of estimated concentration gradients and the simulated (or measured) concentrations at selected monitoring locations.

The optimal monitoring locations obtained as solutions are a subset of the specified potential monitoring locations. As required for multiple objective optimization, the two selected objective functions are conflicting in the sense that the objective function for finding the well locations with the weighted maximum concentration gradient conflicts with the objective function of finding well locations that minimize the normalized error between the simulated and estimated concentrations at potential well locations. A trade-off between these conflicting objectives exists at the Pareto optimal set of solutions. One of the objectives must be compromised in order to improve the other objective and vice versa. As a result, a multi-objective optimization model was formulated to design an optimal monitoring

network that achieves these two objectives at optimal level. The multi-objective optimization model is solved using the constrained method (Datta and Peralta, 1986). This method optimizes one single objective subject to the other objective being defined as an implicit constraint to be satisfied (the constrained method). The level of the implicit constraint to be satisfied is varied iteratively to obtain the Pareto optimal solution set. The upper limit specified for the number of monitoring wells selected from the specified set of all potential well locations is essentially governed by budgetary constraints. The two objectives of the multi-objective optimization model proposed by Prakash and Datta (2014) and Esfahani and Datta (2018) for designing an optimal monitoring network that can accurately identify pollution sources are defined by Equations (5.2) and (5.3).

The multi-objective optimization formulae can be mathematically expressed as:

$$F1 = Minimize\left(Maximum\left(\frac{C^*_{i,j} - C_{i,j}\,\text{int}}{C^*_{i,j} + N} \right)\left(f_{i,j}\right) \right) \forall i, j \tag{5.2}$$

$$F2 = Maximize\left(\sum f_{i,j} C^*_{i,j} \left\{ \frac{\left|C^*_{1-1,j} - C^*_{i,j}\right| + \left|C^*_{1+1,j} - C^*_{i,j}\right|}{dx} + \frac{\left|C^*_{1-1,j} - C^*_{i,j}\right| + \left|C^*_{1+1,j} - C^*_{i,j}\right|}{dy} \right\} \right) \tag{5.3}$$

$$\mathrm{C}_{max} \geq C^*_{i,j} \geq \mathrm{C}_{min} \;\forall i, j \tag{5.4}$$

$$C_{i,j}\,\text{int} = \text{INTRP}\,(f(i,j)*\text{C}*(i,j)) \;\forall i, j \tag{5.5}$$

$$\sum f_{i,j} \leq \text{K(upper limit on number of monitoring wells)} \;\forall i, j \tag{5.6}$$

The binary variable ($f_{i,j}$) represents a "yes or no" decision on selecting a potential monitoring location. The constraint formulated in Equation (5.5) describes the spatially-interpolated concentrations at all the candidate monitoring locations selected for that monitoring network design optimization iteration. INTRP (f (i,j)*C*(i,j)) represents the spatially-kriged interpolated concentration values at node i,j, which are interpolated based on the designated monitoring locations given by $f_{i,j} = 1$.

A value of $(f_{i,j}) = 1$ represents a decision to install a monitoring well at location $i, j,$ while 0 represents a decision to not install a monitoring well. Basically, the optimal locations designated by fi,j are the decision variables. Other variables, such as those interpolated based on the concentrations, can be designated as decision variables. The potential locations coincide with the nodes, and the potential locations are only a subset

of the nodes. The constraining method of solving a two-objective optimization model is used. This method involves minimizing one of the objective functions (F1), with an additional constraint of ensuring a minimum level of satisfaction of the second objective function (F2). The constraint solution method is formulated as shown in Equation (5.7):

$$\sum f_{i,j} C^*_{i,j} \left\{ \frac{\left|C^*_{1-1,j} - C^*_{i,j}\right| + \left|C^*_{1+1,j} - C^*_{i,j}\right|}{dx} + \frac{\left|C^*_{1-1,j} - C^*_{i,j}\right| + \left|C^*_{1+1,j} - C^*_{i,j}\right|}{dy} \right\} - \gamma \geq 0 \qquad (5.7)$$

where γ is the minimum level of satisfaction of the second objective function F2, also termed the *trade-off constant*. Therefore, the resulting model can be solved iteratively as a single-objective optimization model for different satisfaction levels of γ; thus, a Pareto-optimal solution set is generated. The second objective function is incorporated as a new implicit constraint. The upper limit of γ is defined by the maximum value of the second objective function F2 when solved as a single-objective optimization with constraints (Equation 5.8) defining the lower limit of F2. The limit of γ is the value of the second objective function F2 corresponding to the minimum value of the first objective function F1, when the optimization model is solved as a single-objective model, with F1 as the only objective (Equation 5.9).

$$F2 \geq \gamma \qquad (5.8)$$

$$F2 \leq F_{2MaxF1} \qquad (5.9)$$

where F_{2MaxF1} is the maximum feasible value of the objective function F2 corresponding to the minimum (best) value of the first objective function F1 when solved as a single-objective model. All solutions obtained on a Pareto-optimal front correspond to a different Pareto-optimal monitoring network.

5.2.3 Linked Simulation-Optimization Model for Optimal Contaminant Source Identification

One of the intended key benefits of monitoring network design is to improve the accuracy of source characterization. Characterising sources based on the magnitude and location of contaminant release has, over the past three decades, been done successfully using a linked simulation-optimization approach (Datta et al., 2015). A linked simulation-optimisation model is made up of two parts: (a) a groundwater numerical model that simulates physicochemical processes and (b) an optimization model that estimates optimal source characteristics based on an optimization based

decision model. The binding constraint in linked simulation-optimization models is the groundwater simulation model. Hence, for groundwater source characterisation to be feasible, the optimization management model must fulfil the conditions of the flow and transport models. The benefit of a linked simulation-optimisation approach is its ability to link complex numerical models with real field models in an optimization model. In this work, a linked simulation-optimization-based approach is used to characterise contaminant sources. It comprises numerical flow and transport models that describe the transport of multiple reactive chemical species in an aquifer beneath a mine site contaminated by waste rock dumps and mineral waste deposits.

In this work, a numerical groundwater flow and transport simulation model developed based on hydrogeochemical parameters of the study area is used. It is linked to an ASA optimization algorithm within a linked simulation-optimization model. It is used to characterise the sources of contaminant in terms of location, magnitude and timing of activity based on information obtained from a monitoring network. In the performance evaluation reported here, the numerical simulation model is a three-dimensional finite element-based flow and transport model run in HYDROGEOCHEM 5.0 software (Yeh et al., 2004; Steefel et al., 2015).

5.2.4 Groundwater Flow and Transport Simulation Model

A three-dimensional groundwater simulator was used to simulate the flow and transport processes of the study area. The details of the governing equations of the flow and transport process are described in Chapter 3, Section 3.3. (Yeh et al., 2004; Steefel et al., 2015).

$$\frac{D}{Dt}\int_v \theta\, C_i\, dv = -\int_\Gamma n \bullet (\theta\, C_i)\; V_i d\,\Gamma - \int_\Gamma J_i d\,\Gamma + \int_v \theta\, r_i dv + \int_v M_i dv,\ \mathrm{i} \in \mathrm{M} \qquad (3.7)$$

The general transport equation incorporates advection, dispersion/diffusion, source/sink, and biogeochemical reactions. The flow and reactive transport modelling with this numerical model were described in Chapter 3.

5.2.5 Developing a Numerical Simulation Model Linked to a Source Identification Model

This study uses a source identification model that incorporates a numerical simulation model with an ASA optimization algorithm. It is used to determine groundwater distributed source characteristics in a contaminated former mine site. The optimization model solve for candidate solutions of source concentrations based on

the concentrations of different chemical species obtained from the numerical simulation model. The simulation model is coupled to the ASA algorithm. The simulation model uses the source concentration candidate solutions generated by the optimization algorithm to estimate contaminant concentrations at monitoring locations. Then, the optimization algorithm evaluates the objective function. The objective function is a function of the differences between measured concentrations and those estimated by the numerical simulation model. Optimal source characterization is achieved by solving the optimization model to minimize the objective function. Mahar and Datta (2001) defined the objective function of a simulation-optimization model for source characterization as:

$$\text{Minimize F1} = \sum_{k=1}^{n_k} \sum_{iob=1}^{n_{ob}} \left(C^k_{obs_{iob}} - C^k_{sim_{iob}} \right)^2 \bullet w^k_{iob} \qquad (5.10)$$

$$\text{Subject to } C^k_{sim_{iob}} = f(x, y, z, C_{sim}) \qquad (5.11)$$

$$\text{The weight } w^k_{iob} \text{ can be described as } w^k_{iob} = \frac{1}{\left(C^k_{obs_{iob}} + \eta\right)^2} \qquad (5.12)$$

The value of η is a constant, appropriately large such that errors at low concentrations of $C^k_{obs_{iob}}$ (cobs) do not dominate. This study chose a value of η within the range 100–1000 depending on the concentrations of contaminant species and the source locations. The constraints set (Equation 5.11) represent the simulation model for flow and transport processes.

The present study uses the ASA optimization algorithm to solve the optimal source characterization model. This algorithm was selected as offers a statistical guarantee of global convergence to an optimal point. Additionally, ASA was found to regularly attain a global minimum and be efficient in attaining regular minima at each comparable number of generated states.

5.3 PERFORMANCE EVALUATION

The performance of the proposed monitoring network design methodology was evaluated for a heterogeneous, anisotropic, complex aquifer system. The evaluation determined how the monitoring network design and subsequent concentration measurements obtained from it improved the characterization of distributed sources.

For the purposes of performance evaluation (to test the robustness and dependability of the method), this study used simulated concentrations from the groundwater simulation model. In real applications, it is necessary to measure concentrations in the field and interpolate them according to each node on a gridded map. This is because source concentrations will be unknown and the simulation model will not be used for simulating synthetic concentration values. However, for the purposes of this performance evaluation, it was important to use synthetic measurements acquired by specifying known sources and simulating concentrations using the numerical model. These synthetic measurements were used solely to assess whether the source characterization method is capable of identifying and recovering source characteristics for use in simulating concentration measurements.

Another important reason to use synthetic concentrations in the evaluation is that it avoids unreliable field data and model calibration issues. For the concentrations used for performance evaluation, it was assumed that the concentrations from the sources were constant over a stress period within a simulation period of 730 days. Six possible source concentrations (S1, S2, S3, S4, S5 and S6) were considered as explicit unknown variables for identifying sources using the source characterization model.

Contaminated aquifers where several chemically-reactive species are involved in groundwater transport, and where several distributed sources are involved in the system, are deemed highly complex. It is, therefore, important to design an effective monitoring network for the efficient characterisation of unknown distributed sources, as the networks provide effective monitoring of contaminant species. The source locations considered were the four main waste rock dumps and two open-pit lakes. The physical study area is described in Chapter 3.

To evaluate the performance of the proposed monitoring network in the characterization of distributed pollution sources using linked simulation-optimization, an illustrative, heterogeneous, anisotropic aquifer beneath a former uranium mining site was used. The Rum Jungle Mine in the Northern Territory, Australia, is a former uranium-copper project instigated to mine and export uranium for use in nuclear weapons. It made major contributions to the economy of the Northern Territory. The mine caused prevalent and ongoing environmental pollution, which reached many kilometres downstream. The Rum Jungle Mine site is located in the tropical wet-dry climatic region of northern Australia (Figure 5.2). The region is characterized by a

tropical savannah-like climate and typically receives about 1500 mm of annual rainfall. Most of this rainfall (90% or more) occurs during a wet season that lasts from November to April, with no sustained rainfall occurring from May to September. Rum Jungle was a major source of pollution of the Finniss river due to tailings and liquid waste discharges, and also due to acid mine drainage (AMD) from tailings and, especially waste rock dumps.

The geological zone of the Rum Jungle Mine site is within the Pine Creek Geosyncline. The regional geology includes Palaeoproterozoic metasediments unconformably overlying the Rum Jungle Complex. The Rum Jungle Complex consists mainly of granites and occurs primarily along the south-eastern side of the Giant's Reef Fault, whereas the Mount Partridge Group occurs north of the fault and consists of the Crater Formation, Geolsec Formation, Coomalie Dolostone and Whites Formation. The Rum Jungle mineral field contains numerous polymetallic ore deposits associated with the Rum Jungle Mine. Groundwater flow and solute transport models were constructed using the hydrogeological parameters listed in Table 5.1. The reaction equations and reaction coefficients incorporated in the reactive transport model are defined in Table 5.2.

Table 5.1: The Study Aquifer's Physical and Hydrogeological Properties

Aquifer parameter	Unit	Value
Area of mine site	km^2	12
Thickness of mine site	m	150
Number of nodes	-	6587
Number of elements	-	10,704
Horizontal hydraulic conductivity, K		
Layer 1	m/d	0.13
Layer 2	m/d	1.21
Layer 3	m/d	0.44
Layer 4	m/d	0.65
Layer 5	m/d	0.11
Layer 6	m/d	0.04
Vertical hydraulic conductivity, K		
Layer 1	m/d	0.12
Layer 2	m/d	1.21
Layer 3	m/d	0.44
Layer 4	m/d	0.65
Layer 5	m/d	0.11
Layer 6	m/d	0.04
Longitudinal dispersivity, αL	m	10
Transverse dispersivity, αT	m	0.1
Vertical dispersivity (αV):	m	0.01
Average rainfall	mm/year	2372
Effective porosity, θ	-	0.28

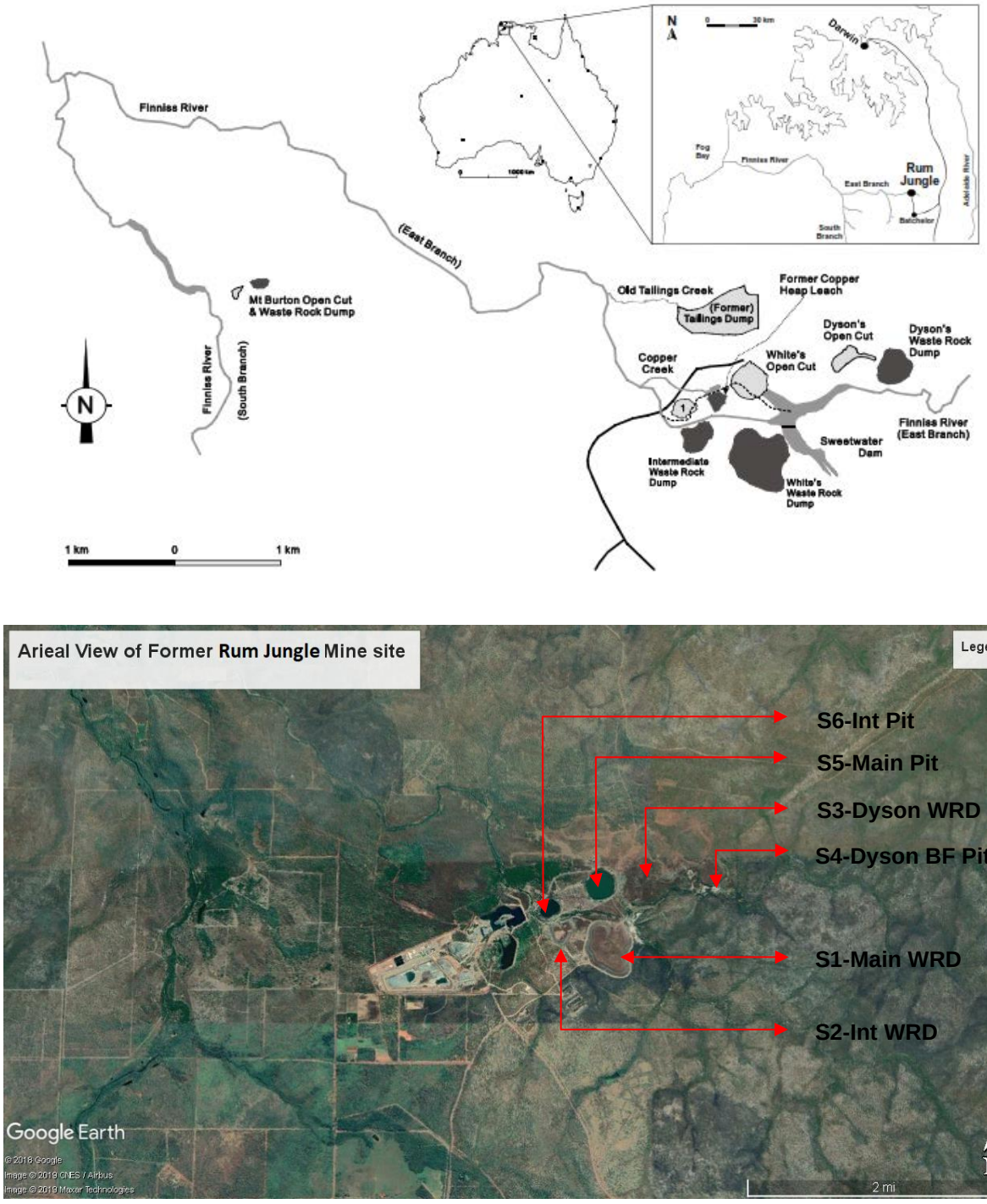


Figure 5.2: Map of the Rum Jungle study area (top) with locations of central mining zones (modified from Mudd & Patterson, 2010). Satellite image of the Rum Jungle site showing source locations (bottom).

The six potential sources are shown in Figure 5.2 (S1, S2, S3, S4, S5 and S6). The hydrogeological parameters used in the groundwater flow and solute transport model are given in Table 5.1. Similarly, the geochemical reactions and reaction coefficients used in the solute transport modelling are listed in Table 5.2. The two open pits (main and intermediate) were assumed to be contaminant sources. A boundary condition was assigned to the perimeters of the pits as an unknown constant concentration boundary condition. The four waste dumps were also considered to be contaminant sources of unknown concentration that drive the release and transport of several chemical species. Due to the complexities of the study area, when evaluating the performance of the proposed method, the unknown source concentrations were considered to be constant over the study period at each location.

