

THE APPLICATION OF OPTIMIZATION AND STOCHASTIC
METHODS TO ANALYTIC TRANSPORT MODELING

by

Clarissa Hansen

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Civil and Environmental Engineering

Brigham Young University

April 2002

BRIGHAM YOUNG UNIVERSITY
GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Clarissa Hansen

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date

Norman L. Jones, Chair

Date

Wayne C. Downs

Date

A. Woodruff Miller

Date

T. Prabhakar Clement

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Clarissa Hansen in its final form and have found that (1) its format, citations and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

Norman L. Jones
Chair, Graduate Committee

Accepted for the Department

Wayne C. Downs
Graduate Coordinator

Accepted for the College

Douglas M. Chabries
Dean, College of Engineering and Technology

ABSTRACT

THE APPLICATION OF OPTIMIZATION AND STOCHASTIC METHODS TO ANALYTIC TRANSPORT MODELING

Clarissa Hansen

Department of Civil and Environmental Engineering

Master of Science

This thesis describes the addition of several new capabilities to ART3D, a three-dimensional analytic reactive transport groundwater model developed by Dr. T. Prabhakar Clement of the University of Western Australia. In its original form, the model was a simple, but useful model built on the Domenico (1987) solution strategy and capable of modeling any number of complicated reaction pathways. As part of this project, an interface to GMS (Groundwater Modeling System) was built to allow easier data entry and better visualization of the results. The capability to analyze specific points in the model domain (observation points) and compare these to field data was added without the need for interpolation because of the analytic nature of the solver. An optimization code, called the PORT library, was also inserted to allow inverse modeling in ART3D. Finally, the model was developed to allow for stochastic simulations and threshold analysis assessments of the stochastic output. To aid ART3D users, a literature

survey of published first order decay constants for chlorinated ethenes is presented with mean and standard deviation values, which can be used in a stochastic simulation.

The optimization code was tested with a hypothetical case study whose findings are presented here. All of the additions were then tested on a case study based on the Brooklawn site in Baton Rouge, Louisiana. This site had previously been analyzed by Clement, et al. (2002) with BIOCHLOR, a predecessor to ART3D.

The hypothetical case study showed that the optimization routine is functional and can successfully optimize the parameters when there are sufficient observation points and when these points are placed in critical sections of the plumes.

The Louisiana case study showed that both the optimization routine and the stochastic code are functional and useful in the screening model. The addition of the optimization code made the calibration much easier and faster than in BIOCHLOR, which does not have internal optimization capabilities. The stochastic output was very useful in showing which areas of the aquifer are at most risk for dangerous contamination levels.

ACKNOWLEDGEMENTS

Thanks to Dr. Norm Jones and Dr. Prabhakar Clement for providing helpful suggestions, direction and analysis of my work. The project was funded by the Environmental Modeling Research Laboratory (EMRL) at Brigham Young University. The EMRL and the Centre for Water Research (CWR) at the University of Western Australia funded my trip to Australia to work with Dr. Clement. Thanks also to the GMS managers at EMRL, Alan Lemon, Mike Kennard and Jonathan Green for assistance with the programming aspects of the project and to Cristhián Quezada at CWR for help with the solution algorithm.

Table of Contents

List of Figures.	ix
List of Tables.	xii
Chapter 1 - Introduction.	1
Current Modeling Approaches.	3
ART3D.	7
Overview of Research.	9
Chapter 2 - Description of ART3D.	11
Equations.	19
Solution Technique.	20
Chapter 3 – Automated Parameter Estimation.	23
Observation Data.	24
Parameter Estimation.	27
Chapter 4 – Stochastic Simulations.	33
Chapter 5 – Review of Decay Constants.	37
Chapter 6 – Hypothetical Application.	47
Chapter 7 – Case Study.	57
Data Collection.	57
Forward Run and Inverse Runs.	62
Determination of the Effect of Natural Attenuation.	68
Determination of Steady State Plume.	70
Stochastic Simulation.	73
Chapter 8 - Conclusions.	81

Parameter Estimation.81
Stochastic Capabilities.82
References.	83
Appendix A – ART3D File Format.89
Appendix B – GMS Interface.	95
Step 1: Building the Grid.95
Step 2: Defining the Species.	95
Step 3: Defining Observation Data.	96
Step 4: Run Options and Output Control.99
Step 5: Entering the Effective Yield Matrix.	102
Step 6: Entering the Aquifer and Specie Parameters.103
Step 7: Running the Model.	106
Step 8: Reading the Solution.	110

Figures

Figure 1: A sequential reaction of a single parent specie (P) producing a single daughter product (D).	11
Figure 2: A reversible reaction between a parent specie (P) and a daughter product (D).	12
Figure 3: Parallel reactions between parent species (P) and daughter products (D).	12
Figure 4: A divergent reaction between a single parent specie (P) and multiple daughter products (D)	13
Figure 5: A convergent reaction between multiple parent species (P) and a single daughter product (D)	13
Figure 6: The effect of the starting point on the outcome of a parameter estimation run.	30
Figure 7: Flowchart for choosing a random number in a normal distribution (Sun, 1997)	36
Figure 8: The first arrangement of points is illustrated here with the four species plumes. (Simulations were also run with three points and with nine points according to this pattern.)	51
Figure 9: A comparison of the error at the observation points for hypothetical case study. Each colored bar represents a different starting point as shown in Table 14. Each set of five bars represents the five tests run with each set of observation points.	53
Figure 10: A comparison of the error in the parameters returned by the optimization code for hypothetical case study. Each colored bar represents a different starting point as shown in Table 14. Each set of five bars represents the five tests run with each set of observation points.	54

Figure 11: A comparison of the error in the grid datasets returned by the optimization code for hypothetical case study. Each colored bar represents a different starting point as shown in Table 14. Each set of five bars represents the five tests run with each set of observation points.	55
Figure 12: Possible pathways for anaerobic decay of chlorinated ethenes and ethanes. .	58
Figure 13: The centerline plots for each of the three calibrations at t = 25 years, plotted against the observed data.	66
Figure 14: The effects of sorption and decay on the model output at t = 25 years.	69
Figure 15: Centerline plots with and without natural attenuation (NA) after 25, 50 and 200 years, shown with field data.	71
Figure 16: Movement of the vinyl chloride plume through 500 years with natural attenuation.	72
Figure 17: Threshold analysis result for TCA; MCL = 0.005 mg/L. Grid length is 600 feet.	76
Figure 18: Threshold analysis result for DCA; MCL = 0.005 mg/L. Grid length is 600 feet.	76
Figure 19: Threshold analysis result for PCE; MCL = 0.005 mg/L. Grid length is 600 feet.	77
Figure 20: Threshold analysis result for TCE; MCL = 0.005 mg/L. Grid length is 600 feet.	77
Figure 21: Threshold analysis result for 1,1-DCE; MCL = 0.007 mg/L. Grid length is 600 feet.	78
Figure 22: Threshold analysis result for cis 1,2-DCE; MCL = 0.07 mg/L. Grid length is 600 feet.	78
Figure 23: Threshold analysis result for trans 1,2-DCE; MCL = 0.1 mg/L. Grid length is 600 feet.	79
Figure 24: Threshold analysis result for VC; MCL = 0.002 mg/L. Grid length is 600 feet.	79
Figure 25: ART3D sample input text file.	89
Figure 26: ART3D Define Species dialog.	96

Figure 27: Observation Coverage Attributes dialog.	97
Figure 28: Adding a timeseries to an observation point.	99
Figure 29: ART3D Run Options dialog.	100
Figure 30: ART3D Output Control dialog.	101
Figure 31: ART3D Effective Yield Matrix dialog.	102
Figure 32: ART3D Parameters dialog for a forward run.	104
Figure 33: ART3D Parameters dialog for a parameter estimation run.	105
Figure 34: ART3D Parameters dialog for a stochastic simulation.	106
Figure 35: ART3D output window for a forward simulation.	107
Figure 36: ART3D output window for a parameter estimation run.	109
Figure 37: ART3D output window for a stochastic simulation.	110

Tables

Table 1: Summary of the data included in the ART3D input file.	14
Table 2: Summary of the output files generated by ART3D.	17
Table 3: Anaerobic first order decay constants for PCE.	38
Table 4: Aerobic and transition first order decay constants for PCE.	42
Table 5: Anaerobic first order decay constants for TCE.	42
Table 6: Aerobic and transition first order decay constants for TCE.	43
Table 7: Anaerobic first order decay constants for DCE.	44
Table 8: Aerobic and transition first order decay constants for DCE.	44
Table 9: Anaerobic first order decay constants for VC.	45
Table 10: Aerobic and transition first order decay constants for VC.	45
Table 11: Summary of anaerobic first order decay constants for chlorinated ethenes. . .	45
Table 12: Adjusted analysis of anaerobic first order decay constants for PCE.	46
Table 13: Parameters for benchmark simulation.	47
Table 14: Parameter ranges and starting values. (Starting value colors correspond to Figures 9 through 11.).	52
Table 15: Model simulation time, regional dimensions and source dimensions.	59
Table 16: Aquifer and constituent parameter values from the BIOCHLOR model.	60
Table 17: Source concentration values.	61
Table 18: Yield constants used in both the BIOCHLOR simulation and the ART3D	

simulation.61
Table 19: Field observation data.	62
Table 20: Standard deviation and weight values applied to each chemical specie.	64
Table 21: Adjusted weight values for second and third calibrations.65
Table 22: Optimized parameter values for each of the three calibrations.67
Table 23: Varied parameters in stochastic simulation.74
Table 24: EPA Maximum Contaminatn Level (MCL) for water pollutants at the Brooklawn site.75
Table 25: ART3D input file card format.	90

Chapter 1 – Introduction

Reactive transport modeling is a branch of groundwater modeling concerned with the mapping of the movements and reactions of chemical species in the saturated zone of an aquifer, and in the vadose zone. This type of modeling must take into account several physical and chemical processes acting on the chemical compounds and the groundwater system. For example, advection is the movement of dissolved species through the aquifer as they move with the flow of groundwater. This process can cause an entire plume to move from one location to another but does not affect the concentration of the species or the size of the plume. As it moves, another physical process, dispersion, causes the dissolved particles of the plume to spread out, consequently affecting a greater volume of the aquifer. This occurs when some particles move at higher or lower velocities than others or when some take longer flow paths around soil particles or other obstacles. Another physical effect causing changes in the plume is diffusion, the movement of solutes from an area of high concentration to one of low concentration. Diffusion can affect a chemical plume even if there is no groundwater flow. The scattering of individual chemical particles in a plume through dispersion and diffusion causes a general decrease in concentration throughout the plume while simultaneously yielding an increase in the volume of affected aquifer.

Another complicating issue in transport modeling is the fact that more than one chemical compound can be present in a source. In addition, chemical reactions and biodegradation cause some chemical species to break down into “daughter products”. So, a useful reactive transport model must track all of the species present and monitor their chemical reactions with one another and their surroundings. A common example of this is the reactions of chlorinated compounds. An initial contamination of a groundwater system could be caused by a spill of trichloroethylene (TCE), but biodegradation will eventually result in the dechlorination of some of the plume to one of several types of dichloroethylene (DCE). Some of the DCE, in turn will lose an additional chlorine atom to form vinyl chloride (VC). A useful model of such a system would need to track the dechlorination of these compounds as well as their movement in the aquifer.

Differing dissolution rates also add complexity to a transport simulation. A contaminant’s dissolution rate depends on the chemical characteristics of the groundwater and on the surrounding conditions. A constituent will dissolve until a chemical equilibrium between the dissolved and nondissolved portions is reached. The equilibrium concentration can change with variation in pressure and temperature below the surface of the earth. The presence of similar chemical compounds already dissolved in the water can also alter the equilibrium concentration.

There are also a few processes that result in the slowing of the movement of a constituent through an aquifer. As a plume travels, it often interacts with the aquifer material through adsorption, chemisorption or ion exchange. Adsorption takes place when solutes are diffused into the aquifer material and attach to the interior surfaces of the material, where they are essentially removed from the main fluid flow.

Chemisorption involves a chemical reaction, which causes the solute to be integrated into the aquifer material. Ion exchange occurs when an ion in the water has a greater attraction to the aquifer material than an ion already attached to the material. The difference in attraction causes the ions to be exchanged, resulting in a modification of the groundwater chemistry. These processes, which can result in the slowing of the movement of a contaminant plume, are often grouped together under the term “retardation”.

Groundwater transport models are used for several different applications in environmental engineering. They can be used at the early stages of a project to determine the size, shape and location of a contaminant plume based on known aquifer conditions and source characteristics. This should be followed up with some subsurface investigations to calibrate and confirm the accuracy of the model. Transport models can also be used to find the source of a contamination plume or to determine the risk of contamination of surface water or wells. A calibrated transport model can be applied to the design of a remediation system by determining the feasibility of a certain remediation technique. Transport models are also practical in exploring the possibility of natural attenuation as a remediation option. In addition, a model can be used to compare several different remediation methods by predicting the time and money required and the expected level of success for each option.

Current Modeling Approaches

There are two basic approaches to computer groundwater modeling. Both begin with complicated partial differential equations describing the physical and chemical changes occurring in the system. A numerical model solves the equations by discretizing

the model domain into a grid or mesh. An approximate solution is then obtained by expressing the governing equations in terms of the local grid or mesh geometry and applying either the finite difference or the finite element method. This results in a solution at each of the grid cells or mesh nodes. The main advantage to this type of model is that it can be used to simulate a wide variety of problems with complex domains and boundary conditions. However, the accuracy of the result depends on the density of the grid or mesh. Furthermore, the process of creating the grid or mesh and preparing the problem input can be tedious and expensive. These models can also require extensive amounts of time to run. If the grid or mesh cells are very small compared to the size of the problem domain, then the solution is more accurate, but the computation time is greatly increased. With the advent of faster computers, the computation time is becoming a less important consideration in smaller scale models.

Analytic solution methods begin with the same partial differential equations describing the physical and chemical processes in the aquifer, but then, a direct solution to the governing equations is found in the form of a system of equations. This allows for an exact solution to the equations at any point in the system without the need to discretize the problem domain. These types of models are usually much more computationally efficient than numerical methods, but are generally limited to models with simple problem domains and boundary conditions.

There are several analytic transport models available today. Brief descriptions of several are listed here. BIOSCREEN and BIOCHLOR are common analytic transport models (USEPA *BIOCHLOR*, 2002 and, USEPA *BIOSCREEN*, 2002). Both are very simple screening tools built in a Windows Excel environment. BIOSCREEN is used to

model BTEX degradation, while BIOCHLOR is used to model reductive dechlorination of ethenes. Both were developed by Groundwater Services, Inc. (GSI).

Another analytic reactive transport model is called Kyspill and was designed and written by Dr. Sergio E. Serrano, a professor at the University of Kentucky (Hydroscience). Kyspill is advertised as a fast, easy transport model with simple input. It can model point, non-point, stationary or transient sources. The solution includes a 3-dimensional dispersion plume in the soil and a 2-dimensional plume in the unconfined aquifer below. It also considers aquifer heterogeneity statistically. Kyspill can include reactive and degradable chemical components.

RBCA Tier 2 Analyzer is another analytic transport model and claims to be simple and fast and to have the ability to incorporate complex geometry and multiple boundary conditions (Waterloo, 2000). The model consists of two main parts. The automatic plume generator uses interpolation to convert field data to a concentration plume, which can then be used as the starting point for a biodegradation simulation or to evaluate the effectiveness of a pump and treat system. The second part is a contaminant transport model. It models advective-dispersive transport with a variety of sorption and desorption options. The model can handle single or multiple species, including sequential decay or BTEX reactions.

Another transport model is called SOLUTRANS (Software Spotlight, 1999 and Fitts Geosolutions). Similar to ART3D, it assumes homogenous material and one-dimensional flow in the positive x direction. It also allows for three-dimensional dispersion with separate constants for each direction. Further, SOLUTRANS can model

equilibrium sorption or several types of kinetic sorption. There are four different basic shapes that can be used for the source, and irregular shapes or transient sources can be modeled through the superposition of several shapes in time and space.

An EPA model, HSSM (Hydrocarbon Spill Screening Model) has been developed to model the movement of LNAPLs (Light Non-Aqueous-Phase Liquids) from a surface spill through the vadose zone to the water table (The Scientific Software Group, HSSM, 1998). Because the model assumes homogeneity and ignores some phenomena, it should be considered a screening model and is only expected to be accurate within an order of magnitude.

AT123D (analytic, transport, 1-, 2-, and 3-D), tracks the fate of radioactive, chemical or heat pollution in the saturated zone (The Scientific Software Group, AT123D, 1998). The source can be defined as a point, line, area or volume source, with instantaneous, continuous, or transient releases. The model assumes an isotropic, horizontal and homogeneous aquifer and considers advection, dispersion, decay, and adsorption.

Another EPA model, PESTAN (Pesticide Transport), is designed to track the movement of organic contaminants through the vadose zone to the water table (The Science Software Group, PESTAN, 1998). Linear isotherm sorption, first order decay and dispersion are used to determine the movement of the species through the unsaturated zone. Like many other models described here, PESTAN assumes homogeneous soil conditions.

N3DADE (Non-equilibrium, Three-Dimensional Advection-Dispersion Equation) assumes homogeneity and steady state, unidirectional groundwater flow (US Salinity Laboratory, 1997). It accounts for advection, dispersion, first-order decay and retardation effects.

The USGS has developed several transport models, including ANALGWST, a suite of programs designed to model solute transport in one, two or three dimensions with a variety of boundary conditions (US Geological Survey, 2002). These programs can handle advection, dispersion and first order decay.

All of these models share some components in common with ART3D, which assumes homogeneity, prismatic aquifer shape, first order decay, linear isotherm adsorption and unidirectional groundwater flow in the positive x direction. ART3D's advantages include the ability to model more complicated reactions and the use of three separate dispersion constants for the three Cartesian directions. The main disadvantage in the current release of ART3D is the requirement that all the species have the same retardation coefficient. However, Dr. Clement has overcome this obstacle in a new version of ART3D that is in the final stages of development at the time of this printing.

ART3D

This research project involved the addition of several new functionalities to ART3D, a three-dimensional analytic reactive transport model designed by Dr. T. Prabhakar Clement of the University of Western Australia

The evolution of ART3D began some years ago when the Air Force Center for Environmental Excellence contracted with Groundwater Services, Inc. (GSI) to create a

two-dimensional analytic model to simulate BTEX degradation in an aquifer setting. The result was an Excel spreadsheet with several Visual Basic (VB) macros and was called BIOSCREEN.

The success of BIOSCREEN was followed by the development of BIOCHLOR, a similar program used to model the sequential degradation of chlorinated ethenes (PCE (perchloroethylene)→TCE→DCE→VC→Ethene). Dr. Clement (who was working at Battelle Pacific Northwest Laboratory at the time) was asked to develop the analytic engine to do the calculations.

Later, Dr. Clement created his own stand-alone version of BIOCHLOR and called it BC3D. This three-dimensional model is written in FORTRAN and has a simple input text file. The output includes a three-dimensional grid and several two-dimensional views of the plume formatted for viewing in GMS (Groundwater Modeling System). Another output file is a text file listing the centerline concentrations of each species at each grid cell. This can be easily read into Excel or any other spreadsheet program for analysis and display. Like BIOCHLOR, BC3D can only handle simple sequential reactions – normally chlorinated solvents. The code was also later translated to Visual Basic by Dr. Norm Jones of Brigham Young University.

Dr. Clement followed up with another version of the program, ART3D, in the summer of 2001. This release allows more complex reaction sequences, such as reversible reactions, multiple parent species producing the same daughter product, or multiple daughter products resulting from the same parent species. The additions and changes described in this thesis were implemented in this version of ART3D

Overview of Research

This research project involved the implementation of several new capabilities to the ART3D model. These additions included the calculation of exact concentration values at observation points, automated parameter estimation, and stochastic simulation capabilities. The project also included a literature review to find first order decay constants for the successive dechlorination of PCE, TCE, DCE, VC and ethene. A graphical user interface to the new ART3D code was built in GMS to simplify data entry and allow for better visual analysis of the results. GMS is developed by the Environmental Modeling Research Laboratory (EMRL) at Brigham Young University (BYU). The project culminated with an application of the finished code to a hypothetical test case and an actual Superfund site. These case studies served to test the new features and demonstrate the value of this new version of ART3D.

The additions made to ART3D as part of this project have extended its usefulness and simplified its operation. The addition of observation points allows a user to see constituent concentrations at any important area such as at a well or along a riverbank. When field data is used, it can be directly compared to exact solutions at that point. Through the use of the new parameter estimation capabilities, a user can more easily and quickly calibrate the model to field data, thereby producing a more reliable model. Stochastic simulations are valuable in determining the statistical reliability of the results and in performing statistical risk analyses.

A literature review of decay constants for chlorinated ethenes is provided to aid the user in the determination of a realistic set of decay values for a stochastic simulation

and to demonstrate the wide variation in these values and the difficulty in determining them with accuracy.

This thesis is organized as follows: Chapter 2 is a brief discussion of the theory behind the ART3D solution engine. The discussion is drawn directly from the paper by Dr. Clement on the subject (2001). The third chapter discusses the addition of observation points and parameter estimation to the model. It discusses both the theory and the implementation of these new capabilities. Chapter 4 describes the theory of stochastic modeling and its application to and implementation in ART3D. The literature review of decay constants is listed in Chapter 5. Following the review, a description of the design and results of the hypothetical case study is presented along with the conclusions drawn from this exercise. Chapter 7 describes the application of the new capabilities of ART3D to a Superfund site in Louisiana. Because this site had previously been used to demonstrate the usefulness of BIOCHLOR, it allowed a nice comparison of the results before and after the current project additions were made. The final conclusions are presented in Chapter 8.

Chapter 2 – Description of ART3D

ART3D is a three-dimensional analytic groundwater model applied to multi-species transport problems. It can track several types of complex reactions between any number of contaminant species. The reaction types include reversible, parallel, divergent and converging reactions in addition to sequential reactions handled by BC3D. These reaction types are shown below in Figure 1 through Figure 5. A sequential reaction (Figure 1) is a simple reaction involving one parent species and one daughter product. The reaction proceeds in only one direction.

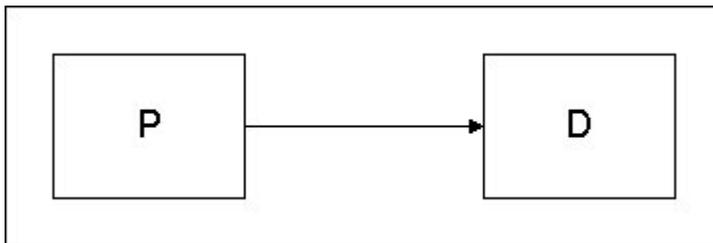


Figure 1: A sequential reaction of a single parent species (P) producing a single daughter product (D).

A reversible reaction (Figure 2) also includes only one parent species and one daughter product, but the reaction can proceed in either direction until an equilibrium point is reached.

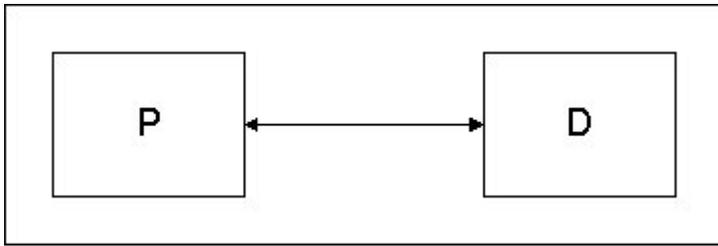


Figure 2: A reversible reaction between a parent specie (P) and a daughter product (D).

In a parallel reaction (Figure 3), two reactions occur simultaneously in the same system and can be unaffected by each other.

