2-9-2010

Hydrogeological engineering approaches to investigate and characterize heterogeneous aquifers

Dylan Harp

Dylan Robert Harp
Candidate

Department of Civil Engineering
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Bruce M Thomson, Chairperson

Hydrogeological Engineering Approaches to Investigate and Characterize Heterogeneous Aquifers

by

Dylan Robert Harp

A.A.S., Forestry Technology, Southeastern Illinois College, 1994
B.S., Civil Engineering, University of New Mexico, 2004
M.S., Civil Engineering, University of New Mexico, 2005

DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy
Engineering

The University of New Mexico

Albuquerque, New Mexico

December, 2009

Acknowledgments

This dissertation required the input and guidance of many people. First and foremost, I am indebted to Velimir Vesselinov for his mentoring, guidance, contributions, and insights. I extend my gratitude to Zhenxue Dai, who provided guidance during the early part of the research, generously sharing his concepts and research ideas. Constructive comments were provided by many people during the research, including, to name a few: Kay Birdsell, Andy Wolfsberg, Jasper Vrugt, Zhiming Lu, Bruce Robinson, George Zyvoloski, Ed Kwicklis, Carl Gable, David Broxton, Mahmoud Reda Taha, and Timothy Ross. I would like to thank Zora Dash and Terry Miller, who generously provided their expertise on technical issues I encountered during the research. I am grateful for constructive comments and insights provided by my advisory committee, Bruce Thomson, Velimir Vesselinov, Gary Weissmann, and John Stormont. I am also indebted to my father, Lowell Harp, for editorial comments on drafts of several of the chapters. The majority of this research was conducted while the author was supported as a Graduate Research Assistant by the Los Alamos National Laboratory.

Hydrogeological Engineering Approaches to Investigate and Characterize Heterogeneous Aquifers

by

Dylan Robert Harp

ABSTRACT OF DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Engineering

The University of New Mexico

Albuquerque, New Mexico

December, 2009

Hydrogeological Engineering Approaches to Investigate and Characterize Heterogeneous Aquifers

by

Dylan Robert Harp

A.A.S., Forestry Technology, Southeastern Illinois College, 1994

B.S., Civil Engineering, University of New Mexico, 2004

M.S., Civil Engineering, University of New Mexico, 2005

Ph.D., Environmental Engineering, University of New Mexico, 2009

Abstract

This dissertation presents a compilation of five stand-alone manuscripts (Chapters 2 through 5 and Appendix A). Chapters 2 through 5 present hydrogeological analysis approaches, while Appendix A is utilized within the dissertation introduction as an example of a non-physically based modeling approach, albeit demonstrated on a non-hydrogeologically based application. Chapter 2 presents an inverse approach to decompose pumping influences from water-level fluctuations observed at a monitoring location. Chapter 3 presents an inferencing approach to identify effective aquifer properties at the interwell scale that can be applied to highly transient datasets. Chapter 4 introduces the use of a Markov-chain model of spatial correlation to an automated geostatistical inverse framework, demonstrating the approach on a 2-D two-stratigraphic-unit synthetic aquifer. Chapter 5 utilizes the inverse framework

introduced in Chapter 4 to develop a stochastic analysis approach to identify the most plausible geostatistical model given the available data. The dissertation introduction reconciles these hydrogeological engineering approaches within the context of the current hydrogeological perspective, discussing where these approaches fit within the often conflicting goals of providing operational decision support based on modeling and advancing the science of hydrogeology beyond its current limitations.

Contents

List of Figures

List of Tables

List of Tables

Chapter 1

Dissertation Introduction

The science of hydrogeology appears to be at a crux. Equipped with mathematical models that fail to adequately represent physical processes at a scale of hydrogeological interest, and measurement techniques that can at best provide extremely sparse samplings of state properties, the best science that hydrogeology can offer at present is the calibration of a theoretically inadequate model that will provide a behaviorally adequate simulation of observations when given details of a particular site and nature of system forcings. Therefore, the hydrogeologist often assumes the role of a calibrationist, unclear or unconcerned about the relationship between the adjustable model parameters and the system properties they are intended to represent. This lack of hydrogeological knowledge and information hinders the application of hydrogeological science in the development and use of technological methods needed to provide decisions for pressing water resources problems. The necessity to develop techniques to provide modeling support to urgent water resources problems has left the hydrogeologist with a collection of disconnected technological approaches, many lacking a strong basis in hydrogeological knowledge (*Klemes*, 1988). As a result, we have multiple perspectives on hydrogeology, providing partial, and sometimes conflicting, explanations of hydrogeological phenomena.

A comprehensive solution to the current state of affairs in hydrogeological science is not currently evident, while the need for informed and defensible decisions based on models is becoming increasingly important for water resource management. Therefore, results of hydrogeological investigations utilizing current approaches must be tempered with the recognition of the limited quantity and quality of information that can be expected, given the existing theoretical and data collection limitations. Furthermore, researchers must be cognizant of these limitations and forthright in the dissemination of this information. In this chapter, I will discuss the current state of hydrogeology, reconciling the approaches developed in Chapters 2 through 5 within the context a current hydrogeological perspective.

1.1 Hydrogeological models

The current theoretical, computational, and measurement limitations in hydrogeology have led to the development of a multitude of modeling approaches. These approaches can be classified into methods that employ (1) physically-based models and (2) black-box models. The current use of a physically-based model in a hydrogeological investigation is typically associated with the inference of system properties that cannot be directly measured. Black-box models provide an efficient framework to calibrate simulations to observations, lacking a physical representation of the system (*Ross*, 2004). As the approaches discussed in Chapters 2 through 5, belong to the former category, this type of approach will be discussed here. Appendix A presents an example of the latter approach with a non-hydrogeological application (i.e. structural bond strength of masonry) using a fuzzy-rule base.

The basis of the current physically-based, hydrogeological model is the Navier-Stokes equation, derived to describe physical processes at the pore scale. Current, widely-used hydrogeological models present an attempt to upscale (average) these

small-scale processes to hydrogeologically relevant scales. To address this upscaling in a direct manner would require the specification of the pore-scale boundary conditions, an impractical task for even the smallest hydrogeological investigation (*Beven*, 2002). This has led to the representation of fluid as a continuum, employing Darcy's equation in an attempt to integrate pore-scale processes over a representative elementary volume (REV) (*Bear*, 1972).

This framework is popular as it provides a compact, elegant description of fluid flow through porous media. However, it has been criticized as it neglects large-scale processes resulting from fluid flow in heterogeneous media (*Klemes*, 1988; *Grayson et al.*, 1992; *Beven*, 1993; *Kirchner*, 2006; *McDonnell et al.*, 2007). The physics of these large-scale processes cannot be captured in the current models as they are dependent on nonlinear and emergent properties, such as threshold behavior and competitive feedbacks, that manifest themselves only at a sufficiently large scale that incorporates the collective behavior of multiple pores and heterogeneities (*McDonnell et al.*, 2007). As a result, the calibration process of hydrogeological models attempt to fit observations made at large scales by estimation of parameters designed to describe pore-scale processes, creating ambiguous relationships between model parameters and the system properties they are intended to represent.

Physically-based models can be classified into analytical versus numerical models. An analytical model implies that a closed-form solution of the groundwater flow equation exists to define the model, while a numerical model implies that a numerical approximation scheme (e.g. finite difference method, finite element method) is utilized to solve the groundwater flow equation, allowing analytically intractable boundary conditions and distributed-parameter specifications (parameterizations) to be considered. Chapters 2 and 3 present examples of the use of an analytical model, while Chapters 4 and 5 present examples utilizing a numerical modeling approach.

1.1.1 Analytical models

Theis (1935) presented the general equation for transient flow in a porous media, based on a solution for heat conduction, for a fully-penetrating well in a confined, homogeneous aquifer of infinite extent. Chapter 2 utilizes the Theis solution to decompose pumping influences at monitoring locations utilizing hydraulic influence-response data. The approach is demonstrated to be useful in identifying cross-hole connections within an aquifer in the presence of spatially and temporally variable pumping influences. Parameter estimates provide cross-hole hydraulic characteristics, however, they cannot be considered to directly represent aquifer properties within the context of the Theis solution.

Cooper and Jacob (1946) present an approximation to the Theis solution valid at late times when pressure gradients have become steady, although pressures remain transient. This approximate solution has been found to be useful in the inference of aquifer properties (Cooper-Jacob method) as early-time portions of the drawdown curve, which are affected by pumping rate variance, well storage, well skin, and inter-well heterogeneities, are omitted from consideration. One of the limitations of this approach is that it requires an adequate drawdown record after quasi-steady state conditions have developed.

Chapter 3 presents an approach to infer aquifer properties from hydraulic influence-response data irregardless of whether quasi-steady state has been reached by accounting for the early-time behavior of Theis solution parameters. The parameter values can be considered estimates of effective aquifer properties at the inter-well scale subject to the assumptions implicit in the use of the Theis solution. As discussed above, the use of the governing groundwater flow equation implies a continuum approach based on the description of small-scale processes with modeling assumptions that are generally not valid in applications. As a result, the significance of the parameters

must be cast within the context of the Theis solution. Within this context, Chapter 3 demonstrates that the parameter estimates provide information on the effective transmissivity at the inter-well scale and indications of inter-well connectivity.

1.1.2 Numerical models

The two most common numerical schemes applied to solve the groundwater flow equation are the methods of finite differences and finite elements. The method of finite differences approximates the groundwater flow equation using an analog where the approximation error can be reduced by refining the spatial and temporal discretization (*Bear*, 1972). The method of finite differences, while easily implemented, is limited to orthogonally structured grids, limiting its ability to represent complicated boundary conditions and parameterizations. Alternatively, the method of finite elements solves the groundwater flow equation using concepts from the calculus of variations, associating a functional with the groundwater flow equation (*Bear*, 1972). The method of finite elements allow the use of unstructured grids, enabling the representation of complicated geometries in the boundary conditions and parameterizations.

Chapters 4 and 5 utilize a control volume finite element (CVFE) approach (*Zyvoloski*, 2007). However, as the grids are orthogonally structured in both papers (2-D in Chapter 4; 3-D in Chapter 5), the CVFE approximation degenerates to an equivalent finite difference approximation.

1.2 Measurement scale and sampling sparsity

Grayson et al. (1992) stated that given the limitations in measurement techniques, hydrologists are currently unable to appropriately design, operate, and verify their

models. Unfortunately, approaching two decades later, many of these limitations still exist. Many of the current practical limitations in hydrogeological modeling stem from an inability to adequately and properly characterize hydrogeological phenomena. The lack of our ability to properly measure system properties at a hydrogeologically relevant scale is the reason model calibration is performed, resorting to the estimation of properties based on data that can be measured (e.g. water-level elevations). Ideally, system properties at hydrogeologically relevant scales would be measured directly from the field for input into a model, thereby allowing truly incisive testing of the model assumptions free of the inherent model uncertainty associated with parameter estimation. However, the relatively short duration of hydrogeological tests versus the inherently slow speed of hydrogeologic processes suggest that this may be an intrinsic limitation to hydrogeological investigations. Therefore, the inference of system properties by parameter estimation will be an integral component to hydrogeological investigations for the foreseeable future.

In the data-limited field of hydrogeology, methods that can extract the available hydrogeologic information from existing data are imperative. The approaches presented in Chapters 2 and 3 demonstrate the ability to identify and decompose pumping influences at monitoring locations and infer aquifer properties from existing pumping and water-level records. The use of these existing datasets can alleviate the burden and/or expense of collecting data solely for hydrogeological analysis. Existing datasets collected from water-supply networks are often neglected as the water-level fluctuations can be highly transient due to the spatial and temporal variations in pumping influences. Chapters 2 and 3 present approaches that allow the utilization of such datasets by considering water-level fluctuations as the cumulative effect of the existing temporally and spatially transient pressure-influence sources (e.g. a cycling pumping well network).

In order to deal with the sparseness of data in hydrogeology, geostatistical meth-

ods are often employed to provide a probabilistic assessment of spatial variability of aquifer properties. Geostatistical methods are generally formulated using a covariance-based set of normal equations (*Deutsch and Journel*, 1992). The definition of the variogram of an aquifer property provides perhaps the earliest means to consider that a pattern may exist in the heterogeneity of an aquifer (*de Marsily et al.*, 2005). A stratigraphic representation of spatial variability can be accommodated in geostatistics using indicator functions. This approach has utility in cases where a stratigraphic layering of soil and/or rock types is assumed or observed.

The pilot-point approach (*de Marsily et al.*, 1984) has been perhaps the most widely applied inverse geostatistical approach to numerical groundwater flow models. This approach designates the value of an aquifer property (typically hydraulic conductivity) at specified locations throughout the model as adjustable parameters, referred to as pilot points. The spatial variation of the field is determined by kriging the pilot points and any observed values that may exist. By systematically adjusting these pilot points, the plausibility of various parameterizations can be evaluated by comparing flow simulations to observed hydraulic data. This approach is flexible as the number and location of pilot points can be adjusted, however, the dimensionality of the problem can hinder the inversion as each pilot point introduces an additional dimension to the parameter space. Regularization methods have been introduced to increase the mathematical tractability of over-parameterized inverse problems, a typical result of the pilot point approach (*Doherty*, 2003). While this approach may enable the inversion to reduce residuals, it does not address the issue that the available information may not be able to constrain the degrees of freedom introduced by the pilot points. This approach is also restricted in an automated inverse framework as the nature of spatial variations must be designated *a priori* by a variogram type.

Markov-chain-based geostatistics present an alternative formulation of the geostatistical normal equations using auto- and cross-transition probabilities of strati-

graphic units (*Carle and Fogg*, 1996). A transition-probability framework emphasizes the sequential juxtapositioning of stratigraphic units, whereas the variogram emphasizes the statistical tendency of a system property value to deviate with distance. Although the incorporation of indicator functions has made covariance-based geostatistics amenable to modeling discrete zones, the fact that the underlying basis of this approach is designed for continuous variations of properties limits its ability to represent stratigraphic representation. For example, a Markov-chain-based geostatistical approach has the following advantages over covariance-based indicator geostatistics: complicated facies juxtapositional tendencies (e.g. asymmetric spatial correlations), such as those produced by depositional processes, are easily represented; the representation of the transition probability function by a sum of exponentials allows significant variability in spatial correlations, eliminating the need to define the spatial-correlation functional form *a priori*; model parameters are geometrically/geologically interpretable; and stratigraphic sequences observed in well-bore logs lend themselves to interpretation by transition probabilities. Markov-chain-based geostatistics are limited to cases where exhaustively defined, mutually exclusive stratigraphic units can be identified (*Carle and Fogg*, 1996).

Chapters 4 and 5 present an alternative to the pilot point approach, designating geostatistical parameters from a Markov-chain model of spatial variability as adjustable parameters. Chapter 4 presents an inverse framework for identifying aquifer structure and stratigraphic unit conductivities using pressure and flow data as calibration targets and well-bore logs for geostatistical conditioning data. The approach utilizes a single stochastic realization to represent the spatial variability for each combination of Markov-chain model parameters. The approach is demonstrated on a 2-D synthetic aquifer composed of two stratigraphic units with uniform hydraulic conductivities where the vertical and lateral Markov-chain model parameters and stratigraphic unit hydraulic conductivities are set as adjustable parameters. The use of this approach is contingent on the existence of spatially and temporally distributed

water-level elevation observations and geologic conditioning data, such as well-bore logs. Issues with over-parameterization are not expected to be as prevalent with this approach as with the pilot point approach as the adjustable structural parameters are limited to the Markov-chain model parameters. Of course, it is also expected that this approach will not be as flexible as the pilot point approach, however, it is theoretically possible to extend the flexibility of the approach by introducing pilot points in the conditioning data and adjustable parameter set. The availability of conditioning data to constrain the inversion will have a significant impact on the approach.

Stochastic simulation can be used to produce an infinite set of geostatistically equally-probable realizations of spatial variability. The hydrogeological characteristics and uncertainty associated with a given geostatistical model can be explored by performing flow simulations on these realizations. This provides a framework for performing a stochastic analysis of the plausibility of various geostatistical models based on hydraulic and geologic data. Chapter 5 extends the approach presented in Chapter 4 to fully utilize stochastic simulation within the inverse framework. The approach characterizes the hydraulic response of a geostatistical model, associated with a combination of Markov-chain model parameters, using convergent statistical averages of pressure and flow predictions from a set of realizations. In this case, the stochastic inversion searches for the most hydraulically plausible Markov-chain model constrained by observed geology. This approach recognizes that the available information in typical hydrogeological investigations is not sufficient to constrain a geostatistical inversion to a single stratigraphy, and that a more realistic approach is to constrain the set of solutions to characteristics (statistical, or otherwise) describing the stratigraphy. The cost of this enhanced recognition of hydrogeological uncertainty is the computational effort required to sufficiently characterize each geostatistical model.

1.3 Hydrogeological inverse modeling

The method of inverse modeling can be described as an effort to reduce residuals by adjusting model parameters. When inverting physically-based models, the process becomes more than mere calibration, as the models are derived under the expectation that the parameters represent system properties. Therefore, inverse modeling can provide the ability to evaluate the conceptual basis of models, including theoretical underpinning, boundary conditions, parameterizations, etc. As the capacity to develop complex models has increased with advances in computational capabilities, so has the necessity to recognize certain pitfalls in the application and interpretation of results in inverse modeling. This section will discuss the inverse approaches presented in Chapters 2 through 5 in the context of the available inverse approaches, indicating the potential pitfalls.

Prior to the development of inverse methods, parameter estimation was carried out by trial and error, where parameter values are substituted into a model until an adequate fit between observations and simulations is achieved. Charles Theis was perhaps the first hydrogeologist to develop a tool (i.e. the Theis-type curve approach) for parameter estimation with a graphical approach utilizing the Theis solution (*Theis*, 1935) described in *Jacob* (1940). The concept of inverse modeling to obtain parameter estimates emerged with the introduction of direct inverse methods, where a formal boundary value problem is derived with the model parameters as spatially dependent variables, as classified by *Neuman* (1973). Direct inverse methods require head variations and derivatives at all locations for input into the groundwater flow equation, thereby creating a partial differential equation where model parameters are spatially dependent variables. Given the typically sparse collection of head observations at a site, unsampled locations within the model domain must be interpolated. Errors associated with the interpolated head values, as well as the measurement error of the head observations, introduces significant uncertainty

in the parameter estimation (*Carrera et al.*, 2005).

A more recent inverse approach is the use of indirect inverse methods defined by *Neuman* (1973) as methods that approach the inverse problem with the concept of minimizing residuals between available observations and their corresponding model predictions by adjusting model parameters. In this way, knowledge of the pressure response of the aquifer at locations without observations is not assumed. While indirect methods do provide a more realistic approach considering the uncertainty due to the sparsity of information, these methods often require a substantial computational effort to deal with nonlinearities and nonconvexities (*Yeh*, 1986). As the approaches presented in Chapters 2 through 5, as well as Appendix A, all utilize indirect inverse methods, the rest of this section will focus on these methods.

The application of indirect inverse methods to hydrogeological models presents certain challenges as these models can be composed of large systems of equations, consideration of transients is often essential, knowledge of heterogeneities and boundary conditions is typically limited, important parameters (other than hydraulic conductivity for example) are often neglected in order to simplify the problem, hydrogeologic properties are scale dependent, and model parameters often have low sensitivity to state variables (e.g. pressures) (*Carrera et al.*, 2005). These difficulties often result in poor matches between observations and simulations.

When a model is found to be unable to provide a sufficient fit to observations, it has become common practice to add more parameters to provide the model with more degrees of freedom. At some point, as the degrees of freedom increases, it becomes inevitable that a mathematical model will be produced with the ability to behaviorally simulate the observations. This does not necessarily indicate that the model is providing a realistic representation of the system, but merely demonstrates that a mathematical model with enough degrees of freedom can be fitted to a set of observations (*Beven*, 2006; *Grayson et al.*, 1992). In other words, the sophistication

of measurement techniques and development of hydrogeological theory has not kept pace with our ability to produce complicated, distributed-parameter models (*Beven*, 2000).

While over-parameterized models often demonstrate the ability to simulate the behavior of a given set of observations, their inability to provide information on the physical properties of a system is evident by the number of parameter sets that provide a behavioral fitting. The non-uniqueness, or equifinality (*Beven*, 1993), of solutions is often considered an algorithmic challenge caused by the ill-posedness of the calibration that must be overcome in order to obtain the "correct" solution, and thereby identify the "true" properties of the system. It should be realized that instead, equifinality is an intrinsic characteristic of over-parameterization as the number of degrees of freedom cannot be constrained by current calibration data sets. Additional causes of equifinality are the inappropriate theoretical underpinnings of current models (*Kirchner*, 2006) and the lack of ability to sufficiently define the system within the model (e.g. boundary conditions) (*Klemes*, 1988).

Hydrogeological investigations that use models that cannot fail lack a framework for testing hypotheses about a system, instead providing a black-box representation of a system that is limited to simulation and possibly prediction of behavior given similar system forcings. If this is the goal of the investigation, it may be more efficient to simply use a model that is designed as a black box, such as an artificial neural network (*Ross*, 2004), where the number of nodes and layers, which have no pretense to represent system properties, can be increased to provide a behavioral model. While this black box approach, whether intentional or not, may be a valid approach to solve a particular problem at hand, it will not provide insights into the intrinsic properties or underlying physics of the system. Appendix A provides an example of this type of approach with the application of modeling the bond strength of masonry.

When acquiring knowledge about the processes within a system is a research objective, the use of models that can fail, and therefore provide information on the validity of model assumptions, must be considered. In general, the less parameterized a model is, the less immune it is to being proven wrong (*Kirchner*, 2006). Therefore, minimally parametrized models can provide the means to reject hypotheses based on the available observations, even if the parameterization reduces the model to an extreme simplification of the system. In this way, model complexity can be added stepwise, subject to the validity of assumptions that are verified by comparing observations to simulations from models of reduced complexity.

The approaches presented in subsequent chapters can be evaluated with respect to their ability to deal with the equifinality of hydrogeological models. Chapter 2 and 3 utilize a minimally parameterized model in the form of an analytical solution to the groundwater flow equation. The use of an analytical model provides a simple and efficient characterization of the system, although it requires the use of assumptions that are not valid in typical applications, as discussed in Section 1.1. The use of this model provides the means to incisively test the conceptual basis of the model and the information content of the data as issues of equifinality are minimized. However, this approach cannot consider details of the heterogeneity of the system, and is therefore limited to exploring the general characteristics of the system.

Chapters 4 and 5 provide examples of inverse modeling approaches designed to identify details of the system heterogeneity. In Chapter 4, the strategy of the approach is to identify an optimal hydrostratigraphy, while in Chapter 5, the goal is to identify the most plausible statistical description of the hydrostratigraphy.

Chapter 4 presents an approach to identify an optimal (in a least squares sense) hydrostratigraphy by modifying geostatistical model parameters and considering a single stochastic realization. This approach is perhaps slightly less prone to issues of equifinality than the pilot point approach as geostatistical model parameters are

being modified, as opposed to a spatially distributed collection of aquifer property values. The approach presented in Chapter 5 provides a more explicit consideration of equifinality, attempting to identify the most plausible geostatistical model by characterizing each geostatistical model with hydraulically convergent sets of stochastic realizations. The fact that the possible hydrostratigraphies are grouped into statistical characterizations defined by geostatistical model parameters constrains the inversion, reducing issues of equifinality while acknowledging the potential limitations in the amount of information available from observations.

Chapter 2

Identification and Analysis of Long-Term Water-Level Fluctuations Due to Spatially and Temporally Variable Water-Supply Pumping[1]

Abstract

Identification of the pumping influences at monitoring wells caused by spatially and temporally variable water-supply pumping can be a challenging, yet important hydrogeological task. The information that can be obtained can be critical for conceptualisation of the hydrogeological conditions, identification of aquifer properties,

[1]Submitted for publication in Ground Water: Harp, D.R. and V.V. Vesselinov (2009), Identification and Analysis of Long-Term Water-Level Fluctuations Due to Spatially and Temporally Variable Water-Supply Pumping, Manuscript submitted for publication.

and estimation of the zone of influence of the individual pumping wells. However, the pumping influences are often intermittent and small in magnitude with variable production rates from multiple pumping wells. Further complications can be caused by water level fluctuations caused by other hydrogeologic mechanisms (e.g. barometric pressure, recharge, poroelastic effects, etc.). While these difficulties may support an inclination to abandon the existing dataset and conduct a traditional cross-hole pumping test, that option can be challenging and expensive to coordinate and execute. This chapter presents a method that utilizes a simple analytical modeling approach for analysis of a long-term water-level record to identify pumping wells influencing the water-level fluctuations and identification of the effective cross-hole aquifer properties associated with pumping and observation wells. Thus the analysis provides an efficient and cost-effective alternative to designed and coordinated cross-hole pumping tests. We demonstrate this method on a dataset from the Los Alamos National Laboratory site. Our analysis also provides (1) an evaluation of the information content of the transient water-level data, (2) indications of potential large-scale structures of the aquifer heterogeneity, and (3) guidance for the development of more complicated models requiring detailed specification of the aquifer heterogeneity.

2.1 Introduction

Identification of the pumping influences at a monitoring well due to pumping at water-supply wells and respective estimation of the aquifer properties are typically performed by analysis of a series of coordinated cross-hole pumping tests. However, the planning and execution of these tests can be expensive and challenging. For example, it is often difficult to exclude, or take into account, all the potential hydrogeological factors influencing the pressure transients during the test (e.g. *Neuman*

and Witherspoon (1972); *Walthall and Ingram* (1984); *Dawson and Istok* (1991); *Rasmussen and Crawford* (1997); *McLin* (2005, 2006a,b)). Also in many cases, it is logistically infeasible to cease water-supply pumping in the entire aquifer to conduct a traditional pumping test (which includes pre- and post-pumping recovery periods) to eliminate influences from nearby water-supply wells. As advocated by *Yeh and Lee* (2007), existing datasets from monitoring well networks recorded during long-term pumping of water-supply wells provide an alternative to datasets generated by traditional pumping test. Such datasets are frequently collected in monitoring-well networks established near contamination sites and municipal water-supply wells (*Barnett et al.*, 2003; *Gross*, 2007; *Mason et al.*, 2005; *Hix*, 2007; *Koch and Schmeer*, 2009). However, the pumping influences are often intermittent and small in magnitude with water level fluctuations caused by other hydrogeologic mechanisms (for example, recharge transients), causing the identification of the pumping influences due to a complex spatially and temporally variable water-supply pumping regime to be difficult. The analysis may require the use of complicated computational models and involve large data sets that are challenging to process. Nevertheless, when compared to traditional pumping tests, this approach provides some important advantages. First, the collected data are representative of the aquifer properties during existing water-supply conditions, while the aquifer properties obtained by pumping-test interpretations may need to be upscaled to be applied for simulation of the flow conditions under water-supply pumping. Second, the aquifer is typically stressed more intensively with pressure influences affecting larger areas, providing better identification of pumping influences with small magnitudes. Third, measurement errors in the collected water levels and pumping records have the potential to be reduced due to the large number of observations and by repeated pumping cycles often present in the long-term data record; as a result, the uncertainties in the estimated aquifer parameters are also expected to be reduced. Last, interpretation of transient water-level data at multiple monitoring wells influenced by transient pumping at multiple

water-supply wells may provide information about the large-scale aquifer structures; furthermore, the analyses can be extended to provide a tomographic characterization of aquifer properties (e.g. *Neuman* (1987); *Vesselinov et al.* (2001); *Straface et al.* (2007)). The identification of the pumping influences at the monitoring wells can also be critical for conceptualization of the hydrogeological conditions at the site, and provide indications of the extent of the zone of influence of the individual pumping wells.

