# Using an Ensemble of Models to Design a Well Field Considering Regional Hydrologic Uncertainty

by

Stephen A. Hundt

———————————

A Thesis Submitted to the Faculty of the

## Department of Hydrology and Water Resources

In Partial Fulfillment of the Requirements
For the Degree of

## Master of Science
## With a Major in Hydrology

In the Graduate College

## The University of Arizona

2 0 1 4

## STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: _____ STEPHEN A. HUNDT _____

## APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

| | |
|---|---|
| Dr. Paul A. Ferré | May 5, 2014 |
| Professor of Hydrology | |

# Acknowledgements

# TABLE OF CONTENTS

Table of Contents—*Continued*

# List of Figures

# LIST OF TABLES

# ABSTRACT

Groundwater models are often developed as tools for environmental decision-making. However, sparse data availability can limit a models effectiveness by confounding efforts to select a single structural representation of a system or to find a unique and optimal set of model parameters. As a result, estimates of prediction uncertainty and the value that further data collection may provide can be important results of a modeling effort. The Discrimination/Inference to Reduce Expected Cost Technique (DIRECT) is a new method for developing an ensemble of models that collectively define prediction uncertainty in a manner that supports risk-based decision making and monitoring network design optimization. We apply aspects of DIRECT to a modeling investigation of an aquifer system in Central Utah where a major Coalbed Methane gas field is located and a new approach for stimulating gas production is being explored. In the first stage of this study we develop an ensemble of regional MODFLOW models and calculate their relative likelihood using a set of observation data. These regional results and likelihoods are then transferred to a regional MT3D residence time model and to a local advective transport model to provide further information for the gas stimulation design. A cost function is applied to the advective transport results to assess the relative expected costs of several proposed well field designs. The set of hydrologic results and associated likelihoods from the ensemble are combined into cost curves that allow for the selection of designs that minimize expected costs. These curves were found to be a useful tool for visualizing the ways that design decisions and hydrologic results interact to generate costs. Furthermore, these curves reveal different ways in which uncertainty can add to the cost of implementing a design. A final analysis explored the cost added by uncertainty with varying magnitudes of uncertainty by applying and manipulating synthetic likelihood distributions to the advective transport results. These results suggest, in a general way, the value that may be added by reducing uncertainty through data collection. Overall, the application of DIRECT was found to provide a rich set of information that is not available when ensemble methods and cost consideration are not included in a modeling study.

Chapter 1

# Introduction

Groundwater models have become important tools in the fields of planning, environmental management, and mining, where they are typically used to predict the response of a groundwater system to future applied stresses and environmental conditions. These conditions may be a combination of changes in the environment, such as changes in evapotranspiration or recharge, or human induced stresses, such as increased pumping. Ultimately, the predictions made by groundwater models are used to help inform the decision making process.

Despite many advances in hydrologic modeling and measurement methods, the ability of a model to provide useful results for decision making is limited by uncertainty in model predictions. Structural (or conceptual model) uncertainty, input data uncertainty, and parameter uncertainty are inherent in any model and transfer into uncertainty in the predictions made by the model. Structural model uncertainty arises from the inability to know which of the possible simplified representations of the system is most accurate. Input data uncertainty arises from the limited precision of physical quantities that are used as input to the model. Parameter uncertainty arises from the inability to identify the single set of parameter values that allow the model to best represent the behavior of the physical system. While the resulting prediction uncertainty is always present, an understanding of its sources, magnitude, and the impacts it may have on decision making can help the modeler communicate model results more meaningfully and assess strategies aimed at reducing uncertainty. At a minimum, groundwater modelers should make an effort to define their prediction uncertainties in ways that are most immediately useful for decision makers.

Several different approaches have been proposed for constructing and using models

under the presence of uncertainty. The Discrimination/Inference to Reduce Expected Cost Technique (DIRECT ) is a framework for developing an ensemble of models that collectively define prediction uncertainty in a manner that supports risk-based decision making and monitoring network design optimization. DIRECT may involve using both multiple conceptual models and multiple parameter realizations to make predictions. Risk-based decision making is applied by using each model variation to predict the costs of different management actions and weighting these costs by the likelihood of their associated models. This calculation results in an estimate of the expected cost of each action. By basing the comparison on expected costs (or benefits), competing actions can be assessed in a way that acknowledges the presence of uncertainty in the hydrologic models. DIRECT can also suggest how additional data collection can be designed to have the greatest likelihood of reducing the components of the expected costs that are due to discrepancies in model predictions. That is, decisions and monitoring network design can be made from a more informed position. This study will apply aspects of DIRECT to a real world groundwater modeling problem, and will include additional hypothetical exercises to explore the utility of DIRECT.

The major purpose of this study is to investigate hydrologic conditions related to optimizing coal-bed methane (CBM) production. To serve this end, a groundwater flow model is developed and explored for the Ferron Sandstone aquifer system in the Castle Valley of north central Utah. The Ferron Sandstone contains coal deposits that have been important and heavily developed sources of coal bed methane over the last two decades. This region has been the subject of several studies including a 2003 study by the USGS (Anna 2003) in which a MODFLOW groundwater flow model was developed. Annas 2003 model serves as the starting point for this study. However, the goals of this study differ from Annas and the conceptual model has been modified to suite these goals. Additional interests of this study are to explore regional residence time distributions, localized transport dynamics, and the sources and propagation of

uncertainties and economic costs. These interests are addressed by constructing local-scale transport models that are informed by the regional groundwater flow model.

This study is well suited to take advantage of many of the capabilities of DIRECT. First, parameter and conceptual model uncertainty are present and very little data are available for constraining the model. As a result, model predictions can be expected to suffer from large uncertainty. Second, the study is framed around choosing between competing strategies of injection and extraction. Therefore, the hydrologic models can give direct insight into the likely performance of different strategies with consideration of the hydrologic uncertainty. DIRECT can provide a number of benefits under this difficult decision-making environment. Its multi-model approach can incorporate identifiable sources of uncertainty, and then allow for risk-based decision making to be performed to select a strategy that minimizes expected costs. In addition, by nesting a local hydrologic model within a regional model, the propagation of uncertainties from the regional to the local scale can be considered in the choice of an optimal operations strategy.

Chapter 2

# Background and Site Description

## 2.1   Geography

The study area encompasses roughly 4500 square miles of the Castle Valley area between the Wasatch Plateau and the San Rafael Swale of Northern Utah. The lower elevations of the valley receive 5-9 inches of precipitation annually, and the Wasatch Plateau receives around 25-29 inches annually (PRISM Climate Group 2012). Several streams flow eastward from the Wastach Plateau, and intersect the outcrop of the Ferron Sandstone on the eastern edge of the study area. Records from stream flow gages (U.S. Geological Survey 2011) near the eastern edge of the study area indicate that flow is perennial those stream reaches. Coalbed methane gas extraction is the most significant economic activity in the area and is the reason for most of the development of the Ferron Sandstone aquifer, especially in the northern portion of the basin. The rough boundaries of the Helper and Drunkards Wash gas fields are outlined on figure 2.1. The town of Emery in the southern portion of the study area draws its municipal supply from the aquifer. In this area, the aquifer is shallower, nearer to the outcrop, and has better water quality than in the portions within the coalbed methane gas fields.

## 2.2   Geology

The Ferron Sandstone Member of the Mancos Shale is overlain by the Bluegate Shale and underlain by the Tununk Shale. It consists of sedimentary rocks that were deposited in advancing and retreating deltas located along the Cretaceous interior seaway (Rice 2003). It is comprised of series of clastic wedges of marine and shoreline
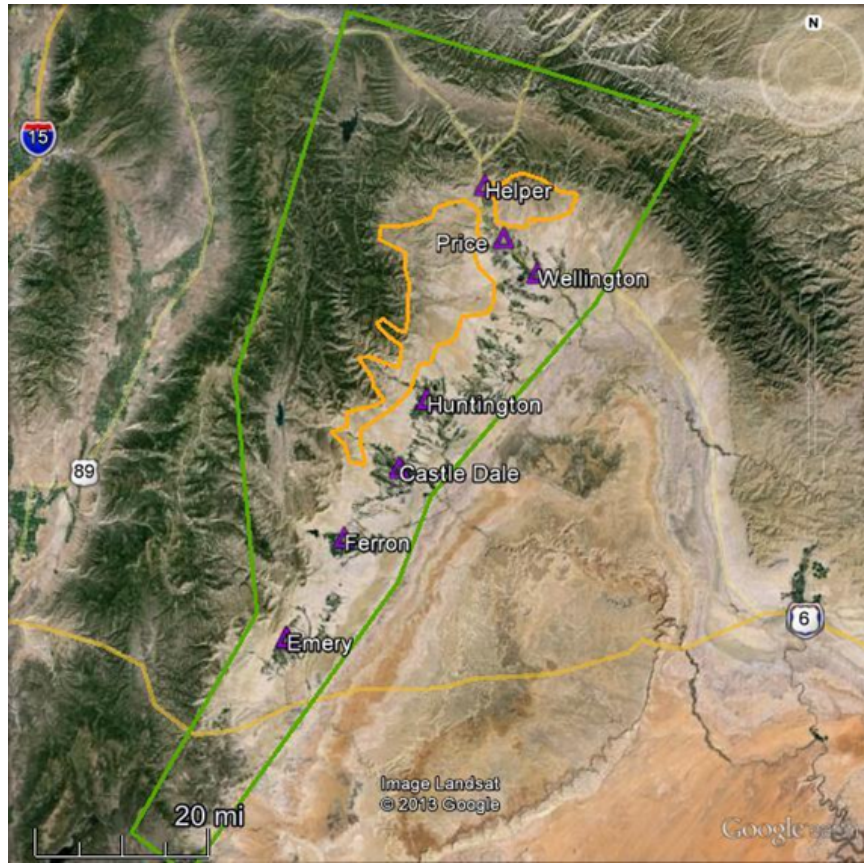
FIGURE 2.1. Study Area. The orange lines mark the rough boundary of the Helper and Drunkards Wash gas fields. The purple triangle marks the location of towns. The green line outlines the model domain.

sediments including sandstones, siltstones, shales, and coals. Ferron Coals occur as sequences of thin, separated layers, with a typical sequence consisting of 3 to 6 seams. Individual layers have thickness of 2 to 7m within sequences that total 45 to 60 m in thickness (Anna 2003, Lamarre 2003). A generalized stratigraphy is shown on figure 2.2. The Ferron thins from west to east, from a thickness of around 150m at the western edge of the study are to zero where it pinches out to the east of the study area (Anna 2003). An extensive outcrop of the Ferron is present along the eastern edge of Castle Valley with the unit dipping West at 45m/km as it approaches the Wasatch Plateau (Anna 2003). In the northern portion the Ferron plunges to the

north. The north-trending Joes Valley fault system punctuates the Ferron along the Eastern edge of the Wasatch plateau.
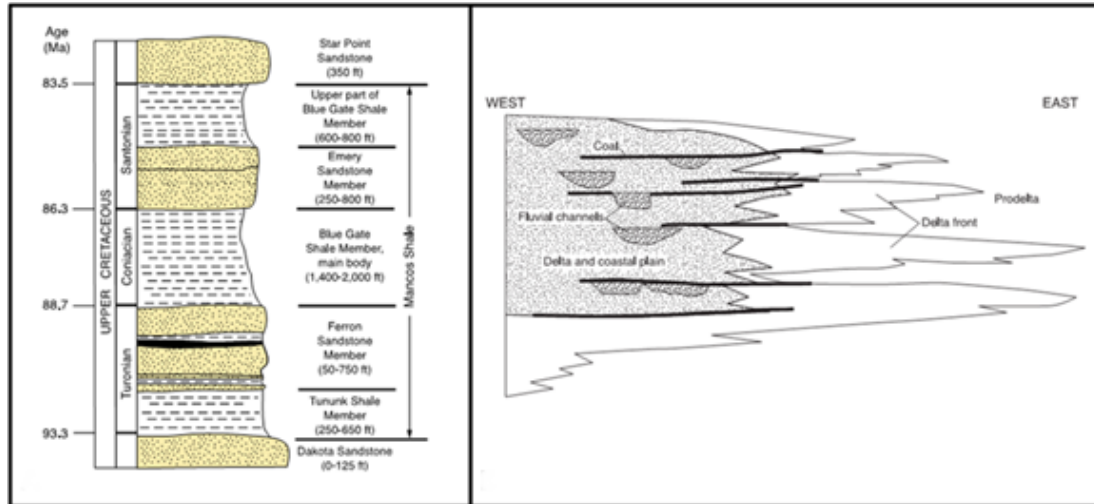


FIGURE 2.2. Generalized Stratigraphy of the Mancos Shale (left) and the Ferron Sandstone (right)

## 2.3 Hydrogeology and previous modeling efforts

The conceptual and numerical flow models used in this study are based upon a prior investigation undertaken by the USGS (Anna 2003). For one component of this study, a regional flow model was constructed in MODFLOW (Harbaugh 2005) and used to simulate drawdowns around the coalbed methane fields. The Ferron Sandstone was assumed to be fully confined by the Bluegate and Tunank shales and the stratigraphy was simplified into 7 distinct hydrogeologic units with five flow units included as model layers. The Ferron Coals were located in the middle model layer. The system was assumed to be at a steady state prior to 1979. For this period of the model, water enters the system along the western boundary as recharge through the Joes Valley fault system and exits near the eastern boundary at seven drains where major streams intersect the Ferron outcrop. A post-1979 transient period was modeled in

which water also exits as pumping from the well fields. A no flow condition was set along the northern boundary where the Ferron dips into the Uintah Basin and along the southern boundary where the Joes Valley Fault System converges with the Ferron outcrop. A generalized depiction of this conceptual model is shown in figure 2.3.
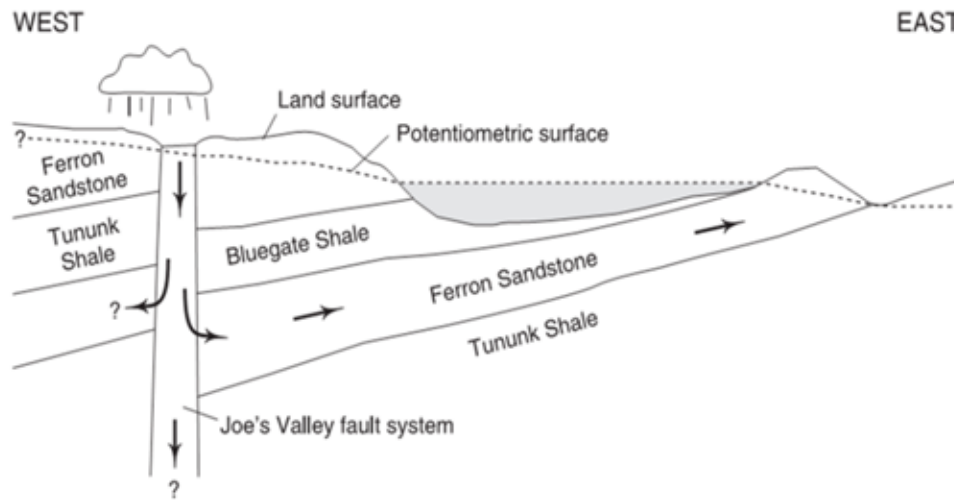


FIGURE 2.3. Generalized Cross-Section of Anna (2003) Conceptual Model of the Ferron Sandstone Hydrogeology

## 2.4 Conceptual model

The above conceptual model is adopted in this study with some modifications. The stratigraphy within the sandstone is simplified further to include only two zones of distinct hydraulic properties within three layers. These layers are composed of homogeneous sandstone with zones of coal included in the middle layer. The influence of faulting on permeability is introduced by defining fault zones in every layer in cells that lay below any fault that is well-identified at the surface. In these zones, the hydraulic properties differ from the rest of the layer only in their anisotropy. The outflow locations are also simplified. Because stream records indicate that the major streams flow perennially at the outcrop, the outflow locations were implemented

as constant head nodes rather than drains. This change removed seven unknown parameters from the model.
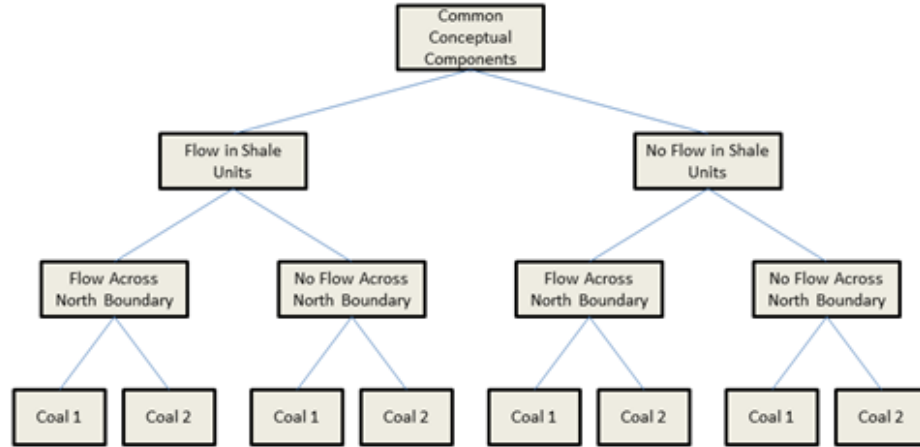


FIGURE 2.4. Tree Digram of Alternate Conceptual Models

For three of the conceptual model features - the behavior of the shale units, the Northern boundary condition, and the coal zone extent - a choice of two alternative conceptualizations are considered to be plausible for each feature. These alternative conceptualizations are included alongside the set of parameter realizations as components of the model ensemble. Two alternatives are proposed for the vertical extent of the model. The first adopts the assumption that flow is fully confined by the shale units. The second expands the flow domain to include the lower portion of the Bluegate Shale and upper portion of the Tununk Shale units. The shale units were included in order to account for the important contribution that thick low-permeability units have on average groundwater ages. It also allows us to compare the head and age results of these conceptualizations. In the second conceptualization, the shale materials are considered to have identical hydraulic properties. Additional modifications were made to the boundary conditions of the model. Two conceptualizations of the northern boundary are included. The first sets a no-flow boundary along the entire

fault. The second allows flow into or out of the system, with rates that are considered unknown. Two conceptualizations of the extent of the coal zone are also included. The first takes the locations and thickness from a geologic model constructed in ArcGIS. The second adjusts these layers to be consistent with mapped Ferron Coal thicknesses (Tabet, Hucka, Sommer, 1995). The eight combinations resulting from these alternate conceptual decisions are charted on figure 2.4.

Several notable assumptions are present in the conceptual models that were adopted. These include: evaporation at the outflow locations is not important; the aquifer system remains saturated and hydraulically connected to the streams; the streams are perennial and have a constant stage; gas flow is unimportant; and all pressure declines in the wells are from the extraction of water.

## 2.5   The Ferron coalbed methane play

The Ferron coalbed methane play spans a 70 mile length of the Castle Valley from the Southeast of Emery to the North of Price, Utah. The play consists of three fields: the Buzzards Bench to the west of Castle Dale, the Drunkards Wash west of Price, and the Helper State north of Price. Overall the field contained over 613 wells in 2012 and produced around 2.7 x107 MCF of gas and 5.4x106 BBL of coproduced water. Peak gas production of over 8.6x107 MCF occurred in 2002 and peak water production of 2.4 x107 BBL occurred in 2001 (Utah Department of Natural Resources, 2011). The Drunkards Wash is the oldest and largest field in the area and is the proposed location of the stimulation system discussed in this study.

Coalbed methane is a source of natural gas within coalbeds that is adsorbed to the solid matrix of the coals. Gas is extracted from coalbeds by depressurizing the aquifer with large withdrawals of water and allowing methane to desorb from the coals. The gas and water then travel toward the well through fractures in the coalbed. At the surface, the gas is collected and the water discarded. (Keith et. al. 2003) The

coproduced water in the Drunkards Wash field is too saline to be disposed of at the surface and must be re-injected to the deeper Navajo Sandstone aquifer (Randall 2009).

Coalbed methane gas can be thermogenic or biogenic. Biogenic gas is produced within the coalbeds by active microbial communities, and there has been a recent interest in stimulating this type of methane production by altering the microbial environment (Akob et. al. 2013). The gas in the Drunkards Wash field is believed to be partially biogenic, and with gas yields decreasing in this field, a plan is being considered to inject a nutrient solution into the coalbeds in an attempt to stimulate the production of the biogenic gas.

## 2.6    Injection and extraction system

The system being proposed for stimulating gas production in Drunkards Wash would inject a nutrient solution into a well and extract what remains through one or more nearby extraction wells. The two designs considered in this study are the dipole and five-spot designs that are commonly used in oil and gas extraction. The dipole consists of a single up-gradient injection well and a down-gradient extraction well. The five-spot includes four equally spaced extraction wells around a single injection well. The extraction wells are located  mile away from the inject ion well in both designs. This spacing is consistent with BLM regulations (Matava, pers. comm.).