Table 5.2: Chemical Reactions used in the Contaminant Transport Modelling

Reaction equation	Rate constant (Log K)
$H_2O(aq) \leftrightarrow H^+ + OH^-$ (R1)	-13.99
$H^+ + SO_4 \leftrightarrow HSO_4^-$	1.99
$Cu^{2+} + H_2O \leftrightarrow Cu(OH)^+ + H^+$	-9.19
$Cu^{2+} + SO_4^{2-} \leftrightarrow CuSO_4$	2.36
$Cu^{2+} + 2H_2O \leftrightarrow Cu(OH)_2 + 2H^+$	-16.19
$Cu^{2+} + 3H_2O \leftrightarrow Cu(OH)_3^- + 3H^+$	-26.9
$Fe2+ + H_2O \leftrightarrow H+ + FeOH^+$	-9.50
$Fe^{2+} + SO_4^{2-} \leftrightarrow FeSO_4$	2.20
$Fe^{2+} + 2H_2O \leftrightarrow 2H^+ + Fe(OH)_2$ (aq)	-20.57
$Fe^{2+} + 3H_2O \leftrightarrow 3H^+ + Fe(OH)_3^-$	-31.00
$Fe^{2+} + 4H_2O \leftrightarrow 4H^+ + Fe(OH)_4^{2-}$	-46.00
$Mn^{2+} + SO4^- \leftrightarrow MnSO_4$	2.26
$Mn^{2+} + H_2O \leftrightarrow MnOH^+ + H$	-10.59
$Mn^{2+} + 3H_2O \leftrightarrow Mn(OH)_3^- + 3H$	-34.08
$UO_2^{2+} + SO_4^{2-} \leftrightarrow UO_2SO_4$	3.15
$UO_2^{2+} + SO_4^{2-} \leftrightarrow UO_2(SO4)_2^{2-}$	4.14
$UO2^{+2} + 2H_2O \leftrightarrow UO_2(OH)_2$ aq + $2H^+$	12.15

5.3.1 Fractal Singularity Mapping of Potential Monitoring Well Locations

In this section, the application of fractal/multifractal modelling using singularity analysis for mapping potential well locations is discussed. The local singularity

mapping method was used to delineate the contamination plumes in the groundwater aquifer. The modelled groundwater contamination plume movement for this complex heterogeneous case study was obtained from flow and transport processes with interactions of reactive chemical species. The accuracy of the proposed method is evaluated for the case of a complex aquifer in a mine site area.

Fractal singularity analysis was applied to delineate the contamination plumes in the aquifer at a former mine site. For this case study, the contaminant sources were assumed to be active and, due to the closeness of the four waste dumps and two open pits, cross-sectional mixing or overlapping of contaminant plumes was expected.

As described in the Methodology section, when calculating the singularity index, some parameters are input earlier for the fractal modelling to proceed. Several sequences of square windows (representing concentration cells) with half-window sizes measuring were used to initiate the singularity analysis. The singularity analysis resulting in singularity index (α) values, were $\alpha < 2$ signifies the inside regions of a plume, while $\alpha > 2$ signifies outside regions. The singularity distribution map based on the singularity index illustrates that the contour for $\alpha = 2$ represents a plume boundary. This suggests that the singularity index can categorize potential monitoring locations effectively. Figure 5.3(a &a) shows a fractal singularity distribution map with estimates of singularity index α. Figure 5.4 shows potential well locations based on the singularity index analysis, as well as arbitrary potential well locations for comparison with the fractal monitoring network design. The red circles are singularity index-based potential well locations and blue circles are the arbitrary well locations.

It is important to note that, in this study, there were already some existing wells. Hence, the main aim was to use the network design to select optimal wells and determine optimal locations for new wells that could improve source characterization and identification.

Well locations were determined by 1) singularity distribution maps, 2) a monitoring network designed without singularity mapping, and 3) concentration measurements at arbitrary well locations. These results were used to assess the efficiency of the fractal singularity distribution-based design and to verify if the monitoring network design improved the characterization of contaminant sources.

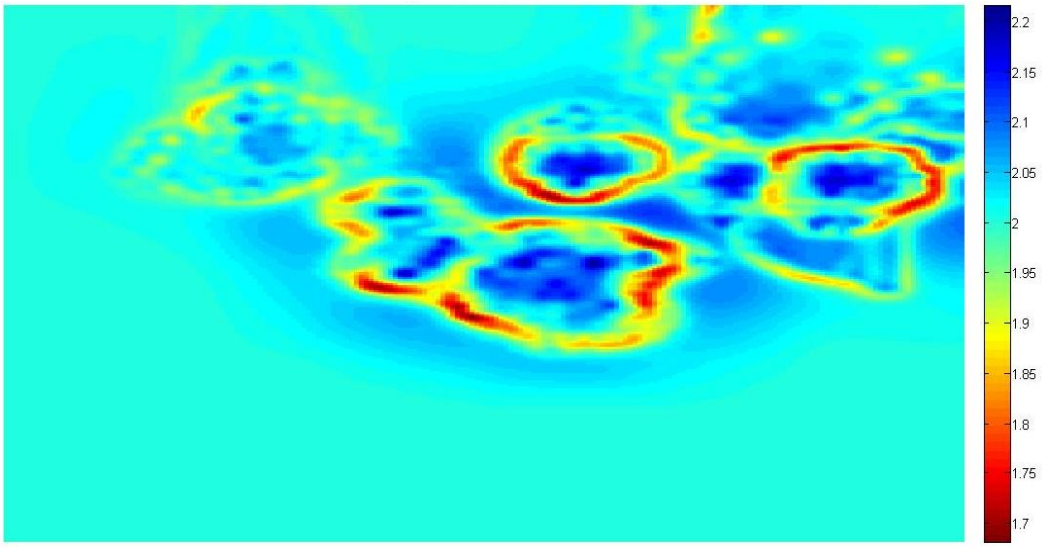


Figure 5.3(a): Singularity analysis showing the release of contaminant species 1 from distributed sources

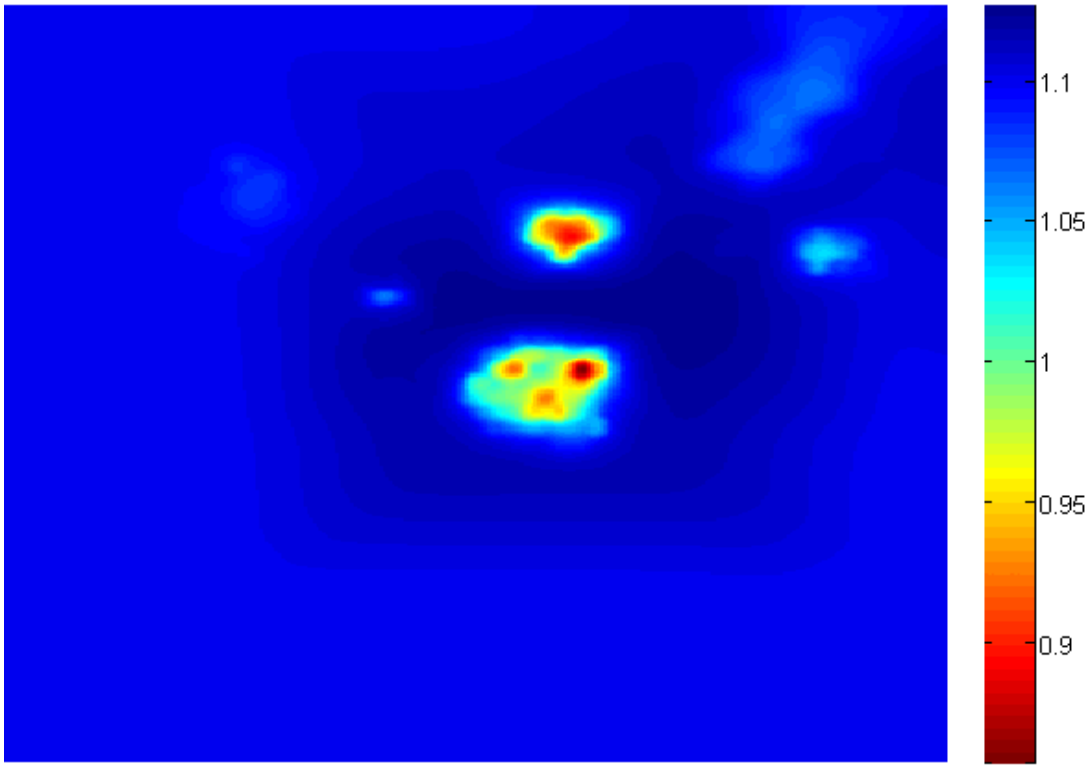


Figure 5.3(b): Singularity analysis showing the release of contaminant species 2 from distributed sources

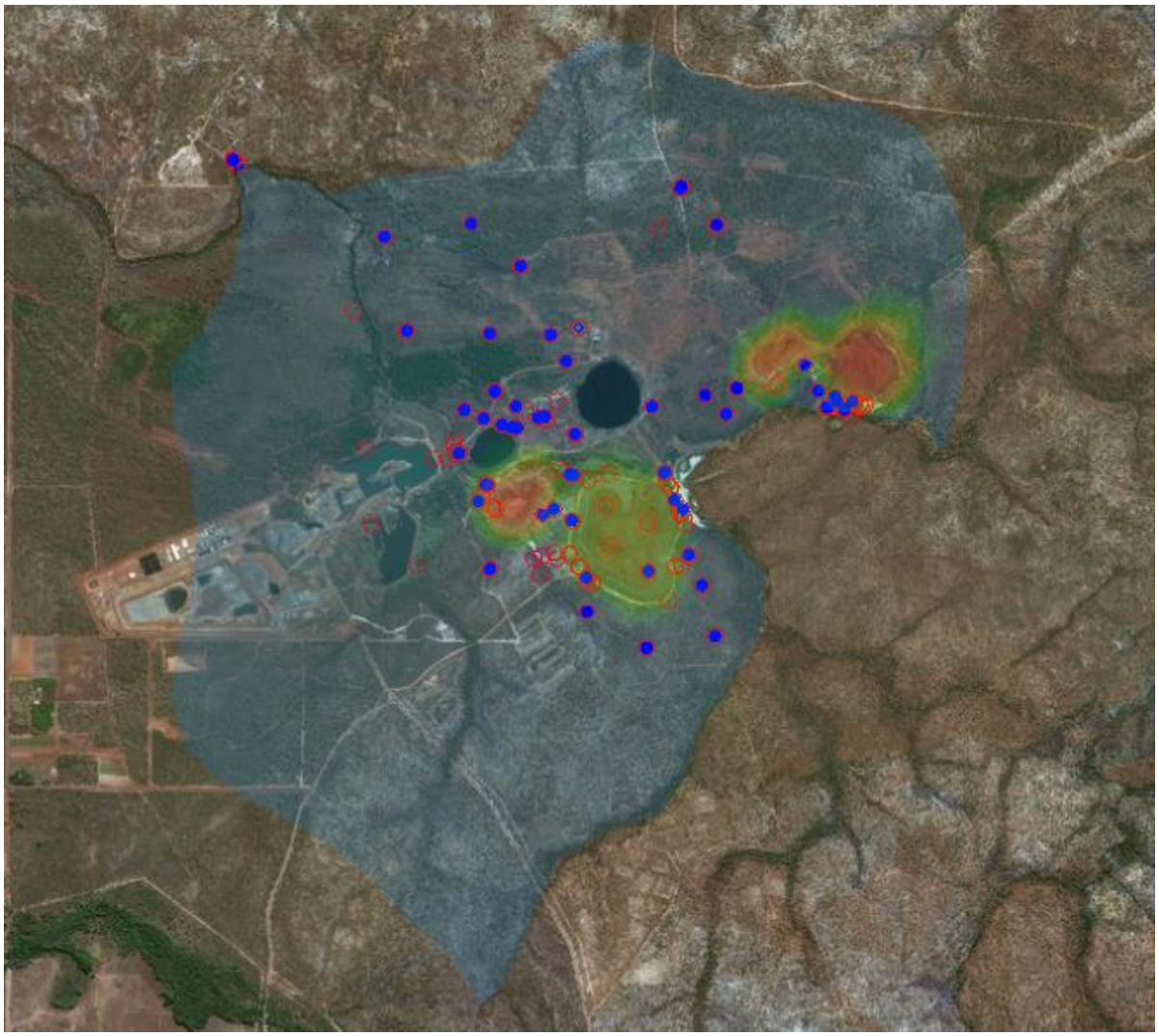


Figure 5.4: Aerial view of plume boundaries and potential well locations based on fractal singularity mapping and arbitrary designation Red circles are singularity index-based potential well locations and blue circles are the arbitrary well locations

5.3.2 Comparison of Optimal and Arbitrary Monitoring Networks

In order to obtain a monitoring network that has well locations suitable for improving source identification, optimally-designed and arbitrary monitoring networks were produced. Their effectiveness was then tested and compared. The optimal monitoring network was designed based on the Pareto-optimal solutions of the two objective function formulations F1and F2 (given in Equations 5.2 and 5.3). Kriging was applied in the design to spatially interpolate contaminant concentrations. The use of kriging is important as the monitoring locations and contaminant concentrations measurements were limited and spatially distributed. For the performance evaluation, the concentration values used were synthetic estimates

(simulation results) of the flow and transport simulation model. Two sets of six unique Pareto-optimal solutions were obtained as different sets of values representing F1and F2. One set of Pareto-optimal monitoring network solutions (labelled FD1 to FD6) was based on fractal singularity mapping and the other set (NFD1 to NFD6) was not. The first monitoring network group (FD1–FD6) showed a satisfactory Pareto-optimal value of the second objective F2 varying from a minimum of 2.35 to a maximum of 3.80. The second group of monitoring networks (NFD1–NFD6) obtained Pareto-optimal values for objective functions F2 with minimum and maximum values of 2.15 and 3.71, respectively, using Equations 5.1 and 5.2. Overall, 15 monitoring wells were chosen, which constitute the monitoring network for each Pareto-optimal solution.

Table 5.3. Pareto-Optimal Monitoring Network Design Solutions using Fractal Modelling (FD denotes a fractal mapping design.)

Monitoring network	Objective function value F1 (minimization)	Objective function value F2 (maximization)
FD1	0.82	2.35
FD2	0.84	2.83
FD3	0.89	3.21
FD4	1.01	3.50
FD5	1.13	3.70
FD6	1.215	3.80

Table 5.4. Pareto-Optimal Monitoring Network Design Solutions without Fractal Modelling (NFD denotes a non-fractal mapping design).

Monitoring network	Objective function value F1 (minimization)	Objective function value F2 (maximization)
NFD1	0.83	2.15
NFD2	0.89	2.57
NFD3	0.91	3.07
NFD4	1.019	3.38
NFD5	1.15	3.55
NFD6	1.29	3.71

5.3.3 A Simulation-Optimization Model for Characterisation of Unknown Sources Using Pareto-Optimal and Arbitrary Monitoring Networks

A linked simulation-optimisation model was used to characterise unknown contaminant sources using the objective function formulation in Equation 5.10 and the constraint set of Equations 5.11 and 5.12. Concentrations estimated by the simulation-optimization model were used to evaluate the performance of the Pareto-optimal monitoring network design. Two sets of six Pareto-optimal monitoring networks were designed; one with potential well locations based on singularity mapping (NFD1–NFD6) and the other without this basis (NF1–NF6).

Based on contaminant concentrations obtained from the two Pareto-optimal monitoring networks, the contaminant concentrations at different well locations were used to estimate source concentrations. These estimated source concentrations were based on concentrations obtained from the monitoring network. They were used to compare the efficiency of monitoring networks 1) based on singularity mapping subject to the two objective functions discussed earlier (Section 5.3), 2) designed without singularity mapping and 3) arbitrarily designed.

For evaluation purposes, a calibrated groundwater flow and transport simulation model was used to simulate the aquifer's responses, in terms of spatial and temporal variations in contaminant concentrations, with relevant initial and boundary conditions specified. The calibrated simulation model estimates contaminant concentrations at monitoring locations and the linked simulation-optimization model minimizes the error by computing the difference between observed and simulated values. Then, the concentration data observed within a specific monitoring-management period is used to evaluate the objective function of the network design optimization model. In real situations where a monitoring network is used to monitor contaminant concentrations, identifying the contamination sources using the proposed methodology requires field-measured concentration data.

In this study, to evaluate the methodology developed for real contamination situation, in terms of identifying contaminant sources and improving source identification, erroneous concentration measurements are essential for testing purposes. Concentration data obtained from the simulation model were perturbed to represent the effects of random measurement error and uncertainty. Hence, the observed contaminant concentrations were incorporated with randomized

measurement errors and used as synthetic concentration data to obtain the source characteristics. The perturbed concentrations were obtained from the simulated concentrations by using Equation 5.13. A maximum value of 0.1 (10%) was utilized for variable *a* in Equation 5.13.

$$pertC^{k}_{sim_{iob}} = C^{k}_{sim_{iob}} + \varepsilon \bullet a \bullet C^{k}_{sim_{iob}} \tag{5.13}$$

Where:

$pertC^{k}_{sim_{iob}}$ = perturbed simulated concentration;

$C^{k}_{sim_{iob}}$ = simulated concentration;

ε = error matrix with normally-distributed error terms with mean = 0 and standard deviation = 1; and

a = a fraction such that $0 \leq a \leq 1$.

To demonstrate the improvement in source identification efficiency after applying fractal singularity mapping, the method was applied to an illustrative study area. It considers an optimisation algorithm and three monitoring network scenarios. These three network scenarios are i) a fractal singularity mapping-based Pareto-optimal monitoring network, ii) a network without the influence of fractal singularity mapping, and iii) an arbitrary monitoring network. Additionally, contaminant sources identified by the three monitoring networks were compared. Then, the linked simulation-optimization model identified sources based on concentration measurements obtained from networks that were designed by fractal singularity mapping or were arbitrary. Two sets of 15 monitoring wells were used in each design, which form part of the source identification model. Source characterisations were achieved by estimating source concentrations using error-free and erroneous data (concentration data perturbed with random errors, as shown in Equation 5.13).

5.4 RESULTS AND DISCUSSION

Fractal/multifractal modelling using the singularity mapping index technique was used to determine the contamination plume boundaries. These guided the selection of potential well locations (i.e., the larger set) within the study site. Selection of well

locations according to fractal singularity index mapping increases the efficiency of a monitoring network, thereby improving the accuracy of source identification, particularly with multiple distributed sources. To evaluate the performance of the monitoring network based on fractal singularity mapping, sources were characterised by the three types of monitoring network design: fractal singularity mapping-based, without fractal singularity mapping, and arbitrary networks.