The decomposition of pressure influences requires a model with the ability to characterize the hydraulic response at a monitoring well due to transient pumping at the water-supply wells. Adequate characterization of the water-level transients requires calibration of the model in the form of parameter estimation. If the model is complicated with a large number of adjustable parameters, the calibration can become computationally demanding. As a result, the optimal parameter estimates may be difficult to identify and the parameter estimation may not have a unique solution (i.e. the inverse problem can become ill-posed) (*Carrera et al.*, 2005). To avoid this, we attempt to use the simplest possible model that can be satisfactorily applied. We choose to use analytical methods here for simulating pumping influences at the observation wells. The use of analytical methods makes the analysis consistent with pumping-test interpretations where analytical type-curve methods are commonly applied (*Freeze and Cherry*, 1979). The adjustable model parameters of analytical models represent the hydrogeologic properties of the aquifer. As a result, the use of an analytical solution for pressure influence decomposition can provide additional hydrogeologic insights.

Theis (1935) introduced an analytical solution of the general equation for flow of a Newtonian fluid in porous media for non-steady conditions (Theis solution), such as those encountered as a result of pumping an aquifer. The Theis solution is valid for simplified hydrogeologic scenarios assuming a constant pumping rate, horizontal flow,

transmissivity and storativity homogeneity, uniform thickness, and infinite lateral extents of the aquifer. The Theis type-curve method (Theis method), developed by Theis and described by *Jacob* (1940), was developed from this work as a means to graphically infer hydrogeologic properties from pumping test data. *Cooper and Jacob* (1946) simplified this approach using an approximation to the Theis solution valid at late pumping times when a quasi-steady state regime is established (Jacob's method), eliminating the use of a Theis type curve. At quasi-steady state (also referred to as steady-shape), pressure gradients are spatially and temporally steady, while pressures remain transient as second order terms become insignificant.

Wu et al. (2005) investigated the behavior of effective property estimates (Theis solution hydrogeologic parameters), concluding that effective parameters will be time dependent at early times. Based on numerical experiments using Gaussian transmissivity and storativity fields, *Wu et al.* (2005) demonstrate that the effective transmissivity will be time dependent at early times, with estimates from different locations converging (decreasing from larger values) towards a similar value at late times. They also demonstrate a time dependency for effective storativity, with values converging (increasing at some locations, decreasing at others) towards distinct values relatively quickly. This late-time convergent behavior corresponds with research by *Meier et al.* (1998) and *Sanchez-Vila et al.* (1999), who investigated the meaning of hydrogeologic parameter estimates obtained from Jacob's method numerically and analytically, respectively. *Straface et al.* (2007) evaluated hydrogeologic parameter inference methods using the Theis solution on a dataset from Montalto Uffugo Alto, Italy. Based on their results, they question the validity of hydrogeologic property inference based on the Theis solution. However, they do state that the Theis solution parameter estimates can be used as first estimates of hydrogeological parameters for a tomographic analysis.

We employ the Theis solution as our groundwater model in order to maintain a

simple and efficient pressure-source identification approach. In doing so, we recognize that the parameter estimates will be affected by the early-time pre-stabilization period, and cannot be considered as accurate estimates of effective hydrogeologic properties. Instead, these estimates can be considered as effective cross-hole parameters that characterize the hydraulic response at a monitoring location due to pumping a well, analogous to parameters that would be obtained from standard cross-hole pumping tests often used to characterize the hydrogeology of an aquifer.

This chapter presents an approach to (1) fingerprint transient water-level variations to the pumping regime of individual water-supply wells and (2) estimate hydrogeologic parameters using a computationally efficient analytical approach. Interpretation of the quantitative results from this approach can provide (1) indications of the large-scale structure of the aquifer heterogeneity, (2) an evaluation of the information content in the calibration data, and (3) guidance for the development of more complicated and less computationally efficient models possessing the ability to explicitly consider heterogeneity. If the development of a simple analytical model is successful, the calibrated model can be used to predict future water-levels under different pumping scenarios. The calibrated model can be applied to address measurement uncertainties; for example, substantial deviations between model predicted and observed water-levels may indicate technical problems in the data collection.

As computational resources have become increasingly more powerful, the complexity and computational demand of models has proportionally increased. The concept of model parsimony, as the principle of Occam's razor proposes, is often lost or neglected in the quest to develop elaborate models that capture increasingly refined details of complexity. While complex models are required in certain applications, in other cases, this approach can mask fundamental insights that become obvious when the data are analyzed with models of minimal complexity. An example of this is the estimation of effective porosity as an indicator of transport connectivity as

described by *Trinchero et al.* (2008), where a full or partial specification of porosity heterogeneity would embed the transport connectivity information in the distributed parameter. The research presented here demonstrates an analysis of pumping and water-elevation records using a relatively simple model that provides fundamental insights into the aquifer pressure response and is a first step toward development of more complicated aquifer models that aim to characterize the groundwater flow complexity and aquifer heterogeneity utilizing the same data.

We demonstrate the proposed method using some of the pressure and water-supply pumping records from the regional aquifer at the Los Alamos National Laboratory (LANL) site located in north-central New Mexico, U.S.A.

2.2 Methodology

The two major goals of the analysis are to (1) fingerprint transient water-level variations to the transients in the pumping regime of individual water-supply wells and (2) estimate effective cross-hole aquifer properties. To do this, we need a model that can simulate potential pumping influences at the monitoring wells. The simplest model that can be applied is the Theis solution, defined as

$$\hat{s}(t) = \frac{Q}{4\pi T} W(u) = \frac{Q}{4\pi T} W\left(\frac{r^2 S}{4Tt}\right) \tag{2.1}$$

where $\hat{s}(t)$ is the predicted drawdown at time t since the pumping commenced, Q is the pumping rate, T is the transmissivity, $W(u)$ is the negative exponential integral ($\int_u^\infty e^{-y}/y\, dy$) referred to as the well function, u is a dimensionless variable, r is radial distance from the pumping well, and S is the storativity. The assumption of homogeneity implicit in the Theis solution, discussed above, is apparent by the constant hydrogeologic parameters, T and S, in equation (2.1). It is important to note

that more complicated analytical solutions accounting for partial well penetration, leakage effects, or three-dimensional flow could have been applied in our analyses as well, if the Theis solution had failed to reproduce the observed water-level data.

In order to include multiple pumping wells and variable rate pumping periods in the Theis solution, the principle of superposition is invoked as

$$\hat{s}(t) = \sum_{i=1}^{N} \sum_{j=1}^{M_i} \frac{Q_{i,j} - Q_{i,j-1}}{4\pi T_i} W\left(\frac{r_i^2 S_i}{4T_i(t - t_{Q_{i,j}})}\right) \tag{2.2}$$

where N is the number of pumping wells (sources), M_i is the number of pumping periods (i.e. the number of pumping rate changes) for pumping well i, $Q_{i,j}$ is the pumping rate of the ith well during the jth pumping period, and $t_{Q_{i,j}}$ is the time when the pumping rate changed at the ith well to the jth pumping period. The drawdown calculated by equation (2.2) represents the cumulative influence of the N pumping wells at a monitoring well. Note that T_i and S_i are effective cross-hole properties of the aquifer that characterize the influence of the ith pumping well at the observation well. It is important to note that if traditional pumping tests were conducted at each of the pumping wells separately, and the drawdowns are monitored and analyzed at the observation well, the interpretation of the cross-hole test data will produce estimates that are theoretically equivalent to T_i and S_i in equation (2.2).

As the calibration targets in the model inversions presented here are water elevations as opposed to drawdowns, we define the predicted water elevation $\hat{h}(t)$ at time t as

$$\hat{h}(t) = \hat{h}_o - \hat{s}(t) \tag{2.3}$$

where $\hat{h}_o = \hat{h}(0)$ and is defined as the initial predicted water elevation at the obser-

vation well at the time the pumping begins.

Model calibration is performed using a Levenberg-Marquardt approach (*Levenberg*, 1944; *Marquardt*, 1963) where the objective function is defined as

$$\mathbf{\Phi}(\theta) = \sum_{i=1}^{n} [h(t_i) - \hat{h}(t_i)]^2 \tag{2.4}$$

where θ contains the cross-hole estimates of T and S associated with each pumping well and $\hat{h}_o$ associated with the monitoring location, and n is the number of head observations, $h(t_i)$, included as calibration targets where i is an observation time index.

The simulation of the drawdowns is performed using the WELLS code (written by *V.V. Vesselinov*, 1992), which implements equation (2.2). The calibration is performed using *PEST* (*Doherty*, 2004).

2.3 Site Data

Due to concerns related to the migration of potential LANL-derived contaminants in the subsurface, a complex monitoring network is established in the regional aquifer beneath LANL. The network includes 92 regional monitoring wells with a total of 336 monitoring screens (*Allen and Koch*, 2008). At each screen, water-level fluctuations are automatically monitored using pressure transducers. In addition, water samples are collected for geochemical analysis. The aquifer beneath LANL is an important source of water for LANL and neighboring municipalities. There are 7 water-supply wells in close vicinity to the study area; 18 more water-supply wells are located nearby. The ultimate goal is to incorporate all these data in the development and calibration of the regional aquifer model. Here we analyze only a subset of the data.

The pressure and water-supply pumping records considered here are collected from 3 monitoring wells (R-11, R-15 and R-28) and 7 water-supply wells (PM-1, PM-2, PM-3, PM-4, PM-5, O-1, and O-4) located within the LANL site. Figure 2.1 displays a map of the spatial location of the wells and Table 2.1 tabulates the distances between monitoring and water-supply well pairs. Figure 2.2 presents the pressure and production records for the monitoring wells and water-supply wells, respectively.

The regional aquifer beneath the LANL site is a complex stratified hydrogeologic structure which includes unconfined zones (under phreatic conditions near the regional water table) and confined zones (the deeper zones) (*Vesselinov*, 2004a,b). The three monitoring wells considered in this analysis are screened near the top of the aquifer with an average screen length of 11 meters. The water-supply wells partially penetrate the regional aquifer with screens that begin near the top of the aquifer, but penetrate deeper with an average screen length of 464 meters. Nevertheless, field tests demonstrate that most of the groundwater supply is produced from a relatively narrow section of the regional aquifer that is about 200-300 m below the regional water table (*Los Alamos National Laboratory*, 2008a). Implicit in the use of the Theis solution is that the groundwater flow is confined and two-dimensional. We assume that this is a justifiable assumption here given the small magnitude of observed drawdowns (less than 1 m at the monitoring wells and less than 20 m at the water-supply wells), the relatively long distances between supply and monitoring wells (more than 1 km; Table 2.1) compared to the effective aquifer thickness (about 200-300 m). Future analyses will address the three-dimensionality of the groundwater flow and complex hydrostratigraphy of this aquifer.

The water-level observation data considered here span approximately three years, commencing on or shortly after the date of installation of pressure transducers (May 4, 2005 for R-11; December 23, 2004 for R-15; January 7, 2005 for R-28), including records up to November 29, 2007. An additional year of data, terminating on

November 18, 2008, is utilized here for subsequent analysis to evaluate forward model predictions and parameter uniqueness and convergence. The barometric pressure fluctuations are removed using constant coefficient methods with 100% barometric efficiency (*Los Alamos National Laboratory*, 2008b) for all monitoring wells. Although the pressure transducers collect observations every 15 minutes, this dataset is reduced to single daily observations by using the earliest recorded measurement for each day. Some daily observations have been excluded due to equipment failure. The barometric-corrected water levels fluctuate over the initial three year period approximately 0.8 meters for R-11 (940 daily records), 1.7 meters for R-15 (1072 daily records), and 1.0 meters for R-28 (1045 daily records). Seasonal trends are apparent in the water level data showing a general increase in the rate of decline during the summer months and recovery during the winter. Similarities are evident for water-level observations at R-11 and R-28 providing an initial indication that there is a region of similar hydrogeological properties around these two monitoring wells.

Considered pumping records for all pumping wells begin on October 8, 2004 and terminate on November 29, 2007. This record is extended to November 18, 2008 for a subsequent analysis, as discussed above for the water-level observations. The pumping record precedes the water-level calibration data to account for water-level transients due to pumping variations before the water-level data collection commenced. The number of pumping-rate changes for each well are: PM-1 – 1809; PM-2 – 1720; PM-3 – 1141; PM-4 – 164; PM-5 – 1615; O-1 – 39; and O-4 – 2035. The dominant water producers are PM-2, PM-5, and O-4, with significant seasonal pumping for PM-3 and intermittent pumping for PM-4 (refer to Figure 2.2).

Drawing correlations between pressure and pumping transients from a visual comparison of the plots in Figure 2.2 is difficult, except perhaps an apparent influence of PM-4 pumping on monitoring well R-15. Therefore it is essential to fingerprint the water level transients to the pumping records in order to determine the hydraulic

	PM-1	PM-2	PM-3	PM-4	PM-5	O-1	O-4
R-11	2399.8	2902.7	803.6	1929.9	2439.5	3007.2	1367.7
R-15	3787.7	2434.7	2252.2	1081.0	986.0	4460.3	1566.7
R-28	2666.7	2522.4	1154.3	1506.3	2103.8	3384.8	1500.2

Table 2.1: Distances between pumping and monitoring well pairs in meters, where the row headings indicate the monitoring wells and column headings indicate the pumping wells.

connections within the aquifer.

In the applied computational framework, forward model run times for predicting water elevations at R-11, R-15, and R-28 for approximately four years (from October 8, 2004 to November 18, 2008) are each approximately 9 seconds on a 3.0 GHz Intel processor. Inversions initiated with uniform initial parameter values require approximately 600 model runs and, using a single processor, are performed for approximately 1 hour and 40 minutes. Inversions initiated with best parameter estimates from previous inversions require approximately 250 model runs and, using a single processor, are performed for approximately 35 minutes. The water-supply pumping records (i.e. the recorded daily volumetric production), pumping well rates (i.e. the flow rate capacity of each water-supply well), water elevations, temperatures, barometric pressures, well coordinates, and well radii for a desired time interval necessary for each model inversion are extracted from a MySQL® database and output into model and inversion input files automatically using a preprocessor.

2.4 Results and Discussion

Superposition is used to include pressure influences from all pumping wells to predict the pressure transients at a single monitoring well (refer to equation (2.2)). The resulting calibrated pressure transient predictions from the model inversions are

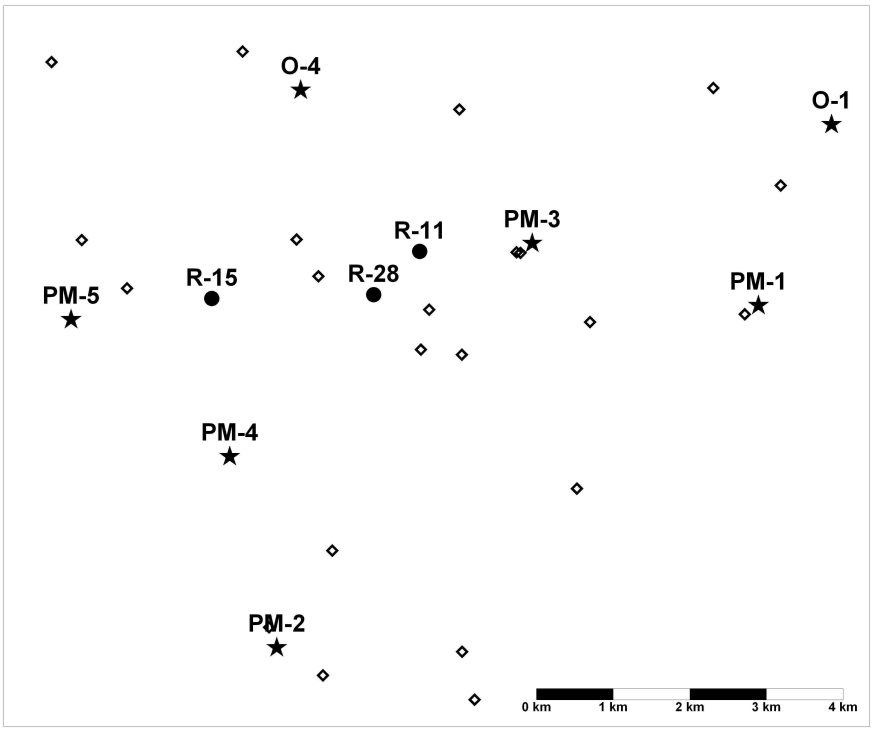


Figure 2.1: Map of observation wells (circles) and water-supply wells (stars) included in the analysis. Locations of newly completed and planned monitoring wells are indicated by open diamonds.

presented along the diagonal of the matrix of plots presented in Figure 2.3. The off-diagonal plots present pressure transient predictions utilizing parameters estimated from model inversions on one monitoring well (diagonal plots) to predict pressure transients at an alternate monitoring well. The row headings denote the monitoring well where pressure transients are being predicted, while the column headings denote the monitoring well associated with the parameter estimates applied in the model. For example, the plot on the row labeled "R-28" and column labeled "R-28", presents pressure transient predictions for R-28 using parameter estimates obtained by inverting the R-28 model. In contrast, the plot on row labeled "R-11" and column labeled "R-28", presents pressure transient predictions for R-11 using parameter es-

timates obtained by inverting the R-28 model. In other words, the diagonal plots are calibration results, while the off-diagonal plots demonstrate the appropriateness of utilizing one well's calibrated parameters to predict an alternate well's pressure transients. The numbers of calibration targets in each inversion are equal to the number of water elevation observations for each monitoring well (i.e. R-11 – 940; R-15 – 1072; R-28 – 1045). In this case, each model inversion has 15 adjustable parameters as it includes two parameters associated with the cross-hole transmissivity and storativity characterizing hydraulic connection between the respective monitoring well and each of the seven pumping wells and the initial water elevation of the monitoring well.

The diagonal plots in Figure 2.3 demonstrate the model's ability to simulate the hydraulic responses at the monitoring locations. The off-diagonal plots indicate that there is likely some large-scale aquifer zone with similar aquifer properties in the area between R-11 and R-28 as hydrogeologic parameters estimated for one of these wells predict pressure transients reasonably well for the other well. However, it is apparent that aquifer properties change to the west of R-11/R-28 near R-15 as parameter estimates for R-11 and R-28 over predict water-level observations at R-15.

Figures 2.4, 2.5, and 2.6 present the decomposed drawdown contributions from the water-supply wells predicted in the calibrations presented in Figure 2.3 (diagonal plots) for monitoring wells R-11, R-15, and R-28, respectively. The associated water-supply pumping record is plotted along with each drawdown contribution to illustrate the predicted pressure influence at the monitoring wells attributed to each water-supply well. The observed and predicted pressure transients from Figure 2.3 for the associated monitoring well are re-plotted along the top of Figures 2.4, 2.5, and 2.6 for reference.

The model identifies a temporal trend of groundwater decline for wells R-11 and R-28 presented in Figures 2.4 and 2.6, respectively. This declining trend is needed in addition to the drawdown contributions from the individual supply wells

to adequately predict the overall drawdown at wells R-11 and R-28. The cause of this temporal trend has not been identified, but it may be related to factors not directly related to the water-supply pumping (e.g. reduction in infiltration recharge). It is apparent that the inversions identify, or fingerprint, the pumping records from PM-2, PM-3, and PM-4 as influencing the water elevation observations at each of the monitoring wells, while in Figure 2.5, PM-5 pumping record is identified to influence R-15. In addition, even though PM-3 is located about two times closer to R-11/R-28 than R-15 (refer to Figure 2.1 and Table 2.1), the magnitude of PM-3 drawdowns are about two times larger at R-15 than at R-11/R-28. These differences between R-11/R-28 and R-15 pumping transients suggest contrasting aquifer properties in the zones near R-15 and R-11/R-28 and a complex structure of the three-dimensional aquifer flow.

This analysis also suggests that there is a lack of pumping influence of O-4 on the monitoring wells. This is somewhat surprising considering the well locations and the substantial water production at O-4. This indicates a large-scale heterogeneous aquifer structure: either (1) a hydraulic barrier between O-4 and the monitoring wells that obstructs propagation of pumping transients or (2) a highly transmissive aquifer zone in the vicinity of O-4 that diminishes the pressure influence of pumping transients. It appears that similar hydrogeologic conditions may exist to the east of PM-3, given the lack of pressure influence attributed to PM-1. In contrast, the substantial drawdown contribution attributed to PM-2 at the monitoring wells predicted by the models when compared to other closely located pumping wells (refer to Figure 2.1 and Table 2.1) indicates that a large highly-permeable north-south aquifer structure is plausible. This conclusion is also supported by relatively high hydraulic diffusivity estimates between PM-2 and the monitoring wells (see discussion below). These conclusion are symbolically illustrated on a map of the well locations in Figure 2.7. These aquifer features will be investigated further with more complex models capable of explicitly considering spatial aquifer heterogeneity and

three-dimensionality of groundwater flow. Nevertheless, our analyses provide important insights about the zone of influence and the zone of capture of the individual pumping wells.

Table 2.2 contains the parameter estimates for the log (base 10) effective cross-hole transmissivity and log (base 10) effective cross-hole storativity obtained from the calibrations presented in Figures 2.3, 2.4, 2.5, and 2.6. As already discussed in the previous section, parameter estimates represent cross-hole effective properties between pumping and monitoring wells. Similar estimates are expected to be obtained if dedicated cross-hole pumping tests were conducted pumping each water-supply well separately. The calculated log (base 10) effective hydraulic diffusivity is also tabulated. The sample (arithmetic) mean and sample variance for each parameter are tabulated as well. Since the parameters are log-transformed, the arithmetic mean represents the geometric mean of the untransformed parameter values.

To capture the information about aquifer connectivity from both the effective cross-hole transmissivity and storativity estimates, the effective cross-hole hydraulic diffusivities are calculated. By inspecting Tables 2.2 and 2.3, it is apparent that the diffusivities are the highest for R-15 with respect to the monitoring wells, indicating that R-15 has a higher level of connectivity to the pumping wells than R-11 and R-28. This result is due to the higher drawdowns at R-15 when compared to R-11 and R-28, as evident in Figure 2.2. With respect to the pumping wells, PM-2 has the highest diffusivity, indicating that it has a higher level of connectivity to the monitoring wells than the other pumping wells. This result is due to the significant drawdowns predicted to be produced by PM-2 at the monitoring wells, as presented in Figures 2.4, 2.5, and 2.6.

Jacob's approximation is only valid at a sufficient length of time after commencement of the pumping. Since pumping regimes are highly transient, we find that the Jacob's constraint ($r^2S/4Tt < 0.03$) is not dominating our analyses in spite of the

long-term water-level records. As a result, the hydrogeological parameter estimates presented here cannot be interpreted strictly according to the methodology presented in *Meier et al.* (1998) and *Sanchez-Vila et al.* (1999). Our estimates are optimal parameters considering both late-time and early-time drawdown, describing the behavior of the drawdown during both steady- and non steady-shape regimes. Consistent with existing theory and previous analysis (*Meier et al.*, 1998; *Sanchez-Vila et al.*, 1999), the transmissivity estimates are less variable than the storativity estimates (Table 2.3). Still, the variability of the transmissivity estimates is not negligible. Deviations from the domain of a valid Jacob's approximation at least partially explain the larger transmissivity variance and smaller storativity variance obtained here. In our case, the transmissivity estimate must accommodate both early- and late-time drawdown, requiring it to contain not only information about the large-scale (i.e. in the general area affected by the water-supply pumping) effective properties of the regional aquifer, but also information on the small-scale heterogeneity structures controlling hydraulic connectivity between the water-supply and monitoring wells. As a result, it can be expected that the variability in the storativity estimates may not only be indicative of local heterogeneities in the transmissivity field (as in the case of a purely steady-shape flow regime) but may also be affected by other complexities of the regional aquifer (e.g. mixture of unconfined and confined flow conditions, leakage from the deeper sections of the regional aquifer, three-dimensionality of groundwater flow, etc.). Future analyses will further investigate the information about the large- and small-scale aquifer heterogeneities in this dataset.

2.5 Model testing

The individual calibrated models for R-11, R-15, and R-28 (using parameter estimates from Table 2.2) are applied to make blind forward predictions of pressure

transients past November 29, 2007 to November 18, 2008. Figure 2.8 presents these predictions along with the calibrated pressure transients (until November 29, 2007; calibrations were presented previously in Figure 2.3). Figure 2.8 also shows the observed water-level data. It is clear that the models successfully predict major features in the water level transients caused by changes in pumping regimes. Still, the model predictions do not provide perfect representations of the data; the models slightly underpredict the drawdowns at R-15 (by 0.2 m), and deviate by less than 0.1 m from the R-11 and R-28 water levels. Recalibrating the models including the additional data (from November 29, 2007 to November 18, 2008) substantially improves the model predictions, as demonstrated by the calibration results presented in Figure 2.9. Table 2.3 lists the updated parameter estimates from these calibrations. By comparing Tables 2.2 and 2.3, it is apparent that there is minimal differences in the overall parameter estimates, although individual cross-hole properties differ in some cases. The means of the parameter values are unchanged, with the single difference being a reduction in the variance for the log storativity. These small differences are likely due to the pre-stabilization time dependency of the transmissivity and storativity parameters. By referencing Table 2.2 with respect to Figure 2.8 and Table 2.3 with respect to Figure 2.9, it is evident that small changes to the cross-hole parameter estimates allow the model to characterize the transient records more accurately. In this way, the calibrated models have been successfully applied to identify measurement errors in new water-level data by comparison to forward model predictions (not presented here). A model calibrated against water-level transients at particular monitoring well can be successfully applied to predict the water-levels in the same well; however, the calibrated model is not expected to be able to predict the water levels at other observation points in the aquifer.

2.6 Conclusions

The approach described in this chapter allows the identification and decomposition of pressure-influence sources at a monitoring location utilizing existing long-term pumping and water-elevation records. This type of dataset is often available from monitoring-well networks established near municipal water-supply well fields. The approach provides (1) fingerprinting of pumping influences in pressure transients to identify drawdown contributions from individual water-supply wells (2) the estimation of effective cross-hole hydrogeological properties between the respective pumping and observation wells, and (3) information about the zone of influence of individual pumping wells. The presented analysis is computationally efficient due to the utilization of a simple analytical model, which facilitates the processing of large amounts of data associated with long-term records. The same analysis will be computationally very demanding and potentially not effective using more complex models representing details of the aquifer heterogeneity. Utilization of such datasets provides several advantages over conducting traditional cross-hole pumping tests, including the ability to consider long-term records with multiple variable pumping regimes. Interpretation of the results can allow (1) the identification of large-scale hydrogeologic structures within the aquifer, (2) an evaluation of the information content in the calibration data, (3) guidance for the development of more complicated models requiring detailed knowledge of aquifer heterogeneity.