The performance of each design (number of wells and injection and extraction rates) is best assessed by modeling the fate and transport of the injection solution at the local scale. Specifically, it is important to assess the ability of a design to sweep a relatively large area while capturing all of the injection solution. At this scale, however, no natural boundaries are present. To address this, the regional scale model, for which natural boundaries and observation data are present, is modeled at a coarser resolution and the calculated head at the location of the boundaries

of the smaller model are assumed to be equal to the regional model results. This telescoping modeling approach allows us to include regional hydrologic uncertainties in the assessment of treatment designs. However, this approach requires that the small model be large enough that it is reasonable to assume that the changes in head due to pumping and extraction occurring within the smaller model do not propagate to the model boundaries.

## 2.7   Observation data

Hydraulic head measurements in the aquifer are limited to a small number taken within a relatively small area before major development of the aquifer began. For this study, 100 data points were used, with only 11 of those collected after the first transient stress period of 1979. Streamflow measurements were collected from six current and historical gages near six of the eight constant head nodes (U.S. Geological Survey, 2011). These data were not used to calibrate the model by matching baseflow rates to the flow rates that occur between the aquifer and stream at the models constant head nodes. Although this procedure is common, it was omitted because none of the streams had gages upstream and downstream of the outcrop as would be necessary to calculate the contribution of flow from the outcrop to the total stream flow. However, because there was still a need to constrain flow rates in some way to relieve the insensitivity and non-uniqueness that is especially present when head observations are used alone (Anderman, Hill, Poeter, 1996) the data were used to create flow rate upper bounds. These bounds filtered the ensemble of models by removing any model which simulated a flow from any constant head node into its corresponding stream with rates greater than twice the average annual flow rate measured in the stream just downstream of the Ferron outcrop.

The data available are scarce, especially during the transient model periods. These data will primarily constrain the alternative conceptual models and steady-state hy-

draulic parameters. The transient storage parameters have little supporting data and the transport parameters have none. The bases for choosing parameter ranges are discussed below. Under these conditions it is especially important to explicitly consider prediction uncertainty in decision-making, as DIRECT does. While it is not an element of this study, this would also be an ideal circumstance to employ the Discrimination/Inference component of DIRECT to optimize future data collection efforts.

CHAPTER 3

# METHODS

## 3.1   DIRECT and multi-model approach

Each stage of modeling described in the sections below employs aspects of DIRECT. Under DIRECT, results are presented using the concept of expected outcome. This is the outcome that is suggested by an ensemble of models, where all members are included and weighted by likelihood. Each stage of modeling considers results from this perspective. In addition, the local flow and transport modeling that is used to investigate the alternative circulation system designs introduces the concept of economic costs and seeks a design that minimizes expected costs. The generation of expected costs, and the importance of modeling in this way, is discussed in more detail in the cost section below.

The use of the multi-model approach employed by DIRECT allows for the inclusion of both structural and parameter uncertainty in the modeling process. These sources of uncertainty are included in each of the separate modeling steps described below by generating ensembles of models that include combinations of alternate conceptual models and sets of parameter realizations. Results are collected by running a separate simulation for each of the thousands of models included in the ensemble. Except for the transport model, for which no applicable data exists, the predictions of each model run are compared to observation data to compute measures of relative model likelihood. The uncertainty in the model predictions are viewed as a combination of the range of results produced by the ensemble and their relative likelihoods. All transport scenarios are considered equally likely, adopting a uniform (uninformed) prior. Details of the model ensembles are outlined in the sections below.

## 3.2   Telescoping model approach

Modeling was divided into three distinct steps based upon the physical process and geographic area being investigated. These steps were: (1) the regional flow model, (2) the regional groundwater residence time model, and (3) the local advective transport model. The advective transport and residence time models were both transport simulations that required flow fields to be specified in advance. For both models, results from the regional flow model were used as inputs. The residence time model used the results from the steady state period and from the entire model area. The advective transport model collected the results from the end of the transient simulation and from the geographic subarea that was covered by the advective transport model.

The models were constructed and run in sequence using the open source MATLAB codes mflab (Olsthoorn 2011). These codes have been designed to build, run, and manipulate different aspects of models that use MODFLOW (Harbaugh 2005) and its companion software. These codes were edited and built upon specifically for this study to create a system for automating the construction and execution of large model ensembles, and for coordinating the sharing of files between the three interrelated modeling steps.

## 3.3   Regional flow model

The conceptual flow models described in the previous modeling efforts section above were translated into MODFLOW numerical flow models using mflab. These models were used to simulate flow conditions near the proposed injection and extraction system and to simulate groundwater residence time distributions for the entire aquifer system.

A geological model of the Ferron Sandstone aquifer system was produced and used to map hydrogeologic units and define the vertical discretization of the model. The surfaces of the Bluegate Shale, Upper Ferron Sandstone, Ferron Coal, Lower Ferron

Sandstone, Tununk Shale, and Navajo Sandstone (bottom of the Tununk Shale) were generated by interpolating between records of formation top depths measured in boreholes. Formation top data was gathered from the Utah Oil and Gas Data Research Center (Utah Department of Natural Resources, 2011), and from a dataset gathered for a previous Ferron Sandstone hydrogeologic study (Anna 2003). Overlapping of the generated layers was eliminated by setting minimum thicknesses to each hydrogeologic layer and adjusting the generated layer elevations from the surface downward. In places where the Bluegate Shale was interpolated above the ground surface, its elevation was capped at the ground surface and a minimum thickness was not applied. Surface elevation data was obtained using the National Elevation Dataset for Utah (U.S. Geological Survey, 2011). Surface faults were collected from geologic maps retrieved from the Utah Geological Survey (Weiss, Witkind, Cashion, 1990; Witkind, Weiss, 1991; Witkind, Weiss, Brown, 1987; Witkind, 1988; Doelling, Hellmut, 2004; Doelling, 2002) and included in the geologic model.

The domain was divided horizontally into a grid of 400 by 170, mile by mile cells. The vertical dimension was divided into 5 layers matching the stratigraphy described above. The active flow region was bounded on all sides by no flow cells. Recharge cells of prescribed inflow rates were located along the western side of the domain in cells that were intersected by the Joes Valley fault system. This region was further divided into a north and south region, and the recharge rates in the southern region were related to the rates in the northern region by multiplier parameter. The major model features are displayed on figure 3.1, with a coarser grid shown on the figure for a clearer illustration.

Two conceptualizations of the northern boundary were included, (1) no flow and (2) prescribed flow inward or outward. These were implemented by either (1) leaving active cells bounded by no-flow cells at the boundary or (2) introducing a well at every cell along the northern boundary, prescribing a total inflow or outflow rate, and dividing that rate evenly between every boundary cell. Seven constant head cells

were defined along the eastern boundary where major streams intersect the Ferron Sandstone. For steady-state simulations with no outward flow along the northern boundary, the constant head cells are the systems only outflow locations. The southern boundary remains no-flow in all conceptualizations.
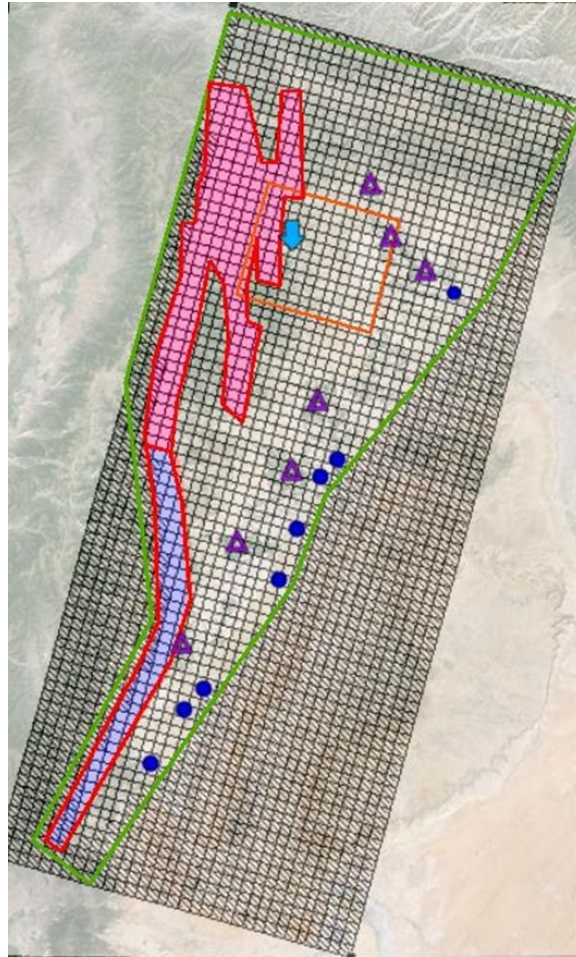


FIGURE 3.1. The Regional Flow Model Domain. The green line outlines the active region. The red polygon outlines the recharge area that is divided into a north and south sub region. The blue circles mark the location of the constant head nodes. The blue arrow marks the location of the proposed circulation system. The orange box locates the sub region of interest. The purple triangle marks the location of towns.

A steady-state condition with no pumping was assumed to exist prior to 1979 and used to model the pre-development head distribution. This distribution was used to

define a flow field for the residence time simulations and to provide an initial condition for the transient simulation. The transient flow simulation covered the years 1979 2011 and was divided into 33 annual stress periods. The stress period length of one year assumes that no seasonal variations take place within the groundwater system. Only four of the hydraulic head observations were collected after 1979, and these were not given greater weight than any other observation. As a result, the transient period has little influence on the assessment of goodness of fit for each member of the model ensemble.

Rates of water withdrawals from CBM wells for the years 1978-2012 were obtained from the Utah Division of Oil, Gas, and Mining (Utah Department of Natural Resources, 2011). The monthly rates were converted to annual and used to specify pumping rates from the wells included in the simulation. The records for each well included the name of the geologic formation in which the wells were screened. These formations were matched to model layers and the simulated wells were placed in the corresponding model layer.

The distribution of hydraulic properties was modeled by choosing three aquifer material types with different hydraulic properties, and applying these materials to distinct zones in the model. The materials are shale, sandstone, and coal. The top and bottom layers of the model were defined as shale zones, the second and fourth layers sandstone zones, and the third layer was a sandstone zone with several smaller coal zones within it. In addition to these zones, fault zones were overlaid upon each layer to represent the Joes Valley fault system. The fault zones were given different vertical and horizontal anisotropy than the zones around them, but were otherwise given the same hydraulic properties. This setup is shown in the example cross section in figure 3.2 below. The blue layers represent the shale layers, the yellow represent the Ferron Sandstone layers, the magenta represents the zones of coal within the Ferron Sandstone, and the red represent the fault zones that cut vertically across all layers wherever there is fault at the surface. The gray area is outside of the model.
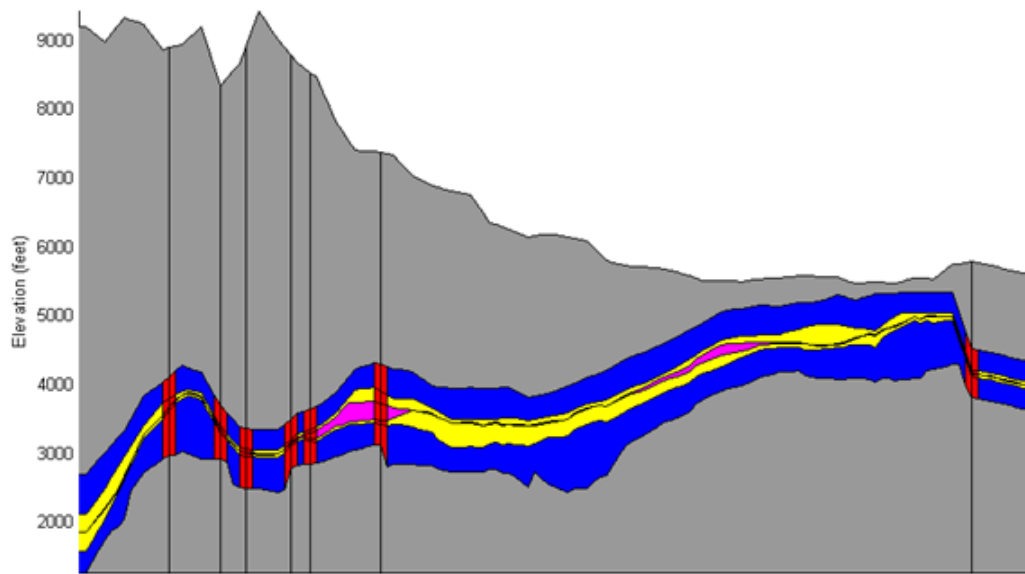
FIGURE 3.2. Hydraulic Property Zones for an Example Cross-Section of the Model. The blue area represents the shale zones, the yellow represents the Ferron Sandstone, the magenta represents the Ferron Coal, and the red represents the fault zones. The gray area is outside of the model domain.

Table 3.1 lists the 13 adjustable parameters and 2 switches that were included in the construction of the regional flow model ensemble. The parameters consisted of specific storage, hydraulic conductivity, horizontal anisotropy, and vertical anisotropy for each of the three material zones; vertical anisotropy, and horizontal anisotropy for the fault zones; a recharge rate for the northern recharge region, and a multiplier for the southern recharge region; and a northern boundary inflow rate. Each parameter was given a range of possible values that was considered to be wide enough to include all realistic values (Domenico and Schwartz 1990; Heath 1983; Batu 1998). Switches were included to control the presence of flow or no-flow at the northern boundary, and to control which of the two mapping techniques for the coal zones was used.

Table 3.1: Regional Flow Model Parameters and Ranges

| Name | Description | Range of Values | Units | Unique Realizations |
|---|---|---|---|---|
| HK1 | Horizontal Hydraulic Conductivity - Sandstone | $10^{-4} - 10^{0}$ | feet per day | 3000 |
| HK2 | Horizontal Hydraulic Conductivity - Coal | $10^{-5.5} - 10^{1.5}$ | feet per day | 3000 |
| HK3 | Horizontal Hydraulic Conductivity - Shale | $10^{-8} - 10^{-0.5}$ | feet per day | 3000 |
| HKA1 | Horizontal Anisotropy - Sandstone, Coal, Shale | $1/3 - 3$ | - | 3000 |
| HKA4 | Horizontal Anisotropy - Fault | $10^{-2} - 10^{2}$ | - | 3000 |
| VK1 | Vertical Anisotropy - Sandstone, Coal, Shale | 1/3 - 3 | - | 3000 |
| VK4 | Verticall Anisotropy - Fault | $10^{-1} - 10^{1}$ | - | 3000 |
| SS1 | Specific Storage - Sandstone | $5 \times 10^{-6}$ - $5 \times 10^{-5}$ | $\text{feet}^{-1}$ | 3000 |
| SS2 | Specific Storage - Coal | $5 \times 10^{-6}$ - $5 \times 10^{-3}$ | $\text{feet}^{-1}$ | 3000 |
| SS3 | Specific Storage - Shale | $10^{-6}$ - $10^{-2}$ | $\text{feet}^{-1}$ | 3000 |
| RECH1 | Recharge Rate - Northern | 0 - 1 | feet per year | 3000 |
| RECH2 | Recharge Rate - Southern | 0 - 3 | feet per year | 3000 |
| Q1 | Northern Boundary Flow Rate | $-0.5 \times 10^{3} - 10^{2}$ | $\text{feet}^{3}$ per day | 3000 |
| SW2 | Northern Boundary Flow Switch | 0,1 | - | 2 |
| SW3 | Coal Zone Extent Switch | 0,1 | - | 2 |

An ensemble of 12,000 models was formed by sampling from the full model space. The full model space is made up of every possible combination of parameter values in the model. It encompasses the full range of behaviors that are possible for the model to exhibit. A sample of 3,000 realizations was drawn from the space of the 17 parameters, and these parameter sets were applied to each of the 4 combinations of settings of the two switches. The sets of values for the parameters that were not used

as switches were drawn using Latin hypercube sampling, a technique for generating multivariate samples that stratifies the distribution of each input variable to ensure that their entire range is represented in the sample (McKay, Beckman, Conover, 1979). For each of these non-switch parameters, the range over which samples were drawn are shown under the Range of Values column in table 3.1.

Each model in the ensemble was run and a variety of results were collected and analyzed. Simulated heads from the entire model domain were collected at the initial steady-state period and the final (2011) transient stress period. The steady-state results were used as inputs for the regional groundwater residence time model, and the 2011 results were used to produce inputs for the local advective transport model. The simulated equivalents of head observations were collected by saving the results from each model cell and stress period that corresponded to an observation.

The performance of each member of the ensemble was assessed in two steps. First, simulated flow rates at the constant head nodes were compared to average annual streamflow rates, and any model with a simulated rate greater than 1.5 times the observed rate was removed from consideration (assigned a likelihood of zero). With the remaining models, a relative likelihood measure was assigned by calculating the normalized, inverse root-mean-squared-error (RMSE) of each member:

$$RMSE_i = \sqrt{\frac{1}{no} \sum_{i=1}^{no} (h_{ji} - \hat{h}_{ji})^2} \qquad (3.1)$$

$$L_j = \frac{RMSE_j^{-1}}{\sum_{j=1}^{nm} RMSE_j^{-1}} \qquad (3.2)$$

Where: $no$ is the number of observations, $h$ is an observation, $\hat{h}$ is the simulated equivalent, $nm$ is the number of models, and $L$ is the likelihood. These likelihood measures sum to one and are used as a weight for subsequent ensemble based predictions from the residence time model and advective transport model.

## 3.4   Groundwater residence time simulation

The focus of this study is on the optimization of pumping/extraction designs. But, the age and travel paths of groundwater in the coalbed are also of interest in understanding the timeframes and conditions under which methane is produced (Bates, McIntosh, Lohse, Brooks 2011) and whether methane production can be stimulated.

A technique for directly simulating average groundwater ages with numerical modeling is proposed by Goode (Goode 1996). This technique uses solute transport as an analogy to the movement and mixing of groundwater of different ages, allowing one to model the processes of diffusion, dispersion, and (possibly) adsorption along with advection. The age-mass technique tracks the constant aging of water by calculating the accumulation of age as a 0th order reaction, in which one unit of age is generated per unit of time.

The commonly used 3D groundwater transport modeling code MT3DMS (Zheng, 2010) was used to simulate groundwater ages using the steady state flow field results from the regional model. Two important options are included in MT3D to allow this type of simulation. First, a zeroth order reaction type is included that allows the solute phase to accumulate (or decay) at a constant rate. This allows the age-mass phase to gain one year of age-mass per year. The second option is the steady-state transport option for simulating long-term transport conditions where a state of equilibrium is reached in concentration. Although this same result can be achieved with a transient transport simulation run until concentrations do not change from step to step, the MT3D approach is more computationally efficient, making this analysis feasible for the large model ensemble used in this study.

The groundwater residence time simulations were implemented by running steady-state flow simulations in MODFLOW 2005 in sequence with steady-state transport simulations in MT3DMS 5.3. These were run for an ensemble of models that was similar to that of the regional flow modeling described above, but with a few important

differences. The additional transport parameters of porosity and longitudinal diffu- sivity were applied to each zone in the same manner as hydraulic conductivity and specific storage were applied in the regional flow model. Table 3.2 lists the additional transport parameter and the range of values over which they were sampled. Each parameter was given a range of possible values that was considered wide enough to include all realistic values (Morris and Johnson 1967; Gelhar, Welty, Rehfledt 1992) and specific parameter values were chosen in a manner similar to that described above.

Table 3.2: Residence Time Model Transport Parameters and Ranges

| Name | Description | Range of Values | Units | Unique Realizations |
|---|---|---|---|---|
| PEFF1 | Effective Porosity - Sandstone | 0.05 - 0.3 | - | 6 |
| PEFF2 | Effective Porosity - Coal | 0.01 - 0.1 | - | 6 |
| PEFF3 | Effective Porosity - Shale | 0.05 - 0.3 | - | 6 |
| PEFF4 | Effective Porosity - Fault | 0.05 - 0.5 | - | 6 |
| AL1 | Longitudinal Dispersivity - Sandstone | $0.5 - 10^3$ | feet | 6 |
| AL2 | Longitudinal Dispersivity - Coal | $0.05 - 10^4$ | feet | 6 |
| AL3 | Longitudinal Dispersivity - Shale | $0.5 - 10^3$ | feet | 6 |
| AL4 | Longitudinal Dispersivity - Fault | $0.05 - 10^4$ | feet | 6 |

The ensemble included an additional conceptual model in which flow (and age- mass transport) was confined entirely within the sandstone layers, and the shale units had no influence on results. This conceptual choice was included to investigate the impact that waters from the large low-permeability shale units have on average groundwater ages in the more permeable materials of the adjacent aquifer. These layers are typically treated as confining units and excluded from flow simulations, yet Bethke (2002) proposed that mixing between large low and high permeability layers has a significant influence on the average groundwater ages in both layers. It was stated that omitting these layers and the phenomena of diffusion and dispersion

could lead to a large underestimate of ages in the high permeable units. Dispersion and diffusion were included in every simulation of flow, but the inclusion or omission of the shale units were treated as two viable alternatives whose validity would be tested through comparison to data.

An ensemble of 3000 flow, each of which had six transport models applied, was produced and run in sequence as described above. The flow simulation resulted in a steady state head distribution and values of flow at each constant head node. These results were used as described above to filter the ensemble and to calculate likelihoods. The transport simulation resulted in a distribution of average groundwater ages for the entire domain of every ensemble member. The ensemble-wide results were summarized by calculating likelihood weighted age distributions for the entire domain. The likelihoods in this calculation were derived from the flow model results; all transport models were considered equally likely. In addition, binned likelihood (histogram) charts were produced for several individual cells in the domain to produce an image of the distributions of ages throughout the model ensemble. The influence of shale was investigated by dividing the results into shale and no-shale groups and comparing the likelihood weighted age distributions. In addition, for each group a distribution of the likelihood values of the models included in that group were assembled. These likelihood distributions were then compared to assess whether either group did a better job, on the whole, of reproducing observations of head and outflow at the constant head nodes.

