UNIVERSITY OF LA CORUÑA

CIVIL ENGINEERING SCHOOL
DEPARTMENT OF
CONSTRUCTION TECHNOLOGY

INVERSE PROBLEM OF WATER FLOW AND REACTIVE SOLUTE TRANSPORT IN VARIABLY SATURATED POROUS MEDIA

Ph.D. Dissertation by

ZHENXUE DAI

as partial fulfillment of Ph.D. requirements within the
Civil Engineering Ph.D. Program

Advisor:

JAVIER SAMPER CALVETE University of La Coruña

La Coruña. October, 2000

UNIVERSITY OF LA CORUÑA

CIVIL ENGINEERING SCHOOL
DEPARTMENT OF
CONSTRUCTION TECHNOLOGY

INVERSE PROBLEM OF WATER FLOW AND REACTIVE SOLUTE TRANSPORT IN VARIABLY SATURATED POROUS MEDIA

Ph.D. Dissertation by

ZHENXUE DAI

as partial fulfillment of Ph.D. requirements within the
Civil Engineering Ph.D. Program

Advisor:

JAVIER SAMPER CALVETE University of La Coruña

La Coruña. October, 2000

Department of Construction Technology
Civil Engineering School. University of La Coruña

INVERSE PROBLEM OF WATER FLOW AND REACTIVE SOLUTE TRANSPORT IN VARIABLY SATURATED POROUS MEDIA

Ph.D. Dissertation Submitted

by *ZHENXUE DAI*

as partial fulfillment of Ph. D requirements within the
Civil Engineering Ph.D. Program

Advisor:

JAVIER SAMPER CALVETE, University of La Coruña

La Coruña. October, 2000

Departamento de Tecnología de la Construcción
Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos
Universidad de La Coruña

EL PROBLEMA INVERSO DEL FLUJO DE AGUA Y EL TRANSPORTE DE SOLUTOS REACTIVOS EN MEDIOS POROSOS PARCIALMENTE SATURADOS

Tesis Doctoral Presentada por

ZHENXUE DAI

Para cumplir los requisitos del
Programa de Doctorado en Ingeniería Civil

Dirigida por:

JAVIER SAMPER CALVETE, Universidad de La Coruña

La Coruña - Octubre de 2000

RESUMEN

El estudio de la calidad química de las aguas subterráneas y la cuantificación de su contaminación, así como la evaluación de los efectos que producen los almacenamientos de residuos en el subsuelo requiere el uso de modelos numéricos. Estos modelos deben considerar conjuntamente el flujo de agua, el transporte de calor y de especies disueltas, junto con sus complejas interacciones con las fases sólidas y gaseosas. En esta tesis se presenta una formulación matemática y numérica para resolver el problema inverso del flujo de agua, transferencia de calor y transporte de un sistema multicomponente de solutos reactivos en medios parcialmente saturados.

Se han desarrollado tres códigos que resuelven el problema inverso basándose en tres métodos diferentes: método de la sección áurea, interpolación cuadratica y Gauss-Newton-Levenverg-Marquardt. Los dos primeros no requieren calcular el Jacobiano y son especialmente apropiados para resolver problemas con pocos parámetros. Sin embargo, para más de diez parámetros estos métodos no son eficientes y se debe de recurrir al método de Gauss-Newton-Levenverg-Marquardt que es más potente, robusto y eficiente que los otros dos. Este método requiere calcular las derivadas de las variables (niveles, concentraciones, caudales y contenido de humedad) con respecto a los parámetros del modelo. Para ello se utiliza un método de diferencias finitas (esquema hacia adelante y centrado). Para parámetros que varían en escala logarítmica el código realiza una transformación logarítmica que generalmente aumenta la velocidad de convergencia y proporciona valores no negativos. Para evaluar la precisión de los parámetros estimados, el código calcula el error de los valores estimados, así como medidas estadísticas de la bondad del ajuste. También proporciona la matriz de varianza-covarianza y la de correlación de los parámetros estimados, los autovalores y autovectores de la matriz de covarianza, así como la aproximación de los intervalos de confianza de los parámetros estimados. Todo esto se ha incorporado en la versión final del código INVERSE-CORE^{2D}, que incluye el modelo directo e inverso. El modelo directo esta basado en CORE^{2D} (Samper et al., 1999) e incluye: 1) Flujo de agua en 2D, confinado o no confinado, saturado o no saturado,

estacionario o transitorio y con condiciones de contorno generales, 2) Transporte de calor transitorio debido a procesos de conducción, dispersión y convección, 3) Transporte de solutos por procesos de advección, difusión molecular y dispersión hidrodinámica, 4) Reacciones químicas: acido-base, redox, complejación acuosa, adsorción, intercambio ionico, disolución/precipitación de minerales, disolución/exolución de gases y 5) Desintegración radiactiva.

Para resolver el problema inverso se tienen en cuenta cinco tipos de datos; niveles piezométricos, concentración de especies químicas, concentración total en fase liquida y sólida, caudales y contenido de humedad, así como información previa de los parámetros. Los datos se generalizan por medio de coeficientes de peso. INVERSE-CORE^{2D} permite la estimación de los siguientes parámetros de flujo y transporte de solutos reactivos, los cuales pueden variar para cada zona de material del modelo: 1) porosidad total, 2) conductividad hidráulica saturada (K_x y K_y), 3) coeficiente de almacenamiento, 4) los parámetros m, n, y α de las curvas de retención y las funciones de permeabilidad relativa, 5) coeficiente de difusión molecular, 6) dispersividad, 7) coeficiente de reparto, 8) porosidad accesible, que difiere de la total en el caso de considerar exclusión anionica, 9) concentraciones iniciales de las especies químicas incluyendo pH y pE, 10) concentraciones de contorno de las especies químicas, 11) coeficientes de selectividad para el intercambio cationico, 12) capacidad de cambio de los cationes (CEC), 13) superficie especifica de los minerales, 14) energía de activación de las reacciones controladas cinéticamente, 15) constantes cinéticas, y 16) los exponentes del efecto catalítico y del estado de saturación en la ecuación cinética que define la velocidad de disolución/precipitación de minerales.

Se han utilizado tres series de datos sintéticos para verificar la convergencia, unicidad y estabilidad del algoritmo inverso. Estos ejemplos muestran que la existencia de información previa de los parámetros reduce la covarianza y la correlación entre parámetros, reduciendo, por tanto, la incertidumbre en esos parámetros. La incorporación de la información previa mejora también la eficiencia del proceso de estimación. Los códigos que resuelven el problema inverso se han utilizado satisfactoriamente para interpretar diferentes tipos de experimentos de laboratorio: 1) infiltración, 2) difusión y 3) flujo "a través" en bentonitas y granitos, en los que los parámetros de flujo y transporte se

han estimado bajo diferentes condiciones. Un experimento en columna, publicado por Appelo et al. (1990), ha sido utilizado para examinar la metodología inversa en un problema real de transporte reactivo. Se estimaron los valores de la dispersividad, coeficientes de selectividad, CEC, así como de las concentraciones iniciales y de contorno de los principales componentes químicos. La solución del problema inverso fue de gran utilidad para la identificación de los procesos geoquímicos más relevantes que tienen lugar durante el experimento.

Esta metodología se aplicó también a dos casos de campo, el acuítardo del Delta del Llobregat (Barcelona) y el acuífero de Aquia (Maryland, USA). En ambos casos se produce una mezcla de agua salada con agua dulce. Los valores de la conductividad hidráulica, dispersividad, coeficientes de selectividad, CEC y concentraciones iniciales y de contorno de los principales componentes químicos se estimaron satisfactoriamente para diferentes hipótesis sobre las condiciones redox. Los resultados de la estimación indican que los procesos redox son procesos geoquímicos importantes en el acuítardo del Delta del Llobregat. Además, se ha cuantificado la incertidumbre asociada a los parámetros y se han calculado los intervalos de confianza aproximados.

En general, los resultados obtenidos con los ejemplos sintéticos y los casos reales indican que INVERSE-CORE^{2D} es una herramienta de optimización flexible y robusta para resolver el problema inverso de flujo y transporte de solutos reactivos.

ABSTRACT

Understanding natural groundwater quality patterns, quantifying groundwater pollution and assessing the effects of waste disposal require modeling tools accounting for water flow, and transport of heat and dissolved species as well as their complex interactions with solid and gaseous phases. This dissertation presents a mathematical and numerical methodology for solving the coupled inverse problem of water flow, heat transfer and multi-component reactive solute transport in variably saturated media.

Three inverse codes have been developed which are based on: (1) Golden section search, (2) quadratic interpolation, and (3) Gauss-Newton-Levenverg-Marquardt method. The first two methods do not require computing the Jacobian matrix. They are especially suited for the estimation of a few parameters. They are not efficient for problems involving more than ten parameters. In such cases, one must resort to the Gauss-Newton-Levenberg-Marquardt method which is more powerful, robust and efficient. The latter is a gradient-based optimization algorithm in which derivatives of observations (heads, solute concentrations, total concentrations, cumulative water inflow and water content) with respect to model parameters are calculated using either forward or central finite differences. For possitively-valued parameters and for those which vary in a logarithmic scale the code performs a log transformation which generally enhances the rate of convergence and ensures non-negative estimates. To evaluate the accuracy of the estimated parameters, the code computes parameter estimation errors and statistical measures of goodness-of-fit. It also computes the covariance and correlation matrices of the estimated parameters, the eigenvalues and eigenvectors of the covariance matrix, as well as approximate confidence intervals of the estimated parameters. All these features are incorporated in INVERSE-CORE^{2D}, the most updated and complete of the three inverse codes. The forward modeling part of INVERSE-CORE^{2D} is entirely based on CORE^{2D} , a code developped at the University of La Coruña by Samper et al. (1999) which accounts for: 1) 2-D confined or unconfined, saturated or unsaturated steady or transient groundwater flow with general boundary conditions, 2) Transient heat transport including conduction, heat dispersion and

advection processes, 3) Solute transport including advection, molecular diffusion and mechanical dispersion, 4) Chemical reactions including: acid-base, redox, aqueous complexation, surface adsorption, ion exchange, mineral dissolution-precipitation, gas dissolution-exsolution, and retardation, and 5) Radioactive decay.

Up to five different types of data can be taken into account for the solution of the inverse problem, including: hydraulic heads, concentrations of dissolved chemical components, total concentrations (including liquid and solid phases), water fluxes and water contents, as well as parameter prior information. These data have been generalized in a weighted least square criterion by a set of weighting coefficients. INVERSE-CORE^{2D} can estimate a wide range of flow and reactive transport parameters which may be different in different parameter zones. These parameters include: 1) total porosity, 2) components of the saturated hydraulic conductivity (K_x and K_y), 3) specific yield, 4) parameters m, n, and α which are used for defining retention curves and relative permeability functions, 5) molecular diffusion coefficient, 6) solute dispersivity, 7) distribution coefficient K_d of a sorbing chemical, 8) accessible porosity, which may differ from total porosity when anion exclusion takes place, 9) initial concentrations of chemical components, including initial pH and pE, 10) boundary concentrations of chemical components, 11) selectivity coefficients of cation exchange, 12) cation exchange capacity (CEC), 13) specific surface of minerals, 14) apparent activation energy of kinetically-controlled reactions, 15) kinetic rate constant, and 16) exponents of catalytic terms and those of the ratio of the ionic activity products and the equilibrium constants for kinetically-controlled mineral dissolution/precipitation.

Three sets of synthetic data have been used to verify the codes and study the convergence, uniqueness and stability of the inverse algorithms. These examples illustrate that parameter prior information reduces parameter correlation and uncertainty. Incorporating prior information leads also to a significant improvement of the numerical efficiency of the estimation process.

The inverse codes developped in this dissertation have been successfully used to interpret different types of laboratory experiments. Infiltration, diffusion, and permeation experiments in bentonite and granite have been used to estimate flow and transport

parameters. A column experiment reported by Appelo et al. (1990) has been used to test the inverse methodology on reactive transport data. Solute dispersivity, selectivity coefficients, CEC, and initial and boundary concentration of the main chemical components have been estimated. The solution of the inverse problem was useful in this case for identifying relevant geochemical processes taking place during the experiment. The methodology has also been applied to field case studies at the Llobregat Delta Aquitard (Barcelona, Spain) and the Aquia aquifer (Maryland, USA). Both of them deal with salt water leaching with fresh water. Hydraulic conductivities, dispersivities, selectivity coefficients, CEC, initial and boundary concentrations of the main chemical components were successfully estimated for different redox hypotheses. Estimation results indicate that redox processes are important geochemical processes in the Llobegat Delta aquitard. The uncertainty of the estimated parameters has been quantified and approximate confidence intervals were computed.

Overall, the results obtained with synthetic and real examples indicate that INVERSE-CORE^{2D} is a flexible and robust optimization tool for the solution of the inverse problem of flow and reactive solute transport.

RESUMO

O estudio da calidade química das augas subterráneas e a cuantificación da súa contaminación, así como a avaliación dos efectos que producen os almacenamentos de refugallos no subsuelo require o uso dos modelos numéricos. Estos modelos deben considerar conxuntamente o fluxo da auga, o transporte da calor e das especies disoltas, xunto coas súas complexas interaccións coas fases sólidas e gaseosas. Nesta tese de doutoramento preséntase unha formulación matemática e numérica para resolve-lo problema inverso do fluxo de auga, transferencia da calor e transporte dun sistema multicomponente dos solutos reactivos nos medios parcialmente saturados.

Desenroláronse tres códigos que resolven o problema inverso baseándose en tres métodos diferentes: método da sección áurea, interpolación cuadrática e Gauss-Newton-Levenverg-Marquardt. Os dous primeiros non requiren calcula-lo Jacobiano e son especialmente apropiados para resolve-los problemas con poucos parámetros. Sen embargo, para máis de dez parámetros estos métodos non son eficientes e débese recurrir ó método de Gauss-Newton-Levenverg-Marquardt que é máis potente, robusto e eficiente que os outros dous. Este método require calcula-las derivadas das variables (niveis, concentracións, caudais e contido de humidade) con respecto ós parámetros do modelo. Para isto utilízase un método de diferencias finitas (esquema hacia diante e centrado). Para parámetros que varían na escala logarítmica o código realiza unha transformación logarítmica que xeralmente aumenta a velocidade da convergencia e proporciona valores non negativos. Para evalua-la precisión dos parámetros estimados, o código calcula o erro dos valores estimados, así como as medidas estadísticas da bondade do axuste. Tamén proporciona a matriz da varianza-covarianza e a da correlación dos parámetros estimados, os autovalores e autovectores da matriz de covarianza, así como a aproximación dos intervalos de confianza dos parámetros estimados. Todo isto incorporouse na versión final do código INVERSE-CORE^{2D}, que inclúe o modelo directo e inverso. O modelo directo básase no CORE^{2D} (Samper et al., 1999) e inclúe: 1) Fluxo de auga en 2D, confinado ou non confinado, saturado ou non saturado, estacionario ou transitorio e con condicións de

contorno xerais, 2) Transporte da calor transitorio debido a procesos de conducción, dispersión e convección, 3) Transporte de solutos por procesos de advección, difusión molecular e dispersión hidrodinámica, 4) Reaccións químicas: ácido-base, redox, complexación acuosa, adsorción, intercambio iónico, disolución/precipitación de minerais, disolución/exolución de gases e 5) Desintegración radiactiva.

Para resolve-lo problema inverso téñense en conta cinco tipos de datos; niveis piezométricos, concentracións das especies químicas, concentración total na fase líquida e sólida, caudais e contido da humidade, así como a información previa dos parámetros. Os datos xeneralízanse por medio dos coeficientes de peso. INVERSE-CORE^{2D} permite a estimación dos seguintes parámetros de fluxo e transporte dos solutos reactivos, os cales poden variar para cada zona de material do modelo: 1) porosidade total, 2) conductividade hidráulica saturada (K_x e K_y), 3) coeficiente de almacenamento, 4) os parámetros m, n, e α das curvas de retención e as funcións da permeabilidade relativa, 5) coeficiente da difusión molecular, 6) dispersividade, 7) coeficiente do reparto, 8) porosidade accesible, que difiere da total no caso de considera-la exclusión aniónica, 9) concentracións iniciais das especies químicas incluindo o pH e pE, 10) concentracións de contorno das especies químicas, 11) coeficientes de selectividade para o intercambio catiónico, 12) capacidade de cambio dos catións (CEC), 13) superficie especifica dos minerais, 14) enerxía de activación das reaccións controladas cinéticamente, 15) constantes cinéticas, e 16) os exponentes do efecto catalítico e do estado de saturación na ecuación cinética que define a velocidade de disolución/precipitación de minerais.

Utilizaronse tres series de datos sintéticos para verifica-la converxencia, unicidade e estabilidade do algoritmo inverso. Estos exemplos amosan que a existencia da información previa dos parámetros reduce-la covarianza e a correlación entre parámetros, reducindo, polo tanto, a incertidume nesos parámetros. A incorporación da información previa mellora tamén a eficiencia do proceso de estimación. Os códigos que resolven o problema inverso utilizáronse satisfactoriamente para interpreta-los diferentes tipos de experimentos de laboratorio: 1) infiltración, 2) difusión e 3) fluxo "a través" en bentonitas e granitos, nos que os parámetros de fluxo e transporte estimáronse baixo diferentes condicións. Un experimento en columna, publicado por Appelo et al. (1990), utilizouse para examina-la

metodoloxía inversa nun problema real de transporte reactivo. Estimaronse os valores da dispersividade, coeficientes de selectividade, CEC, así como das concentracións iniciais e de contorno das principais compoñentes químicas. A solución do problema inverso foi de gran utilidade para a identificación dos procesos xeoquímicos máis relevantes que teñen lugar durante o experimento.

Esta metodoloxía aplicouse tamén a dous casos de campo, o acuitardo do Delta do Llobregat (Barcelona) e o acuífero de Aquia (Maryland, USA). En ámbolos dous casos prodúcese unha mixtura da auga salada coa auga doce. Os valores da conductividade hidráulica, dispersividade, coeficientes de selectividade, CEC e concentracións iniciais e de contorno das principais compoñentes químicas estimáronse satisfactoriamente para diferentes hipóteses sobre as condicións redox. Os resultados da estimación indican que os procesos redox son procesos xeoquímicos importantes no acuítardo do Delta do Llobregat. Además, cuantificouse a incertidume asociada ós parámetros e calculáronse os intervalos de confianza aproximados.

En xeral, os resultados obtidos cos exemplos sintéticos e os casos reais indican que INVERSE-CORE^{2D} é unha ferramenta de optimización flexible e robusta para resolve-lo problema inverso do fluxo e transporte dos solutos reactivos.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to Dr. Javier Samper for his effective supervision and constant support in the course of the development of this dissertation, and also for his dedication and patience in solving the complex administrative matters of a foreign graduate student. I am very grateful to all kinds of support provided by the Escuela de Ingenieros de Caminos, Canales y Puertos of La Coruña University. My appreciation is extended to Dr. Jordi Delgado and Dr. Luis Montenegro, for their help, constructive discussions, and review of geochemical aspects of the inverse problem. Thanks are given to Dr. Jesús Carrera, Dr. Ne-Zheng Sun, Dr. Tianfu Xu, Dr. Yunwei Sun, Dr. John Doherty, Dr. Lehua Pan, and Dr. C. A. J. Appelo for their helpful suggestions to the development of our inverse codes.

I thank all my colleagues, in particular to Dr. Jorge Molinero, Dr. Ricardo Juncosa, Dr. Francisco Padilla, Josefa Pilar, Llorenç Huguet, Ana Vázquez, Ángel Ruiz Picó, Guoxiang Zhang, Gemma Soriano, Nuria Cuéllar, Javier Gómez and Gonzalo Mosqueira, for their help. I am grateful to Dr. Marisol Manzano and Dr. Tianfu Xu for providing basic data as well as interesting ideas and prior information of the related parameters for the Llobregat aquitard problem, to Miguel García and María Victoria del Villar of CIEMAT for providing the data of the infiltration, diffusion and permeation expriments of the FEBEX Project, to Dr. C. A. J. Appelo for providing the column experiment data as well as hints and suggestions, and to Francis H. Chapelle for providing the Aquia aquifer data.

The financial support of this dissertation was mostly provided by the Spanish Nuclear Waste Company (ENRESA) within the framework of the FEBEX Research Project and a Project for the development of Inverse Methodologies for reactive transport. Both projects were funded by a Research Grant signed with the University of La Coruña and the Civil

Engineering Foundation of Galicia (Contracts # 703231 and 703336). The FEBEX Project as a whole was funded by the Commission of the European Community (Project FI4W-CT95-0006) within the Nuclear Fission Safety Programme. Funding for this dissertation was also provided by a CICYT Reseach Project of the Spanish Ministry of Education (Project HID98-0282). My research work during 2000 was funded by a research scholarship awarded by the University of La Coruña.

I should thank my wife, Liying, for her encouragement and understanding which made it possible for me to concentrate on my research work, my parents, my parents-in-law and my son. This dissertation is dedicated to them.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1. MOTIVATION

The assessment of the performance of waste disposal facilities requires a quantitative analysis of the migration of toxic substances away from the repository. This in turn requires the use of modeling tools which are able to analyze both the transport of dissolved species as well as their complex interactions with the solid phases of the engineering and geological barriers. Computer simulations based on numerical models have been increasingly used for these purposes, a trend that undoubtedly will continue as more sophisticated models are being developed and computer costs keep decreasing. Significant efforts and attempts have been made during recent years towards the development of such tools.

However, with greater forward model sophistication the need comes for more accurate data requirements. The real improvement in precision will eventually hinge on our ability to determine accurately the required model parameters, such as saturated permeability, porosity, relative permeability, diffusion and dispersion coefficients, initial and boundary concentrations, initial pH and pE, selectivity coefficients for cation exchange, cation exchange capacity (CEC), the kinetic parameters such as specific surface of mineral, apparent activation energy, kinetic rate constant, the exponent of the catalytic term and the exponent of ratio of the ionic activity product and the equilibrium constants. Unfortunately, over the past 20 years, it appears that the ability to fully characterize the parameter systems of the reactive transport models has not kept pace with the numerical and modeling expertise. Difficulties in model calibration are nowhere more evident than in the analysis of flow and reactive solute transport in the unsaturated zone. There is a clear need for the development of general purpose and numerically efficient codes that can estimate

simultaneously some or all the relevant hydrodynamic, solute transport and hydrochemical parameters by robust optimization techniques.

1.2. STATE-OF-THE-ART

Building a model for a groundwater flow and reactive solute transport system requires solving two problems, the forward problem (simulation) and its inverse (calibration) (Sun, 1994). The former predicts unknown system states by solving appropriate governing equations, while the latter determines unknown physical and chemical parameters and initial and boundary conditions of the system. One must first solve the inverse problem to find appropriate model structure and model parameters, and then solve the forward problem to obtain prediction results.

1.2.1. FORWARD MODELING OF REACTIVE SOLUTE TRANSPORT

In the past decade, coupled models accounting for complex hydrological and chemical processes, with varying degrees of sophistication, have been developed. Samper et al. (1995) presented a state-of-the-art on hydrogeochemical, water flow and reactive solute transport models. Existing models of reactive transport employ two basic sets of equations: the transport equation and chemical equation. Two major numerical approaches have been proposed to solve the coupled problem (Yeh and Tripathi, 1989; Xu et al, 1999): (1) Direct substitution approach (DSA) according to which chemical equations are directly substituted into the transport equations and (2) sequential iteration approach (SIA) in which transport and chemical equations are solved separately in a sequential manner and following an iterative procedure. DSA leads to a system of coupled highly non-linear transport equations. It was used by Carnahan (1990), Steefel and Lasaga (1994), White (1995), Ayora et al. (1998), Saaltink et al. (1998) and Saaltink (1999). Its main advantage is high accuracy and quadratic rate of convergence. Especially for chemically difficult (that is, highly non linear) cases, when the SIA requires very small time steps leading to excessive computation times, the DSA is consistently very robust and does not show this inconvenience (Saaltink, 1999).

However, For chemically simple cases but with grids of many nodes, the SDA tends to be less favorable because of the size of the set of equations to be solved. It is very demanding in terms of computing time and memory. This can seriously limit its use for large problems (Yeh and Tripathi, 1989). On the contrary, the sets of equations that are solved simultaneously in the SIA are much smaller than in the DSA, and therefor, larger systems with larger sets of chemical species can be handled with SIA. This approach was used by many investigators, such as Liu and Narasimhan (1989), Nienhuis et al (1991), Yeh and Tripathi (1991), Delgado et al. (1999), Engesgaard and Kipp (1992), Samper and Ayora (1994), Samper et al. (1998), Samper et al. (1999 a, b), Lensing et al. (1994), Simunek and Suarez (1994), Walter et al. (1994a, b), Zysset et al. (1994a, b), Neretnieks et al. (1997), and Xu et al. (1999).

A mixed model containing ion exchange and surface complexation is included in the PHREEQM code (Appelo and Postma, 1993). It consists of a one-dimensional transport code coupled to the PHREEQE (Parkhurst et al., 1980) geochemical code. The solution of the advection-dispersion equation is based on the mixing cell concept. For each time step and for each cell, the modified PHREEQE version computes aqueous speciation and mass transfer to/from minerals. Ion exchange reactions are incorporated into the PHREEQE database as half reactions of the cations with a fictitious species representing the sorbent substrate. The thermodynamic activities of sorbed species are calculated from equivalent fractions following the Gaines-Thomas convention. PHREEQM allows also calculating H^+ adsorption/desorption following a surface complexation model with the electrostatic term computed from a constant capacitance model. PHREEQM has been successfully applied to the description of the zonation of the chemical composition of groundwater in saline intrusion processes and refreshening of saline aquifers (Appelo, 1994).

Walter et al. (1994) presented the MINTRAN code which consists on coupling the finite element transport model PLUME2D and MINTEQA2 equilibrium geochemistry model. These authors claim that making use of the local equilibrium assumption, the inherent nonlinearity is confined to the chemical domain. This linearizes the coupling between physical and chemical processes and leads to a simple and efficient two-step sequential solution algorithm (without iteration). The program can handle realistic aquifer properties and boundary conditions. The

geochemical equilibrium model MINTEQA2 includes aqueous complexation, acid-base reactions, oxidation-reduction by either fixing the solution pE value or using the activities of electroactive species, cation exchange by constant charge model, adsorption via surface complexation by the constant capacitance, the diffuse layer and the triple-layer models. Adsorption isotherm models (K_d approach) are also included in the model. Activity coefficients are calculated using the extended Debye-Hückel and Davies equations. The Newton-Rapson iterative technique is used for solving the chemical equations.

Version 2 of PHREEQC (Parkhurst, 1999) is a computer program written in C that is designed to perform a wide variety of low-temperature aqueous geochemical calculations. It is based on an ion-association aqueous model and has capabilities for: (1) speciation and saturation-index calculations, (2) batch-reaction and one-dimensional (1D) transport calculations involving reversible reactions, which include aqueous, mineral, gas, solid-solution, surface-complexation, and ion-exchange equilibria, and irreversible reactions, which include specified mole transfers of reactants, kinetically controlled reactions, mixing of solutions, and temperature changes, and (3) inverse modeling, which finds sets of mineral and gas mole transfers that account for differences in composition between waters, within specified compositional uncertainty limits. Additional features in version 2 of PHREEQC relative to version 1 (Parkhurst, 1995) include capabilities to: (1) simulate dispersion (or diffusion) and stagnant zones in 1D-transport calculations, (2) model kinetic reactions with user-defined rate expressions, (3) model the formation or dissolution of ideal, multicomponent or nonideal, binary solid solutions, fixed-volume gas phases in addition to fixed-pressure gas phases, (4) allow the number of surface or exchange sites to vary with the dissolution or precipitation of minerals or kinetic reactants, (5) include isotope mole balances in inverse modeling calculations, and (6) automatically use multiple sets of convergence parameters. However, the capacities of this model to simulate flow and solute transport are limited.

Samper et al. (1999) developed a 2-D finite element multi-components reactive transport code, CORE^{2D}, (a COde for modeling water flow saturated or unsaturated, heat transport and multicomponent REactive solute transport under local chemical equilibrium or kinetic conditions), which can use either SIA or SNIA (Sequential non-iterative approach)

for the purpose of improving numerical efficiency. General water flow, heat transfer and solute transport boundary conditions are considered under fully or variably saturated media. The code accounts for a wide range of chemical processes such as aqueous complexation, acid-base, redox, mineral dissolution/precipitation (kinetic), gas dissolution/ exsolution, cation exchange, and adsorption via surface complexation. This code has been widely used in several research projects dealing with radioactive waste disposal at laboratory and field scales (Samper et al., 1998, 1999 and Delgado et al., 1999).

1.2.2. INVERSE MODELING OF REACTIVE SOLUTE TRANSPORT

As a whole, the solution of the inverse problem is still limited to flow and conservative solute transport models. In most of case studies, very fine reactive transport models are usually calibrated by trial-and-error method. Sophisticated inverse models are only available for estimating flow and solute transport parameters. A considerable effort has been devoted to estimation of flow parameters, and the number of studies of the groundwater flow inverse problem is large. Several different methods have been applied. Field-scale applications are also numerous (Neuman et al., 1980; Cooley, 1983; Carrera and Neuman, 1986c). Comprehensive reviews of the inverse problem of aquifer flow are provided by Yeh (1986), Carrera (1987), Sun (1994) and Samper and García-Vera (1998).

1.2.2.1. Inverse analysis of flow in variably saturated media

Over the past 15 years, interest has arisen in the feasibility of using inverse algorithms to analyze soil hydraulic properties from transient infiltration events (Kool et al., 1985; Parker et al., 1985; Kool and Parker, 1988; Russo, 1988; Yeh and Harvey, 1990; Russo et al., 1991; Kabala and Milly; 1991; Zayani et al. 1991; Eching et al., 1994). Morel-Seytoux et al. (1996) studied the parameter equivalence for the Brooks-Corey and van Genuchten soil characteristics and provided a simple way for the conversion between these two sets of parameters. Simunek and van Genuchten (1996) estimated the hydraulic properties of unsaturated soils via numerical inversion using a tension disc infiltrometer.

Inoue et al. (1998) demonstrated the potential application of the soil water extraction method for estimating soil water retention and unsaturated hydraulic conductivity parameters in the field. The method of simulated annealing is a technique that has attracted significant attention as suitable for optimization problems, especially for problems in which a desired global extremum is hidden among many, poorer, local extrema (Press et al, 1992). Pan and Wu (1998) developed a hybrid global optimization method which incorporates simulated annealing strategies into a classical downhill simplex method. They used the annealing-simplex method for the estimation of hydraulic parameters. The computing efficiency of this method, however, is greatly limited by the number of parameters.

1.2.2.2. Coupled inverse modeling of flow and solute transport

The coupled inverse approach provides a way whereby measurements of state are used to determine unknown flow and transport parameters by fitting the model output with the measurements. Wagner and Gorelic (1987) considered optimal groundwater quality management under parameter uncertainty in which a coupled flow-mass transport model was identified, and the parameter uncertainty was estimated. Chu et al. (1987) discussed data requirements for the calibration of a flow-mass transport model to determine unknown transmissivities and dispersivities. Woodbury et al. (1987) and Woodbury and Smith (1988) used temperature measurements to improve the estimation of hydraulic conductivities, in which a coupled inverse problem of flow and heat transport was solved. Kool and Parker (1987) and Mishra and Parker (1989) considered the inverse problem for coupled unsaturated flow and mass transport. A general framework and description of the coupled inverse problems of flow and transport is given by Sun and Yeh (1990). They defined the objective function based on an L_2 norm for each problem and coupled them with weighting coefficients to form a scalar objective function and the Jacobian matrix was computed by the adjoint state method. Samper et al. (1990), used an automatic calibration technique to construct a groundwater flow and solute transport model for the quantitative assessment of groundwater uranium pollution from a mill tailings pond in Southern Spain. They also performed sensitivity and uncertainty analyses related to flow and solute transport parameters. Gailey et al. (1991) demonstrated a field application of the model presented by

Wagner and Gorelick (1987). Based on steady state flow data and time series concentration data, Gailey et al. (1991) estimated log conductivity, recharge, and porosity. Further, the worth of head and concentration data was analyzed, showing that parameter estimates derived from head and concentration data were more reliable than estimates obtained from using only concentration data. Wagner (1992) presented an inverse flow and transport model based on nonlinear maximum likelihood theory. Using hypothetical aquifer 'measurements' corrupted with normally distributed random noise (steady state head data and transient concentration data), log conductivity, boundary recharge, dispersivities, porosity, and source disposal flux were estimated. The synthetic example illustrated the possibility of estimating the value of the statistical parameters partly describing the errors associated with the measured data and partly acting as weighting factors between the head and the concentration norms. Xiang et al. (1993) estimated permeabilities, dispersivities, porosities and concentrations of source terms in a steady flow and transport inverse problem by using an L_1 norm to define the objective function. They applied the sensitivity equation method to compute the Jacobian matrix. Medina and Carrera (1996) minimized the objective function of a coupled flow and transport inverse problem by using Marquardt's method and computing head and concentration derivatives by means of the sensitivity equation method.

The computation of the Jacobian matrix is one of the key steps in gradient-based optimization methods. The entries of this matrix are the derivatives of observations with respect to parameters which are usually called sensitivity coefficients. There are three approaches for computing sensitivity coefficients: the finite difference method (or the influence coefficient method), the sensitivity equation method and the adjoint state method (or variational method). Some comparisons of the three approaches can be found in work by, for example, Yeh (1986), Li et al. (1986), and Sun and Yeh (1990a). The sensitivity equation and adjoint state methods are widely used to compute the Jacobian matrix of the coupled inverse problem of flow and solute transport. Samper and Neuman (1986) derived the adjoint state equations for flow and advective-dispersivity transport using two approaches, continuous and discrete approaches, under both steady and transient conditions. Carrera and Neuman (1986), Sun and Yeh (1990), Sun (1994), Medina and Carrera (1996) and other researchers used these two methods to compute the Jacobian

Matrix, respectively. None of these authors considered chemical reactions. The computation of the Jacobian matrix for multi-components reactive transport becomes very difficult.

The finite difference method, on the contrary, does not require complex mathematical derivations. Generally, two kinds of finite difference methods can be used to compute the Jacobian matrix: forward and central difference methods. The forward difference method requires solving the forward numerical modeling once for each parameter to be estimated. A small increment is added to the parameter value prior to the run. The finite difference method is generally considered to be less accurate than other two methods because it is difficult to find appropriate parameter increments (Yeh, 1986, and Sun and Yeh, 1990a). However, with the improvement of forward modeling accurate and the accumulation of experience, this method has been successfully used to solve complex inverse problems for which adjoint state and sensitivity methods are difficult to be derived. Mishra and Parker (1989) used forward finite difference method to compute the Jacobian matrix for estimating flow and transport parameters of soils from transient unsaturated flow and tracer experiments by a combined simulation-optimization approach. Barlebo et al. (1998) used the central difference method to evaluate the Jacobian matrix with a perturbation of 10^{-4} times the parameter estimate for the inverse problem of a three-dimensional flow and transport model. Sonnenborg et al. (1996) presented also an application of an inverse flow and transport model to a contaminated aquifer using such method. Under complex conditions, the finite difference method leads to a substantial saving of computation time for calculating the Jacobian matrix.

1.2.2.3. Coupled inverse problem of reactive solute transport

There is no reported inverse model for the estimation of geochemical parameters within the context of multi-component reactive transport. However, some model-independent parameter estimation techniques such as PEST (Doherty et al., 1994) and UCODE (Poeter and Hill, 1998) are purported to be universal inversion codes which can be used to estimate the parameters of any model. These techniques require running the model as many times as needed while adjusting its parameters until the discrepancies between

model output and field or laboratory data are reduced to a minimum in the weighted least squares sense. These estimation techniques implement a robust variant of the Gauss-Marquardt-Levenberg method of nonlinear parameter estimation in which the Jacobian matrices are computed with finite differences. This method allows one to estimate parameters of complex systems for which the computation of sensitivity coefficients with the sensitivity equation or the adjoint state methods is not straight-forwards. There are no reported examples in which these inverse tools are used to estimate geochemical parameters.

Trial and error methods (Anderson and Woessner, 1992 and Sun 1994) are commonly used to estimate the parameters of reactive transport models. Appelo and Postma (1993) and Appelo et al. (1990) used PHREEQM to model column experiments. They claim that there are some deviations between measured and computed concentrations because of the variation of the exchange coefficients that had not been identified by some inverse models. Latterly, Appelo et al. (1998) reported the hydrogeochemical modeling of a complicated suite of reactions that took place during the oxidation of pyrite in a marine sediment by using an adapted version of PHREEQC (Parkhurst, 1995). The key kinetic parameter, the reciprocal of the specific surface (ratio *A/V* of surface over volume), was adjusted to fit the data by trial and error.

Keating and Bahr (1998a, b) proposed a trial and error method for calibrating a groundwater flow model at a site in Northern Wisconsin using both head and concentration data in a system with kinetically-controlled mineral dissolution processes. They used a two-stage approach. First, they tested the consistency of residence times computed with a flow model with solute concentrations predicted by a reaction-path geochemical model. Then, for a more robust comparison, they included the effects of advection and mixing on solute concentrations using a reactive transport model. They got satisfactory results with this approach. Recently, Saaltink et al. (1998) and Saaltink (1999) applied his reactive transport model to simulate deep well recharge of oxic water into an anoxic sandy pyrite-bearing aquifer at Langerak (Stuyfzand and Timmer, 1999). They used a trial and error method to estimate chemical and hydrological parameters and obtained a reasonable agreement between calculated and measured concentration data.

The trial and error procedure, however, is time consuming, especially when the number of estimated parameters is large. This method does not provide an accurate solution for the inverse problem. Different researchers may obtain different results (Sun, 1994). Moreover, the process is incomplete and it is hard to evalute the quality of the estimated results. Therefore, there is a need to develop more sophisticated and objective calibration methods based on the solution of the inverse problem.

1.3. LIMITATIONS OF EXISTING MODELS

Existing inverse models can only cope with the identification of the water flow and conservative solute transport parameters. Model-independent parameter estimation techniques such as PEST (Doherty et al., 1994) and UCODE (Poeter and Hill, 1998) which are purported to be universal inversion codes have not been used so far to calibrate reactive solute transport models because these inversion codes are complex and usually require much CPU time. The extremely large computer requirement almost makes it impossible their application to the inverse problem of complex reactive transport models. Geochemical parameters are very special, such as kinetic reaction parameters, CEC and selectivity coefficients of the cation exchange, initial and boundary concentrations of different components. They are arranged in a different parameterization method from each other. Furthermore, the calibration of reactive solute transport model involves more types of observation data, such as heads, inflow rates, water contents, and concentrations of dissolved and exchanged chemical species. Solving the inverse problem of multi-component reactive transport requires a systematic and optimum weighting method which at present is not yet implemented in PEST.

1.4. MAIN CAPABILITIES AND FEATURES OF MODELS CONSTRUCTED IN THIS DISSERTATION

This dissertation deals with inverse modeling of reactive solute transport in variably saturated media. Three inverse programs have been developed based on the following optimization techniques: Golden section search, quadratic interpolation and Gauss-Newton-Levenverg-Marquardt methods. The first two methods use linear and nonliear search algorithms and do not require computing the Jacobian matrix. They are useful for estimating a few parameters. For a large number of parameters, they are not economical. The more popular and powerful method of Gauss-Newton-Levenberg-Marquardt method has been developed to solve large inverse problems. This method has been implemented in the final version of the inverse code, INVERSE-CORE^{2D}, which solves both the forward and inverse problems. The forward modeling part is based on CORE^{2D} (Samper et al, 1999) and includes:

1) 2-D confined or unconfined, saturated or unsaturated steady-state or transient groundwater flow. General boundary conditions are allowed.
2) Transient heat transport considering conduction, heat dispersion and advection processes.
3) Solute transport including advection, molecular diffusion and mechanical dispersion.
4) Chemical equilibrium reactions including:
   a) Acid-base,
   b) Redox,
   c) Aqueous complexation,
   d) Surface adsorption via surface complexation and K_d-approach,
   e) Ion exchange,
   f) Mineral dissolution-precipitation,
   g) Gas dissolution-exsolution,
5) Chemical kinetics for dissolution/precipitation of minerals.
6) Radioactive decay.

Five different types of data can be taken into account for the inverse problem, which include: hydraulic heads, solute concentrations of different chemical components, total concentration including liquid and solid phases, cumulative water inflows and water contents. Prior information of the parameters can be included in the objective function. These data have been generalized in a weighted least squares criterion by means of a set of weighting coefficients.

INVERSE-CORE^{2D} uses the Gauss-Newton-Levenberg-Marquardt algorithm to minimize the objective function following an iterative process. It is necessary to supply a set of initial parameter values at the beginning of a run. They are used to start the first optimization iteration. Derivatives of observations with respect to parameters are calculated using a forwards finite difference method. During each optimization iteration, the model is run once for each estimated parameter and a small parameter increment is added to the parameter value prior to the run. The resulting changes in predicted responses (heads, concentrations, water inflows and water contents) are divided by the parameter increment in order to calculate sensitivity coefficients. This is repeated for each parameter. Derivatives calculated with the forwards finite difference method may not always be accurate. To alleviate the inaccuracy of these derivatives, the code allows these derivatives be calculated using the central difference method which requires two runs for each parameter. For the first run, an increment is added to the current parameter value, while for the second run, the increment is subtracted.

By comparing parameter changes and objective function improvement achieved through the current iteration with those achieved in previous iterations, the algorithm can tell whether it is worth undertaking another optimization iteration. If so, the whole process is repeated. Five or six optimization iterations are usually required. In some cases convergence may be slow, requiring more iterations. Finally, the program stops when a convergent solution is attained or the specified maximum number of iterations is exceeded.

Error estimates for the estimated parameters as well as a statistical measure of the goodness-of-fit are computed to evaluate the accuracy of parameter estimates. The code calculates the variance-covariance and correlation matrices of the estimated parameters,

performs the eigenanalysis of the covariance matrix, and computes approximate confidence intervals of the estimated parameters.

INVERSE-CORE^{2D} can estimate the following 16 types of parameters which may take different values at different parameter zones:

1) Total porosity,
2) Components of the saturated hydraulic conductivity (K_x and K_y),
3) Specific yield,
4) Parameters m, n, and α of van Genuchten retention curve,
5) Molecular diffusion coefficient in water,
6) Solute dispersivity,
7) Distribution coefficient K_d,
8) Accessible porosity (which may differ from total porosity when anion exclusion takes place),
9) Initial concentrations of chemical components, including initial pH and pE,
10) Boundary concentrations of chemical components, including pH and pE,
11) Selectivity coefficients for cation exchange,
12) Cation Exchange Capacity (CEC),
13) Specific surface of minerals,
14) Apparent activation energy of overall reaction process,
15) Rate constants of kinetic reactions,
16) Exponents of the catalytic terms and those of the ratio of the ionic activity product and the equilibrium constant (for kinetic processes).

Three sets of synthetic data have been used to verify the codes and study the convergence, uniqueness and stability of the algorithms. The codes have been used successfully to interpret diffusion, permeation, infiltration, and reactive transport experiments (Appelo et al., 1990). The inverse model has also been applied to field data from the Llobregat Delta Aquitard (Barcelona, Spain) and the Aquia aquifer in Maryland (USA).

1.5. SCOPE

Chapter 2 presents a brief description of the mathematical and the numerical aspects of the forwards problem of water flow and reactive solute transport in variably media. The most relevant physical processes involved on groundwater flow and solute transport and heat transport in aquifers and variably saturated porous media are described in this chapter together with their mathematical formulation. Quantification of these processes for conservative solutes leads to an advection-dispersion equation. The analysis of hydrochemical interactions requires the definition and characterization of the geochemical system. To this purpose, a general formulation for reactions in complex chemical systems is also presented. The main features of homogeneous and heterogeneous reactions are discussed. Contrary to static systems where the chemical equilibrium condition is largely justified, in dynamic systems some reactions may progress slowly and be kinetically-controlled. Kinetics of mineral dissolution-precipitation is taken into account. Numerical methods for the solution of the forwards reactive solute transport problem are also described.

The formulation and the algorithms of the inverse problem of reactive solute transport are discussed in Chapter 3. The first section of this chapter covers the issues of model structure, and parameters, ill-posedness, non-uniqueness and parameter identifiability. The formulation of the coupled inverse problem is developed in the second section of the chapter. Methods for solving the inverse problem such as trial and error, Golden section search, quadratic interpolation, Newton's method, and Gauss-Newton-Levenberg-Marquardt method are described in Section 3. A posteriori error analysis of estimated parameter is presented in Section 4. Statistic methods are provided to evaluate the quality of the estimated results which include the variance-covariance and the correlation matrices of parameter estimates, the eigenvalues and eigenvectors of the covariance matrix and approximate confidence regions and intervals of the estimated parameters.

A description of the codes and their verification is presented in Chapter 4. Solution methods, code structure and capabilities are provided for the three inverse codes, INVGS-CORE, INVQI-CORE and INVERSE-CORE^{2D}. A computer code is a set of commands or statements that are intended to solve the numerical equations derived from the original mathematical problem. Prior to any application of the code, one must check that it solves correctly the equations it is intended to solve. Checking that computer calculations are correct is known as code verification. Three synthetic examples dealing with water flow and solute transport (EJ1), cation exchange reactions (EJ2), and cation exchange and mineral kinetic dissolution/precipitation (EJ2KIN) have been used to verify the codes, using genuine data and data corrupted with white-noise error.

Chapter 5 presents several examples of application of the inverse methodology to real case studies which demonstrate its potential to cope with real problems. The first case deals with the interpretation of infiltration experiments performed on unsaturated bentonite for the purpose of estimating unsaturated flow parameters. Parameters of the retention curve and relative permeability of the bentonite have been successfully estimated using cumulative water infiltration and final water content data. Section 2 describes the interpretation of through-diffusion, in-diffusion and permeation experiments in compacted bentonite, using different tracers such as tritium, strontium, cesium and selenium. In addition to the estimation of flow and transport parameters, suggestions are provided to improve experiment design on the basis of the uncertainty analysis of the estimated results. The interpretation of two permeation experiments performed on granite samples is presented in Section 3. In this case, porosity, diffusion coefficient and dispersivity of the granite are estimated.

The inverse code INVERSE-CORE^{2D} has been used to interpret a column experiment performed by Appelo et al. (1990) which deals with flushing the cation exchange complex of a sediment taken from the Ketelmeer aquifer with $SrCl_2$. Appelo et al. (1990) presented the numerical interpretation of this experiment using a trial and error method. Using their results as initial parameter estimates, the solution of the inverse problem provides optimal estimates of geochemical parameters which match the measured data better than Appelo's results. As described in the fourth section of this chapter, the inverse analysis has been

useful also to investigate the role of relevant geochemical processes taking place in the experiment.

Section 5 contains the inverse analysis of the Llobregat Delta aquitard data previously modeled by Manzano (1993) and Xu et al. (1999). Selectivity coefficients, CEC, hydraulic conductivity, dispersivity, and initial and boundary concentrations of some chemical components are estimated simultaneously for several hypotheses regarding redox processes. The uncertainties of the estimated parameters have been quantified by means of 90% confidence intervals. Again, the solution of the inverse problem provides: additional insight into the role of transport and geochemical processes, a better fit to measured data and optimal estimates of transport and geochemical parameters.

The solution of the inverse problem in a complex field case study in Aquia aquifer (Maryland, USA) is presented in Section 6. Available geochemical data have been used to estimate hydraulic conductivities, leakage rates, dispersivities, selectivity coefficients, CEC, and initial and boundary concentrations of several chemical components. The following chemical reactions are involved: aqueous complexation, acid-base, cation exchange, proton exchange which is simulated by surface sorption, mineral precipitation-dissolution, and redox reactions. This case study was simulated numerically by Appelo et al. (1990) using PHREEQM. Several hypotheses have been tested regarding leakage into the overlying aquifer, the role of proton exchange, and chemical equilibrium versus calcite kinetics. A good fit of measured concentrations is attained also in this case.

Chapter 6 includes the main conclusions of this dissertation and some recommendations for future research.

Appendix 1 contains a list of the main variables used in this dissertation. Appendix 2 introduces the algorithm for generation of the synthetic data which are corrupted by log-normal errors. Appendix 3 includes details of the iterative processes and estimation results of example EJ2. To illustrate the role of prior information, Appendix 4 presents the iterative processes and estimation results of EJ2 example in which prior information is considered, which improves the convergence rate of the inverse code. Appendix 5 describes the input

data required to run the code INVERSE-CORE^{2D} and the output results. This appendix is part of the User's Manual of INVERSE-CORE^{2D}.

CHAPTER 2

MATHEMATICAL AND NUMERICAL FORMULATION OF FORWARD MODELING

This chapter presents the mathematical formulation of the forward problem for groundwater flow, heat transport and multicomponent reactive transport. This formulation has been adapted from that presented by Samper et al. (1995) and Xu (1996) which was used to develop CORE^{2D} (Samper et al., 1998; 2000). It includes the formulation of saturated water flow (Section 2.1.1.1), unsaturated flow (Section 2.1.1.2), solute transport (Section 2.1.2), heat transport (Section 2.1.3), and chemical reactions (Section 2.2). The main features of the finite element numerical methods used to solve this mathematical problem are also described. Further details can be found in Samper et al. (2000).

2.1. GROUNDWATER FLOW, HEAT TRANSFER AND SOLUTE TRANSPORT

2.1.1. GROUNDWATER FLOW

2.1.1.1. Aquifer flow

Water flow through porous media is governed by Darcy's Law which relates water flux **q** to the gradient of water pressure p and elevation z through

$$\mathbf{q} = -\frac{\mathbf{k}}{\mu}(\nabla p + \rho g \nabla z) \qquad (2.1.1)$$

where ρ is water density (mass per unit volume), μ is dynamic viscosity, **k** is intrinsic permeability tensor and g is gravity acceleration. When density changes are negligible, Darcy's Law can be written in terms of hydraulic head h as

$$\mathbf{q} = -\mathbf{K}\nabla h \tag{2.1.2}$$

where

$$\mathbf{K} = \frac{\mathbf{k}\rho g}{\mu} \tag{2.1.3}$$

and

$$h = \frac{p}{\rho g} + z \tag{2.1.4}$$

Here **K** is the hydraulic conductivity tensor. By combining Darcy's Law and the mass balance equation one has

$$\nabla \cdot (\mathbf{K}\,\nabla h) + w = S_s \frac{\partial h}{\partial t}$$

where w represents fluid sink/sources per unit volume of medium and S_s is storativity (or specific storage coefficient) defined as the volume of water delivered per unit time and unit volume of medium in response to a unit change of hydraulic head.

When analyzing groundwater flow through aquifers, it is usual to integrate the flow equation over the entire aquifer thickness, b, defined as the difference between the top and bottom elevations, z_t and z_b, respectively. By so doing, the flow equation becomes

$$\nabla \cdot (\mathbf{T}\nabla h) + r = S \frac{\partial h}{\partial t} \tag{2.1.5}$$

where **T** is the transmissivity tensor, and S is the storage coefficient, which are defined as

$$\mathbf{T}(x,y) = \int_{z_b}^{z_t} \mathbf{K}(x,y,z)dz = \overline{\mathbf{K}}b$$

$$S(x,y) = \int_{z_b}^{z_t} S_s(x,y,z)dz = \overline{S}_s b$$

Here $\overline{K}$ and $\overline{S}_s$ represent vertically averaged values of hydraulic conductivity and storativity, respectively. The term r represents a fluid sink/source term per unit surface area. In unconfined aquifers the upper aquifer limit coincides with the water table. In these aquifers, the transmissivity T is given by $\overline{K}\left(h - z_b\right)$ and depends on hydraulic head. In addition, the storage coefficient S becomes.

$$S = \overline{S}_s\left(h - z_b\right) + S_y$$

where the specific yield S_y is usually several orders of magnitude greater than the elastic term $\overline{S}_s\left(h - z_b\right)$.

The governing flow equation (2.1.5) must be complemented with appropriate initial and boundary conditions. Initial heads $h_0(x, y)$ are either known or satisfy steady-state conditions. In the latter case, h_0 satisfies the following equation

$$\nabla \cdot \left(\mathbf{T}\nabla h_0\right) + r_0 = 0 \tag{2.1.6}$$

where r_0 represents the steady-state source term.

At the boundary Γ of the domain Ω, either the head or its gradient (water flux) is known. Possible boundary conditions include (Samper et al., 2000):

(1) Dirichlet condition on Γ_1

$$h(x, y, t)\big|_{\Gamma_1} = H \tag{2.1.7}$$

(2) Neuman condition on Γ_2

$$\mathbf{T}\nabla h \cdot \mathbf{n}\big|_{\Gamma_2} = Q \tag{2.1.8}$$

(3) Mixed condition on Γ_3

$$\mathbf{T}\nabla h \cdot \mathbf{n}\big|_{\Gamma_3} = \alpha(H - h) \tag{2.1.9}$$

where **n** is a unit vector orthogonal to Γ and pointing outwards, H and Q are specified heads and fluxes which may vary in space and time, and α is a leakage coefficient $\left(LT^{-1}\right)$, and Γ_1, Γ_2, and Γ_3 are pieces of Γ. By convention, Q is positive when water flows into the aquifer.

2.1.1.2. Variably saturated flow

In variably saturated media, hydraulic head is written as the sum of the pressure head ψ and elevation z (a form slightly different to Equation 2.1.4),

$$h = \psi + z \tag{2.1.10}$$

Pressure head takes on positive values in saturated regions and negative values in partly saturated regions.

Similar to groundwater flow in aquifers, the governing equation for flow in variably saturated porous media can be derived from mass conservation and Darcy's Law (Voss, 1984; Galarza, 1993):

$$\nabla \cdot [K_r \mathbf{K} \nabla(\psi + z)] + w = (\phi \frac{\partial S_w}{\partial \psi} + S_w S_s) \frac{\partial \psi}{\partial t} \tag{2.1.11}$$

where the hydraulic conductivity is the product of relative conductivity K_r (a function of pressure head) and saturated conductivity K. The term multiplying the time derivative of ψ in (2.1.11) is the storage capacity which depends on pressure head. The first term is important in unsaturated regions while the elastic term is relevant mostly in saturated regions. Water saturation S_w is defined as the ratio of volumetric water content θ to porosity ϕ,

$$S_w = \frac{\theta}{\phi} \tag{2.1.12}$$

Equation (2.1.11) is highly nonlinear because both hydraulic conductivity and moisture content are functions of pressure head. These functions are characteristic of the porous medium and are usually derived by fitting known functions to experimental data. There are a number of such functions in the literature. Among them, van Genuchten (1980) equations are the most commonly used. These functions are given by

$$S_w(\psi) = S_r + (1 - S_r)\left[1 + (-\alpha\psi)^n\right]^{-m} \tag{2.1.13a}$$

$$K_r(\psi) = \sqrt{S_e}\left[1 - (1 - S_e^{1/m})^m\right]^2 \tag{2.1.13b}$$

where m and S_e are defined as

$$m = 1 - \frac{1}{n} \tag{2.1.14}$$

$$S_e = \frac{S_w - S_r}{1 - S_r} \tag{2.1.15}$$

S_r is the residual water saturation, m, n and $\alpha\ (1/L^{-1})$ are constants which are obtained by fitting these functions to experimental data. For values of m close to 1, parameter α can be interpreted as the reciprocal of the critical suction head Ψ_c (air-entry pressure head):

$$\alpha = \frac{1}{\Psi_c}$$

which can be derived from porosity and effective particle diameter d_{10} (Bear, 1972).

There are in the literature several expressions for the relative permeability as a function of saturation. The van Genuchten expression of K_r presented in Equation

(2.1.13b) is obtained by using Mualem (1976) model combined with the retention curve of Equation (2.1.13a) (Samper, 1991). Other functions include that proposed by Irmay (1954)

$$K_r(S_e) = S_e^3 \tag{2.1.16}$$

which generally leads to good fits to experimental data for soils of uniform size.

Equation (2.1.11) is not only valid for unsaturated regions, but also for saturated regions where both water saturation and relative hydraulic conductivity are equal to 1.

The solution of Equation (2.1.11) requires knowing initial pressure heads ψ_0 (at t = 0):

$$\psi(x, z, t = 0) = \psi_0(x, z) \tag{2.1.17}$$

which vary with horizontal coordinate x and elevation z (for a 2-D vertical cross-section). At the boundaries of the flow domain, either the pressure head or its gradient must be known. Possible boundary conditions include:

(1) Dirichlet condition on Γ_1

$$\psi(x, z, t)\big|_{\Gamma_1} = \psi_1 \tag{2.1.18}$$

(2) Neuman condition on Γ_2

$$K_r \mathbf{K} \nabla(\psi + z) \cdot \mathbf{n}\big|_{\Gamma_2} = Q \tag{2.1.19}$$

(3) Mixed condition on Γ_3

$$K_r \mathbf{K} \nabla(\psi + z) \cdot \mathbf{n}\big|_{\Gamma_3} = \alpha[H - (\psi + z)] \tag{2.1.20}$$

where ψ_1, H and Q are specified pressure heads, hydraulic heads and water fluxes, respectively (they may vary in space and time); $\mathbf{n}$ is a unit vector orthogonal to Γ and pointing outwards; and α is leakage coefficient.

(3) Free drainage (unit vertical hydraulic gradient) condition. This condition is commonly used as a bottom outflow boundary condition where the water table lies far below the domain of interest. It is a particular case of a Neuman condition where the flux is not specified, but updated as simulation progresses.

2.1.2. TRANSPORT OF CONSERVATIVE SOLUTES

Dissolved species in saturated media are subject to transport and chemical processes. Transport processes include: (1) advection, (2) molecular diffusion and (3) hydrodynamic dispersion. Each of them produces a solute mass flux **F** (mass of solutes crossing a unit surface area of medium per unit time).

Advection refers to solute migration associated to water flow. Solutes move with water. If water flows at a specific discharge **q** (volumetric water flux), the advective solute flux $\mathbf{F}_A$ is given by

$$\mathbf{F}_A = \mathbf{q}c \tag{2.1.21}$$

where c is solute concentration, usually expressed as solute mass (gram or mol in reactive solute transport) per unit fluid volume. Solutes migrate at an average velocity **v** given by

$$\mathbf{v} = \frac{\mathbf{q}}{\theta} \tag{2.1.22}$$

where θ is volumetric water content which is equal to porosity ϕ for saturated media. From (2.1.21) and (2.1.22) it follows that

$$\mathbf{F}_A = \theta \mathbf{v} c \tag{2.1.23}$$

Molecular diffusion is a transport mechanism related to the continuous Brownian motion of solute and fluid molecules. Molecular diffusion in pure water produces a mixing effect which obeys Fick's first law which states that solute flux due to diffusion $\mathbf{F}_D$ is proportional to the concentration gradient ∇c :

$$\mathbf{F}_D = -D_0 \nabla c$$

where D_0 is the molecular diffusion coefficient in water. Solutes in porous media can only diffuse along fluid pores following tortuous paths. This means that the effective molecular diffusion coefficient for porous media D_e is smaller than that for pure water. Usually D_e is related to D_0 through

$$D_e = \theta \tau D_0 \tag{2.1.24}$$

where τ is tortuosity. For partially saturated porous media, tortuosity is related to water content through relationships such as (Simunek and Suares, 1993)

$$\tau = \frac{\theta^{7/3}}{\phi^2}$$

Therefore, the diffusive flux in porous media is given by:

$$\mathbf{F}_D = -\theta \tau D_0 \nabla c \tag{2.1.25}$$

In addition to molecular diffusion there is a mixing phenomenon known as hydrodynamic dispersion which is caused by medium heterogeneities and produces both longitudinal and transverse solute spreading. Actual flow paths are highly irregular. Some water particles move faster than the average velocity while others displace more slowly. The overall effect of all heterogeneities is a solute spreading in all directions.

Laboratory and field evidence indicates that this phenomenon can be described by Fick's Law. Therefore, the hydrodispersive flux $\mathbf{F}_H$ can be described as

$$\mathbf{F}_H = -\theta\, \mathbf{D}_h \nabla c \tag{2.1.26}$$

where $\mathbf{D}_h$ is the hydrodynamic or mechanical dispersion tensor. Its principal directions coincide with the flow direction and its normals. The component along the flow direction D_L is the largest and is given by

$$D_L = \alpha_L |\mathbf{v}| \tag{2.1.27}$$

while the smallest components D_T occur along the transverse directions and are given by

$$D_T = \alpha_T |\mathbf{v}| \tag{2.1.28}$$

where $|\mathbf{v}|$ is the modulus of the velocity vector $\mathbf{v}$, and α_L and α_T are the longitudinal and transverse dispersivities which are characteristic parameters of the medium. They have dimensions of length and measure the scale of the spatial heterogeneities.

$\mathbf{D}_h$ is a symmetric tensor whose components in two dimensions are:

$$D_{xx} = \frac{\alpha_L v_x^2 + \alpha_T v_y^2}{|\mathbf{v}|} \tag{2.1.29a}$$

$$D_{yy} = \frac{\alpha_L v_y^2 + \alpha_T v_x^2}{|\mathbf{v}|} \tag{2.1.29b}$$

$$D_{xy} = D_{yx} = (\alpha_L - \alpha_T)\frac{v_x v_y}{|\mathbf{v}|} \tag{2.1.29c}$$

For practical purposes, the effects of molecular diffusion and hydrodynamic dispersion are usually lumped in a single dispersion tensor $\mathbf{D}$ which takes the form:

$$\theta \mathbf{D} = \mathbf{I} D_e + \theta \mathbf{D}_h \tag{2.1.30}$$

where **I** is the identity tensor.

The equation governing solute transport through porous media is derived from the principle of mass conservation. This principle states that for any reference elementary volume of medium, the net flux plus sink/source terms must be equal to the time rate of change of the solute mass contained in the reference volume. Solute mass per unit volume of medium is equal to θc. The net solute flux is given by minus the divergence of the total flux vector. Therefore, mass conservation leads to the following equation:

$$-\nabla \cdot (\mathbf{F}_A + \mathbf{F}_D + \mathbf{F}_H) = \frac{\partial(\theta c)}{\partial t} \tag{2.1.31}$$

where $\nabla \cdot (\ )$ is the divergence operator which when applied to a vector **F** of components (F_x, F_y, F_z) is equal to

$$\nabla \cdot (F) = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \tag{2.1.32}$$

Substituting mass fluxes $\mathbf{F}_A$ (Equation 2.1.21), $\mathbf{F}_D$ (Equation 2.1.25) and $\mathbf{F}_H$ (Equation 2.1.26) into the continuity Equation (2.1.31) and taking into account (2.1.30) leads to

$$\nabla \cdot (\theta \mathbf{D} \nabla c) - c \nabla \cdot \mathbf{q} - \nabla c \cdot \mathbf{q} = \frac{\partial(\theta c)}{\partial t}$$

Solute sinks and sources are added to the left-hand-side of this equation. For a fluid source of water flux w (per unit volume of medium) having a concentration c^*, and

a solute sink/source term R (solute mass added per unit time and unit fluid volume) the transport equation becomes (Samper et al., 1998),

$$\nabla \cdot (\theta \mathbf{D} \nabla c) - \mathbf{q} \cdot \nabla c + w(c^* - c) + \theta R = \theta \frac{\partial c}{\partial t} \tag{2.1.33}$$

where the following identity, which derives from the flow equation (by disregarding time variations in water density), has been taken into account

$$\left(-\nabla \cdot \mathbf{q} + w - \frac{\partial \theta}{\partial t} \right) c = 0 \tag{2.1.34}$$

The solution of the transient solute transport equation requires knowing:

(1) transport parameters which include: porosity, ϕ, molecular diffusion coefficient D_0, tortuosity, τ, and dispersivities (α_L, α_T).

(2) sink/sources: w, c^* and R.

(3) initial conditions: $c_0(x, y)$ at $t = t_0$

(4) boundary conditions

The initial condition c_0 is either known or can be obtained from the solution of a steady-state transport problem such as

$$\nabla \cdot (\theta_0 \mathbf{D} \nabla c_0) - \mathbf{q}_0 \nabla c_0 + w_0 (c_0^* - c_0) + \theta_0 R_0 = 0 \tag{2.1.35}$$

At the boundary of the domain, either concentration or a function of its gradient must be known. Possible types of boundary conditions include (Samper et al., 2000):

(1) Dirichlet condition. The points lying at this part of the boundary, Γ_1, satisfy the following condition

$$c\big|_{\Gamma_1} = \tilde{c} \tag{2.1.36}$$

where $\tilde{c}$ is a known and specified concentration, which may vary in space and time.

(2) Neumann condition. Let **n** be the unit vector normal to Γ_2, the part of Γ at which the dispersive flux $\tilde{F}_D$ is known. This type of condition states that the component of the dispersive flux normal to the boundary is given by

$$-\theta \mathbf{D}\nabla c \cdot \mathbf{n}\Big|_{\Gamma_2} = \tilde{F}_D \tag{2.1.37}$$

This condition is commonly used at impervious boundaries where $\tilde{F}_D$ is equal to zero.

(3) Cauchy mixed condition. Some parts of the boundary, Γ_3, may have a condition in terms of the total mass flux:

$$(-\theta \mathbf{D}\nabla c + \mathbf{q}c)\cdot \mathbf{n}\Big|_{\Gamma_3} = F_0 \tag{2.1.38}$$

The imposed flux is given by $\mathbf{F}_0$. Usually $\mathbf{F}_0$ is taken equal to the advective flux $c\mathbf{q}\cdot\mathbf{n}$.

The solute mass flux at outflow boundaries is usually taken as the product of the water flux $\mathbf{q}\cdot\mathbf{n}$ times the concentration c of the flowing water. In this case, $F_0 = c\mathbf{q}\cdot\mathbf{n}$ and therefore the boundary condition reduces to

$$-\theta \mathbf{D}\nabla c \cdot \mathbf{n}\Big|_{\Gamma_3} = 0 \tag{2.1.39}$$

which is a particular case of (2.1.37) with $\tilde{F}_D = 0$.

All types of boundary conditions can be written in the following general closed form:

$$-\theta \mathbf{D}\nabla c \cdot \mathbf{n} = \beta(\tilde{c} - c) + \tilde{F}_D \tag{2.1.40}$$

where β is a parameter that controls the type of condition. For Neumann conditions $\beta = 0$. While for Dirichlet conditions $\beta \to \infty$ and $\tilde{F}_D = 0$. At outflow boundaries $\tilde{F}_D = 0$ and $\beta = 0$ while at inflow boundaries $\tilde{F}_D = 0$ and $\beta = \mathbf{q} \cdot \mathbf{n}$.

When the transport equation is integrated over aquifer thickness, b, the result is

$$\nabla \cdot (\theta b \mathbf{D} \nabla c) - b\mathbf{q}\nabla c + r(c^* - c) + b\theta R = b\theta \frac{\partial c}{\partial t} \qquad (2.1.41)$$

where r is the fluid source term per unit surface area.

2.1.3. TRANSPORT OF DECAYING SOLUTES

Dissolved species in porous media can be subject to chemical and biological degradation as well as to radioactive decay. In this case, it is necessary introduce these processes into the transport equation. The formulation described here refers to radioactive decay chains although it can be used for other linearly decaying processes.

In a radioactive chain a species decays producing another species. This new species can in turn decay into another species. In general, one has

$$P \xrightarrow{\lambda_P} D \xrightarrow{\lambda_D} E \xrightarrow{\lambda_E} \ldots \to F$$

where P is the original or father species and D and E are the first and second species of the chain. The time variation of species D is given by:

$$\frac{dc_D}{dt} = \lambda_P c_P - \lambda_D c_D \qquad (2.1.42)$$

where c_p is the concentration of the father species. In the absence of other transport and chemical processes, c_p obeys the following equation:

$$c_p = c_p^\circ \exp(-\lambda_p t) \tag{2.1.43}$$

where c_p° is the initial concentration and λ_P is the decay constant of the father species. Introducing radioactive decay given by (2.1.42) into the transport equation (2.1.41) of species D, one has the following transport equation for a radioactive species D:

$$\nabla \cdot (\theta b \mathbf{D} \nabla c_D) - b\mathbf{q}\nabla c_D + r(c_D^* - c_D) - \lambda_D b\theta c_D + \lambda_P b\theta c_P + b\theta R = b\theta \frac{\partial c_D}{\partial t} \tag{2.1.44}$$

where λ_D is the decay constant of species *D*.

2.1.4. TRANSPORT OF SOLUTES SUFFERING EXCLUSION

Ion exclusion is a process by which negatively charged ions are repelled from the surface of clay particles. These anions are excluded from a significant part of the pores, and especially from the smallest pores. This process together with the mechanism of exclusion caused by size effects has led to the definition of diffusion accessible porosity. These exclusion effects are known to be relevant for Cl^-, and I^-. Since anion exclusion leads to a reduction on accessible porosity, its effect on a system with advective transport is that the anion tends to move faster than the average water velocity. This is termed sometimes as a negative sorption. The fraction of total porosity accessible to the solute is called accessible porosity ϕ_a. For a species suffering exclusion one has,

$$\theta c = \theta_a c_a \tag{2.1.45}$$

where θ_a is the water content accessible to the solute and c_a is the concentration in the accessible water. In a partly saturated media, θ_a can expressed as:

$$\theta_a = \theta - \phi + \phi_a \tag{2.1.46}$$

where ϕ_a is the accessible porosity. This expression assumes that θ_a varies linearly with θ. The transport equation for a species suffering exclusion is given by:

$$\nabla \cdot (\theta_a b\mathbf{D}\nabla c_a) - b\mathbf{q}\nabla c_a + r(c_a^* - c_a) + b\theta_a R = b\theta_a \frac{\partial c_a}{\partial t} \tag{2.1.47}$$

Accessible porosity is a solute-dependent parameter which must be determined from advection and dispersion experiments. The tortuosity of species suffering exclusion is given by:

$$\tau_a = \frac{(\theta_a)^{7/3}}{\phi^2} \tag{2.1.48}$$

2.1.5. HEAT TRANSPORT

Heat transport in porous media is governed by four separate mechanisms: (1) conduction in the solid matrix, (2) transport by the fluid phase (advection), (3) conduction in the fluid phase, and (4) heat exchange between the two phases depending on their temperature difference. The first phenomenon would produce a heat equation relative to the mean temperature T_S of the solid. The second and third would resemble the dispersion equation for the fluid with the fluid temperature T playing the role of the concentration. The fourth would be related to the exchange mechanisms between the solid and the liquid phase. However, in practice, except for a few cases, it can be safely assumed that the temperature of the solid and that of the fluid become identical almost at once, and that there is only one temperature T in the porous medium. All the theory on solute transport can then be applied to heat transfer in porous media. A single temperature is calculated for the porous medium. Heat transport processes include advection in a manner similar to that of the solutes and dispersion. Pure conduction in

the two phases, solid plus liquid, takes the place of molecular diffusion, while the heterogeneity of the real velocity gives rise to an anisotropic "fictitious thermal conductivity", equivalent to the mechanical dispersion, which experience shows to be a linear function of the modulus of the velocity vector (de Marsily, 1986).

The principle of heat conservation (analogous to the solute mass balance) for a saturated porous media is:

$$\nabla \cdot (\lambda \nabla T - \rho_w c_w \mathbf{q} T) = \phi \rho_w c_w \frac{\partial T}{\partial t} + (1-\phi)\rho_s c_s \frac{\partial T_s}{\partial t} = \rho_m c_m \frac{\partial T}{\partial t} \qquad (2.1.49)$$

where λ is the thermal conductivity tensor, T is temperature, ϕ is porosity, ρ_w is water density, c_w is specific heat of the water, ρ_s and c_s are the mass per unit volume and specific heat of the solid, respectively, and ρ_m and c_m are the mass per unit volume and specific heat of the whole porous medium (water plus solid) so that $\rho_m c_m = \phi \rho_w c_w + (1-\phi)\rho_s c_s$.

The tensor of thermal conductivity λ combines the isotropic conductivity λ_0 of the porous medium (water plus solid) in the absence of flow and a term for the heat dispersivity linked to the heterogeneity of the velocity, which is a linear function of this velocity. de Marsily (1986) suggests using Darcy's velocity $\mathbf{q}$ multiplied by the volumetric heat capacity of the water $\rho_w c_w$ so that the proportionality coefficient has the dimension of length like the macrodispersivity in the case of dispersion. In the longitudinal and transverse axes linked to the velocity, one has

$$\lambda_L = \lambda_0 + \beta_L \rho_w c_w |q| \qquad \lambda_T = \lambda_0 + \beta_T \rho_w c_w |q| \qquad (2.1.50)$$

where β_L and β_T are the longitudinal and transversal thermal dispersivities, respectively. At other directions, the components of the λ tensor are given by

$$\lambda_{xx} = \lambda_0 + \rho_w c_w \frac{\beta_L q_x^2 + \beta_T q_y^2}{|q|} \tag{2.1.51a}$$

$$\lambda_{yy} = \lambda_0 + \rho_w c_w \frac{\beta_T q_x^2 + \beta_L q_y^2}{|q|} \tag{2.1.51b}$$

$$\lambda_{xy} = \lambda_{yx} = \rho_w c_w (\beta_L - \beta_T) \frac{q_x q_y}{|q|} \tag{2.1.51c}$$

2.2. CHEMICAL REACTIONS

2.2.1. MATHEMATICAL FORMULATION OF CHEMICAL REACTIONS

A chemical system is made up of a set of atomic constituents or elements. A chemical species is defined as any chemical entity distinguishable from the rest due to (1) its elemental composition, and (2) by the phase at which it is present. For instance, gaseous CO_2 is a different species than dissolved CO_2.

Not all species are needed to describe fully the chemical system. The subset of species which is strictly necessary is made up of what are known as components. These components can be chosen arbitrarily among all species. Although the N_E atomic constituents could serve as a set of components, they are never used as such because the constituents themselves are rarely present in aqueous phases. For this reason, it is more convenient to select as components a subset of N_C chemical species. These species are also known as the primary species.

INVERSE-CORE^{2D} allows defining a set of aqueous species in the database as a primary species. The code can check that this set of species are initially independent and are able to form together with the sorbed primary species the rest of species (such as secondary species or aqueous complexes, minerals, gases and surface complexes) as a linear combination of primary species (Samper et al., 1995).

$$Q_j^s = \sum_{i=1}^{N_C} \nu_{ji} Q_i^P \qquad j = 1....N_S \qquad (2.2.1)$$

where Q_j^s and Q_i^p are the chemical formulae of the j-th secondary species and the i-th primary species respectively; ν_{ij} is the stoichiometric coefficient of the i-th primary species in the dissociation reaction of the j-th species and N_s is the total number of secondary species.

2.2.2. CHEMICAL EQUILIBRIUM

At a given pressure and temperature, the Gibbs free energy of the system reaches a minimum at local equilibrium, and the system cannot spontaneously carry out any chemical work. The principle of chemical equilibrium leads to the well known Mass-Action law expression (Denbigh, 1987):

$$a_j^{-1} \prod_{i=1}^{N_T} a_i^{\nu_{ji}} = K_{j(p,T)} \qquad (2.2.2)$$

where K_j is the equilibrium constant which depends on the pressure and temperature of the system, a_i and a_j are the thermodynamic activities of the i-th and j-th species, respectively, N_T is the total number of chemical species and ν_{ji} is the stoichiometric coefficient of the i-th primary species in the dissociation reaction of the j-th species. As a convention the stoichiometric coefficients are positive for species on the right hand side of the reaction, and negative for those on the left hand side. The reactions are always written as the dissociation of one mole of secondary species.

2.2.3. AQUEOUS COMPLEXATION REACTIONS

The continuous motion of dissolved ions together with their large number per unit volume cause numerous collisions making possible the formation of ion pairs and/or dissolved complexes which usually have an ephimerous live (on the order of 10^{-10} s). Since these reactions are almost instantaneous, they can be effectively considered as equilibrium reactions. The equilibrium constant relates the average number of ions pairs or complexes which are being formed. Applying the Mass-Action Law to the dissociation of the j-th secondary species, one has:

$$K_j = a_j^{-1} \prod_{i=1}^{N_c} a_i^{\nu_{ji}} \tag{2.2.3}$$

where a_i and a_j are the thermodynamic activities of the species j and i, respectively, and ν_{ji} is the stoichiometric coefficient of the i-th primary species on the j-th species. This equation allows one expressing the concentration of secondary species x_j in terms of primary species concentrations c_i according to:

$$x_j = K_j^{-1} \gamma_j^{-1} \prod_{i=1}^{N_c} c_i^{\nu_{ji}} \gamma_i^{\nu_{ji}} \tag{2.2.4}$$

where x_j and c_i are molal concentrations (mol/kg water) and γ_j and γ_i are thermodynamic activity coefficients.

2.2.3.1. Activity coefficients of aqueous species

For non-concentrated solutions (less than 1 mol/kg), the activity coefficient of the i-th aqueous species can be calculated according to the extended Debye-Hückel formula:

$$\log \gamma_i = -\frac{Az_i^2 (I)^{\frac{1}{2}}}{1 + Ba_i (I)^{\frac{1}{2}}} + bI \qquad (2.2.5)$$

where I is the ionic strength of the solution; z_i and a_i are the electric charge and the ionic radius in solution of the i-th species, respectively; A and B are constants which depend on temperature and dielectric constant of water, and b is a constant determined from fitting experimental data. The values of A, B and b at different temperatures are tabulated in Helgeson and Kirkham (1974). The value of the ionic strength is calculated as:

$$I = \frac{1}{2} \sum_{i=1}^{N_T} c_i z_i^2 \qquad (2.2.6)$$

The activity of water can be calculated according to the approximation of Garrels and Christ (1965):

$$a_{H_2O} = 1 - 0.018 \sum_{i=2}^{N_T} c_i \qquad (2.2.7)$$

where i includes all the dissolved species except water. The expressions of log γ_i and a_{H_2O} are valid only for diluted solutions ($I \leq 0.1$ M).

2.2.3.2. Total solute concentration

Total dissolved concentration of a given component, C_k, can be written in an explicit form as a function of the concentration of the N_C primary species:

$$C_k = c_k + \sum_{j=1}^{N_x} \nu_{jk} x_j = c_k + \sum_{j=1}^{N_x} \nu_{jk} \left(\mathbf{K}_j^{-1} \gamma_j^{-1} \prod_{i=1}^{N_c} c_{ji}^{\nu_{ji}} \gamma^{\nu_{ji}} \right) \qquad (2.2.8)$$

where N_x is the number of secondary species. Notice the difference between the concentration of primary species c_k and the total concentration C_k.

As shown before, the chemical composition of an aqueous system containing N_E species can be expressed in terms of the concentrations of N_C components (primary species). This is of great relevance for reactive solute transport modeling because instead of N_E transport equations only N_C equations have to be solved. The concentration of the N_x secondary species can be explicitly computed from the concentrations of primary species. This results in a significant reduction of computing time.

2.2.4. ACID-BASE REACTIONS

These reactions include those involving the transfer of protons H^+. Proton concentration in the solution is obtained by defining a new variable measuring the proton excess or total proton concentration, C_H, which is defined as:

$$C_H = c_H + \sum_{j=1}^{N_x} \nu_{jH}\, x_j \tag{2.2.9}$$

2.2.5. REDOX REACTIONS

The transfer of electrons between two different atoms changes their chemical valence. This transfer is known as an oxidation-reduction reaction.

The redox potential of a chemical system can be described by means of redox pairs such as O_2/H_2O, SO_4^{2-}/H_2S, Fe^{3+}/Fe^{2+}, CO_2/CH_4, etc. Usually, the redox potential is governed by the most abundant redox pair. Even though this approach seems to be the most adequate, it is rarely used due to the difficulty of obtaining the analytical concentrations of the two species of a redox pair.

Among the different alternative approaches to describe redox reactions the external approach considers hypothetical electron activity as an aqueous component or a master species. Contrary to the protons, which exist in reality as dissolved species, the electron concentration is a hypothetical variable. The definition of this virtual concentration is useful because it allows to complete the redox half-reactions and treat them as the rest of the chemical reactions in the aqueous phase. Each half redox reaction is completed by adding electrons as transferable species. The activity coefficient of this hypothetical species is assumed to be equal to one. It is possible then to define the total electron concentration C_e as:

$$C_e = c_e + \sum_{j=1}^{N_x} \nu_{je} x_j \qquad (2.2.10)$$

where c_e is the "free electron concentration" and ν_{je} is the stoichiometric coefficient of the electron in j-th half-redox reaction.

Similar to acid-base reactions, the total concentration C_e represents the net electron balance or "electron excess" for all redox reactions. Thus, C_e may take on positive or negative values. In this way, redox reactions can be treated in the same way as the rest of equations.

Once the "free electron concentration", C_e, is known, the calculation of other redox indicators, such as the counter-part species of a redox pair, the concentration of dissolved oxygen ($O_{2(aq)}$), or the fugacity of oxygen gas ($O_{2(g)}$), can be easily calculated by considering them as secondary species (Equation 2.2.4.). On the other hand, if the concentration of dissolved oxygen ($O_{2(aq)}$) is known, it also can be taken as a primary species. Then the "free electron concentration" (C_e), as well as pE can be easily calculated too. Redox processes in INVERSE-CORE^{2D} are formulated either using the external approach (with the hypothetical electron activity) or in terms of $O_{2(aq)}$ transfer. Although the user can choose between one of these approaches, the latter is recommended for natural groundwater system.

2.2.6. CATION EXCHANGE

Cation exchange takes place when free cations in solution exchange with interlayer cations. This process can be described as an equilibrium reaction between an exchangeable cation and an exchange site. The equilibrium constant is usually known as the exchange coefficient because its value depends on the ionic strength of the solution. A general expression for cation exchange reactions according to the Gaines-Thomas convention is (Appelo and Postma, 1993):

$$\frac{1}{\nu_i}S_i + \frac{1}{\nu_j}\left(X_{\nu_j} - S_j\right) \Leftrightarrow \frac{1}{\nu_i}\left(X_{\nu_i} - S_i\right) + \frac{1}{\nu_j}S_j \qquad (2.2.11)$$

where ν_i and ν_j are the stoichiometric coefficients (equal to their charges) of dissolved and interlayer cations, respectively; S_i and S_j denote dissolved cationic species and $\left(X_{\nu_i} - S_i\right)$ and $\left(X_{\nu_j} - S_j\right)$ represent exchange sites or exchange interlayer cations. The equilibrium equation for cation exchange is obtained from the Mass-Action Law:

$$K_{ij}^{*} = \frac{\overline{w}_i^{1/\nu_i} \cdot a_j^{1/\nu_j}}{\overline{w}_j^{1/\nu_j} \cdot a_i^{1/\nu_i}} \qquad (2.2.12)$$

where K_{ij}^{*} is the exchange coefficient or selectivity, a_j is the activity of the j-th dissolved species and $\overline{w}_i$ is the activity of the i-th exchanged species. Activities of dissolved cations are related to concentrations according to the Debye-Hückel theory as described in Section 2.2.3.1. Activities of interlayer cations are approximated by their equivalent fractions of the number of exchange sites. Thus, the activity of the interlayer cation $\overline{w}_i$ is assumed to be equal to its equivalent fraction β_i, and is calculated as:

$$\overline{w_i} \cong \beta_i = \frac{w_i}{\sum_{i=1}^{N_w} w_i} \tag{2.2.13}$$

where w_i is the concentration of the i-th interlayer cation and N_w is the total number of such interlayer cations. The sum of concentrations of surface sites or interlayer cations is the so-called cation exchange capacity (CEC). Substituting Equation (2.2.13) into (2.2.12) yields the general equation for cation exchange:

$$K_{ij}^{*} = \frac{\beta_i^{1/\nu_i} \cdot (c_j\gamma_j)^{1/\nu_j}}{\beta_j^{1/\nu_j} \cdot (c_i\gamma_i)^{1/\nu_i}} \tag{2.2.14}$$

where the activity of each dissolved species a_i has been expressed as the product of its concentration c_i time its activity coefficients γ_i. From this equation, the equivalent fraction of the j-th interlayer cation can be expressed as

$$\beta_j = \left(K_{ij}^{*}\right)^{-\nu_j} c_j\gamma_j \left(\frac{\beta_i}{c_i\gamma_i}\right)^{\nu_j/\nu_i} \qquad j = 1,2,\Lambda\,,N_w \tag{2.2.15}$$

From the definition of equivalent fraction, one has:

$$\sum_{j=1}^{N_w} \beta_j = 1 \tag{2.2.16}$$

Substituting Equation (2.2.15) into (2.2.16) results in:

$$\sum_{j=1}^{N_w} \left(K_{ij}^{*}\right)^{-\nu_j} c_j\gamma_j \left(\frac{\beta_i}{c_i\gamma_i}\right)^{\nu_j/\nu_i} = 1 \tag{2.2.17}$$

Once the equivalent fraction β_i is known, the rest of exchange fractions can be calculated from Equation (2.2.15). In general, selectivity coefficients are not constant and depend on equivalent fractions. They must be calculated from experimental exchange isotherms. These isotherms are used to define polynomial expressions relating the Neperian logarithm of selectivity coefficients to equivalent fractions:

$$\ln K_{ij}^{*} = a + b\beta_i + c\beta_i^2 + d\beta_i^3$$

where a, b, c and d are coefficients derived from exchange isotherms (Huertas et al., 2000). In this case, Equation 2.2.17 must be solved with a Newton-Raphson iterative method.

The concentration of the i-th exchanged cation w_j (in moles per liter of fluid) can be obtained from the i-th equivalent fraction according to Equation (2.2.13) through:

$$w_j = \beta_j \, CEC \, \rho_s \, z_j \frac{(1-\phi)}{100\phi} \qquad (2.2.18)$$

where CEC is the cation exchange capacity (usually expressed as meq of cations per 100 grams of solid), ϕ is porosity, ρ_s is density of solids (Kg of solids per dm^3 of solids) and z_j is cation charge.

There are other different conventions for cation exchange which are also included in INVERSE-CORE^{2D}, such as that of Gapon in which the activities of exchanged species are taken equal to their molar fractions (see Appelo and Postma, 1993).

2.2.7. ADSORPTION

Many minerals such as metal oxides, hydroxides and layered silicates (Al_2O_3, TiO_2, FeOOH, SiO_2, kaolinite, etc.) exhibit electrically charged surfaces in the presence of natural waters. These surfaces contain ionizable functional groups such as silanol groups in hydrated silica (Si-OH) which are responsible for chemical reactions at the surface.

2.2.7.1. The surface electric potential

The sorption of solutes at the solid surfaces can be described by means of chemical reactions between aqueous species and surface sites (surface complexation). These surface reactions include proton exchange, cation binding and anion binding via ligand exchange at surface hydroxil sites (represented here as XOH to avoid confusion with other chemical species). For example, the sorption of a metal could be represented as:

$$XOH + M^{z+} = XOM^{(Z-1)+} + H^+$$

At equilibrium the sorption reactions can be described by:

$$K_{intr} = \frac{[XOM^{(Z-1)+}]a_{H^+}}{[XOH]a_{M^{z+}}} \qquad (2.2.19)$$

where K_{intr} is the equilibrium constant related to the chemical reaction, usually referred to as the intrinsic constant (Dzombak and Morel, 1990), a denotes thermodynamic activity of aqueous species and terms within brackets represent the concentration of surface complexes (mol/kg).

In aqueous complexation reactions the electric charge is assumed to be homogeneously distributed in the solution. However, surface reactions take place on a fixed charged surface which creates an electrostatic field. An additional energetic term is required, accounting for the work needed for the aqueous species to travel across the surface electric field:

$$K_{ads} = K_{intr} e^{\frac{\Delta z F \psi_0}{RT}} \qquad (2.2.20)$$

where R is the gas constant (8.314 J/mol/K) and T the absolute temperature (K), K_{ads} is the apparent equilibrium constant of the overall adsorption process, and Δz is the change of the reaction. K_{intr} does not depend on the surface charge. Although the contribution of the electrostatic term to the total sorption energy is not determined experimentally, it is useful to take into account the variations of surface charge effects on surface complexation reactions. In Equation (2.2.20) the thermodynamic activities of the surface complexes are substituted by their concentrations. This is based on the assumption that all nonideal behavior of the surface species is attributed to the mean surface potential and is accounted for by the electrostatic term.

By considering the electrostatic term in the equilibrium formulation, a new unknown, the mean surface potential, ψ_0, is introduced. A new equation to find the solution of the problem is, therefore, required.

There are several theoretical models which describe the value of the capacitance of the surface, i.e., the relationship between the mean surface potential and the charge density. In increasing order of complexity, these models are: the constant capacitance model, the diffuse double layer model and the triple layer model. A diffuse layer model is implemented in INVERSE-CORE^{2D}.

2.2.7.2. Diffuse layer model

The diffuse layer model has been fully described by Dzombak and Morel (1990) and applied to adsorption of metals on iron oxide surfaces. In this model the solid-water interface is composed of two layers: a layer of surface-binded complexes and a diffuse layer of counter ions in solution. All sorbed ions are assigned to a surface layer, whose surface charge is calculated according to:

$$\sigma_0 = \frac{F}{A} \sum_{k=1}^{N_y} z_k y_k \qquad (2.2.21)$$

where F is the Faraday constant (C/mol), A is the specific surface area (dm^2/dm^3 solution), z_k and y_k are the charge and the concentration of the k-th adsorbed species (mol/dm^3), respectively; and N_y is the number of adsorbed species.

The relationship between the surface charge and the potential can be calculated from the Guy-Chapman theory.

$$\sigma_d = (8RT\varepsilon\varepsilon_0 I)^{1/2} \sinh(F\psi/2RT) \qquad (2.2.22)$$

where ε is the dielectric constant of water, ε_0 the permitivity of free space ($8.854x10^{-13}$ C/V/dm), and I the ionic strength of the solution (mol/dm^3).

For low values of the potential Ψ, Equation (2.2.22) can be linearized as

$$\sigma_d = \varepsilon\varepsilon_0 \kappa\psi \qquad (2.2.23)$$

where 1/k (dm) is the double-layer thickness which is given by:

$$\frac{1}{\kappa} = \left(\frac{\varepsilon\varepsilon_0 RT}{20 F^2 I}\right)^{\frac{1}{2}} \qquad (2.2.24)$$

Therefore, in the difuse-layer model the value of the capacitance C relating the surface charge and the potential can be calculated based on theoretical considerations instead of being an experimental fitting parameter.

2.2.7.3. Mathematical formulation of adsorption reactions

Lets us consider the following surface desorption reaction:

$$XOM^{(Z-1)+} = XOH + M^{Z+} - H^{+}$$

According to Equations (2.2.19) and (2.2.20), the equilibrium constant of the reaction will be:

$$K_{ads} = \frac{[XOH]a_{M^{z+}}}{[XOM^{(Z-1)+}]a_{H^{+}}} e^{\left(\frac{F\psi_0}{RT}\right)} \qquad (2.2.25)$$

In order to represent the adsorbed species, an additional set of N_S sorbed primary species has to be added to the initial set of N_C aqueous primary species. The new primary species are (for $N_S = 2$):

$$s_1 = C_{XOH}$$

$$s_2 = \exp(-F\psi_0 / RT)$$

The concentration of a surface complex, y_j (mol/dm^3) can be expressed in terms of the concentration of the two sets of components according to:

$$y_j = K_j^{-1} \prod_{i=1}^{N_C} c_i^{\nu_{ji}^{y}} \gamma_i^{\nu_{ji}^{y}} \prod_{k=1}^{N_S} s_k^{\nu_{jk}^{s}} \qquad j = 1...N_Y \qquad (2.2.26)$$

where K_j is the equilibrium constant of the desorption reaction, N_Y the number of surface complexes, and ν_{ji}^{s} is the stoichiometric coefficient of the j-th complex on the i-th additional sorbed primary species, which takes the following values:

$$\nu_{j1}^{s} = \nu_{j,XOH}^{s}$$

$$\nu_{j2}^{s} = z_{j}^{0}$$

where z_j denotes the electrostatic charge of the surface complex.

2.2.7.4. Adsorption via distribution coefficients

Sorption can be modeled using the K_d approach which assumes that there is a fast and instantaneous chemical equilibrium between dissolved and sorbed species. In addition, the concentrations of sorbed, s, and dissolved species, c, are assumed to be related through the distribution coefficient K_d according to

$$s = K_d\ c \tag{2.2.27}$$

In this case, sorption can be included directly into the transport equation by introducing the retardation coefficient R_e by means of:

$$\nabla \cdot \left(\theta\ b\mathbf{D}\nabla c\right) - b\mathbf{q}\nabla c + r\left(c^{*} - c\right) + b\theta\, R = b\theta\, R_e \frac{\partial c}{\partial t} \tag{2.2.28}$$

where R_e is given by:

$$R_e = \left(1 + \frac{\rho_d K_d}{\theta}\right) \tag{2.2.29}$$

where ρ_d is the dry density of porous media which is related to density of the solids ρ_r through:

$$\rho_d = \rho_r (1 - \phi) \tag{2.2.30}$$

If a species suffers simultaneously linear sorption and exclusion, the retardation coefficient is given by:

$$R_{ea} = \left(1 + \frac{\rho_d K_d}{\theta_a}\right) \tag{2.2.31}$$

2.2.8. DISSOLUTION-PRECIPITATION REACTIONS

Under equilibrium conditions, dissolution-precipitation reactions can be described by the Law of Mass Action which states that

$$X_m \lambda_m K_m = \prod_{i=1}^{N_c} c_i^{\nu_{mi}} \gamma^{\nu_{mi}} \tag{2.2.32}$$

where X_m is the molar fraction of the m-th solid phase, λ_m is its thermodynamic activity coefficient (X_m and λ_m are taken equal to 1 for pure phases), c_i and γ_i are the concentration and activity coeficient of the i-th species, ν^p_{mi} its stoichiometric coefficient in the dissolution reaction of the m-th solid phase, and K_m is the corresponding equilibrium constant.

The equilibrium condition provides a relationship among the concentrations of the involved aqueous species. The mass transfer needed to achieve this condition is not specified. In fact, the equation does not include the concentration of the m-th solid phase, and therefore the amount of dissolved/precipitated mineral cannot be computed explicitly. This is a unique feature of these reactions.

2.2.9. REACTIONS WITH OTHER FLUID PHASES.

For reactions involving species in aqueous and gaseous phases, the Mass Action law states that

$$p_f \Gamma_f K_f = \prod_{i=1}^{N_c} c_i^{\nu_{fi}} \gamma_i^{\nu_{fi}} \qquad (2.2.33)$$

where p_f is the partial pressure of the f-th species in the gaseous phase, Γ_f is its activity coefficient, c_i and γ_i are the concentration and activity coefficient of the i-th dissolved primary species, respectively, ν_{fi} is the stoichiometric coefficient of the f-th gas on the i-th species and K_f is the equilibrium constant of the reaction. For low pressures (in the range of atmospheric pressure), the gaseous phase behaves like an ideal mixture. Under these conditions, the activity coefficient Γ_f is equal to 1 and the partial pressure p_f is equal to the total pressure times the molar fraction X_f. The partial pressure p_f can also be computed as $p_f = [f]RT$, where [f] is the molar concentration of the f-th species in the gaseous phase.