The estimates of the aquifer properties characterize the hydraulic response at an observation well due to pumping of a well, and do not represent the effective properties of the entire aquifer. Similar estimates of aquifer properties could be obtained, if a series of coordinated cross-hole pumping tests were conducted between each pumping and observation well pairs. A model calibrated against water-level transients at a particular monitoring well can be successfully applied to predict the water-levels in the same well; however, the calibrated model is not expected to be

able to predict the water levels at other observation points in the aquifer.

Utilizing this approach on a dataset from the LANL site has indicated that (1) relatively small magnitude water-level transients (less than 1 m) do not preclude our ability to identify the pumping influences associated with individual wells, (2) water-levels at some of the wells exhibit a temporal trend that cannot be directly attributed to any of the pumping wells, (3) the regional aquifer appears to be highly heterogeneous. Future work will include more complicated analytical solutions that can account for partial penetration of pumping and observation wells, aquifer anisotropy, flow three-dimensionality, and leakage from overlying strata. Future work will also include data from additional monitoring wells, coupled inversions (i.e. inversions including data from multiple monitoring wells simultaneously), spatial analysis of aquifer heterogeneity based on tomographic techniques, and characterization of the three-dimensional structure of groundwater flow.

The results also provide guidance for development of more complicated numerical models of the site. Our analyses suggest that numerical models characterizing the aquifer heterogeneity will benefit substantially if the long-term pumping and water-level records are incorporated in the calibration process. The spatial representation of the aquifer heterogeneity should be (1) capable to represent the identified large-scale aquifer structures and (2) with resolution sufficient to represent the differences in the water-level transients at R-15 and R-11/R-28. The model should also be capable of accounting for water-level declines that may not be directly associated with pumping transients. The results show that it is critical to account for the three-dimensional structure of the groundwater flow.

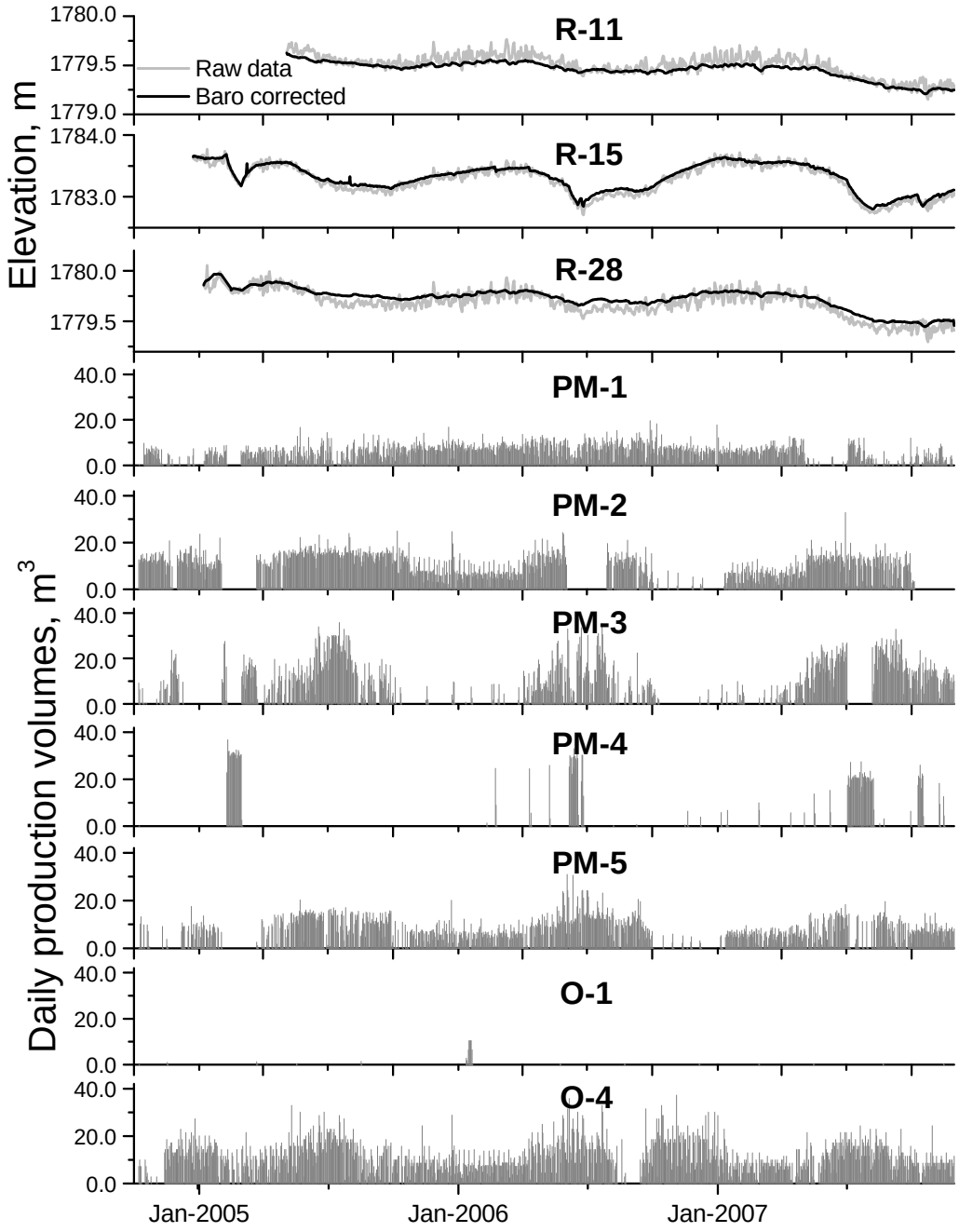


Figure 2.2: Water elevations at monitoring wells and production records for water-supply wells.

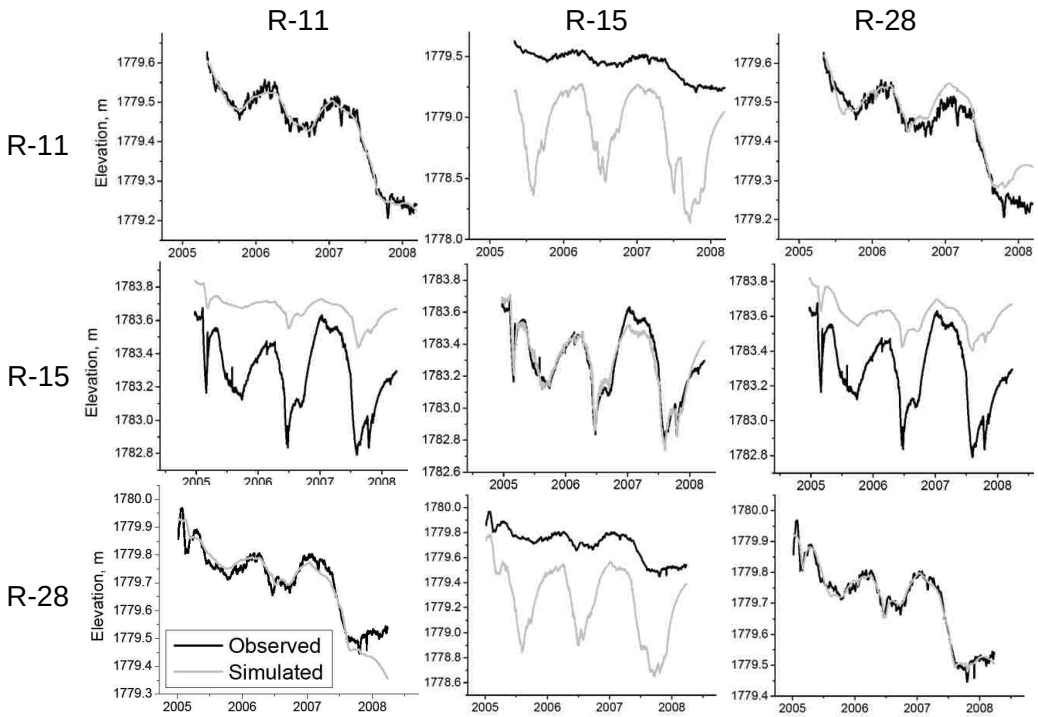


Figure 2.3: Water elevation predictions (gray) using parameters from uncoupled calibrations (considering a single monitoring well and all pumping wells). Observed water elevations (black) are included for reference. Row headings denote the well associated with the water elevation predictions in each plot, while column headings denote the well associated with the model inverted to obtain the hydrogeological properties.

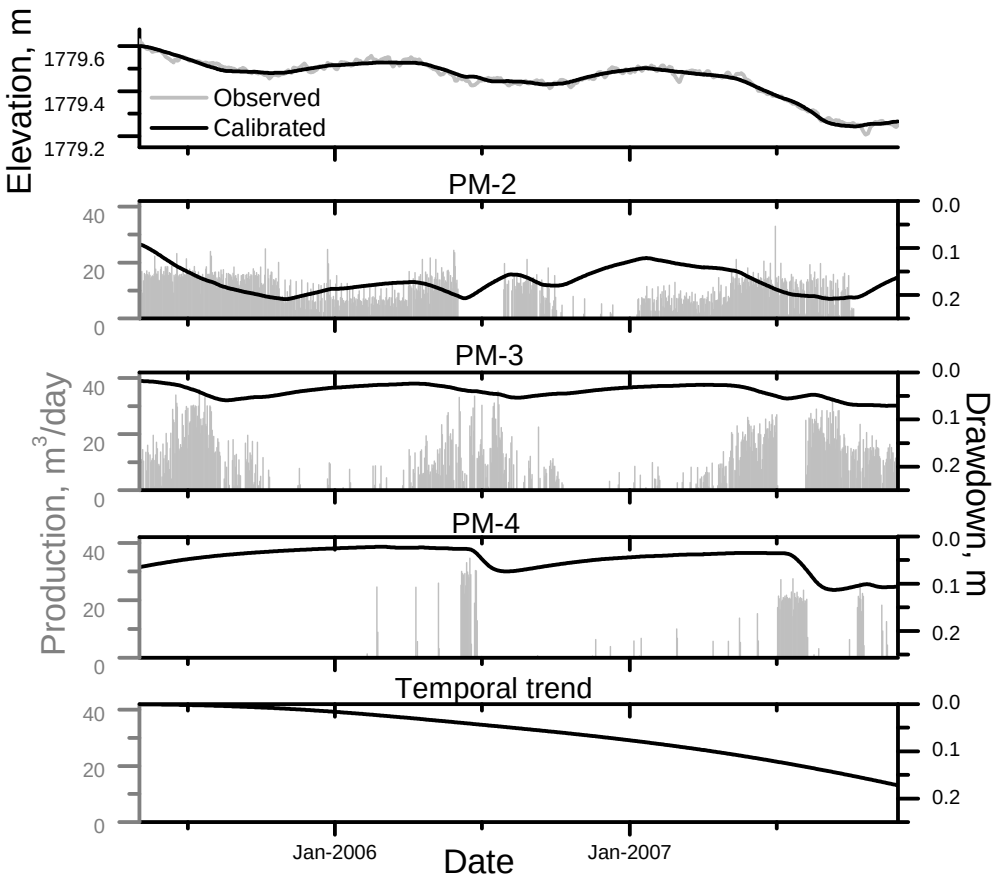


Figure 2.4: Top plot: predicted (black) and observed (gray) water elevations for R-11 model inversion. Bottom plots: predicted drawdown contributions (black lines) from individual pumping wells, plotted with their associated pumping record (gray bars), and temporal trend required to reproduce the total predicted drawdown at R-11.

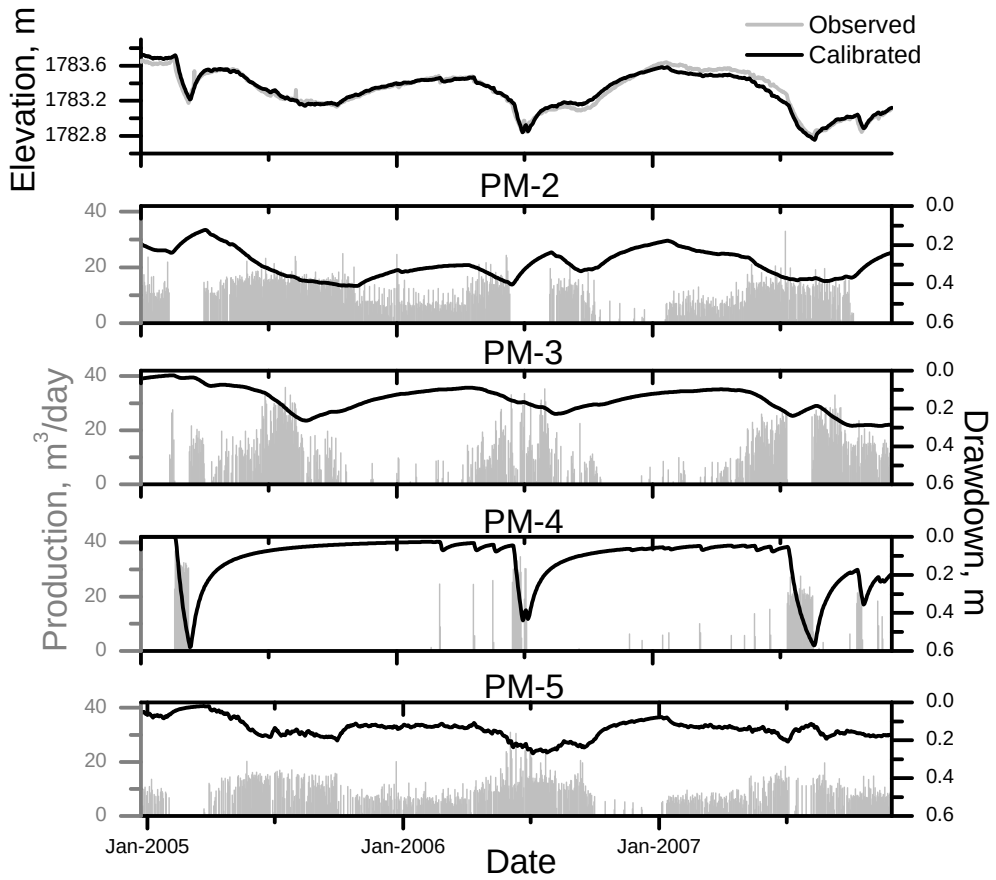


Figure 2.5: Top plot: predicted (black) and observed (gray) water elevations for R-15 model inversion. Bottom plots: predicted drawdown contributions (black lines) from individual pumping wells, plotted with their associated pumping record (gray bars), required to reproduce the total predicted drawdown at R-15.

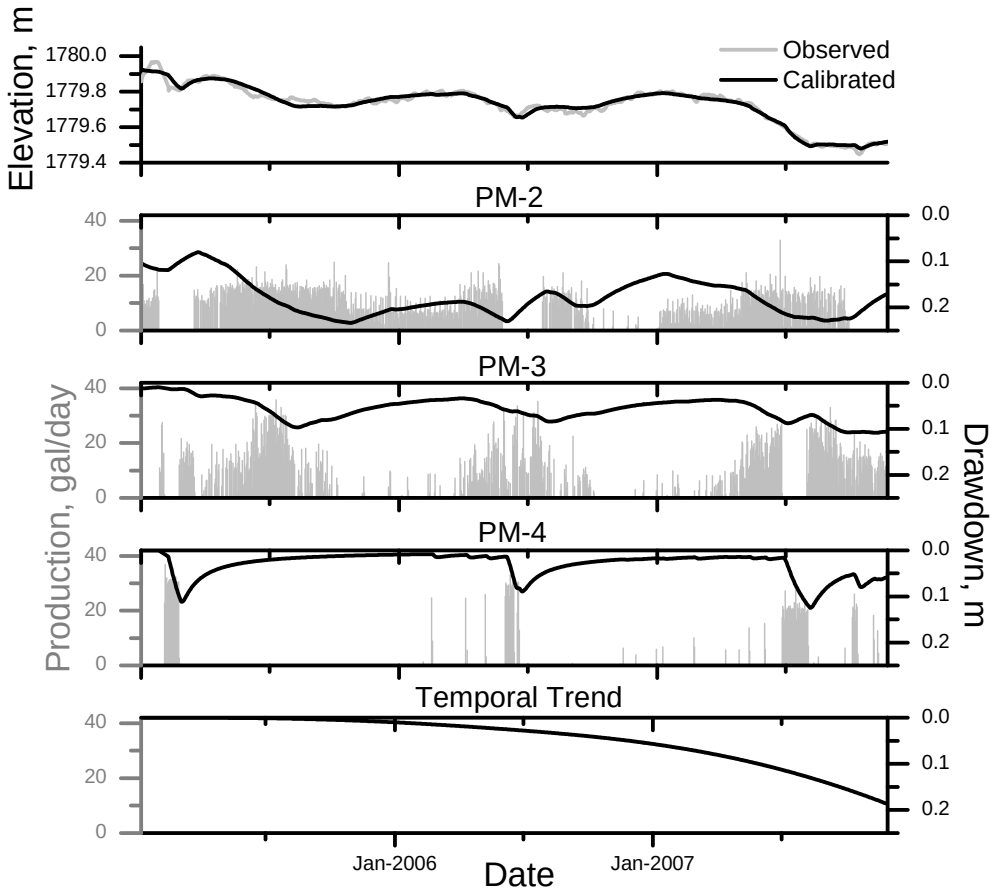


Figure 2.6: Top plot: predicted (black) and observed (gray) water elevations for R-28 model inversion. Bottom plots: predicted drawdown contributions (black lines) from individual pumping wells, plotted with their associated pumping record (gray bars), and temporal trend required to reproduce the total predicted drawdown at R-28.

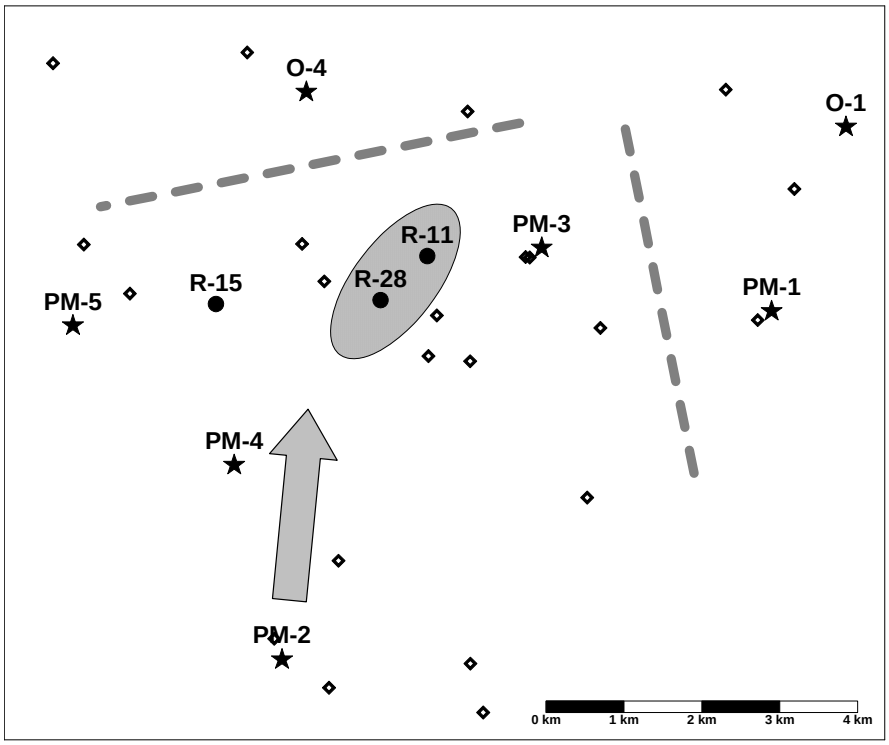


Figure 2.7: Map of observation wells (circles) and water-supply wells (stars) included in the analysis, along with symbols to illustrate the analysis conclusions. Dashed lines indicate locations of apparent hydrogeologic barriers, separating O-4 and PM-1 from the monitoring wells. The arrow illustrates the strong hydraulic connection between PM-2 and the monitoring wells. The grey-filled oval illustrates the apparent existence of a region of similar hydrogeologic properties around R-11 and R-28. Locations of newly completed and planned monitoring wells are indicated by open diamonds.

Hydrogeologic	Monitoring	Pumping well				Sample statistics	
Property	Well	PM-2	PM-3	PM-4	PM-5	Mean	Variance
Log	R-11	3.5	4.0	3.2	-		
Transmissivity	R-15	3.3	3.4	3.2	3.7	3.5	0.1
$log_{10}(T_{eff}[m^2/d])$	R-28	3.5	3.8	3.7	-		
Log	R-11	-1.5	0.1	-1.1	-		
Storativity	R-15	-1.9	-1.5	-1.7	-1.5	-1.2	0.4
$log_{10}(S_{eff}[-])$	R-28	-1.5	-0.4	-1.2	-		
Log	R-11	5.0	3.8	4.3	-		
Diffusivity	R-15	5.2	4.9	4.9	5.2	4.7	0.2
$log_{10}(D_{eff}[m^2/d])$	R-28	4.9	4.3	4.8	-		

Table 2.2: Effective cross-hole parameter estimates and sample statistics from model inversions using data from October 8, 2004 to November 29, 2007.

Hydrogeologic	Monitoring	Pumping well				Sample Statistics	
Property	Well	PM-2	PM-3	PM-4	PM-5	Mean	Variance
Log	R-11	3.5	3.8	3.4	-		
Transmissivity	R-15	3.5	3.5	3.0	3.5	3.5	0.1
$log_{10}(T_{eff}[m^2/d])$	R-28	3.4	3.8	3.9	-		
Log	R-11	-1.3	-0.1	-1.0	-		
Storativity	R-15	-2.0	-1.4	-1.7	-1.3	-1.2	0.3
$log_{10}(S_{eff}[-])$	R-28	-1.3	-0.7	-1.2	-		
Log	R-11	4.8	3.9	4.3	-		
Diffusivity	R-15	5.5	4.8	4.8	4.8	4.7	0.2
$log_{10}(D_{eff}[m^2/d])$	R-28	4.7	4.5	5.2	-		

Table 2.3: Effective cross-hole parameter estimates and sample statistics from model inversions using data from October 8, 2004 to November 18, 2008.

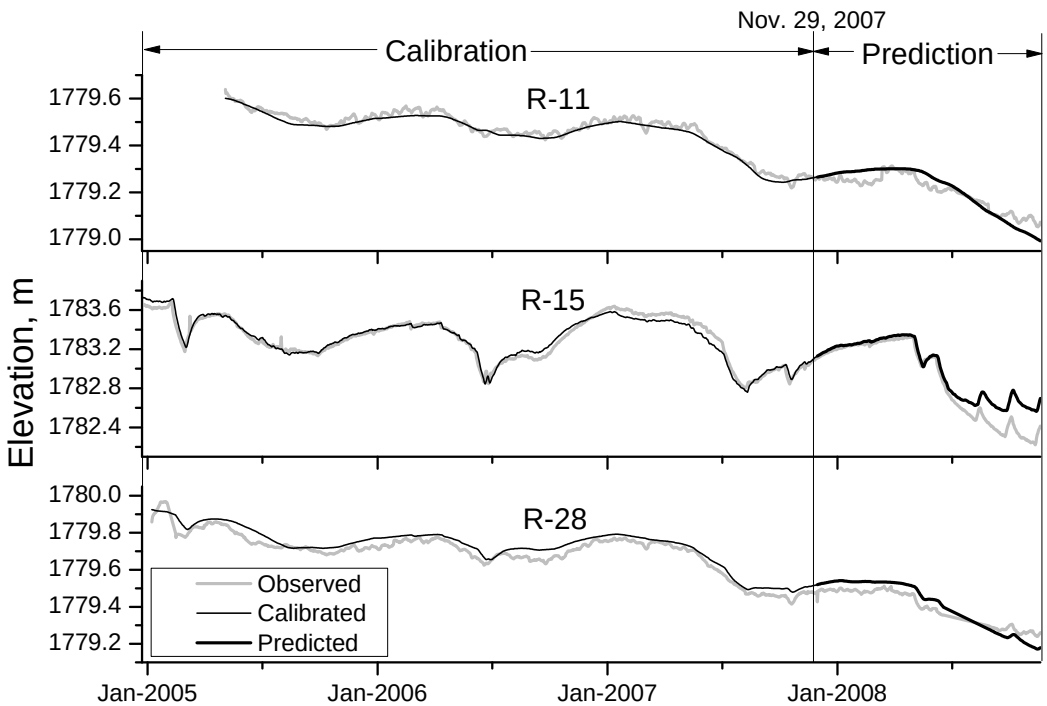


Figure 2.8: Blind forward water-elevation predictions (Predicted) using calibrated parameters from Table 2.2. The best-fit water elevation predictions (Calibrated) associated with the parameters in Table 2.2 are presented also, along with the observed water elevations (Observed).

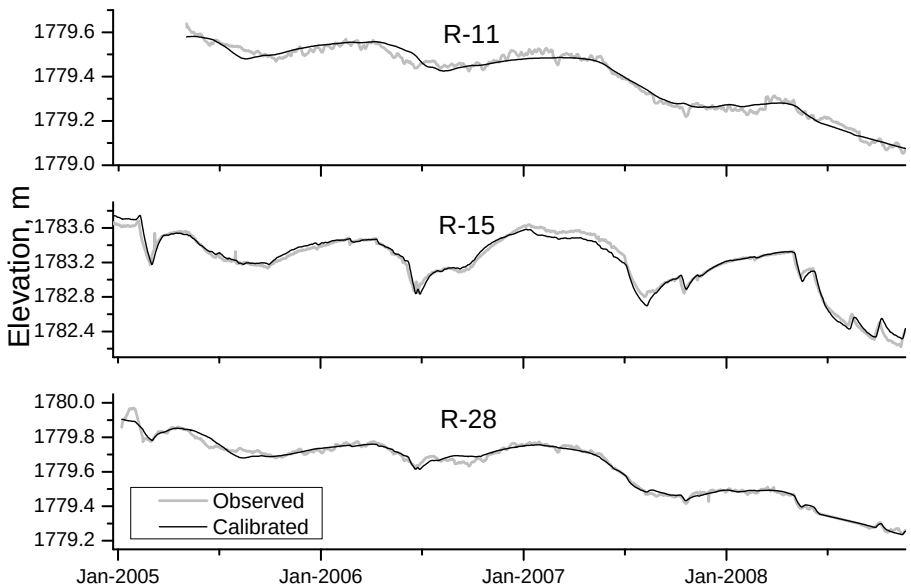


Figure 2.9: Best-fit water-elevation predictions (Predicted) associated with the parameters from Table 2.3 plotted along with the observed water elevations (Observed).