## 3.5   Advective transport model

Predictions of the fate and transport of the nutrient solution injected at the proposed circulation system were simulated with an advective transport model written specifically for this study. The model covered a subset of the regional flow model and used simplified results from the regional flow models to define initial and boundary condi-

tions. The advective transport model was comprised of the Theis analytical solution of drawdown due to pumping and a particle tracking algorithm. The MATLAB code is provided in Appendix A.

The advective transport model was applied to a three by three mile area centered at the injection well of the circulation system. The vertical domain was simplified into a single layer that was given the properties of uniform depth, material homogeneity, and full vertical confinement that are assumed in the Theis solution. A grid was established with a cell of size of 1/16 1/16 mile and values for depth, hydraulic conductivity, specific storage, porosity, and aquifer thickness were specified. In addition, values of time, pumping rate, injection rates, and were assigned. Table refAdvectiveParameters lists all parameters and values used.

Table 3.3: Drawdown Parameters of Advective Transport Model

| Name | Value |
|---|---|
| Cell Length | 330 ft (1/16 mi) |
| Hydraulic Conductivity | 1 ft/yr |
| Specific Storage | $10^{-5}$ ft$^{-1}$ |
| Aquifer Thickness | 100 ft |
| Injection Rate | 500 ft$^3$/day |
| Extraction Rate | 62.5 - 2000 ft$^3$/d |
| Time | 50 yrs |

The physical modeling consists of a series of six steps. (1) A uniform background flow field was defined by specifying the direction and magnitude of the gradient based on the aggregated results of the regional model. (2) Wells were introduced in either a 5-spot or dipole arrangement, given pumping rates, and the Theis solution was used to calculate the drawdown or mounding induced by each well at every cell in the grid. (3) The resulting hydraulic head distribution was calculated by superimposing the drawdown and mounding of every well onto the background head distribution, assuming that the saturated thickness is great enough to allow for superposition. (4) The gradient field was calculated from the head distribution, multiplied by the

effective porosity parameter, and reversed in direction to produce the flow velocity field. (5) A ring of 60 particles were placed evenly around the injection well and tracked through the flow field until they terminated at the boundary or one of the local minima produced by the extraction wells. (6) Capture efficiencies were calculated as the percentage the total particles whose paths terminated at an extraction well.

The regional flow model and the advective transport model were linked by using results from the flow model to generate the background flow conditions of the advective transport model. This was done by collecting the final (2012) transient head results from the flow model from the 3 mi by 3 mi area around the circulation systems injection well and calculating gradients from this distribution. Rather than using these exact results as the background flow field, the results were summarized by calculating the mean direction and mean magnitude of the gradient field. This direction and gradient were then used to generate a uniform background flow field for the advective transport model.

By translating the detailed regional flow results into two simplified parameters, a distribution summary approach could be employed for transferring the ensemble-wide results from the flow model to the transport model without referring to any specific member of the flow model ensemble. Instead, probability density functions of the background gradient direction and magnitude were approximated based upon the flow model results and used to generate synthetic background conditions and likelihoods. These probability functions were made by first taking the average gradient direction and average gradient magnitude in the area of interest, along with their associated likelihood values, and producing binned likelihood plots. A normal probability function was then fit to each of these binned likelihood plots.

With this approach, a new and smaller ensemble of background flow conditions was generated by sampling uniformly over both the direction and magnitude distributions and by taking all combinations of those sets of values to produce the full range of behavior seen in the regional flow models. Individual probability values of each

member of this ensemble were calculated by obtaining a pair of probability values for both the direction and magnitude from their respective probability density functions and taking their product. Likelihoods were then calculated by normalizing the probability of each member of the ensemble by the sum of all probabilities from the entire ensemble, thereby generating likelihood values.

Twenty circulation system designs were proposed and the model was run to calculate capture efficiencies. The twenty designs were made up of the dipole and 5-spot each trying ten extraction to injection ratios between  and 8 (extraction rate/injection rate). For each well design, the model was run for every member in the ensemble of background flow conditions, and the capture efficiency was calculated. Each of these, resulting capture efficiencies was paired with the likelihood of its background flow condition, and these likelihoods were used as weights for summarizing the results.

This analysis simplifies the hydrology in several ways. First, the use of the Theis solution assumes uniform depth and hydraulic properties, fully penetrating wells, fully confined flow, and an infinite horizontal extent. While these conditions were not present exactly for the regional model, for the small sub area analyzed, they are assumed to be approximately true. Secondly, using the hydraulic gradient to determine the flow field assumes a horizontally isotropic aquifer. Third, dispersion is omitted although it may be significant. Fourth, the drawdowns are calculated for the specific elapsed time of 50 years and the particle tracking technique assumes this is a steady-state head distribution. Fifth, a single set of hydraulic conductivity, specific storage, and porosity values set were specified despite the parameter values being uncertain.

Of the shortcomings listed above, it is believed that the fourth and the fifth points are the most significant simplifications. Each point, however, could be addressed by modifying the setup of the advective transport model. The required adjustments would be relatively straightforward, but would greatly expand the scope of the analysis and the computational effort. Alternatively, a more complete design optimization

could be performed given site-specific solute transport property values.

The approach taken in this study was tailored to address the question of how uncertainty and bias in regional flow conditions translates into local uncertainty, and how DIRECT can be used to guide decision making. It was decided that this approach, despite the above shortcomings, was better suited for our interests than a more complex numerical scheme, such as those provided by MT3D, for several reasons: (1) adding additional complexity to the ensemble, through all of the additional transport parameters, was not justified by data, (2) simulation runs were too computationally expensive to allow the explore much of parameter space at a resolution needed to achieve acceptably good representations drawdown and solute transport, (3) this simplification allowed the effect of many parameters to be boiled down into a few characteristics of the flow field, and (4) these savings allowed for more effort to be spent exploring the decision aspects of this problem.

## 3.6  Cost, bias, and uncertainty

No model is a perfect representation of a real hydrogeologic system. Therefore, it is important that a model or set of models be proposed that can approximately reproduce the full range of plausible behaviors for every aspect that is of interest to an investigation. Different behaviors of the system are reproduced by selecting different values for the parameters included in the model and different model conceptualizations, which together form an ensemble of models. It is likely that there will be several parameter and conceptual model realizations that produce similar predictions of the system behavior (non-uniqueness). But, these models may produce different predictions of specific outcomes of interest. DIRECT assesses the relative plausibility of the models within the ensemble by quantifying the degree of disagreement between the measurements of the system and the corresponding predicted values from each model. This is designed to be an objective, systematic, and reproducible measure

of model likelihood. The model likelihoods are then applied to the specific model predictions of interest to define the expected outcomes.

In practice, a variety of different approaches are taken to building a good model and to using models to make predictions and analyze scenarios of interest. Historically, models were built and calibrated largely through intuition and experience of the modeler. More recently, automated calibration and uncertainty assessment tools such as PEST (Doherty 2005), UCODE (Hill et. al. 2005), and SCEM-UA (Vrugt et. al. 2003) have become popular. But, however sophisticated and thorough the modeling approach, there is almost always a division drawn between the physical results of the model and the economic or other non-hydrological considerations that face a decision maker when choosing a course of action. Too often, the full suite of information that is encompassed in the model ensemble and associated likelihoods is condensed into a simplified summary of results that is communicated to a decision maker who uses it in conjunction with several other considerations. DIRECT is designed to avoid this step of consolidating the model information before combining it with economic considerations. Rather, DIRECT carries each model prediction forward, through a cost model or set of cost models, to an economic prediction. The contention is that this provides richer and more meaningful results that include predicted economic costs (or benefits) alongside the likelihood or probability of their occurrence. Providing model results in these terms will allow for the use of risk-based decision making, an understanding of the costs that are incurred due to uncertainty, and the evaluation of the worth of additional proposed data collection efforts.

This study focuses on the use of risk based decision making and on the costs associated with model and parameter uncertainty. Under risk based decision making, one seeks to find the optimal action to take under conditions in which outcomes are uncertain. This is done by calculating the expected cost of each action  the sum of all costs weighted by their probability  and selecting the action with the minimum expected cost (or maximum benefit). Such a calculation can only be undertaken if

both the probability (or likelihood) and the economic costs of each result are available. This is especially important when some models predict high cost outcomes, but with low probability. These possible outcomes are overlooked in most standard approaches to hydrologic modeling for decision support.

The advective transport model described above is used to test different possible circulation system designs by calculating a cost of each result. The capture efficiency described above is the important hydrological result derived from the model, but two additional factors were considered to have important impacts on costs. First, costs associated with well installation, which would be higher for the 5-spot than for the 2-well dipole, was included. Second, the cost of extracting and disposing of the pumped water, which varies with the extraction to injection ratio of the design, was included. To compare the tradeoffs of these three factors, a cost function is applied to the results of each well design, pumping rate, and capture efficiency result and is expressed as a financial or economic cost. These cost functions and values are artificial and for demonstration purposes only. That is, they were not designed to imply anything about actual costs that would be incurred by installing and operating these systems and are not informed by communication with potential operators of the facility. We also do not make any claim about the impact that the injected solution would or would not have on the quality of the groundwater. The cost function summed the three individual costs of installing wells, pumping and disposing of water, and migration of injected particles beyond the circulation system. The function had the following form:

$$C_t = C_w N + C_p Q_i (1 - CE) + C_e (1 + R) Q_i \tag{3.3}$$

Where: $C_t$ is total costs, $C_w$ is the installation cost per well, $N$ is the number of wells, $C_p$ is the cost of particle loss per unit of injected water (concentrations ignored), $Q_i$ is the injection rate, $CE$ is the capture efficiency, $C_e$ is the extraction and disposal

cost per unit of injected water, and $R$ is the extraction to injection ratio.

A further analysis again for demonstration purposes only was undertaken to explore the separate impacts that bias and uncertainty in model predictions have on the economic costs of a circulation system design. Two sets of cost curves were created to show how costs vary with pumping rate and choice of well arrangement. The first set of curves is shown only for the results for the maximum likelihood model; these being the single set of results that are associated with the model with the highest likelihood. The second set calculated expected costs by weighting the costs from each model by that models likelihood. These two curves were then compared in various ways to explore the ways that additional costs may be incurred by (1) selecting a conservative model that accounts for the uncertainty in predictions but would be sub-optimal for the actual system behavior; (2) being over-confident and choosing a design based on the maximum likelihood results.

The analysis was expanded further to compare the effect of errors in the estimation of the magnitude and direction of the background flow fields. The effect of different degrees of uncertainty in these two model components was explored by generating an ensemble of synthetic likelihood distributions that differed in the degree of uncertainty present in each component. This was done by adjusting the standard deviation parameters of the individual probability distribution of each component. These were compared to a reference case with a lower degree of uncertainty. Bias was analyzed in a similar manner. First, a single unbiased location for the maximum likelihood model was chosen. Then the medians of the individual probability distributions were adjusted away from the unbiased cases by different degrees. The costs under each of these cases were compared to the bias-free reference case.

CHAPTER 4

# RESULTS

## 4.1   Regional flow model

Figure 4.1 below shows the objective function (RMSE) values for the simulated flow
rates from the constant head nodes. If a simulated flow rate from a constant head node
was more than twice the average flow rate observed downstream in the intersecting
stream, then the RMSE was calculated; if not, a value of zero was assigned. Any
model with a flow objective function of zero had no simulated flow rates greater than
1.5 times the observed streamflows in the corresponding streams. It was decided that
allowing the simulated values to be up to twice the average observed values was a
loose enough criteria that the models with objective function values of zero could
be selected as the only passing models. There were 3626 models with flow objective
function values of zero, from the ensemble of 12,000 models.

The subsequent analysis was done only on these 3626 models. The values of the
objective function of head observations for this subset of models was used to calculate
each models likelihood value. Figure 4.2 shows the hydraulic head objective function
values ordered by objective function value. There was significant overlap between
the models that passed the flow filter and the models with the lowest head objective
function value. Of the 3626 models that passed the flow filter, 43% of them were also
among the 3626 models with the lowest head objective function value.
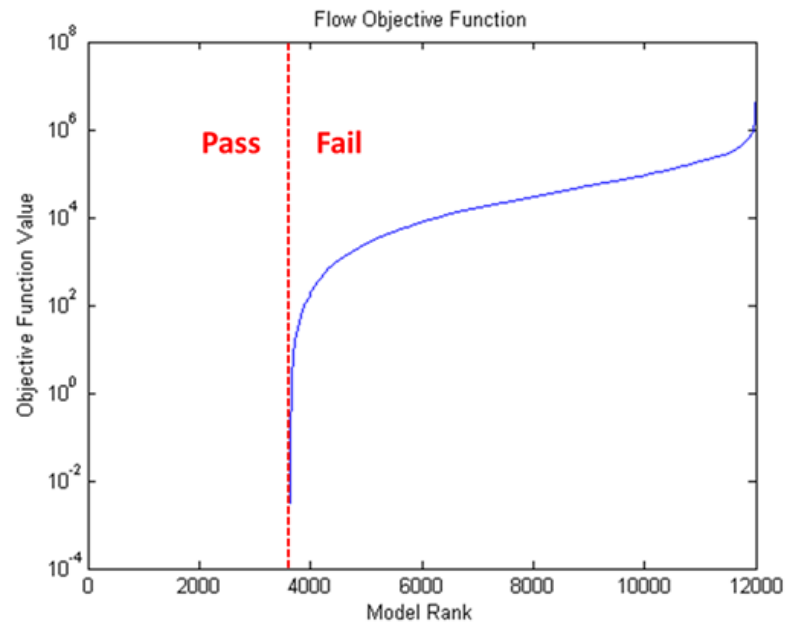
FIGURE 4.1. Distribution of Outflow Objective Function Values for the Regional Flow Model Ensemble
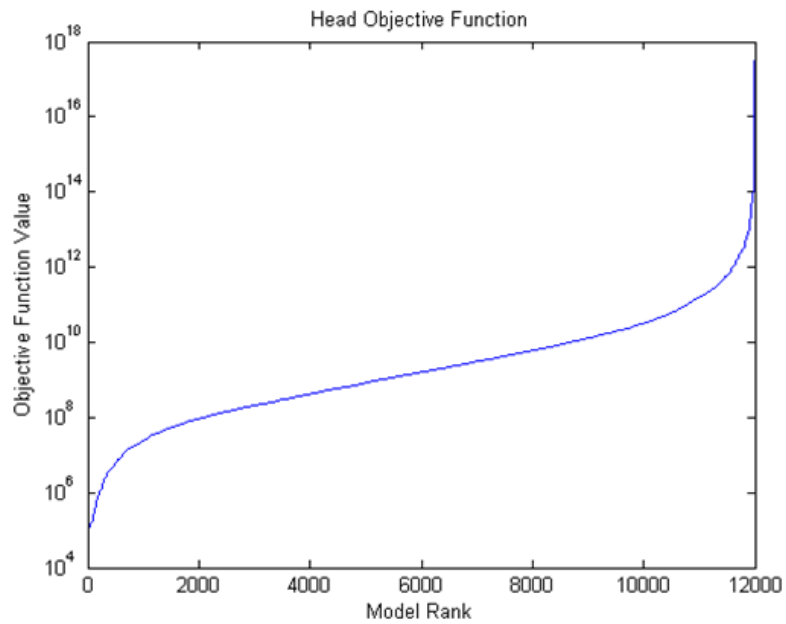


FIGURE 4.2. Distribution of Head Objective Function for the Regional Flow Model Ensemble

The distributions of the average direction and magnitude of the flow fields around the proposed circulation system are summarized in binned likelihood plots on figure **??** and figure 4.6. The likelihoods for the direction of the flow field are not distributed normally, but they are nearly symmetric and tightly distributed around the mean of 0.65 degrees with a standard deviation of 2.5 degrees. The zero degree direction is along the rows of the finite difference grid, and about 104 degrees from North. The distribution of likelihood for the magnitude has a greater spread than the direction, and is not quite a log normal distribution, but is nearly symmetric in log-space, with a mean of -2.7 and a (log) standard deviation of -2.5.



FIGURE 4.3. Likelihood Distribution of Mean Flow Field Magnitude over the Subregion Around the Proposed Circulation System
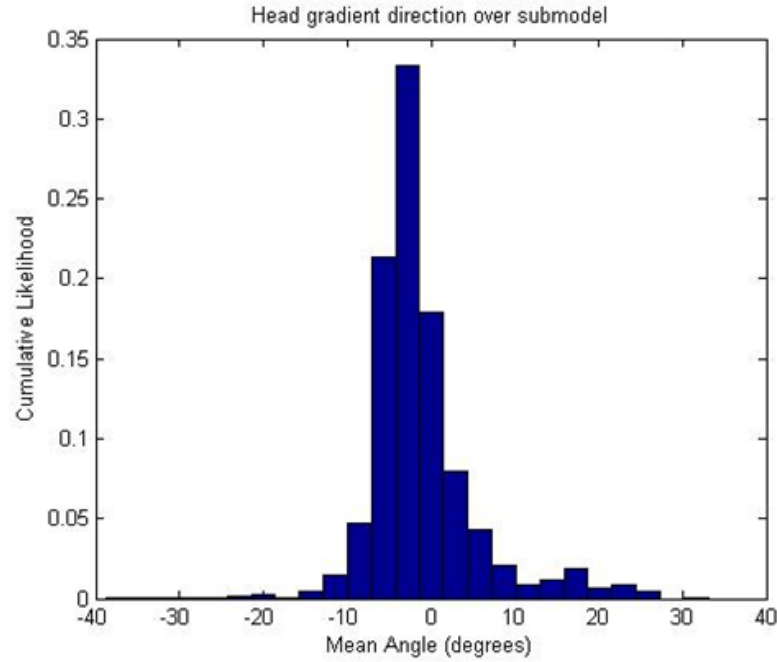
FIGURE 4.4. Likelihood Distribution of Mean Flow Field Direction over the Subregion Around the Proposed Circulation System

The flow field around the circulation system is not predicted to be perfectly uniform in either direction or magnitude for any model. In addition to summarizing the general flow field by calculating the average direction and magnitudes (shown for all models on figure ?? and figure 4.6), the variability within a single models flow field was also assessed for each model. Figure ?? and figure ?? below display binned likelihood plots of statistics that summarize the variability within the flow fields of individual model results. Figure ?? shows the standard deviation of the magnitudes, and 4.6 shows the standard deviation of the directions. The mean standard deviation of the magnitudes is about an order of magnitude smaller than the mean magnitude, and the mean standard deviation in the directions is less than 10 degrees. It was decided that these are sufficiently small variations to support the simplified assumption of uniform background flow that underlies the advective transport model.
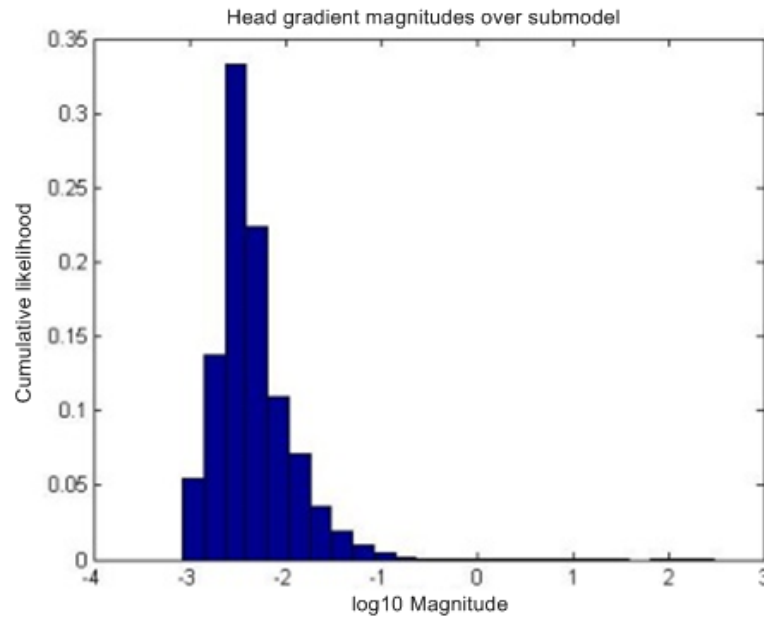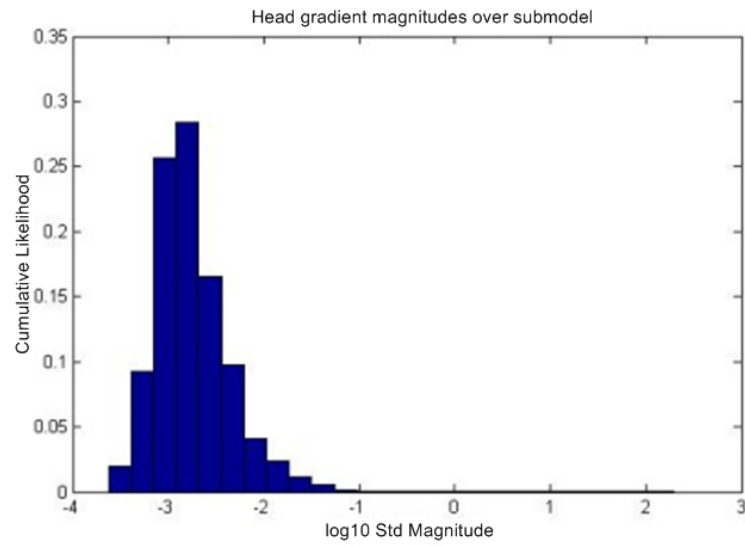
FIGURE 4.5. Likelihood Distribution of Standard Deviation in Flow Field Magnitude over the Subregion Around the Proposed Circulation System
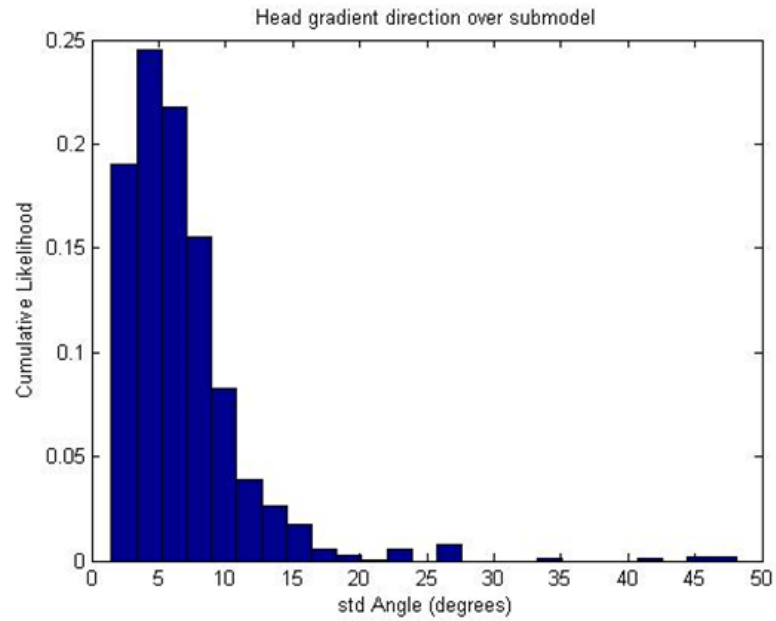