5.4.1 Pareto-optimal solutions

The two-objective optimal monitoring network design model was solved and the Pareto-optimal solutions for the two objective criteria of F1: minimization, and F2: maximization were obtained. Figures 5.5 and 5.6 show the Pareto-optimal front for the monitoring network design. Figure 5.5 shows the set of Pareto-optimal solutions for the two-objective problem. It is evident, as expected, that due to the conflicting nature of the two objectives, there is a trade-off between the two objective functions. Each non-dominated solution set represents a Pareto-optimal solution.

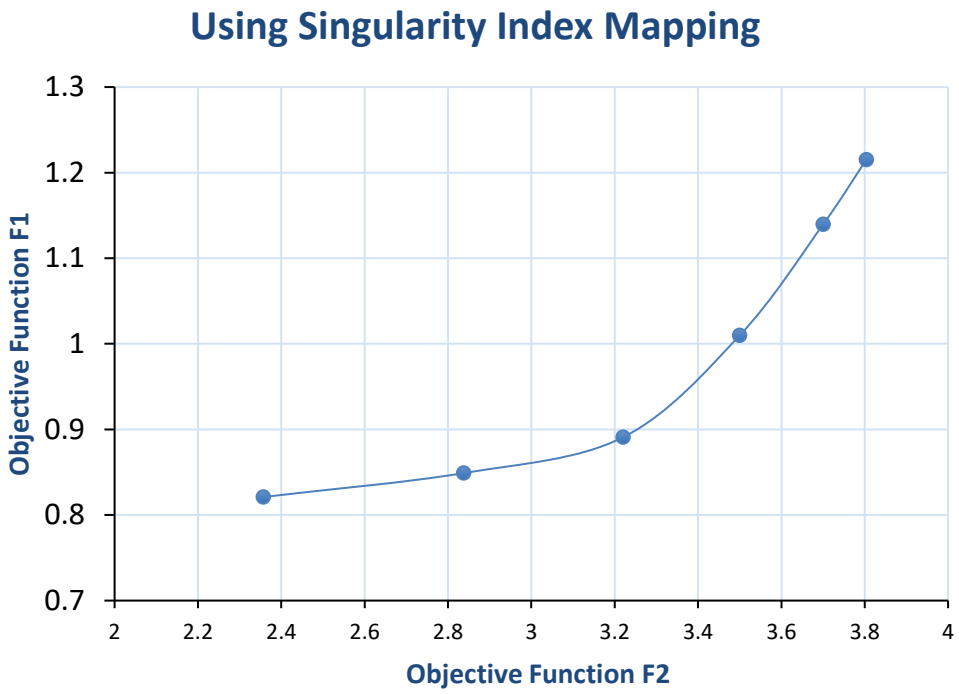


Figure 5.5: Pareto-optimal monitoring networks based on fractal singularity mapping

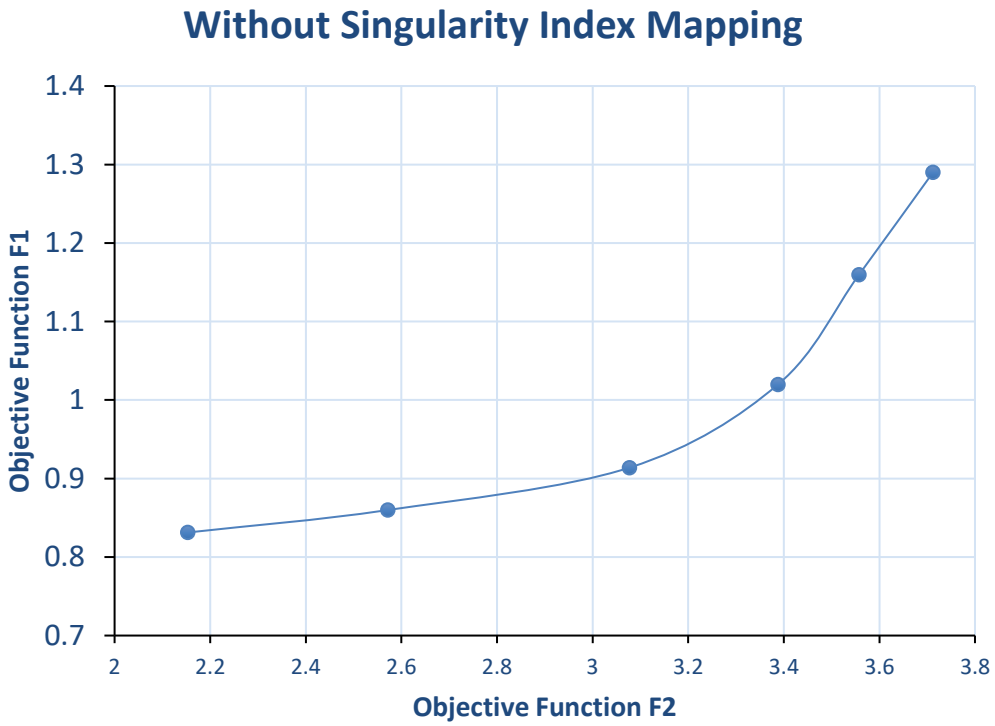


Figure 5.6: Pareto-optimal monitoring networks with no fractal singularity mapping

Two sets of six Pareto-optimal networks were chosen for different values of the objective function F2 (Tables 5.3 and 5.4). To design the monitoring networks, fractal singularity mapping was used to select potential well locations FD1 to FD6, while arbitrary potential well locations (no fractal singularity mapping) were utilized to optimally design monitoring locations NFD1–NFD6. The non-inferior solutions show the conflicting nature of the two objective functions. Objective function F1 is represented by the *y*-axis in Figure 5.5. Objective function F2 is represented by the *x*-axis in the same figure as well as in Figure 5.6. As the value of F2 (the maximization objective function) increases, the value of the other objective function F1 (the minimization objective function) increases, and vice versa.

The non-dominated set of solutions obtained by solving the two objective problem show that, in order to improve one objective function, the level achieved for the other objective function(s) becomes worse. Thus, the objective function value of one of the objective functions F1 (to be minimized) improves at the cost of the other objective function F2. F2 is to be maximized. This property shows that the two optimization models are suitable for solution using multiple objective optimization technique, as the two objective functions are conflicting in nature. There are trade-offs (Datta and Peralta, 1986) that exist between the conflicting objectives at the noninferior or nondominated (Pareto optimal) solutions. One more observation that can be made from these solution results is that the level of one of the objective function values, for a given value of the other objective function is comparatively better, when using fractal singularity mapping information in the optimization model for the monitoring network design.

It can be also inferred that the possibility of choosing an efficient monitoring network which may improve the accuracy of identifying unknown contaminant source characteristics.

5.4.2 Source Identification Solutions

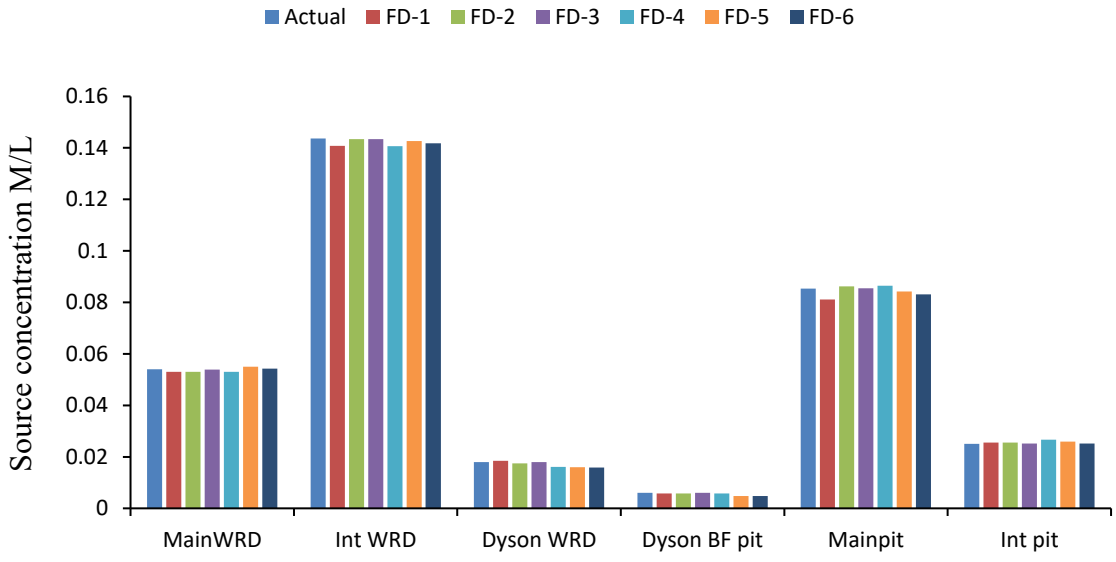
This subsection discusses the source identification results obtained using a linked optimization-based optimization model to obtain a distributed source characterization in terms of magnitudes and locations. The source characterization results for all 12 Pareto-optimal monitoring networks using fractal singularity index information (FD1 to FD6) or without it (NFD1 to NFD6) were compared. Figures 5.7 and 5.8 illustrate these comparisons using error-free and erroneous data, respectively, for the first set of six monitoring networks (FD1 to FD6) obtained using the fractal singularity mapping method. Similarly, comparisons of the results using error-free and perturbed data for the second set of six potential well monitoring networks (NFD1 to NFD6) are shown in Figures 5.9 and 5.10, respectively. The *x*-axis shows the unknown potential contaminant sources, namely, the Main WRD, Int WRD, Dyson WRD, Dyson BF Pit, Main Pit and Int Pit, previously shown in Figure 5.3 as S1, S2, S3, S4, S5and S6, respectively. The *y*-axis shows source concentrations in moles per litre (M/L). The bars in Figures 5.7 to 5.10 correspond to source concentrations, both actual and those estimated by multi-objective optimal monitoring networks.

In both Pareto-optimal networks, the source concentration identification results for all 12 Pareto-optimal monitoring network designs are very close to the actual concentrations when solved using error-free data. However, when the concentration data are perturbed with random errors to simulate measurement error, the results deviate more from the actual concentrations in all 12 Pareto-optimal monitoring networks when compared to the results based on error-free data.

The average errors for both error-free scenarios of source identification (Figures 5.7 and 5.9) by networks FD1–FD6 and NFD1–NFD6 were calculated and compared to establish the accuracy of the methodology. The average errors with error-free measurements for networks FD1, FD2, FD3, FD4, FD5 and FD6 were 0.001522129, 0.00054828, 0.000330518, 0.00144489, 0.001203354 and 0.001302412, respectively. Those for NFD1, NFD2, NFD3, NFD4, NFD5 and NFD6 were 0.002703406, 0.001938888, 0.002624391, 0.002353263, 0.003324196 and 0.001820891, respectively. These error values clearly indicate the closeness of the actual source concentration values to those estimated by the multi-objective optimal monitoring networks.

Similarly, the average errors for both erroneous scenarios of source identification (Figures 5.8 and 5.10) of networks FD1–FD6 and NFD1–NFD6 were calculated and compared for each network to establish the amount of deviation from the actual concentration values in all 12 Pareto-optimal monitoring networks when compared to the results with error-free data. The average errors for erroneous measurements for networks FD1, FD2, FD3, FD4, FD5 and FD6 were 0.001975668, 0.004062984, 0.001303269, 0.004318811, 0.002062606 and 0.002371128, respectively. Also, the average errors for error-free measurement for NFD1, NFD2, NFD3, NFD4, NFD5 and NFD6 were 0.013509675, 0.016350228, 0.009822113, 0.01144313, 0.008843035 and 0.005469973, respectively. These error values confirm the effect of using concentration data that are perturbed with random errors on the results of source identification. This shows a greater amount of deviation from the actual concentration values due to larger errors in all 12 Pareto-optimal monitoring networks when compared to the results based on error-free data. These results show that for the efficient determination of source concentrations, it is important to have the right balance between monitoring well locations and concentration measurement data.

Figures 5.7 to 5.8 shows the performance of Pareto optimal network designs with fractal singularity mapping



Source concentrations for all the six different sources

Figure 5.7: Source identification results using error-free concentration data from Pareto-optimal monitoring networks based on fractal singularity modelling

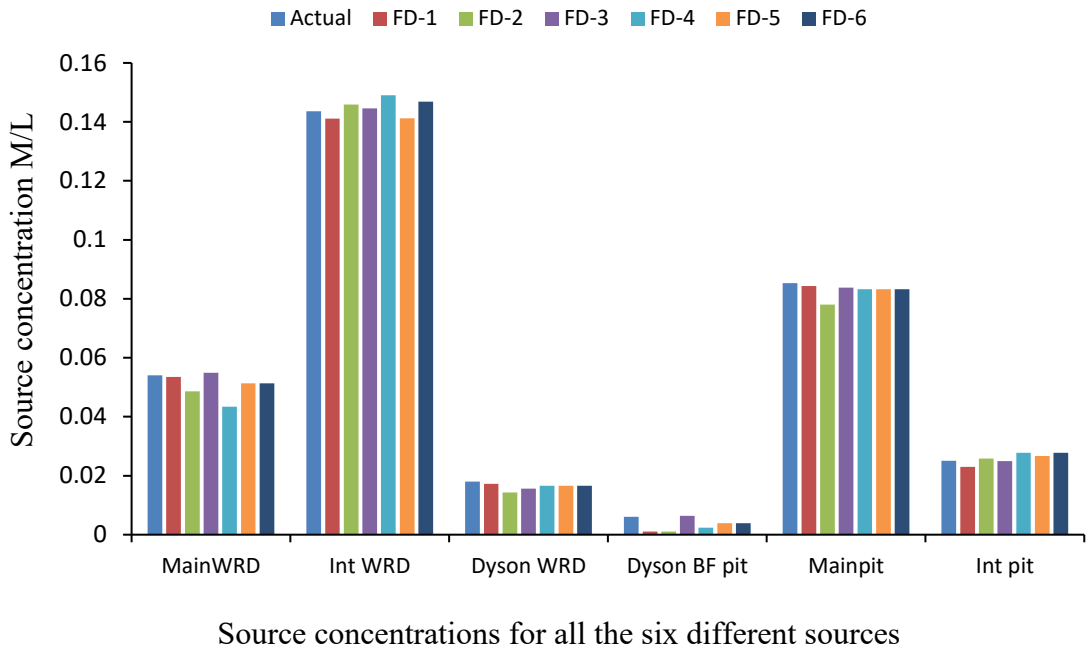


Figure 5.8: Source identification results using erroneous concentration data from Pareto-optimal monitoring networks based on fractal singularity modelling

Figures 5.9 to 5.10 shows the performance of Pareto optimal network designs with no fractal singularity mapping

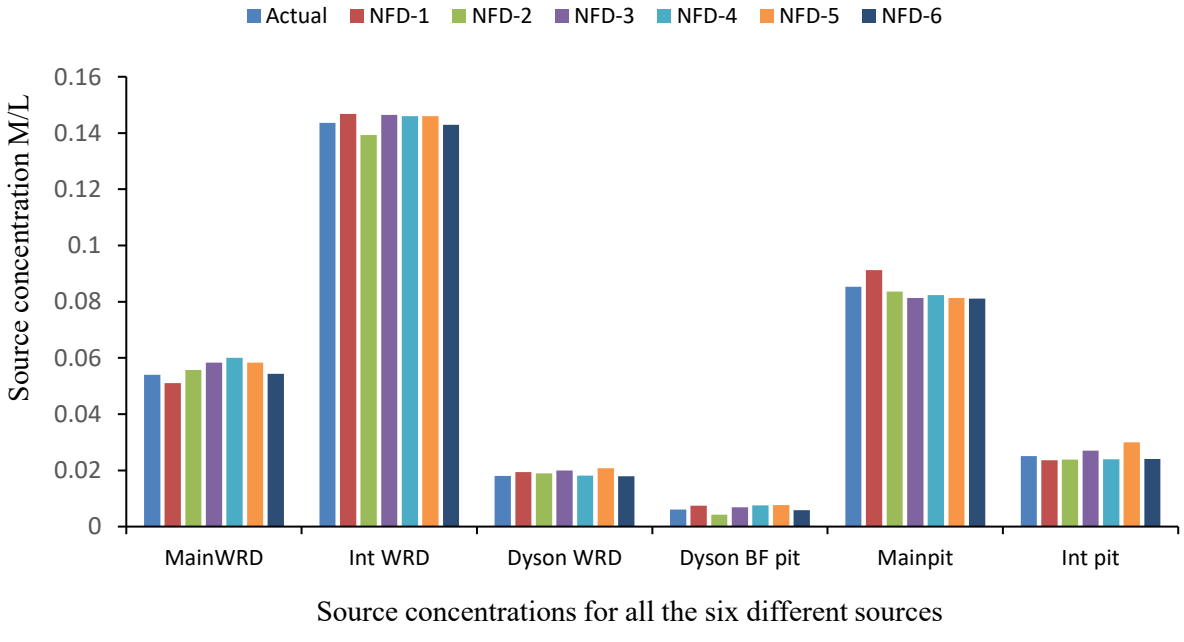


Figure 5.9: Source identification results using error-free concentration data from Pareto-optimal monitoring networks with no fractal singularity modelling

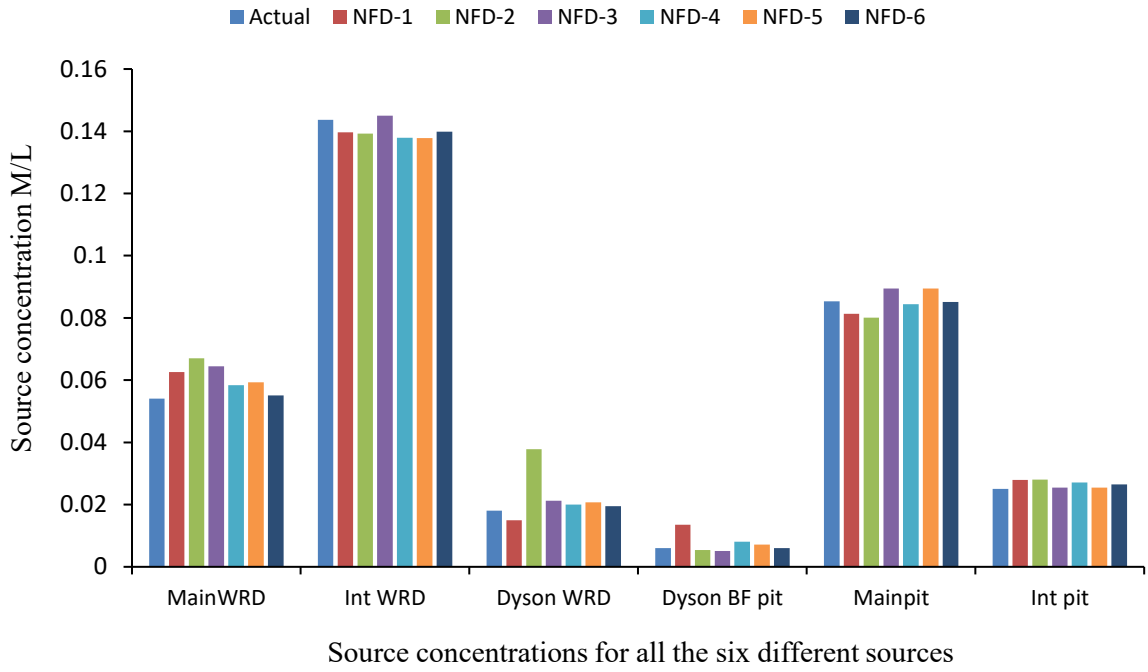


Figure 5.10: Source identification results using erroneous concentration data from Pareto-optimal monitoring networks with no fractal singularity modelling

To select the most efficient monitoring network out of the 12 Pareto-optimal monitoring networks (FD1–FD6 and NFD1– NFD6), the absolute difference between the actual and estimated source concentrations was calculated.