Chapter 3

Hydrogeologic Property Inference Accounting for Early-Time Behaviour of Theis Solution Parameter Estimates

Abstract

It has been demonstrated that during a pumping test in heterogeneous media, drawdown data from different time periods collected at a single location produce different estimates of aquifer properties (e.g. early-time Theis type-curve vs late-time Copper-Jacob methods) and that Theis type-curve inferences are more variable than late-time Cooper-Jacob inferences. This suggests that as the cone of depression from a well propagates towards an observation location, drawdown behavior is affected by many factors (e.g. skin effects, wellbore storage, well friction, and interwell heterogeneities), while after the cone of depression has passed the observation location and quasi-steady state drawdown has been established, convergent hydrogeologic parameters

can be inferred. It has been demonstrated that transmissivity inferences from early-time drawdown data produce time-dependent results, while late-time data inferences are typically found to converge towards a single value. In order to obtain estimates of hydrogeologic properties from highly transient drawdown data at a stationary location, it is necessary to account for the temporal dependency of hydrogeologic parameters. We present an approach that accounts for these effects by proposing an exponential functional representation of the potential temporal dependencies. This is a novel approach for aquifer property inference utilizing transient drawdown data. To better characterize temporal dependency of stationary inferences, we utilize a multi-year dataset containing multi-well transient water-level observations due to transient water-supply pumping; the long-term highly-transient record produces multiple drawdown responses consistent with early-time (transient) and late-time (quasi-steady state) regimes. It is demonstrated that adequate drawdown calibrations can be achieved for multiple monitoring wells considering pressure influences from multiple pumping wells by superposition of the Theis solution using exponential hydrogeologic parameter functions which converge to a single value for transmissivity and are allowed to converge to distinct values for storativity. The convergent transmissivity parameter provides a first estimate of the effective transmissivity of the aquifer, while the distinct values for the late-time convergent storativities provide indicators of inter-well connectivity.

3.1 Introduction

Hydrogeologic property inferences obtained using the Theis type-curve method (*Jacob*, 1940) (Theis method) and the Cooper-Jacob straight-line approximation method (*Cooper and Jacob*, 1946) (Cooper-Jacob method) at a given location have been observed to differ (*Ramey*, 1982; *Butler*, 1990). More recent investigations into this

phenomenon have found that Theis solution parameters (*Theis*, 1935) estimated at a stationary location at various times during a pumping test have been observed to vary at early times converging to stable values at late-times (*Wu et al.*, 2005; *Straface et al.*, 2007). This implies that at early times while the cone of depression is approaching the observation location, the drawdown is affected by many factors, such as: skin effects; well losses; and heterogeneities encountered as the cone of depression extends laterally (*Butler*, 1990), complicating the estimation of stable parameters. However, at late times when quasi-steady state conditions have developed (i.e. when pressure gradients have reached steady state but pressures are transient), the stable parameter estimates are consistent with hydrogeologic property inferences that would be obtained using the Cooper-Jacob method. This implies that the late-time parameter estimates provide "effective" hydrogeologic properties representative of the support scale defined by the distance between the pumping and monitoring wells (*Neuman*, 1990; *Neuman and Di Federico*, 2003).

Variable model parameters indicate the inadequacy of a model to represent the observed phenomenon, as parameters are intended to represent invariant intrinsic properties of the system (homogeneous transmissivity and storativity in the case of the Theis solution). The inadequacy of applying the Theis solution to model typical pumping tests is not a matter of debate, as its limitations are readily apparent by the assumptions required in its derivation (*Theis*, 1935) (e.g. fully penetrating well, infinite lateral extents, homogeneous properties, unperturbed initial conditions, confined aquifer). Recognizing these limitations, the question becomes whether or not the model is useful. We agree with previous researchers that the Theis solution is useful for obtaining apparent hydrogeologic properties that characterize the groundwater transport if late-time drawdown data is used consistent with the Cooper-Jacob method (*Butler*, 1990; *Meier et al.*, 1998; *Sanchez-Vila et al.*, 1999; *Knudby and Carrera*, 2006; *Trinchero et al.*, 2008). As noted by *Butler* (1990) in reference to the use of the Cooper-Jacob method, the advantage of drawing inferences from late-time

drawdown data is that the estimated parameters will be independent of the numerous early-time effects that can influence the drawdown at the initial stages of expansion of the cone of depression.

Making late-time hydrogeologic inferences is not always possible, however, as it may not be feasible to continue a pumping test for a sufficient duration to allow quasi-steady state conditions to develop. Or, in cases where an existing water-supply and water-level elevation dataset is available from a municipal water supply network, quasi-steady state may not be reached due to a high frequency of cycling multiple supply wells on and off to: meet shifting demand; to take advantage of lower cost off-peak electrical rates; and perform well maintenance and/or repair. In this paper, we present an approach that allows late-time quasi-steady state inferences to be obtained from transient drawdown datasets by considering the behaviour of Theis solution parameters at early times.

The proposed approach is demonstrated on a long-term highly-transient drawdown record collected at the Los Alamos National Laboratory (LANL) site where the water-level transients result from multi-well municipal water-supply pumping. The pumping regimes are highly transient, cycling diurnally and seasonally, including variations due to maintenance, repair, and shifting supply loads within the network. As a result, the drawdown at monitoring wells within the network cannot fully attain quasi-steady state as new pressure influences (cones of depression and impression) begin to propagate through the aquifer as the pumping wells cycle on and off (*Harp and Vesselinov*, 2009). The use of a long-term dataset containing multiple pressure influence cycles has certain advantages, such as: reduction of measurement errors; improved characterization of the hydraulic response allowing the refinement of hydrogeologic inferences; and the lack of the expense and coordination of a conventional pumping test. We demonstrate the inference of hydrogeologic properties from this dataset by considering the transient early-time behaviour of Theis solution parame-

ters.

As the approach presented here relies on observations, numerical experiments, and analytical investigations of many previous researchers, a review of these bodies of research will be presented in the background section. The proposed approach for accounting for early time parameter variations will be presented in the methodology section. The approach will be demonstrated on the LANL dataset in the results section.

3.2 Background

Theis (1935) introduced the inference of hydrogeologic properties from transient drawdown observations using an analytical solution derived by applying the analogous concepts of heat conduction to groundwater flow. The Theis type-curve method (Theis method), postulated by Theis and described by *Jacob* (1940), was developed from this work. *Cooper and Jacob* (1946) simplified this approach using an approximation to the Theis solution valid at late times when a quasi-steady state regime is established (Cooper-Jacob method), eliminating the use of a Theis type curve.

It has been recognized that hydrogeologic property inferences based on the Theis method and Cooper-Jacob method differ (*Ramey*, 1982). This is due to the fact that the inference methods emphasize properties in different regions of the aquifer. The Theis method considers the entire drawdown curve, often causing the inadvertent emphasis of the interval of greatest curvature located during early times. As indicated by *Butler* (1990), drawdown at early times can be affected by many factors, including local heterogeneities near the pumping well and well skin and pumping storage affects, creating greater variability in Theis method inferences. The Cooper-Jacob method ignores early times, providing information on the properties of the aquifer within a ring formed by the outward moving front of the cone of depression

during the time interval under consideration. At late time, when the Cooper-Jacob approximation is valid, the region included in this ring can be large. *Butler* (1990) demonstrates that the difference between inferences obtained from the Theis and Cooper-Jacob methods depend on the level of aquifer heterogeneity and the distance between the pumping well and the observation location. The inferences become more similar as the level of heterogeneity decreases and the distance increases.

Meier et al. (1998) explore the use of the Cooper-Jacob approximation to infer effective transmissivity (T_{eff}) from the estimated transmissivity parameter $\hat{T}$ and provide indications of hydraulic connectivity by evaluating the estimated storativity parameter $\hat{S}$ in heterogeneous aquifers. Consistent with theoretical findings of *Butler* (1990), *Meier et al.* (1998) present cases where field data demonstrate that although small-scale (point) estimates of transmissivity T are highly variable, values of $\hat{T}$ obtained from the Cooper-Jacob method are relatively constant. Furthermore, *Meier et al.* (1998) demonstrate that corresponding values of $\hat{S}$ are typically highly variable, even though the storativity in the field is believed to be relatively constant. *Meier et al.* (1998) investigate this phenomena by performing numerical experiments with heterogeneous transmissivity fields and homogeneous storativity fields, producing similar values for $\hat{T}$ and variable values for $\hat{S}$ consistent with the field cases.

The reason for this paradoxical result can be explained by examining the equation for estimating T from the Cooper-Jacob method; $\hat{T} = 2.3Q/4\pi I$, where Q is a constant pumping rate and I is the slope of the late-time drawdown with respect to the (base 10) log of time (i.e. $I = (s_2 - s_1)/(log\ t_2 - log\ t_1)$, where s is drawdown and t is time). This equation demonstrates that $\hat{T}$ is dependent on the rate of drawdown decline, which is dependent on the choice of t_2 and t_1. Considering only the late-time drawdown where the data approximate a straight line with respect to log time, in accordance with the Cooper-Jacob method, means that $\hat{T}$ will approximate T_{eff} described by the rate of drawdown after the drawdown cone of depression has passed

the monitoring well. Storativity estimates using the Cooper-Jacob method (defined as $\hat{S} = 2.25Tt_0/r^2$, where r is the distance from the pumping well to the observation point and t_0 represents the time-axis intercept of a line drawn through the late-time drawdown), on the other hand, are dependent on the variability of T between the pumping well and the front of the cone of depression. Although the heterogeneity between the pumping and monitoring well does not affect the slope of the late-time drawdown used to determine $\hat{T}$, it can affect $\hat{S}$ as the time-axis intercept (t_0) is dependent on the arrival of the cone of depression at the monitoring well. If a region of high T connects the monitoring well and the pumping well, the value of t_0 will be relatively small and vice-versa. As noted by *Sanchez-Vila et al.* (1999), the Cooper-Jacob method interprets an early/late arrival of a drawdown cone of depression as low/high storativity. This explains the high variability of $\hat{S}$ in the presence of T heterogeneity between the pumping well and the monitoring well, even in cases where S is known to be constant.

Research by *Meier et al.* (1998) demonstrate from a numerical analysis that $\hat{T}$ estimated from a simulated pumping test (radial flow) is close to T_{eff} for parallel flow for an area of influence for multilognormal stationary (geostatistically homogeneous) T fields (the S field is assumed uniform in all cases). While *Meier et al.* (1998) also demonstrated that this can be true for nonmultigaussian fields, *Sanchez-Vila et al.* (1996) demonstrate that this is not true in general. Similar to findings by *Butler* (1988), who demonstrated that $\hat{S}$ depends on transmissivities between the pumping well and the front of the cone of depression, *Meier et al.* (1998) find that $\hat{S}$ depends on transmissivities between and nearby the well and the observation point.

Sanchez-Vila et al. (1999) verify these conclusions using an analytical approximation to the groundwater flow equation. They demonstrate analytically that $\hat{T}$ is independent of spatial location. They also demonstrate that storativity estimates will provide an indication of the local deviations of T from its large-scale geometric

mean (denoted as T_G) representing the equivalent geostatistically homogeneous T field. If T in a specific location is less than T_G, $\hat{S}$ will be larger than the true value of S and vice-versa. They also show that the geometric mean of $\hat{S}$ values is an unbiased estimator of S.

Recognizing that individual values of $\hat{S}$ estimated using the Cooper-Jacob approximation are not representative of S, *Knudby and Carrera* (2006) refer to the value as apparent storativity, S_a. They demonstrate that the apparent diffusivity, defined as $D_a = \hat{T}/S_a$, is affected in a similar way as S_a to aquifer heterogeneity and contains information on general characteristics of the system due to its dependence on $\hat{T}$. Furthermore, *Knudby and Carrera* (2006) demonstrate that D_a correlates well with indicators of flow and transport connectivity, thereby encapsulating information on these two process dependent forms of connectivity in a single parameter.

Trinchero et al. (2008) explored the relationship between the flow and transport connectivity indicators described by *Knudby and Carrera* (2006), deriving an analytical relation between the estimated effective homogeneous porosity $\hat{\phi}$ (transport connectivity indicator) and $\hat{S}$ (flow connectivity indicator) of a heterogeneous T field. They determined that $\hat{\phi}$ depends on a weighted function of actual transmissivity T and the estimated storativity parameter $\hat{S}$ along the flow line.

In contrast to *Meier et al.* (1998), *Sanchez-Vila et al.* (1999), *Knudby and Carrera* (2006), and *Trinchero et al.* (2008), *Wu et al.* (2005) explore the effect of the homogeneity assumption of the Theis solution on parameter estimates at both early and late times for cases with both heterogeneous T and S fields. Conceptualizing T and S as spatial stochastic processes in the equation of flow, *Wu et al.* (2005) derive the mean flow equation of a heterogeneous aquifer as

$$T_{eff}\nabla^2\overline{h} = S_{eff}\frac{\partial\overline{h}}{\partial t} \tag{3.1}$$

where $\overline{h}$ is the ensemble mean head, t is the time, T_{eff} is the ensemble effective transmissivity defined as

$$T_{eff} = \overline{T} + E\langle T'\nabla h'\rangle[\nabla\overline{h}]^{-1}, \tag{3.2}$$

and S_{eff} is the ensemble effective storativity defined as

$$S_{eff} = \overline{S} + E\langle S'\nabla h'\rangle[\nabla\overline{h}]^{-1}. \tag{3.3}$$

where the over bar and prime denote the mean and perturbation of the variable, respectively. T_{eff} and S_{eff} are denoted as ensemble effective parameters as they will produce the ensemble mean head $\overline{h}$ as a convergent average for a set of realizations of heterogeneity based on the stochastic parameters $T = \overline{T} + T'$ and $S = \overline{S} + S'$. As indicated by *Wu et al.* (2005), in order for the ensemble mean head to equal the spatially averaged head of a single realization of heterogeneity, the field must contain an adequate sampling of the heterogeneity (i.e. the field must be ergodic). As traditional pumping tests typically estimate T_{eff} and S_{eff} based on one or a small number of point estimates of head, which will not equal the spatially averaged head in general, $\hat{T}$ and $\hat{S}$ will not provide estimates of effective properties in an ensemble sense in general.

Wu et al. (2005) performed numerical experiments using synthetic aquifers with multilognormal heterogeneous T and S fields. They observe that at early time, $\hat{T}$ estimates at different locations are highly variable, while at large times (when the Cooper-Jacob approximation is valid) values of $\hat{T}$ converge to a value close to T_G as the cone of depression expands for the multilognormal fields considered. The late-time convergent values of $\hat{T}$ presented by *Wu et al.* (2005) are consistent in character

with Cooper-Jacob transmissivity estimates presented by *Butler* (1990), *Meier et al.* (1998), and *Sanchez-Vila et al.* (1999). As the considered transmissivity field in *Wu et al.* (2005) is multilognormal, $T_G = T_{eff}$. In cases where the transmissivity is nonmultigaussian, the significance of $\hat{T}$ is less certain (*Sanchez-Vila et al.*, 1999), however, we assume that it is a good first estimate of T_{eff}. Values of $\hat{S}$ do not converge to a single value, but stabilize relatively quickly to values predominantly dependent on the heterogeneity between the pumping well and the given monitoring location. These distinct late-time convergent values for $\hat{S}$ are consistent with the variable Cooper-Jacob estimates presented in *Meier et al.* (1998) and *Sanchez-Vila et al.* (1999). As the late-time convergent values for $\hat{T}$ and $\hat{S}$ presented by *Wu et al.* (2005) are equivalent to values that would be obtained by the Cooper-Jacob method and are consistent with the behavior of Cooper-Jacob property estimates presented in *Meier et al.* (1998) and *Sanchez-Vila et al.* (1999), these values can be interpreted based on the discussions of *Meier et al.* (1998) and *Sanchez-Vila et al.* (1999) summarized above.

Utilizing a cross-correlation analysis, *Wu et al.* (2005) determine that at early times, the head at a monitoring location will be strongly correlated to S between the monitoring location and the pumping well, and only weakly negatively correlated to the T in the same region. Furthermore, at late times, the spatial distribution of observed head and S correlation is similar, albeit greatly diminished, while the correlation between T and observed head is greatly increased throughout the domain, with the largest (positive) correlations in regions behind the monitoring location and pumping well (*Wu et al.*, 2005, Figure 10). This indicates that it is likely that $\hat{T}$ at late times can be affected by anomalies and boundaries.

In contrast to the four field cases discussed by *Meier et al.* (1998) (i.e. Grimsel test site, Switzerland (*Frick*, 1992); El Cabril site, Spain (Bureau de Recherches Géologiques et Miniéres); Horkheimer Insel site, Germany (*Schad and Teutsch*, 1994);

and Columbus Air Force Base, U.S.A. (*Herweijer and Young*, 1991)), *Straface et al.* (2007) observe a lack of similar slope for drawdown vs log time at late times from pumping tests near Montalto Uffugo Scalo, Italy, indicating that the Cooper-Jacob straight-line approximation for late-time drawdown will not be valid in all cases. Based on their analysis of these pumping tests, *Straface et al.* (2007) question the validity of using traditional pumping tests to estimate meaningful hydrogeological parameters, but do acknowledge that these results can provide quick inexpensive first estimates. Furthermore, they suggest that these first estimates can be useful as starting parameters for a tomographic inversion of the same dataset.

Harp and Vesselinov (2009) demonstrate an approach to identify and decompose the pressure influences at a monitoring location using the Theis solution. Their approach is demonstrated on the same dataset as in the current research. As the objective of the research in *Harp and Vesselinov* (2009) is the decomposition of pressure influences, attempts are not made to account for early time behaviour of the Theis solution parameters, and constant and distinct values are applied to pumping/monitoring well pairs. Therefore, the parameter estimates are not considered representative of the hydrogeologic properties of an aquifer, but are considered to characterize the hydraulic response at a monitoring location due to pumping a single well, analogous to parameter estimates that would be obtained from conventional pumping tests. Therefore, constant and distinct interwell Theis solution parameters are estimated concurrently in a model inversion.

The current research has extended the inverse framework presented in *Harp and Vesselinov* (2009) for the purpose of inferring late-time hydrogeologic properties considering the extensive body of research presented above. While the current approach is demonstrated on the long-term dataset from the LANL site, providing the decomposition of pressure influences similar to the approach presented in *Harp and Vesselinov* (2009), the current approach could also be applied to a conventional

pumping test to more appropriately account for the early time behaviour of the Theis solution parameters. Furthermore, this could be particular useful to obtain late time hydrogeologic inferences from conventional pumping tests that were not conducted for a sufficient length of time to establish quasi-steady state conditions.

3.3 Methodology

We develop the proposed approach here considering the existence of multiple pumping wells influencing multiple monitoring wells within an aquifer. In this way, the methodology applies the to application presented here utilizing the LANL dataset. In more simplified cases (e.g. a conventional cross-hole interference test), the presented equations will reduce to simplified non-superposed solutions.

We begin with the Theis solution of the flow equation ($T\nabla^2 h = S\partial h/\partial t$), defined as

$$s(t) = \frac{Q}{4\pi T} W(u) = \frac{Q}{4\pi T} W\left(\frac{r^2 S}{4Tt}\right), \tag{3.4}$$

where $s(t)$ is the predicted drawdown at time t since the pumping commenced (i.e. $h(t) - h(0)$), Q is the pumping rate, T is the transmissivity, $W(u)$ is the negative exponential integral ($\int_u^\infty e^{-y}/y\, dy$) referred to as the well function, u is a dimensionless variable, r is radial distance from the pumping well, and S is the storativity. Multiple pumping wells and variable rate pumping periods can be included in the Theis solution by employing the principle of superposition as

$$s(t) = \sum_{i=1}^{N} \sum_{j=1}^{M_i} \frac{Q_{i,j} - Q_{i,j-1}}{4\pi T} W\left(\frac{r_i^2 S}{4T(t - t_{Q_{i,j}})}\right), \tag{3.5}$$

where N is the number of pumping wells (sources), M_i is the number of pumping periods (i.e. the number of pumping rate changes) for pumping well i, $Q_{i,j}$ is the pumping rate of the ith well during the jth pumping period, and $t_{Q_{i,j}}$ is the time when the pumping rate changed at the ith well to the jth pumping period. The drawdown calculated by equation (3.5) represents the cumulative influence at a monitoring location of the N pumping wells at distances $r_i,\ i = 1, \ldots, N$ from the monitoring location.

Equations (3.4) and (3.5) are only valid under the assumption of homogeneity. If a system is homogeneous, then T and S in equations (3.4) and (3.5) will be equivalent to T_{eff} and S_{eff}, respectively. If the system is heterogeneous, this will only be true in an ensemble mean sense. In this case, the Theis solution can be expressed as

$$\overline{s}(t) = \frac{Q}{4\pi T_{eff}} W(u) = \frac{Q}{4\pi T_{eff}} W\left(\frac{r^2 S_{eff}}{4T_{eff}t}\right) \tag{3.6}$$

which is the solution to equation (3.1) (*Wu et al.*, 2005) where $\overline{s}(t)$ is the ensemble mean drawdown. Invoking superposition with equation (3.6), an ensemble mean drawdown equation analogous to equation (3.5) can be expressed as

$$\overline{s}(t) = \sum_{i=1}^{N} \sum_{j=1}^{M_i} \frac{Q_{i,j} - Q_{i,j-1}}{4\pi T_{eff}} W\left(\frac{r_i^2 S_{eff}}{4T_{eff} * (t - t_{Q_{i,j}})}\right). \tag{3.7}$$

As water elevations recorded at monitoring wells in an aquifer system are merely point samples from a single realization of heterogeneity, and not ensemble mean values of multiple realizations or spatial averages of an ergodic field, application of equations (3.6) and (3.7) are invalid for cross-hole interference tests. Recognizing this theoretical limitation of applying the Theis solution (or the Cooper-Jacob approximation) to heterogeneous fields to infer representative effective properties,

researchers have investigated what information is contained in the hydrogeologic parameter estimates (*Meier et al.*, 1998; *Sanchez-Vila et al.*, 1999; *Wu et al.*, 2005; *Knudby and Carrera*, 2006; *Trinchero et al.*, 2008). We propose that although the Theis solution parameters will not provide precise representations of hydrogeological properties, the analytical framework of the Theis solution can provide initial estimates of the effective transmissivity and indications of connectivity.

Recognizing that datasets containing early-time drawdown outside of the Cooper-Jacob domain require consideration of the behaviour of parameter estimates at early times, we define an estimated drawdown $\hat{s}(t)$ as

$$\hat{s}(t) = \sum_{i=1}^{N} \sum_{j=1}^{M_i} \frac{Q_{i,j} - Q_{ij-1}}{4\pi \hat{T}_i} W\left(\frac{r_i^2 \hat{S}_i}{4\hat{T}_i * (t - t_{Q_{i,j}})}\right), \tag{3.8}$$

where we propose an exponential functional form for $\hat{T}_i$ and $\hat{S}_i$ to account for the early-time behavior of Theis solution parameters as

$$\hat{T}_i(t) = T_a e^{c_{T,i}/(t - t_{Q_{i,j}})} \quad c_T \geq 0, \tag{3.9}$$

and

$$\hat{S}_i(t) = S_{a,i} e^{c_{S,i}/(t - t_{Q_{i,j}})}, \tag{3.10}$$

where T_a is the apparent transmissivity, providing a first estimate for T_{eff}, $S_{a,i}$ is the apparent storativity associated with the ith pumping well providing an indication of connectivity between the ith pumping well and the monitoring location (*Knudby and Carrera*, 2006), and $c_{T,i}$ and $c_{S,i}$ are constants describing the time dependent

slope of the transmissivity and storativity functions, respectively, with respect to the ith pumping well. T_a and S_a provide late time inferences consistent with values that would be obtained using the Cooper-Jacob method. Based on results of numerical experiments performed by *Wu et al.* (2005), we constrain $c_T \geq 0$, indicating that $\hat{T}$ values from early time portions of drawdown data are expected to be higher than late time convergent values. This may be explained by the early-time negative correlation between head and transmissivity (*Wu et al.*, 2005) and/or the existence of high conductivity interwell pathways as described by *Herweijer* (1996). Other possible explanations for the need for time-dependent hydrogeologic parameters are well-bore storage and leakage effects known to exist at the site (*McLin*, 2005, 2006a,b). We demonstrate in this Chapter that the exponential functional forms for $\hat{T}$ and $\hat{S}$ proposed here in equations (3.9) and (3.10) provide adaquate calibrations of model simulations to observed water levels and can provide late-time Cooper-Jacob parameter estimates from highly transient datasets.

Substituting equations (3.9) and (3.10) into equation (3.8) produces

$$\hat{s}(t) = \sum_{i=1}^{N} \sum_{j=1}^{M_i} \frac{Q_{i,j} - Q_{ij-1}}{4\pi T_a e^{c_{T,i}/(t-t_{Q_{i,j}})}} W\left(\frac{r_i^2 S_{a,i} e^{c_{S,i}/(t-t_{Q_{i,j}})}}{4T_a e^{c_{T,i}/(t-t_{Q_{i,j}})} * (t - t_{Q_{i,j}})} \right), \tag{3.11}$$

As the calibration targets in the model inversions presented here are water elevations as opposed to drawdowns, we define the predicted water elevation $\hat{h}(t)$ at time t as

$$\hat{h}(t) = \hat{h}_o - \hat{s}(t) \tag{3.12}$$

where $\hat{h}_o = \hat{h}(0)$ and is defined as the initial predicted water elevation at the monitoring well. As defined by the Theis solution, h_o is the head at the time that a

perturbation commences. As pumping of the regional aquifer began at the LANL site over 50 years ago, it is reasonable to assume that the influence of the earlier pumping has propagated through the system and/or dissipated. However, more recent pumping rate changes preceding pressure transient records at monitoring locations need to be considered. In order to account for residual effects of pumping prior to monitoring, simulations are started in advance of pressure transient records, thereby setting $t = 0$ at the beginning of the simulation, and not when monitoring commenced.

Model calibration is performed using a Levenberg-Marquardt approach (*Levenberg*, 1944; *Marquardt*, 1963) where the objective function can be defined as

$$\mathbf{\Phi} = \sum_{i=1}^{m} \sum_{j=1}^{n_i} [h_i(t_j) - \hat{h}_i(t_j)]^2 \tag{3.13}$$

where m is the number of monitoring wells considered, n_i is the number of head observations for the ith monitoring well, and $h_i(t_j)$ are the head observations for the ith monitoring well included as calibration targets where j is an observation time index.

The simulation of the drawdowns is performed using the WELLS code (written by *V.V. Vesselinov*, 1992), which implements equation (3.11). The calibration is performed using *PEST* (*Doherty*, 2004).

3.4 Site Data

The regional aquifer beneath the LANL site is a complex stratified hydrogeologic structure which includes unconfined zones (under phreatic conditions near the regional water table) and confined zones (deeper zones) (*Vesselinov*, 2004a,b). The

three monitoring wells considered in this analysis are screened near the top of the aquifer with an average screen length of 11 meters. The water-supply wells partially penetrate the regional aquifer with screens that also begin near the top of the aquifer, but penetrate deeper with an average screen length of 464 meters. Nevertheless, field tests demonstrate that most of the groundwater supply is produced from a relatively narrow section of the regional aquifer that is about 200-300 m below the regional water table (*Los Alamos National Laboratory*, 2008a). Implicit in the use of the Theis solution is the assumption that groundwater flow is confined and two-dimensional. We assume that this is a justifiable assumption here given the small magnitude of observed drawdowns (less than 1 m at the monitoring wells and less than 20 m at the water-supply wells), the relative long distances between supply and monitoring wells (more than 1 km; Table) compared to the effective aquifer thickness (about 200-300 m).