FIGURE 4.6. Likelihood Distribution of Standard Deviation in Flow Field Direction over the Subregion Around the Proposed Circulation System

## 4.2 Groundwater residence time modeling

The groundwater residence time model provided a good opportunity to explore how greatly the maximum likelihood model results may differ from the likelihood weighted summary of all model results. Figure 4.7 displays a contour map of simulated average groundwater ages, expressed in log10 years, for the third layer for all models weighted by likelihood (left) and the maximum likelihood model (right). Ages for the likelihood weighted results range from 104 years near the fault to 106.2 years in a region of stagnation in the northeast corner of the model. The results of the maximum likelihood model have lower average residence times than the likelihood weighted results showing ages from 103.2 to 105.1 years. In both cases groundwater ages in the other model layers (not shown) display similar values and distribution patterns.
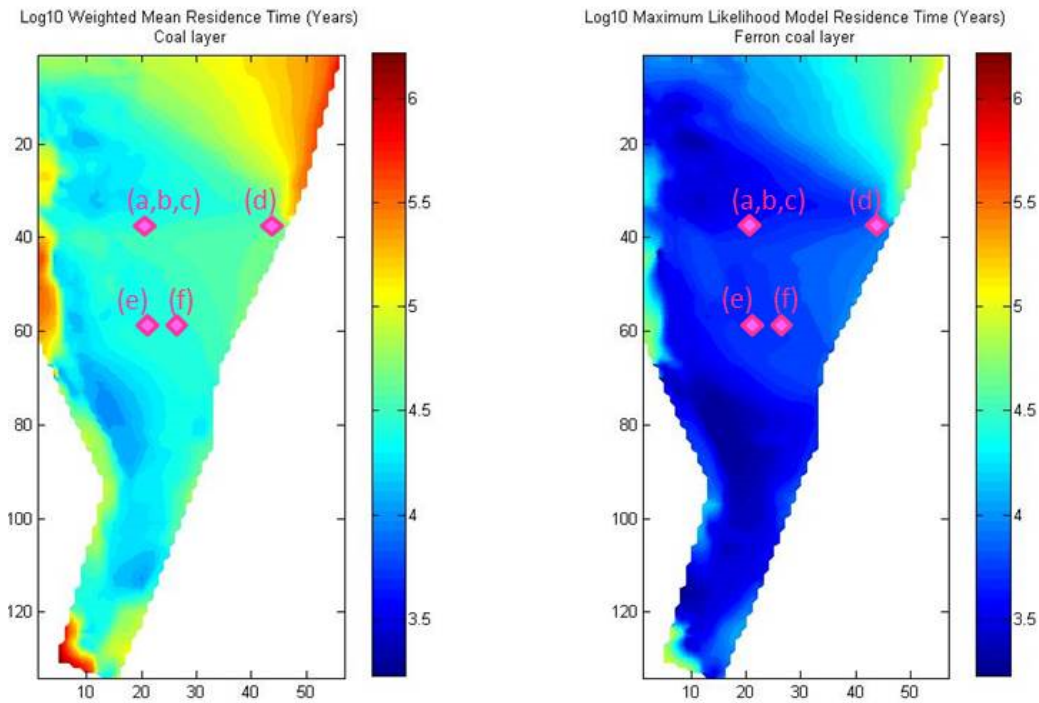


FIGURE 4.7. Log10 Mean Residence Times in Years Over the Middle Layer of the Model. The left figure shows likelihood weighted results of the entire enseble. The right figure shows results for the maximum likelihood model. The labelled diamond markers display the locations of the residence time observations points.

The large difference (roughly two orders of magnitude) between the maximum likelihood and likelihood weighted results may suggest a possible pitfall of the way the model ensemble was generated. This large difference may in part be a consequence of the method used to sample parameter space and of the relatively limited size of the model ensemble. The binned likelihood plots shown below on figure 4.8 reveal roughly log-normal distributions of the ensemble ages at different points in the model domain. Assuming that the maximum likelihood model age lies near the mode of the distribution, it is expected that the likelihood weighted age, which is equivalent to the mean of that distribution, will have a larger value than the maximum likelihood age. Being a log distribution, this difference has the potential to span a several orders of magnitude. The difference between these points may have been further exacerbated by using the quasi-uniform latin-hypercube sampling technique with very wide parameter ranges but only a relatively small 8000 samples. It may have only been within a narrow band of each parameters range that the resulting flow models produced a good fit to the data. However, with only 8000 samples the latin hypercube sampling may have only hit the sweet spot of good models a small number of times. This small set of good models may not have produced enough high likelihood values to lower the relative likelihood of poor and awful model results beyond having an influence. The effect of this on the ensemble age distributions would be to lower the magnitude near the mode and to raise those of the tails. The use of more models in the ensemble, tightening the parameter ranges, or using a sampling technique that generated a greater density of high likelihood models may have produced a less striking difference between the likelihood weighted and maximum likelihood results.

A few locations were chosen as points at which to collect and view predictions of average groundwater ages from the entire model ensemble. The locations chosen were (a) in the third layer where the injection well is proposed to be located; (b) the second layer directly above (a); (c) the fourth layer directly below (a); (d) in the third layer near the drain and down gradient of the injection location in our conceptual model of

flow; (e) in a fault zone in the 3rd layer, in a different part of the system than (a-d); and (f) off of the fault zone in 3rd layer near and (conceivably) down gradient of (d). Points (a), (b), and (c) were chosen to see how ages vary vertically. Point (d) was chosen near the outlet that was thought to be where flow from the injection well may eventually terminate. Points (e) and (f) were chosen to see how different ages would be on and off of the fault and recharge zone.

The results of the ensemble simulated ages at these six locations are summarized in binned likelihood plots in figure 4.8, with age being expressed in log10 years. These plots are produced by dividing the entire range of ages simulated by the ensemble into a number of bins, and then for every bin summing up the likelihoods of those models whose predicted ages fall within the bin. These plots show the distribution of ages predicted by the ensemble in a way that accounts for model likelihoods. The mean ages at each location are: 2.2x104 years at (a), 2.1x104 at (b), 2.1x104 at (c), 7.1x104 at (d), 2.6x104 at (e), and 3.1x104 at (f). The distribution for locations (b) and (c) appear very similar to (a) but with a shift towards younger ages in (b) and older ages in (c). It can be seen by comparing (d) to (a, b, c) and (f) to (e) that ages, as expected, are shifted towards larger values away from recharge cells. In general there is a wide spread in ages among the models in the ensemble, with ages at location (a) ranging from 100s to almost millions of years, with the best estimate of ages centering around 1000s to 10,000s.
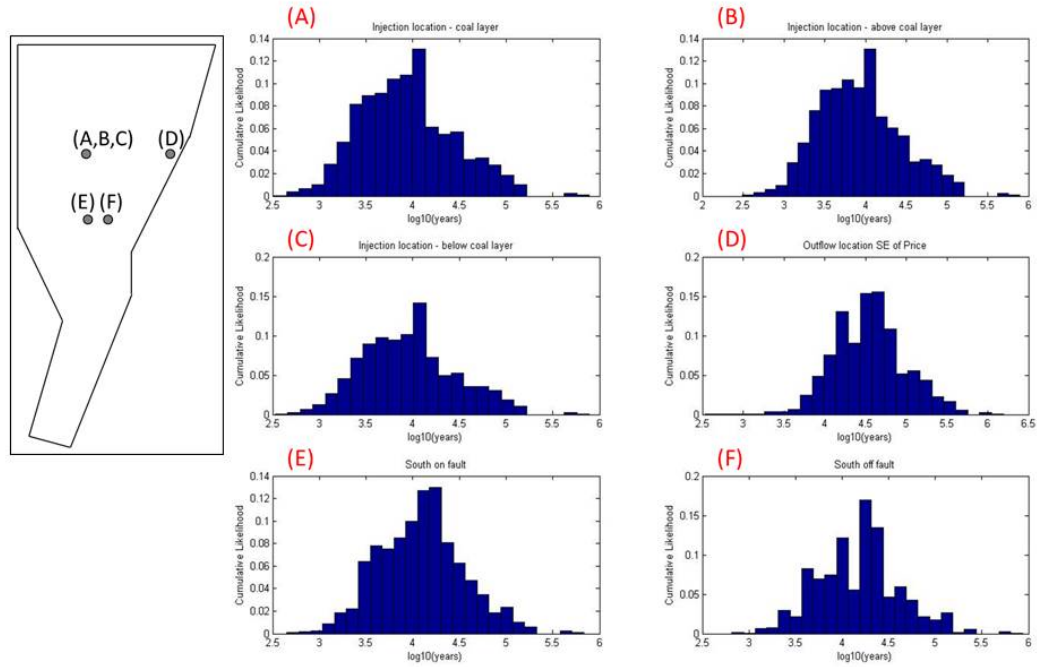
FIGURE 4.8. Residence Time Distributions at Six Model Cells

A comparison between models that included flow in the (Tununk and Bluegate) shale layers and those that treated the sandstone as completely confined reveal a clear difference in the distribution of ensemble ages. Figure 4.9 and Figure 4.10 display the binned likelihood plots of ages for the group with the shale and the group without the shale at the location (a) and (d). It can be seen that differences of an order of magnitude in age are typical and the differences appear greater at point (d) near the outlet. This comparison confirmed the expectation that large, low permeability confining units have a great influence on average groundwater ages; raising them significantly
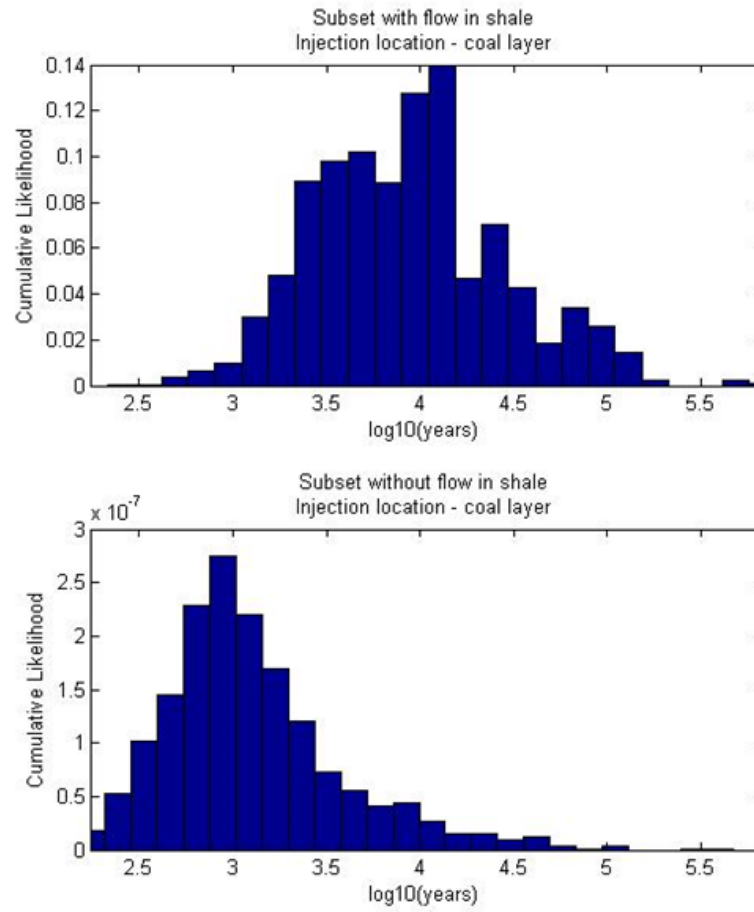
FIGURE 4.9. Comparison of Likelihood Distributions of Resdidence Times at Observation Point A for the Group of Models that Include Flow in the Shale Units (Top) and the Group of Models that Do Not Include Flow in the Shale Units (Bottom)
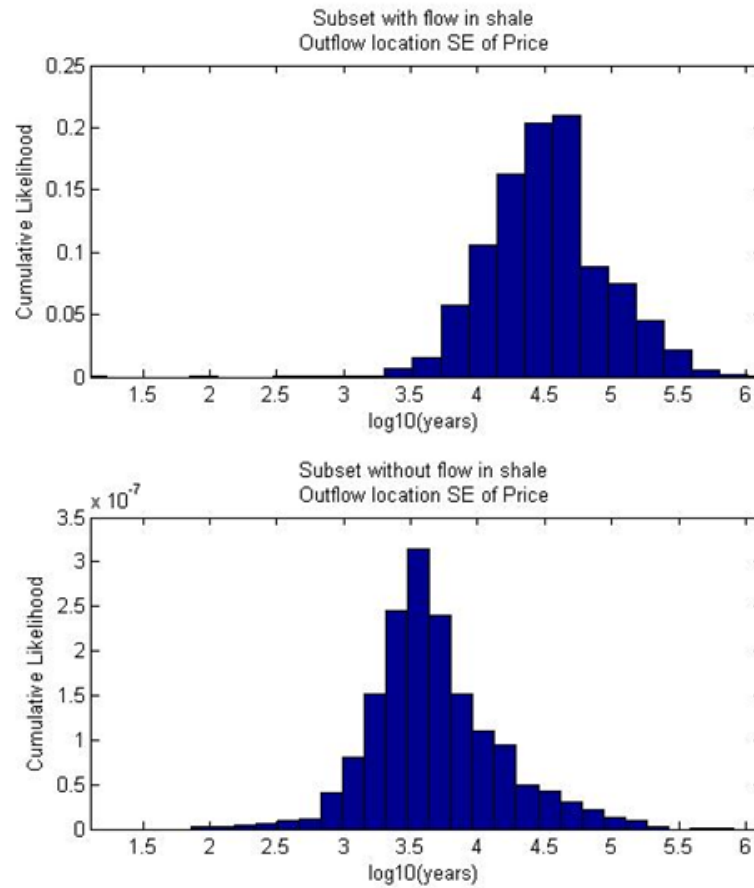
FIGURE 4.10. Comparison of Likelihood Distributions of Resdidence Times at Observation Point D for the Group of Models that Include Flow in the Shale Units (Top) and the Group of Models that Do Not Include Flow in the Shale Units (Bottom)

It was also found that there was a significant difference in how well the shale and non-shale groups of models reproduced observations of both flow and hydraulic head. Figure 4.11 compares the likelihoods of the group of models that included flow in the shale units and the group that is treated as confined between the shale units. Each plot is a histogram of the likelihood values, with the horizontal axis marking intervals of likelihood value and the vertical bars showing how many models from each group have likelihoods within each interval. The no shale models (top) have a mean likelihood of 3.8x10-4 with a standard deviation of 6.0x10-4 and the

shale group (bottom) has a mean likelihood of 1.2x10-9 with a standard deviation of 1.2x10-11. Along with the greatly different average likelihoods there is very little overlap in likelihood between these two groups of models, with almost every member of the shale group having higher likelihoods than the best of the non-shale group. This indicates that the models including the shale units do a more accurate job of representing the steady-state head distribution than those without.
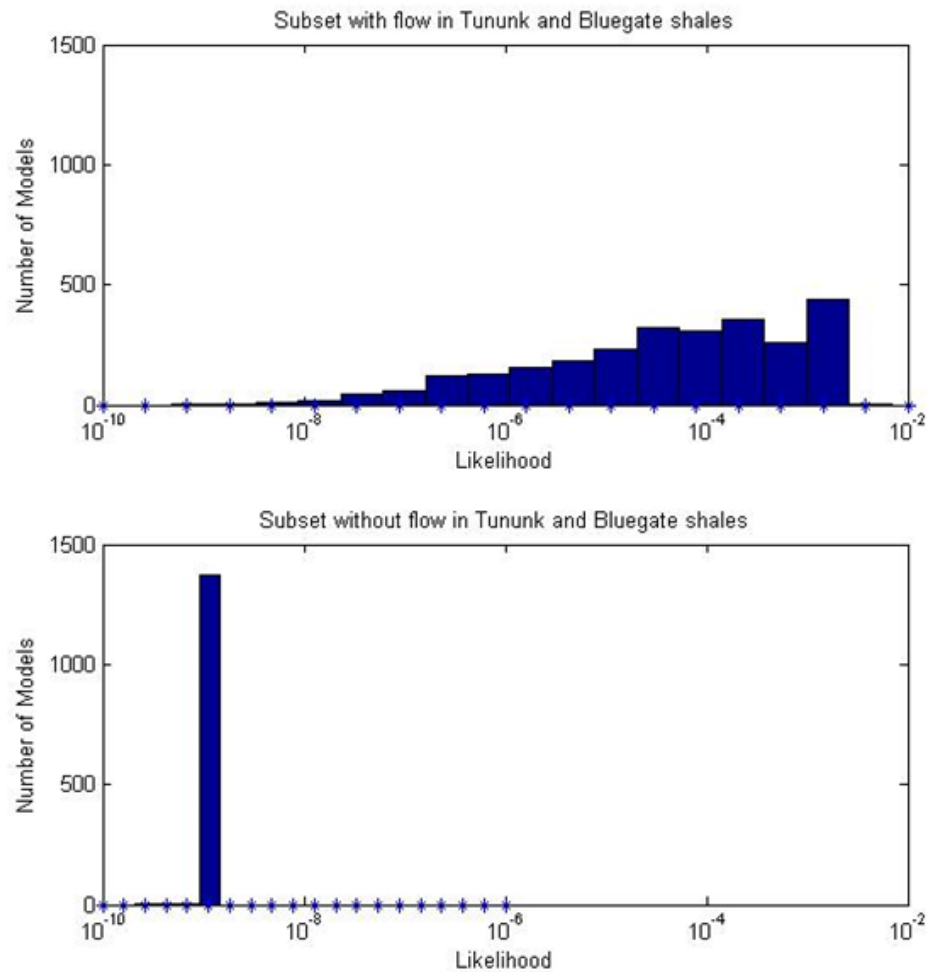


FIGURE 4.11. Comparison of Distributions of Likelihoods for the Group of Models that Include Flow in the Shale Units (Top) and the Group of Models that Do Not Include Flow in the Shale Units (Bottom)

## 4.3   Advective transport model

Figure 4.12 below displays four examples of results obtained from the adjective transport model. The solid colors represent the head values resulting from superposition of the background flow field and the drawdowns and mounding caused by the circulation system wells. The large arrow indicates the direction of the uniform background flow field. The lines in the foreground show the paths of every particle as they travel through this flow field. The particles originate around the injection well, at locations marked by a pink circle, and travel until they are captured by one of the extraction wells or exit at the boundary as escaped mass. The locations of the extraction wells are shown with pink squares, with the dipole having one well to the right of the injection well and the 5-spot having four at each cardinal direction around the injection well.