Figures 5.11 and 5.12 show the absolute difference between actual source concentrations and those estimated using the singularity index mapping method with the Pareto-optimal network (FD1–FD6). Also, Figures 5.13 and 5.14 show the absolute difference between actual source concentrations and those estimated for scenarios without singularity index mapping (NFD1–NFD6). These figures also show the average absolute difference of all six unknown source concentrations, which is represented by the black dashed lines. The terms Main WRD, Int WRD, Dyson WRD, Dyson BF Pit, Main pit and Int Pit refer to source locations S1, S2, S3, S4, S5 and S6, respectively, as shown in Figure 5.2. The *y*-axis represents the absolute error between the actual and estimated source concentrations, while the *x*-axis represents the six monitoring networks based on fractal singularity design (FD) and no fractal singularity

design (NFD). The absolute difference between actual and estimated source concentrations for all 12 Pareto-optimal monitoring networks (FD1–FD6 and NFD1–NFD6) using error-free and perturbed data shows similar trends.

Figures 5.11 and 5.12 show that monitoring network FD3 recorded the lowest average absolute difference between the actual and estimated source concentrations, with both error-free and erroneous data. The absolute errors suggest that monitoring network FD3 is optimal, based on it having lower errors than the other networks.

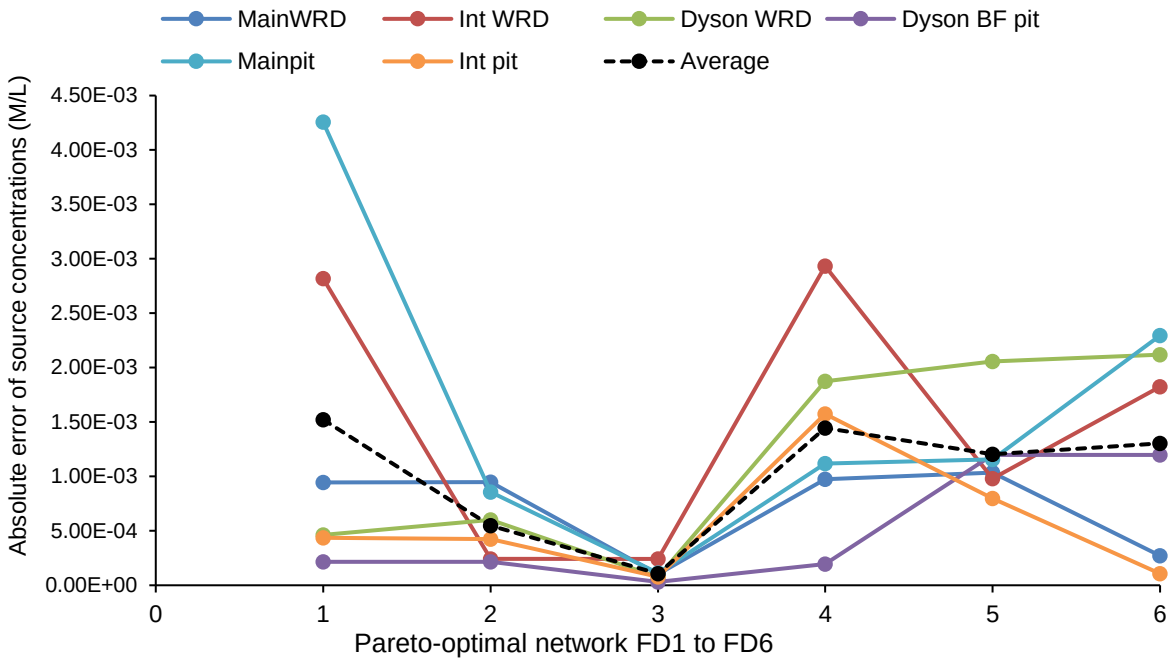


Figure 5.11: Absolute differences between actual and estimated source concentrations with error-free concentration data

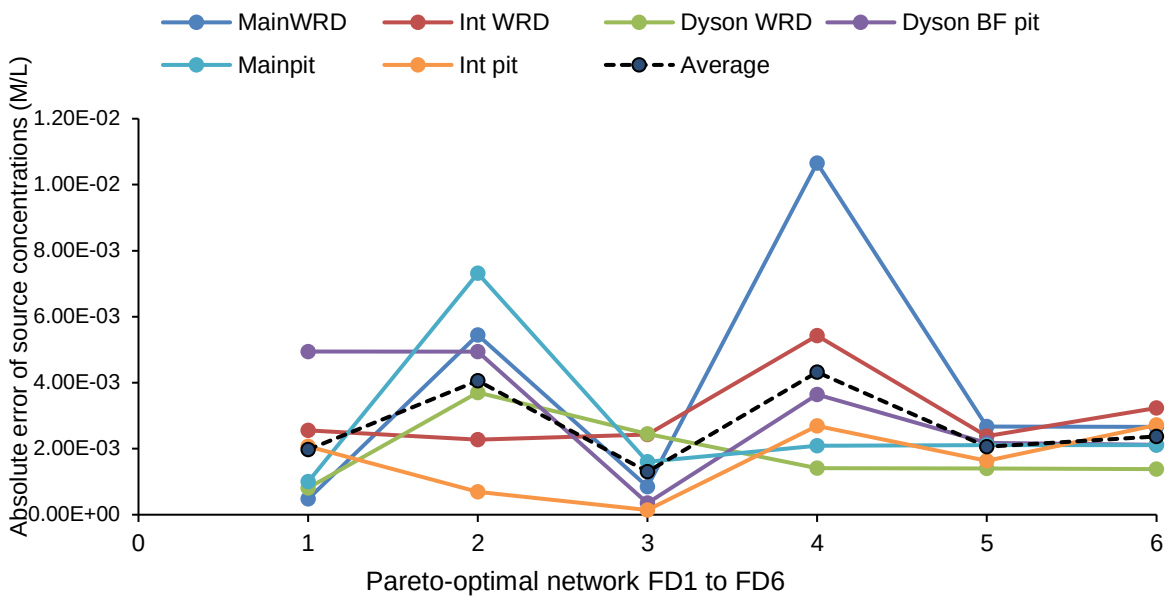


Figure 5.12: Absolute difference between actual and estimated source concentrations with erroneous concentration data

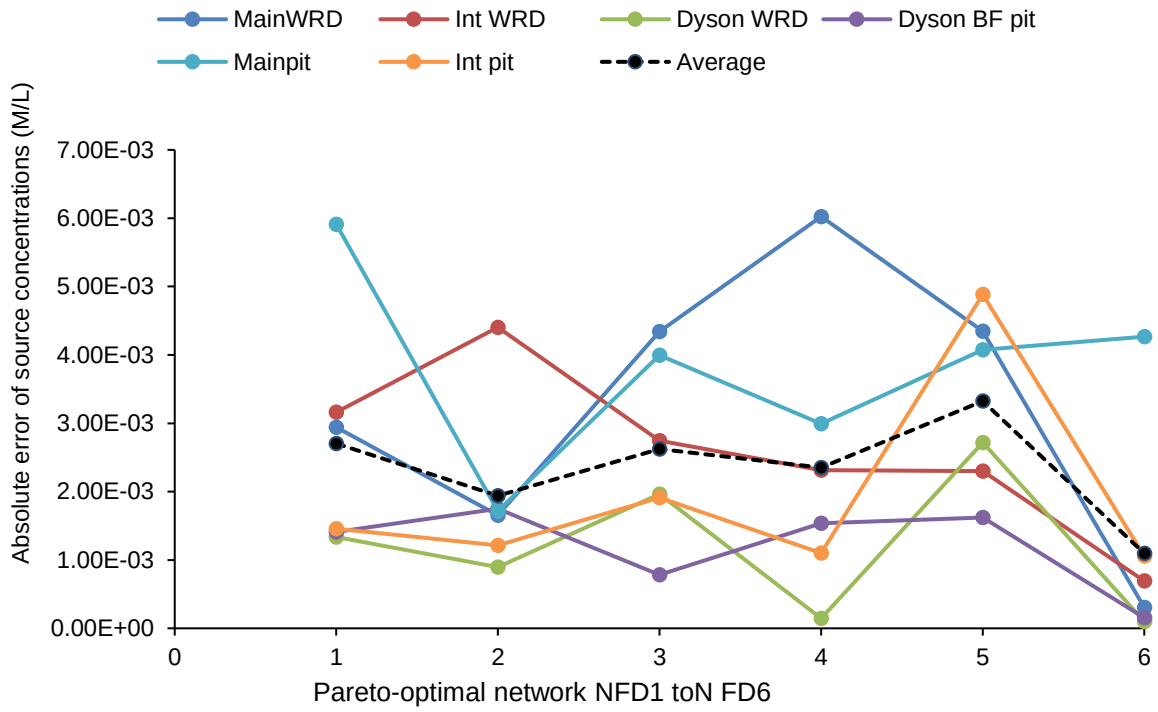


Figure5.13: Absolute difference between actual and estimated concentrations with error-free concentration data

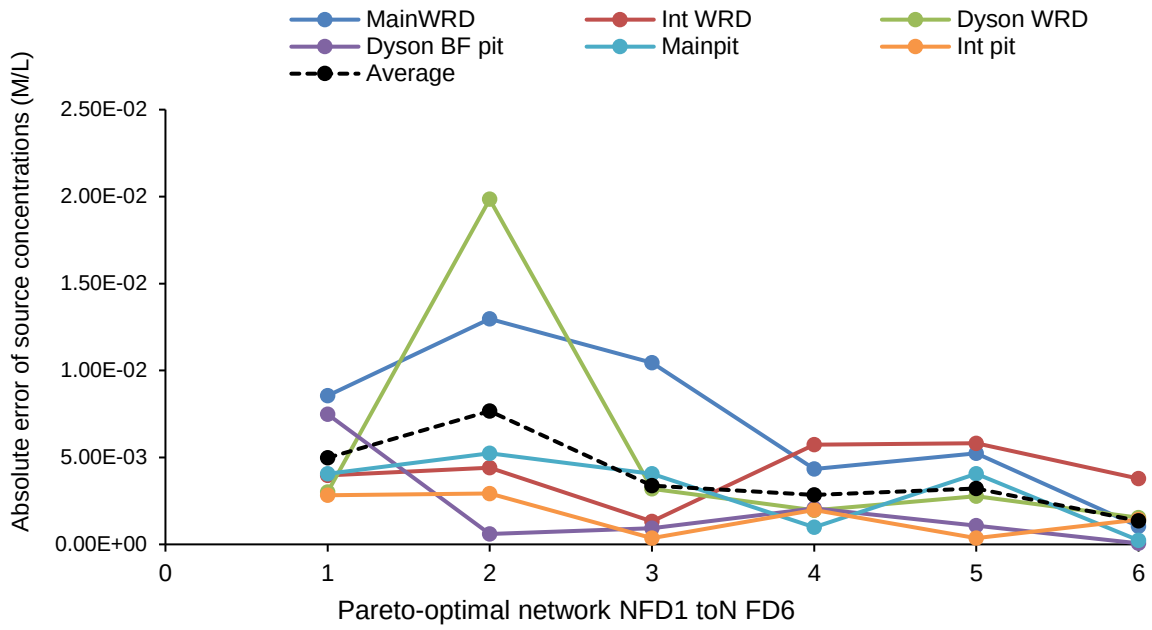


Figure 5.14: Absolute difference between actual and estimated source concentrations with erroneous concentration data

Figures 5.13 and 5.14 also show the absolute errors obtained using error-free and erroneous data for networks NFD1–NFD6. The minimum absolute average values of 0.001096 and 0.001347, respectively, were obtained for error-free and erroneous data from NFD1–NFD6. In both scenarios, network NFD 6 recorded the lowest error values, implying that the source concentration estimates using network NFD 6 closely match the actual concentrations. Hence, NFD6 is the best choice of optimal monitoring network for source identification. These results show that for effective and efficient source characterization, Pareto-optimal monitoring networks FD3 and NFD6 resulted in the minimum average absolute difference between the actual and estimated source concentrations for each Pareto-optimal monitoring network, with both error-free and erroneous data. However, more accurate source identification results and lower absolute errors with both error-free and erroneous data were achieved using the singularity mapping method. Comparing how singularity affects the Pareto-optimal network results, minimum absolute error values obtained with both fractal singularity and no fractal singularity were compared. The fractal singularity-based network FD3 showed a minimum absolute error of 0.00107 with error-free data, compared to a minimum absolute error of 0.001096 for non-fractal singularity network NFD6. Similarly, with erroneous data, FD3 showed a minimum absolute error of 0.001058,

compared to 0.001347 for network NFD6. Although both absolute error values are very low, the network based on fractal singularity shows a much lower error value, hence, there is a greater possibility that FD3 is the best monitoring network design for source identification.

To evaluate the effects of using fractal singularity mapping, the results of source characterisation using concentration data from an arbitrary monitoring network (ARBMN) and Pareto-optimal monitoring networks with (FD) and without fractal singularity design (NFD) were compared. The comparison results for the arbitrary monitoring network, and the Pareto-optimal monitoring networks FD and NFD with error-free and perturbed data are shown in Figures 5.15 and 5.16, respectively. The comparison of absolute error for arbitrary networks (ARBMN) versus networks FD-3 and NFD-6 using error-free and erroneous data are shown in Figures 5.17 and 5.18, respectively.

For all three monitoring networks (ARBMN, FD and NFD), the absolute difference between the actual and estimated concentrations obtained using the arbitrary network (ARBMN) was larger compared to using the other monitoring networks (FD and ND). It is clear from Figure 5.15 that arbitrary monitoring networks had greater errors in concentration estimation at sources Main WRD, Int WRD, Dyson WRD and Main Pit, than those obtained using formally designed monitoring networks (FD3 and NFD6). The source concentration estimates obtained using monitoring network FD3 were better at all source concentrations when using error-free data. However, arbitrary network ARBMN performed better in source identification than NFD6 when error-free data were used. ARBMN showed a much smaller difference between the actual and estimated concentrations at the Dyson BF pit and Int pit sources (with absolute error values of 0.001211 and 0.00168, respectively) compared to the Pareto-optimal network NFD6 (error values of 0.002691 and 0.002057; Figure 5.17).

Similarly, in Figure 5.16, when erroneous data were used to estimate source concentrations, the absolute difference between the actual and estimated concentrations obtained using the arbitrary network ARBMN was larger than when using the other monitoring networks (FD and ND). However, ARBMN performed better than NFD6 (in terms of a smaller difference between the actual and estimated concentrations) when estimates of concentrations at the Int pit source were compared.

For the Int pit source, the absolute error obtained by ARBMN was 0.008988, as compared to 0.020083 obtained by NFD6, as shown in Figure 5.18.

Comparison of the source concentration estimates obtained using concentration data from optimally-designed and arbitrary networks shows that it is most efficient to use an optimally-designed network based on fractal singularity mapping. Comparison of the source identification results shows that the use of a fractal singularity-designed monitoring network provides good results in terms of source concentration estimation. Use of such a monitoring network is critical in scenarios where aquifers are contaminated by multiple distributed sources and multiple species. In such situations, positioning monitoring wells arbitrarily may produce satisfactory characterisation of some sources; however, characterisation may be very poor in complex aquifers with overlapping plumes from multiple distributed sources and multiple species. This is because the measurement of the concentration of contaminants at arbitrary observation locations is not representative of all the sources present in the aquifer system. The observed measurements can only representthe effect of a subset of sources.