Water-level fluctuations (pressure transients) are automatically monitored using pressure transducers. The pressure and water-supply pumping records considered here are collected from 3 monitoring wells (R-11, R-15 and R-28) and the 7 water-supply wells (PM-1, PM-2, PM-3, PM-4, PM-5, O-1, and O-4) located within the LANL site. Figure 3.1 displays a map of the location of the wells. Figure 3.2 presents the pressure and production records for the monitoring wells and water-supply wells, respectively.

The water-level observation data considered here span approximately four years, commencing on or shortly after the date of installation of pressure transducers (May 4, 2005 for R-11; December 23, 2004 for R-15; January 7, 2005 for R-28), all terminating on November 18, 2008. The barometric pressure fluctuations are removed using constant coefficient methods using 100% barometric efficiency for all monitoring wells (*Los Alamos National Laboratory*, 2008b). Although the pressure transducers collect observations every 15 minutes, this dataset is reduced to single daily observations by

using the earliest recorded measurement for each day. Some daily observations have been excluded due to equipment failure. Considered pumping records for all pumping wells begin on October 8, 2004 and terminate on November 18, 2008. The pumping record precedes the water-level calibration data to account for water-level transients due to pumping variations before the water-level data collection commenced. For additional information on the site and dataset, refer to *Harp and Vesselinov* (2009).

In the applied computational framework, forward model run times for predicting water elevations at R-11, R-15, and R-28 for approximately four years (from October 8, 2004 to November 18, 2008) are each approximately 9 seconds on a 3.0 GHz Intel processor. Inversions initiated with uniform initial parameter values require approximately 600 model runs and, using a single processor, are performed for approximately 1 hour and 40 minutes.

3.5 Results and Discussion

Figure 3.3 presents the calibrated drawdown from the water-supply wells for monitoring wells R-11, R-15, and R-28. These plots indicate that adequate calibrations can be achieved by accounting for the early time behaviour of parameter estimates using exponential functional forms, conceptually consistent with numerical investigations by *Wu et al.* (2005) and parameter inferences from field collected hydrographs by *Straface et al.* (2007).

Figure 3.4 presents the percentage of superposed drawdown calculations performed within the Cooper-Jacob domain as the length of the data record increases. In this way, each superposed calculation is weighted equally in the determination of the percentage. The Cooper-Jacob approximation is assumed to be valid when $u < 0.03$, or when the approximation error is less than 1%. This figure demonstrates that as the data record length increases, the percentage of the superposed drawdown

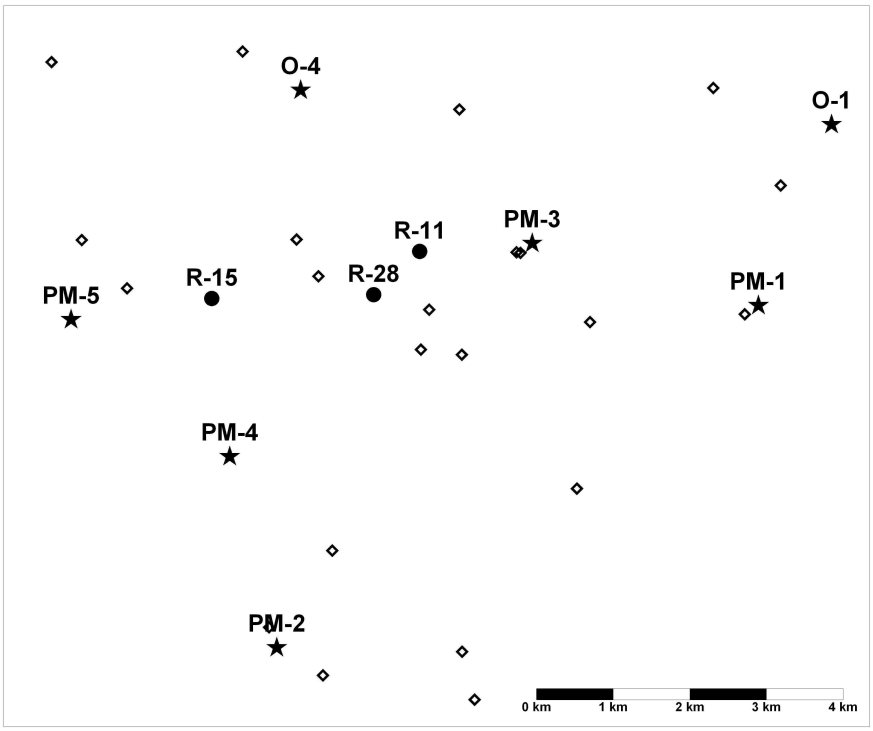


Figure 3.1: Map of monitoring wells (circles) and water-supply wells (stars) included in the analysis. Locations of newly completed and planned monitoring wells are indicated by open diamonds.

calculations performed within the domain of a valid Cooper-Jacob approximation is generally increasing as longer time intervals since the initiation of pumping perturbations are included in the superposed drawdown calculation. By inspecting Figure 3.4, it is apparent that the R-15 superposed drawdown calculations are dominated by late-time drawdown data, while R-11 and R-28 are still dominated by calculations outside of the Cooper-Jacob domain. We can therefore conclude that it is necessary to consider the early-time pre-stabilization of the hydrogeologic parameters, especially for R-11 and R-28. This could have also been evaluated as the percentage of superposed drawdown calculated within the Cooper-Jacob domain, however, this would apply heavier weight to the calculations within the Cooper-Jacob domain, as

larger drawdowns occur at later times. As each component of the superposed drawdown calculation will affect the parameter estimation, irregardless of the magnitude, basing the percentage on the number of superposed calculations is appropriate.

Figure 3.5 presents the estimated transmissivity functions (equation (3.9)) for R-11, R-15, and R-28. The functions are plotted for values slightly greater than four years to include parameter values used in the model runs. It is apparent that for all three monitoring wells, $\hat{T}(t)$ estimates are converging towards a single value, T_a, as constrained by the inversion. Figure 3.6 presents the estimated storativity functions (equation (3.10)) on a single plot. It is apparent that in general the storativity functions converge quickly to distinct values.

Table 3.1 presents the estimated parameters associated with the transmissivity and storativity functions plotted in Figures 3.5 and 3.6. As the apparent transmissivity and storativity values are equivalent to late-time Cooper-Jacob estimates, they can be interpreted by methods discussed in *Meier et al.* (1998) and *Sanchez-Vila et al.* (1999). As constrained in the inversion, all transmissivity functions converge to a single value ($1 \times 10^{3.57}$ m/d), which can be considered a first estimate of T_{eff} (*Meier et al.*, 1998; *Sanchez-Vila et al.*, 1999; *Wu et al.*, 2005). This value is similar to the average of the transmissivity estimates ($1 \times 10^{3.5}$ m/d) obtained in *Harp and Vesselinov* (2009). While previous pumping tests were performed at PM-2 and PM-4 (*McLin*, 2005, 2006b), cross-hole transmissivity estimates were not obtained as pressure influence was not observed at R-11, R-15, or R-28 during these tests. This may have been due to the duration of the PM-2 and PM-4 pumping tests, 25 and 21 days, respectively, and the shallow penetration of the monitoring wells compared to the deeper penetration of the pumping wells considering the leaky-confined nature of this aquifer. As a result, we are unable to compare our parameter estimates with cross-hole pumping test estimates.

Values of S_a indicate the level of connectivity between the monitoring and pump-

ing well (*Meier et al.*, 1998; *Sanchez-Vila et al.*, 1999). Large/small values of S_a indicate a connected path or region of relatively low/high transmissivity between the monitoring and pumping wells. This is evaluated further by calculating the time of arrival of the cone of depression at the monitoring well, t_0, from the Cooper-Jacob storativity estimation equation, also tabulated in Table 3.1. Values of t_0 provide indications of interwell connectivity that are not normalized by interwell distance. These values indicate the quick pressure response of R-15 to PM-2, PM-4, and PM-5, further demonstrating the high level of connectivity between these wells.

As expected, we find values of S_a to be more similar for R-11 and R-28 than for R-15 given the relative locations of these monitoring wells. This is consistent with the findings of *Harp and Vesselinov* (2009), based on an evaluation of constant cross-hole parameter estimates from a related approach. Figure 3.6 and Table 3.1 indicate that the highest connectivity is between R-15 and PM-2, and R-15 and PM-4. This is similar to the findings of *Harp and Vesselinov* (2009), where high connectivity paths were indicated between the monitoring wells and PM-2. Values of S_a for PM-3 decrease with distance, suggesting that connectivity increases with distance in a generally westerly direction from PM-3. A map with schematic representations of these conclusions is presented in Figure 3.7.

A decomposition of the pressure influences from the pumping wells at the monitoring wells also resulted from this research. These results are similar to the decomposition analysis of this dataset presented in *Harp and Vesselinov* (2009), and therefore are not presented.

3.6 Conclusions

This paper demonstrates an approach to obtain late-time hydrogeologic property inferences consistent with the Cooper-Jacob method from transient datasets by uti-

lizing exponential parameter functions. The methodology is motivated by numerical experiments by *Wu et al.* (2005) and analysis of field-collected hydrographs by *Straface et al.* (2007). The hydrogeologic inferences are evaluated based on a large body of research into the meaning of late-time hydrogeologic property inferences (*Butler*, 1990; *Neuman*, 1990; *Meier et al.*, 1998; *Sanchez-Vila et al.*, 1999; *Neuman and Di Federico*, 2003; *Wu et al.*, 2005; *Knudby and Carrera*, 2006).

Utilizing this approach on a dataset from the LANL site has indicated that adequate water-level calibrations can be achieved within the constraints of the inversion: a single value of T_a is applied to all pumping/monitoring well pairs; $\hat{T}(t)$ decreases towards a constant value ($c_T > 0$); S_a is allowed to take distinct values and is allowed to increase or decrease towards convergent values. T_a provides an initial estimate of the effective transmissivity at the support scale characterized by the distances between the pumping and observation wells (*Neuman*, 1990; *Neuman and Di Federico*, 2003). In accordance with *Meier et al.* (1998), *Sanchez-Vila et al.* (1999), and *Knudby and Carrera* (2006), S_a is recognized as an indicator of interwell connectivity, indicating the degree in which pumping and monitoring well pairs are hydraulically connected.

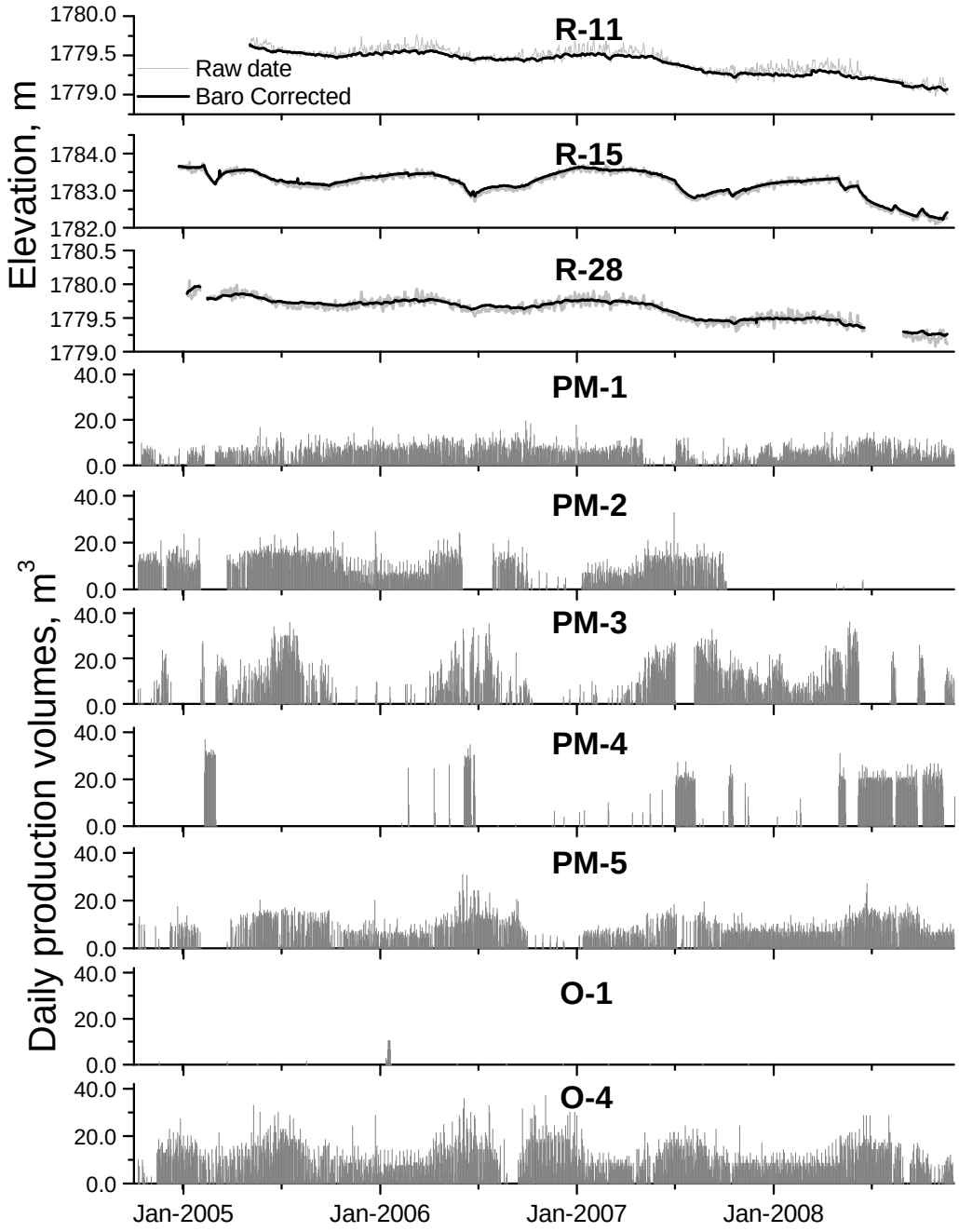


Figure 3.2: Water elevations at monitoring wells and production records for water-supply wells.

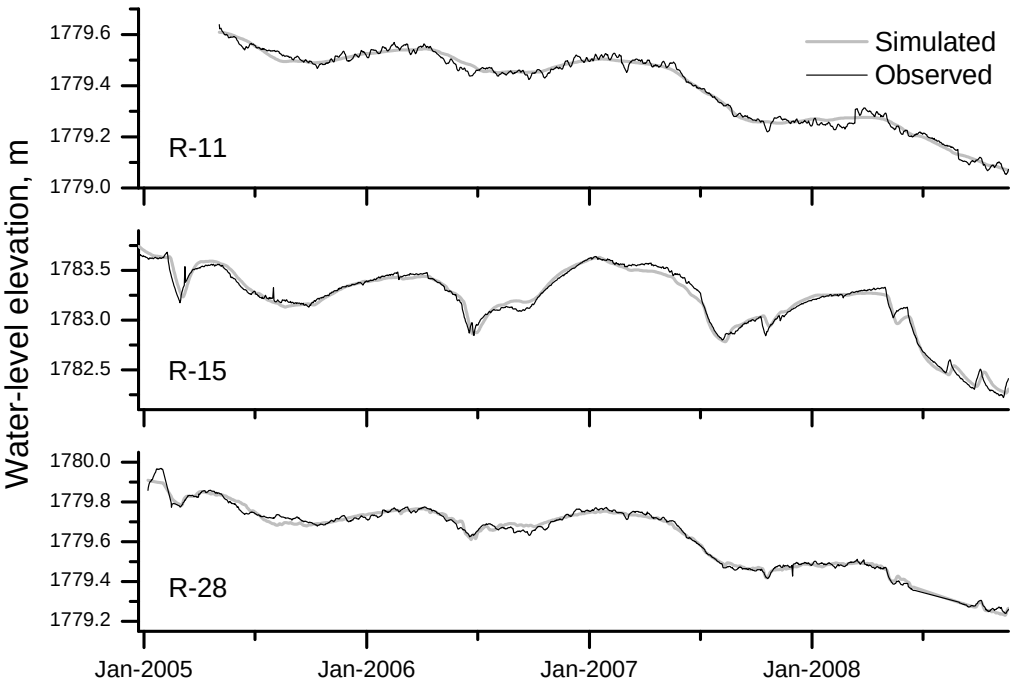


Figure 3.3: Top plot: calibrated (black) and observed (gray) water elevations for R-11 model inversion. Bottom plots: simulated drawdown contributions (black lines) from individual pumping wells, plotted with their associated pumping record (gray bars), and temporal trend required to reproduce the total predicted drawdown at R-11.

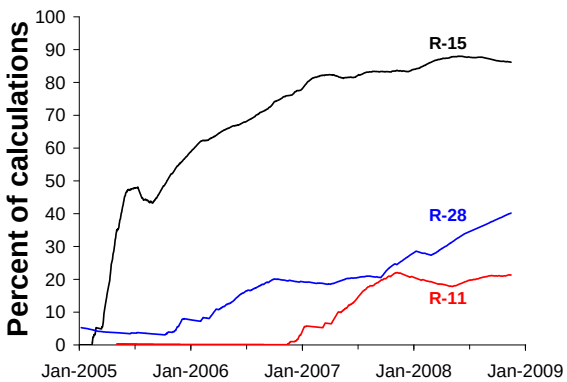


Figure 3.4: Percentage of superposed drawdown calculations performed within the valid Cooper-Jacob approximation domain ($u < 0.03$) as the length of the data record increases.

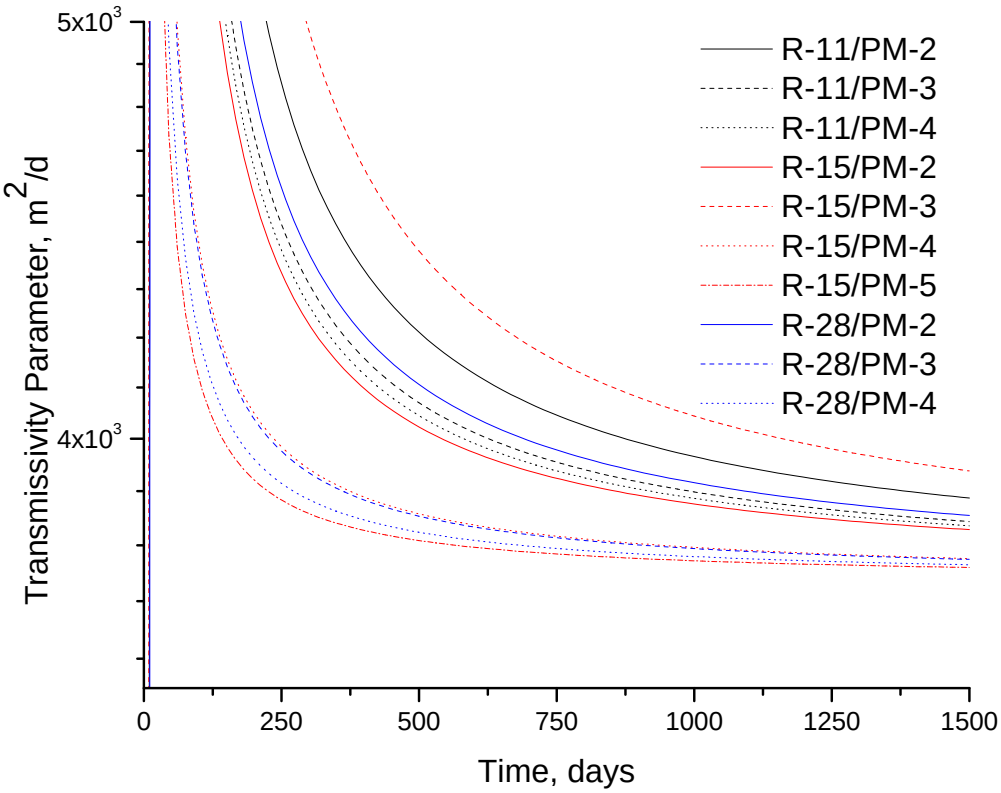


Figure 3.5: Estimated transmissivity parameter $\hat{T}$ functions

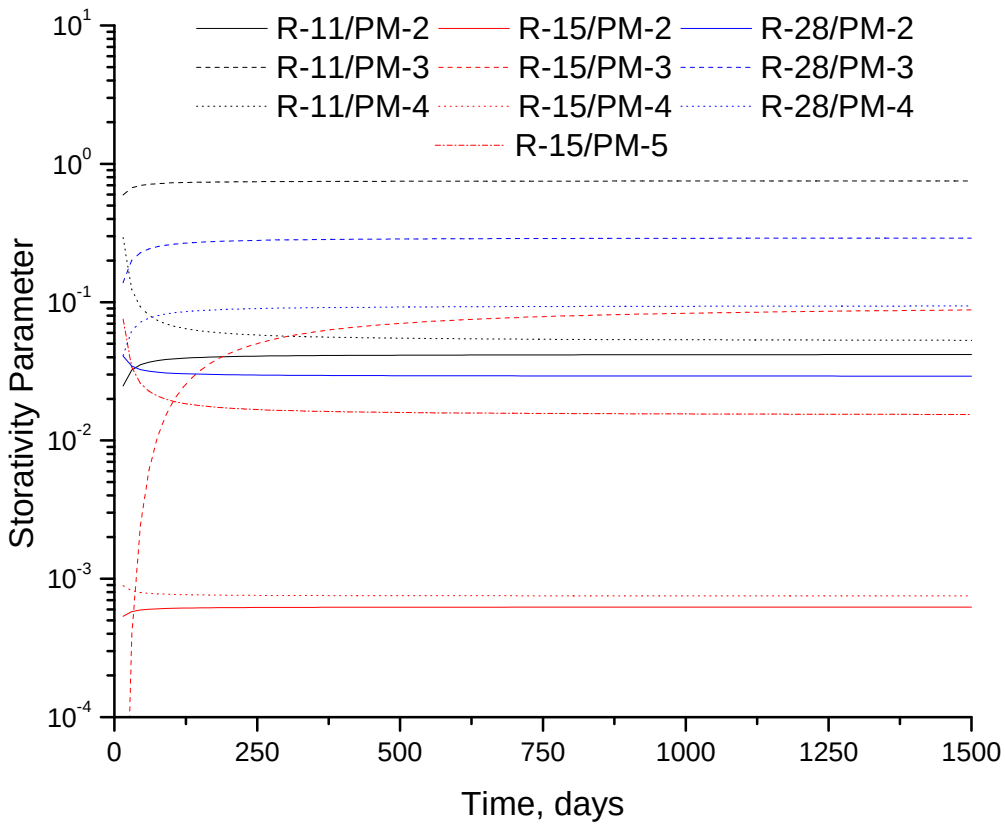


Figure 3.6: Estimated storativity parameter $\hat{S}$ functions

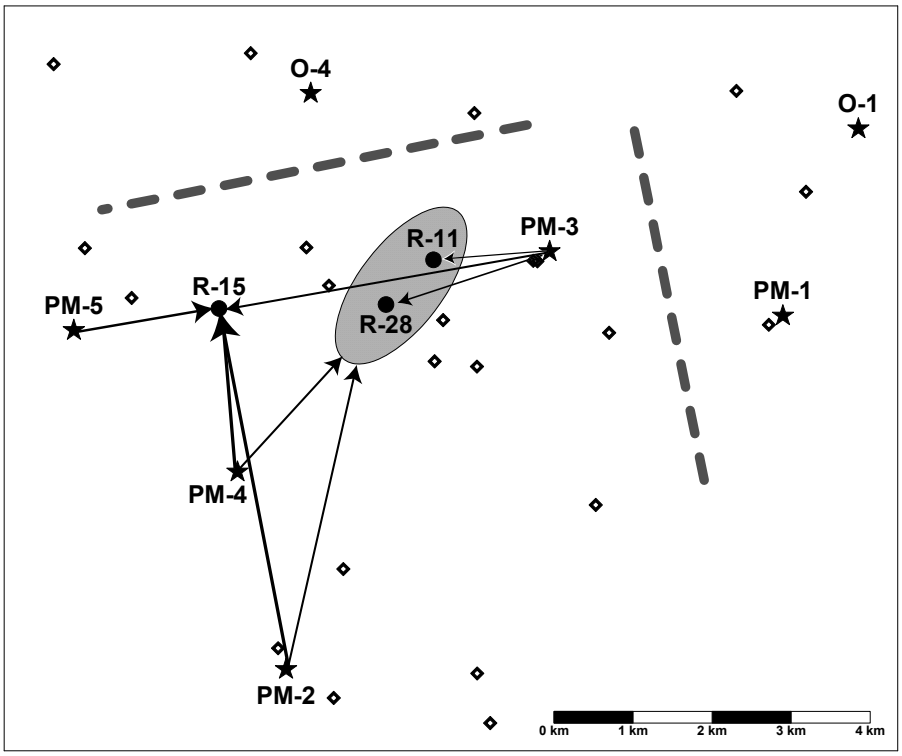


Figure 3.7: Map of observation wells (circles) and water-supply wells (stars) included in the analysis, along with symbols to illustrate the analysis conclusions. Dashed lines indicate locations of apparent hydrogeologic barriers, separating O-4 and PM-1 from the monitoring wells. The arrow size indicates the level of connectivity; for example, the large arrow connecting PM-4 and PM-2 with R-15 indicates the apparent high connectivity between these wells. The grey-filled oval illustrates the apparent existence of a region of similar hydrogeologic properties around R-11 and R-28. Locations of newly completed and planned monitoring wells are indicated by open diamonds.

Monitoring Well	Pumping Well	$\log(\hat{T}_G[m^2/d])$	c_T	$\log(\hat{S}_a)$	c_S	t_0 [d]	r [m]
R-11	PM-2	3.57	66.6	-1.38	-8.0	42.5	2902.7
	PM-3		47.7	-0.12	-3.5	58.4	803.6
	PM-4		44.2	-1.28	26.2	23.3	1929.9
R-15	PM-2		41.2	-3.20	-2.4	0.4	2434.7
	PM-3		88.3	-1.01	-168.1	59.8	2252.2
	PM-4		18.0	-3.12	2.6	0.1	1081.0
	PM-5		10.8	-1.82	24.3	1.8	986.0
R-28	PM-2		52.6	-1.54	5.1	22.2	2522.4
	PM-3		17.4	-0.53	-11.4	46.8	1154.3
	PM-4		13.1	-1.02	-12.7	25.7	1506.3

Table 3.1: Transmissivity function $\hat{T}(t)$ and storativity function $\hat{S}(t)$ parameter estimates from coupled model inversion. The Cooper-Jacob time-axis intercept ($t_0 = \hat{S}r^2/(2.25\hat{T})$, where r is the inter-well distance) is also tabulated indicating the arrival time of pumping rate change perturbations at the monitoring wells. Inter-well distances are tabulated for reference.