2.2.10. NUMBER OF PHASES IN THE SYSTEM

The description of the chemical system requires knowing the number of phases which besides the aqueous phase include all possible minerals. Some minerals are known to be initially present. Others, however, may form due to the hydrochemical evolution of the system. The formation of the latter must obey the principle of minimum free energy G. The change of the free energy G with the progress of the reaction ξ_m is related to the saturation ratio Ω_m through

$$\frac{\partial G}{\partial \xi_m} = 2.303 RTS_m \tag{2.2.34}$$

where the saturation ratio S_m is the ratio between the ion activity product and the solubility constant:

$$\Omega_m = \frac{1}{K_m} \prod_{i=1}^{N_C} c_i^{\nu^P_{mi}} \gamma_i^{\nu^P_{mi}} \tag{2.2.35}$$

The $\log_{10}$ of Ω_m is known as the saturation index SI_m. The system reaches the minimum free energy at equilibrium when $\Omega_m = 1 (\text{or } SI_m = 0)$. Saturation ratios Ω_m greater than 1 ($SI_m > 0$) indicate that the system will evolve so that the m-th mineral precipitates. On the contrary, when $\Omega_m < 1 (SI_m < 0)$ the mineral will tend to dissolve. This means that mineral phases that can participate are those already present initially plus other potential minerals having Ω_m values greater than 1. However, the phase rule states that at given pressure and temperature the number of phases cannot exceed the number of components N_c. In such cases, only the minerals having the largest Ω_m values are retained as phases. Similarly, when a mineral being dissolved is exhausted it must be excluded from the list of mineral phases.

2.2.11. KINETICS OF DISSOLUTION-PRECIPITATION

When dissolution-precipitation reactions are slow compared to other relevant transport and geochemical processes they have to be analyzed kinetically. In this case, these reactions cannot be assumed to be at local equilibrium.

The general kinetic rate expression of mineral dissolution/precipitation is given by:

$$r_m = \varsigma_m e^{-\frac{Ea_m}{RT}} \sum_{K=1}^{N_K} K_{mk} \prod_{i=1}^{N_C+N_X} a_i^{p_{mki}} (\Omega_m^{\theta_{mk}} - 1)^{\eta_{mk}} \tag{2.2.36}$$

where:

r_m: is the dissolution/precipitation rate (mol/m^2/ s).

ς_m: is a function which can take a value of 1 or −1 depending on whether the saturation index defined in Eq. (2.2.35) is greater or smaller than 1. It is equal to 1 for dissolution and −1 for precipitation. At equilibrium (Ω_m=1), and r_m=0

Ω_m: is the ratio between the ion activity product and the equilibrium constant (as defined in Eq. 2.2.35)

$e^{-\frac{Ea_m}{RT}}$: is a thermodynamic factor, which takes into account the apparent activation energy of the overall reaction Ea_m (KJ/mol). R (KJ/mol·K) and T (K) are the gas constant and absolute temperature, respectively.

N_k: is the number of kinetic terms for the m-th mineral.

K_{mk}: is the kinetic rate constant of the k-th kinetic term of the m-th mineral (mol/m^2/ s).

$\prod_{i=1}^{N_C+N_X} a_i^{p_{mki}}$: is a factor which accounts for the catalytic effect of some species (usually for H^+). If the catalytic effect is disregarded this term is equal to 1, $p^{mki} \equiv 0$ where p^{mki} is a dimensionless exponent for the i-th aqueous species for the m-th mineral in the k-th kinetic term, which is usually determined from experimental data and a_i is the activity of the i-th aqueous species.

N_C, N_X : are the number of primary and secondary species, respectively.

θ_{mk}, η_{mk} : are exponents for the k-th kinetic term for the m-th mineral, which are usually determined from experimental data.

For the i-th primary aqueous species, the source/sink term r_i (mol/m^3/s) associated to kinetically-controlled reactions is given by:

$$r_i = \sum_{m=1}^{N_p} \nu_{mi}^p r_m A_m \qquad (2.2.37)$$

where r_m is the mineral dissolution/precipitation rate (mol/m^2/s) of the m-th mineral as defined in Eq. 2.2.36 and A_m is the specific surface of the m-th mineral (usually expressed as the surface of mineral per unit volume of medium.

2.3. NUMERICAL SOLUTION OF THE FORWARD PROBLEM

A 2-D Galerkin finite element method is used to solve groundwater flow, solute and heat transport equations. Coupled water flow and reactive transport equations are solved in the following manner (Xu, 1996):

(1) Solution of flow equation. It is solved in terms of hydraulic heads for flow in saturated media (confined or unconfined aquifers). Unconfined aquifer flow equation is solved iteratively using a predictor-corrector scheme. Flow in variably saturated media is solved in terms of pressure heads by a Newton-Raphson iterative method. For steady-state flow, the equation is solved only at the first time step.

(2) Computation of water velocity. Water velocities, which are needed to evaluate advective and dispersive solute and heat fluxes, are computed from nodal head values by direct application of Darcy's Law to the finite element solution.

(3) Solution of heat transport. The solution of the heat transport equation is used to update temperature-dependent chemical parameters (chemical equilibrium constants, parameters for the computation of activity coefficients). The solution of heat transport shares some subroutines of the solute transport solution because the structure of both equations is similar.

(4) Solution of solute transport equations. A transport equation is solved for each chemical component in terms of its total dissolved concentration C. Since chemical

sink/source terms are assumed known (they are evaluated at the previous iteration), each transport equation can be solved separately. In addition, transport matrices are the same for all chemical components because it is assumed that all components have the same dispersion tensor. Only the right-hand-side independent terms change from one component to another. A band LU solver is used because only one decomposition is needed at each time step. This decomposition (which takes up much of the computing time required by a LU solver) is carried out only for the first component. After LU decomposition, only a forwards and a backwards substitution are needed for each component at each iteration. In addition, when the flow is steady and the time step remains constant, the transport coefficient matrix does not change and therefore the LU decomposition is performed only once at the first time step and for the first chemical component. This leads to a substantial saving of computational time for solving solute transport equations.

(5) Solution of hydrochemical equations. The following chemical reactions are included: acid-base, aqueous complexation, redox, mineral dissolution/ precipitation, gas dissolution/ex-solution, cation exchange and adsorption. There is a mass balance equation for each chemical component at each node. Total analytical concentrations T are the primary dependent variables in the chemical submodel. A Newton-Rapson method is used to solve the chemical equilibrium equations. Once the concentrations of primary species are obtained, the concentrations of secondary species, the amount of minerals dissolved or precipitated and the concentrations of sorbed and exchanged species are computed in a straightforward manner. Hydrochemical reactions are solved in a node-wise manner.

The transport equation of each chemical component has a chemical source/sink term which represents the mass transfer between aqueous and other phases (by mineral dissolution/precipitation, cation exchange and adsorption processes). These source terms render the problem highly non-linear. An iterative procedure between steps (4) and (5) is required. In the beginning of the iterative process, it is assumed that the reaction source term is equal to zero. Then, transport equations are solved. The resulting dissolved

concentrations are substituted into the chemical submodel to get the mass transfer between aqueous and solid phases which are used to update the chemical source and the transport equations are solved again. This process is repeated until some prescribed convergence criteria are satisfied.

CHAPTER 3

INVERSE MODELING OF REACTIVE SOLUTE TRANSPORT

3.1. BASIC CONCEPTS

3.1.1. MODEL STRUCTURE AND MODEL PARAMETERS

A model is a simplified representation of a real system. A groundwater flow and reactive transport model aims at simulating water flow and reactive solute transport through a geological formation. Once a simulation model is established for a reactive solute transport system, the forward numerical prediction problem can be solved, i.e., one can forecast the response (states **u**) or performance of the system for different excitations or under variable conditions. As a result, different management decisions can be compared and an optimal decision can be selected based on specified criteria. A requirement for the application of this procedure to real problems is that the input-output relation of the simulation model must be consistent with the excitation-response relation of the original system. A good simulation model should meet the following requirements:

- Model structure should be a suitable description of the system.
- The forward numerical model should provide an accurate solution for the mathematical model.
- Model parameters $\mathbf{P} = (p_1, p_2, \ldots, p_M)$ should be exactly known.

In practice, however, it is very difficult to construct an accurate simulation model for a flow and reactive solute transport system. First, the governing equations of the constructed model may not be an exact description of the original system. Second, the properties of the system, such as geometry of the model region, flow and reactive transport parameters, initial and boundary conditions, sink and source terms, are difficult to measure accurately in the field.

Fortunately, the state variables of a flow and reactive solute transport system can be easily measured. These measurements usually include water content, hydraulic head, solute concentration, total concentration (including liquid and solid phases) and cumulative water flux. These data can be collected either from historical records or from laboratory or field experiments. The input-output relation of a correct simulation model should reproduce the observed responses. Therefore, it is possible to determine indirectly the properties of the system by using measured data (Sun, 1994).

3.1.2. MODEL CALIBRATION AND INVERSE PROBLEM

Model calibration involves adjustments of the structure and parameters of a numerical model simultaneously or sequentially so as to make the input-output relation of the model fit observed excitation-response data of the real system (Sun, 1994). This process can be performed by trial and error or using automatic methods.

If the model structure is known, the problem of determining model parameters from the observed system states and other available information is called parameter identification. In a certain sense, parameter identification is an inverse of the forward problem (FP). System state **u** or the output of FP now becomes the input, while model parameter **P** becomes the output. Therefore, parameter identification problems are often called inverse problems (IP).

During the calibration or inverse analysis, model selection is very important and usually performed using qualitative criteria (Samper et al., 1998). Conceptual models that fail to reproduce measured data are regarded as inappropriate. Quite often, this allows one to be left with a single model. However, in some cases, several models match equally well the observed data. In such cases, the uncertainty in the conceptual models must be carried out into the prediction process. The parameters obtained during calibration are uncertain. This uncertainty can be taken into account by perturbing the values of the parameters with respect to the best estimates and examining how the simulated predictions change. If the simulated results are acceptably close to their measured counterparts, then the

corresponding parameter perturbation should be accepted as feasible. Otherwise, it should be rejected. Uncertainty can be evaluated in a more systematic manner, usually by means of a posteriori statistical analyses such as variance-covariance matrix, correlation matrix, eigenvalues and eigenvectors of the covariance matrix, as well as confidence intervals of the parameters. Various approaches were reviewed by Yeh (1986) and Carrera and Neuman (1986a).

Once the model is calibrated, predictions are made by appropriately changing boundary conditions and internal sinks and sources in order to reproduce the various hypotheses about the future. Therefore, the solution of inverse problems is indeed a key step in flow and reactive transport modeling. Unfortunately, there are some difficulties related to inverse problems which will be discussed in the following section.

3.1.3. ILL-POSEDNESS, NON-UNIQUENESS AND PARAMETER IDENTIFIABILITY

Intuitively, the existence of an inverse solution appears to be no problem at all, since the physical reality must be a solution (Sun, 1994). In practice, however, the observation error of state variables and the structure error of the model cannot be avoided. As a result, an accurate solution of an inverse problem may not exist. However, existence is not a major difficulty in the solution of inverse problems because generally there is an approximate solution to the physical reality. The major difficulty of the inverse problems is ill-posedness.

An inverse problem is ill-posed when it has either no solution at all, no unique solution, or the solution is unstable (Kool, et al., 1989). Non-uniqueness occurs when there are multiple parameter vectors that correspond to extrema of roughly equal magnitude in the objective functions, making it impossible to determine the correct solution. As Yeh (1986) pointed out, the uniqueness problem is closely related to the notion of parameter identifiability, i.e. whether or not it is possible to obtain accurate estimates of the parameter in the mathematical model from available data. Identifiability thus depends on both the assumed model and the experimental data used. When parameters are highly correlated, a change in one parameter is balanced by a corresponding change in the correlated parameter,

with the result that neither can be determined accurately. On the other hand, even when model parameters are completely independent, available data may lead to an objective function that lacks sensitivity to one or more of the parameters, again with the result that these parameters will have large estimation variances. Instability, on the other hand, occurs when the estimated parameters are excessively sensitive to changes in data. Relatively small measurement errors can then lead to significant errors in estimated parameter values.

3.2. FORMULATION OF COUPLED INVERSE PROBLEMS

3.2.1. WEIGHTED LEAST SQUARE OBJECTIVE FUNCTIONS

The objective function is one of the most important parts of the inverse problems. Carrera (1984), Yeh (1986), Carrera and Neuman (1986a, b), Samper et al. (1990), Sun (1994) and Samper and García-Vera (1998) made widely review of the parameter estimation methods, as well as the corresponding objective functions.

The most common tool for parameter identification is the generalized output least squares (OLS) criterion (Neuman and Yokowitz, 1979; Clifton and Neuman, 1982; Sun, 1994). Let $\mathbf{P} = (p_1, p_2, p_3, ..., p_M)$ be the unknown parameter vector defined by a parameterization method, then the least squares criterion E(p) can be expressed as,

$$\begin{aligned} E(p) &= \sum_{i=1}^{NE} W_i E_i(p) \\ &= \sum_{i=1}^{NE} W_i \sum_{l=1}^{L_i} w_l^2 [u_l(p) - F_l]^2 \\ &= \sum_{i=1}^{NE} W_i \sum_{l=1}^{L_i} w_l^2 r_l^{\,2}(p) \end{aligned} \qquad (3.2.1)$$

where NE is the number of different types of data such as hydraulic head, dissolved concentration, total concentration, water flux, water content and prior information of the parameters to be estimated; $E_i(p)$ is the least-squares for each type of data (i = 1 for

hydraulic heads, i = 2 for dissolved concentrations, ...); W_i are weighting coefficients for each type of data and prior information which are defined according to Equation (3.2.2); w_l are weighting coefficients for different observation points. Their values depend on the accuracy of observations and are generally equal to 1; $\mathbf{u}_l(p)$ is the value of the variable computed with the numerical model; F_l are observed values and prior information of the parameters; L_i is the number of observation points in the space-time domain; and $\mathbf{r}_l$ are residuals or differences between model outcomes and actual field or laboratory measurements.

Equation (3.2.1) can be solved by different optimization methods such as Newton's method, Gauss-Newton method, modified Gauss-Newton method (Yeh and Yoon, 1976; Sun, 1994), conjugate gradient method (Carrera and Neuman, 1986a, b), and Golden section search method (Sun, 1994).

3.2.2. OBSERVATION DEFINITION AND WEIGHTING

3.2.2.1. Weighting of different types of data

Equation (3.2.1) is a weighted multiobjective optimization criterion. It is not a trivial task to designate optimal weighting coefficients. Several methods were proposed by Neuman and Yakowitz (1979) and Carrera and Neuman (1986). Generally, the determination of optimal weights should be an iterative procedure. Carrera and Neuman (1986) and Sun (1994) propose the following equation

$$W_i = \frac{W_{0i}}{\frac{E_i(p)}{L_i}} \qquad (i = 1, 2, ..., NE) \qquad (3.2.2)$$

where W_{0i} are user-defined initial weighting coefficients of different types of observation data. Generally, they are equal to 1. However, from the point of view of optimization, the selection of W_i depends upon the preference of a modeler by using different values of W_{0i} based on the importance of different types of observation data. It should be noticed that W_{0i} are dimensionless. However, W_i have dimensions which are reciprocal to those of $E_i(p)$

which are generally the square of the units of each type of data (m^2 for heads, $(mol/l)^2$ for dissolved concentrations, ...)

Different types of observation data are used in a generalized least squares criterion which incorporates a set of weighting coefficients. Once the initial weights W_{0i} (i = 1, 2, ..., NE) are given, the optimal weights are updated automatically during the iterative process according to Equation (3.2.2). It should be noticed that If NE = 1, $W = W_0$. Weights w_l for different observation points in (3.2.1) must be specified in advance. If all data points are equally reliable, $w_l = 1$ for all l = 1, 2, ..., L_i. If some data are judged to be unreliable, they are assigned smaller weights in order to prevent their pernicious effect on the optimization process.

3.2.2.2. Weighting of concentrations

The group of concentration data requires a different treatment because concentration of different chemical species usually vary within different ranges. For this reason the structure of the weights w_l for concentration is slightly than those of other types of data (heads, water fluxes, water contents and total concentration). The weight for the l-th data of the k-th species, w_l^k is computed as

$$w_l^k = \frac{w_{0l}}{\sigma_k} \tag{3.2.3a}$$

where, $$\sigma_k = \sqrt{(N-1)^{-1}\sum_{i=1}^{N}(F_i^k - \overline{F_k})^2} \tag{3.2.3b}$$

$$\overline{F_k} = N^{-1}\sum_{i=1}^{N} F_i^k \tag{3.2.3c}$$

where w_{0l} is a user-supplied initial weight for the l-th data point (generally, w_{0l} is equal to 1); F_i^k is the i-th measured datum of the k-th species; N is the number of concentration data of the k-th species; $\overline{F}_k$ is the average measured concentration of the k-th species and σ_k is the standard deviation of measured concentrations of the k-th species. This weighting method ensures that no concentration data are either "drowned" by other information, or dominate the inversion process.

3.2.3. PARAMETER DEFINITION AND PRIOR INFORMATION

3.2.3.1. Parameter definition and terminology

Parameterization is a technique used to simplify the structure of an infinite or $\overline{N}$-dimensional parameter space such that it can be represented approximately by a M-dimensional space ($M << \overline{N}$), where M is the dimension of parameterization. According to Sun (1994), there are more than five methods of parameterization, such as piecewise constant or zonation method, polynomial approximation method, interpolation method, distance weighting method, stochastic field method and geological structure method.

Zonation is the method most commonly used for parameterization, which was applied to an alluvial aquifer by Samper et al. (1990) and to a semiarid evaporitic closed basin of Los Monegros by Samper and García-Vera (1998). In this method, parameters are assumed constant within each parameter zone, which consists of a set of elements. During calibration, the number and the outline of zones are modified to minimize the residual sum of squared differences between computed and measured data.

Identification numbers are given to parameters in order to define and identify them. Sixteen types of parameters can be estimated with the inverse model developed in this dissertation. They include:

1. Total porosity,
2. Saturated hydraulic conductivity (K_x and K_y),
3. Specific yield or storage coefficient,
4. Parameters m, n, and α of the van Genuchten retention curve,
5. Molecular diffusion coefficient,
6. Solute dispersivity,
7. Distribution coefficient K_d,
8. Accessible porosity (it may differ from total porosity when anion exclusion takes place),
9. Initial concentrations of chemical components, and initial pH and pE,

10. Boundary concentrations of chemical components,
11. Cation selectivity coefficients,
12. Cation exchange capacity (CEC),
13. Specific surface of minerals,
14. Apparent activation energy of kinetically-controlled reactions,
15. Rate constants of kinetic reactions, and
16. Exponents of the catalytic and saturation index terms in kinetically-controlled chemical reactions.

These parameters can be changed in different parameter zones or for different chemical reactions, minerals and different chemical components. For example, the identification number of the porosity in the first parameter zone (POR(1)) is 1 and in the 10th zone it (POR(10)) is 10. For the conductivity of the first parameter zone (PK1(1)), the identification number is 41 and in the 10th zone it (PK1(10)) is 50. The identification number of kinetic rate constant of the first mineral at first kinetic reaction is 301. The program is could estimate, if needed, up to 800 model parameters. In practice, one only needs to input the identification number of the parameters to be estimated. The full description of the identification number of all of the parameters that can be estimated by INVERSE-CORE^{2D} is listed the User's Manual (Dai and Samper, 1999) and in Appendix 5.

3.2.3.2. Prior information of parameters

Independent prior information about parameters can be incorporated in the objective function and used to define lower and upper bounds, as well as initial values for the optimization process. This information is very important, for it provides a range of permissible values that a parameter can take within its minimum and maximum during the optimization process. Lower and upper parameter bounds should be chosen wisely to avoid numerical dispersion and instabilities. This is the case for diffusion coefficient, solute dispersivity and kinetic rate constants.

Moreover, incorporating prior information into the objective function improves the structure of the covariance matrix and alleviates the ill-posedness of the inverse problem. In some cases, some parameters can be varied simultaneously with no (or very little) effect in the objective function. This can lead to nonunique parameter estimates. This nonuniqueness sometimes can even lead to numerical instability and failure of the estimation process. However, if something is known about at least one of the members of such a parameter group, this information, if included in the estimation process, may overcome nonuniqueness and provide stability.

Parameter estimates will also be nonunique if there are less observations than parameters. However, the inclusion of prior information, being mathematically equivalent to taking extra measurements, may alter the numerical predominance of parameters over observations and thus provide the system with the ability to supply a unique set of parameter estimates. A set of weights must be given to each piece of prior information. When the number of parameter zones is one, this weight is inversely proportional to the prior information itself. When the number of zones is equal to or greater than 2, this weight is inversely proportional to the standard deviation of the prior information. Weighting coefficients for different types of parameters are different and they are calculated by the same Equation (3.2.3) as used for the observations. In practice, by giving different initial weighting coefficients, one can assign the weights in accordance with the extent to which one wishes each article of prior information to influence the parameter estimation process.

It is sometimes helpful to view the inclusion of prior parameter information in the estimation process as the introduction of one or more "penalty functions". The aim of the estimation process is to lower the objective function defined by Equation (3.2.1) to its minimum possible value; this is done by adjusting parameter values until a set of parameters is found for which the objective function can be lowered no further. If there is no prior information, the objective function is defined solely in terms of the discrepancies between model outcomes and laboratory or field measurements. However with the inclusion of prior information, minimizing the discrepancy between model calculations and experimental measurements is no longer the sole aim of the parameter estimation process.

To the extent that a piece of prior information is not satisfactory, (if it is accurate enough, we would not require to estimate the parameters), the discrepancy between the prior information and currently estimated values of the parameters is multiplied by the squared weight pertaining to that piece of prior information into the objective function. It can, more or less, lead the final estimation results close to the prior information. Therefore, with the improvement of the parameter optimization, one should reduce initial weighting coefficients gradually to limit the influence of prior information on estimated results.

3.2.4. CONVERGENCE CRITERIA

There are several stopping criteria to check the convergence of the iterative process (Sun, 1994).

Criterion 1. The norm of the parameter displacement ($P_{k+1} - P_k$) becomes very small after k+1 times of iterations:

$$\|P_{k+1} - P_k\| < \varepsilon_1 \tag{3.2.4}$$

Criterion 2. The norm of the relative change in parameters between two consecutive iterations is less than a given small positive value, or

$$\left\| \frac{P_{k+1} - P_k}{0.5(P_{k+1} + P_k)} \right\| < \varepsilon_2 \tag{3.2.5}$$

Criterion 3. The value of objective function E(p) has no significant change between two consecutive iterations,

$$\|E(p_{k+1}) - E(p_k)\| < \varepsilon_3 \tag{3.2.6}$$

Criterion 4. The norm of the relative change in objective function between two consecutive iterations is less than a given small positive value, ε_4

$$\left\| \frac{E(p_{k+1}) - E(p_k)}{0.5(E(p_{k+1}) + E(p_k))} \right\| < \varepsilon_4 \tag{3.2.7}$$

In our inverse programs criteria 2 and 4 are used as stopping criteria. The program also stops when the specified maximum number of iterations is reached.

3.3. NUMERICAL METHODS OF THE INVERSE PROGRAMS

3.3.1. TRIAL AND ERROR METHODS

The trial and error procedure is the simplest method for solving an inverse problem. It only requires: (1) some observation data on state variables, (2) a forward modeling program and (3) an expert who is familiar with the conceptual model, the numerical model and the field or experiment conditions of the problem. The expert can provide good initial guesses for unknown model parameters. After analyzing the model output, he or she knows how to modify model parameters in order to improve the fit between observed data and model results. This step is repeated until a fitting is judged to be satisfactory. Otherwise, the model structure is modified for further trial runs.

The trial and error method has at least three advantages (Sun, 1994): (1) it is not necessary to write any new program for inverse problem solutions, (2) it can be used to solve any type of inverse problems (actually, as a method of sensitivity analysis, any model uncertainty can be analyzed by this procedure), and (3) the judgement of an expert can be easily incorporated into the search procedure. This may help to overcome ill-posed problems. For instance, when two sets of model parameters generate similar output of the state variables, the expert can select one according to his or her judgement. These advantages can explain why this primitive method is still extensively used in practice (Anderson and Woessner, 1992).

However, the trial and error procedure is time consuming, especially when the number of unknown parameters is large. It does not provide optimum solutions for inverse problems. Different researchers may obtain different results (Sun, 1994). Moreover, this method fails to provide quantitative error estimates or statistical measures of the quality of estimated parameters. Therefore, it is necessary to develop more sophisticated inverse models.

3.3.2. NON-GRADIENT INVERSE SOLUTION METHODS

3.3.2.1. Golden section search

Generally, search methods define a bracket by three points $a < c < b$, satisfying the condition $E(c) \leq \min[E(a), E(b)]$, where $E(\ )$ is the objective function in (3.2.1). Obviously, the minimum must be located in this bracket. By taking a new point d ($a < c < d < b$) in the bracket, where the function takes a value E(d), one can define a new bracket based on the value of E(d) relative to E(c). If $E(d) > E(c)$, [a, c, d] forms a new bracket, and the interval [d, b] can be cut off without missing the minimum point. Otherwise, [c, d, b] can be taken as the new bracket and interval [a, c] can be left out. This procedure is then repeated until the length of the maintained bracket is so small that a convergence criterion is satisfied. Golden section search is a way to determine the new search points, or,

$$c_1 = a_1 + (1-\tau)(b_1 - a_1)$$
$$d_1 = b_1 - (1-\tau)(b_1 - a_1) \qquad (3.3.1)$$

where $\tau = 0.618$

The main advantage of Golden section search is that it does not require the calculation of derivatives and the number of iterations does not have to be determined in advance. The iterative process can be terminated at any iteration by any criterion. The iterative process terminates either when the interval of uncertainty falls below a given tolerance error, when

the relative changes in objective function values or parameter values are less than the stopping criteria, or when the designed number of iterations is reached.

3.3.2.2. Quadratic interpolation method

The quadratic interpolation method is another way for determining a new search point, in which c is the middle point of the bracket and x is the minimum of a quadratic polynomial defined by three points (a, E(a)), (c, E(c)), (b, E(b)):

$$x = \frac{1}{2}\frac{(b^2 - c^2)E_a + (c^2 - a^{2)}E_b + (a^2 - b^2)E_c}{(b-c)E_a + (c-a)E_b + (a-b)E_c} \tag{3.3.2}$$

where $E_a = E(a)$, $E_b = E(b)$ and $E_c = E(c)$.

The objective function is evaluated at x. One of the outer points a or b must be replaced by x to begin the next iteration in such a way that the three new points are disposed over an interval bracketing a minimum. The iterative process terminates either when the interval of uncertainty falls below a given tolerance error or when the relative changes in objective function or parameter values are less than some specified stopping criteria.

3.3.2.3 Methodology for multiple parameter identification

For multiple parameter identification, Golden section and quadratic interpolation must be repeated in a cyclic manner in order to attain an overall minimum solution. Within two consecutive cycles, the weighting coefficients are modified automatically according to Equation (3.2.2). At a given cycle, parameters are estimated one by one.

The iterative process for a given parameter terminates when the interval of uncertainty falls below a given tolerance error, when the relative change in objective function value is less than some a designed local convergent criterion, or when the number of iterations is greater or equal to the given maximum number of iterations. Then, the algorithm continues with the next parameter. The program terminates once the objective function value is less than a specified overall convergence criterion or when the maximum number of cycles has been reached. Therefore, a set of stable minimum solutions (local minimal or overall minimal) can always be identified. Another advantage of these algorithms is that they do not require the calculation of sensitivity coefficients. These features make them practical algorithms for the inverse analysis of complex reactive transport models involving a few parameters. Although these algorithms have been found to be effective for solving inverse problems with 4 or 5 unknown parameters (Dai and Samper, 1999a), its convergence rate is generally very slow. On the other hand, gradient solution methods, such as the Gauss-Newton-Marquardt method, usually have a faster convergence rate, are more robust and more efficient, especially when the number of parameters is large, for example, greater than 10.

3.3.3. GRADIENT SOLUTION METHODS

3.3.3.1. Newton method

The starting point for specialized non-linear least square algorithms is Newton's method which is designed to achieve quadratic termination with a positive definite Hessian matrix in a single iteration given the first and second derivatives of the function. Let $\mathbf{P}_k$ be the vector of parameter values at the k-th iteration and $\mathbf{D}_k$ a vector of parameter increments. The Taylor series gives the expansion of the gradient of a quadratic function at

$$\mathbf{P}_{k+1} = \mathbf{P}_k + \mathbf{D}_k \tag{3.3.3}$$

as

$$\mathbf{g}_{k+1} = \mathbf{g}(\mathbf{P}_k + \mathbf{D}_k) = \mathbf{g}_k + \mathbf{H}\mathbf{D}_k \tag{3.3.4}$$

where **g** is the gradient vector and **H** is the Hessian matrix. If $\mathbf{P}_{k+1}$ is to be the minimum of the objective function, according to the necessary condition for a minimum, then $\mathbf{g}_{k+1}$ must be zero, then $\mathbf{D}_k$ is given by

$$\mathbf{D}_k = - \mathbf{H}^{-1} \mathbf{g}_k \tag{3.3.5}$$

It is more economical to obtain $\mathbf{D}_k$ by the solution of a system of linear equations than by matrix inversion and so Equation (3.3.5) is better expressed as

$$\mathbf{H}_k \mathbf{D}_k = - \mathbf{g}_k \tag{3.3.6}$$

Updated parameter values are computed as,

$$\mathbf{P}_{k+1} = \mathbf{P}_k - \mathbf{H}_k^{-1} \mathbf{g}_k \tag{3.3.7}$$

Fast convergence is the major advantage of Newton's method. However, there are several disadvantages associated with this method. First, it is not economical to compute accurately the second order Hessian matrix. Furthermore, it is impractical to calculate its inverse in each iteration when the objective function is obtained from a numerical model. Second, when $\mathbf{P}_k$ is not close to the minimum, $\mathbf{H}_k$ may not be positive definite. As a result, there is no guarantee that the value of the objective function keeps decreasing. In other words, the Newton's search sequence (3.3.7) may not be convergent. In practice, the Newton's method is seldom used for parameter identification in groundwater flow and transport modeling (Sun, 1994).

3.3.3.2. Gauss-Newton method

The Gauss-Newton algorithm is specially designed for minimizing least squares objective functions. Most inverse models fall into this category of nonlinear problems. This algorithm is developed from Newton's method. The m-th component of the gradient vector ($\mathbf{g}_m$) of the objective function in (3.2.1) can be easily computed by taking derivatives with respect to $\mathbf{P}_m$ (m = 1, 2, . . ., M)

$$\mathbf{g}_m = 2\sum_{i=1}^{NE} W_i \sum_{l=1}^{L_i} w_l^2 \mathbf{r}_l(\mathrm{p}) \frac{\partial \mathbf{u}_l}{\partial \mathbf{P}_m} \tag{3.3.8}$$

Let **J** be the Jacobian matrix defined as

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \mathbf{u}_1}{\partial \mathrm{p}_1} & \frac{\partial \mathbf{u}_1}{\partial \mathrm{p}_2} & \cdots & \frac{\partial \mathbf{u}_1}{\partial \mathrm{p}_M} \\ \frac{\partial \mathbf{u}_2}{\partial \mathrm{p}_1} & \frac{\partial \mathbf{u}_2}{\partial \mathrm{p}_2} & \cdots & \frac{\partial \mathbf{u}_2}{\partial \mathrm{p}_M} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{u}_{NL}}{\partial \mathrm{p}_1} & \frac{\partial \mathbf{u}_{NL}}{\partial \mathrm{p}_2} & \cdots & \frac{\partial \mathbf{u}_{NL}}{\partial \mathrm{p}_M} \end{bmatrix} \tag{3.3.9}$$

where M is the number of the parameters to be estimated and NL the total number of observations. Then, the gradient vector **g** can be written as

$$\mathbf{g}(\mathrm{p}) = 2\mathbf{J}^{\mathrm{T}} \mathbf{W}\, \mathbf{r}(\mathrm{p}) \tag{3.3.10}$$

where **W** is a (NL×NL) diagonal weighting matrix, its entries being equal to the products of W_i in Equation (3.2.2) and w_l^2 (see Equations 3.2.1 and 3.2.3). The second order derivative matrix **H** can be obtained by taking the derivative of Equation (3.3.8) with respect to $\mathbf{P}_j$,

$$\mathbf{H}_{mj} = 2\sum_{i=1}^{NE} W_i \sum_{l=1}^{L_i} w_l^2 \left(\mathbf{r}_l(\mathrm{p}) \frac{\partial^2 \mathbf{u}_l}{\partial \mathrm{P}_m \partial \mathrm{P}_j} + \frac{\partial \mathbf{u}_l}{\partial \mathrm{P}_m} \frac{\partial \mathbf{u}_l}{\partial \mathrm{P}_j} \right) \tag{3.3.11}$$

Let $\mathbf{T}_l$ be the Hessian matrix of $\mathbf{r}_l(\mathrm{p})$,

$$\mathbf{T}_l(\mathrm{p}) = \nabla^2 \mathbf{r}_l(\mathrm{p}) \tag{3.3.12}$$

Then, the complete Hessian matrix **H**(p) of the objective function can be written as

$$\mathbf{H}(\mathrm{p}) = 2\mathbf{J}^{\mathrm{T}}\mathbf{W}\,\mathbf{J} + 2\sum_{i=1}^{NE} W_i \sum_{l=1}^{L_i} w_l^2 \mathbf{r}_l(\mathrm{p})\, \mathbf{T}_l(\mathrm{p}) \tag{3.3.13}$$

Furthermore, defining the second order residual error **S**(p) as

$$\mathbf{S}(\mathrm{p}) = \sum_{i=1}^{NE} W_i \sum_{l=1}^{L_i} w_l^2 \mathbf{r}_l(\mathrm{p})\, \mathbf{T}_l(\mathrm{p}) \tag{3.3.14}$$

One has:

$$\mathbf{H}(\mathrm{p}) = 2\mathbf{J}^{\mathrm{T}}\mathbf{W}\,\mathbf{J} + 2\mathbf{S}(\mathrm{p}) \tag{3.3.15}$$

By neglecting the term **S**(p) in (3.3.15) together with Equation (3.3.10), the equation of Newton's method (3.3.7) becomes now

$$\mathbf{P}_{k+1} = \mathbf{P}_k - (\mathbf{J}^{\mathrm{T}}\mathbf{W}\,\mathbf{J})^{-1}\mathbf{J}^{\mathrm{T}}\mathbf{W}\,\mathbf{r}(\mathrm{p}) \tag{3.3.16}$$

The basic iteration Equation (3.3.16) is called Gauss-Newton method. The search direction of the Gauss-Newton method is determined only by first order derivatives. When **r**(p) is a linear function, it coincides with the Newton direction because **S**(p) is 0. When **r**(p) is a nonlinear function, the deviation of the Gauss-Newton direction from the Newton direction depends on the residual error **S**(p). Therefore, for small residual problems, the Gauss-Newton sequence will converge rapidly. However, the inverse problem of the flow and reactive transport modeling often belongs to the case of large residuals. In the generation of a Gauss-Newton sequence for inverse solution, the following difficulties are often encountered (Sun, 1994):

a) Matrix $\mathbf{J}^{\mathrm{T}}\mathbf{J}$ is near singular, and Equation (3.3.16) has no solution.
b) The search sequence does not converge, i.e., $E(p_{k+1}) > E(p_k)$ for some k.
c) Parameter values $\mathbf{P}_{k+1}$ computed with (3.3.16) may not be feasible.

In order to avoid these difficulties, it is necessary to modify the Gauss-Newton algorithm. One of the most effective methods is the Marquardt-Levenberg method.

3.3.3.3. Gauss-Newton-Levenberg-Marquardt method

The Levenberg-Marquardt method incorporates a technique to deal with the singularity of matrix $\mathbf{J}^T\mathbf{J}$. It is a robust algorithm for small residual problems. Equation (3.3.16) is modified according to

$$\mathbf{P}_{k+1} = \mathbf{P}_k + \Delta\mathbf{P} = \mathbf{P}_k - (\mathbf{J}^T\mathbf{W}\,\mathbf{J} + \alpha\mathbf{I})^{-1}\mathbf{J}^T\mathbf{W}\,\mathbf{r}(p) \qquad (3.3.17)$$

where $\mathbf{I}$ is the $n \times n$ identity matrix and α is the Marquardt parameter, named after Marquardt (1963), although the use of this parameter was in fact pioneered by Levenberg (1944).

If $\alpha = 0$ in Equation (3.3.17), then $\mathbf{P}$ is the Gauss-Newton vector. When α is given a sufficiently large value, the term $\alpha\mathbf{I}$ dominates over $\mathbf{J}^T\mathbf{J}$, so that the parameter increment is close to $\alpha^{-1}\mathbf{J}^T\mathbf{r}$, which represents an infinitesimal step in the steepest descent direction ($\mathbf{g}$). Marquardt (1963) improved the efficiency of the algorithm by proposing a better strategy for selecting α. It is set initially to some positive value (say, 0.01) and a factor $\beta>1$ (say, 10) by which α may be increased or decreased is set. At the beginning of each iteration, α is reduced by a factor β in an attempt to push the algorithm closer to the Gauss-Newton method. If this fails to give a reduction in the objective function, α is repeatedly increased by factor β until a reduction in such function is obtained. By comparing parameter changes and objective function improvement achieved in the current iteration with those achieved in previous iterations, the algorithm can tell whether it is worth undertaking another optimization iteration. If so, the whole process is repeated. Finally, the algorithm stops when a convergent solution is met (Figure 3.3.1) or when the maximum number of iterations is attained.

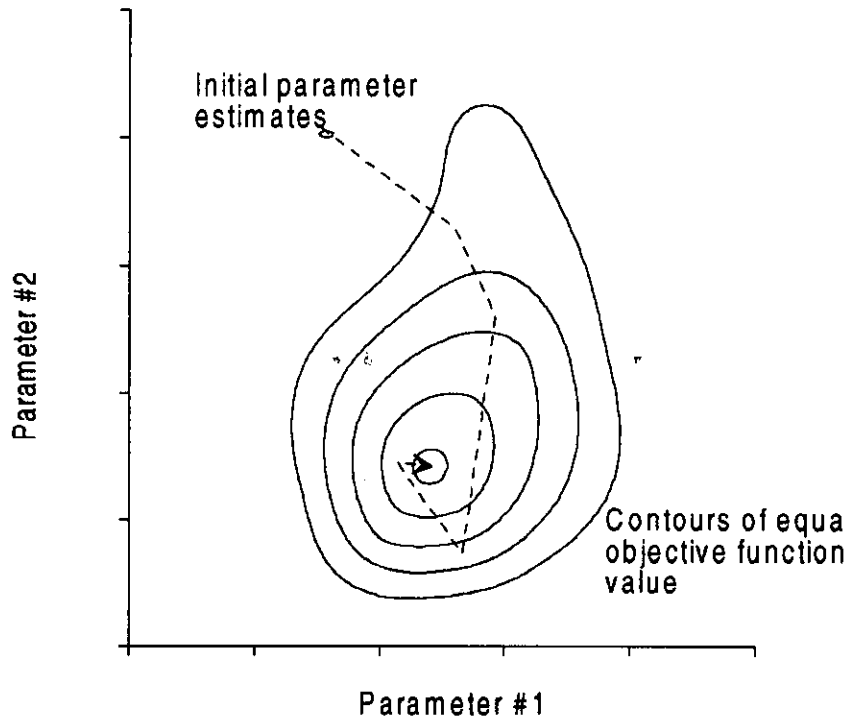


Figure 3.3.1. Objective function contour values and iterative improvement from initial parameter values towards a global minimum (adapted from Doherty et al., 1994)

3.3.3.4. Optimum Length of the Parameter Upgrade Vector

Inclusion of the Marquardt parameter in Equation (3.3.17) has the desired effect of rotating the parameter upgrade vector $\Delta \mathbf{P}$ towards the negative of the gradient vector. However, while the direction of $\Delta \mathbf{P}$ may now be favourable, its magnitude may not be optimum.

If $\mathbf{P}_k$ is the current parameter set, the updated set is calculated from

$$\mathbf{P}_{k+1} = \mathbf{P}_k + \beta \Delta \mathbf{P} \tag{3.3.18}$$

where $\Delta \mathbf{P}$ is given by Equation (3.3.17) and β is obtained from a Taylor series expansion of the objective function in terms of β (Doherty et al., 1994)

$$\beta = \frac{\sum_{i=1}^{NE} W_i \sum_{l=1}^{L_i} w_l^2 r_l(p_k)\gamma_l}{\sum_{i=1}^{NE} W_i \sum_{l=1}^{L_i} w_l^2 \gamma_l^2} \tag{3.3.19}$$

where γ_l is given by,

$$\gamma_l = \sum_{j=1}^{M} \Delta \mathbf{P}_j \frac{\partial \mathbf{u}_l}{\partial \mathbf{P}_j} \tag{3.3.20}$$

i.e.

$$\gamma = \mathbf{J}\,\Delta\mathbf{P} \tag{3.3.21}$$

and **J** is the Jacobian matrix.

3.3.3.5. Calculation of the Jacobian matrix

The Jacobian matrix contains derivatives of observations with respect to parameters (see Equation 3.3.9) which are usually called sensitivity coefficients. There are three approaches for computing sensitivity coefficients: (1) the finite difference method (or the influence coefficient method), (2) the sensitivity equation method and (3) the adjoint state method (or variational method) (Yeh, 1986, and Sun and Yeh, 1990a). The sensitivity equation and adjoint state methods are widely used to compute the Jacobian matrix for coupled inverse problems of flow and solute transport. Samper and Neuman (1986) derived adjoint state equations for flow and advective-dispersive transport using two approaches (continuous and discrete approaches), under both steady and transient conditions. Carrera and Neuman (1986), Sun and Yeh (1990), Sun (1994), Medina and Carrera (1996) and others have used these two methods to compute the Jacobian matrix for conservative solute transport. The application of these methods to multi-component reactive transport, however, entails a large complexity.

The finite difference method does not require complex mathematical derivations. Generally, two kinds of finite difference methods can be used to compute the Jacobian matrix: forward and central difference methods. The forward problem must be solved once for each parameter to be estimated during each optimization iteration when the forward difference method is used. A small increment is added to the parameter value prior to each run of the forward problem. The finite difference method is generally considered to be less accurate because it is difficult to find appropriate parameter increments (Yeh, 1986, and Sun and Yeh, 1990a). However, with the improvement of forward computing accuracy and the accumulation of experience, this method has been successfully used to complex inverse problems for which the derivation of the adjoint state and sensitivity equations is difficult. Mishra and Parker (1989) used a forward finite difference method with a relative parameter increment of 10^{-2} for the estimating of soil hydraulic and transport parameters from transient unsaturated flow and tracer experiments. Barlebo et al. (1998) used a central difference method to evaluate the Jacobian matrix with a relative parameter perturbation of 10^{-4} for the inverse problem of three-dimensional flow and solute transport. Sonnenborg et al. (1996) also used a central difference method for inverse flow and solute transport model within the context of aquifer contamination. The finite difference method is very economical and practical for computing the Jacobian matrix for these complex conditions.

To increase the accuracy of the derivatives, parameter perturbations $\delta \mathbf{P}_i$ are computed as increments for different parameters by

$$\delta \mathbf{P}_i = \mathrm{P}_{i+1} - \mathrm{P}_i = \alpha_i \mathrm{P}_i \qquad (\mathrm{i} = 1, 2, \ldots, \mathrm{M}) \tag{3.3.22}$$

where α_i is a user-defined dimensionless coefficient which generally ranges from 10^{-2} to 10^{-5}. Its optimum value should be derived by a try and error procedure. This technique to compute derivatives is referred to as the forward difference method. It can be expressed as

$$\frac{\partial \mathbf{u}}{\partial \mathbf{P}} \cong \frac{\mathbf{u}(\mathrm{P}_i + \alpha_i \mathrm{P}_i) - \mathbf{u}(\mathrm{P}_i)}{\alpha_i \mathrm{P}_i} \qquad (\mathrm{i}= 1, 2, \ldots, \mathrm{M}) \tag{3.3.23}$$

Obviously, this expression provides an approximation of the true derivatives. If the increment is too large, the approximation may be poor; if the increment is too small, however, roundoff errors may badly affect the accuracy of the derivatives. Both of these effects may degrade the performance of the optimization algorithm.

A more accurate estimation of the derivatives can be obtained by resorting to a central difference method. This method requires solving the forward problem twice per parameter. At the first run, an increment is added to the current parameter value, while for the second run the increment is subtracted. It can be expressed by the following equation

$$\frac{\partial \mathbf{u}}{\partial \mathbf{P}} \cong \frac{\mathbf{u}(P_i + \alpha_i P_i) - \mathbf{u}(P_i - \alpha_i P_i)}{2\alpha_i P_i} \qquad (i= 1,2,\ldots, M) \qquad (3.3.24)$$

It is wise to start the optimization process with the more economical forward difference method (3.3.23). Near the optimum, one can switch to the central difference method (3.3.24) to refine the results.

3.4. ERROR ANALYSIS

There are important issues that go beyond the mere finding of best-fit parameters. Data are generally not exact. They are subject to measurement and interpretation errors (called noise in the context of signal-processing). Thus, the model never fits exactly the data, even when the model structure is correct. It is needed to assess whether or not the model is appropriate. Goodness-of-fit statistical tests are useful for that purpose.

One needs to know the uncertainty of the estimated parameters. In other words, it is needed to know the likely error of the best-fit parameters. Parameter fitting is not the sole purpose of parameter estimation. To be genuinely useful, a fitting procedure should provide (1) parameter estimates, (2) estimation errors, and (3) a statistical measure of goodness-of-fit. If the latter shows a poor fit, then items (1) and (2) are probably worthless.

Unfortunately, some practitioners of parameter estimation never proceed beyond item (1). They deem a fit acceptable if a graph of data and model responses "looks good". This kind of approach is known as "chi-by-eye". Luckily, its practitioners get what they deserve (Press et al., 1992).

In the next section the statistical measures of goodness-of-fit as well as measures of parameter uncertainty (approximate confidence intervals) are computed from covariance and correlation matrices.

3.4.1. VARIANCE-COVARIANCE MATRIX

Optimum parameter values **P** in Equation (3.3.18) are the "best unbiased" estimators of the set of the true system parameters appearing in Equation (3.2.1). As estimators, they are one particular realization of a M-dimensional random vector **P** estimated by Equation (3.3.18), from the NL-dimensional random vector **F** of experimental observations. According to the equation given by Mishra and Parker (1989), and Doherty et al. (1994), the variance of the estimated error σ^2 can be calculated as

$$\sigma^2 = E(p)/(NL - M) \tag{3.4.1}$$

where E(p) is the final least-squares error (as given by Equation (3.2.1). The term (NL - M) is the difference between the number of observations and the number of parameters to be estimated and represents the number of "degrees of freedom" of the parameter estimation problem. Equation (3.4.1) shows that σ^2 is directly proportional to the objective function value and thus varies inversely with the goodness of fit between experimental observation **F** and the state variable **u** calculated on the basis of the optimal parameter set.

A first-order approximation of the covariance matrix of **P**, **C**(p), can be derived from (Mishra and Parker, 1989)

$$\mathbf{C}(p) \cong \sigma^2(\mathbf{J}^T\mathbf{W}\mathbf{J})^{-1} \tag{3.4.2}$$

where, **W** is the weighting matrix. Notice that, the covariance matrix is symmetric, that is, $C_{ij} = C_{ji}$. For i = j, then $C_{ij} = C_{ii} = \sigma_i^2$. If parameters P_i and P_j are uncorrelated, then $C_{ij} = 0$.

In general, **C**(p) is not a diagonal matrix. In fact, in many parameter estimation problems, parameters are strongly correlated. In such cases some parameter variances (the diagonal entries of **C**(p)) may be large even though the objective function E(P) may be reasonably low. When parameters are highly correlated, the matrix ($\mathbf{J}^T \mathbf{WJ}$) in Equation (3.4.2) may become near singular and parameter estimation may become highly unstable.

There are two matrices, the correlation matrix and the normalized eigenvector matrix, both of which are derived from the covariance matrix **C**, but measure better the parameter correlation and uncertainty than **C**(p) itself.

3.4.2. CORRELATION MATRIX

The entries of the correlation matrix, ρ_{ij}, are obtained from those of the covariance matrix according to:

$$\rho_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} C_{jj}}} \tag{3.4.3}$$

where C_{ij} represents the element at the *i*'th row and *j*'th column of **C**. The entries of the correlation matrix have the same signs as those of the covariance matrix. It is also symmetric. Furthermore, the diagonal elements of the correlation coefficient matrix are always 1; off-diagonal elements range between -1 and 1. The closer are these off-diagonal elements to 1 or -1, the higher the correlation. If ρ_{ij} is close to +1, then P_i and P_j are highly positively correlated. On the other hand, if ρ_{ij} is close to -1, then P_i and P_j are highly negatively correlated. The correlation matrix measures the correlation between parameters, but tells nothing about the uncertainty of the estimated parameters.

3.4.3. EIGENANALYSIS OF THE COVARIANCE MATRIX

The eigenvalues and eigenvectors of the covariance matrix can be used for evaluating the uncertainty of the estimated parameters. The $M \times M$ covariance matrix $\mathbf{C}$ is said to have an eigenvector $\mathbf{X}$ and corresponding eigenvalue λ if

$$\mathbf{C} \cdot \mathbf{X} = \lambda \mathbf{X} \tag{3.4.4}$$

From Equation (3.4.4) it follows that

$$\det|\mathbf{C} - \lambda \mathbf{I}| = 0 \tag{3.4.5}$$

This equation defines an M-th degree polynomial in λ whose roots are the eigenvalues of $\mathbf{C}$. Since $\mathbf{C}$ is semi-definite positive, it always has M positive eigenvalues (not necessarily different) and each eigenvalue has a set of corresponding eigenvectors. In Equation (3.4.4) the eigenvector $\mathbf{X}$ is called right eigenvector because it is multiplied to the right of the matrix $\mathbf{C}$. On the other hand, if it is multiplied to the left of the matrix $\mathbf{C}$, such as

$$\mathbf{X} \cdot \mathbf{C} = \lambda \mathbf{X} \tag{3.4.6}$$

it is called left eigenvector. From the above equations we can see that the corresponding left and right eigenvalues of matrix $\mathbf{C}$ are identical. Moreover, the covariance matrix $\mathbf{C}$ is symmetric, then the left and right eigenvectors are just transposes of each other, that is, have the same numerical values as components. Eigenvalues and normalized eigenvectors of the covariance matrix $\mathbf{C}$ are computed by a subroutine obtained from the IMSL library (Press et al., 1992).

Generally, large eigenvalues of the covariance matrix correspond to linear combinations of parameters that are poorly estimated and highly uncertain with the available observation data. The coefficients of each linear combination are the components of the corresponding eigenvector. On the other hand, small eigenvalues correspond to linear

combinations that are estimated with more confidence (Carrera, 1984). If an eigenvector is dominated by one element, individual parameter values are well resolved by the estimation process. However, if predominance within each eigenvector is shared between a number of elements (especially for those eigenvectors whose eigenvalues are largest), the corresponding parameters are highly correlated and uncertain.

Another measure of particular interest is the condition number which is defined as the ratio between the largest and the smallest eigenvalues of the covariance matrix (Carrera, 1984). It is the most widely used measure of ill-conditioning in estimation problems.

3.4.4. CONFIDENCE REGIONS AND CONFIDENCE INTERVALS

In this chapter, statements have been made about the variance-covariance matrix and the correlation matrix of estimated parameters, as well as eigenvalues and eigenvectors of the covariance matrix. In this section, More precise quantitative measures of parameter uncertainty, in terms of the confidence limits are presented. Confidence limits include confidence regions and confidence intervals.

3.4.4.1. Confidence region

A confidence region is a region of an M-dimensional space (hopefully a small region) that contains a certain (hopefully large) percentage of the total probability distribution. There is a, for example, 90 or 95% chance that the true parameter values fall within this region around the estimated values. In one dimension, the convention is to use a line segment centered on the estimated values; in higher dimensions confidence regions are ellipses or ellipsoids.

When the method (Equation (3.2.1)) used to estimate the parameters $\mathbf{P}^*$ is chi-square minimization $x^2 = E(p)$, then there is a natural choice for the shape of confidence region,

whose use is almost universal. For the observed data set **F**, the value of E(p) is a minimum at P^*. Let $E(p)_{min}$ be this minimum value. If the vector **P** of parameter values is perturbed away from P^*, then E(p) increases. The region within which E(p) increases by no more than a given amount $\Delta x^2 = \Delta E(p)$ defines an M-dimensional confidence region around $\mathbf{P}^*$. If Δx^2 is set to a large number, the region will be large, and conversely. Somewhere in between there will be choices of Δx^2 that will cause the region to contain various amounts (68, 90 or 95%) of probability distribution. These regions are taken as the confidence regions for the parameters $\mathbf{P}^*$.

3.4.4.2. Confidence interval

Very frequently one is interested not in the full M-dimensional confidence region, but in the confidence interval of each parameter taken separately. In Equation (3.4.2) the diagonal elements of the covariance matrix **C** contain the individual parameter variance. From them and the appropriate value of Student's t distribution with ν = (NL-M) degrees of freedom, Mishra and Parker (1989) given a equation to determine the individual parameter confidence intervals,

$$Pr\,[|P_i - P_i^*| \leq t_\alpha \sqrt{C_{ii}}\,] = 1-\alpha \qquad (3.4.7)$$

where P_i^* is the estimated value of the parameter P_i, C_{ii} is the parameter variance, and t_α is the value of Student's t distribution of ν degrees of freedom for confidence level (1- α). However, Equation (3.4.7) is strictly applicable for linear regression and under the assumption of that the observed data set consists of NL independent and identically distributed data points. Although it has been shown to be reasonably accurate in nonlinear least square problems (Donaldson and Schnabel, 1986), it may be suboptimal because it neglects some information.

According to the theorem given by Press et al. (1992) and Samper and Neuman (1989a, b), let $\Delta x^2 = \Delta E(p) = E(p^*) - E(p^0)$ be distributed as a chi-square distribution with M degrees of freedom, if P^* is drawn from the universe of simulated data sets with actual

parameters P^0. This theorem makes the connection between particular values of Δx^2 and the fraction of the probability distribution. As a modification to Equation (3.4.7), substituting the *t* distribution with a chi-square distribution, one obtains,

$$Pr\left[\ |P_i - P_i^*| \le \sqrt{x_\alpha^2}\sqrt{C_{ii}}\ \right] = 1-\alpha \qquad (3.4.8)$$

where x_α^2 is the chi-square statistic corresponding to M degrees of freedom and a confidence interval of $100(1-\alpha)$. Values of x_α^2 are given by Press et al. (1992). For other values, they can be computed from the inverse of chi-square probability distribution function (Press et al., 1992). Our code incorporates Equation (3.4.8) with 90% or 95% confidence interval values (i.e., α =0.1 or 0.05) to quantify the uncertainty of estimated parameters.

CHAPTER 4

COMPUTER CODES

4.1. DESCRIPTION OF THE CODES

Numerical methods of forward and inverse solutions were described in previous chapters. The forward modeling part is based on CORE^{2D} (Samper et al., 1999), a code thoroughly verified by Xu (1996), Samper et al. (1998), Xu et al. (1999) and Samper et al. (1999). From the published literature on the topic of inverse modeling, it is clear that there is no single approach that is best in all of possible conditions. For this reason, three codes, which share the same forward modeling subroutines, were developed. INVGS-CORE is based on the linear search or Golden section search method to estimated flow and reactive transport parameters. INVQI-CORE is an inverse code which uses the quadratic interpolation method. The third, INVERSE-CORE^{2D}, is the most powerful inverse code and uses the Gauss-Newton-Levenberg-Marquart method which is especially designed for the optimization of least square problems. The main features of these codes are described in Section 4.1. Code verification is the process by which one ensures that codes solve the inverse problem correctly. It is presented in Section 4.2.

4.1.1. GOLDEN SECTION SEARCH METHOD: INVGS-CORE

The Golden section search method is primarily developed for one-dimensional inverse problems. However, it was extended for the estimation of multiple parameters by adopting a multi-cycle iterative process to attain the overall minimum solution. At a given cycle parameters are estimated one by one according to the following sequence:

input a_1, b_1 , x*tol, ftol, Nitermax*

set $c_1 = a_1 + (1-\tau)(b_1-a_1)$, $E_c = E(c_1)$

$d_1 = b_1 - (1-\tau)(b_1-a_1)$, $E_d = E(d_1)$

Niteration=0

for k = 1, 2, ... repeat

Niteration= Niteration +1

if $E_c < E_d$ **then**

set $a_{k+1} = a_k$, $b_{k+1} = d_k$, $d_{k+1}=c_k$

$c_{k+1} = a_{k+1} + (1-\tau)(b_{k+1}-a_{k+1})$

$E_d = E_c$, $E_c = E(c_{k+1})$

else

set $a_{k+1} = c_k$, $b_{k+1} = b_k$, $c_{k+1}=d_k$

$d_{k+1} = b_{k+1} - (1-\tau)(b_{k+1}-a_{k+1})$

$E_c = E_d$, $E_d = E(d_{k+1})$

end

repeat until $|b_{k+1}-a_{k+1}| <$ *xtol* or,

$|E(c_k)-E(c_{k+1})|/|E(c_k)+ E(c_{k+1})|*0.5 <$ *ftol* or,

Niteration $\geq$ *Nitermax*

The iterative process terminates either when the interval of uncertainty falls below *xtol* or when the relative change in the objective function value is less than *ftol* or when the number of iterations is greater or equal to a specified value *Nitermax.*

The main advantage of Golden section search method is that it does not require computing the Jacobian matrix. Besides, the process can be terminated at any iteration by any criteria. However, its performance decreases when the number of parameters increases. Therefore, this method is mostly effective for solving inverse problems of water flow and reactive solute transport with a few parameters.

4.1.2. QUADRATIC INTERPOLATION METHOD: INVQI-CORE

As described in chapter 3, the quadratic interpolation method provides another way to determine new search points for the inverse problem. The algorithm is stated as follows:

input a_1, b_1, c_1, *xtol*, *ftol*

set $E_a = E(a_1)$, $E_b = E(b_1)$, $E_c = E(c_1)$

for k = 1, 2, ... **repeat**

set $$d = \frac{1}{2}\frac{(b^2 - c^2)E_a + (c^2 - a^{2)}E_b + (a^2 - b^2)E_c}{(b-c)E_a + (c-a)E_b + (a-b)E_c}$$

$E_d = E(d)$

if $d < c_k$ and $E_d < E_c$ **then**

set $a_{k+1} = a_k$, $b_{k+1} = c_k$, $c_{k+1} = d$

$E_b = E_c$, $E_c = E_d$

else if $d > c_k$ and $E_d > E_c$ **then**

set $a_{k+1} = a_k$, $c_{k+1} = c_k$, $b_{k+1} = d$

$E_b = E_d$

else if $d < c_k$ and $E_x > E_c$ **then**

set $a_{k+1} = d$, $c_{k+1} = c_k$, $b_{k+1} = b_k$

$E_a = E_d$

else

set $a_{k+1} = c_k$, $b_{k+1} = b_k$, $c_{k+1} = d$

$E_a = E_c$, $E_c = E_d$

end

repeat until $|b_{k+1}-a_{k+1}| <$ *xtol* or $|E(c_k)-E(c_{k+1})|/|E(c_k)| <$ *ftol*

The iterative process terminates either when the interval of uncertainty falls below *xtol* or when the relative change in the objective function is less than *ftol*. Similar to the Golden section search method, this method does not require computing the Jacobian matrix, and its effectiveness is restricted to inverse problems with a few parameters.

4.1.3. GAUSS-NEWTON-LEVENBERG-MARQUART: INVERSE-CORE^{2D}

4.1.3.1. Solution method

The Gauss-Newton-Levenberg-Marquart method is an effective and powerful algorithm for minimizing least square objective functions. It incorporates a technique to deal with problems related to the singularity of matrix $\mathbf{J}^T\mathbf{J}$ in a manner described in Equations 3.3.17 and 3.3.18. For a sufficiently large value of α ($\alpha \geq 0$), matrix $(\mathbf{J}^T\mathbf{J} + \alpha \mathbf{I})$ is positive definite and $\Delta\mathbf{P}$ is then a descent direction. As $\mathbf{P}_{k+1}$ approaches the optimum, however, it is required that $\alpha \rightarrow 0$ so that the method acquires the asymptotic rate of convergence of the Gauss-Newton method. Various strategies for selecting the value of α at each iteration are discussed below.

When $\alpha = 0$, $\Delta\mathbf{P}$ is the Gauss-Newton vector. As $\alpha \rightarrow \infty$, the effect of the term $\alpha \mathbf{I}$ increasingly dominates that of $\mathbf{J}^T\mathbf{J}$ so that $\Delta\mathbf{P} \rightarrow -\alpha^{-1}\mathbf{J}^T\,\mathbf{r}(\mathrm{P})$ which represents an infinitesimal step in the steepest descent direction. Between these two extremes, both $\| \Delta P \|$ and the angle between $\Delta\mathbf{P}$ and $-\boldsymbol{g}$ (the steepest descent direction) decrease monotonically as α increases. This property is useful because, while the magnitude of the Gauss-Newton vector is a rough indication of an acceptable step length, increasing the bias of $\Delta\mathbf{P}$ towards the steepest descent direction makes $\| \Delta P \|$ more and more likely to be too large a step to give a reduction in function value. The set of all points $\mathbf{P}_{k+1} = \mathbf{P}_k + \mu\,\Delta\mathbf{P}$, $0 < \mu \leq 1$, as α varies from 0 to ∞ defines part of a hyperplane in the space of the variables sometimes known as the region of trust.

The computation of the Jacobian matrix $\mathbf{J}$ which contains the derivatives of observations with respect to parameters is the main part of this optimization algorithm. It is calculated by using finite differences. At each iteration the forward model is run once for each adjustable parameter and a small parameter increment is added to the parameter value prior to the run. The resulting numerical changes are divided by this increment in order to compute an approximation of derivatives. This is repeated for each parameter. This

technique of derivative calculation is referred to as the method of "forward differences". Derivatives calculated in this way are only approximate. If the increment is too large, the approximation will be poor. If the increment is too small, roundoff errors will affect the accuracy of the derivatives. Both of these effects will degrade optimization performance. To overcome this problem, the model allows derivatives to be calculated using a "central differences" method. It requires two forward model runs to calculate derivatives with respect to a parameter. For the first run an increment is added to the current parameter value, while for the second run the increment is subtracted.

Each iteration of the Gauss-Newton-Levenberg-Marquart method requires running the forward model (M+1) times (M is the number of the parameters to be estimated) if the forward difference method is used to compute the Jacobian matrix. If the central difference method is used, the number of runs is (2M+1).

4.1.3.2. Main features of INVERSE-CORE^{2D}

The main structure of INVERSE-CORE^{2D} is shown in the flowchart of Figure 4.1.1. The program starts by reading all the relevant data for inverse modeling which includes control parameters, lower and upper bounds, initial values, and prior information of the parameters, as well as all kinds of measured data. Then, it continues reading the data for the forward modeling of flow, heat and solute transport. Thermal data are only needed if heat transport is to be simulated. At the beginning of a run of Gauss-Newton-Marquardt-Levenberg iteration, it is necessary to supply a set of initial parameter values. These are the values that the algorithm uses to start the first optimization iteration. At each iteration, the Jacobian matrix is calculated using finite differences (forward or central difference method). After solving the Gauss-Newton-Marquardt-Levenberg equation, a set of updated parameters is obtained. Then subroutine TRANQIN is called to calculate an improved objective function value. By comparing parameter changes and objective function improvement achieved through the current iteration with those achieved in previous iterations, the algorithm can tell whether it is worth undertaking another optimization

iteration. If so, the whole process is repeated. Finally, the program stops when a convergent solution is attained or when maximum number of iterations is exceeded.

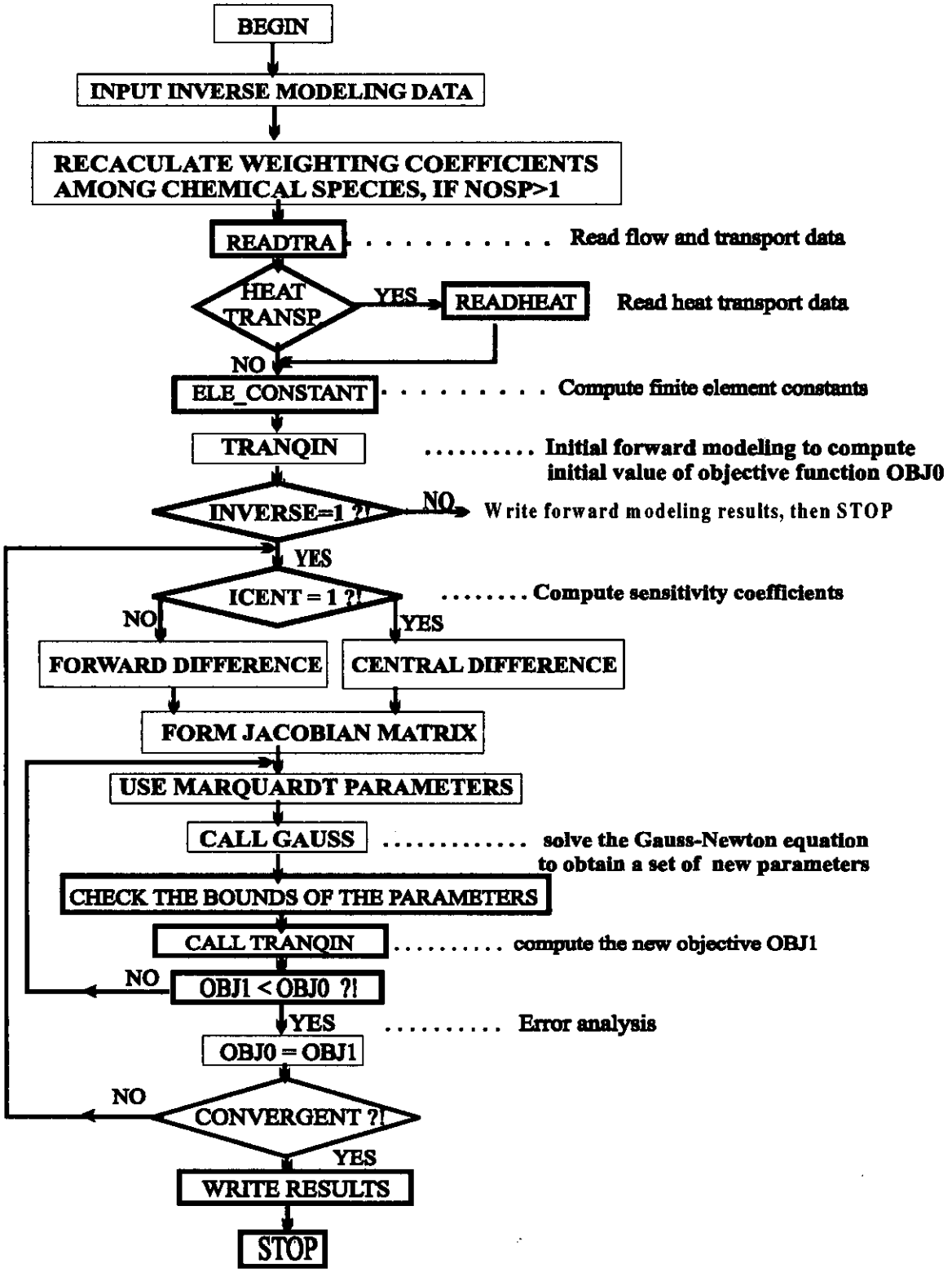


Figure 4.1.1. Flowchart of INVERSE-CORE^{2D} main program

4.1.3.3. Flowcharts of the forwards modeling program

The first subroutine of INVERSE-CORE^{2D} is TRANQIN which is a modified version of TRANQUI3, the main program of the forward modeling code (CORE^{2D}) (see Figure 4.1.2). This subroutine starts by updating the parameters, according to the parameter identification numbers, which are estimated by the main program. Then, hydrochemical data are read and initialized. Once all data are read, the code constructs the finite element flow and transport matrices which remain unchanged for steady flow and constant time increment. The logical sequence for solving the equations is the following: first the flow equation is solved and nodal hydraulic heads (pressure heads for variably saturated media) are computed (see Figure 4.1.3 for details of the flow solution subroutine). Then water fluxes are computed by applying Darcy's Law to the finite element solution. These fluxes are later used to construct solute and heat transport matrices. Computation of heat and solute transport matrices is similar to that of flow matrices. As one can see in Figure 4.1.4, dispersive terms are computed first, then convective terms, and finally boundary condition terms. Both heat and solute transport finite element matrices are solved by LU decomposition. Then, reactive transport equations are solved by following the sequential iterative scheme shown in Figure 4.1.5. The solution of the hydrochemical equations is performed by a Newton-Raphson method. The detailed calculations needed to obtain the solution of the chemical equilibrium equations are depicted in Figure 4.1.6. Finally, the subroutine TRANQIN calculates the current objective function and updates the weighting coefficients.

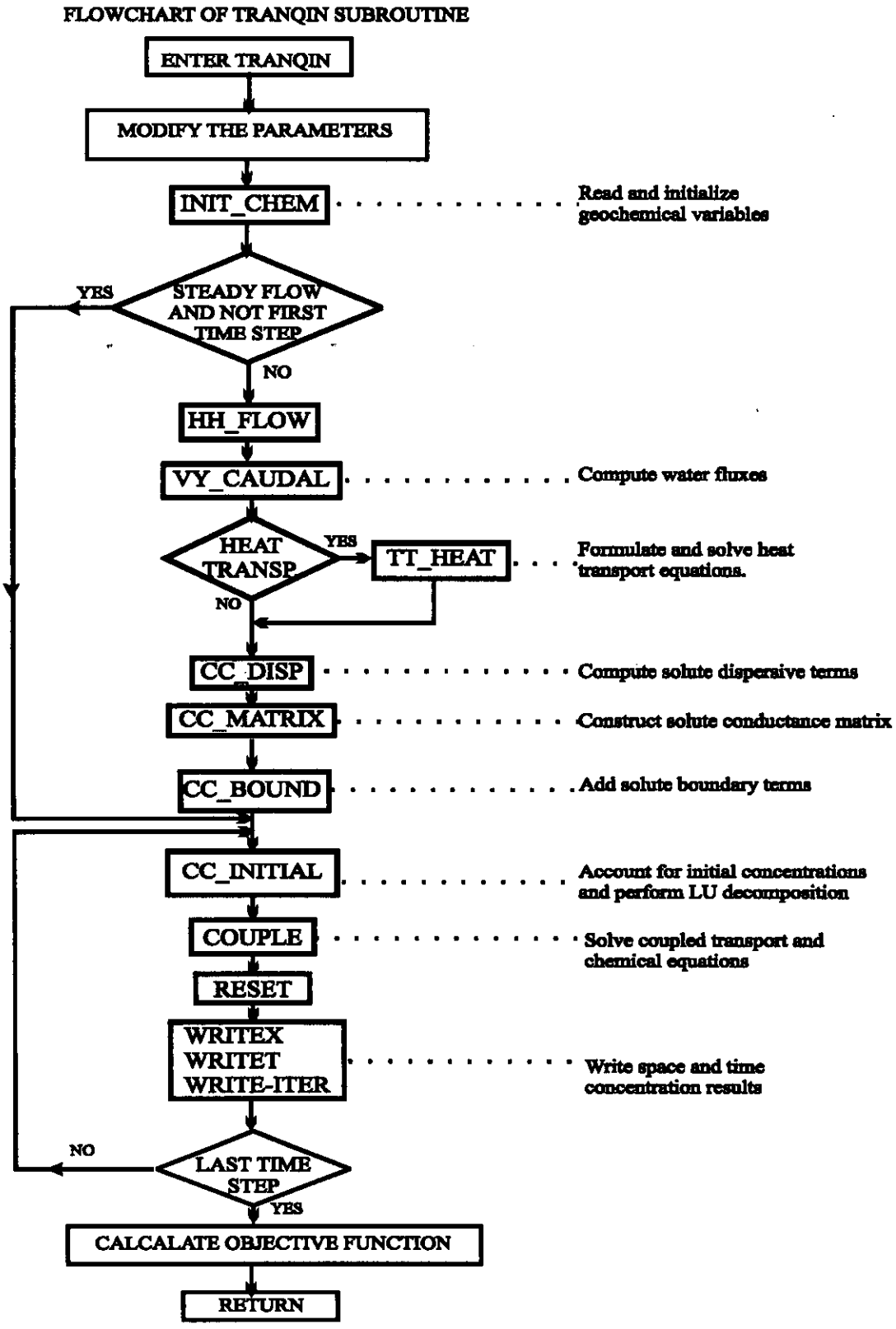


Figure 4.1.2. Flowchart of subroutine TRANQIN (adapted from Xu, 1996)

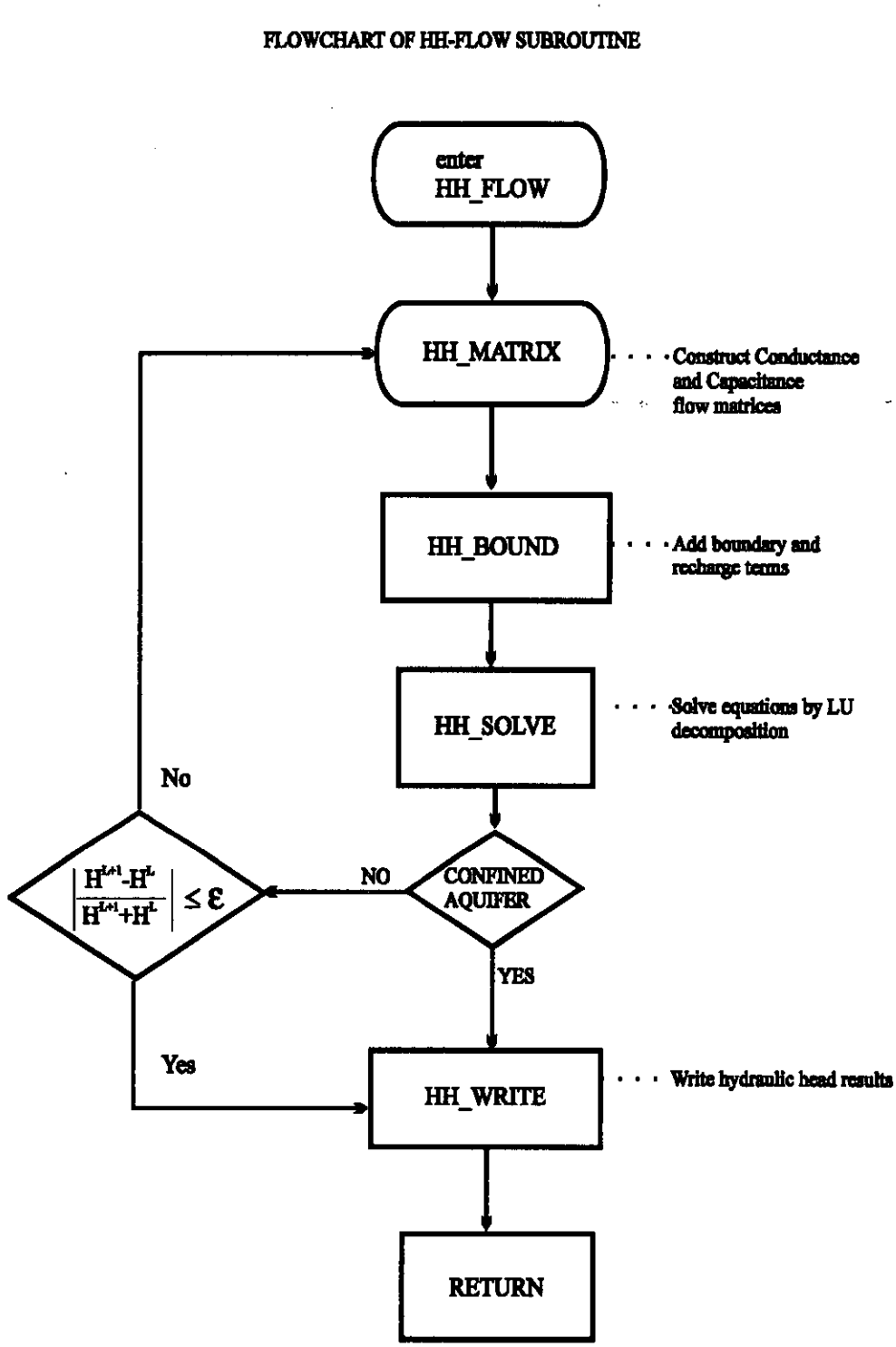


Figure 4.1.3. Flowchart of the HH_FLOW subroutine which solves the groundwater flow equation (Xu,1996)

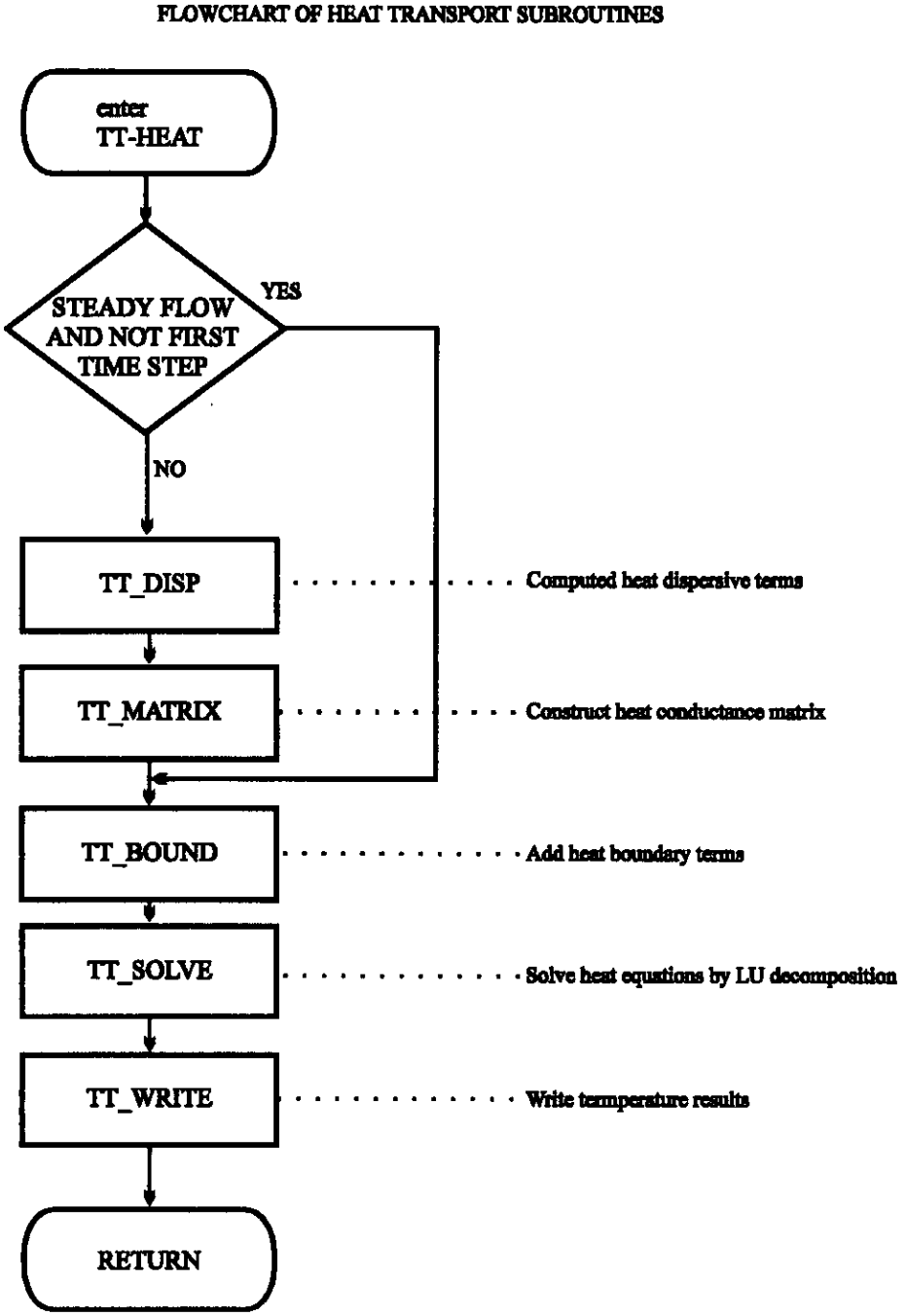


Figure 4.1.4. Flowchart of the HH_HEAT subroutine which solves the heat transport equation (Xu,1996)

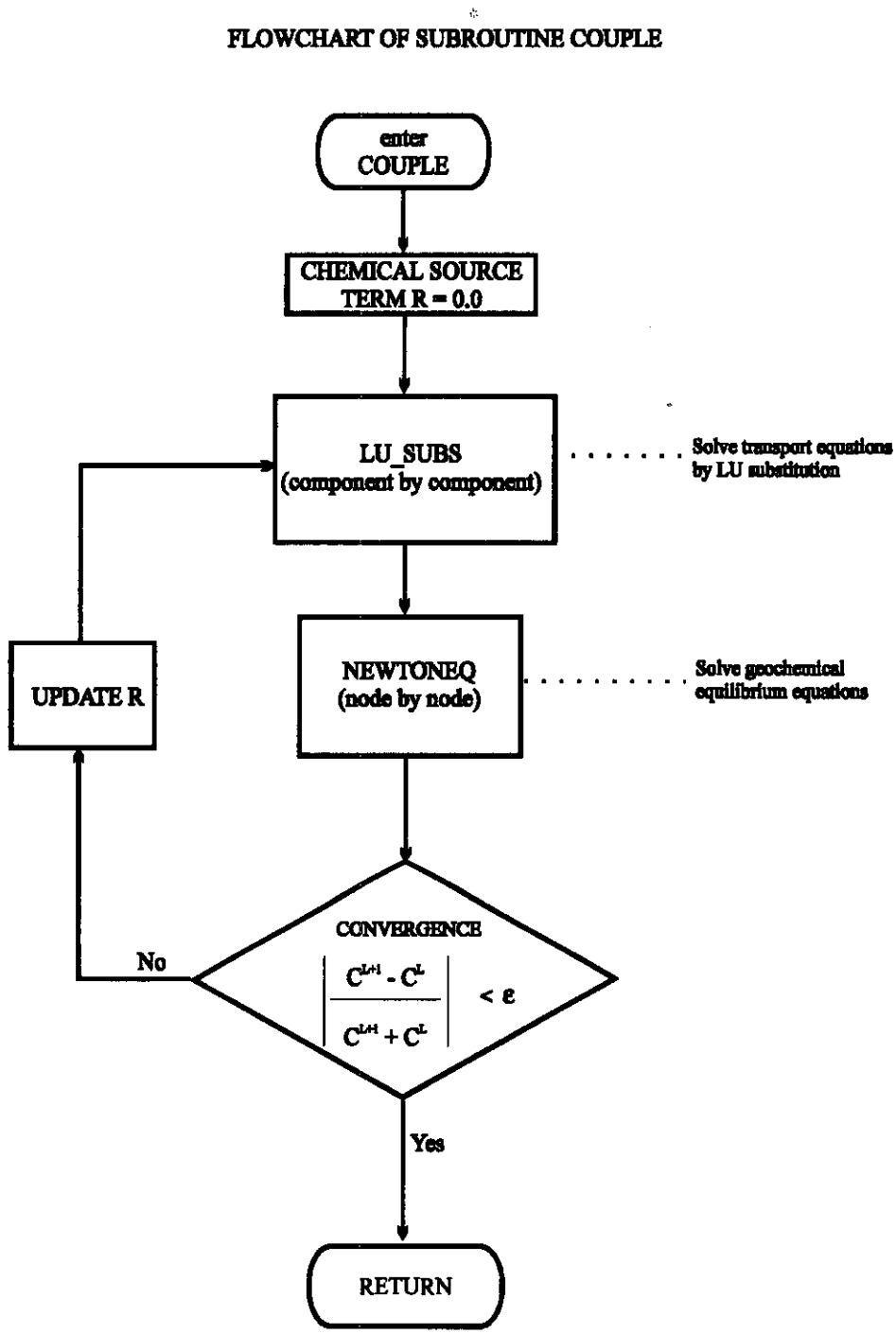


Figure 4.1.5. Flowchart of the subroutine COUPLE which performs the sequential iterative process for solving reactive solute transport (adapted from Xu,1996)

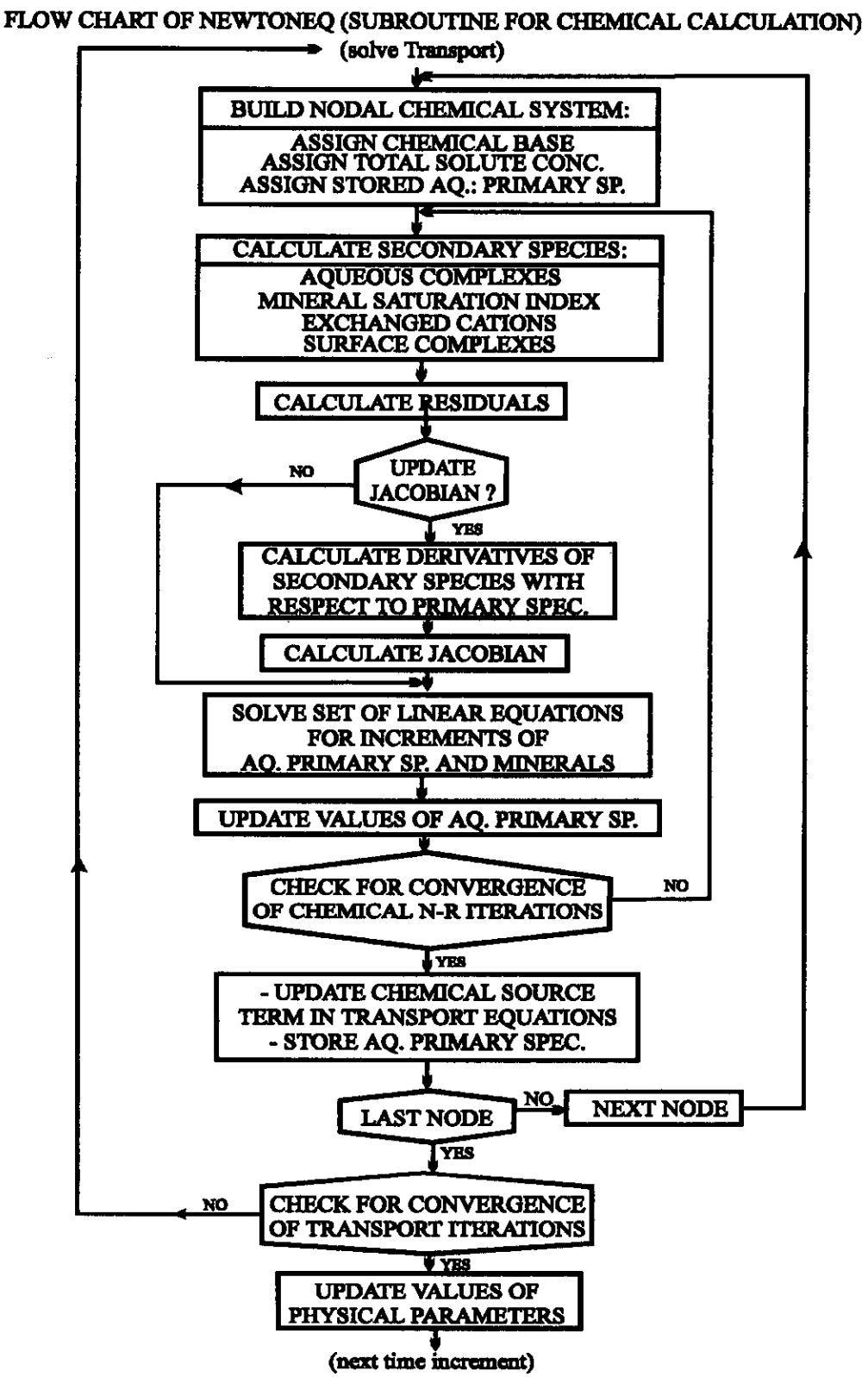


Figure 4.1.6. Flowchart of the NEWTONEQ subroutine which solves the chemical equilibrium equations by a Newton-Raphson method (adapted from Xu, 1996)

4.2. CODE VERIFICATION

Verification consists on testing the proper performance of the code, i.e., checking that the code solves properly the equations it is intended to solve for. For that purpose, synthetic examples, in which the true model parameters are known exactly, are used for demonstrating the accuracy of the estimated parameters. Modeling the forward problem has been extensively verified as part of the development of CORE^{2D} (see Xu, 1996, Samper et al., 1998, Xu et al., 1999, and Samper et al., 1999). The following synthetic examples were used to verify INVERSE-CORE^{2D}: (1) unsaturated flow and conservative solute transport (EJ1), (2) reactive solute transport with cation exchange (EJ2), and (3) reactive solute transport with kinetic mineral dissolution/precipitation reactions (EJ2KIN). Synthetic observation data for examples EJ2 and EJ2KIN were obtained from a forward modeling by adding different levels of random noise. Noise-corrupted synthetic data for these examples are included in Appendix 2.

4.2.1. SYNTHETIC EXAMPLE OF WATER FLOW AND SOLUTE TRANSPORT

This section is mainly devoted to the verification of codes INVGS-CORE and INVQI-CORE. A procedure of increasing complexity was followed by increasing the number of parameters to be estimated. The effectiveness of Golden section search and quadratic interpolation methods has been tested under different conditions. Different sets of initial values of parameters were used to test the capabilities of the inverse code INVGS-CORE. This example was also used to verify code INVERSE-CORE^{2D}. By using noise-free data, the true parameters were obtained with just a few iterations. To avoid repetitions, this part of work will not presented here.

4.2.1.1. Formulation of synthetic example (EJ1)

This example deals with the transport of iodide through a bentonite barrier. It corresponds to run M1-I-SW-OUT-1 which involves the transport of I^- added at the externd

surface of the cylindrical clay barrier of the FEBEX experiment (Samper et al., 1997). It is a radial 1-D numerical model with 200 triangular elements and 202 nodes. Tracer iodide was put at the outer boundary and it is transported into the unsaturated FEBEX bentonite due to water injection at a constant pressure head of 55 m. All other boundaries are close for water flow and solute transport. The model has only one parameter zone.

The following transport parameters are estimated: (1) Total porosity ϕ, (2) Molecular diffusion coefficient D_0, and (3) Distribution coefficient K_d. The prior information, and the lower and upper bounds of the parameters are listed in Table 4.2.1. Breakthrough curves of concentrations at three nodes, which were obtained from forwards numerical modeling without noise, are taken as measured data. In this case, it is assumed that model structure, observation error, as well as parameterization errors, are equal to zero.

Table 4.2.1. Prior information of the parameters (three-parameter case)

Parameter	D_0 (m^2/s)	K_d (ml/g)	ϕ
True value	$9.83 \cdot 10^{-11}$	0	0.41
Lower bound	$2.9 \cdot 10^{-11}$	0	0.21
Upper bound	$2.9 \cdot 10^{-10}$	1.5	0.51
Initial guess	$9.4 \cdot 10^{-11}$	0.001	0.32
Tolerance error	$2 \cdot 10^{-7}$	10^{-6}	10^{-4}

4.2.1.2. Estimation of one parameter

It is assumed that the diffusion coefficient is unknown, although its value is located within the interval ($2.9 \cdot 10^{-11}$, $2.9 \cdot 10^{-10}$) m^2/s. Both search algorithms (INVGS-CORE and INVQI-CORE) were used to estimate this parameter. Results are presented in Figures 4.2.1 and 4.2.2. Figure 4.2.3 shows the shape of the objective function, which indicates that the inverse problem for estimating the diffusion coefficient has a single optimal solution.

A convergence tolerance for the objective function (relative change between two consecutive iterations *ftol*) of 10^{-6} was used in all cases. The maximum number of iterations was fixed to 15.

Both methods provide accurate solutions (the true value of the diffusion coefficient is $9.83 \cdot 10^{-11}$ m^2/s). However, the quadratic interpolation method is more efficient than the Golden section search. The former requires 9 consecutive iterations plus three additional iterations for computing the objective function values at the three initial points. Golden section requires 22 iterations. Therefore, from this example of noise-free data, it can be concluded that the quadratic interpolation method is more efficient than the Golden section search method.