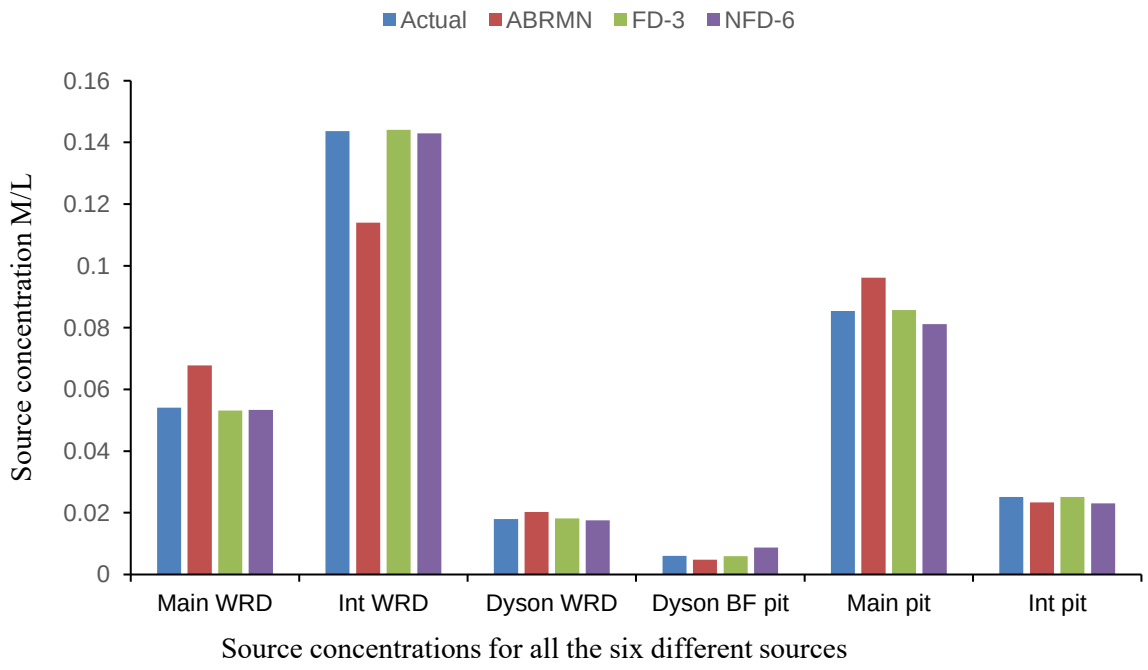


Figure 5.15: Comparison of source identification results using error-free measurement data

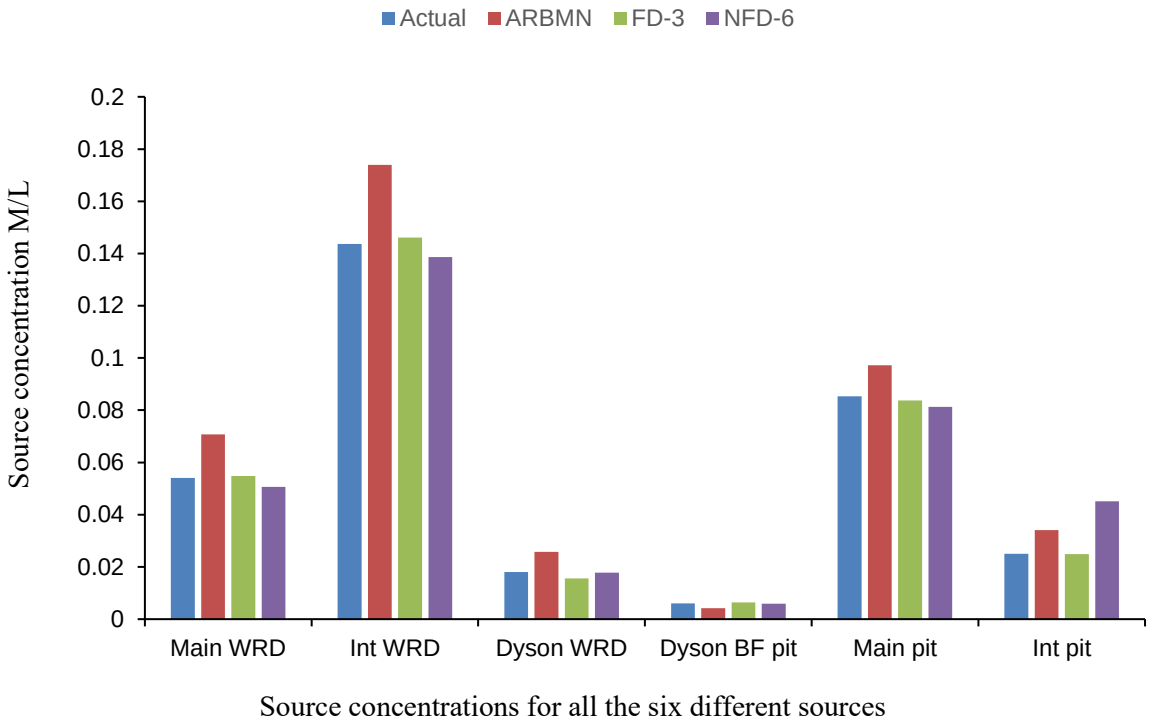


Figure 5.16: Comparison of source identification results using erroneous measurement data

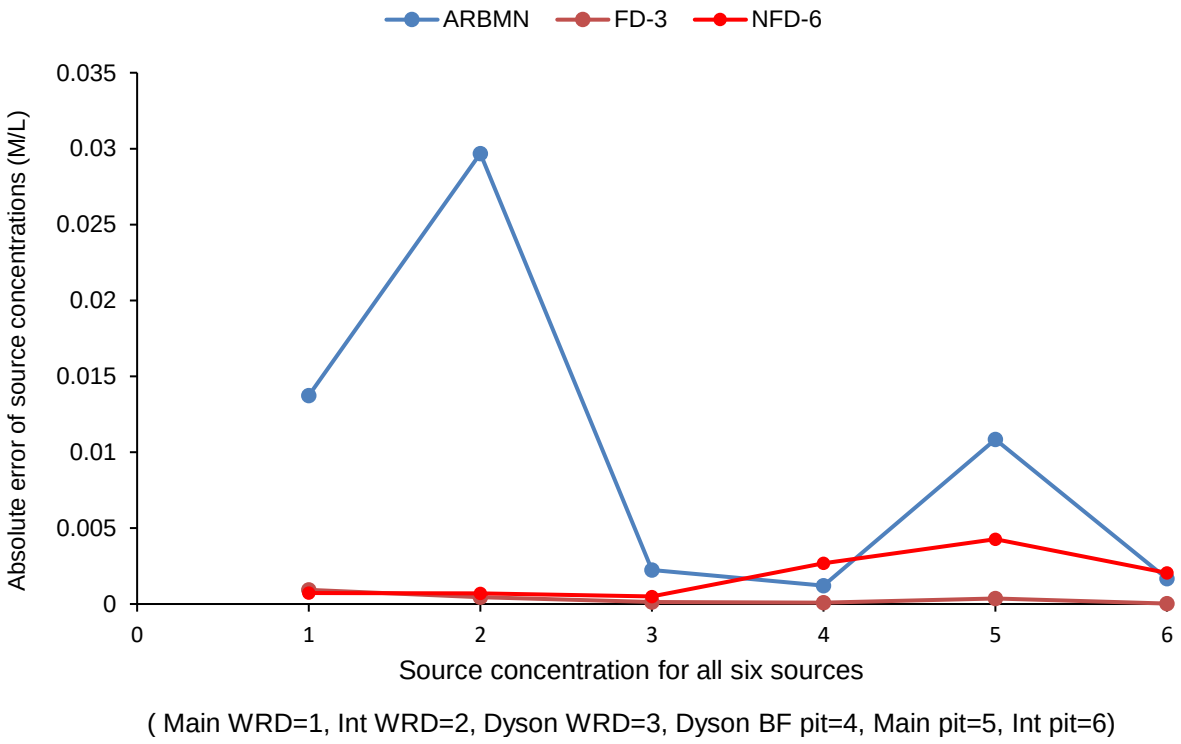


Figure 5.17: Comparison of absolute error for arbitrary networks (ARBMN) and Pareto-optimal monitoring networks FD-3 and NFD-6 using error-free measurement data

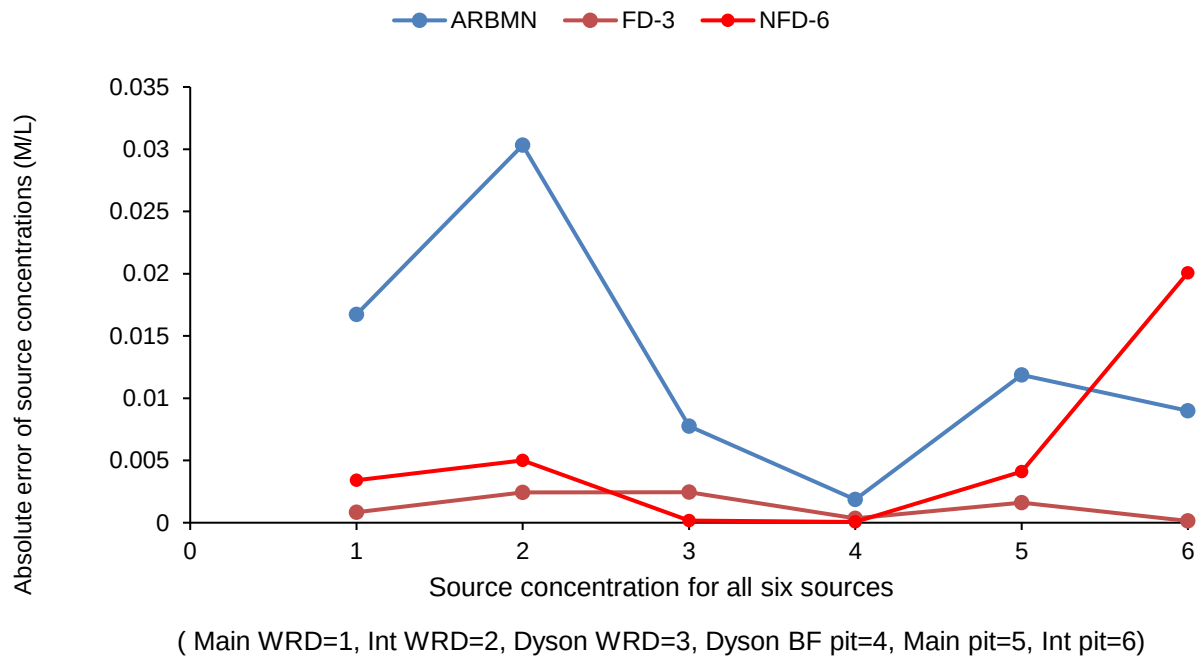


Figure 5.18: Comparison of absolute error for arbitrary networks (ARBMN) and Pareto-optimal monitoring networks FD-3 and NFD-6 using erroneous measurement data

5.5 CONCLUSIONS

This study presented a multi-objective Pareto-optimal design of a groundwater monitoring network that provides improved groundwater monitoring and source identification. The multi-objective Pareto-optimal design combines flow and reactive transport simulations, an ASA optimization algorithm and fractal singularity index mapping. The technique was demonstrated by applying it to a complex contaminated aquifer at a former uranium mine site where monitoring locations and sources were optimized using two objective function formulations in a linked simulation-optimization approach.

This work demonstrates that using multi-objective Pareto-optimal network designs incorporating fractal singularity mapping is a highly adjustable approach to monitoring network design and source identification. Additionally, this study illustrates the ability of the monitoring network design methodology to optimize the locations of observation wells by designing new monitoring networks and redesigning and/or expanding existing ones. Fractal singularity mapping was used to divide the study area into contaminated and clean zones according to singularity indices. Singularity index values of two were used to delineate contamination plume boundaries, which is useful information for monitoring decisions. The multi-objective

optimization formulations provide ways to discover multiple locations and trade-offs that may be overlooked using traditional single-objective formulations.

This chapter used a real case study to provide an illustrative example of how to use the multi-objective optimisation methodology to design optimal monitoring networks. With this example, 12 monitoring networks from a set of Pareto-optimal networks were designed. Out of the 12 Pareto-optimal networks, six sets of networks with fractal singularity designs achieved accurate source concentration estimates in cases where both error-free and erroneous data were used in the source characterization process. In all, network FD3 showed errors of < 5%, providing excellent source estimations compared with the other Pareto-optimal monitoring networks. The performance of network FD3 was better than those of the non-fractal-designed Pareto-optimal networks (NFD) and arbitrary networks (ARBMN) because fractal singularity mapping improved its design.

This methodology, once it is fully tested and modified for site-specific applications, will play a significant role in the selection of optimal numbers and locations of monitoring wells at contaminated sites. Ultimately, this will improve aquifer rehabilitation and prevent groundwater contamination.

The next chapter summarizes various aspects of the research undertaken for this thesis. The main conclusions are stated, and limitations and future research directions are discussed.

Chapter 6: Summary and Conclusions

This chapter contains summaries of the thesis chapters, conclusions and limitations, and makes recommendations for how this field of research may be extended.

6.1 SUMMARY

The purposes of this study were to 1) develop optimization-based methods to characterise unknown groundwater contamination sources in a hydrogeologically- and geochemically-complex heterogeneous anisotropic aquifer comprising transport of multiple reactive species and 2) develop monitoring network design models that improve the efficiency of source characterization.

The methodologies developed and implemented consist of numerical simulation models of three-dimensional flow and transport processes in a complex aquifer that address different reaction types that contribute to the reactive transport of multiple species. This was a novel attempt to design a model that incorporates both kinetic and equilibrium reactions to represent ongoing acid mine drainage issues at a former mine site, and to identify the contaminants affecting groundwater quality.

The developed flow and transport simulation model was linked to an optimization algorithm to obtain a management model. The management optimization model, referred to as a linked simulation-optimization model, was used to achieve optimal source characterization by solving a new optimization formulation proposed for individual species from distributed sources using the adaptive simulated annealing optimization algorithm. The optimization algorithm improved the search for an optimal solution to the problem of source concentration identification for contamination by five reactive contaminants originating from several distributed sources.

Additionally, an optimal monitoring network was designed based on a multi-objective optimisation formulation. This method involves multifractal modelling using fractal singularity indices to map the study area and define contaminated and clean areas (according to singularity indices greater or less than 2). The singularity mapping

results were used as prior information for the selection of potential monitoring locations. This approach to monitoring network design with fractal modelling showed better source characterization than monitoring networks designed without fractal modelling or with an arbitrary design. The Pareto-optimal monitoring networks proved to be reliable when spatial and temporal concentration data obtained from these networks were used in the source characterization process.

The performance of the linked simulation model in source characterization was evaluated using a real former uranium mine site in the Northern Territory, Australia. Limited field concentration measurements were used for the performance evaluation of characterization of multiple contaminant sources. In this case, the contaminants were assumed to be reactive.

In all, several methodologies were used in this thesis. The methodology used to develop and enhance the efficiency of source characterization comprised numerical simulation models, fractal/multi-fractal modelling, adaptive simulated annealing (ASA) optimization algorithms and statistical interpolation techniques.

6.2 CONCLUSIONS

Firstly, this study demonstrated the potential applicability of a numerical simulation model that simulates the transport of multiple chemically-reactive species in a complex contaminated aquifer. The case study of an abandoned mine site with uncertain input data was used. The high complexity of the area and a sparsity of available data were the main challenges in calibrating the flow and transport processes within the numerical simulation model. The limitations in implementing adequately calibrated flow and transport simulation models for such a hydrogeochemically-complex contaminated aquifer site with very limited measurement data were also highlighted.

This study demonstrates that the developed reactive transport modelling approach can contribute to a more integrated understanding of geochemical reactions and transport controls on contaminants. Hence, it is a step towards more efficient use of modelling in water resource management. In addition, the results of this reactive modelling approach provide a benchmark for the prediction of contaminant transport to assess its effects on groundwater quality. The primary conclusions in regard to the numerical simulations of groundwater processes are as follows:

1. A three-dimensional flow and reactive transport model was designed and implemented. It is primarily based on a calibrated flow model.
2. The flow and reactive transport model was successfully validated using two years of limited data, which is all that was available. However, it was not possible to validate or calibrate the transport model as the actual contaminant sources were not known.
3. The developed model increases our understanding of the reaction processes and interactions that affect groundwater quality.
4. Problems of acid mine drainage were addressed by incorporating several geochemical reactions to model the transformations of species in the subsurface.
5. The simulation results show that the proposed model can help understand the current state of mine rock waste contamination at the Rum Jungle Mine site. The model may also be applied to similar complex real-life sites with reactive contaminants.
6. A case study highlighted the utility of the approach in assessing kinetic and equilibrium reaction controls on groundwater contaminants.
7. The current work provides a basis for supplementary predictive flow and reactive transport modelling of the study site to predict the effects of acid mine drainage and to provide effective rehabilitation strategies.

Secondly, a multiple species source characterization optimization formulation was demonstrated. This optimization formulation was tested to characterise four species with distributed sources. Performance evaluations of the developed optimization model showed its potential applicability to the characterisation of groundwater contaminant sources involving reactive transport of multiple species. One of the main advantages of this source characterisation model is its ability to link any complex groundwater simulation model to an optimization model. The applicability of the source characterisation model was demonstrated using error-free measurements and those with random errors. This model showed the capability of simultaneously estimating aquifer parameters and characterizing contaminant sources at multiple

potential locations. The results show the potential for application of the methodology when real field concentration measurements are available. The results also show the applicability of this methodology to large, complex heterogeneous systems and real-life scenarios where multiple reactive species interact. The primary conclusions related to source characterization are as follows:

1. The proposed method of characterizing unknown sources of groundwater contamination can successfully identify sources irrespective of the number of sources /contaminants existing or where they are located.
2. The source concentration estimates showed errors of < 10% for individual contaminants, regardless of the magnitude of contaminant concentration at distributed source locations.
3. With a percent average estimation error (PAEE) of < 7%, the solutions for the multiple contaminant sources case were considered accurate. The PAEE values changed little and remained within an acceptable error range when perturbed datasets were used.

Thirdly, a multi-objective optimization approach to the design of optimal groundwater monitoring networks was presented. The proposed monitoring network optimization model is based on two optimization objectives within a Pareto-optimal design. These objectives are 1) minimizing the maximum normalized error between contaminant concentrations estimated by interpolating concentrations from candidate monitoring locations and actual contaminant concentration data at potential monitoring well locations, and 2) maximizing the summation of the product of estimated concentration gradients and the simulated (or measured) concentrations at selected concentration monitoring locations.