Figure 4.12 compares the behavior of two treatment methods applied under two different background flow fields with the same magnitude but different directions. The images on the left show the performance of the dipole arrangement which performs very well if the extraction well is placed down gradient of the injection well (top) - much better than the 5-spot under the same circumstances. If, however, the well is instead placed off-angle from the flow direction, the plume can largely miss the capture area of the dipole extraction well. The images on the left show that while the 5-spot has poorer performance than the best-located dipole that could be placed if one had high confidence in the flow direction, it is more robust and maintains good performance even when the flow is directed between its extraction wells. These images indicate that under conditions where the direction is unknown, or changes through time, the selection of a treatment option would benefit from robustness or versatility, while in conditions with greater certainty a more precise and effective treatment can be applied. These images also show the danger of following a deterministic modeling approach where only a single maximum likelihood model, or best calibrated model,

is used to choose a treatment strategy. In this case, if the maximum likelihood model was represented by the top row examples, one may choose the dipole due to its (slight) advantage over the 5-spot. This decision, however, would be made despite the fact that there may be other valid models that, although less likely according to data, are more like the model represented in the bottom row examples. Under these conditions the dipole performs very poorly while the 5-spot performance only drops a small amount.
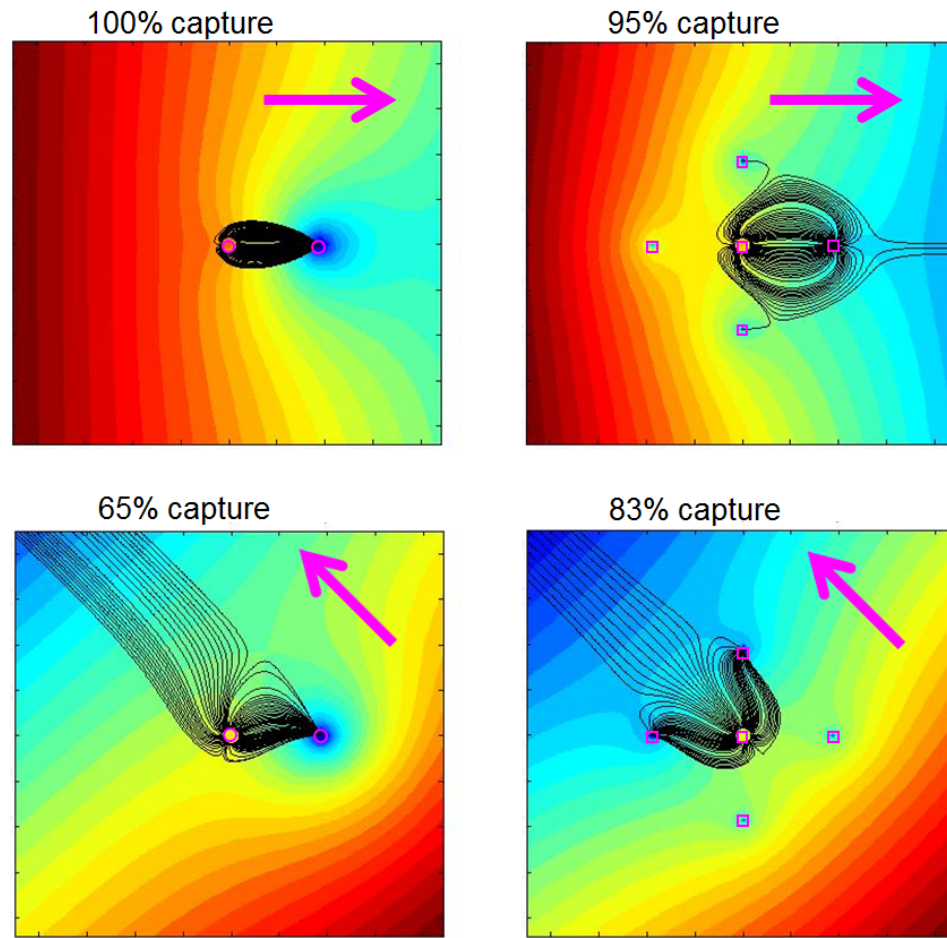
FIGURE 4.12.  Paths Simulated by the Advective Transport Model for Particles
Introduced at the Injection Well for the Dipole (Left Column) and 5-Spot (Right
Column) Circulation System Designs Under Two Different Flow Field Directions (Top
Row and Bottom Row). The color fill represents hydraulic head. The direction of the
flow fields are shown with the pink arrows and the percent of the particles captured
at the wells are shown above each panel.

The performance of each well arrangement under the maximum likelihood model
results and the likelihood weighted results are shown in the performance curves in
figure 4.13.  The blue curves show the results for the dipole and the green for the
5-spot. Dashed lines are results for the maximum likelihood background flow and the
solid lines are a likelihood weighted summary of all results.  Each curve shows the

percent of the particles that are captured by the extraction wells for each extraction rate that was simulated. One unsurprising result shown by these curves is that more of the particles are captured as pumping increases. From the 5-spot maximum likelihood (dashed green) curve, for example, it can be seen that the capture efficiency is zero at extraction to injection ratios between 0 and 2.5, that it rises between 2.5 and 4.5, and that it achieves 100% capture at extraction to injection ratios above 4.5. These curves also demonstrate how the models results, and the management decisions that they suggest, can vary substantially depending on which approach is taken to defining optimality. For these sets of curves it can be seen that under the maximum likelihood model, the dipole performs better than the 5-spot at all pumping rates, and that performance increases suddenly at extraction ratios between two and three. For this model, 100% capture can be achieved by operating the dipole with an extraction to injection ratio of 2.5 For the likelihood weighted performance curves, however, the designs perform similarly at extraction ratios below four and the 5-spot performs slightly better at higher extraction rates. The likelihood weighted results also show the performance of either design only improving gradually with increasing pumping rates; with only about 65% of particles captured with the highest extraction to injection ratio of 8.
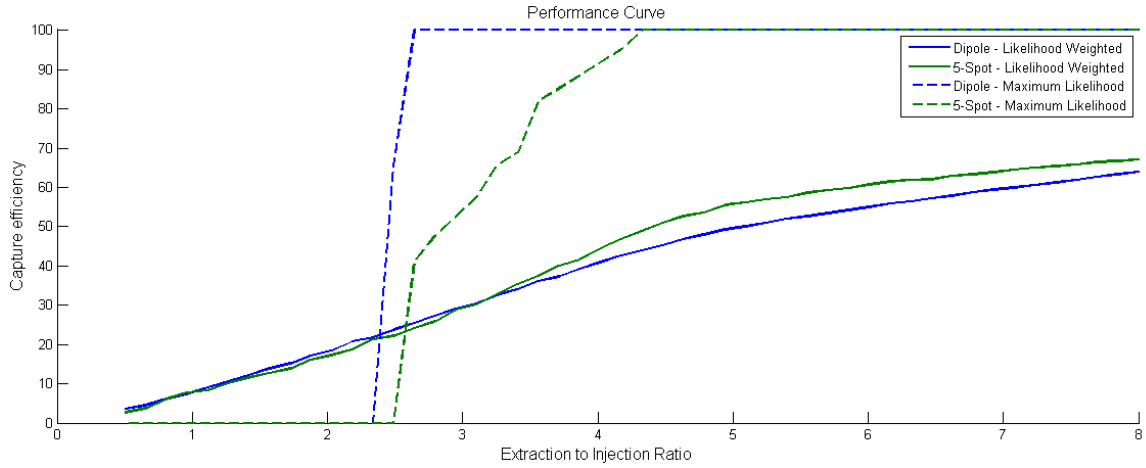
FIGURE 4.13. Performance Curves Plotting the Capture Efficiency of a the Extraction Wells Against the Extraction to Injection Ratio. The dipole performance is shown in blue and the 5-spot performance is shown in green. Results of the the maximum likelihood model is signified with the dashed curves and the likelihood weighted results of the ensemble with the solid curves.

## 4.4   Cost

The performance curves from figure 4.13 are converted to the cost curves of figure 4.14 below by applying the cost function to the results of each advective transport model. This cost function was created for demonstrating one of the important features of DIRECT. For each model result, three factors were used to collect costs: the capture efficiency, the number of wells (5 for 5-spot, 2 for dipole), and the rate of extraction. Each of these factors can be seen on the performance curves above, with the vertical axis showing capture, the horizontal axis showing extraction, and the color of the curve indicating the well design and, in turn, the number of wells.
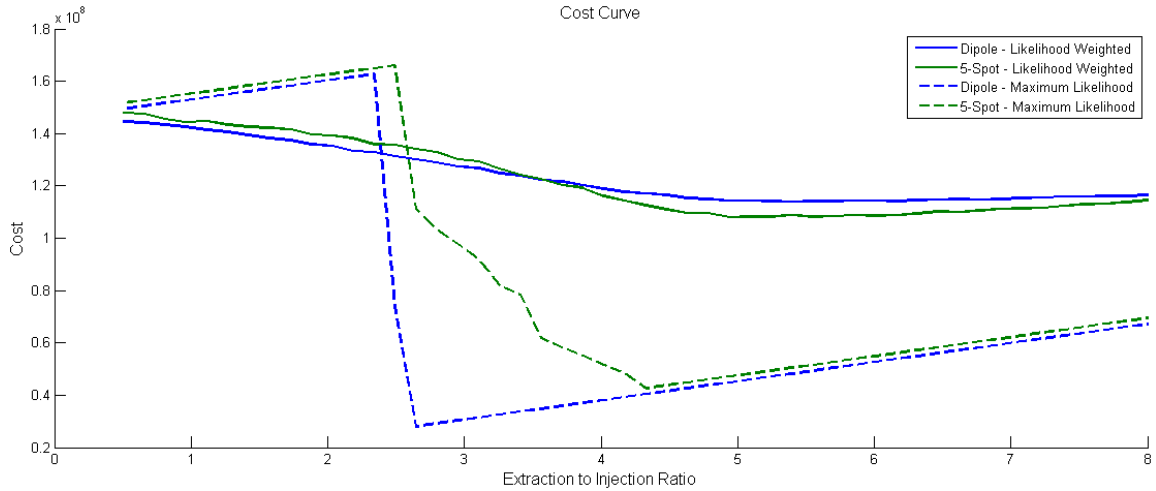
FIGURE 4.14. Cost Curves Plotting the Economic Cost of a the Extraction Wells Against the Extraction to Injection Ratio. The dipole cost is shown in blue and the 5-spot cost is shown in green. Results of the the maximum likelihood model are signified with the dashed curves and the likelihood weighted results of the ensemble with the solid curves.

The cost curves of figure 4.14 above are read similarly to the performance curves in figure 4.13, except that the vertical axis now represents total cost. The blue curves show the results for the dipole and the green for the 5-spot. Dashed lines are results from the maximum likelihood background flow and the solid lines are a likelihood weighted summary of all results. The influence of each of the three components of the cost can be identified most easily in the behavior seen in the maximum likelihood (dashed) curves. First, the differing cost of wells for the dipole and 5-spot can be seen at early and late times when the curves are slightly separated and travelling straight and parallel. At these times the costs from the other two components are the same for each design and the cost curves only differ by the higher fixed costs of the three additional wells that the 5-spot requires. The cost of extracting (and disposing of) water can also be seen during these early and late time periods when the other two factors dont vary with changes in the pumping rate and the constant increase in costs is due entirely to the cost per unit pumping. Finally, the cost of not

capturing the (solute) particles can be seen in the middle sections of the curves where the costs begin dropping as the systems begin achieving greater capture efficiencies with increase pumping rates. This continues until they reach 100% and the costs increase again due to the cost of increasing pumping. The behavior of the middle section of each curve can be compared easily to the behavior of the corresponding sections of the performance curves.

Although the costs functions for this problem are not derived from any true financial or economic costs, these curves demonstrate the simple benefit that using a cost function can provide. They combine the important factors that enter into a decision and convert the physical results, whose importance may be ambiguous or difficult to compare, into a single metric for identifying the least cost option. The example above is straightforward and is the result of the simple cost function that was applied. Given that the maximum likelihood model correctly predicted the results, the least costly choice for this example was the dipole arrangement with the lowest pumping rate that achieved 100% capture. Although this choice may seem obvious, it was not inevitable. More complexity could have been added to the cost function to represent such things as a tiered system of penalty costs for contamination or per unit costs of pumping that rose with the total extraction rate to account for the added difficulty of disposing of larger and larger quantities of wastewater. Such changes to the cost function would alter the behavior of the cost curves and produce a different optimal solution. It is even possible that the optimal design choice may have been one that achieved less than 100% capture.

The likelihood weighted cost curves in figure 4.14 behave differently than the maximum likelihood curves. These curves are the combinations of the cost curves that result from every model in the ensemble, with the influence of each being weighted by the likelihood of that model. The design with the lowest likelihood weighted costs, or least expected cost, is the five-spot with an extraction to injection ratio of about five to one. This optimal choice, however, does not appear greatly different than

designs that use the dipole or different extraction to injection ratios. The influences that shape these curves can be understood by trying to imagine all of the curves that would be drawn for each member of the ensemble. Each of these would look more like the maximum likelihood curves from above with the middle area, where the costs drop, dropping at different rates and the drop taking place over different ranges of extraction rates. The exact shape of the likelihood weighted cost curve that results from this ensemble of curves will depend upon how much variation there is in the model ensemble and how the likelihoods are spread among the models. The results seen in the figure above, which are much smoother and less responsive to changes in extraction than the individual cost curves, come about from a collection of individual curves that have a lot of variation and for which likelihoods are not dominated by a few members of the ensemble. If a successful data collection or model improvement effort were undertaken, it would have the effect of altering the shape of the likelihood weighted curves to be more like that of the maximum likelihood curves. This would be done by reducing the overall uncertainty of the model results and elevating the relative likelihood of the (hopefully correct) maximum likelihood model to give it a greater influence over the aggregated results.

## 4.5    Uncertainty

The analysis below explores how much more costly the implementation of a plan may become when a only single model is used to make predictions even though there is uncertainty in how to build a model to best represent the real groundwater (or any other natural) system. This form of error is always present in some degree in any modeling task  it is not merely the result of an unsure modeler  and it can result in outcomes that are different than those that are predicted by the model that is used for making decisions.

When there is uncertainty about which set-up of a model is best or correct for

representing a system there will also be uncertainty about the outcome of any action taken in that system. If, as in the example explored in this paper, a chemical solution is injected into an aquifer and we are uncertain about the direction of groundwater flow, then we will also be uncertain about where the solution will be carried. It may be problematic, and costly, if one designs a system that would capture all of the solution under the most likely flow conditions, yet different flow conditions are present that carry the solution in an unexpected direction. This kind of potential harm caused by uncertainty can be guarded against by trading off efficiency in a design for robustness. In this example, the trade-off between efficiency and robustness may be understood as a switch from the simpler and cheaper dipole design to the larger and more expensive 5-spot design, and/or from a lower to a higher pumping rate.

The common approach to using models is to identify only the single most likely setup of the model and to make predictions without considering alternatives. With this approach one cannot identify a more robust design and the design may be vulnerable to unexpected and costly results. The chance that this type of undesirable outcome will occur, and the amount of added cost that it may introduce, both depend upon the degree of uncertainty that is present. It also depends upon what aspect of the system is uncertain. In some cases the design chosen for the most likely conditions may perform well even if some aspect of the system is not well known, while in other cases unexpected results may cause huge problems. An analysis of how much uncertainty is present in a model is not necessarily easy to find out, nor is collecting data to reduce this uncertainty, so it may be helpful to first estimate whether, or how much of a threat exists by ignoring uncertainty.

The analysis below is an example of how this type of information can be approximated. In this example uncertainty can originate from two aspects of the groundwater system that are represented in the model: (1) the magnitude and (2) the direction of the background flow field. If no uncertainty is present, the exact magnitude and direction of the flow field will be known and the most efficient design can be chosen.

If there is uncertainty in either the direction or magnitude of flow, then the chosen design may result in a variety of different possible outcomes, each one having different costs. These possible costs are summarized as an expected cost and this expected cost will be higher than the outcome that was anticipated when choosing the design. To estimate how much higher this expected cost is with different degrees of uncertainty, a single flow direction and magnitude were chosen (this would be from the most likely model) and a design was selected that performed best under these conditions. This design is a dipole with an extraction rate roughly three times greater than the extraction rate. Artificial uncertainty was then added (as described in greater detail below) by considering conditions that were more and more different than those that exist in the chosen model and looking at how the chosen design would perform under those conditions. This was done for the direction and magnitude individually and in combination.

The results are shown as a contour map of additional costs calculated from these different levels of uncertainty in both the direction and magnitude of flow. This contour map is shown on figure 4.16 and discussed in greater detail below. These results show that, for this example, a significant amount of additional cost may be incurred by ignoring uncertainty in either the direction or magnitude of flow. It appears, however, that being uncertain about the direction can cause much greater harm than being uncertain about the magnitude of flow. In addition, while the largest added expected costs are incurred under very large, and possibly unrealistic levels of uncertainty, even modest levels of uncertainty can raise the expect costs by 50 to 100 percent. This result may help justify a further investigation of where uncertainty is actually present in the model and how one may act to reduce this uncertainty.

The presence of bias in model results is a hazard similar to uncertainty which may also complicate a models use in decision making. The choice to proceed with model results without considering whether or not there is a bias may lead to the same type of regrettable outcomes as are described above for uncertain model results. The

ways in which additional costs may be incurred as a result of biased model results is explored in greater detail following the uncertainty section below.

## 4.6    Uncertainty analysis

In the context of uncertainty, the likelihood weighted cost curves of figure 4.14 are curves of expected costs. For this study, the source of uncertainty that is considered is caused by the background flow conditions not being perfectly known. Each of the models that were considered in the ensemble produced background flow conditions that, if they were the true conditions, would results in particular cost curves of operating the circulation system. If each of these conditions, and associated cost curves, are considered possible, and the relative likelihood of each is expressed by the measure of likelihood that was calculated above, then a single cost curve can be created that accounts for all of the different possible results. This is found by calculating the likelihood weighted mean of every cost curve - producing the curve of expected cost. Finding the point on this curve that produces the lowest possible expected cost is how one can choose a robust design that performs fairly across the many possible true flow conditions. This design, however, is unlikely to produce the best results for the actual conditions  it is simply a balanced choice made when the actual conditions are not known. The actual cost of uncertainty can only be determined by comparing the cost of the strategy chosen based on expected cost to the true optimal design with perfect knowledge of the true system.

The alternate strategy of choosing the optimal design from the single cost curve that is the most likely to be true may provide the best design for the actual conditions, assuming that one of the models produces true results. This will only happen, however, if there is a true model in the ensemble and if it is correctly (and uniquely) identified as the model with the maximum likelihood. If there is uncertainty about this then there are many different ways in which this design could end up perform-

ing very poorly, depending upon which alternative background conditions are truly present.

Differences between these two approaches (maximum likelihood or likelihood weighted optimal design) and the type of costs that may result from following either one are demonstrated on figure 4.15. This figure shows the same two sets of cost curves as figure 4.14 with a few important designs and costs highlighted. These points are selected and compared with the purpose of showing the expected costs that could be incurred by ignoring uncertainty and the costs that could be incurred by having uncertainty. First, if the optimal point from the maximum likelihood curve is used to select a design (point I), the cost that this design would have if this model were the true model is cost A. This may be the cost that is anticipated if the maximum likelihood model is expected to be true, but if there is uncertainty in actual conditions then this design may have very different costs depending upon which alternate conditions are actually present. The expected cost of this design across all of these possible conditions (that were considered) is cost B. The difference between costs B and A is the expected additional cost over what is anticipated from the maximum likelihood results.

The design shown on point II minimizes the expected cost curve with the cost shown as cost C. The difference between cost A and costs C is the expected additional cost that is incurred by ignoring (or grossly underestimating) uncertainty and selecting the best design from the maximum likelihood cost curve rather than the robust design from the expected cost curve.

These two sets of curves can also demonstrate how the robust design that minimizes the expected cost may be non-optimal for the actual true conditions. This is done by assuming for the moment that the dashed cost curves actually represent the costs that would occur under true conditions. Point II, the robust design which results in the minimum expected cost of C, would end up producing the actual costs of D under the true conditions. The difference between points D and C is a cost

63

that results from selecting a design that is robust under uncertainty, but non-optimal for the actual conditions. This cost results from the need to trade off efficiency for robustness when a choosing a design that guards against acknowledged uncertainty.