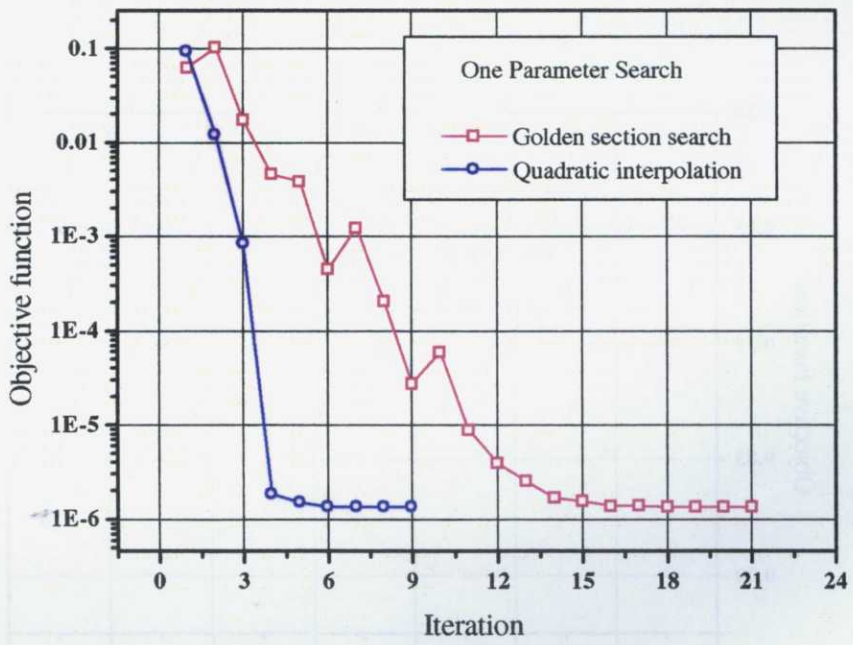


Figure 4.2.1. Objective function values as a function of iteration number for the two search algorithms

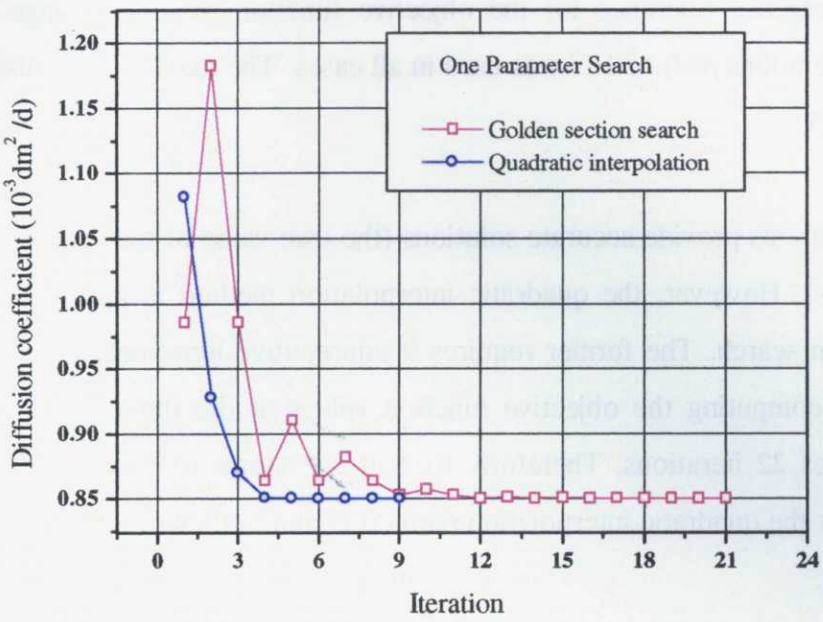


Figure 4.2.2. Parameter values as a function of iteration number for the two search algorithms

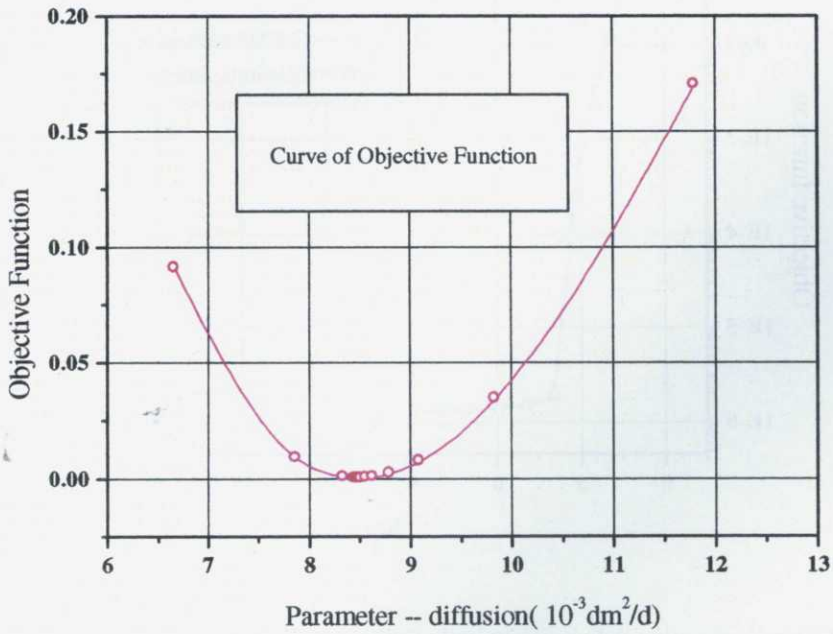


Figure 4.2.3A. Objective function as a function of the parameter (diffusion coefficient)

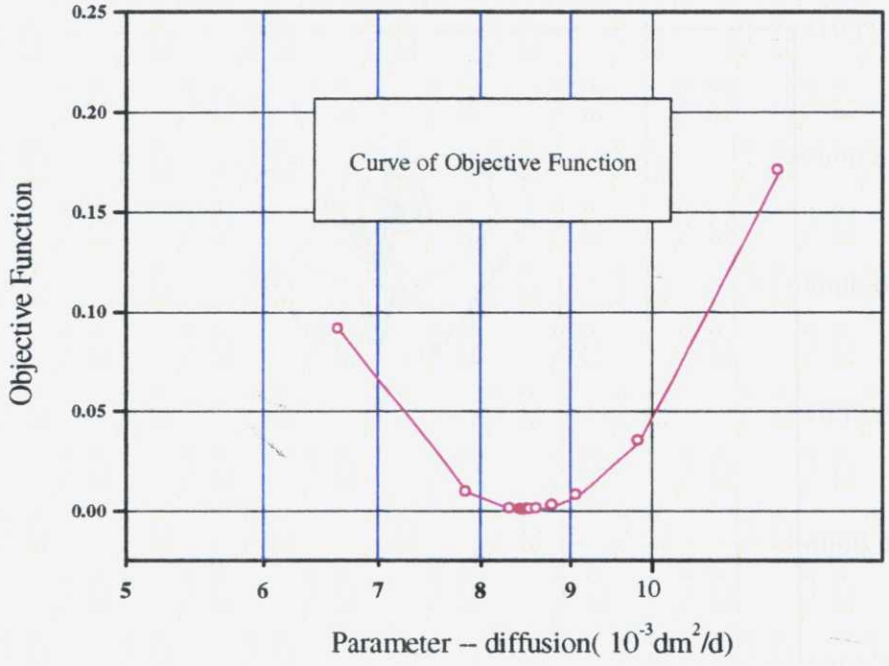


Figure 4.2.3B. Objective function as a function of log- parameter (diffusion coefficient)

4.2.1.3. Estimation of more than one parameter

In the second stage, both porosity and diffusion coefficient are estimated. The shape of the objective function is shown in Figure 4.2.4. It can be seen that this inverse problem is well-posed and has a single optimal solution. In fact, the minimum of the function in this plot is attained at the true parameter values.

In the third stage, the distribution coefficient is also estimated in addition to diffusion coefficient and porosity. Information on their initial values and upper and lower bounds are presented in Table 4.2.1.

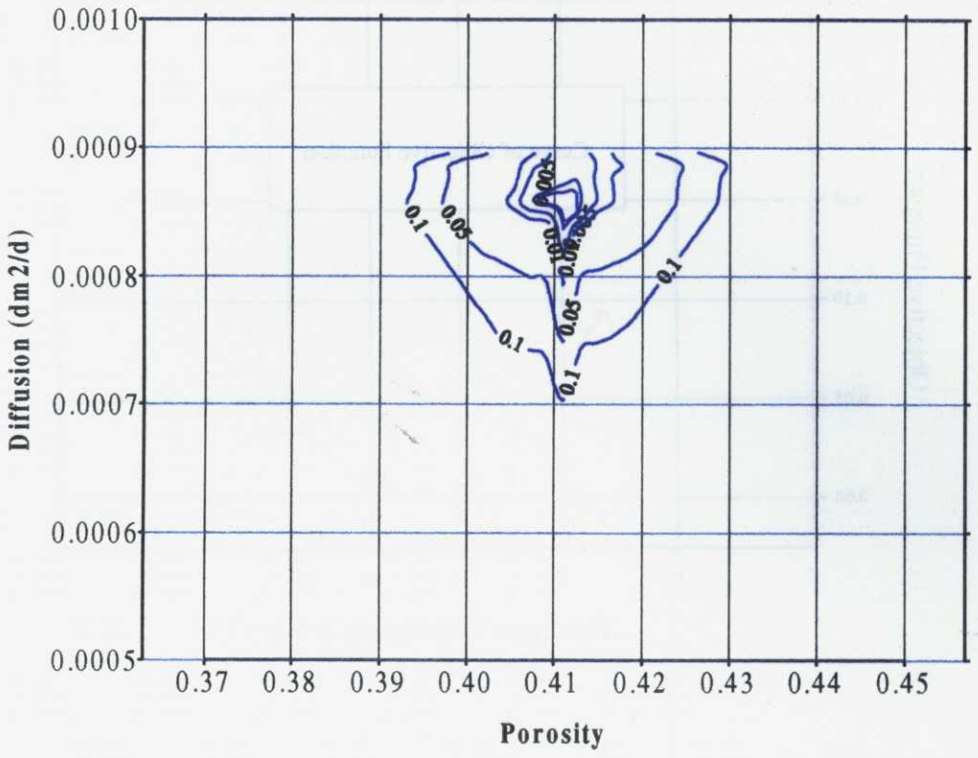


Figure 4.2.4. Objective function contour plot when two parameters are estimated

The two search methods were used to estimate the three parameters. Quadratic interpolation failed to provide an optimal solution possibly because it is very sensitive to the initial values of the parameters. Errors in initial parameter values can make this algorithm invalid. Another reason for the poor behavior of the quadratic method is that the true value of K_d is located at the lower bound of the interval [0, 1.5] which does not satisfy the interpolation condition $E(c) \leq \min [E(a), E(b)]$. However, the Golden section search method provides accurate parameter estimates after four iterative cycles. Final results are shown in Table 4.2.2. Figure 4.2.5 illustrates how the objective function varies during the four iterative cycles. During a given cycle, each parameter is estimated independently from the others. In order to ensure a global optimal solution, at the next cycle the parameters are searched on the same intervals as those of the previous cycles. Obviously, this procedure is time-consuming.

Table 4.2.2. Estimation results for the three-parameter estimation case using the Golden section search method

Parameter	D_0 (m^2/s)	K_d (ml/g)	ϕ	Obj. Function
True value	$9.83 \cdot 10^{-11}$	0	0.41	
Estimated value	$9.84 \cdot 10^{-11}$	$6.13 \cdot 10^{-5}$	0.41	$1.37 \cdot 10^{-6}$
Relative error(%)	0.11	/	0	

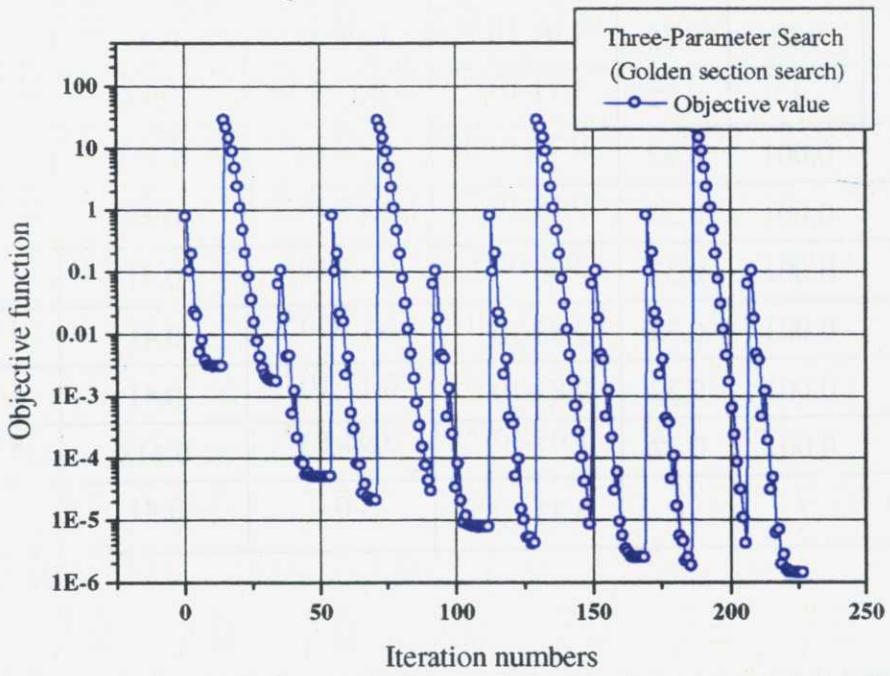


Figure 4.2.5. Reduction in objective function values achieved with the Golden section search for the three-parameter case during four iterative cycles

4.2.1.4. Sensitivity to initial parameter values

The ability of the inverse algorithm to provide accurate parameter estimates for different choices of initial parameter values has been tested by trying with different sets of

initial values of K_d and D_0. Golden section search with 5 iterative cycles was used in all cases. Estimated values are listed in Table 4.2.3.

Table 4.2.3. Sensitivity of estimation results to initial parameter values

Initial parameter values			Estimated values			
D_0 (m^2/s)	K_d (ml/g)	ϕ	D_0 (m^2/s)	K_d (ml/g)	ϕ	Objective
$8.68 \cdot 10^{-11}$	0	0.32	$9.83 \cdot 10^{-11}$	$6.13 \cdot 10^{-5}$	0.41	$1.37 \cdot 10^{-6}$
$8.68 \cdot 10^{-11}$	0.001	0.32	$9.84 \cdot 10^{-11}$	$6.13 \cdot 10^{-5}$	0.41	$1.47 \cdot 10^{-6}$
$8.68 \cdot 10^{-11}$	0.01	0.32	$1.07 \cdot 10^{-10}$	$4.04 \cdot 10^{-3}$	0.42	$1.5 \cdot 10^{-4}$
$8.68 \cdot 10^{-11}$	1	0.32	$1.71 \cdot 10^{-10}$	$2.8 \cdot 10^{-2}$	0.48	$7.89 \cdot 10^{-3}$
$2.89 \cdot 10^{-11}$	0.001	0.32	$9.71 \cdot 10^{-11}$	$6.13 \cdot 10^{-5}$	0.408	$3.54 \cdot 10^{-4}$
$6.37 \cdot 10^{-11}$	0.001	0.32	$9.79 \cdot 10^{-11}$	$6.13 \cdot 10^{-5}$	0.40	$4.44 \cdot 10^{-5}$
$7.52 \cdot 10^{-11}$	0.001	0.32	$9.8 \cdot 10^{-11}$	$6.13 \cdot 10^{-5}$	0.41	$1.91 \cdot 10^{-5}$
$8.68 \cdot 10^{-11}$	0.001	0.32	$9.83 \cdot 10^{-11}$	$6.13 \cdot 10^{-5}$	0.41	$4.69 \cdot 10^{-6}$
$9.49 \cdot 10^{-11}$	0.001	0.32	$9.83 \cdot 10^{-11}$	$6.13 \cdot 10^{-5}$	0.41	$1.47 \cdot 10^{-6}$
$1.1 \cdot 10^{-10}$	0.001	0.32	$1.05 \cdot 10^{-10}$	$2.94 \cdot 10^{-3}$	0.418	$8.08 \cdot 10^{-5}$
True values	/	/	$9.83 \cdot 10^{-11}$	0	0.41	/

It is found that model results are not sensitive to small changes in the initial values of the diffusion coefficient. A satisfactory solution is obtained almost always. If we run the model for more cycles, it is possible to achieve a more accurate solution. Estimation results, on the other hand, are very sensitive to the initial value of K_d. A poor initial value can make the search algorithm to converge to a local minimum. When the initial K_d=1, one cannot get a satisfactory solution.

4.2.2. SYNTHETIC EXAMPLE OF CATION EXCHANGE

This section is mainly devoted to the verification of INVGS-CORE and INVERSE-CORE^{2D} codes. INVGS-CORE was used first to estimate selectivity coefficients. If in addition to these parameters, one tries to estimate initial and boundary concentrations, the rate of convergence of the estimation procedure is a very slow. This case was used to test and verify INVERSE-CORE^{2D} using data with different amounts of noise.

4.2.2.1. Problem formulation (EJ2)

This case (EJ2) corresponds to a 0.1 m long laboratory column documented by Appelo and Postma (1993). The column, initially filled with a solution containing 1 mM $NaNO_3$ and 0.2 mM KNO_3, is flushed with a 0.6 mM $CaCl_2$ solution. This case illustrates the chromatographic separation of Na^+ and K^+. Na^+ is weakly adsorbed, and is eluted first. K^+ is more tenaciously held than Na^+, and is retarded in the column effluent.

Dispersivity is fixed to a value of 0.001 m while diffusion coefficient is assumed to be equal to 0. The flushing discharge is 0.003 m/h. The flow velocity is 0.01 m/h. Porosity is 0.3 and cation exchange capacity in terms of pore water is CEC = 1.1 meq/l. Nitrate is used as the main anion. Redox reactions which could affect NO_3^- are neglected.

Gaines-Thomas convention is used for cation exchange. In this convention, activities of exchanged cations are equal to equivalent fractions. Na^+ was selected as the reference cation. Therefore, Na^+ selectivity is equal to 1. Cation exchange reactions can be written as:

$$Na^+ + K\text{-}X = K^+ + Na\text{-}X \qquad K_{Na/K} = 0.1995$$

$$Na^+ + 0.5Ca\text{-}X = 0.5Ca^{2+} + Na\text{-}X \qquad K_{Na/Ca} = 0.3981$$

where $K_{Na/K}$ and $K_{Na/Ca}$ are potassium and calcium selectivities, respectively, with respect to sodium. The lower the selectivity, the higher the exchange capacity. The input of INVERSE-CORE^{2D} requires cation exchange capacity, CEC, expressed as meq per 100 g of solids, which can be derived from the following expression:

$$\text{CEC(meq/100g solid)} = \frac{\phi}{10\rho_s(1-\phi)} \cdot \text{CEC(meq/l)} \qquad (4.2.1)$$

where ρ_s is the density of the solids (Kg/l) which is equal to 2.65. By substituting these parameters, one has for this case CEC = 0.0213 meq/100g solid. Initial and boundary total dissolved concentrations are listed in Table 4.2.4.

Table 4.2.4. Initial and boundary total dissolved component concentrations (mol/l) (Appelo and Postma, 1992, and Samper *et* al., 1998).

Components	Initial	Boundary
H_2O	1	1
Na^+	10^{-3}	10^{-7}
K^+	$2 \cdot 10^{-4}$	10^{-7}
Ca^{+2}	10^{-7}	$6 \cdot 10^{-4}$
Cl^-	10^{-7}	$1.2 \cdot 10^{-3}$
NO_3^-	$1.2 \cdot 10^{-3}$	$1.2 \cdot 10^{-7}$

A grid of 102 nodes and 100 triangles was used. In this case, the breakthrough curves of dissolved Na^+, K^+, Ca^{2+}, and Cl^- at node 67, which were obtained by solving the forwards model, were taken as measured data. 25 data are available for each chemical component.

4.2.2.2. Estimation of selectivity coefficients

It is assumed that the two selectivity coefficients ($K_{Na/K}$, $K_{Na/Ca}$) are unknown and their values are within the interval (0.1, 1). INVGS-CORE was used to estimate these parameters. In the first run, synthetic data contain no noise and no prior information on the

parameters is included in the objective function. Eight iterations are required to achieve the optimum solution listed in Table 4.2.5.

Table 4.2.5. Estimation results of selectivity coefficients

Parameter	Lower bound	Upper bound	Initial guess	Tolerance error	True value	Estimated value
$K_{Na/K}$	0.1	1	0.12	10^{-5}	0.1995	0.1995
$K_{Na/Ca}$	0.1	1	0.45	10^{-5}	0.3981	0.3981

Plots of the sensitivity of the objective function to changes in selectivity coefficients are shown in Figures 4.2.6 and 4.2.7. One can see clearly that the objective function is a convex function near the optimum solution. Therefore, a unique minimum of the objective function can be found.

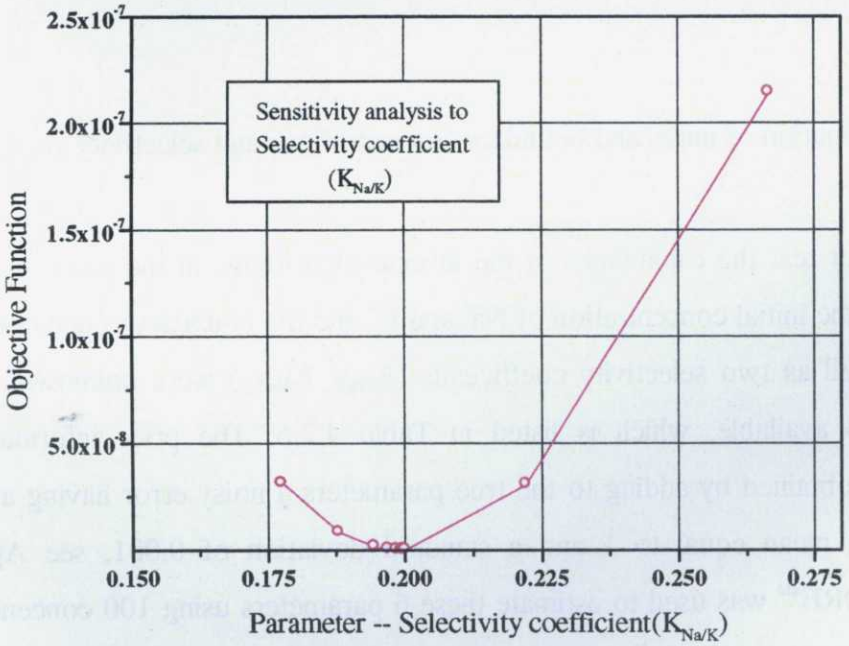


Figure 4.2.6. Sensitivity of the objective function to changes in the selectivity coefficient $K_{Na/K}$

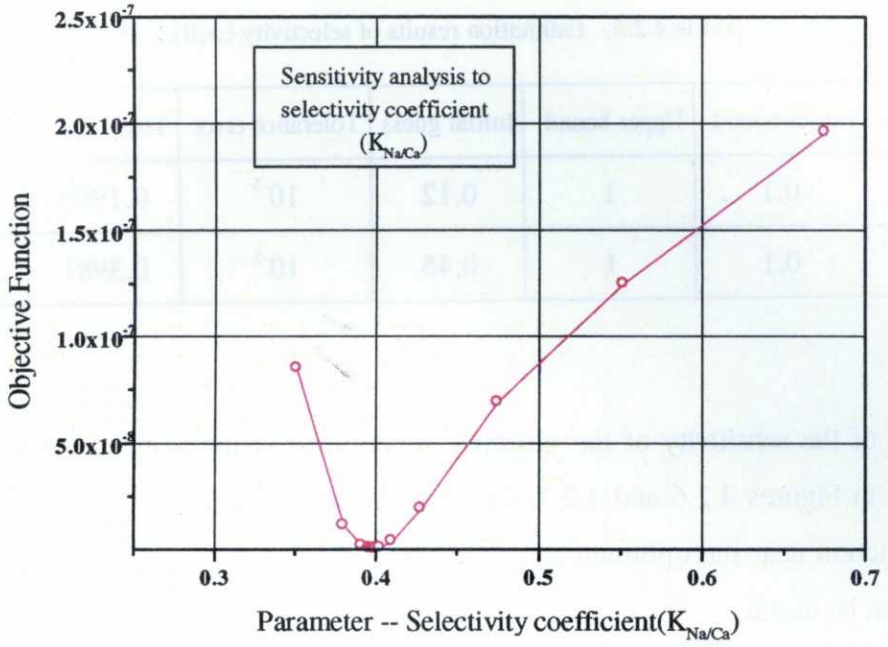


Figure 4.2.7. Sensitivity of the objective function to changes in the selectivity coefficient $K_{Na/Ca}$

4.2.2.3 Estimation of initial and boundary concentrations and selectivity coefficients

To further test the capabilities of the inverse algorithms, in the second stage, it was assumed that the initial concentration of Na^+ and K^+ and the boundary concentration of Ca^{+2} and Cl^-, as well as two selectivity coefficients ($K_{Na/K}$, $K_{Na/Ca}$) were unknown. Some prior information is available, which is listed in Table 4.2.6. The prior information of the parameters is obtained by adding to the true parameters a noisy error having a log-normal distribution (a mean equal to 1 and a standard deviation of 0.001, see Appendix 2). INVERSE-CORE^{2D} was used to estimate these 6 parameters using 100 concentration data of four species (Na^+, K^+, Ca^{2+} and Cl^-) at node 67 which were corrupted with different levels of random noise (see Appendix 2). Estimation results are listed in Tables 4.2.6 through 4.2.9, Further details are provided in Appendices 3 and 4.

Table 4.2.6. Information of the parameters to be estimated (C_0 and C_B are initial and boundary concentrations, respectively expressed as mol/l)

Parameter	Lower bound	Upper bound	Initial guess	True value	Prior information
$K_{Na/K}$	0.1	0.3	0.12	0.1995	0.1994
$K_{Na/Ca}$	0.25	0.42	0.41	0.3981	0.3982
$C_{0\,Na}$	$8\cdot10^{-4}$	$8\cdot10^{-3}$	$3\cdot10^{-3}$	10^{-3}	$1.002\cdot10^{-3}$
$C_{0\,K}$	10^{-4}	10^{-3}	$3\cdot10^{-4}$	$2\cdot10^{-4}$	$1.997\cdot10^{-4}$
$C_{B\,Ca}$	$3\cdot10^{-4}$	$8\cdot10^{-4}$	$5.5\cdot10^{-4}$	$6\cdot10^{-4}$	$5.996\cdot10^{-4}$
$C_{B\,Cl}$	$5\cdot10^{-4}$	$5\cdot10^{-3}$	$1.4\cdot10^{-4}$	$1.2\cdot10^{-3}$	$1.201\cdot10^{-3}$

Table 4.2.7. Estimation results using synthetic concentration data containing different degrees of noise (mean μ_l=1 and error σ_l). No prior information about parameters is considered (concentrations are expressed as mol/l)

Parameter		Estimated Results for different error (σ_l)				
Name	True value	σ_l=0	σ_l=0.01	σ_l=0.05	σ_l=0.1	σ_l=0.5
$K_{Na/K}$	0.1995	0.1995	0.1986	0.1951	0.1910	0.1644
$K_{Na/Ca}$	0.3981	0.3980	0.4003	0.4077	0.4149	0.3849
$C_{0\,Na}$	10^{-3}	10^{-3}	$9.975\cdot10^{-4}$	$9.886\cdot10^{-4}$	$9.797\cdot10^{-4}$	$9.920\cdot10^{-4}$
$C_{0\,K}$	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$1.993\cdot10^{-4}$	$1.971\cdot10^{-4}$	$1.947\cdot10^{-4}$	$1.868\cdot10^{-4}$
$C_{B\,Ca}$	$6\cdot10^{-4}$	$6\cdot10^{-4}$	$6.007\cdot10^{-4}$	$6.032\cdot10^{-4}$	$6.058\cdot10^{-4}$	$6.108\cdot10^{-4}$
$C_{B\,Cl}$	$1.2\cdot10^{-3}$	$1.2\cdot10^{-3}$	$1.201\cdot10^{-3}$	$1.204\cdot10^{-3}$	$1.208\cdot10^{-3}$	$1.194\cdot10^{-3}$
Number of iterations		6	5	7	7	6
Objective function		$5.34\cdot10^{-5}$	0.44	10.85	42.58	619.81

Table 4.2.8. Estimation results using synthetic concentration data containing different degrees of noise (mean μ_l=1 and error σ_l). Parameter prior information is considered in the objective function (w_p=1) (concentrations are expressed as mol/l)

Parameter		Estimated Results for different error (σ_l)				
Name	True value	σ_l=0	σ_l=0.01	σ_l=0.05	σ_l=0.1	σ_l=0.5
$K_{Na/K}$	0.1995	0.1994	0.1994	0.1991	0.1989	0.1977
$K_{Na/Ca}$	0.3981	0.3982	0.3982	0.3983	0.3984	0.3984
$C_{0\,Na}$	$1\cdot10^{-3}$	$1\cdot10^{-3}$	$9.977\cdot10^{-4}$	$9.892\cdot10^{-4}$	$9.806\cdot10^{-4}$	$9.811\cdot10^{-4}$
$C_{0\,K}$	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$2.001\cdot10^{-4}$	$2.009\cdot10^{-4}$	$2.002\cdot10^{-4}$	$2.160\cdot10^{-4}$
$C_{B\,Ca}$	$6\cdot10^{-4}$	$6\cdot10^{-4}$	$6.002\cdot10^{-4}$	$6.006\cdot10^{-4}$	$6.015\cdot10^{-4}$	$6.136\cdot10^{-4}$
$C_{B\,Cl}$	$1.2\cdot10^{-3}$	$1.2\cdot10^{-3}$	$1.201\cdot10^{-3}$	$1.204\cdot10^{-3}$	$1.208\cdot10^{-3}$	$1.194\cdot10^{-3}$
Number of iterations		5	5	5	5	4
Objective function		$1.21\cdot10^{-4}$	0.44	11.04	43.25	625.25

Table 4.2.9. Estimation results using synthetic concentration data containing different degrees of noise (mean μ_l=1 and error σ_l). Parameter prior information is considered in the objective function (w_p=10) (concentrations are expressed as mol/l)

Parameter		Estimated Results for different error (σ_l)				
Name	True value	σ_l=0	σ_l=0.01	σ_l=0.05	σ_l=0.1	σ_l=0.5
$K_{Na/K}$	0.1995	0.1994	0.1994	0.1994	0.1994	0.1992
$K_{Na/Ca}$	0.3981	0.3982	0.3982	0.3982	0.3982	0.3982
$C_{0\,Na}$	$1\cdot10^{-3}$	$1\cdot10^{-3}$	$9.979\cdot10^{-4}$	$9.902\cdot10^{-4}$	$9.807\cdot10^{-4}$	$9.825\cdot10^{-4}$
$C_{0\,K}$	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$2.002\cdot10^{-4}$	$2.012\cdot10^{-4}$	$2.020\cdot10^{-4}$	$2.182\cdot10^{-4}$
$C_{B\,Ca}$	$6\cdot10^{-4}$	$6\cdot10^{-4}$	$6.001\cdot10^{-4}$	$6.007\cdot10^{-4}$	$6.012\cdot10^{-4}$	$6.134\cdot10^{-4}$
$C_{B\,Cl}$	$1.2\cdot10^{-3}$	$1.2\cdot10^{-3}$	$1.201\cdot10^{-3}$	$1.204\cdot10^{-3}$	$1.207\cdot10^{-3}$	$1.195\cdot10^{-3}$
Number of iterations		5	3	5	3	3
Objective function		$2.29\cdot10^{-4}$	0.44	11.06	43.31	625.81

One can see in Table 4.2.7 that when concentration data are free of noise, 6 iterations are needed to obtained parameter estimates which coincide with the true values. Computed concentrations of Na^+, K^+, Ca^{2+} and Cl^- match perfectly the concentrations data (Figure 4.2.8). Sensitivity analysis of the objective function to changes in the Ca^{2+} boundary concentration is shown in Figure 4.2.9. It is shows clearly that the objective function is very sensitive to changes in Ca^{2+} boundary concentration.

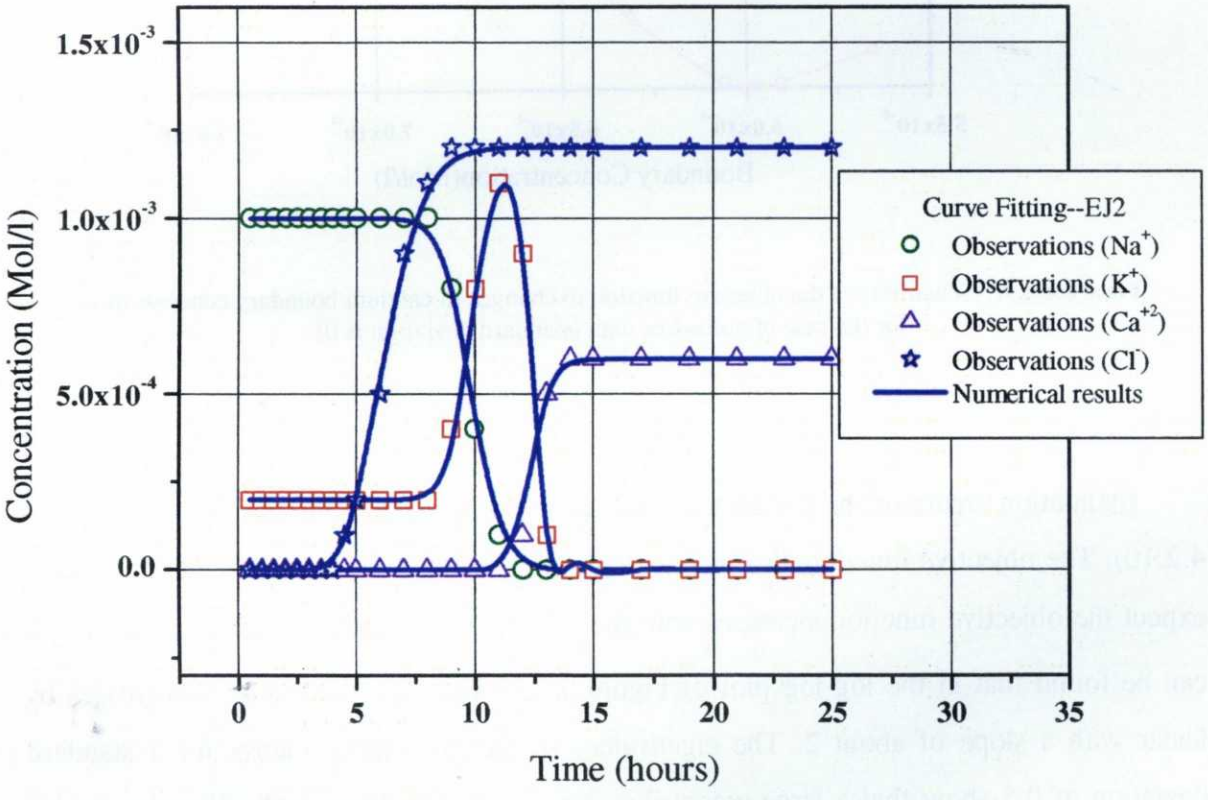


Figure 4.2.8. Concentration evolution and best fit for the cation exchange verification case using noise-free data (standard deviation is 0)

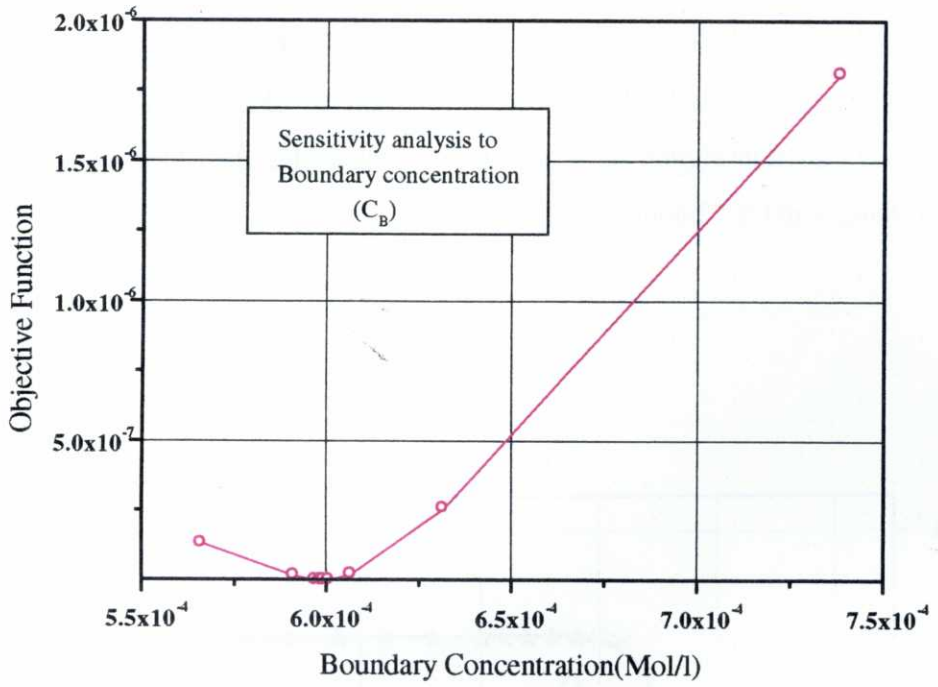


Figure 4.2.9. Sensitivity of the objective function to changes in calcium boundary concentration for the case of noise-free data (standard deviation is 0)

Estimation errors of the parameters increase with increasing amount of noise (Figure 4.2.10). The objective function also increases greatly (Figure 4.2.11). In fact, as one would expect the objective function increases with the square of the standard deviation σ_l. So, it can be found that in the log-log plot of Figure 4.2.11 the objective function increases by linear with a slope of about 2. The eigenvalues of the covariance matrix for a standard deviation of 0.5 show that a large eigenvalue is associated to $K_{Na/K}$, indicating that in this case this parameter is the most uncertain.

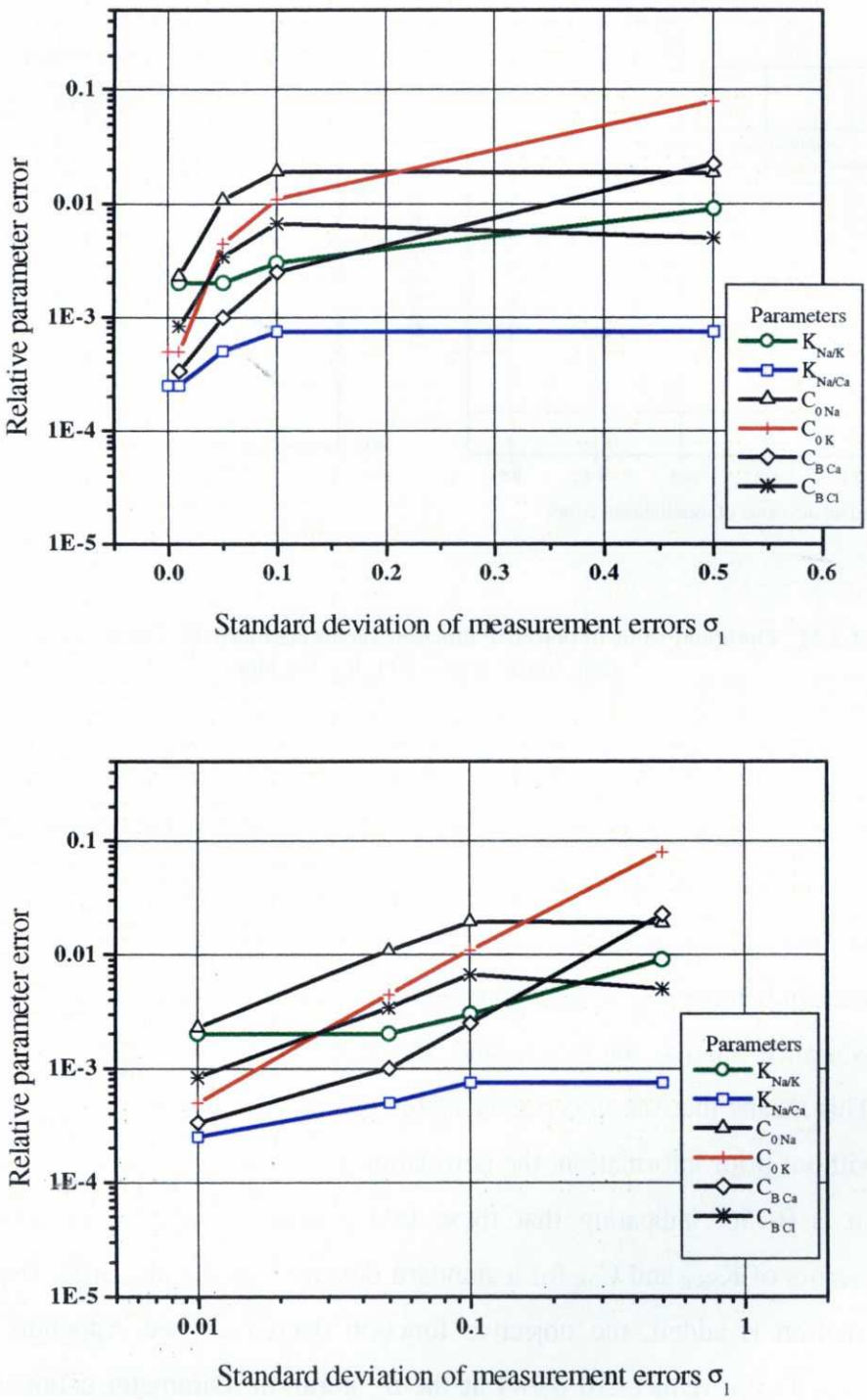


Figure 4.2.10. Relative error of parameter estimation versus standard deviations of measurement error (σ_1) without parameter prior information. Semi-log plot (above) and log-log plot (below)

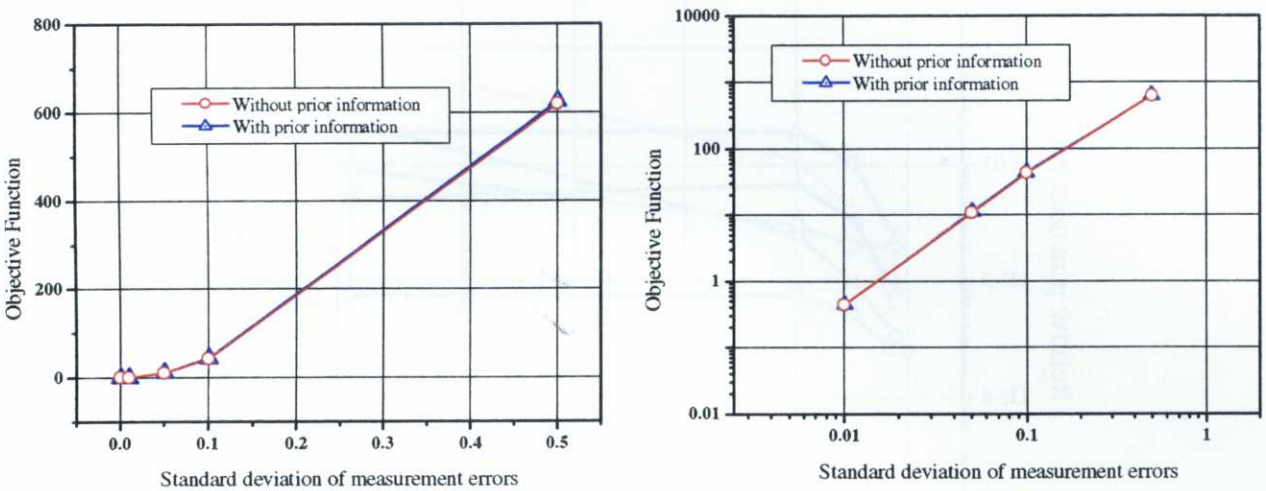


Figure 4.2.11. Optimum value of objective function versus standard deviation of observation data (left, linear scale; right, log-log plot)

Using the same data as those in Table 4.2.7 but adding prior parameter information also corrupted with log-normal distributed errors, one obtains similar results (Table 4.2.8) but with less iterations (Figure 4.2.12). The computing efficiency is increased by adding prior information. Moreover, by including prior information of the parameters the structure of the covariance matrix improves, and the correlation coefficients and eigenvalues decrease. This means that the uncertainty of the estimated parameters is also reduced. For example, without prior information, the correlation coefficient between $K_{Na/K}$ and C_{0K} at the 6^{th} iteration is 0.756, indicating that these two parameters are highly related. So, the estimation errors of $K_{Na/K}$ and C_{0K} for a standard deviation of 0.5 are large. However, when prior information is added, the objective function decreases (see Appendix 4), and the correlation coefficient reduces to 0.284 at the 5^{th} iteration. Parameter estimates, when the standard deviations are greater than 0 (Table 4.2.8), are much better than those obtained without prior information (Table 4.2.7).

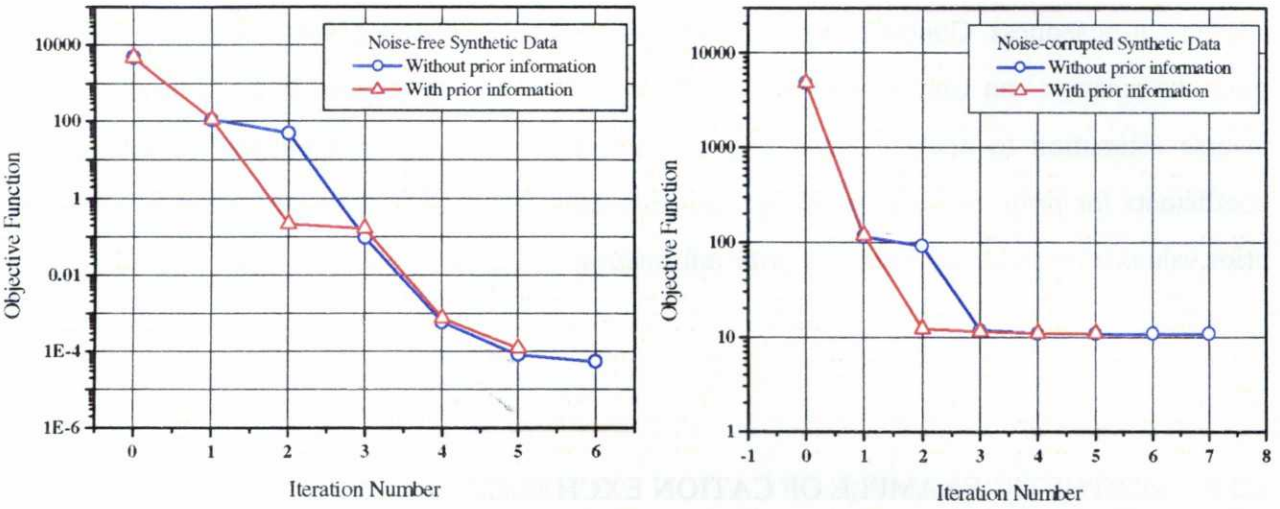


Figure 4.2.12. Reduction of objective function with iteration number for different conditions (noise-free synthetic data on the left and noise-corrupted data with a standard deviation of 0.05 on the right)

Appendices 3 and 4 contain the details of the statistical analyses and optimization process. One can see that initial and boundary concentrations converge to their true values faster than the selectivity coefficients. This is due to the fact that the objective function is quadratic in initial and boundary concentrations. Its dependence on selectivity coefficients, however, is highly nonlinear. In order to improve the convergence rate and optimization efficiency for these highly-nonlinear parameters, our inverse codes provide the possibility of performing a logarithmic transformation to decrease the nonlinearity of the problem. This option is discussed further in Chapter 5 in the context of the application to real cases.

When the weighting coefficient of the prior information (w_p) in the objective function is increased from 1 (Table 4.2.8) to 10 (Table 4.2.9), estimated parameters tend to be close to their prior information. This result indicates that if the weighting coefficient is too large, the estimated parameters tend to be close to their prior estimates and may deviate from the true values.

It can be concluded from this synthetic example that reliable prior information with suitable weighting can reduce estimation variances, correlation coefficients and eigenvalues of the covariance matrix, and improve the efficiency of the optimization algorithms and alleviate ill-posedness. Contrary to the opinion of Ginn and Cushman (1990) who state that prior information can only make 'worse' the model, it is very important for a suitable inverse estimation to analyze carefully prior information and select adequate weighting coefficients for prior information. Otherwise, estimation results might be very close to the prior values or even be corrupted by prior information.

4.2.3. SYNTHETIC EXAMPLE OF CATION EXCHANGE AND MINERAL DISSOLUTION-PRECIPITATION

4.2.3.1. Problem formulation (EJ2KIN)

This example (EJ2KIN) is based on the cation exchange example model (EJ2) descripted in the previous section. Kinetically-controlled calcite dissolution/ precipitation is included in addition to all the chemical processes considered there. Calcite dissolution is described by

$$\text{Calcite} + H^+ = Ca^{2+} + HCO_3^-$$

This problem contains the following 8 primary aqueous species: H_2O, Na^+, K^+, Ca^{2+}, Cl^-, HCO_3^-, NO_3^- and H^+. Chemical speciation of the initial water performed with EQ3 (Wolery, 1983) indicates that the following 11 species must be considered as secondary species, corresponding to aqueous dissociation reactions (pK = -log (K) at 25 ^{0}C):

$OH^- = H_2O - H^+$	pK = -13.995
$CO_3^{2-} = HCO_3^- - H^+$	pK = -10.329
$CO_2(aq) = HCO_3^- + H^+ - H_2O$	pK = 6.345
$NaHCO_3(aq) = Na^+ + HCO_3^-$	pK = 0.1541
$CaHCO_3^+ = Ca^{2+} + HCO_3^-$	pK = 1.0467
$NaCO_3^-(aq) = Na^+ + HCO_3^- - H^+$	pK = -9.814

$CaCO_3(aq) = Ca^{2+} + HCO_3^- - H^+$	pK = -7.002
$NaCl(aq) = Na^+ + Cl^-$	pK = -0.777
$NaOH(aq) = Na^+ + OH^-$	pK = -14.175
$KOH(aq) = K^+ + OH^-$	pK = -14.46
$CaOH^+ = Ca^{+2} + OH^-$	pK = -12.85

Initial and boundary total dissolved concentrations are listed in Table 4.2.10. Model discretization and its parameters are the same as those of the previous example. Breakthrough curves of Na^+, K^+, Ca^{2+}, Cl^-, and HCO_3^- at node 67 were obtained by solving the forwards problem. These curves were taken as measured data for the solution of the inverse problem.

Table 4.2.10. Initial and boundary total dissolved concentrations (mol/l) of the components of problem EJ2KIN

Components	Initial	Boundary
H_2O	1	1
Na^+	10^{-3}	10^{-7}
K^+	$2 \cdot 10^{-4}$	10^{-7}
Ca^{+2}	10^{-7}	$6 \cdot 10^{-4}$
Cl^-	10^{-7}	$1.2 \cdot 10^{-3}$
NO_3^-	$1.2 \cdot 10^{-3}$	$1.2 \cdot 10^{-7}$
HCO_3^-	$1.2 \cdot 10^{-5}$	$1.2 \cdot 10^{-5}$
H^+	$1.2 \cdot 10^{-8}$	$1.2 \cdot 10^{-7}$

4.2.3.2 Parameter estimation

The following nine parameters were assumed to be unknown and therefore were estimated: selectivity coefficients ($K_{Na/K}$ and $K_{Na/Ca}$), Na and K initial concentrations ($C_{0\,Na}$ and $C_{0\,K}$), Ca and Cl boundary concentrations ($C_{B\,Ca}$ and $C_{B\,Cl}$), specific surface (A) and kinetic rate constant (R_k) of calcite, and the exponent of the ratio of the ionic activity product to the equilibrium constant (θ). In Table 4.2.11, the initial guess of the parameters

are arbitrary values between the lower and upper bounds and the prior information of the parameters was obtained by adding to the true parameters an error having a log-normal distribution (mean equal to 1 and standard deviation of 0.001). INVERSE-CORE^{2D} has been used to estimate these parameters with and without parameter prior information. Estimation results are listed in Tables 4.2.12 and 4.2.13.

Table 4.2.11. Information of the parameters to be estimated. (C_0 and C_B are initial and boundary concentrations, $K_{Na/K}$ and $K_{Na/Ca}$ are the selectivity coefficients, A is specific surface of calcite (dm^2/dm^3), R_k is the rate constant ($mol/m^2/s$) and θ is the exponent of the ratio of the ion activity product to the equilibrium constant)

Parameter	Lower bound	Upper bound	Initial guess	True value	Prior information
$K_{Na/K}$	0.1	0.3	0.12	0.1995	0.1994
$K_{Na/Ca}$	0.25	0.42	0.41	0.3981	0.3997
$C_{0\,Na}$	$8\cdot10^{-4}$	$8\cdot10^{-3}$	$3\cdot10^{-3}$	10^{-3}	$1.001\cdot10^{-3}$
$C_{0\,K}$	10^{-4}	10^{-3}	$3\cdot10^{-4}$	$2\cdot10^{-4}$	$2.002\cdot10^{-4}$
$C_{B\,Ca}$	$3\cdot10^{-4}$	$8\cdot10^{-4}$	$5.5\cdot10^{-4}$	$6\cdot10^{-4}$	$5.997\cdot10^{-4}$
$C_{B\,Cl}$	$5\cdot10^{-4}$	$5\cdot10^{-3}$	$1.4\cdot10^{-4}$	$1.2\cdot10^{-3}$	$1.199\cdot10^{-3}$
A	16	22	18	20	19.88
R_k	10^{-8}	10^{-7}	$6\cdot10^{-8}$	$5\cdot10^{-8}$	$4.995\cdot10^{-8}$
θ	0.8	1.2	0.9	1	1.001

One can see in Table 4.2.12 that by ignoring prior information the accuracy of the estimated parameters decreases with increasing amount of noise. The estimated values of the first six parameters ($K_{Na/K}$, $K_{Na/Ca}$, $C_{0\,Na}$, $C_{0\,K}$, $C_{B\,Ca}$ and $C_{B\,Cl}$) are close to their true values. The estimation errors for the kinetic parameters (A, R_k and θ) are large except for free-of-noise data (the standard deviation is 0). Only in this case the estimated results of these three parameters are close to their true values. When the standard deviation of the noise is greater than 0.05, the estimates of A and θ reach their bounds. Their corresponding

eigenvalues in the covariance matrix are large. This means that the estimates of kinetic parameters (A, R_k and θ) have large uncertainty. In fact, from Figure 4.2.13 one can see that the specific surface (A) and the rate constant (R_k) are highly correlated. When one of them increases, the other decreases, and their product is close to the product of their true values (see Figure 4.2.13.).

Table 4.2.12. Estimated results using synthetic concentration data with different levels of noise (μ_l=1) without prior information (w_p=0). See Table 4.2.11 for the definition of the other parameters.

Parameter		Estimated Results for different error (σ_l)				
Name	True value	σ_l=0	σ_l=0.01	σ_l=0.05	σ_l=0.1	σ_l=0.5
$K_{Na/K}$	0.1995	0.1994	0.2002	0.2089	0.2158	0.25
$K_{Na/Ca}$	0.3981	0.3988	0.4006	0.4144	0.4106	0.3943
$C_{0\,Na}$	10^{-3}	$9.998\cdot10^{-4}$	$9.954\cdot10^{-4}$	$9.769\cdot10^{-4}$	$9.662\cdot10^{-4}$	$9.326\cdot10^{-4}$
$C_{0\,K}$	$2\cdot10^{-4}$	$1.998\cdot10^{-4}$	$1.989\cdot10^{-4}$	$1.967\cdot10^{-4}$	$1.976\cdot10^{-4}$	$2.049\cdot10^{-4}$
$C_{B\,Ca}$	$6\cdot10^{-4}$	$6.001\cdot10^{-4}$	$6.008\cdot10^{-4}$	$6.054\cdot10^{-4}$	$6.052\cdot10^{-4}$	$5.959\cdot10^{-4}$
$C_{B\,Cl}$	$1.2\cdot10^{-3}$	$1.2\cdot10^{-3}$	$1.201\cdot10^{-3}$	$1.204\cdot10^{-3}$	$1.206\cdot10^{-3}$	$1.184\cdot10^{-3}$
A	20	20.55	16	22	16	16
R_k	$5\cdot10^{-8}$	$4.779\cdot10^{-8}$	$6.\cdot109\cdot10^{-8}$	$4.396\cdot10^{-8}$	$6.15\cdot10^{-8}$	$6.793\cdot10^{-8}$
θ	1	1.07	1.172	1.2	1.2	1.2
Number of iterations		13	13	6	5	5
Objective function		$9.62\cdot10^{-3}$	0.90	22.65	86.36	1066.9

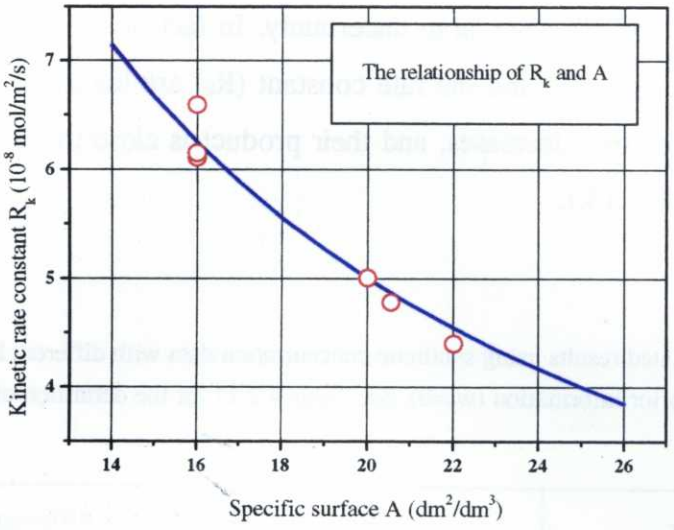


Figure 4.2.13. Estimates of specific surface (A) and kinetic rate constant (R_k) showing a negative correlation (symbols are estimated parameters taken from Table 4.2.12 and the line is the function R_k A = 10^{-6}).

The accuracy of the estimated parameters decreases with increasing amount of noise (see Figure 4.2.14) except one point of the parameter $K_{Na/cCa}$ at a standard deviation of 0.05 in which the estimated value is close to prior parameter information. The objective function increases with increasing standard deviation (see Figure 4.2.15). However, comparing with estimated parameter values listed in Table 4.2.12, better results are obtained when parameter prior information is included into the objective function (see Table 4.2.13). The correlation between kinetic parameters (A and R_k) is greatly reduced. Their estimates are good even when the standard deviation of the noise level is greater than 0.05. Their corresponding eigenvalues are also reduced. Therefore, a conclusion can be made that by adding the prior information into the objective function with some proper weight (w_p=1) one can improve the structure of the objective function, as well as the corresponding covariance matrix, reduce the correlation among troublesome parameters, increase the stability of the minimization algorithms, and, once again, reduce the number of iterations required to reach the optimum.

The optimum relative parameter increment for computing the Jacobian matrix by finite difference has been found to be 10^{-5}. With this value the Gauss-Newton-Levenberg-Marquardt algorithm converges fast. However, when it is changed between 10^{-3} and 10^{-6}, it has very little influence on the estimated results, although sometimes it may require a few more iterations. It has been found that the Gauss-Newton-Marquardt algorithm is a flexible and robust optimization tool which can overcome the problem of small errors in the sensitivity coefficients which were computed by finite differences with different relative parameter increments. When this parameter is out of range, the accuracy of the sensitivity coefficients is poor, and it is difficult to get a convergent solution.

Table 4.2.13. Estimated results using synthetic concentration data with different levels of noise (μ_l=1) and considering prior parameter information (w_p=1). See Table 4.2.11 for the definition of the parameters

Parameter		Estimated Results for different error (σ_l)				
Name	True value	σ_l=0	σ_l=0.01	σ_l=0.05	σ_l=0.1	σ_l=0.5
$K_{Na/K}$	0.1995	0.1995	0.1991	0.1977	0.1960	0.1888
$K_{Na/Ca}$	0.3981	0.3982	0.3987	0.3982	0.3955	0.3184
$C_{0\,Na}$	10^{-3}	10^{-3}	$1.001\cdot10^{-3}$	$1.007\cdot10^{-3}$	$1.013\cdot10^{-3}$	$1.05\cdot10^{-3}$
$C_{0\,K}$	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$2.005\cdot10^{-4}$	$2.025\cdot10^{-4}$	$2.051\cdot10^{-4}$	$2.265\cdot10^{-4}$
$C_{B\,Ca}$	$6\cdot10^{-4}$	$6\cdot10^{-4}$	$6\cdot10^{-4}$	$5.992\cdot10^{-4}$	$5.976\cdot10^{-4}$	$5.675\cdot10^{-4}$
$C_{B\,Cl}$	$1.2\cdot10^{-3}$	$1.2\cdot10^{-3}$	$1.201\cdot10^{-3}$	$1.204\cdot10^{-3}$	$1.206\cdot10^{-3}$	$1.179\cdot10^{-3}$
A	20	19.98	19.98	19.98	19.98	19.978
R_k	$5\cdot10^{-8}$	$4.949\cdot10^{-8}$	$4.96\cdot10^{-8}$	$5\cdot10^{-8}$	$5.068\cdot10^{-8}$	$5.694\cdot10^{-8}$
θ	1	1	1	1.001	1.001	1.003
Number of iterations		6	7	8	6	6
Objective function		$2.26\cdot10^{-2}$	1	23.15	88.40	1081.9

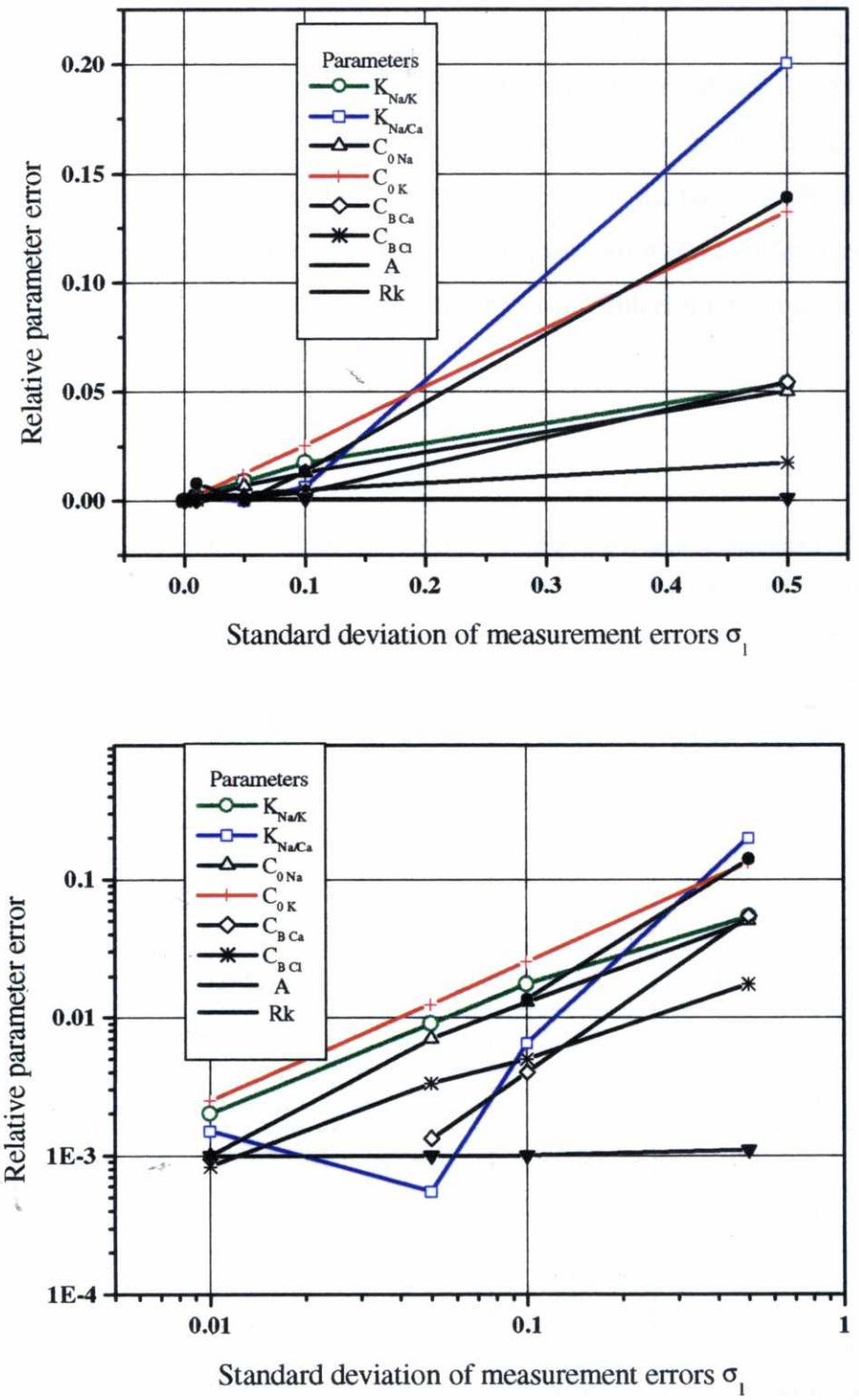


Figure 4.2.14. Relative error of parameter estimation versus standard deviations of measurement error (σ_1) with parameter prior information. Semi-log plot (above) and log-log plot (below)

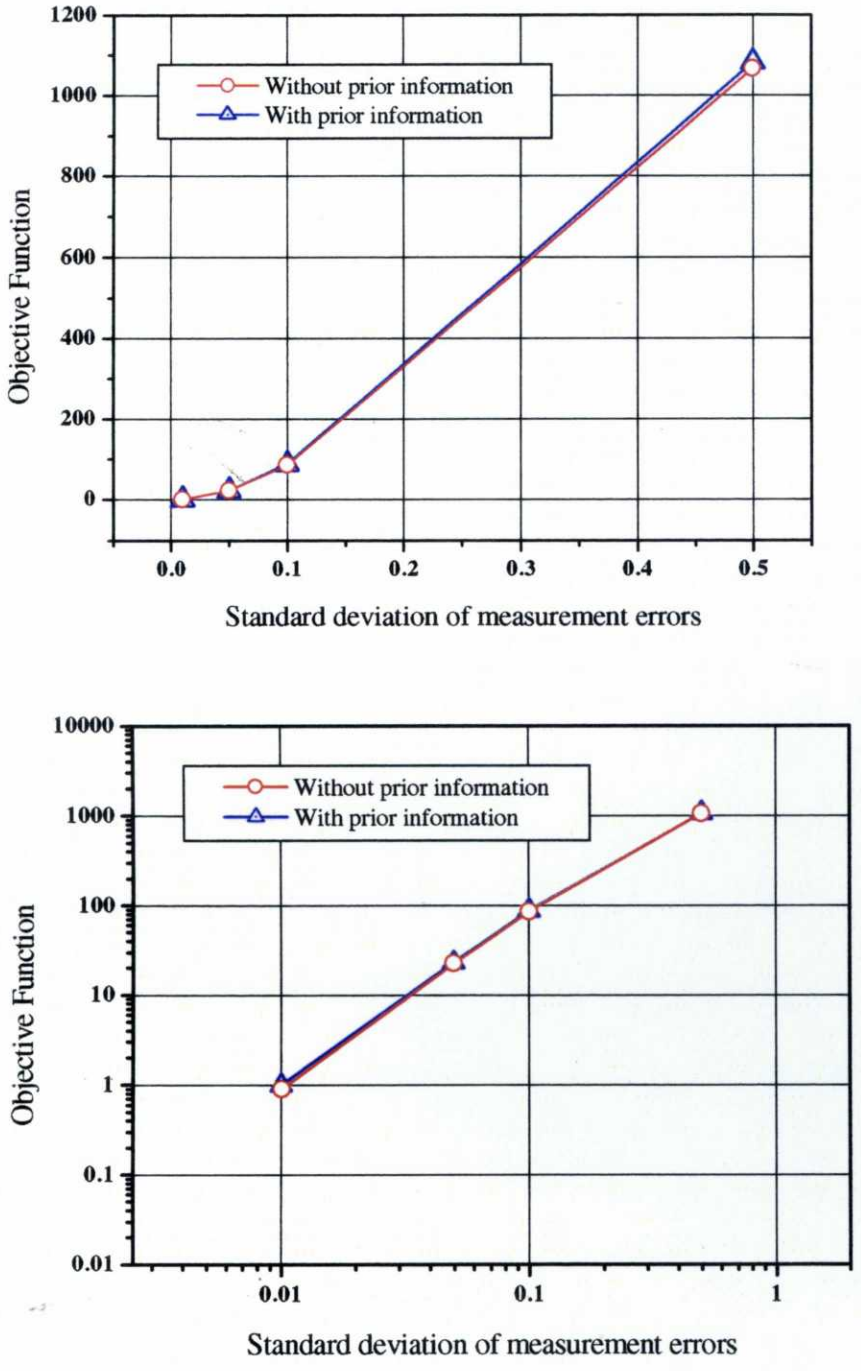


Figure 4.2.15. Optimum value of objective function versus standard deviation of observation data (above, linear scale; below, log-log plot)

CHAPTER 5

APPLICATIONS TO LABORATORY AND FIELD CASES

This chapter contains the application of the inverse analysis to several real cases involving laboratory and field data. These cases demonstrate the capabilities of the inverse modeling codes to deal with laboratory and field scale problems.

5.1. INTERPRETATION OF INFILTRATION EXPERIMENT DATA

5.1.1. INFILTRATION EXPERIMENTS

A series of infiltration experiments on compacted FEBEX bentonite were performed by CIEMAT (Villar and Cuevas, 1996). The FEBEX bentonite is also called "Serrata" clay which was extracted from the Cortijo de Archidona deposit, exploited by Minas de Gádor, S. A., in the zone of Serrata de Níjar (Almería, Spain). This deposit was selected by ENRESA prior to the FEBEX project as the most suitable material for the backfilling and sealing of a HLW repository. The reasons for this selection are its very high content of montmorillonite, large swelling pressure, low permeability, acceptable thermal conductivity, good retention properties and ease of compaction for the fabrication of blocks. Furthermore, compacted bentonite was packaged as homogeneous as possible. Homogenization reduces the uncertainties in modeling, in laboratory results and in the final interpretation of the entire test.

These infiltration experiments include 5 tests (SAT1, SAT2, SAT3, SAT4 and SAT5) performed in steel cylindrical cells of 5 cm of inner diameter and 2.5 cm of length. Compacted bentonite samples were confined between two blocks of porous sinters. They were hydrated from the upper end under a constant water pressure of 1 MPa. Infiltration

rate was recorded with time. At the end of the test, the sample was taken out of the cell. Final water content (w) and dry density (ρ_d) were measured at five sections located at different distances to the hydration front (Table 5.1.1).

Table 5.1.1. Summary of the infiltration experiment data

Test	SAT1	SAT2	SAT3	SAT4	SAT5
ρ_d (g/cm^3)	1.56	1.58	1.65	1.64	1.58
Initial w (%)	14.2	14.2	13.9	14.1	14.1
Test duration (d)	29	15	10	10	59
Final w (%)	29.14	25.8	24.3	23.98	23.98

5.1.2. SELECTION OF RELATIVE PERMEABILITY FUNCTIONS

It has been recognized that selecting a suitable model for relative permeability function is needed to establish a well posed inverse model for the estimations of unsaturated parameters. Using a robust optimization methodology can increase the possibility of getting stable and convergent solutions. Golden section search and Gauss-Newton-Marquardt methods have been used to estimate unsaturated flow parameters of compacted bentonite from infiltration data. Originally, only two relative permeability functions were included in INVERSE-CORE^{2D} (Equations 5.1.1 and 5.1.3) corresponding to those proposed by Irmay (1954) and van Genuchten (1980). Initial results indicated that the widely used and recommended function of van Genuchten (van Genuchten, 1980, Russo, 1988 and Russo et al., 1991) denoted as vG-M0 did not perform better than Irmay model. With the intention of making up for the weakness of vG-M0 model, two other more complex models, vG-M1 (Equation. 5.1.4) and vG-B0 (Equation 5.1.5) (van Genuchten, 1980) were introduced, even though van Genuchten (1980) suggested that these two models do not present an attractive alternative. The results obtained from these four functions, that of Irmay (1954) and the three of van Genuchten (1980), have been analyzed and compared.

Assuming that the resistance to flow offered by the solid matrix is proportional to the solid-liquid interfacial area and for an approximate cubic arrangement of spherical grains, Irmay (1954) derived an equation of the type

$$K_r(S_e) = S_e^{n_i} \tag{5.1.1}$$

where n_i is the power coefficient of the Irmay model (which usually takes a value of 3) and S_e is the effective saturation degree defined as

$$S_e = \frac{S_w - S_r}{1 - S_r} \tag{5.1.2}$$

Based on Mualem's theory (1976) for predicting the $K_r(\psi)$ function from soil-water retention data, van Genuchten (1980) developed two relative permeability functions corresponding to different relationships between m and n. The parameter m is related to n by $k = m\text{-}1 + n^{-1}$, where k is an integer. For $k = 0$ (i.e., $m = 1 - n^{-1}$), van Genuchten obtained the commonly used expression:

$$K_r(\psi) = \sqrt{S_e}[1 - (1 - S_e^{1/m})^m]^2 \qquad (m = 1\text{-}1/n,\ n = (1\text{-}m)^{-1}) \tag{5.1.3}$$

For $k = 1$ (i.e., $m=2 - n^{-1}$), the following relative permeability function is obtained,

$$K_r(S_e) = \sqrt{S_e}[1 - m(1 - S_e^{1/m})^{m-1} + (m-1)(1 - S_e^{1/m})^m]^2 \quad (m= 2\text{-}1/n,\ n = (2\text{-}m)^{-1}) \tag{5.1.4}$$

Based on Burdine's theory (1953), van Genuchten derived another model for relative permeability,

$$K_r(S_e) = S_e^2[1 - (1 - S_e^{1/m})^m] \qquad (m = 1\text{-}2/n,\ n = 2(1\text{-}m)^{-1}) \tag{5.1.5}$$

Here, the functions in Equations (5.1.1), (5.1.3), (5.1.4) and (5.1.5) are denoted as Irmay, vG-M0, vG-M1 and vG-B0 models, respectively. In order to compare these relative permeability functions and test their suitability for compacted bentonite, they were included in INVGS-CORE and INVERSE-CORE^{2D} and their parameters were estimated from several infiltration experiments.

5.1.3. PARAMETER IDENTIFICATION

A 1-D numerical model was used to simulate the infiltration experiments and estimate unsaturated flow parameters. The real porosity may vary slightly with time in the experiments due to bentonite swelling. However, it was simplified to be a constant parameter in the inverse model. Five parameters were estimated, K_s (saturated conductivity), m, α, n_i, and ϕ (porosity) when the Irmay model was used. When the van Genuchten equation was used, the estimated parameters included K_s, m, α and ϕ.

Prior information, initial guess values, upper and lower bounds for the parameters were derived from Villar and Martin (1996), Villar et al. (1996), Villar and Pelayo (1996) and Dai and Samper (1999) and are listed in Table 5.1.2.

Table 5.1.2. Prior information, lower and upper bounds of the parameters (initial guess values of the parameters are equal to prior information. Values with * apply only to vG-M1 model)

Parameter	K_s (m/d)	m	n_i	α (m^{-1})	ϕ
Lower bound	$2.5 \cdot 10^{-9}$	0.2/1.2*	1.5	$3 \cdot 10^{-4}$	0.35
Upper bound	$8.5 \cdot 10^{-9}$	0.95/1.95*	4	$3 \cdot 10^{-3}$	0.45
Prior information	$5.1 \cdot 10^{-9}$	0.35/1.35*	3	$1.3 \cdot 10^{-3}$	0.41
Tolerance error	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}

Parameter estimation was performed in two steps. In the first one, INVGS-CORE (Golden section search method) was used to obtain preliminary estimates of model parameters. The results of INVGS-CORE were taken as initial parameter values for

INVERSE-CORE^{2D} which was used to refine parameter estimates. Final results obtained for SAT1, SAT2, SAT3, SAT4 and SAT5 tests are listed in Tables 5.1.3, 5.1.4, 5.1.5, 5.1.6 and 5.1.7, respectively. The fit between observation data and modeling results is shown in Figures 5.1.1 and 5.1.2. With these estimated parameters, relative permeability curves and retention curves have been plotted in Figures 5.1.3 and 5.1.4.

Table 5.1.3. Estimation results for test SAT1 (E_{inf} and E_{con} are the objective functions of water inflow and final water content, respectively)

Parameter	K_s (m/d)	m	n_i	α (m^{-1})	ϕ	E_{inf}	E_{con}
Irmay model	$4.41 \cdot 10^{-9}$	0.361	2.84	$1.393 \cdot 10^{-3}$	0.43	$4.38 \cdot 10^{-8}$	$1.95 \cdot 10^{-3}$
vG-M0 model	$4.47 \cdot 10^{-9}$	0.936	/	$1.328 \cdot 10^{-3}$	0.439	$8.16 \cdot 10^{-7}$	$1.21 \cdot 10^{-3}$
vG-M1 model	$4.42 \cdot 10^{-9}$	1.930	/	$1.23 \cdot 10^{-3}$	0.439	$8.16 \cdot 10^{-7}$	$1.21 \cdot 10^{-3}$
vG-B0 model	$5.48 \cdot 10^{-9}$	0.586	/	$1.365 \cdot 10^{-3}$	0.432	$7.89 \cdot 10^{-8}$	$1.75 \cdot 10^{-3}$

Table 5.1.4. Estimation results for test SAT2

Parameter	K_s (m/d)	m	n_i	α (m^{-1})	ϕ	E_{inf}	E_{con}
Irmay model	$5.3 \cdot 10^{-9}$	0.307	3.48	$1.32 \cdot 10^{-3}$	0.439	$6.57 \cdot 10^{-8}$	$6.47 \cdot 10^{-4}$
vG-M0 model	$5.99 \cdot 10^{-9}$	0.811	/	$1.005 \cdot 10^{-3}$	0.439	$7.08 \cdot 10^{-8}$	$9.44 \cdot 10^{-4}$
vG-M1 model	$6.62 \cdot 10^{-9}$	1.836	/	$1.003 \cdot 10^{-3}$	0.439	$9.04 \cdot 10^{-8}$	$9.44 \cdot 10^{-4}$
vG-B0 model	$5.99 \cdot 10^{-9}$	0.589	/	$1.007 \cdot 10^{-3}$	0.439	$8.14 \cdot 10^{-8}$	$9.44 \cdot 10^{-4}$

Table 5.1.5. Estimation results for test SAT3

Parameter	K_s (m/d)	m	n_i	α (m^{-1})	ϕ	E_{inf}	E_{con}
Irmay model	$5.07 \cdot 10^{-9}$	0.251	2.79	$1.194 \cdot 10^{-3}$	0.418	$2.85 \cdot 10^{-8}$	$1.3 \cdot 10^{-4}$
vG-M0 model	$7.2 \cdot 10^{-9}$	0.905	/	$1.005 \cdot 10^{-3}$	0.42	$1.07 \cdot 10^{-7}$	$6.83 \cdot 10^{-4}$
vG-M1 model	$7.19 \cdot 10^{-9}$	1.921	/	$1.005 \cdot 10^{-3}$	0.42	$1.69 \cdot 10^{-7}$	$6.83 \cdot 10^{-4}$
vG-B0 model	$7.2 \cdot 10^{-9}$	0.805	/	$1.005 \cdot 10^{-3}$	0.419	$1.1 \cdot 10^{-7}$	$6.82 \cdot 10^{-4}$

Table 5.1.6. Estimation results for test SAT4

Parameter	K_s (m/d)	m	n_i	α (m^{-1})	ϕ	E_{inf}	E_{con}
Irmay model	$4.89 \cdot 10^{-9}$	0.217	2.54	$1.446 \cdot 10^{-3}$	0.416	$8.66 \cdot 10^{-8}$	$3.33 \cdot 10^{-4}$
vG-M0 model	$7.2 \cdot 10^{-9}$	0.908	/	$1.008 \cdot 10^{-3}$	0.415	$2.05 \cdot 10^{-7}$	$7.3 \cdot 10^{-4}$
vG-M1 model	$7.2 \cdot 10^{-9}$	1.924	/	$1.005 \cdot 10^{-3}$	0.419	$2.35 \cdot 10^{-7}$	$9.49 \cdot 10^{-4}$
vG-B0 model	$7.2 \cdot 10^{-9}$	0.814	/	$1.005 \cdot 10^{-3}$	0.416	$2.05 \cdot 10^{-7}$	$7.52 \cdot 10^{-4}$

Table 5.1.7. Estimation results for test SAT5

Parameter	K_s (m/d)	m	n_i	α (m^{-1})	ϕ	E_{inf}	E_{con}
Irmay model	$5.59 \cdot 10^{-9}$	0.209	3(fixed)	$1.896 \cdot 10^{-3}$	0.404	$3.72 \cdot 10^{-6}$	$1.01 \cdot 10^{-3}$
Irmay model	$5.23 \cdot 10^{-9}$	0.195	3.49	$1.839 \cdot 10^{-3}$	0.404	$3.68 \cdot 10^{-6}$	$1.04 \cdot 10^{-3}$
vG-M0 model	$8.2 \cdot 10^{-9}$	0.307	/	$3.022 \cdot 10^{-4}$	0.398	$2.76 \cdot 10^{-5}$	$1.88 \cdot 10^{-3}$
vG-M1 model	$8.43 \cdot 10^{-9}$	1.456	/	$3.005 \cdot 10^{-4}$	0.408	$3 \cdot 10^{-5}$	$6.83 \cdot 10^{-4}$
vG-B0 model	$8.47 \cdot 10^{-9}$	0.175	/	$6.624 \cdot 10^{-4}$	0.398	$2.99 \cdot 10^{-5}$	$2.09 \cdot 10^{-3}$

From Tables 5.1.3 to 5.1.7 and Figures 5.1.1 and 5.1.2, one can see that Irmay function gives always less estimation variance and a fit to measured data better than those of the van Genuchten functions. The fit of final water content to measured data near the water inflow boundary (at the right of Figure 5.1.2) is not good. The possible reason is that porosity may vary a little with time due to clay swelling, but these parameters are assumed to be constant in our models.

Figure 5.1.3 shows the relative permeability functions computed for all 5 tests. There are significant differences between different models, especially in the case of test SAT5. The saturated hydraulic conductivities obtained with Irmay model are lower than those derived from van Genuchten models. Available measurements of saturated hydraulic conductivity for the FEBEX bentonite (dry density of 1.6 g/cm^3) (Villar and Cuevas, 1996) vary from $2.7 \cdot 10^{-9}$ m/d to $7.2 \cdot 10^{-9}$ m/d, with an arithmetic average of $5.1 \cdot 10^{-9}$ m/d. So, the estimated saturated hydraulic conductivities obtained with Irmay model are closer to measured values. Those obtained with van Genuchten models are usually larger than the average measured value. As shown in Figure 5.1.4, the retention curves obtained with Irmay model are also more coherent with measured suction-moisture data than those of van Genuchten models.

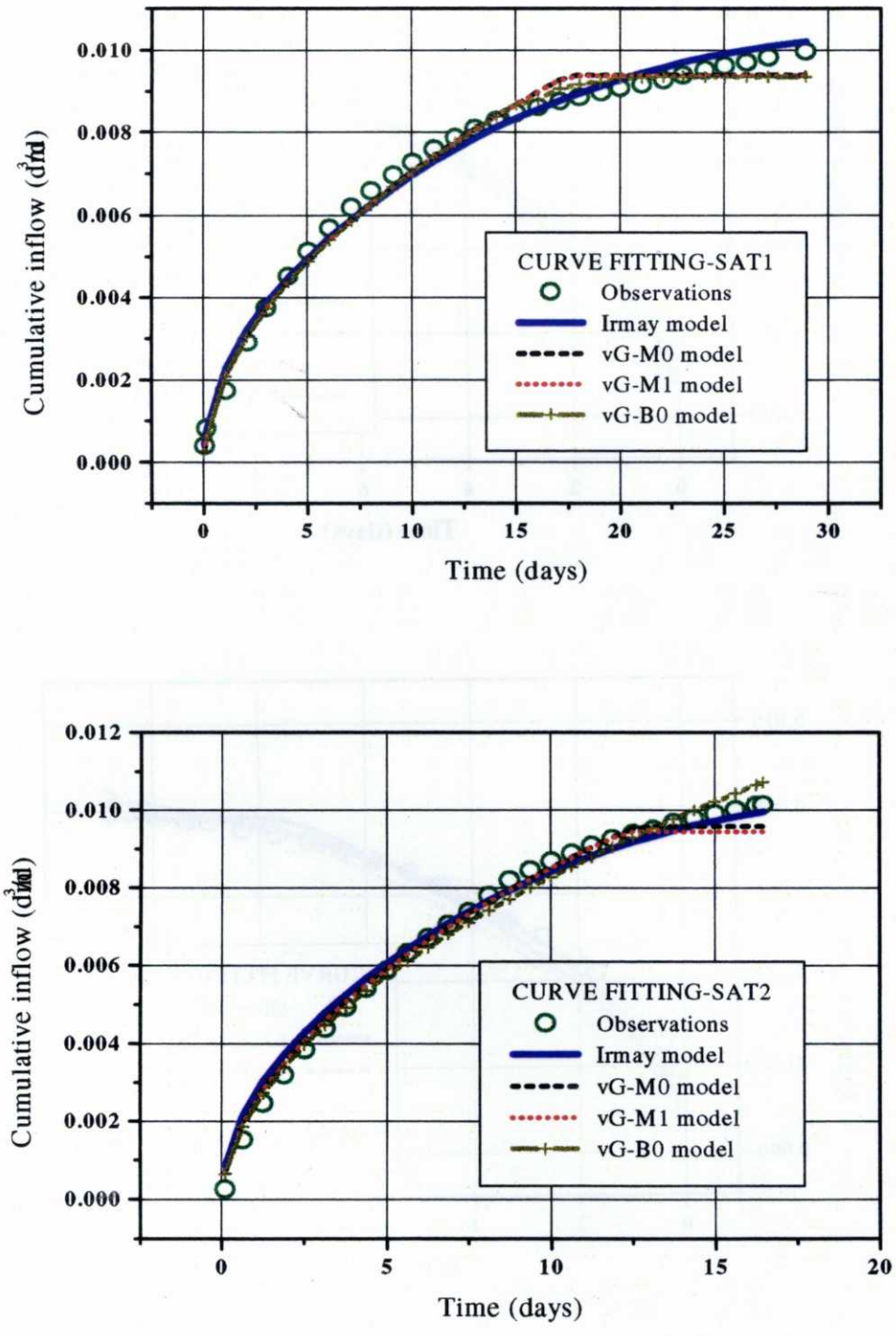


Figure 5.1.1A. Measured cumulative water inflow and best fit obtained with different relative permeability functions in tests SAT1 and SAT2

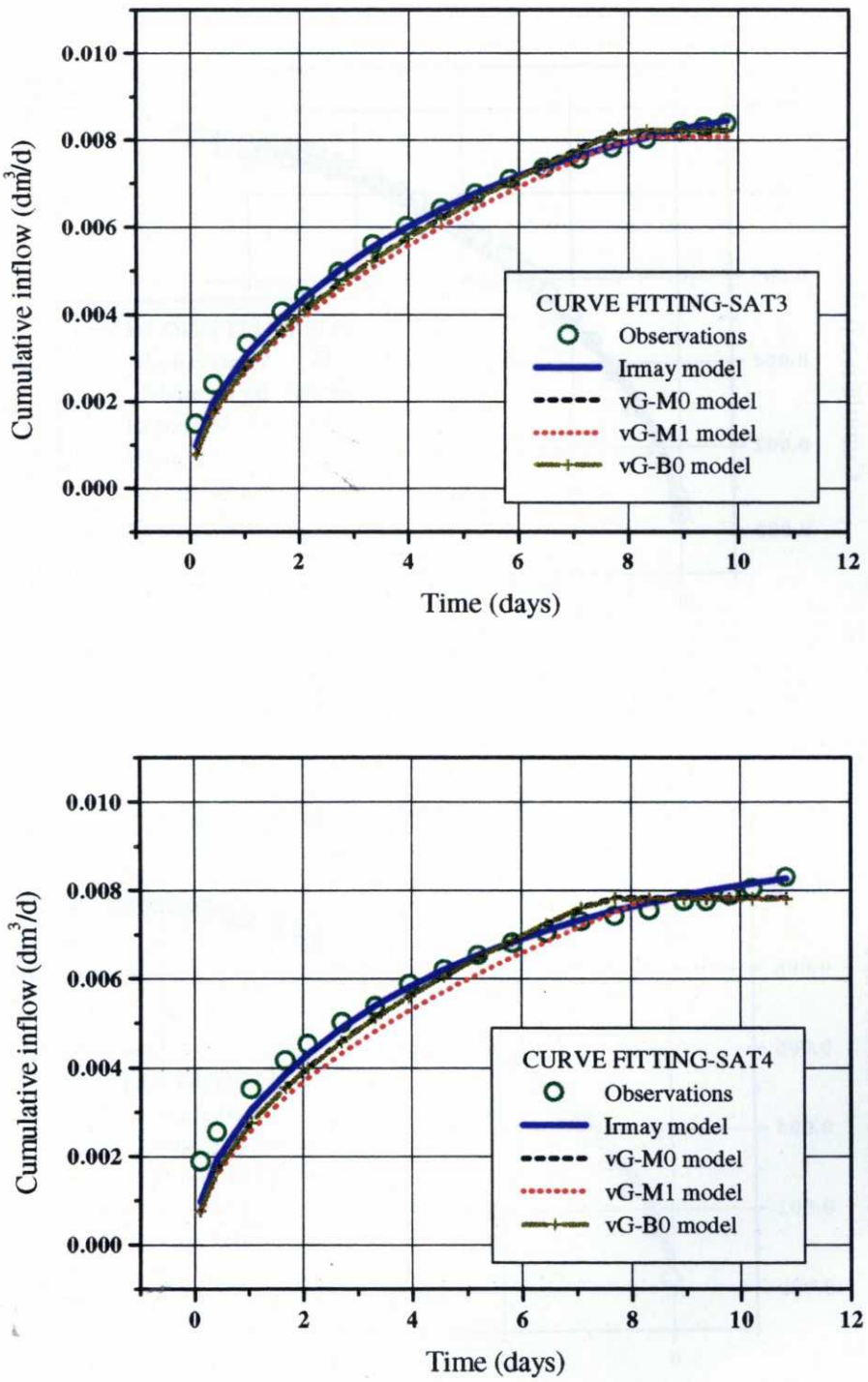


Figure 5.1.1B. Measured cumulative water inflow and best fit obtained with different relative permeability functions in tests SAT3 and SAT4

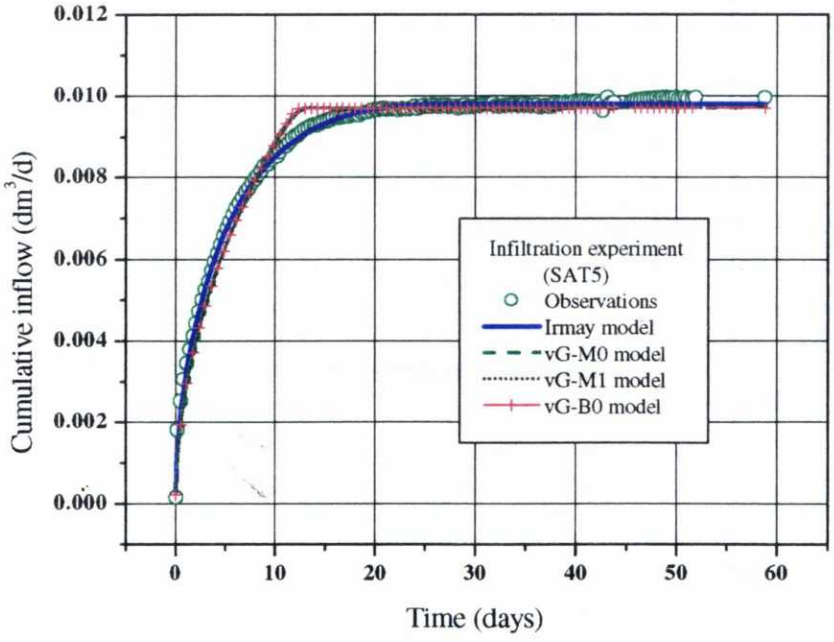


Figure 5.1.1C. Measured cumulative water inflow and best fit obtained with different relative permeability functions in test SAT5

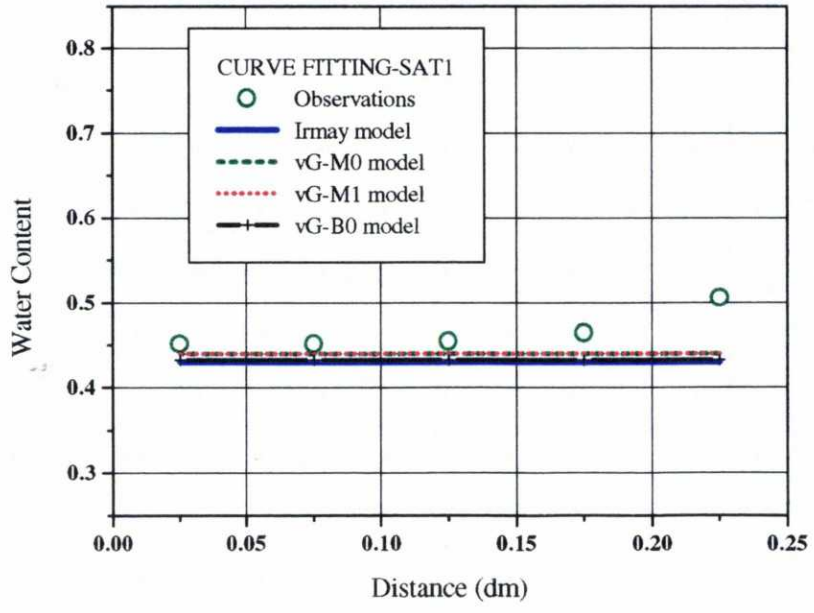
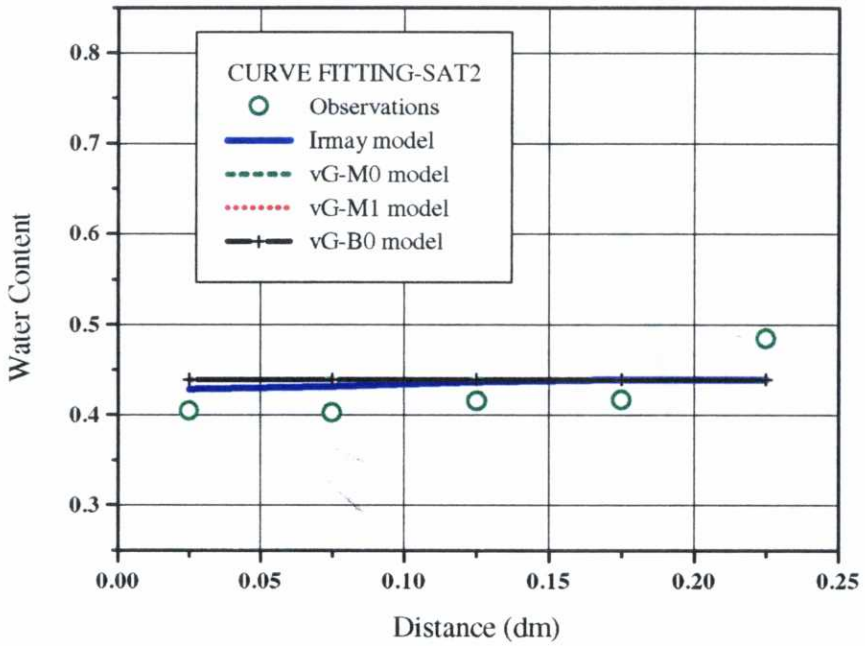


Figure 5.1.2A. Measured final water content distribution and best curve fit obtained with different relative permeability functions in tests SAT1



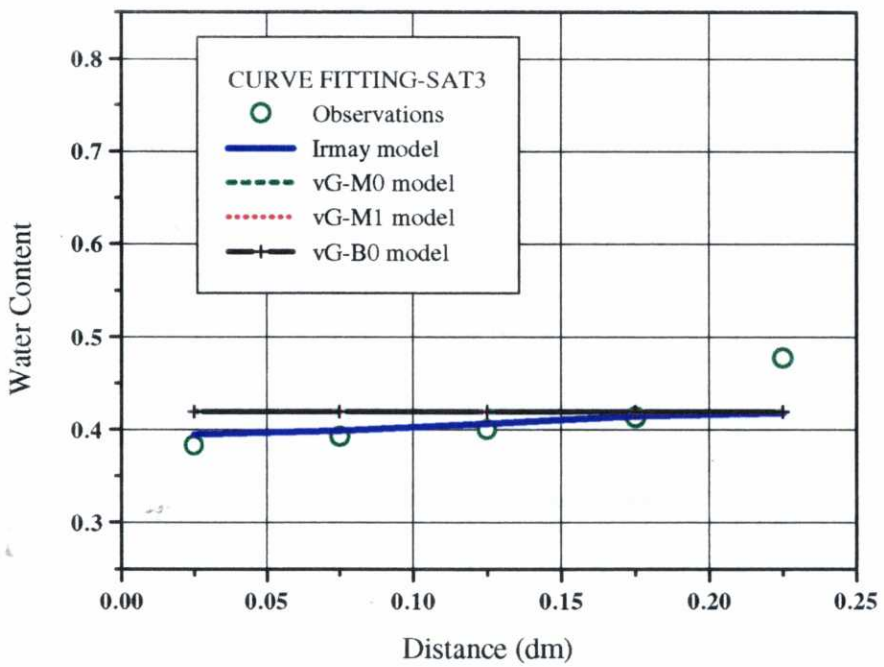


Figure 5.1.2B. Measured final water content distribution and best curve fit obtained with different relative permeability functions in tests SAT2 and SAT3

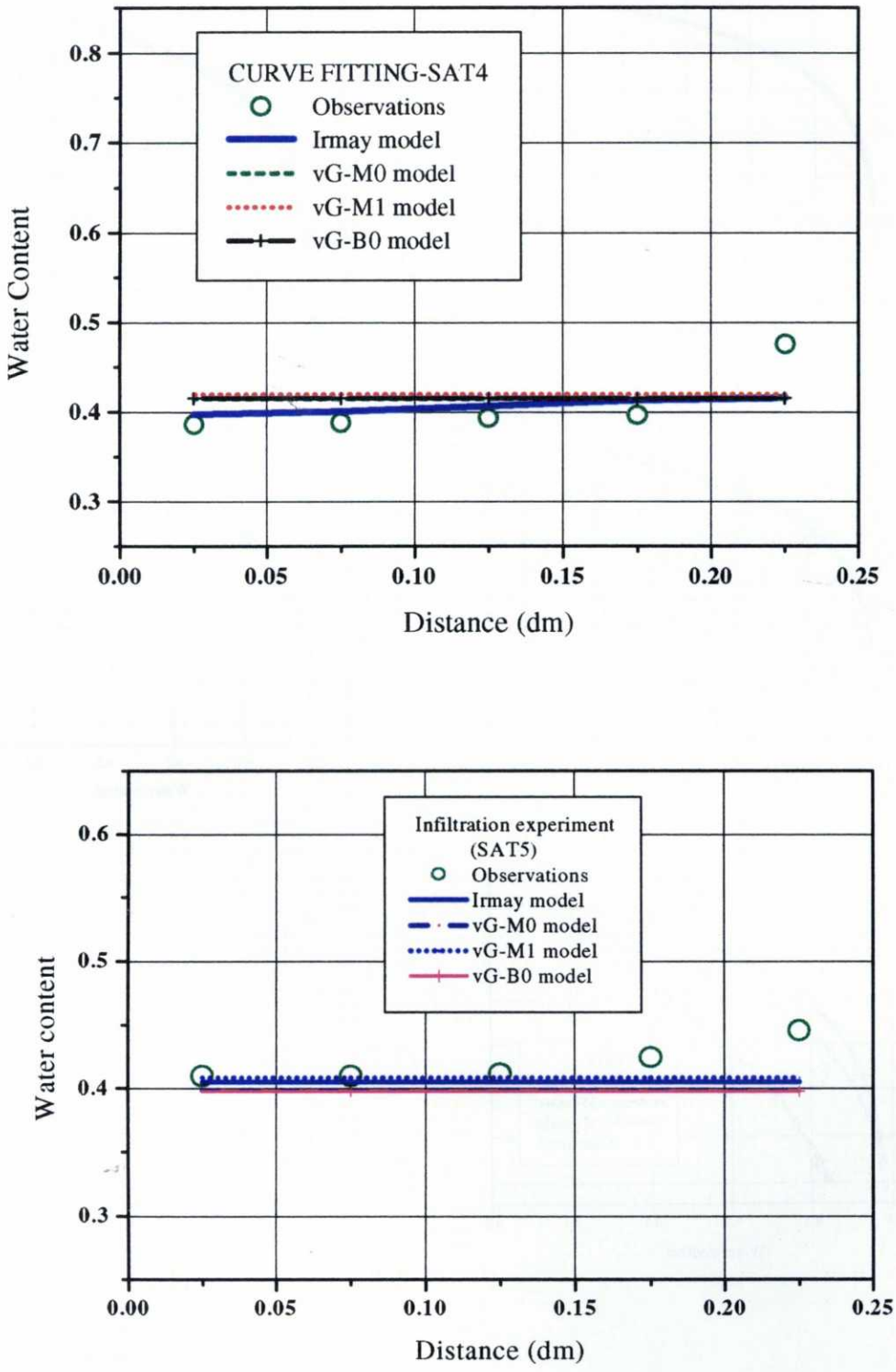


Figure 5.1.2C. Measured final water content distribution and best curve fit obtained with different relative permeability functions in tests SAT4 and SAT5

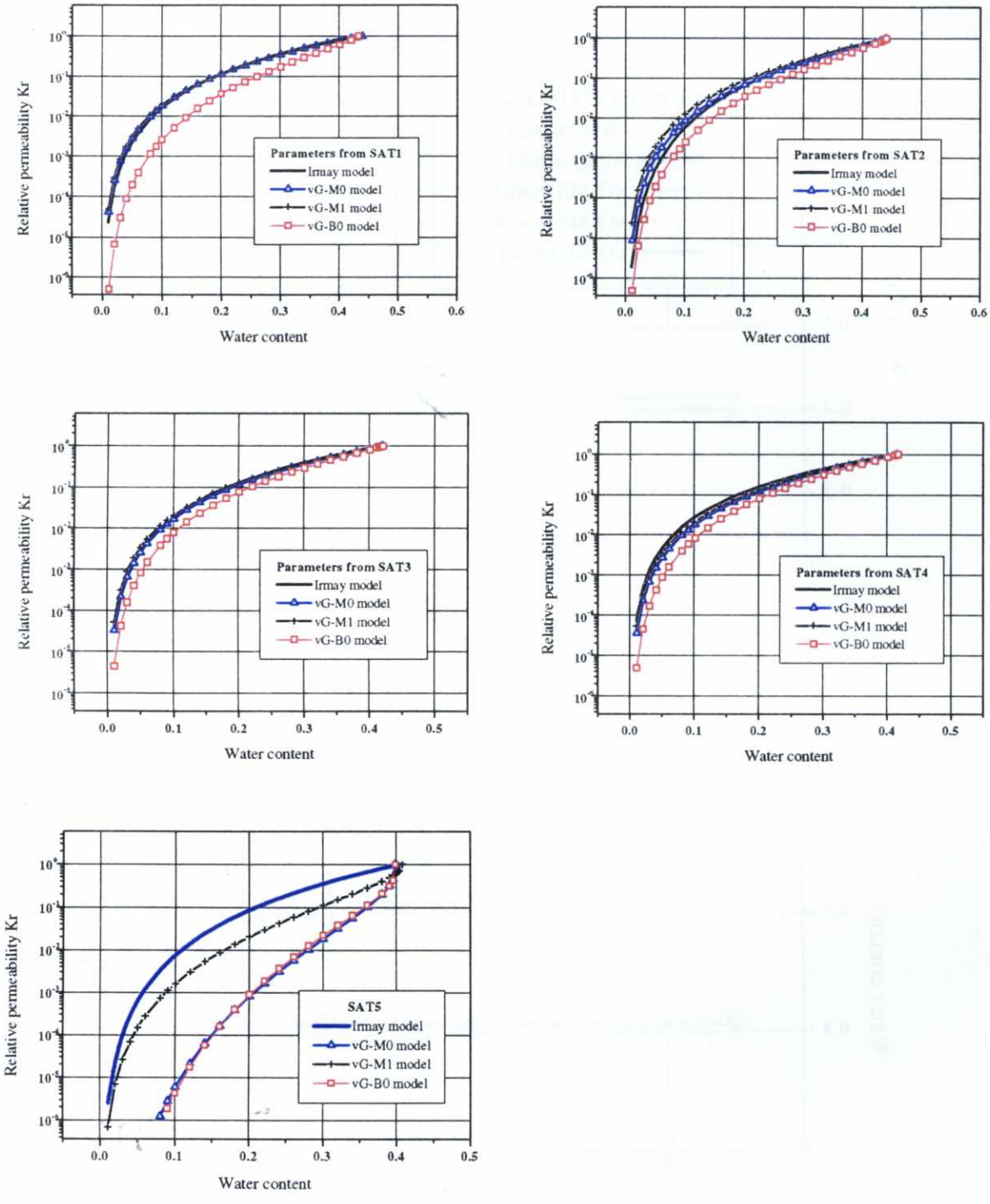


Figure 5.1.3. Estimated relative permeability functions for tests SAT1, SAT2, SAT3, SAT4 and SAT5

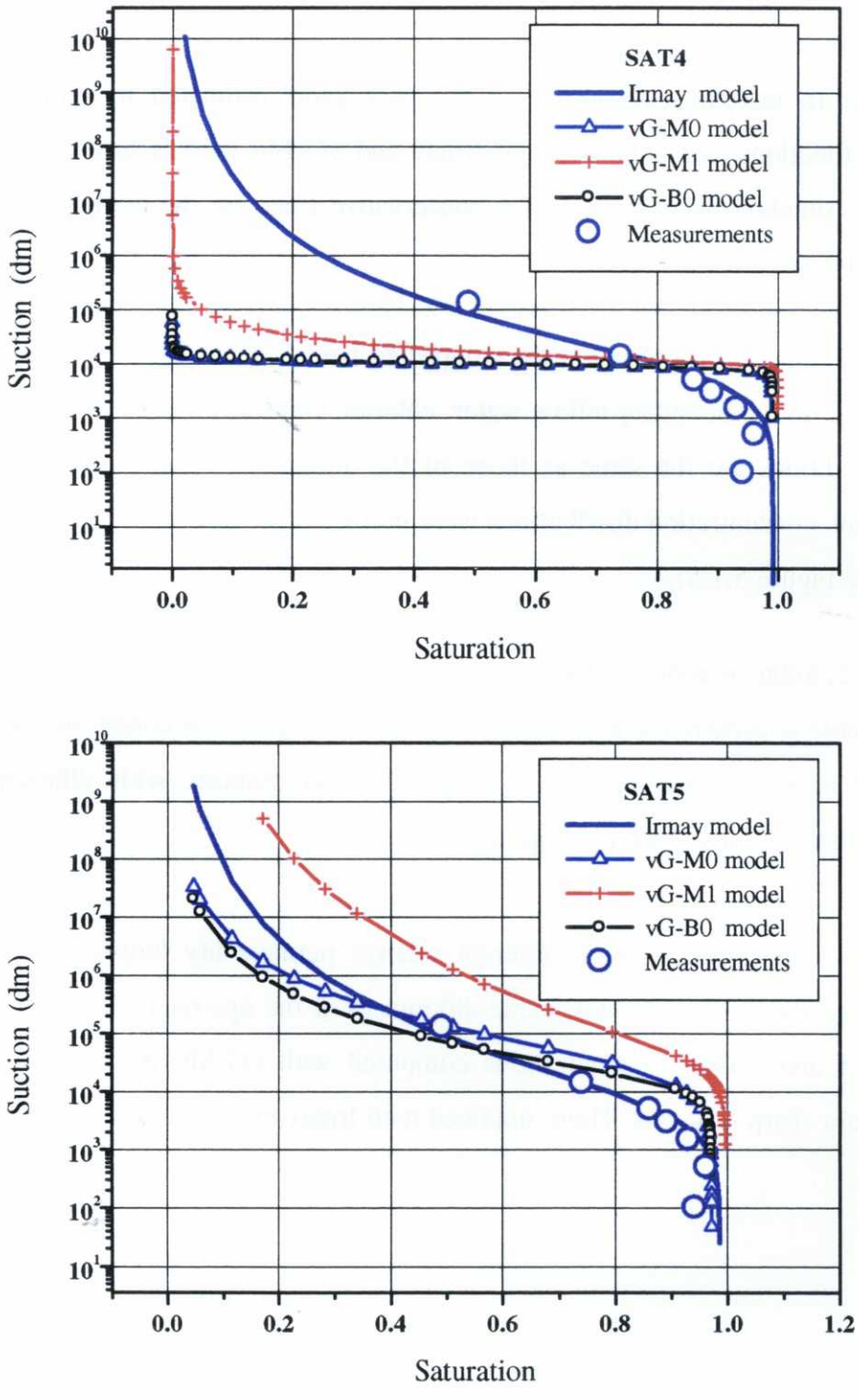


Figure 5.1.4. Measured (circles) and computed retention curves for tests SAT4 and SAT5

5.1.4. IMPLICATIONS FOR SOLUTE TRANSPORT

In order to test the equivalence of the parameters estimated by different relative permeability functions, the parameters of Irmay and vG-M0 models listed in Table 5.1.7 were used to simulate the transport of a conservative tracer in the compacted bentonite. Two different cases are considered.