For an optimal monitoring network design to improve source characterisation accuracy, it is essential to obtain appropriate concentration measurements from relevant monitoring wells. To achieve this, multifractal modelling using the singularity index technique was applied to provide prior information for selecting potential monitoring locations that will achieve an optimal monitoring network design. The results of source characterisation are more accurate when monitoring network designs are based on prior information from fractal modelling compared to non-fractal-

influenced networks and arbitrary networks. With this multi-objective optimization methodology, the monitoring network design can be optimized and current strategies for monitoring groundwater quality can be improved to reduce redundancies and cost. The primary conclusions of the research into multi-objective optimal monitoring network design are as follows:

1. This approach can be used to design a new monitoring network with existing monitoring locations and to add additional monitoring locations to design an optimal monitoring network.
2. Source characterization results are more accurate when singularity index information is utilised in choosing potential well locations during the optimal design of the monitoring network.
3. Singularity index mapping was shown to be useful and effective in providing prior information for selecting potential monitoring network locations.
4. The outcomes of the developed methodology demonstrate that there is a trade-off involved in selecting optimal monitoring locations, in terms of minimizing the maximum normalized error between assumed and interpolated concentration values and maximizing the sum of the estimated concentration gradient and concentrations at locations. Hence, the ideal level of trade-off needs to be determined and may depend on site-specific conditions and stakeholder policies.
5. The developed methodology produces Pareto-optimal groundwater monitoring network designs. The outcomes demonstrate that there is a trade-off in selecting optimal monitoring locations in terms of two objective functions. Hence, there is a need to perform trade-off studies to determine ideal levels, which may depend on site conditions and management policies.

6.3 RECOMMENDATIONS

Calibration and validation of the flow and reactive transport model according to site conditions are essential for strengthening the reliability of source identification.

It is recommended that expert knowledge of site scenarios and active site-specific conditions should be combined with the outcomes of this study to plan appropriate monitoring strategies.

The source characterisation results were attained using a linked simulation-optimisation approach and limited data. The monitoring network designs were subject to a limited number of monitoring wells being available. The source characterisation method described in this research has the potential to achieve effective source characterisation in study areas with contamination by multiple species and limited observational concentration data. The performance evaluation results for all the developed methodologies demonstrate their possible applications in field studies, particularly with complex aquifer systems and distributed sources. While there may be some limitations associated with the methodologies discussed in this research, the following limitations can serve as guidelines for future studies. The key limitations are:

1. The availability of adequate field-measured data and consistent monitoring are decisive in calibrating a suitable model.
2. The methods developed are sensitive to uncertainties in hydrogeological parameters and random heterogeneity; hence, these must be incorporated into the models.
3. Some of the performance evaluations assume that the calibrated model represents actual field conditions as closely as possible. However, the performance evaluation will depend on the accuracy of the calibration/validation.
4. Further advancement can be made in fractal modelling to improve monitoring network design, especially in situations where network designs are based on several contaminant concentrations.
5. In multi-objective monitoring network design, further studies are required to establish guidelines for obtaining ideal trade-offs considering site-specific conditions and stakeholder preferences.

References

Afzal, P., Alghalandis, Y. F., Khakzad, A., Moarefvand, P., & Omran, N. R. (2011). Delineation of mineralization zones in porphyry Cu deposits by fractal concentration–volume modeling. *Journal of Geochemical Exploration*, *108*(3), 220–232. https://doi.org/10.1016/j.gexplo.2011.03.005

Aguiar e Oliveira Junior, H., Ingber, L., Petraglia, A., Rembold Petraglia, M., & Augusta Soares Machado, M. (2012) Adaptive Simulated Annealing. In *Stochastic global optimization and its applications with fuzzy adaptive simulated annealing.* Intelligent Systems Reference Library (vol 35). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-27479-4_4

Ahlfeld, D. P., Mulvey, J. M., Pinder, G. F., & Wood, E. F. (1988). Contaminated groundwater remediation design using simulation, optimization, and sensitivity theory: 1. Model development. *Water Resources Research*, *24*(3), 431–441. https://doi.org/10.1029/wr024i003p00431

Ahlfeld, D. P., Mulvey, J. M., & Pinder, G. F. (1988). Contaminated groundwater remediation design using simulation, optimization, and sensitivity theory: 2. Analysis of a field site. *Water Resources Research*, *24*(3), 443–452. https://doi.org/10.1029/wr024i003p00443

Alapati, S., & Kabala, Z. J. (2000). Recovering the release history of a groundwater contaminant using a non-linear least-squares method. *Hydrological Processes*, *14*(6), 1003–1016. https://doi.org/10.1002/(sici)1099-1085(20000430)14:6<1003::aid-hyp981>3.0.co;2-w

Allison, J. D., Brown, D. S., & Novo-Gradac, K. J. (1991). MINTEQA2/PRODEFA2, a geochemical assessment model for environmental systems: Version 3.0 user's manual. United States.

Alpers, C. N., & Nordstrom, D. K. (1997). Geochemical modeling of water-rock interactions in mining environments. *Pubs.Er.Usgs.Gov*, *6*(1), 289–324. https://doi.org/10.5382/Rev.06.14

Amirabdollahian, M., & Datta, B. (2013). Identification of contaminant source characteristics and monitoring network design in groundwater aquifers: An overview. *Journal of Environmental Protection*, *4*(5), 26–41. https://doi.org/10.4236/jep.2013.45a004

Amirabdollahian, M. & Datta, B. (2014). Identification of pollutant source characteristics under uncertainty in contaminated water resources systems using adaptive simulated annealing and fuzzy logic. *International Journal of Geomate*. https://doi.org/10.21660/2014.11.3258

Antoniou, M., Theodossiou, N., & Karakatsanis, D. (2017). Coupling groundwater simulation and optimization models using MODFLOW and Harmony Search Algorithm. *Desalination and Water Treatment*, *86*, 297–304. https://doi.org/10.5004/dwt.2017.20993

Aral, M. M., Guan, J., & Maslia, M. L. (2001). Identification of contaminant source location and release history in aquifers. *Journal of Hydrologic Engineering*, *6*(3), 225–234. https://doi.org/10.1061/(asce)1084-0699(2001)6:3(225)

Atmadja, J., & Bagtzoglou, A. (2001). State of the art report on mathematical methods for groundwater pollution source identification. *Environmental Forensics*, *2*(3), 205–214. https://doi.org/10.1006/enfo.2001.0055

Atmadja, J., & Bagtzoglou, A. C. (2001). Pollution source identification in heterogeneous porous media. *Water Resources Research*, *37*(8), 2113–2125. https://doi.org/10.1029/2001wr000223

Aubertin, M., Bussière, B., & Bernier, L. (2002). Environnement et gestion des rejets miniers. *Cédérom publié par Les Presses Internationales de Polytechnique* (à paraître automne 2002).

Ayaz, M. (2017). Groundwater pollution source identification using genetic algorithm based optimization model. *International Journal of Computer Sciences and Engineering*, *5*(10), 65–72. https://doi.org/10.26438/ijcse/v5i10.6572

Ayvaz, M. T. (2010). A linked simulation-optimization model for solving the unknown groundwater pollution source identification problems. *Journal of Contaminant Hydrology*, *117*(1–4), 46–59. https://doi.org/10.1016/j.jconhyd.2010.06.004

Ayvaz, M. T. (2016). A hybrid simulation-optimization approach for solving the areal groundwater pollution source identification problems. *Journal of Hydrology*, *538*, 161–176. https://doi.org/10.1016/j.jhydrol.2016.04.008

Bacon, D. H., White, M. D., & McGrail, B. P. (2000). Subsurface Transport Over Reactive Multiphases (STORM): A general, coupled, nonisothermal multiphase flow, reactive transport, and porous medium alteration simulator, Version 2 user's guide, report, March 7, 2000; Richland, Washington. (https://digital.library.unt.edu/ark:/67531/metadc704629/: accessed March 20, 2020).

Bagtzoglou, A. C., & Baun, S. A. (2005). Near real-time atmospheric contamination source identification by an optimization-based inverse method. *Inverse Problems in Science and Engineering*, *13*(3), 241–259. https://doi.org/10.1080/10682760412331330163

Bagtzoglou, A. C., Dougherty, D. E., & Tompson, A. F. B. (1992). Application of particle methods to reliable identification of groundwater pollution sources. *Water Resources Management*, *6*(1), 15–23. https://doi.org/10.1007/bf00872184 Balances, February 2011.

Banwart, S. A., & Malmström, M. E. (2001). Hydrochemical modelling for preliminary assessment of mine water pollution. *Journal of Geochemical Exploration*, *74*(1–3), 73–97. https://doi.org/10.1016/s0375-6742(01)00176-5

Bencala, K. E., and Ortiz, R. F. (1999). Theory and (or) reality: Analysis of sulfate mass-balance at Summitville, Colorado, poses process questions about the estimation of metal loadings. In D.W. Morganwalp & H.T. Buxton (Eds.), U.S. Geological Survey Toxic Substances Hydrology Program. Proceedings of the Technical Meeting, Charleston, South Carolina, March 8-12, 1999. Volume 1 of 3. Contamination from Hardrock Mining: U.S. Geological Survey Water-Resources Investigations Report 99-4018A, p. 119-122.

Blanning, R. W. (1975). The construction and implementation of metamodels. *Simulation*, *24*(6), 177–184. https://doi.org/10.1177/003754977502400606

Blowes, D., Ptacek, C., Jambor, J., & Weisener, C. (2003). The geochemistry of acid mine drainage. *Treatise on Geochemistry*, 149–204. doi: 10.1016/b0-08-043751-6/09137-4

Borah, T., & Bhattacharjya, R. K. (2016). Development of an improved pollution source identification model using numerical and ANN based simulation-optimization model. *Water Resources Management*, *30*(14), 5163–5176. https://doi.org/10.1007/s11269-016-1476-6

Brown, P. L., Ritchie, I. M., & Bennett, J. W. (2000). Geochemical kinetic modelling of acidic rock drainage. In: *Proceedings of ICARD 2000, 5th International Conference on Acid Rock Drainage*, Denver, CO, USA, May 21-24, 2000, vol. 1. pp. 289-296.

Chadalavada, S., & Datta, B. (2007). Dynamic optimal monitoring network design for transient transport of pollutants in groundwater aquifers. *Water Resources Management*, *22*(6), 651–670. https://doi.org/10.1007/s11269-007-9184-x

Chadalavada, S., Datta, B., & Naidu, R. (2010). Uncertainty based optimal monitoring network design for a chlorinated hydrocarbon contaminated site. *Environmental Monitoring and Assessment*, *173*(1–4), 929–940. https://doi.org/10.1007/s10661-010-1435-2

Chadalavada, S., Datta, B., & Naidu, R. (2011). Optimisation approach for pollution source identification in groundwater: An overview. *International Journal of Environment and Waste Management*, *8*(1/2), 40. https://doi.org/10.1504/ijewm.2011.040964

Chadalavada, S., Datta, B., & Naidu, R. (2012). Optimal identification of groundwater pollution sources using feedback monitoring information: A case study. *Environmental Forensics*, *13*(2), 140–153. https://doi.org/10.1080/15275922.2012.676147

Chen, G., Cheng, Q., Zuo, R., Liu, T., & Xi, Y. (2014). Identifying gravity anomalies caused by granitic intrusions in Nanling mineral district, China: A multifractal perspective. *Geophysical Prospecting*, *63*(1), 256–270. https://doi.org/10.1111/1365-2478.12187

Chen, Z., Cheng, Q., Chen J., & Xie, S. (2007). A novel iterative approach for mapping local singularities from geochemical data. *Nonlinear Processes in Geophysics*, *14*(3), 317–324. https://doi.org/10.5194/npg-14-317-2007

Cheng, Q. (1999). Spatial and scaling modelling for geochemical anomaly separation. *Journal of Geochemical Exploration*, *65*(3), 175–194. https://doi.org/10.1016/s0375-6742(99)00028-x

Cheng, Q. (2008). A combined power-law and exponential model for streamflow recessions. *Journal of Hydrology*, *352*(1–2), 157–167. https://doi.org/10.1016/j.jhydrol.2008.01.017

Cheng, Q. (2012). Singularity theory and methods for mapping geochemical anomalies caused by buried sources and for predicting undiscovered mineral deposits in covered areas. *Journal of Geochemical Exploration*, *122*, 55–70. https://doi.org/10.1016/j.gexplo.2012.07.007

Cheng, Q., Agterberg, F. P., & Ballantyne, S. B. (1994). The separation of geochemical anomalies from background by fractal methods. *Journal of Geochemical Exploration*, *51*(2), 109–130. https://doi.org/10.1016/0375-6742(94)90013-2

Cheng, Q., Agterberg, F. P., & Bonham-Carter, G. F. (1996). A spatial analysis method for geochemical anomaly separation. *Journal of Geochemical Exploration*, *56*(3), 183–195. https://doi.org/10.1016/s0375-6742(96)00035-0

Cheng, Q., Xu, Y. & Grunsky, E. (2000). Integrated spatial and spectrum method for geochemical anomaly separation. *Natural Resources Research 9*, 43–52. https://doi.org/10.1023/A:1010109829861

Claire Streten-Joyce, Judy Manning, Karen S. Gibb, Brett A. Neilan, & David L. Parry (2012). The chemical composition and bacteria communities in acid and metalliferous drainage from the wet–dry tropics are dependent on season, Science

of The Total Environment, Volume 443, 2013, Pages 65-79, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2012.10.024.

Clement, T.P. (1997). RT3D - A modular computer code for simulating reactive multi-species transport in 3-dimensional groundwater aquifers. *PNNL-11720 Pacific Northwest National Laboratory, Richland, Washington*

Das, A., & Datta, B. (1999). Development of management models for sustainable use of coastal aquifers. *Journal of Irrigation and Drainage Engineering*, *125*(3), 112–121. https://doi.org/10.1061/(asce)0733-9437(1999)125:3(112)

Datta, B., & Peralta, R. C. (1986). Interactive computer graphics-based multiobjective decision-making for regional groundwater management. Agricultural Water Management, 11(2), 91-116. https://doi:10.1016/0378-3774(86)90023-5

Datta, B. (2002). Discussion of "Identification of contaminant source location and release history in aquifers" by Mustafa M. Aral, Jiabao Guan, and Morris L. Maslia. *Journal of Hydrologic Engineering*, *7*(5), 399–400. https://doi.org/10.1061/(asce)1084-0699(2002)7:5(399)

Datta, B. & Chakrabarty, D. (2003). Optimal identification of unknown pollution sources using linked optimization simulation methodology. In *Proceedings of Symposium on Advances in Geotechnical Engineering (SAGE 2003)*, pp. 368–379. Kanpur, India: Indian Institute of Technology.

Datta, B., & Kourakos, G. (2015). Preface: Optimization for groundwater characterization and management. *Hydrogeology Journal*, *23*(6), 1043–1049. https://doi.org/10.1007/s10040-015-1297-3

Datta, B., Amirabdollahian, M. R., Zuo, R., & Prakash, O. (2016). Groundwater contamination plume delineation using local singularity mapping technique. *International Journal of Geomate*, *11*. https://doi.org/10.21660/2016.25.5157

Datta, B., Beegle, J. E., Kavvas, M. L. & Orlob, G. T. (1989). Development of an expert system embedding pattern recognition techniques for pollution source identification. *Completion report for U.S.G.S. Grant No. 14–08-0001-G1500.* Davis, CA: University of California.

Datta, B., Chakrabarty, D., & Dhar, A. (2008). Optimal dynamic monitoring network design and identification of unknown groundwater pollution sources. *Water Resources Management*, *23*(10), 2031–2049. https://doi.org/10.1007/s11269-008-9368-z

Datta, B., Chakrabarty, D., & Dhar, A. (2009). Simultaneous identification of unknown groundwater pollution sources and estimation of aquifer parameters. *Journal of Hydrology*, *376*(1–2), 48–57. https://doi.org/10.1016/j.jhydrol.2009.07.014

Datta, B., Chakrabarty, D., & Dhar, A. (2011). Identification of unknown groundwater pollution sources using classical optimization with linked simulation. *Journal of Hydro-Environment Research*, *5*(1), 25–36. https://doi.org/10.1016/j.jher.2010.08.004

Datta, B., Petit, C., Palliser, M., Esfahani, H. K., & Prakash, O. (2017). Linking a simulated annealing based optimization model with PHT3D simulation model for chemically reactive transport processes to optimally characterize unknown contaminant sources in a former mine site in Australia. *Journal of Water Resource and Protection*, *09*(05), 432–454. https://doi.org/10.4236/jwarp.2017.95028

Datta, B., Prakash, O., & Sreekanth, J. (2014). Application of genetic programming models incorporated in optimization models for contaminated groundwater systems management. *Advances in Intelligent Systems and Computing*, 183–199. https://doi.org/10.1007/978-3-319-07494-8_13

Datta, B., Prakash, O., Campbell, S., & Escalada, G. (2013). Efficient identification of unknown groundwater pollution sources using linked simulation-optimization incorporating monitoring location impact factor and frequency factor. *Water Resources Management*, *27*(14), 4959–4976. https://doi.org/10.1007/s11269-013-0451-8

Dhar, A., & Datta, B. (2007). Multiobjective design of dynamic monitoring networks for detection of groundwater pollution. *Journal of Water Resources Planning and Management*, *133*(4), 329–338. https://doi.org/10.1061/(asce)0733-9496(2007)133:4(329)

Dhar, A., & Datta, B. (2009). Saltwater intrusion management of coastal aquifers. I: Linked simulation-optimization. *Journal of Hydrologic Engineering*, *14*(12), 1263–1272. https://doi.org/10.1061/(asce)he.1943-5584.0000097

Dhar, A., & Datta, B. (2010). Logic-based design of groundwater monitoring network for redundancy reduction. *Journal of Water Resources Planning and Management*, *136*(1), 88–94. https://doi.org/10.1061/(asce)0733-9496(2010)136:1(88)

Dold, B. (2017). Acid rock drainage prediction: A critical review. *Journal of Geochemical Exploration*, *172*, 120–132. doi: 10.1016/j.gexplo.2016.09.014

Druhan, J., Christophe Tournassat, America, & Geochemical Society. (2019). *Reactive transport in natural and engineered systems*. Chantilly: Virginia Mineralogical Society of America.