Chapter 4

Aquifer Structure Identification Using Stochastic Inversion[1]

Abstract

This study presents an inverse method for aquifer structure identification using sparse geophysical and hydraulic response data. The method is based on updating structure parameters from a transition probability model to iteratively modify the aquifer structure and parameter zonation. The method is extended to the adaptive parameterization of facies hydraulic parameters by including these parameters as optimization variables. The stochastic nature of the statistical structure parameters leads to nonconvex objective functions. A multi-method genetically adaptive evolutionary approach (AMALGAM-SO) was selected to perform the inversion given its search

[1]An edited version of this paper was published by AGU. Copyright 2008 American Geophysical Union: Harp D.R., Z. Dai, A.V. Wolfberg, J.A. Vrugt, B.A. Robinson, and V.V. Vesselinov (2008), Aquifer Structure Identification Using Stochastic Inversion, *Geophysical Research Letters, 35*, L08404, doi:10.1029/2008GL033585. To view the published open abstract, go to http://dx.doi.org and enter the DOI.

capabilities. Results are obtained as a probabilistic assessment of facies distribution based on indicator cokriging simulation of the optimized structural parameters. The method is illustrated by estimating the structure and facies hydraulic parameters of a synthetic example with a transient hydraulic response.

4.1 Introduction

Determining field-scale parameters in sufficient detail to capture aquifer heterogeneities is one of the greatest challenges for predicting flow and contaminant transport in large-scale subsurface systems. Uncertainty in the form of conceptual model bias can be introduced in modeling groundwater flow and contaminant transport when aquifer structure is fixed based on sparse or incomplete geophysical data (*Chen and Rubin*, 2003). Once the structure has been identified in this way, focus is placed on model calibration of zoned parameters through parameter estimation. This type of inverse method has been applied in many instances to estimate flow and transport parameters at various spatial scales (e.g. *Cooley* (1983); *Carrera and Neuman* (1986); *Doherty* (1994); *Kitanidis* (1996); *Dai and Samper* (2006)). *Ye et al.* (2004) indicate that using a predefined, deterministic aquifer structure can introduce larger bias and uncertainty into a model than an inappropriate choice of facies hydraulic parameters. Therefore, it may be more appropriate to assume reasonable values for hydraulic parameters (e.g., permeabilities), and allow the geometry of the aquifer structure to be evaluated during the inverse process.

Given uncertainty in both structure and parameters, *Sun* (2005) suggested "adaptive parameterization" to couple structure identification and parameter estimation in contaminant transport model calibration. This provides a more complete inversion of the model, allowing the optimization to explore combinations of structure geometry and hydraulic parameters. *Sun and Yeh* (1985) considered this approach

theoretically, posing the problem as a combinatorial optimization problem. *Eppstein and Dougherty* (1996) modified the extended Kalman filter to dynamically determine and refine structure. *Heredia et al.* (2000) proposed the refinement of an initial pattern using available prior information. *Ben Ameur et al.* (2002) presented a method to adaptively identify structure by iteratively determining whether a zone should be refined or coarsened. The use of Voronoi tessellation and Delaunay triangulation in a natural neighbor interpolation was proposed by *Tsai et al.* (2005) to perform adaptive parametrization. *Lu and Robinson* (2006) presented the level set method for identifying zones of high or low permeability by updating level set functions used to characterize zonal boundaries. The previously mentioned studies usually assume that the aquifer parameter zonations are randomly distributed. However, recent geological and geostatistical studies indicate that aquifer facies distributions are spatially correlated (*Carle and Fogg*, 1997; *Ritzi et al.*, 2004).

We propose a structure identification method that accounts for spatial correlation by means of a stochastic inversion of a transition probability model, in an analytical framework (*Dai et al.*, 2007), describing facies volume proportions, mean lengths, and juxtapositioning. The transition probability model provides a nonparametric, Markov chain approach to indicator geostatistics that is well suited to applications with sparse information (*Carle and Fogg*, 1997). The analytical solution allows structure identification to be cast as a conventional inverse modeling problem, using statistical structure parameters (such as facies volume proportions and mean lengths) to iteratively update the transition probability model. The facies proportions and mean lengths define the transition probability matrix. Indicator cokriging simulation produces the aquifer facies distributions, ensuring the statistical properties defined by the transition probability model are maintained. In this way, the aquifer structure is updated in the inversion, while the information provided by the conditional data (the sparse geological and geophysical data used to describe the facies distribution in boreholes) is honored. The optimization of the model inversion is performed using

a genetically-adaptive multimethod search algorithm called AMALGAM-SO. This method was chosen as it combines the strengths of several different evolutionary search approaches and has been shown to achieve good efficiency across a range of difficult synthetic benchmark problems (*Vrugt et al.*, 2008). While other optimization algorithms could potentially be implemented to drive the stochastic inversion, AMALGAM-SO was selected to illustrate the stochastic inversion methodology, without comparing its performance to other optimization algorithms. This decision was partly based on the assumption that, although an analytical solution of the transition probability model is utilized here, gradient-based methods would still have difficulty given the stochastic nature of the structural variables, which serve as inputs to the stochastic simulation. The analytical solution of the transition probability model provides the computational efficiency necessary for the large number of model evaluations required by the stochastic inversion. The combination of the analytical solution of the transition probability model and AMALGAM-SO provides a robust, computationally efficient model inversion with the ability to deal with complex fitness response surfaces. While in the past, the stochastic inversion described here was not possible due to computational and algorithmic limitations, we show that through the use of modern computers and analytically derived structure parameters (*Dai et al.*, 2007), this type of inversion can be realized.

4.2 Facies Transition Probability Model

Transition probability models have been used by geologists to describe sediment facies distributions for a few decades [e.g. *Agterberg* (1974); *Carle and Fogg* (1997)]. Recently, *Ritzi et al.* (2004) and *Dai et al.* (2005) incorporated the work of *Carle and Fogg* (1997) to relate the structure of the indicator random variables to proportions, geometry, and pattern of the aquifer facies. *Dai et al.* (2005) derived an analytical

solution for the transition probability model using two assumptions: (1) the cross-transition probabilities depend on facies volumetric proportions only, and (2) the juxtapositional tendencies between categories k and j are assumed symmetric in the direction ϕ. This solution can be expressed as

$$t_{ki}(\mathbf{h}_\phi) = p_i + (\delta_{ki} - p_i)e^{-\frac{\mathbf{h}_\phi}{\lambda_I}} \quad (k,\ i = \overline{1, N}), \tag{4.1}$$

where $t_{ki}(\mathbf{h}_\phi)$ is the probability of transitioning from facies k to facies i in lag distance $\mathbf{h}$ in direction ϕ, p_i is the proportion of facies i, δ_{ki} is the Kronecker delta, λ_I is the indicator correlation length, and N is the number of facies. By taking the partial derivative of the auto-transition probability (equation (4.1) with $k = i$) with respect to $\mathbf{h}$, the mean length ($\overline{L}_{k,\phi}$) can be related to λ_I as (*Dai et al.*, 2007)

$$\left.\frac{\partial t_{kk}(\mathbf{h}_\phi)}{\partial \mathbf{h}_\phi}\right|_{\mathbf{h}_\phi \to 0+} = -\frac{1}{\lambda_I}(1 - p_k) = -\frac{1}{\overline{L}_{k,\phi}}. \tag{4.2}$$

Equation (4.2) defines the relationship between indicator correlation length and the statistical parameters of facies proportion and mean lengths. Using equations (4.1) and (4.2), the continuous-lag transition probability matrix $\mathbf{T}$ in direction ϕ can be defined by the facies proportions and mean lengths as $\mathbf{T}(\mathbf{h}_\phi) = (t_{ki}(\mathbf{h}_\phi))_{N\times N}$. The transition probability matrix can be used for indicator simulation of aquifer facies using the indicator cokriging method (*Carle and Fogg*, 1997).

4.3 Stochastic Inverse Method

The transition probability model establishes a bridge between aquifer statistical parameters and aquifer facies distributions. By estimating these statistical parameters,

we are able to formulate a model inversion to identify aquifer structure. A flow diagram of the stochastic inversion method is presented in Figure 4.1.

The transition probability model is updated by calculating the transition probability matrix, using values of facies lengths and proportions generated as offspring of the previous generation of solutions. The structure is updated by indicator cokriging simulation using the updated transition probability model, where a single realization is used to represent the collection of equally probable realizations of the given transition probability model. Updated facies hydraulic parameters are applied to the structure zonation. Flow simulation is performed using the Finite Element Heat and Mass transfer (FEHM) code (*Zyvoloski et al.*, 1997), employing observed or assumed flow and boundary conditions, producing simulated transient head data.

The inverse modeling is performed with the goal of minimizing residuals of hydraulic head, where the objective function, J, can be defined as

$$J = \min_{\beta \in \mathbf{B}} \sum_{i=1}^{M} (\hat{h}_i(\beta) - h_i)^2, \tag{4.3}$$

where $\hat{h}(\beta)$ is an estimated head using parameter values in the vector β constrained by $\mathbf{B}$, which is defined by the upper and lower parameter bounds, where $\mathbf{B} \subseteq \Re^p$, p being the number of parameters, h is a measured head, and M is the number of measured heads.

Equation (4.3) defines the fitness function optimized with AMALGAM-SO (*Vrugt et al.*, 2008). In general, AMALGAM-SO allocates the number of offspring at each population size, $N^l = \{N_1^l, \ldots, N_q^l\}$, to q algorithms using a weighting scheme based on each algorithm's previous performance, where l is the population size index. In this way, AMALGAM-SO is able to exploit the individual strengths of selected algorithms at various stages of the optimization. AMALGAM-SO employs a population

incrementing restart strategy as its basis for collecting algorithm performance information used to update the offspring allocation (*Vrugt et al.*, 2008). This method has the advantage of combining individual algorithm strengths by allowing algorithms to exchange search information, and by adaptively distributing preference to algorithms exhibiting superior performance. In the current study, Covariance Matrix Adaptation (CMA), Genetic Algorithm (GA), and Particle Swarm Optimization (PSO) strategies were selected for the $q = 3$ optimization algorithms, as this combination has shown improved performance over other combinations (*Vrugt et al.*, 2008). The sequence of population sizes used in the current research was $l = 5, 10, 15, 20$. For more details on the settings of AMALGAM-SO, refer to *Vrugt et al.* (2008).

The uncertainty associated with the resulting optimized transition probability model is evaluated using conditional simulation and presented as the final result in the form of a structural probability map, which in the present case defines the shape and location of clay facies within a background of sand. A map of clay probability is produced by approximating the one-location marginal probabilities for clay by the sample mean of the clay indicator spatial function $I_{clay}(\mathbf{x})$ with respect to the optimized structure parameters as

$$p_{clay}(\mathbf{x}) = Pr\{I_{clay}(\mathbf{x}) = 1\} = \frac{1}{R}\sum_{i=1}^{R} I_{clay,i}(\mathbf{x}), \tag{4.4}$$

where R is the number of realizations.

4.4 Synthetic Example

The stochastic inversion method is illustrated by means of a synthetic example of a confined aquifer with conditional geophysical data and transient head measurements

from a synthetic pump test. The distribution of clay and sand in cross-section, where clay facies are embedded within a background of sand, is illustrated in Figure 4.2. In this example, the permeability of sand is $10^{-10}m^2$ and the permeability of clay is $10^{-13}m^2$. The synthetic structure was generated by conditional simulation using prespecified structure parameters. The proportion of sand (p_s) in the synthetic example is set at 0.7, while the proportion of clay (p_c) is 0.3. The mean lengths of the clay in the x (length) and y (thickness) directions are $300m$ and $20m$, respectively. Conditional data, which comprises the indicator data of the facies distribution, is collected as continuous bore log data from observation wells defined along the transect at $x = 0, 250, 500, 750$, and $1000m$ indicated by the vertical lines in Figure 4.2. The well at $x = 500m$ is set as the pumping well with a constant flow rate of $Q = 10.2kg/s$. The boundaries at $x = 0m$ and $x = 1000m$ are set as constant head boundaries with heads of $100m$ and $95m$, respectively. The top and bottom boundaries at $y = 0m$ and $y = 200m$ are no-flow boundaries for this confined model formulation. Measured transient heads were collected at 20 discrete times over a one year simulation, where the size of the time step increased over the simulation. The measurements were collected at 8 locations along the three central wells indicated by white dots in Figure 4.2, where it has been assumed that the pumping well has the ability to measure pressures while pumping.

This synthetic example was constructed to emulate typical applications encountered in practice, while still allowing a complete evaluation of the robustness and efficiency of the stochastic inversion with respect to a known structure and parameterization.

4.5 Results and Discussion

Figure 4.3(a) presents a plot of the objective function as a function of the number of model evaluations, where it is apparent by the step-like decreases in the objective function that AMALGAM-SO continues to locate improved solutions throughout the inversion. Figures 4.3(b) and 4.3(c) present the distributions resulting at the lower and upper bounds of the parameters, respectively, defined in the first two rows of Table 4.1. These two scenarios represent two points along the convex hull of the solution space $\mathbf{B} \subseteq \Re^p$ that the stochastic inversion is required to explore. Inspection of these two plots indicates the diversity of structures considered in the inversion. Figures 4.3(d), 4.3(e), and 4.3(f) present distributions at key points during the progression of the inversion, while Figure 4.3(g) presents the distribution of the optimal parameters. By inspecting these distributions, the transformation towards the synthetic distribution is apparent. These stages of the optimization are indicated in Figure 4.3(a) by their subfigure letter.

Corresponding tabular information on these distributions, and the synthetic distribution (Figure 4.2), are presented in Table 4.1, including iteration number, parameter values, and objective function values. By inspecting the decrease of the objective function plotted in Figure 4.3(a) and listed in Table 4.1, it is apparent that dramatic improvements are made during the course of the optimization, resulting in an extremely small objective function value for the optimized solution. This indicates that an estimate of the aquifer structure has been obtained that closely mimics the hydraulic response of the aquifer. Although an excellent fit to the observed aquifer response has been achieved, it is apparent in Table 4.1 that there is some discrepancy between the parameters used to generate the synthetic aquifer response, and those estimated with AMALGAM-SO. This is explained by the stochastic nature of this inverse problem. Multiple realizations for the same parameter combination result in widely varying values of the objective function. For instance, when evaluating 100

realizations generated using the optimized parameters, objective function values are obtained that vary between 0.5 and 200, with one outlier around 600. The standard deviation of these objective function values is approximately 68.9. This stochasticity allows different parameter combinations to generate nearly similar responses of the aquifer. The interest is therefore not so much in the exact values of the parameters, but on the optimized structure that has been identified. The latter has been successfully achieved, considering the very small value of the objective function, and the close similarity between the true and inversely estimated facies distribution. Furthermore, 95.5% of the optimized facies grid is assigned to the correct facies with respect to the synthetic example.

The final result is presented in a probability map of clay facies distribution in Figure 4.3(h) where the estimated one-location marginal probabilities of clay (equation (4.4)) were calculated from 100 realizations based on the optimized structure parameters. While this does provide the uncertainty of the structure based on the optimized transition probability model, it does not indicate the uncertainty of the structure with respect to the aquifer response. Future work will expand this uncertainty analysis by identifying the set of plausible structures with regard to aquifer response. It is apparent from a comparison of the true distribution (Figure 4.2) and the resulting clay probability map (Figure 4.3(h)) that the large structural features are captured in the analysis. These results demonstrate that the stochastic inversion is able to reduce the objective function to a reasonably low value (indicating that the response of the aquifer is modeled accurately) and that the large structural features of the aquifer are identified adequately given the limited amount of information provided in the synthetic data.

4.6 Conclusions

A method has been presented for aquifer structure identification using stochastic inversion. The method is based on updating structure parameters of a facies transition probability model to identify aquifer structure. This approach allows the problem to be formulated in an inverse modeling framework that can be extended to adaptive parameterization of facies hydraulic parameters by taking advantage of the efficiency in model evaluations with the analytical solution of the facies transition probability model. The inversion is driven by a multi-method genetically adaptive evolutionary optimization approach (*Vrugt et al.*, 2008), providing a robust inversion capable of traversing the complicated fitness landscape. Results are obtained as a probabilistic estimate of the existence of facies at a particular location given the optimized statistical structure parameters. This method can be applied to pump-test datasets, where some geophysical data are available, to provide probabilistic estimates of the facies distributions based on the hydraulic responses of the aquifer.

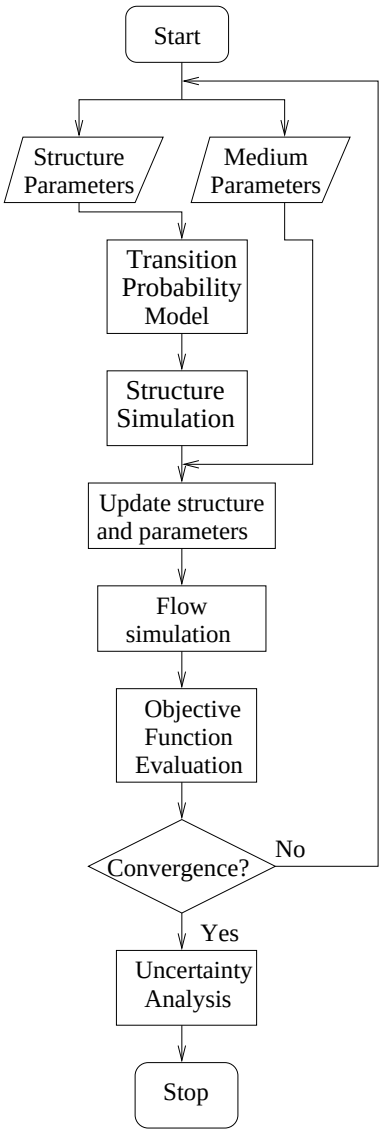


Figure 4.1: Stochastic inversion flow diagram.

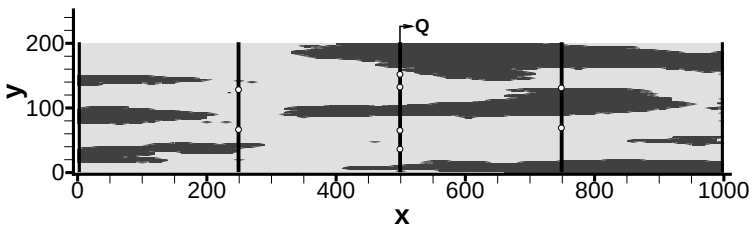


Figure 4.2: Synthetic distribution where light grey represents sand and dark grey represents clay. The black vertical lines indicate observation and pumping wells, as well as locations where facies indicator data has been collected. The white dots indicate head observation locations.

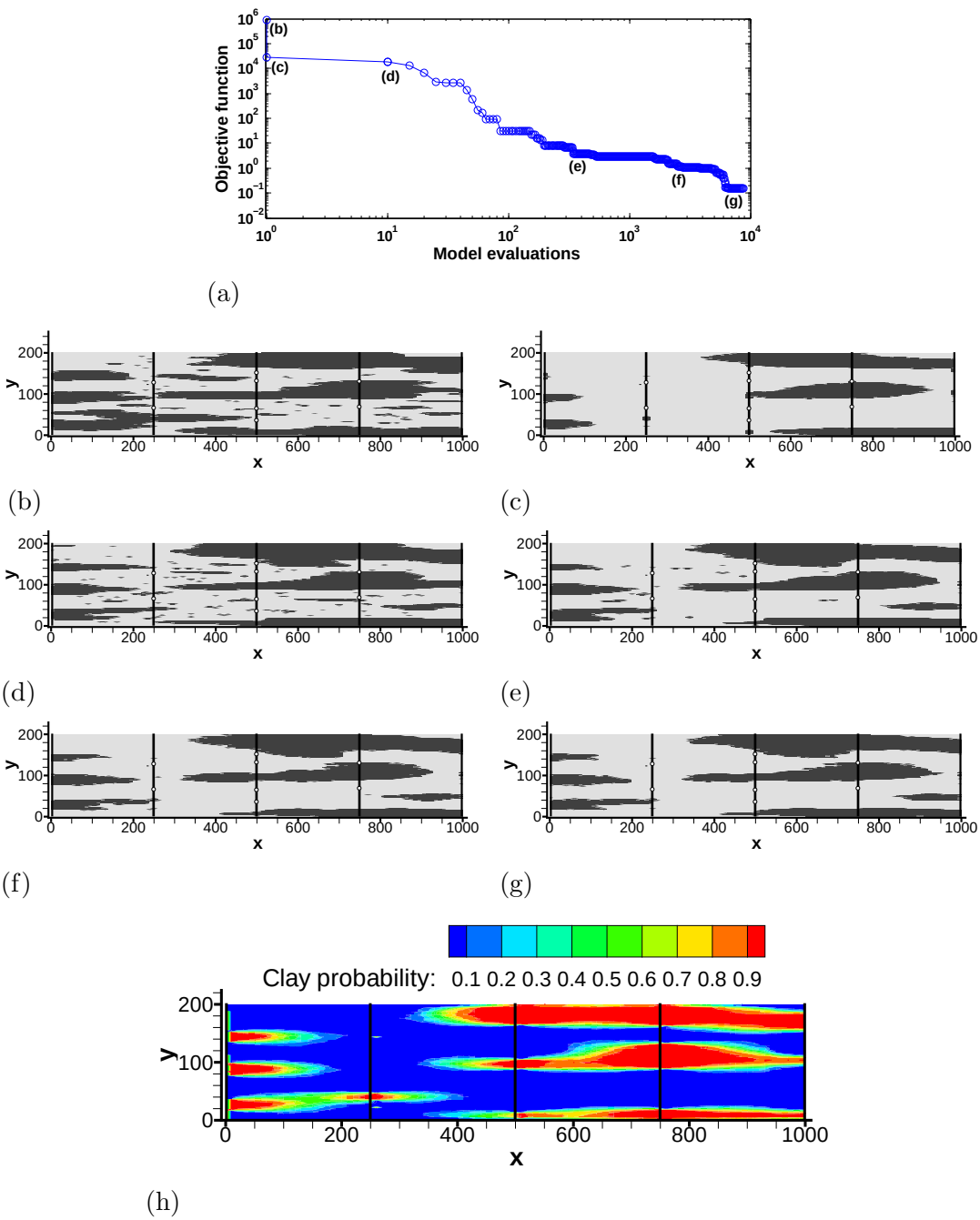


Figure 4.3: A plot of the objective function versus the number of model evaluations is presented in subfigure (a). The aquifer structure as different stages of the stochastic inversion are presented in subfigures (b) through (g), where their corresponding locations are noted in subfigure (a). Refer to Table 4.1 for detailed information on these aquifer structures. Subfigure (h) presents the clay probability map produced by stochastic simulation of the optimized structural parameters.

Table 4.1: Structure and hydraulic parameters at the lower (1a) and upper (1b) parameter bounds, at various stages of the model inversion, and for the synthetic example. Refer to the figures identified in column 1 for plots of the distributions.

Figure	Iteration	Proportion		Clay Mean	Clay Mean	Perm., log(m^2)		Objective
	Number	Sand	Clay	Length, m	Thickness, m	Sand	Clay	Function
4.3(b)	1a	0.6	0.4	280	15	-11	-14	876109
4.3(c)	1b	0.8	0.2	320	25	-9	-12	28459
4.3(d)	10	0.63	0.37	285.1	15.9	-9.33	-12.31	18440
4.3(e)	345	0.67	0.33	310.5	22.3	-9.99	-12.35	3.63
4.3(f)	2500	0.69	0.31	309.3	22.0	-10.00	-12.68	1.15
4.3(g)	6900	0.69	0.31	319.4	23.2	-10.00	-12.89	0.15
4.2	True	0.7	0.3	300	20	-10	-13	0

Chapter 5

Stochastic Inverse Method for Estimation of Geostatistical Representation of Hydrogeologic Stratigraphy Using Borehole Logs and Pressure Observations[1]

Abstract

An approach is presented for identifying statistical characteristics of stratigraphies from borehole and hydraulic data. The approach employs a Markov-chain based geostatistical framework in a stochastic inversion. Borehole data provide information on

[1]Submitted for publication in Stochastic Environmental Research and Risk Assessment: Harp, D., V.V. Vesselinov (2009), Stochastic Inverse Method for Estimation of Geostatistical Representation of Hydrogeologic Stratigraphy Using Borehole Logs and Pressure Observations, Manuscript submitted for publication.

the stratigraphy while pressure and flux data provide information on the hydraulic performance of the medium. The use of Markov-chain based geostatistics as opposed to covariance-based geostatistics can provide a more easily interpreted model geologically and geometrically. The approach hinges on the use of mean facies lengths (negative inverse auto-transition rates) and mean transition lengths (inverse cross-transition rates) as adjustable parameters in the stochastic inversion. Along with an unconstrained Markov-chain model, simplifying constraints to the Markov-chain model, including (1) proportionally random and (2) symmetric spatial correlations, are evaluated in the stochastic inversion. Sensitivity analyses indicate that the simplifying constraints can facilitate the inversion at the cost of spatial correlation model generality. Inverse analyses demonstrate the feasibility of this approach, indicating that despite some low parameter sensitivities, all adjustable parameters do converge towards their "true" values during the inversions for all models (proportionally random, symmetric, and unconstrained).

5.1 Introduction

Aquifers are often comprised of stratigraphic units with distinct geologic properties and compositions. In cases where the effective hydrogeologic properties of these stratigraphic units are distinct as well (e.g. *LaBolle and Fogg* (2001)), information pertaining to the stratigraphy can be inferred from the hydrogeologic behavior of the aquifer. If the stratigraphy is known at some locations (e.g. based on borehole logs or outcrops), this information can be used to constrain the set of possible aquifer structures using various geostatistical techniques (e.g. covariance-based (*Deutsch and Journel*, 1992) and Markov-chain-based (*Carle and Fogg*, 1996) indicator co-kriging). However, given the typically sparse hydraulic and geologic observations available from most aquifers, the number of possible structures concurrent with the available

data and the applied geostatistical model is generally large. Therefore, estimation of the stratigraphy must be posed in a stochastic framework. A set of geologically-equally-probable stratigraphies can be represented by a single geostatistical model of spatial variability. This allows the problem of aquifer structure identification to be simplified to the identification of a set of structures exhibiting highly-probable hydraulic characteristics conditioned on the available data.

In order to model the aquifer structure as a composition of distinct hydrostratigraphic units, we employ a multidimensional continuous-lag Markov-chain model of the spatial variability of categorical variables. In this case, aquifer heterogeneity is assumed to be composed of spatially-discontinuous structures characterized by distinct hydrostratigraphic units with uniform properties. The Markov-chain model is represented in a transition-probability framework describing the auto- and cross-transition probabilities of the categorical variables at specific lag distances (*Carle and Fogg*, 1997).

This approach contrasts the use of geostatistical techniques that model hydrogeologic heterogeneity as spatially continuous. This continuous approach has proven to be useful in many applications (*de Marsily et al.*, 1984; *LeVenue et al.*, 1995; *Ramarao et al.*, 1995; *Vesselinov et al.*, 2001; *Doherty*, 2003). We propose that these two approaches are each well suited for different hydrogeologic scenarios; distinct hydrostratigraphic units with contrasting hydrogeologic properties in one case and smooth continuous transitions in the aquifer heterogeneity in the other. For cases in between these two extremes, hybrids of these approaches have been proposed (*Gégo et al.*, 2001; *Lu and Zhang*, 2002).