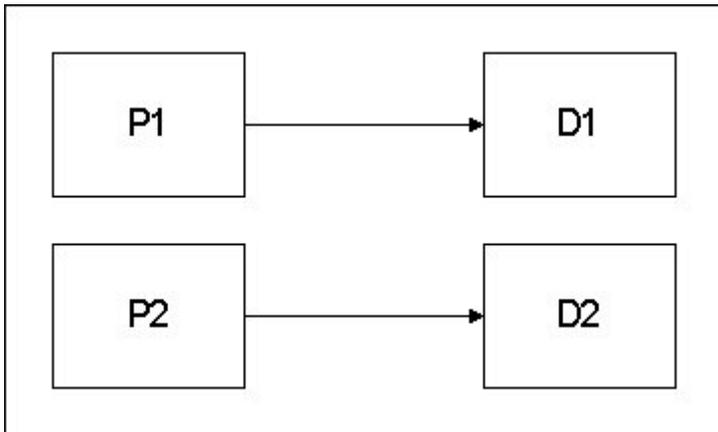


Figure 3: Parallel reactions between parent species (P) and daughter products (D).

A divergent reaction (Figure 4) occurs when a single parent specie produces more than one daughter product.

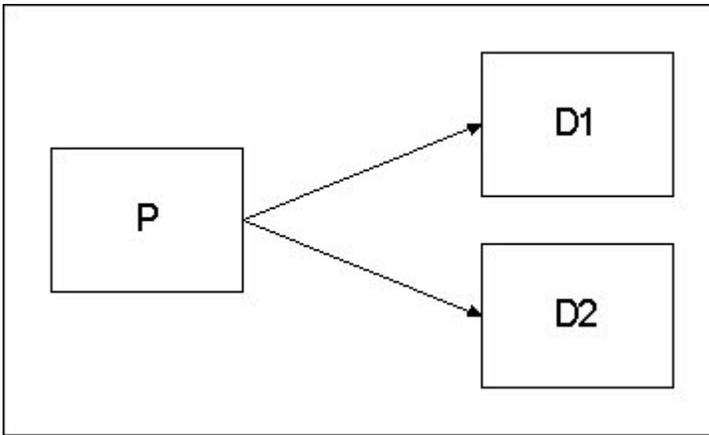


Figure 4: A divergent reaction between a single parent specie (P) and multiple daughter products (D).

In a convergent reaction (Figure 5), a single daughter product may be produced from the decay of two separate parent species.

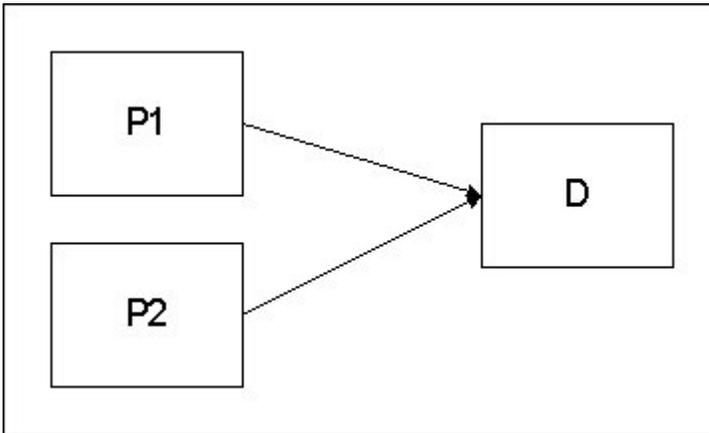


Figure 5: A convergent reaction between multiple parent species (P) and a single daughter product (D).

Many complex systems involve a combination of these reaction types. ART3D is designed to easily manage any of these reactions.

ART3D is a simple model that assumes a one-dimensional flow field and constant (homogenous) hydraulic and transport properties. All reactions are assumed to be first

order. The main drawback to this version of ART3D is the requirement that all of the retardation coefficients be the same.

A key advantage of ART3D is its simplicity. Often, the modeling objectives and the project budget prohibit the use of a more expensive numerical simulation. The data needed to run a basic ART3D model is minimal and uncomplicated. The time and money saved by the simplicity of input, creation and analysis make ART3D a good choice for a low budget, preliminary, or screening study.

To allow for the new additions to the code, the input file format was altered. Some of the data from the original file format is included, although in a different form and a great deal of new information has been added to the file. The data included in the new file format is summarized in Table 1. A more detailed description of each element of the file is found in Appendix A.

Table 1: Summary of the data included in the ART3D input file.

Information similar to the original ART3D input file:	Source dimensions, time data, and the number of species are included in the input file similarly to the original ART3D input file. The length of each time step is calculated from the total simulation time and the number of time steps input by the user. These time steps are not for discretization, but to define the times at which the solution will be output for post-processing
---	--

Table 1: Summary of the data included in the ART3D input file. (continued)

GMS information:	This information includes the name of the GMS gridfile, observation coverage, the units to be used and the reference time information. Except for the grid file, these lines of the input file are not used by ART3D, but are important when a simulation is read into GMS. The gridfile is opened by ART3D to obtain the dimensions of the grid and the number of cells in each direction. ART3D does not use the grid for discretization, but to determine the resolution of the output for post-processing.
Yield matrix:	The original ART3D input file required the user to calculate and enter the reaction coefficient matrix, a complicated combination of the yield values and the decay constants. The new file format is simpler, in that the user enters the effective yield coefficients in matrix form and ART3D uses this information with the decay data to calculate the reaction coefficient matrix required by the ART3D solver.
Parameter Estimation data:	Stopping tolerances used by the optimization routine are included in the new input file. These include the maximum number of iterations and function calls allowed and the relative and absolute function convergence tolerance. The new file format also requires the user to enter bounds on each parameter, which is to be optimized. This data is described in more detail in Chapter 3.

Table 1: Summary of the data included in the ART3D input file. (continued)

Stochastic data:	In the new file format, the user must enter the bounds, type of distribution and (with a normal distribution) the standard deviation for each varied parameter. The number of stochastic simulations to be run is also included. This data is explained in Chapter 4.
Observation points:	The new input file includes information on the location of any number of observation points as well as their observed species concentrations at any time. The addition of observation points makes it possible for ART3D to compare the computed and observed concentration values at each point and to report this residual. To aid in this comparison, a matrix of weighting values is included in the input text file. Weighting values will be explained further in Chapter 3.

Several different types of files are written out at the end of an ART3D simulation.

A brief description of each file type is listed below in Table 2.

Table 2: Summary of the output files generated by ART3D.

File type	Extension	Description
Data set file	*.dat	These are binary files listing the species concentration at each point in the grid. A separate file is written out for each defined specie. They are differentiated by the addition of the name of the specie at the end of the filename as in example_pce.dat and example_tce.dat.
Time series file	*.ats	This file lists each of the observation points and their calculated values at each time step in the simulation. One important use for this file is to plot the change in concentration over time for a particular point in the simulation
Output file	*.out	The output file echoes all of the input information, includes a list of the output files and their names, and lists the calculated and observed concentration values for each observation point at each time step. The file also shows the residual error, weighted residual error and several total error calculations. This file is useful in determining the accuracy of the model by comparing it to observed data.

Table 2: Summary of the output files generated by ART3D. (continued)

Centerline file	*.ctr	This output file lists the concentrations of each species along the top centerline of the grid for each time step. This file can be directly opened in Excel or a similar spreadsheet program to view the concentration profile at different points along the grid.
Data set superfile	*.dss	This file lists each of the data set files (*.dat). It is used by GMS to load the data sets into the project for viewing.
Stochastic super file	*.sto	This file lists each of the data set superfiles associated with a single stochastic run. It is used by GMS to read in each of the solutions to the project.
Parameter file	*.par	The parameter file lists the final, optimized values for each parameter at the end of a parameter estimation run. This file can be read by GMS to load these values to the <i>Parameters</i> dialog.
Iteration record file	*.itr	This file contains a list of each of the parameter values after each iteration of a parameter estimation simulation.

Equations

The governing equation solved by ART3D is as follows:

$$R_i \frac{\partial c_i}{\partial t} + v \frac{\partial c_i}{\partial x} - D_x \frac{\partial^2 c_i}{\partial x^2} - D_y \frac{\partial^2 c_i}{\partial y^2} - D_z \frac{\partial^2 c_i}{\partial z^2} = \sum_{j=1}^{i-1} y_{ij} k_j c_j - k_i c_i + \sum_{j=i+1}^n y_{ij} k_j c_j \quad (1)$$

where

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

c = the i th species concentration,

y_{ij} = the effective yield factor describing the mass of species i produced per mass of species j used,

k_i = the first order reaction rate for the disappearance of species i ,

R_i = the retardation coefficient,

v = seepage velocity,

D_x, D_y, D_z = the dispersion coefficients in each orthogonal direction, and

n = the total number of species in the reaction series.

This equation can be converted to matrix notation as follows:

$$R\hat{c}' + v\hat{c}_x^1 - D_x\hat{c}_x^2 - D_y\hat{c}_y^2 - D_z\hat{c}_z^2$$

$$R\hat{c}' + v\hat{c}_x^1 - D_x\hat{c}_x^2 - D_y\hat{c}_y^2 - D_z\hat{c}_z^2 = K\hat{c} \quad (2)$$

where \mathbf{K} is the reaction coefficient matrix, a combination of the effective yield coefficients and the decay coefficients for each species. The matrix is arranged in the following manner:

$$\begin{bmatrix} -k_1 & y_{2 \rightarrow 1}k_2 & y_{3 \rightarrow 1}k_3 & y_{4 \rightarrow 1}k_4 \\ y_{1 \rightarrow 2}k_1 & -k_2 & y_{3 \rightarrow 2}k_3 & y_{4 \rightarrow 2}k_4 \\ y_{1 \rightarrow 3}k_1 & y_{2 \rightarrow 3}k_2 & -k_3 & y_{4 \rightarrow 3}k_4 \\ y_{1 \rightarrow 4}k_1 & y_{2 \rightarrow 4}k_2 & y_{3 \rightarrow 4}k_3 & -k_4 \end{bmatrix}$$

where $y_{a \rightarrow b}$ is the effective yield factor for the reaction of constituent a going to constituent b and k_a is the first order destruction rate for constituent a . For example, the following matrix is for the simple sequential reaction, $\text{PCE} \rightarrow \text{TCE} \rightarrow \text{DCE} \rightarrow \text{VC}$

$$\begin{bmatrix} -k_{\text{PCE}} & 0 & 0 & 0 \\ y_{\text{PCE} \rightarrow \text{TCE}}k_{\text{PCE}} & -k_{\text{TCE}} & 0 & 0 \\ 0 & y_{\text{TCE} \rightarrow \text{DCE}}k_{\text{TCE}} & -k_{\text{DCE}} & 0 \\ 0 & 0 & y_{\text{DCE} \rightarrow \text{VC}}k_{\text{DCE}} & -k_{\text{VC}} \end{bmatrix}$$

The effective yield coefficient, denoted by y , is the product of the stoichiometric yield fraction of the reaction and the ratio of the molecular mass values of the reactants and products. The final units are in mass of product produced / mass of reactant used. To meet the law of conservation of mass, the sum of all the yield fractions for a given parent compound should equal unity.

Solution Technique

The main difficulty in solving the matrix equation (2) is that \mathbf{K} is cross-coupled, except in the special case where \mathbf{K} is a diagonal matrix. In order to solve the equation, \mathbf{K} must be transformed to a diagonal matrix. The transformation will also affect the other matrices in the equation.

The transformation can be accomplished through the use of a matrix \mathbf{S} such that $\mathbf{S}^{-1}\mathbf{K}\mathbf{S} = \tilde{\mathbf{K}}$ and $\tilde{\mathbf{K}}$ is a diagonal matrix. The \mathbf{S} matrix can be found using a similarity transform. If \mathbf{S} is chosen so that the columns of \mathbf{S} are the eigenvectors of \mathbf{K} , the resulting $\tilde{\mathbf{K}}$ will be a diagonal matrix whose diagonal values are the eigenvalues of the original matrix \mathbf{K} .

The solution strategy can be best explained by defining two separate domains for all of the matrices in the equation. The “ c ” domain is defined as the original domain before any transformation takes place. Premultiplication of all the terms by \mathbf{S}^{-1} transforms all of the matrices into the “ b ” domain. In this domain, the \mathbf{K} matrix is diagonal and uncoupled. The solution can then be found through simple matrix operations. Once the equation has been solved, it must be transformed back into the “ c ” domain through postmultiplication by the \mathbf{S} matrix.

The use of this strategy allows the incorporation of very complicated reaction sequences through the definition of the yield matrix. The new input files for ART3D requires that the user enter the effective yield matrix including only the off-diagonal yield values (the product of the molecular weight ratios and the stoichiometric yield fractions). ART3D then calculates the reaction coefficient matrix by multiplying by a \mathbf{K} vector and

setting the diagonal values. A full description of this solution strategy can be found in Clement (2001).

Chapter 3 – Automated Parameter Estimation

One of the main modifications made to the ART3D code as part of this research was the addition of a new option for automated parameter estimation. In computer modeling, the selection of parameter values is challenging, since they are often difficult to measure in the field or in the lab. To get around this problem, many models are calibrated to observation data by manually adjusting the input parameters until the output matches that observed in the field. This can be a tedious process, but can be made easier through the use of an automated parameter estimation tool, which can compare the calculated and observed data and calculate an error value. It can then adjust the input parameter values until the error value drops to a minimum.

In order to add automated parameter estimation to ART3D, it was necessary to add the capability to read observation data and compare that data to the simulated values. This makes it possible to compute residual errors for selected species and a global weighted error that serves as the objective function to be minimized in the parameter estimation process. As an analytic model, ART3D was especially adaptable to computing residuals at observation wells because of its ability to compute concentrations at precise locations, instead of having to interpolate these values from surrounding grid cells. At each iteration of the parameter estimation process, the solution is only

computed at the observation points, rather than at the grid nodes. This causes the iteration process to be extremely fast.

Observation Data

In any forward run, ART3D calculates the concentrations at each observation point and for each time step. This includes the time steps assigned to the grid and those assigned to any of the observation points. These concentration values can be obtained from the output file (*.out) written after an ART3D simulation. This file lists all of the observation points with their observed and calculated values and some statistical information to aid in the analysis of the model's reliability.

First, the output file lists the residual error and weighted error for each point. The residual error, as shown in equation 3, is a simple difference between the observed and calculated data.

$$r_i = c_i - o_i \tag{3}$$

where

r_i = residual error for point i , and

c_i = computed value for point i ,

o_i = observed value for point i .

Weighting values can be used to equalize the contribution of different magnitudes of residual error on the objective function and are defined in the weighting matrix of the ART3D input file. Their importance is illustrated in the following example. Suppose the

observed concentration of specie A was 20 mg/L while the concentration observed for specie B at the same point was 2 mg/L. The calculated value was off by 20% in each case, resulting in calculated values of 24 mg/L and 2.4 mg/L, respectively. Without a weighting factor, the error contributed by specie A would be 4 and that contributed by specie B would be 0.4. The optimizer would place higher importance on reducing the higher error on specie A, and would pay little attention to specie B. If, however, specie A is assigned a weight value of 0.0693 and specie B receives a weighting value of 0.694, the resulting weighted residual values are 0.277 and 0.278, respectively. In this case, the optimizing routine will place equal emphasis on improving both residual values.

It is suggested that the weighting value be the inverse of the square of the standard deviation of each measured value as in equation 4 below. When running ART3D from GMS, the user enters the standard deviation and the weighting value is automatically calculated using this equation. When running ART3D outside of the GMS interface, the user must calculate the weighting values and list them in the input file.

$$weight = \frac{1}{stdev^2} \quad (4)$$

The weighted error is the residual error multiplied by the weighting factor described above.

$$wr_i = w_i(r_i) \quad (5)$$

where

wr_i = weighted residual for point i , and

w_i = weighting value for point i .

A set of global error norms are also presented, including total error, mean error, mean absolute error and root mean square error for the residual error and weighted residual error as shown in the equations below. In the following equations the number of data points refers to the number of observation points multiplied by the number of time steps and the number of species.

$$te = \sum_{i=1}^n r_i, \quad wte = \sum_{i=1}^n wr_i \quad (6)$$

where

te = total error or sum,

wte = total weighted error, and

n = number of data points

$$me = \frac{1}{n} \sum_{i=1}^n r_i, \quad wme = \frac{1}{n} \sum_{i=1}^n wr_i \quad (7)$$

where

me = mean error, and

wme = weighted mean error

$$mae = \frac{1}{n} \sum_{i=1}^n |r_i|, \quad wmae = \frac{1}{n} \sum_{i=1}^n |wr_i| \quad (8)$$

where

mae = mean absolute error, and

$wmae$ = mean absolute weighted error.

$$rmse = \sqrt{\left(\frac{1}{n} \sum_{i=1}^n (r_i)^2\right)}, \quad wrmse = \sqrt{\left(\frac{1}{n} \sum_{i=1}^n (wr_i)^2\right)} \quad (9)$$

where

$rmse$ = root mean square error, and

$wrmse$ = weighted root mean square error.

These error norms are useful in determining the accuracy of one model run vs. another and in calibrating the model.

Parameter Estimation

Once the observation data capabilities were added, an automated parameter estimation algorithm was implemented. There are two ways to include parameter estimation in a computer model. It can be done through the use of an external optimization utility, or through the addition of an internal algorithm. One common example of an external parameter estimation utility is PEST, written by John Doherty (Doherty, 2000). PEST is a general purpose optimization utility that takes control of a model and systematically adjusts the input parameter values in an attempt to minimize the residual between observed and calculated data. The use of an internal algorithm has

the advantage that running the model in inverse mode can be much simpler and the code can be more efficient since there is no need to pass data text files between the utility and the model. In ART3D, we opted to use a built-in optimization algorithm in the form of a set of optimization subroutines called the PORT library and designed by David M. Gay at Bell Laboratories (Gay, 1990). The use of the PORT library required the writing of a subroutine to calculate the objective function.

In a parameter estimation run, the user enters all of the field observed concentrations for each specie at any number of observation points. The modeler also chooses which parameters to optimize, sets their bounds and assigns a starting value for each parameter. The user must also define the stopping points for the routine: the maximum number of iterations to run, the maximum number of calls to the objective function and the epsilon values, describing the accuracy required in the final results. All of this data is then passed to the PORT library, which calls a subroutine in ART3D to calculate the value of the objective function (or weighted residual error) using the user-defined starting values. To calculate this error, the objective function routine runs a quick forward simulation of ART3D calculating the concentration values at the observation points but not at the grid points. Then it calculates the resulting error as a weighted difference between the calculated and observed values. Equation 10 is used to calculate the error:

$$error = \sum_{j=1}^m \sum_{i=1}^n w_{i,j} |c_c - c_o|_i \quad (10)$$

where

n = number of observation points,

m = number of species,

$w_{i,j}$ = weight value for species j at observation point i ,

c_c = the calculated concentration at observation point i , and

c_o = the observed concentration at observation point i .

Once the objective function has been computed, its value is passed back to the optimization routine. The routine then changes each of the parameters slightly, one at a time, and calls the objective function again. In this way, the gradient of the objective function with respect to each of the parameters can be estimated. Once the gradient values are known, the optimization routine changes all of the parameter values down gradient and computes the objective function again. This continues until, (a) the error value returned is less than the absolute function convergence tolerance, (b) the improvement in error is less than the relative function convergence tolerance, (c) the maximum number of iterations is reached, (d) the maximum number of function evaluations is reached, or (e) an error is encountered. All of these stopping tolerances can be defined by the user.

If the optimization routine has difficulty obtaining a sufficiently accurate solution, there are often several things the user can do to improve the success of the routine. Often, changing the bounds or the starting points on the parameters will result in a better

outcome. The importance of trying a variety of starting points for the parameters is illustrated in Figure 6.

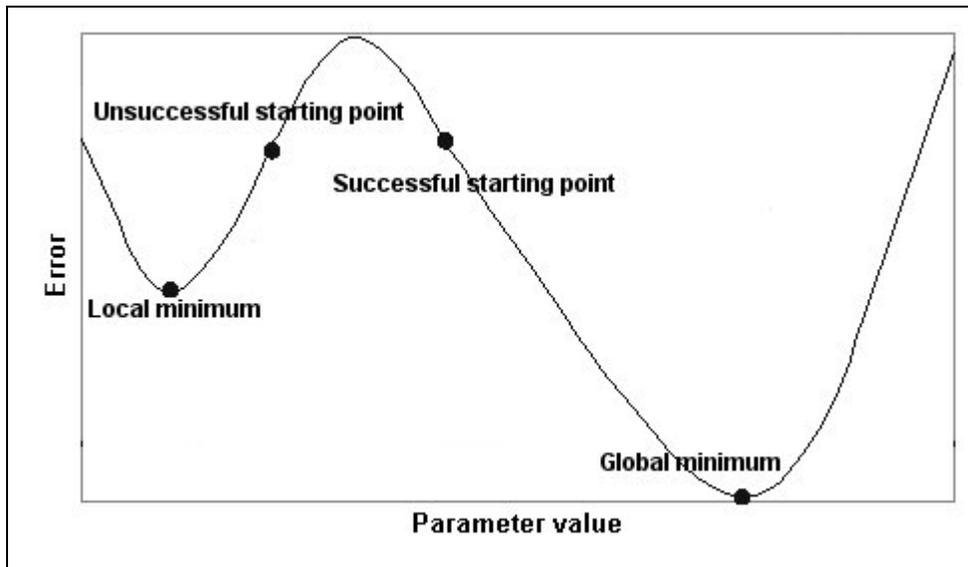


Figure 6: The effect of the starting point on the outcome of a parameter estimation run.

Figure 6 represents a hypothetical relationship between the total error and one parameter value. The point at the bottom right of the figure is the “global minimum” or the parameter value which will minimize the difference between the observed and calculated values on observation points. Sometimes, however, the optimization routine will find a local minimum value. If the point marked “unsuccessful starting point” is used as a starting point, the routine will find the gradient at that point and then reduce the parameter value to reduce the error value. When the parameter value reaches the local minimum, the gradient will be zero and the routine will stop.

If, however, the point marked as “successful starting point” is used as a starting point, the gradient will be calculated and the routine will successively increase the parameter value until it converges on “the global minimum” where the gradient is zero.

Chapter 4 – Stochastic Simulations

A second component of this research was to implement stochastic modeling capabilities in the ART3D code using a simple, built-in Monte Carlo approach. This new capability is described in this chapter.

In a typical modeling project, a common approach is to select a set of input parameters representing the “best estimate” and to present the resulting model solution as the most likely answer to the problem being studied. The problem with this approach is that it does not accurately portray the uncertainty in the solution that results from the uncertainty inherent in the input values. It is often very difficult to settle on one “correct” value for any parameter. In addition, ART3D is a very simple model and does not allow for regions of different aquifer characteristics. This means that the modeler must settle on an average value for the entire aquifer, which will best duplicate the behavior of the real aquifer. Even aquifers that are considered to be homogenous have pockets and small areas with different characteristics from the main aquifer.

One example of this difficulty is mechanical dispersivity, whose value is sensitive to the length of travel. As a particle travels further and further from its source, the probability that it will encounter an obstacle which will cause it to spread away from other particles, becomes higher and higher. These obstacles can include sudden changes in porosity or hydraulic conductivity, especially in small pockets. At a certain distance,

the probability ceases to rise and dispersion becomes a constant with respect to travel length. Because the dispersivity values are so changeable, it is difficult to know which coefficient to use in the model.

Another difficulty in settling on a single parameter value is that they are usually difficult to measure. Often, field tests (such as pump tests) are used because they give a fairly accurate average of the characteristics over a large area. However, they do not account for small areas with different characteristics, which largely affect dispersion. Conversely, when lab tests are run on small soil samples, the values can be accurately measured for that one point, but that value cannot always be extrapolated over a larger area.