In case 1, chloride is present in the compacted bentonite initially with a concentration of $3.6 \cdot 10^{3}$ mg/l. In the boundary inflow water, chloride concentration is 10^{-4} mg/l. The flow boundary conditions are the same as those of the infiltration experiments. By using the forward model, concentration distributions were predicted for different relative permeability functions (see Figure 5.1.5).

In case 2, iodine is added to the boundary inflow water with a concentration of 2 mg/l. The initial solute concentration in the bentonite is $2 \cdot 10^{-2}$ mg/l. Flow conditions are the same as those of infiltration experiments. Computed concentrations with different relative permeability functions are shown in Figure 5.1.6.

Computed concentrations with different relative permeability functions show similar trends although there are some noticeable differences at the upstream side of tracer fronts (Figures 5.1.5 and 5.1.6). Concentrations computed with vG-M0 model always show a surprisingly too sharp behavior. Those obtained with Irmay model show smoother patterns.

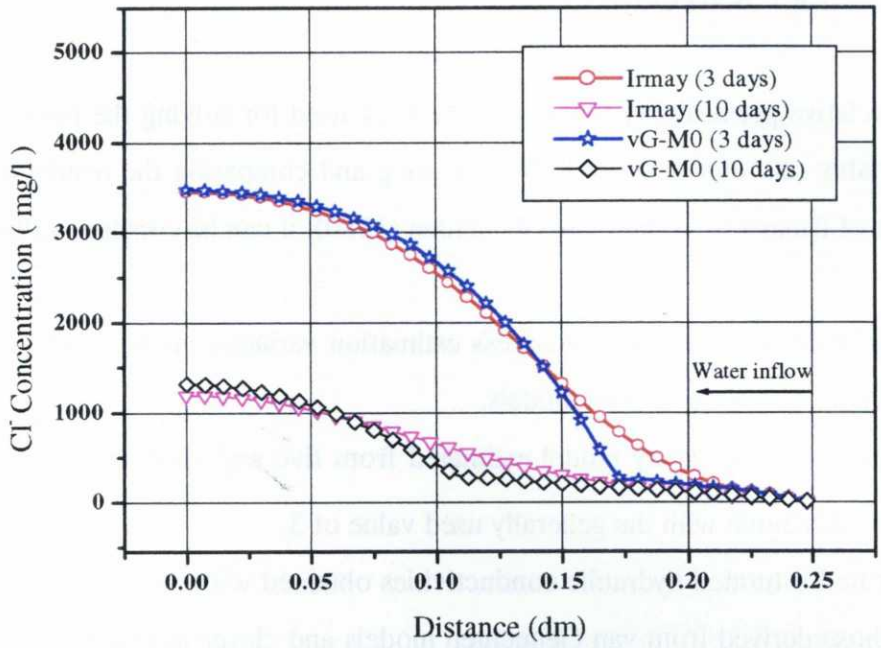


Figure 5.1.5. Plot of Cl^- concentrations computed with different relative permeability functions in Test SAT5

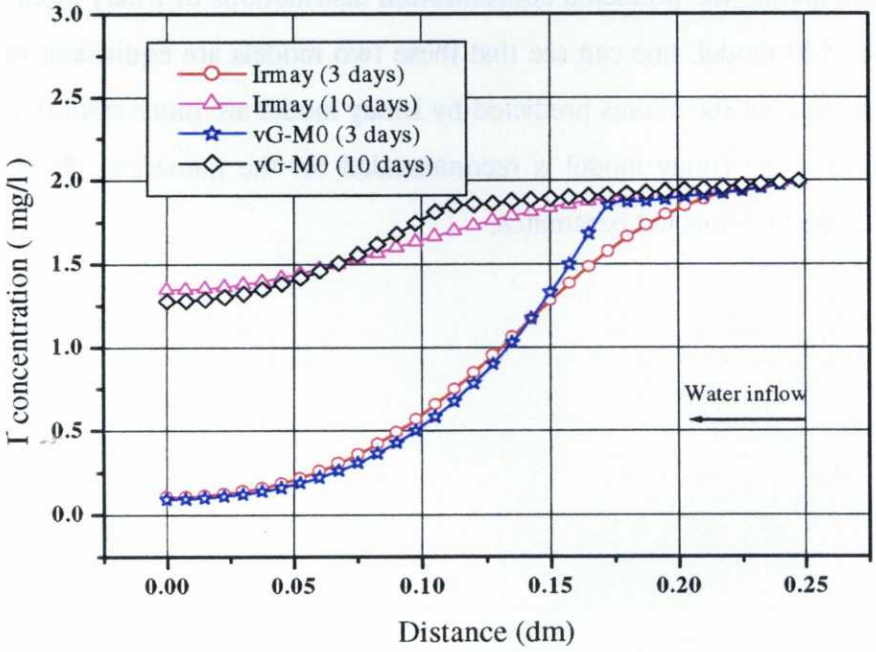


Figure 5.1.6. Plot of I^- concentrations computed with different relative permeability functions in Test SAT5

5.1.5. SUMMARY AND CONCLUSION

Four relative permeability functions have been used for solving the inverse problem of flow in variably saturated bentonite. By analyzing and comparing the results obtained with the models of Irmay (1954) and van Genuchten (1980), it can be concluded that

(1) Irmay model gives always much less estimation variances and provides better results than those of van Genuchten models.

(2) Parameters n_i of Irmay model estimated from five experiments range from 2.54 to 3.49, which agree with the generally used value of 3.

(3) Estimated saturated hydraulic conductivities obtained with the Irmay model are lower than those derived from van Genuchten models and closer to independently measured values.

(4) Retention curves obtained with Irmay model are more coherent to measured suction-moisture data than those of van Genuchten models.

(5) By comparing the predicted concentration distributions of Irmay model with those of the vG-M0 model, one can see that these two models are equivalent in most parts of the curves, but the curves predicted by Irmay model are more smooth and reasonable. Therefore, the Irmay model is recommended for the numerical modeling of flow in unsaturated compacted bentonites.

5.2. INTERPRETATION OF DIFFUSION EXPERIMENT IN COMPACTED FEBEX BENTONITE

The measurement of sorption and diffusion coefficients in porous media can be carried out in laboratory conditions using different experimental techniques. Compacted bentonites have a LARGE sorption capacity and an extremely low permeability. Except for the early stages of bentonite hydration (when advective solute transport dominates), diffusion is the main mechanism of radionuclide transport through bentonites. Diffusive transport of radionuclides in clay porewater depends on the charge, ionic size, dipolar momentum and polarizability of the diffusing species.

Figure 5.2.1 illustrates the methods which have been used by CIEMAT within the FEBEX Project (Garcia et al., 1998). There are three types of diffusion experiments made by CIEMAT: In-diffusion, through-diffusion and permeation experiments. Diffusion and permeation experiments are needed to obtain sorption and transport parameters which include: (1) distribution coefficient, (2) diffusion coefficient, and (3) transport porosity.

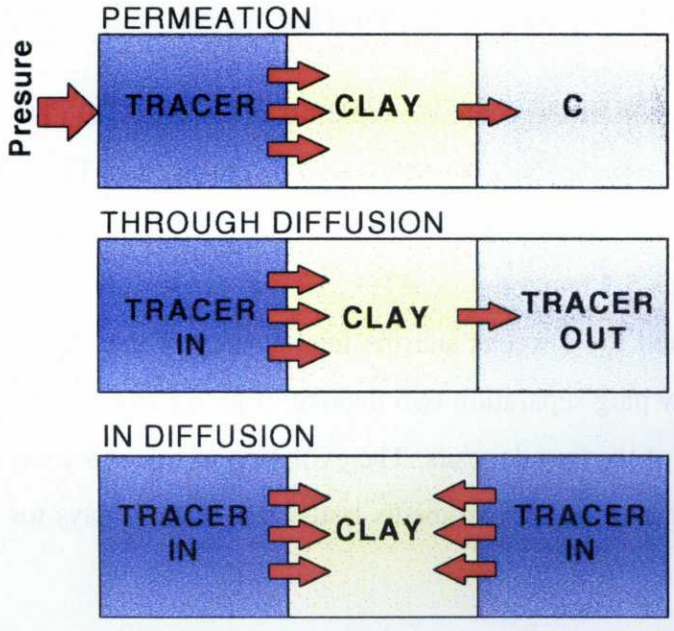


Figure 5.2.1. Methods used for determining sorption and diffusion parameters (Garcia et al., 1998)

Since the interpretation of diffusion and permeation experiments in compacted bentonite requires estimation of two or three parameters, codes INVGS-CORE and INVQI-CORE was used to estimate transport parameters. The process of parameter identification is complex. The optimal solution cannot be found in a single run of the inverse program. The trial and error method is used to define initial values, and upper and lower bounds of the parameters. After analyzing the fit and the values of the objective function, initial values and bounds are updated. At the beginning, only 2 or 3 cycles of iterations are used. They are increased until a convergent solution is obtained. The maximum number of cycles is 10 and the convergence criterion of the objective function (relative error between two consecutive iterations, *ftol*) is 10^{-5}. The maximum number of iterations for each parameter is 15 for all the experiments. The interpretation of these experiments was carried out with early versions of the inverse code (INVGS-CORE and INVQI-CORE) which were available at the time that such interpretation was needed for the purpose of providing transport parameters for the FEBEX bentonite.

5.2.1. THROUGH-DIFFUSION EXPERIMENTS

5.2.1.1. Experiment description and the numerical models

Through-diffusion experiments were performed using HTO (TD-1, TD-2 and TD-3) and strontium (Sr^{2+}) experiments (TD-Sr5, TD-Sr6, TD-Sr7 and TD-Sr8).

Bentonite plugs 5.3 mm long for HTO (8.3 mm for strontium) and with a diameter of 50 mm were saturated for 4 weeks and the final density of the plugs was 1.18 g/cm^3. The cells contain the clay plug separation two deposit of groundwater. The tracer was added to the solutions in one of the two deposits. The evolution of tracer activity is measured in both reservoirs at regular intervals. Experiments lasted less than 54 days for HTO and near 200 days for Sr^{2+}.

The numerical models for through-diffusion experiments consist of 5 parts, the first reservoir, first sinter, clay plug, second sinter and second reservoir. The porosity of the reservoirs is 1 and 0.5 for the sinters in all the experiments. The diffusion coefficient in the

reservoirs is large enough to ensure that the nodes within the reservoir have the same activity. The initial activity in the left reservoir is equal to the initial tracer activity. In the sinters, clay and the right reservoir, they are equal to zero.

Two grids were used for the numerical simulation of through diffusion experiments. That of HTO (TD-1, TD-2 and TD-3) includes 200 triangular elements and 202 nodes. The grid for strontium experiments (TD-Sr5, TD-Sr6, TD-Sr7 and TD-Sr8) has 114 triangular elements and 116 nodes.

The simulation time horizon covers the expected duration of the experiments. For TD-1,TD-2 and TD-3, the total time is 54 days, which was discretized with 14 time periods and 800 time increments. For TD-Sr5, TD-Sr6, TD-Sr7 and TD-Sr8, the total experiment time 193 days is discretized in 12 time periods and 1800 time increments. Each time period has different number of time increments and from the first time period to the last, they increase gradually.

5.2.1.2. Interpretation of HTO experiments

5.2.1.2.1. Estimated results

Golden section search and quadratic interpolation method were used for the interpretation. Both lead to the same results. Porosity and diffusion coefficient (D_0) of the sinters are prescribed to be 0.5 and $2.454 \cdot 10^{-9}$ m^2/s. Parameters' upper and lower bounds and initial values for HTO were derived from published data (Samper et al, 1998, García et al., 1998, Yuan-Hui Li et al, 1974, and Yu and Neretnieks, 1996) and are listed in Table 5.2.1.

Table 5.2.1. The prior information of the parameters of the through-diffusion experiments

Parameter	Diffusion coefficient (D_0, m^2/s)	Porosity (ϕ)
Lower bound	$5.5 \cdot 10^{-11}$	0.2
Upper bound	$5.5 \cdot 10^{-8}$	0.65
Initial guess	$2.1 \cdot 10^{-9}$	0.56
Tolerance error	$2 \cdot 10^{-14}$	10^{-4}

García et al. (1998) employed an approximate analytical equation (Wolfrum et al., 1988) to estimate diffusion coefficients from experiments TD-1, TD-2, and TD-3. Estimated values from analytical and numerical methods are presented in Table 5.2.2. In this table, clay porosity is fixed to 0.56, and only the diffusion coefficient is estimated. The tortuosity factor is the same for all samples and equal to 0.8243.

Table 5.2.2. Estimated diffusion coefficient by different methods (porosity is equal to 0.56; analytical solution is provided by García et al. (1998); t_* is a characteristic time defined in Equation (5.2.1))

TD Test	Analytical method	Numerical method			
	D_e (m^2/s)	D_0 (m^2/s)	D_e (m^2/s)	t_* (day)	Objective
TD-1	$1.03 \cdot 10^{-10}$	$3.061 \cdot 10^{-10}$	$1.413 \cdot 10^{-10}$	23.01	0.0932
TD-2	$8.89 \cdot 10^{-11}$	$2.684 \cdot 10^{-10}$	$1.239 \cdot 10^{-10}$	26.24	0.0922
TD-3	$9.80 \cdot 10^{-11}$	$2.550 \cdot 10^{-10}$	$1.177 \cdot 10^{-10}$	27.62	0.0935

t_* is the time needed for the tracer to diffuse through the clay plug. It can be computed approximately from:

$$t_* = 10 \cdot L^2 / D_e \qquad (5.2.1)$$

where, L is sample length and D_e is the effective diffusion coefficient.

The thickness of the sinters is 3.35 mm. They cannot be ignored comparing to the thickness of clay plug (5.3 mm). The sinters used in the experiments have an effective

diffusion coefficient of $1.227 \cdot 10^{-9} m^2/s$. The diffusion time ($t_*$) for a sinter, calculated with Equation 5.2.1, is 1.059 days. It means that the sinters play an important role for the interpretation of through-diffusion experiments.

The results obtained with analytical and numerical methods show some differences. In general, the analytical method underestimates the diffusion coefficient because it ignores the role of the sinters. The tracer must diffuse through the sinters before it does through the bentonite. Therefore, the sinters introduce a retardation in the diffusion process and this is why the analytical method always provides diffusion coefficients smaller than the numerical method which accounts explicitly for both sinters as well as for the time variation in the activity of the tracer in the reservoirs. Since the numerical method reproduces the experimental conditions better than the analytical method, it can be concluded that parameters estimated with the numerical method are more reliable.

Since transport porosity may differ from total porosity, in the next stage, both of porosity and diffusion coefficient of each sample were estimated. The results are listed in Table 5.2.3. Estimated porosities are smaller than total porosity, although the estimated effective molecular diffusion coefficients are almost identical to those estimated previously with fixed porosity. As expected, when both porosity (ϕ) and diffusion coefficient (D_0) are estimated, the fit improves (Figures 5.2.2, 5.2.3 and 5.2.4) and the objective function is reduced slightly (comparing Tables 5.2.2 and 5.2.3). However, the improvement is not spectacular. The fact that similar fits can be obtained with different values of ϕ and D_0 but similar values of D_e indicates that in this case the experimental data allow only the estimation of effective diffusion ($D_e = \phi D_0 \tau$), but not the separate estimation of both parameters.

A summary of published effective diffusion coefficients of tritiated water in clay was presented by Yu and Neretnieks (1996). D_e values range from $3.6 \cdot 10^{-12}$ to $1.5 \cdot 10^{-10}$ m^2/s. The values in Tables 5.2.2 and 5.2.3 are consistent with this range.

Table 5.2.3. Estimated diffusion coefficient and porosity by numerical methods

TD Test	Porosity	Tortuosity	D_0 (m^2/s)	De (m^2/s)	t_* (day)	Objective
TD-1	0.4986	0.7929	$3.550 \cdot 10^{-10}$	$1.409 \cdot 10^{-10}$	23.07	0.0896
TD-2	0.4515	0.7671	$3.566 \cdot 10^{-10}$	$1.235 \cdot 10^{-10}$	26.33	0.07817
TD-3	0.4397	0.7604	$3.508 \cdot 10^{-10}$	$1.173 \cdot 10^{-10}$	27.72	0.08737

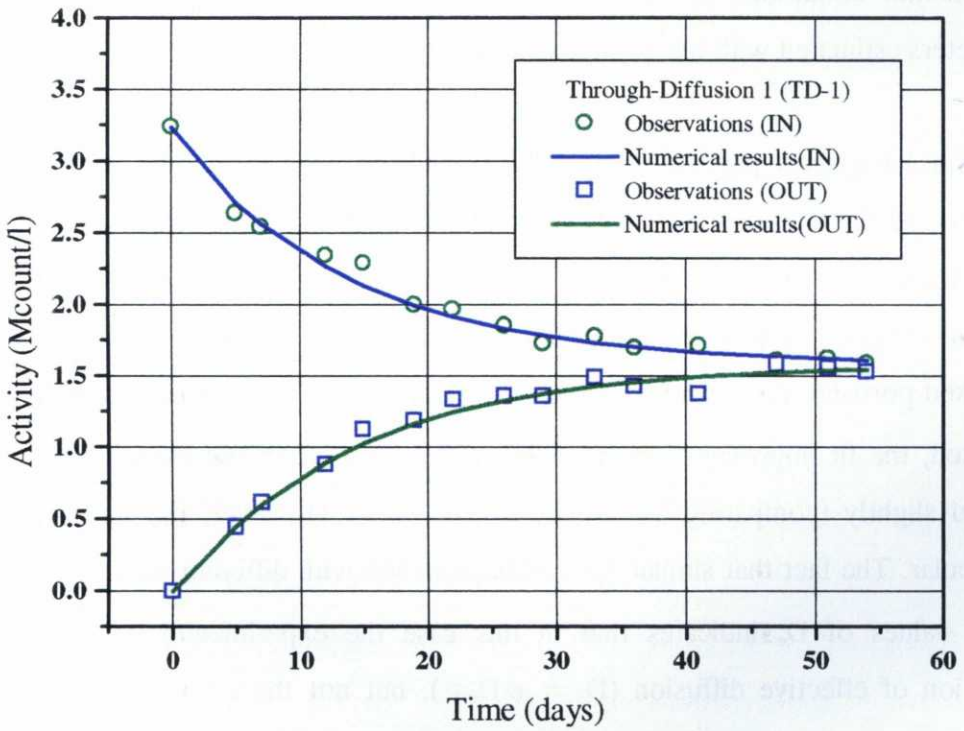


Figure 5.2.2. Activity evolution and best fit of IN and OUT reservoirs of through-diffusion experiment TD-1

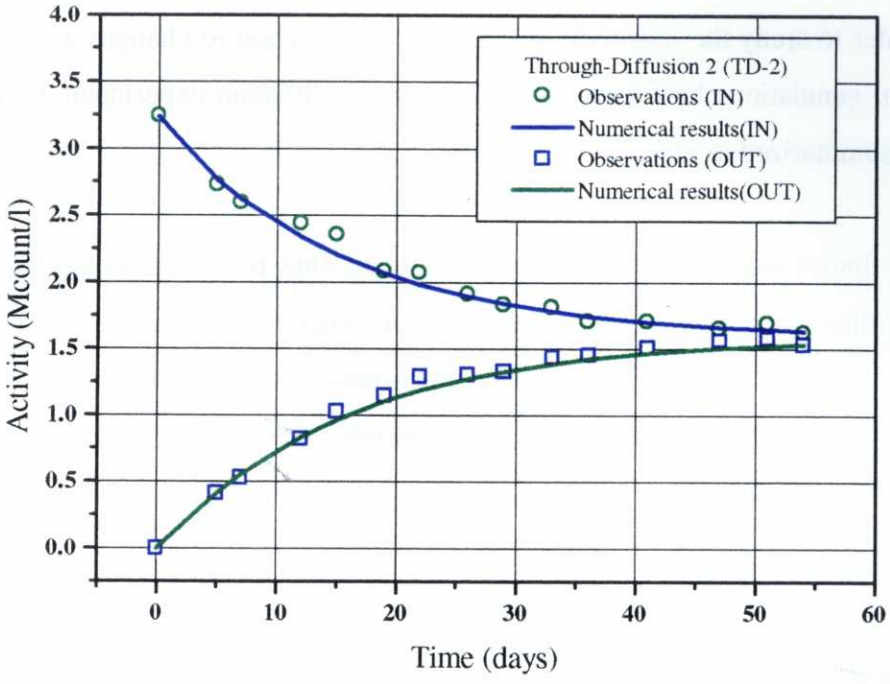


Figure 5.2.3. Activity evolution and best fit of IN and OUT reservoirs of through-diffusion experiment TD-2

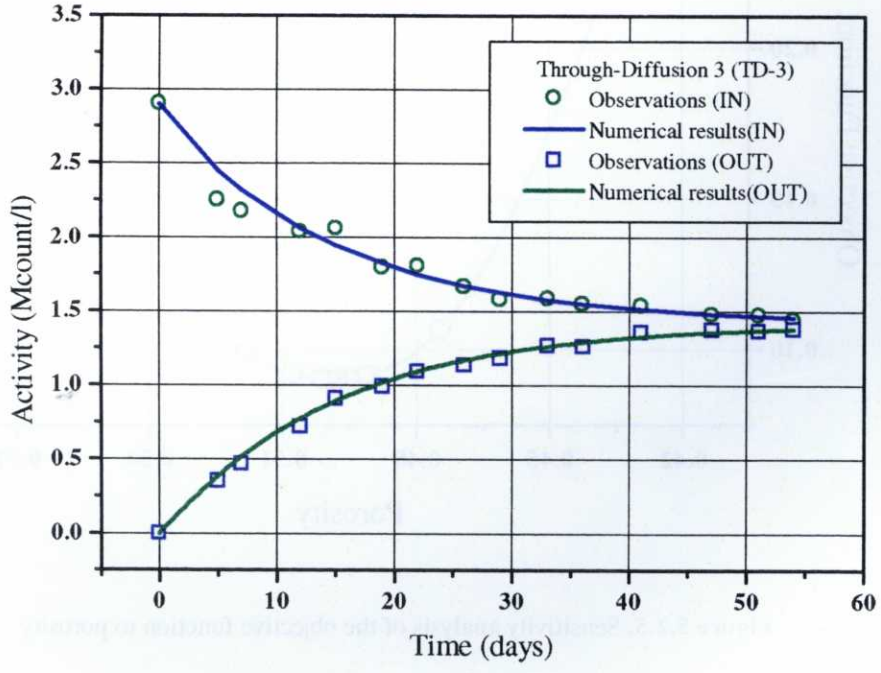


Figure 5.2.4. Activity evolution and best fit of IN and OUT reservoirs of through-diffusion experiment TD-3

5.2.1.2.2. Sensitivity analysis

In order to study the sensitivity of the objective function to changes in the parameters, a number of simulations have been done for through-diffusion experiment 1 (TD-1). The results are summarized in Figures 5.2.5 and 5.2.6.

These figures show that computed activities and objective function are highly sensitive to one-at-a-time changes in porosity and diffusion. Such high sensitivity ensures a reliable estimation of these parameters when they are estimated separately. However, this result does not ensure that they can be estimated accurately when both are estimated at the same time.

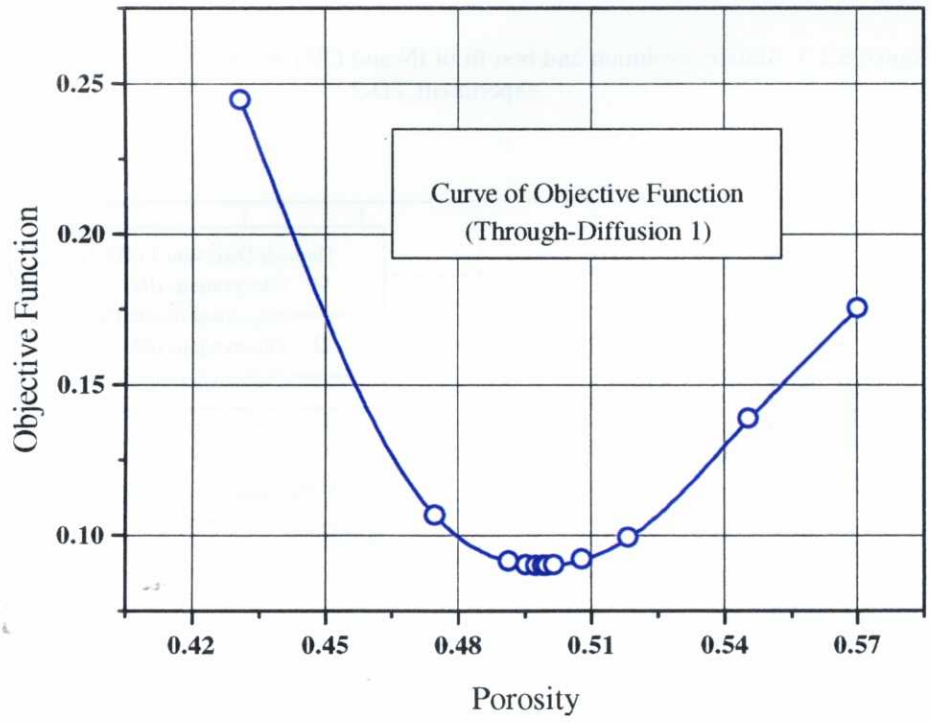


Figure 5.2.5. Sensitivity analysis of the objective function to porosity

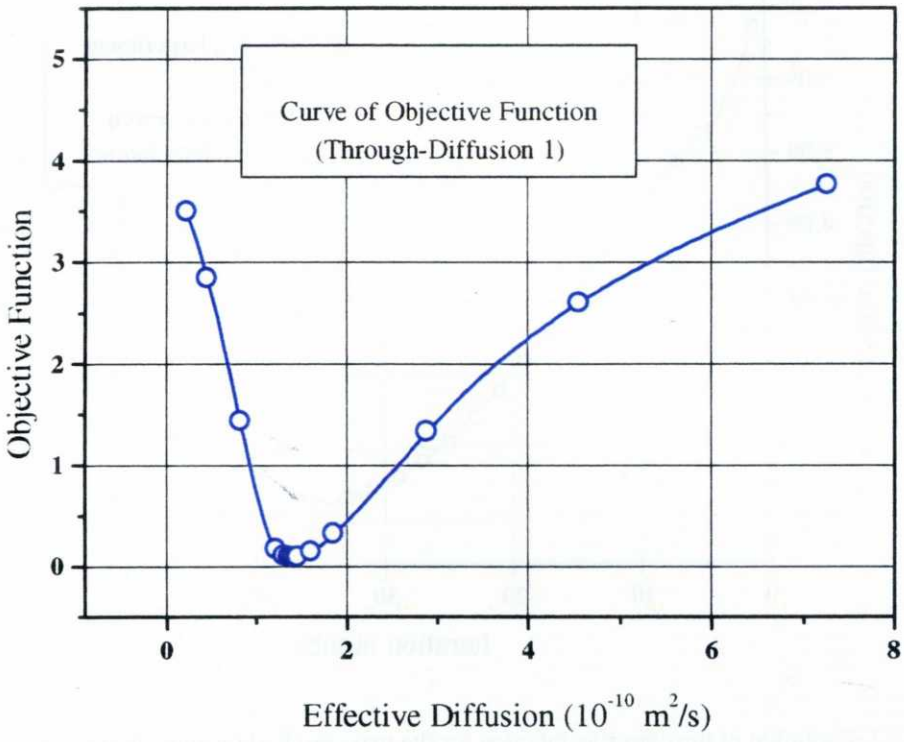


Figure 5.2.6. Sensitivity analysis of the objective function to effective diffusion coefficient

5.2.1.2.3. Comparison of estimation algorithms

Golden section search (INVGS-CORE) and quadratic interpolation methods (INVQI-CORE)) were used for solving the inverse problem of HTO through-diffusion. Both give the same results. The Golden section method requires 22 iterations, while the quadratic interpolation needs 60 iterations (Figures 5.2.7 and 5.2.8). Even though the quadratic interpolation algorithm was more efficient than the Golden Section search algorithm for one parameter identification with noise-free synthetic data, in this real case with data containing measurement noise, the Golden section search method is more efficient. For this reason, the Golden section search was used for the interpretation of the rest of the experiments.

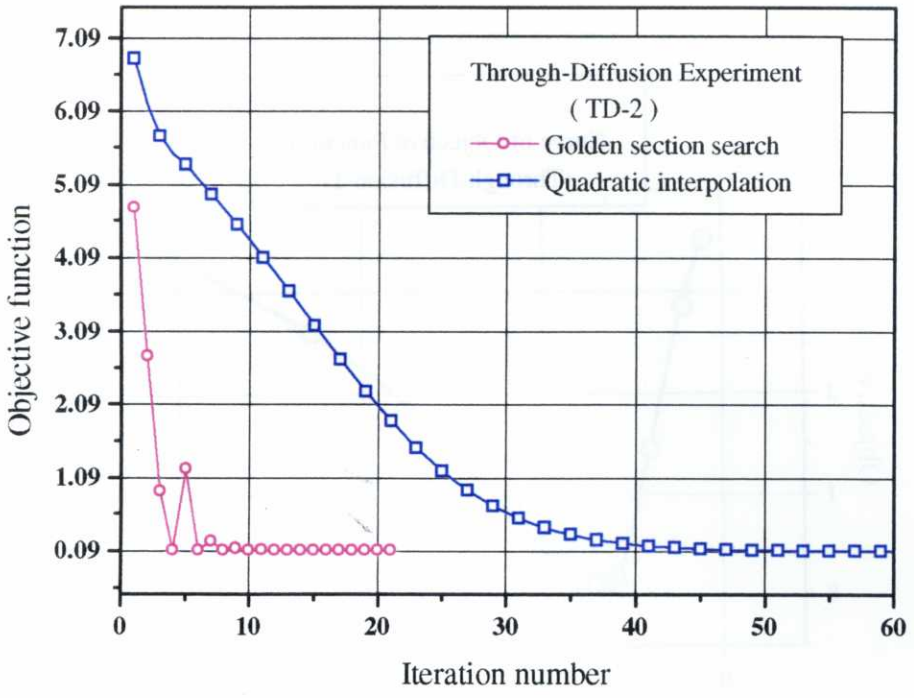


Figure 5.2.7. Evolution of the objective function for the two search algorithms in the interpretation of through-diffusion experiment TD-2.

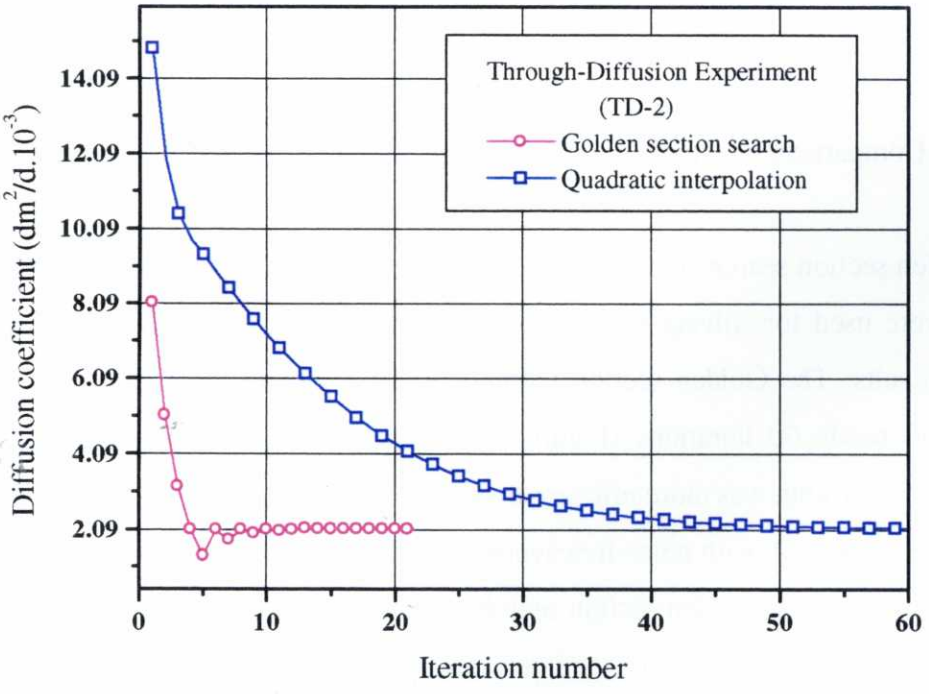


Figure 5.2.8. Evolution of parameter values for the two search algorithms in the interpretation of through-diffusion experiment TD-2.

5.2.1.3. Interpretation of strontium experiments

Strontium has multiple chemical forms, but, in the bentonite porewater it prevails as Sr^{2+} (Brandberge and Skagius, 1991). Strontium presents a significant adsorption on geological materials. Through-diffusion experiments (TD-Sr5, TD-Sr6, TD-Sr7, and TD-Sr8) (García et al., 1998) were performed in order to estimate diffusion and retardation coefficients of strontium in compacted saturated FEBEX bentonite.

The porosity and diffusion coefficient (D_0) of the sinters of the experiments (TD-Sr5, TD-Sr6, TD-Sr7, and TD-Sr8) are prescribed to be 0.5 and $4.051 \cdot 10^{-10}$ m^2/s respectively. Upper and lower bounds and initial parameter values for strontium (Sr^{2+}) in clay plugs were derived from published results (Yuan-Hui Li et al., 1974; Yu and Neretnieks, 1996) and are listed in Table 5.2.4.

Table 5.2.4. Prior information of strontium parameters

Parameter	Diffusion coefficient (D_0, m^2/s)	Porosity (ϕ)	K_d (cm^3/g)
Lower bound	$5.5 \cdot 10^{-11}$	0.2	100
Upper bound	$5.5 \cdot 10^{-8}$	0.65	1500
Initial guess	$2.1 \cdot 10^{-9}$	0.56	500
Tolerance error	$2 \cdot 10^{-14}$	10^{-4}	10^{-4}

Strontium through-diffusion experiments (TD-Sr5, TD-Sr6, TD-Sr7, and TD-Sr8 of García et al., 1998) have been used to estimate porosity, K_d, and diffusion coefficient (D_e). First, porosity is fixed to the measured value of 0.56 and only K_d and diffusion coefficient are estimated. Then the three parameters are estimated simultaneously.

Estimation results are listed in Table 5.2.5. The best fit is plotted in Figure 5.2.9. In Table 5.2.5, D_a is the apparent diffusion coefficient, given by

$$D_a = \frac{D_e}{\phi + \rho K_d} \qquad (5.2.2)$$

where,

D_e = effective diffusion coefficient (m^2/s)

ϕ = porosity available for diffusion

K_d = distribution coefficient (cm^3/g)

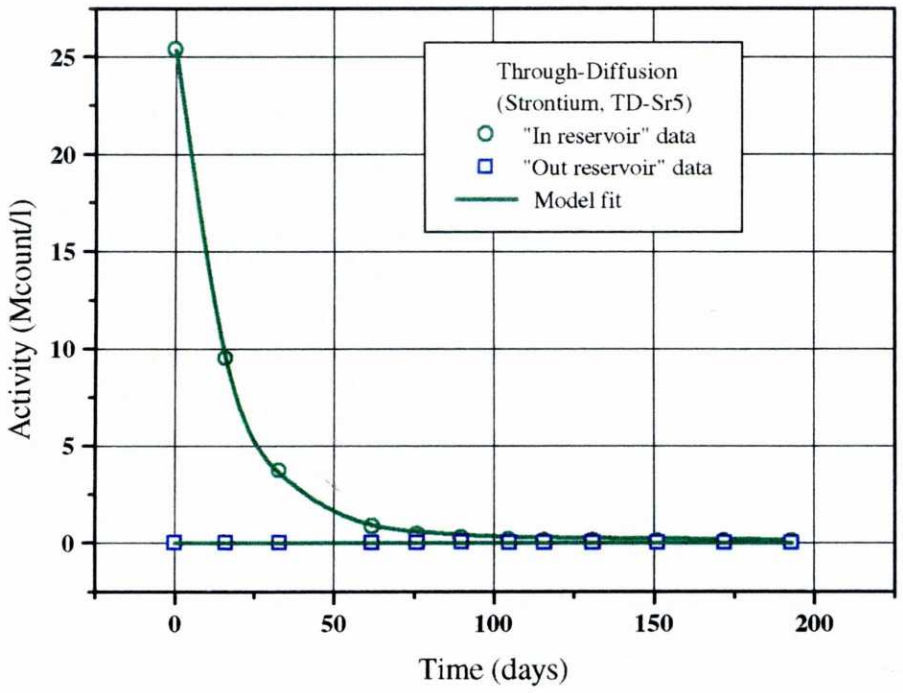
ρ = bulk dry density of the bentonite (g/cm^3)

Table 5.2.5A. Estimates of diffusion coefficient (D_0, D_e and D_a) and K_d from strontium through-diffusion experiments (porosity fixed).

Test	Porosity	K_d (cm^3/g)	Diffusion coefficients (m^2/s)			Objective
			D_0	D_e	D_a	
TD-Sr5	0.56	626	$2.422 \cdot 10^{-9}$	$1.118 \cdot 10^{-9}$	$6.864 \cdot 10^{-13}$	0.3006
TD-Sr6	0.56	557.4	$1.64 \cdot 10^{-9}$	$7.57 \cdot 10^{-10}$	$5.222 \cdot 10^{-13}$	0.4692
TD-Sr7	0.56	975.4	$3.763 \cdot 10^{-9}$	$1.737 \cdot 10^{-9}$	$6.847 \cdot 10^{-13}$	0.6434
TD-Sr8	0.56	594	$1.843 \cdot 10^{-9}$	$8.505 \cdot 10^{-10}$	$5.505 \cdot 10^{-13}$	0.3491
TD-Sr5	0.42	633.5	$3.609 \cdot 10^{-9}$	$1.135 \cdot 10^{-9}$	$6.887 \cdot 10^{-13}$	0.1977

Table 5.2.5B. Estimates of diffusion coefficient (D_0, D_e and D_a) , K_d and porosity from strontium through-diffusion experiments.

Test	Porosity	K_d (cm^3/g)	Diffusion coefficients (m^2/s)			Objective
			D_0	D_e	D_a	
TD-Sr5	0.62	621.3	$2.094 \cdot 10^{-9}$	$1.107 \cdot 10^{-9}$	$6.888 \cdot 10^{-13}$	0.1978
TD-Sr6	0.64	549.9	$1.379 \cdot 10^{-9}$	$7.603 \cdot 10^{-10}$	$5.383 \cdot 10^{-13}$	0.4561
TD-Sr7	0.52	995.1	$4.089 \cdot 10^{-9}$	$1.71 \cdot 10^{-9}$	$6.689 \cdot 10^{-13}$	0.2369
TD-Sr8	0.63	609.1	$1.545 \cdot 10^{-9}$	$8.345 \cdot 10^{-10}$	$5.273 \cdot 10^{-13}$	0.314



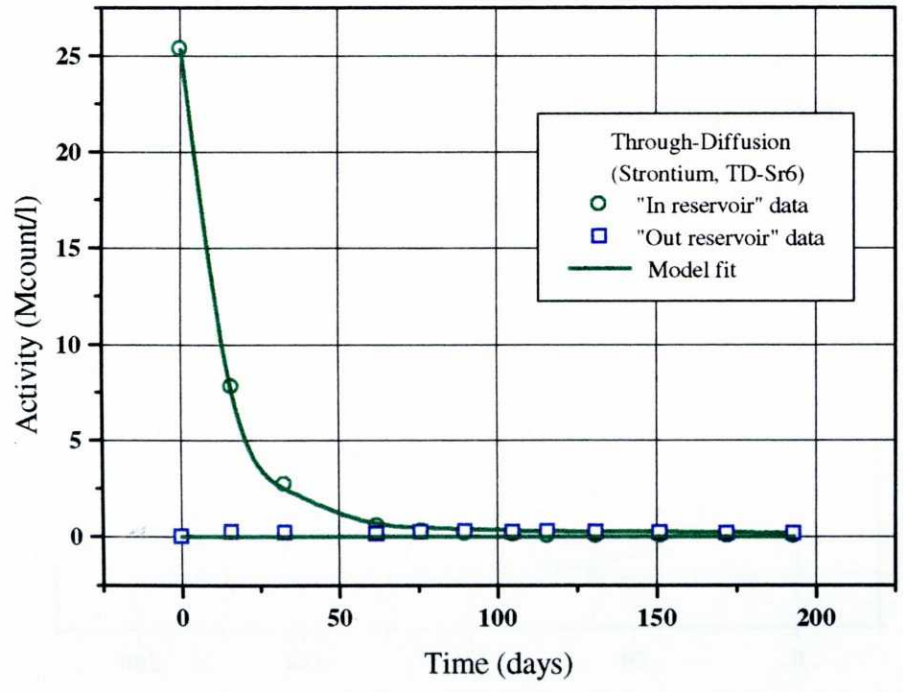
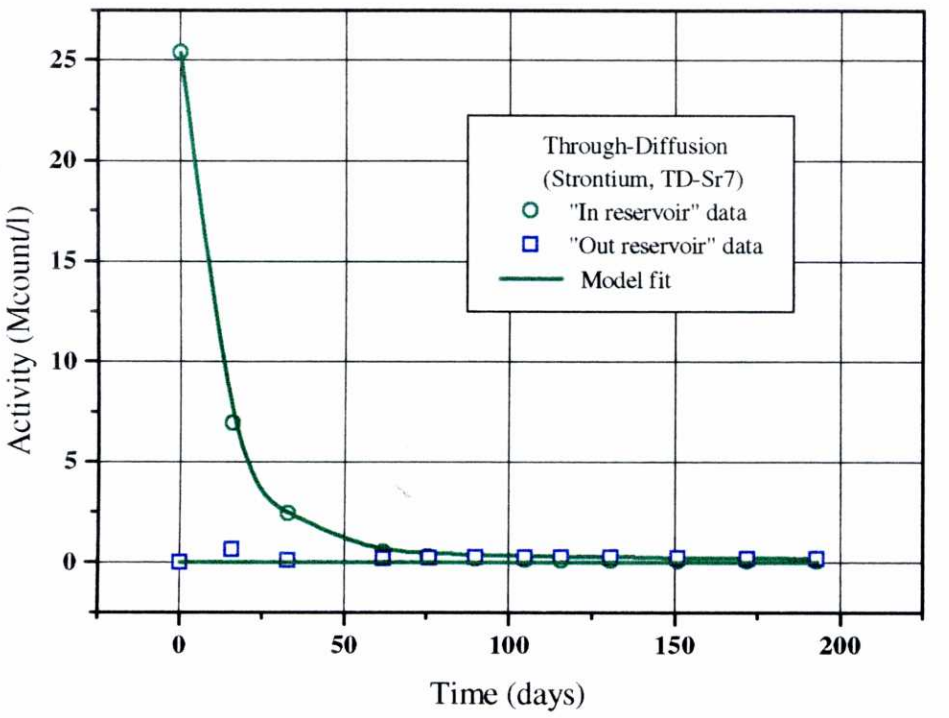


Figure 5.2.9A. Activity evolution and best fit of IN and OUT reservoirs of the through-diffusion experiments TD-Sr5 and TD-Sr6 (corresponding to the estimates of Table 5.2.5B)



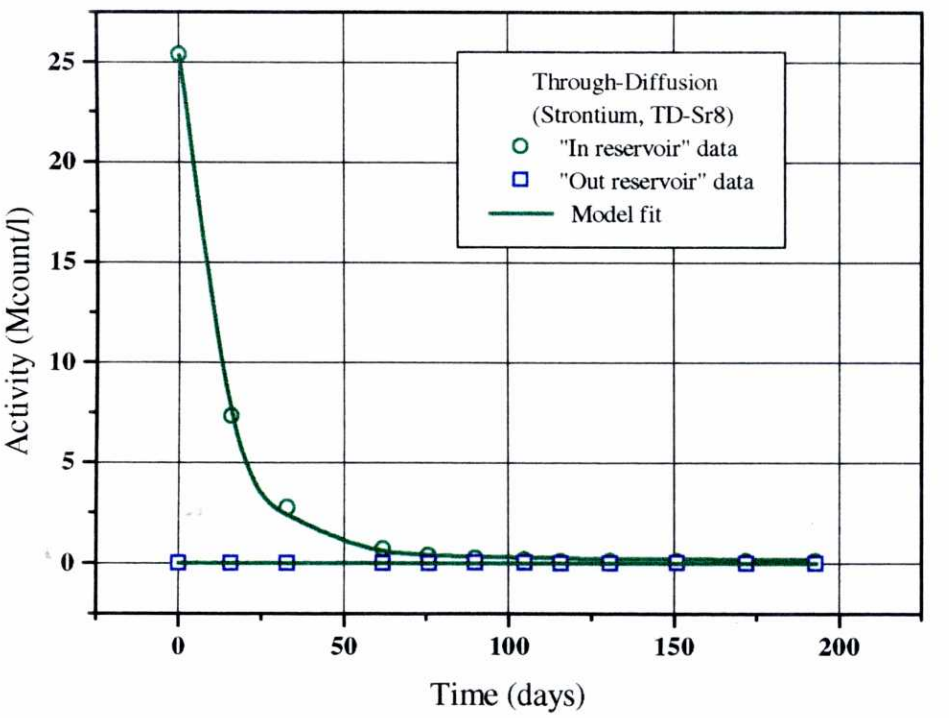


Figure 5.2.9B. Activity evolution and best fit of IN and OUT reservoirs of the through-diffusion experiments TD-Sr7 and TD-Sr8 (corresponding to the estimates of Table 5.2.5B)

Estimated porosities are generally greater than 0.56. Their mean is 0.6025. K_d values range from 549.9 to 995.1 cm^3/g. The mean is 693.85 cm^3/g. The effective diffusion coefficients range from $7.603 \cdot 10^{-10}$ to $1.71 \cdot 10^{-9}$ m^2/s with a mean of $1.103 \cdot 10^{-9}$ m^2/s. The apparent diffusion coefficients vary between $5.273 \cdot 10^{-13}$ and $6.888 \cdot 10^{-13}$ m^2/s with a mean of $6.058 \cdot 10^{-13}$ m^2/s.

From batch sorption experiments Brandberge and Skagius (1991) obtained strontium K_d values in the range 600-1300 cm^3/g in bentonite. A realistic K_d of 500 cm^3/g was proposed by Yu and Neretnieks (1996).

A summary of published effective and apparent diffusion coefficients for strontium cab be found in Yu and Neretnieks (1996). They gave the following intervals for the two parameters: $(2.4 \cdot 10^{-11}, 6.7 \cdot 10^{-9})$ and $(2.3 \cdot 10^{-13}, 4.8 \times 10^{-11})$ m^2/s, respectively. The estimated values of the parameters listed in Table 5.2.5 are consistent with those of Brandberge and Skagius (1991) and Yu and Neretnieks (1996).

One can see from Table 5.2.5 that the values of effective diffusion coefficient follow the same trend as those of K_d. In fact, the greater the value of D_e (experiments TD-Sr5 and TD-Sr7), the greater the K_d. An additional run with a fixed porosity of 0.42 was performed for experiment TD-Sr5 (see Table 5.2.5A). Even though the fit is better than that obtained with a fixed porosity of 0.56, the estimates of D_e and K_d are similar. Moreover, an equally good fit is obtained when porosity is estimated (ϕ = 0.62). Again, the estimates of D_e and K_d are almost identical. These results indicate clearly that the data of through-diffusion experiment TD-Sr5 allows one to estimate D_e and K_d, but not porosity or D_0. The estimates of D_e and K_d for TD-Sr5 show some correlation (see Figure 5.2.10), although the range of variation of these parameters is small.

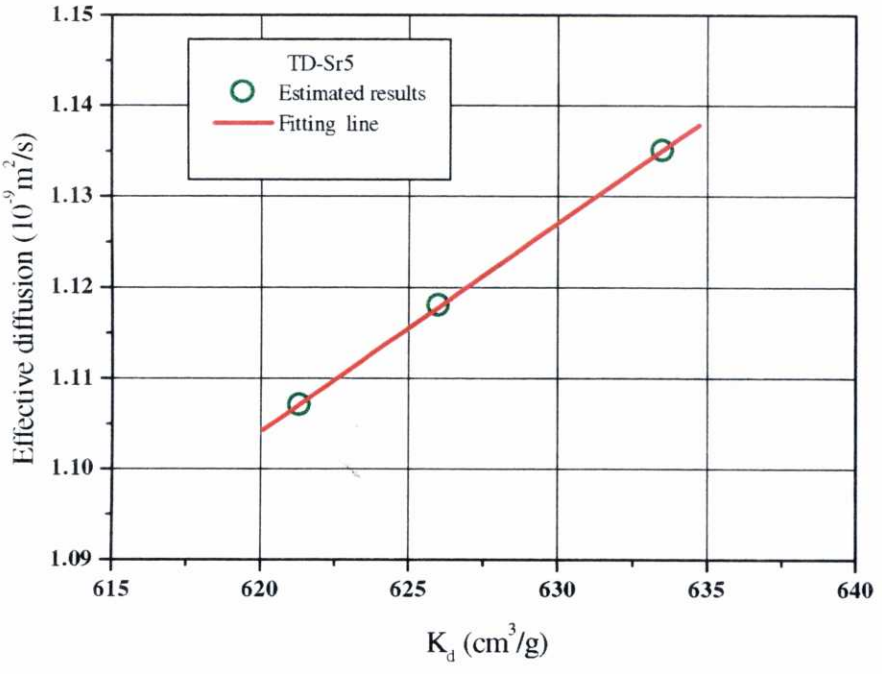


Figure 5.2.10. Linear relationship of the estimates of effective diffusion coefficient and K_d in strontium through-diffusion experiment TD-Sr5.

The activity in the upstream reservoir decreases with time reaching a near-zero value after about 80 days. After that time all the measured data are near zero at both reservoirs. This means that

1) data from the upstream reservoir after 80 days are worthless
2) the experiment was not long enough to allow the tracer to reach the downstream reservoir. Additional information on transport parameters could have been gained from this reservoir if the experiment had lasted longer
3) only 3 measurements were carried out in the upstream reservoir within the first 50 days. Parameter estimates are highly sensitive to errors affecting these measurements. This remark could explain why the estimates of experiment TD-Sr7 are so different from the others.

As a recommendation for future experiments, the measurement frequency should be increased in the first 50 days to measure better the activity evolution during the early time.

The other shortcut of these experiments is that the measured activities in the OUT reservoir are very low. If the thickness of the clay plug is reduced, for example, from 8.3 mm to 5.3 mm, it would be possible to measure non-zero activities in the downstream reservoir.

5.2.2. IN-DIFFUSION EXPERIMENTS

5.2.2.1. Experiment description and numerical models

In-diffusion experiments are designed for estimating diffusion and distribution coefficients of sorbing tracers in compacted FEBEX bentonite. Cesium and selenium were used in these experiments.

The clay was compacted in a stainless steel ring prior to emplacement into a cylindrical cell. Cells were immersed in reservoirs containing synthetic granite groundwater and maintained there until water saturation was achieved (García et al., 1998). Then the tracer was added to the reservoir solution. Since the cells were confined by sintered discs, placed on both sides of the sample with an open cap, the tracer could enter the clay from both sides. Clay plugs have 38 mm of diameter and either 25 or 50 mm of length.

Tracer activity in the reservoirs was monitored continuously. At the end each experiment the clay plug was sliced and the activity profile along the bentonite was obtained. Activity profiles are not always symmetric because differences were observed between activities measured at the upper and lower faces, most likely due to a non-uniform compaction of the clay plug. It should be pointed out that in this type of in-diffusion experiments it is possible to calculate the distribution coefficient of compacted clay directly from the steady-state activity if the experiment is long enough to obtain a uniform activity profile.

There are two kinds of numerical models in this kind of experiments, one for tracer cesium (In-Cs-9 and In-Cs-10) with 600 triangular elements and 602 nodes, and the other

for selenium (In-Se-17, In-Se-18) with 200 triangular elements and 202 nodes. The simulation time for both tracers (cesium and selenium) is 439 days and it is discretized in 16 time periods. The time increments increase gradually. Total increments for each of two cesium experiments are 630 and 370 for selenium experiment.

One of the observations in in-diffusion experiments is the final total activity along the clay plug. This activity includes the activity in the liquid and solid phases, in Mcount/per gram of sample (clay + water). In order to fit the measurements, it was needed to translate modeling results from liquid activity to total activity by means of

$$C_T = C_L \frac{\phi + \rho_d K_d}{\phi \rho_w + \rho_d} \tag{5.2.3}$$

where, C_T = total activity (Mcount/g)

C_L = liquid activity (Mcount/cm^3)

ϕ = porosity of the clay plug

K_d = distribution coefficient (cm^3/g)

ρ_d = bulk dry density of the porous medium (g/cm^3)

ρ_w =density of water (g/cm^3)

5.2.2.2. Cesium in-diffusion experiments

Cesium is largely present as unhydrated Cs^+ ion under most groundwater conditions (Yu and Neretnieks, 1996). In order to obtain accurate estimates of sorption and diffusion parameters of cesium in Spanish bentonite, CIEMAT performed several in-diffusion experiments. The bentonite was compacted in a 5 cm of length stainless steel ring with a final bulk density of 1.57g/cm^3. Two cells (ID-Cs-9 and ID-Cs-10) were totally immersed in 1250 ml of solution containing ^{137}Cs ($2.6 \cdot 10^{-9}$M). When the cell is immersed, the ^{137}Cs diffuses through the sintered stainless steel and into the pores of the clay, where it is sorbed. In fact, the activity in the contacting solution decreases. The activity of ^{137}Cs in solution was

monitored periodically by gamma counting. When the activity of ^{137}Cs in solution became constant, a steady-state condition was established. In that moment the experiment was over.

At the end of the test, the ^{137}Cs activity in the immersion solution became near zero (0.044Mcount/l), and the clay plug was sliced for measuring the ^{137}Cs activity profile. Cesium has entered through both faces of the clay approximately 15 mm. In the center of the clay plug, there is a zone in which the activity of cesium is almost zero. This indicates that the clay samples have a very large sorption capacity for cesium.

5.2.2.2.1. Interpretation using the analytical method

García et al (1998) estimated the apparent diffusion coefficients by using an approximate analytical method (Put and Henrion, 1992):

$$C_b(x,t) = \eta R C_0 \exp(Bx + B^2 Dt)\operatorname{erfc}\frac{x + 2BDt}{2\sqrt{Dt}} \qquad (5.2.4)$$

where:

C = activity in the porous material

η = diffusion accessible porosity

R = retardation factor

C_0 = initial activity in the solution

Q_0 = mass of tracer added initially at source position

S = specific surface

$B = \eta R C_0 S / Q_0$

x = position coordinate

D = diffusion coefficient

t = time

The estimated apparent diffusion coefficients were shown in Table 5.2.6:

Table 5.2.6. Cesium diffusion coefficients estimated with the analytical method (García et al., 1998)

Diffusion Coefficient (D_a, m^2/s)			
ID-Cs-9		ID-Cs-10	
Left	Right	Left	Right
$3.07 \cdot 10^{-13}$	$2.67 \cdot 10^{-13}$	$3.51 \cdot 10^{-13}$	$2.51 \cdot 10^{-13}$

5.2.2.2.2. Interpretation using numerical methods

Since the clay plugs were compacted with a hydraulic press, the lower part of the clay sample is better compacted than the upper part (García et al., 1998). Then there is some difference in diffusion between the two faces. In experiment ID-Cs-9, the difference is very small (Figure 5.2.11) and it is assumed that the clay plug is homogeneous. However, in experiment ID-Cs-10, the difference of the activity distribution between the two faces is noticeable. For the modeling purpose, the clay plug was divided into two parts (ID-Cs-10A and ID-Cs-10B) with different values of the parameters. The first measurement in the left-hand-side of ID-Cs-10 is much lower than the second one (see Figure 5.2.11). This is an anomalous datum that was taken out from the curve fitting.

A numerical model, accounting for the reservoirs, sinters and clay plug was constructed to simulate and calibrate the in-diffusion experiments ID-Cs-9 and ID-Cs-10A, and B. The estimated model parameters are porosity, diffusion and retardation coefficients.

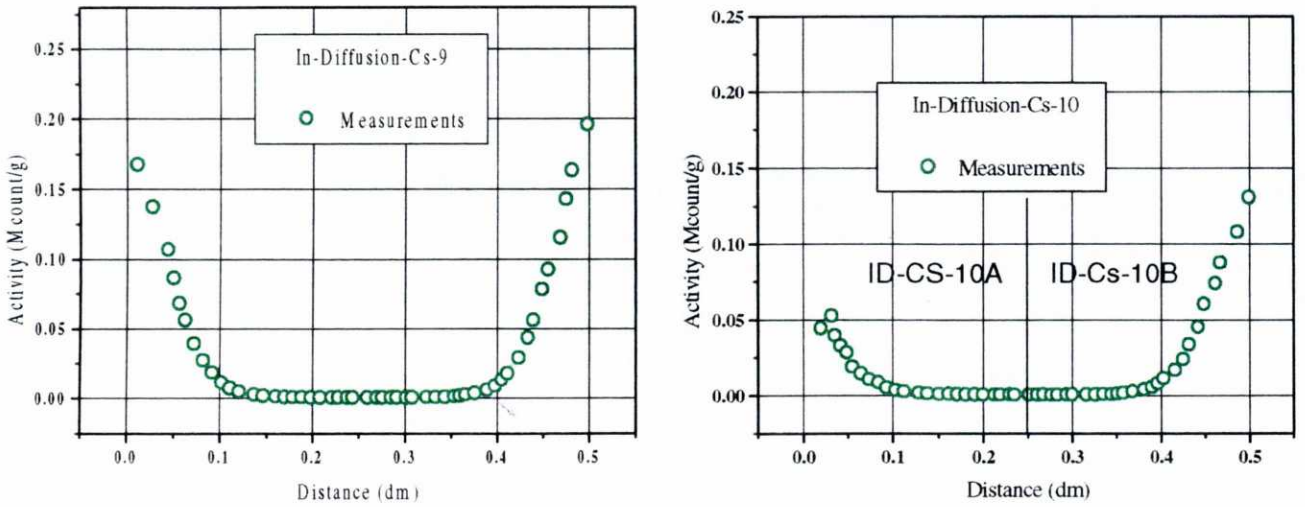


Figure 5.2.11. Measured data for experiments of ID-Cs-9 (left) and ID-Cs-10 (right)

Porosity and diffusion (D_0) of the sinters for experiments ID-Cs-9 and ID-Cs-10 are prescribed to be 0.5 and $1.6 \cdot 10^{-4}$ m^2/day, respectively. Parameters' upper and lower bounds as well as initial values are listed in Table 5.2.7. This information was derived from published data (Samper et al, 1998 and Yu and Neretnieks, 1996).

Table 5.2.7. Parameter prior information for cesium in-diffusion experiments

Parameter	Diffusion coefficient (D_0, m^2/s)	Porosity (ϕ)	K_d (cm^3/g)
Lower bound	$8.5 \cdot 10^{-11}$	0.2	200
Upper bound	$8.5 \cdot 10^{-8}$	0.65	2000
Initial guess	$2.5 \cdot 10^{-9}$	0.42	800
Tolerance error	$2 \cdot 10^{-14}$	10^{-4}	10^{-4}

Estimated parameter values are listed in Table 5.2.8 and the final model fits are plotted in Figures 5.2.12, 5.2.13, and 5.2.14.

Table 5.2.8A. Estimates of diffusion coefficient (D_0, D_e and D_a) and K_d from cesium in-diffusion experiments (porosity fixed).

ID-Cs Test	Porosity	K_d (cm^3/g)	Diffusion coefficients (m^2/s)			Objective	
			D_0	D_e	D_a	E_1	E_2
ID-Cs-9	0.42	925.1	$2.554 \cdot 10^{-9}$	$8.034 \cdot 10^{-10}$	$3.34 \cdot 10^{-13}$	6.831	$7.375 \cdot 10^{-6}$
ID-Cs-10[a]	0.42	924.1	$1.049 \cdot 10^{-9}$	$3.299 \cdot 10^{-10}$	$1.373 \cdot 10^{-13}$	4.46	$2.027 \cdot 10^{-4}$
ID-Cs-10B	0.42	907.9	$2.407 \cdot 10^{-9}$	$7.572 \cdot 10^{-10}$	$3.207 \cdot 10^{-13}$	3.045	$2.178 \cdot 10^{-5}$

Table 5.2.8B. Estimates of diffusion coefficient (D_0, D_e and D_a) , K_d and porosity from cesium in-diffusion experiments.

ID-Cs Test	Porosity	K_d (cm^3/g)	Diffusion coefficients (m^2/s)			Objective	
			D_0	D_e	D_a	E_1	E_2
ID-Cs-9	0.486	931.8	$2.106 \cdot 10^{-9}$	$8.049 \cdot 10^{-10}$	$3.322 \cdot 10^{-13}$	6.248	$6.948 \cdot 10^{-6}$
ID-Cs-10[a]	0.494	924.4	$8.817 \cdot 10^{-10}$	$3.443 \cdot 10^{-10}$	$1.423 \cdot 10^{-13}$	4.321	$9.562 \cdot 10^{-5}$
ID-Cs-10B	0.449	999.6	$2.563 \cdot 10^{-9}$	$8.813 \cdot 10^{-10}$	$3.391 \cdot 10^{-13}$	2.482	$1.054 \cdot 10^{-5}$

Note: E_1 = Objective function 1, sum of square difference between observations and modeling activities in the reservoir;
E_2 = Objective function 2, sum of square difference between observations and modeling final activities in the clay plug.

The estimated K_d of cesium ranges from 907.9 to 999.6 cm^3/g while the apparent diffusion coefficient is between $1.373 \cdot 10^{-13}$ and $3.391 \cdot 10^{-13}$ m^2/s. The latter are in agreement with the values obtained with the analytical method (Table 5.2.6).

Sorption and diffusion behavior of cesium has been extensively studied. From batch sorption experiments, Brandberge and Skagius (1991) reported K_d values for cesium in compacted bentonite in the range 100 to 4173 cm^3/g. A realistic K_d of 500 cm^3/g is proposed by Yu and Neretnieks (1996). A summary of published effective and apparent diffusion coefficients for cesium in compacted bentonites was presented by Yu and

Neretnieks (1996) according to which, these parameters are in the following ranges: ($1.3 \cdot 10^{-11}$, $4.5 \cdot 10^{-9}$) and ($1.7 \cdot 10^{-13}$, $4 \cdot 10^{-11}$ m^2/s), respectively. The estimated parameters in Table 5.2.8 are consistent with those reported by Brandberge and Skagius (1991) and Yu and Neretnieks (1996).

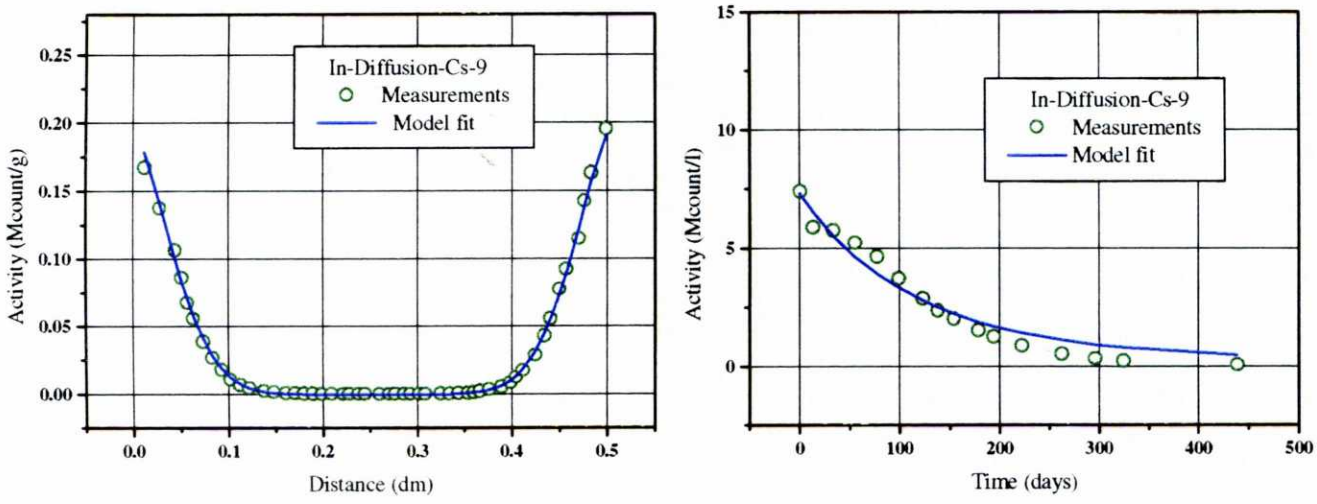


Figure 5.2.12. Best fit of activities in clay plug (left) and reservoir (right) of ID-Cs-9

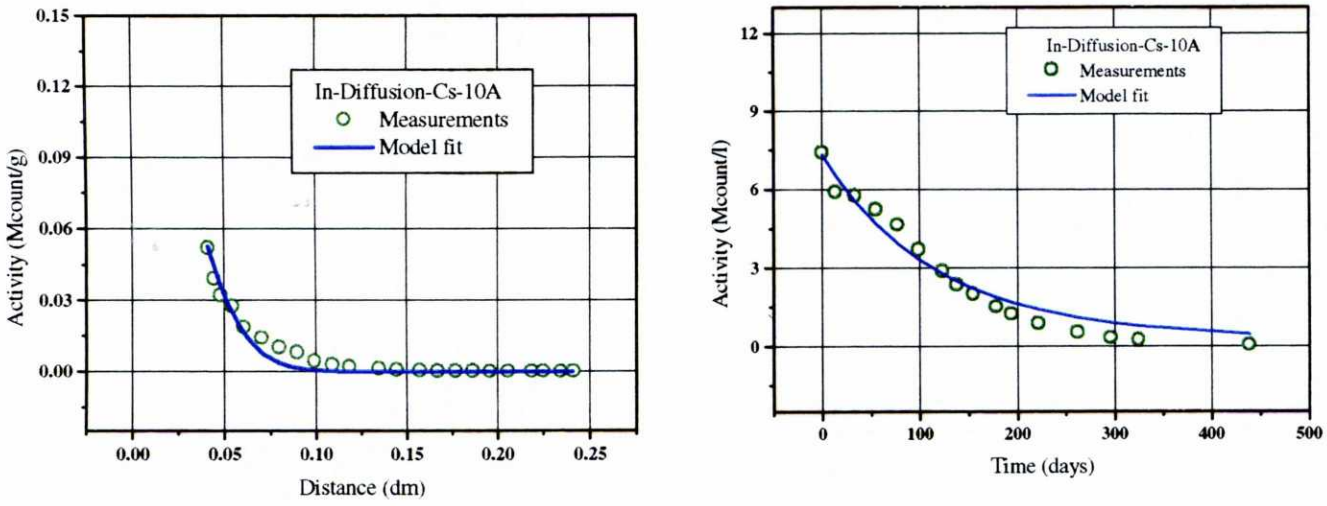


Figure 5.2.13. Best fit of activities in clay plug (left) and reservoir (right) of ID-Cs-10A

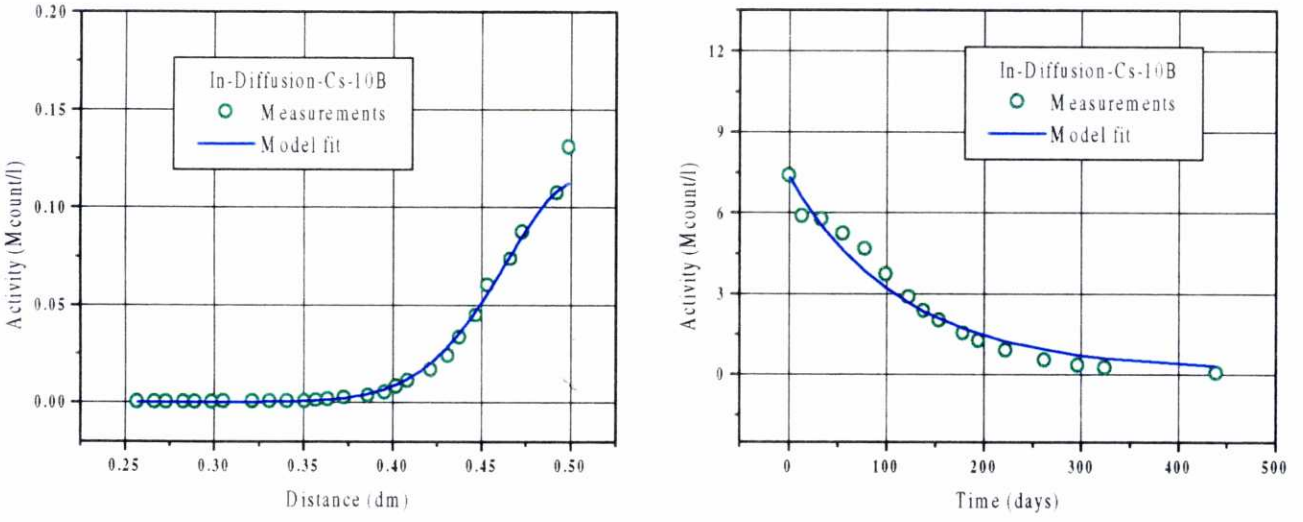


Figure 5.2.14. Best fit of activities in clay plug (left) and reservoir (right) of ID-Cs-10B

5.2.2.3. Selenium in-diffusion experiments

Two selenium in-diffusion experiments (ID-Se-17 and ID-Se-18) were performed by CIEMAT. Different from cesium experiments, the clay was compacted in a 2.5 cm length stainless steel ring. Selenium (IV) in the oxic water (pH over 7.5) has the form of the selenite $SeO3^{2-}$. Contrary to selenate, selenite is sorbed in bentonites. Moreover, it decays with a half life of 118 days.

A numerical model, accounting for radioactive decay and retardation, was carried out to interpret experiments ID-Se-17 and ID-Se-18. Porosity and diffusion coefficient (D_0) of the sinters of the in-diffusion experiments (ID-Se-17 and ID-Se-18) were prescribed to be 0.5 and $7.5 \cdot 10^{-5}$ m^2/day, respectively. The parameters' upper and lower bounds and initial values are listed in Table 5.2.9. These values were derived from published data (Samper et al, 1997 and Yu and Neretnieks, 1996).

Table 5.2.9. Parameter prior information for selenium in-diffusion experiments

Parameter	Diffusion coefficient (D_0, m^2/s)	Porosity (ϕ)	K_d (cm^3/g)
Lower bound	$1.5 \cdot 10^{-13}$	0.2	0.5
Upper bound	$8.5 \cdot 10^{-9}$	0.65	20
Initial guess	$2.5 \cdot 10^{-10}$	0.42	5
Tolerance error	$2 \cdot 10^{-15}$	10^{-4}	10^{-4}

Estimated results obtained under different conditions are listed in Table 5.2.10. The best fit is obtained by estimating ϕ, D_0 and K_d (see Figures 5.2.15 and 5.2.16).

Table 5.2.10A. Estimates of diffusion coefficient (D_0, D_e and D_a) and K_d from selenium in-diffusion experiments (porosity fixed).

Test	Porosity	K_d (cm^3/g)	Diffusion coefficients (m^2/s)			Objective	
			D_0	D_e	D_a	J_1	J_2
ID-Se-17	0.42	2.41	$1.383 \cdot 10^{-12}$	$4.35 \cdot 10^{-13}$	$6.507 \cdot 10^{-14}$	2.325	$3.486 \cdot 10^{-7}$
ID-Se-18	0.42	2.17	$4.41 \cdot 10^{-13}$	$1.387 \cdot 10^{-13}$	$2.288 \cdot 10^{-14}$	2.609	$4.892 \cdot 10^{-8}$

Table 5.2.10B. Estimates of diffusion coefficient (D_0, D_e and D_a) , K_d and porosity from selenium in-diffusion experiments.

Test	Porosity	K_d (cm^3/g)	Diffusion coefficients (m^2/s)			Objective	
			D_0	D_e	D_a	J_1	J_2
ID-Se-17	0.21	1.87	$1.412 \cdot 10^{-12}$	$1.726 \cdot 10^{-13}$	$3.404 \cdot 10^{-14}$	1.39	$1.831 \cdot 10^{-7}$
ID-Se-18	0.25	2.02	$7.804 \cdot 10^{-13}$	$1.229 \cdot 10^{-13}$	$2.234 \cdot 10^{-14}$	2.581	$4.713 \cdot 10^{-8}$

According to Brandberge and Skagius (1991), the K_d of selenium in compacted bentonites ranges from 1 to 2 cm^3/g at pH between 8.1 and 9.1. Sorption data of selenium on Ca-bentonite was also compiled by Oscarson et al. (1994). At a pH of 7, ionic strength of about 0.09 M and Eh of –160 mV, a K_d of 3 cm^3/g was reported. A realistic K_d of 3 cm^3/g is proposed by Yu and Neretnieks (1996). Estimated K_d values in Table 5.2.2 (from

1.87 to 2.41 cm^3/g) are consistent with those reported by Brandberge and Skagius, (1991), Oscarson et al. (1994) and Yu and Neretnieks (1996).

Estimated effective diffusion coefficients for selenium are between $1.229 \cdot 10^{-13}$ and $4.35 \cdot 10^{-13}$ m^2/s while apparent diffusion coefficients are between $2.234 \cdot 10^{-14}$ and $6.507 \cdot 10^{-14}$ m^2/s. These values are lower than those of cesium even though the K_d of selenium is 400 times smaller than that of cesium (see Table 5.2.3).

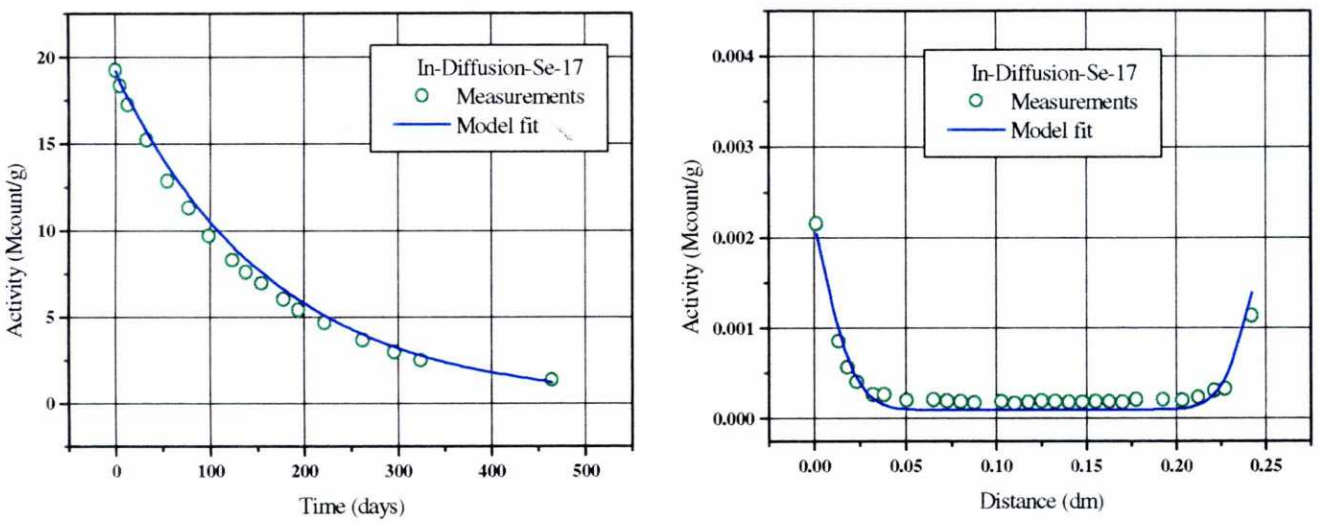


Figure 5.2.15. Best fit of activity evolution in reservoir (left) and final profile activity in clay plug (right) of ID-Se-17

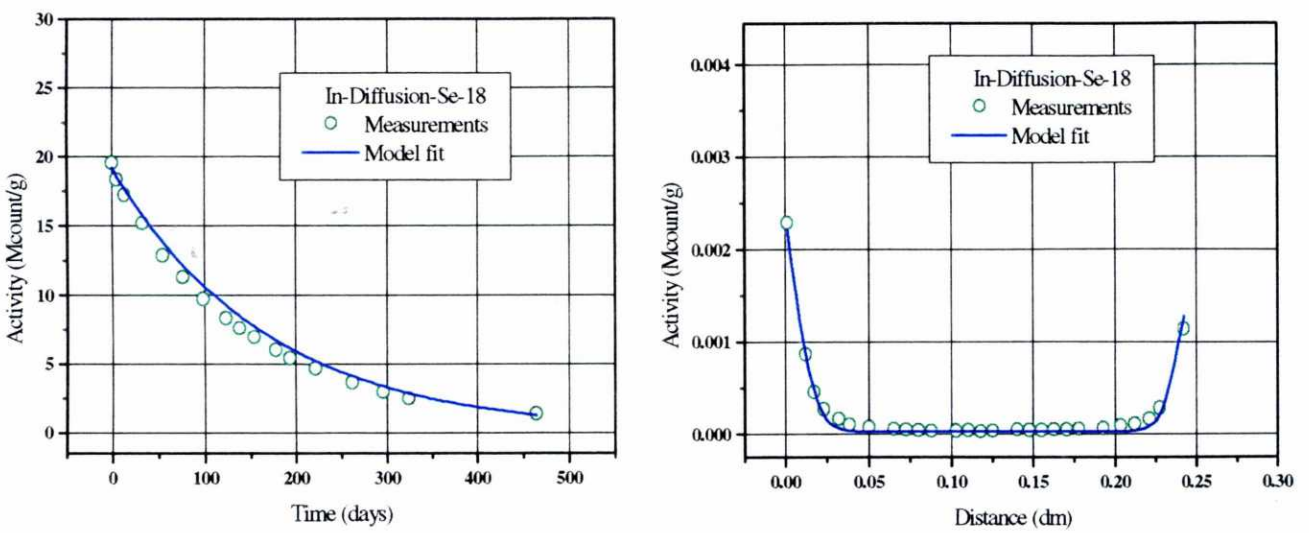


Figure 5.2.16. Best fit of activity evolution in reservoir (left) and the final profile activity in clay plug (right) of ID-Se-18

A more detailed sensitivity analysis of the role of the reservoir volume and other experimental conditions is described in a FEBEX Technical Report (Dai et al., 1999).

5.2.3. PERMEATION EXPERIMENTS

5.2.3.1. Permeation experiments and numerical models

A high-pressure stainless steel cell was used for permeation experiments. The clay plug of 5.3 mm in length and 50 mm of diameter was sealed between two 3.35 mm thick stainless steel sinters. Granite synthetic groundwater (García et al., 1998) was introduced into the sample with a low driving pressure (6 bar) and flow regulation (between 2.7 and 4.0 ml/day). After 4 weeks' saturation, a small volume (as a pulse) of tracer was injected into the fluid pathway and the breakthrough activity was monitored continuously by a fraction collector. The hydrodynamic characterization of the clay used in these migration experiments was performed by means of tritiated water (HTO).

Two types of numerical models were used for modeling these experiments, single and double porosity models. The former only requires a one-dimensional model. For the latter, a 2-D grid was used which consists of a mobile and an immobile part. Advection, molecular diffusion and hydrodynamic dispersion take place at the mobile part. In the immobile part there is only molecular diffusion (see Figure 5.2.17).

Tracer injection is simulated with a time-dependent function. Since tracer was injected as a pulse of unknown duration, a trial and error method is used to derive such time function. The initial activity is zero in all the nodes.

For the single porosity model a 1-D grid with 200 triangular elements and 202 nodes was used (P-1 and P-2). This model failed to match the observations. Then, a double porosity 2-D model (P-1d and P-2d) with 560 triangular elements and 324 nodes was used (Figure 5.2.17).

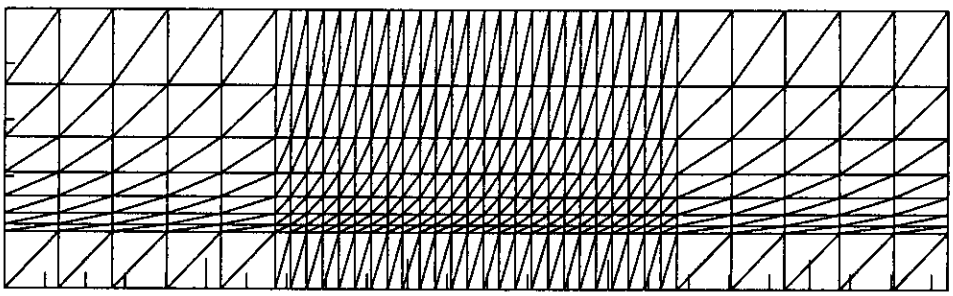


Figure 5.2.17. The finite element mesh of the two-D double-porosity model

The simulation time is 3.78 days which was discretized with 88 time periods with gradually-increasing time increments.

5.2.3.2. Numerical interpretation

5.2.3.2.1. Single porosity model

A single porosity numerical model with two parameter zones, sinter and clay, was first used to model flow and solute transport. The Golden section search method was used to estimate porosity, molecular diffusion and dispersivity. Porosity and diffusion coefficient (D_0) of the sinters were prescribed to be 0.5 and $2.454 \cdot 10^{-9}$ m^2/s. Information of initial values and upper and lower bounds of parameters is listed in Table 5.2.11.

Table 5.2.11. The prior information of the parameters of the permeation experiments

Parameter	Diffusion coefficient (D_0, m^2/s)	Porosity (ϕ)
Lower bound	$2.5 \cdot 10^{-11}$	0.2
Upper bound	$2.5 \cdot 10^{-8}$	0.56
Initial guess	$8.1 \cdot 10^{-10}$	0.41
Tolerance error	$2 \cdot 10^{-14}$	10^{-4}

Dispersivity was fixed to be $5 \cdot 10^{-4}$ m after a few trial runs. The identification of the time function for the tracer injection pulse was a hard task. A trial and error method was used to derive this time function. After more than 100 trials, a time function for experiments P-1 and P-2 was selected (see function 3 in Figure 5.2.22).

For permeation experiment 1 (P-1), a local minimum was obtained (see Figure 5.2.18). For permeation experiment 2 (P-2), the porosity is fixed to a value similar to that obtained from P-1 (0.35). But, the algorithm failed to attain a convergent solution. The best results for this case are shown in Figure 5.2.19. The estimated results are listed in Table 5.2.12.