A final comparison can be made to show the total cost that is added by uncertainty, even when it is recognized. If we assume that the dashed curves are the true conditions, that one of the models in the ensemble predicts these true conditions exactly, and that this model would have a likelihood value of one given perfect data, then this model would allow us to choose the truly optimal design of point I. Comparing the cost A of this best design with the cost of the robust design cost of C (which is chosen from the expected cost curve without knowing what the true conditions are), can be thought of as both an added cost of having uncertainty and the expected possible benefit of reducing uncertainty to zero. This is the additional possible savings that could be achieved by improving the certainty of the model through data collection.
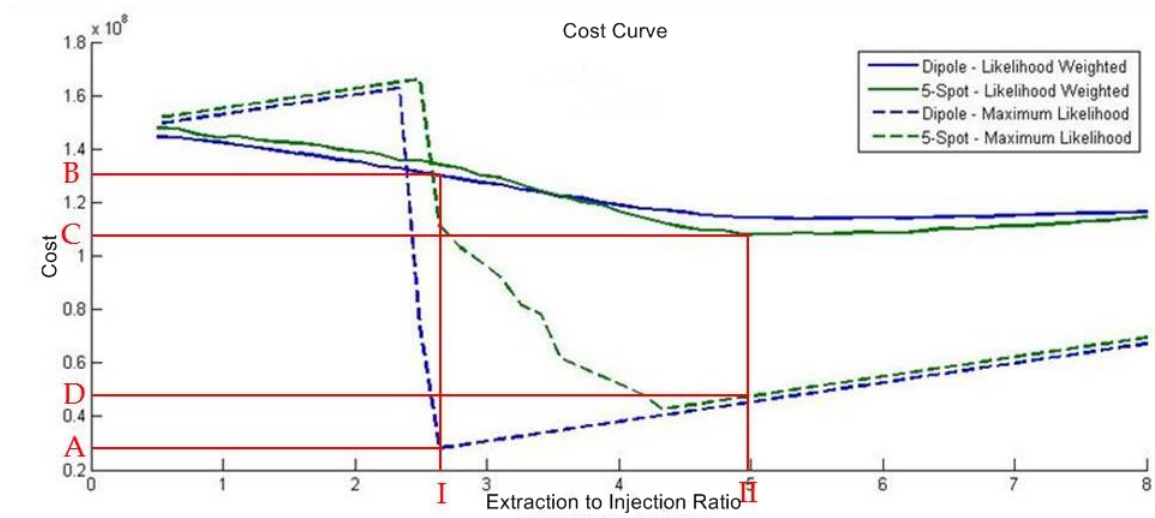


FIGURE 4.15. Costs Curves with Notable Design Decisions and Associated Costs

In this study, uncertainty in the regional model was transferred to the advective transport model through the magnitude and direction of the background flow field.

As described in the methods sections, the impact of each of these two sources of uncertainty were explored by creating synthetic likelihood functions for each component of the uniform background flow, perturbing the variance in these functions by different degrees, and calculating the added cost of ignoring uncertainty for each synthetic likelihood realization. Figure 4.16 below is a contour plot of these results, normalized by the minimum possible cost that could be achieved if true conditions were perfectly represented by one of the choices of background flow conditions in the ensemble. For these calculations, the background flow conditions that were selected to represent true conditions are a flow magnitude of 0.1 and angle of 10 degrees. These were artificial conditions chosen because they provided a clear demonstration of the impact that uncertainty can have on costs. These conditions are selected as true for this exercise only. The results of the flow model suggest a different mean angle and magnitude. It is not believed that the actual flow system is completely uniform over such a large area or that it was perfectly represented by any of the members of the regional flow model ensemble. In the context of the costs shown on figure 4.15 above, the calculation being performed is: (B-A)/A. The horizontal axis marks the standard deviation in the synthetic likelihood function of the flow fields magnitude, and the vertical axis marks the same for the direction. Costs rise in both directions, and are greater for the highest uncertainties in the direction that were considered than they were for the highest uncertainties in magnitude that were considered. Costs always rise with increases in direction uncertainty, but they do not increase monotonically with magnitude when uncertainties are high in both. This is likely due to the flattening of the expected cost curve that occurs with high uncertainty as was discussed in the cost section and shown on figure 4.14 and figure 4.15.

It is important to note that these results are for the particular background conditions chosen (magnitude of 0.1 and angle of 10 degrees), for which the selected design (I) is the dipole. A different set of true background conditions with a different least cost design would generate a different set of results. If, for example, the background

conditions chosen had indicated that the five-spot would be a better design, the features of figure 4.16 would likely show less sensitivity to the uncertainty in the angle of the background gradient. This is because the five-spot is a more robust design under the full range of possible flow directions.
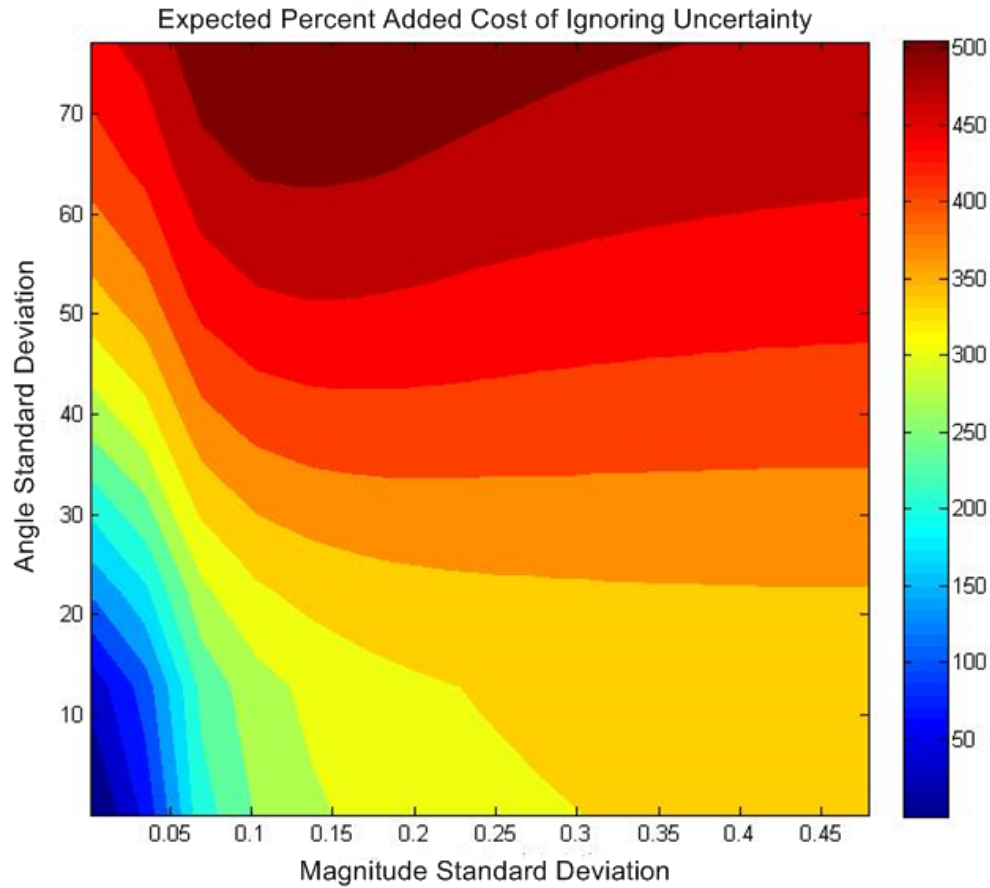


FIGURE 4.16. A Contour Plot of the Cost Increase of Ignoring Uncertainty Normalized by The Minimum Possible Costs with No Uncertainty. Results are plotted against the uncertainty in the magnitude and angle of the background flow field.

## 4.7 Bias analysis

The above exercise explored a sythetic example of a model ensemble for which the available data could only provide an imprecise identification of the best model of the

physical system. Though uncertain, it was assumed that the likelihood distribution was centered around the true or best model; that there was, in other words, no bias in the parameter and model identification. The presence of bias in the likelihood distribution is a possibility that, like uncertainty, can lead to the misidentification of the minimum expected cost design and to an added actual cost of implementation.

The effects of bias upon design costs were explored using a similar approach to that employed in the uncertainty analysis above. In this case, however, the synthetic likelihood functions of the background flow fields direction and magnitude were given fixed levels of uncertainty and it was the mean values of these functions that were perturbed to generate different levels of bias in each component. A single likelihood distribution was identified as the unbiased case, and every other one was deemed biased.The uncertainty was fixed and designs were selected by choosing the robust design from the minimum of the expected cost curves.

Figure 4.17 shows two sets of cost curves: one set generated from the unbiased likelihood distribution and one set generated from one of the biased distribution. The solid curves represent unbiased and the dashed curves represent biased. The maximum likelihood model curves are not used in this analysis and are not shown. The expected additional cost created by bias was calculated by identifying the least cost design (point I) from the biased curves, finding the expected cost of this design for the unbiased curves (point A) and comparing it to the mimimum expected cost of the unbiased curves (point B).
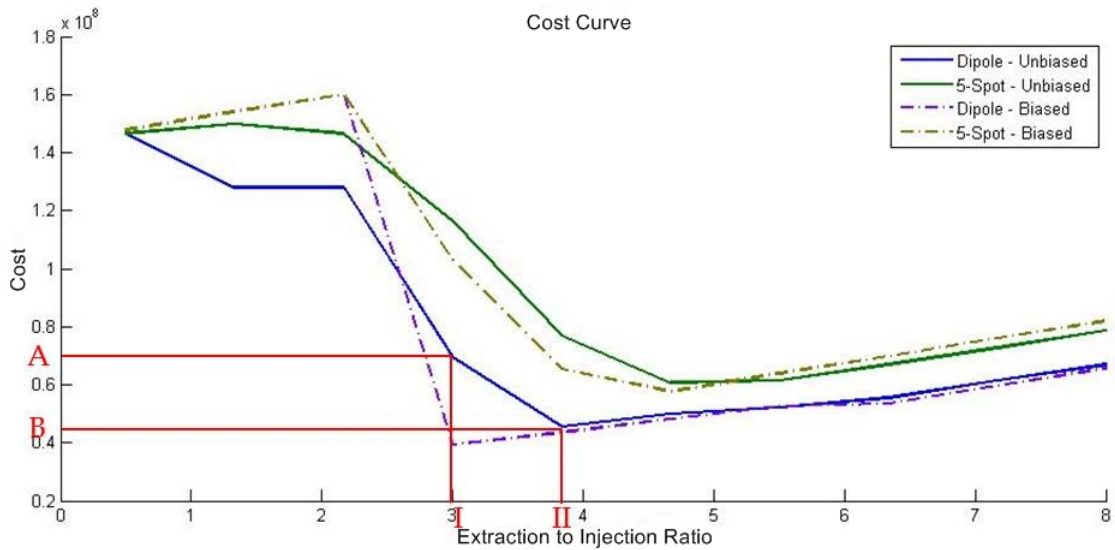
FIGURE 4.17. Biased and Unbiased Cost Curves with Notable Design Decisions and Associated Costs

This calculation was done for every artificially biased likelihood distribution that was generated by perturbing the mean of the flow field magnitude and direction likelihood distributions (individually and in combination) by different degrees. Figure 4.18 below is a contour plot of these results, normalized by the minimum expected cost with no bias, and expressed as a percent. The horizontal axis marks the perturbation in the mean of the synthetic likelihood function of the flow fields magnitude, and the vertical axis marks the same for the direction. The unbiased case is located in the center of the horizontal axis and the bottom of the vertical axis, and can be seen to have an added cost of zero.
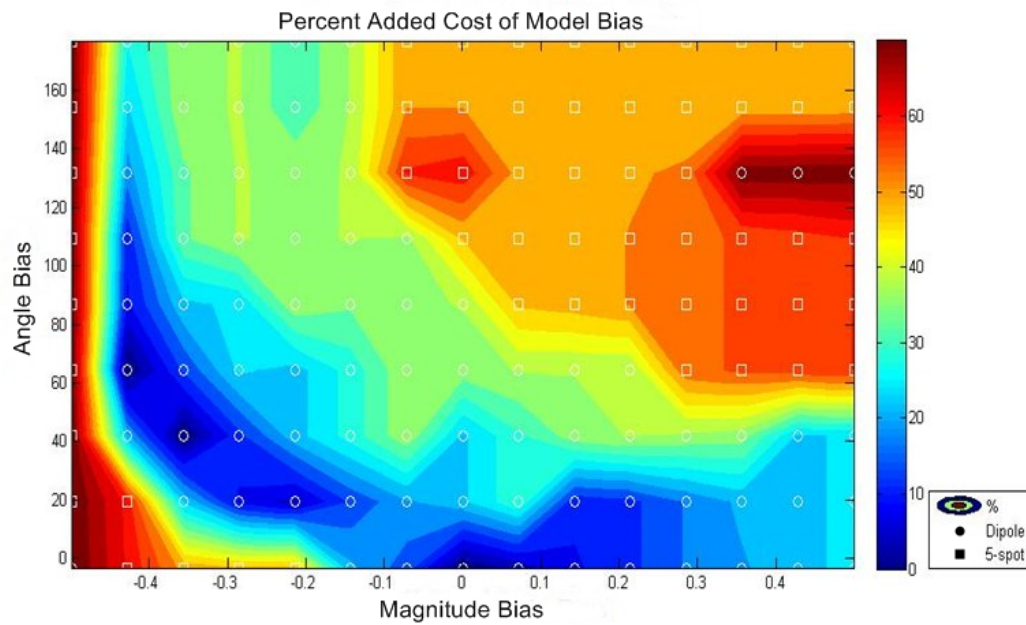
FIGURE 4.18. A Contour Map of the Cost Increase of Model Bias Normalized by The Expected Cost of the Unbiased Model. Results are plotted against the bias in the magnitude and angle of the background flow field.

The pattern and overall magnitude of the costs below are dependent upon the levels of uncertainty set in the synthetic likelihood functions. For the case above, the uncertainties were high and robust designs were chosen most of the time. As a result, the added costs that were calculated did not become very great as the robust designs performance is much less sensitive to differences in the background from conditions. If the uncertainty were lessened, more efficient designs would have been selected with a greater vulnerability to biased results, and much greater variation in costs would have been seen. It is generally expected that the greatest added costs will be experienced under low levels of uncertainty and high levels of bias. No analysis was undertaken to identify the degree of bias in the regional flow model results.

These costs differ in several ways from the expected additional costs of ignoring uncertainty that were shown in figure 4.16 above. Because of these differences figure 4.16 and figure 4.18 should not be used for direct comparison. In each case, the figure

is generated to explore how the effect of bias or uncertainty in the degree and angle of the background flow compare to one another and to see how they interact. With true cost curves, these and similar plots may be generated to diagnose how the model may be improved, or to decide how comfortable one is with the levels of uncertainty and bias in their cost estimates.

CHAPTER 5

CONCLUSION

The Discrimination/Inference to Reduce Expected Cost Technique (DIRECT) is a new modeling framework designed to allow risk-based decision making and monitoring network design optimization to be included in modeling studies that are beset by prediction uncertainty. We applied aspects of DIRECT to a set of regional and local groundwater models with a focus on the how uncertainty is transferred from the regional scale to the local scale and the impact that this uncertainty has on the design of a system of injection and extraction wells. Throughout the study, the utility of DIRECT was considered by exploring the unique set of information it provides and by comparing its modeling approach to the traditional single-model approach.

Two important aspects of DIRECT were explored in this study. First, model predictions were made using multiple models from an ensemble with varying conceptual models and many parameter realizations. The summarized results of the ensemble were compared to the results of the single model that was identified to produce the most accurate results given the observation data. Second, an economic cost function was applied to modelled predictions of interest and the potential circulation system designs were compared based on their predicted costs.

## 5.1  Typical approach vs DIRECT

The inclusion of multiple models in the analysis of potential circulation system designs was used to compare decisions made using two different approaches: (1) the typical approach of selecting a design based upon the single model that best matches observed conditions; and (2) an approach that considers the results from an entire ensemble of model results based upon their relative likelihood (as in DIRECT). The comparison

of these two approaches revealed a potential for reaching widely different conclusions.

It was demonstrated that the typical approach may select a design that performs very poorly if conditions differ from those predicted by the model. Because this approach draws conclusions based on only one model, the potential that the conditions set by the modeler differ from those in the actual system can produce inaccurate results. The more robust DIRECT guards against this danger by using many different models to capture the inherent uncertainty in the system. By expanding the range of the ensemble of models, we have a greater chance of capturing both the reality of the situation and uncovering possibilities that may be less-likely given the data, but ultimately more important in terms of the possibly costly results they predict. However, in using a summary of many results, the design chosen with DIRECT may perform sub-optimally compared to a design selected for known actual conditions. Because we do not have the means to perfectly predict hydrologic system responses, DIRECT is generally a better way of selecting designs under uncertainty.

Both of these points suggest the importance of reducing uncertainty by collecting additional data. This will lessen the chance that the typical approach selects a poorly performing model while also allowing the robust design chosen with DIRECT to perform better under true conditions.

## 5.2 Importance of economic cost functions

Including economic costs in the modeling process is a critical aspect of DIRECT that enriches the information that models can provide for environmental decision making. This study included several analyses that took advantage of the additional layer of information that the cost function revealed.

Using an economic cost function condenses many important design considerations into a common metric that can be used to compare potential actions. A cost function can be difficult to define, but it offers the modeler an opportunity to identify the role

the model ultimately has in decision making. In this study, the economic cost function for the potential circulation systems included terms for total pumping and disposal, well installation, and a penalty for the loss of the injected solution. The form of the function itself was artificial, but the primary considerations were realistic concerns. Furthermore, the interactions among the different factors was clearly revealed by the cost curves, which varied as the designs were changed to increase pumping, add wells, and capture more solution. These simple results offered utility in transferring the hydrological results into a form more directly useful for decision making.

## 5.3   Uncertainty

The two aspects of DIRECT that were applied in this study  modeling with an ensemble and including economic costs  provided a flexible way to explore predictive uncertainty and how it impacts decisions and costs. The probabilistic nature of the ensemble provided a way to introduce structural and parameter uncertainty into the analysis and to estimate prediction uncertainty. The inclusion of the cost function allowed us to explore not just where uncertainty exists in model predictions, but how these uncertainties affect the relative and absolute costs of designs.

When testing potential actions using a model, DIRECT advocates choosing a design that minimizes expected costs. While the artificial nature of the cost function used in this study prevented a specific best design from being selected for the real system, the framework allowed for a useful comparison and demonstration of the ways in which prediction uncertainty can add to the cost of a design. The most useful result from this study for demonstrating uncertainty costs was the cost curves. These show the cost of every design option as they are predicted by the typical single-model approach and by the expected costs summary of the entire ensemble. The figures clearly demonstrate how (1) the designs and costs differ when using the typical single model approach versus DIRECT; (2) the expected cost of an efficient

design chosen with the single model can be higher than is predicted by the single model; (3) the lowest expected cost can be higher than the lowest cost predicted by the single model; (4) the lowest expected cost design may perform sub-optimally under actual conditions; and (5) the designs and costs may all converge upon the true optimal design as uncertainty is reduced if the ensemble includes an accurate prediction of the actual system. If made with real cost functions, these comparisons would allow a modeler to explicitly view the impact that uncertainty can have on decisions, and to decide whether or how much to compromise a designs efficiency to include robustness.

The final aspect of DIRECT - that of defining the worth of new data and devising data acquisition strategies - was not included in this study. However, the technique of transferring uncertainty from the regional flow model to the local transport model through simplified statistical models provided an alternate way of addressing the question of data worth. Combining these simple models with cost curves provided a flexible tool for modeling the relationship between uncertainty in model inputs and costs. The example analysis included in this study estimated the additional costs that could be expected to beset the efficient designs when different degrees of uncertainty are present in model inputs. This result could be used to aid data collection strategies.

Ultimately, the analyses performed in this study provide a glimpse into level of information that can be provided with DIRECT and the information that is lost when ensemble methods and cost consideration are omitted from a modeling project.

## 5.4 Practical Considerations

It should be emphasized that the expected cost of a specific design refers to a prob-abilistic expected value. This is the average cost that would be achieved if cost is considered to be a random variable and the design were implemented innumerable times. Importantly, the expected cost is not the cost that one should expect to occur

for one implementation of a design. As a result, implementation of DIRECT offers no guarantee of making a better decision for any single application. Rather, it is aimed at using all available information, quantitatively, to balance a design to account for both the more likely, but less costly, outcome and rare but potentially disastrous conditions or events.

The possibility that risk-based decision making may choose designs that are frequently found to be unnecessarily cautious may make risk-based decision making a difficult strategy to defend for fields in which it is not already accepted. Not only will one be citing a cost (expected cost) that will probably not be experienced, but under most situations one will also be choosing a design that underperforms the more straight forward maximum likelihood design. It is in the larger view that this design strategy can be seen to have its advantage. Countless decisions are being made every day without full knowledge of what will occur. Without a consideration of risks many of these decisions will fail and be found to have recklessly ignored low probability but highly undesirable outcomes. With such a large number of decisions being made in every discipline, risk-based decision making is the strategy that will produce the best average outcomes overall. While it may be difficult for any individual to see the incentive to spearhead this change of thinking, the overall advantage it offers is difficult to dispute. This study touched upon the benefit of risk based decision making for contaminant treatment design, which is an application of subsurface hydrology where the strategy may be particularly beneficial. Specifically, the extra price of a robust design may be easy to justify to decision makers who understand the extremely high cost that can result from a release of contaminants into the environment.