One approach to dealing with the uncertainty associated with parameter selection is the use of stochastic modeling. This allows the user to enter a range of values for each parameter and information on the statistical distribution with respect to the mean value. The uncertainty of the parameters, as defined through their standard deviations, can then be projected into the model using a stochastic method. A stochastic simulation in ART3D is accomplished by running the model multiple times, each time with a slightly different combination of parameter values. The parameter values are chosen randomly based in the statistical distribution data supplied by the user. Each set of parameter values, then can be assumed to be equally probable. The result of this process is a large number of equally possible solutions that can be viewed individually or used in groups to determine the probability or risk that a given outcome will occur. For example, stochastic solutions can be used to generate maps illustrating the probability that a given threshold concentration will be exceeded at any point in the problem domain. This

allows the user to make some conclusions about the reliability of the model solution and the statistical probability of a certain condition.

The stochastic modeling capability was implemented by allowing the user to choose any combination of input parameters to be varied. For each parameter, the user then enters the upper and lower bounds, the mean, the distribution type (linear or normal), and in the case of a normal distribution, the standard deviation. The user also selects the number of simulations to be run. ART3D runs repeatedly with a different set of parameter values each time. The new parameter values for each run are selected through the generation of a random number that honors the provided range and distribution data. The selection of these values is discussed below.

The stochastic modeling code added to ART3D was designed to allow for both linear and normal distributions. For a linear distribution, ART3D uses the `RANDOM_NUMBER` function (in FORTRAN) to find a random number between zero and one. This value is then transformed to the proper range through the following equation.

$$Value = ((Max - Min)Randomval) + Min \quad (11)$$

In the case of a normal distribution, ART3D chooses two independent random numbers using the `RANDOM_NUMBER` function and converts them to a single parameter value within the defined bounds and faithful to the normal distribution data

provided by the user. The parameter is generated using the process and equations described in the following flow chart (Figure 7):

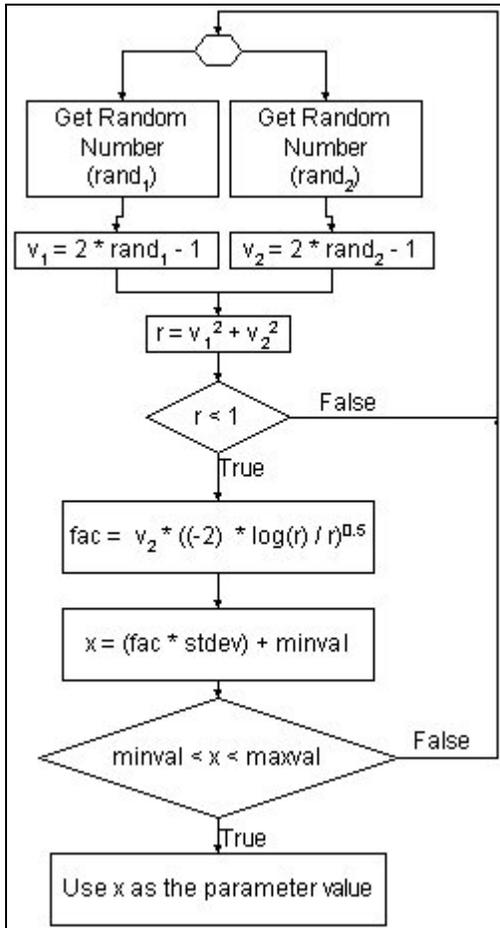


Figure 7: Flowchart for choosing a random number in a normal distribution. (Sun, 1997)

The output for a stochastic simulation can be large. A new set of output files is written for each individual run. They are differentiated by the addition of an underscore followed by a sequential number to each filename as stochrun1_pce.dat, stochrun1_tce.dat, stochrun2_pce.dat, stochrun2_tce.dat, etc.

Chapter 5 – Review of Decay Constants

Another important part of this research project was an extensive literature search to find first order decay rate constants for the dechlorination of chlorinated ethenes (PCE, TCE, DCE, VC). Because ART3D is a simple model, which allows no heterogeneity or complicated boundary conditions, it can only be used as a screening model or on a project with a small budget. Those modelers with large budgets or who need high accuracy may start with ART3D to get an initial idea of the plume, but will want to use a more intricate model such as RT3D. Modelers on a small budget will not have the means to carry out the expensive tests that will be required to determine aquifer characteristics such as retardation, dispersion and velocity or species characteristics such as decay rates and source concentrations. With the addition of stochastic capabilities, the modeler can run ART3D multiple times with varying parameter values each time to come up with several possible plumes. The difficulty in a stochastic simulation is in the defining of the bounds and standard deviation values for each of the parameters. However, the literature search outlined below shows average values and standard deviations for the decay constants across all of the journal articles searched. These values can be directly used in a stochastic simulation.

The first order decay constants from the literature search are shown in Tables 3-11. Each value has a short description explaining the conditions under which the values

were observed. They are arranged with separate tables for each of the chlorinated ethene species and, within each table, they are ordered with the smallest values first and the largest last. The values are also separated with one table for anaerobic conditions and another for values that are from aerobic conditions or a combination of anaerobic and aerobic conditions. Averages and standard deviations are shown for the anaerobic condition values. Many of the values have been calculated from data presented in the paper. The great variance in the values is obvious even at a quick glance. Many of the higher values come from tests involving the use of bacterial cultures developed specifically to dechlorinate these compounds. Natural conditions in an aquifer will seldom match these highly specific lab conditions.

Table 3: Anaerobic first order decay constants for PCE.

Parent species	Rate constant (day ⁻¹)	Source	Notes
PCE	0.00031	Clement <i>et al.</i> , 1999	lab scale microcosm, anaerobic
PCE	0.00032	Clement <i>et al.</i> , 1999	calibrated values, anaerobic zone
PCE	0.0004	Clement <i>et al.</i> , 1999	calibrated values, anaerobic zone
PCE	0.0012	Rügge <i>et al.</i> , 1999	anaerobic, overall value
PCE	0.002	Roberts <i>et al.</i> , 1982 quoted in Schaeerlaekens <i>et al.</i> , 1999	sand aquifer, anaerobic conditions
PCE	0.003	Roberts <i>et al.</i> , 1982 quoted in Schaeerlaekens <i>et al.</i> , 1999	sand aquifer, anaerobic conditions
PCE	0.0038	Rügge <i>et al.</i> , 1999	anaerobic, maximum value
PCE	0.0039 - 0.0077	Hardy <i>et al.</i> , 1999	anaerobic, reference book (calculated from half-life)
PCE	0.0075 - 0.07	Bagley and Gosset, 1990 quoted in Schaeerlaekens <i>et al.</i> , 1999	anaerobic conditions, sulfate-reducing conditions
PCE	0.0079	Bouwer <i>et al.</i> , 1981 quoted in Schaeerlaekens <i>et al.</i> , 1999	anaerobic conditions

Table 3: Anaerobic first order decay constants for PCE. (continued)

Parent species	Rate constant (day ⁻¹)	Source	Notes
PCE	0.0008 - 0.0198	Hardy <i>et al.</i> , 1999	anaerobic, field study (calculated from half life)
PCE	0.64	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 6.0µM and flow rate = 2 mL/hr), anaerobic
PCE	1.24	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 6.0µM and flow rate = 4 mL/hr), anaerobic
PCE	1.72	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 6.0µM and flow rate = 6 mL/hr), anaerobic
PCE	2.16	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 65.1µM and flow rate = 70 mL/hr), anaerobic
PCE	1.78 - 2.56	Carter and Jewell, 1993	degraded by anaerobic attached-films; value calculated from µ _{max} , volatile solid concentration and initial PCE concentration values provided
PCE	2.42	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 91.2µM and flow rate = 70 mL/hr), anaerobic
PCE	1.25 - 4.1	Carter and Jewell, 1993	degraded by anaerobic attached-films; value calculated from µ _{max} , volatile solid concentration and initial PCE concentration values provided

Table 3: Anaerobic first order decay constants for PCE. (continued)

Parent species	Rate constant (day ⁻¹)	Source	Notes
PCE	2.68	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 6.0µM and flow rate = 8 mL/hr), anaerobic
PCE	3.24	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 6.0µM and flow rate = 10 mL/hr), anaerobic
PCE	3.64	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 6.0µM and flow rate = 16 mL/hr), anaerobic
PCE	3.73	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 32.6µM and flow rate = 70 mL/hr), anaerobic
PCE	4.0	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 6.0µM and flow rate = 12 mL/hr), anaerobic
PCE	0.0058-8.32	Hardy <i>et al.</i> , 1999	recent literature (calculated from half-life)
PCE	4.36	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 6.0µM and flow rate = 25 mL/hr), anaerobic
PCE	4.39 - 5.88	Carter and Jewell, 1993	degraded by anaerobic attached-films; value calculated from µ _{max} , volatile solid concentration and initial PCE concentration values provided

Table 3: Anaerobic first order decay constants for PCE. (continued)

Parent species	Rate constant (day ⁻¹)	Source	Notes
PCE	5.4	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 6.0µM and flow rate = 20 mL/hr), anaerobic
PCE	3.92 - 6.9	Carter and Jewell, 1993	degraded by anaerobic attached-films; value calculated from µmax, volatile solid concentration and initial PCE concentration values provided
PCE	5.6	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 6.0µM and flow rate = 30 mL/hr), anaerobic
PCE	8.0	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 6.0µM and flow rate = 70 mL/hr), anaerobic
PCE	8.28	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 6.0µM and flow rate = 50 mL/hr), anaerobic
PCE	9.14	Fathepure and Tiedje, 1994	degraded by a chlorobenzoate-enriched biofilm reactor; calculated from dechlorination rate with PCE concentration = 5.25µM and flow rate = 70 mL/hr), anaerobic
Average:	2.2792		
Standard Deviation:	2.2612		

Table 4: Aerobic and transition first order decay constants for PCE.

Parent species	Rate constant (day ⁻¹)	Source	Notes
PCE	0.0001	Clement <i>et al.</i> , 1999	calibrated values, transition (aerobic/anaerobic) zone
PCE	0.00035 - 0.0007	Clement <i>et al.</i> , 1999	field estimated
PCE	0.001705	Gerritse <i>et al.</i> , 1995	degraded by a combination of anaerobic dechlorinating and aerobic methanotrophic enrichment cultures; calculated from dechlorination rate with PCE concentration = 0.2 M
PCE	0.0021	Praamstra, 1996 quoted in Schaerlaekens <i>et al.</i> , 1999	

Table 5: Anaerobic first order decay constants for TCE.

Parent species	Rate constant (day ⁻¹)	Source	Notes
TCE	0.000067 - 0.000012	Clement <i>et al.</i> , 1999	lab scale microcosm, anaerobic
TCE	0.0001 - 0.0003	Martin and Imbrigiotta, 1994 quoted in Schaerlaekens <i>et al.</i> , 1999	Microcosm aquifer, anaerobic conditions
TCE	0.0003	Rügge <i>et al.</i> , 1999	anaerobic, overall value
TCE	0.00045	Clement <i>et al.</i> , 1999	calibrated values, anaerobic zone
TCE	0.0009	Clement <i>et al.</i> , 1999	calibrated values, anaerobic zone
TCE	0.0013	Rügge <i>et al.</i> , 1999	anaerobic, maximum value
TCE	0.001 - 0.003	Wilson <i>et al.</i> , 1994 quoted in Schaerlaekens <i>et al.</i> , 1999	in situ aquifer, anaerobic
TCE	0.0021	Lowry and Reinhard, 2000	anaerobic, catalyzed by Pd-on- γ -Al ₂ O ₃ , in the presence of 659 mg/L CO ₃ ²⁻ , pH = 10.9
TCE	0.0028	Lowry and Reinhard, 2000	anaerobic, catalyzed by Pd-on- γ -Al ₂ O ₃ , in the presence of 580 mg/L H ₂ CO ₃ , pH = 4.4
TCE	0.0033	Kleopfer <i>et al.</i> , 1985 quoted in Schaerlaekens <i>et al.</i> , 1999	Microcosm aquifer, anaerobic conditions
TCE	0.0037	Lowry and Reinhard, 2000	anaerobic, catalyzed by Pd-on- γ -Al ₂ O ₃ , in the presence of 1003 mg/L Cl ⁻ , pH = 9.6
TCE	0.005	Lowry and Reinhard, 2000	anaerobic, catalyzed by Pd-on- γ -Al ₂ O ₃ , in the presence of 660 mg/L HCO ₃ ⁻ , pH = 8.6

Table 5: Anaerobic first order decay constants for TCE. (continued)

Parent species	Rate constant (day ⁻¹)	Source	Notes
TCE	0.0056	Lowry and Reinhard, 2000	anaerobic, catalyzed by Pd-on- γ -Al ₂ O ₃ , in de-ionized water
TCE	0.0039 - 0.0077	Hardy <i>et al.</i> , 1999	anaerobic, reference book (calculated from half-life)
TCE	0.006	Bouwer <i>et al.</i> , 1981 quoted in Schaerlaekens <i>et al.</i> , 1999	Anaerobic conditions
TCE	0.0062	Barrio-Lage <i>et al.</i> , 1987 quoted in Schaerlaekens <i>et al.</i> , 1999	Microcosm, loamy sand, sulfate reducing
TCE	0.0076	Lowry and Reinhard, 2000	anaerobic, catalyzed by Pd-on- γ -Al ₂ O ₃ , in the presence of 690 mg/L SO ₄ ²⁻ , pH = 6.1 - 5.0
TCE	0.008	Wilson <i>et al.</i> , 1996 quoted in Schaerlaekens <i>et al.</i> , 1999	Microcosm, silt/clay/sand, anaerobic conditions
TCE	0.0009 - 0.0198	Hardy <i>et al.</i> , 1999	anaerobic, recent literature (calculated from half-life)
TCE	0.0039 - 0.099	Hardy <i>et al.</i> , 1999	anaerobic, field study (calculated from half life)
TCE	0.1	Anderson and McCarty, 1994	by methanotrophic biofilms
TCE	0.424	Lowry and Reinhard, 2000	anaerobic, catalyzed by Pd-on- γ -Al ₂ O ₃ , in the presence of 0.4 mg/L HS ⁻ , pH = 8.1
Average	0.0294		
Standard Deviation	0.0910		

Table 6: Aerobic and transition first order decay constants for TCE.

Parent species	Rate constant (day ⁻¹)	Source	Notes
TCE	0.000004	Clement <i>et al.</i> , 1999	calibrated values, transition (aerobic/anaerobic) zone
TCE	0.00001	Clement <i>et al.</i> , 1999	calibrated values, aerobic zone
TCE	0.0001125	Clement <i>et al.</i> , 1999	calibrated values, transition (aerobic/anaerobic) zone
TCE	0.00015 - 0.00054	Clement <i>et al.</i> , 1999	field estimated
TCE	0.004	Poulsen <i>et al.</i> , 1996 quoted in Schaerlaekens <i>et al.</i> , 1999	In situ, no addition

Table 7: Anaerobic first order decay constants for DCE.

Parent species	Rate constant (day ⁻¹)	Source	Notes
DCE	0.0000062 - 0.0000014	Clement <i>et al.</i> , 1999	lab scale microcosm, anaerobic
DCE	0.00065	Clement <i>et al.</i> , 1999	calibrated values, anaerobic zone
DCE	0.000845	Clement <i>et al.</i> , 1999	calibrated values, anaerobic zone
DCE	0.0017	Wilson <i>et al.</i> , 1994 quoted in Schaerlaekens <i>et al.</i> , 1999	In situ, aquifer, anaerobic conditions
DCE	0.0014 - 0.002	Wilson <i>et al.</i> , 1994 quoted in Schaerlaekens <i>et al.</i> , 1999	In situ, aquifer, anaerobic conditions
DCE	0.001 - 0.0026	Wilson <i>et al.</i> , 1994 quoted in Schaerlaekens <i>et al.</i> , 1999	In situ, aquifer, anaerobic conditions
DCE	0.004	Barrio-Lage <i>et al.</i> quoted in Schaerlaekens <i>et al.</i>	Microcosm sediment, anaerobic conditions
DCE	0.0052	Barrio-Lage <i>et al.</i> , 1986 quoted in Schaerlaekens <i>et al.</i> , 1999	Microcosm sediment, anaerobic conditions
DCE	0.006	Barrio-Lage <i>et al.</i> , 1986 quoted in Schaerlaekens <i>et al.</i> , 1999	Microcosm sediment, anaerobic conditions
DCE	0.007	Schaerlaekens <i>et al.</i> , 1999	Microcosm sediment, anaerobic conditions
DCE	0.0084	Wilson <i>et al.</i> , 1996 quoted in Schaerlaekens <i>et al.</i> , 1999	Alluvial silt/clay/sand, anaerobic conditions
DCE	0.0058 - 0.0231	Hardy <i>et al.</i> , 1999	reference book (calculated from half-life)
DCE	0.033	Hardy <i>et al.</i> , 1999	anaerobic, field study (calculated from half life)
DCE	0.0006 - 0.231	Hardy <i>et al.</i> , 1999	anaerobic, recent literature (calculated from half-life)
Average:	0.0143		
Standard Deviation:	0.0304		

Table 8: Aerobic and transition first order decay constants for DCE.

Parent species	Rate constant (day ⁻¹)	Source	Notes
DCE	0.0001625	Clement <i>et al.</i> , 1999	calibrated values, transition (aerobic/anaerobic) zone
DCE	0.00044 - 0.00064	Clement <i>et al.</i> , 1999	field estimated

Table 8: Aerobic and transition first order decay constants for DCE. (continued)

Parent species	Rate constant (day ⁻¹)	Source	Notes
DCE	0.0016	Clement <i>et al.</i> , 1999	calibrated values, transition (aerobic/anaerobic) zone
DCE	0.004	Clement <i>et al.</i> , 1999	calibrated values, aerobic zone

Table 9: Anaerobic first order decay constants for VC.

Parent species	Rate constant (day ⁻¹)	Source	Notes
VC	0.0005 - 0.002	Wilson <i>et al.</i> , 1994 quoted in Schaerlaekens <i>et al.</i> , 1999	In situ, aquifer, anaerobic conditions
VC	0.004	Clement <i>et al.</i> , 1999	calibrated values, anaerobic zone
VC	0.011 - 0.0037	Clement <i>et al.</i> , 1999	lab scale microcosm, anaerobic
VC	0.008	Clement <i>et al.</i> , 1999	calibrated values, anaerobic zone
Average:	0.0052		
Standard Deviation:	0.0031		

Table 10: Aerobic and transition first order decay constants for VC.

Parent species	Rate constant (day ⁻¹)	Source	Notes
VC	0.0008	Clement <i>et al.</i> , 1999	calibrated values, transition (aerobic/anaerobic) zone
VC	0.00082	Clement <i>et al.</i> , 1999	field estimated
VC	0.001	Clement <i>et al.</i> , 1999	calibrated values, transition (aerobic/anaerobic) zone
VC	0.002	Clement <i>et al.</i> , 1999	calibrated values, aerobic zone

Table 11: Summary of anaerobic first order decay constants for chlorinated ethenes.

	PCE	TCE	DCE	VC
Minimum Value:	3.10E-04	6.70E-05	6.20E-06	5.00E-04
Maximum Value:	9.14	0.424	0.231	0.008
Average Value:	2.2792	0.0294	0.0143	0.0052
Standard Deviation:	2.2612	0.0910	0.0304	0.0031

As mentioned above, many of the cited articles involve tests on specific bacteria, which have been specifically engineered to make them proficient at dechlorinating the chlorinated ethenes. These types of tests sometimes result in unnaturally high decay rates, which would not normally be found naturally at contamination sites. In the values presented for PCE, there is a distinct difference between these types of tests and those run with in situ conditions and microbes. It is useful to re-evaluate the PCE values without including these high values. Table 12 shows the summary for PCE without including these unnaturally high values.

Table 12: Adjusted analysis of anaerobic first order decay rates for PCE.

	PCE
Minimum Value:	3.10E-04
Maximum Value:	0.07
Average Value:	0.00384
Standard Deviation:	0.00348

Chapter 6 – Hypothetical Application

A series of tests were performed with a hypothetical model in order to validate the parameter estimation subroutines added to the ART3D code. These tests also served to determine how well the ART3D code operates under an optimization environment and to gain insight into the inverse modeling process when applied to an analytic transport model.

The testing process began by defining the model solution that would serve as the benchmark or “field observations” for the inverse model. A normal forward run was randomly designed. The simulation had four species defined: c1, c2, c3, and c4. For simplicity, the reaction was constrained to be sequential and only one time step was considered. Table 13 shows the parameters used in this simulation. Because this is a theoretical simulation, no units are included

Table 13: Parameters for benchmark simulation.

Grid length, x direction:	40	Retardation coefficient:	0.8
Grid length, y direction:	10	Velocity	0.5
Grid length, z direction:	4	Dispersion, x direction:	0.08
Number of grid cells, x dir.:	40	Dispersion, y direction:	0.009
Number of grid cells, y dir.:	10	Dispersion, z direction:	0.0
Number of grid cells, z dir.:	1	Decay coefficient, c1:	0.075
Source dimension, y direction:	4	Decay coefficient, c2:	0.05
Source dimension, z direction:	4	Decay coefficient, c3:	0.02
Simulation time:	50	Decay coefficient, c4:	0.045
Yield coefficients:		Source concentration, c1:	100.0
c1 → c2:	0.75	Source concentration, c2:	0.0
c2 → c3:	0.5	Source concentration, c3:	0.0
c3 → c4:	0.8	Source concentration, c4:	0.0

Once the parameters had been chosen, a group of observation points was added to the project and the simulation was run in forward mode. As part of the output, ART3D provided the concentrations of each of the species at each observation point. ART3D also provided four data sets showing the concentration values at each grid point. This solution was used as the benchmark and the concentration values at the observation points were used as the observed values in the subsequent model runs in inverse mode. For simplicity's sake, only five of the parameters were designated as "unknown": the four decay coefficients and the dispersion value in the x direction. Multiple tests were conducted with different numbers of observation points and with different geometries.

There were several objectives of this case study. The first was to test the PORT library and find out if it could return the parameter values from the benchmark solution when the starting values were changed. Another purpose was to determine how many observation points are needed to obtain an accurate result from the optimization routine. Further, several different geometries of points were to be tested to learn about the optimal arrangement of data points.

The rule of thumb normally used in parameter estimation is that there must be no more unknowns than there are observation points. Using the rule of thumb, five observation points should be required to obtain a good result since there are five values being optimized. It was expected that the difference between the optimized solution and the benchmark solution would be highest if there was only one observation point in the simulation. That error was expected to gradually shrink as more observation points were added to the simulation, reaching an acceptably low value when there were five points. It was also anticipated that, although the error might continue to decrease as the number of

observation points increased beyond five, there might be some oscillation caused by the optimization routine.

However, during the running of these tests, we realized that since each observation point has four known values, each concentration value represents an independent observation. Thus, one point with four concentration values can be used to solve for four unknown parameters. As a result, we were able to achieve an excellent match for the parameter values even with only one or two observation points.

With regard to determining the best layout for the observation points, we concluded that this model and this type of hypothetical case study were not adaptable to this kind of analysis. With a model as simple as ART3D, it is usually best to have all of the observation points along the centerline of the plume since the flow is one dimensional and the aquifer is assumed to be homogenous. Because neither the vertical nor transverse dispersion were allowed to vary and because the source dimension in the z direction was the same as the depth of the grid, points off of the centerline did not add significant accuracy to the optimization. These difficulties were compounded by the fact that the hypothetical case study was based on a previous benchmark ART3D simulation. In a real scenario, the true plume will probably not exactly follow the mathematical equations used to run ART3D. Even more importantly, the requirement of homogeneity and averaging of several of the parameters over the entire area makes it very unlikely that even with the best parameter values, the model could exactly predict the plume size, location and concentration in a real modeling situation. In this case, where the optimized parameters exactly predicted the benchmark plumes, the optimization was very easily

obtained. The difference between the optimized solutions and the benchmark solutions were not significantly large in any of the inverse simulations.