Table 5.2.12. The estimated results of single porosity models (dispersivity=$5 \cdot 10^{-4}$ m)

Name of Test	Porosity	Tortuosity	D_0 (m^2/s)	De (m^2/s)	Objective
P-1	0.35	0.705	$1.898 \cdot 10^{-9}$	$4.682 \cdot 10^{-10}$	98.348
P-2	0.35(fixed)	0.705	$2.558 \cdot 10^{-9}$	$6.309 \cdot 10^{-10}$	238.109
Mean	0.35	0.705	$2.228 \cdot 10^{-9}$	$5.496 \cdot 10^{-10}$	168.229

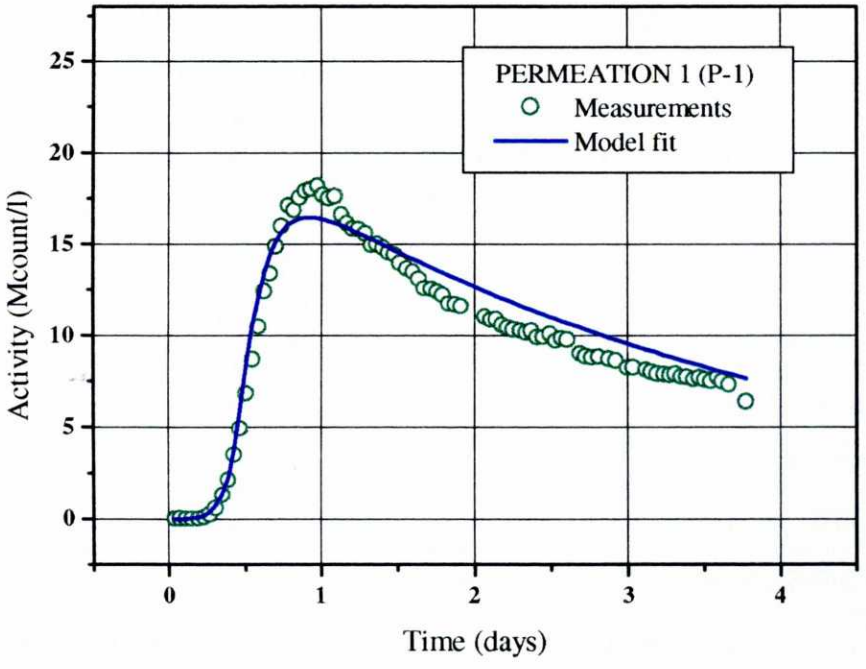


Figure 5.2.18. Activity evolution and best fit of permeation experiment 1 (P-1)

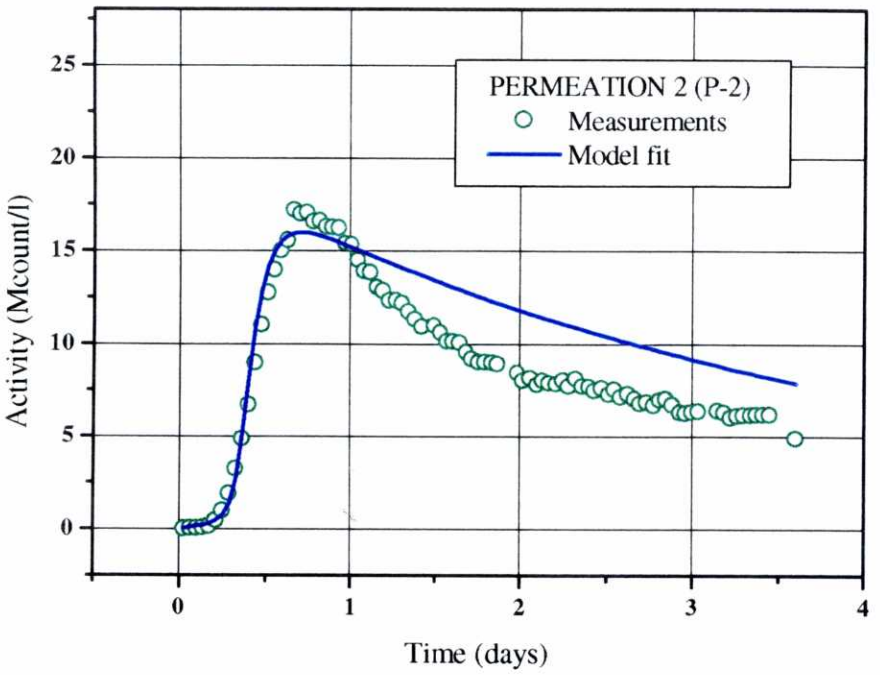


Figure 5.2.19. Activity evolution and best fit of permeation experiment 2 (P-2)

Although many trial runs were made to fit the measurements of P-1 and P-2, the results shown in Figures 5.2.18 and 5.2.19 could not be improved. Lack of agreement between measured and computed activities is usually interpreted to mean that the model structure is unreliable.

Compacted bentonite used in the permeation experiments has a great porosity (0.56) with only a part of the pores being effective for water flow and tracer advection. A significant part of porosity contains unconnected or dead-end pores which do not participate in advection (Bear, 1975). Therefore, a single porosity model cannot simulate the double porosity behavior of compacted bentonite.

5.2.3.2.2. Double porosity model

Porosity is a key transport parameter in clays. The literature on modeling solute transport through clays indicates that the porosity structure plays a major role on solute transport (Carrera et al., 1991; ENRESA, 1992).

Effective, flowing and kinematic porosity are terms used synonymously to describe the part of the total porosity along which fluid flows. Many porous media contain unconnected or dead-end pores which do not participate in advection (Bear, 1975). In clays, it seems likely that most of the pores are actually interconnected, but many interparticle spaces are so small that viscous drag, surface adsorption and other effects render the water in these voids effectively immobile (Horseman et al., 1996). The double porosity (or dual porosity) approach in clays is equivalent to the concept of matrix diffusion in fractured media.

Surface charge phenomena also affect the accesible porosity. In negatively charged clays, anions tend to be repelled from the particle surfaces, suffering what is known as anion exclusion. The accesible porosity of anions is therefore smaller than that of uncharged and possitively charged species. This means that porosity is also a solute-dependent parameter which must be determined from advection and dispersion experiments at different pH and salinity values.

The double porosity model includes two parts, mobile and immobile parts. An increased level of complexity can be built into the single porosity model by assuming that equilibrium is not reached between the mobile part (large pores) and the immobile part (smaller pores which may not be well connected). Furthermore, it is assumed that almost all of the water will flow through the mobile part but a mass transfer (by molecular diffusion) will exist between mobile and immobile parts.

A double porosity numerical model including three parameter zones, one zone for the sinter, two zones for the clay, (mobile and immobile zones,) has been used for modeling flow and solute transport in both permeation experiments (models P-1d and P-2d).

The Golden section search method was used to estimate porosity, molecular diffusion coefficient and dispersivity in the mobile and immobile zones. Information of their initial guess values and upper and lower bounds is presented in Table 5.2.13. The porosity and diffusion coefficient (D_0) of the sinters are prescribed to be 0.5 and $2.454 \cdot 10^{-9}$ m^2/s.

Table 5.2.13. Prior information of the parameters for double-porosity model

Parameter	Diffusion coefficient (D_0, m^2/s)		Porosity (ϕ)	
	mobile	immobile	mobile	immobile
Lower bound	$2.5 \cdot 10^{-11}$	$2.5 \cdot 10^{-14}$	0.2	0.1
Upper bound	$2.5 \cdot 10^{-8}$	$2.5 \cdot 10^{-10}$	1.25	0.85
Initial guess	$8.1 \cdot 10^{-10}$	$1.1 \cdot 10^{-12}$	0.56	0.56
Tolerance error	$2 \cdot 10^{-14}$		10^{-4}	

After several trial runs, dispersivity was fixed to $5 \cdot 10^{-4}$ m to reduce the number of estimated parameters. The trial and error method was used to derive the time function for tracer injection pulse. The best results achieved with the function (see Figure 5.2.22) are plotted in Figures 5.2.20 and 5.2.21. The final estimated results are presented in Table 5.2.14. They are much better than those obtained with the single porosity model.

Table 5.2.14. Estimated results of permeation experiments using a double porosity model (* denotes model porosities which are translated into sample porosities)

Test	Porosity				D_0 (m^2/s)		D_e (m^2/s)		Obj.
	mobile		immobile		mobile	immobile	mobile	immobile	
P-1d	0.19	0.969*	0.44	0.547*	$3.651 \cdot 10^{-9}$	$2.54 \cdot 10^{-12}$	$3.502 \cdot 10^{-9}$	$1.136 \cdot 10^{-12}$	12.21
P-2d	0.2	1.016*	0.6	0.753*	$2.277 \cdot 10^{-9}$	$2.747 \cdot 10^{-12}$	$2.326 \cdot 10^{-9}$	$1.88 \cdot 10^{-12}$	39.16
Mean	0.2	0.993*	0.52	0.65*	$2.941 \cdot 10^{-9}$	$2.678 \cdot 10^{-12}$	$2.914 \cdot 10^{-9}$	$1.508 \cdot 10^{-12}$	25.69

Values of porosity denoted with a star * (in Table 5.2.14) represent model porosity. The relationship between model porosity and physical porosity is given by

$$\phi = (\phi_m^* L_m + \phi_{im}^* L_{im})/(L_m + L_{im}) \quad (5.2.5)$$

where,

$\phi_m = \phi_m^* L_m/(L_m + L_{im})$

$\phi_{im} = \phi_{im}^* L_{im}/(L_m + L_{im})$

ϕ = total porosity, $\phi = \phi_m + \phi_{im}$

ϕ_m = physical porosity in mobile part

ϕ_{im} = physical porosity in immobile part

ϕ_m^* = porosity of the mobile part of the numerical model

ϕ_{im}^* = porosity of the immobile part of the numerical model

L_m= the length of mobile part in the model (L_m = 0.2 mm)

L_{im} = the length of immobile part in the model (L_{im} = 0.8 mm)

According to Equation 5.2.5, total porosity for P-1d is 0.6314 and 0.8052 for P-2d. The effective diffusion coefficient in the mobile part is three orders of magnitude greater than that of the immobile part.

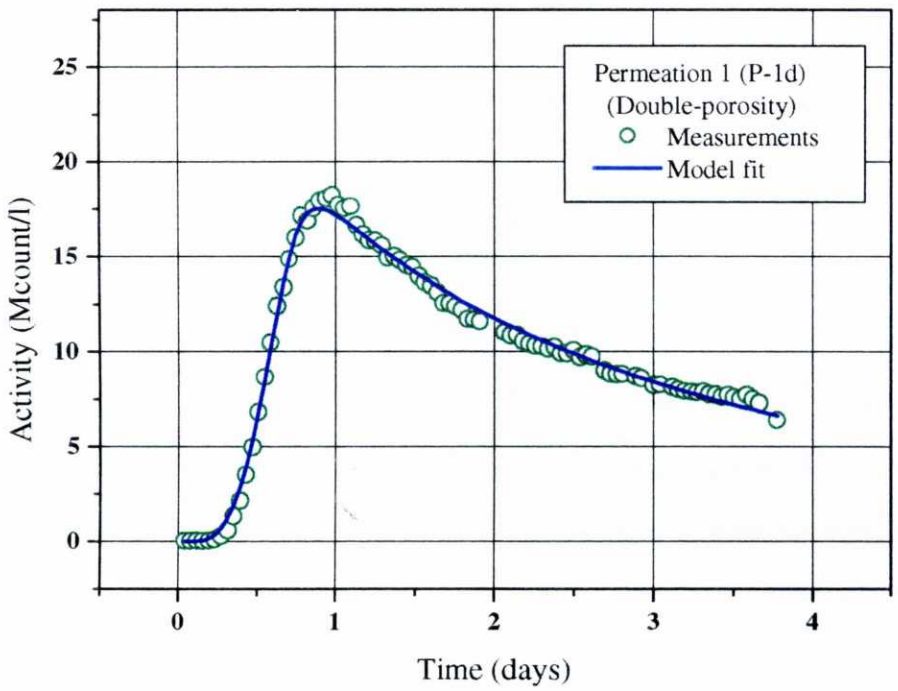


Figure 5.2.20. Actvity evolution and curve fit of permeation experiment 1 (P-1d) using a double porosity model

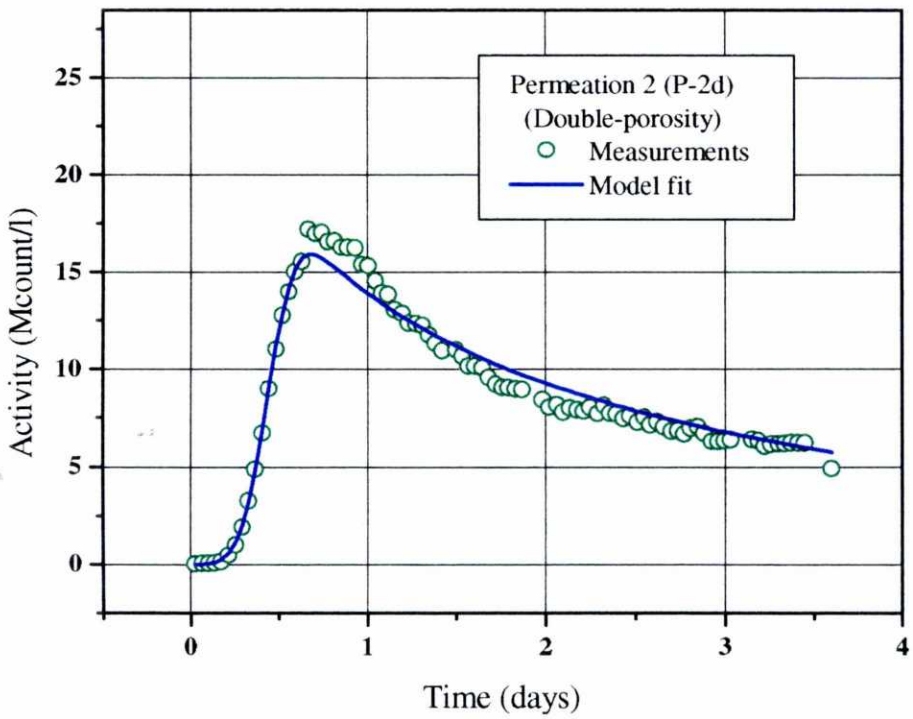


Figure 5.2.21. Actvity evolution and curve fit of permeation experiment 2 (P-2d) using a double porosity model

5.2.3.3. Experiment analysis

The interpretation of these permeation experiments was much more difficult than diffusion experiments.. It took a great quantity of time to identify model structure and boundary conditions. In this case modifications of model structure and boundary conditions had much more effect on curve fit than changes in parameters. A reliable conceptual model is essential to ensure good parameter estimation results. Our inverse model can automatically estimate the optimum model parameters, but cannot account for the identification of model structure and boundary conditions. Such identification must be done manually by trial and error.

On the other hand, a well-designed experiment can decrease the problem of identifying model structure and boundary conditions. In the permeation experiments, some experiment conditions, including the process and speed of trace injection are uncertain. Even though we have done a large number of sensitivity analyses to model initial and boundary conditions, we cannot completely simulate the time-space variations of these uncertainties.

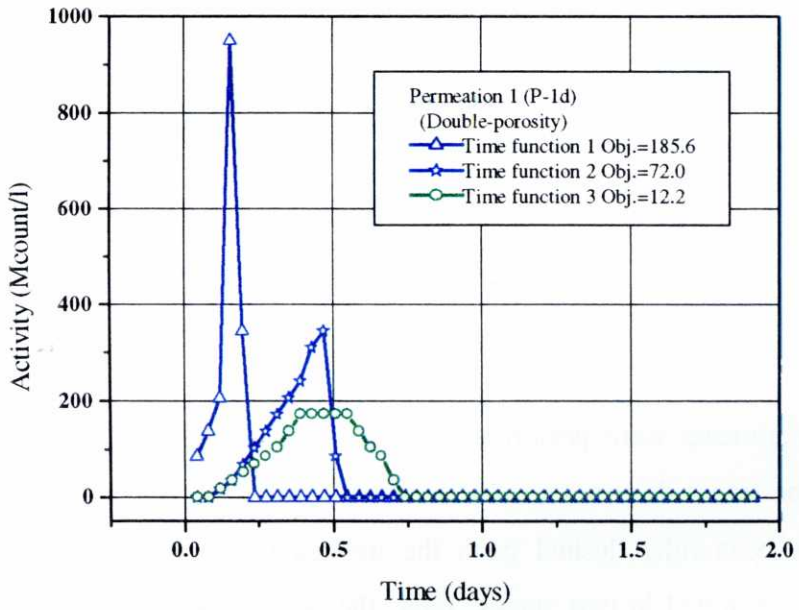


Figure 5.2.22. Time functions used to define the tracer pulse of the inflow water. Also shown are the values of the objective function for each time function. One can see the large sensitivity of the fit to the selected time function.

5.2.4. SUMMARY AND CONCLUSIONS

5.2.4.1. Through-diffusion experiments

Two types of diffusion experiments have been interpreted. The first one is the so-called through-diffusion experiment, in which the clay sample is placed between two water reservoirs. A tracer is initially added only to one of the reservoirs (the upstream reservoir). Neither the clay nor the sinters contain this tracer. The gradient in tracer activity induces a diffusive flux from the upstream to the downstream reservoir through the clay sample. Several sets of through-diffusion (TD) experiments were performed on clay samples having a total porosity of 0.56 which amounts to a dry density of 1.18 g/cm^3 using tritiated water (HTO) and strontium.

HTO was used in three TD experiments which lasted about 50 days, a time long enough for the experiment to achieve steady activities. HTO experiments were used to test the estimation algorithms and to explore the role of the sinters. Our results indicate that the sinters should be taken into account for a proper interpretation of the experiments. Otherwise, bentonite diffusion coefficients can be underestimated by a factor of 3. Based on this finding, the sinters were considered for the interpretation of all the experiments. In general, the numerical solution is able to reproduce quite well the observed tracer data in both reservoirs. Effective diffusion coefficients for tritium range from 1.17 to $1.4 \cdot 10^{-10}$ m^2/s, which are within the range of published values for this tracer in bentonites. The best fit is obtained with diffusive porosities ranging from 0.44 to 0.5 which are slightly smaller than the total porosity. Tortuosity factor ranges from 0.76 to 0.79.

Four experiments were performed using strontium, Sr^{2+}. This tracer has a large sorption. In fact, the tracer does not reach the downstream reservoir after almost 200 days, even though it has been fully flushed from the upstream reservoir after that time. These experiments were interpreted in two stages. First, the porosity was fixed to a value of 0.56 while K_d and D_o were estimated. In the second stage, porosity was also estimated, resulting in an excellent fit to measured data. The best fit is obtained with diffusive porosities slightly greater than the total porosity. Mean diffusive porosity is 0.6. Effective diffusion coefficients

D_e range from $7.6 \cdot 10^{-10}$ to $1.71 \cdot 10^{-9}$ m^2/s, with a mean value of $1.1 \cdot 10^{-9}$ m^2/s. Estimated values of K_d show also a wide range (from 555 to 995.1 cm^3/g) with a mean of 693.5 cm^3/g. Apparent diffusion coefficients, which take into account the effect of sorption, show a much less scatter (from $5.27 \cdot 10^{-13}$ to $6.88 \cdot 10^{-13}$ m^2/s) with a mean value of $6.05 \cdot 10^{-13}$ m^2/s). These values are within the range of published values for this tracer in bentonites. Estimated values of K_d and D_e show a positive correlation, i.e., the larger D_e the greater K_d. The fact that the apparent diffusion coefficients, D_a, obtained with different experiments are quite stable leads us to conclude that TD experiments for sorbing tracers only allow the estimation of D_a. The estimates of K_d and D_e may contain a significant uncertainty.

5.2.4.2. In-diffusion experiments

The second type of diffusion experiments is the so-called in-diffusion (ID) experiment, in which two clay samples of identical dimensions are immersed in a reservoir containing a tracer which is initially present neither on the clay nor on the sinters. The gradient in tracer activity induces a diffusive flux from the reservoir into the clay samples. Several sets of ID experiments were performed on clay samples having a total porosity of 0.42 which amounts to a dry density of 1.57 g/cm^3 using radioactive Cesium (^{137}Cs) and selenite, SeO_3^{2+}. Tracer activity in the reservoir was monitored during the duration of the experiment. At the end of the experiment clay samples were sliced and total tracer activity was measured in each slice.

Two Cesium experiments were available (ID-Cs-9 and ID-Cs-10). In one of them, the behavior of the sample is not symmetric, i.e. the tracer diffuses more in one of the faces. In this case, data from each side of the sample were interpreted separately. The experiments were interpreted in two stages. First, the porosity was fixed to a value of 0.42 while K_d and D_o were estimated. In the second stage, porosity was also estimated, resulting in an excellent fit to measured final activity data. Some discrepancies are observed in the time evolution of tracer activity in the reservoir which are attributed to uncertainties in the initial activity and the lack of well-mixed conditions in the reservoir. A detailed sensitivity analysis was performed to evaluate the effects of several sources of uncertainty which include: (1) the initial activity in the sinters, (2) the appropriate boundary condition for the reservoir, (3) the initial activity in the

reservoir, and (4) the effective volume of the reservoir. The results of this analysis indicates that estimated parameters are very sensitive to the properties and conditions prevailing at the reservoir. Therefore, for the proper interpretation of this type of experiments we suggest that: (1) only a clay sample should be located in each reservoir, (2) the initial activity should be measured as accurately as possible, (3) special care should be taken in measuring the activities of the slices located near the faces of the clay samples.

A better fit is obtained with diffusive porosities slightly greater than the total porosity. Diffusive porosities range from 0.45 to 0.49. Effective diffusion coefficients D_e range from $3.44 \cdot 10^{-10}$ to $8.81 \cdot 10^{-10}$ m^2/s. Estimated values of K_d range from 924.4 to 999.6 cm^3/g. Apparent diffusion coefficients, which take into account the effect of sorption show some scatter. They range from $1.43 \cdot 10^{-13}$ to $3.39 \cdot 10^{-13}$ m^2/s. These values are within the range of published values for this tracer in bentonites.

Two selenite, SeO_3^{2-} experiments were performed on clay samples of 2.5 cm, half the length of the Cs cells. Again, the experiments were interpreted in two stages. First, the porosity was fixed to a value of 0.42 while K_d and D_o were estimated. In the second stage, porosity was also estimated, resulting in an excellent fit to both final activity data and time evolution of tracer activity in the reservoir. A better fit is obtained with diffusive porosities much smaller than the total porosity. Diffusive porosities are equal to 0.21 and 0.25. Effective diffusion coefficients D_e range from $1.23 \cdot 10^{-10}$ to $1.72 \cdot 10^{-10}$ m^2/s. Estimated values of K_d range from 1.87 to 2.02 cm^3/g. Apparent diffusion coefficients, which take into account the effect of sorption range from $2.23 \cdot 10^{-13}$ to $3.40 \cdot 10^{-13}$ m^2/s. Estimated values of K_d, D_e and D_a are within the range of published values for selenium in bentonites.

5.2.4.3. Permeation experiments

Contrary to in- and through-diffusion experiments, permeation experiments involve advective solute transport in addition to diffusive processes. These experiments are intended to provide information on dispersivities and kinematic porosity, which measures the volume of well-connected pores. Several permeation experiments were performed using tritiated water and Se. Here we only report the results of the interpretation of HTO permeation experiments.

The results of the Se experiments show unpredictable and hardly explainable patterns. Even though a large number of trials has been made in order to interpret Se permeation experiments, no conclusive results have been found yet. These experiments will be included in future FEBEX reports. Following the parsimony principle, the classic single-porosity model was tested first. For the first permeation experiment (P-1) a convergent but suboptimal solution is obtained with a porosity of 0.35. The fit of the second experiment is very poor, in fact, the estimation algorithm fails to attain a convergent solution. Given the limitations of the single-porosity model, a more complex double-porosity model was also considered. With this model the fit of both tritium experiments is excellent. Mobile or kinematic porosity is 0.2 while immobile porosity ranges from 0.44 to 0.6. The effective diffusion coefficient in the mobile phase ranges from 2.32 to $3.50 \cdot 10^{-9}$ m^2/s, while that of the immobile phase is almost three orders of magnitude smaller.

5.2.4.4. Uncertainty analysis

The numerical interpretation of diffusion and permeation experiments has been useful to identify the following key issues which should be analyzed more in depth in the future:

1) The role and properties of the sinters. Their effective diffusion properties should be measured
2) the initial tracer activity in the reservoirs
3) the effective volume of the reservoir in in-diffusion experiments
4) the activities of the slices located near the faces of the clay samples in in-diffusion experiments
5) the appropriate duration of the experiment which should depend on the sorbing properties of the tracer

Some suggestions for future experiments are:

(1) Only a clay sample should be located in each reservoir for in-diffusion experiments
(2) the initial tracer activity should be measured as accurately as possible
(3) special care should be taken in measuring the activities of the slices located near the faces of the clay samples

(4) the duration of the tracer pulse should be much smaller than the duration of permeation experiments. Otherwise, the activity of the inflow water should be carefully measured

(5) additional diffusion experiments under unsaturated conditions are needed to explore the variation of anion exclusion (excluded porosity) with water content.

(6) radial diffusion experiments in order to identify better the retardation coefficients of compacted bentonite.

5.2.4.5. Comparison of numerical and analytical solutions

Three types of experiments have been interpreted in this section. The results are summarized in Table 5.2.15, in which the analytical solutions obtained by CIEMAT (Huertas et al., 2000) are compared to these derived from numerical interpretation.

The comparison leads to the following conclusions:

1) The analytical method tends to underestimate diffusion coefficients because it does not account for the role of the sinters. This is most clearly shown in HTO through-diffusion experiments which have a duration shorter than the rest of experiments

2) The estimates of D_a and K_d for cesium and selenium obtained by both methods with in-diffusion experiment data are generally in the same range. The experiments have so long duration that the influence of the sinters to the estimated results is not significant

3) There are siglificant differences in the estimates of D_a and K_d for strontium in through-diffusion experiments. The analytical method provides K_d values which are double of those obtained numerically while the estimates of D_a are three times larger than the numerical estimates. These differences are attributed to the large uncertainty of parameter estimates derived from these experiments. Thinner clay samples and more frequent sampling would have provided more accurate parameter estimates.

Table 5.2.15. Summary of interpretation of CIEMAT diffusion experiments performed on samples of compacted FEBEX bentonite. Analytical results were obtained by CIEMAT (Huertas et al., 2000) (ϕ denotes diffusion-accesible porosity; D_e is effective diffusion coefficient; D_a is apparent diffusion coefficient and K_d is distribution coefficient in ml/g). * means that the parameter is fixed and therefore not estimated.

Species	Test number	Numerical				Analytical		
		ϕ	D_e (m^2/s)	D_a (m^2/s)	K_d (ml/g)	D_e (m^2/s)	D_a (m^2/s)	K_d (ml/g)
HTO	TD-1	0.56*	$1.41\cdot10^{-10}$	$2.52\cdot10^{-10}$		$1.03\cdot10^{-10}$		
		0.5	$1.41\cdot10^{-10}$	$2.82\cdot10^{-10}$				
	TD-2	0.56*	$1.24\cdot10^{-10}$	$2.21\cdot10^{-10}$		$8.89\cdot10^{-11}$		
		0.45	$1.24\cdot10^{-10}$	$2.74\cdot10^{-10}$				
	TD-3	0.56*	$1.18\cdot10^{-10}$	$2.1\cdot10^{-10}$		$9.80\cdot10^{-11}$		
		0.44	$1.17\cdot10^{-10}$	$2.67\cdot10^{-10}$				
	P-1d	0.44/0.19	$3.5\cdot10^{-9}$					
	P-2d	0.6 /0.2	$2.33\cdot10^{-9}$					
Strontium	TD-Sr5	0.56*	$1.12\cdot10^{-9}$	$1.51\cdot10^{-12}$	626.0		$(4.8\pm0.81)\cdot10^{-12}$	1365±130
		0.62	$1.11\cdot10^{-9}$	$1.51\cdot10^{-12}$	621.3			
	TD-Sr6	0.56*	$7.57\cdot10^{-10}$	$1.15\cdot10^{-12}$	557.4			
		0.64	$7.61\cdot10^{-10}$	$1.17\cdot10^{-12}$	549.9			
	TD-Sr7	0.56*	$1.74\cdot10^{-9}$	$1.46\cdot10^{-12}$	975.4			
		0.52	$1.71\cdot10^{-9}$	$1.46\cdot10^{-12}$	995.1			
	TD-Sr8	0.56*	$8.5\cdot10^{-10}$	$1.21\cdot10^{-12}$	594.0			
		0.63	$8.35\cdot10^{-10}$	$1.16\cdot10^{-12}$	609.1			
Cesium	ID-Cs-9	0.42*	$8.03\cdot10^{-10}$	$3.34\cdot10^{-13}$	925.1		$3.07\cdot10^{-13}$	823±121
		0.49	$8.05\cdot10^{-10}$	$3.32\cdot10^{-13}$	931.8			
	ID-Cs-10A	0.42*	$3.3\cdot10^{-10}$	$1.37\cdot10^{-13}$	924.1		$2.51\cdot10^{-13}$	
		0.49	$3.44\cdot10^{-10}$	$1.43\cdot10^{-13}$	924.4			
	ID-Cs-10B	0.42*	$7.57\cdot10^{-10}$	$3.21\cdot10^{-13}$	907.9		$3.51\cdot10^{-13}$	
		0.45	$8.81\cdot10^{-10}$	$3.39\cdot10^{-13}$	999.6			
Selenium	ID-Se-17	0.42*	$4.35\cdot10^{-13}$	$6.51\cdot10^{-14}$	2.41		$(6.25\pm2.9)\cdot10^{-14}$	3.35±1.5
		0.21	$1.73\cdot10^{-13}$	$3.4\cdot10^{-14}$	1.87			
	ID-Se-18	0.42*	$1.39\cdot10^{-13}$	$2.29\cdot10^{-14}$	2.17			
		0.25	$1.23\cdot10^{-13}$	$2.23\cdot10^{-14}$	2.02			

5.3. INTERPRETATION OF PERMEATION EXPERIMENTS ON GRANITE

5.3.1 PERMEATION EXPERIMENTS

Two permeation experiments (P-1G and P-2G) were carried out by CIEMAT on granite columns of 0.15 m length and 0.045 m diameter. There is a fracture through the center of each sample. It is the major pathway for solute transport. A nearly constant flux (steady state) is introduced to the sample with pressure and flow regulation. At the initial phase, a small volume of tritiated water is injected into the fluid pathway and the concentration is measured continuously at the outlet by a fraction collector. Since the fracture aperture is small enough, one can assume that the flow in the fracture is linear and satisfies Darcy's Law.

A 1-D numerical grid with 200 triangular elements and 202 nodes, was used for simulating HTO transport through granite samples.

5.3.2. PARAMETER ESTIMATION

The inverse code has been used to estimate molecular diffusion, dispersivity and porosity. Prior information as well as upper and lower bounds of transport parameters derived from published results (Samper et al., 1998 and Yu et al., 1996) are presented in Table 5.3.1.

Table 5.3.1. Prior information and upper and lower bounds of the parameters

Parameter	Diffusion (D_0, m^2/s)	Dispersivity (m)	Porosity (ϕ)
Lower bound	$2.5 \cdot 10^{-11}$	0.005	0.001
Upper bound	$2.5 \cdot 10^{-8}$	0.02	0.05
Initial guess	$8.1 \cdot 10^{-10}$	0.01	0.01

These permeation experiments with granite samples were performed in a relatively short time, P-1G 43 minutes and P-2G 165 minutes. Since advective transport dominates, the diffusion coefficient (D_0) was fixed to $2.558 \cdot 10^{-9}$ m^2/s after several trial runs. Then, the inverse model only estimated porosity and dispersivity. The final results are plotted in Figures 5.3.1 and 5.3.2 and listed in Table 5.3.2.

Table 5.3.2. Estimation results for permeation experiments in granite samples

Test	Porosity	Dispersivity (m)	Objective
P-1G	0.00582	0.01809	0.2095
P-2G	0.00655	0.01380	186.38

Tracer activities in Test P-2G are much greater than those of P-1G. This is the reason that the optimum objective function of P-2G is much greater. Estimated porosity and dispersivity values are consistent with published values (Yu et al., 1996). The successful interpretation of advective-dominated permeation experiments in fracture granite indicates that the inverse model can be used to estimate flow and transport parameters for different materials.

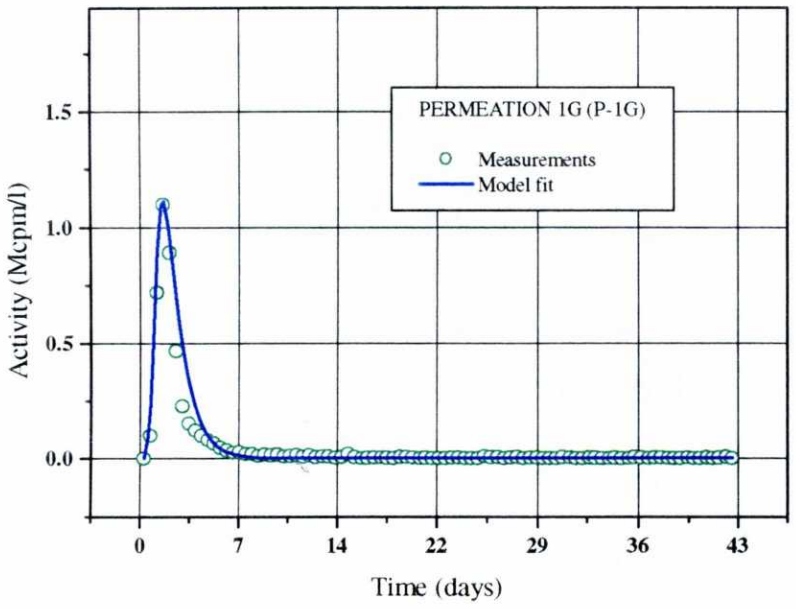


Figure 5.3.1. Best fit of HTO breakthrough curve in permeation experiment P-1G on granite

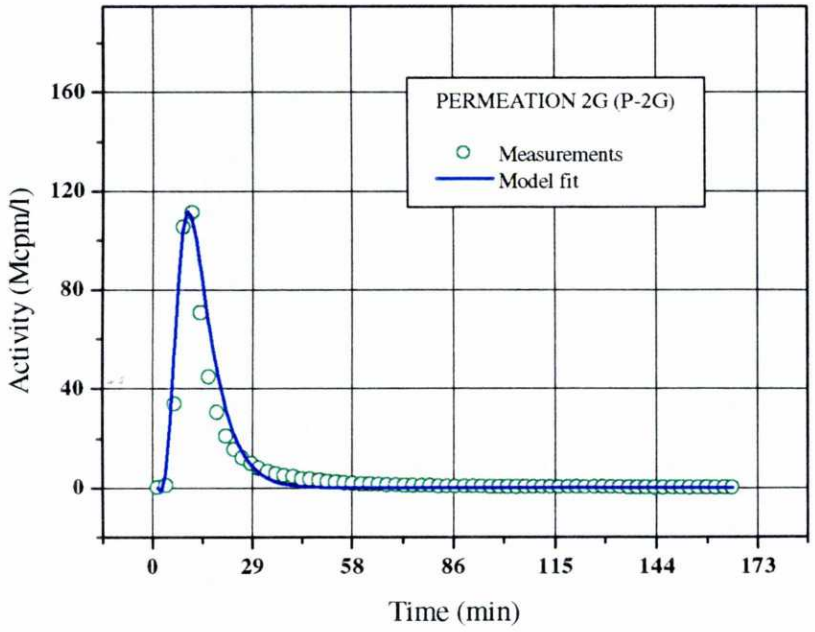


Figure 5.3.2. Best fit of HTO breakthrough curve in permeation experiment P-1G on granite

5.4. INTERPRETATION OF COLUMN EXPERIMENT FOR FLUSHING A CATION EXCHANGE COMPLEX WITH $SrCL_2$

5.4.1. COLUMN EXPERMENT

This real case corresponds to a laboratory column experiment published by Appelo et al. (1990). According to their description, a laboratory column was constructed as a receptacle for the sediment core. The column is made of stainless steel internally lined with teflon. Teflon base and top plates with a spiral groove ensure an even distribution of inlet and outlet water. The groove is covered with a stainless steel filter (double, of 160 μm mesh) which prevents the sediment from entering and blocking the groove. The column is closed with a piston of stainless steel to which the top plate of teflon is fixed.

The sediment is clay with relatively high cation exchange capacity (CEC) from 'Ketelmeer', deposited in the brackish to salty former inland sea, the 'Zuiderzee' in Netherlands. At present, there is a lake which receives heavily polluted mud from the Rhine River. This sediment has been eluted with $SrCl_2$ to obtain information about the mobility of adsorbed heavy metals. The displacing solutions ($SrCl_2$) are pumped into the column with a HPLC pump (LKB 2150) with double pump-head to ensure a smooth and continuous flow. The column core is 5.8 cm in diameter and 7.5 cm in length. The sediment pore water is mainly Rhine River water. The large CEC then made necessary a long elution time with a total of 2.5L of 0.05N $SrCl_2$ solution.

5.4.2. FORMULATION OF THE MODEL

Exchangeable cations are routinely determined by removing the ions, such as Na, K, Ca, Mg and NH_4, with a cation, such as Sr, which is not present in the soil solution. This case illustrates the chromatographic separation of Na, K, Ca, Mg and NH_4. Na is weakly

adsorbed, and is eluted first. Other cations are more tenaciously held than Na, and are retarded in the column effluent.

Gaines-Thomas convention was used for cation exchange. In this convention, selectivity coefficients are calculated by using equivalent fractions of exchanged cations as their activities. It should be pointed out that the cation selectivity is a relative concept. Here we choose Na^+ as the reference cation. Therefore, Na^+ selectivity is equal to 1. Cation exchange reactions can be written as:

$$Na^+ + K\text{-}X = K^+ + Na\text{-}X$$
$$Na^+ + 0.5Ca\text{-}X = 0.5Ca^{2+} + Na\text{-}X$$
$$Na^+ + 0.5Mg\text{-}X_2 = 0.5Mg^{2+} + Na\text{-}X$$
$$Na^+ + NH_4\text{-}X = NH_4^+ + Na\text{-}X$$
$$Na^+ + 0.5Sr\text{-}X_2 = 0.5Sr^{2+} + Na\text{-}X$$

The selectivity coefficients of the cation exchange in above reaction equations, such as $K_{Na/K}$, $K_{Na/Ca}$, $K_{Na/Mg}$, $K_{Na/NH4}$, and $K_{Na/Sr}$, are defined with respect to sodium, respectively. The lower the selectivity, the higher the exchange capacity. A divalent cation, in general, is more exchangeable than a monovalent cation.

The finite element grid has 42 nodes and 40 elements and one parameter zone. Flow and transport parameters are given by Appello et al. (1990). They include: hydraulic conductivity = $3 \cdot 10^{-4}$ m/d, porosity = 0.3, longitudinal dispersivity α_L = 0.0098 m, flow velocity = 22 m/year and cation exchange capacity CEC = 10.1 meq/100g. Molecular diffusion can be neglected.

In this case, computed breakthrough curves at node 41 (x=7.5 cm) are compared to measured data provided by Appelo et al. (1990). Concentrations of Na^+, K^+, Ca^{2+}, Mg^{2+}, Sr^{2+}, NH_4^+, and Cl^- were measured at times ranging from 0.143 to 17.004 porewater volumes. Pore volumes were transformed to actual times according to the input data file (p10.48 to p10.50, Appelö and Postma, 1992). These data were used for the inverse model. The time domain of 2.8 days was discretized into 64 time steps. Cl^- is also used as a

conservative solute to display conservative transport during the initial runs. Initial and boundary total dissolved concentrations of chemical components are listed in Table 5.4.1. Appelo et al. (1990) reported only the boundary concentrations of Sr^{2+}, H^+ and Cl^-. Since the numerical model requires non-zero concentrations, a very low concentration of 10^{-8} mol/l was used for the rest of species.

Table 5.4.1. Initial and boundary total dissolved concentrations (mol/l) of components in Ketelmeer sediment (Appelo et al. 1990). (*) indicates that these values were not measured and a small value of 10^{-8} is assumed. σ is the standard deviation of measured concentrations.

Component	Initial	Boundary	σ	Weight
pH	6.3	7	/	/
Na^+	3.2510^{-3}	10^{-8*}	$1.81\cdot10^{-4}$	5513
K^+	$3.94\cdot10^{-4}$	10^{-8*}	$3.52\cdot10^{-5}$	28420
Ca^{+2}	$1.12\cdot10^{-3}$	10^{-8*}	$4.35\cdot10^{-4}$	2299
Mg^{+2}	$1.24\cdot10^{-3}$	10^{-8*}	$3.62\cdot10^{-4}$	2764
NH_4^+	$1.02\cdot10^{-3}$	10^{-8*}	$8.7\cdot10^{-5}$	11490
Sr^{+2}	$3.3\cdot10^{-5}$	$2.5\cdot10^{-2}$	$1.15\cdot10^{-3}$	872
Cl^-	$4.54\cdot10^{-3}$	$5.0\cdot10^{-2}$	$1.64\cdot10^{-3}$	608
HCO_3^-	$4.74\cdot10^{-3}$	10^{-8*}	/	/
SO_4^{-2}	$7\cdot10^{-5}$	10^{-8*}	/	/

σ (in Table 5.4.1) is the standard deviation of measured concentrations. The weighting coefficients for each component were calculated from Equation (3.2.3).

In order to interpret the experiment and estimate the relevant parameters, the following step-wise procedure was used:

1. Case 0: Estimate dispersivity using Cl^- data. Its estimated value was fixed in the following stages.
2. Case 1: Estimate selectivity coefficients using Na^+, K^+, Ca^{+2}, Mg^{+2}, NH_4^+, Sr^{+2} and Cl^- data. Only cation exchange is considered and CEC is fixed to be 10.1 meq/100g.

3. Case 2: Estimate selectivity coefficients and CEC, other conditions being the same as in Case 1.
4. Case 3: Same as Case 2, but adding calcite dissolution/precipitation and aqueous complexation in the geochemical system.
5. Case 4: Same as Case 2, but adding H^+ exchange reaction and estimating the selectivity coefficient $K_{Na/H}$.
6. Case 5: Same as Case 3, but adding H^+ exchange reaction and estimating the selectivity coefficient $K_{Na/H}$. The sensitivity of this parameter is analyzed.
7. Case 6: Estimate selectivity coefficients, CEC and initial concentrations of some cations considering only cation exchange (on the basis of Case 2).
8. Case 7: Same as Case 6, plus the estimation of some boundary concentrations.
9. Case 8: Same as Case 6, but adding kinetically-controlled calcite dissolution/ precipitation and aqueous complexation in the geochemical system.

5.4.3. ESTIMATION OF DISPERSIVITY AND SELECTIVITY COEFFICIENTS

In the first stage (Case 0), only conservative Cl^- concentration data were considered for estimating dispersivity using INVERSE-CORE^{2D}. The estimated value (α_L = 0.0121 m) is slightly greater than the value (0.0098 m) used by Appelo et al. (1990). It should be noticed that the fit to Cl^- data is excellent (Figure 5.4.1). Then, dispersivity was fixed for the estimation of selectivity coefficients using all the concentration data and considering only cation exchange reactions (Case 1). Parameter prior information obtained from Appelo et al. (1990) and Appelo and Postma (1992) is listed in Table 5.4.2. It was used to define lower and upper bounds for the parameters to avoid the parameters taking extremely large or small values during the estimation process. Model parameters must lie within the intervals defined by the lower and upper bounds. If a parameter reaches a bound, the bound is slightly moved in order to improve the optimization process. Estimation results are presented in Table 5.4.3. The best fit to measured chemical data of components is shown in Figure 5.4.1.

Table 5.4.2. Parameter prior information data (obtained from Appelo et al., 1990 and Appelo and Postma, 1992)

Parameter	Lower bound	Upper bound	Prior information
$K_{Na/K}$	0.01	0.9	0.1170
$K_{Na/Ca}$	0.01	0.9	0.2415
$K_{Na/Mg}$	0.01	0.9	0.4046
$K_{Na/Sr}$	0.01	0.9	0.3162
$K_{Na/NH4}$	0.01	0.9	0.2323
CEC (meq/100g)	2.0	15.0	10.1

Table 5.4.3. Estimated results for selectivity coefficients (Case 1) considering only cation exchange reactions and a fixed CEC = 10.1meq/100g (objective function=$1.721 \cdot 10^{3}$)

Parameter	Initial guess	Estimated value	Estimation Variance	90% confidence interval	
				Lower bound	Upper bound
$K_{Na/K}$	0.1170	0.1082	$2.729 \cdot 10^{-6}$	0.1031	0.1132
$K_{Na/Ca}$	0.2415	0.0873	$3.946 \cdot 10^{-8}$	0.8667	0.8787
$K_{Na/Mg}$	0.4046	0.2280	$6.647 \cdot 10^{-7}$	0.2255	0.2305
$K_{Na/Sr}$	0.3162	0.2912	$5.246 \cdot 10^{-7}$	0.2890	0.2934
$K_{Na/NH4}$	0.2323	0.2125	$6.319 \cdot 10^{-6}$	0.2049	0.2202

The CEC has been fixed in a value of 10.1 meq/100g. Five selectivity coefficients have been estimated as well as the lower and upper bounds of 90% confidence intervals. Estimation variances are very low. However, the objective function, which measures the goodness of fit, is still very high. The fit to measured data in Figure 5.4.1A, B is not good, especially that of Ca^{+2}. It indicates that there may be more model parameters needed to be modified or other chemical processes not yet considered. Therefore, in the next stage cation exchange capacity (CEC) is also estimated.

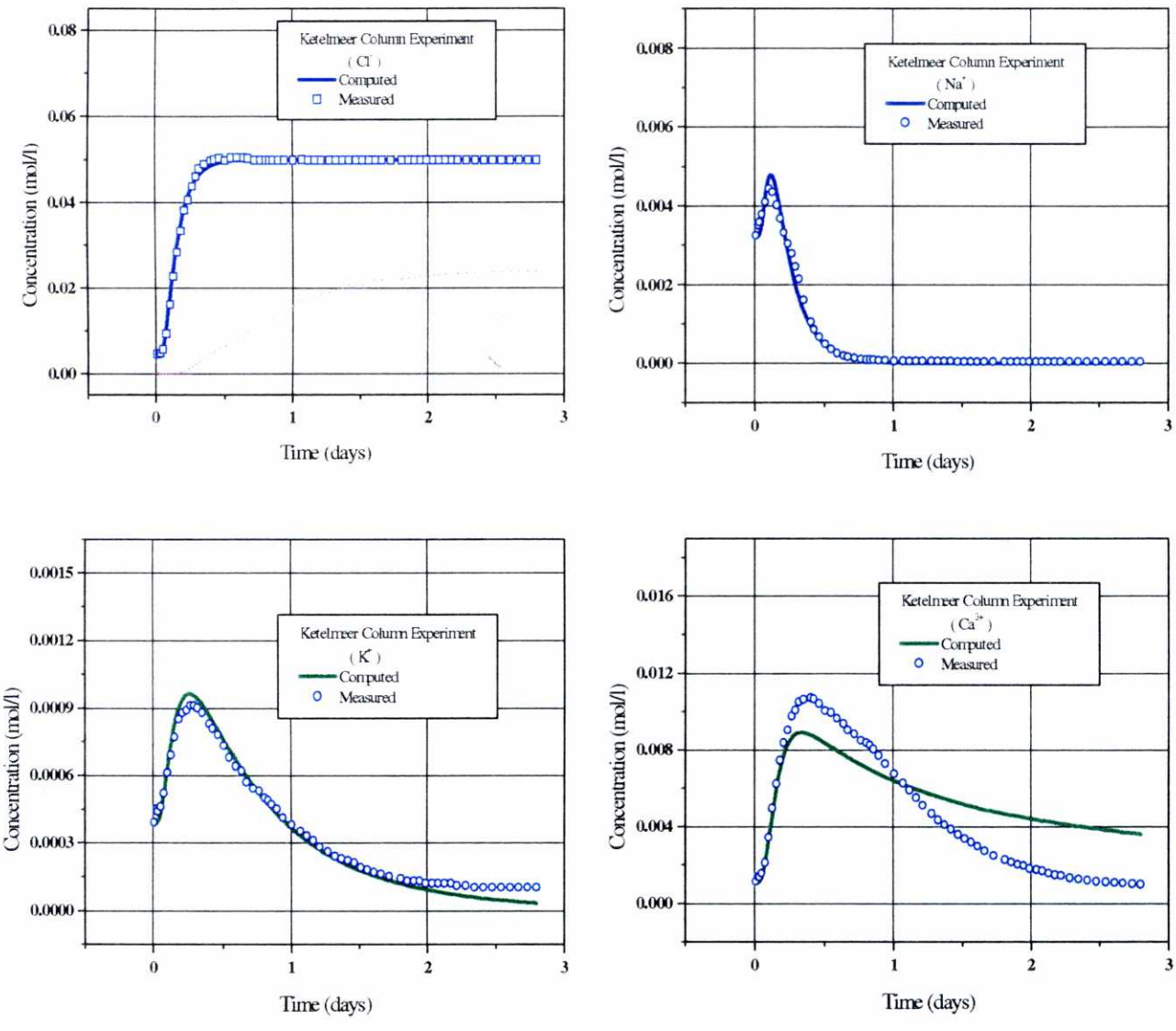


Figure 5.4.1A. Measured (symbols) and computed (lines) concentration breakthrough curves in Case 1 when only exchange reactions are considered. (five selectivity coefficients are estimated and CEC is fixed to 10.1 meq/100g).

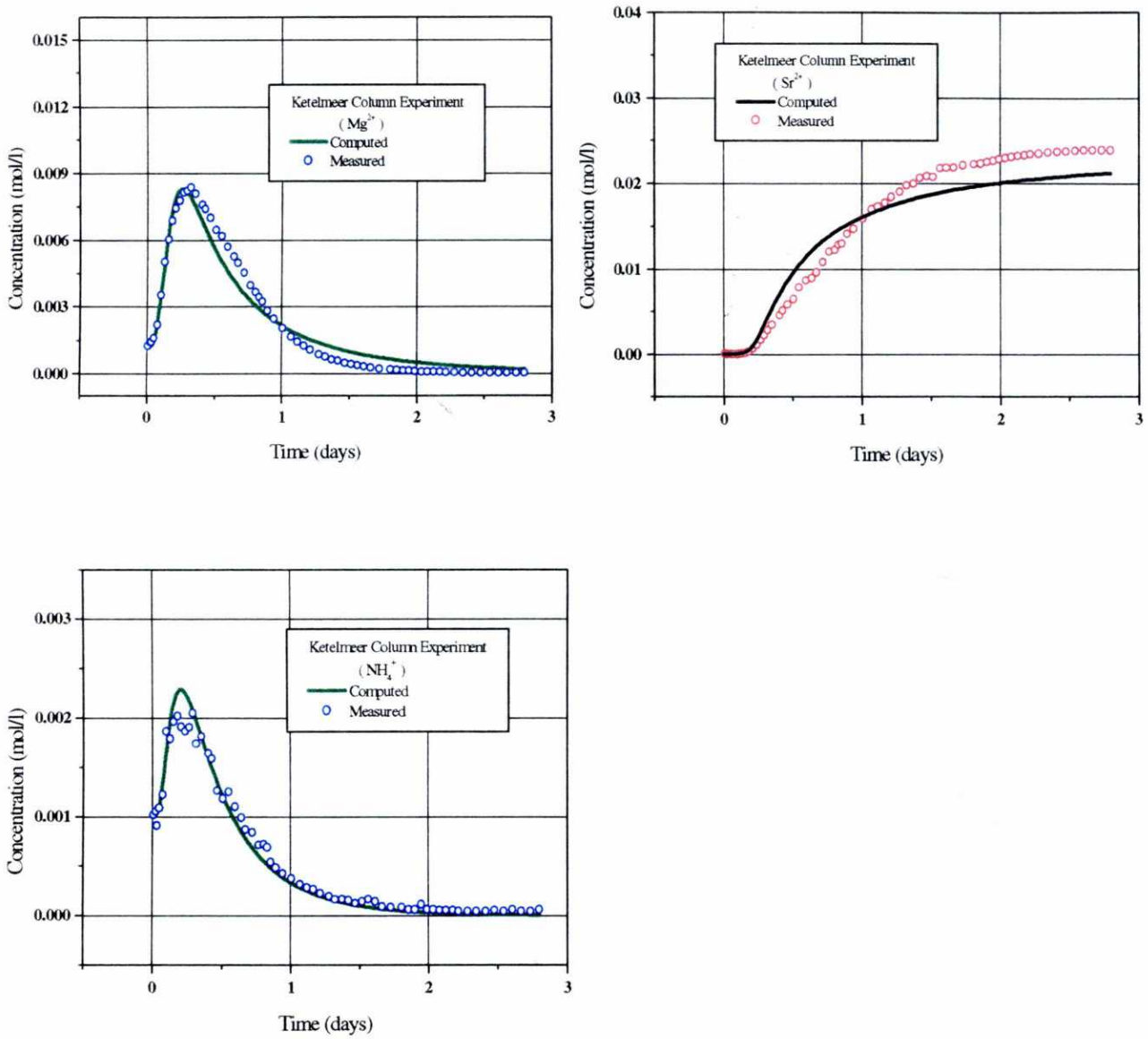


Figure 5.4.1B. Measured (symbols) and computed (lines) concentration breakthrough curves in Case 1 when only exchange reactions are considered. (five selectivity coefficients are estimated and CEC is fixed to 10.1 meq/100g).

5.4.4. ESTIMATION OF SELECTIVITY COEFFICIENTS AND CEC

Taking the estimates of the selectivity coefficients of the previous stage (see Table 5.4.3) as the initial parameter values, these coefficients as well as CEC have been estimated (Case 2). Estimation results are listed in Table 5.4.4. The best fit for different components are plotted in Figure 5.4.2.

Table 5.4.4. Estimation results for Case 2 (estimation of selectivity coefficients and CEC considering only cation exchange reactions, objective function = $9.397 \cdot 10^2$)

Parameter	Initial guess	Estimated value	Estimation Variance	90% confidence interval	
				Lower bound	Upper bound
$K_{Na/K}$	0.1082	0.1241	$1.583 \cdot 10^{-6}$	0.12	0.1282
$K_{Na/Ca}$	0.0873	0.236	$2.203 \cdot 10^{-7}$	0.2344	0.2375
$K_{Na/Mg}$	0.228	0.3762	$7.175 \cdot 10^{-7}$	0.3735	0.379
$K_{Na/Sr}$	0.2912	0.334	$6.409 \cdot 10^{-7}$	0.3314	0.3366
$K_{Na/NH4}$	0.2125	0.2312	$3.618 \cdot 10^{-6}$	0.225	0.2374
CEC	10.1	4.262	$5.496 \cdot 10^{-5}$	4.238	4.286

Compared to Case 1 (Figure 5.4.1), the fit to measured data in Figure 5.4.2 improves greatly, especially for Ca^{2+} and Sr^{2+}. In fact, the objective function reduces almost to a half of that in Case 1. The Na-Ca selectivity coefficient ($K_{Na/Ca}$) increases from 0.0873 in Case 1 to 0.2360 in Case 2. The latter seems more reasonable and is closer to the prior information value. On the other hand, the estimate of CEC is about half of the value used by Appelo et al.(1990). After discussing this result with Dr. Appelo, he sent us more information about CEC which originally was 256.4 meq/ per liter of pore water. According to Equation (4.2.1), this value amounts to a CEC of 4.15 meq/100g solid. So, the additional information indicates that our estimated value of CEC is reasonable. Afterwards, the prior information value of the CEC was changed from 10.1 to 4.15 meq/100g solid.

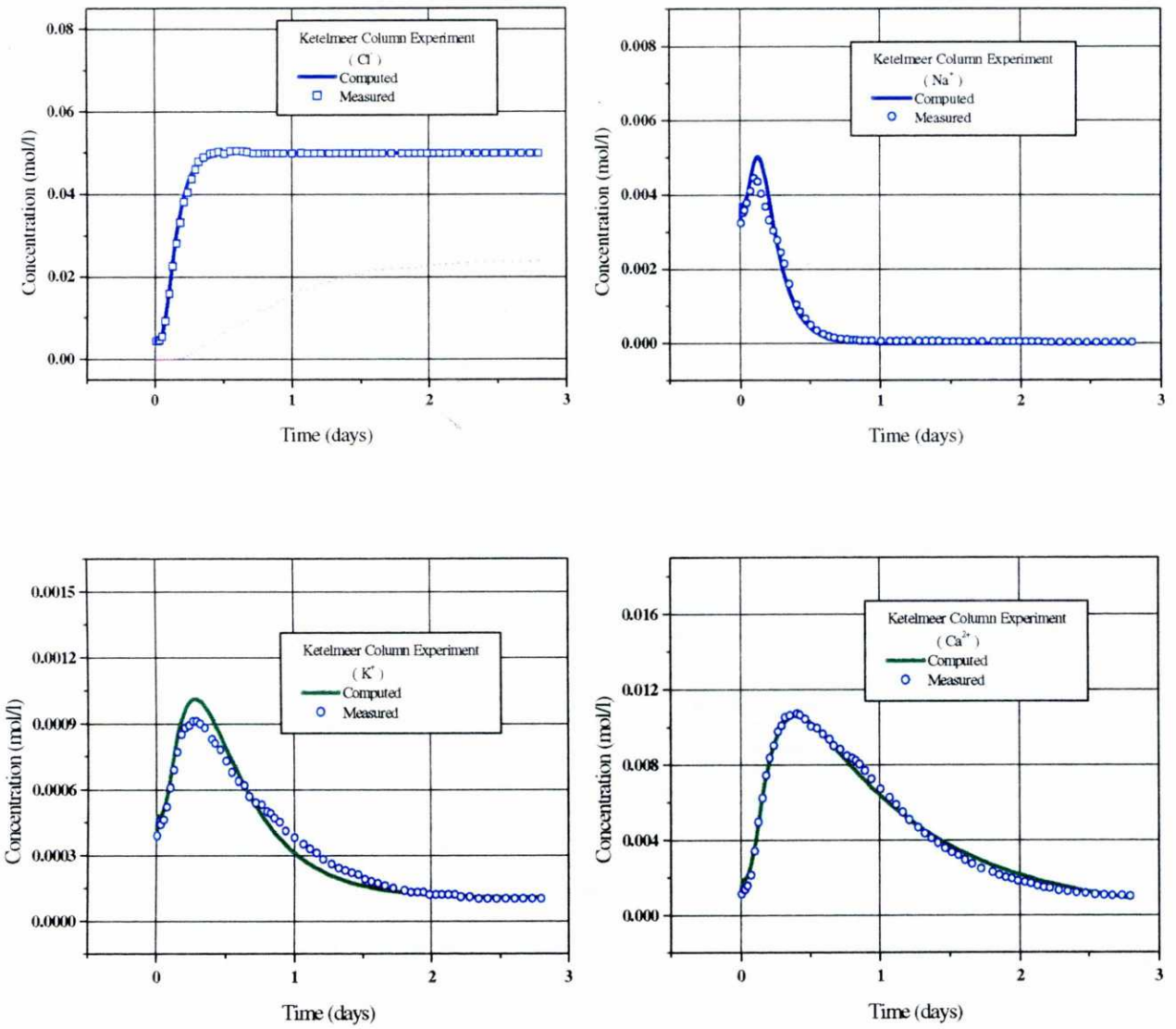


Figure 5.4.2A. Measured (symbols) and computed (lines) concentration breakthrough curves when only exchange reactions are considered. (Case 2: Five selectivity coefficients and CEC are estimated)

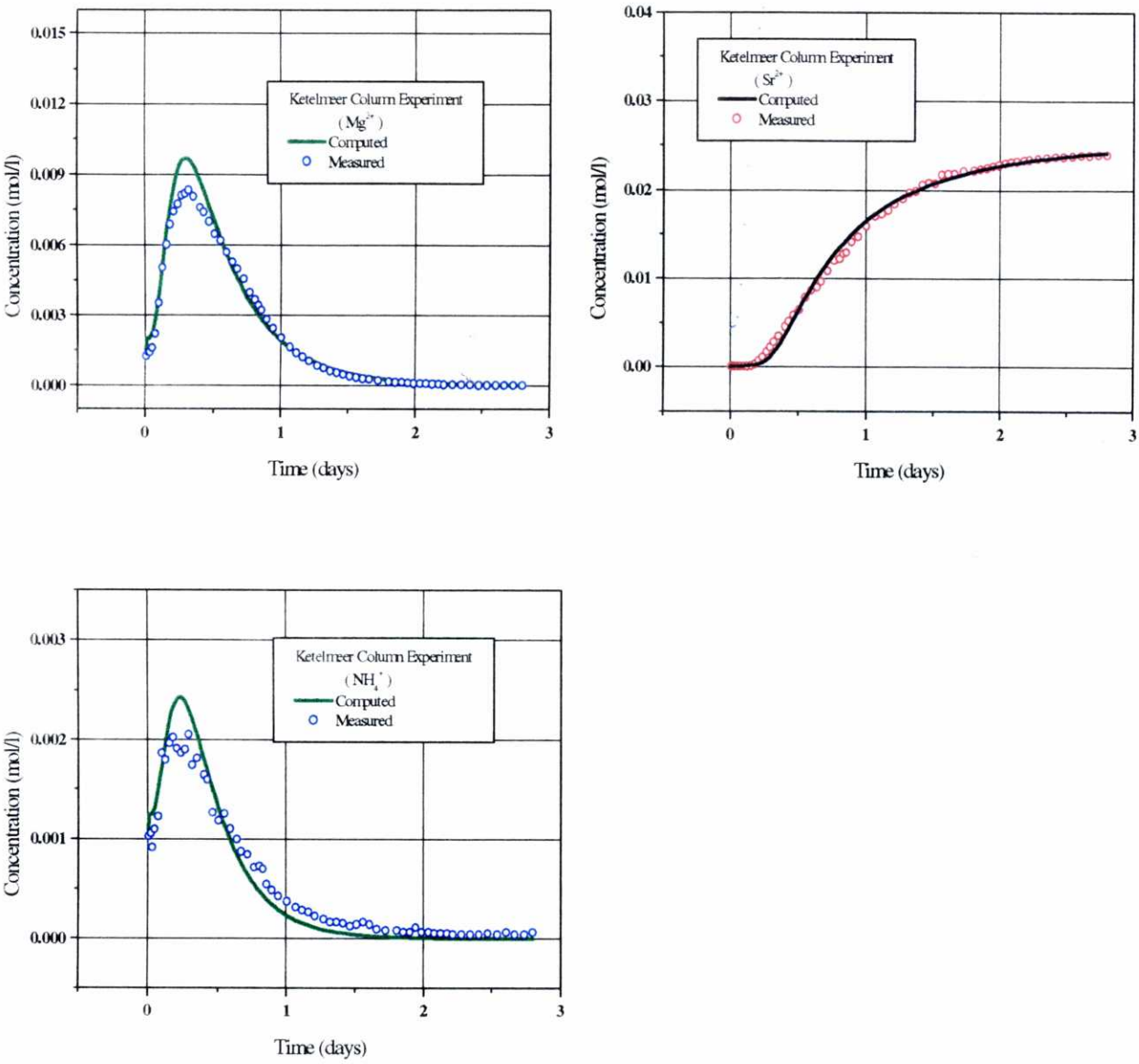


Figure 5.4.2B. Measured (symbols) and computed (lines) concentration breakthrough curves when only exchange reactions are considered. (Case 2: Five selectivity coefficients and CEC are estimated)

5.4.5. CASE 3: ESTIMATION OF SELECTIVITY COEFFICIENTS AND CEC CONSIDERING CALCITE DISSOLUTION AND AQUEOUS COMPLEXATION

The results in Figure 5.4.2 improve due to the estimation of cation exchange capacity (CEC). However, there is still some space to improve the results by considering additional chemical reactions in the geochemical system. In Case 3, aqueous dissociation reactions and calcite dissolution/precipitation are added to the geochemical system. The following 11 species are taken as the primary aqueous species: H_2O, Na^+, K^+, Ca^{2+}, Mg^{2+}, Sr^{2+}, NH_4^+, Cl^-, HCO_3^-, SO_4^{-2} and H^+. By using EQ3 the initial geochemical system has been modeled, the following 16 secondary species are selected. Aqueous dissociation reactions as well as calcite precipitation-dissolution reaction are listed in Table 5.4.5.

Table 5.4.5. List of additional chemical reactions considering in Case 3

Geochemical reactions	$Log_{10}(K)$ at 25 °C
Aqueous dissociation:	
$OH^- = H_2O - H^+$	13.995
$CO_3^{2-} = HCO_3^- - H^+$	10.329
$CO_2(aq) = HCO_3^- + H^+ - H_2O$	-6.3447
$CaHCO_3^+ = Ca^{2+} + HCO_3^-$	-1.0467
$MgHCO_3^+ = Mg^{2+} + HCO_3^-$	-1.0357
$CaCO_3(aq) = Ca^{2+} + HCO_3^- - H^+$	7.0017
$MgCO_3(aq) = Mg^{2+} + HCO_3^- - H^+$	7.3499
$NaHCO_3(aq) = Na^+ + HCO_3^-$	-0.1541
$CaSO_4(aq) = Ca^{2+} + SO_4^{2-}$	-2.1111
$MgSO_4(aq) = Mg^{2+} + SO_4^{2-}$	-2.309
$NaSO_4^- = Na^+ + SO_4^{2-}$	-0.082
$KSO_4^- = K^+ + SO_4^{2-}$	-0.8796
$HSO_4^- = SO_4^- + H^+$	-1.9791
$NaCO_3^-(aq) = Na^+ + HCO_3^- - H^+$	9.8144
$SrCO_3(aq) = Sr^{2+} + HCO_3^- - H^+$	7.4635
$SrSO_4(aq) = Sr^{2+} + SO_4^{2-}$	-2.3
Mineral dissolution-precipitation:	
Calcite $= Ca^{2+} + HCO_3^- - H^+$	1.8487

Estimation results are listed in Table 5.4.6. When more chemical reactions are included in the inverse model, the fit is worse than that in Case 2. The objective function increases from $9.397 \cdot 10^2$ to $1.028 \cdot 10^3$. This indicates that the chemical system may not be in equilibrium with calcite.

Table 5.4.6. Estimation results for Case 3: Estimation of selectivity coefficients and CEC considering calcite dissolution / precipitation and aqueous dissociation reactions (objective function=$1.028 \cdot 10^3$)

Parameter	Initial guess	Estimated value	Estimation Variance	90% confidence interval	
				Lower bound	Upper bound
$K_{Na/K}$	0.1241	0.1283	$2.941 \cdot 10^{-5}$	0.1106	0.1460
$K_{Na/Ca}$	0.2360	0.2325	$4.473 \cdot 10^{-5}$	0.2106	0.2543
$K_{Na/Mg}$	0.3762	0.3750	$1.128 \cdot 10^{-4}$	0.3404	0.4097
$K_{Na/Sr}$	0.3340	0.3407	$8.770 \cdot 10^{-5}$	0.3101	0.3712
$K_{Na/NH4}$	0.2312	0.2371	$7.862 \cdot 10^{-5}$	0.2082	0.2661
CEC	4.262	4.594	$1.085 \cdot 10^{-3}$	4.486	4.701

In Case 4, H^+ exchange reaction is also considered. This reaction is treated as an ion exchange reaction:

$$Na^+ + H\text{-}X = H^+ + Na\text{-}X$$

Calcite dissolution/precipitation and aqueous complexation chemical reactions are taken out from the geochemical system. Six selectivity coefficients and CEC have been estimated and the results are listed in Table 5.4.7.

Table 5.4.7. Estimation results for Case 4: Estimation of six selectivity coefficients and CEC and including H^+ exchange reaction (objective function=$9.411 \cdot 10^2$)

Parameter	Initial guess	Estimated value	Estimation Variance	90% confidence interval	
				Lower bound	Upper bound
$K_{Na/K}$	0.1283	0.1264	$1.654 \cdot 10^{-6}$	0.1220	0.1309
$K_{Na/Ca}$	0.2325	0.2399	$2.288 \cdot 10^{-7}$	0.2382	0.2416
$K_{Na/Mg}$	0.3750	0.3826	$7.468 \cdot 10^{-7}$	0.3796	0.3856
$K_{Na/Sr}$	0.3407	0.3402	$6.677 \cdot 10^{-7}$	0.3374	0.3431
$K_{Na/NH4}$	0.2371	0.2362	$3.802 \cdot 10^{-6}$	0.2294	0.2429
$K_{Na/H}$	0.1000	0.0998	$2.324 \cdot 10^{-4}$	0.0469	0.1527
CEC	4.594	4.257	$5.411 \cdot 10^{-5}$	4.231	4.283

By adding H^+ exchange, the objective function increases slightly from $9.397 \cdot 10^2$ in Case 2 to $9.411 \cdot 10^2$ in Case 4. This means that proton exchange can be neglected in the geochemical system.

However, since proton exchange may affect pH which in turn can influence calcite dissolution/precipitation and other aqueous chemical reactions, in Case 5, all these chemical reactions are included in the inverse model. Six selectivity coefficients and CEC are estimated. The results are listed in Table 5.4.8.

Results of Case 5 are worse than those of Case 2 and Case 4 and slightly better than those of Case 3. The estimation variance and eigenvalue of $K_{Na/H}$ are the largest of all seven parameters. Its 90% confidence interval is very wide, meaning that this parameter is highly uncertain mostly because the objective function is not sensitive to it. Including proton exchange, calcite dissolution/precipitation and other aqueous chemical reactions does not help to get a better estimation. Therefore, these chemical reactions can be neglected.

Table 5.4.8. Estimation results for Case 5: Estimation of six selectivity coefficients and CEC (meq/100g) considering cation exchange, calcite dissolution / precipitation, and aqueous complexation (objective function=$9.738 \cdot 10^{2}$)

Parameter	Initial guess	Estimated value	Estimation Variance	90% confidence interval	
				Lower	Upper
$K_{Na/K}$	0.1264	0.1295	$1.104 \cdot 10^{-5}$	0.118	0.141
$K_{Na/Ca}$	0.2399	0.2368	$1.715 \cdot 10^{-6}$	0.2323	0.2413
$K_{Na/Mg}$	0.3826	0.3826	$5.635 \cdot 10^{-6}$	0.3744	0.3908
$K_{Na/Sr}$	0.3402	0.3635	$5.767 \cdot 10^{-6}$	0.3551	0.3718
$K_{Na/NH4}$	0.2362	0.2435	$2.565 \cdot 10^{-5}$	0.226	0.2611
$K_{Na/H}$	0.0998	0.1158	$1.379 \cdot 10^{-2}$	0	0.5228
CEC	4.257	4.539	$5.01 \cdot 10^{-4}$	4.462	4.617

5.4.6. ESTIMATION OF INITIAL AND BOUNDARY CONCENTRATIONS, SELECTIVITY COEFFICIENTS AND CEC

In later stages, the initial concentrations of some main cations were also estimated. Taking the estimated values in Table 5.4.4 corresponding to Case 2 as initial values, five selectivity coefficients, CEC, and four initial concentrations were estimated. Estimation results are listed in Table 5.4.9. The best fit for different components is plotted in Figure 5.4.3.

Table 5.4.9. Estimation results for Case 6: Estimation of five selectivity coefficients, CEC (meq/100g) and initial concentrations (C_0, mol/l) only considering exchange reactions (objective function=$4.05 \cdot 10^2$)

Parameter	Initial guess	Estimated value	Estimation Variance	95% confidence interval		Associated eigenvalue
				Lower	Upper	
$K_{Na/K}$	0.1241	0.1204	$4.221 \cdot 10^{-7}$	0.1176	0.1232	$7.43 \cdot 10^{-6}$
$K_{Na/Ca}$	0.236	0.2838	$1.124 \cdot 10^{-6}$	0.2793	0.2884	$5.96 \cdot 10^{-6}$
$K_{Na/Mg}$	0.3762	0.4158	$2.38 \cdot 10^{-6}$	0.4092	0.4224	$4.54 \cdot 10^{-7}$
$K_{Na/Sr}$	0.334	0.3892	$2.064 \cdot 10^{-6}$	0.383	0.3953	$2.23 \cdot 10^{-7}$
$K_{Na/NH4}$	0.2312	0.2139	$1.105 \cdot 10^{-6}$	0.2094	0.2184	$8.12 \cdot 10^{-8}$
CEC	4.262	4.393	$7.081 \cdot 10^{-6}$	4.381	4.404	$2.78 \cdot 10^{-8}$
$C_{0\,Na}$	$3.25 \cdot 10^{-3}$	$3.07 \cdot 10^{-3}$	$3.819 \cdot 10^{-11}$	$3.043 \cdot 10^{-3}$	$3.096 \cdot 10^{-3}$	$1.84 \cdot 10^{-11}$
$C_{0\,K}$	$3.94 \cdot 10^{-4}$	$3.528 \cdot 10^{-4}$	$8.985 \cdot 10^{-13}$	$3.488 \cdot 10^{-4}$	$3.569 \cdot 10^{-4}$	$2.46 \cdot 10^{-12}$
$C_{0\,Ca}$	$1.12 \cdot 10^{-3}$	$1.353 \cdot 10^{-3}$	$3.345 \cdot 10^{-12}$	$1.346 \cdot 10^{-3}$	$1.361 \cdot 10^{-3}$	$1.6 \cdot 10^{-12}$
$C_{0\,NH4}$	$1.02 \cdot 10^{-3}$	$8.871 \cdot 10^{-4}$	$3.969 \cdot 10^{-12}$	$8.786 \cdot 10^{-4}$	$8.957 \cdot 10^{-4}$	$5.47 \cdot 10^{-13}$

By estimating the initial concentrations of the main cations, the objective function reduces drastically from $9.397 \cdot 10^2$ in Case 2 to $4.05 \cdot 10^2$. The fit to observed breakthrough curves shows also a significant improvement. Compared to measured values (the initial values in the Table 5.4.9), estimated initial concentrations of Na, K and NH_4 are slightly lower while that of Ca is higher. Estimates of selectivity coefficients and CEC also change slightly. Approximate 95% confidence intervals and the associated eigenvalues in Table 5.4.9 indicate that the estimated parameters have small uncertainty.

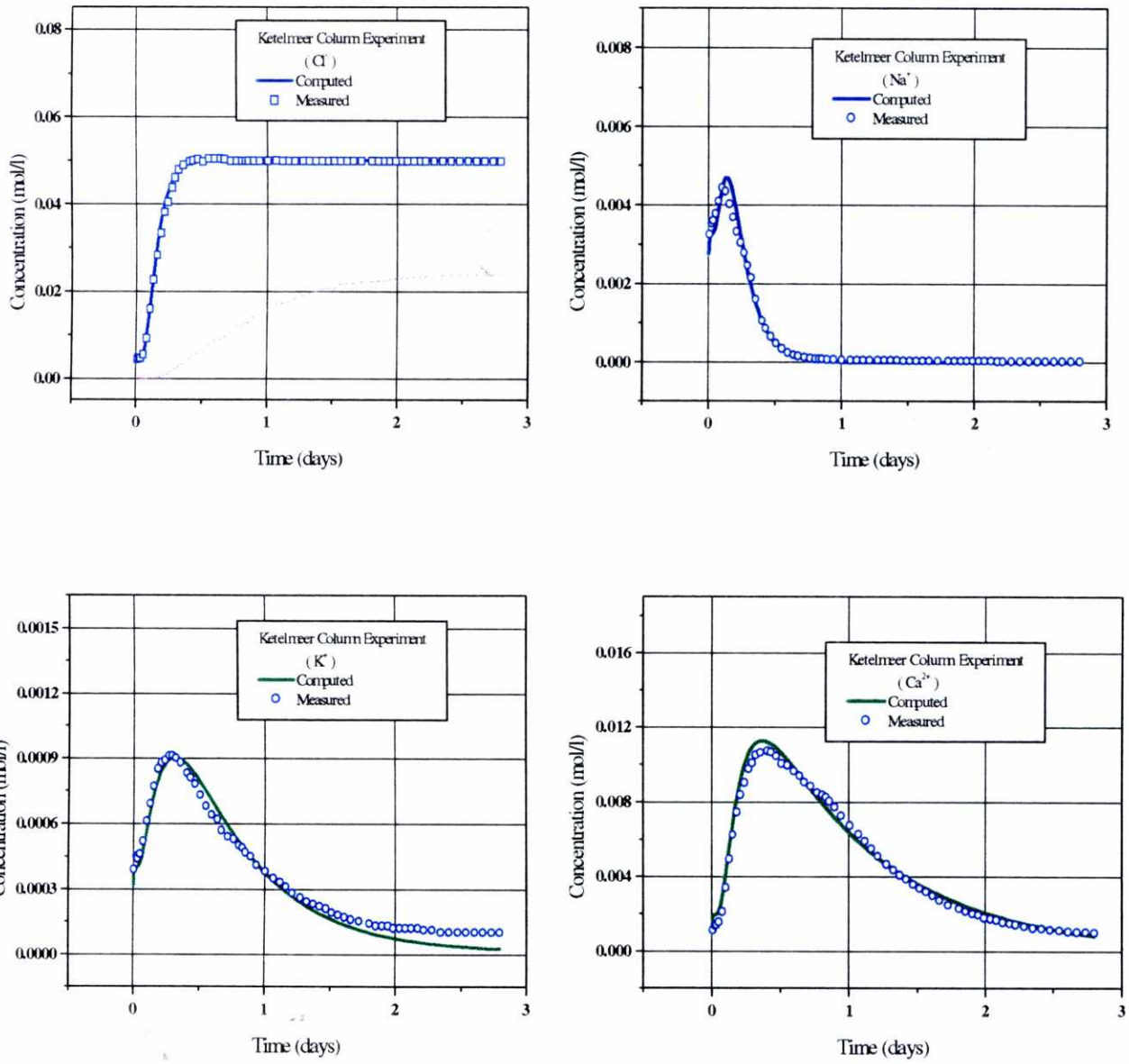


Figure 5.4.3A. Measured (symbols) and computed (lines) concentration breakthrough curves when only exchange reactions are considered. (Case 6: Estimation of five selectivity coefficients, CEC, and four initial concentrations)

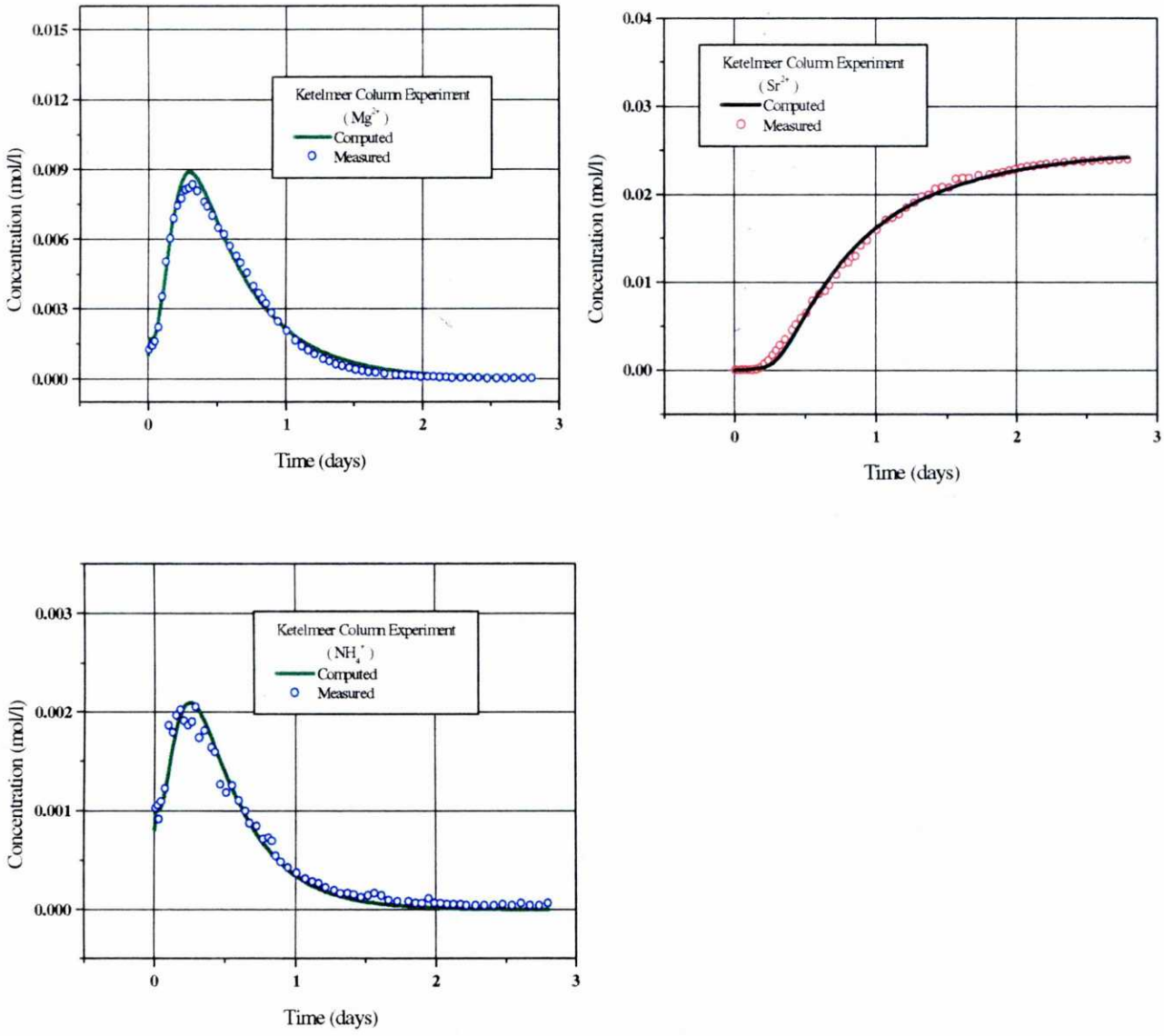


Figure 5.4.3B. Measured (symbols) and computed (lines) concentration breakthrough curves when only exchange reactions are considered. (Case 6: Estimation of five selectivity coefficients, CEC, and four initial concentrations)

Boundary concentrations of some components are assumed to be very low (10^{-8} mol/l) in Table 5.4.1. In order to test the influence of this 'assumed' value, Case 7 was designed to estimate the boundary concentrations of four components (Na, K, Ca, NH_4) based on Case 6. Boundary concentrations listed in Table 5.4.1 are taken as prior information values. However, these values have a large uncertainty, so, a very low weighting coefficient 10^{-3} was given to this information. Estimation results are listed in Table 5.4.10 and the best fit for different components is shown in Figure 5.4.4.

Both the fit to measured data and the objective function improve by estimating boundary concentrations (compare Figures 5.4.3 and 5.4.4). The estimated boundary concentration of K^+ is $9.065 \cdot 10^{-5}$ mol/l, a value much larger than 10^{-8} mol/l. It leads to a K^+ breakthrough curve which fits the data much better than that in Figure 5.4.3A. Table 5.4.11 presents the values of the objective function as well as its contributions of each chemical component for Cases 6, 7 and 8. It shows that in case 7 the contributions of K and NH_4 to the objective function decrease much more than those of other components. This is clearly seen in their fit improvement of breakthrough curves in Figure 5.4.4.

However, the fit to measured breakthrough curves should not be the only criterion used to identify parameters. Estimated parameters should be consistent with experimental conditions. Such a large boundary concentration for K^+ should be checked with the authors of the experiment. The results of the inverse problem indicate that the breakthrough curve of K^+ is consistent with a boundary K^+ concentration on the order of 10^{-4} mol/l. Whether this hypothesis is plausible or not should be tested by checking experimental conditions. Other hypotheses for the long tail of the K^+ breakthrough curve such as kinetic processes or double porosity behavior cannot be discarded.

The estimated 95% confidence interval of Ca boundary concentration in Table 5.4.10 has a wide range. Because the estimated Ca boundary concentration is very low (10^{-8} mol/l), the objective function is not sensitive to it. So, this estimated value has relatively high uncertainty.

Table 5.4.10. Estimation results for Case 7: Estimation of five selectivity coefficients, CEC (meq/100g), initial and boundary concentrations (C_0 and C_B, mol/l) considering only exchange reactions (objective function=$2.85 \cdot 10^2$)

Parameter	Initial guess	Estimated value	Estimation variance	95% confidence interval		Associated eigenvalue
				Lower	Upper	
$K_{Na/K}$	0.1204	0.1456	$7.299 \cdot 10^{-7}$	0.1414	0.1497	$5.58 \cdot 10^{-6}$
$K_{Na/Ca}$	0.2838	0.2866	$8.953 \cdot 10^{-7}$	0.282	0.2912	$4.53 \cdot 10^{-6}$
$K_{Na/Mg}$	0.4158	0.4179	$1.869 \cdot 10^{-6}$	0.4113	0.4246	$5.3 \cdot 10^{-7}$
$K_{Na/Sr}$	0.3892	0.3931	$1.635 \cdot 10^{-6}$	0.3869	0.3994	$4.53 \cdot 10^{-7}$
$K_{Na/NH4}$	0.2139	0.2228	$1.066 \cdot 10^{-6}$	0.2178	0.2278	$5.87 \cdot 10^{-8}$
CEC	4.393	4.376	$4.976 \cdot 10^{-6}$	4.366	4.387	$1.98 \cdot 10^{-8}$
$C_{0\,Na}$	$3.25 \cdot 10^{-3}$	$3.058 \cdot 10^{-3}$	$2.714 \cdot 10^{-11}$	$3.033 \cdot 10^{-3}$	$3.083 \cdot 10^{-3}$	$1.34 \cdot 10^{-11}$
$C_{0\,K}$	$3.94 \cdot 10^{-4}$	$3.673 \cdot 10^{-4}$	$7.93 \cdot 10^{-13}$	$3.629 \cdot 10^{-4}$	$3.716 \cdot 10^{-4}$	$4.81 \cdot 10^{-12}$
$C_{0\,Ca}$	$1.12 \cdot 10^{-3}$	$1.359 \cdot 10^{-3}$	$2.382 \cdot 10^{-12}$	$1.352 \cdot 10^{-3}$	$1.367 \cdot 10^{-3}$	$3.5 \cdot 10^{-12}$
$C_{0\,NH4}$	$1.02 \cdot 10^{-3}$	$8.955 \cdot 10^{-4}$	$3.02 \cdot 10^{-12}$	$8.870 \cdot 10^{-4}$	$9.039 \cdot 10^{-4}$	$1.83 \cdot 10^{-12}$
$C_{B\,Na}$	10^{-8}	$4.114 \cdot 10^{-5}$	$5.391 \cdot 10^{-12}$	$2.984 \cdot 10^{-5}$	$5.244 \cdot 10^{-5}$	$1.71 \cdot 10^{-12}$
$C_{B\,K}$	10^{-8}	$9.065 \cdot 10^{-5}$	$3.007 \cdot 10^{-12}$	$8.221 \cdot 10^{-5}$	$9.909 \cdot 10^{-5}$	$1.13 \cdot 10^{-12}$
$C_{B\,Ca}$	10^{-8}	10^{-8}	$2.792 \cdot 10^{-14}$	0.0	$8.232 \cdot 10^{-8}$	$4.32 \cdot 10^{-13}$
$C_{B\,NH4}$	10^{-8}	$4.8 \cdot 10^{-5}$	$4.488 \cdot 10^{-12}$	$3.769 \cdot 10^{-5}$	$5.831 \cdot 10^{-5}$	$2.63 \cdot 10^{-14}$

Table 5.4.11. Objective function and the contribution of each component for Cases 6, 7 and 8

Case	E(p)	E_{Cl}	E_{Na}	E_K	E_{Ca}	E_{Mg}	E_{Sr}	E_{NH_4}
6	405	10.1	64.2	143	37	36.9	11.7	102
7	285.4	10.1	61.5	37.5	39.4	36.9	12.1	88
8	407.5	10.1	62.1	136	46.2	43.1	11.1	98.9

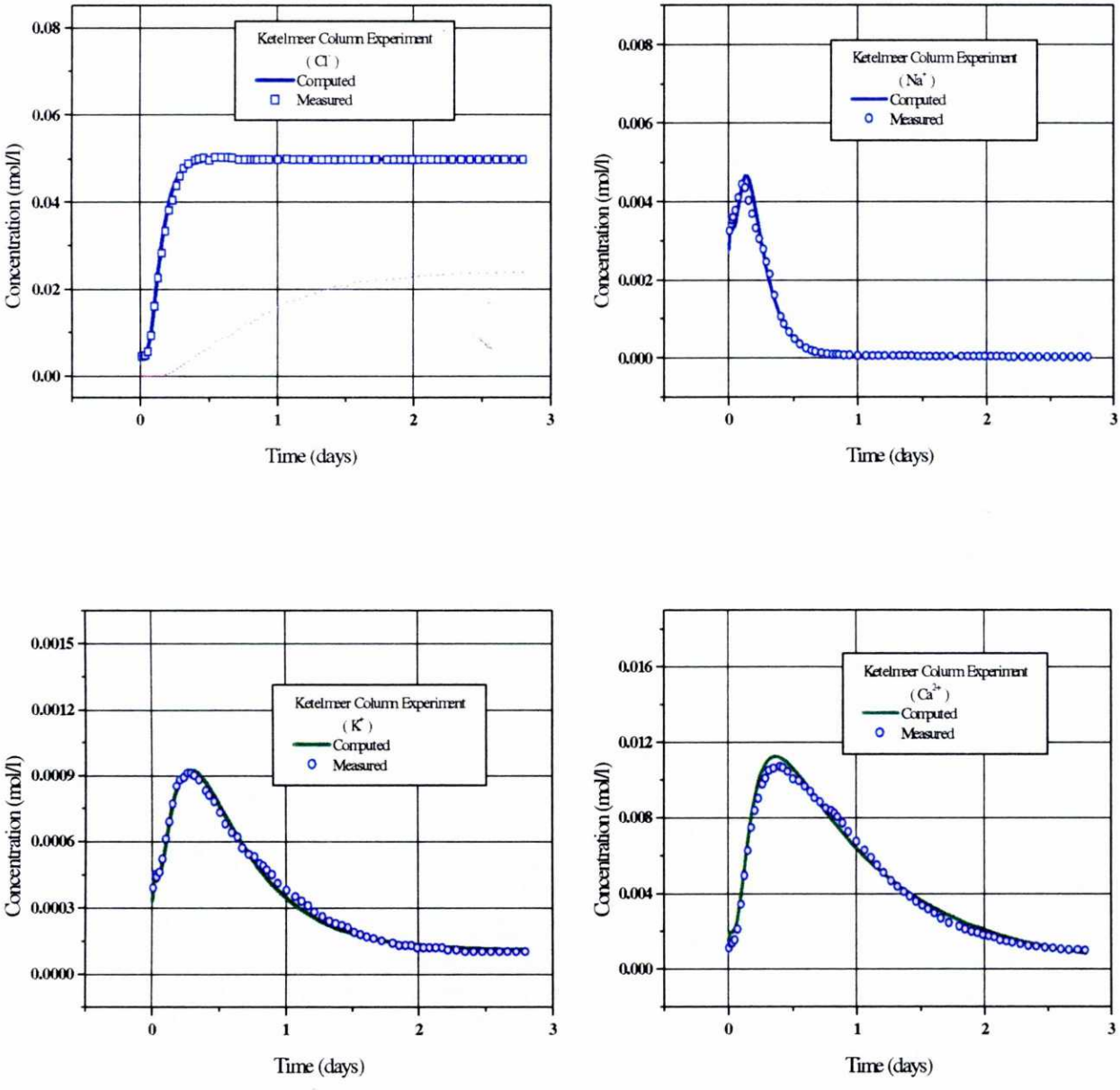


Figure 5.4.4A. Measured (symbols) and computed (lines) concentration breakthrough curves when only exchange reactions are considered. (Case 7: Estimating the selectivity coefficients, CEC, initial and boundary concentrations)

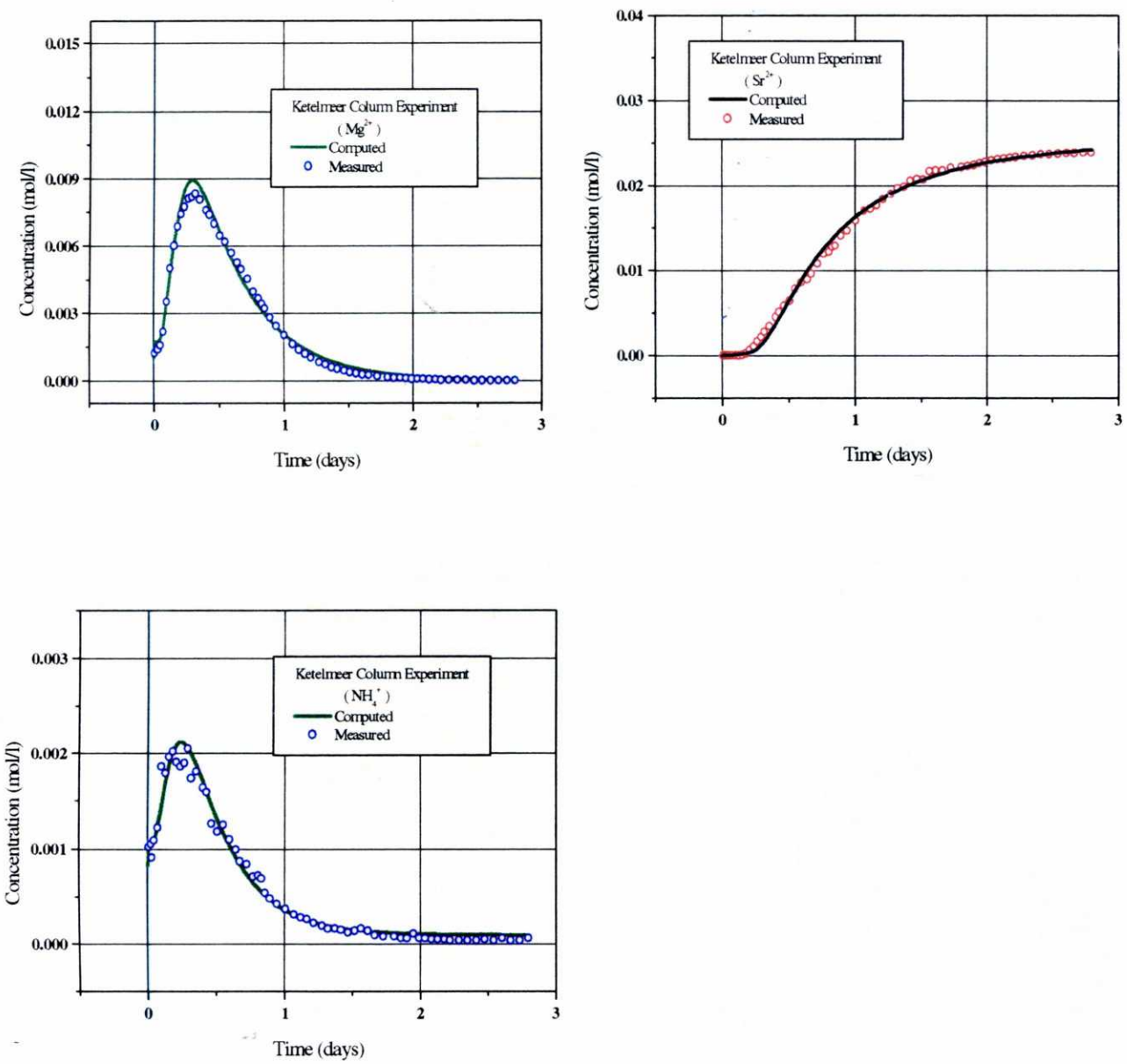


Figure 5.4.4B. Measured (symbols) and computed (lines) concentration breakthrough curves when only exchange reactions are considered. (Case 7: Estimating the selectivity coefficients, CEC, initial and boundary concentrations)

5.4.7. ESTIMATION OF INITIAL CONCENTRATIONS, SPECIFIC SURFACE, SELECTIVITY COEFFICIENTS AND CEC

In Case 3 calcite dissolution/precipitation was assumed at equilibrium. The estimation results indicate that this reaction is not relevant in this experiment possibly because the experimental time is too short (2.8 days) to make the reaction reach equilibrium. The chemical reaction characteristic time (t_{eq}) can be calculated approximately by equation,

$$t_{eq} = \frac{C_{Ca}^{eq}}{k \cdot A}$$

where, C is Ca^{2+} equilibrium concentration, k is the kinetic rate constant and A is specific surface.

Taking C = $5 \cdot 10^{-3}$ mol/l from the average measured Ca^{2+} concentration, k = $5 \cdot 10^{-8}$ mol/m^2/s and A = 20 dm^2/dm^3 from the database of CORE^{2D} respectively, t_{eq} is equal to 5.8 days. This means that calcite dissolution has not reached equilibrium after the 2.8-day duration of the column experiment. Therefore, in the last stage, kinetically-controlled calcite dissolution/ precipitation is included in the geochemical system. The corresponding aqueous complexation reactions are the same as those listed in Table 5.4.5. The kinetic rate constant is fixed to $5 \cdot 10^{-8}$ mol/m^2/s. The initial guess of specific surface is 20 dm^2/dm^3 and its lower and upper bounds are 5 and 100 dm^2/dm^3, respectively. Prior information of other parameters is the same as that in Case 6. Initial concentrations, specific surface (A, dm^2/dm^3), selectivity coefficients and CEC are estimated by INVERSE-CORE^{2D}. Estimation results are listed in Tables 5.4.11 and 5.4.12, and the best fit for different chemical components is shown in Figure 5.4.5.