A given geostatistical model defines spatial characteristics of a stratigraphy using statistical parameters. Therefore, any geostatistical model defines an infinite set of stratigraphies exhibiting defined statistical spatial characteristics. Equally-probable realizations of this infinite set of stratigraphies can be generated using conditional

simulation. While the equally-probable realizations of a geostatistical model may exhibit deviations from the defined spatial characteristics, statistical inference of a given characteristic are expected to converge to the defined parameter values at a sufficiently large set of realizations. Similarly, we demonstrate that convergent statistical inferences can be obtained that define the hydraulic response of an aquifer given a prescribed geostatistical model. This is fundamental to the structure identification approach presented here, providing a necessary link between stratigraphic spatial correlation and aquifer hydraulic response. We utilize this link to calibrate observed hydraulic responses to geostatistically inferred hydraulic responses.

If the available geological data from an aquifer is obtained from vertical boreholes, characteristics of the vertical spatial correlation of stratigraphic units (e.g. juxtapositional tendencies and mean lengths) can be estimated directly (*Zhang and Li*, 2008). However, given the sparse and discontinuous nature of such data in the horizontal plane, direct computation of the horizontal juxtapositional tendencies and mean lengths is not always possible (*Weissmann and Fogg*, 1999). While mean lengths can be estimated indirectly by inference based on discontinuous geologic observations, these estimates can be highly uncertain. Recognizing this situation, we have conducted an analysis assuming that the vertical spatial correlation model is known, focusing on demonstrating the ability to complete the spatial-correlation model by inferring the horizontal spatial correlation using hydraulic data.

Carle et al. (1998) demonstrate the use of a trial-and-error approach to modify cross-transition rates for horizontal Markov-chain spatial correlation models to obtain geologically-plausible juxtapositional relationships. The calibration of a conditional Markov-chain model to hydraulic data was proposed by Zhenxue Dai in a personal communication in 2007. This proposal provided the impetus for the research presented in *Harp et al.* (2008), where a two-dimensional two-stratigraphic unit aquifer was analyzed utilizing the concept of a representative realization of a

spatial correlation model. *Harp et al.* (2008) inverted structural parameters from a two-dimensional two-stratigraphic unit Markov-chain model (e.g. facies mean lengths and volumetric facies proportions) obtaining head predictions from a single realization for each Markov-chain model. It was assumed that the hydraulic response from the realization is representative of the hydraulic response for the infinite set of geologically-equally-probable realizations for the Markov-chain model.

In contrast, the current research considers a three-dimensional three-stratigraphic unit aquifer utilizing the concept of representative hydraulic sample statistics of a spatial correlation model. In this case, a flow simulation is performed on a convergent set of geologically-equally-probable stratigraphies from the current Markov-chain model and averaging is performed on the model-predicted hydraulic characteristics (i.e. pressures and fluxes). This provides a true characteristic hydraulic response of the current Markov-chain model, improving the inversion performance as averaging the hydraulic response of a set realizations provides smoother variations with changes in structural parameters than the use of a single representative realization. This paper demonstrates the feasibility of the proposed structure identification approach utilizing a three-dimensional three-stratigraphic unit synthetic model by (1) a sensitivity analysis of the adjustable parameters defining spatial correlation and hydraulic properties and (2) example model inversions demonstrating the performance of a gradient-based (Levenberg-Marquardt) optimization on this inversion framework.

5.2 Theoretical Discussion of Stochastic Representations of Stratigraphy

Boolean models present the earliest approach for modeling discontinuous spatial variability (*Matheron*, 1967), providing the ability to model adjacently-located distinct stratigraphic units. While Boolean models presented an advantage over continuous geostatistical approaches in considering aquifer connectivity, it is difficult to honor conditional data in this framework and the Boolean object geometries must be defined a priori (*de Marsily et al.*, 2005). The introduction of the indicator kriging approach by *Journel* (1983) provided an alternative to Boolean models, modeling distinct stratigraphic units using indicator functions in a geostatistical framework. Indicator kriging is easily conditioned to observed geologic data and produces more realistic structures based on an inferred or imposed covariance model as opposed to assumed geometric objects utilized in Boolean models.

Carle and Fogg (1997) extended 1-D Markov-chain models, previously demonstrated to successfully model embedded occurrences of individual strata (*Vistelius*, 1949), to multidimensional continuous-lag Markov-chain models by interpolation of transition-rate matrices defining the 1-D Markov-chain models. The multi-dimensional Markov-chain model provides an alternative representation of spatial correlation to the indicator cross-covariance (or cross-variogram) conventionally used in indicator geostatistics.

Transition probabilities can be inferred from a sufficient number of geologic observations by determining the relative frequency of stratigraphic unit transitions at various lag distances. A transition-probability based representation of a Markov-chain model can then be empirically derived by determining transition rates that optimally relate transition probabilities to distance (h [L]) according to the following

matrix exponential functional form:

$$\mathbf{T}(h) = \exp(\mathbf{R}h), \tag{5.1}$$

where $\mathbf{T}(h)$ is a transition-probability matrix defined as $\mathbf{T}(h) = (t_{ij}(h)),\ i,j = 1,\ldots,N$ and $\mathbf{R}$ is a transition-rate matrix defined as $\mathbf{R} = (r_{ij}),\ i,j = 1,\ldots,N$, where N is the number of mutually-exclusive exhaustively-defined stratigraphic units (categories) considered in the Markov-chain model. Equation (5.1) cannot be evaluated component-wise (i.e. $t_{ij}(h) \neq \exp(r_{ij}h)$), but can be exactly determined using an eigensystem analysis as

$$\mathbf{T}(h) = \sum_{k=1}^{N} \mathbf{Z}_k \exp(\lambda_k h) \tag{5.2}$$

where λ_k is the kth eigenvalue and $\mathbf{Z}_k$ is the kth spectral component matrix of $\mathbf{R}$ (*Carle*, 1999) (alternatively, equation (5.1) could be approximated using the infinite series definition of an exponential). In this way, individual transition probability functions can be empirically derived by fitting measured transition probabilities at discrete lag distances to the following functional form:

$$t_{ij}(h) = \sum_{k=1}^{N} z_{ij,k} \exp(\lambda_k h), \tag{5.3}$$

where $z_{ij,k}$ is the ijth component of the kth spectral component matrix.

By inspecting equation (5.3), it is apparent that the functional form for the transition-probability is a summation of N exponentials. As a result, a Markov-chain model is not restricted to purely "exponential-looking" spatial correlation functions, allowing the modeling of complex juxtapositional patterns.

This empirical derivation of a spatial-correlation model is analogous to the inference of auto- and cross-variograms in covariance-based indicator geostatistics, differing in that a variogram relates a statistical tendency of deviation in property or indicator values to separation distance, while the Markov-chain model relates probabilistic stratigraphic juxtapositioning to separation distance (i.e. relative-locational tendencies of stratigraphic units) (*Carle and Fogg*, 1996). Therefore, a Markov-chain model can provide a more interpretable stochastic representation of a stratigraphic distribution than the covariance-based approach.

Similar to covariance-based indicator geostatistics, tranisition-probability geostatistics are based on an assumption of stationarity. This implies that facies volumetric proportions should be spatially invariant. As such, this approach is best suited to modeling aquifers, or portions of aquifers, which have developed from processes that result in stationary stratigraphies, such as depositional processes.

As the Markov-chain approach directly models juxtapositional tendencies, asymmetric spatial correlations, such as fining or coarsening cycles, are easily represented. Additionally, given the necessary adherence to laws of probability concerning the stratigraphic transitions requires the simultaneous calibration of the components of the full transition-probability matrix, ensuring that the Markov-chain-model is internally consistent. This is not necessarily the case for the covariance-based approach as individual auto- and cross-variograms are calibrated independently (*de Marsily et al.*, 2005).

Considering the previous discussion, the Markov-chain model presents certain advantages to modeling categorical spatial variability, while still being amenable to implementation into an indicator co-kriging interpolation scheme. While much effort has been focused on developing inversion strategies for covariance-based spatial-correlation models, little has been applied to the Markov-chain based approach (e.g. *Harp et al.* (2008)). This paper presents a strategy that implements a Markov-chain

spatial-correlation framework into a stochastic inversion.

5.3 Methodology

Computation of the model evaluated by the stochastic inverse method presented here requires a multi-step approach: (1) generation of multiple geologically-equally-probable realizations of hydrostratigraphy from a geostatistical model, (2) flow simulation on the hydrostratigraphic fields assigning distinct uniform hydraulic properties to the stratigraphic units, and (3) inference of hydraulic-response characteristics of the geostatistical model. These steps are discussed in detail below along with a discussion of the applied stochastic inverse approach.

5.3.1 Generation of Stratigraphic Realizations of a Markov-Chain-Based Geostatistical Model

A 1-D continuous-lag Markov-chain model can be represented by a transition-rate matrix (introduced in equation (5.1)) composed of auto- and cross-transition rates as

$$\mathbf{R} = (r_{ij}), \quad i, j = 1, \ldots, N, \tag{5.4}$$

where r_{ij} is an auto-transition rate when $i = j$ and a cross-transition rate when $i \neq j$. As cross-transition rates are not as readily inferred from geologic observations as cross-transition probabilities (*Carle and Fogg*, 1997), it is generally desirable to compute the transition-rate matrix from an inferred discrete-lag transition-probability

matrix $\mathbf{T}(\Delta h)$ as

$$\mathbf{R} = \frac{\ln(\mathbf{T}(\Delta h))}{\Delta h}, \tag{5.5}$$

where Δh is a discrete lag (*Carle and Fogg*, 1997). If a stratigraphic pattern has been identified in the Markov-chain model (i.e. in the transition probabilities), similar results will be obtained for arbitrary selections of Δh. As presented in equation (5.1), the transition-rate matrix defines the continuous-lag transition-probability matrix $\mathbf{T}(h)$.

Adherence to probability theory constrains the Markov-chain model as

$$0 \leq t_{ij}(h) \leq 1 \quad \forall i, j; \tag{5.6}$$

$$\sum_{i=1}^{N} p_i = 1; \tag{5.7}$$

$$\sum_{j=1}^{N} t_{ij}(h) = 1 \quad \forall i; \tag{5.8}$$

$$\sum_{i=1}^{N} p_i t_{ij}(h) = p_j \quad \forall j, \tag{5.9}$$

where p_i is the one-location marginal probability or volumetric proportion of the ith stratigraphic unit, commonly referred to simply as the proportion. Equation (5.6)

ensures that $t(h)$ remains within the bounds of a probability measure. Equation (5.7) ensures that categories are mutually exclusive and exhaustively defined. Equation (5.8) indicates that an auto- or cross-transition must occur at lag h. Equation (5.9) requires that the summation of the probabilities of categories transitioning to the jth category, weighted by their respective proportions, must sum to the proportion of the jth category.

Using the definition of a transition rate (*Carle and Fogg*, 1997),

$$r_{ij} = \frac{\partial t_{ij}(0)}{\partial h} \quad \forall i, j, \tag{5.10}$$

inequality (5.6) and equations (5.8) and (5.9) can be expressed as

$$r_{ii} \leq 0 \quad \forall i \quad \text{and} \quad r_{ij} \geq 0 \quad \forall i, j \neq i, \tag{5.11}$$

$$\sum_{j=1}^{N} r_{ij} = 0 \quad \forall i, \tag{5.12}$$

$$\sum_{i=1}^{N} p_i r_{ij} = 0 \quad \forall j, \tag{5.13}$$

respectively.

Adherence to the equations and inequalities numbered (5.6) to (5.13) ensures that the Markov-chain model is internally consistent and ergodic (i.e. $\lim_{h\to\infty} t_{ij}(h) = p_j$ (*Ross*, 1993)). As a result, the auto- and cross-transition rates associated with one of the stratigraphic units (i.e. rates along the row and column associated with a

stratigraphic unit), usually denoted as the background category, can be automatically determined from the other transition rates. Therefore, it is only necessary to determine $(N-1)^2$ transition rates to define a transition-probability representation of a Markov-chain model.

A Markov-chain model can be completely modified by adjusting the transition rates comprised in a transition-rate matrix $\mathbf{R}$ (equation (5.4)). A two-category Markov-chain model is completely defined by specifying a single transition rate, assuming that the proportions are available, utilizing equations (5.12) and (5.13) as

$$\mathbf{R} = \begin{bmatrix} r_{11} & -r_{11} \\ -p_1/p_2 r_{11} & p_1/p_2 r_{11} \end{bmatrix}, \tag{5.14}$$

where category two has been set as the background and r_{11} is the only transition rate requiring specification (alternatively, category one could have been set as the background, thereby only requiring the specification of r_{22}). As proportions are independent of direction, proportions may be available with reasonable certainty from continuous bore logs (assuming stationarity and lack of bias in borehole locations) for use in directions with discontinuous data. It can be verified that the spatial correlations represented in equation (5.14) are symmetric and random with respect to proportions (proportionally random) (as will be discussed in more detail below). Therefore, a two-category Markov-chain model is necessarily symmetric and proportionally random.

When utilizing Markov-chain models with more than two categories, the spatial correlations are not required to be symmetric or proportionally random, allowing more diversity in the geostatistical model. A three-category transition-rate matrix

(designating category three as background) illustrates this as

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & -r_{11} - r_{12} \\ r_{21} & r_{22} & -r_{21} - r_{22} \\ -p_1 r_{11} - p_2 r_{21} & -p_1 r_{12} - p_2 r_{22} & p_1(r_{11} + r_{12}) + p_2(r_{21} + r_{22}) \end{bmatrix}, \quad (5.15)$$

where it is apparent that four transition rates (i.e. r_{11}, r_{12}, r_{21}, and r_{22}) require specification.

It is possible to impose symmetry and proportional randomness on spatial correlations within a Markov-chain model with greater than two categories either for reference or when such simplifying assumptions are justified. For instance, for spatial-correlation symmetry between two categories in a given direction, it is possible to relate diagonally-opposing cross-transition probabilities of the two categories as

$$t_{ij}(h) = \frac{p_j}{p_i} t_{ji}(h). \quad (5.16)$$

Using equation (5.10), equation (5.16) can be expressed in terms of transition rates as

$$r_{ij} = \frac{p_j}{p_i} r_{ji}. \quad (5.17)$$

Using the assumption of symmetry between non-background categories, a three-category Markov-chain model requires the specification of three transition rates ($\sum_{i=1}^{N-1} i$, in general), reducing the number of rates needing specification by one ($\sum_{i=1}^{N-2} i$, in general). This is illustrated utilizing equation (5.17) to impose sym-

metric spatial correlation in the non-background categories (categories 1 and 2) in equation (5.15) as

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & -r_{11} - r_{12} \\ p_1/p_2 r_{12} & r_{22} & -p_1/p_2 r_{12} - r_{22} \\ -p_1 r_{11} - p_1 r_{12} & -p_1 r_{12} - p_2 r_{22} & p_1 r_{11} + 2p_1 r_{12} + p_2 r_{22} \end{bmatrix}, \tag{5.18}$$

where it is apparent that it is only necessary to specify r_{11}, r_{12}, and r_{22} (alternatively, r_{21} could have been specified as opposed to r_{12}).

If the occurrence of stratigraphic units is assumed to be proportionally random for non-background categories (i.e. spatially uncorrelated juxtapositional tendencies), specification of the Markov-chain model is simplified by assuming that cross-transition probabilities of the non-background categories can be determined by weighting the collective probability of any cross-transition (not including an auto-transition) from a particular category by the relative proportion of the considered transition category (*Carle and Fogg*, 1997) as

$$t_{ij}(h) = (1 - t_{ii}(h)) \frac{p_j}{1 - p_i} \quad \text{for } j \neq i. \tag{5.19}$$

This can be expressed with respect to transition rates utilizing equation (5.10) as

$$r_{ij} = -r_{ii} \frac{p_j}{1 - p_i} \quad \text{for } j \neq i. \tag{5.20}$$

The proportionally random assumption for non-background category spatial correlations will reduce the number of transition rates needing specification for a three-category Markov-chain model to two ($N - 1$, in general), as it is only necessary

to specify the auto-transition rates, excluding the background auto-transition rate. This is illustrated by utilizing equation (5.20) to impose proportional randomness in the non-background categories in equation (5.15) as

$$\mathbf{R} = \begin{bmatrix} r_{11} & -r_{11}\frac{p_2}{(1-p_1)} & r_{11}\left(\frac{p_2}{1-p_1} - 1\right) \\ -r_{22}\frac{p_1}{1-p_2} & r_{22} & r_{22}\left(\frac{p_1}{1-p_2} - 1\right) \\ p_1\left(r_{22}\frac{p_2}{1-p_2} - r_{11}\right) & p_2\left(r_{11}\frac{p_1}{1-p_1} - r_{22}\right) & p_1 r_{11} + p_2 r_{22} - p_1 p_2\left(\frac{r_{11}}{1-p_1} + \frac{r_{22}}{1-p_2}\right) \end{bmatrix}, \tag{5.21}$$

where it is apparent that it is only necessary to specify r_{11} and r_{22}.

It is important to note that proportional randomness does not imply symmetric correlations (*Turk*, 1982). In fact, spatial correlations are proportionally random and symmetric only if

$$r_{ii} = \frac{1-p_i}{1-p_j} r_{jj} \quad \text{for } j \neq i, \tag{5.22}$$

which can be derived by substituting equation (5.20) in equation (5.17). Equation (5.22) indicates that symmetric proportionally-random spatial correlations require that transition rates are related by their volumetric proportions. It can be demonstrated that this is necessarily the case for a two-category Markov-chain model as $1-p_1 = p_2$ and $1-p_2 = p_1$. Using these relations for a two-category Markov-chain model and setting $i = 2$ and $j = 1$ in equation (5.22) results in $r_{22} = p_1/p_2 r_{11}$, which is equivalent to the formula for r_{22} presented in equation (5.14), further demonstrating that 2-category Markov-chain models are necessarily composed of symmetric and proportionally random spatial correlations.

As indicated by *Dai et al.* (2007), the eigenvalues of a symmetric proportionally-random Markov-chain model will be $\eta_1 = 0$ and $\eta_i = (p_1 - 1)/r_{11}$ for $i = 2, \ldots, N$.

The numerical eigenvalue decomposition utilized in the current research requires the computation of spectral component matrices $\mathbf{Z}_i$ (*Carle and Fogg*, 1997) as

$$\mathbf{Z}_i = \frac{\prod_{j \neq i} (\eta_j \mathbf{I} - \mathbf{R})}{\prod_{j \neq i} (\eta_j - \eta_i)} \quad \forall i = 1, \ldots, N \tag{5.23}$$

where $\mathbf{I}$ is the $N \times N$ identity matrix. As noted by *Dai et al.* (2007), it is apparent that cases where $\eta_i \approx \eta_j$ will cause numerical instability in equation (5.23). We avoid symmetric proportionally-random spatial correlations here as this case results in 2 repeating eigenvalues for a three-category Markov-chain model ($(N-1)$ in general). Note that this issue is irrelevant for two-category Markov-chain models as there are no repeating eigenvalues in this case (i.e. $N - 1 = 1$). *Dai et al.* (2007) presents an analytical expression for transitional probability matrix coefficients t_{ij} as a function of the lag distance. This expression is derived under the assumption of symmetric proportionally-random spatial correlations and can be applied only for this special case.

To facilitate a geometric representation of structures generated from a Markov-chain model, facies mean lengths have been defined (*Carle and Fogg*, 1996) as

$$\bar{l}_i = -\frac{1}{r_{ii}} = \left[\frac{\partial t_{ij}(0)}{\partial h} \right]^{-1} \tag{5.24}$$

where $\bar{l}_i$ is the facies mean length of the ith stratigraphic unit in a particular direction. We utilize inverse transition rates as the adjustable parameters in the stochastic inversion as we have found they scale more appropriately in the inversion than transition rates. In order to facilitate the use of inverse transition rates, we extend

equation (5.24) to define mean facies lengths as

$$\bar{l}_{ii} = -\frac{1}{r_{ii}} = -\left[\frac{\partial t_{ii}(0)}{\partial h}\right]^{-1} \tag{5.25}$$

and mean transition lengths as

$$\bar{l}_{ij} = \frac{1}{r_{ij}} = \left[\frac{\partial t_{ij}(0)}{\partial h}\right]^{-1} \quad i \neq j \tag{5.26}$$

where $\bar{l}_{ii}$ is equivalent to $\bar{l}_i$ in equation (5.24) and $\bar{l}_{ij}$ is the mean length of the ith stratigraphic unit between transitions from the ith to the jth stratigraphic unit. To clarify this linguistically,

$$\bar{l}_{ii} = \frac{\text{total length of the } i\text{th unit}}{\text{number of transition from the } i\text{th unit to any other unit}} \tag{5.27}$$

while

$$\bar{l}_{ij} = \frac{\text{total length of the } i\text{th unit}}{\text{number of transitions from the } i\text{th to the } j\text{th unit}} \quad i \neq j. \tag{5.28}$$

In a two-category Markov-chain model, $\bar{l}_{ii}$ will be equivalent to $\bar{l}_{ij}$ as the number of transitions from the ith unit to any other unit will equal the number of transitions from the ith to jth unit. This is also apparent by inspecting equation (5.14) considering equations (5.25) and (5.26), which can be expressed as

$$\mathbf{R} = \begin{bmatrix} -1/\bar{l}_{11} & 1/\bar{l}_{11} \\ p_1/(p_2\bar{l}_{11}) & -p_1/(p_2\bar{l}_{11}) \end{bmatrix}. \tag{5.29}$$

However, in a Markov-chain model with more than two categories, this will not be the case, as the number of transitions from the ith unit to any other unit will not equal the number of transitions from the ith to the jth unit. The relationship between $\bar{l}_{ii}$ and $\bar{l}_{ij}$ can be understood by considering the individual lengths of the ith unit identified along a column placed end for end, leaving indicators at locations where transitions occur. The mean facies length will provide the mean length of all segments, considering all transitions. However, the mean transition length will only consider the i to j transitions, providing the mean length along the ith unit between i to j transitions. For further details, refer to Appendix 5.A for derivations of $\bar{l}_{ii}$ and $\bar{l}_{ij}$.

Utilizing equations (5.25) and (5.26), we can express equations (5.15), (5.18), and (5.21) in terms of mean lengths as

$$\mathbf{R} = \begin{bmatrix} -\frac{1}{\bar{l}_{11}} & \frac{1}{\bar{l}_{12}} & \frac{1}{\bar{l}_{11}} - \frac{1}{\bar{l}_{12}} \\ \frac{1}{\bar{l}_{21}} & -\frac{1}{\bar{l}_{22}} & \frac{1}{\bar{l}_{22}} - \frac{1}{\bar{l}_{21}} \\ \frac{p_1}{\bar{l}_{11}} - \frac{p_2}{\bar{l}_{21}} & \frac{p_2}{\bar{l}_{22}} - \frac{p_1}{\bar{l}_{12}} & \frac{p_1}{\bar{l}_{12}} - \frac{p_1}{\bar{l}_{11}} + \frac{p_2}{\bar{l}_{21}} - \frac{p_2}{\bar{l}_{22}} \end{bmatrix}, \tag{5.30}$$

$$\mathbf{R} = \begin{bmatrix} -\frac{1}{\bar{l}_{11}} & \frac{1}{\bar{l}_{12}} & \frac{1}{\bar{l}_{11}} - \frac{1}{\bar{l}_{12}} \\ \frac{p_1}{p_2\bar{l}_{12}} & -\frac{1}{\bar{l}_{22}} & \frac{1}{\bar{l}_{22}} - \frac{p_1}{p_2\bar{l}_{12}} \\ \frac{p_1}{\bar{l}_{11}} - \frac{p_1}{\bar{l}_{12}} & \frac{p_2}{\bar{l}_{22}} - \frac{p_1}{\bar{l}_{12}} & -\frac{p_1}{\bar{l}_{11}} + \frac{2p_1}{\bar{l}_{12}} - \frac{p_2}{\bar{l}_{22}} \end{bmatrix}, \tag{5.31}$$

and

$$\mathbf{R} = \begin{bmatrix} -\frac{1}{\bar{l}_{11}} & \frac{p_2}{\bar{l}_{11}(1-p_1)} & \frac{1}{\bar{l}_{11}}\left(1 - \frac{p_2}{1-p_1}\right) \\ \frac{p_1}{\bar{l}_{22}(1-p_2)} & -\frac{1}{\bar{l}_{22}} & \frac{1}{\bar{l}_{22}}\left(1 - \frac{p_1}{1-p_2}\right) \\ \frac{p_1}{\bar{l}_{11}} - \frac{p_1p_2}{\bar{l}_{22}(1-p_2)} & \frac{p_2}{\bar{l}_{22}} - \frac{p_1p_2}{\bar{l}_{11}(1-p_1)} & -\frac{p_1}{\bar{l}_{11}} - \frac{p_2}{\bar{l}_{22}} + \frac{p_1p_2}{\bar{l}_{11}(1-p_1)} + \frac{p_1p_2}{\bar{l}_{22}(1-p_2)} \end{bmatrix}, \tag{5.32}$$

respectively.

In the research presented here, we utilize equations (5.30), (5.31), and (5.32) to demonstrate the performance of the stochastic inverse approach on spatial correlation models with different levels of complexity using mean facies lengths and mean transition lengths as adjustable parameters.

Geologically-equally-probable realizations of the Markov-chain models are generated using simulated quenching (*Carle et al.*, 1998) utilizing initial fields (images) generated by transition-probability based indicator co-kriging. The generation of initial fields by transition-probability based indicator co-kriging is analogous to the process of utilizing covariance-based indicator co-kriging in sequential indicator simulation (SIS) (*Deutsch and Journel*, 1992), except that the weighting coefficients are determined using a transition-probability based co-kriging set of equations as opposed to a covariance-based co-kriging set of equations (*Carle*, 1999). It is important to note that while the methods for computation and inference of a Markov-chain and indicator cross-covariance models are fundamentally different, *Carle and Fogg* (1996) demonstrate that they are related as $C_{ij}(h) = p_i(t_{ij}(h) - p_j)$, where $C_{ij}(h)$ is the indicator cross-covariance of the ith and jth indicator variables at lag h.

Simulated quenching randomly cycles through the nodes of a kriged field changing the category of a node if this produces transition probabilities closer to the prescribed transition probabilities defined by the Markov-chain model. In this way, simulated quenching can be considered a zero-temperature form of simulated annealing, as the most desirable category is always selected, as opposed to probabilistically accepting some undesirable category changes (*Carle et al.*, 1998). As the initial fields are produced by transition-probability based indicator co-kriging, a full implementation of simulated annealing is not required (*Deutsch and Journel*, 1992, p. 160). Nodes with categories specified by conditional data are excluded from consideration. Adherence of the realization to the prescribed Markov-chain model is evaluated using

an objective function Φ_{SQ} defined as

$$\Phi_{SQ} = \sum_{l=1}^{M} \sum_{i=1}^{N} \sum_{j=1}^{N} (t_{ij}(\mathbf{h}_l) - \hat{t}_{ij}(\mathbf{h}_l))^2, \tag{5.33}$$

where M is the number of specified lag vectors, $\mathbf{h}_l$ is the lth lag vector, $t_{ij}(\mathbf{h}_l)$ is the transition probability from the ith category to the jth category at lag vector $\mathbf{h}_l$ prescribed by the given Markov-chain model, and $\hat{t}_{ij}(\mathbf{h}_l)$ is the measured transition probability from the ith to the jth category for the realization at lag vector $\mathbf{h}_l$.