In the end, the results shown here are from one geometric arrangement of the points (shown in Figure 8). The points were spaced two units apart. The test began with only one point near the center of the important portion of the plumes. The next test was run with three points: the same point used in the first test, plus two extra points two units away in each direction. The testing continued until there were eleven points. Figure 8 shows the points for the tests with one, five, seven and eleven points. According to this pattern, tests were run with three and nine points also.

Figure 8 also shows the output plumes from the benchmark simulation. It allows the reader to see that most of the points were well within the plume boundaries. It also nicely illustrates the output from a simple solution with each plume being placed slightly further from the source. (As with all ART3D simulations, the source is placed at the left end of the grid, and at the midpoint between the side boundaries. The groundwater flow is assumed to be from left to right.)

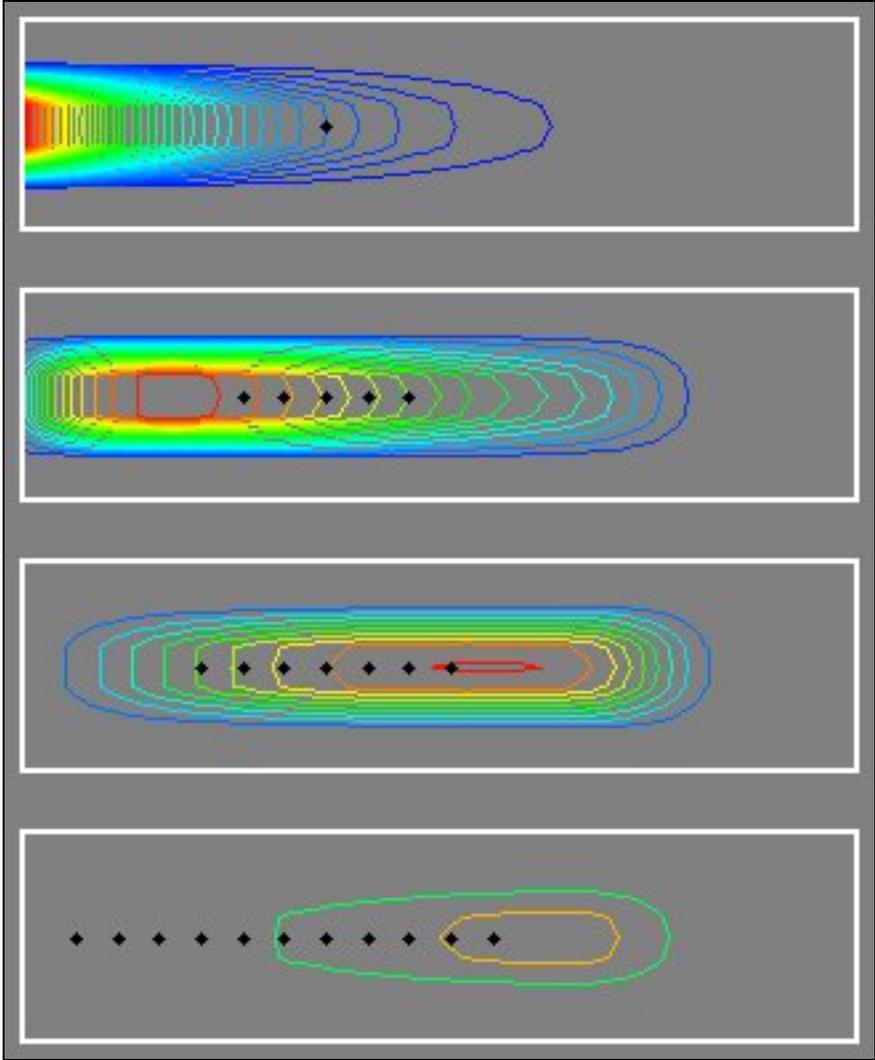


Figure 8: The first arrangement of points is illustrated here with the four species plumes. (Simulations were also run with three points and with nine points according to this pattern.)

The ranges for the parameters were set at 20% below and above the benchmark simulation value for that parameter. As explained in Chapter 3, the starting point can be very important in the accuracy of the final result of a parameter estimation run. To allow for this, five different starting values were randomly chosen for each unknown parameter. Table 14 below shows the five variables, their ranges and the five random start values.

Table 14: Parameter ranges and starting values. (Starting value colors correspond to Figures 9 through 11.)

Variable	Minimum	Maximum	Starting Values				
							
Alpha-x	0.064	0.096	0.069	0.082	0.067	0.074	0.072
k_c1	0.06	0.09	0.074	0.072	0.072	0.074	0.079
k_c2	0.04	0.06	0.046	0.041	0.052	0.057	0.043
k_c3	0.016	0.024	0.016	0.020	0.016	0.019	0.020
k_c4	0.036	0.054	0.052	0.050	0.038	0.044	0.044

When the tests were run, the error values turned out to be extremely small in all cases. The results from the first point arrangement is shown and described below:

The error between the optimized solution and the benchmark solution was determined in three different ways. First, the error in the calculation of the observation point concentrations was gathered from the output (*.out) file. Under normal circumstances, this would be the only error analysis available to a modeler. The graph in Figure 9 shows the mean absolute residual error averaged across all the observation points. Each bar color represents a different starting point as described in Table 14. Each group of bars represents the error obtained using a different number of observation points according to the geometry in Figure 8.

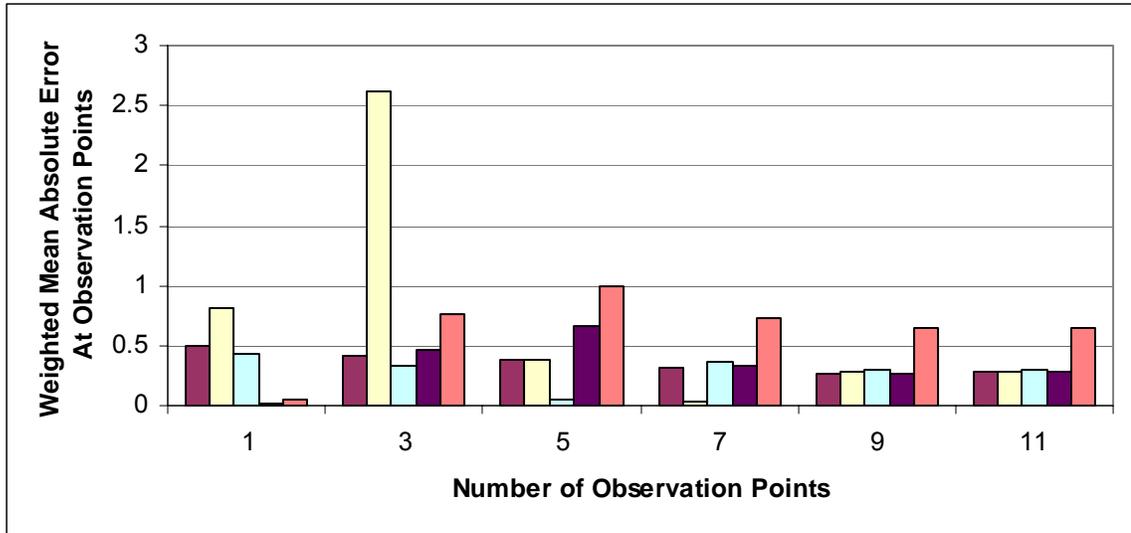


Figure 9: A comparison of the error at the observation points for hypothetical case study. Each colored bar represents a different starting point as shown in Table 14. Each set of five bars represents the five tests run with each set of observation points.

The second method of determining the success of the PORT library was to compare the benchmark parameters (the goal of the optimizer) with the optimized parameter values determined from an inverse ART3D run. The percent error was calculated by dividing the difference between the two values by the benchmark value and multiplying by 100. The errors found on the same simulations are summarized in Figure 10. As before, each different color bar represents a different randomly selected starting point. Each group of bars represents the test run with a different number of observation points.

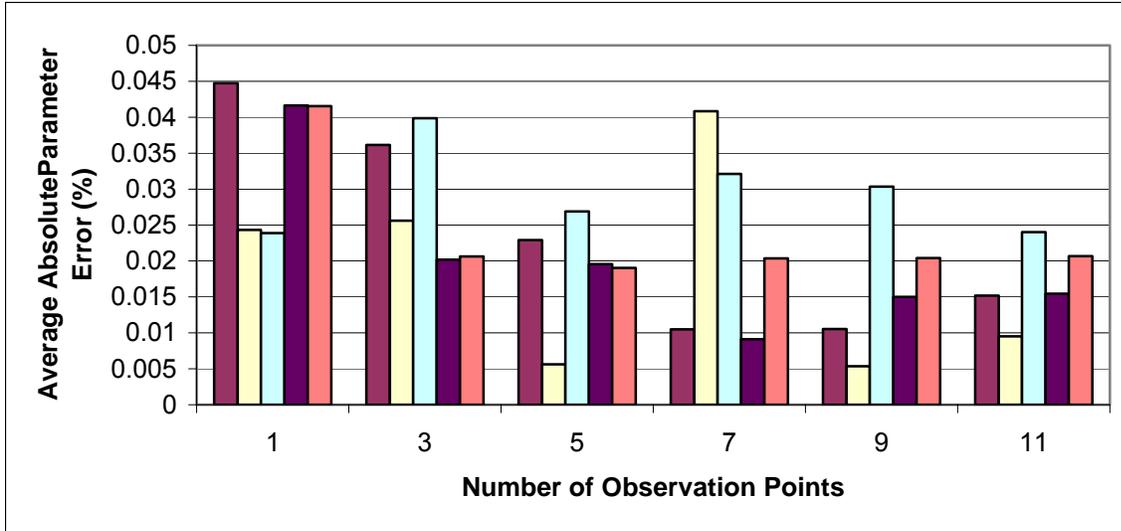


Figure 10: A comparison of the error in the parameters returned by the optimization code for hypothetical case study. Each colored bar represents a different starting point as shown in Table 14. Each set of five bars represents the five tests run with each set of observation points.

The third error calculation was made on the grid as a whole. This calculation was made by comparing the grid data sets from the benchmark simulation and the optimized data sets. Using the *Data Calculator* in GMS, four new data sets were created by taking the absolute value of the difference between the benchmark data sets and the “optimized” data sets for each specie. To make this calculation, GMS takes the absolute value of the difference of each grid point in the two data sets and returns a new grid whose scalar values at each grid point are equal to that difference. The *Data Set Info* dialog then allows access to statistical information on the new data set, including the mean value. To calculate the total error across all the species, the four mean values (one for each specie) were averaged. These averages are presented in the graph below, again with a different color bar representing a different starting point, and each group of bars representing a different number of observation points.

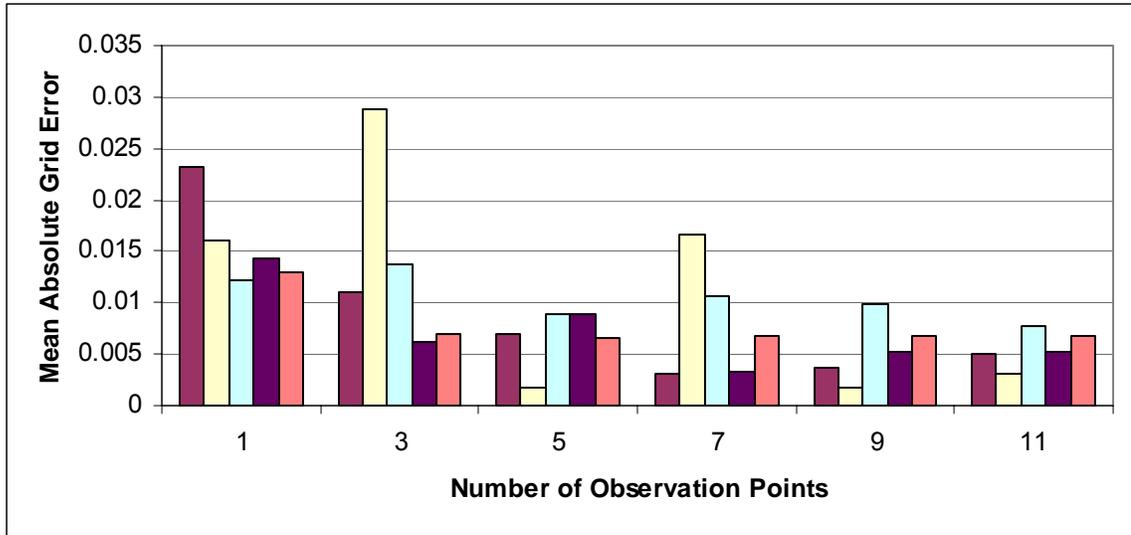


Figure 11: A comparison of the error in the grid data sets returned by the optimization code for hypothetical case study. Each colored bar represents a different starting point as shown in Table 14. Each set of five bars represents the five tests run with each set of observation points.

Although this hypothetical testing did not yield purposeful results describing the optimum number of points or geometry needed to obtain a good optimized set of values, some interesting conclusions can be drawn from the data obtained. Although the general trend in the graphs seems to show that more points can yield better accuracy, it is interesting that that is not necessarily true for a given starting point. For example, in Figure 11, the cyan bar yielded the best results for a simulation with one observation point, the purple bar for three points, the yellow bar for five points, and so on. This shows that there is not one starting point that is best in all cases and the user should try several starting points within the specified range to be sure the best result is obtained.

Another interesting observation is that the parameter error and observation point error, while showing a general trend of lower error with more points, show a lot of oscillation and noise and actually do not have a significant increase in accuracy with the increase in the number of points. The grid error graph, however, does show a nice

downward trend with less noise and oscillation. One reason for this is that the increased number of points with this geometry results in the points covering a larger portion of the grid and the major portions of all the plumes. This means there can be greater overall accuracy in the grid data sets without an accompanying increased precision in the observation point values. The lesson is that observation points should be scattered over the entire plume area and not concentrated in a single area.

The final conclusion from this hypothetical case study is that the observation point error does not closely mirror the grid or parameter error plots. In the case of a real groundwater study, the only error value available to the user for analysis, is the error associated with the observation point. But, sometimes a low observation point error occurs when there is a high grid error, making it difficult to know which calibration is actually the most accurate. This challenge makes a strong case for the use of stochastic modeling to show statistical results rather than one “best guess” result. Because there is no perfect way to know when or if the model exactly shows the true nature of the aquifer characteristics and contaminant flow, it is best to run several simulations with equally likely parameters and consider a range of possibilities as the result.

It should be noted that, regardless of how the errors were computed, they were extremely small for all inverse model runs using the hypothetical model. As mentioned above, this is due to the fact that this is a comparison of the optimized values from an analytic model with a previous solution of the same analytic model. The errors are so small that most of the solutions may be considered “exact”. This is encouraging, as it shows that the optimization routine works favorably on this type of model.

Chapter 7 – Case Study

As a final assessment of the capabilities of ART3D, the code was tested by applying it to a case study involving a real contaminated site. A Superfund site in Baton Rouge, Louisiana was selected for the analysis. The area was a waste dumping site for the Petro-Processors, Inc., Brooklawn Site and had been contaminated with various hazardous wastes, including several types of chlorinated solvents. This site is the basis of a paper illustrating the use of BIOCHLOR as one of the steps in the US EPA Monitored Natural Attenuation Screening Protocol for chlorinated solvents (Clement, et al 2002). The fact that the site had already been modeled using BIOCHLOR, made it a good choice for a case study since ART3D evolved from BIOCHLOR. The comparison of the two models allows an investigation of the usefulness of improvements made in ART3D.

Data Collection

Most of the data for the case study came from Dr. Clement's paper and included aquifer characteristics, the known source species and their characteristics as well as basic information used to run BIOCHLOR. Although not included in the paper, Dr. Clement provided field concentration data obtained from four wells at the site. The ART3D model was designed using the same conceptual model presented in the BIOCHLOR paper. However, some minor changes were made to the model design.

One of the main difficulties with BIOCHLOR, which has been eliminated by ART3D, was the requirement that reactions be sequential. BIOCHLOR cannot handle complex, convergent or divergent reactions. The main contaminants at the Brooklawn site were chlorinated ethenes and chlorinated ethanes. (The chlorinated ethanes include Tetrachlorethane - TeCA, Trichloroethane - TCA, dichloroethane - DCA, and chloroethane - CA.) The graphic below describes the reactions that occur with these species and shows that there are several complex pathways that can be taken, few of which can be modeled as simply sequential.

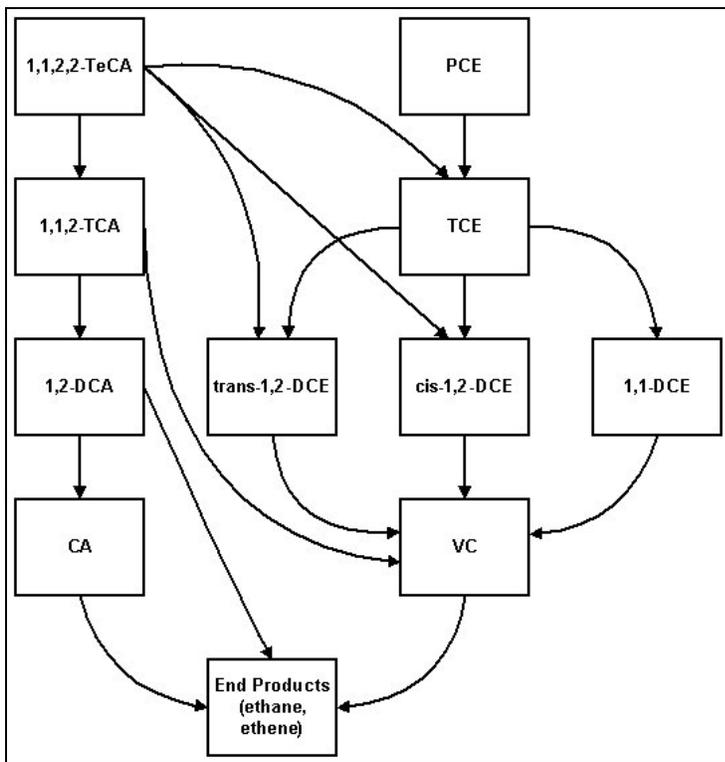


Figure 12: Possible pathways for anaerobic decay of chlorinated ethenes and ethanes.

In order to model this site using BIOCHLOR, the ethane and ethene corridors had to be analyzed separately and then the contribution of the ethane line to the ethene line was approximated and a third simulation was run of the ethane line. This resulted in a rough approximation that fairly closely matched the field data. Using ART3D, however, both corridors were modeled simultaneously and some of the approximation became unnecessary.

Some simplifications to the BIOCHLOR model were carried over into the ART3D model. As in Clement's paper (2002), all three DCE isomers were combined together. Also, the degradation to ethene or ethane was not included in the model because of a lack of data.

The conceptual model was used in Clement et al. (2002) to determine the size and dimensions of the grid, the time variables and the source dimensions. These values are summarized in Table 15 below.

Table 15: Model simulation time, regional dimensions and source dimensions.

Property	Value
Source dimension, y direction	1000 ft
Source dimension, z direction	45 ft
Total simulation time	25 years, or 9125 days
Model depth	45 ft
Model width	2000 ft

The BIOCHLOR model length was set at 5000 ft, the distance to the Mississippi River. In the ART3D model, this was shortened to 600 feet, except when the 200-year plume was calculated, when the length was extended to 1200 feet. An even larger grid

was used to look at the long-term placement of the vinyl chloride plume. This was sufficient to show the main section of the plume, but resulted in faster calculations than the larger grid.

Data describing the characteristics of the aquifer and the chemical species were also obtained from the Clement et al. (2002) paper. Because separate BIOCHLOR simulations were run with ethane and ethene corridors, two retardation values were used. They were averages of values derived for each constituent. Because ART3D can only take one retardation coefficient, the value shown below is an average of the two presented in the BIOCHLOR paper.

Table 16: Aquifer and constituent parameter values from the BIOCHLOR model.

Property	Value
Retardation coefficient	3.0
Velocity	0.05 ft/d
Longitudinal dispersivity	50 ft
Dispersivity ratio, α_v/α_x	0.1
Dispersivity ratio, α_z/α_x	0.01
Decay constant, TeCA	0.014 d ⁻¹
Decay constant, TCA	0.013 d ⁻¹
Decay constant, DCA	0.001 d ⁻¹
Decay constant, CA	0.014 d ⁻¹
Decay constant, PCE	0.005 d ⁻¹
Decay constant, TCE	0.005 d ⁻¹
Decay constant, DCE	0.005 d ⁻¹
Decay constant, VC	0.0006 d ⁻¹

In the original BIOCHLOR simulation, the source concentrations were calculated in two ways. First, they were measured from samples taken at the location of the contamination source and then, they were calculated from a solubility analysis. The

direct measurement was done with equipment whose detection limit was set at 25 mg/L.

Table 17 lists these values:

Table 17: Source concentration values.

Specie	Measured value (mg/L)	Calculated value (mg/L)
TeCA	57	52
TCA	337	210
DCA	600	184
CA	<25	0.0
PCE	<25	3.5
TCE	<25	6.2
DCE	<25	0.8
VC	<25	1.0

Clement (2002) also provided the yield constants for each of the possible decay pathways in units of moles/moles. Using these values and the molecular weight of each specie, the yield values were calculated in g/g. These values are shown below:

Table 18: Yield constants used in both the BIOCHLOR simulation and the ART3D simulation.

Reaction	Yield constant (mole/mole)	Yield constant (g/g)
PCE → TCE	1.0	0.792
TCE → DCE	1.0	0.738
DCE → VC	1.0	0.645
TeCA → TCA	0.35	0.278
TeCA → TCE	0.02	0.016
TeCA → DCE	0.63	0.364
TCA → DCA	0.2	0.148
TCA → VC	0.8	0.375
DCA → CA	0.7	0.456

The last part of the data used in the ART3D simulation was the observation point data. At four points in the plume, boreholes were used to determine the concentration of some of the constituents. Although the exact locations of each point are not given, all were within the source dimension and it is a good approximation to assume (as was done in the BIOCHLOR model) that they were on the centerline of the plume. The field data and point locations are given in Table 19.

Table 19: Field observation data.

Point	1	2	3	4
Distance from source (ft):	150	300	400	500
Constituent	Concentration (mg/L)			
TeCA	No Data	0.01	No Data	0.007
TCA	No Data	0.02	No Data	0.02
DCA	0.01	No Data	0.1	0.3
CA	No Data	No Data	No Data	No Data
PCE	0.05	0.02	No Data	No Data
TCE	0.04	No Data	No Data	No Data
DCE	0.013	No Data	0.5	0.02
VC	37	No Data	12	3.5

Forward Run and Inverse Runs

The first simulation run on the data presented in the previous section was a normal forward run. This run gave an idea of the general locations and concentrations of the plumes. It quickly became obvious that the observation points were placed too far from the source of the contamination. They were far from the measurable part of the plumes, near the edge. At these locations, they could not pick up the significant section of the plumes. This made parameter estimation difficult.

Once the forward run had been completed, a series of inverse runs using the PORT library were completed. Table 19 shows that there are 16 known values of concentration. From the hypothetical case study described in the previous chapter, it is apparent that the optimization algorithm can only solve for 16 of the parameters in order to prevent an under-constrained equation. Because there are no observation values for chloroethane, there is no purpose in solving for its source concentration, since that value will have no effect on the value of the objective function. In addition, the simulation was not set to solve for the source concentration for vinyl chloride, the two dispersivity ratios or the velocity. Occasionally during the calibration of the model, these values were varied by hand in an effort to improve the correlation between the model output and the field data. Also, while the PORT library did not solve for the dispersivity ratios, the transverse and vertical dispersivity values did change whenever the longitudinal dispersivity value changed. In this way, the solver could optimize all three values without adding extra degrees of freedom.