Engesgaard, P., & Kipp, K. L. (1992). A geochemical transport model for redox-controlled movement of mineral fronts in groundwater flow systems: A case of nitrate removal by oxidation of pyrite. *Water Resources Research*, *28*(10), 2829–2843. https://doi.org/10.1029/92wr01264

Esfahani, H. K., & Datta, B. (2016). Linked optimal reactive contaminant source characterization in contaminated mine sites: Case study. *Journal of Water Resources Planning and Management*, *142*(12), 04016061. https://doi.org/10.1061/(asce)wr.1943-5452.0000707

Esfahani, H. K., & Datta, B. (2018). Fractal singularity-based multiobjective monitoring networks for reactive species contaminant source characterization. *Journal of Water Resources Planning and Management, 144*(6), 04018021. https://doi.org/10.1061/(asce)wr.1943-5452.0000880

Evangelou, V. P. (Bill), & Zhang, Y. L. (1995). A review: Pyrite oxidation mechanisms and acid mine drainage prevention. *Critical Reviews in Environmental Science and Technology, 25*(2), 141–199. https://doi.org/10.1080/10643389509388477

Fang, Y., Yeh, G.-T., & Burgos, W. D. (2003). A general paradigm to model reaction-based biogeochemical processes in batch systems. *Water Resources Research, 39*(4). https://doi.org/10.1029/2002wr001694

Fen, C.-S., Chan, C., & Cheng, H.-C. (2009). Assessing a response surface-based optimization approach for soil vapor extraction system design. *Journal of Water Resources Planning and Management*, *135*(3), 198–207. https://doi.org/10.1061/(asce)0733-9496(2009)135:3(198)

Ferguson, P. R., Wels, C., & Fawcett, M. (2012). Current water quality conditions at the historic Rum Jungle Mine Site, northern Australia. *Proceedings of 9th International Conference on Acid Rock Drainage (ICARD), May 20-26, 2012,* Ottawa, Canada.

Fethi, B. J., Loaiciga, A. H., & Marino, A.M. (1994). Multivariate geostatistical design of groundwater monitoring networks. *Journal of Water Resource Planning and Management ASCE, 120*, 505-522. http://dx.doi.org/10.1061/(ASCE)0733-9496(1994)120:4(505)

Forrester, A. I. J., & Keane, A. J. (2009). Recent advances in surrogate-based optimization. *Progress in Aerospace Sciences, 45*(1–3), 50–79. https://doi.org/10.1016/j.paerosci.2008.11.001

Giannakoglou, K. C. (2002). Design of optimal aerodynamic shapes using stochastic optimization methods and computational intelligence. *Progress in Aerospace Sciences, 38*(1), 43–76. https://doi.org/10.1016/s0376-0421(01)00019-7

Goffe, W. L. (1996). SIMANN: A global optimization algorithm using simulated annealing. *Studies in Nonlinear Dynamics & Econometrics, 1*(3). https://doi.org/10.2202/1558-3708.1020

Gorelick, S. M., Evans, B., & Remson, I. (1983). Identifying sources of groundwater pollution: An optimization approach. *Water Resources Research*, *19*(3), 779–790. https://doi.org/10.1029/wr019i003p00779

Guan, J., & Aral, M. M. (1999). Optimal remediation with well locations and pumping rates selected as continuous decision variables. *Journal of Hydrology*, *221*(1–2), 20–42. https://doi.org/10.1016/s0022-1694(99)00079-7

Gurarslan, G., & Karahan, H. (2015). Solving inverse problems of groundwater-pollution-source identification using a differential evolution algorithm. *Hydrogeology Journal*, *23*(6), 1109–1119. https://doi.org/10.1007/s10040-015-1256-z

Hazrati Y., S. (2017). Self-organizing map based surrogate models for contaminant source identification under parameter uncertainty. *International Journal of Geomate*, *13*(36). https://doi.org/10.21660/2017.36.2750

He, L., Huang, G. H., & Lu, H. W. (2009). A coupled simulation-optimization approach for groundwater remediation design under uncertainty: An application to a petroleum-contaminated site. *Environmental Pollution*, *157*(8–9), 2485–2492. https://doi.org/10.1016/j.envpol.2009.03.005

He, L., Huang, G., Zeng, G., & Lu, H. (2008). An integrated simulation, inference, and optimization method for identifying groundwater remediation strategies at petroleum-contaminated aquifers in western Canada. *Water Research*, *42*(10–11), 2629–2639. https://doi.org/10.1016/j.watres.2008.01.012

Hemker, T., Fowler, K. R., Farthing, M. W., & von Stryk, O. (2008). A mixed-integer simulation-based optimization approach with surrogate functions in water resources management. *Optimization and Engineering*, *9*(4), 341–360. https://doi.org/10.1007/s11081-008-9048-0

Huang, Y. F., Li, J. B., Huang, G. H., Chakma, A., & Qin, X. S. (2003). Integrated simulation-optimization approach for real-time dynamic modeling and process control of surfactant-enhanced remediation at petroleum-contaminated sites. *Practice Periodical of Hazardous, Toxic, and Radioactive Waste Management*, *7*(2), 95–105. https://doi.org/10.1061/(asce)1090-025x(2003)7:2(95)

Hudak, P. F., & Loaiciga, H. A. (1993). An optimization method for monitoring network design in multilayered groundwater flow systems. *Water Resources Research*, *29*(8), 2835–2845. https://doi.org/10.1029/93wr01042

Ingber, L. (1989). Very fast simulated re-annealing. *Mathematical and Computer Modelling*, *12*(8), 967–973. https://doi.org/10.1016/0895-7177(89)90202-1

Ingber, L. (1993). Simulated annealing: Practice versus theory. *Mathematical and Computer Modelling*, *18*(11), 29–57. doi: 10.1016/0895-7177(93)90204-c

Ingber, L., & Rosen, B. (1992). Genetic algorithms and very fast simulated reannealing: A comparison. *Mathematical and Computer Modelling*, *16*(11), 87–100. https://doi.org/10.1016/0895-7177(92)90108-w

Jambor, J. L., Blowes, D. W., & Ritchie, A. I. M. (2003). *Environmental aspects of mine wastes*. Mineralogical Association of Canada.

Jha, M. K., & Datta, B. (2011). Simulated annealing based simulation-optimization approach for identification of unknown contaminant sources in groundwater aquifers. *Desalination and Water Treatment*, *32*(1–3), 79–85. https://doi.org/10.5004/dwt.2011.2681

Jha, M., & Datta, B. (2013). Three-dimensional groundwater contamination source identification using adaptive simulated annealing. *Journal of Hydrologic Engineering*, *18*(3), 307–317. https://doi.org/10.1061/(asce)he.1943-5584.0000624

Jha, M., & Datta, B. (2015). Application of dedicated monitoring-network design for unknown pollutant-source identification based on dynamic time warping. *Journal of Water Resources Planning and Management*, *141*(11), 04015022. https://doi.org/10.1061/(asce)wr.1943-5452.0000513

Jha, M., & Datta, B. (2015). Application of unknown groundwater pollution source release history estimation methodology to distributed sources incorporating surface-groundwater interactions. *Environmental Forensics*, *16*(2), 143–162. https://doi.org/10.1080/15275922.2015.1023385

Jin, Y. (2003). A comprehensive survey of fitness approximation in evolutionary computation. *Soft Computing*, *9*(1), 3–12. https://doi.org/10.1007/s00500-003-0328-5

.

Karlsson, T., Räisänen, M. L., Lehtonen, M., & Alakangas, L. (2018). Comparison of static and mineralogical ARD prediction methods in the Nordic environment. *Environmental Monitoring and Assessment*, *190*(12). https://doi.org/10.1007/s10661-018-7096-2

Ketabchi, H., & Ataie-Ashtiani, B. (2015). Review: Coastal groundwater optimization—advances, challenges, and practical solutions. *Hydrogeology Journal*, *23*(6), 1129–1154. https://doi.org/10.1007/s10040-015-1254-1

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, *220*(4598), 671–680. https://doi.org/10.1126/science.220.4598.671

Kollat, J. B., Reed, P. M., & Maxwell, R. M. (2011). Many-objective groundwater monitoring network design using bias-aware ensemble Kalman filtering, evolutionary optimization, and visual analytics. *Water Resources Research*, *47*(2). https://doi.org/10.1029/2010wr009194

Kourakos, G., & Mantoglou, A. (2009). Pumping optimization of coastal aquifers based on evolutionary algorithms and surrogate modular neural network models. *Advances in Water Resources*, *32*(4), 507–521. https://doi.org/10.1016/j.advwatres.2009.01.001

Kraatz, M. (2004). Rum Jungle rehabilitation site scoping study, environmental issues and considerations for future management. *Report to the Department of Industry, Tourism and Resources.* M4K Environmental Consulting, February 2004.

Leichombam, S., & Bhattacharjya, R. K. (2016). Identification of unknown groundwater pollution sources and determination of optimal well locations using ANN-GA based simulation-optimization model. *Journal of Water Resource and Protection*, *08*(03), 411–424. https://doi.org/10.4236/jwarp.2016.83034

Lensing, H. J., Vogt, M., & Herrling, B. (1994). Modeling of biologically mediated redox processes in the subsurface. *Journal of Hydrology*, *159*(1–4), 125–143. https://doi.org/10.1016/0022-1694(94)90252-6

Lichtner, P. C. (1996). Chapter 1. Continuum formulation of multicomponent-multiphase reactive transport. *Reactive Transport in Porous Media*, 1–82. doi: 10.1515/9781501509797-004

Liu, C., & Ball, W. P. (1999). Application of inverse methods to contaminant source identification from aquitard diffusion profiles at Dover AFB, Delaware. *Water Resources Research*, *35*(7), 1975–1985. https://doi.org/10.1029/1999wr900092

Loaiciga, H. A. (1989). An optimization approach for groundwater quality monitoring network design. *Water Resources Research*, *25*(8), 1771–1782. https://doi.org/10.1029/wr025i008p01771

Loaiciga, H. A., Charbeneau, R. J., Everett, L. G., Fogg, G. E., Hobbs, B. F., & Rouhani, S. (1992). Review of ground-water quality monitoring network design. *Journal of Hydraulic Engineering*, *118*(1), 11–37. doi: 10.1061/(asce)0733-9429(1992)118:1(11)

Luo, Q., Wu, J., Yang, Y., Qian, J., & Wu, J. (2016). Multi-objective optimization of long-term groundwater monitoring network design using a probabilistic Pareto genetic algorithm under uncertainty. *Journal of Hydrology*, *534*, 352–363. https://doi.org/10.1016/j.jhydrol.2016.01.009

Mahar, P. S., & Datta, B. (1997). Optimal monitoring network and ground-water-pollution source identification. *Journal of Water Resources Planning and Management*, *123*(4), 199–207. https://doi.org/10.1061/(asce)0733-9496(1997)123:4(199)

Mahar, P. S., & Datta, B. (2001). Optimal identification of ground-water pollution sources and parameter estimation. *Journal of Water Resources Planning and Management*, *127*(1), 20–29. https://doi.org/10.1061/(asce)0733-9496(2001)127:1(20)

Mahar, P.S., & Datta, B. (2000). Identification of pollution sources in transient groundwater systems. *Water Resources Management* **14,** 209–227 https://doi.org/10.1023/A:1026527901213

Mahinthakumar, G. (Kumar), & Sayeed, M. (2005). Hybrid genetic algorithm—local search methods for solving groundwater source identification inverse problems. *Journal of Water Resources Planning and Management*, *131*(1), 45–57. https://doi.org/10.1061/(asce)0733-9496(2005)131:1(45)

Mandelbrot, B. B. (1989). Multifractal measures, especially for the geophysicist. *Pure and Applied Geophysics*, *131*(1–2), 5–42. https://doi.org/10.1007/bf00874478

Md. Azamathulla, H., Wu, F.-C., Ghani, A. A., Narulkar, S. M., Zakaria, N. A., & Chang, C. K. (2008). Comparison between genetic algorithm and linear programming approach for real time operation. *Journal of Hydro-Environment Research*, *2*(3), 172–181. https://doi.org/10.1016/j.jher.2008.10.001

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, *21*(6), 1087–1092. https://doi.org/10.1063/1.1699114

Meyer, P. D., & Brill, E. D. (1988). A method for locating wells in a groundwater monitoring network under conditions of uncertainty. *Water Resources Research*, *24*(8), 1277–1282. https://doi.org/10.1029/wr024i008p01277

Meyer, P. D., Valocchi, A. J., & Eheart, J. W. (1994). Monitoring network design to provide initial detection of groundwater contamination. *Water Resources Research*, *30*(9), 2647–2659. https://doi.org/10.1029/94wr00872

Michalak, A. M., & Kitanidis, P. K. (2004). Estimation of historical groundwater contaminant distribution using the adjoint state method applied to geostatistical inverse modeling. *Water Resources Research*, *40*(8). https://doi.org/10.1029/2004wr003214

Mirghani, B. Y., Tryby, M. E., Ranjithan, R. S., Karonis, N. T., & Mahinthakumar, K. G. (2010). Grid-enabled simulation-optimization framework for environmental

characterization. *Journal of Computing in Civil Engineering*, *24*(6), 488–498. https://doi.org/10.1061/(asce)cp.1943-5487.0000052

Molson, J. W., Aubertin, M., Frind, E. O., & Blowes, D. (2006). *POLYMIN, User Guide, A 2D multicomponent reactive transport model with geochemical speciation and kinetic sulphide oxidation.* Department of Civil, Geological and Mining Engineering, Ecole Polytechnique, Montreal.

Morrison, R. D. (2000). Application of forensic techniques for age dating and source identification in environmental litigation. *Environmental Forensics*, *1*(3), 131–153. https://doi.org/10.1006/enfo.2000.0015

Mudd, G. M., & Patterson, J. (2010). Continuing pollution from the Rum Jungle U–Cu project: A critical evaluation of environmental monitoring and rehabilitation. *Environmental Pollution*, *158*(5), 1252–1260. https://doi.org/10.1016/j.envpol.2010.01.017

Neupauer, R. M., & Wilson, J. L. (1999). Adjoint method for obtaining backward-in-time location and travel time probabilities of a conservative groundwater contaminant. *Water Resources Research*, *35*(11), 3389–3398. https://doi.org/10.1029/1999wr900190

Neupauer, R. M., & Wilson, J. L. (2005). Backward probability model using multiple observations of contamination to identify groundwater contamination sources at the Massachusetts Military Reservation. *Water Resources Research*, *41*(2). https://doi.org/10.1029/2003wr002974

Neupauer, R. M., Borchers, B., & Wilson, J. L. (2000). Comparison of inverse methods for reconstructing the release history of a groundwater contamination source. *Water Resources Research*, *36*(9), 2469–2475. https://doi.org/10.1029/2000wr900176

Nordstrom, D. K. (2015). Aqueous pyrite oxidation and the consequent formation of secondary iron minerals. *SSSA Special Publications Acid Sulfate Weathering*, 37–56. doi: 10.2136/sssaspecpub10.c3

Nordstrom, D. K., & Alpers, C. N., (1999). Geochemistry of acid mine waters. In G. S. Plumlee & M. J. Logsdon (Eds.), *The Environmental Geochemistry of Mineral Deposits*. Rev. Econ. Geol. Vol. 6a, Soc. Econ. Geol. Inc., Littleton, CO (pp. 133-160).

Parasuraman, K., & Elshorbagy, A. (2008). Toward improving the reliability of hydrologic prediction: Model structure uncertainty and its quantification using ensemble-based genetic programming framework. *Water Resources Research*, *44*(12). https://doi.org/10.1029/2007wr006451

Parker, G. K., & Robertson, A. (1999). *Acid drainage.* Melbourne, Australia: Australian Minerals & Energy Environment Foundation.

Parkes, R. J. (1998). Geomicrobiology: Interactions between microbes and minerals. In J. H. Banfield & K. H. Nealson (Eds.), (Mineralogical Society of America: Reviews in Mineralogy Vol 35), 1997, 448 pp. *Mineralogical Magazine*, *62*(5), 725–726. https://doi.org/10.1180/minmag.1998.062.5.01

Parkhurst. D. L., & Appelo, C. A. J. (1999). User's guide to PHREEQC (Version 2): A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. *Technical Report 99-4259*, *US Geological Survey Water-Resources Investigations Report* https://doi.org/10.3133/wri994259

Prakash, O. (2014). Optimal monitoring network design for efficient identification of unknown groundwater pollution sources. *International Journal of Geomate*. https://doi.org/10.21660/2014.11.3248

Prakash, O., & Datta, B. (2012). Sequential optimal monitoring network design and iterative spatial estimation of pollutant concentration for identification of unknown groundwater pollution source locations. *Environmental Monitoring and Assessment*, *185*(7), 5611–5626. https://doi.org/10.1007/s10661-012-2971-8

Prakash, O., & Datta, B. (2014). Characterization of groundwater pollution sources with unknown release time history. *Journal of Water Resource and Protection*, *06*(04), 337–350. https://doi.org/10.4236/jwarp.2014.64036

Prakash, O., & Datta, B. (2014). Multiobjective monitoring network design for efficient identification of unknown groundwater pollution sources incorporating genetic programming-based monitoring. *Journal of Hydrologic Engineering*, *19*(11), 04014025. https://doi.org/10.1061/(asce)he.1943-5584.0000952

Prakash, O., & Datta, B. (2015). Optimal characterization of pollutant sources in contaminated aquifers by integrating sequential-monitoring-network design and source identification: Methodology and an application in Australia. *Hydrogeology Journal*, *23*(6), 1089–1107. https://doi.org/10.1007/s10040-015-1292-8

Pruess, K. (1991). *TOUGH2: A general-purpose numerical simulator for multiphase fluid and heat flow*. https://doi.org/10.2172/5212064

Ptacek, C. J., & Blowes, D. W. (2003). Geochemistry of concentrated waters at mine waste sites. In J. L. Jambor, D. W. Blowes, & A. I. M. Ritchie (Eds.) *Environmental Aspects of Mine Wastes*, *Short Course Volume 31,* Mineralogical Association of Canada (pp. 239-252).