Case 8 does not improve the results of Case 6. This means that cation exchange reactions dominate the geochemical processes in this experiment and calcite dissolution/precipitation plays a minor role and can be neglected in the inverse model.

Table 5.4.12. Estimation results for Case 8 estimating specific surface (A, dm^2/dm^3), five selectivity coefficients, CEC (meq/100g) and initial concentrations (C_0, mol/l) (objective function=$4.075 \cdot 10^2$)

Parameter	Initial guess	Estimated value	Estimation Variance	95% confidence interval		Associated eigenvalue
				Lower	Upper	
$K_{Na/K}$	0.1241	0.1134	$2.4 \cdot 10^{-8}$	0.1127	0.114	$7.89 \cdot 10^{-4}$
$K_{Na/Ca}$	0.2360	0.2696	$1.377 \cdot 10^{-8}$	0.2691	0.2701	$2.64 \cdot 10^{-6}$
$K_{Na/Mg}$	0.3762	0.3984	$3.612 \cdot 10^{-8}$	0.3976	0.3993	$6.14 \cdot 10^{-8}$
$K_{Na/Sr}$	0.3340	0.3758	$3.23 \cdot 10^{-8}$	0.375	0.3766	$4.22 \cdot 10^{-8}$
$K_{Na/NH4}$	0.2312	0.2027	$6.177 \cdot 10^{-8}$	0.2016	0.2038	$2.67 \cdot 10^{-8}$
CEC	4.262	4.436	$2.656 \cdot 10^{-6}$	4.429	4.443	$2.07 \cdot 10^{-8}$
$C_{0\,Na}$	$3.25 \cdot 10^{-3}$	$3.06 \cdot 10^{-3}$	$7.714 \cdot 10^{-12}$	$3.048 \cdot 10^{-3}$	$3.072 \cdot 10^{-3}$	$1.13 \cdot 10^{-8}$
$C_{0\,K}$	$3.94 \cdot 10^{-4}$	$3.397 \cdot 10^{-4}$	$1.859 \cdot 10^{-13}$	$3.378 \cdot 10^{-4}$	$3.416 \cdot 10^{-4}$	$7.31 \cdot 10^{-12}$
$C_{0\,Ca}$	$1.12 \cdot 10^{-3}$	$1.337 \cdot 10^{-3}$	$6.121 \cdot 10^{-13}$	$1.333 \cdot 10^{-3}$	$1.340 \cdot 10^{-3}$	$8.18 \cdot 10^{-13}$
$C_{0\,NH4}$	$1.02 \cdot 10^{-3}$	$8.596 \cdot 10^{-4}$	$8.64 \cdot 10^{-13}$	$8.555 \cdot 10^{-4}$	$8.637 \cdot 10^{-4}$	$4.77 \cdot 10^{-13}$
A	20	18.85	$7.892 \cdot 10^{-4}$	18.72	18.97	$1.54 \cdot 10^{-13}$

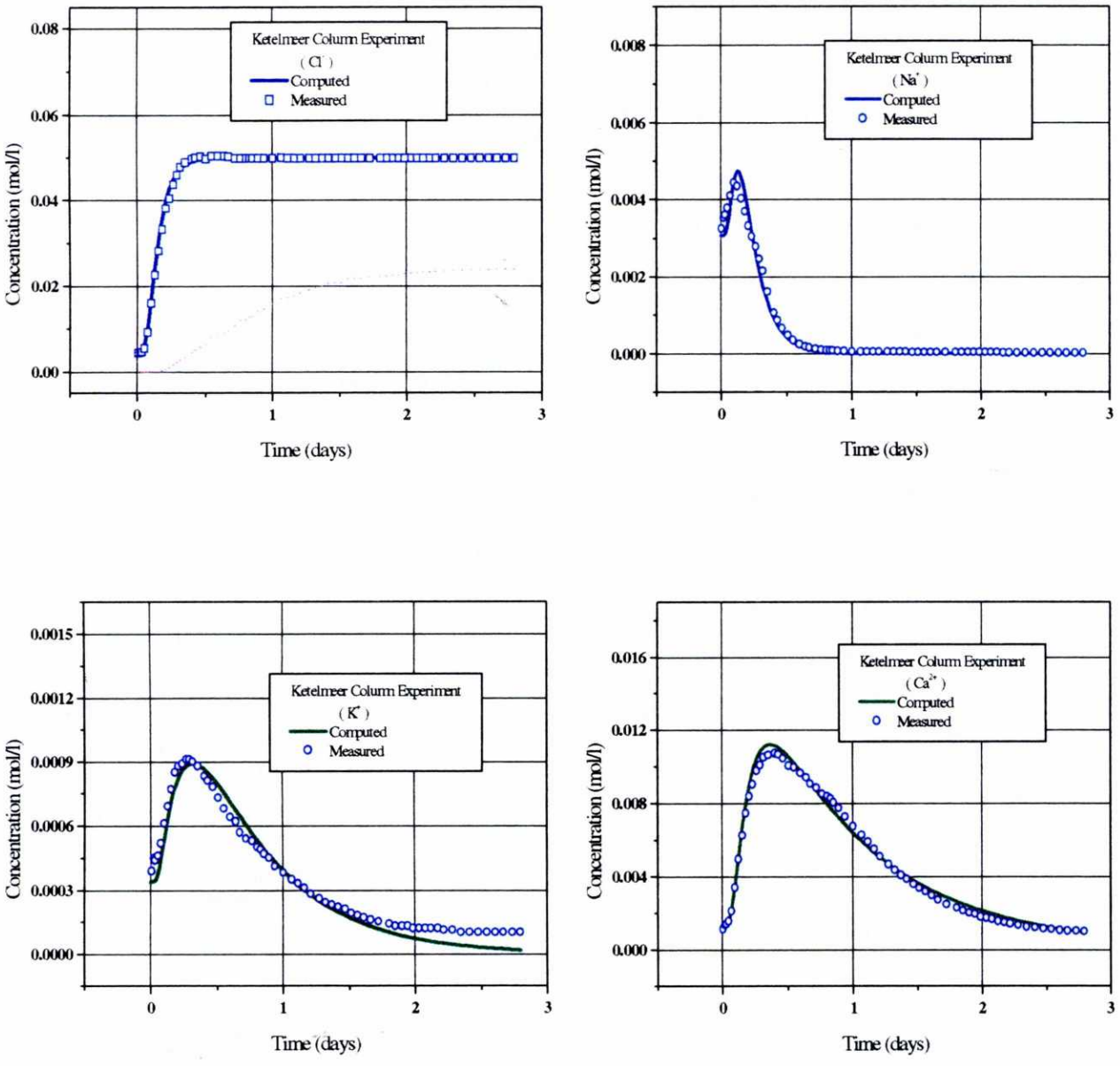
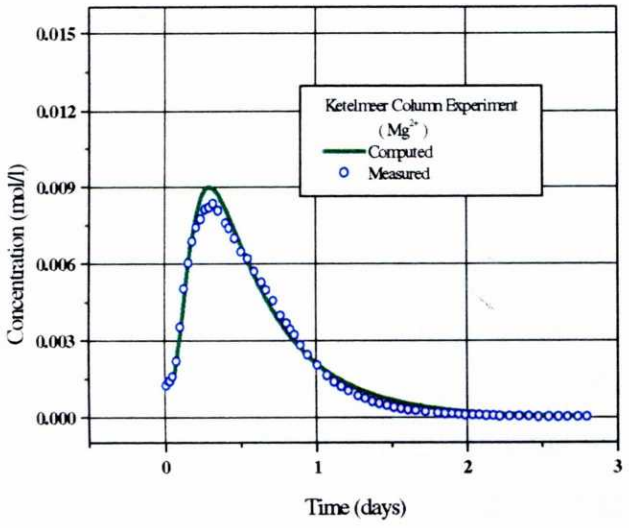
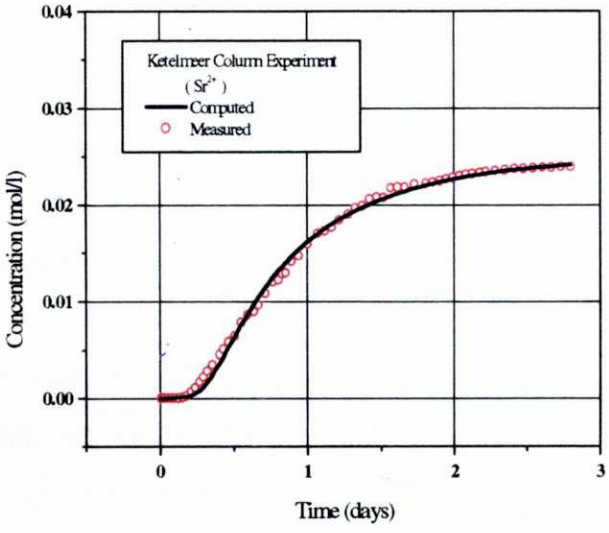


Figure 5.4.5A. Measured (symbols) and computed (lines) concentration breakthrough curves. (Case 8: Estimating the specific surface, selectivity coefficients, CEC, and initial and boundary concentrations)





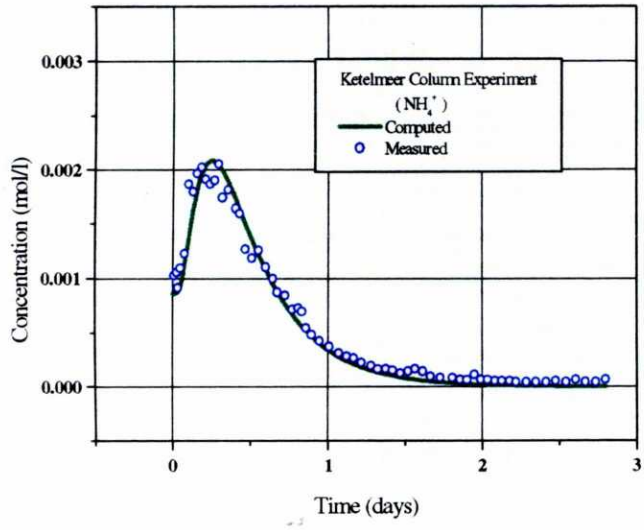


Figure 5.4.5B. Measured (symbols) and computed (lines) concentration breakthrough curves. (Case 8: Estimating the specific surface, selectivity coefficients, CEC, and initial and boundary concentrations)

5.4.8. SUMMARY

This real case study dealt with the interpretation of a column experiment performed on clay sediments from Ketelmeer. Five selectivity coefficients, CEC, four initial and four boundary concentrations have been estimated using INVERSE-CORE^{2D} under different conditions. It can be concluded that:

1. Breakthrough curves are highly sensitive to CEC. Our estimates are about a half the value reported by Appelo et al. (1990). Additional information received from Dr. Appelo has demonstrated that the estimated CEC value is reasonable.
2. Estimated selectivity coefficients are similar to published values (Appelo et al., 1990). The objective function is very sensitive to these parameters. Slight changes of these parameters lead to significant improvements in curve matching.
3. The objective function increases and the fit gets worse when H^+ exchange, aqueous dissociation reactions and equilibrium or kinetically-controlled calcite dissolution/ precipitation are added to the geochemical system. It indicates that the above-mentioned chemical reactions are not relevant for this column experiment.
4. Estimates of initial concentrations of most components almost coincide with measured data. However, there might be some observation errors in the measured data. The inverse model neglected some chemical components such as Fe and Mn, which could affect some of the estimates of initial concentrations.
5. Boundary concentrations of cations have been estimated. A significant reduction in the objective function is achieved. These results, however, need to be tested with additional information about experimental conditions.

5.5. INVERSE ANALYSIS OF LLOBREGAT DELTA AQUITARD HYDROCHEMICAL DATA

5.5.1. PROBLEM STATEMENT

This problem was originally investigated by Manzano (1993), Manzano and Custodio (1995), and Xu et al. (1999). All of the data used here were basically taken from previous investigators. The origin of saline water in the Llobregat Delta aquitard (Barcelona, Spain) can be explained using chromatographic theory. The cation content of interstitial water in the delta aquitard shows the typical distribution of saline water in equilibrium with the soil which is displaced by fresh water flow. According to the work of Xu et al. (1999), now measured chromatographic pattern is established in about 3200 year.

In addition to hydrological transport processes, the model accounts for aqueous complexation, cation exchange and calcite dissolution-precipitation. Simulated aqueous concentrations were compared to their corresponding measurements for the purpose of estimating selectivity coefficients. The sensitivity of the objective function to changes in selectivities and initial and boundary concentrations was explained. Finally, redox reactions were included in addition to aqueous complexation, cation exchange and calcite dissolution.

Main hydrodynamic parameters include: porosity = 0.411, dispersivity = 2.04 m, and Darcy's velocity = $6.5 \cdot 10^{-6}$ m/day ($2.373 \cdot 10^{-3}$ m/year), which were estimated by Manzano (1993) with the conservative Cl^- concentrations.

The aquitard column was initially filled with high concentrated saline water. Table 5.5.1 summarizes the main chemical reactions considered in the model. The fresh water comes from deep aquifer which is treated here as the bottom boundary (see Table 5.5.2). Gaines-Thomas convention was used for cation exchange. With Na^+ as the reference cation, measured selectivity coefficients are listed in Table 5.5.3.

Table 5.5.1. List of geochemical reactions considered for the Llobregat Delta aquitard

Geochemical reactions	$Log_{10}(K)$ at 25 °C
Aqueous dissociation:	
$OH^- = H_2O - H^+$	13.995
$CO_3^{2-} = HCO_3^- - H^+$	10.329
$CO_2(aq) = HCO_3^- + H^+ - H_2O$	-6.3447
$CaHCO_3^+ = Ca^{2+} + HCO_3^-$	-1.0467
$MgHCO_3^+ = Mg^{2+} + HCO_3^-$	-1.0357
$CaCO_3(aq) = Ca^{2+} + HCO_3^- - H^+$	7.0017
$MgCO_3(aq) = Mg^{2+} + HCO_3^- - H^+$	7.3499
$NaHCO_3(aq) = Na^+ + HCO_3^-$	-0.1541
$CaSO_4(aq) = Ca^{2+} + SO_4^{2-}$	-2.1111
$MgSO_4(aq) = Mg^{2+} + SO_4^{2-}$	-2.309
$NaSO_4^- = Na^+ + SO_4^{2-}$	-0.082
$KSO_4^- = K^+ + SO_4^{2-}$	-0.8796
Reduction-oxidation:	
Methane (aq) $= HCO_3^- - 3H_2O + 9H^+ + 8e^-$	-27.655
$HS^- = SO_4^{2-} - 4H_2O + 9H^+ + 8e^-$	-33.48
$H_2S(aq) = SO_4^{2-} - 4H_2O + 10H^+ + 8e^-$	-40.467
Cation exchange:	
$Na^+ + 0.5Ca\text{-}X_2 = 0.5Ca^{2+} + Na\text{-}X$	
$Na^+ + 0.5Mg\text{-}X_2 = 0.5Mg^{2+} + Na\text{-}X$	
$Na^+ + K\text{-}X = K^+ + Na\text{-}X$	
$Na^+ + NH_4\text{-}X = NH_4^+ + Na\text{-}X$	
Mineral precipitation-dissolution:	
Calcite $= Ca^{2+} + HCO_3^- - H^+$	1.8487

Table 5.5.2. Initial and bottom boundary total dissolved chemical component concentrations (mmol/l) (from Xu,1996)

Component	Initial	Boundary
Na	521.9	51.67
K	12.82	2.74
Ca	5	1.8
Mg	32.6	3
NH_4	15	1.9
pH	7.9	6.83
Cl	613.17	5.01
HCO_3	32	32
SO_4	8.5	0.9
$O_{2(aq)}$	$3.71 \cdot 10^{-75}$	$1.08 \cdot 10^{-59}$
pE	-5	0

Table 5.5.3. Cation selectivity coefficients(Gaines-Thomas convention) and CEC values initially calculated from experimental data (according to Manzano, 1993); K(Na/cation) indicates cation selectivity with respect to Na^+

Depth (m)	K(Na/K)	K(Na/Ca)	K(Na/Mg)	CEC(meq/100g)
1	0.1217	0.2979	0.2606	14.16
5	0.1743	0.2541	0.4645	15
21	0.1684	0.257	4.2106	15
33	0.0787	0.0649	0.0698	17

5.5.2. ESTIMATION OF SELECTIVITY COEFFICIENTS AND INITIAL CONCENTRATIONS

Measured concentrations of different chemical species at several depths are listed in Table 5.5.4. The estimation process was performed first leaving out redox processes. Later, these reactions were also added to the inverse model.

Table 5.5.4. Measured concentrations at well CM (mol/l) (after Manzano,1993) and the weighting coefficients for each chemical species computed from Equation 3.2.3.

Depth (m)	Na	K	Ca	Mg	HCO_3	Cl	SO_4*	NH_4
3.5	0.522	0.0101*	0.005	0.0326	0.032*	0.573	0.0085	0.015
8	0.431	0.0116	0.007*	0.026	0.03	0.489	0.0018	0.0142
12	0.337	0.0098	0.0044	0.0205	0.034	0.411	0	0.0085
16	0.28	0.0089	0.0024	0.0105	0.036	0.286	0.0005	0.0085
19	0.211	0.0076	0.003*	0.0062	0.038	0.208	0.0013	0.0119*
23	0.15	0.0047	0.0004	0.0036	0.04	0.127	0	0.0048
27	0.101	0.0041	0.0008	0.003	0.038	0.079	0	0.0051
31	0.052	0.0027	0.0018	0.003	0.032	0.034	0	0.002
Weight	22	1091	2127	303	135	18	/	853

(*: Values and species marked with (*) are not considered for the estimation because they are considered highly uncertain)

5.5.2.1. Estimation without redox reactions (Case 1)

Leaving out redox reactions, four selectivity coefficients and five initial concentrations (H^+, Na^+, K^+, Ca^{2+}, Mg^{2+}) were estimated. After more than 20 trial runs with different initial values and different increments of parameters, a convergent solution was achieved after seven iterations. Prior information and bounds for each parameter are listed in Table 5.5.5. Estimation results of selectivity coefficients and initial concentrations are listed in Table 5.5.6.

Table 5.5.5. Lower and upper bounds and prior parameter information for the case without redox reactions (C_0 denotes initial concentration, mol/l and CEC, meq/100g)

Parameters	Lower bound	Upper bound	Prior information
$K_{Na/K}$	0.01	0.9	0.1543
$K_{Na/Ca}$	0.01	0.9	0.2697
$K_{Na/Mg}$	0.01	0.9	0.3626
$K_{Na/NH4}$	0.01	0.9	0.5
CEC	5	25	15
$C_{0\,H}$	10^{-8}	10^{-6}	$1.2 \cdot 10^{-7}$
$C_{0\,Na}$	0.3	0.8	0.56
$C_{0\,K}$	0.01	0.03	0.016
$C_{0\,Ca}$	0.001	0.01	0.006
$C_{0\,Mg}$	0.01	0.1	0.0326

Table 5.5.6. Initial guess, estimated values, estimation variance and confidence intervals for the case without redox reactions. See Table 5.5.5 for the definition of the parameters (objective function value =54.05)

Parameter	Initial guess	Estimated value	Estimation variance	90% confidence interval	
				Lower bound	Upper bound
$K_{Na/K}$	0.1543	0.0712	$1.587 \cdot 10^{-3}$	0.01	0.2305
$K_{Na/Ca}$	0.2697	0.2444	$8.061 \cdot 10^{-3}$	0.01	0.6034
$K_{Na/Mg}$	0.3626	0.014	$7.827 \cdot 10^{-7}$	0.0104	0.0175
$K_{Na/NH4}$	0.5	0.3332	$5.645 \cdot 10^{-3}$	0.0328	0.6336
CEC	15	15.12	0.36	12.72	17.52
$C_{0\,H}$	$1.2 \cdot 10^{-7}$	$9.268 \cdot 10^{-8}$	$2.593 \cdot 10^{-11}$	10^{-8}	$2.045 \cdot 10^{-5}$
$C_{0\,Na}$	0.56	0.5293	$8.241 \cdot 10^{-7}$	0.5257	0.533
$C_{0\,K}$	0.016	0.0139	$1.974 \cdot 10^{-8}$	0.0134	0.0145
$C_{0\,Ca}$	0.006	0.014	$3.419 \cdot 10^{-8}$	0.0133	0.0148
$C_{0\,Mg}$	0.0326	0.0355	$6.283 \cdot 10^{-8}$	0.0345	0.0365

The final curve fitting results are plotted in Figure 5.5.1. The fit is good for all species, except for HCO_3^-. Estimated values of selectivity coefficients are smaller than measured values, especially for $K_{Na/Mg}$. The confidence intervals of these parameters, as well as that of the H^+ initial concentration, are wide in Table 5.5.6. It means that their estimates are relatively uncertain. For improving the estimated results, it is necessary to consider more chemical reactions in the geochemical system.

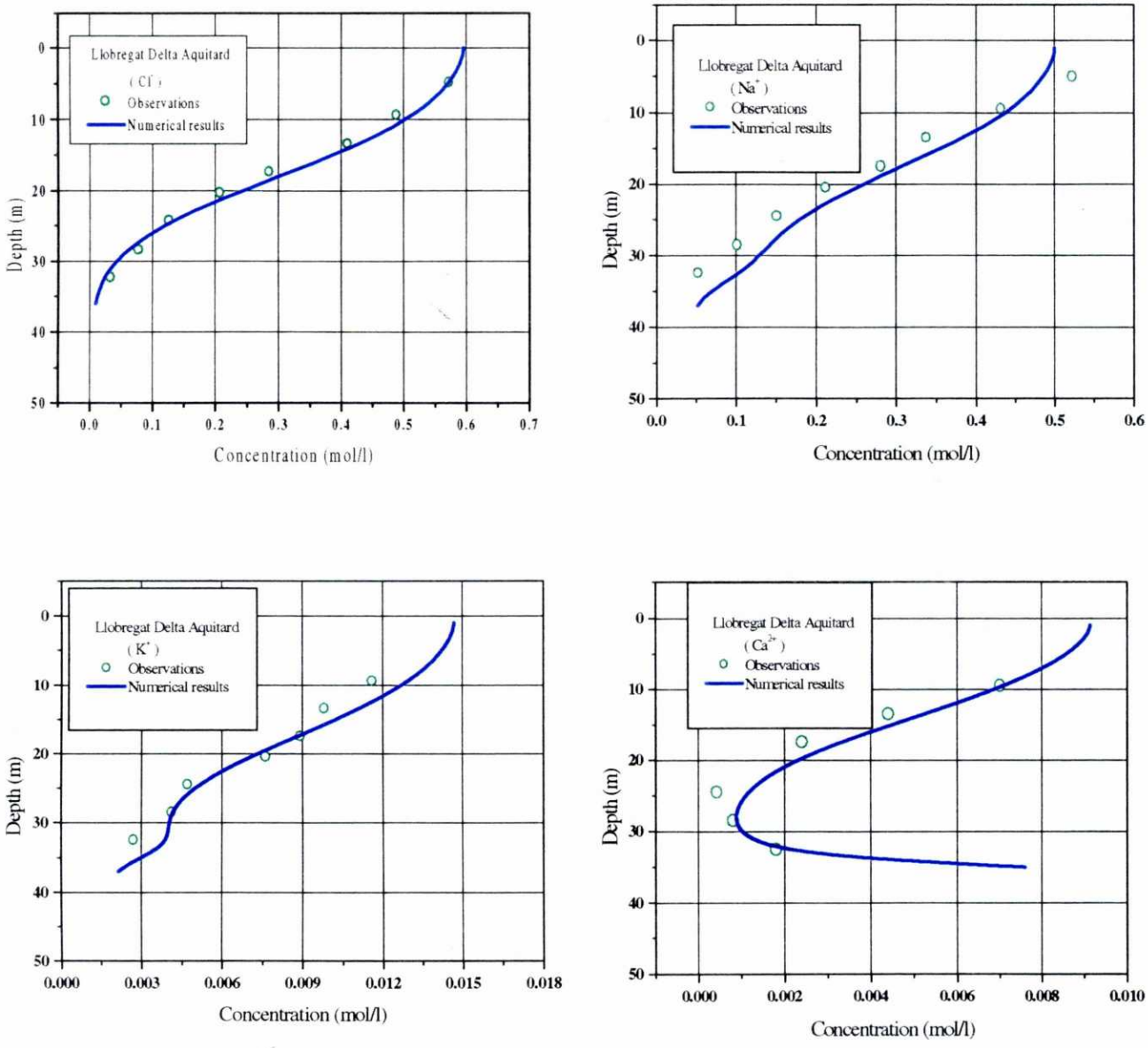


Figure 5.5.1A. Computed (lines) and measured (symbols) concentrations of dissolved species after 3200 years when estimating 10 parameters (Case 1: without redox reactions)

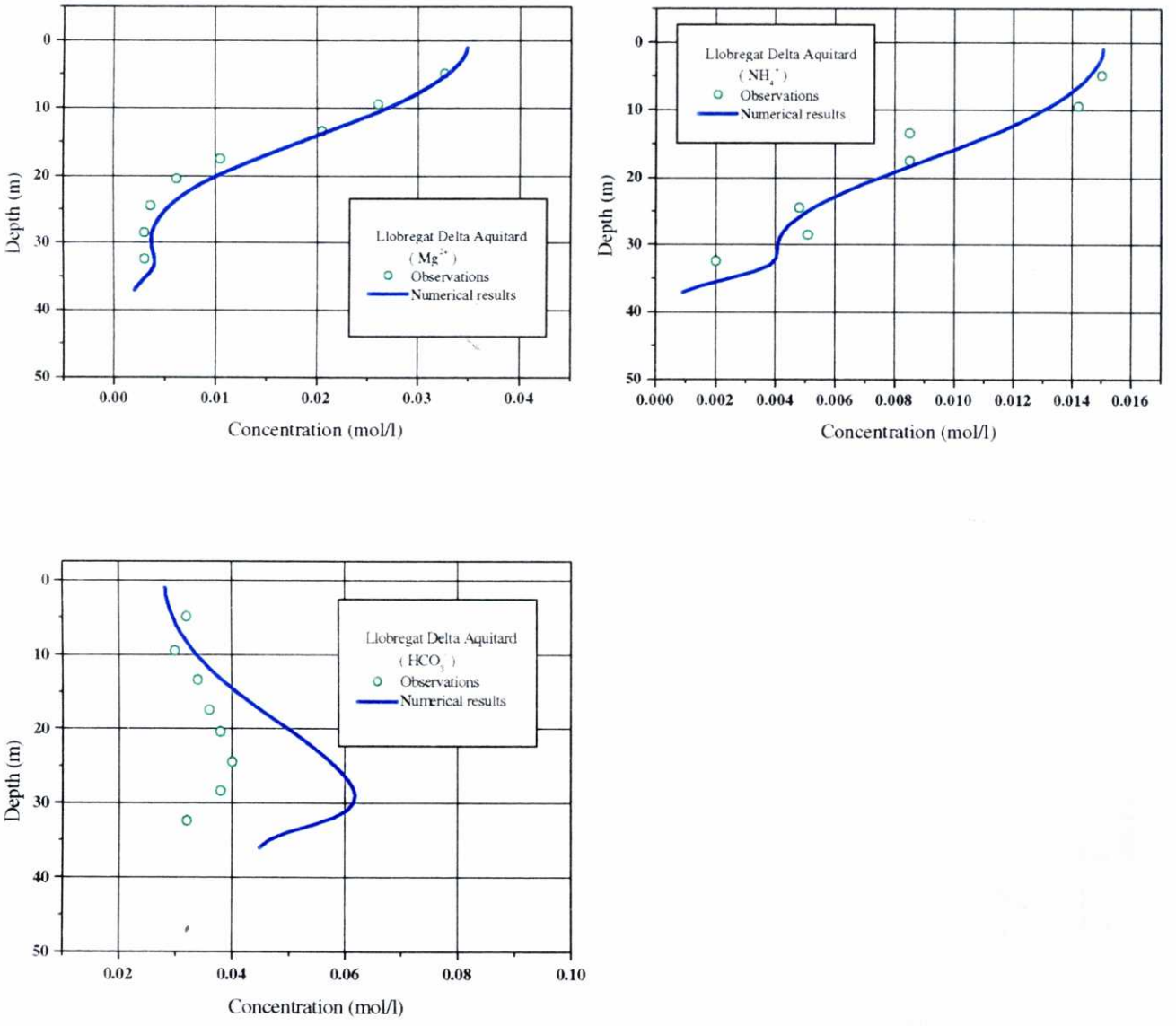


Figure 5.5.1B. Computed (lines) and measured (symbols) concentrations of dissolved species after 3200 years when estimating of 10 parameters (Case 1: without redox reactions)

5.5.2.2. Estimation with redox processes

In this case (Case 2) all the chemical reactions listed in Table 5.5.2 are considered including redox reactions. In addition to dispersivity and hydraulic conductivity, four selectivity coefficients and five initial concentrations (H^+, Na^+, K^+ , Ca^{2+}, Mg^{2+}) are estimated. Prior information and parameter intervals are listed in Table 5.5.7. After a large number of trial runs with different initial values and different increments of parameters, a convergent solution is achieved after six iterations. Estimation results are shown in Table 5.5.8.

Table 5.5.7. Lower and upper bounds and prior parameter information for the case with redox reactions (C_0 denotes initial concentration, mol/l and CEC, meq/100g)

Parameters	Lower bound	Upper bound	Prior information
$K_{Na/K}$	0.01	0.9	0.1543
$K_{Na/Ca}$	0.01	0.9	0.2697
$K_{Na/Mg}$	0.01	0.9	0.3626
$K_{Na/NH4}$	0.01	0.9	0.5
CEC	5	25	15
$C_{0\,H}$	10^{-8}	10^{-6}	$1.2\cdot10^{-7}$
$C_{0\,Na}$	0.3	0.8	0.56
$C_{0\,K}$	0.01	0.03	0.016
$C_{0\,Ca}$	0.001	0.01	0.006
$C_{0\,Mg}$	0.01	0.1	0.0326
Dispersivity (m)	1	5	2.04
Conductivity (m/d)	$2.74\cdot10^{-5}$	$1.37\cdot10^{-4}$	$6.5\cdot10^{-5}$

Table 5.5.8. Initial guess, estimated values, estimation variance and confidence intervals for the case with redox reactions (Case 2). See Table 5.5.5 for the definition of the parameters (objective function value = 30.96)

Parameter	Initial guess	Estimated value	Estimation variance	90% confidence interval	
				Lower bound	Upper bound
$K_{Na/K}$	0.1543	0.0708	$1.786 \cdot 10^{-4}$	0.0133	0.1284
$K_{Na/Ca}$	0.2697	0.2895	$1.685 \cdot 10^{-3}$	0.1127	0.4663
$K_{Na/Mg}$	0.3626	0.0136	$8.280 \cdot 10^{-8}$	0.0123	0.0148
$K_{Na/NH4}$	0.5	0.2671	$1.415 \cdot 10^{-3}$	0.1051	0.4291
CEC	15	14.72	$5.398 \cdot 10^{-2}$	13.72	15.72
$C_{0\,H}$	$1.2 \cdot 10^{-7}$	$9.163 \cdot 10^{-8}$	$1.683 \cdot 10^{-12}$	10^{-8}	$5.679 \cdot 10^{-6}$
$C_{0\,Na}$	0.56	0.5290	$9.185 \cdot 10^{-8}$	0.5277	0.5303
$C_{0\,K}$	0.016	0.0135	$2.187 \cdot 10^{-9}$	0.0133	0.0137
$C_{0\,Ca}$	0.006	0.0174	$4.385 \cdot 10^{-9}$	0.0171	0.0177
$C_{0\,Mg}$	0.0326	0.0355	$6.898 \cdot 10^{-9}$	0.0351	0.0358
Dispersivity (m)	2.04	2.272	$3.261 \cdot 10^{-3}$	2.247	2.296
Conductivity (m/d)	$6.5 \cdot 10^{-5}$	$6.71 \cdot 10^{-5}$	$8.138 \cdot 10^{-9}$	$6.68 \cdot 10^{-5}$	$6.71 \cdot 10^{-5}$

In Table 5.5.8 the objective function value reduces from 54.05 in Case 1 to 30.96 due to the consideration of redox reactions and the estimation of dispersivity and hydraulic conductivity. Estimated parameters are slightly different from those listed in Table 5.5.6. Their confidence intervals reduce greatly, meaning that their uncertainty has decreased.

In order to check the influence of initial pE or dissolved oxygen concentration ($C_{0\,O2(aq)}$) and boundary concentrations (C_B) of some species (Na^+, K^+, Ca^{+2}, and Mg^{+2}), in the following inverse run (Case 3), the dispersivity and hydraulic conductivity were fixed and the above mentioned five parameters were estimated. So, 15 parameters were estimated in this case. During the estimation process, the concentrations of H^+ and $O_{2(aq)}$ were log-transformed to improve the convergence rate. Estimated parameters and confidence

intervals are listed in Table 5.5.9. A slight reduction in the objective function is achieved. The estimated H^+ initial concentration ($9.185 \cdot 10^{-8}$ mol/l) corresponds to a pH of 7.04 and $O_{2(aq)}$ initial concentration ($3.7 \cdot 10^{-74}$ mol/l) corresponds to a pE of –5.05. The estimation confidence interval of initial $O_{2(aq)}$ concentration has a wide range, which means that this parameter is relatively uncertain because there are no measured data of pE or $O_{2(aq)}$. The final fits are plotted in Figure 5.5.2. The results are excellent for all species except for HCO_3^-. The fit obtained with redox reactions is better than that without redox, which indicates that redox reactions must be considered for the Llobregat Delta aquitard.

Table 5.5.9. Initial guess, estimated values, estimation variance and confidence intervals for the case with redox reactions (Case 3). See Table 5.5.5 for the definition of the parameters (objective function value = 29.98)

Parameter	Initial guess	Estimated value	Estimation variance	90% confidence interval	
				Lower bound	Upper bound
$K_{Na/K}$	0.1543	0.0396	$4.132 \cdot 10^{-4}$	10^{-2}	0.1356
$K_{Na/Ca}$	0.2697	0.2798	$5.358 \cdot 10^{-2}$	10^{-2}	1.373
$K_{Na/Mg}$	0.3626	0.019	$6.385 \cdot 10^{-6}$	$7.035 \cdot 10^{-3}$	0.0309
$K_{Na/NH4}$	0.5	0.3557	$5.738 \cdot 10^{-2}$	10^{-2}	1.487
CEC	15	15.14	1.925	8.583	21.69
$C_{0\,H}$	$1.2 \cdot 10^{-7}$	$1.2 \cdot 10^{-7}$	$2.929 \cdot 10^{-8}$	10^{-8}	8.08410^{-4}
$C_{0\,Na}$	0.56	0.5334	$5.092 \cdot 10^{-6}$	0.5227	0.544
$C_{0\,K}$	0.016	0.0142	$1.174 \cdot 10^{-7}$	$1.26 \cdot 10^{-2}$	$1.584 \cdot 10^{-2}$
$C_{0\,Ca}$	0.006	0.0150	$2.635 \cdot 10^{-7}$	$1.255 \cdot 10^{-2}$	$1.74 \cdot 10^{-2}$
$C_{0\,Mg}$	0.0326	0.0365	$4.064 \cdot 10^{-7}$	$3.346 \cdot 10^{-2}$	$3.948 \cdot 10^{-2}$
$C_{0\,O2(aq)}$	$3.708 \cdot 10^{-75}$	$3.7 \cdot 10^{-74}$	$9.03 \cdot 10^{-76}$	10^{-90}	$1.419 \cdot 10^{-37}$
$C_{B\,Na}$	$5.167 \cdot 10^{-2}$	$3.94 \cdot 10^{-2}$	$4.159 \cdot 10^{-6}$	$2.977 \cdot 10^{-2}$	$4.903 \cdot 10^{-2}$
$C_{B\,K}$	$2.74 \cdot 10^{-3}$	$7 \cdot 10^{-4}$	$8.691 \cdot 10^{-7}$	10^{-4}	$5.103 \cdot 10^{-3}$
$C_{B\,Ca}$	$1.8 \cdot 10^{-3}$	$1.073 \cdot 10^{-3}$	$5.771 \cdot 10^{-7}$	$8 \cdot 10^{-4}$	$4.66 \cdot 10^{-3}$
$C_{B\,Mg}$	$3 \cdot 10^{-3}$	$4.682 \cdot 10^{-3}$	$5.84 \cdot 10^{-7}$	$1.073 \cdot 10^{-3}$	$8.291 \cdot 10^{-3}$

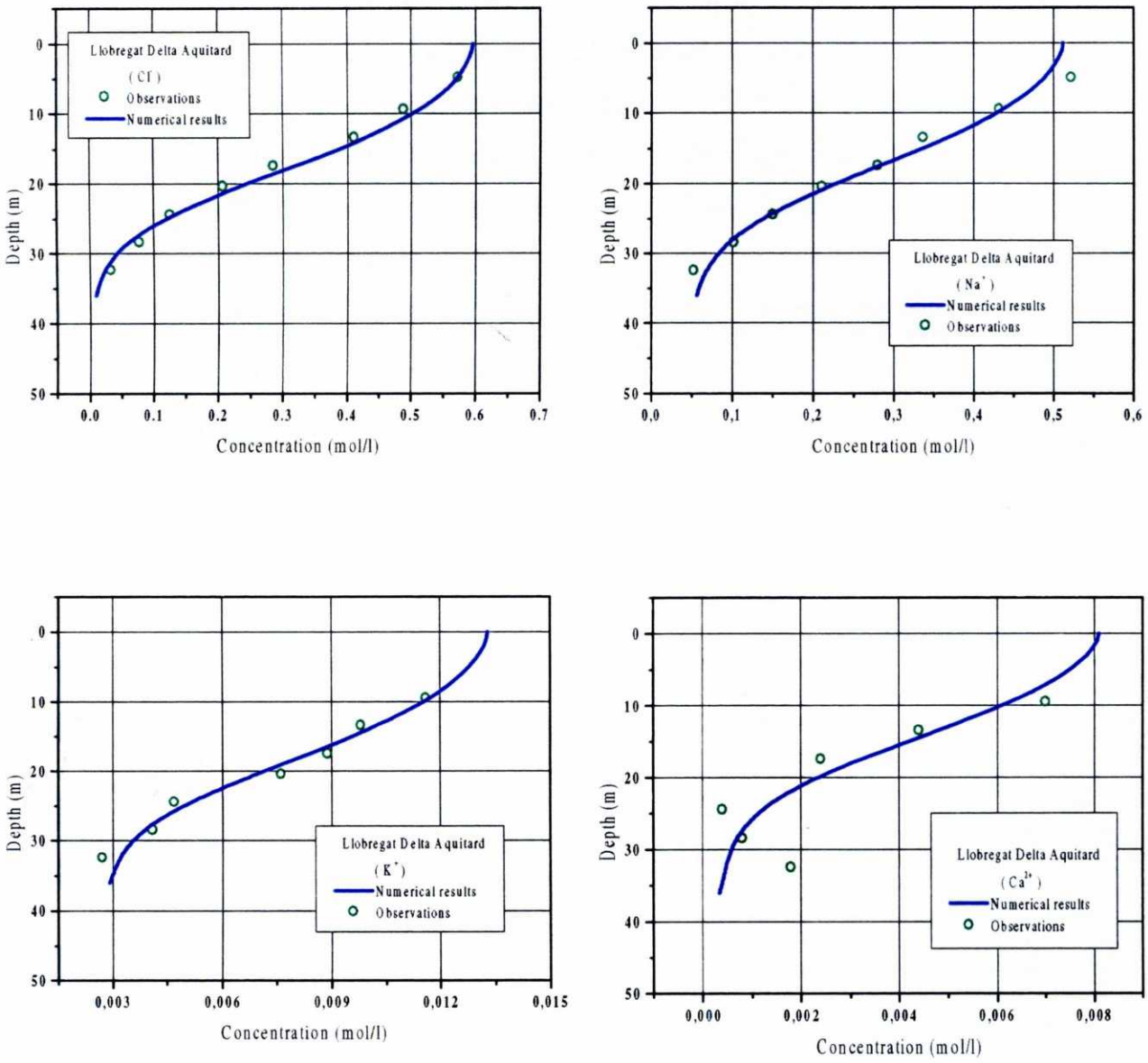


Figure 5.5.2A. Computed (lines) and measured (symbols) concentrations of dissolved species after 3200 years when estimating 15 parameters (Case 3: with redox reactions)

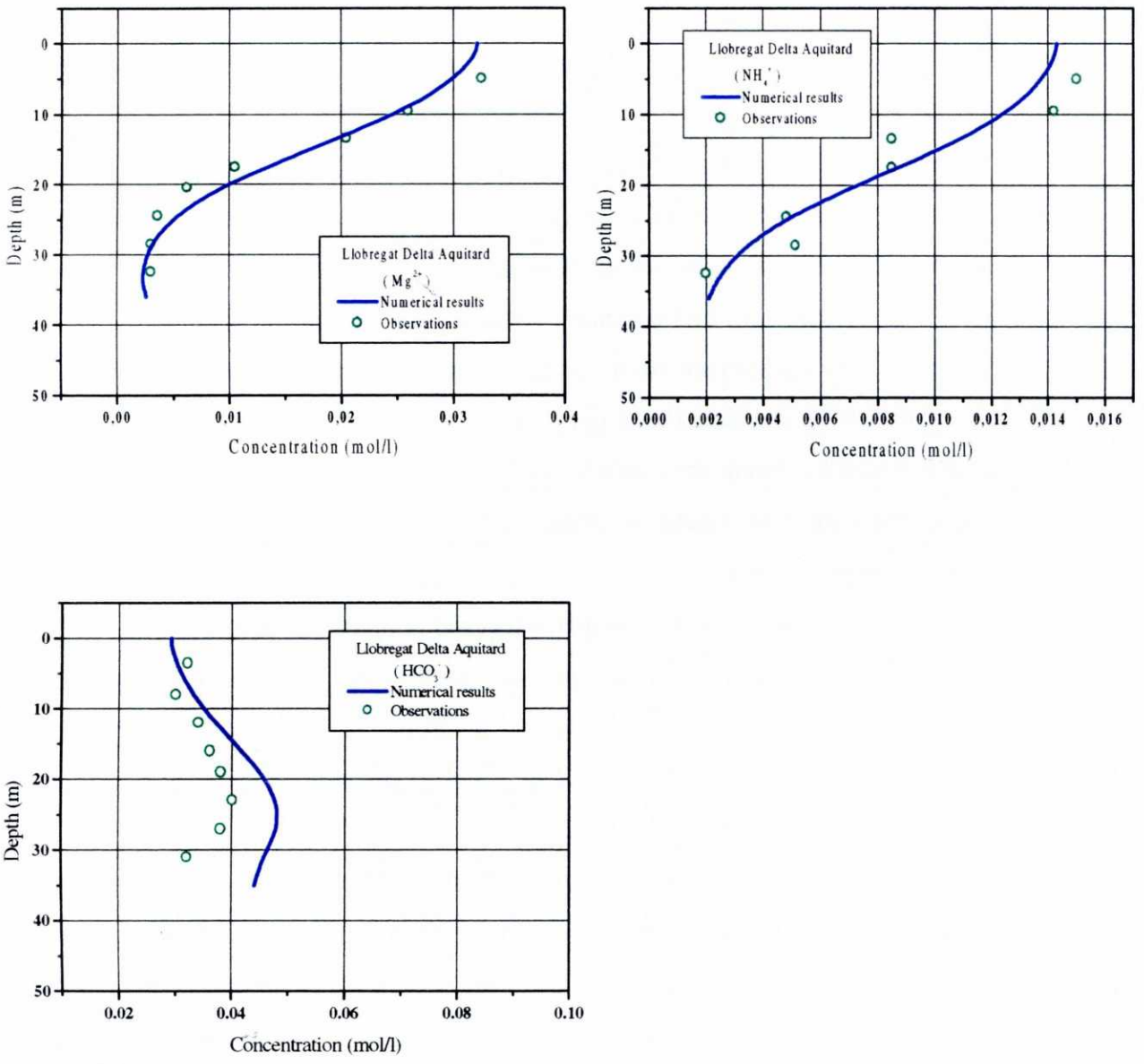


Figure 5.5.2B. Computed (lines) and measured (symbols) concentrations of dissolved species after 3200 years when estimating of 15 parameters (Case 3: with redox reactions)

5.5.3. SUMMARY

By using INVERSE-CORE^{2D}, selectivity coefficients, CEC, hydraulic conductivity, dispersivity, initial and boundary concentrations (total 17 parameters) were estimated. Compared to previous results obtained by Xu (1996) and Xu et al. (1999), we conclude that:

1) Estimated CEC and selectivity coefficients of $K_{Na/Ca}$ and $K_{Na/NH4}$ are close to the values that Xu (1996) and Xu et al. (1999) obtained by trial and error. However, our estimated selectivity coefficients $K_{Na/K}$ and $K_{Na/Mg}$ are smaller.
2) Initial (3200 years ago) and boundary concentrations of chemical components in Llobregat Delta aquitard are not accurately known. Manzano (1993), Manzano and Custodio (1995), and Xu (1996) gave the values according to their analyses and predicted results. Using their results as parameter prior information and initial values, our estimated results of initial and boundary concentrations lead to a significant improvement.
3) Similar to Xu (1996), it can be concluded that redox processes have a strong effect on the transport of reactive cations, and such effect is exerted through a pronounced change in pH as well as on a change in total dissolved carbon. Accounting for redox processes leads to a significant improvement in matching measured Na, K, Mg and Ca concentrations.
4) Estimated conductivity and dispersivity are slightly different from those estimated by Manzano (1993). She only used the concentrations of Cl^- to estimate these parameters. Our inverse model, on the other hand, used the concentration data of 7 chemical components (Cl^-, Na^+, K^+, Ca^{+2}, Mg^{+2}, NH_4^+, and HCO_3^-). So, our estimated results have a larger confidence and reliability.
5) The capabilities of our inverse model to estimate various types of parameters and account for a wide range of chemical processes in the Llobregat aquitard can be taken as an indication of partial validation in the sense that the code is a useful tool that can be practically used as an inverse modeling tool for complex natural geochemical systems.

5.6. INTERPRETATION OF HYDROCHEMICAL DATA IN THE AQUIA AQUIFER, MARYLAND

5.6.1. PROBLEM STATEMENT

The Aquia Formation consists of a medium-to coarse-grained, medium-to well-sorted glauconitic-quartz sand. Carbonate shell debris is abundant and makes up approximately 5 to 20 percent of the aquifer material. Quartz sand grains make up 50 to 70 percent of the aquifer material. Although the Aquia formation is predominantly sand, some zones occur which contain significant quantities of clay and silt-sized particles (Chapelle and Drummond, 1983). Sodium-bicarbonate waters in the coastal plain aquifers of the eastern United States have been related to freshening of the aquifer (Chapelle and Knobel, 1 983). The major ions in Aquia groundwater are calcium, magnesium, sodium, potassium and bicarbonate, which collectively account for about 95% of all dissolved constituents. These investigators depict major cation patterns as a function of flow length in the aquifer. Water quality shows zonal bands with changes in concentrations of major cations that have been attributed to cation exchange and calcite dissolution/precipitation (Appelo, 1994). The observed water quality pattern was previously simulated by Appelo (1994) with PHREEQM. Xu et al. (1999) also simulated this problem for the purpose of verifying TOUGHREACT. Appelo (1994) fitted solute transport and exchange parameters as well as initial and boundary concentrations by try and error.

Here, hydrodynamic and chemical parameters will be estimated using the automatic calibration techniques developed in this dissertation. The application of INVERSE-CORE^{2D} to Aquia aquifer is useful to validate the inverse model for the estimation of flow and reactive solute transport parameters under complex field scale conditions. Estimated parameters include hydraulic conductivity, leakage rate, dispersivity, selectivity coefficients, cation exchange capacity (CEC), as well as initial and boundary concentrations of main ions. The hydrological conditions considered are similar to those proposed by Appelo (1994). Figure 5.6.1 shows an outline of the Aquia aquifer in Maryland with an estimated prepumping head distribution, which suggests a confined aquifer in the upstream part and a gradual loss of water in the downstream

part of the aquifer (Chapelle and Drummond, 1983). Figure 5.6.2 shows a schematic cross section along a flow path. The aquifer is bounded to the east by a change in facies. Leakage probably occurs via Pleistocene channels that cut through the confining beds. According to the estimation of Appelo (1994), the observed chromatographic pattern in the aquifer has been established during the last 10^5 years.

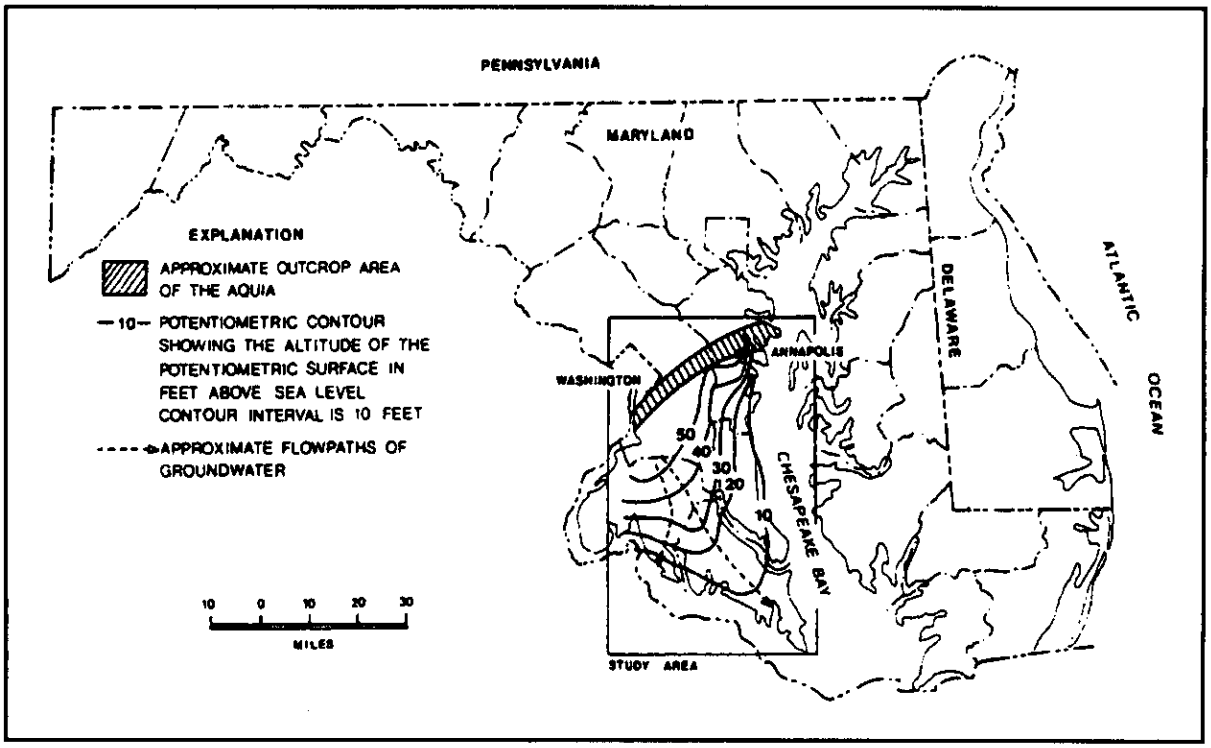


Figure 5.6.1. Outline of the Aquia aquifer in Maryland with estimated prepumping head distribution (adapted from Chapelle and Knobel, 1983) (1 foot = 0.3046 m, 1 mile = 1.608 km)

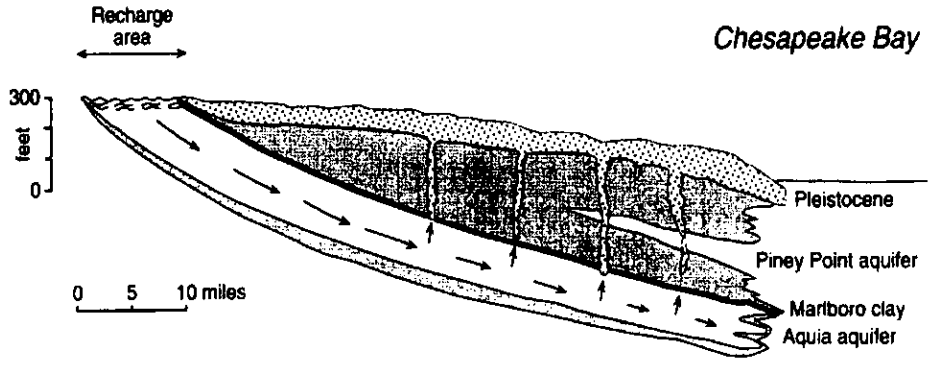


Figure 5.6.2. Schematic cross section of the Aquia aquifer. Recharge occurs in the outcrop of the formation; discharge takes place evenly in the downstream half (adapted from Appelo, 1994)

5.6.2. FORMULATION OF THE MODEL

The hydrological conditions have been modeled assuming a one-dimensional flow tube with recharge at x = 0, and with leakage into the confining layers evenly distributed over the second half of the flow tube. The flow tube has been discretized with 70 finite elements and 72 nodes. Two parameter zones have been considered: the upstream zone #1 and the downstream zone #2. Water flow is assumed to be at steady state. Molecular diffusion can be safely disregarded. Porosity is 0.3.

Cation exchange is the main chemical reaction in the aquifer. This case illustrates the chromatographic separation of Na^+, K^+, Ca^{2+}, and Mg by freshwater displacement. The displacing cation is usually Ca^{2+}, which has a higher affinity for the exchanger than the displaced saltwater ions Na^+, K^+, and Mg^{2+}. An ordered sequence develops for the cations, whereby first Na^+, then K^+, and finally Mg^{2+} are displaced.

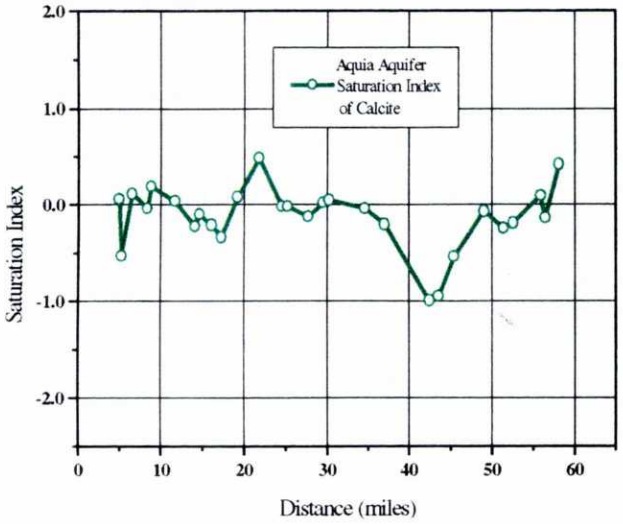
Gaines-Thomas convention is used for cation exchange. In this convention, selectivity coefficients are calculated by using the equivalent fraction of the exchanged cations for their activities. Here we choose Na^+ as the reference cation. Therefore, Na^+ selectivity is equal to 1. Selectivity coefficients for other cation, $K_{Na/K}$, $K_{Na/Ca}$, and $K_{Na/Mg}$, are defined with respect to sodium. Cation exchange reactions can be written as:

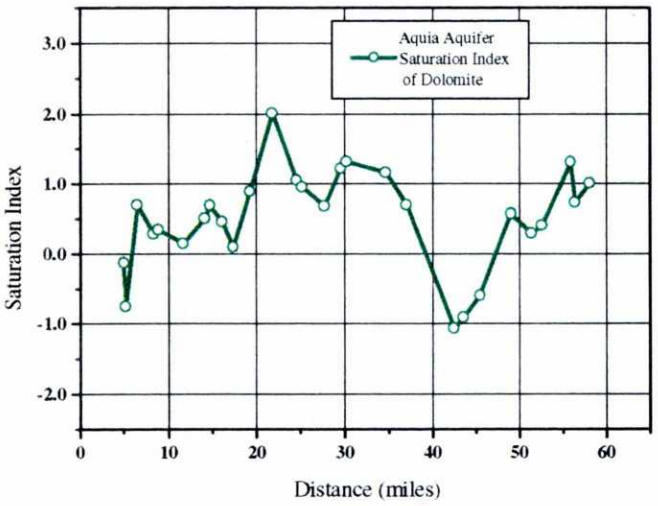
$$Na^+ + K\text{-}X = K^+ + Na\text{-}X$$
$$Na^+ + 0.5Ca\text{-}X = 0.5Ca^{2+} + Na\text{-}X$$
$$Na^+ + 0.5Mg\text{-}X_2 = 0.5Mg^{2+} + Na\text{-}X$$

Other chemical reactions include aqueous dissociation reactions and calcite and dolomite dissolution/precipitation. The following 9 species have been taken as primary aqueous species: H_2O, Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^-, HCO_3^-, SO_4^{-2} and H^+. EQ3 was used to generate the initial geochemical compositions. The following 14 secondary species were considered in our inverse model, their dissociation reactions being:

Aqueous dissociation reactions:	$Log_{10}(K)$ at 25 °C
$OH^- = H_2O - H^+$	13.995
$CO_3^{2-} = HCO_3^- - H^+$	10.329
$CO_2(aq) = HCO_3^- + H^+ - H_2O$	-6.3447
$CaHCO_3^+ = Ca^{2+} + HCO_3^-$	-1.0467
$MgHCO_3^+ = Mg^{2+} + HCO_3^-$	-1.0357
$CaCO_3(aq) = Ca^{2+} + HCO_3^- - H^+$	7.0017
$MgCO_3(aq) = Mg^{2+} + HCO_3^- - H^+$	7.3499
$NaHCO_3(aq) = Na^+ + HCO_3^-$	-0.1541
$CaSO_4(aq) = Ca^{2+} + SO_4^{2-}$	-2.1111
$MgSO_4(aq) = Mg^{2+} + SO_4^{2-}$	-2.309
$NaSO_4^- = Na^+ + SO_4^{2-}$	-0.082
$KSO_4^- = K^+ + SO_4^{2-}$	-0.8796
$HSO_4^- = SO_4^{2-} + H^+$	-1.9791
$NaCO_3^-(aq) = Na^+ + HCO_3^- - H^+$	9.8144
Mineral dissolution-precipitation:	
$Calcite = Ca^{2+} + HCO_3^- - H^+$	1.8487
$Dolomite = Ca^{2+} + Mg^{2+} + 2HCO_3^- - 2H^+$	2.5135

The saturation index (SI) of calcite, magnesite and dolomite at different points were calculated by using EQ3 with measured data of Na^+, K^+, Ca^{2+}, Mg^{2+}, HCO_3^- and pH. The results are plotted in Figure 5.6.3. According to Chapelle and Knobel (1983), the Aquia aquifer is recharged by meteoric water that is relatively low mineralized and subsaturated with respect to calcite. When this water enters the aquifer it rapidly dissolves calcite until equilibrium is reached. In Figure 5.6.3 the SI of calcite varies between 0.1 and –0.1 except at a few points. It means that water analyses are either close to saturation or slightly supersaturated (average SI = $\pm$ 0.1). Dolomite generally has saturation indexes between 0.1 and 1.1. Although it is supersaturated under natural conditions, it may dissolve at the recharge area because recharge water is depleted in Mg. Magnesite saturation indexes are well below zero (–0.5 to –1.5) meaning that it is clearly undersaturated. Figure 5.6.3 indicates the existence of low SI for all these minerals at a distance of 35 and 48 miles. It is probably due to the loss of Ca^{2+} ions to the cation exchange reaction and the leakage to the confining layers.





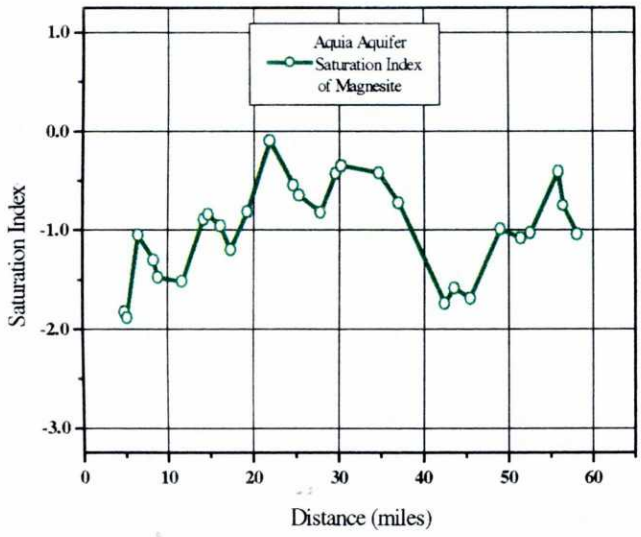


Figure 5.6.3. Calcite, dolomite and magnesite saturation indexes calculated from chemical data of Aquia aquifer using EQ3.

Calculations were performed at a temperature of 25 °C. The concentration distribution curves of Na^+, K^+, Ca^{2+}, Mg^{2+}, HCO_3^- and pH are fitted with measured concentration data. It was assumed that the initial water was brackish as a result of mixing of seawater with fresh water during the deposition of the overlying Marlboro clay, which is a brackish-water clay. Initial and boundary total dissolved concentrations of the components as proposed by Appelo (1994) are listed in Table 5.6.1.

Table 5.6.1. Initial and boundary total dissolved concentrations (mol/l) in Aquia aquifer (Appelo, 1994)

Component	Initial	Boundary
pH	6.8	7.57
Na^+	$8.74 \cdot 10^{-2}$	10^{-4}
K^+	$1.9 \cdot 10^{-3}$	$5 \cdot 10^{-5}$
Ca^{+2}	$4.38 \cdot 10^{-3}$	$1.4 \cdot 10^{-3}$
Mg^{+2}	$9.92 \cdot 10^{-3}$	0
Cl^-	0.1018	10^{-4}
HCO_3^-	$1.55 \cdot 10^{-2}$	$2.8 \cdot 10^{-3}$
SO_4^{-2}	$2.70 \cdot 10^{-4}$	0

In order to interpret field data, parameters were estimated in a step-wise manner with the following procedure:

1. Case 1: Hydraulic heads at the two ends of the aquifer were fixed to be 15.2 and 1.5 m based on Figure 5.6.1 (Chapelle and Drummond, 1983 and Appelo, 1994). Two conductivities, two dispersivities, CEC, three selectivity coefficients, initial concentrations of Na^+, and K^+, and boundary concentrations of Ca^{+2} and Mg^{+2} were estimated. Other conditions are the same as the basic model described above.

2. Case 2: Proton exchange, modeled as a standard cation exchange reaction, $Na^+ + H\text{-}X = H^+ + Na\text{-}X$, is also considered. The initial concentration of Mg^{+2}, and boundary concentration of K^+ are also estimated. Other conditions are the same as Case 1.
3. Case 3: Hydrodynamic conditions are modified and hydraulic head is fixed only at the downstream end to a value of 1.5 m. At the upstream boundary a constant recharge (164.38 l/d) is specified at x = 0. Uniform leakage into the confining aquitard is assumed along the second half of the aquifer. Appelo (1994) assumed that all the recharge leaks out from the second half of the aquifer. In order to test the plausibility of this assumption, we also estimate groundwater discharge. In this case, proton exchange was simulated with a surface sorption model by adding XOH as a primary species. Related surface dissociation reactions are expressed as,

$$XOH_2^+ = XOH + H^+ \qquad \log_{10}K = -7.4 \text{ (at 25 °C)}$$
$$XO^- = XOH - H^+ \qquad \log_{10}K = 9.24 \text{ (at 25 °C)}$$

4. Case 4: Similar to Case 3, but adding methane oxidation in order to try to improve the fit of HCO_3^-

$$Methane(aq) = HCO_3^- - 3H_2O + 9H^+ + 8e^- \qquad \log_{10}K = -27.655 \text{ (at 25 °C)}$$

In this case, additional initial and boundary concentrations are estimated.

5. Case 5: Similar to Case 4, but leaving out dolomite.

5.6.3. PARAMETER ESTIMATION

Case 1

In the first stage, two hydraulic conductivities K_1 and K_2, two dispersivities α_1 and α_2, three selectivity coefficients $K_{Na/K}$, $K_{Na/Ca}$ and $K_{Na/Mg}$, CEC (meq/100g), two initial concentrations $C_{0\,Na}$ and $C_{0\,K}$ (mol/l) and two boundary concentrations $C_{B\,Ca}$ and $C_{B\,Mg}$ (mol/l) were estimated.

Prior information data was obtained from Chapelle and Knobel (1983) and Appelo (1994) and is listed in Table 5.6.2. Estimation results are presented in Table 5.6.3, while the fit to measured data is plotted in Figure 5.6.3.

Table 5.6.2. Prior information data obtained from Appelo et al. (1990) and Chapelle and Knobel (1983) (* 10^{-9} is used for zero concentration values)

Parameter	Lower bound	Upper bound	Prior information
K_1 (m/a)	$2 \cdot 10^{3}$	$5 \cdot 10^{3}$	$3.5 \cdot 10^{3}$
K_2 (m/a)	$2 \cdot 10^{3}$	$5 \cdot 10^{3}$	$3.5 \cdot 10^{3}$
α_1 (m)	$2 \cdot 10^{3}$	$8 \cdot 10^{3}$	$3.22 \cdot 10^{3}$
α_2 (m)	$2 \cdot 10^{3}$	$8 \cdot 10^{3}$	$3.22 \cdot 10^{3}$
CEC	0.1	15	3.235
$K_{Na/K}$	0.001	1	0.112
$K_{Na/Ca}$	0.01	1	0.242
$K_{Na/Mg}$	0.01	1	0.405
$C_{0\,Na}$	$2 \cdot 10^{-2}$	$5 \cdot 10^{-1}$	$8.74 \cdot 10^{-2}$
$C_{0\,K}$	$5 \cdot 10^{-4}$	10^{-2}	$1.9 \cdot 10^{-3}$
$C_{B\,Ca}$	10^{-2}	$1.4 \cdot 10^{-3}$	$1.4 \cdot 10^{-3}$
$C_{B\,Mg}$	10^{-9*}	$5 \cdot 10^{-3}$	10^{-9*}

Table 5.6.3. Parameter estimates for Case 1 (objective function = $2.62 \cdot 10^4$)

Parameter	Initial guess	Estimated value	Estimation Variance	95% confidence interval	
				Lower bound	Upper bound
K_1 (m/a)	$3.5 \cdot 10^3$	$3.116 \cdot 10^3$	$6.218 \cdot 10^3$	$2.755 \cdot 10^3$	$3.478 \cdot 10^3$
K_2 (m/a)	$3.5 \cdot 10^3$	$3.756 \cdot 10^3$	$5.72 \cdot 10^3$	$3.409 \cdot 10^3$	$4.103 \cdot 10^3$
α_1 (m)	$3.22 \cdot 10^3$	$4.208 \cdot 10^3$	$1.272 \cdot 10^4$	$3.691 \cdot 10^3$	$4.725 \cdot 10^3$
α_2 (m)	$3.22 \cdot 10^3$	$5.973 \cdot 10^3$	$1.84 \cdot 10^4$	$5.351 \cdot 10^3$	$6.595 \cdot 10^3$
CEC	3.235	9	12.73	0.11	25.36
$K_{Na/K}$	0.112	0.0013	$5.149 \cdot 10^{-6}$	10^{-3}	$1.172 \cdot 10^{-2}$
$K_{Na/Ca}$	0.242	0.7441	$5.390 \cdot 10^{-2}$	10^{-2}	1.809
$K_{Na/Mg}$	0.405	0.0384	$2.492 \cdot 10^{-4}$	10^{-2}	0.0111
$C_{0\,Na}$	$8.74 \cdot 10^{-2}$	$3.098 \cdot 10^{-2}$	$3.37 \cdot 10^{-5}$	$4.36 \cdot 10^{-3}$	$5.76 \cdot 10^{-2}$
$C_{0\,K}$	$1.9 \cdot 10^{-3}$	$1.799 \cdot 10^{-3}$	$2 \cdot 10^{-7}$	$5 \cdot 10^{-4}$	$3.85 \cdot 10^{-3}$
$C_{B\,Ca}$	$1.4 \cdot 10^{-3}$	$1.196 \cdot 10^{-4}$	$2.308 \cdot 10^{-8}$	$4.993 \cdot 10^{-4}$	$1.893 \cdot 10^{-3}$
$C_{B\,Mg}$	10^{-9}	$3.328 \cdot 10^{-4}$	$1.705 \cdot 10^{-9}$	10^{-9}	$9.316 \cdot 10^{-4}$

Computed concentrations show significant deviations from measured data, especially for K^+, Mg^{2+} and HCO_3^- (Figure 5.6.4). Besides, most parameter estimates are far away from their prior information. The 95% confidence intervals of CEC, $K_{Na/K}$, $K_{Na/Ca}$ and $K_{Na/Mg}$ have a wide range (Table 5.6.3), which indicate that estimated parameters are highly uncertain, possibly because the flow conditions and geochemical reactions of the numerical model are not realistic.

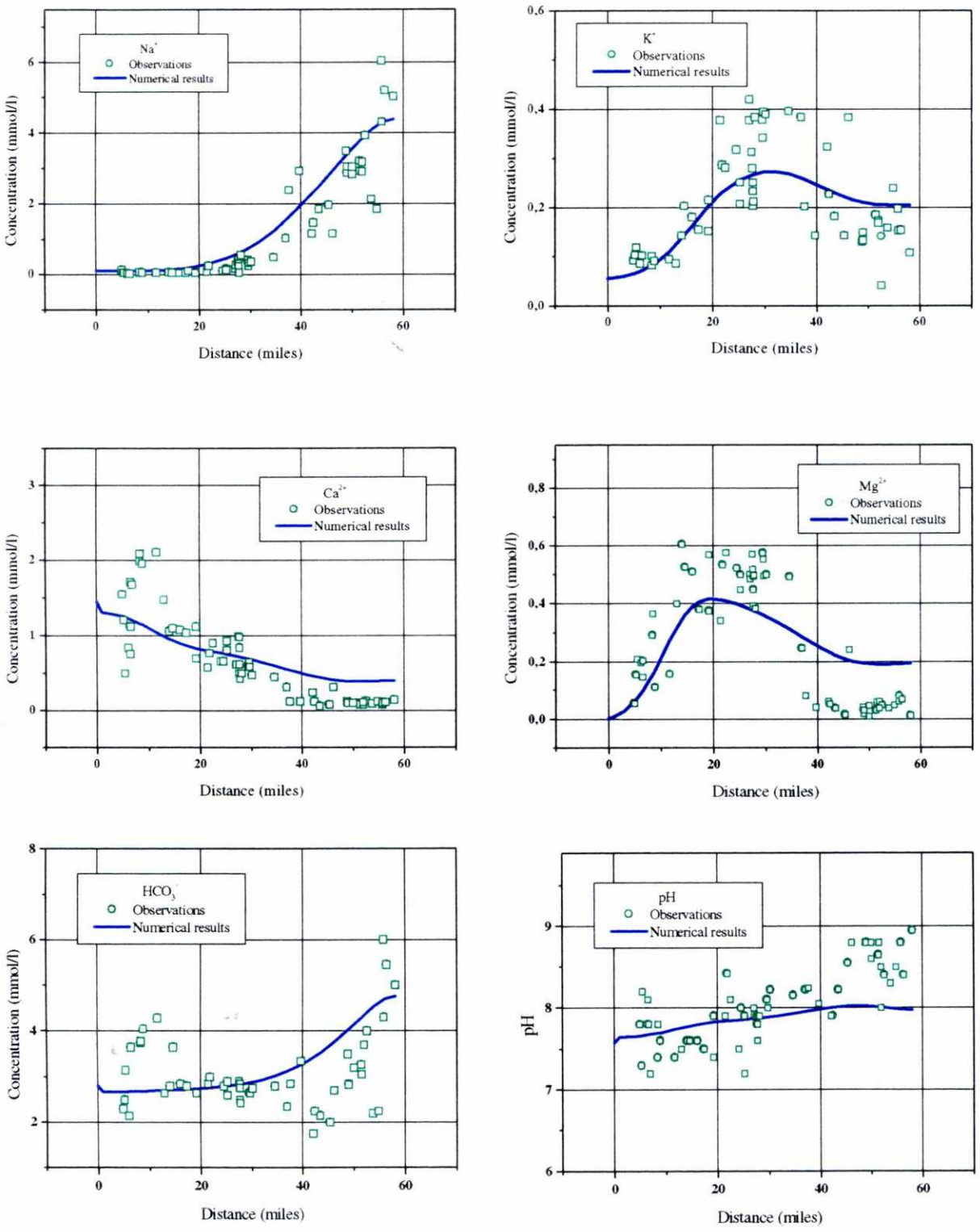


Figure 5.6.4. Measured (symbols) and computed (lines) concentrations at Aquia aquifer (Case 1)

Case 2

In case 2, proton exchange is added to the geochemical system and Mg initial and K boundary concentrations are also estimated. Taking the estimation results in Table 5.6.3 as initial parameter values, 14 parameters are estimated. Estimation results are listed in Table 5.6.4. The best fit for different components are plotted in Figure 5.6.5.

Table 5.6.4. Estimation results of Aquia aquifer for Case 2 (objective function = $3.33 \cdot 10^3$)

Parameter	Initial guess	Estimated value	Estimation Variance	95% confidence interval	
				Lower bound	Upper bound
K_1 (m/a)	$3.116 \cdot 10^3$	$3.453 \cdot 10^3$	$4.878 \cdot 10^1$	$3.419 \cdot 10^3$	$3.487 \cdot 10^3$
K_2 (m/a)	$3.756 \cdot 10^3$	$4.132 \cdot 10^3$	$4.553 \cdot 10^1$	$4.099 \cdot 10^3$	$4.164 \cdot 10^3$
α_1 (m)	$4.208 \cdot 10^3$	$4.199 \cdot 10^3$	$8.383 \cdot 10^1$	$4.155 \cdot 10^3$	$4.244 \cdot 10^3$
α_2 (m)	$5.973 \cdot 10^3$	$5.885 \cdot 10^3$	$1.173 \cdot 10^2$	$5.832 \cdot 10^3$	$5.937 \cdot 10^3$
CEC	9	8.071	$6.103 \cdot 10^{-6}$	8.059	8.083
$K_{Na/K}$	0.0013	0.1059	$2.338 \cdot 10^{-4}$	$3.152 \cdot 10^{-2}$	0.1803
$K_{Na/Ca}$	0.7441	0.02394	$5.037 \cdot 10^{-6}$	$1.302 \cdot 10^{-2}$	$3.487 \cdot 10^{-2}$
$K_{Na/Mg}$	0.0384	0.1358	$5.121 \cdot 10^{-5}$	0.101	0.1707
$C_{0\,Na}$	$3.098 \cdot 10^{-2}$	$2.031 \cdot 10^{-2}$	$8.533 \cdot 10^{-9}$	$1.986 \cdot 10^{-2}$	$2.076 \cdot 10^{-2}$
$C_{0\,K}$	$1.799 \cdot 10^{-3}$	$5.845 \cdot 10^{-4}$	$9.300 \cdot 10^{-11}$	$5.375 \cdot 10^{-4}$	$6.314 \cdot 10^{-4}$
$C_{0\,Mg}$	$1.4 \cdot 10^{-3}$	$1.819 \cdot 10^{-3}$	$7.72 \cdot 10^{-10}$	$1.683 \cdot 10^{-3}$	$1.954 \cdot 10^{-3}$
$C_{B\,K}$	$5 \cdot 10^{-5}$	$5.824 \cdot 10^{-5}$	$2.738 \cdot 10^{-11}$	$3.277 \cdot 10^{-5}$	$8.370 \cdot 10^{-5}$
$C_{B\,Ca}$	$1.196 \cdot 10^{-4}$	$1.455 \cdot 10^{-3}$	$6.484 \cdot 10^{-11}$	$1.416 \cdot 10^{-3}$	$1.495 \cdot 10^{-3}$
$C_{B\,Mg}$	$3.328 \cdot 10^{-4}$	$5.172 \cdot 10^{-6}$	$4.911 \cdot 10^{-12}$	$1.000 \cdot 10^{-9}$	$1.596 \cdot 10^{-5}$

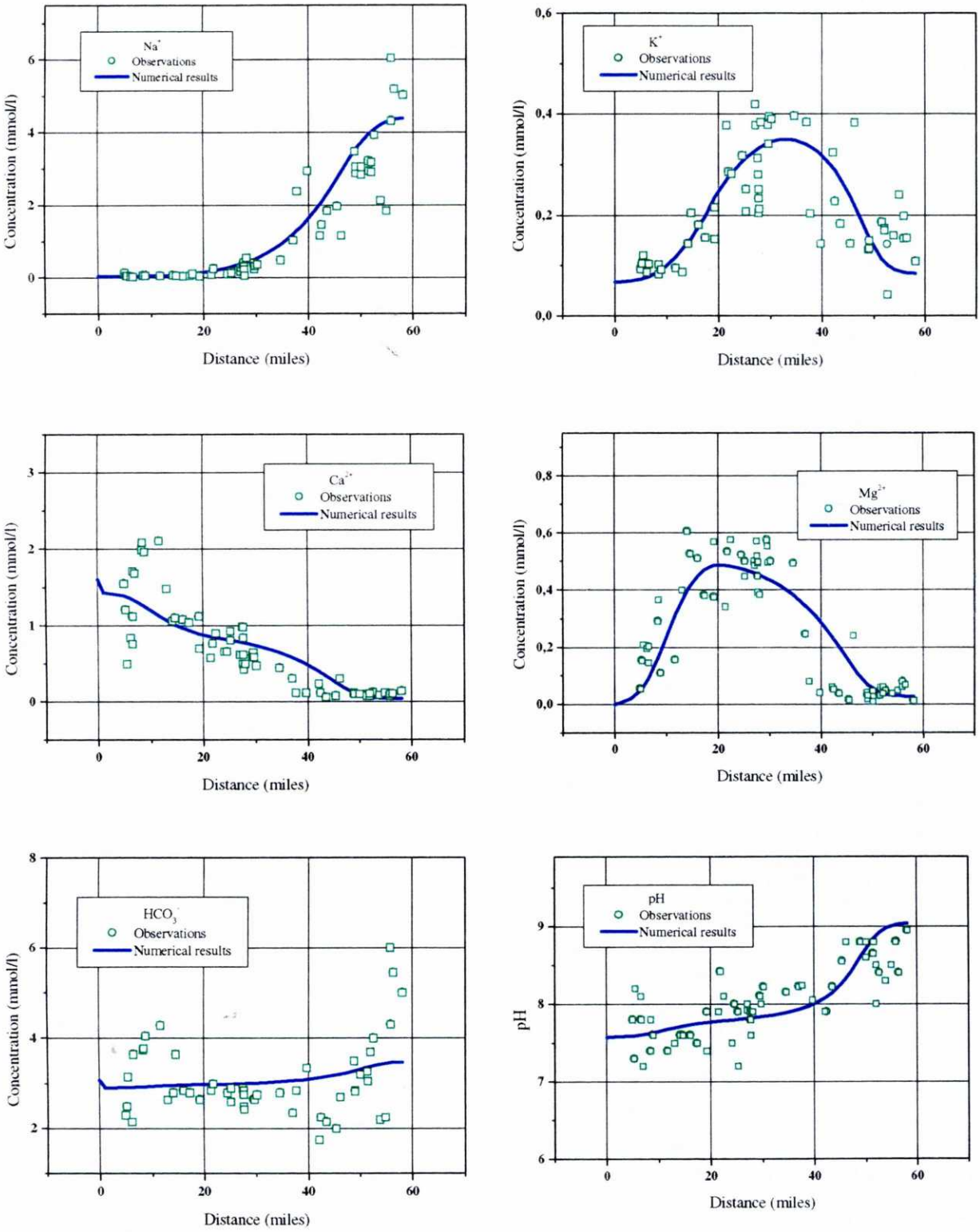


Figure 5.6.5. Measured (symbols) and computed (lines) concentrations at Aquia aquifer (Case 2)

Adding proton exchange and estimating more initial and boundary concentrations results in a better fit (compare Figures 5.6.5 and 5.6.4) and leads to a significant improvement in objective function which decreases from $2.62 \cdot 10^4$ to $3.33 \cdot 10^3$. The fit of pH, K and Mg improves noticeably although those of HCO_3^- and other cations are still bad. So, it is necessary to adjust flow conditions and some chemical reactions.

Case 3

In case 3, proton exchange is simulated as a surface complexation reaction. Related chemical parameters were obtained from Xu (1996). Sorption site concentration was taken equal to 0.05 mol/dm^3 and specific adsorbent surface per unit volume of solution is 1000 m^2/kg. XOH is added as a sorption primary species. Its initial concentration was estimated while its boundary concentration is assumed to be very low (10^{-9} mol/l). Conductivities and dispersivities were fixed to the values listed in Table 5.6.4.

After modifying the hydraulic conditions according to the numerical model of Appelo (1994), it was assumed that all of the recharge leaks out from the second half of the aquifer, so at the downstream end there is no discharge at all. However, a few trial runs indicated that this assumption was not reasonable. In order to determine how much water leaks to the confining layers over the second half of the flow tube, discharge rate was also estimated. In this case, the following parameters were estimated: CEC, three selectivity coefficients, node discharge rate, and four initial and three boundary concentrations. Taking the estimation results in Table 5.6.3 as prior information and initial parameter values, 12 parameters were estimated. Estimation results are listed in Table 5.6.4. Curve fits for different components are plotted in Figure 5.6.6.

Table 5.6.4. Estimation results of Aquia aquifer for Case 3 (objective function = $1.53 \cdot 10^3$)

Parameter	Initial guess	Estimated value	Estimation Variance	95% confidence interval	
				Lower bound	Upper bound
CEC	9	0.5459	$9.04 \cdot 10^{-6}$	0.5321	0.5597
$K_{Na/K}$	0.0013	0.1134	$4.244 \cdot 10^{-4}$	0.0189	0.2079
$K_{Na/Ca}$	0.7441	0.5969	$5.149 \cdot 10^{-5}$	0.564	0.6298
$K_{Na/Mg}$	0.0384	0.5642	$4.548 \cdot 10^{-5}$	0.5333	0.5951
Discharge (l/a)	$1.8 \cdot 10^3$	$1.079 \cdot 10^3$	39.29	$1.05 \cdot 10^3$	$1.107 \cdot 10^3$
$C_{0\,Na}$	$2.031 \cdot 10^{-2}$	$5.203 \cdot 10^{-2}$	$2.725 \cdot 10^{-7}$	$4.963 \cdot 10^{-2}$	$5.442 \cdot 10^{-2}$
$C_{0\,K}$	$5.845 \cdot 10^{-4}$	$9.664 \cdot 10^{-4}$	$6.26 \cdot 10^{-10}$	$8.517 \cdot 10^{-4}$	$1.081 \cdot 10^{-3}$
$C_{0\,Mg}$	$1.819 \cdot 10^{-3}$	$3.894 \cdot 10^{-3}$	$8.968 \cdot 10^{-9}$	$3.459 \cdot 10^{-3}$	$4.328 \cdot 10^{-3}$
$C_{0\,XOH}$	$4.715 \cdot 10^{-2}$	$4.088 \cdot 10^{-2}$	$2.547 \cdot 10^{-3}$	10^{-2}	$2.723 \cdot 10^{-1}$
$C_{B\,K}$	$5 \cdot 10^{-5}$	$5.682 \cdot 10^{-5}$	$9.707 \cdot 10^{-11}$	$1.164 \cdot 10^{-5}$	$1.02 \cdot 10^{-4}$
$C_{B\,Ca}$	$1.4 \cdot 10^{-3}$	$1.561 \cdot 10^{-3}$	$1.648 \cdot 10^{-10}$	$1.502 \cdot 10^{-3}$	$1.62 \cdot 10^{-3}$
$C_{B\,Mg}$	10^{-9}	10^{-9}	$3.868 \cdot 10^{-12}$	0	$9.019 \cdot 10^{-6}$

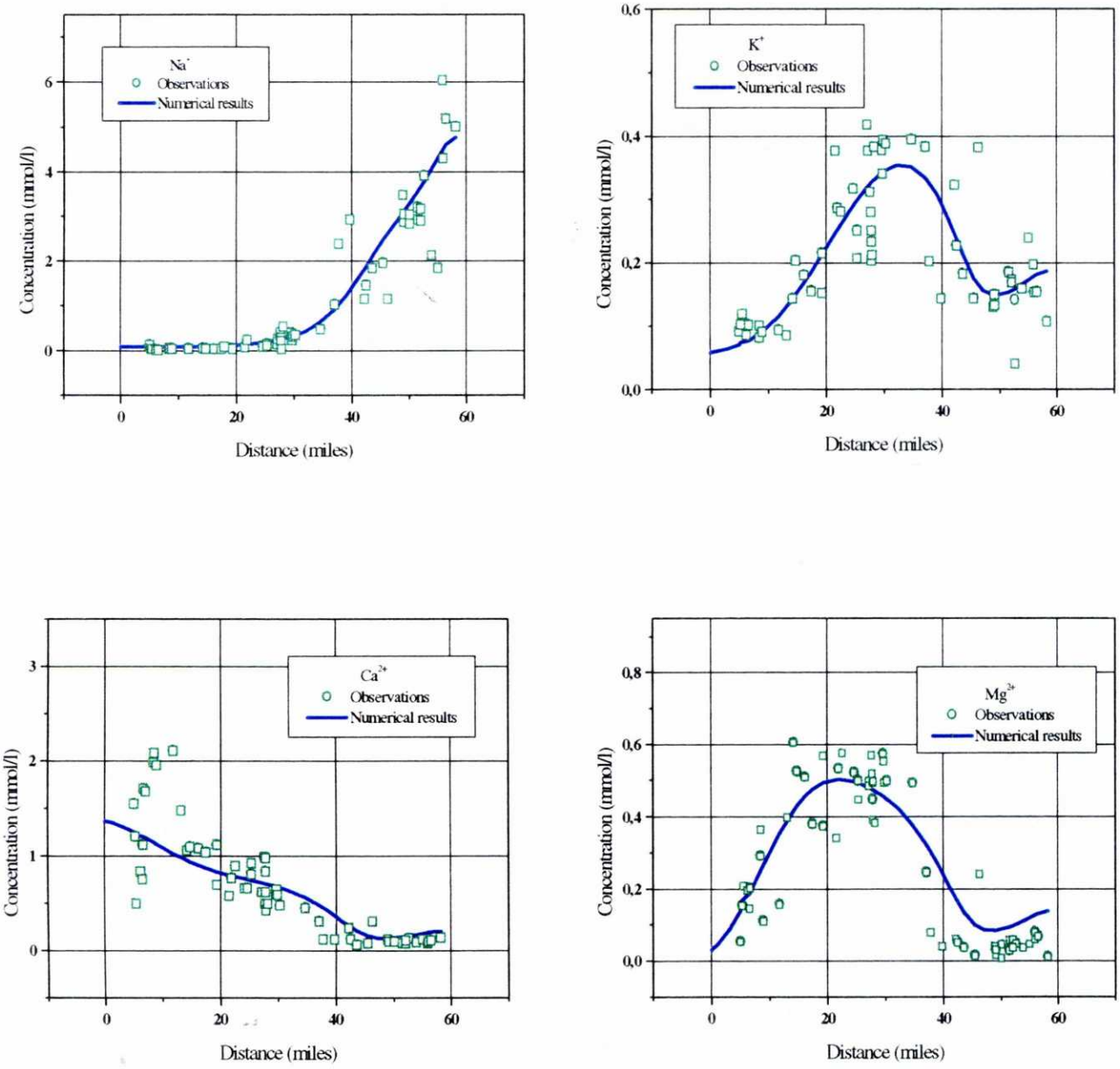


Figure 5.6.6A. Measured (symbols) and computed (lines) concentrations at Aquia aquifer (Case 3)

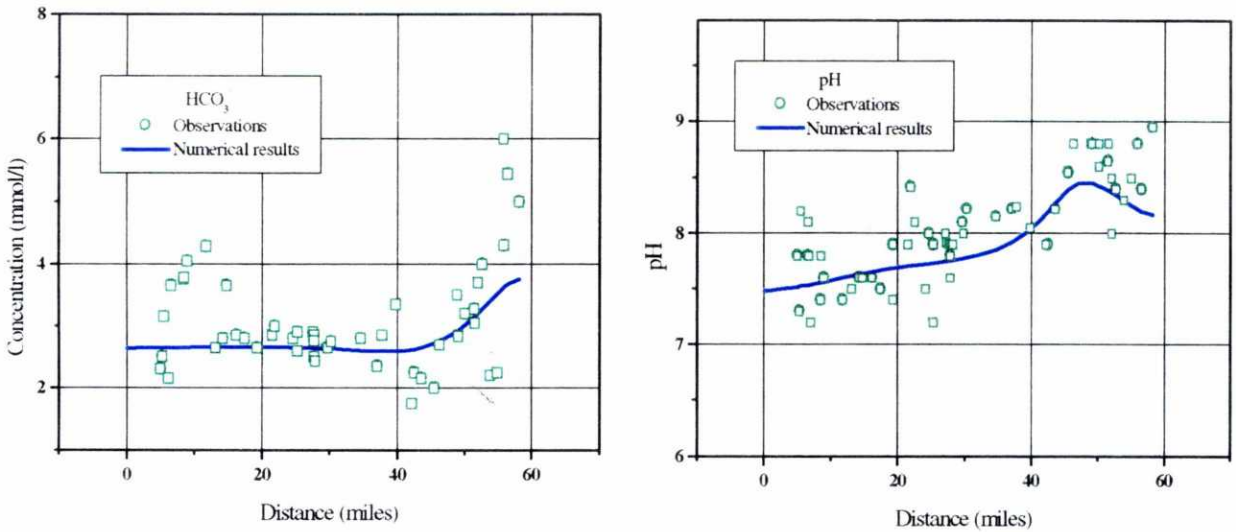


Figure 5.6.6B. Measured (symbols) and computed (lines) concentrations at Aquia aquifer (Case 3)

The objective function dereases from $3.33 \cdot 10^3$ in Case 2 to $1.53 \cdot 10^3$ in Case 3 and the curve fits (Figure 5.6.6) are better than those of Case 2. The estimated parameters such as selectivity coefficients, initial and boundary concentrations are more reasonable and closer to their prior information.

The estimated discharge rate in Table 5.6.4 is the value at one node. There is discharge evenly distributed along 32 nodes over the second half of the flow tube. Then, the total discharge rate is $1079 \times 32 = 34528$ l/a. The recharge rate at the upstream end is 60000 l/a (164.38 l/d). So, there is 58% of the recharge water leaked to the confining layers over the second half of the flow tube. It means that there is still 42% of the recharge water flow out from the downstream end in this case. However, the fit of HCO_3^- is still not satisfactory because its concentration at the downstream end is much lower than the measured value. It is necessary to set up another case for adjusting further the geochemical system. Therefore, the above conclusion about the discharge rate still needs to be proved further.

Case 4

In Case 4, methane redox reaction is taken into account in order to improve the fit of HCO_3^-. Since the prior estimate of the Mg^{2+} boundary concentration is 10^{-9} mol/l, according with that of Appelo (1994), in this case, it is fixed while the boundary concentration of HCO_3^- is estimated. With the inclusion of redox reaction into the geochemical system, $O_{2\,(aq)}$ is added as primary species and its initial and bounadry concentrations are obtained after a few trial runs. The initial pE is taken equal to -5 and the boundary pE is 0. 12 parameters are estimated. Estimation results are listed in Table 5.6.5. The best fit for different components are plotted in Figure 5.6.7.