The observed concentration values were used to determine a standard deviation and a weighting value to be applied to each specie measured in the observation boreholes. This standard deviation was used as described in Chapter 3 to determine the weighting value used in the calculation of the objective function. The standard deviation for each specie was assumed to be 20% of the average measurement of that specie across those wells that detected it. Table 20 shows the standard deviation used for each chemical constituent and the weight values. These same weight values were applied to every measurement of a specie, regardless of the location of the point.

Table 20: Standard deviation and weight values applied to each chemical specie.

Constituent	Standard Deviation	Weight
TeCA	0.0004	6.25×10^6
TCA	0.0001	1.00×10^8
DCA	0.007	2.04×10^4
PCE	0.002	2.50×10^5
TCE	0.002	2.50×10^5
DCE	0.009	1.23×10^5
VC	0.893	1.25

The calibration was achieved by repeatedly running inverse simulations with different randomly selected starting points. Later, the automatic optimization was combined with hand calibration as random starting values were replaced with starting values specifically chosen in an effort to better match the field data. After setting some of the starting values to numbers that were expected to give a better match, the optimization routine was used to polish the hand calibration.

During the calibration, it quickly became obvious that it was quite easy to match up the concentration values measured at the two points closest to the source(1 & 2), but it was harder to get the model results at the two further points (3 & 4), to match the field data. The model consistently output values several orders of magnitude smaller than the field observed values at these more distant points. This can be explained, as the microbes near the source should be better adapted to break down the chlorinated compounds since they have been exposed to the contaminants at higher concentrations and for a longer period of time. Those microbes located 400 or 500 feet from the source, where the concentration is smaller and where the plume has arrived relatively recently have not had as much time to adapt to the contaminated environment. Thus, a single decay constant

for each specie is not sufficient since the microbes in some areas break down the chemicals faster than those in other areas. The differing decay constants can also be caused by differences in the availability of carbon, oxygen or other necessary elements.

To deal with this problem, three separate calibrations were prepared. The first was done with the above weight values as described previously; a second calibration was made with high weight values applied to the two closest points; and a third with high weight values on the two farthest points. These standard deviations and weight values are shown in Table 21. The low standard deviations (for calculating a high weight) were half of those presented in Table 20, while the high values were double that originally calculated.

The plots in Figure 13 show the centerline concentrations for each calibration of the plumes plotted with the observed data.

Table 21: Adjusted weight values for second and third calibrations.

Constituent	High Weight Points		Low Weight Points	
	Standard Deviation	Weight	Standard Deviation	Weight
TeCA	0.0002	2.50×10^7	0.0008	1.56×10^6
TCA	0.00005	4.00×10^8	0.0002	2.50×10^7
DCA	0.0035	8.16×10^4	0.014	5.10×10^3
PCE	0.001	1.00×10^6	0.004	6.25×10^4
TCE	0.001	1.00×10^6	0.004	6.25×10^4
DCE	0.0045	4.94×10^4	0.018	3.09×10^3
VC	0.4465	5.02	1.786	0.31

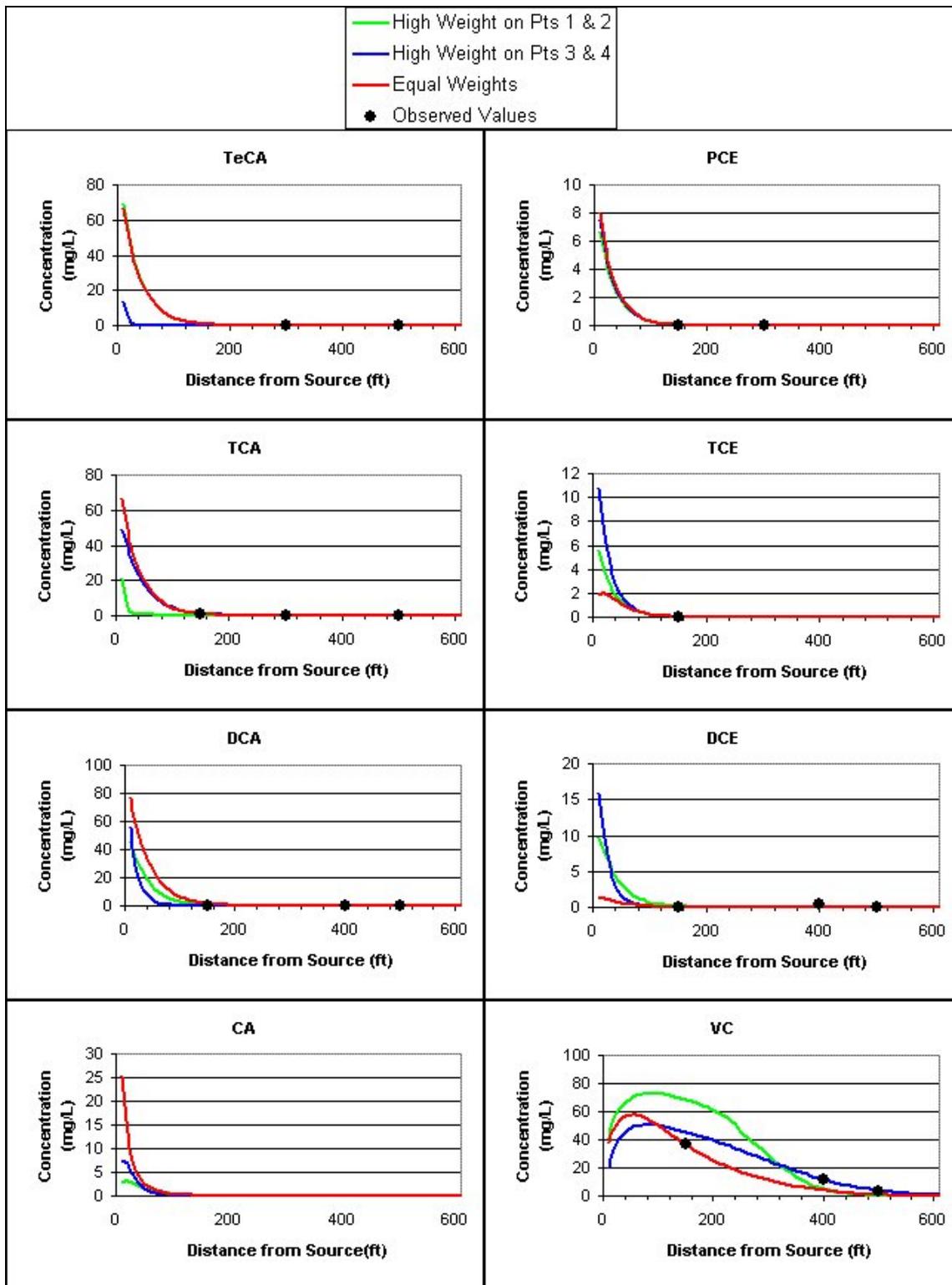


Figure 13: The centerline plots for each of the three calibrations at t=25 years, plotted against the observed data

The plots in Figure 13 show that computed concentration values do not differ very much from one calibration to the next. If the y-axis scale were smaller, some variation in the values would be seen, but it is very insignificant at such low values. The main differences between the calibrations occur in the first 100 feet from the source. Observation data in this area would greatly help the accuracy of the model. This shows the importance of obtaining field data in the significant areas of the plumes, not near the edges as has been done here.

The only plot in Figure 13 with significant variation between the three calibrations is the vinyl chloride (VC) plot. This plot seems to show that either the equal weighting calibration or the calibration with high weights on points 3 and 4 could be chosen as the best calibration. An analysis of the parameters returned by the PORT library (Table 22) shows that the parameters for the equal weighting calibration are the most likely. The third calibration requires extremely low source concentrations and decay constants, which are not likely, considering the measured values presented above. Thus, the first calibration, with equal weighting on all the points, will be used in the rest of this analysis.

Table 22: Optimized parameter values for each of the three calibrations.

Parameter	Calibration 1 (Equal weighting)	Calibration 2 (High weight on the closest points)	Calibration 3 (High weight on the furthest points)
Retardation	2.24	1.5	1.50
Velocity (ft/day)	0.072	0.05	0.05
Dispersivity (x) (ft)	10.00	30.9	99.67
Dispersivity (y) (ft)	1.00	3.09	9.967
Dispersivity (z) (ft)	0.10	0.309	0.9967
k TeCA (day ⁻¹)	0.0028	0.0030	0.0004
k TCA (day ⁻¹)	0.031	0.0025	0.0005

Table 22: Optimized parameter values for each of the three calibrations. (continued)

Parameter	Calibration 1 (Equal weighting)	Calibration 2 (High weight on the closest points)	Calibration 3 (High weight on the furthest points)
k DCA (day ⁻¹)	0.0031	0.010	0.0002
k CA (day ⁻¹)	0.014	0.019	0.000011
k PCE (day ⁻¹)	0.0034	0.0037	0.010
k TCE (day ⁻¹)	0.0076	0.0080	0.0005
k DCE (day ⁻¹)	0.012	0.015	0.005
k VC (day ⁻¹)	0.0001	0.0001	0.00001
C _o TeCA (mg/L)	93.19	47.31	0.013
C _o TCA (mg/L)	100.0	55.26	0.027
C _o DCA (mg/L)	50.0	107.39	0.0001
C _o CA (mg/L)	0.0	0.0	0.0
C _o PCE (mg/L)	9.39	10.72	0.058
C _o TCE (mg/L)	8.23	16.86	0.0015
C _o DCE (mg/L)	11.12	20.44	0.004
C _o VC (mg/L)	0.0	0.0	0.0

Determination of the Effect of Natural Attenuation

The purpose of the original BIOCHLOR paper was to determine if natural attenuation was occurring and if it could be relied on to clean up the spill before the contaminants reached the Mississippi River, about 5000 feet from the source location. This was done by running the BIOCHLOR simulation twice more using the same parameters, but disregarding decay in one simulation, and disregarding decay and sorption in the other. In this way, the effect of natural attenuation on the plume could be discerned. The same process was followed here by running normal forward simulations with the optimized parameters. Sorption was disregarded by setting the retardation constant to one; decay was ignored by setting each of the decay constants to zero. All other parameters were left at their optimized values. The results are displayed in Figure 14. Both chloroethane (CA) and vinyl chloride (VC) show concentrations of zero when

there is no decay. This is because their source concentrations were optimized at 0.0 mg/L and without decay, none of the other species can react to form either CA or VC.

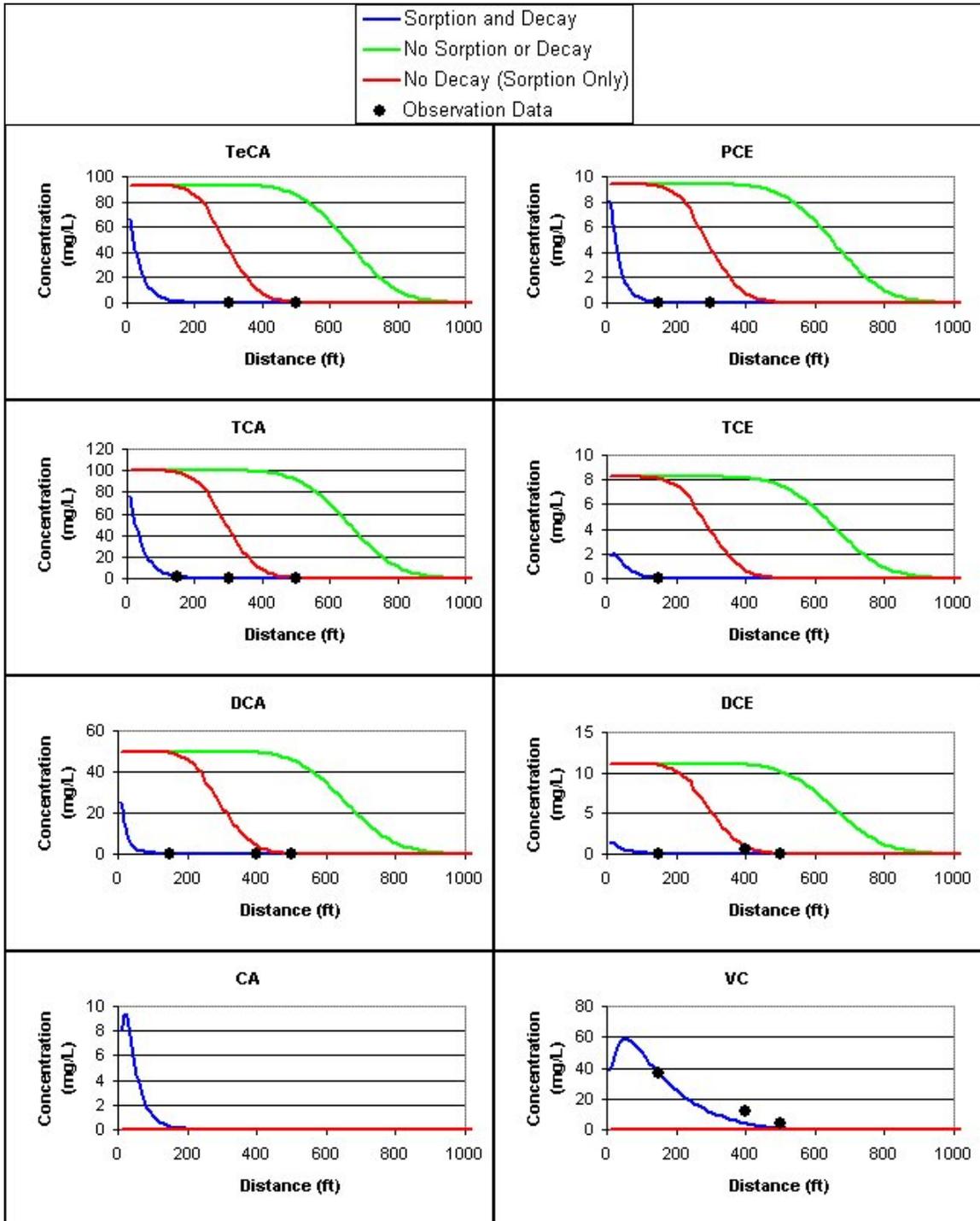


Figure 14: The effects of sorption and decay on the model output at t = 25 years.

The plots in Figure 14 show that natural attenuation is occurring in the aquifer. The observed data closely coincides to the lines including decay and sorption, but is far from those not including decay and sorption. Even accounting for model errors, it is unlikely that the calculations without decay or sorption could ever be similar to the field data.

The plots in Figure 14 also show that, even without natural attenuation, after 25 years of movement, the plumes would only have moved 1000 feet from the source. There are another 4000 feet to go before the Mississippi River is encountered. That could take another 100 years at this rate. The risk of contaminating the river is even lower when decay and sorption are included.

Determination of Steady State Plume

Another objective of the original BIOCHLOR simulation was to determine when and where the plumes would reach steady state. To repeat this analysis, ART3D was run with a simulation time length of 200 years and with a time step every 25 years. The plots below show the resulting centerline plots after 25, 50 and 200 years with and without natural attenuation. For those simulations without natural attenuation, both retardation and decay were ignored by setting the decay constants to zero and the retardation factor to one. As in the previous plots, CA and VC without natural attenuation are displayed as flat lines at zero since their initial source concentrations are zero.

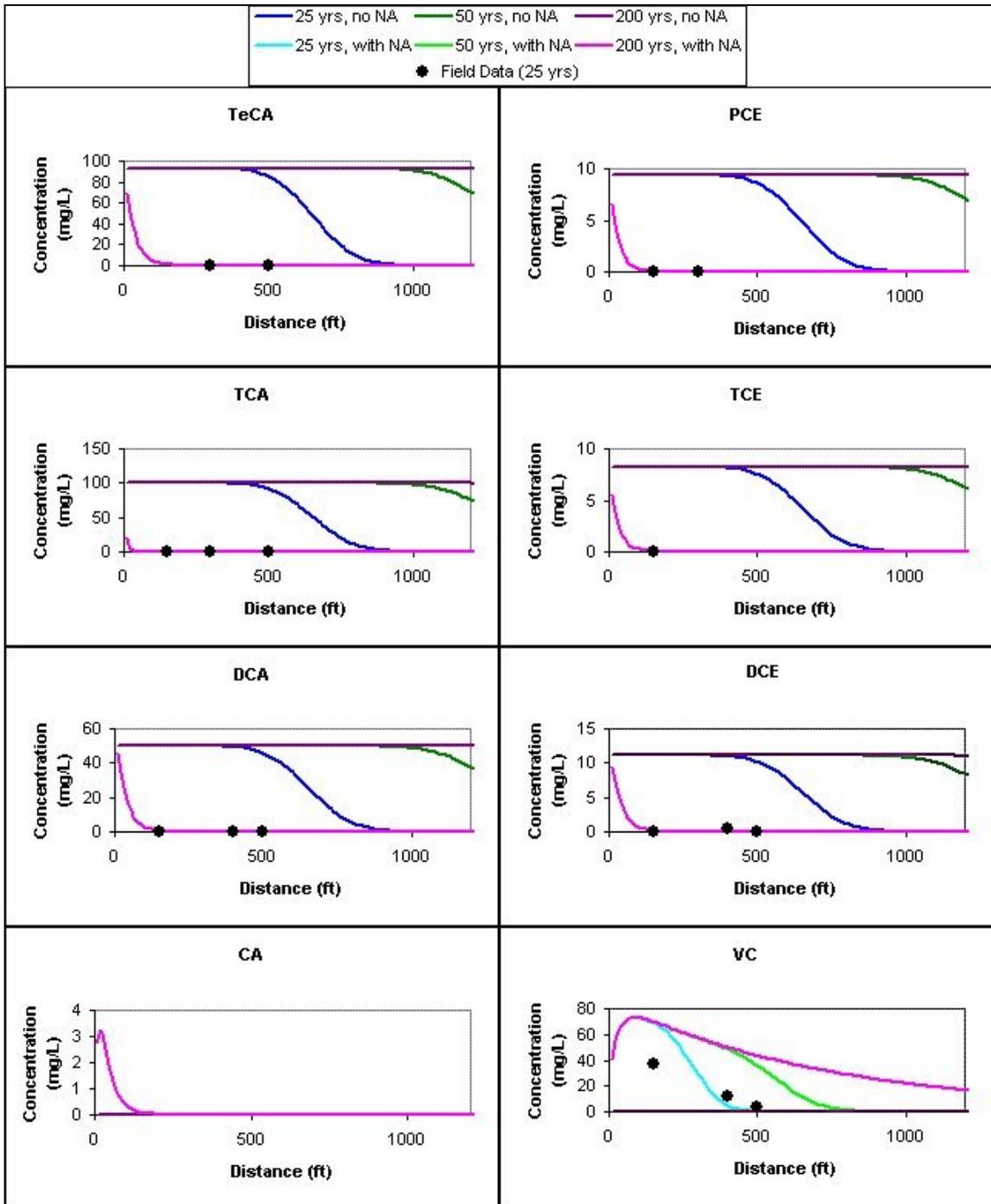


Figure 15: Centerline plots with and without natural attenuation (NA) after 25, 50 and 200 years, shown with field data

In the plots above (Figure 15), all species except vinyl chloride have reached a steady state by 25 years from the start of the simulation when natural attenuation is

considered. This is revealed by the fact that the natural attenuation lines are equivalent for each of the three time steps. In the vinyl chloride plot, the three time steps do not coincide with each other and the plume continues moving towards the Mississippi River after 200 years.

Without natural attenuation, the plume moves steadily away from the source with no reduction in concentration. The good match between the natural attenuation plots and the field data, however, shows that this is not a serious concern since the concentration levels of the field data indicate that attenuation is occurring.

The VC plot is shown again in Figure 16 with an extended grid and time period to show the time and location of the steady state plume.

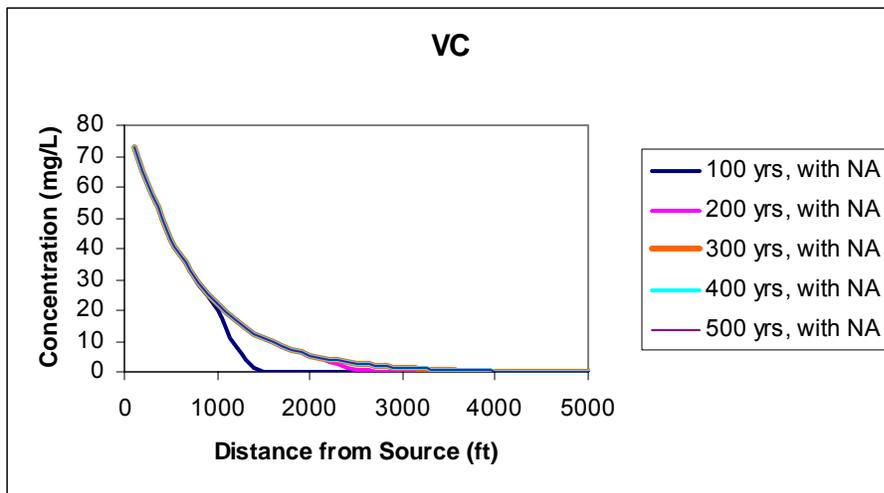


Figure 16: Movement of the vinyl chloride plume through 500 years with natural attenuation

Figure 16 indicates that the VC plume will reach steady state at about 300 years from the time of the initial contamination. It also shows that this plume will reach undetectable levels about 1500 feet before reaching the Mississippi River.

Stochastic Simulation

The addition of stochastic capabilities to ART3D allows this analysis of the Brooklawn site to go beyond that available in a BIOCHLOR simulation. In the stochastic simulation, selected parameters were assigned distribution data and ART3D ran a number of simulations, each time with different randomly chosen parameters. For this simulation, the ranges and distribution of the four ethene decay values (PCE, TCE, DCE, VC) were chosen based on the literature search described in Chapter 5. The calculated mean values, minimum values, maximum values and standard deviations were entered into ART3D as listed in Tables 11 and 12. The PCE values used were those calculated without using the high values from tests involving specially developed bacteria (Table 12). All other parameters, except for the two dispersion ratios and the chlorinated ethane (CA) source concentration were allowed to vary with a normal distribution selected to cover the range of acceptable values. These statistical values are shown below in Table 23. All other input parameters such as the yield coefficient matrix remained the same as in the previous simulations. The simulation time length was set at 25 years, to allow an analysis of current conditions, although the same procedure could be used to predict the future.

Table 23: Varied parameters in the stochastic simulation.

Parameter	Mean value	Maximum value	Minimum value	Standard deviation
Retardation	1.75	1.0	6.5	2.0
Velocity	0.072	0.0	0.1	0.03
Longitudinal dispersivity	11.18	0.0	50	12.0
TeCA decay constant	0.0029	0.0001	0.05	0.02
TCA decay constant	0.0032	0.0001	0.05	0.02
DCA decay constant	0.01	0.0001	0.05	0.015
CA decay constant	0.0048	0.0001	0.05	0.018
TeCA source concentration	90.17	0.0	100	40.0
TCA source concentration	100.0	0.0	400.0	100.0
DCA source concentration	50.0	0.0	700.0	250.0
PCE source concentration	11.46	0.0	25.0	5.0
TCE source concentration	0.63	0.0	25.0	9.0
DCE source concentration	0.002	0.0	25.0	9.0
VC source concentration	25.0	0.0	25.0	10.0

When the data had been entered, ART3D was set to run 100 times. Each run was completed with a new set of randomly chosen values according to the user-defined bounds and distributions. Using GMS, then, the 100 simulations could be used to complete a threshold analysis.

In a threshold analysis, the user selects a condition whose probability is desired. For example, a user might choose to view all locations where the PCE concentration is greater than 0.005 mg/L. GMS will run through all of the grid cells and data sets and compute the percentage of data sets that have a value greater than 0.005 mg/L at each point in the grid. The output from a threshold analysis is a data set showing percentages at every grid cell. This is the percentage of the data sets where the condition was true. If enough simulations are included in the analysis and if the parameter bounds and distributions have been well chosen, these percentages can represent the percent likelihood of the condition being true, based on the uncertainty of the input parameters.