APPENDIX A

# ADVECTIVE TRANSPORT MODEL MATLAB CODE

```
% MainAdvectiveTransport
   ====================================================
% Code for using the Theis solution and particle tracking to
   evaluate the
% performance of different well orientations under different
   hydrogeological
% conditions

%% Initialization

% USER INPUTS
   ----------------------------------------------------------------
O.Loop1=0;                % Hydrologic model loop
if O.Loop1
    clear all
    close all
    O.Loop1=1;
    D.fig=0;
end
O.Loop2=0;                % Grouping, summarizing, and likelihood loop
O.Initialize=0;

if O.Initialize;
    % Define grid and solution options
    O.DistCrit0 = 2.5;       % (cell) distance away from wells in
        particle 'capture' zone
    O.LMax = 1.25*5280;      % (ft) define size of domain
    O.LRes = 1320/4;         % (ft) cell length
    O.PartDist = 1;          % (cell) distance from injection well to
        place particles
    O.PartNum = 61;          % number of particles to place
    O.ModelNumber=0;         % choose the model numbers, 0 is all, if
        higher than number of models it is final
    O.ModelNumberL=0;        % choose a Likelihood model number, 0 is
        for all
    O.GroupBy={'Scenario','QRat'};

    % Select which sets of figures to produce
    O.Plot1=0;                % head contours with streamlines
    O.Plot2=0;                % pumping effiency contour maps
    O.Plot3=1;                % Qrat vs. pumping effiency & cost
```

```
O.Plot4=0;                    % Bias analysis plots
O.Plot5=0;                    % Variance analysis plots
O.Plot6=0;                    % Variance analysis 'cost of ignoring
   uncertainty'

% Define costs and cost function
P.CostW=0.05e5;                  % Cost of installing wells
P.CostP=160;                  % Cost of pumping and injecting water
P.CostC=200;                  % Cost of 'contamination' (percent
   boundary)
Cost1 = inline('CostWell.*NumWell+CostCont.*(QInj.*PerLost)+
   CostPump.*(QInj+QExt)');

% Define parameters and input variables
P.Time = 100;                         % (yr)
P.Q = 125*365;%500*365;                % (ft3 yr-1)
P.Scenario = [2 4]';                  % Well scenario (defined below)
P.QRat = linspace(1/2,8,10)';%linspace(1/2,10,40)';   % Total
   extraction to injection ratio
P.K = 10^0;                           % (ft yr-1)
P.Ss = 10^-5;                         % (ft-1)
P.GradAng = linspace(0,pi,5)';  % Direction of gradient
P.GradMag = linspace(10^-3,1,10)';%linspace(10^-3,1,30)';% (ft
   ft-1) Magnitude of gradient
P.Thick = 100;                 % (ft)

% Define parameter probability distribution parameters (if
   uniform, don't define)
P.Prob=[];

P.Prob(end+1).Name='GradAng';
P.Prob(end).Distribution='Normal';
P.Prob(end).mu=P.GradAng(end);%P.GradAng;
P.Prob(end).sig=linspace(1e-4,pi/2,5)';%linspace(1e-4,pi/2,15)';
P.Prob(end).true.mu=P.GradAng(end);
P.Prob(end).true.sig=pi/8;

P.Prob(end+1).Name='GradMag';
P.Prob(end).Distribution='Normal';
P.Prob(end).mu=median(P.GradMag);%P.GradMag;
P.Prob(end).sig=linspace(1e-2,1.5*std(P.GradMag),5)';%linspace(1
   e-3,1.5*std(P.GradMag),15)';
P.Prob(end).true.mu=median(P.GradMag);
P.Prob(end).true.sig=0.005;

% Define the well scenarios
O.Scenario(1).Name='Single Pumping Well';
O.Scenario(1).WellType=[-1];
O.Scenario(1).WellLoc=[0 0];     %(ft)
```

```
    O.Scenario (2).Name='Dipole ';
    O.Scenario (2).WellType =[-1;1];
    O.Scenario (2).WellLoc=sqrt (2)*[0 0;0.25 0]*5280;     %(ft)
    O.Scenario (3).Name='5-spot ';
    O.Scenario (3).WellType =[-1;1;1;1;1];
    O.Scenario (3).WellLoc=[0 0;-.25 -.25;-.25 .25;.25 -.25;.25
       .25]*5280;     %(ft)
    O.Scenario (4).Name='5-spot shifted pi/4';
    O.Scenario (4).WellType =[-1;1;1;1;1];
    O.Scenario (4).WellLoc=sqrt (2)*[0 0;-.25 0;0 .25;.0 -.25;.25
       0]*5280;     %(ft)

    % CREATE VARIABLES BEFORE MAIN LOOP
       ------------------------------------------

    % Initialize variables
    D.PartLoc=O.LRes*O.PartDist*[cos(linspace(0,2*pi,O.PartNum))',
       sin(linspace(0,2*pi,O.PartNum))'];
    D.DistCrit0=O.LRes*O.DistCrit0;
    D.k0=0;

    % Create the model ensemble
    M.All=allcomb(P.K,P.Ss,P.Q,P.QRat,P.GradAng,P.GradMag,P.Scenario
       ,P.Time,P.Thick);
    M.AllNames ={'K','Ss','Q','QRat','GradAng','GradMag','Scenario','
       Time','Thick'};
    for i0=1:size(M.AllNames ,2)
        eval(['M.',M.AllNames{i0},'=M.All(:,',num2str(i0),');']);
    end

    % Post processing of model ensemble
    %SubPostProcessing;

    % Use O.Model number to define a single model to run, or run all
        if = 0
    if O.ModelNumber ==0
        D.Models =(1:size(M.All,1))';
    elseif O.ModelNumber >size(M.All,1)
        D.Models=size(M.All,1);
    else
        D.Models=O.ModelNumber;
    end
    D.DispFreq=max (1,round(length(D.Models)/25));

end %if O.Initialize
%% Calculation
tic
% LOOP OVER THE MODEL ENSEMBLE
    --------------------------------------------------
```

```
if O.Loop1

    % Initialize other variables
    M.NExt=zeros(size(M.All,1),1);
    M.NInj=zeros(size(M.All,1),1);
    R.PerBound=zeros(size(M.All,1),1);
    R.PerWells=zeros(size(M.All,1),1);
    % Find largest number of wells from O.Scenario for determining
        the # of
    % columns for R.PerWellsInd
    D.A0=0;
    for j0=P.Scenario'
        D.A0=max(D.A0,size(O.Scenario(j0).WellLoc,1)-1);
    end
    R.PerWellsInd=zeros(size(M.All,1),D.A0);

    % BEGIN LOOP --------------
    for i0 = D.Models'
        D.k0=D.k0+1;

        % Define the Scenario
        D.WellType=O.Scenario(M.Scenario(i0)).WellType;
        D.A0=sum(D.WellType==-1); %Number of injection wells in
            scenario
        D.A1=max(sum(D.WellType==1),1); %Number of extraction wells
            in scenario
        D.QRat=-1*(D.WellType==-1)+(D.WellType==1)*M.QRat(i0)*D.A0/D
            .A1;
        D.Q=M.Q(i0)*D.QRat;
        D.WellLoc=O.Scenario(M.Scenario(i0)).WellLoc;
        D.ExtWellLoc=D.WellLoc;
        D.ExtWellLoc(1,:)=[];

        M.NExt(i0)=D.A1;
        M.NInj(i0)=D.A0;

        % Create the grid
        D.A0=ceil(O.LMax+max(abs(D.WellLoc(:))));
        D.x=-D.A0:O.LRes:D.A0;D.x=sort(unique([D.x,D.WellLoc(:,1)'])
            );
        D.y=(D.A0:-O.LRes:-D.A0)';D.y=flipud(sort(unique([D.y;D.
            WellLoc(:,2)])));
        [D.X1,D.Y1]=meshgrid(D.x,D.y);
        D.rr=zeros([size(D.X1),size(D.WellLoc,1)]);
        D.QQ=zeros([size(D.X1),size(D.WellLoc,1)]);
        % Put some of the Theis parameters on spatial grid
        for j0=1:size(D.WellLoc,1);
            D.rr(:,:,j0)=sqrt((D.X1-D.WellLoc(j0,1)).^2+(D.Y1-D.
                WellLoc(j0,2)).^2);
```

```
      D.QQ(:,:,j0)=D.Q(j0)*ones(size(D.X1));
end

% Create background head distribution
Sub3DPlane=inline('m*cos(angle)*X+m*sin(angle)*Y');
D.InitialHead = Sub3DPlane(D.X1,D.Y1,M.GradAng(i0),M.GradMag
    (i0));

% Calculate drawdown with Theis
D.Drawdown = TheisSolution(D.QQ,M.Time(i0),D.rr,M.Ss(i0),M.K
    (i0),M.Thick(i0));
% Eliminate infinite values at wells
D.Drawdown(D.Drawdown==Inf)=max(D.Drawdown(D.Drawdown~=Inf))
    ;
D.Drawdown(D.Drawdown==-Inf)=min(D.Drawdown(D.Drawdown~=-Inf
    ));

% Sum drawdown and background head distribution
D.FinalHead = D.InitialHead-D.Drawdown;

% Calculate gradient
[D.GradX D.GradY]=gradient(D.FinalHead,D.x,D.y);

% Caclulate streamlines placed around injection well
D.Streamlines=stream2(D.x,D.y,-D.GradX,-D.GradY,D.PartLoc
    (:,1),D.PartLoc(:,2));

% Collect the streamline destinations, classify as well or
    boundary, and calculate percentages

% Loop over each streamline
for j0=1:size(D.Streamlines,2)
    % Collects the destination by extracting the last row
        from D.Streamlines that is
    % not NaN
    D.Dest(j0,:)=(D.Streamlines{j0}(find(isnan(sum(D.
        Streamlines{j0},2))==0,1,'last'),:));
    % Assign a code for destination
    % Default is to remain at initial position (code:-1)
    D.Code(j0)=-1;
    % Well
    for j1=1:size(D.ExtWellLoc,1)
        if sum((abs(D.Dest(j0,:)-D.ExtWellLoc(j1,:))<D.
            DistCrit0))==2
             D.Code(j0)=j1;
        end
    end
    % Boundary
    if max((abs(D.Dest(j0,:)-[max(D.x) max(D.y)])<D.
```

```
                    DistCrit0)+
(abs(D.Dest(j0,:)-[min(D.x) min(D.y)])<D.DistCrit0))>0
                    D.Code(j0)=0;
            end
        end %j0

        % Calculate percentage of particles ending at different
            locations
        R.PerBound(i0)=100*sum(D.Code==0)/sum(D.Code~=-1);
        R.PerWells(i0)=100*sum(D.Code>0)/sum(D.Code~=-1);
        for j0=1:size(D.ExtWellLoc,1)
            R.PerWellsInd(i0,j0)=100*sum(D.Code==j0)/sum(D.Code~=-1)
                ;
        end

        % Save any other desired variables to the the results
            structure R

        % Display the time and the percent completed
        if rem(D.k0,D.DispFreq)==0
            display([num2str(100*D.k0/length(D.Models)),'% complete
                '])
            toc
        end
    end %i0
end %if O.Loop1
%END LOOP --------------------------------

% Calculate Cost of models
R.Cost1=Cost1(P.CostC,P.CostP,P.CostW,(M.NInj+M.NExt),R.PerBound,M.Q
    .*M.NInj.*
M.QRat,M.NInj.*M.Q);
R.Cost1Wells=Cost1(0,0,P.CostW,(M.NInj+M.NExt),R.PerBound,M.Q.*M.
    NInj.*M.QRat,M.NInj.*M.Q);
R.Cost1Pump=Cost1(0,P.CostP,0,(M.NInj+M.NExt),R.PerBound,M.Q.*M.NInj
    .*M.QRat,M.NInj.*M.Q);
R.Cost1Cont=Cost1(P.CostC,0,0,(M.NInj+M.NExt),R.PerBound,M.Q.*M.NInj
    .*M.QRat,M.NInj.*M.Q);

% GROUP RESULTS BY O.GroupBy AND ANALYZE RESULTS OF GROUPS
    -----------------
% Create fields in structure that group() will be looking for (that
    function could be improved)
D.all=M.All; D.all_names=M.AllNames;

% Find groups
R.Groups = group(D,O.GroupBy,[]);

% CREATE LIKELIHOOD "ENSEMBLE", LOOP OVER ENSEMBLE AND COLLECT
```

```
    RSummary
% ARRAYS
% Create the likelihood ensemble
D.LikeParTypes=fields(P.Prob(1).true);
P.Like.All={};
D.k0=0;
for i0=1:size(D.LikeParTypes,1)
    P.Like.Trues{i0}=[];
    P.Like.TruesType{i0}=D.LikeParTypes{i0};
    for i1=1:size(P.Prob,2)
        D.k0=D.k0+1;
        P.Like.Name{D.k0}=[D.LikeParTypes{i0},P.Prob(i1).Name];
        P.Like.TruesName{i0}{i1}=P.Like.Name{D.k0};
        P.Like.Type{D.k0}=D.LikeParTypes{i0};
        P.Like.ParName{D.k0}=P.Prob(i1).Name;
        % Get full vectors and true values of likelihood model
            parameters
        eval(['D.LikeArray=P.Prob(i1).',P.Like.Type{D.k0},';'])
        eval(['D.LikeTrue=P.Prob(i1).true.',P.Like.Type{D.k0},';'])
        % Change them to column vectors if they aren't
        if size(D.LikeArray,1)==1
            D.LikeArray=D.LikeArray';
        end
        if size(D.LikeTrue,1)==1
            D.LikeTrue=D.LikeTrue';
        end
        % If the true value is empty, make the true value equal to
            the full
        % vector
        if isempty(D.LikeTrue)
            D.LikeTrue=D.LikeArray;
        else
            D.LikeArray=unique([D.LikeTrue;D.LikeArray]);
        end
        % Make true values set for each parameter type
        P.Like.Trues{i0}=RowCombinations(P.Like.Trues{i0},D.LikeTrue
            );
        % Make an array of all likelihood parameter values made up
            of all
        % combinations of the parameter vectors, will be trimmed
            later
        P.Like.All=RowCombinations(P.Like.All,D.LikeArray);
    end
end
D.A0=zeros(size(P.Like.All,1),1);
D.c1=0;
% Filter P.Like.All to only include those that have one of the 'true
    ' sets
% of each likelihood parameter type
```

```matlab
for i0=1:size(P.Like.Trues,2)
    D.c0=D.c1+1;
    D.c1=D.c0+size(P.Like.Trues{i0},2)-1;
    D.A0=D.A0+ismember(P.Like.All(:,D.c0:D.c1),P.Like.Trues{i0},'
        rows');
end
P.Like.All=P.Like.All((D.A0>0),:);

% LOOP over likelihood "ensemble"
D.DispFreq=max(1,round(size(P.Like.All,1)/25));
D.k0=0;
tic
D.LikeLoop=1:size(P.Like.All,1);
if O.ModelNumberL~=0
    D.LikeLoop=D.LikeLoop(O.ModelNumberL);
end
if O.Loop2
    display('Likelihood Loop')
    for i0=D.LikeLoop
        D.k0=D.k0+1;
        % Calculate model likelihoods for current likelihood "model"
        D.A0=ones(size(M.All,1),1);
        for i1=1:size(P.Prob,2)
            D.Type=P.Prob(i1).Distribution;
            eval(['D.A1=M.',P.Prob(i1).Name,';']);
            switch D.Type
                case 'Normal'
                    D.c0=find(ismember(P.Like.ParName,P.Prob(i1).
                        Name).*ismember(P.Like.Type,'mu'));
                    D.mu=P.Like.All(i0,D.c0);
                    D.c0=find(ismember(P.Like.ParName,P.Prob(i1).
                        Name).*ismember(P.Like.Type,'sig'));
                    D.sig=P.Like.All(i0,D.c0);
                    D.A0=D.A0.*normpdf(D.A1,D.mu,D.sig);
            end
        end %i1
        R.Likelihood=D.A0/sum(D.A0);

        % CREATE STRUCTURE RGrouped TO HOLD GROUPED RESULTS AND
            MODEL INFORMATION--
        RGrouped = SubRGrouped(R,M,O.GroupBy);

        % CREATE STRUCTRE RSummary to combine results of the
            different pumping
        % arrangements
        D.ScenNum=unique(M.Scenario);
        D.Fields={'QRat','LWPerWells','StdLWPerWells','LWPerBound','
            StdLWPerBound'...
             ,'LWCost1','StdLWCost1','BestFitCost1','BestFitPerWells
```

```
                       '};
            RSummary = SubRSummary(RGrouped,D.Fields,D.ScenNum,0);


            % COLLECT SUMMARY DATA in --- structure
            for i1=1:size(RSummary,2)
                RSummaryLikeEns(i1).ScenarioName=RSummary(i1).
                    ScenarioName;
                RSummaryLikeEns(i1).Scenario=RSummary(i1).Scenario;
                for i2=1:size(P.Like.All,2)
                    eval(['RSummaryLikeEns(i1).',P.Like.Name{i2},'=P.
                        Like.All(:,i2);'])
                end
                D.fields=fields(RSummary);
                D.fields(ismember(D.fields,{'ScenarioName','Scenario'}))
                    =[];
                for i2=1:size(D.fields,1)
                    if i0==1
                         eval(['RSummaryLikeEns(i1).',D.fields{i2},'=
zeros(size(P.Like.All,1),length(RSummary(i1).QRat(:)));'])
                    end
                    eval(['RSummaryLikeEns(i1).',D.fields{i2},'(i0,:)=
                        RSummary(i1).',D.fields{i2},';'])
                end
                RSummaryLikeEns(i1).MinLWCost1=min(RSummaryLikeEns(i1).
                    LWCost1,[],2);
                D.A0=RSummaryLikeEns(i1).MinLWCost1*ones(1,size(
                    RSummaryLikeEns(i1).LWCost1,2));
                D.A1=(D.A0==RSummaryLikeEns(i1).LWCost1);
                RSummaryLikeEns(i1).QRatMinLWCost1=sum(RSummaryLikeEns(
                    i1).QRat.*D.A1,2);
            end %i1
            % Display the time and the percent completed
            if rem(D.k0,D.DispFreq)==0
                display([num2str(100*D.k0/size(P.Like.All,1)),'%
                    complete'])
                toc
            end
        end

    end

    for i2=D.LikeLoop
        D.MaxL=zeros(size(RSummary,2),size(RSummary(1).QRat,1));
        %CHeck
        %[r,c]=find(D.MaxL==min(D.MaxL(:)),1,'first');
        D.LW=zeros(size(RSummary,2),size(RSummary(1).QRat,1));
        for i1=1:size(RSummary,2)
            D.MaxL(i1,:)=RSummaryLikeEns(i1).BestFitCost1(i2,:);
            D.LW(i1,:)=RSummaryLikeEns(i1).LWCost1(i2,:);
```

```
            end
        D.A0=min(D.LW(:));
        if i2==D.LikeLoop(1)
            D.overallLWMin=D.A0;
        elseif D.A0<D.overallLWMin
            D.overallLWMin=D.A0;
        end
    end

    for i2=D.LikeLoop
        % Get cost of ignoring uncertainty
        D.MaxL=zeros(size(RSummary,2),size(RSummary(1).QRat,1));
        %CHeck
        %[r,c]=find(D.MaxL==min(D.MaxL(:)),1,'first');
        D.LW=zeros(size(RSummary,2),size(RSummary(1).QRat,1));
        for i1=1:size(RSummary,2)
            D.MaxL(i1,:)=RSummaryLikeEns(i1).BestFitCost1(i2,:);
            D.LW(i1,:)=RSummaryLikeEns(i1).LWCost1(i2,:);
        end
        [D.r,D.c]=find(D.MaxL==min(D.MaxL(:)),1,'first');
        [D.r1,D.c1]=find(D.LW==min(D.LW(:)),1,'first');
        for i1=1:size(RSummaryLikeEns,2)
            RSummaryLikeEns(i1).DetermCost1(i2,1)=D.LW(D.r,D.c)-min(
                D.LW(:));
            RSummaryLikeEns(i1).PerDetermCost1(i2,1)=(D.LW(D.r,D.c)-
                min(D.LW(:)))./D.overallLWMin;
            RSummaryLikeEns(i1).PerDetermCost1a(i2,1)=(D.LW(D.r,D.c)
                -min(D.LW(:)))./min(D.LW(:));
            RSummaryLikeEns(i1).DetermChoice(i2,1)=D.r;
            RSummaryLikeEns(i1).CostLWTreat2Det(i2,1)=D.MaxL(D.r1,D.
                c1)-min(D.MaxL(:));
            RSummaryLikeEns(i1).PerCostLWTreat2Det(i2,1)=
(D.MaxL(D.r1,D.c1)-min(D.MaxL(:)))./min(D.MaxL(:));
            RSummaryLikeEns(i1).DetermCost2(i2,1)=D.LW(D.r,D.c)-min(
                D.MaxL(:));
            RSummaryLikeEns(i1).PerDetermCost2(i2,1)=(D.LW(D.r,D.c)-
                min(D.MaxL(:)))./min(D.MaxL(:));
            RSummaryLikeEns(i1).PerDetermCost2a(i2,1)=(D.LW(D.r,D.c)
                -min(D.MaxL(:)))./D.overallLWMin;
        end
    end

    if length(D.LikeLoop(:))>1
        % RUN VARIANCE ANALYSIS
        SubVarianceAnalysis;

        % RUN BIAS ANALYSIS
        SubBiasAnalysis;
    end
```

```
end %if O.Loop2
\end{verbatim}

%% Figures


\begin{verbatim}
% SubRGrouped
    ========================================================
function RR = SubRGrouped(R,M,GroupBy)
% Function that takes the results from R, separates them by the
    fields
% specified and calculates new arrays for the grouped results

%% Initialization

% Preallocate RGrouped
RGrouped(size(R.Groups,2)).Group=[];

% Get the field names from M and R and filter out grouped by and non
    -parameter fields
D.MFieldNames=fieldnames(M);
D.RFieldNames=fieldnames(R);
D.A2=ones(size(D.MFieldNames));
D.A3=ones(size(D.RFieldNames));

%% Calculations

% Find those fields that aren't nModels in length
for i1=1:size(D.MFieldNames,1)
    D.A0=D.MFieldNames{i1};
    eval(['D.A1=size(M.',D.A0,',1);']);
    D.A2(i1)=(D.A1==size(M.All,1));
end
for i1=1:size(D.RFieldNames,1)
    D.A0=D.RFieldNames{i1};
    eval(['D.A1=size(R.',D.A0,',1);']);
    D.A3(i1)=(D.A1==size(M.All,1));
end
% Eliminate those that grouping was based upon
D.A2=D.A2-ismember(D.MFieldNames,GroupBy);
% Filter field list
D.MFieldNames=D.MFieldNames(D.A2==1);
D.RFieldNames=D.RFieldNames(D.A3==1);

% Loop through number of groups
for i0=1:size(R.Groups,2)
    % List group members by position in full ensemble in R
    RGrouped(i0).Group=R.Groups{i0};
```

```
% Create fields from GroupBy to values of the groups shared
% characteristics
RGrouped(i0).GroupBy=GroupBy;
for i1=1:size(GroupBy,2)
    D.A0=GroupBy{i1};
    eval(['RGrouped(i0).',D.A0,'=unique(M.',D.A0,'(R.Groups{i0})
        );']);
end %i1

% Get parameter values from fields
RGrouped(i0).AllNames=M.AllNames;
for i1=1:size(D.MFieldNames,1)
    D.A0=D.MFieldNames{i1};
    eval(['RGrouped(i0).M.',D.A0,'=M.',D.A0,'(R.Groups{i0});']);
end
for i1=1:size(D.RFieldNames,1)
    D.A0=D.RFieldNames{i1};
    eval(['RGrouped(i0).R.',D.A0,'=R.',D.A0,'(R.Groups{i0});']);
end

% Re-normalize likelihoods
RGrouped(i0).R.Likelihood=RGrouped(i0).R.Likelihood/sum(RGrouped
    (i0).R.Likelihood);

% Get Likelihood Weighted Results
D.A0=fieldnames(RGrouped(i0).R);
D.A0(ismember(D.A0,'Likelihood'))=[];
for i1=1:size(D.A0,1)
    if isempty(strfind(D.A0{i1},'LW')) && isempty(strfind(D.A0{
        i1},'BestFit'))
        D.A4=RGrouped(i0).R.Likelihood;
        eval(['D.A1=RGrouped(i0).R.',D.A0{i1},';']);
        D.A2=sum(D.A1.*D.A4);
        D.A3=sqrt(sum(D.A4.*(D.A1-mean(D.A1)).^2));
        eval(['RGrouped(i0).R.LW',D.A0{i1},'=D.A2;']);
        eval(['RGrouped(i0).R.StdLW',D.A0{i1},'=D.A3;']);
    end
end %i1

% Get best-fit model results
D.A0=fieldnames(RGrouped(i0).R);
D.A0(ismember(D.A0,'Likelihood'))=[];
for i1=1:size(D.A0,1)
    if isempty(strfind(D.A0{i1},'BestFit')) && isempty(strfind(D
        .A0{i1},'LW'))
        D.A4=find(RGrouped(i0).R.Likelihood==max(RGrouped(i0).R.
            Likelihood));
        eval(['D.A1=RGrouped(i0).R.',D.A0{i1},';']);
```

```
            D.A2=mean(D.A1(D.A4));
            eval(['RGrouped(i0).R.BestFit',D.A0{i1},'=D.A2;']);
        end
    end %i1

    % Get Cost1(LW results)
    D.NInj=unique(M.NInj(RGrouped(i0).Group));
    D.NExt=unique(M.NInj(RGrouped(i0).Group));
    D.PerBound=RGrouped(i0).R.LWPerBound;
    D.Q=unique(M.Q(RGrouped(i0).Group));
    D.QRat=unique(M.QRat(RGrouped(i0).Group));
%       RGrouped(i0).R.Cost1LWResults=Cost1(O.CostC,O.CostP,O.CostW,(D
    .NInj+D.NExt),D.PerBound,D.Q.*D.NInj.*D.QRat,D.NInj.*D.Q);

    % Create matrices of Results (PerWells,PerBound,etc.) versus
        parameter
    % values
    RGrouped(i0).ResMats=ResultsVsParams(RGrouped(i0).R,RGrouped(i0)
        .M,size(RGrouped(i0).Group,1));

end %i0

%% Output

RR = RGrouped;

%SubRSummary==============================================================
function RR = SubRSummary(RGrouped,fields,ScenNum,O)
% Function that summarizes the results from RGrouped

%% Initialization

% Create a table for indexing the scenario numbers from 1 to
    nScenario
numScen=length(ScenNum);
D.ScenNum=[ScenNum,cumsum(ones(numScen,1))];
D.k1=zeros(size(D.ScenNum,1),1);
D.n1=size(RGrouped,2)/size(D.k1,1);

% Preallocate RSummary and fields
RSummary(size(D.k1,1)).ScenarioName=[];
RSummary(size(D.k1,1)).Scenario=0;
for i0=1:size(fields,2)
    eval(['RSummary(size(D.k1,1)).',fields{i0},'=zeros(D.n1,1);']);
end

%% Calculations

for i0=1:size(RGrouped,2)
```

```
    % Get scenario number
    D.A0=RGrouped(i0).Scenario;
    % Get index number of RSummary , based on scenario number , and
        advance
    % D.k1 counter
    D.A1=find(D.ScenNum(:,1)==D.A0);
    D.k1(D.A1)=D.k1(D.A1)+1;
    % Define scenario and name
    RSummary(D.A1).Scenario=D.A0;
    RSummary(D.A1).ScenarioName=O.Scenario(D.A0).Name;
    % Get results
    for i1=1:size(fields,2)
        if strcmp(fields{i1},'QRat')
            eval(['RSummary(D.A1).',fields{i1},'(D.k1(D.A1),1)=
                RGrouped(i0).',fields{i1},';']);
        else
            eval(['RSummary(D.A1).',fields{i1},'(D.k1(D.A1),1)=
                RGrouped(i0).R.',fields{i1},';']);
        end
    end
end %i0

%% Output
RR = RSummary;

%SubBiasAnalysis
    ========================================================
% SubVarianceAnalysis
% A subroutine that uses the capture efficiencies and costs from the
    flow
% models calculated in MainTheisAnalysis , and makes changes in the
% likelihood fields to analyze how natural variability , bias , and
    uncertainty
% affect design decisions and the costs that arise from actions
    taken under
% uncertainty

%% Initialization

% Input Structure
R0=RSummaryLikeEns;

% Specify fields
D.ScenarioFields={'Scenario','ScenarioName'};

%% Calculation

% Create a matrix to be used by group for select groupings
D.all=P.Like.All;
```

```
D.all_names=P.Like.Name;
% Find the fields to group by
D.GroupBy=P.Like.Name(strncmpi('sig',P.Like.Name,2));
% Create cell arrays of different categories of fields in
    RSummaryLikeEns
D.AllFields=fields(R0);
D.MFields=D.AllFields(ismember(D.AllFields,P.Like.Name));
D.RFields=D.AllFields((ismember(D.AllFields,P.Like.Name)+ismember(D.
    AllFields,D.ScenarioFields))==0);
D.RTemp=rmfield(R0,D.ScenarioFields);
% Create a structure of the 'Result' arrays
D.R=rmfield(D.RTemp,D.MFields);
% Create a structure of the likelihood 'model ensemble' arrays
D.M=rmfield(D.RTemp,D.RFields);
% Find groups
D.Groups=group(D,D.GroupBy,'');
% Only keep those with more than one entry
D.Groups=D.Groups(cellfun(@length,D.Groups)>1);

% Loop over the number of elements in R0
for i0 = 1:size(R0,2)
    D.A0=find(strcmp(P.Like.TruesType,'mu'));
    R1(i0).Trues=P.Like.Trues{D.A0};
    R1(i0).TruesName=P.Like.TruesName{D.A0};
    for i1=1:size(D.ScenarioFields,2)
        eval(['R1(i0).', D.ScenarioFields{i1},'=R0(i0).',D.
            ScenarioFields{i1},';'])
    end
    for i1=1:size(D.Groups,2)
        for i2=1:size(D.GroupBy,2)
            eval(['R1(i0).Grouped(i1).',D.GroupBy{i2},'=unique(R0(i0
                ).',D.GroupBy{i2},'(D.Groups{i1}));']);
        end
        R1(i0).Grouped(i1).GroupBy=D.GroupBy;
        R1(i0).Grouped(i1).Group=D.Groups{i1};
        for i2=1:size(D.MFields,1)
            eval(['R1(i0).Grouped(i1).M.',D.MFields{i2},'=R0(i0).',D
                .MFields{i2},'(D.Groups{i1});']);
        end
        for i2=1:size(D.RFields,1)
            eval(['R1(i0).Grouped(i1).R.',D.RFields{i2},'=R0(i0).',D
                .RFields{i2},'(D.Groups{i1},:);']);
        end
        D.QQ=R1(i0).Grouped(i1).R.QRat;
        D.QMin=R1(i0).Grouped(i1).R.QRatMinLWCost1*ones(1,size(D.QQ
            ,2));
        D.True=ones(size(D.QQ,1),1);
        for i2=1:size(R1(i0).Trues,2)
            D.val=R1(i0).Trues(i2);
```

```
                    D.nam=R1(i0).TruesName{i2};
                    eval(['D.A0=R1(i0).Grouped.M.',D.nam,'==D.val;']);
                    D.True=D.True.*D.A0;
            end
            D.TrueCost=ones(size(D.QQ,1),1)*R1(i0).Grouped(i1).R.LWCost1
                (D.True==1,:);
            R1(i0).Grouped(i1).R.TrueCaseLWCost1=min(D.TrueCost(:));
            R1(i0).Grouped(i1).R.ActionCost=sum(D.TrueCost.*(D.QQ==D.
                QMin),2);
            R1(i0).Grouped(i1).R.BiasExtraCost=
-R1(i0).Grouped(i1).R.TrueCaseLWCost1+R1(i0).Grouped(i1).R.
    ActionCost;
            %R1(i0).Grouped(i1).ResMats=
%ResultsVsParams(R1(i0).Grouped(i1).R,R1(i0).Grouped(i1).M,size(D.QQ
    ,1));
        end %i1
end %i0


% Get BIAS cost 2 (the cost of guessing the wrong model (fixed) for
    every
% possible correct model)

for i0=1:size(R1(1).Grouped,2)
    % Find the row of the 'true' or selected model
    D.a0=[R1(1).Grouped(i0).muGradAng,R1(1).Grouped(i0).muGradMag];
    D.a1=[R1(1).Grouped(i0).M.muGradAng,R1(1).Grouped(i0).M.
        muGradMag];
    D.a2=find(sum(ones(size(D.a1,1),1)*D.a0==D.a1,2)==2,1,'first');
    % Get the treatment details (as a row, column location) as the
        least
    % cost from the selected model
    D.A0=zeros(size(R1,2),size(R1(1).Grouped(i0).R.QRat,2));
    for i2=1:size(R1,2)
        D.A0(i2,:)=R1(i2).Grouped(i0).R.LWCost1(D.a2,:);
    end
    [D.r,D.c]=find(D.A0==min(D.A0(:)),1,'first');

    for i1=1:size(R1(1).Grouped(i0).R.QRat,1);
        D.A0=zeros(size(R1,2),size(R1(1).Grouped(i0).R.QRat,2));
        for i2=1:size(R1,2)
            D.A0(i2,:)=R1(i2).Grouped(i0).R.LWCost1(i1,:);
        end
        D.BiasCost(i1,1)=D.A0(D.r,D.c);
    end %i1
    for i2=1:size(R1,2)
        R1(i2).Grouped(i0).R.BiasCost=D.BiasCost;
    end

end %i0
```

```
for i0 = 1:size(R0,2)
    for i1=1:size(D.Groups,2)
        R1(i0).Grouped(i1).ResMats=
ResultsVsParams(R1(i0).Grouped(i1).R,R1(i0).Grouped(i1).M,size(D.QQ
    ,1));
    end %i1
end %i0


%% Output
RBias=R1;

%SubVarianceAnalysis
    ==========================================================
% SubVarianceAnalysis
% A subroutine that uses the capture efficiencies and costs from the
    flow
% models calculated in MainTheisAnalysis, and makes changes in the
% likelihood fields to analyze how natural variability, bias, and
    uncertainty
% affect design decisions and the costs that arise from actions
    taken under
% uncertainty

%% Initialization

% Input Structure
R0=RSummaryLikeEns;

% Specify fields
D.ScenarioFields={'Scenario','ScenarioName'};

%% Calculation

% Create a matrix to be used by group for select groupings
D.all=P.Like.All;
D.all_names=P.Like.Name;
% Find the fields to group by
D.GroupBy=P.Like.Name(strncmpi('mu',P.Like.Name,2));
% Create cell arrays of different categories of fields in
    RSummaryLikeEns
D.AllFields=fields(R0);
D.MFields=D.AllFields(ismember(D.AllFields,P.Like.Name));
D.RFields=D.AllFields((ismember(D.AllFields,P.Like.Name)+
ismember(D.AllFields,D.ScenarioFields))==0);
D.RTemp=rmfield(R0,D.ScenarioFields);
% Create a structure of the 'Result' arrays
D.R=rmfield(D.RTemp,D.MFields);
```

```
% Create a structure of the likelihood 'model ensemble' arrays
D.M=rmfield(D.RTemp,D.RFields);
% Find groups
D.Groups=group(D,D.GroupBy,'');
% Only keep those with more than one entry
D.Groups=D.Groups(cellfun(@length,D.Groups)>1);

% Loop over the number of elements in R0
for i0 = 1:size(R0,2)
    for i1=1:size(D.ScenarioFields,2)
        eval(['R1(i0).', D.ScenarioFields{i1},'=R0(i0).',D.
            ScenarioFields{i1},';'])
    end
    for i1=1:size(D.Groups,2)
        R1(i0).Grouped(i1).Group=D.Groups{i1};
        R1(i0).Grouped(i1).GroupBy=D.GroupBy;
        for i2=1:size(D.GroupBy,2)
            eval(['R1(i0).Grouped(i1).',D.GroupBy{i2},'=unique(R0(i0
                ).',D.GroupBy{i2},'(D.Groups{i1}));']);
        end
        for i2=1:size(D.MFields,1)
            eval(['R1(i0).Grouped(i1).M.',D.MFields{i2},'=R0(i0).',D
                .MFields{i2},'(D.Groups{i1});']);
        end
        for i2=1:size(D.RFields,1)
            eval(['R1(i0).Grouped(i1).R.',D.RFields{i2},'=R0(i0).',D
                .RFields{i2},'(D.Groups{i1},:);']);
        end
    end %i1
    R1(i0).Grouped(i1).ResMats=
ResultsVsParams(R1(i0).Grouped(i1).R,R1(i0).Grouped(i1).M,size(D.
    Groups{i1},1));
end %i0

%% Output
RVariance=R1;
```