Qin, X. S., Huang, G. H., Chakma, A., Chen, B., & Zeng, G. M. (2007). Simulation-based process optimization for surfactant-enhanced aquifer remediation at heterogeneous DNAPL-contaminated sites. *Science of the Total Environment*, *381*(1–3), 17–37. https://doi.org/10.1016/j.scitotenv.2007.04.011

Razavi, S., Tolson, B. A., & Burn, D. H. (2012). Review of surrogate modeling in water resources. *Water Resources Research*, *48*(7). https://doi.org/10.1029/2011wr011527

Reed, P. M., & Kollat, J. B. (2013). Visual analytics clarify the scalability and effectiveness of massively parallel many-objective optimization: A groundwater monitoring design example. *Advances in Water Resources*, *56*, 1–13. https://doi.org/10.1016/j.advwatres.2013.01.011

Robertson GeoConsultants Inc RGC. (2010a). *Phase 1 report: Initial review and data gap analysis, June 2010.*

Robertson GeoConsultants Inc RGC. (2010b). *Technical specifications for 2010 Rum Jungle drilling program, September 2010.*

Robertson GeoConsultants Inc RGC. (2011a). *Phase 2 report: Detailed water quality review and preliminary contaminant load.*

Robertson GeoConsultants Inc RGC. (2011b). *Phase 3 (stage 1 report): Development of conceptual flow model for the Rum*

Robertson GeoConsultants Inc RGC (2011b), Phase 3 (Stage 1 Report) – Development of Conceptual Flow Model for the Rum

Robertson GeoConsultants Inc. (2012). *Phase 3 (stage 2 report): Groundwater flow model for the Rum Jungle Mine*. RGC Report Submitted to NT DOR, May 2012.

Robertson GeoConsultants Inc. (2012). Phase 3 (Stage 2 Report) - *Groundwater Flow Model for the Rum Jungle Mine*, RGC Report submitted to NT DoR, May 2012.

Romano, C. G., Ulrich Mayer, K., Jones, D. R., Ellerbroek, D. A., & Blowes, D. W. (2003). Effectiveness of various cover scenarios on the rate of sulfide oxidation of mine tailings. *Journal of Hydrology*, *271*(1–4), 171–187. https://doi.org/10.1016/s0022-1694(02)00348-7

Sadeghi, B., Madani, N., & Carranza, E. J. M. (2015). Combination of geostatistical simulation and fractal modeling for mineral resource classification. *Journal of Geochemical Exploration*, *149*, 59–73. https://doi.org/10.1016/j.gexplo.2014.11.007

Salomons, W. (1995). Environmental impact of metals derived from mining activities: Processes, predictions, prevention. *Journal of Geochemical Exploration*, *52*(1–2), 5–23. https://doi.org/10.1016/0375-6742(94)00039-e

Schaerlaekens, J., Mertens, J., Van Linden, J., Vermeiren, G., Carmeliet, J., & Feyen, J. (2006). A multi-objective optimization framework for surfactant-enhanced remediation of DNAPL contaminations. *Journal of Contaminant Hydrology*, *86*(3–4), 176–194. https://doi.org/10.1016/j.jconhyd.2006.03.002

Sidauruk, P., Cheng, A.-D., & Ouazar, D. (1998). Ground water contaminant source and transport parameter identification by correlation coefficient optimization. *Ground Water*, *36*(2), 208–214. https://doi.org/10.1111/j.1745-6584.1998.tb01085.x

Singh, R. M., & Datta, B. (2004). Groundwater pollution source identification and simultaneous parameter estimation using pattern matching by artificial neural network. *Environmental Forensics*, *5*(3), 143–153. https://doi.org/10.1080/15275920490495873

Singh, R. M., & Datta, B. (2006). Identification of groundwater pollution sources using GA-based linked simulation optimization model. *Journal of Hydrologic Engineering*, *11*(2), 101–109. https://doi.org/10.1061/(asce)1084-0699(2006)11:2(101)

Singh, R. M., Datta, B., & Jain, A. (2004). Identification of unknown groundwater pollution sources using artificial neural networks. *Journal of Water Resources Planning and Management*, *130*(6), 506–514. https://doi.org/10.1061/(asce)0733-9496(2004)130:6(506)

Skaggs, T. H., & Kabala, Z. J. (1994). Recovering the release history of a groundwater contaminant. *Water Resources Research*, *30*(1), 71–79. https://doi.org/10.1029/93wr02656

Snodgrass, M. F., & Kitanidis, P. K. (1997). A geostatistical approach to contaminant source identification. *Water Resources Research*, *33*(4), 537–546. https://doi.org/10.1029/96wr03753

Sreekanth, J., & Datta, B. (2010). Multi-objective management of saltwater intrusion in coastal aquifers using genetic programming and modular neural network based surrogate models. *Journal of Hydrology*, *393*(3–4), 245–256. https://doi.org/10.1016/j.jhydrol.2010.08.023

Sreekanth, J., & Datta, B. (2011). Optimal combined operation of production and barrier wells for the control of saltwater intrusion in coastal groundwater well fields. *Desalination and Water Treatment*, *32*(1–3), 72–78. https://doi.org/10.5004/dwt.2011.2680

Sreekanth, J., & Datta, B. (2015). Review: Simulation-optimization models for the management and monitoring of coastal aquifers. *Hydrogeology Journal*, *23*(6), 1155–1166. https://doi.org/10.1007/s10040-015-1272-z

Srivastava, D., & Singh, R. M. (2014). Breakthrough curves characterization and identification of an unknown pollution source in groundwater system using an artificial neural network (ANN). *Environmental Forensics, 15*(2), 175–189. https://doi.org/10.1080/15275922.2014.890142

Steefel C. (2008) Geochemical kinetics and transport. In S. Brantley, J. Kubicki, & A. White (Eds.) *Kinetics of water-rock interaction*. New York, NY: Springer https://doi.org/10.1007/978-0-387-73563-4_11

Steefel, C. I., Yabusaki, S. B., & Mayer, K. U. (2015). Reactive transport benchmarks for subsurface environmental simulation. *Computational Geosciences*, *19*(3), 439–443. https://doi.org/10.1007/s10596-015-9499-2

Steefel, C. I., & Yabusaki, S. B. (1996). *OS3D/GIMRT software for multicomponent-multidimensional reactive transport: User's manual and programmer's guide.* Pnl-11166. Richland, WA: Pacific Northwest National Laboratory. 99352.

Sun, A. Y., Painter, S. L., & Wittmeyer, G. W. (2006). A constrained robust least squares approach for contaminant release history identification. *Water Resources Research*, *42*(4). https://doi.org/10.1029/2005wr004312

Tamer Ayvaz, M. (2009). Application of Harmony Search algorithm to the solution of groundwater management models. *Advances in Water Resources*, *32*(6), 916–924. https://doi.org/10.1016/j.advwatres.2009.03.003

Therrien, R., & Sudicky, E. A. (1996). Three-dimensional analysis of variably-saturated flow and solute transport in discretely-fractured porous media. *Journal of Contaminant Hydrology*, *23*(1–2), 1–44. https://doi.org/10.1016/0169-7722(95)00088-7

Turcotte, D. L. (2002). Fractals in petrology. *Lithos*, *65*(3–4), 261–271. https://doi.org/10.1016/s0024-4937(02)00194-9

Ulrich, T., Golding, S. ., Kamber, B. ., Khin Zaw, & Taube, A. (2003). Different mineralization styles in a volcanic-hosted ore deposit: The fluid and isotopic signatures of the Mt Morgan Au-Cu deposit, Australia. *Ore Geology Reviews*, *22*(1–2), 61–90. https://doi.org/10.1016/s0169-1368(02)00109-9

Unger, C. (2020, March 4). What should we do with Australia's 50,000 abandoned mines? Retrieved from http://theconversation.com/what-should-we-do-with-australias-50-000-abandoned-mines-18197

Wagner, B. J. (1992). Simultaneous parameter estimation and contaminant source characterization for coupled groundwater flow and contaminant transport modelling. *Journal of Hydrology*, *135*(1–4), 275–303. https://doi.org/10.1016/0022-1694(92)90092-a

Wang, G., Carranza, E. J. M., Zuo, R., Hao, Y., Du, Y., Pang, Z., Sun, Y., & Qu, J. (2012). Mapping of district-scale potential targets using fractal models. *Journal of Geochemical Exploration*, *122*, 34–46. https://doi.org/10.1016/j.gexplo.2012.06.013

Wang, J., & Zuo, R. (2015). A MATLAB-based program for processing geochemical data using fractal/multifractal modeling. *Earth Science Informatics*, *8*(4), 937–947. https://doi.org/10.1007/s12145-015-0215-5

Wang, W., Zhao, J., & Cheng, Q. (2013). Fault trace-oriented singularity mapping technique to characterize anisotropic geochemical signatures in Gejiu mineral district, China. *Journal of Geochemical Exploration*, *134*, 27–37. https://doi.org/10.1016/j.gexplo.2013.07.009

Woodbury, A. D., & Ulrych, T. J. (1996). Minimum relative entropy inversion: Theory and application to recovering the release history of a groundwater contaminant. *Water Resources Research,* *32*(9), 2671–2681. https://doi.org/10.1029/95wr03818

Wunderly, M. D., Blowes, D. W., Frind, E. O., & Ptacek, C. J. (1996). Sulfide mineral oxidation and subsequent reactive transport of oxidation products in mine tailings impoundments: A numerical model. *Water Resources Research, 32*(10), 3173–3187. https://doi.org/10.1029/96wr02105

Xiao, F., Chen, J., Agterberg, F., & Wang, C. (2014). Element behavior analysis and its implications for geochemical anomaly identification: A case study for porphyry Cu-Mo deposits in Eastern Tianshan, China. *Journal of Geochemical Exploration, 145*, 1–11. https://doi.org/10.1016/j.gexplo.2014.04.008

Xiao, Y., Whitaker, F., Tianfu, Xu., & Steefel, C. (2018). *Reactive transport modeling: Applications in subsurface energy and environmental problems*. John Wiley & Sons. DOI: 10.1002/9781119060031

Xu, T., Sonnenthal, E., Spycher, N., & Pruess, K. (2006). TOUGHREACT—A simulation program for non-isothermal multiphase reactive geochemical transport in variably saturated geologic media: Applications to geothermal injectivity and CO2 geological sequestration. *Computers & Geosciences, 32*(2), 145–165. https://doi.org/10.1016/j.cageo.2005.06.014

Yan, S., & Minsker, B. (2011). Applying dynamic surrogate models in noisy genetic algorithms to optimize groundwater remediation designs. *Journal of Water Resources Planning and Management, 137*(3), 284–292. https://doi.org/10.1061/(asce)wr.1943-5452.0000106

Yang, C., Samper, J., Montenegro, L.: CORE2D V4: A code for water flow, heat and solute transport and geochemical reactions: Simulations of chemical interactions of clays and concrete. *International Workshop 'Modelling Reactive Transport in Porous Media'*, Strassbourg, France, 21–24 January 2008.

Yeh, G. T., Siegel, M. D., & Li, M.-H. (2001). Numerical modeling of coupled variably saturated fluid flow and reactive transport with fast and slow chemical reactions. *Journal of Contaminant Hydrology*, *47*(2–4), 379–390. https://doi.org/10.1016/s0169-7722(00)00164-9

Yeh, G. T., Carpenter, S. L., Hopkins, P. L., & Siegel, M. D. (1995a). *Users' manual of LEGHC: A Lagrangian–Eulerian finite element model of hydrogeochemical transport through saturated-unsaturated media, version 1.0.* SAND93–7081. Sandia National Laboratory, Albuquerque, NM.

Yeh, G. T., Carpenter, S. L., Hopkins, P. L., & Siegel, M. D. (1995b). *Users' manual of LEGHC: A Lagrangian–Eulerian finite element model of hydrogeochemical*

transport through saturated–unsaturated media, version 1.1. SAND95–1121. Sandia National Laboratory, Albuquerque, NM.

Yeh, G. T., Cheng, J. R., & Cheng, H. P. (1993). *2DFEMFAT: User's manual of a 2-dimensional finite element model of flow and transport through saturated-unsaturated media.* Course notes on modeling of flow and contaminants in the subsoil. IGMC, Delft, The Netherlands, June 28–July 2.

Yeh, G. T., Li, M. H., & Siegel, M. D. (1999). *Users' manual for LEHGC: A Lagrangian-Eulerian finite element model of coupled fluid flows and hydrogeochemical transport through variably saturated media—version 2.0.* Department of Civil and Environmental Engineering, Penn State University, University Park.

Yeh, G. T., Li, Y., Jardine, P. M., Burger, W. D., Fang, Y. L., Li, M. H., & Siegel, M. D. (2007). *HYDROGEOCHEM 4.1: A coupled model of fluid flow, thermal transport, and hydrogeochemical transport through saturated unsaturated media version 4.1*. Department of Civil and Environmental Engineering, University of Central Florida, Orlando.

Yeh, G. T., & Luxmoore, R. J. (1983). Modeling moisture and thermal transport in unsaturated porous media. *Journal of Hydrology (Amst.) 64*, 299–309. doi:10.1016/0022-1694(83)90074-4

Yeh, G. T., Salvage, K., & Choi, W. H. (1996). Reactive multispecies-multicomponent chemical transport controlled by both equilibrium and kinetic reactions. *Proc. XI Int. Conf. on Numerical Methods in Water Resources*, pp. 585–592, Cancun, Mexico, 22–26 July 1996.

Yeh, G. T., Salvage, K. M., Gwo, J. P., Zachara, J. M., & Szecsody, J. E. (1998). *HYDROBIOGEOCHEM: A coupled model of hydrologic transport and mixed biogeochemical kinetic/equilibrium reactions in saturated-unsaturated media.* Oak Ridge National Laboratory, Oak Ridge.

Yeh, G. T., Salvage, K. M. (1997). *HYDROGEOCHEM 2.0: A coupled model of hydrologic transport and mixed geochemical kinetic/equilibrium reactions in*

saturated unsaturated media. Technical Report, Department of Civil and Environmental Engineering, Penn State University, University Park.

Yeh, G. T., Siegel, M. D., Li, M. H. (2001). Numerical modeling of coupled fluid flows and reactive transport including fast and slow chemical reactions. *J. Contam. Hydrol*., *47*, 379–390. doi:10.1016/S0169-7722(00)00164-9

Yeh, G. T., Sun, J. T., Jardine, P. M., Burger, W. D., Fang, Y. L., Li, M .H., and Siegel, M. D. (2004). *HYDROGEOCHEM 4.0: A coupled model of fluid flow, thermal transport, and hydrogeochemical transport through saturated-unsaturated media: version 4.0.* ORNL/TM-2004/103. Oak Ridge National Laboratory, Oak Ridge.

Yeh, G. T., Sun, J. T., Jardine, P. M., Burger, W. D., Fang, Y. L., Li, M. H., Siegel, M. D. (2004). *HYDROGEOCHEM 5.0: A three-dimensional model of coupled fluid flow, thermal transport, and hydrogeochemical transport through variably saturated conditions—version 5.0.* Oak Ridge National Laboratory, Oak Ridge.

Yeh, G. T., Sun, J. T., Jardine, P. M., Burger, W. D., Fang, Y. L., Li, M. H., & Siegel, M. D. (2007). *HYDROGEOCHEM 5.1: A three dimensional model of coupled fluid flow, thermal transport, and hydrogeochemical transport through variably saturated conditions, version 5.1*. Dept. of Civil and Environ. Engineering, University of Central Florida, Orlando.

Yeh, G. T., & Tripathi, V. S. (1989). A critical evaluation of recent developments in hydrogeochemical transport models of reactive multichemical components. *Water Resour. Res. 25*(1), 93–108. doi:10.1029/WR025i001p00093

Yeh, G. T., & Tripathi, V. S. (1991). A model for simulating transport of reactive multispecies components: Model development and demonstration. *Water Resour. Res. 27*(12), 3075–3094. doi:10.1029/91WR02028

Yeh, G. T., & Tripathi, V. S. (1990). *HYDROGEOCHEM: A coupled model of hydrological transport and geochemical equilibrium of multi-component systems*. Oak Ridge National Laboratory, Oak Ridge.

Yeh, G.T. (1999). *Computational subsurface hydrology fluid flows*. Kluwer, New York.

Yeh, G.T. (2000). *Computational subsurface hydrology reactions, transport, and fate of chemicals and microbes*. Kluwer, New York.

Yeh, M.-S., Lin, Y.-P., & Chang, L.-C. (2006). Designing an optimal multivariate geostatistical groundwater quality monitoring network using factorial kriging and genetic algorithms. *Environmental Geology*, *50*(1), 101–121. https://doi.org/10.1007/s00254-006-0190-8

Yeh, W. W.-G. (1986). Review of parameter identification procedures in groundwater hydrology: The inverse problem. *Water Resources Research*, *22*(2), 95–108. https://doi.org/10.1029/wr022i002p00095

Zechman, E., M. Baha, G. Mahinthakumar, and S. R. Ranjithan (2005). A genetic programming based surrogate model development and its application to a groundwater source identification problem. ASCE Conf. Proc.173, 341

Zhang, F., Yeh, G.-T., Parker, J. C., Brooks, S. C., Pace, M. N., Kim, Y.-J., Jardine, P. M., & Watson, D. B. (2007). A reaction-based paradigm to model reactive chemical transport in groundwater with general kinetic and equilibrium reactions. *Journal of Contaminant Hydrology*, *92*(1–2), 10–32. https://doi.org/10.1016/j.jconhyd.2006.11.007

Zhao, Y., Lu, W., & An, Y. (2015). Surrogate model-based simulation-optimization approach for groundwater source identification problems. *Environmental Forensics*, *16*(3), 296–303. https://doi.org/10.1080/15275922.2015.1059908

Zhao, Y., Lu, W., & Xiao, C. (2016). Mixed integer optimization approach to groundwater pollution source identification problems. *Environmental Forensics*, *17*(4), 355–360. https://doi.org/10.1080/15275922.2016.1230906

Zuo, R. (2011). Decomposing of mixed pattern of arsenic using fractal model in Gangdese belt, Tibet, China. *Applied Geochemistry*, *26*, S271–S273. https://doi.org/10.1016/j.apgeochem.2011.03.122

Zuo, R., & Wang, J. (2016). Fractal/multifractal modeling of geochemical data: A review. *Journal of Geochemical Exploration*, *164*, 33–41. https://doi.org/10.1016/j.gexplo.2015.04.010