For the threshold analysis of the Brooklawn site case study, the thresholds were chosen to be the US EPA’s Maximum Contaminant Level (MCL) obtained from the EPA website (USEPA, *Current Drinking Water Standards*, 2002). Values were not available for TeCA or CA, because these contaminants are more commonly air pollutants. Because the model was simplified by combining the three DCE isomers, three separate threshold analyses were run on the combined data set from the ART3D model. Each scenario is conservative since the concentration of a given isomer will be less than the total concentration of all three. Table 24 lists these contaminant levels.

Table 24: EPA Maximum Contaminant Level (MCL) for water pollutants at the Brooklawn site.

Contaminant	MCL (mg/L)
TCA	0.005
DCA	0.005
PCE	0.005
TCE	0.005
1,1-DCE	0.007
cis 1,2-DCE	0.07
trans 1,2-DCE	0.1
VC	0.002

Figures 17 through 24 show these threshold analysis outputs. The scales are in decimal percentages; the red sections are where 100% of the simulations showed that the MCL was exceeded, while the blue sections are where none of the simulations exceeded the MCL value. The figures only show the first 600 feet of the grid.

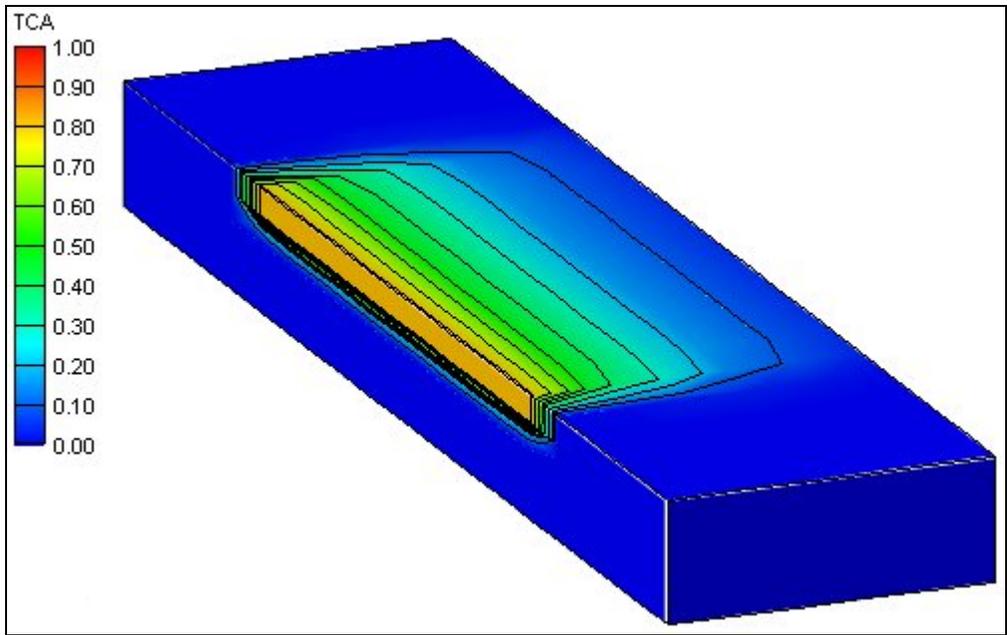


Figure 17: Threshold analysis result for TCA. MCL = 0.005 mg/L. Grid length is 600 ft.

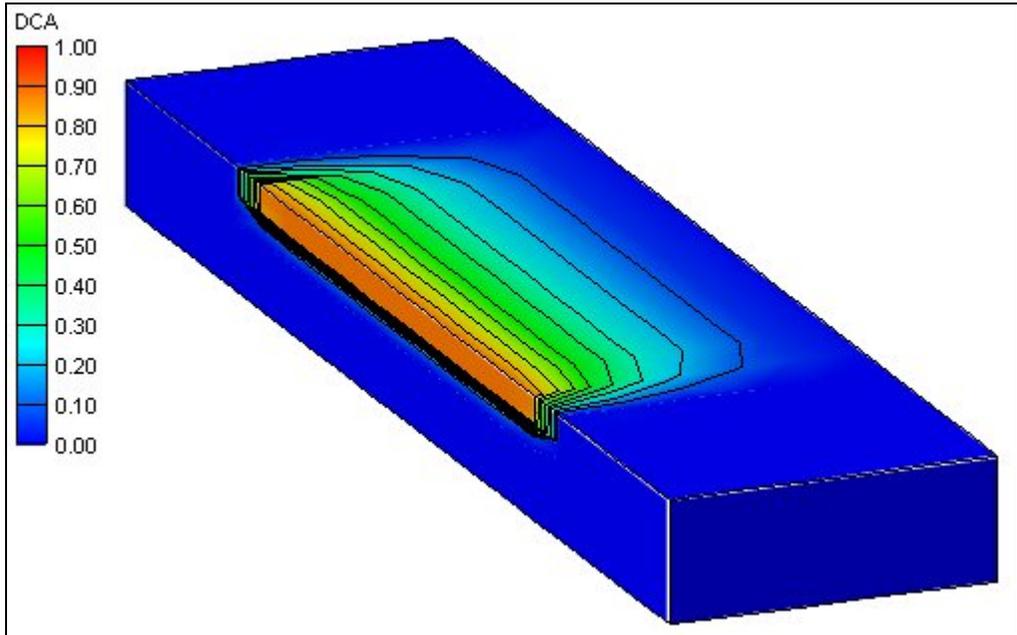


Figure 18: Threshold analysis result for DCA; MCL = 0.005 mg/L. Grid length is 600 ft.

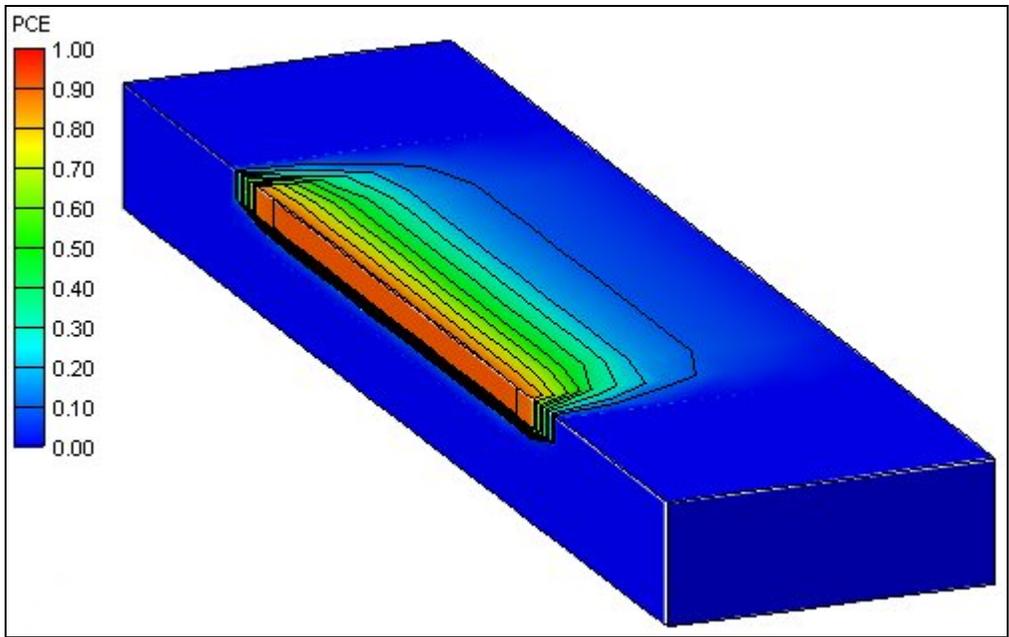


Figure 19: Threshold analysis result for PCE; MCL = 0.005 mg/L. Grid length is 600 ft.

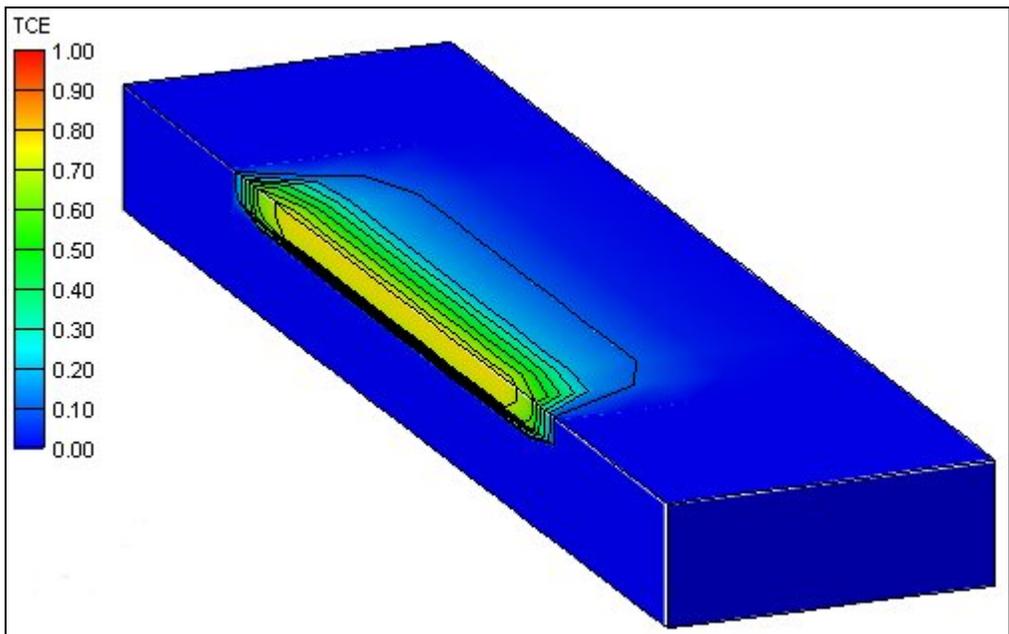


Figure 20: Threshold analysis result for TCE; MCL = 0.005 mg/L. Grid length is 600 ft.

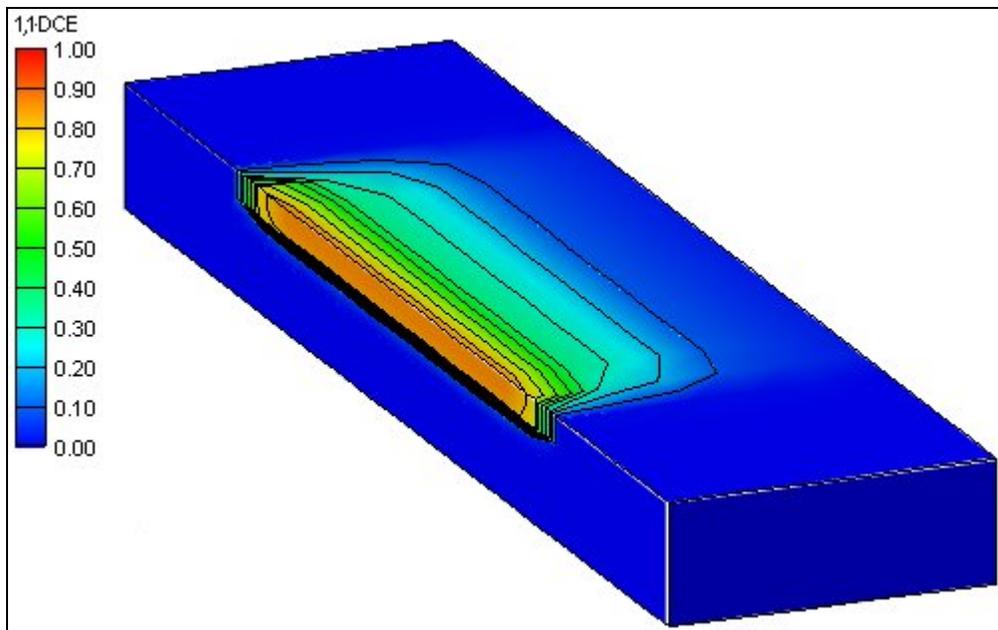


Figure 21: Threshold analysis result for 1,1-DCE; MCL = 0.007 mg/L. Grid length is 600 ft.

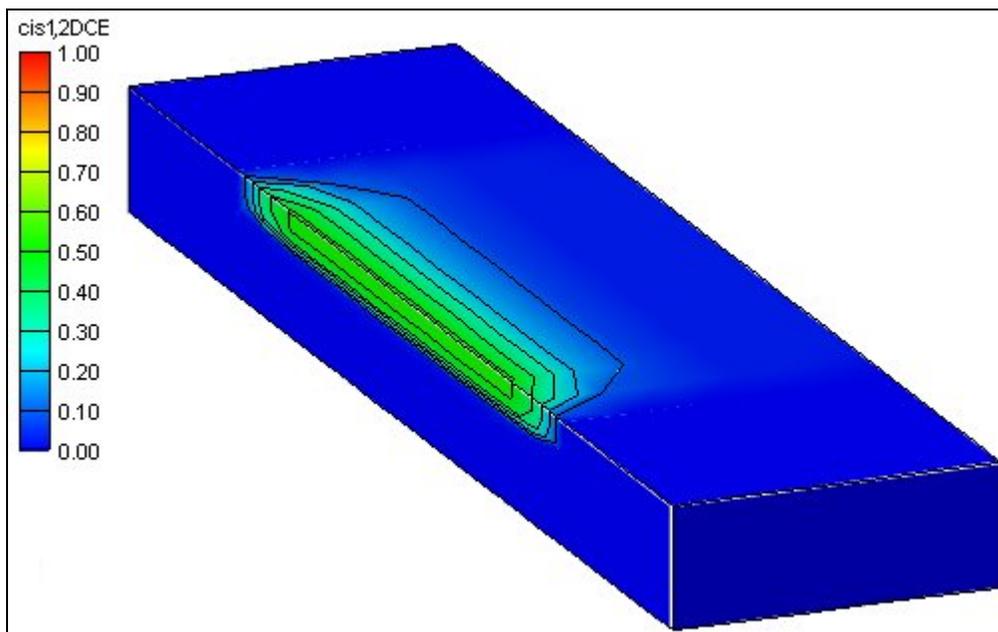


Figure 22: Threshold analysis result for cis 1,2-DCE; MCL = 0.07 mg/L. Grid length is 600 ft.

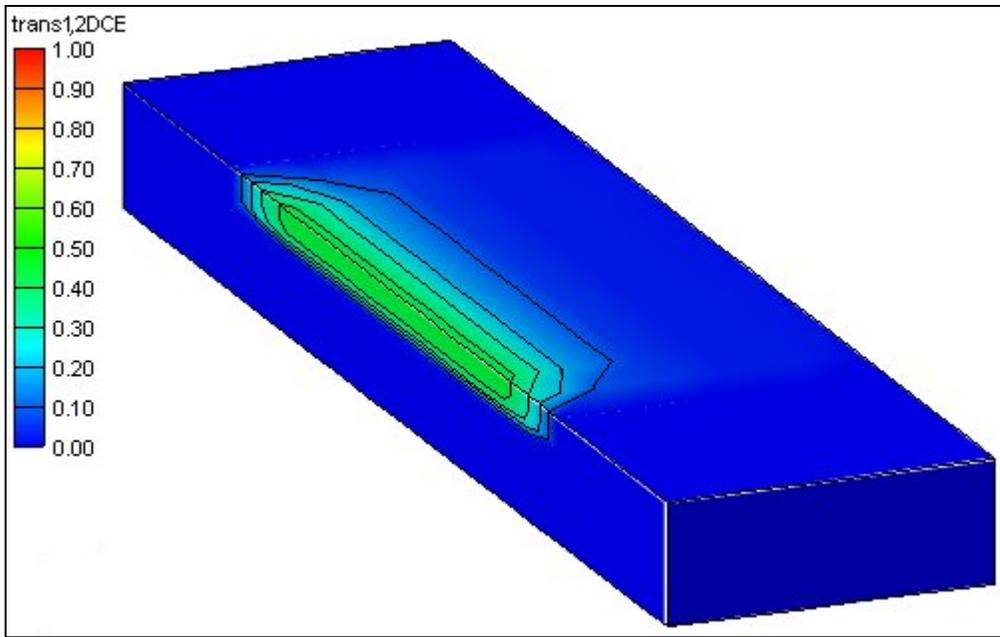


Figure 23: Threshold analysis result for trans 1,2-DCE; MCL = 0.1 mg/L. Grid length is 600 ft.

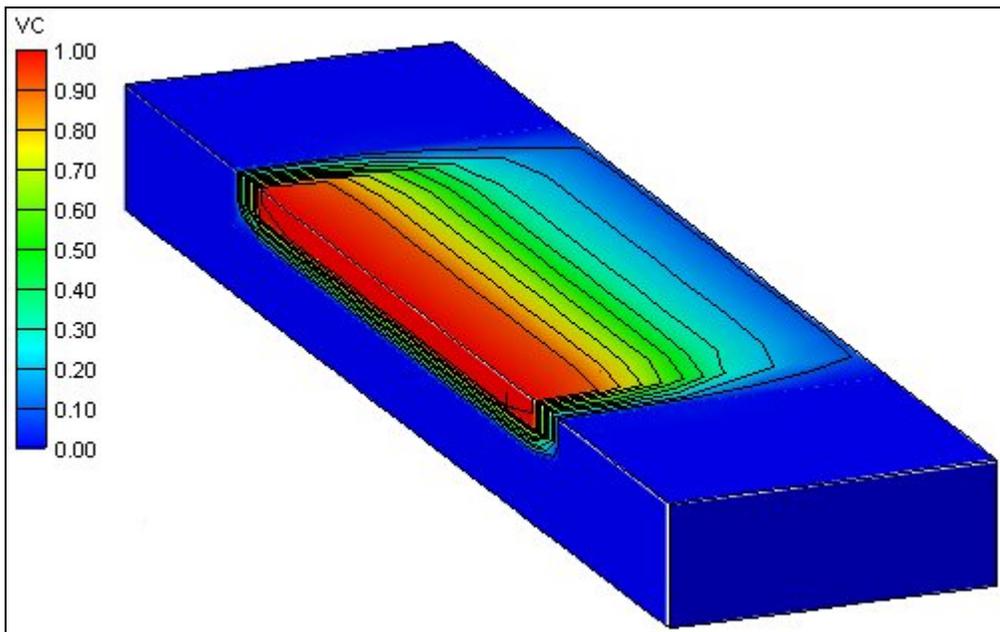


Figure 24: Threshold analysis result for VC; MCL = 0.002 mg/L. Grid length is 600 ft.

Obviously the most important threat from this analysis is the vinyl chloride concentration. All of the random simulations showed VC concentration values above the

MCL near the source and a good portion of the simulations show high concentrations continuing out 400 or 500 feet from the source. Because the VC plume will not reach equilibrium for a few hundred years, this plume will likely get worse before natural attenuation is able to reduce the concentrations. It will grow longer and its concentration will increase as other chlorinated compounds reduce to VC. Clearly, this contaminant is the most dangerous one in the plume and the one that should be most heavily considered in the design of remediation efforts.

Most of the other plumes are only likely to be a problem in the first 100 or 200 feet from the source. After this point, the percentages drop lower and the chance that the contamination exceeds EPA limits is small. Analysis of DCE is difficult due to the three separate isomers, but it seems that DCE may not be a big problem. The percentages for each of the isomers is low even close to the source, and the model output is for a combination of all three isomers, so these percentages would be even lower if the model accounted for the three separate types of DCE.

The stochastic simulation shows where drinking water could be contaminated. Any drinking water wells in the high probability regions will need to be tested for contamination. The advantage of this simulation over the previous forward simulations is that it does not output only a single result. Because it accepts the possibility of error in the parameter values, it prevents the information from being treated as absolute truth and allows policy-makers to make better decisions about the contamination at the site.

Chapter 8 – Conclusions

The main objectives of this project were accomplished successfully. The interface to GMS makes data input and results visualization much easier. The extra capabilities have made modeling easier, less time consuming, and therefore, less expensive. They have also extended the application of the model by providing more options for analysis.

Parameter Estimation

The parameter estimation engine successfully optimizes the parameters to yield an accurate output, closely matching the observed data when observation wells are well placed in time and space and when the model assumptions are accurate for the site. The success of the PORT library is increased when the number of observed concentration values (not necessarily observation points) exceeds the number of parameters being solved for.

The main difficulty in parameter estimation in ART3D is that the basic model assumptions are not always accurate for a given site. The requirement that all retardation values be the same and the inability to model any type of heterogeneity are its main drawbacks and hinder the accuracy of the optimized parameters.

It is important that the observations be placed in the main section of the plume to obtain the best optimized results. The Brooklawn site case study involved observation

points that were near the edges of the plumes. This made optimization very difficult and highly objective. A better way to handle a project would be to initially run a simple forward run with approximate parameters to determine the likely location of each plume and then place the boreholes or wells in the main sections of those plumes where the data will be above the detection limit and so the model can be calibrated to the more significant sections of the plume.

Stochastic Capabilities

The stochastic simulation option, when coupled with a threshold analysis is a useful addition to the ART3D code since it allows a statistical analysis of a number of equally likely results. The difficulty and expense in accurately defining a set value to any of the parameters makes a strong case for the definition of a range of values and the use of stochastic methods to show the parameter uncertainty in the results.

The literature review of first order decay constants for chlorinated ethenes, while helpful in the other two modes of ART3D, is especially valuable in the definition of distribution data for a stochastic run. It helps obtain practical results without the expense of carefully analyzing the aquifer and specie characteristics.

In the Brooklawn site case study, the threshold analysis helped determine which species were most likely to exceed EPA requirements for drinking water and where this excess would most likely occur.

References

- Action Hydrology Software. <http://www.hydrology-software.com/>, 1999.
- Agency for Toxic Substances and Disease Registry. *ToxFAQs for 1,1,2,2-Tetrachloroethane*. <http://atsdr1.atsdr.cdc.gov/tfacts93.html>, 1997.
- Agency for Toxic Substances and Disease Registry. *ToxFAQs for Chloroethane*. <http://atsdr1.atsdr.cdc.gov/tfacts105.html>, 1999.
- Anderson, James E. and Perry L. McCarty. Model for Treatment of Trichloroethylene by Methanotrophic Biofilms, *Journal of Environmental Engineering*, 120(2), 379-400, 1994.
- Carter, Sean R. and William J. Jewell. Biotransformation of Tetrachloroethylene by anaerobic attached-films at low temperatures, *Water Resources*, 27(4), 607-615, 1993.
- Clement, Prabhakar, et al. Modeling Natural Attenuation of Chlorinated Solvent Plumes at the Dover Air Force Base Area-6 Site, in *Natural Attenuation of Chlorinated Solvents, Petroleum, Hydrocarbons and other Organic Compounds: The Fifth International In Situ & On-Site Bioremediation Symposium, San Diego, California, April 1999*, pp 29-34, Battelle Press, Columbus, 1999.
- Clement, T. Prabhakar. Generalized solution to Multispecies Transport Equations Coupled with a First-Order Reaction Network, in *Water Resources Research*, 37(1), 157-163, January 2001.
- Clement, T. Prabhakar, Michael J. Truex, and Peter Lee. A Case Study for Demonstrating the Application of U.S. EPA's Monitored Natural Attenuation Screening Protocol at a Hazardous Waste Site, submitted to *Journal of Contaminant Hydrology*, February 2002.
- Doherty, John. PEST: Model Independent Parameter Estimation. Watermark Numerical Computing, 2000.
- Domenico, P. A., An Analytical Model for Multidimensional transport of a Decaying Contaminant Species. *J. Hydrol.*, 91, 49-58, 1987.
- Fathepure, Babu Z., and James M. Tiedje. Reductive Dechlorination of Tetrachloroethylene by a Chlorobenzoate-Enriched Biofilm Reactor. *Environ. Sci. Technol.*, 28(4), 746-752, 1994.
- Fitts Geosolutions. *Solutrans*. <http://www.fittsgeosolutions.com/solutran.htm>.