Table 5.6.5. Estimation results of Aquia aquifer for Case 4 (objective function = $1.15 \cdot 10^3$)

Parameter	Initial guess	Estimated value	Estimation Variance	95% confidence interval	
				Lower bound	Upper bound
CEC	0.5459	0.4837	$8.342 \cdot 10^{-9}$	0.4833	0.4841
$K_{Na/K}$	0.1134	0.1080	$5.095 \cdot 10^{-7}$	0.1047	0.1113
$K_{Na/Ca}$	0.5969	0.6058	$4.290 \cdot 10^{-8}$	0.6048	0.6067
$K_{Na/Mg}$	0.5642	0.6388	$5.464 \cdot 10^{-8}$	0.6377	0.6399
Discharge (l/a)	$1.8 \cdot 10^3$	$1.332 \cdot 10^3$	65.01	$1.296 \cdot 10^3$	$1.368 \cdot 10^3$
$C_{0\,Na}$	$5.203 \cdot 10^{-2}$	$5.794 \cdot 10^{-2}$	$3.142 \cdot 10^{-10}$	$5.786 \cdot 10^{-2}$	$5.802 \cdot 10^{-2}$
$C_{0\,K}$	$9.664 \cdot 10^{-4}$	$1.074 \cdot 10^{-3}$	$1.167 \cdot 10^{-12}$	$1.069 \cdot 10^{-3}$	$1.079 \cdot 10^{-3}$
$C_{0\,Mg}$	$3.894 \cdot 10^{-3}$	$4.149 \cdot 10^{-3}$	$1.226 \cdot 10^{-11}$	$4.133 \cdot 10^{-3}$	$4.165 \cdot 10^{-3}$
$C_{0\,XOH}$	$4.088 \cdot 10^{-2}$	$4.383 \cdot 10^{-2}$	$3.747 \cdot 10^{-5}$	$1.576 \cdot 10^{-2}$	$7.190 \cdot 10^{-2}$
$C_{B\,K}$	$5.682 \cdot 10^{-5}$	$5.235 \cdot 10^{-5}$	$1.179 \cdot 10^{-13}$	$5.078 \cdot 10^{-5}$	$5.393 \cdot 10^{-5}$
$C_{B\,Ca}$	$1.561 \cdot 10^{-3}$	$1.615 \cdot 10^{-3}$	$2.350 \cdot 10^{-13}$	$1.613 \cdot 10^{-3}$	$1.617 \cdot 10^{-3}$
$C_{B\,HCO3}$	$2.8 \cdot 10^{-3}$	$3.068 \cdot 10^{-3}$	$7.337 \cdot 10^{-13}$	$3.064 \cdot 10^{-3}$	$3.072 \cdot 10^{-3}$

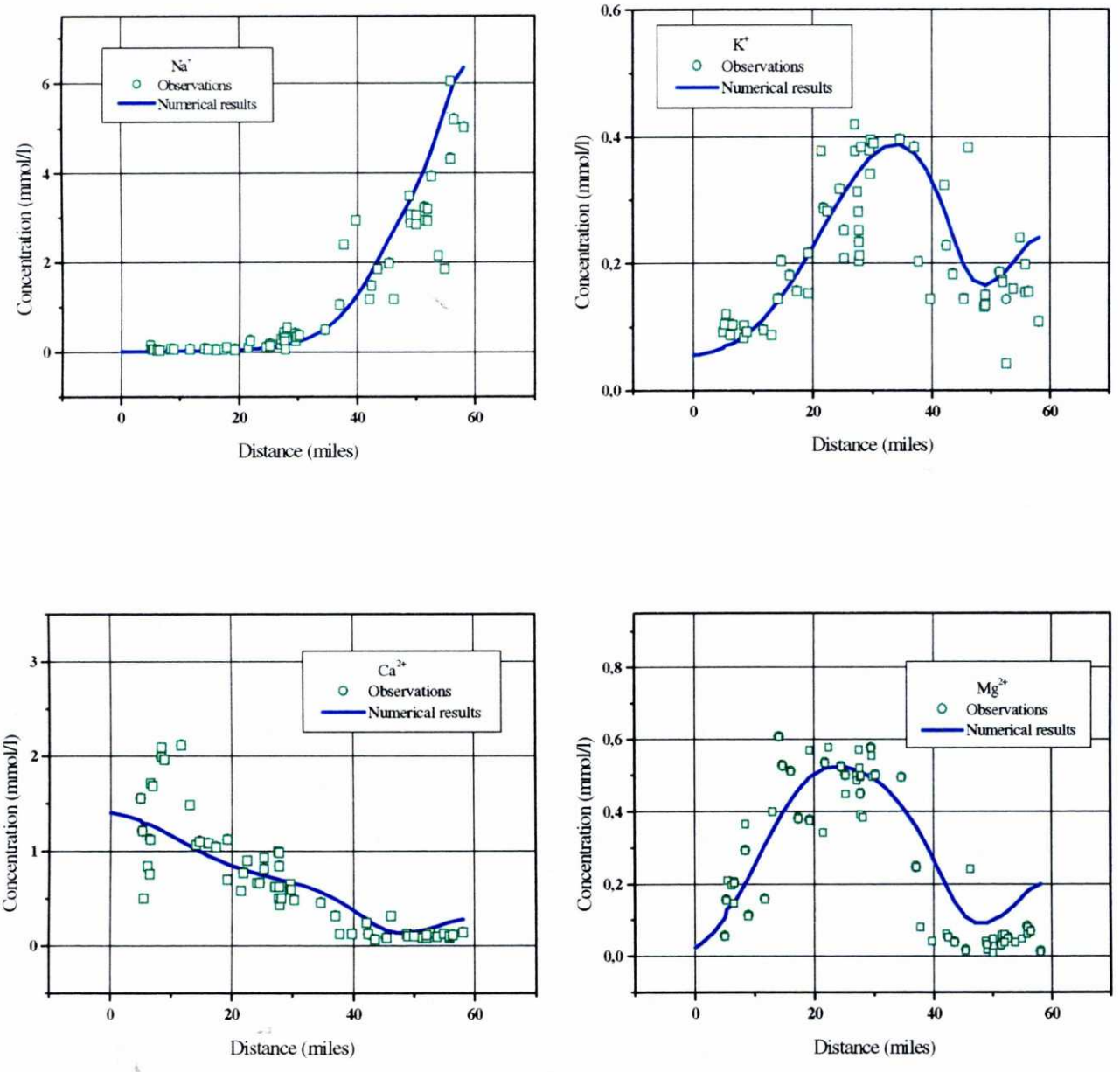


Figure 5.6.7A. Measured (symbols) and computed (lines) concentrations at Aquia aquifer (Case 4)

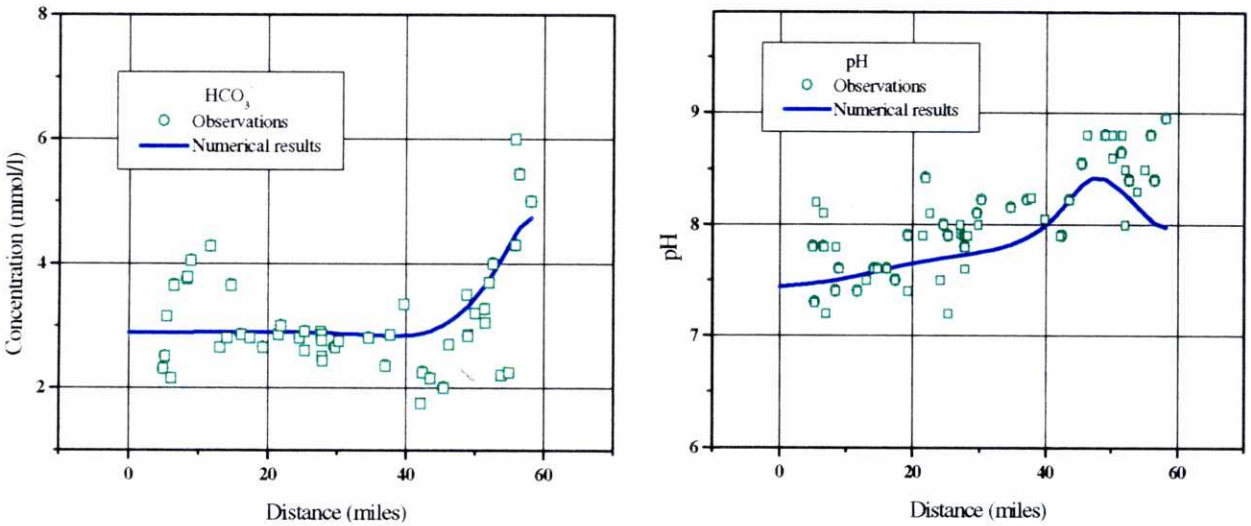


Figure 5.6.7B. Measured (symbols) and computed (lines) concentrations at Aquia aquifer (Case 4)

In Table 5.6.5, the estimated nodal discharge rate is $1.332 \cdot 10^3$ l/a. Then the total discharge to the confining layers is $1332 \times 32 = 42624$ l/a. The recharge rate at the upstream end is still 60000 l/a. So, 71% of the recharge water leaks to the confining layers over the second half of the aquifer. It means that there is still 29% of the recharge water flow out from the downstream end.

By adding methane redox reaction into the geochemical system and more leakage to the confining layers, the concentration of HCO_3^- at the downstream end increases to 4.743 mmol/l. The correponding fit of HCO_3^- is better than before. However, the value of objective function does not reduce a lot and CEC and selectivities $K_{Na/Ca}$ and $K_{Na/Mg}$ are far away from their prior estimates. Case 5 is designed to improve the estimated parameters.

Case 5

In Case 5, dolomite is taken out from the geochemical system and the initial weights for the prior parameter information of CEC and selectivities $K_{Na/Ca}$ and $K_{Na/Mg}$ are increased. Estimated parameters are the same as those in Case 4. Estimation results are listed in Table 5.6.6. The best fit for different components is plotted in Figure 5.6.8. Concentrations of exchanged cations are shown in Figure 5.6.9. Dissolved Cl^- is plotted in Figure 5.6.10 while calcite dissolution is shown in Figure 5.6.11.

Table 5.6.6. Estimation results of Aquia aquifer for Case 5 (objective function = $1.37 \cdot 10^3$)

Parameter	Initial guess	Estimated value	Estimation Variance	95% confidence interval	
				Lower bound	Upper bound
CEC	3.235	3.09	$5.622 \cdot 10^{-2}$	2.003	4.178
$K_{Na/K}$	0.112	0.1282	$2.263 \cdot 10^{-3}$	$2 \cdot 10^{-2}$	0.3463
$K_{Na/Ca}$	0.242	0.0845	$6.012 \cdot 10^{-2}$	10^{-3}	0.1209
$K_{Na/Mg}$	0.405	0.2888	$1.444 \cdot 10^{-2}$	10^{-2}	0.8397
Discharge (l/a)	$1.8 \cdot 10^3$	$1.38 \cdot 10^3$	$1.447 \cdot 10^2$	$1.325 \cdot 10^3$	$1.435 \cdot 10^3$
$C_{0\,Na}$	$8.74 \cdot 10^{-2}$	$3.175 \cdot 10^{-2}$	$2.765 \cdot 10^{-6}$	$2.413 \cdot 10^{-2}$	$3.938 \cdot 10^{-2}$
$C_{0\,K}$	$1.9 \cdot 10^{-3}$	$6.758 \cdot 10^{-4}$	$1.209 \cdot 10^{-9}$	$5.164 \cdot 10^{-4}$	$8.353 \cdot 10^{-4}$
$C_{0\,Mg}$	$9.92 \cdot 10^{-3}$	$1.679 \cdot 10^{-3}$	$9.608 \cdot 10^{-8}$	$1.230 \cdot 10^{-3}$	$2.129 \cdot 10^{-3}$
$C_{0\,XOH}$	$4.383 \cdot 10^{-2}$	$5 \cdot 10^{-2}$	0.5117	10^{-2}	0.333
$C_{B\,K}$	$5 \cdot 10^{-5}$	$5.648 \cdot 10^{-5}$	$3.367 \cdot 10^{-10}$	10^{-5}	$1.406 \cdot 10^{-4}$
$C_{B\,Ca}$	$1.4 \cdot 10^{-3}$	$1.579 \cdot 10^{-3}$	$6.916 \cdot 10^{-10}$	$1.459 \cdot 10^{-3}$	$1.7 \cdot 10^{-3}$
$C_{B\,HCO3}$	$2.8 \cdot 10^{-3}$	$2.977 \cdot 10^{-3}$	$1.238 \cdot 10^{-9}$	$2.815 \cdot 10^{-3}$	$3.138 \cdot 10^{-3}$

In Table 5.6.6 the estimated CEC (3.09 meq/100g solid) is very close to prior parameter information (3.235 meq/100g solid) and other estimated parameter values such as selectivities, initial and boundary concentrations are also close to the puplished values (Appelo, 1994).

Uncertainties in estimated parameters measured by their 95% confidence intervals (Table 5.6.6) are small except parameter $C_{0\,XOH}$. Even through the objective function value of this case is slightly greater than that of Case 4, the estimated results are reasonable and accepted.

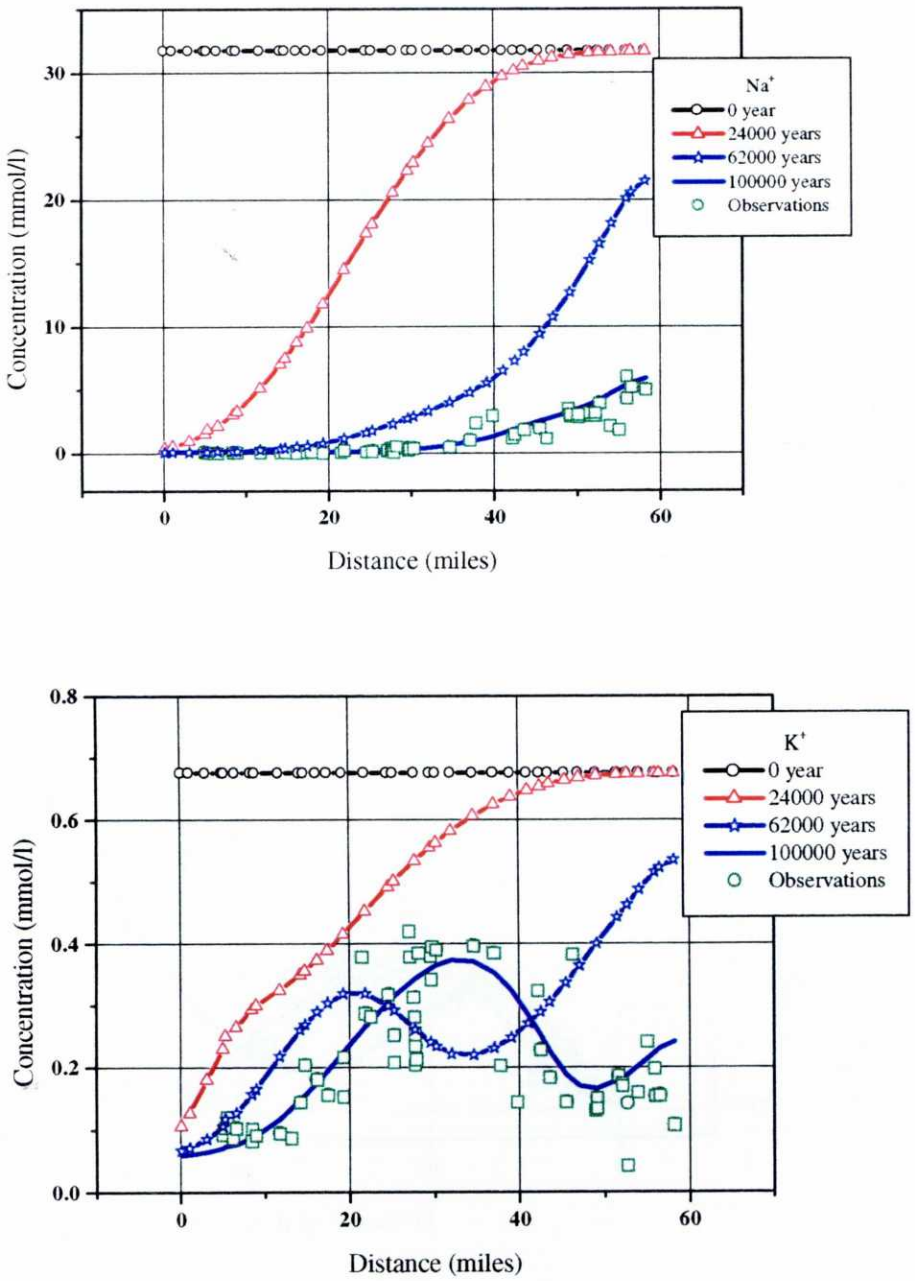


Figure 5.6.8A. Measured (square) and computed (at different times) concentrations of Na and K at the Aquia aquifer (Case 5)

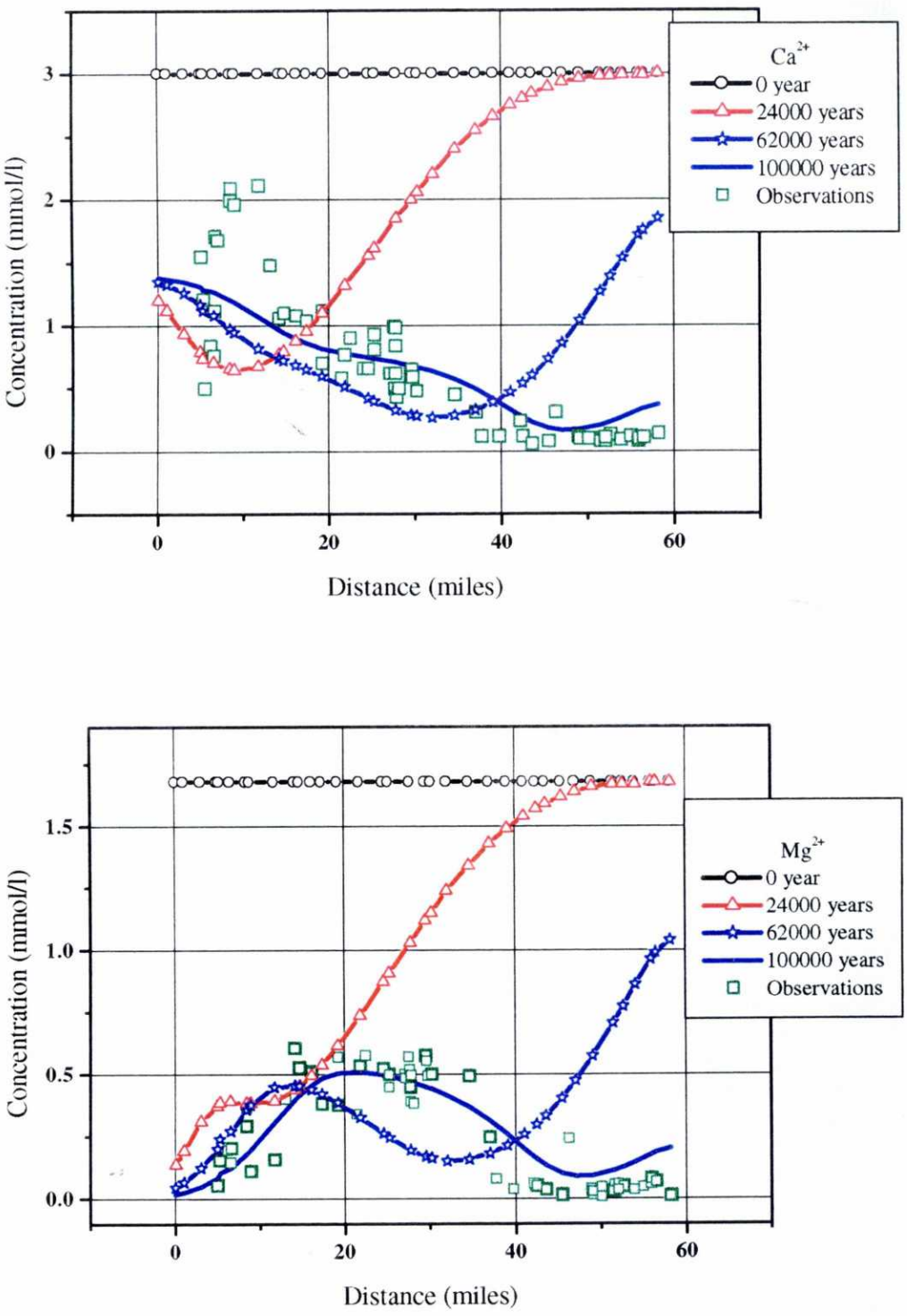


Figure 5.6.8B. Measured (square) and computed (at different times) concentrations of Ca and Mg at the Aquia aquifer (Case 5)

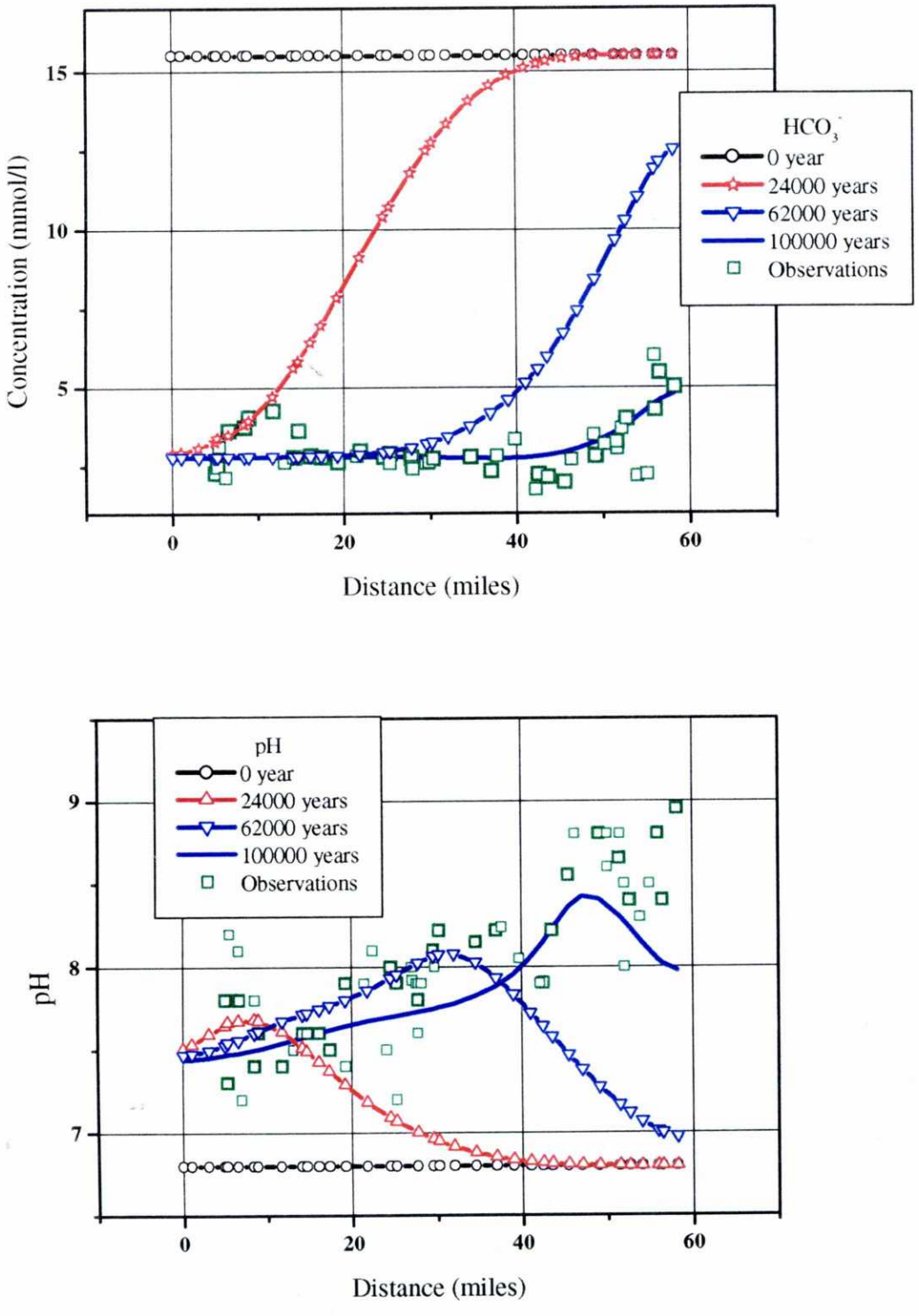


Figure 5.6.8C. Measured (square) and computed (at different times) concentrations of HCO_3 and pH at the Aquia aquifer (Case 5)

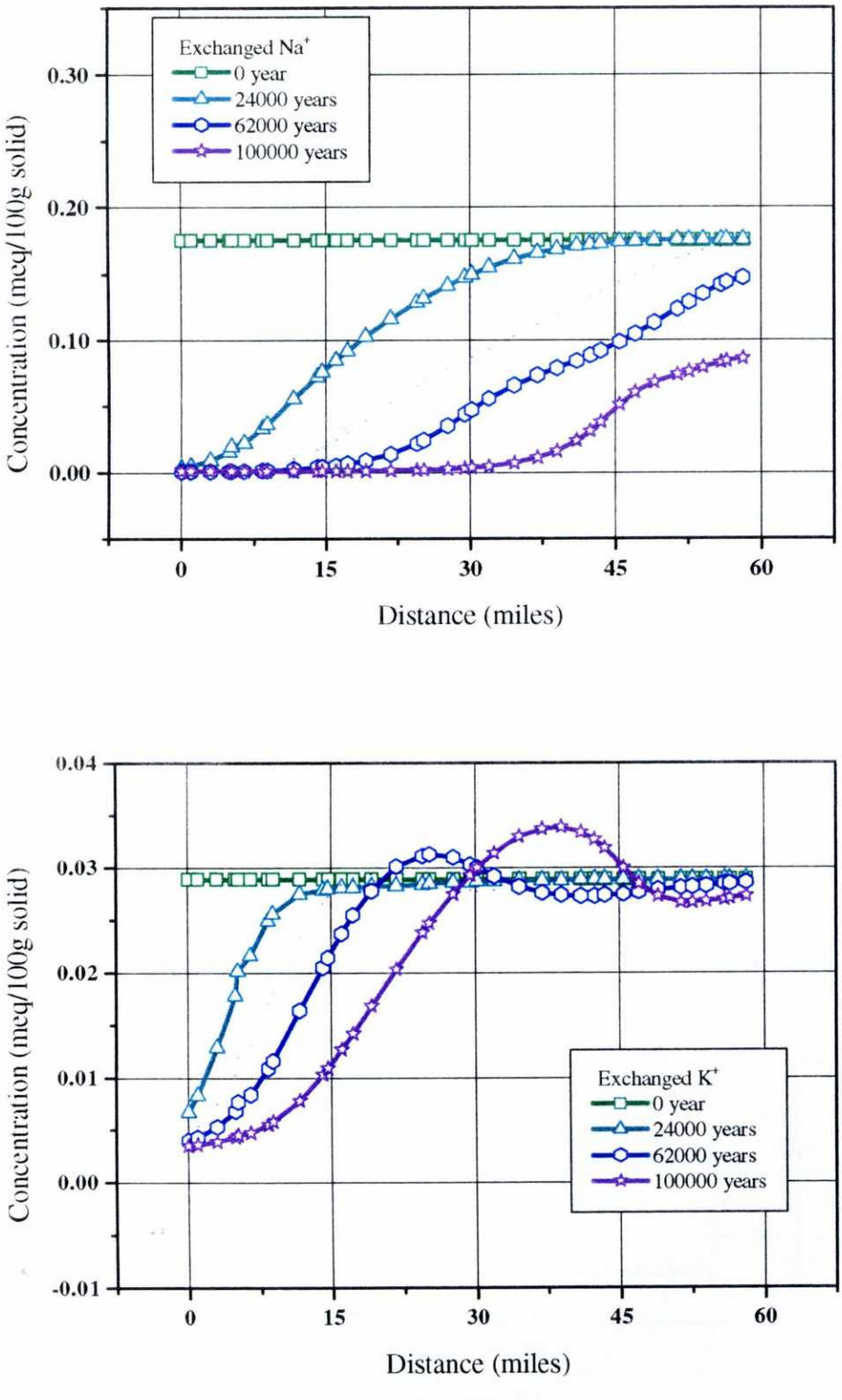


Figure 5.6.9A. Computed exchanged cation concentrations of Na and K (at different times) at the Aquia aquifer (Case 5)

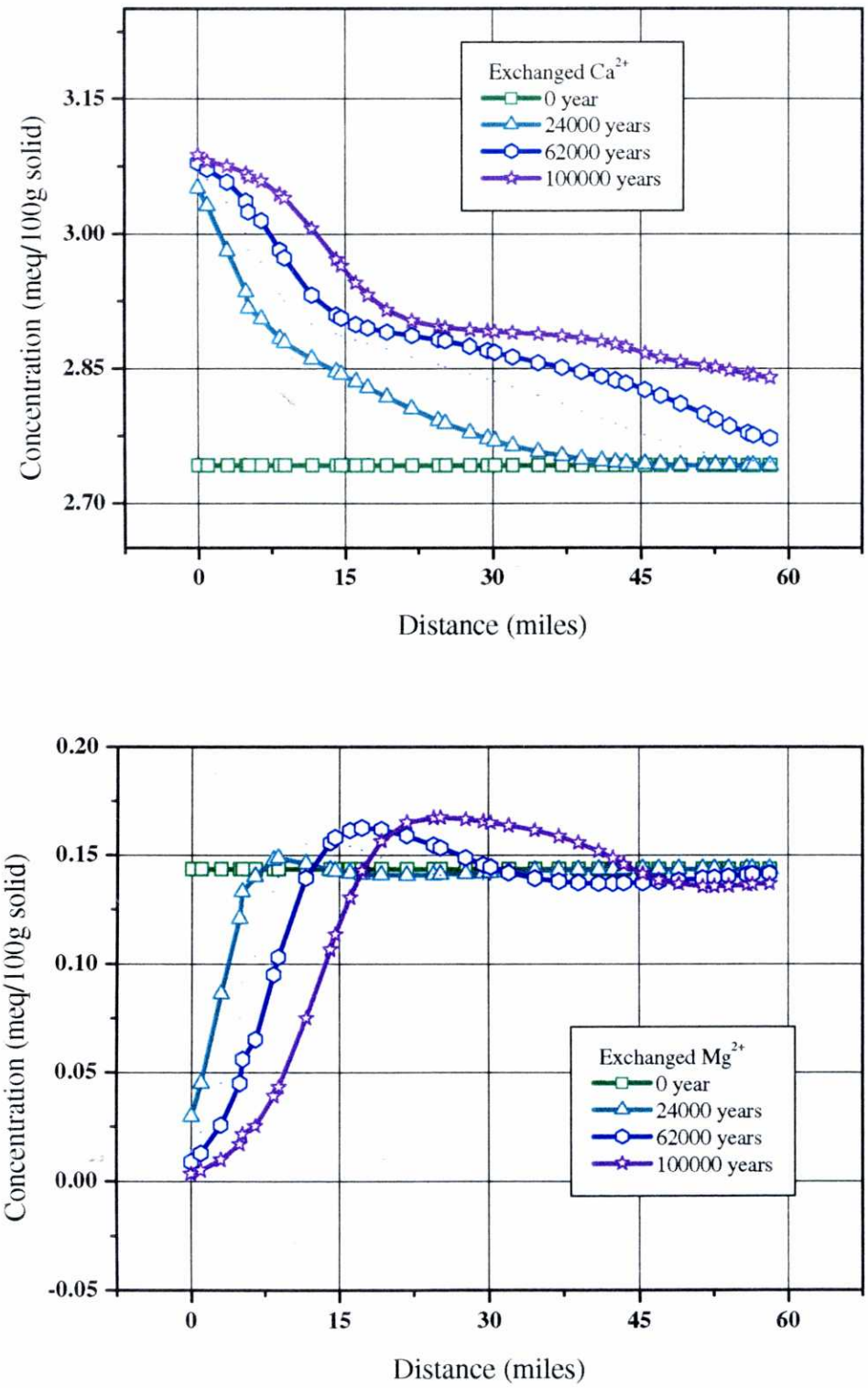


Figure 5.6.9B. Computed exchanged cation concentrations of Ca and Mg (at different times) at the Aquia aquifer (Case 5)

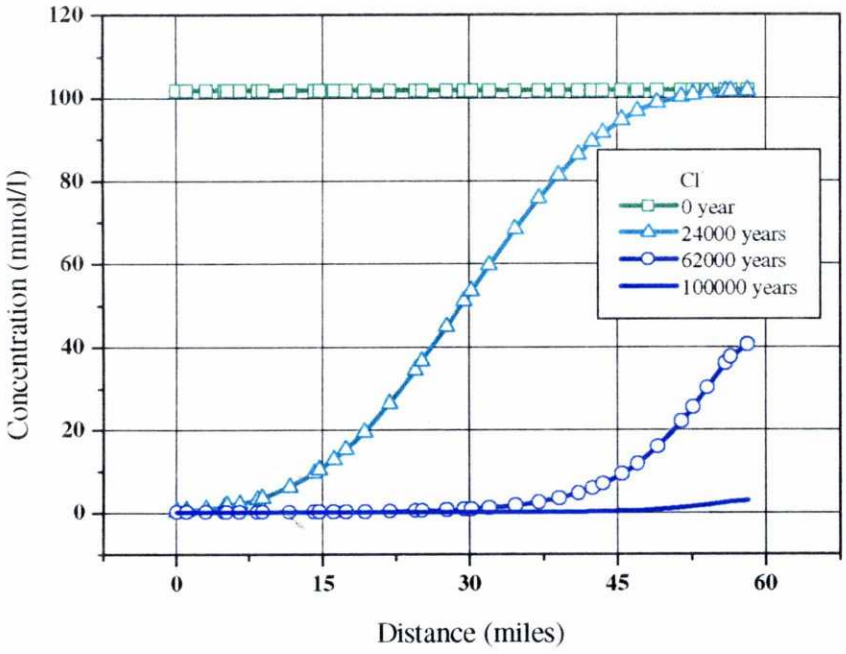


Figure 5.6.10. Computed aqueous Cl^- concentrations (at different times) at the Aquia aquifer (Case 5)

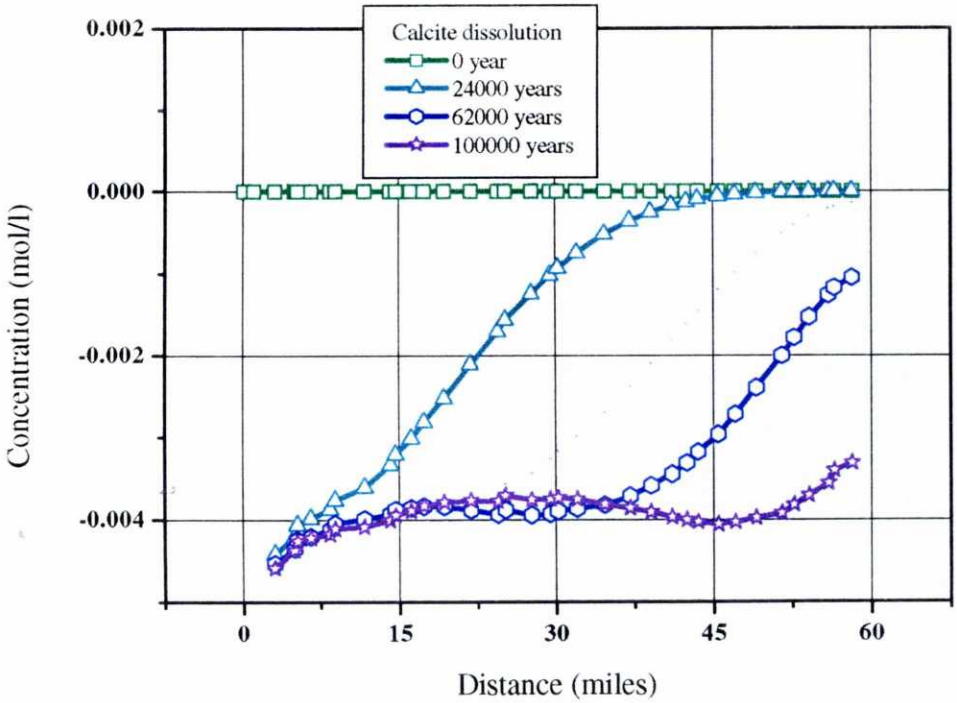


Figure 5.6.11. Calcite dissolution (negative values) (at different times) at the Aquia aquifer (Case 5)

Figures 5.6.8, 5.6.9 and 5.6.10 illustrate the chromatographic separation of ions during displacement of freshwater in Aquia aquifer. Figure 5.6.8 shows the evolution of species and the final fit (at 100000 years) to measured data. Computed values match measured data for all six components except at the downstream end of the flowpath where computed concentrations of Ca^{2+} and Mg^{2+} are higher than measured data. This discrepancy could be due to a too low groundwater discharge from the downstream end. The estimated nodal leakage is $1.38 \cdot 10^3$ l/a, then the total leakage to the confining layers is $1380 \times 32 = 44160$ l/a. The recharge rate at the upstream end is still 60000 l/a. So, there is 74% of the recharge water leaked to the confining layers and 26% of the recharge water flow out from the downstream end. Therefore, water flow velocity in the downstream end is small. It leads to a week cation exchange reaction in this part and the tails of the concentration curves of Ca^{2+} and Mg^{2+} are higher than measured data. Appelo (1994) used a no-flow boundary at the downstream end and he got much higher tails of concentration curves of K^+, Ca^{2+} and Mg^{2+}.

Figure 5.6.9 indicates that in this case Ca^{2+} is the displacing cation, which has a higher affinity for the exchanger than Na^+, K^+ and Mg^{2+}. The concentration of exchanged Ca^{2+} increases with time while those of Na^+, K^+ and Mg^{2+} decrease gradually. An important aspect is that appreciable amounts of K^+ and Mg^{2+} are adsorbed on the exchanger in addtion to Na^+. An order sequence develops for the cations, whereby first Na^+, then K^+, and finally Mg^{2+} is displaced. Figure 5.6.10 shows the conservative ion, Cl^-, which is flushed from the aquifer. Its concentration reduces greatly with evolution of time.

Figure 5.6.11 shows the evolution of the cumulative amount of calcite dissolution. Negative values indicate calcite dissolution. Comparing to Figure 5.6.8C, an increasing salinity is accompanied by a decrease of pH, and this enhances calcite dissolution.

5.6.4. SUMMARY

This case study dealt with hydrogeochemical data interpretation of Aquia aquifer (Maryland). By using INVERSE-CORE^{2D}, the inverse problem has been solved in five stages. Estimated parameters include conductivity, dispersivity, discharge rate, selectivity coefficient, CEC and initial and boundary concentrations. The main conclusions of our inverse analyses are:

1. The chromatographic separation of ions during displacement of different water types in Aquia aquifer has been demonstrated for the case of freshwater displacement. The observed water chemistry displays the expected sequential peaks of Mg^{2+}, K^+ and Na^+ and has been modeled accurately. Chromatographic patterns should be considered as one of the common features of groundwater chemistry. Ion exchange reactions are always dynamic by their nature and will give heterogeneous patterns of ion concentrations and side reactions along a flow line. The dynamic nature of ion exchange allows for a link with historic flow conditions and can be used to obtain paleohydrological information about the aquifer system.
2. Estimated CEC and selectivity coefficients almost coincide with published values (Appelo et al., 1990 and Appelo, 1994). However, these parameters are very sensitive to changes in the objective function. A little change of the parameters can lead to a substantial improvement of the curve fit.
3. Estimated boundary concentrations, conductivities and dispersivities are similar to what Appelo (1994) used. But, Appelo (1994) assumed all the recharge from upstream leaked out from the second half of the aquifer, so at the downstream end there is no discharge at all. Our estimated results do not agree with his assumption. The discharge should be divided into two parts, one part (74%) leaks into the confining layers over the second half of the flow tube, the other (26%) flows out from the downstream end.
4. Estimated values of initial concentrations indicate that the initial water in Aquia aquifer is not pure seawater which is the asme as what reported by Appelo (1994).
5. There are other processes which affect the hydrochemical evolution along the Aquia aquifer (Chapelle and Knobel, 1983). For example, the role of the temperature to the chemical reaction patterns increases with depth. Lack of related data now, this study can be done in future.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

6.1. CONCLUSIONS

The inverse problem of water flow and reactive solute transport in variably saturated media is a new topic which involves a variety of parameters and geochemical processes. Existing optimization algorithms and theories of inverse problems in groundwater modeling have been extended in this dissertation to estimate flow and reactive transport parameters and identify geochemical processes.

6.1.1. FORMULATION OF FORWARD AND INVERSE MODELING

Inverse problems are established on the basis of forward modeling. Effective forward models such as TRANQUI (Xu, 1996 and Xu, et al., 1999) and CORE^{2D} (Samper et al., 1999), are needed for the inverse model. To this purpose, the most relevant physical processes involved on groundwater flow and solute and heat transport in variably saturated porous media have been described. From the principle of mass conservation, a general procedure is presented for the selection of the N_C primary species or components of a geochemical system having N_T species which are involved in N_R chemical reactions. Among the set of chemical reactions one has to distinguish those which only involve aqueous species (homogeneous reactions) from those which take place at the solid-liquid interface (heterogeneous reactions). A general formulation for reactions in complex chemical systems has been presented. Contrary to static systems where the chemical equilibrium condition is largely justified, in dynamic systems some reactions may progress slowly and kinetically controlled. Thus, the kinetics of such reactions has to be taken into account. In dynamic

systems both hydrodynamic and reactive processes take place simultaneously and thus they should be jointly analyzed. The derivation of the physical and chemical formulation of all processes affecting the transport of multicomponent species systems is based on two basic principles: (1) mass conservation of all species present in the liquid and solid phases and (2) the continuity of mass transfer between both phases. General equations for reactive transport have been derived and expressed in a closed form.

The aim of this dissertation is to develop algorithms for solving the inverse problem that can be applied to interpret laboratory experiments and field case studies. Basic concepts, theories and optimization methods of the inverse problems in flow and reactive solute transport have been described in Chapter 3. A generalized least square criterion which includes five types of data has been formed by a set of weighting coefficients. The criterion or objective function can take into account hydraulic heads, solute concentrations of different chemical components, total concentration including liquid and solid phases, water fluxes and water contents, as well as parameter prior information. Weighting coefficients for different types of data are updated iteratively.

The importance of independent prior information of parameters is proposed for the inverse problems. This information can provide a range of permissible values that a parameter can take and defines the maximum and minimum values which a parameter is allowed to assume during the optimization process. Moreover, it can be incorporated in the objective function with some special weighting coefficients. It provided a way of improving the structure of the covariance matrix and alleviating the ill-posedness of the inverse problem. Lack of prior information can lead to nonunique parameter estimates because certain pairs or groups of parameters, if varied in a linear or nonlinear combination, may lead to similar values of the objective function. In some cases this nonuniqueness can even lead to numerical instability and failure of the estimation process. However, if something is known about at least one of the parameters of such troublesome parameter group, this information, if included in the estimation process, may overcome the nonuniqueness and provide stability. Moreover, parameter estimates will also be nonunique if the observations are fewer than parameters. Furthermore, the inclusion of prior information, being mathematically equivalent to taking extra measurements, may alter the numerical

predominance of parameters over observations and thus provide the system with the ability to supply a unique set of parameter estimates. In practice, by giving different initial weighting coefficients to different pieces of prior information, one can assign the weights in accordance with the extent to which one wishes each piece of prior information to influence the parameter estimation process. On the other hand, to the extent that any piece of prior information is not satisfied, it can, more or less, lead the final estimation results close to the prior information. Therefore, with the improvement of the parameter optimization, one should reduce the initial weighting coefficients gradually to limit the influence of the prior information on the estimated results.

6.1.2. MINIMIZATION ALGORITHMS

The trial and error method is a model calibration technique performed manually. It has been traditionally preferred probably because of the uniqueness and identifiability problems that were supposed to plague automatic parameter estimation. In most cases these difficulties are either intrinsic to the problem (insufficient data), so that they are present whether calibration is performed manually or automatically, or can be easily overcome by state-of-the-art estimation methodologies (Samper et al. 1990). However, the trial and error procedure is time consuming, especially, when the number of unknown parameters is large. Accurate solutions for the inverse problem cannot be found using this procedure, and different researchers may obtain different results. Moreover, this method cannot provide quantitative error estimates and statistical measures of goodness-of-fit of the estimated parameters.

The non-gradient methods or search algorithms, such as Golden section search and quadratic interpretation methods, do not require the calculation of sensitivity coefficients. They belong to automatic calibration methods and are faster than trial and error. Their most interesting aspect is that they free the modeler from the most time consuming part of modeling, thus allowing one to concentrate on the conceptual side and to explore more alternative conceptual models. This algorithm is efficient for inverse problems involving a few (4 or 5) unknown parameters. Its convergence rate is generally slow. Since these kinds

of methods do not compute Jacobian matrix, they are not convenient for the evaluation of the estimated parameters on the statistical framework.

On the other hand, gradient solution methods, such as the Gauss-Newton-Marquardt method, usually have a faster convergence rate and are more economical and robust, especially when the number of parameters is large. The other advantage of the gradient-depend methods is that they can easily provide estimates of the variance-covariance matrix by using the existing Jacobian matrix or sensitivity coefficient matrix, and then compute the correlation matrix, eigenvalues and eigenvector components, as well as the confidence intervals for the error analysis. In INVERSE-CORE^{2D}, the Jacobian matrix is computed with a finite difference method. The central difference method is more accurate than the forward difference method, but requires more computing time (2M+1 forward runs, M is the number of the parameters). The forward difference method is faster and cheaper (M+1 forward runs). Therefore, for computing the Jacobian matrix it is suggested to start with the forward difference method. Then, when the solution is close to the optimum, switch to the central difference method for obtaining an accurate inverse solution.

6.1.3. ERROR ANALYSIS

Besides obtaining optimum parameters, one usually needs to know parameter uncertainty. In other words, one needs to know the likely error of the best-fit parameters. Data are generally not exact. They are subject to measurement and interpretation errors. Thus, the model never fit exactly the data, even when the model structure is correct. It is needed to assess whether or not the model is appropriate. Goodness-of-fit statistical tests are useful for that purpose. So, parameter fitting is not the sole purpose of parameter estimation. To be genuinely useful, a fitting procedure should provide (1) parameter estimates, (2) estimation errors, and (3) a statistical measure of goodness-of-fit. If the latter shows a poor fit, then items (1) and (2) are probably worthless.

For evaluating the estimated results, a series of methods have been incorporated in INVERSE-CORE^{2D}, including the computation of covariance and correlation matrices, the

eigenanalysis of the covariance matrix, and the calculation of confidence intervals of the estimated parameters. By using the error analysis tools, one can evaluate the uncertainty of the estimated parameters, may try several conceptual models, and choose the one that best fits available quantitative and qualitative data.

6.1.4. COMPUTER CODES

Three computer codes have been developed. These codes are INVGS-CORE (Golden section search method), INVQI-CORE (Quadratic interpretation method) and INVERSE-CORE^{2D} (Gauss-Newton-Levenberg-Marquardt method). The first two codes were verified with a synthetic example of the flow and solute transport in compacted bentonite (EJ1). Their efficiencies were compared under different conditions. INVGS-CORE was demonstrated more robust and practical.

INVERSE-CORE^{2D} was verified with two synthetic examples with noise-corrupted data. One example deals with solute transport with cation exchange reactions (EJ2). The other involves solute transport with cation exchange and kinetic dissolution (EJ2KIN). The code successfully estimated selectivity coefficients, CEC, initial and boundary concentrations, kinetic rate constant and specific surface of the mineral using synthetic data containing different levels of noise. With large amounts of noise, accurate solutions cannot be found without adding the prior information of the parameters to the objective function. On the other hand, when it is taken account into the objective function, estimated results improve significantly, which demonstrates the importance of the prior information of the parameters.

6.1.5. APPLICATION TO LABORATORY AND FIELD CASES

In this dissertation, five types of experiments and two sets of field data (Llobregat Delta Aquitard, Barcelona and Aquia aquifer, Maryland) have been successfully interpreted. The first real case is the interpretation of infiltration experiments. Five infiltration tests were

interpreted using four types of relative permeability functions. Even though these functions have some equivalence under some conditions, the Irmay model always has less estimated variance and it is recommended for the relative permeability of unsaturated compacted bentonite.

The second real case deals with through-diffusion experiments. Two types of tracers (HTO and strontium) and a total of seven tests have been interpreted. For HTO, the estimation procedure is quite easy and the porosity and the diffusion coefficient of the compacted bentonite are estimated. Four tests were performed for strontium. Since this tracer has a large sorption, it does not reach the downstream reservoir after almost 200 days, even though it has been fully flushed from the upstream reservoir after that time. Thus, only the activity observation data in the upstream reservoir are effective for parameter estimation, and the estimates of K_d and D_e contain a significant uncertainty. These two parameters are not identifiable at the same time. Only apparent diffusion coefficients are identifiable.

The third case is the so-called in-diffusion experiment. Two cesium and two selenium tests were interpreted. Porosity, K_d and D_e were identified by the inverse model. An important recommendation derived from the numerical interpretation of this kind of experiments is that only one clay sample should be located in each reservoir.

Two HTO permeation experiments in bentonite were interpreted. The classic single-porosity model was tested first. For the first permeation experiment (P-1) a convergent but suboptimal solution was obtained with a porosity of 0.35. The fit of the second experiment is very poor, in fact, the estimation algorithm failed to attain a convergent solution. Given the limitations of the single-porosity model, a more complex double-porosity model was considered. The fit of both tritium experiments is excellent with this model. Mobile or kinematic porosity is 0.2 while immobile porosity ranges from 0.44 to 0.6. The effective diffusion coefficient in the mobile phase ranges from 2.32 to $3.50 \cdot 10^{-9}$ m^2/s, while that of the immobile phase is almost three orders of magnitude smaller. This case indicates the importance of model selection.

Two other permeation experiments with granite samples were also interpreted. The mean porosity of the granite samples is 0.0062 and the dispersivity is about 0.016 m. The successful interpretation of advective-dominated permeation experiments in bentonite and granite indicates that the inverse model can be used to estimate flow and solute transport parameters in various experiments with different materials.

The fifth real case study deals with the interpretation of a column experiment with the clay sediment from Ketelmeer. By using INVERSE-CORE^{2D}, several parameters including five selectivity coefficients, CEC, four initial and four boundary concentrations were estimated under different conditions. Conceptual model selection and identification of geochemical processes was carried out during the estimation procedure. With our automatic inverse code, we get better fits to concentration breakthrough curves than those obtained by Appelo et al. (1990) by trial and error. CEC is the most sensitive parameter. Our result is about half the value of Appelo et al. (1990). Additional information from Dr. Appelo confirmed that our estimated CEC value is correct.

The Llobregat Delta Aquitard is one of the two field examples. By using INVERSE-CORE^{2D}, selectivity coefficients, CEC, hydraulic conductivity, dispersivity, initial and boundary concentrations (total 17 parameters) were estimated. Estimated CEC and selectivity coefficients of $K_{Na/Ca}$ and $K_{Na/NH4}$ are close to the values that Xu (1996) and Xu et al. (1999) obtained by trial and error. However, our estimated selectivity coefficients $K_{Na/K}$ and $K_{Na/Mg}$ are smaller. Initial and boundary concentrations of chemical components in Llobregat Delta aquitard are not accurately known. Manzano (1993), Manzano and Custodio (1995), and Xu (1996) gave the values according to their analyses and predicted results. Using their results as parameter prior information and initial values, our estimated results of initial and boundary concentrations lead to a significant improvement. Similar to Xu (1996), it can be concluded that redox processes have a strong effect on the transport of reactive cations, and such effect is exerted through a pronounced change in pH as well as on a change in total dissolved carbon. Accounting for redox processes leads to a significant improvement in matching measured Na, K, Mg and Ca concentrations. Estimated conductivity and dispersivity are slightly different from those estimated by Manzano (1993). She only used the concentrations of Cl^- to estimate these parameters. Our inverse model, on

the other hand, used the concentration data of 7 chemical components (Cl^-, Na^+, K^+, Ca^{+2}, Mg^{+2}, NH_4^+, and HCO_3^-). So, our estimated results have a larger confidence and reliability.

The other more complex field case is the Aquia aquifer (Maryland). By using INVERSE-CORE^{2D}, the inverse problem has been solved in five stages. Estimated parameters include conductivity, dispersivity, discharge rate, selectivity coefficient, CEC and initial and boundary concentrations. The chromatographic separation of ions during displacement of different water types in Aquia aquifer has been demonstrated for the case of freshwater displacement. The observed water chemistry displays the expected sequential peaks of Mg^{2+}, K^+ and Na^+ and has been modeled accurately. Chromatographic patterns should be considered as one of the common features of groundwater chemistry. Ion exchange reactions are always dynamic by their nature and give heterogeneous patterns of ion concentrations and side reactions along a flow line. The dynamic nature of ion exchange allows for a link with historic flow conditions and can be used to obtain paleohydrological information about the aquifer system. Estimated values of initial concentrations indicate that the initial water in Aquia aquifer is not pure seawater, which is consistent with what was reported by Appelo (1994). Estimated boundary concentrations, conductivities and dispersivities are similar to what Appelo (1994) used. But, Appelo (1994) assumed all the recharge from upstream leaked out from the second half of the aquifer, so at the downstream end there is no discharge at all. Our estimated results do not agree with his assumption. The discharge should be divided into two parts, one part (74%) leaks into the confining layers over the second half of the aquifer, the remaining part (26%) flows out from the downstream end.

The capabilities of our inverse models to estimate various types of parameters and account for a wide range of chemical processes in the laboratory and field cases demonstrate that they are very useful and robust tools. They can be practically used as inverse modeling tools for complex natural geochemical systems.

6.1.6. PRACTICAL ASPECTS AND LESSONS LEARNED IN SOLVING INVERSE PROBLEMS

The theory and applications of automatic parameter estimation of flow and reactive solute transport have been presented in this dissertation. However, as Carrera (1984) stated, whereas the formal term is "automatic parameter estimation", not everything in the approach is truly automatic, nor should it be. Trial and error is carried out throughout the estimation procedure. The ultimate decision to select a good conceptual model and fitted model parameters belongs to the modeler. A general procedure of constructing a flow and reactive solute transport model is summarized as following steps:

Step 1. Collect existing geological, hydrogeological and geochemical data.

Step 2. Establish conceptual models based on careful analysis of collected data.

Step 3. Parameterization. Zonation method has been implemented by our inverse codes because it is the most commonly used framework for geologically based conceptual models. Different parameters such as flow, transport and geochemical parameters may have different zone domains.

Step 4. Discretization. It involves subdividing the spatial domain to a set of finite elements and the time domain to a set of time intervals. Time intervals should coincide with the observation intervals. Each time interval can have variable time increments to satisfy convergent requirements of forward modeling.

Step 5. Select prior parameter information for different conceptual models. Each piece of prior information must have an attached initial weight. The more reliable the information, the larger the weight given. Each parameter to be estimated requires a piece of prior information. If it is not available, one should provide an educated guess and use a very low weight.

Step 6. Input files preparation. Three input files are needed to run INVERSE-CORE^{2D}:

(1) Root_TRA.INP: flow and transport input data file.

(2) Root_CHE.INP: geochemical reaction input data file

(3) Root_PAR.INP: parameter estimation input data file

The detail descriptions of these input files can be found in Appendix 5.

Step 7. Calibrate conceptual models using existing data. This stage is the most important and usually performed using qualitative criteria or an objective function (Equation

3.2.1). Conceptual models that fail to reproduce measured data are regarded as inappropriate. During this stage, identification of geochemical processes is very difficult. So far, it can be carried out by trial and error on different conceptual models with different chemical reaction patterns. However, in some cases, several models match equally well the observed data. Then, the parameters obtained during calibration are uncertain. Uncertainty can be evaluated by means of a posterior statistical analysis such as variance-covariance matrix, correlation matrix, eigenvalues and eigenvectors of the covariance matrix, as well as confidence intervals of the parameters. Inclusion of prior parameter information in the estimation process may overcome the nonuniqueness and provide stability to the solutions. A comparison of the objective function values obtained with different conceptual models and statistical analyses of estimation error can help to select the best model.

Step 8. Prediction. Use the identified conceptual model and the estimated parameters for forward prediction. This procedure can test further the effectiveness of the estimated parameters and finally determinate the model structure and parameters.

It is obvious that to finish the above 8 steps is not an easy job. There are difficulties associated with each step. As to an inverse analysis of a geochemical system, now observed chromatographic pattern in aquifer usually established during the last thousands years. Insufficient observation data or information is the main reason for improper conceptual model, unreliable prior parameter information, improper initial weights for measured data and prior information and it may finally lead to poor convergence of the minimization algorithm. At this situation, one needs to have much patience and confidence to keep trying with some systematically designed cases or conceptual models for overcoming the difficulties. In our experience, we have always ended up with satisfactory estimations.

6.2. RECOMMENDATIONS

The inverse problem of reactive solute transport in variably saturated media involves sixteen types of different parameters. They are surely not always independent from each other and the correlation among them is very complex. The identifiability, stability and uniqueness of the inverse problem are necessary to study in more detail.

Model selection and the identification of geochemical process are some topics which require future work. In our application examples, this work was done mainly on the basis of the objective functions and by trial and error. Although this method is a useful tool to identify model structure and geochemical processes, some special model selection criteria have to be developed and tested. This topic can lead to further research on experiment design and network design for the inverse problem of reactive solute transport.

Another line of future work is the derivation of the adjoint state equation for computing the Jacobian matrix for coupled flow and reactive solute transport in variably saturated media. This work can improve the accuracy of sensitivity coefficients and make the gradient-depend optimization algorithms more effective and economical.

The codes developed can estimate flow and reactive solute transport parameters in variably saturated media, which is sufficient for most cases. However, in some cases such as radioactive waste disposal and aquifer remediation, multiphase flow (including water, NAPL and gas) model should be considered. In these cases it would be required to estimate parameters related to gas flow and heat transport. Therefore, with the improvement of the forward model, a more complex inverse model could be solved.

REFERENCES

Allen, A. J., Baston, A.H., Bourke, P.J. and Jefferies, N.L., Small angle neutron scattering studies of diffusion and permeation through pores in clays; Nirex Safety Studies. Rept. NSS/R160, Harwell Laboratory, UK, 1988.

Anderman, E.R., M.C. Hill, and E.P. Poeter, Two-Dimensional Advective Transport in Ground-Water Flow Parameter Estimation, *Ground Water*, 34 (6), 1001-1009, 1996.

Anderson, M.P. and W.W. Woessner, Applied Groundwater Modeling: Simulating Flow and Advective Transport, Academic Press, San Diego, 381 pp., 1992.

Appelo, C. A. J., and D. Postma, Geochemistry, Groundwater and pollution, (Ed. Balkema, Rotterdam), 1993.

Appelo, C. A. J., A. Willemsen, H. E. Beekman and J. Groffioen, Geochemical calculations and observations on salt water intrusions. II. Validation of a geochemical model with laboratoty experiments, *J. Hydrol.*, 120, 225-250, 1990.

Appelo, C. A. J., Cation and proton exchange, pH variations and carbonate reactions in a freshening aquifer, *Water Resour. Res.*, 30(10), 2793-2805, 1994.

Appelo, C. A. J., E. Verweij and H. Schafer, A hydrogeochemical transport model for an oxidation experiment with pyrite/calcite/exchangers/organic matter containing sand, *Applied Geochemistry*, 13, 257-268, 1998.

Ayora, C., J. Samper, N. Cuéllar, and J. Cama, Métodos numéricos para la simulación del transporte de solutos reactivos en medios porosos, *V Congreso de Geoquímica*, 287-292, 1993.

Ayora, C., C. Taberner, and J. Samper, Modelización del transporte reactivo: aplicación a la dedolomitización, *Estudios Geológicos*, 50, 397-410, 1994.

Ayora, C., J. Samper, and T. Xu, Reactive solute transport modeling: problem statement and hydrochemical model, *VI Simposio de Hidrogeología y Recursos Hidráulicos*, Vol XIX, pp., 769-784, 1995.

Ayora, C., C. Taberner, M. W. Saaltink and J. Carrera, The genesis of dedolomites: a discussion based on reactive transport modeling, *J. Hydrol.*, 209(1-4), 346-365, 1998.

Barlebo, H. C., M. C. Hill, D. Rosbjerg and K. H. Jensen, Concentration data and dimensionality in groundwater models: Evaluation using inverse modeling, *Nordic Hydrology*, 29, 149-178, 1998.

Bear, J., Dynamics of fluids in porous media, Elsevier, New York, 764 pp., 1972.

Bear, J., Hydraulics of Groundwater, McGraw-Hill, New York, 569 pp., 1979.

Brandberg F. and Skagius K., Porosity, sorption and diffusivity data compiled for the SKB 91 study, Swedish Nuclear Fuel and Waster Management Co, SKB TR 91-16, 1991.

Bruno, J., C. Ayora, J. Delgado, L. Duro, M. J. Gimeno, C. Linklaker, and J. Peña, J., Blind prediction modeling at El Berrocal, Central Spain, INTERA Spain, *Cerdanyola*, 132 pp., 1995.

Bryant, S. L., R. S. Schechter, and L. W. Lake, Interactions of precipitation/dissolution waves and ion exchange in flow through permeable media, *AIChE J.*, 32 (5), 751-764, 1986.

Burdine, N.T., Relative permeability calculation from size distribution data, *Trans. Am. Inst. Min. Metall. Pet. Eng.*, 198, 71-78, 1953

Carnahan, C. L., Coupling of precipitation/dissolution reactions and mass diffusion via porosity changes, In: Chemical Modeling of Aqueous Systems II, edited by D. C, Melchior and R.L. Bassett, 234-242. *American Chemical Society*, Washington, D.C., 1990.

Carrera, J., Estimation of aquifer parameters under transient and steady-state conditions, Ph.D. dissertation, Dep. of Hydrol. and Water Resour., University of Arizona, Tucson, 1984.

Carrera, J., State of the art of the inverse problem applied to the flow and solute transport equations, *NATO ASI Ser., Ser. C*, 224, 549-584,1987.

Carrera, J , Samper J., J. Galarza, G. and Medina, A., An approach to process identification: Application to Solute transport through clays, International Conference on Calibration and Reliability in groundwater Modeling, IAHS, Publ. No. 195, 231-240, The Hague, 1991.

Carrera, J and Samper J., Hydrological Modeling for Safety Assessment of Radioactive Waste Disposal, In: Safety Assessment of Radioactive Waste Repositories, OCDE, Paris, 481-500, 1989.

Carrera, J. and S.P. Neuman, Estimation of aquifer parameters under steady-state and transient condition: 1. Background and Statistical framework, *Water Resour. Res.*, 22, 2, 199-210,1986a.

Carrera, J. And S.P. Neuman, Estimation of aquifer parameters under steady-state and transient condition: 2. Uniqueness, satbility, and solution algorithms, *Water Resour. Res.*, 22, 2,PP. 211-227, 1986b.

Carrera, J. and S.P. Neuman, Estimation of aquifer parameters under steady-state and transient condition: 3. Application to synthetic and field data., *Water Resour. Res.*, 22, 2, 228-242,1986c

Cederberg, G. A., R. Street, and J. O. Leckie, A groundwater mass transport and equilibrium chemistry model for multicomponent systems, *Water Resour. Res.*, 21(8), 1095-1104, 1985.

Chapelle, F. H., and D. D. Drummond, Hydrogeology, digital simulation, and geochemistry of the Aquia and Piney Point-Nanjemoy aquifer system in southern Maryland, Rep. Invest. 38, Md. Geol. Surv., Baltimore, Maryland, 100 pp., 1983.

Chapelle, F. H., and L. L. Knobel, Aqueous geochemistry and exchangeable cation composition of glauconite in the Aquia aquifer, Maryland, *Ground water*, 21, 343-352,1983.

Chu, W.-S., E. W. Strecker and D.P. Lettenmaier, An evaluation of data requirements for groundwater contaminant transport modeling, *Water Resour. Res.*, 23(3), 408-424, 1987.

CIEMAT, Ensayos THM para el Proyecto FEBEX, 70-IMA-L-5-51, Madrid, 51 pp., 1989.

Clifton, P. M. and S. P. Neuman, Effects of Kriging and inverse modeling on conditional simulation of the Avra Valley aquifer in southern Arizona, *Water Resour. Res.*, 18(4), 1215-1234, 1982.

Cooley, R. L., Incorporation of prior information on parameters into nonlinear regression groundwater flow models, 1, Thory, *Water Resour. Res.*, 18(4), 965-976, 1982.

Cooley, R. L., Some new procedures for numerical solution of variably saturated flow problems. *Water Resour. Res.*, 19(5), p1271-1285, 1983.

Cooley, R. L., A comparison of several methods of solving nonlinear regression groundwater flow problems, *Water Resour. Res.*, 21(10), 1525-1538, 1985.

Dagan, G. and Y. Rubin, Stochastic identification of recharge, transmissivity, and storativity in aquifer transient flow: A quasi-steady approach, *Water Resour. Res.*, 24(10), 1698-1710, 1988.

Dai, Z., J. Samper and J. Molinero, Numerical interpretation of diffusion and permeation experiments, FEBEX Report, 70-ULC-L-5-011, University of La Coruña, 84 pp., 1999.

Dai, Z. and J. Samper, Estimación automática de parámetros de flujo no saturado mediante ensayos de infiltracón, *IV Jornadas sobre Investigación en la Zona no Saturada del Suelo, Tenerife*, ISBN 84-699-1258-5, 175-180, 1999a.

Dai, Z. and J. Samper, INVERSE-CORE^{2D}: A Code for inverse problem of water flow and reactive solute transport, Users Manual, Version 0, *University of La Coruña*, 240 pp., December, 1999b.

de Marsily, G., Quantitative hydrogeology, Academic Press Inc., San Diego, 440 pp., 1986.

Delgado, J., Montenegro, L., Samper, J., Vázquez, A. and Juncosa, R., Reactive solute transport modeling of thermally-perturbed and irradiated Boom clay, *Geochemistry of the Earth's Surface*. H. Ármannsson (Ed.), Balkema, Rotterdam, 445-448, 1999.

Denbigh, K., The principles of chemical equilibrium, Cambridge, Univ. Press, 494 pp., 1981.

Doherty, J., L. Brebber and P. Whyte, PEST, Model-independent parameter estimation, *Watermark Computing*, 1994.

Dzombak, D. A., and F. M. M. Morel, Surface Complexation Modeling, *Wiley Interscience*, New York, 1990.

Eching, S.O., J. W. Hopmans y W. W. Wallender, Estimation of in situ unsaturated soil hydraulic functions from scaled cumulative drainage data, *Water Resour. Res.*, 30(8), 2387-2394, 1994.

Engesgaard, P., and K, L. Kipp, 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 Resour. Res.*, 28(10), 2829-2843, 1992.

ENRESA, FEBEX, Full-scale engineered barriers experiment in crystalline host rock, *Pre-operational Stage Summary Report,* Publicación técnica 1/98, 1998.

ENRESA, TRANQUI Users Manual. Informe para la Empresa Nacional de Residuos Radiactivos (ENRESA) J. Samper, C. Ayora, T. Xu, N. Cuéllar. La Coruña, 1995.

Gailey, R. M., S. M. Gorelick, and A. S. Crowe, Coupled process parameter estimation and prediction uncertainty using hydraulic head and concentration data, *Adv. Water Resour.*, 14(5), 301-314, 1991.

Galarza, G. A., Calibración automática de parámetros en problemas no lineales de flujo y transporte, Ph.D. Dissertation, Universidad Politécnica de Cataluña, Barcelona, Spain, 1993.

García-Gutiérrez M, T. Missana, M. Mingarro, J. Samper, Z. Dai and J. Molinero, Solute transport properties of compacted Ca-bentonite used in FEBEX Project, *accepted by J. Contaminant Hydrology, March, 2000.*

García, M., Missana, T., Yllera, A. and Mingarro, M., Diffusion and sorption experiment in compacted FEBEX bentonite. Technical Report 70-IMA-L-0-50. CIEMAT, 1998.

García, M., Martín, P.L., Celdas y técnicas para los estudios de difusión y coeficiente de distribución en arcillas compactadas, CIEMAT/IMA/54B50/5/96, 1996.

Garrels, R. M. and C. L. Christ, Solutions, minerals and equilibria, Freeman Cooper, San Francisco, 450 pp., 1965.

Ginn, T. R., and J. H. Cushman, Inverse methods for subaurface flow: A critical review of stochastic techniques, *Stochastic Hydrol. Hydraul.*, 4, 1-26, 1990.

Helgeson, H. C., and D. H. Kirkham, Theoretical prediction of the thermodynamic behaviour of aqueous electrolytes at high pressures and temperatures: II. Debye-Hückel parameters for activity coefficients and relative partial molal properties, *Am. J. Sci.*, 274, 1199-1261, 1974.

Herzer, J., and W. Kinzelbach, Coupling of transport and chemical processes in numerical transport models, *Geoderma*, 44, 115-127, 1989.

Hill, M. C., A Computer Program (MODFLOWP) for Estimating Parameters of a Transient, Three-Dimensional, Ground-Water Flow Model using Nonlinear Regression. U. S. Geological Survey Open-File Report 91-484, 1992.

Huertas, F. J., P. Carretero, J. Delgado, J. Linares and J. Samper, Ion exchange behavior of the FEBEX bentonite. 1. Na/K, Na/Mg and Na/Ca experimental exchange isotherms, Sydney, 2000.

Huertas, F., J. L. Fuentes-cantillana, F. Jullien, P. Rivas, J. Linares, P. Fariña, M. Ghoreychi, N. Jockwer, W. Kickmaier, M. A. Martínez, J. Samper, E. Alonso and F. J. Elorza, Full-scale engineered barriers experiment for a deep geological repository for high-level radioactive waste in crystalline host rock (FEBEX project), Final report, EUR 19147, European Communities, 362 pp., 2000.

Inoue, M., J. Simunek, J. W. Hopmans and V. Clausnitzer, In situ estimation of soil hydraulic functions using a multistep soil-water extraction technique, *Water Resour. Res.*, 34(5), 1035-1055, 1998.

Irmay, S., On the hydraulic conductivity of saturated soils, *Trans. Amer. Geophys. Union* 35, 463-468, 1954.

Juncosa, R., Modelos de flujo multifasico no isotermo y de transporte reactivo multicomponente en medios porosos, Ph.D. Dissertation, Universidad Politécnica de Madrid, 1999.

Kabala, Z.J. and P. C. D. Milly, Sensitivity analysis of infiltration, exfiltration, and drainage in unsaturated Miller-Similar porous media, *Water Resour. Res.*, 27(10) 2655-2666, 1991.

Keating, E.H., and J. M. Bahr, Using reactive solutes to constrain groundwater flow models at a site in northern Wisconsin, *Water Resour. Res.*, 34(12), 3561-3571, 1998a.

Keating, E.H., and J. M. Bahr, Reactive transport modeling of redox geochemistry: Approaches to chemical disequilibrium and reaction rate estimation at a site in northern Wisconsin, *Water Resour. Res.*, 34(12), 3573-3584, 1998b.

Keidser, A. and D. Rosbjerg, A comparison of four inverse approaches to groundwater flow and transport parameter identification, *Water Resour. Res.*, 27(9) 2219-2232, 1991.

Kool, J. B., Parker, J. C. y van Genuchten, Determining soil hydraulic properties from one-step outflow experiments by parameter estimation, 1. Theory and numerical studies, *Soil Sci. Soc. Am. J.*, 49, 1348-1354, 1985.

Kool, J. B., Parker, J. C. and van Genuchten, Parameter estimation for unsaturated flow and transport model—A review, *J. Hydrol.*, 91, 255-293, 1987.

Kool, J. B. and Parker, J. C., Analysis of inverse problem for transient unsaturated flow, *Water Resour. Res.*, 24(8), 817-830, 1988.

Law, A. M. and W. D. Kelton, Simulation modeling and analysis, New York, 398pp, 1980.

Lensing, H. J., M. Vogt, and B. Herrling, Modeling of biologically mediated redox processes in the subsurface, *J. Hydrol.*, 159, 125-143, 1994.

Levenberg, K., A method for the solution of certain non-linear problems in least squares. *Q. Appl. Math.*, v. 2, 164-168, 1944.

Li, J., N.-Z. Sun, and W. W.-G. Yeh, A comparative study of sensitivity coefficient calculation methods in groundwater flow, in *Proceedings of the 6th International Conference in Finite Elements in Water Resources*, Lisbon, Portugal, 1986.

Liu, C. W., and T. N. Narasimhan, Redox-controlled multiple species reactive chemical transport, 1. Model development, *Water Resour. Res.*, 25(5), 869-882, 1989a.

Liu, C. W., and T. N. Narasimhan, Redox-controlled multiple species reactive chemical transport, 2. Verification and Application, *Water Resour. Res.*, 25(5), 883-910, 1989b.

Manzano, M., Génesis del agua intersticial del acuitardo del delta del Llobregat: origen de los solutos y transporte reactivo con el medio sólido, Ph.D. Dissertation, Universidad Politécnica de Cataluña, Barcelona, Spain,1993.

Manzano, M. and E. Custodio, Origen de las aguas salobres en sistemas acuíferos deltaicos: applicacion de la teoría de la cromatográfia iónica al acuitardo del delta del Llobregat, VI Simposio de Hidrogeología y Recursos Hidráulicos. Vol. XX. pp., 179-201,1995.

Marquardt, D. W., An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society of Industrial and Applied Mathematics*, v11, no. 2, 431-441, 1963.

McLaughlin, Dennis and Lloyd R. Townley, A reassessment of the groundwater inverse problem, *Water Resour. Res.*, 32(5), 1131-1161, 1996.

McNab, W. Jr. and T. N. Narasimhan, Modeling reactive transort of organic compounds in groundwater using a partial redox disequilibrium approach, *Water Resour. Res.*, 30(9), 2619-2635, 1994.

Medina,A. and J. Carrera, Coupled estimation of flow and solute transport parameters, *Water Resour. Res.*, 32(10), 3063-3076, 1996.

Melchior, D. C., and R. L. Bassett (Eds.), Chemical Modeling of Aqueous Systems II, *American Chemical Society*, Washington, D. C., 1990.

Miller, C. W., and L. V. Benson, Simulation of solute transport in a chemically reactive heterogeneous system: Model development and application, *Water Res. Res.*, 19(2), 381-391, 1983.

Mishra, S. and S. C. Parker, Parameter estimation for coupled unsaturated flow and transport, *Water Resour. Res.*, 25(3), 385-396, 1989.

Miyahara, K., Ashida, T., Kohara, Y., Yusa, Y., Sasaki, N., Effectof bulk density on diffusion of cesium in compacted sodium bentonite, *Radiochim. Acta* 52/53, 293-297, 1991.

Montenegro, L., G. Zhang, J. Samper and J. Delgado. Documento de Verificación del código CORE^{2D}. Informe Técnico, E.T.S. Ingenieros de Caminos. Universidad de A Coruña, 179 pp., 1999.

Morel-Seytoux, Hubert J., P. D. Meyer, M. Nachabe, J. Touma, M. T. van Genuchten and R. J. Lenhard, Parameter equivalence for the Brooks-Corey and van Genuchten soil characteristics: Preserving the effective capillary drive, *Water Resour. Res.*, 32(5), 1251-1258, 1996.

Mualem, Y.; A new model for predicting the hydraulic conductivity of unsaturated media; *Water Resour. Res.*, 12(3), 513-522, 1976.

Narasimhan, T. N., A. F. White, and T. Tokunaga, Groundwater Contamination from an inactive uranium mill tailings pile, 2. Aplication of a dynamic mixing model, *Water Resour. Res.*, 22(13), 1820-1834, 1986.

Neretnieks, I., Diffusivities of some dissolved constituents in compacted clay MX-80 and the impact on radionuclide migration in the buffer, KBS Technical Report 82-27, 1982.

Neretnieks, I., Yu, J. W. and Liu, J., An efficient time scaling technique for coupled geochemical and transport models, *J. Contam. Hydrology*, 26, 269-277, 1997.

Neuman, S.P. and S. Yakowitz, A statistical approach to the inverse problem of aquifer hydrology, 1. Theory: *Water Resour. Res.*, 15, 845-860, 1979.

Neuman, S.P., A statistical approach to the inverse problem of aquifer hydrology, 3. Improved solution method and added perspective, *Water Resour. Res.*,16(2),331-346, 1980.

Neuman, S. P., Universal scaling of hydraulic conductivities and dispersivities in geologic media, *Water Resour. Res.*, 26 (8), 1749-1758, 1990.

Nienhuis, P., C. A. T. Appelo, and A. Willemsen, Program PHREEQM: Modified from PHREEQE for use in mixing cell flow tube, Free University, Amsterdam, The Netherlands, 1991.

Oscarson, D.W., Hume, H.B., King, F., Sorption of cesium on compacted bentonite, *Clay and clay Minerals*, Vol.42, N° 6, 731-736, 1994.

Pan, L. and Wu, L. A hybrid global optimization method for inverse estimation of hydraulic parameters: Annealing-simplex method, *Water Resour. Res.*, 34(9), 2261-2269, 1998.

Paniconi, C., and M. Putti, A comparison of Picard and Newton iteration in the numerical solution of multidimensional variably saturated flow problems, *Water Resour. Res.*, 30 (12), 3357-3374, 1994.

Parker, J. C., Kool, J. B. y van Genuchten, Determining soil hydraulic properties from one-step outflow experiments by parameter estimation, II. Experimental studies, *Soil Sci. Soc. Am. J.*, 49, 1354-1359, 1985.

Parkhurst, D. L., D. C. Thorstenson, and L. N. Plummer, PHREEQE: A computer program for geochemical calculations, US Geol. Surv. Water Resour. Invest. 80-96, 174 pp., 1980.

Parkhurst, D. L., User's guide to PHREEQC—a computer program for speciation, reaction-path, advective-transport, and inverse geochemical calculations, *U.S. Geol. Surv. Water Resour. Inv.*, 95-4227, 1995.

Parkhurst, D.L., User's guide to PHREEQC (version 2)--A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations: U.S. Geological Survey Water-Resources Investigations Report 99-4259, 312 pp., 1999.

Poeter, E.P., and M.C. Hill, Unrealistic Parameter Estimates in Inverse Modeling: A Problem or a Benefit Calibration?, MODELcare, Colorado School of Mines, Golden, Colorado, 1996.

Poeter, E.P. and M.C. Hill, Inverse Methods: A Necessary Next Step in Groundwater Modeling, *Ground Water*, v. 35, no. 2, 250-260, 1997.

Poeter, E.P. and Hill, M.C., Documentation of UCODE, a computer code for universal inverse modeling: U.S. Geological Survey Water-Resources Investigations Report 98-4080, 116pp., 1998.

Press, W. H., S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, Numerical Recipes - The Art of Scientific Computing, Cambridge, Cambridge University Press (second edition), 963 pp., 1992.

Put, M. and Henrion, P., Modelling of radionuclide migration and heat transport from an HLW-repository in Boom clay, EUR-14156, 1992.

Put, M.J., An improved mathematical model for the interpretation of the Flow-Through type diffusion test with influence of filterplates, *Modelling of radionuclide migration and heat transport from an HLW-repository in Boom clay*, EUR-14156, 1992.

Rubin, J., Solute transport with multisegment, equilibrium-controlled reactions: A feed-forward simulation method, *Water Resourc. Res.*, 26, 2029-2055, 1990.

Rubin, J., Solute transport with multisegment, equilibrium-controlled classical reactions: Problems solvability and feed forward method's applicability for complex segments of at most binary participants, *Water Resour. Res.*, 28 (6), 1681-1702, 1992.

Russo, David, Determining soil hydraulic properties by parameter estimation: On the selection of a model for the hydraulic properties, *Water Resour. Res.*, 24(3), 453-459, 1988.

Russo, David, E. Bresler, U. Shani, y J. C. Parker, Analysis of infiltration events in relation to determining soil hydraulic properties by inverse problem methodology, *Water Resour. Res.*, 27(6), 1361-1373, 1991.

Saaltink, M. W., C. Ayora and J. Carrera, A mathematical formulation for reactive transport that eliminates mineral concentrations, *Water Resour. Res.*, 34(7), 1649-1656, 1998.

Saaltink, M. W., On the approaches for incorporating equilibrium and kinetic chemical reactions in transport models, Ph.D. Dissertation, Universidad Politécnica de Cataluña, Barcelona, Spain, 1999.

Samper, F.J., Statistical methods of analyzing hydrological, hydrochemical and isotopic data from aquifers, Ph.D. dissertation, Dep. Of Hydrol. and Water Resour., University of Arizona, Tucson, 1986.

Samper, F.J. and S.P. Neuman, Adjoint state equations for advective-dispersive transport, in *Proceedings of the 6th International Conference in Finite Elements in Water Resources*, Lisbon, Portugal, 423-437, 1986.

Samper, F.J., J. Carrera, G. Galarza, and A. Medina, Application of an automatic calibration technique to modeling an alluvial aquifer, ModelCARE 90: *Calibration and reliability in groundwater modeling*, IAHS Publ. no. 195, 87-95, 1990.

Samper, F.J., Flujo de agua y transporte de masa en el medio no saturado: Estado del arte. En: Hidrogeología, Estado Actual and Prospectiva., C.I.H.S., Ed. F. Anguita, I. Aparicio, L. Candela and M. Zurbano, 79-100, 1991.

Samper, F.J., and C. Ayora, Modeling groundwater chemical evolution in a closed-basin evaporitic environment, In Grounwater Quality Management, IAHS Publ. 220, 173-182, 1994.

Samper, F.J., C. Ayora, T. Xu, and N. Cuéllar, Reactive solute transport modeling: numerical formulation and computer codes, VI Simposio de Hidrogeología. Hidrogeología y Recursos Hidráulicos, Vol XIX, 785-800, 1995.

Samper, F.J. and M. A. García-Vera, Inverse modeling of groundwater flow in the semiarid evaporitic closed basin of Los Monegros, Spain, *Hydrogeology Journal*, 6, 33-49, 1998.

Samper, F.J. and S.P. Neuman, Estimation of spatial covariance structure by adjoint state maximum likelihood cross validation, 1, Theory, *Water Resour. Res.*, 25(3),351-362, 1989a.

Samper, F.J. and S.P. Neuman, Estimation of spatial covariance structure by adjoint state maximum likelihood cross validation, 2, Synthetic experiments, *Water Resour. Res.*,25(3),363-371, 1989b.

Samper, F.J. and S.P. Neuman, Estimation of spatial covariance structure by adjoint state maximum likelihood cross validation, 3, Application to Hydrochemical and Isotopic Data, *Water Resour. Res.*, 25(3), 373-384, 1989c.

Samper, F.J., R. Juncosa, J. Delgado and L. Montenegro, CORE-LE-2D: Users Manual, Universidad de La Coruña, 207 pp., 1998.

Samper, F.J., R. Juncosa, J. Delgado and L. Montenegro,. CORE^{2D}: A code for non-isothermal water flow and reactive solute transport, Users manual version 2, Universidad de La Coruña, Publ. ENRESA, 131 pp., 2000.

Samper, F.J., J. Delgado, R. Juncosa, G. Monlinero and Z. Dai, Pre-operational THG modeling of FEBEX large scale experiments, *International Workshop on Key Issues in Waste Isolation Research, Barcelona*, 1998.

Samper, F.J., J. Delgado, R. Juncosa, P. Carretero, Z. Dai, A. Vázquez, G. Monlinero, and A. Ruiz, Modelización Termo-Hidrogeoquímica del Proyecto FEBEX, *Final Report of FEBEX Project*, Volumen I, 70-ULC-L-5-001, 1999a.

Samper, F.J., J. Delgado, R. Juncosa, P. Carretero, Z. Dai, A. Vázquez, G. Monlinero, and A. Ruiz, Modelización Termo-Hidrogeoquímica del Proyecto FEBEX, *Final Report of FEBEX Project*, Volumen II, 70-ULC-L-5-002, 1999b.

Samper, F.J., J. Delgado, J. Molinero and R. Juncosa, Predictive numerical modeling of artificial tracers migration in the large-scale heating and hydration experiments of the FEBEX Project, *Scientific Basis for Nuclear Waste Management*, I. McKinley y Ch. McCombie; *Mat. Res. Soc. Symp. Proc.* Vol.506, 1998.

Samper, F.J., R. Juncosa, V. Navarro, J. Delgado, L. Montenegro and A. Vázquez, Coupled thermo-hydro-geochemical models of engineered barrier systems: The FEBEX Project, Sydney, accepted 2000.

Simunek, J. and D. L. Suares, Two-dimensional transport model for variably saturated porous media with major ion chemistry. U.S. Salinity Laboratory Agricultural Research Service, U.S. Department of Agriculture, Riverside, California, U.S.A. Riverside Group, 1993.

Simunek, J., and D. L. Suares, UNSATCHEM-2D: Code for simulating two-dimensional variably saturated water flow, heat transport, carbon dioxide production and transport, and multicomponent solute transport with major ion equilibrium and kinetic chemistry, User's manual, U.S. Salinity Laboratory, Agricultural Research Service, U.S. Department of Agriculture, Riverside, California, 1993.

Simunek, J., and D. L. Suares, Two-dimensional transport model for variably saturated porous media with major ion chemistry, *Water Resour. Res.*, 30(4), 1115-1133, 1994.

Simunek, J., M. T. Van Genuchten, and D. L. Suares, Modeling multiple solute transport in variably saturated soils, In: Groundwater Quality: Remediation and Protection, IAHS Publ. 225, 311-318, 1995.

Simunek, J., and M. T. van Genuchten, Estimating unsaturated soil hydraulic properties from tension disc infiltrometer data by numerical inversion, *Water Resour. Res.*, 32(9), 2683-2696, 1996.

Sonnenborg, T. O., P. Engesgaard and D. Rosbjerg, Contaminant transport at a waste residue deposit, 1. Inverse flow and nonreactive transport modeling, *Water Resour. Res.*, 32(4), 925-938, 1996.

Sposito, G., and S. V. Mattigod, GEOCHEM: a computer program for the calculation of chemical equilibria in soil solutions and other natural water systems, Dept. Soil and Environ. Sci., Univ. California, Riverside, 92 pp., 1980.

Sposito, G., The chemistry of soils, Oxford University Press, New York, 277 pp., 1989.

Steefel, C. I., 1DREACT: One dimensional reaction-transport model, User's manual and programmer's guide, Batelle Pacific Northwest Laboratories, 1993.

Steefel, C. I., and A. C. Lasaga, A coupled model for transport of multiple chemical species and kinetic precipitation/dissolution reactions with applications to reactive flow in single phase hydrothermal system, *Am. J. Sci.*, 294, 529-592, 1994.

Stuyfzand, P.J., and Timmer, H., Deep well injection at the Langerak and Nieuwegein sites in the Netherlands: chemical reactions and their modeling, report no SWE 96.006, Kiwa Research and Consultancy, Nieuwegein (the Netherlands), 1999.

Sun, N-Z. And W. W-G. Yeh, Identification of parameter Structure in Groundwater Inverse Problems, *Water Resour. Res.*, 21(6), 869-883, 1985.

Sun, N-Z, Inverse Problems in Groundwater Modeling, *Kluwer Academic Publishers*, the Netherland, 364 pp., 1994.

Sun, N-Z., y W. W.-G. Yeh, Coupled Inverse Problems in Groundwater Modeling, 1 Sensitivity Analysis and Parameter Identification, *Water Resour. Res.*, 26(10), 2507-2525, 1990a.

Sun, N-Z., y W. W.-G. Yeh, Coupled Inverse Problems in Groundwater Modeling, 2. Identifiability and Experimemtal Design, *Water Resour. Res.*, 26(10), 2527-2540, 1990b.

Sun, N-Z., Mathematical Modeling of Groundwater pollution, Springer Verlag., New York, 377 pp., 1996.

Sun, N-Z. M-C. Jeng, and W. W-G. Yeh, A proposed geological parameterization method for parameter identification in three-dimensional groundwater modeling, *Water Resour. Res.*, 31(1), 89-102, 1995.

Sun, N-Z., and W. W-G. Yeh, A stochastic Inverse Solution for Transient Groundwater Flow: Parameter Identification and Reliability Analysis, *Water Resour. Res.*, 28(12): 3269-3280, 1992.

van Genuchten, M. Th, A closed-form equation for predicting the hydraulic conductivity of unsaturated soils; Soil Sci. Soc. Amer. J. 44, 892-898, 1980.

van Genuchten, M. T., and W. J. Alves, Analytical solutions of the one-dimensional convective-dispersive solute transport equation, U. S. Department of Agriculture, Technical Bulletin No. 1661, 151 pp., 1982.

Villar, M.V. and P.L. Martin, Bentonite Buffer: Summary of properties, performed and ongoing studies and processing of the material and block manufacturing. Second FEBEX workshop, Toledo, April 911, 70-IKA-L-0-07, CIEMAT, 1996.

Villar, M.V. and Pelayo, M., Characterization of La Serrata clay used in the FEBEX Project, FEBEX report 70-IMA-L-3-15, 1996.

Villar, M.V., Cuevas, J. and Fernández, A.M., Caracterizacion geoquimica de bentonita compactada, Efectos producidos por flujo termohidraulico, FEBEX report 70-IMA-L-0-2, 1996.

Voss, C., SUTRA, A finite element simulation model for saturated-unsaturated, fluid-density-dependent groundwater flow with energy transport or chemically-reactive single-species solute transport, U.S.G.S., Water-Resources Investigations Report 84-4369, 409 pp., 1984.

Wagner, B. J., and S. M. Gorelick, Optimal groundwater quality management under parameter uncertainty, *Water Resour. Res.*, 23(7), 1162-1174, 1987.

Wagner, B. J., Simultaneous parameter estimation and contamination source characterization for coupled groundwater flow and contaminant transport modeling, *J. Hydrol.*, 135, 275-303, 1992.

Walsh, M. P., S. L. Bryant, and L. W. Lake, Precipitation and dissolution of solids attending flow through porous media, *AIChE J.*, 30 (2), 317-328, 1984.

Walter, A. L., E. O. Frind, D. W. Blowes, C. J. Ptacek, and J. W. Molson, Modeling of multicomponent reactive transport in groundwater, 1, Model development and evaluation, *Water Resour. Res.*, 30 (11), 3137-3148, 1994a.

Walter, A. L., E. O. Frind, D. W. Blowes, C. J. Ptacek, and J. W. Molson, Modeling of multicomponent reactive transport in groundwater, 2, Metal mobility in aquifers impacted by acidic mine tailings discharge, *Water Resour. Res.*, 30 (11), 3149-3158, 1994b.

Wang, J. S. Y., and N. T. Narasimhan, Fluid flow in partially saturated, welded-nonwelded tuff units, *Geoderma*, 46, 155-168, 1990.

White, S. P., Multiphase non-isothermal transport of systems of reacting chemicals, *Water Resour. Res.*, 31(7), 1761-1772, 1995.

Wilson, J., and P. J. Miller, Two Dimensional Plume in Uniform Ground Water Flow, *ASCE Journal Hidraulics Div.*, HY4, 503-514, 1978.

Wolery T. J., Calculation of chemical equilibrium between aqueous solution and minerals: the EQ3/6 software package, Lawrence Livermore Laboratory, 1979.

Wolery, T. J., EQ3NR: a computer program for geochemical aqueous speciation-solubility calculations: User's guide and documentation, Lawrence Livermore Laboratory URCL-53414, 1983.

Wolery, T. J., EQ6: a computer program for reaction-path modeling of aqueous geochemical systems, User's guide and documentation, Lawrence Livermore Lab., 251 pp., 1984.

Woodbury, A. D., L. Smith, and W. S. Dunbar, Simultaneous inversion of hydrogeologic and thermal data, 1, Theory and application using hydraulic head data, *Water Resour. Res.*, 23(8), 1586-1606, 1987.

Woodbury, A. D., L. Smith, and W. S. Dunbar, Simultaneous inversion of hydrogeologic and thermal data, 2, Incorporation of thermal data, *Water Resour. Res.*, 24(3), 356-372, 1988.

Xiang, Y., J. F. Sykes, and N. R. Thompson, Composite L_1 parameter estimator for model fitting in groundwater flow and solute transport simulations, *Water Resour. Res.*, 29(6), 1661-1673, 1993.

Xu, T., J. Samper, C. Ayora, M. Manzano and E. Custodio, Modeling of non-isothermal multicomponent reactive transport in field scale porous media flow systems, *J. of Hydrol.*, Vol. 214, 144-164, 1999.

Xu, T., K. Pruess and G. Brimhall, An improved equilibrium-kinetics speciation algorithm for redox reactions in variably saturated subsurface flow systems, *Computers and Geosciences*, 25, 655-666, 1999.

Xu, T., Modeling Nonisothermal Multi-component Reactive Solute Transport through Variably Saturated Porous Media, Tesis Doctoral, La Coruña, E.T.S.I.C.C.P., U.D.C., 1996.

Yeh, G. T., and V. S. Tripathi, HYDROGEOCHEM: A coupled model of hydrological and geochemical equilibrium of multicomponent systems, Rep. ORNL-6371, 312 pp., Oak Ridge Natl. Lab., Oak Ridge, Tenn., 1990.

Yeh, G.T., and V.S. Tripathi, A critical evalution of recent developments of hydrogeochemical transport models of reactive multi-chemical components, *Water Resour. Res.*, 25(1), 93-108, 1989.

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

Yeh, T.-C. Jim and Donald J. Harvey, Effective unsaturated hydraulic conductivity of layered sands, *Water Resour. Res.*, 26(6), 1271-1279, 1990.

Yeh, W. W-G., Review of parameter identification procedures in ground-water hydrology: The inverse problem, *Water Resour. Res.*, 22(2), 95-108, 1986.

Yu, J-W. and Neretnieks, I., Diffusion and sorption properties of radionuclides in compacted bentonite, SKB, TR-97-12, 1996.

Zayani, K., Tarhouni,J., Vachaud, G. and Kutilek, M., An inverse method for estimating soil core water characteristics. *J. Hydrol.*, 122, 1-13, 1991.

Zijlstra, Jan and J. H. Dane, Identification of hydraulic parameters in layered soils based on a quasi-Newton method, *J. of Hydrol.*, 181, 233-250, 1996.

Zysset, A., F. Stauffer, and T. Dracos, Modeling of chemically reactive groundwater transport, *Water Resour. Res.*, 30(7), 2217-2228, 1994a.

Zysset, A., F. Stauffer, and T. Dracos, Modeling of reactive groundwater transport governed by biodegradation, *Water Resour. Res.*, 30(8), 2423-2434, 1994 b.