# References

[1] Akob, D., Barnhart, E., Orem, B., Clark, A., Ruppert, L., 2013, USGS Energy Program Review: Origin and Controls on Microbial Gas Accumulations (Task 7 Geochemistry of Solid Fuels Project).

[2] Anderman, E.R., Hill, M.C., Poeter, E.P., 1996, Two-Dimensional Advective Transport in Ground-Water Flow Parameter Estimation Groundwater, v. 34, i. 6,p. 10011009.

[3] Anna, L.O., 2003, Groundwater flow associated with coalbed gas production, Ferron Sandstone, east-central Utah.

[4] Bates, B.L., McIntosh, J.C., Lohse, K.A., Brooks, P.D., 2011, Influence of groundwater flowpaths, residence times and nutrients on the extent of microbial methanogenesis in coal beds: Powder River Basin, USA, Chemical Geology, v. 284, i. 12, p. 45-61, Available at URL: http://dx.doi.org/10.1016/j.chemgeo.2011.02.004.

[5] Batu, V., 1998. Aquifer Hydraulics: A Comprehensive Guide to Hydrogeologic Data Analysis, John Wiley & Sons, New York, 727p.

[6] Bauder, K.K., Wheaton, J., 2003, Frequently Asked Questions: Coal Bed Methane (CBM) . MSU Extension Water Quality, Available at URL: http://waterquality.montana.edu/docs/methane/cbmfaq.shtml#whatiscoalbedmethane

[7] Bethke, C.M., and Johnson, T.M., 2002, Paradox of groundwater age. Geology, v.30, p. 385-388.

[8] Chummiao Zheng, 2010, MT3DMS v. 5 i.3 Supplemental Users GuideTechnical Report to the U.S. Army Engineer Research and Development Center, Department of Geological Sciences, University of Alabama, 51 p.

[9] Doelling, H.H., 2002, Interim geologic map of the San Rafael Desert 30'x60' quadrangle, Emery and Grand Counties, Utah, Utah Geological Survey OFR 404.

[10] Doelling, H. H., 2004, Interim geologic map ot the east half of the Salina 30' x 60' quadrangle, Emery, Sevier, and Wayne Counties, Utah, Utah Geological Survey OFR 438.

[11] Doherty, J., 2005, PEST: Model-Independent Parameter Estimation. Watermark Numerical Sompuating.

[12] Domenico, P.A. and Schwartz, F.W., 1990, Physical and Chemical Hydrogeology, John Wiley & Sons, New York, 824 p.

[13] Gelhar, L.W., Welty, C., Rehfeldt, K. R., 1992, A Critical Review of Data on Field-Scale Dispersion in Aquifers, Water Resources Research, v. 28. n. 7. p. 1955-1974.

[14] Goode, D.J., 1996. Direct simulation of groundwater age. Water Resources Research, v. 32, p. 289-296.

[15] Harbaugh, A.W., 2005, MODFLOW-2005, The U.S. Geological Survey modular ground-water model – the Ground-Water Flow Process: U.S. Geological Survey Techniques and Methods 6-A16, variously p.

[16] Heath, R.C., 1983, Basic ground-water hydrology, U.S. Geological Survey Water-Supply Paper 2220, 86p.

[17] McKay, M. D., Beckman, R. J., Conover, W. J., 1979, A Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics, v. 21, no. 2, p. 239-245.

[18] Morris, D.A. and Johnson, A.I., 1967, Summary of hydrologic and physical properties of rock and soil materials as analyzed by the Hydrologic Laboratory of the U.S. Geological Survey, U.S. Geological Survey Water-Supply Paper 1839-D, 42p.

[19] Olsthoorn, T.N., 2011, User guide for mflab, 2011.

[20] Poeter, E.P., Hill, M.C., Banta, E.R., Mehl, S., Christensen, S., 2005, UCODE_2005 and Six Other Computer Codes for Universal Sensitivity Analysis, Calibration, and Uncertainty Evaluation: U.S. Geological Survey Techniques and Methods 6-A11, 283p.

[21] PRISM Climate Group, 2012, PRISM Climate Data. Available at: http://www.prism.oregonstate.edu/ (accessed November 2012).

[22] Randall, K.L., 2009, A Geologic and Hydrochemical Investigation of the Suitability of Central Utahs Navajo Sandstone for the Disposal of Saline Process Water and CO2, M. Sc. Thesis. Utah State University.

[23] Rice, C.A., 2003. Production waters associated with the Ferron coalbed methane fields, central Utah: chemical and isotopic composition and volumes, International Journal of Coal Geology, v. 56, p. 151-169.

[24] Tabet D.E., Hucka, B.P., Sommer, S.N., 1995, Depth vitrinite reflectance and coal thickness maps Ferron Sandstone central Utah, Utah Geological Survey, Open-File Report 329, 3 plates.

[25] Utah Department of Natural Resources, 2011, Utah Oil and Gas Data Research Center. Available at URL: http://oilgas.ogm.utah.gov/Data_Center/DataCenter.cfm

[26] Utah Geological Survey, 2011. Available at URL: http://geology.utah.gov/maps/geomap/30x60/index.htm

[27] U.S. Geological Survey, 2011, National Elevation Dataset (NED) Shaded Relief Direct Download. Available at URL: http://ned.usgs.gov/

[28] U.S. Geological Survey, 2011, Surface-Water Data for the Nation. Available at URL: http://waterdata.usgs.gov/usa/nwis/sw

[29] Vrugt, J.A., Gupta, H.V., Bouten, W., Sorooshian, S.,2003 , A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, Water Resources Research, v. 39, n. 8, p. 1201.

[30] Weiss, M.P., Witkind, I.J., Cashion, W.B., 1990, Geologic map of the Price 30' x 60' quadrangle, Carbon, Duchesne, Uintah, Utah, and Wasatch Counties, Utah: U.S. Geological Survey Map I-1981.

[31] Witkind, I.J., 1988, Geologic map of the Huntington 30' x 60' quadrangle, Carbon, Emery, Grand and Uintah Counties, Utah, U.S. Geological Survey Map I-1764.

[32] Witkind, I.J., Weiss, M.P., 1991, Geologic map of the Nephi 30' x 60' quadrangle, U.S. Geological Survey Map I-1937.

[33] Witkind, I.J.; Weiss M.P.; Brown, T.L., 1987, Geologic map of the Manti 30' x 60' quadrangle, Carbon, Emery, Juab, Sanpete, and Sevier Counties, Utah, U.S. Geological Survey Map I-1631.