- Gay, David M., *Computing Science Technical Report No. 153: Usage Summary for Selected Optimization Routines*. <http://netlib.bell-labs.com/cm/cs/ctr/153.pdf>. October 1990.
- Gerritse, J., et al. Complete degradation of tetrachloroethane by combining anaerobic dechlorinating and aerobic methanotrophic enrichment cultures, *Appl. Microbiol. Biotechnol.*, 43, 920-928, 1995.
- Hardy, Leslie, Ernesto Moeri and Maria Cristina Salvador. Rapid Intrinsic Degradation of Chlorinated Solvents at a Manufacturing Site in Brazil, in *Natural Attenuation of Chlorinated Solvents, Petroleum, Hydrocarbons and other Organic Compounds: The Fifth International In Situ & On-Site Bioremediation Symposium, San Diego, California, April 1999*, pp 19-28, Battelle Press, Columbus, 1999.
- Hydroscience, Inc. *Kyspill*. <http://home.earthlink.net/~hydroscience/kyspill.html>.
- Lowry, Gregory V. and Martin Reinhard. Pd-Catalyzed TCE Dechlorination in Groundwater: Solute Effects, Biological Control, and Oxidative Catalyst Regeneration, *Environ. Sci. Technol.*, 34(15), 3217-3223, 2000
- Rügge, Kirsten, et al. An anaerobic field injection experiment in a landfill leachate plume, Grindsted, Denmark. 1. Experimental setup, tracer movement and fate of aromatic and chlorinated compounds, *Water Resources Research*, 35(4), 1231-1246, 1999.
- Schaerlaekens, J., et al. Numerical simulation of transport and sequential biodegradation of chlorinated aliphatic hydrocarbons using CHAIN_2D, *Hydrological Processes*, 13, 2847-2859, 1999.
- The Scientific Software Group. *AT123D Description*. http://www.scisoftware.com/products/at123d_details/at123d_details.html, 1998.
- The Scientific Software Group. *HSSM Description*. http://www.scisoftware.com/products/hssm_model_details/hssm_model_details.html, 1998.
- The Scientific Software Group. *PESTAN Description*. http://www.scisoftware.com/products/pestan_details/pestan_details.html, 1998.
- Software Spotlight, *Ground Water*, 37(6), 808-809.
- Sun, Anthony. *Anthony's VBA Page*. <http://www.geocities.com/WallStreet/9245/vba4.htm>, 1997.
- United States Environmental Protection Agency. *BIOCHLOR*. <http://www.epa.gov/ada/csmos/models/biochlor.html>, 2002.
- United States Environmental Protection Agency. *BIOSCREEN*. <http://www.epa.gov/ada/csmos/models/bioscrn.html>, 2002.
- United States Environmental Protection Agency. *Current Drinking Water Standards*. <http://www.epa.gov/safewater/mcl.html>, 2002.

United States Geological Survey. *Summary of ANALGWST*. http://water.usgs.gov/cgi-bin/man_wrdapp?analgwst, 2002.

United States Salinity Laboratory. *N3DADE*.
<http://www.ussl.ars.usda.gov/models/n3dade.HTM>, 1997.

Waterloo Hydrogeologic. *RBCA Tier 2 Analyzer*.
<http://www.flowpath.com/software/rbcaanalyzer/rbcaanalyzer.html>, 2000.

Appendix

Appendix A – ART3D Input File Format

The ART3D input file should have the extension *.art and should be arranged in card format according to the example in Figure 25. This example is for a normal forward run. With a parameter estimation or stochastic simulation, some minor changes are made to the file. These changes and all of the cards are explained in detail in the table below the sample file.

```
ART3D
GRIDFILE "bestanswer.3dg"
OBSCOV "obs"
SOURCE 1000 45
NUMSPEC 8
SPECNAMES "TeCA" "TCA" "DCA" "CA" "PCE" "TCE" "DCE" "VC"
YIELD
0 0 0 0 0 0 0 0
0.278 0 0 0 0 0 0 0
0 0.148 0 0 0 0 0 0
0 0 0.456 0 0 0 0 0
0 0 0 0 0 0 0 0
0.016 0 0 0 0.792 0 0 0
0.364 0 0 0 0 0.738 0 0
0 0.375 0 0 0 0 0.645 0
TIME 9125 1
UNITS "ft" "d" "mg" "lb" "mg/l"
REFTIME d 0 1900 1 1 0 0 0
SIMTYPE 0
PARAM retard 1.75
PARAM velocity 0.072
PARAM alphax 11.18
PARAM alphayratio 0.1
PARAM alphazratio 0.01
PARAM kTeCA 0.0029
PARAM kTCA 0.0032
PARAM kDCA 0.01
PARAM kCA 0.0048
PARAM kPCE 0.0036
PARAM kTCE 0.0081
PARAM kDCE 0.05
```

Figure 25: ART3D sample input text file

```

PARAM kVC 0.0006
PARAM cTeCA 90.17
PARAM cTCA 100
PARAM cDCA 50
PARAM cCA 0
PARAM cPCE 11.46
PARAM cTCE 0.63
PARAM cDCE 0.002
PARAM cVC 25
NOBS 4
WEIGHT
6250000.0 1000000.0 20408.2 62500.0 250000.0 250000.0 12345.7 1.25
6250000.0 1000000.0 20408.2 62500.0 250000.0 250000.0 12345.7 1.25
6250000.0 1000000.0 20408.2 62500.0 250000.0 250000.0 12345.7 1.25
6250000.0 1000000.0 20408.2 62500.0 250000.0 250000.0 12345.7 1.25
OBS 4 point_4 500.0 1000.0 45.0 1
9125.0 0.007 0.02 0.3 -999 -999 -999 0.02 3.5
OBS 3 point_3 400.0 1000.0 45.0 1
9125.0 -999 -999 0.1 -999 -999 -999 0.5 12.0
OBS 2 point_2 300.0 1000.0 45.0 1
9125.0 0.01 0.02 -999 -999 0.02 -999 -999 -999
OBS 1 point_1 150.0 1000.0 45.0 1
9125.0 -999 -999 0.01 -999 0.05 0.04 0.013 37.0

```

Figure 25: ART3D sample input text file (continued).

Table 25: ART3D input file card format.

Card	Arguments	Meaning
ART3D	None	Defines the file type
GRIDFILE	“name.3dg” – name of 3D grid file	Tells ART3D which file to look in to find the size and dimensions of the grid and the number of cells. This must be a mesh-centered grid.
OBSCOV	“name” - name of observation coverage holding the observation points	Used only by GMS – tells GMS where to look for the observation points used in the ART3D simulation.
SOURCE	ysourcedimension, z sourcedimension	Defines the size of the source. The source will always be at the left (x = 0) edge of the grid, centered in the y direction, and at the top of the grid in the z direction.
NUMSPEC	number	Tells ART3D how many species to expect.
SPECNAMES	“name1”, “name2”, . . . “namen”	Lists the species names in the order they will be considered in the yield matrix, weight matrix, parameters list and OBS cards. The number of names listed must be equal to the number defined in the NUMSPEC card.

Table 25: ART3D input file card format. (continued)

Card	Arguments	Meaning
YIELD	The yield matrix begins on the following line with one row and one column for each specie.	The columns and rows are ordered as the species were ordered in the SPECNAMES card. See the Equations section of Chapter 2 for a description of the yield matrix. The effective yield value should be placed in the column associated with the product and the row associated with the reactant for each possible reaction. If the species names are ordered with parent species always coming before their daughter products, there should be no values above the diagonal. There should never be any values on the diagonal (nothing reacts to form itself).
TIME	Totaltime numtimesteps	The first argument tells ART3D the total simulation time; the second indicates the number of time steps at which solutions should be reported. The actual time step lengths are calculated by dividing the first by the second; all time steps are equal in length.
UNITS	“length units” “time units” “mass units” “force units” “concentration units”	This card is read only by GMS. The only important units for an ART3D simulation are the length, time and concentration units.
REFTIME	time units (Boolean for the use of reftime) refyear refmonth refday refhour refminute refsecond	This card is read only by GMS. Reference times cannot be used in ART3D. If the user defines a reference time and chooses reference from the combo box in the Output Control Dialog, results can be output with a date and time instead of a length of time from the beginning of the simulation.
SIMTYPE	0 for forward run, 1 for parameter estimation run, 2 for stochastic run	Defines run type.
STOCH	Number of stochastic runs	This card only appears in a stochastic simulation and defines the number of independent solutions which will be calculated using random parameter values
NUMITER	Number of iterations before PORT library stops	This card only appears in a parameter estimation simulation.

Table 25: ART3D input file card format. (continued)

Card	Arguments	Meaning
NUMFUNCALLS	Number of calls to the objective function before the PORT library stops	This card only appears in a parameter estimation simulation.
EPSA	Absolute function convergence tolerance for PORT library	This card only appears in a parameter estimation simulation. If the objective function returns a value smaller than this tolerance, the optimizer will stop.
EPSR	Relative function convergence tolerance for PORT library	This card only appears in a parameter estimation simulation. If the difference between the current objective function value and the optimized value are less than the product of this tolerance and the current objective function value, the optimizer will stop.
PARAM retard PARAM velocity PARAM alphax PARAM alphayratio PARAM alphazratio PARAM k_... (decay constant) PARAM Co_... (source concentration)	Value	In a forward run, only the parameter value is listed.
	Value Min Max Solve(0/1)	In a parameter estimation run, the start value, and the lower and upper bounds are listed, followed by a one if the value is to be optimized or zero if not. If the final value is a zero, the lower and upper bounds values are not important and can be any value. There must be four values in the list regardless.
(the '...' following 'k_' and 'Co_' represent the name of the specie. A 'k_' line and a 'Co_' line must be provided for every defined specie.)	Value Min Max Vary(0/1) Distribution type(0/1) Standard Deviation	In a stochastic simulation, the mean value, and the lower and upper bounds are listed, followed by a one if the parameter is to be varied, or a zero if not. This is followed by a zero for a normal distribution or a one for a linear distribution and finally, the standard deviation. If the parameter is not to be varied, only the initial mean value is important; if the distribution type is linear, the standard deviation is not important. There must be six values in the list regardless.
NOBS	Number of observation points	This represents the number of observation points in the simulation. If this value is zero, the rest of the lines are omitted.

Table 25: ART3D input file card format. (continued)

Card	Arguments	Meaning
WEIGHT	The weighting matrix follows on the next lines. The matrix must be included regardless of the simulation type, unless there are no observation points	The weighting matrix has a row for each observation point (in the order listed below in the OBS card) and a column for each specie (in the order listed in the SPECNAMES card). In a forward run or stochastic simulation, these weighting values are applied to the residuals and presented in the output file (*.out). In a parameter estimation simulation, the weighting values are used in the calculation of the objective function. When running ART3D from inside GMS, these values are automatically calculated from the standard deviation values entered in the Observation Point dialog. However, they can be changed by hand if desired by editing the *.art file.
OBS	id name xvalue yvalue zvalue numtimes	The id and the name should be unique to the point. The xvalue, yvalue and zvalue are relative values measured from the bottom left hand corner of the x-y grid and measured up from the bottom layer in the z direction. The numtimes value represents the number of times that the observation data are sampled. Any number of time points can be included and they do not have to coincide with the grid time steps defined above. The number of time points must be equal to the number of lines of data included under each OBS card.
	time value_species1 value_species2 . . .	The number of measurement lines included after each OBS line should be equivalent to the numtimes value defined in the OBS line. The values should be ordered exactly as in the SPECNAMES line above. If the value for a particular specie is not known or was not measured, enter “-999”. If the exact concentration is required at a certain point and time, but measurements were not made, enter -999 for all the values and the output file (*.out) will show these values. As explained above, the times listed here do not have to correspond to the grid time steps defined in the TIME card above.

All parameters on a given line should be separated by spaces only (not commas or semi-colons). In general, the order of the cards is not important, but the user is urged to keep the order shown above. In some cases, the order becomes very important. For example, the NUMSPEC card must be read before the SPECNAMES, YIELD, WEIGHT or OBS card so that ART3D knows how many values to expect. When running ART3D inside GMS, this file is automatically written out. However, an understanding of the input file format is useful in case changes are to be made or if ART3D is to be run alone.

Appendix B – GMS Interface

Running ART3D from inside the GMS environment makes data input much easier and allows a better visualization of the results. The sections below present the GMS interface and explain the procedure for defining and running an ART3D simulation from GMS. Users unfamiliar with GMS are urged to familiarize themselves with the *Map* and *3D Grid* modules first.

Step 1: Building the Grid

The first step in building an ART3D simulation is to create a grid in GMS by entering the *Grid* module and selecting *Grid/Create Grid*. The grid can be any length and have any number of cells in the x, and z directions, however, there must be an even number of cells in the y direction and all cells must be the same size, so the *Bias* must be set at 1.0. In the *Orientation/Type* drop-down menu, the *FACT/ART3D* option must be selected to create a mesh-centered grid. Because ART3D is a fast, analytic model, the size and density of the grid do not greatly affect the time required to compute a forward run or a parameter estimation simulation, but if the grid is large or dense, a threshold analysis can take several minutes.

Step 2: Defining the Species

Once the grid has been created, the ART3D simulation should be initialized by choosing *ART3D/New Simulation*. This resets the parameters to default values and

allows access to the previously dimmed ART3D menu items. With the simulation initialized, the species can be defined and named by choosing *ART3D/Define Species*. This brings up the dialog shown in Figure 26.

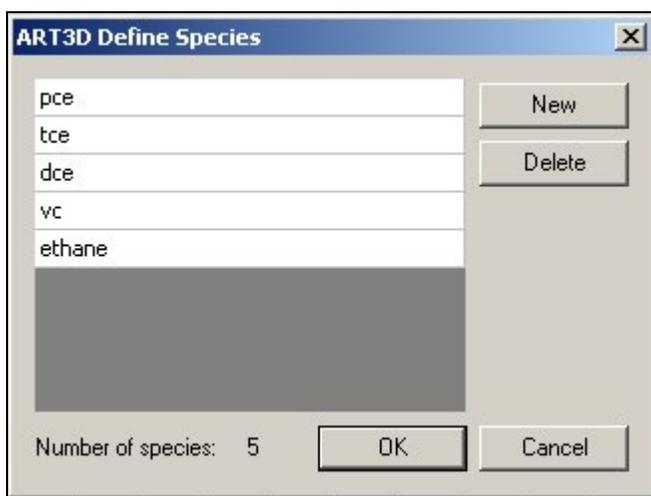


Figure 26: ART3D Define Species dialog

The *New* button should be selected once for each desired specie and its name should be typed into the spreadsheet. A specie can be deleted by highlighting its name in the spreadsheet and selecting the *Delete* button. The text at the bottom of the dialog keeps a running total of the number of species. It is suggested that the species be listed in order with parent species always above all of their daughter products. In the case of reversible reactions, however, this is not possible.

Step 3: Defining Observation Data

Observation data is not required unless a parameter estimation simulation is to be run. However, even if there is no observation data available, points can be defined in areas where a plot of computed concentration values is desired.

Observation data must be entered in the *Map* module using an observation coverage. In the *Map* module, a new coverage can be created by right clicking on the *Map Coverages* folder in the *Data Tree*, and choosing *New Coverage*. The coverage type can then be set by right clicking on the new coverage and choosing *Properties*. The combo box titled *Type* should be set at *Observation*. The name of the coverage can also be defined by right clicking on the new coverage in the *Data Tree* and choosing *Rename*.

Once the observation coverage has been defined, the points can be added using the *Create Point Tool* in the *Map* module. Double clicking on one of these points, or selecting one and choosing *Feature Objects/Attributes* will bring up the *Observation Coverage Attributes* dialog shown in Figure 27.

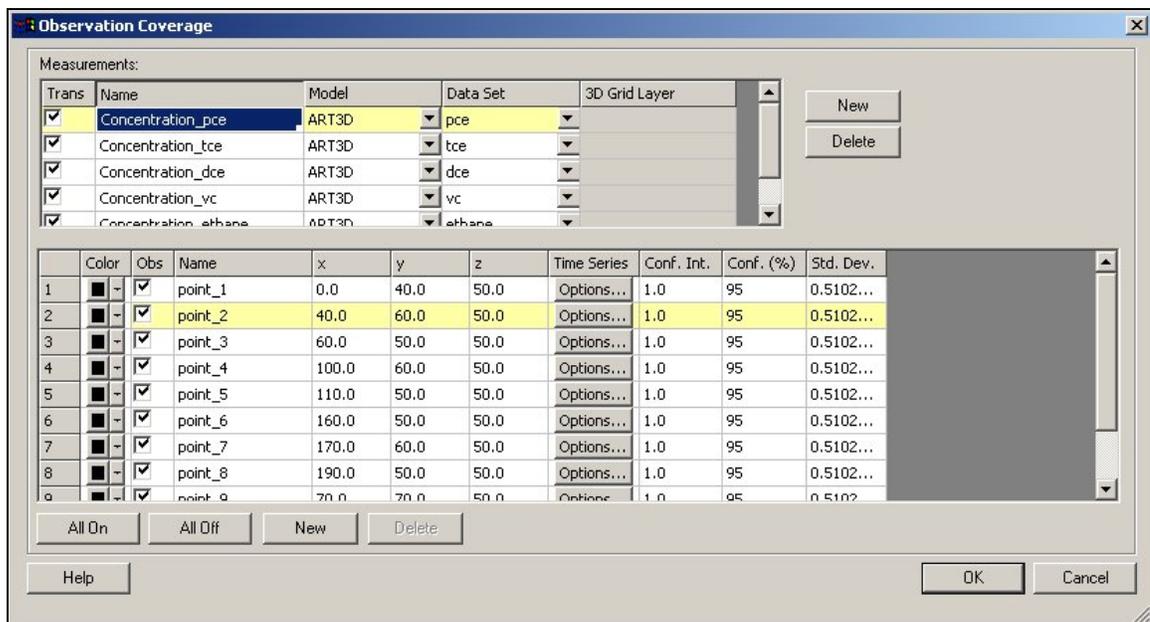


Figure 27: Observation Coverage Attributes dialog

The *Observation Coverage Attributes* dialog is also where the measurements must be defined. If the species have already been defined in the *3D Grid* module, then the measurements can easily be linked to these species in the top spreadsheet of the dialog. The *New* button should be selected once for each specie name and the model (ART3D) and the appropriate data sets should be selected for each measurement. This spreadsheet also has a column for naming the measurement and for choosing to make the measurements transient. All of the measurements in Figure 27 have been selected to have transient data.

The bottom spreadsheet can be used to name the points, define their color and set them as observation points. Here, the x, y, and z coordinates can be defined or adjusted. The user also should define the confidence interval and percentage or the standard deviation. The standard deviation will be used in determining the weight value for each solution value as explained in Chapter 3. Setting the confidence interval and percentage will automatically adjust the standard deviation.

The measurements that are selected to have transient data must have a time series associated with them. The time series can be added by selecting the *Options* button in the *Time Series* column. The dialog which opens up is shown below in Figure 28. The *Import* button can be used to import time step data from another file, or the data can be entered by hand to the chart on the left. The plot field will show the time series visually as it is created.

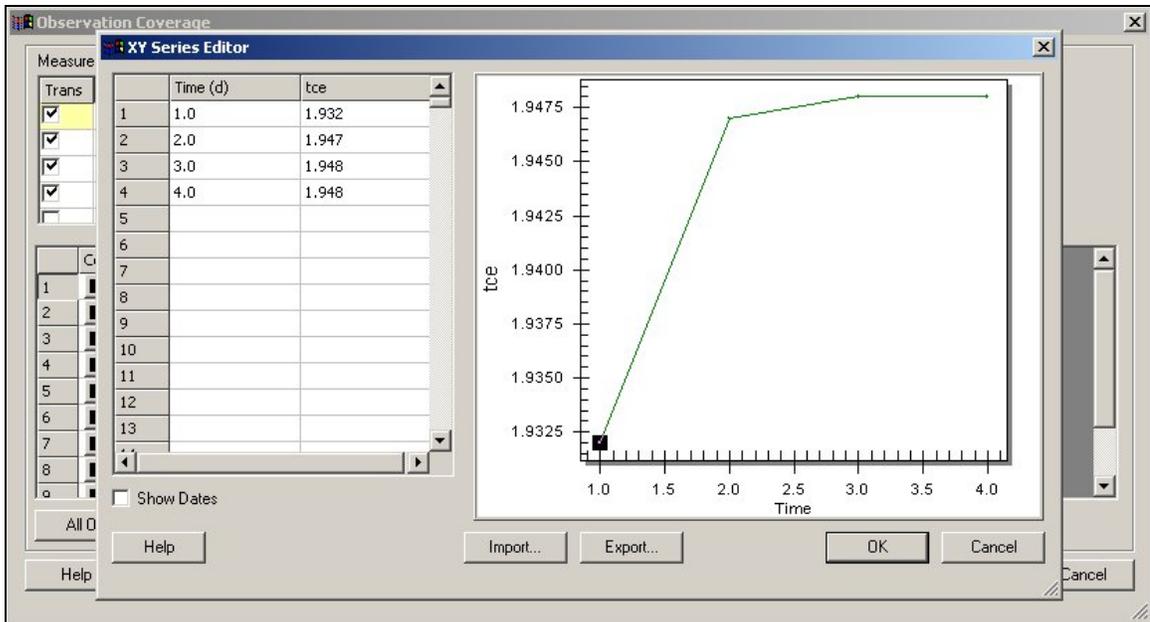


Figure 28: Adding a timeseries to an observation point

If the measurement has not been marked as transient, the *Time Series* button shown in Figure 27 will be replaced by a spreadsheet cell where a single value can be entered. Because ART3D is always transient, this value will be applied to the final grid time step.

Step 4: Run Options and Output Control

Once the observation data has been defined, the user must return to the *3D Grid* module to enter the rest of the data. The first set of data to be entered is found in the dialog which appears on selection of *ART3D/Run Options*. This dialog is shown in Figure 29.

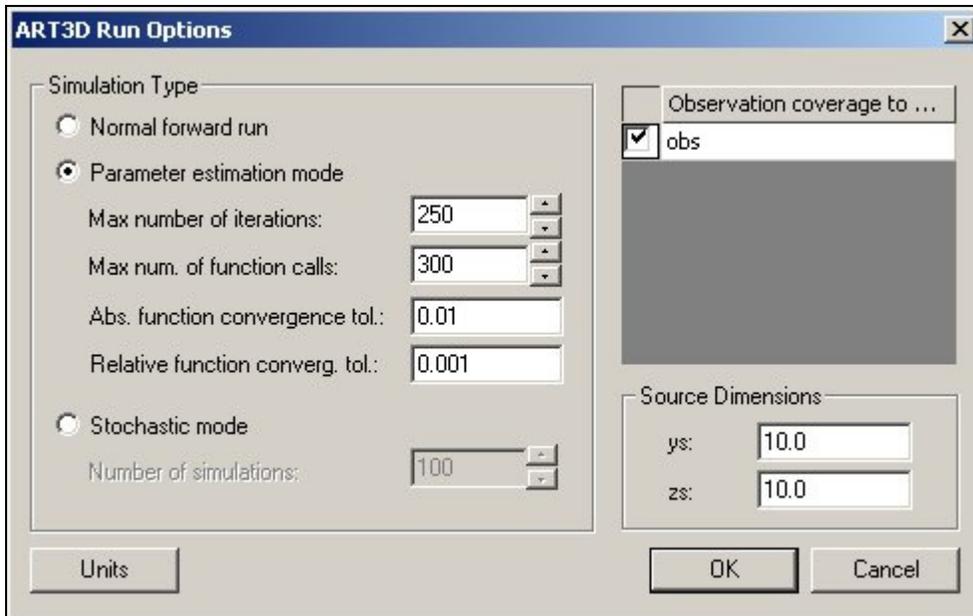


Figure 29: ART3D Run Options dialog

On the left half of the dialog, the user should choose the type of simulation to run. If the parameter estimation option is selected, the edit fields for defining four stopping conditions for the PORT library become available. These four values include the maximum number of iterations, the maximum number of function calls, the absolute function convergence tolerance and the relative function convergence tolerance. These values are described in Chapter 3.