APPENDIX 1

LIST OF TERMS

Greek terms

α	probability of a chi-square variable for computing confidence intervals
α	Marquardt parameter
α	leakage coefficient
α	dispersivity
α	user-defined coefficient or relative parameter increment
α_L, α_T	longitudinal and transverse dispersivities
β	parameter controlling the type of transport boundary condition
β	optimal parameter adjustment coefficient
β	thermal dispersivity
β	equivalent fraction of exchanged species
β_L, β_T	longitudinal and transverse thermal dispersivity (L)
ΔG_{ads}	free entalpy change of the overall adsorption reaction
ΔG_{intr}	enthalpy change due to chemical bonding
ΔG_{coul}	enthalpy change due to electrostatic work
ΔG_R	Gibbs Free energy of reaction
Δz	change in the charge of the surface species
ε	updating damping factor in Newton-Raphson iterations
ε	dielectric constant of water
ε_0	permitivity of free space
Γ	boundary
Γ	activity coefficient of gaseous species
γ	activity coefficient of aqueous species
ϕ	porosity
κ	Debye-Hückel reciprocal length
Λ	activity coefficient of solid species
λ	eigenvalue
λ	equivalent thermal conductivity tensor
λ_L	longitudinal thermal conductivity
λ_T	transverse thermal conductivity
λ_{xx}, λ_{yy}, λ_{xy}	components of tensor λ
λ_0	medium isotropic thermal conductivity
μ	dynamic viscosity

μ	chemical potential
η	order of kinetic reactions
ν_{ji}	stoichiometric coefficient of the i-th species in the j-th reaction
ν_{ij}^{x}	stoichiometric coefficient of the j-th primary species in the dissociation reaction of the i-th aqueous secondary species
ν_{ij}^{p}	stoichiometric coefficient of the j-th aqueous primary species in the dissolution reaction of the i-th precipitated species
ν_{i}	stoichiometric coefficient of dissolved cations
ν_{j}	stoichiometric coefficient of interlayer cations
ν_{ij}^{p}	stoichiometric coefficient of the j-th aqueous species on the i-th precipitated species
ν_{ij}^{s}	stoichiometric coefficient of the j-th adsorption component in the desorption reaction of the i-th surface complex
ν_{ij}^{y}	stoichiometric coefficient of the j-th aqueous primary species in the desorption reaction of the i-th surface complex
ν_{ij}^{w}	stoichiometric coefficient of the j-th aqueous primary species in the i-th exchanged species
θ	volumetric water content
θ	experimental parameter of kinetic reactions
ρ	water density
ρ_{ij}	correlation coefficient
ρ_{m}	bulk density
ρ_{s}	density of the solids
σ	standard deviation of measured concentrations
σ^2	estimation variance
σ_{s}	solid-solution interfacial free energy
ρ_{w}	water density
σ	surface charge
τ	medium tortuosity
τ	tolerance threshold for convergence in Newton-Raphson iterations
ξ	reaction progress variable
ψ	pressure head
Ω	flow domain
Ω	sub/super-saturation
Ω	zero matrix
Ω_{m}	saturation ratio
Ψ	electric potential
$\nabla .( )$	divergence operator
$\nabla ( )$	gradient operator

Latin terms

A	specific surface area of the mineral

A^*	surface area of the solid per unit mass of solids
A_γ, B_γ, b_γ	coefficients for Debye-Hückel extended formula
a	thermodynamic activity
a	parameter lower bound
AC	acetate in solution
a_{ie}	index of electron in the chemical composition of the i-th species
a_{ik}	index of element k in the chemical composition of the i-th species
b	aquifer thickness
b	parameter upper bound
c	parameter middle point
c	solute concentration
C_e	total electron concentration
c_e	free electron concentration
C_H	total proton concentration
c_{H+}	concentration of free proton
C_j	total concentration of the j-th aqueous primary species
$C_{j,inp}$	concentration of j-th component in the input
$C_{j,out}$	concentration of j-th component in the output
c_j	concentration of the j-th aqueous primary species
c_w,c_s,c_m	specific heat of water, solids, porous medium
c_0	initial concentration
C_1	capacitance of the adsorption layer (constant cap. model)
C_2	capacitance of the ß-layer (triple layer model)
$\tilde{c}$	boundary concentration
c^*	sink/source concentration
CEC	cation exchange capacity
C(***P***)	a first-order approximation of the covariance matrix of ***P***
d	new search point of the parameter
D	dispersion tensor
D_a	apprarent diffusion coefficient
D_e	effective diffusion coefficient
D_h	hydrodynamic dispersion tensor
D_k	parameter increment vector of Newton's method
D_I, D_{II}	Dämkohler numbers
D_L, D_T	main components (longitudinal, transversal) of dispersion tensor
D_m	coefficient of molecular diffusion in porous medium
D_{xx}, D_{yy}, D_{xy}	dispersion tensor components
D_0	coefficient of molecular diffusion in water
E	objective function
E_a	activation energy
P'	coefficient matrix
F	Faraday's constant
F	solute mass flux vector
F_1	observation values and the prior information of the parameters
$\bar{F}$	average of the observation concentrations
$\mathbf{F_A}$, $\mathbf{F_B}$, $\mathbf{F_H}$	advection, molecular diffusion and hydrodynamic dispersion mass flux vectors

F_0	specified solute mass flux on boundaries
G	molar Gibbs free enthalpy
g	gravity acceleration
g	vertical vector pointing downwards of modulus equal to gravity acceleration
$\mathbf{g}_k$	gradient vector of the objective function
[g]	concentration of gas species
H	specified head
H(p)	Hessian matrix
I	identity matrix
h	hydraulic head
h_0	initial head
I	ionic strength
I	identity tensor
I	identity matrix
IAP	ion activity product of the solution
J	Jacobian matrix
K	equilibrium constant
K	saturated hydraulic conductivity tensor
K_r	relative hydraulic conductivity
k	intrinsic permeability
k	dissolution/precipitation rate constant
k_b	Boltzmann constant
K_{eq}	equilibrium ion activity product of the solution
K_{intr}	equilibrium constant related to the chemical reaction. Intrinsic constant
K_i	equilibrium constant in the j-th chemical reaction
K^*	coefficient for cation exchange
k^*_m	effective rate constant
L_i	number of observation points in the time-space domain
L( )	linear transport operator which includes both advective and dispersive transport
M	number of parameters to be estimated
n	unit vector normal to the boundary and pointing outwards
N_a	number of mobile primary species
N_C	number of primary species
N_C	number of aqueous primary species
N_E	number of atomic elements
NE	number of different types of observation data
n_i	number of moles of the i-th species
NL	number of observations
N_P	total number of minerals involved in dissol./prec. reactions
N_p	dissol./prec. reactions
N_w	ion exchange reactions
N_y	sorption reactions
N_R	number of chemical reactions
$N_S(J)$	number of primary species for adsorption reactions
N_T	total number of species
N_W	total number of species sorbed as ion exchange

N_X	number of aqueous secondary species
N_y	total number of species sorbed as surface complexes
N_0	number of species sorbed in the 0 layer
N_β	number of species sorbed in the β layer
P	Peclet number
P	parameter vector
P_i^*	estimated value of the parameter P_i
p	water pressure
p_f	partial pressure of the f-th species in the gas phase
Pg	partial pressure of gas species
P_i	concentration of the i-th dissolved/precipitated mineral
p_i	mass (moles) of the i-th mineral species per unit fluid volume
P_j	virtual concentration of the j-th primary species in the N_p minerals
PR	propionate in solution
Q	specified water flux
q	Darcy's velocity
$\mathbf{Q}_i$	chemical formula of the i-th species
$\|\mathbf{q}\|$	Darcy's flow modulus
q_x, q_y	Darcy's flow components
R	gas constant
R	conc. of solute involved in kinetic dis./prec.
R	solute sink/source term (mass added per unit time and unit fluid volume)
R_{diss}	dissolution reaction rate of the mineral
R_{net}	net reaction rate
R_j^p	generation/dissociation of primary species due to dissolution/precipitation reactions
R_j^w	generation/dissociation of primary species due to ion exchange reactions
R_j^y	generation/dissociation of primary species due to adsorption reactions
R_j^{total}	total source term ($R_j^p + R_j^w + R_j^y$)
$\mathbf{r}_t$	the residual between the model outcome and the actual measurement
r	areal fluid sink/source
r	reaction rate
r	vector of chemical potentials, μ_i divided by RT
r_i^p	dissolution/precipitation reaction rate
r_i^x	rate of progress of a reaction
r_i^y	adsorption reaction rate
S	storage coefficient
S	solid concentration
SI	mineral saturation index
SIA	sequential iteration approach
$\mathbf{S_i}$, $\mathbf{S_j}$	dissolved cation species
S_i^x	i-th secondary aqueous species
S_i^p	i-th mineral species
S_j	j-th primary species
S_m	factor for dissolution (-1)/ precipitation (+1) of the m-th reaction
S_s	specific storage coefficient
s_T	total concentration of surface sites in sorption reactions

S_y	specific yield
S(p)	the second order residual error
S_w	water saturation degree
S_r	residual water saturation degree
T	absolute temperature
$\mathbf{T}_l$	the Hessian matrix of $r_l(p)$
T	temperature
T	transmissivity tensor
T	matrix of stoichiometric coefficients of primary species
T	matrix $(N_R \cdot N_T)$ of stoichiometric coefficient of the i-th species in the j-th reaction
t	time
t	t-distribution
T_S	mean temperature of the solid
T_y	concentration of total sorption sites
$\boldsymbol{u}$	state variable computed by the numerical model
V	volume
$\mathbf{v}, \lvert v \rvert, (v_x, v_y)$	real water velocity vector, modulus, and components
v_c	linear growth rate
V_m	molar volume
v_s	volume of unit formula
w	specific fluid sink/source
w_i	concentration of the i-th exchanged species
$\overline{\mathbf{w}}_\mathbf{i}$	activity of the i-th exchanged species
W_j	virtual concentration of the j-th primary species in the N_w ion exchanged species
$\boldsymbol{W}_i$	weighting coefficient for each type of data and prior information
w_l	weighting coefficient for different observation points
X	mol fraction
X	eigenvector
x^2	chi-square minimization or chi-square distribution
x_i	concentration of the i-th aqueous secondary species
X_j	total concentration of the j-th aqueous primary species that appears in all secondary species
X_m	molar fraction of the m-th solid phase
X_{vi} - S_i	exchange sites which are occupied by the i-th cation
X_{vj} - S_j	exchange interlayer cation
y_i	concentration of the i-th sorbed species as surface complex ($L^{-3}M$)
Y_j	(J) virtual concentration of the j-th primary species in the N_y sorbed species (c) concentration of the j-th solute sorbed as surface complexes
z	topographic elevation
z	chemical valence of an element
z	electric charge
z_j	electrostatic charge of the surface complex
z_{ik}	chemical valence of the k-th element in the i-th species
z_k^*	maximum chemical valence of the k-th element
z_t, z_b	top and bottom aquifer elevation

APPENDIX 2

GENERATION OF SYNTHETIC DATA FOR EXAMPLES EJ2 AND EJ2KIN

The synthetic data of examples EJ2 and EJ2KIN in Chapter 4.2 are corrupted by some lognormal distribution noises. Since the values of the concentration data for different components at different time steps are variable from 10^{-2} to 10^{-8} (mol/l), adding simply a Gaussian noise to the true concentrations would lead to some negative concentrations. So log-normal distributed noise with mean μ_l=1 and different standard deviations (σ_l=0.01, 0.05, 0.1, 0.5) is multiplied to the true concentrations which were obtained from the forward modeling to form the synthetic concentration observation data by following steps (Law and Kelton, 1980):

1. Generate U_1 and U_2 as Independent Identically uniform Distribution IID U(0,1) random variables.
2. Set $X_1 = (-2 \ln U_1)^{0.5}\cos 2\pi U_2$ and $X_2 = (-2 \ln U_1)^{0.5}\sin 2\pi U_2$, then X_1 and X_2 are IID N(0,1) random variables.
3. Given $X \sim N(0,1)$, we can compute $Y \sim N(\mu,\sigma^2)$ by setting $Y = \mu + \sigma X$. In this way, we can obtain the desired random variables in pairs.
4. A special property of the lognormal distribution (LN) is that if $Y \sim N(\mu,\sigma^2)$ then $e^Y \sim LN(\mu,\sigma^2)$. Setting $Z = e^Y$, the lognormal distribution noises $Z \sim N(\mu,\sigma^2)$ have been generated with given mean μ_l and standard deviation σ_l. There are formulas for transforming the mean and variance of lognormal distribution to those of normal distribution as $\mu = \ln\left[\mu_l^2/(\sigma_l^2+\mu_l^2)^{0.5}\right]$ and $\sigma^2 = \ln\left[(\sigma_l^2+\mu_l^2)/\mu_l^2\right]$.
5. Multiply the true concentration by the generated random noise to obtain the noise-corrupted synthetic data.

The synthetic data of example EJ2 are listed in the Table A2.1

Table A2.1. True and Noise-Corrupted Concentration Data(μ_l=1)

Na^+(mol/l):

T(hours)	True Con.	σ_l=0.01	σ_l=0.05	σ_l=0.1	σ_l=0.5
0.5	0.10000E-02	0.99363E-03	0.96759E-03	0.93402E-03	0.66282E-03
1	0.10000E-02	0.98918E-03	0.94616E-03	0.89318E-03	0.53634E-03
1.5	0.10000E-02	0.10104E-02	0.10518E-02	0.11032E-02	0.14583E-02
2	0.10000E-02	0.98450E-03	0.92400E-03	0.85191E-03	0.42870E-03
2.5	0.10000E-02	0.98616E-03	0.93178E-03	0.86629E-03	0.46406E-03
3	0.10000E-02	0.10061E-02	0.10297E-02	0.10574E-02	0.11930E-02
3.5	0.10000E-02	0.10018E-02	0.10080E-02	0.10136E-02	0.97610E-03
4	0.10000E-02	0.10078E-02	0.10384E-02	0.10755E-02	0.12924E-02
4.5	0.10000E-02	0.99187E-03	0.95909E-03	0.91772E-03	0.60981E-03
5	0.10000E-02	0.10017E-02	0.10076E-02	0.10126E-02	0.97189E-03
6	0.10000E-02	0.98586E-03	0.93039E-03	0.86372E-03	0.45758E-03
7	0.99930E-03	0.99530E-03	0.97847E-03	0.95578E-03	0.74113E-03
8	0.98800E-03	0.10054E-02	0.10771E-02	0.11709E-02	0.20224E-02
9	0.84690E-03	0.83987E-03	0.81154E-03	0.77586E-03	0.51217E-03
10	0.38620E-03	0.38840E-03	0.39693E-03	0.40691E-03	0.45292E-03
11	0.10100E-03	0.10047E-03	0.98261E-04	0.95370E-04	0.70491E-04
12	0.17790E-04	0.17637E-04	0.17020E-04	0.16245E-04	0.10594E-04
13	0.19140E-05	0.19298E-05	0.19924E-05	0.20685E-05	0.25317E-05
14	0.28140E-06	0.28583E-06	0.30396E-06	0.32743E-06	0.52805E-06
15	0.11790E-06	0.11861E-06	0.12138E-06	0.12464E-06	0.14047E-06
17	0.10020E-06	0.10165E-06	0.10754E-06	0.11510E-06	0.17694E-06
19	0.10000E-06	0.10011E-06	0.10044E-06	0.10063E-06	0.94331E-07
21	0.10000E-06	0.98773E-07	0.93926E-07	0.88022E-07	0.50048E-07
23	0.10000E-06	0.10036E-06	0.10173E-06	0.10323E-06	0.10644E-06
25	0.10000E-06	0.10054E-06	0.10263E-06	0.10506E-06	0.11571E-06

K^+(mol/l):

T(hours)	True Con.	σ_l=0.01	σ_l=0.05	σ_l=0.1	σ_l=0.5
0.5	0.20000E-03	0.19873E-03	0.19352E-03	0.18680E-03	0.13256E-03
1	0.20000E-03	0.19784E-03	0.18923E-03	0.17864E-03	0.10727E-03
1.5	0.20000E-03	0.20207E-03	0.21035E-03	0.22065E-03	0.29167E-03
2	0.20000E-03	0.19690E-03	0.18480E-03	0.17038E-03	0.85740E-04
2.5	0.20000E-03	0.19723E-03	0.18636E-03	0.17326E-03	0.92811E-04
3	0.20000E-03	0.20121E-03	0.20593E-03	0.21149E-03	0.23860E-03
3.5	0.20000E-03	0.20036E-03	0.20161E-03	0.20271E-03	0.19522E-03
4	0.20000E-03	0.20155E-03	0.20768E-03	0.21509E-03	0.25849E-03
4.5	0.20000E-03	0.19837E-03	0.19182E-03	0.18354E-03	0.12196E-03
5	0.20000E-03	0.20034E-03	0.20151E-03	0.20253E-03	0.19438E-03
6	0.20000E-03	0.19717E-03	0.18608E-03	0.17274E-03	0.91515E-04
7	0.20070E-03	0.19990E-03	0.19652E-03	0.19196E-03	0.14885E-03
8	0.21200E-03	0.21574E-03	0.23112E-03	0.25125E-03	0.43395E-03
9	0.35300E-03	0.35007E-03	0.33826E-03	0.32339E-03	0.21348E-03
10	0.81300E-03	0.81764E-03	0.83559E-03	0.85660E-03	0.95345E-03
11	0.10980E-02	0.10922E-02	0.10682E-02	0.10368E-02	0.76632E-03
12	0.80530E-03	0.79836E-03	0.77042E-03	0.73535E-03	0.47957E-03
13	0.94910E-04	0.95695E-04	0.98797E-04	0.10257E-03	0.12554E-03
14	0.10520E-04	0.10686E-04	0.11364E-04	0.12241E-04	0.19741E-04
15	0.12350E-05	0.12425E-05	0.12714E-05	0.13056E-05	0.14714E-05
17	0.11670E-06	0.11839E-06	0.12525E-06	0.13406E-06	0.20608E-06
19	0.10030E-06	0.10041E-06	0.10074E-06	0.10093E-06	0.94614E-07
21	0.10000E-06	0.98773E-07	0.93926E-07	0.88022E-07	0.50048E-07
23	0.10000E-06	0.10036E-06	0.10173E-06	0.10323E-06	0.10644E-06
25	0.10000E-06	0.10054E-06	0.10263E-06	0.10506E-06	0.11571E-06

Ca^{+2}(mol/l):

T(hours)	True Con.	σ_I=0.01	σ_I=0.05	σ_I=0.1	σ_I=0.5
0.5	0.10000E-06	0.99363E-07	0.96759E-07	0.93402E-07	0.66282E-07
1	0.10000E-06	0.98918E-07	0.94616E-07	0.89318E-07	0.53634E-07
1.5	0.10000E-06	0.10104E-06	0.10518E-06	0.11032E-06	0.14583E-06
2	0.10000E-06	0.98450E-07	0.92400E-07	0.85191E-07	0.42870E-07
2.5	0.10000E-06	0.98616E-07	0.93178E-07	0.86629E-07	0.46406E-07
3	0.10000E-06	0.10061E-06	0.10297E-06	0.10574E-06	0.11930E-06
3.5	0.10000E-06	0.10018E-06	0.10080E-06	0.10136E-06	0.97610E-07
4	0.10000E-06	0.10078E-06	0.10384E-06	0.10755E-06	0.12924E-06
4.5	0.10000E-06	0.99187E-07	0.95909E-07	0.91772E-07	0.60981E-07
5	0.10000E-06	0.10017E-06	0.10076E-06	0.10126E-06	0.97189E-07
6	0.10000E-06	0.98586E-07	0.93039E-07	0.86372E-07	0.45758E-07
7	0.10030E-06	0.99898E-07	0.98210E-07	0.95932E-07	0.74388E-07
8	0.10500E-06	0.10685E-06	0.11447E-06	0.12444E-06	0.21493E-06
9	0.17030E-06	0.16889E-06	0.16319E-06	0.15601E-06	0.10299E-06
10	0.50030E-06	0.50315E-06	0.51420E-06	0.52713E-06	0.58673E-06
11	0.78100E-06	0.77687E-06	0.75982E-06	0.73746E-06	0.54508E-06
12	0.18850E-03	0.18687E-03	0.18034E-03	0.17213E-03	0.11226E-03
13	0.55170E-03	0.55626E-03	0.57430E-03	0.59624E-03	0.72975E-03
14	0.59470E-03	0.60407E-03	0.64239E-03	0.69198E-03	0.11160E-02
15	0.59940E-03	0.60302E-03	0.61709E-03	0.63365E-03	0.71415E-03
17	0.60000E-03	0.60867E-03	0.64396E-03	0.68924E-03	0.10595E-02
19	0.60000E-03	0.60065E-03	0.60263E-03	0.60377E-03	0.56599E-03
21	0.60000E-03	0.59264E-03	0.56355E-03	0.52813E-03	0.30029E-03
23	0.60000E-03	0.60218E-03	0.61038E-03	0.61937E-03	0.63864E-03
25	0.60000E-03	0.60325E-03	0.61580E-03	0.63038E-03	0.69427E-03

Cl^{-}(mol/l):

T(hours)	True Con.	σ_I=0.01	σ_I=0.05	σ_I=0.1	σ_I=0.5
0.5	0.10000E-06	0.99363E-07	0.96759E-07	0.93402E-07	0.66282E-07
1	0.10000E-06	0.98918E-07	0.94616E-07	0.89318E-07	0.53634E-07
1.5	0.10000E-06	0.10104E-06	0.10518E-06	0.11032E-06	0.14583E-06
2	0.10160E-06	0.10003E-06	0.93878E-07	0.86554E-07	0.43556E-07
2.5	0.16100E-06	0.15877E-06	0.15002E-06	0.13947E-06	0.74713E-07
3	0.97290E-06	0.97880E-06	0.10017E-05	0.10288E-05	0.11607E-05
3.5	0.64550E-05	0.64666E-05	0.65068E-05	0.65426E-05	0.63007E-05
4	0.28260E-04	0.28480E-04	0.29345E-04	0.30393E-04	0.36524E-04
4.5	0.85850E-04	0.85152E-04	0.82338E-04	0.78786E-04	0.52352E-04
5	0.19570E-03	0.19603E-03	0.19718E-03	0.19817E-03	0.19020E-03
6	0.54740E-03	0.53966E-03	0.50930E-03	0.47280E-03	0.25048E-03
7	0.89230E-03	0.88872E-03	0.87370E-03	0.85344E-03	0.66178E-03
8	0.10920E-02	0.11113E-02	0.11905E-02	0.12942E-02	0.22353E-02
9	0.11700E-02	0.11603E-02	0.11212E-02	0.10719E-02	0.70757E-03
10	0.11930E-02	0.11998E-02	0.12262E-02	0.12570E-02	0.13991E-02
11	0.11990E-02	0.11927E-02	0.11665E-02	0.11322E-02	0.83681E-03
12	0.12000E-02	0.11897E-02	0.11480E-02	0.10958E-02	0.71463E-03
13	0.12000E-02	0.12099E-02	0.12491E-02	0.12969E-02	0.15873E-02
14	0.12000E-02	0.12189E-02	0.12962E-02	0.13963E-02	0.22518E-02
15	0.12000E-02	0.12072E-02	0.12354E-02	0.12686E-02	0.14297E-02
17	0.12000E-02	0.12173E-02	0.12879E-02	0.13785E-02	0.21191E-02
19	0.12000E-02	0.12013E-02	0.12053E-02	0.12075E-02	0.11320E-02
21	0.12000E-02	0.11853E-02	0.11271E-02	0.10563E-02	0.60057E-03
23	0.12000E-02	0.12044E-02	0.12208E-02	0.12387E-02	0.12773E-02
25	0.12000E-02	0.12065E-02	0.12316E-02	0.12608E-02	0.13885E-02

APPENDIX 3

INFORMATION ON THE ITERATIVE PROCESS AND ESTIMATION RESULTS OF SYNTHETIC CASE EJ2 WITHOUT NOISE AND PRIOR INFORMATION

This information is extracted from the main output files of INVERSE-CORE^{2D}:

```
0. WRITE THE INITIAL PARAMETER AND THE OBJECTIVE VALUES ----

 NO. NIP   NAME          INITIAL VALUE       TRUE VALUE
  1  122   EKX(2)          0.1200E+00        0.1995E+00
  2  123   EKX(3)          0.4100E+00        0.3981E+00
  3  132   PC0(1,2)        0.3000E-02        0.1000E-02
  4  133   PC0(1,3)        0.3000E-03        0.2000E-03
  5  184   PCB(1,4)        0.5500E-03        0.6000E-03
  6  185   PCB(1,5)        0.1400E-02        0.1200E-02

--INITIAL VALUE OF OBJECTIVE FUNCTION IS: 0.46781E+04
--THE CONTRIBUTION TO OBJECTIVE OF EACH COMPONENT IS:
  Na+: 0.408E+04  K+: 0.460E+03  Ca++: 0.102E+03  Cl-: 0.407E+02

 1. THE NUMBER OF GAUSS-NEWTON ITERATIONS  1

 *****COVARIANCE MATRIX*****

  0.385E-05 -0.143E-05 -0.107E-08  0.590E-08  0.111E-08 -0.314E-12

 -0.143E-05  0.147E-02 -0.924E-08  0.246E-07  0.363E-07 -0.691E-10

 -0.107E-08 -0.924E-08  0.322E-10  0.521E-12 -0.377E-12  0.260E-15

  0.590E-08  0.246E-07  0.521E-12  0.164E-10  0.117E-11 -0.773E-16

  0.111E-08  0.363E-07 -0.377E-12  0.117E-11  0.263E-11 -0.127E-14

 -0.314E-12 -0.691E-10  0.260E-15 -0.773E-16 -0.127E-14  0.275E-10

 *****CORRELATION MATRIX*****

  0.100E+01 -0.190E-01 -0.963E-01  0.741E+00  0.347E+00 -0.305E-04

 -0.190E-01  0.100E+01 -0.425E-01  0.159E+00  0.585E+00 -0.344E-03

 -0.963E-01 -0.425E-01  0.100E+01  0.226E-01 -0.409E-01  0.873E-05

  0.741E+00  0.159E+00  0.226E-01  0.100E+01  0.178E+00 -0.364E-05
```

```
  0.347E+00  0.585E+00 -0.409E-01  0.178E+00  0.100E+01 -0.150E-03

 -0.305E-04 -0.344E-03  0.873E-05 -0.364E-05 -0.150E-03  0.100E+01
```

****EIGENVALUES ASSOCIATED TO THE EIGENVECTOR COMPONENTS****

EIGENVALUES	EIGENVECTOR COMPONENTS					
0.147E-02:	0.978E-03	-0.100E+01	0.631E-05	-0.168E-04	-0.248E-04	0.472E-07
0.385E-05:	0.100E+01	0.978E-03	-0.281E-03	0.154E-02	0.296E-03	-0.991E-07
0.321E-10:	-0.137E-03	-0.481E-05	-0.996E+00	-0.924E-01	-0.201E-02	0.277E-04
0.275E-10:	0.269E-07	-0.455E-07	-0.201E-04	-0.823E-04	-0.175E-04	-0.100E+01
0.695E-11:	0.146E-02	0.131E-04	0.899E-01	-0.973E+00	0.212E+00	0.746E-04
0.113E-11:	-0.620E-03	-0.285E-04	-0.216E-01	0.211E+00	0.977E+00	-0.340E-04

---- THE PARAMETERS AFTER MARQUARDT ITERATIONS ---- 0

NO.	NIP	NAME	ESTIMATED PARAM.	VARIANCE	CONFIDENCE INTERVAL(95%)
1	122	EKX(2)	0.1544E+00	0.3855E-05	(0.1475E+00, 0.1614E+00)
2	123	EKX(3)	0.5000E+00	0.1465E-02	(0.3642E+00, 0.6358E+00)
3	132	PC0(1,2)	0.9818E-03	0.3224E-10	(0.9617E-03, 0.1002E-02)
4	133	PC0(1,3)	0.2280E-03	0.1641E-10	(0.2136E-03, 0.2423E-03)
5	184	PCB(1,4)	0.6463E-03	0.2627E-11	(0.6406E-03, 0.6521E-03)
6	185	PCB(1,5)	0.1200E-02	0.2753E-10	(0.1181E-02, 0.1219E-02)

```
--MODIFIED OBJECTIVE FUNCTION (TLSQM,TLSQ2): 0.10909E+03 0.10909E+03
--THE CONTRIBUTION TO OBJECTIVE OF EACH COMPONENT IS:
  Na+: 0.138E+02  K+: 0.884E+02  Ca++: 0.687E+01  Cl-: 0.314E-04
```

.....

6. THE NUMBER OF GAUSS-NEWTON ITERATIONS 6

*****COVARIANCE MATRIX*****

```
 0.608E-11  0.399E-11 -0.150E-14  0.457E-14  0.734E-15 -0.199E-18

 0.399E-11  0.932E-10 -0.973E-14  0.838E-15  0.214E-13 -0.218E-16

-0.150E-14 -0.973E-14  0.155E-16  0.118E-17 -0.954E-18  0.225E-22

 0.457E-14  0.838E-15  0.118E-17  0.602E-17  0.695E-18 -0.301E-22

 0.734E-15  0.214E-13 -0.954E-18  0.695E-18  0.530E-17  0.961E-23

-0.199E-18 -0.218E-16  0.225E-22 -0.301E-22  0.961E-23  0.135E-16
```

*****CORRELATION MATRIX*****

0.100E+01	0.168E+00	-0.155E+00	0.756E+00	0.129E+00	-0.220E-04
0.168E+00	0.100E+01	-0.256E+00	0.354E-01	0.964E+00	-0.614E-03
-0.155E+00	-0.256E+00	0.100E+01	0.122E+00	-0.105E+00	0.156E-05
0.756E+00	0.354E-01	0.122E+00	0.100E+01	0.123E+00	-0.334E-05
0.129E+00	0.964E+00	-0.105E+00	0.123E+00	0.100E+01	0.114E-05
-0.220E-04	-0.614E-03	0.156E-05	-0.334E-05	0.114E-05	0.100E+01

****EIGENVALUES ASSOCIATED TO THE EIGENVECTOR COMPONENTS****

EIGENVALUES	EIGENVECTOR COMPONENTS					
0.934E-10:	-0.456E-01	-0.999E+00	0.105E-03	-0.112E-04	-0.230E-03	0.233E-06
0.590E-11:	0.999E+00	-0.456E-01	-0.179E-03	0.768E-03	-0.414E-04	0.135E-06
0.148E-16:	0.497E-04	0.773E-04	0.980E+00	0.173E+00	0.929E-01	-0.131E-02
0.135E-16:	0.584E-07	-0.245E-06	-0.124E-02	-0.233E-03	-0.514E-03	-0.100E+01
0.225E-17:	-0.766E-03	-0.423E-04	-0.189E+00	0.961E+00	0.203E+00	-0.929E-04
0.160E-18:	0.187E-03	-0.236E-03	-0.540E-01	-0.217E+00	0.975E+00	-0.384E-03

---- THE PARAMETERS AFTER MARQUARDT ITERATIONS ---- 0

NO.	NIP	NAME	ESTIMATED PARAM.	VARIANCE	CONFIDENCE INTERVAL(95%)
1	122	EKX(2)	0.1995E+00	0.6083E-11	(0.1995E+00, 0.1995E+00)
2	123	EKX(3)	0.3980E+00	0.9324E-10	(0.3980E+00, 0.3981E+00)
3	132	PC0(1,2)	0.1000E-02	0.1548E-16	(0.1000E-02, 0.1000E-02)
4	133	PC0(1,3)	0.2000E-03	0.6017E-17	(0.2000E-03, 0.2000E-03)
5	184	PCB(1,4)	0.6000E-03	0.5305E-17	(0.6000E-03, 0.6000E-03)
6	185	PCB(1,5)	0.1200E-02	0.1347E-16	(0.1200E-02, 0.1200E-02)

--MODIFIED OBJECTIVE FUNCTION (TLSQM,TLSQ2): 0.53380E-04 0.53380E-04
--THE CONTRIBUTION TO OBJECTIVE OF EACH COMPONENT IS:
Na+: 0.768E-06 K+: 0.189E-04 Ca++: 0.277E-05 Cl-: 0.310E-04

***** END *****

APPENDIX 4

INFORMATION ABOUT ITERATIVE PROCESS AND ESTIMATION RESULTS OF SYNTHETIC CASE EJ2 WITH PRIOR INFORMATION

This information is extracted from the main output files of INVERSE-CORE^{2D} when prior information is taken account into the objective function.

```
0. WRITE THE INITIAL PARAMETER AND THE OBJECTIVE VALUES

 NO. NIP   NAME         INITIAL VALUE   PRIOR INFORMATION
  1  122   EKX(2)         0.1200E+00       0.1994E+00
  2  123   EKX(3)         0.4100E+00       0.3982E+00
  3  132   PC0(1,2)       0.3000E-02       0.1002E-02
  4  133   PC0(1,3)       0.3000E-03       0.1997E-03
  5  184   PCB(1,4)       0.5500E-03       0.5996E-03
  6  185   PCB(1,5)       0.1400E-02       0.1201E-02

--INITIAL VALUE OF OBJECTIVE FUNCTION (TLSQI):  0.46825E+04
--THE CONTRIBUTION TO OBJECTIVE OF EACH COMPONENT IS:
  NA+: 0.408E+04   K+: 0.460E+03   CA++: 0.102E+03   CL-: 0.407E+02

 1. THE NUMBER OF GAUSS-NEWTON ITERATIONS  1

 *****COVARIANCE MATRIX*****

  0.545E-06 -0.211E-09 -0.155E-09  0.828E-09  0.163E-09 -0.557E-13

 -0.211E-09  0.127E-05 -0.831E-11  0.227E-10  0.318E-10 -0.597E-13

 -0.155E-09 -0.831E-11  0.312E-10  0.203E-11  0.127E-12 -0.265E-15

  0.828E-09  0.227E-10  0.203E-11  0.800E-11 -0.921E-12  0.154E-14

  0.163E-09  0.318E-10  0.127E-12 -0.921E-12  0.142E-11  0.526E-15

 -0.557E-13 -0.597E-13 -0.265E-15  0.154E-14  0.526E-15  0.270E-10

 *****CORRELATION MATRIX*****

  0.100E+01 -0.253E-03 -0.375E-01  0.397E+00  0.185E+00 -0.145E-04

 -0.253E-03  0.100E+01 -0.132E-02  0.713E-02  0.237E-01 -0.102E-04
```

```
-0.375E-01 -0.132E-02  0.100E+01  0.128E+00  0.191E-01 -0.911E-05
 0.397E+00  0.713E-02  0.128E+00  0.100E+01 -0.273E+00  0.105E-03
 0.185E+00  0.237E-01  0.191E-01 -0.273E+00  0.100E+01  0.850E-04
-0.145E-04 -0.102E-04 -0.911E-05  0.105E-03  0.850E-04  0.100E+01
```

****THE EIGENVALUES ASSOCIATED TO THE EIGENVECTOR COMPONENTS****

```
EIGENVALUES                        EIGENVECTOR COMPONENTS
 0.127E-05:  0.290E-03 -0.100E+01  0.650E-05 -0.177E-04 -0.250E-04  0.470E-07
 0.545E-06: -0.100E+01 -0.290E-03  0.284E-03 -0.152E-02 -0.299E-03  0.102E-06
 0.314E-10: -0.143E-03 -0.484E-05 -0.996E+00 -0.914E-01 -0.218E-02  0.297E-04
 0.270E-10:  0.206E-07 -0.452E-07 -0.223E-04 -0.816E-04 -0.177E-04 -0.100E+01
 0.679E-11: -0.144E-02 -0.129E-04 -0.889E-01  0.973E+00 -0.213E+00 -0.737E-04
 0.111E-11: -0.619E-03 -0.285E-04 -0.216E-01  0.211E+00  0.977E+00 -0.340E-04
```

---- THE PARAMETERS AFTER MARQUARDT ITERATIONS ---- 0

NO.	NIP	NAME	ESTIMATED PARAM.	VARIANCE	CONFIDENCE INTERVAL(95%)
1	122	EKX(2)	0.1930E+00	0.5446E-06	(0.1904E+00, 0.1956E+00)
2	123	EKX(3)	0.3990E+00	0.1271E-05	(0.3950E+00, 0.4030E+00)
3	132	PCO(1,2)	0.9769E-03	0.3121E-10	(0.9570E-03, 0.9967E-03)
4	133	PCO(1,3)	0.2701E-03	0.8003E-11	(0.2601E-03, 0.2802E-03)
5	184	PCB(1,4)	0.6355E-03	0.1418E-11	(0.6312E-03, 0.6397E-03)
6	185	PCB(1,5)	0.1200E-02	0.2700E-10	(0.1182E-02, 0.1218E-02)

--MODIFIED OBJECTIVE FUNCTION (TLSQM,TLSQ2) ARE: 0.10777E+03 0.10764E+03
--THE CONTRIBUTION TO OBJECTIVE OF EACH COMPONENT IS:
NA+: 0.102E+02 K+: 0.886E+02 CA++: 0.885E+01 CL-: 0.351E-04

.....

5. THE NUMBER OF GAUSS-NEWTON ITERATIONS 5

*****COVARIANCE MATRIX*****

```
 0.679E-12  0.871E-14 -0.799E-16  0.543E-15 -0.718E-17  0.276E-18
 0.871E-14  0.141E-11 -0.200E-15 -0.132E-15  0.413E-15 -0.612E-18
-0.799E-16 -0.200E-15  0.316E-16  0.376E-17  0.351E-17 -0.685E-20
 0.543E-15 -0.132E-15  0.376E-17  0.539E-17  0.154E-17 -0.426E-20
```

```
-0.718E-17  0.413E-15  0.351E-17  0.154E-17  0.156E-17  0.142E-19
 0.276E-18 -0.612E-18 -0.685E-20 -0.426E-20  0.142E-19  0.303E-16


*****CORRELATION MATRIX*****
 0.100E+01  0.890E-02 -0.173E-01  0.284E+00 -0.697E-02  0.608E-04
 0.890E-02  0.100E+01 -0.299E-01 -0.478E-01  0.278E+00 -0.935E-04
-0.173E-01 -0.299E-01  0.100E+01  0.288E+00  0.499E+00 -0.221E-03
 0.284E+00 -0.478E-01  0.288E+00  0.100E+01  0.529E+00 -0.333E-03
-0.697E-02  0.278E+00  0.499E+00  0.529E+00  0.100E+01  0.206E-02
 0.608E-04 -0.935E-04 -0.221E-03 -0.333E-03  0.206E-02  0.100E+01

****THE EIGENVALUES ASSOCIATED TO THE EIGENVECTOR COMPONENTS****
EIGENVALUES                          EIGENVECTOR COMPONENTS
0.141E-11: -0.119E-01 -0.100E+01  0.142E-03  0.889E-04 -0.292E-03  0.431E-06
0.679E-12:  0.100E+01 -0.119E-01 -0.114E-03  0.802E-03 -0.178E-04  0.418E-06
0.325E-16:  0.159E-05  0.117E-03  0.983E+00  0.142E+00  0.120E+00 -0.262E-02
0.303E-16: -0.309E-06  0.576E-06  0.253E-02  0.251E-03  0.823E-03  0.100E+01
0.471E-17:  0.775E-03  0.105E-04  0.170E+00 -0.947E+00 -0.272E+00  0.309E-04
0.685E-18:  0.236E-03 -0.318E-03 -0.748E-01 -0.288E+00  0.955E+00 -0.524E-03

---- THE PARAMETERS AFTER  MARQUARDT ITERATIONS ----  0

 NO. NIP   NAME       ESTIMATED PARAM.    VARIANCE     CONFIDENCE INTERVAL(95%)
  1  122   EKX(2)        0.1994E+00     0.6787E-12   (0.1994E+00, 0.1994E+00)
  2  123   EKX(3)        0.3982E+00     0.1411E-11   (0.3982E+00, 0.3982E+00)
  3  132   PC0(1,2)      0.1000E-02     0.3158E-16   (0.1000E-02, 0.1000E-02)
  4  133   PC0(1,3)      0.1999E-03     0.5391E-17   (0.1999E-03, 0.1999E-03)
  5  184   PCB(1,4)      0.6000E-03     0.1562E-17   (0.6000E-03, 0.6000E-03)
  6  185   PCB(1,5)      0.1200E-02     0.3035E-16   (0.1200E-02, 0.1200E-02)

--MODIFIED OBJECTIVE FUNCTION (TLSQM,TLSQ2) ARE: 0.12110E-03 0.11603E-03
--THE CONTRIBUTION TO OBJECTIVE OF EACH COMPONENT IS:
  NA+: 0.424E-05   K+: 0.784E-04   CA++: 0.242E-05   CL-: 0.310E-04

***** END *****
```

APPENDIX 5

INPUT DATA AND OUTPUT DESCRIPTION OF INVERSE-CORE^{2D}

This appendix contains the description of input data and main output files for INVERSE-CORE^{2D}, a code developed during the course of this dissertation. Data input of this code has the same structure as that of CORE^{2D} (Samper et al., 2000).

A5.1. GROUNDWATER FLOW AND HEAT AND SOLUTE TRANSPORT DATA

Data for groundwater flow and solute transport are read from a file ROOT_TRA.INP where "ROOT" refers to the common roots of all the files used for a given simulation. The name of this file must contain no more than 20 characters.

A5.1.1. GROUNDWATER FLOW AND SOLUTE TRANSPORT DATA

Input data for flow and transport of INVERSE-CORE^{2D} is read from unit 8. The input data are writen into a file (unit 9) ROOT_TRA.OUT which reproduces the content of the input file.

The first card is used for writing the title, some comments and remarks; it is followed by 20 data cards. Some cards must be omitted in cases where data are not required for a particular problem. It should be noticed that there is always a blank line after each card that can be used for auxiliary comments. Some variables in data cards are not required under certain conditions. In such cases one should leave them blank or input any value. Data values must be provided using standard FORTRAN format.

The content of each card is presented, indicating the name and description of each variable as well as its corresponding FORTRAN format.

The internal length unit used by INVERSE-CORE^{2D} is dm. For time and mass units the user is free to use any system units as long as they are consistent among them.

Card_1. Title

Variable: TITLE
Format: A76

TITLE: Title of the problem, comments and remarks. This title is reproduced in the output files.

Card_2. General options for flow and transport

Variables: IOTPA IOFLU IODIM XITA1 XITA2 IHEAT IRESTART
Format: I5 I5 I5 F10.3 F10.3 I5 I5

IOTPA : integer indicator of the type of medium
0 = confined aquifer
1 = unconfined aquifer
2 = variably saturated flow in vertical cross-sections

IOFLU : integer indicator for flow regime
0 = steady state
1 = transient state

IODIM : integer indicator for problem dimensionality
2 = two-dimensional flow
3 = three-dimensional axisymmetric flow. In this case the X coordinate must be the horizontal (radial) coordinate

XITA1 : coefficient for the type of numerical scheme used for solving the flow equation
1.0 = implicit
0.0 = explicit
between 1.0 and 0.0 = mixed scheme
XITA1=1.0 (implicit) is suggested. If IOFLU=0, XITA1 is not required (leave it blank).

XITA2 : coefficient for the type of numerical scheme used for solving transport equations
1.0 = for an implicit scheme (unconditionally stable)
0.0 = for an explicit scheme (conditionally stable)
between 1.0 and 0.0 = mixed scheme
XITA2=1.0 (implicit) is suggested

IHEAT : integer indicator for heat transport simulation
0 = not considered
1 = considered

IRESTART: indicator for using the restart procedure. Computed results are saved in a file named RESTART.DAT which can be used later to restart the simulation.
0 = unused (usually this is the case)
1 = used

Card_3. Options for chemical calculations

Variables: ISPIA INIBOUND
Format: I5 I5

ISPIA: indicator for using a sequential partly-iterative approach (SPIA) according to which after one transport iteration, the chemical model is only called for the nodes not satisfying a prescribed convergence tolerance
0 = not using SPIA. In this case the general sequential iteration approach is used.
1 = using SPIA
ISPIA = 0 is suggested

INIBOUND: indicator used for the definition of the chemical composition of

boundary and recharge inflows. i.e. chemical initialization of the boundary and recharge waters.
If INIBOUND=0 The options of Card_26 of the chemical input file are not activated for boudnary waters, and therefore chemical speciation of these waters is not performed.
If INIBOUND is different than zero, options of Card_26 are activated and therefore chemical speciation of these waters can be performed. Use of INIBOUND=1 is suggested.

Card_4. Input and output file names (except flow and solute transport input and output files)

Variables:

IOWRITE(1), INPUThet	I5, A20
IOWRITE(2), INPUTche	I5, A20
IOWRITE(3), OUTche	I5, A20

Names of files used for input of heat transport data (INPUThet), chemical input data (INPUTche) and chemical output results (OUTche). Each name must be in a separate line.

INPUThet is only needed when IHEAT in Card 2 is equal to 1. This line has to be always present in the input file.
INPUTche is the name of the geochemical input file. This file is always required, therefore a name must be provided always.
OUTche is the name of the geochemical output file. This file is always required, therefore a name must be provided always.

Each line must contain in the first column an integer value which will indicate wether a given file is needed or not. If this integer is 1 the file will be generated otherwise it will not.

Variables:

IOWRITE(4), OUThx: Output of nodal hydraulic head in dm (pressure head for variably saturated media) values at specified time steps (which are specified in Card_19): Integer code and file name. (I5, A20)

IOWRITE(5), OUTht: Output of hydraulic head (in dm) versus time (in day) at selected nodes (which are specified in Card_20): Integer code and file name. (I5, A20)

IOWRITE(6), OUTvt: Output of computed volumes (in liter) of water at selected nodes (which are specified in Card_21) as a function of time: Integer code and file name. (I5, A20)

IOWRITE(7), OUTft: Output file name of water flux (liter per day) versus time (in day) at selected nodes (which are specified in Card_21): Integer code and file name. (I5, A20)

IOWRITE(8), OUTwt: Output file containing the total volume of water

stored in selected material zones (which are specified in Card_19) as a function of time: (I5, A20)

IOWRITE(9), OUTtex: Output file name of temperature nodal values at specified time steps (which are specified in Card_19): (I5, A20)

IOWRITE(10), OUTtet: Output file name of temperature versus time at selected nodes (which are specified in Card_20): (I5, A20)

IOWRITE(11), OUTspx: Output file name of nodal values of total dissolved concentrations (molality) of selected components (which are specified in Card_22) at specified time steps (which are specified in Card_19): (I5, A20)

IOWRITE(12), OUTspt: Output file name of total dissolved concentrations (molality) of selected components (which are specified in Card_22) versus time at selected nodes (which are specified in Card_20): (I5, A20)

IOWRITE(13), OUTphx: Output file name of pH nodal values at specified time steps (which are specified in Card_19 (I5, A20)

IOWRITE(14), OUTpht: Output file name of pH values versus time at selected nodes (which are specified in Card_22): (I5, A20)

IOWRITE(15), OUTmix, KLITER: Output file name of nodal values of cumulative precipitation (accumulated from time equal to zero) of selected mineral phases (which are specified in Card_23) at specified times (which are specified in Card_19). Variable KLITER may have values of 0 or different from 0. In case KLITER equal to 0, the nodal cumulative dissolution/precipitation will be expressed in mol mineral/dm**3 dissolution, which is adequate for saturated conditions. Otherwise in mol mineral/kg dry soil, more appropriate for non-saturated conditions. File OUTmit is congruent with the units specified through KLITER. (I5, A20, I5)

IOWRITE(16), OUTmit: Output file name of cumulative precipitation of selected mineral phases (which are specified in Card_23) versus time at selected nodes (which are specified in Card_20. (I5, A20)

IOWRITE(17), OUTpex: Output file name of pE nodal values at specified times (which are specified in Card_19 (I5, A20)

IOWRITE(18), OUTpet: Output file name of pE value versus time at selected nodes (which are specified in Card_20) (I5, A20)

IOWRITE(19), OUTadx: Output file name of nodal values of total sorbed component concentration at specified times (which are specified in Card_19). (I5, A20)

IOWRITE(20), OUTadt: Output file name of total sorbed component concentration versus time at selected nodes (which are specified in Card_20). (I5, A20).

IOWRITE(21), OUTiter: Output file name containing information on iterative processes. (I5, A20).

IOWRITE(22), OUTcsec: Output file name of computed concentrations (molality) of dissolved primary, secondary species and selected gases at selected nodes (which are specified in Card_20) versus time.-(I5, A20). Notice that for moderately large problems, this file may be exceedingly large.

IOWRITE(23), OUTint: Output file name containing the computed nodal concentrations of exchanged species at selected times (which are specified in Card_19). (I5, A20).

IOWRITE(24), OUTsix: Output file name containing the nodal saturation indices of minerals at selected times (which are specified in Card_19). (I5, A20).

IOWRITE(25), OUTsit: Output file name containing the saturation indices of minerals at selected times for given nodes (which are specified in Card_19). (I5, A20).

The information contained in this file is written in files OUTspx (nodal ionic strength) and OUTcsec (activity coefficients)

IOWRITE(26), OUTmasol: Output file containing the total mass of dissolved components in the whole system versus time. (I5, A20). Computed values include only the mass of dissolved components. Possible mass transfer to/from mineral phases or exchange complex is not taken into account.

IOWRITE(27), OUTmprec: This file remains unused in the current release of INVERSE-CORE^{2D}. That means that no information will be dumped into it even thought its corresponding line must be kept. (I5,A20).

IOWRITE(28), OUTnodbal, INODBAL: Output file containing a summary of solute and water mass (useful for variably saturated media). The file contains: (1) nodal values of mass of aqueous components, (2) total volume of water in the system, (3) total solute mass in the system for selected aqueous components and (4) nodal values of water content, (5) total mass present in the system as minerals, (6) total mass present in the system as exchanged species. For mass balances, units are mol. INODBAL may have values of 0 or different from 0. In case INODBAL is equal to 0, only total mass balance checks will be printed (total masses of selected components in the water, minerals, and exchanger as well as the amount of mass entering and/or leaving the system within a selected time period). Otherwise nodal mass balance are written. In the latter case, OUTnod can be extremely large. (I5, A20, I5)

IOWRITE(32), OUTpeclet.OUT: This file contains the maximum dimensionless Peclet and Courant numbers for a given time period. (I5, A20).

IOWRITE(29), LOG1.OUT, ITERAT, ITERSTOP: This is file is a diagnostic where the values of aqueous species concentrations, activity and activity coefficients are dumped for each chemical iteration for a given node and time step. The file is rewritten

for each node and each time step. So, activating this file translates into a really poor performance of INVERSE-CORE^{2D} in terms of time execution. It can be used to identify: (1) species giving problems in terms of chemical convergence; (2) nodes where chemical problems arise; (3) time periods for which chemistry cannot be properly solved. Variables ITERAT and ITERSTOP are provided to indicate the code the number of time period when start writing the LOG1.OUT file and the last time period when this file need to be written. In general, it is easy to identify the time period in which the code crashes. Therefore this time period minus 1 is a good point to start writing this file. ITERAT may take the value of 0. In this case, the file LOG1.OUT will contain chemical information belonging to time 0 (initial conditions). No matter the filename given to this file, the output file will be called LOG1.OUT. (I5, A20, I5, I5)

IOWRITE(30), LOG2.OUT: In this file are dumped the saturation indices of the minerals of the problem for each chemical iteration in a given node and time period. This file may be as lenghtly as the LOG1.OUT file. If the problem does not contain minerals, an error will be generated and, eventually, the execution will crash. Variables ITERAT and ITERSTOP operates over the time periods when this file is going to be generated. General strategy to use this file is similar to that of the LOG1.OUT file. No matter the filename given to this file, the output file will be called LOG2.OUT (I5, A20).

IOWRITE(31), LOG3.OUT, SELECTR: In this file is dumped the incremental value of the sink/source term (R) or the total solute affected by transport (UT), i.e. the value of array R (or UT) at iteration i minus the same value for iteration i-1. This file has as many blocks as trasnport iterations needed to achieve convergence in all nodes. The whole file is rewritten for each time period. No matter the filename given to this file, the output file will be called LOG3.OUT (I5, A20, 4X, A1).

Card_5. Dimension variables

Variables:	NNOD	NELE	NMA	NTTIME
Format:	I5	I5	I5	I5

NNOD : number of nodes in the finite element grid
NELE : number of triangular elements of the grid
NMA: number of aquifer material zones (parameter zones)
NTTIME: number of time periods

Card_6. Time stepping data

Repeat this card as many times as the number of time periods, NTTIME, defined in Card_5.

Variable: ITI TIMEINT NSTEP
Format: I5 F10.4 I5

ITI : time period number
TIMEINT : duration of time period [T]
NSTEP : number of time steps within this time period.

The duration TIMEINT is divided into NSTEP equal time increments. Flow and transport equations are solved at each one of the intermediate time increments (see Figure 3.1.1).

Card_7. Number of stepwise time functions for boundary and recharge conditions

Variables: NBOUNDFH NBOUNDFC
Format: I5 I5

NBOUNDFH : number of time functions for groundwater recharge and flow boundary conditions; if boundary parameters do not vary with time, NBOUNDFH is equal to zero. In this case Card_8 is not required.

NBOUNDFC : number of time functions for solute and heat transport boundary conditions. If they are constant in time, NBOUNDFC is equal to zero; in this case, Card_9 is not required.

Card_8. Stepwise time functions for areal recharge and water flow boundary conditions

This card must be omitted if NBOUNDFH=0 in Card_7. Repeat this card NBOUNDFH times, one for each time function.

Variables: I (BOUNDFH(I, J), J=1,NTTIME)
Format: I5 F10.0

BOUNDFH(I, J): value of the i-th time function at the j-th time period

Card_9. Stepwise time functions for solute and heat transport

This card must be omitted if NBOUNDFC=0 in Card_7. Repeat this card NBOUNDFC times, one per time function.

Variables: I (BOUNDFC(I, J), J=1,NTTIME)
Format: I5 F10.0

BOUNDFC(I, J): value of the i-th function at the j-th time period. Use the same format as in Card_8.

Card_10. Parameters of aquifer material zones

Repeat this card NMA times, one per aquifer material zone. A material zone is a set of one or more triangular elements within which the parameters take constan values.

Variables: I PK1(I) PK2(I) ANGLE(I) SS(I) POR(I) DFM(I) DSL(I) DST(I) DENSEC(I) PORCIN(I)

Format: I5 10E9.2

I : material zone number

PK1(I) : first principal component of the hydraulic conductivity tensor [L/T]

PK2(I) : second principal component of the hydraulic conductivity tensor [L/T]

ANGLE(I): Angle (in degrees) between the direction of the first component and the x-coordinate (counter-clock wise)

SS(I): specific storage coefficient for confined (IOTPA=0) or unsaturated flow (IOTPA=2) [1/L]; It is equal to the specific yield for unconfined flow when IOTPA=1 [dimensionless]

POR(I): total transport porosity [dimensionless]

DFM(I): molecular difussion coefficient [L^2/T]

DSL(I): longitudinal dispersivity [L]

DST(I): transverse dispersivity [L]

DENSEC(I): dry density [M/L^3]. This datum is used to compute the retardation coefficient. It is only needed for sorbing solutes. Any coherent set of mass units (g, kg) are allowed.

PORCIN(I): accesible porosity [dimensionless]

For steady-state flow the storage coefficient SS is not required. One can leave SS blank or assign any value. One must use 'dm' as the unit of length [L]. That is required because concentrations are expressed as mol/L in the chemical model. Notice that for confined flow or unsaturated conditions SS represents the specific storage coefficient. The program computes the storage coefficient by multiplying SS and the thickness of the aquifer. For unconfined aquifers, SS represents the storage coefficient which approximately coincides with specific yield.

Card_11. Unsaturated flow parameters

This card must be provided only when IOTPA = 2 (unsaturated flow), otherwise it must be omitted. Repeat this card NMA times, one per material zone .

Variables: I , ITPRL, SR(I), SAS(I), CM(I), ALPHA(I)

Format: 2I5 E9.2 E9.2 E9.2 E9.2

I : material zone number

ITPRL: Integer for relative permeability function:

= 1 for van Genuchten equation

= 2 for Irmay equation

SR(I): residual saturation degree

SAS(I): saturation degree at saturated conditions, normally it is assumed to be equal to 1

CM(I): value of parameter m in van Genuchten retention curve

ALPHA(I): value of parameter α in van Genuchten retention curve [L-1]

Notice that the basic unit length in the program is dm.

Card_12. Default values of data defined on triangular elements

Variable: MATDF THICKDF RECHDF IRECHDF

Format: I5 F10.3 F10.3 I5

MATDF, THICKDF, RECHDF and IRECHDF are default values of variables MAT, THICK, RECH and IRECH which are defined in the next card, respectively. Default values are assigned to all elements. The next card allows the user to specify different values of these parameters at selected elements.

Card_13. Data for triangular elements

This card is used to provide the values of the parameters and variables that are defined over the elements. These variables include aquifer thickness and areal recharge. Aquifer parameters are assumed constant within material zones. The number of these zones (NMA) is defined in Card_5 and the parameter values of each zone are specified in Card_10. One must indicate the zone number to which a given element belongs to.

One must provide information for all elements (NELE). However, if the properties of the elements (parameters and geometry) is the same for a string of elements, one only has to input the properties of the first element of the string. The properties of the remaining elements are generated automatically by the code. One should always input the information of at least the first and last elements of the grid. The test examples attached to this manual illustrate this feature.

Variable: IE (NODE(IE,I),I=1,3) MAT(IE) THICK(IE) RECH(IE), IRECH(IE)
Format: I5 3I5 I5 F10.0 F10.0 I5

IE: element number

NODE(IE, 3): the three node (vertice) numbers of element IE (in a counterclock wise manner)

MAT(IE): the number of material zone in which element IE is located

THICK(IE):

= aquifer thickness for flow in 2-D horizontal confined flow
= base elevation for 2-D unconfined flow
= 1 for 2-D vertical confined flow or variably saturated flow
= blank or any value for 3-D axisymmetric flow

RECH(IE): rate of areal recharge (possitive for recharge and negative for discharge) [L/T]

IRECH(IE): the number of recharge time function (listed in Card_8) used for element IE. In the program, the amount of recharge is computed as the product of RECH times the time function, BOUNDFH. Leave IREACH blank if areal recharge is constant in time.

Card_14. Default values for nodal data

Variables: IDBHDF HPDF Q1DF ALFADF IQDF H0DF IDBOCDF

IIWDF IBWDF IRWDF IMIDF IGSDF IADDF IEXDF
Format: I3 3F7.0 I3 F10.0 8I3

IDBHDF, HPDF, Q1DF, ALFADF, IQDF, H0DF, IDBOCDF, IIWDF, IBWDF, IRWDF, IMIDF, IGSDF, IADDF and IEXDF are default nodal values of the variables IDBH, HP, Q1, ALFA, IQ, H0, IDBOC, IZONEIW, IZONEBW, IZONERW, IZONEM, IZONEG, IZONED and IZONEX which are defined in the next card. Default values are assigned automatically to all the nodes of the grid. However, changes to these values can be made with the next card.

Card_15. Nodal data

This card is used to specify the values of properties and parameters which are defined in a nodewise manner. These properties include the nodal coordinates (x,y) as well as nodal boundary conditions and parameters for water flow and solute transport. One must provide these data for all the nodes (NNOD). However, if there is a string of equally spaced nodes having the same properties, the user only has to input the information for the first node of the string. Data for the remaining nodes are generated automatically by the code. Notice that the chemical parameters and properties are defined on a nodewise manner. These properties are assumed constant within zones which consist of a set of grid nodes. There are zones for:

* initial chemical composition
* the chemistry of water flowing into the system along the boundaries
* the chemistry of areal recharge water
* mineral phases gas phases
* sorption processes and properties
* ion exchange processes and properties

Variables: IP XX(IP) Y(IP) IDBH(IP) HP(IP) Q1(IP) ALFA(IP) IQ(IP) H0(IP) IDBOC(IP) IZONEIW(IP) IZONEBW(IP) IZONERW(IP) IZONEM(IP) IZONEG(IP) IZONED(IP) IZONEX(IP)
Format: I5 2F10.0 I3 3F7.0 I3 F7.0 8I3

IP: node number
XX(IP): X coordinate of node IP [L]
Y(IP): Y coordinate of node IP [L]
IDBH(IP): integer indicator of the type of flow boundary condition
1-prescribed head
2-prescribed flux (including no flux). Use this condition for nodes not lying along the boundary.
3-mixed Cauchy condition
4-free drainage (only for variably saturated flow)
HP(IP): when IDBH(IP)=1 represents the prescribed hydraulic head (pressure head for unsaturated flow); when IDBH(IP)=3 represents the external reference hydraulic head; Otherwise, HP is not required, and can be left blank
Q1(IP): prescribed water flux [L^3/T], (positive for inflow); only required

when IDBH(IP)=2; otherwise leave blank or put any value

ALFA(IP): leakage coefficient [L^2/T]; only required when IDBH(IP)=3 , otherwise leave blank or put any value. Notice that ALFA(IP) represents the lumped value of the leakage coefficient for node IP.

IQ(IP): the number of stepwice time function (listed in Card_8) used for the boundary condition of node IP; in the program, BOUNDFH multiplies either HP or Q1 depending on the type of flow boundary condition; if the boundary condition does not change with time, input a zero value

H0(IP): it represents:

1. initial hydraulic head for confined flow
2. initial pressure head for unsaturated flow)
3. prior estimation of head for transient unconfined
4. steady-state hydraulic heads for steady-state flow

In all other cases, leave it blank or input any value.

IDBOC(IP): integer indicator of the type of solute transport boundary condition

1-prescribed concentration

2-solute flux associated with water flux (including no flux)

3-prescribed solute flux

IZONEIW(IP): zone number of the initial chemical compisition.

IZONEBW(IP): zone number for the chemistry of boundary inflows (including inflows at internal nodes). Leave it blank if there are no inflow zones.

IZONERW(IP): zone number for the chemistry of recharge water. Leave it blank if there is no areal recharge.

IZONEM(IP): mineral zone number to which the node belongs to. Leave it blank if there are no mineral zones.

IZONEG(IP): gas zone number to which the node belongs to. Leave it blank if there are no gas zones.

IZONED(IP): adsorption zone number to which the node belongs to. Leave it blank if there are no sorption zones.

IZONEX(IP): ion exchange zone number to which the node belongs to. Leave it blank if there are no exchange zones.

Card_16. Data for free drainage boundary nodes

This card is needed for calculating the free drainage boundary distance. Therefore, it is only needed for unsaturated zones having a free drainage boundary condition. Repeat as many times as nodes having a boundary condition IDBH=4 (see the previous card)

Variables:	NODF	NODF1	NODF2
Format:	I5	I5	I5

NODF : free drainage node number

NODF1 : number of the first neighbour node (on the free drainage boundary) of node NODF

NODF2 : number of the second neighbour node of node NODF

If one neighbour node is out of the free drainage boundary, input a zero value.

Card_17. Data for convergence criteria

Variables:MAXITPFL TOLFL MAXITPTR TOLTR MAXITPCH TOLCH MAXITPAD TOLAD

Format: I5 E10.3 I5 E10.3 I5 E10.3 I5 E10.3

MAXITPFL : maximum allowed number of iterations for solving the water flow equation; It is needed for unconfined and unsaturated flow;

TOLFL: relative tolerance for convergence of head (or pressure head) values. A value between 10^{-3} and 10^{-6} is suggested.

MAXITPTR: maximum allowed number of iterations for solving transport and geochemistry; set MAXITPTR=1 for the sequential non-iterative approach where transport and chemistry are sequentially solved only once.

TOLTR: relative tolerance for convergence of concentration values during transport calculations. A value between 10^{-3} and 10^{-6} is suggested.

MAXITPCH: maximum allowed number of iterations for solving the geochemical equations.

TOLCH: relative tolerance for convergence of concentration values during the solution of the chemical system. A value between 10^{-3} and 10^{-6} is suggested.

MAXITPAD: maximum allowed number of iterations for solving the equations sorption via surface complexation.

TOLAD: relative tolerance of concentration convergence for solving sorption reactions. A value between 10^{-3} and 10^{-6} is suggested.

Card_18. Control variables for Jacobian matrix

Variables: ISWITCH, ITEMP, NJACOB

ISWICHT: base switch option. ISWICHT = 1 when switching is performed otherwise, leave it blank. Use of ISWICHT = 0 is recommended.

ITEMP: Integer for thermal options (= 0 for constant temperature at 25 °C; = 1 for variable temperature)

NJACOB: The Jacobian matrix for chemical speciation is updated every NJACOB iterations. (Use of NJACOB = 1 is recommended).

Card_19. Control variables for output printout

Variables: NWXY NWDIM NWTI NWNOD NWCOM NWMIN INDMAT NVOL NWTV

Format: 10I5

NWXY : time frequency for writing spatial nodal results

NWDIM: options for spatial results

1= results versus distance x for 1-D problems

2= results at all nodes for 2-D problems

NWTI : time frequency for writing the time evolution of results at some selected nodes (NWNOD)

NWNOD : number of nodes at which time evolution will be printed out

NWCOM: number of chemical components for which results will be printed out

NWMIN : number of minerals for which results will be printed out

INDMAT: aquifer material zone number for which the time evolution of water volume will be printed out.

NVOL: number of nodes for which the time evolution of water content will be printed out.

NWTV: time frequency for writing water content

Card_20. Node numbers at which results are printed out

Variables: (IWNOD(I),I=1,NWNOD)

Format: 15I5

IWNOD(I): Vector of node numbers at which time evolution of results are printed out. Omit this card if NWNOD = 0.

Card_21. Node numbers at which results of water content will be printed out variable

Variable: (INVOL(I), I=1, NVOL)

Format: 15I5

INVOL(I): Vector of node numbers.
Omit this card if NVOL = 0.

Card_22. Data related to exchanged cations

Repeat this card as many times as exchanged cations are considered

Variable: NAEXC(K) IMS(K) IEX(K) (AEKX(K,L), L=1,4)

Format: A20 I4 I4 Free

NAEXC(K): Name of the k-th cation of the solution (either primary or aqueous complexes) able to be exchanged with surface cations.
Omit NAEXC if no exchangeable cations are required. However, the species '*' is always needed to indicate the end of the list.

IMS(K): Integer indicator taking a value of 1 for the cation used as reference for the exchange reactions (usually Na) and 0 for the rest of cations.

IEX(K): Exchange convention type used for exchange calculations:
1= Gaines-Thomas; 2= Vanselow; 3= Gapon
The value of IEX(K) must be the same for all the exchanged cations.

AEKX(K,L): Exchange coefficient of the k-th cation with respect to the reference cation or coefficients for the exchange isotherm. In case that isotherms are not used, AEKX(1) stores the selectivity coefficient and the

rest of values of AEKX (2 through 4) must be 0.0d0. For the reference cation AEKX(K,L) =1.0 0.0 0.0 0.0. These coefficients are the parameters define polynomial function in order to calculate the logarithm of exchange coeffient.

Card_23. Minerals for which results will be printed out

Variable: (IWMIN(I),I=1,NWMIN)
Format: 15I5

IWCOM(I): Vector of mineral identifiers for which results will be printed out. Omit this card if NWMIN=0.

A5.1.2. HEAT TRANSPORT DATA

This input file is required if heat transport is selected. Heat transport data are read from unit # 7, and are written to unit # 9. Heat input file consists of 4 data cards. There is a comment line before each data card (that can be left blank).

Card_1. Physical and thermal parameters of water

Variables: RHOW CATW
Format: F10.0 F10.0

RHOW: Water density [g/dm3]
CATW: Specific heat capacity of water which is assumed constant with temperature [cal/g C°]

Card_2. Thermal aquifer parameters

Variables: IA RHOS CATS ITLAMD AMBDA0 AMBD AMBW TIN DSLT(IA) DSTT(IA)
Format: I5 F10.0 F10.0 I5 6F10.0

IA: order of medium parameter zone
RHOS: density of solid phase [g/dm3]
CATS: Specific heat capacity of the solid phase [cal/g C°]
ITLAMB: Integer code used to specify thermal conductivity laws. Its values are:
ITLAMB = 1 for $\lambda = \lambda_0$

ITLAMB = 2 for $\lambda = \lambda_w^{S_e^n} \cdot \lambda_d^{(1-S_e)^{1/n}}$

where S_e is effective water saturation degree.
AMBDA0: Isotropic thermal conductivity, λ_0, [cal/dm T °C]
AMBD: Dry thermal conductivity, λ_d (W/m °C)
AMBW: Saturated thermal conductivity, λ_w (W/m °C)
TIN: Parameter n (dimensionless)
DSLT(IA): longitudinal thermal dispersivity [dm]
DSTT(IA): transverse thermal dispersivity [dm]

Card_3. Default nodal heat transport data

Variables: IDBOTDF TPDF TFLUXDF TALFADF ITFLUXDF
Format: I5 F10.0 F10.0 F10.0 I5

These variables are the default values of IDBOT, TP, TFLUX, TALFA and ITFLUX which are defined in the next card.

Card_4. Nodal heat transport data

Repeat as many times as nodes are in the grid, NNOD
Variables: IP IDBOT(IP) TP(IP) TFLUX(IP) TALFA(IP) ITFLUX(IP)
Format: I5 I5 F10.0 F10.0 F10.0 I5

IP: node number
IDBOT(IP): integer indicator of the type of thermal boundary condition
1-prescribed temperature
2-prescribed heat flux (including no flux)
3-mixed Cauchy condition
TP(IP): represents prescribed temperature when IDBOT(IP)=1;It is the external reference temperature when IDBH(IP)=3; Otherwise it is not required, leave it blank or put any value
TFLUX(IP): heat flux not carried by water [cal./unit time], (positive for inflow); it is required only when IDBOT(IP)=2; Otherwise leave it blank or put any value; The heat flux associated with water flux is automatically accounted for by the program
TALFA(IP): leakage coefficient for mixed boundary; only required when IDBOT(IP)=3, otherwise leave it blank or put any value; Similar to variable ALFA defined in Card_15 for water flow, variable TALFA(IP) represents a lumped value of the leakage coefficient for node IP.
IFLUX(IP): the number of piecewice time function (listed in card_9) used for node IP; The time function BOUNDFC multiplies either TP or TFLUX depending on the thermal boundary condition. If heat boundary conditions does not vary with time, input a zero value.

A5.2. HYDROGEOCHEMICAL DATA

Card-1. Title

Variable: TITLE
Format: A76 (write the content of TITLE within 'single quotes')

TITLE: title, coments and remarks (in one single line)

Card-2. Label

Variable: LABEL
Format: A76 (write the content of LABEL within 'single quotes')

LABEL: comments which will appear in the output file

A5.2.1. DEFINITION OF THE GEOCHEMICAL SYSTEM

These cards contain the information on aqueous species, minerals, gases, surface complexes and exchangeable cations. The names of the aqueous species, minerals and gases must be taken from those listed in *master25.dat* and *mastertemp.dat* (variable temperature) databases.

Card-3. Label

Variable: LABEL
Format: A76 (write the content of LABEL within 'single quotes')

LABEL: comments which will appear in the output file

Card-4. Initial temperature

Variable: TC2
Format: free

TC2: Reference temperature of the system (°C)

A5.2.1.1. Primary aqueous species

Card_5. Label

Variable: LABEL
Format: A76 (write the content of LABEL within 'single quotes')

LABEL: comments which will appear in the output file

Card_6. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments

Card_7. Number of primary species

Variable: NDSI
Format: Free

NDSI: Number of primary species

Card_8. Label

Variable: label name of species
Format: A76 (write LABEL within 'single quotes').

Card_9. Primary aqueous species

Repeat Card_9 as many times as needed to describe all the aqueous components of primary species.
Variable: NAPRI(K) (ALAMDA (K, I), I=1,NDSI)
Format: A20 (write the content of NAPRI within 'single quotes')

NAPRI(k): Name of the k-th primary species as it appears in the databases. The species '*' is needed to indicate the end of the list.
ALAMDA: Decay constant of primary species

A5.2.1.2. Aqueous complexes

Card_10. Label

Variable: LABEL
Format: A76 (write the content of LABEL within 'single quotes')

LABEL: comments which will appear in the output file

Card_11. Aqueous complexes

Repeat Card-8 as many times as needed to describe all the aqueous secondary species (aqueous complexes). These species may contain also dissolved gases such as 'O_2(aq)'.
Variable: NAAQX(K)
Format: A20 (write the content of NAAQX within 'single quotes')

NAAQX(K): Name of the k-th aqueous complex.
Omit NAAQX if no aqueous complexes are present.
However, the species '*' is always needed to indicate the end of the list.

A5.2.1.3. Minerals

Card_12. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_13. Mineral names

One should provide the name of the mineral and the option for either chemical equilibrium or kinetics. Use a line per mineral phase present in the system. If one or more minerals are kinetically controlled, at the end of the list, the user should define the time units used in the problem within single quotes. The following units are allowed: 'second', 'minute', 'hour', 'day', 'month' and 'year'. If the time unit used in the problem is `day', one can skip this line.

Variables: NAMIN(K), INDICATOR(K)
Format: A20 (write NAMIN within 'single quotes'), free integer

NAMIN(K): Name of the K-th mineral phase. Omit NAMIN if no minerals are present. However, the species '*' is always needed to indicate the end of the list.
INDICATOR(K): Integer variable taking the following values
= 0 when the K-th mineral is at local equilibrium
> 0 when the K-th mineral is kinetically controlled.
An example of this card is given below:

```
'MINERALS'
'quartz'   1
'day'
'*'
```

A5.2.1.4. Gases

Card_14. Label
Variable: LABEL
Format: A76 (write LABEL within 'single quotes')
LABEL: comments that will appear in the output file

Card_15. Gases
Repeat this card as many times as needed
Variable: NAGAS(K)
Format: A20 (write NAGAS within 'single quotes')

NAGAS(K): Name of the k-th gas.

Omit NAGAS if no gases are required. However, the species '*' is always needed to indicate the end of the list.

A5.2.1.5. Surface complexes

Card_16. LABEL
Variable: LABEL
Format: A76 (write LABEL within 'single quotes')
LABEL: comments that will appear in the output file

Card_17. Surface complexes
Repeat this card as many times as needed
Variable: NAADS(k)
Format: A20 (write NAADS within 'single quotes')

NAADS(K): Name of the k-th surface complex.
Omit NAADS if no surface complexes are required. However, the species '*'

is always needed to indicate the end of the list.

A5.2.1.6. Exchangeable cations

Card_18. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_19. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments which will appear in the output file

Card_20. Data related to exchanged cations

Repeat this card as many times as exchanged cations are considered
Variable: NAEXC(K) IMS(K) IEX(K) (AEKX(K,L), L=1,4)
Format: A20 I4 I4 Free

NAEXC(K): Name of the k-th cation of the solution (either primary or aqueous complexes) able to be exchanged with surface cations.
Omit NAEXC if no exchangeable cations are required. However, the species '*' is always needed to indicate the end of the list.

IMS(K): Integer indicator taking a value of 1 for the cation used as reference for the exchange reactions (usually Na) and 0 for the rest of cations.

IEX(K): Exchange convention type used for exchange calculations:
1= Gaines-Thomas; 2= Vanselow; 3= Gapon
The value of IEX(K) must be the same for all the exchanged cations.

AEKX(K,L): Exchange coefficient of the k-th cation with respect to the reference cation or coefficients for the exchange isotherm. In case that isotherms are not used, AEKX(1) stores the selectivity coefficient and the rest of values of AEKX (2 through 4) must be 0.0d0. For the reference cation AEKX(K,L) =1.0 0.0 0.0 0.0.

A5.2.2. INITIAL, BOUNDARY AND RECHARGE WATER SOLUTIONS

Card_21. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_22. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_23. Initial and boundary aqueous solutions

Variable: NIWTYPE NBWTYPE NRWTYPE
Format: I4 I4 I4

NIWTYPE: Number of different aqueous solutions initially present in the system
NBWTYPE: Number of boundary solutions (including pumping/injection at the internal nodes)
NRWTYPE: Number of different areal recharge solutions

A5.2.2.1. Chemical composition of initial and boundary aqueous solutions

This describes the way to specify the chemical composition of initial and boundary aqueous solutions. One should repeat Cards 24, 25 and 26 as many times as the total number of initial and boundary aqueous solutions which is given by NIWTYPE + NBWTYPE + NRWTYPE. One should first input the composition of the NIWTYPE initial solutions, then the composition of the NBWTYPE boundary solutions and finally those of the RRWTYPE areal recharge solutions.

Card_24. Identification of the solution

Variables: IWTYPE TC2 ITC2
Format: I4 Free I10

IWTYPE: Number of order of the initial, boundary or areal recharge solution. For initial aqueous solutions, the value of IWTYPE varies from 1 to NIWTYPE.

TC2: Temperature of the aqueous solution (°C)
ITC2: The number of stepwise time function (listed in Card_9 of water flow and transport input file), used for time-varying boundary and areal recharge conditions. For the initial water solution, and for constant (in time) thermal boundary and recharge conditions, one should set ITC2 equal to 0

Card_25. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_26. Chemical composition of the aqueous solutions

Variables:	NAPRI(K)	ICON(K)	CGUESS(K)	CTOT(K)	NADUM(K)	ICTOT(K)
Format:	A20	I4	Free	Free	A20	I10

NAPRI(K): Name of the k-th primary aqueous species.

The name of the species must coincide with those previously listed as primary species in the definition of the chemical system, although the order may change. Names must be included between 'single quotes'. The species '*' indicates the end of the list of species.

ICON(K): Integer flag variable indicating the procedure used for computing the concentration of primary species from the total dissolved concentrations, CTOT.
1 = the concentration of the k-th primary species is constrained by the total concentration, CTOT, except for water which is assumed equal to 1.
2 = the concentration of the k-th primary species is equated to the concentration of another species of opossite charge. For instance, the concentrtion of Na can be fixed to be equal to that of Cl.
3 = the activity of the k-th species is assumed initially equal to CTOT.
4 = the concentration of the k-th species is calculated from the assumption of equilibrium of the solution with respect to the mineral NADUM(K).
5 = the concentration of the k-th primary species is calculated from the assumption of equilibrium with respect to a given partial pressure of the gas NADUM(K).
CGUESS(K): Initial guess for the concentration (mol/kg) of the k-th primary species

CTOT(K): Total dissolved component concentration (mol/kg). For ICON=5, CTOT is equal to the partial pressure of the gas used to constrain the concentration of the species. For a prescribed solute flux boundary condition (IDBOC=3 in Card_14 of the transport input file), CTOT is equal to the solute flux (mol per unit time).
NADUM(K): Name of the mineral (ICON(K)=4) or gas (ICON(K)=5) used to constrain the concentration of the k-th primary species. Names must be included between 'single quotes'. NADUM(K) should coincide with one of those previously listed as minerals or gases in the definition of the chemical system.

ICTOT(K) The number of stepwise time function (listed in Card_9 of water flow and transport input file), used for time-varying boundary and recharge conditions. For initial water solution, and for constant (in time) transport boundary and recharge conditions, one should set ICTOT(K) equal to 0

A5.2.3. INITIAL MINERAL ZONES

This section describes the different mineral zones initially present in the system.

Card_27. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_28. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_29.

Variable: NMTYPE
Format: I4

NMTYPE: Number of mineral zones in the system.

The following three cards must be repeated NMTYPE times.

Card_30.

Variable: IMTYPE
Format: I4

IMTYPE: Number of order of the mineral zone

Card_31. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_32. Composition of mineral zones

Variable: NAMIN(K) VOL(K) AREAM2(K)
Format: A20 Free Free format

NAMIN(K): Name of the k-th mineral involved in the lithology. The name of the mineral must be one of those previously listed in the definition of the chemical system, although the order may change. It is not needed to repeat the complete list. Names must be included between 'single quotes'. The mineral '*' indicates the end of the list of species.
VOL(K): Volume fraction of the k-th mineral.
AREAM2(K): Specific surface of the mineral (dm^2 of mineral surface per dm^3 of porous medium).

A5.2.4. DATA RELATED TO ADSORTION COEFFICIENTS

Card_33. Label
Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

Card_34. Distribution coefficient
Variable: LABEL. Format: A76 (write within 'single quotes')
Variable: XKD. Format: Free

LABEL: comments of XKD
XKD: Distribution coefficient of primary species (NDSI)

A5.2.5. INITIAL GAS ZONES

This section describes the different gas zones initially present in the system.

Card_35. Label
Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_36. Label
Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_37.
Variable: NGTYPE
Format: I4
NGTYPE: Number of gas zones present in the system.

The following three cards must be repeated NGTYPE times.

Card_38.
Variable: IGTYPE
Format: I4

IGTYPE: Number of order of the gas zone

Card_39. Label
Variable: LABEL

Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_40. Composition of gas zones

Variable: NAGAS(k) VOLG(k)
Format: A20 Free

NAGAS(k): Name of the k-th gas involved in the system. The name of the gas must be one of those previously listed in the definition of the chemical system, although the order may change. It is not needed to repeat the complete list. Names must be included between 'single quotes'. The gas '*' indicates the end of the list of species.

VOLG(k): Partial pressure of the k-th gas.

A5.2.6. SURFACE ADSORPTION ZONES

This section describes the characteristics of the zones with different surface adsorption properties.

Card_41. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_42. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_43.

Variable: NDTYPE
Format: I4

NDTYPE: Number of surface adsorption zones.

Card_44. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_45. Data for adsorption zones

This card must be repeated NDTYPE times. Omit this card if NDTYPE is zero. No

'*' is required to indicate the end of the list of adsorption zones.

Variable: IDTYPE SUPADS TSS
Format: I4 Free Free

IDTYPE: Number of order for the surface adsorption zone.

SUPADS: Specific adsorbent surface of the solid phase per unit of volume of solution (m^2 of surface per kg of solid).

TSS: Total adsorption sites per unit volume of solution (mol/dm^3)

A5.2.7. CATION EXCHANGE ZONES

This section describes the characteristics of cation-exchange zones.

Card_46. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_47. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_48.

Variable: NXTYPE
Format: I4

NXTYPE: Number of cation exchange zones.

Card_49. Label

Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_50. Data related to the cation exchange zone

This card and card_51 must be repeated NXTYPE times. Omit this card if NXTYPE is zero. No '*' is required to indicate the end of the list of cation exchange zones.
Variable: IXTYPE
Format: I4

IXTYPE: Number of order for the cation exchange zone.

Card_51. Data related to the cation exchange zone

This card must be repeated NEXC times. Omit this card if NXTYPE is zero. No '*' is required to indicate the end of the list of cation exchange zones.

Variable: NAEXC(K) CECA(K)
Format: Free Free
CECA(K): Exchange capacity of the k-th cation (meq/100 g soil).

A5.2.8. END OF INPUT DATA

This section allows the user to be sure that the chemical data have been read entirely.

Card_52. Label
Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: comments that will appear in the output file

Card_53. Label to check the end of chemical data input
Variable: LABEL
Format: A76 (write LABEL within 'single quotes')

LABEL: This label must be 'end'.

A5.3. INVERSE MODELING INPUT DATA

Data for inverse modeling of groundwater flow and reactive solute transport are read from a file ROOT_PAR.INP where "ROOT" refers to the common roots of all the files used for a given simulation.

A5.3.1. MAIN CONTROL PARAMETERS

The input data for inverse modeling of flow and reactive transport of INVERSE-CORE^{2D} is read from unit number 48. Data values can be provided with free format. The input data are printed into a file (unit 46) ROOT_PAR.OUT which reproduces the content of the input file.

The first card is used for writing the title and some comments and remarks. It is followed by 12 data cards. Some cards must be omitted in cases where data are not required for a particular problem. It should be noticed that there is always a blank line or a line with star symbols after each card that can be used for auxiliary comments. Some variables in data cards are not required under certain conditions. In such cases one only need to input the titles of the cards..

Next we describe the content of each card, indicating the name and description

of each variable as well as its corresponding meanings.

Card_1. Title and output file name

Variable: TITLE ,OUTNAM

Format: A76 (one line for Title),
A6 (OUTNAM),
A76 (another line for descreption or a blank line),

TITLE: Title of the inverse problem, comments and remarks. This title is reproduced in the output files.

OUTNAM: A root name for five output files.

Card_2. General control or options for the inverse optimization process

Variables: NMA, NPA, NTIME, KSPE0, NOSP, INVERS, ITRAN

NMA: parameter zone number or aquifer material zones.

NPA: the parameter number to be estimated.

NTIME: the time step number of the observations.

KSPE0: the species number for which total concentration data (including solid phase) are available.

NOSP: the total number of chemical components for which concentration data are available.

INVERS : integer indicator of FORWARD or INVERSE Modeling
0 = Forward modeling using the initial parameter
1 = Inverse modeling by Gauss-Newton-Marquardt method

ITRAN : integer indicator for ONLY flow modeling or flow and reactive transport modeling
0 = For modeling only the water flow
1 = For water flow and reactive solute transport modeling

Card_3. Iteration contral parameters

Variables: NITMAX, MAQMAX, FUNCER1, FUNCER2, SDob,SDp, IBP, CONF

NITMAX: the maximum number of Gauss-Newton iterations

MAQMAX: the maximum number of Marquardt-Levenberg iterations

FUNCER1:the tolerance error for the relative change in objective function.

FUNCER2:the tolerance error for the relative change in parameter values.

SDob: standard deviation of synthetic observation data.

SDp: standard deviation of prior parameter information.

IBP : integer indicator of optimum parameter upgrade vector
0 = optimizing parameter upgrade vector.
1 = not optimizing parameter upgrade vector.

CONF: percentage of parameter confidence interval, 90% or 95%.

Card_4. The initial weighting coefficients and the observation number

Variables: WCOEF1, WCOEF2, WCOEF3, WCOEF4, WPARINF
NIWH, NIWC, NIWAT

WCOEF1: initial weighting coefficient for the water head data
-1 = not using the water head data
0 = including water head data, but, not taking account them into the objective function.
1 = only water head data in the objective function and the coefficient is constantly and equal to 1.0.
>1 including more than one type of data in the objective function and the weighting coefficient will be modified during the iterative process.

WCOEF2: initial weighting coefficient for the concentration data
-1 = not including concentration data
0 = including concentration data, but, not taking them account into the objective function.
1 = only including concentration data in the objective function and the coefficient is constant and equal to 1.0.
>1 including more than one kind of observations in the objective function and the weighting coefficient will be modified during the iterative process.

WCOEF3: initial weighting coefficient for the water infiltration data
-1 = not including water infiltration data
0 = including water infiltration data, but, not taking account them into the objective function.
1 = only including water infiltration data in the objective function and the coefficient is constant and equal to 1.0.
>1 including more than one kind of observations in the objective function and the weighting coefficient will be modified during the iterative process.

WCOEF4: initial weighting coefficient for the water content data
-1 = not including water content data
0 = including water content data, but, not taking account them into the objective function.
1 = only including water content data in the objective function and the coefficient is constant and equal to1.0.
>1 including more than one kind of observations in the objective function and the weighting coefficient will be modified during the iterative process.

WPARINF: weighting coefficient for prior parameter information (≤ 1)

NIWH: number of nodes at which hydraulic head data are available
NIWC: number of nodes at which concentration data are available
NIWAT:number of nodes at which water content data are available

Card_5. Parameter identification numbers

Variables: (NIP(I), I = 1, NPA)

NIP(I) : identification number of the i-th parameter

NPA: number of parameters to be estimated

The full description of the identification number of all the NPA parameters to be estimated by INVERSE-CORE^{2D} is listed in USER'S MANUAL (Dai & Samper, 1999).

Card_6. Lower and upper bounds, initial values and the maximum relative increments of parametrs

Variable: (PMIN(I), I = 1, NPA)
(PMAX(I), I = 1, NPA)
(PINI(I), I = 1, NPA)
(PINF(I), I = 1, NPA)
(WPINF(I), I = 1, NPA)

(Use five lines, one per variable)

PMIN(I): lower bounds of parameters
PMAX(I):upper bounds of parameters
PINI(I): initial values of parameters
PINF(I): prior parameter information.
WPINF(I): initial weight for each piece of prior information.

Card_7. Options for parameter transformation and difference method for the Jacobian matrix

Variables: (IPLOG(I), I = 1, NPA)
(ICENT(I), I = 1, NPA)
(ALFAP(I), I = 1, NPA)

IPLOG(I): options for parameter transformation to logarithm:
0 = no log transformation
1 = log-parameter transformation is used for the optimization process.

ICENT(I): options to choose forward or central difference methods to calculate the Jacobian matrix:
0 = using forward difference method
1 = using central difference method

ALFAP(I): maximum relative increments of the parameters

One should provide values of IPLOG, ICENT and ALFAP for all the NPA parameters.

A5.3.2. MAIN INPUT OBSERVATION DATA

Card_8. Input of hydraulic head data

Title: Input of hydraulic head data

If WCOEF1 is equal to or greater than zero, it means that there are hydraulic head data to be input in this Card. Otherwise, one needs a title line and a blank or star symbol line.

Variables: (IWH(I), I=1,NIWH)
(HOB(I, J), I=1, NIWH , J=1,NTTIME)
(WHOB(I, J), I=1, NIWH , J=1,NTTIME)

IWH(I): node numbers at which data are available (from one to the NIWH).
HOB(I, J): hydraulic head at the i-th observation point and the j-th time period.
WHOB(I, J): weighting coefficient for hydraulic head at the i-th observation point and the j-th time period.

Card_9. Input of concentration data

Title: Input of observation points and measured concentration data

If WCOEF2 is equal to or greater than zero, it means that there are concentration data to be inputed in this Card. Otherwise, one only needs to include a title line and a blank or star symbol line.

Variables: (IOSP(I), I=1, NOSP)
(IWCON(I), I=1, NIWC)
(((CONOB(I, K, J), I=1, NIWC), K=1, NOSP), J=1, NTTIME)
(((WCONOB(I, K, J), I=1, NIWC), K=1, NOSP), J=1, NTTIME)

IOSP(I): chemical species number for which data are available (from one to the NOSP).
IWCON(I): node numbers at which chemical data are available (from one to the NIWC).
CONOB(I, K, J): concentration at the i-th observation points, k-th chemical species and the j-th time period.
WCONOB(I, K, J): initial weighting coefficient for concentration data at the i-th observation points, k-th chemical species and the j-th time period.

If the number of the chemical species is greater than one, the program will automatically recalculate the weighting coefficients for different species according to equation (3.2.3).

Card_10. Input of water flux data

Title: Input of observation data of water flux

If WCOEF3 is equal to or greater than zero, it means that there are water infiltration data to be input in this Card. Otherwise, one only needs to include a title line and a blank or star symbol line.

Variables:
(WATINOB(J), J=1,NTTIME)
(WWATINOB(J), J=1,NTTIME)

WATINOB(J): cumulative water infiltration data in the j-th time period.
WWATINOB(J): weighting cofficient for the cumulative water infiltration data in the j-th time period.

Card_11. Input of final water content data

Title: Input of observations of final water content data

If WCOEF4 is equal to or greater than zero, it means that there are water content data to be inputed in this Card. Otherwise, one only needs to include a title line and a blank or star symbol line.

Variables: WAT0, PD0
(IWAT(I), I = 1, NIWAT)
(WATCONOB(I), I=1,NIWAT)
(WWATCONOB(I), I=1,NIWAT)
(FINPD(I), I = 1, NIWAT)

WAT0: initial water content of the sample
PD0: initial dry density of the sample
IWAT(I): node number at which water content data are available (from one to the NIWAT).
WATCONOB(I): final water content at the i-th observation point.
WWATINOB(I): weighting cofficient for final water content at the i-th observation point.
FINPD(I): final dry density at the i-th observation point.

A5.3.3. TOTAL CONCENTRATION OBSERVATION DATA

Card_12. Input other control parameters and total concentration data

Title: Input of other control parameters and total concnetration observations

Variables: WCOEF5, IEXCLUSION

WCOEF5: initial coefficient for the total concentration data
-1 = not including total concentration data
0 = including the data, but, not taking them account into the objective function.
1 = only total concentration data included in the objective function and the coefficient is constant and equal to be 1.0.
>1 including more than one kind of observations in the objective function and the weighting coefficient will be modified with the iteration.

IEXCLUSION: integer indicator for anion exclusion
0 = there is no anion exclusion.
1 = including anion exclusion.

If WCOEF5 is equal to or greater than zero, it means that there are total concentration observations, as well as the related data, to be input in here.

Variables: NIWTC, PD1
(IWTC(I), I = 1, NIWTC)
(TCONOB(I), I = 1, NIWTC)
(WTCONOB(I), I = 1, NIWTC)

NIWTC: number of total concentration observation nodes
PD1: dry density of the sample
IWTC(I): node numbers at which total concentration data are available (from one to the NIWTC).
TCONOB(I): total concentration at the i-th observation point.
WTCONOB(I): weighting cofficient for total concentration data at the i-th observation point.

If WCOEF5 is less than zero, it means that there no total concentration data. So, this part is not necessary.

The last line of the input data file can be a blank line or a line with 'end'.

A5.4. OUTPUT DESCRIPTION

INVERSE-CORE^{2D} has up to 36 output files. Three of them are used to check input data. One of them presents the estimated results, as well as the main features of the optimization process. Another one includes the fitting results. The rest of output files are the same as those of CORE^{2D} (Samper et al., 2000).

A5.4.1. OUTPUT FILES FOR CHECKING INPUT DATA

Data check of inverse modeling in unit 48, ROOT_PAR.OUT: It includes control parameters, prior information of the parameters to be identified and the five kinds of observation data, as well as the corresponding weighting coefficients.

A5.4.2. OUTPUT FILES FOR PARAMETER ESTIMATION RESULTS

The program automatically generates the following five output files with the name OUTNAM:

(1) Estimation results in unit 6, OUTNAM_RESLT.OUT: It contains the results of the optimization process and the final results. It starts with the initial parameters and the corresponding initial value of objective function. Then it records all the iterative process, including the parameter modifications, confidence intervals, values of the objective function, weighting coefficients of different kinds of observations and information about the Marquardt-Levenberg iterations.
(2) Fitting results in unit 79, named OUTNAM_FITCN.OUT: It contains measured and calculated concentrations and fitting errors at different observation points.
(3) Fitting results in unit 69, named OUTNAM_FITHH.OUT: It contains measured and calculated hydraulic heads and fitting errors at different observation points.
(4) Estimation process and errors in unit 66, named OUTNAM_ITERA.OUT: It includes variance-covariance matrix, correlation matrix, eigenvalues and eigenvectors, as well as estimated parameter confidence intervals.
(5) Computed sensitivity coefficients in unit 59, named OUTNAM_SENSI.OUT: It contains sensitivity coefficient matrix in each iteration.

A5.4.3. OUTPUT FILES OF FORWARD PROBLEM

These output files are the same as those of CORE^{2D}. They include:

(1) General output of transport in unit 9: Including mainly flow, solute and heat transport input data, general control data and error messages

(2) General output of geochemical data in unit 2

(3) Output of nodal hydraulic head (pressure head for variably saturated media) values at specified time steps in unit 61

(4) Output of hydraulic head versus time at selected nodes in unit 63

(5) Output of water volume vs. Time in unit 71

(6) Output of mass flow in given nodes vs. Time in unit 75

(7) Output of total water flux vs. Time in unit 41

(8) Output of temperature nodal values at specified time steps in unit 81

(9) Output of temperature versus time at selected nodes in unit 83

(10) Output of nodal values of total dissolved component concentration at specified time steps in unit 10

(11) Output of total dissolved component concentration versus time at selected nodes in unit 15

(12) Output of pH nodal values at specified time steps in unit 11

(13) Output of pH value versus time at selected nodes in unit 16

(14) Output of nodal values of total mineral precipitation (accumulated from time zero) at specified time steps in unit 12

(15) Output of total mineral precipitation versus time at selected nodes in unit 17

(16) Output of pE nodal values at specified time steps in unit 14

(17) Output of pE value versus time at selected nodes in unit 19

(18) Output of nodal values of total sorbed component concentration at specified time steps in unit 24

(19) Output of total sorbed component concentration versus time at selected nodes in unit 29

(20) output of modeling progress and iteration information in unit 39

(21) Output of aqueous species concentration vs. Distance in unit 76

(22) Output of exchange complex evolution vs. Distance in unit 72

(23) Output of saturation indexes of minerals vs. Distance in unit 26.

(24) Output of saturation indexes of minerals vs. Time at selected nodes in unit 27.

(25) Output of solute mass balance (kg) in unit 28.

(26) Output of mass of mineral diss./prec. In unit 78

(27) Output of nodal mass at volume balance in unit 88

(28) Output of maximum peclet and courant number in unit 89

(29), (30) and (31) Output of iteration inform

Not all these files are always used. There is no file (4) under state-steady flow conditions. Neither (8) nor (9) are needed if heat transport is not included. Neither (12) nor (13) are needed if H^+ is not involved in the chemical system. Neither (16) nor (17) are needed if electrons e^- are not considered. Files (14) and (15) are not needed if no mineral precipitation or dissolution processes are considered. Similarly, (18) and (19) are not required if there is no ion exchange and adsorption by surface complexation.

Output file (21) is necessary if computed concentrations (molality) of dissolved primary, secondary species and selected gases at selected nodes versus time are wanted. Notice that for moderately large problems, this file may be exceedingly large. Output file (22) is necessary if the computed nodal concentrations of exchanged species at selected times are wanted. Output files (23) and (24) are necessary if the nodal saturation indices of minerals at selected times and the saturation indices of minerals at selected times for given nodes are wanted. Output file (25) contains the total mass of dissolved components in the whole system versus time. Computed values include only the mass of dissolved components. Possible mass transfer to/from mineral phases or exchange complex is not taken into account. The file (26) remains unused in the current release of INVERSE-CORE^{2D}. That means that no information will be dumped into it even thought its corresponding line must be kept. Output file (27) contains a summary of solute and water mass (useful for variably saturated media). The file contains: (1) nodal values of mass of aqueous components, (2) total volume of water in the system, (3) total solute mass in the system for selected aqueous components and (4) nodal values of water content, (5) total mass present in the system as minerals, (6) total mass present in the system as exchanged species. For mass balances, units are mol. In the latter case, (27) can be extremely large. The file (28) contains the maximum dimensionless Peclet and Courant numbers for a given time period.

File (29) is a diagnostic file where the values of aqueous species concentrations, activity and activity coefficients are dumped for each chemical iteration for a given node and time step. The file is rewritten for each node and each time step. So, activating this file translates into a really poor performance of INVERSE-CORE^{2D} in terms of time execution. It can be used to identify: (1) species giving problems in terms of chemical convergence; (2) nodes where chemical problems arise; (3) time periods for which chemistry cannot be properly solved. Variables ITERAT and ITERSTOP (see Chapter 4) are provided to indicate the code the number of time period when start writing the LOG1.OUT file and the last time period when this file need to be written. In general, it is easy to identify the time period in which the code crashes. Therefore this time period minus 1 is a good point to start writing this file. ITERAT may take the value of 0. In this case, the file LOG1.OUT will contain chemical information belonging to time 0 (initial conditions). No matter the filename given to this file, the output file will be called LOG1.OUT

In file (30) the saturation indices of the minerals of the problem are dumped for each chemical iteration in a given node and time period. This file may be as length as the LOG1.OUT file. If the problem does not contain minerals, an error will be generated and, eventually, the execution will crash. Variables ITERAT and ITERSTOP operates over the time periods when this file is going to be generated. General strategy to use this file is

similar to that of the LOG1.OUT file. No matter the filename given to this file, the output file will be called LOG2.OUT.

In file (31) the incremental value of the sink/source term (R) or the total solute affected by transport (UT) is dumped, i.e. the value of array R (or UT) at iteration i minus the same value for iteration i-1. This file has as many blocks as transport iterations needed to achieve convergence in all nodes. The whole file is rewritten for each time period. No matter the filename given to this file, the output file will be called LOG3.OUT.

In addition, whenever execution terminates normally or due to convergence problems (reaching to maximum number of iterations allowed) the results and all necessary information are written to an input/output file in unit 21 named RESTART.DAT. If RESTART.DAT is saved as a permanent file, execution can be resumed at any later date with the hydraulic head, temperature, concentration, precipitation, etc, as computed in the previous run, by adjusting time increment and convergence criteria for termination of convergence problem. Another useful feature of restart is that it can be used to change boundary conditions or other control data in previous input files as many times as one may desire without a pause in execution.