If the stochastic option is selected, the stopping conditions edit fields are dimmed and the edit field for choosing the number of stochastic simulations becomes available.

The right half of the dialog includes a list of all observation coverages in memory. The one containing the observation data for the ART3D simulation should be selected. Only one coverage can be selected at a time, so it is important that all ART3D

observation data be placed in the same coverage. If no coverage is checked, no observation data will be calculated by ART3D.

In the bottom right corner of the dialog, the user inputs the source dimensions to be used. The source is assumed to be at the left ($x = 0$) side of the grid, centered in the y direction and at the top of the grid in the z direction.

The units for the simulation can be defined by using the dialog which appears on selection of the *Units* button in Figure 29

The *Output Control* dialog, shown in Figure 30, allows the user to enter data associated with the output time steps and is accessed by choosing the *Output Control* command from the *ART3D* menu.

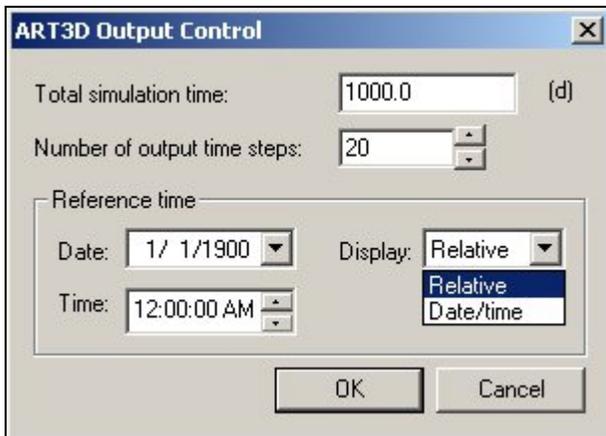


Figure 30: ART3D Output Control dialog

In the example in Figure 30, the user has selected a total simulation time of 1000 days with 20 time steps. This means that ART3D will calculate the grid solution every 50 days up to the total simulation time. The concentration values at the observation

points will also be calculated at each of these times and at all times at which observed data is provided by the user. The grid solution will not be calculated to correspond with the observed value time steps.

Also in the *Output Control* dialog, the user can select either relative time display or date/time display in the combo box titled *Display*. In the case of a relative time display, all times will be shown as the amount of time passed since the beginning of the simulation. If the *Date/Time* option is selected, the user must select a reference time, or the date and time at which the simulation began. Then all times will be displayed as dates and times (for example, “March 18, 2002 15:00:00”).

Step 5: Entering the Effective Yield Matrix

The yield values are entered in matrix form in the dialog which is displayed upon selection of *ART3D/Yield Coefficients*. This dialog is shown in Figure 31.

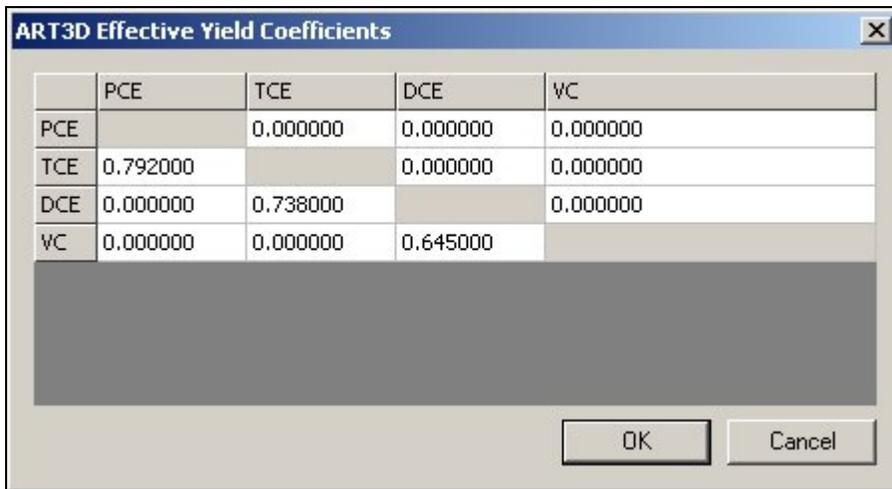


Figure 31: ART3D Effective Yield Matrix dialog

This matrix has one column and one row for each defined specie. Each effective yield coefficient should be placed in the cell corresponding to the column of the parent specie and the row of the daughter specie. If the species have been defined in order with the parent species always above their daughter products, all values will be below the diagonal in the yield matrix. The units of each value should be mass of product / mass of reactant. The diagonal values have been dimmed since no specie will react to form itself.

The effective yield coefficient is the product of the stoichiometric yield fraction and the ratio of the molecular weights of the two species. This is not the same yield matrix required by the original ART3D input file. When using this new input file format, the matrix will be automatically manipulated inside ART3D to create the matrix required in ART3D. This method saves the user some calculations.

Step 6: Entering the Aquifer and Specie Parameters

The rest of the parameters required for an ART3D simulation are entered in the *Parameters* dialog, accessed by choosing *ART3D/Parameters*. This dialog is slightly different depending on the type of simulation selected in the *Run Options* dialog. In the case of a forward run, the dialog looks as shown in Figure 32 and is fairly self-explanatory.

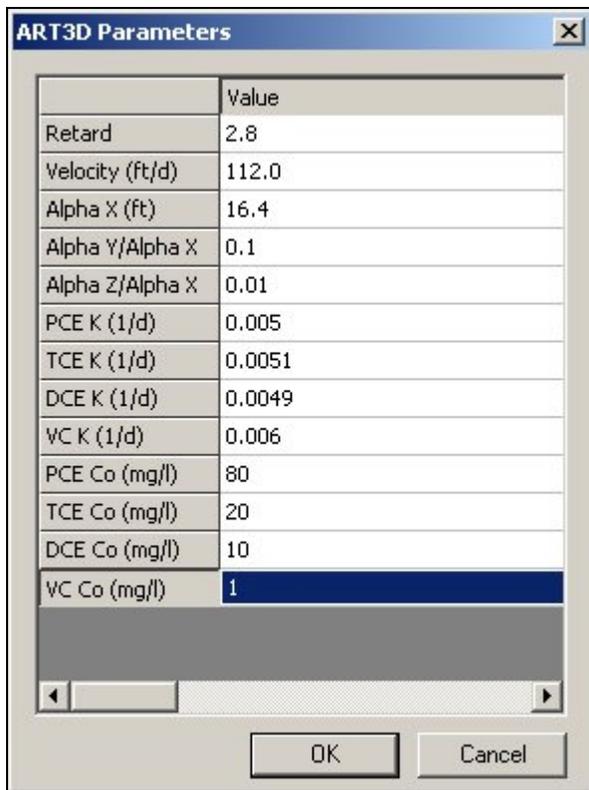


Figure 32: ART3D Parameters dialog for a forward run

If the parameter estimation mode has been selected, the *Parameters* dialog appears as in Figure 33.

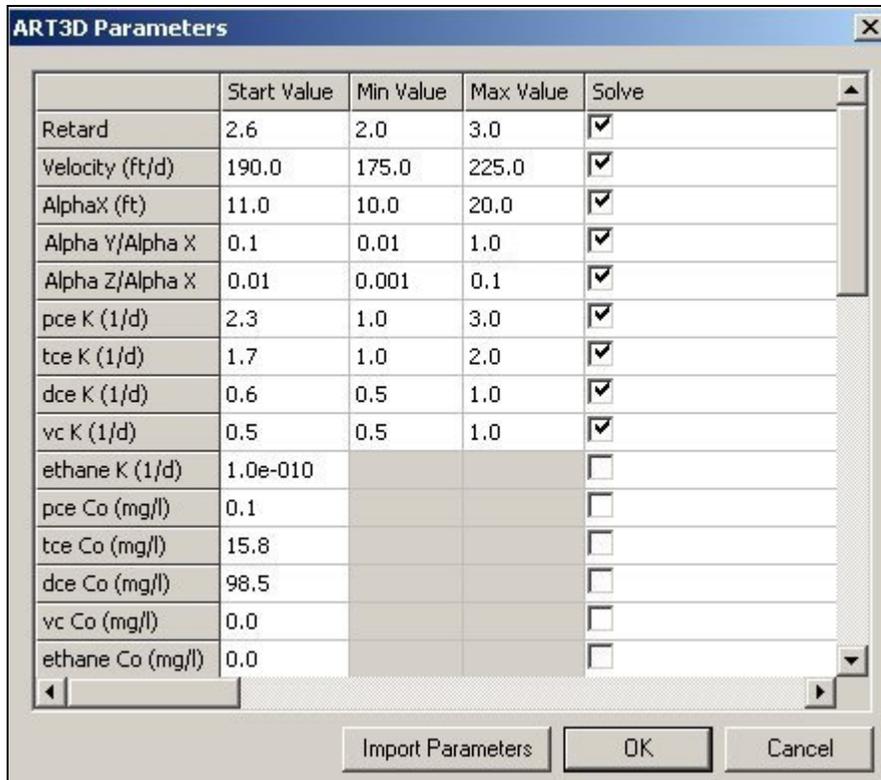


Figure 33: Parameters dialog for a parameter estimation run

In the above example, the user has selected the first ten parameters to be solved for by the optimization routine. Each of those ten parameters has been given a start value, which will be used in the first iteration, and a minimum and maximum value. The optimization routine will not allow the parameters to change beyond the bounds set in this dialog. The five parameters that are not to be optimized, only have a start value. This value is used in every iteration and does not change.

After the simulation has been run, the user can enter this dialog again and choose the *Import Parameters* button and GMS will fill in the start values with the optimized values from the PORT library.

In the case of a stochastic simulation, the *Parameters* dialog will appear as in Figure 34.

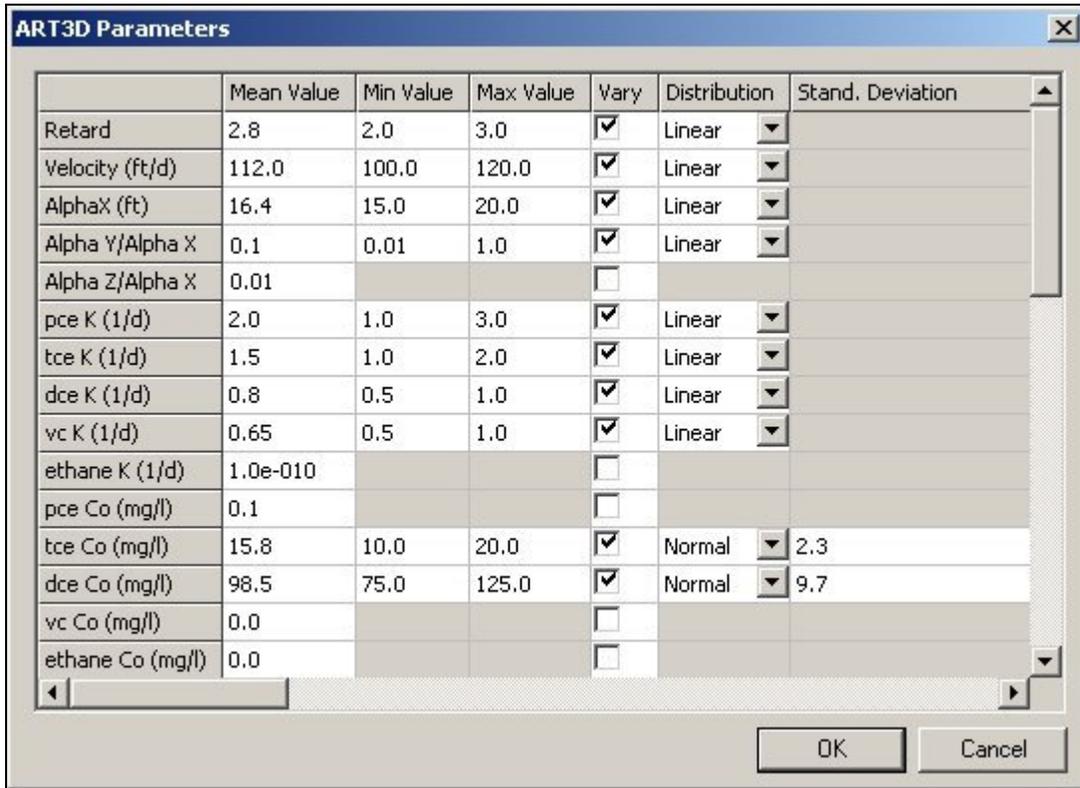


Figure 34: ART3D Parameters dialog for stochastic simulation

Here, the user has chosen ten parameters to be varied. Two will have normal distributions and the rest will have linear distributions. All values which have not been selected for variance will be set at the mean value in every simulation. For parameters with a normal distribution, a standard deviation has also been defined.

Step 7: Running the Model

When all of the data has been entered, the user should choose *File/Save* to save out the ART3D input file. Once this file has been saved, the selection of *ART3D/Run*

Simulation will call the ART3D code and pass it the input file. A dialog will come up in GMS to track the progress of the simulation. This dialog is slightly different depending on the type of simulation.

For a forward simulation, the dialog consists of a single window which lists output information from the model. If an error occurs, its description will be listed in this window. When the simulation has finished, the text, “Simulation completed successfully” will appear in the window and the *Abort* button will change to a *Close* button. This dialog is shown in Figure 35 after the completion of the simulation.



Figure 35: ART3D output window for a forward simulation

In a parameter estimation simulation, the output window shows a little more information. The dialog, shown in Figure 36, includes a plot of the error versus iteration. This plot updates after every iteration and helps visualize the success of the PORT library. In the middle of the dialog, there is a spreadsheet showing the values of each parameter and the error after each iteration. Like the plot window, it is updated after each

iteration, to show the progress of the optimization. A copy and paste command can easily transfer this data to a spreadsheet program where it can be plotted. The bottom message window displays any errors as well as the final error value, the total number of objective function evaluations and the reason for stopping the optimization routine. In the example shown in Figure 36, the routine stopped because it reached the absolute function convergence tolerance. If the user wishes to continue the optimization, this tolerance value will need to be reduced in the *Run Options* dialog.

After the PORT library stops, ART3D is run one more time, this time with the optimized values. Both grid and observation point solutions are calculated.

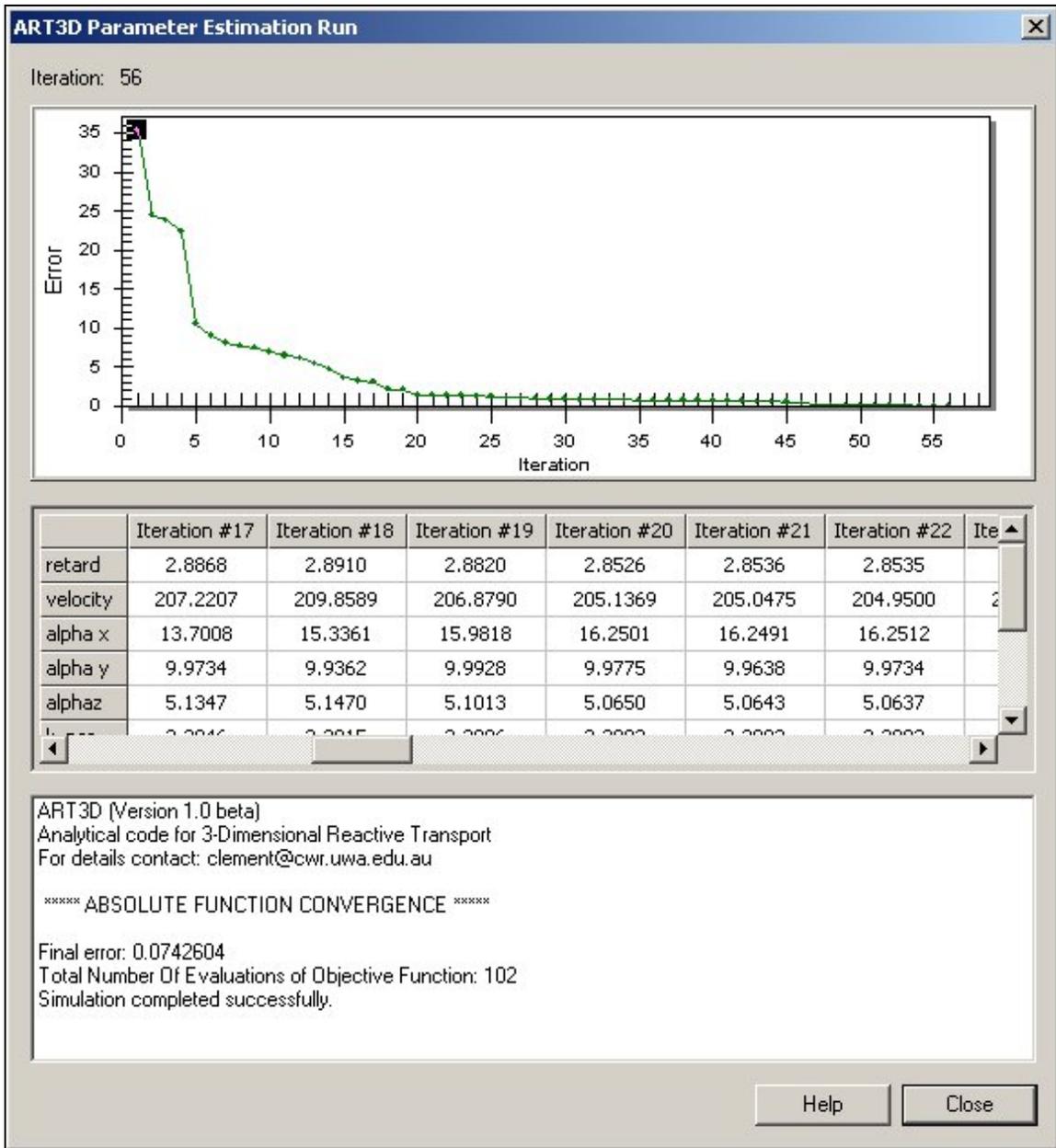


Figure 36: ART3D output window for a parameter estimation simulation

Finally, if a stochastic simulation is being run, the output window will appear as in Figure 37 with a spreadsheet and a text window. The spreadsheet lists the random values used in each simulation and is updated after every simulation. As before, the text

window displays error messages and tells the user when the simulations have finished.

The bar in the center of the dialog moves across the screen as the simulations progress.

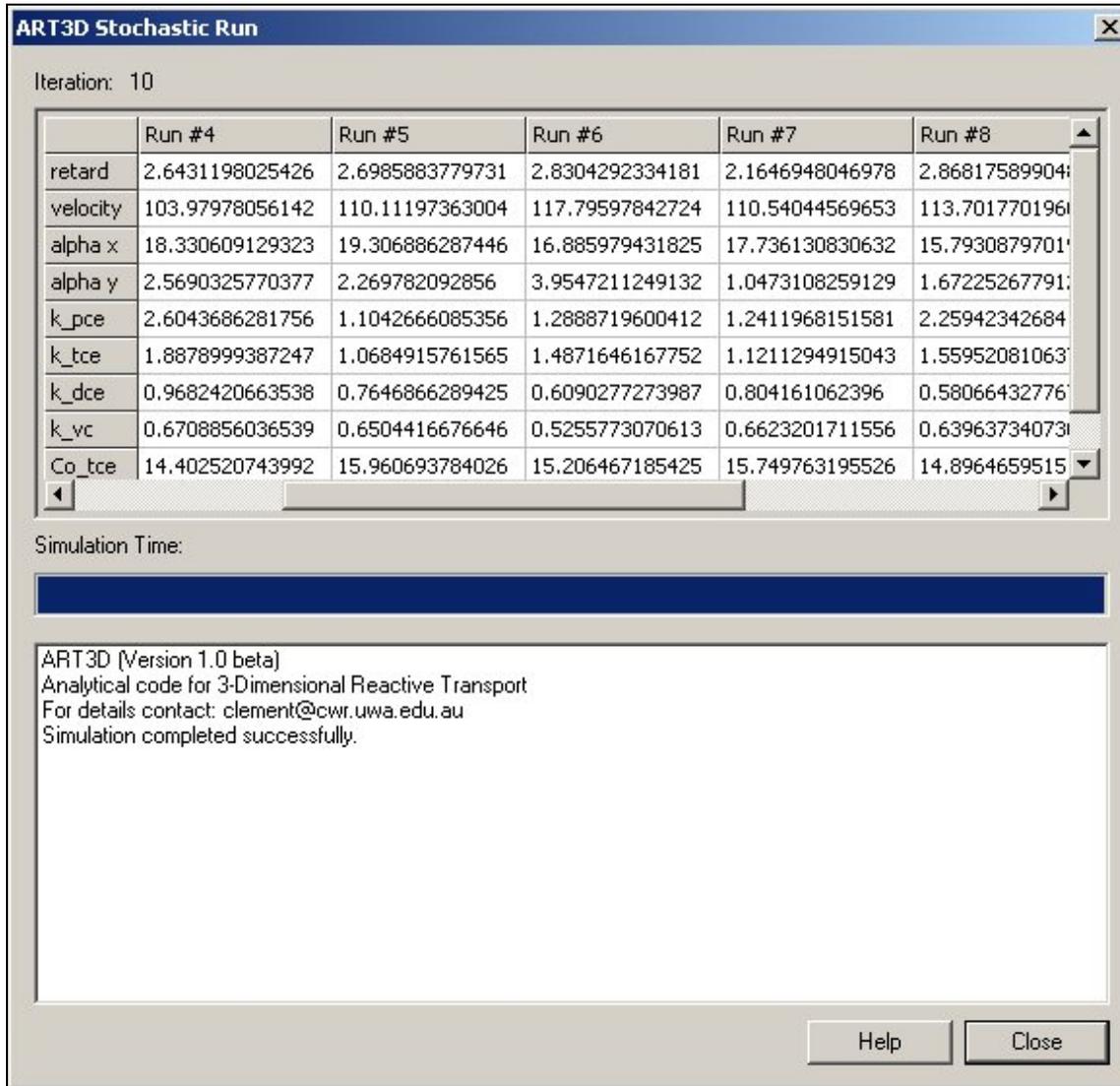


Figure 37: ART3D output window for a stochastic simulation

Step 8: Reading the Solution

After the model has finished and the output window has been closed, the user can view the solutions by choosing *ART3D/Read Solution*. In a forward run, a single

solution, consisting of several data sets will be read into memory and listed in the *Data Tree* at the right of the GMS window. After a parameter estimation run, the solution from the optimized parameter values is read in similarly. These solutions can be viewed by clicking on them in the data tree window. Display options such as contours and orientation can be changed as described in GMS documentation

After a stochastic simulation, all of the solutions will be read in and grouped in a single folder. Any single solution may be viewed individually and solutions can be moved to other folders or deleted as desired using the *Data Tree*. A risk analysis can be run on a folder containing any number of solutions. This analysis can be initiated by right clicking on the folder holding the solutions and selecting *Risk Analysis*. In the first window, *ART3D* should be chosen from the text window, and the *Probabilistic Threshold Analysis* should be selected below the window. In the next dialog, the user can select any of the species, a comparison type and a threshold value. A name can also be assigned to the analysis. When the *Finish* button is selected, GMS will go through all the solutions in the selected folder and find the percentage of these solutions that meet the criteria entered in the dialog. If more than one comparison has been specified, it will find the spatial distribution of the probability of satisfying all criteria. When the process is finished, a new data set will be available in the data tree. This data set will show the percent probability of the entered criteria being true in each grid cell. Again, the reader is directed to GMS documentation for information on post-processing.