Considered lag vectors are selected based on statistical closeness measured using the determinants of the transition-probability matrices (*Carle*, 1999). Lag vectors with greater statistical closeness (associated with transition probability matrices with greater determinant values) than a prescribed determinant cutoff value are included. The transition probability at a particular lag vector for a realization can be calculated as

$$\hat{t}_{ij}(\mathbf{h}_l) = \frac{p_{ij}(\mathbf{h}_l)}{p_i}, \tag{5.34}$$

where $p_{ij}(\mathbf{h}_l)$ is the two-location joint probability determined by calculating the histogram of occurrences of the ith and jth categories separated by $\mathbf{h}_l$ (*Carle and Fogg*, 1996). The simulated quenching is terminated when either the objective function is decreased below a prescribed value or the prescribed maximum number of iterations is exceeded. Equally-probable realizations are generated for a given Markov-chain model by modifying the random number seed used to determine the cycling order of the nodes in the simulated quenching.

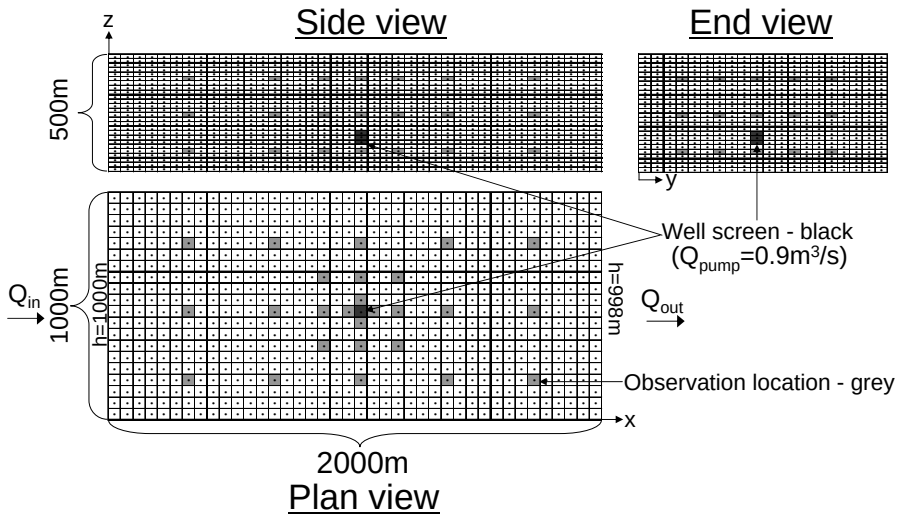


Figure 5.1: Block-centered finite difference grid with points representing nodes and lines indicating the connectivity of the grid. Light grey cells denote observation locations and black cells denote the screen of the pumping well. Constant head boundaries are indicated on the plan view.

5.3.2 Steady-State Flow Simulation

We evaluate the feasibility of the proposed stochastic inverse approach utilizing three-dimensional synthetic models. The models contain three substantially-contrasting stratigraphic units with uniform hydraulic conductivities of $K_1 = 10^{-2}$, $K_2 = 10^{-5}$, and $K_3 = 10^{-8}$ m/s, where the subscript denotes the stratigraphic unit. Note the significant contrast in the facies conductivity; such a contrast is expected for facies with significantly different hydrogeologic properties such as, for example, gravels, silts, and clays. The flow equations are solved on a block-centered finite-difference grid, presented in Figure 5.1, using FEHM (*Zyvoloski et al.*, 1997) with $40 \times 20 \times 25$ nodes in the x, y, and z directions, respectively, where x and y are horizontal directions and z is vertical. Node spacing is 50 m$\times$50 m$\times$20 m, resulting in model dimensions of 2000 m$\times$1000 m$\times$500 m in the x, y, and z directions, respectively.

The flow simulations are performed under steady-state conditions. The flow is driven by a flow gradient in x direction and a well pumping at the approximate horizontal center of the model. Constant-head boundaries are prescribed with a head drop of 2m (hydraulic gradient of $i = 0.1\%$) from $x = 0$ m to $x = 2000$ m. No-flow conditions are assigned for the rest of the model domain boundaries. The well is located at ($x = 1025$ m, $y = 525$ m) with a well screen depth, with respect to the top of the model, from 320 m to 380 m. The well pumping rate, Q_{pump}, is 0.9 m^3/s. Pressure-head data are collected at three depths from 26 observation wells for a total of 78 observation points within the model domain (refer to Figure 5.1). The head data is defined as a drawdown from an initial hydrostatic water level (initial head is $h_{init} = 1000\ m$ everywhere maintaining confined conditions).

It should be noted that since the flow direction is predominantly along the x-axis and the model domain is elongated along the x-axis (which allows for better spatial representation of the facies), it can be expected that in the sensitivity and inverse analyses performed below, the transitional properties of the facies will be better characterized along the x-axis rather than the y-axis.

5.3.3 Inference of Characteristic Hydraulic Response of a Geostatistical Model

The proposed approach hinges on the assumption that alternate geostatistical models of stratigraphy will produce distinct hydraulic-response characteristics in the simulation model. To characterize hydraulic performance of a given Markov-chain model, we generate a series of geologically-equally-probable realizations and perform flow simulation on each realization. Then we compute representative statistics of hydraulic responses. This requires that the representative statistics converge by a reasonable number of realizations. Given this consideration, we use the sample mean

of the simulated drawdowns Δh and the sample median of the model output flux Q_{out}, defined as $Q_{out} = Q_{in} - Q_{pump}$ (refer to Figure 5.1). We use the sample median instead of the sample mean for Q_{out} as it converges more quickly due to its insensitivity to outlier values. We found that outlier values are not an issue with the Δh values, and that the sample mean converges as fast or faster than the sample median in most cases for Δh.

5.3.4 Stochastic Inverse Approach

The goal of the inversion is to identify model parameters that minimize residuals between calibration targets and model predictions. The predictions are inferred from a set of equally-probable geostatistical realizations, and include the sample mean of pressure response at the observation locations ($\overline{\Delta h}$) and the sample median of the down-gradient model output flux ($\overline{Q}_{out}$). The adjustable parameters are the required mean facies lengths and mean transition lengths ($\bar{l}_{ij,\phi}$ $i, j = 1, \ldots, N$, where ϕ indicates direction) defining the geostatistical properties of the realizations. Therefore, the inversion attempts to identify the most plausible geostatistical model constrained by observed geology concurrent with the available hydraulic data. Note that in the inverse process presented here, we assume that the facies conductivities are perfectly known.

The stochastic inversion described here utilizes a Levenberg-Marquardt optimization strategy (*Doherty*, 1994) where the objective function is defined as

$$\Phi(\theta) = \sum_{i=1}^{M} (w_i * (s_i - \hat{s}_i(\theta)))^2, \tag{5.35}$$

where M is the number of calibration targets (78 mean drawdown observations ($\overline{\Delta h}$) and the median output flux ($\overline{Q}_{out}$)), θ is a vector containing the current adjustable

parameter values, w_i is the weight associated with the ith calibration target, s_i is the ith calibration target (observation), and $\hat{s}_i$ is the ith predicted value given θ. Weights are set proportional (scaled between 0 and 1) to the inverse of the variance of the calibration target values from the n realizations for the "true" parameter values in order assign greater weight to calibration targets with lower variance, as given a particular model, pressure responses from certain locations will converge more slowly than others, thereby introducing greater uncertainty in the model.

In order to easily evaluate the mathematical feasibility of the proposed inversion, the calibration targets are collected as the sample mean and sample median of the drawdown predictions and output flux, respectively. This facilitates an analysis of the feasibility of the approach as a global minimum objective function value of zero exists at the true parameters. This differs from an actual application, where the calibration targets will be measurements from a single realization of a random field and the global minimum cannot be as easily identified.

5.4 Facies Geostatistical Models

A model with proportionally random non-background spatial correlations and a model with symmetric non-background spatial correlations were developed as "true" stratigraphic distributions for the sensitivity analyses and model inversion runs. Each model includes 3 facies. An unconstrained model was not developed, as the symmetrically constrained geostatistical model was used as the "true" stratigraphy in this case. It is assumed that the behavior of the unconstrained model inversion will behave similarly if the "true" stratigraphy contains symmetric or non-symmetric spatial correlations, as symmetry is merely a special case of spatial correlation for the unconstrained geostatistical model. Conditional data about the facies distributions are assumed to be available from the pumping well and the 26 observation wells

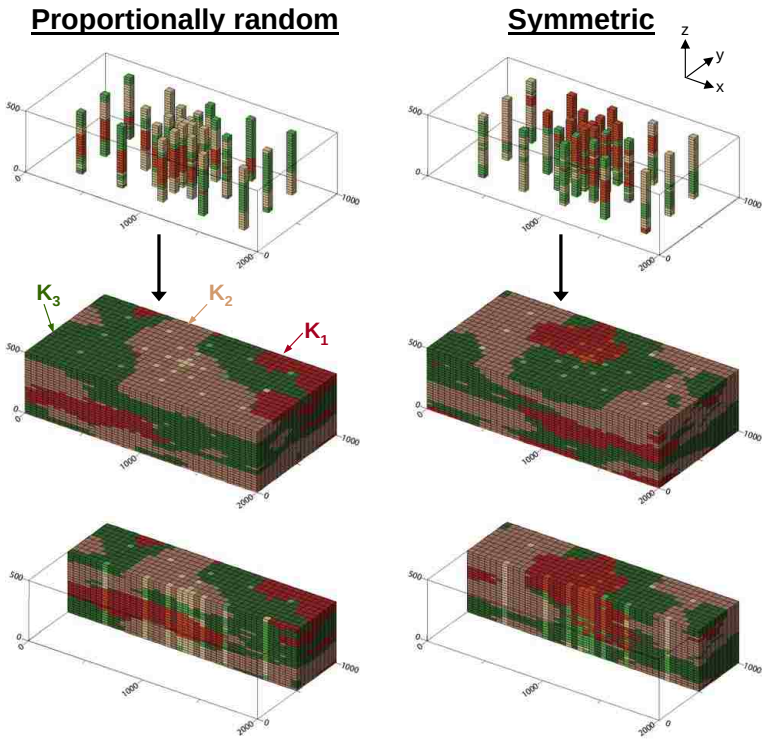


Figure 5.2: Proportionally random and symmetric conditional data and example realization produced from the given geostatistical model and the conditioning data. Stratigraphic units are identified by color indicated as K_1 ($= 10^{-2} m/s$) green, K_2 ($= 10^{-5} m/s$) tan, and K_3 ($= 10^{-8} m/s$) red. Conditioning data are presented in a slightly lighter color than non-conditioning points. Axis labels are in meters.

(Figure 5.1). This produces 675 conditioning data points for each model (27 wells $\times$ 25 cells in the vertical direction). For simplicity sake, the support scales of the conditioning data and the finite-difference grid are equal. At the same 26 observation wells drawdown data are collected from 3 depths in each observation well (Figure 5.1); note that no drawdown is collected at the pumping well. The conditioning data were collected from a representative realization produced from a geostatistical model with the prescribed characteristics (e.g. proportionally random/symmetric spatial correlations for non-background categories) and a reasonable hydraulic response. In order to ensure that the pumping well screen is located in the high conductivity unit, the

well screen was conditioned to K_1 for the development of the representative realization. Figure 5.2 presents the conditioning data for the proportionally random and symmetric models along with example realizations produced from the given geostatistical model and the conditional data in each case. A cutout at $y = 500m$ of each model is also included to display the model interior.

The transition rate matrix for the proportionally random constrained Markov-chain model in both the x- and y-directions in inverse meters is

$$\mathbf{R}_{x,y;rnd} = \begin{bmatrix} -2.00 \times 10^{-3} & 1.00 \times 10^{-3} & 1.00 \times 10^{-3} \\ 3.33 \times 10^{-4} & -1.83 \times 10^{-3} & 1.50 \times 10^{-3} \\ 6.67 \times 10^{-4} & 1.33 \times 10^{-3} & -2.00 \times 10^{-3} \end{bmatrix}, \tag{5.36}$$

in accordance with equation (5.21). The corresponding mean facies and transition lengths in meters for the proportionally random constrained model are

$$\overline{\mathbf{L}}_{x,y;rnd} = \begin{bmatrix} 500 & 1000 & 1000 \\ 3000 & 545 & 667 \\ 1500 & 750 & 500 \end{bmatrix}, \tag{5.37}$$

where $\overline{\mathbf{L}} = (\bar{l}_{ij}),\ i, j = 1, \ldots, N$. Due to the designation of category two as the background category and the constraint of proportional randomness for the non-background categories, equation (5.37) can be completely specified from two mean lengths (refer to equation (5.32)).

The transition rate matrix of the symmetrically constrained Markov-chain model

in both the x- and y-directions in inverse meters is

$$\mathbf{R}_{x,y;sym} = \begin{bmatrix} -2.00 \times 10^{-3} & 1.00 \times 10^{-3} & 1.00 \times 10^{-3} \\ 5.00 \times 10^{-4} & -2.00 \times 10^{-3} & 1.50 \times 10^{-3} \\ 5.00 \times 10^{-4} & 1.50 \times 10^{-3} & -2.00 \times 10^{-3} \end{bmatrix}, \tag{5.38}$$

in accordance with equation (5.18). The corresponding mean facies and transition lengths in meters for the symmetric model are

$$\overline{\mathbf{L}}_{x,y;sym} = \begin{bmatrix} 500 & 1000 & 1000 \\ 2000 & 500 & 667 \\ 2000 & 667 & 500 \end{bmatrix}. \tag{5.39}$$

Due to the designation of category two as the background category and the constraint of symmetry for the non-background categories, equation (5.39) can be completely specified from three mean lengths (refer to equation (5.31), while in the unconstrained case, four mean lengths must be specified (refer to equation (5.30)).

It is important to note the mean facies lengths (diagonal values in equations (5.37) and (5.39)) are equal and approximately 1/4 of the x-direction model domain (2000 m) and 1/2 of the y-direction model domain (1000 m). As a result, the x-direction spatial correlations are more fully represented than the y-direction.

The vertical, or z-direction mean facies and transition lengths are not considered here, as it is assumed that the vertical Markov-chain model can be determined directly from the borehole data. As the volumetric proportions are independent of direction, it is assumed that they are available from the vertical Markov-chain model, and are not included as adjustable parameters. For reference, the z-direction

Markov-chain model transition rates for both models in inverse meters are

$$\mathbf{R}_z = \begin{bmatrix} -1.00 \times 10^{-2} & 5.00 \times 10^{-3} & 5.00 \times 10^{-3} \\ 1.67 \times 10^{-3} & -9.17 \times 10^{-3} & 7.50 \times 10^{-3} \\ 3.33 \times 10^{-3} & 6.67 \times 10^{-3} & -1.00 \times 10^{-2} \end{bmatrix}, \tag{5.40}$$

while the corresponding mean facies and transition lengths in meters are

$$\overline{\mathbf{L}}_z = \begin{bmatrix} 100 & 200 & 200 \\ 600 & 109 & 133 \\ 300 & 150 & 100 \end{bmatrix}. \tag{5.41}$$

5.5 Results and Discussion

5.5.1 Aquifer Response Statistical Convergence

The generation of realizations by simulated quenching constitutes a Monte Carlo sampling of a spatial correlation model. While the realizations generated with respect to a given spatial correlation model are equally probable geologically, this is not necessarily the case hydraulically. Therefore, sample statistics from a set of realizations are used to characterize the aquifer response of a spatial correlation model. An analysis of the convergence of these sample statistics is necessary in order to determine the number of realizations (n) required for a desired or necessary precision. Figure 5.3 presents Δh at the 78 observation locations as n increases, where n is presented on a log scale as suggested by *Ballio and Guadagnini* (2004). Presenting n on a log scale as opposed to a linear scale provides a better visualization of convergence as the uncertainty associated with inferring a statistic decreases at a decreasing rate

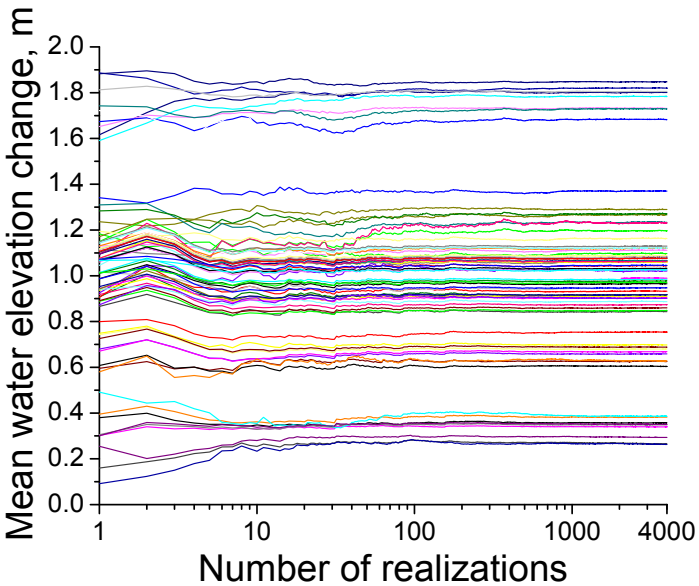


Figure 5.3: Convergence of drawdown from undisturbed conditions (water elevation change) for the 78 observation locations as a function of the number of realizations plotted on log scale.

as n increases. This is apparent by inspecting the root mean square (RMS) of the sample mean of the drawdown predictions $\overline{\Delta h}$ defined as

$$\overline{\Delta h}_{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\overline{\Delta h}_i^2}, \tag{5.42}$$

Equation (5.42) implies that the uncertainty associated with $\overline{\Delta h}$ scales with $1/\sqrt{n}$, a trend that is more easily visualized on a log scale.

We find that values of Δh obtained from the proportionally random spatial correlation model fail the Shapiro-Wilks test for normality (*R Development Core Team*, 2005) at the 5% significance level at all observation locations. Therefore, a conservative estimate (over-estimate) of the confidence interval length for the expected

value of Δh, $\mathrm{E}[\Delta h]$, from n realizations can be obtained considering the Chebyshev inequality as

$$\Pr\left[\overline{\Delta h}_n - k\frac{\sigma}{\sqrt{n}} \leq \mathrm{E}[\Delta h] \leq \overline{\Delta h}_n + k\frac{\sigma}{\sqrt{n}}\right] \geq 1 - \frac{1}{k^2} \tag{5.43}$$

where σ is the standard deviation, and k is the number of standard deviations considered. The right hand side of inequality (5.43) underestimates the probability that $\overline{\Delta h}_n$ is k standard deviations from the $\mathrm{E}[\Delta h]$. Therefore, an overestimate of the confidence limit interval length $|CL|$ can be obtained using the square root of the sample variance $\sqrt{S_n}$ to approximate σ as

$$|CL| = 2k\sqrt{\frac{S_n}{n}}. \tag{5.44}$$

Based on an approximation of σ by $\sqrt{S_{4000}}$, the average 95% $|CL|$ ($k = 4.472$) for $\overline{\Delta h}$ for the 78 observation locations at $n = 1000$ is 0.0294 m with a standard deviation of 0.0109 m. This indicates that we are inferring $\overline{\Delta h}$ within precision typical for field measurements of Δh. This ensures that the stochastic inversion will be limited by the precision of field measurements and not by the statistical inference of the hydraulic response. A larger number of realizations will provide smoother gradients for the objective function at the cost of computational efficiency. We have chosen $n = 1000$ in order to compromise between inference precision and computational efficiency.

A single forward model run (including facies structure generation and flow simulation) takes approximately 30 seconds on a 2.4 GHz AMD Opteron processor. Therefore, a single evaluation of the objective function, or model call (1000 forward model runs), takes approximately 20 wall clock hours on the same processor. As the forward model runs can be performed independently, the model call computa-

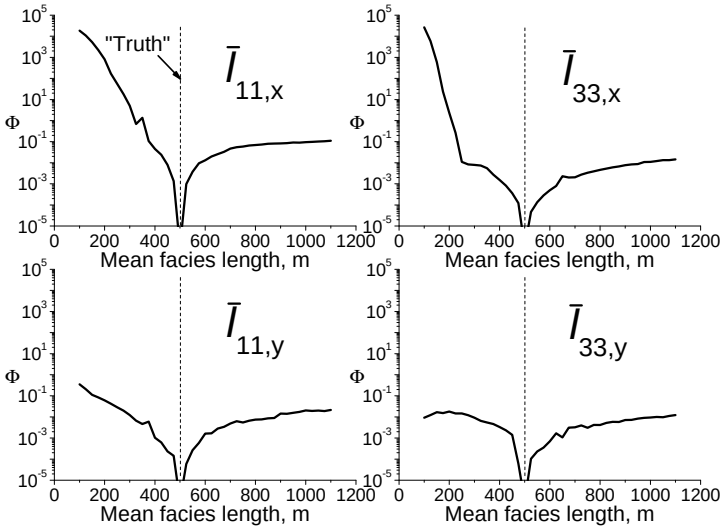


Figure 5.4: Sensitivity analysis for the proportionally random constrained model where $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{11,y}$, and $\bar{l}_{33,y}$ are plotted versus the objective function. The objective function has been evaluated at 25m intervals, half the model grid spacing.

tion time can be greatly reduced utilizing multiple processors to evaluate individual forward model runs concurrently.

5.5.2 Sensitivity analyses

Sensitivity analyses for the proportionally random constrained, symmetrically constrained, and unconstrained model inversions are presented in Figures 5.4, 5.5, and 5.6, respectively. These figures contain plots of the adjustable parameters versus the objective function Φ (equation (5.35)). The "true" parameter values, which are indicated by vertical dashed lines in each plot, indicate the parameter values used to generate the calibration targets. A subscript ϕ will be appended to the mean lengths to indicate the direction of interest as $\bar{l}_{ij,\phi}$, where ϕ can be either x or y in the case presented here.

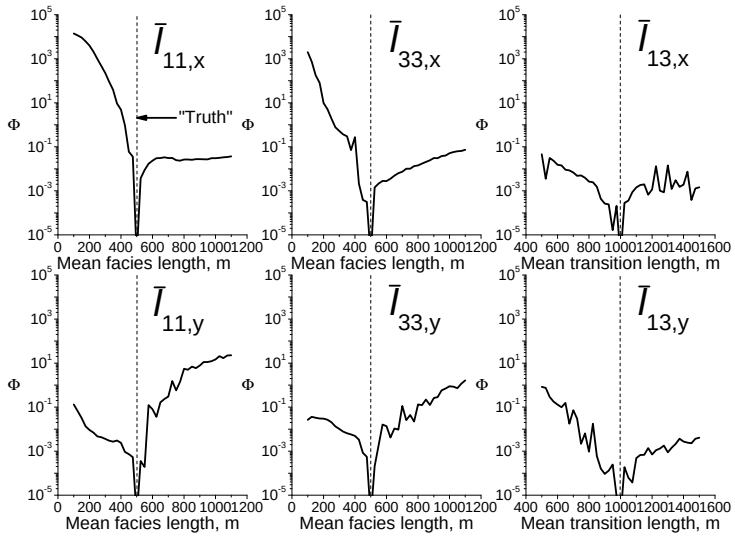


Figure 5.5: Sensitivity analysis for the symmetrically constrained model where $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{13,x}$, $\bar{l}_{11,y}$, $\bar{l}_{33,y}$, and $\bar{l}_{13,y}$ are plotted versus the objective function. The objective function has been evaluated at 25m intervals, half the model grid spacing.

Figure 5.4 contains four plots, presenting the sensitivities of $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{11,y}$, and $\bar{l}_{33,y}$. Due to the designation of category two as the background category and the constraint of proportional randomness for the non-background categories, the 14 remaining mean lengths can be determined from these four mean lengths (refer to equation (5.32)). By inspecting the plots in Figure 5.4, it is apparent that the x-direction mean lengths are more sensitive than the y-direction mean lengths. As discussed in the previous section this behavior was expected. This is due to the fact that the x-direction is parallel to the flow direction and has better representation in the model due to domain size. It is also apparent that the sensitivities appear smooth at this discretization (half the grid spacing), implying that a gradient-based optimization strategy will function well on this problem.

Figure 5.5 presents the sensitivities for the symmetrically constrained model. Plots are included for $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{13,x}$, $\bar{l}_{11,y}$, $\bar{l}_{33,y}$, and $\bar{l}_{13,y}$. Due to the designation

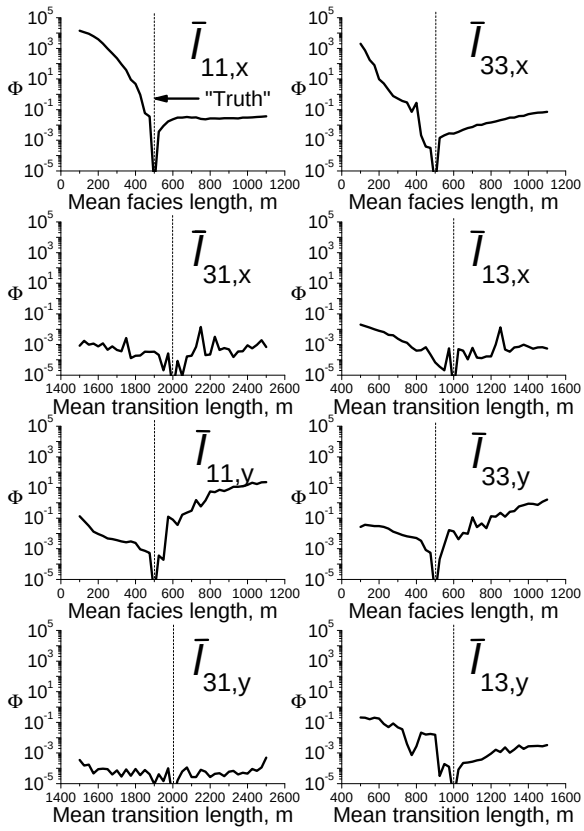


Figure 5.6: Sensitivity analysis for the unconstrained model where $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{31,x}$, $\bar{l}_{13,x}$, $\bar{l}_{11,y}$, $\bar{l}_{33,y}$, $\bar{l}_{31,y}$, and $\bar{l}_{13,y}$ are plotted versus the objective function. The objective function has been evaluated at 25m intervals, half the model grid spacing.

of facies 2 as the background category and the constraint of symmetry on the non-background categories, the 12 remaining mean lengths can be determined from these six mean lengths (refer to equation (5.31)). By comparing the sensitivities for the symmetrically constrained model (Figure 5.5) with those from the proportionally random constrained model (Figure 5.4), it is apparent that while the mean facies length parameters in the x-direction are still relatively smooth, the y-direction mean facies length parameters are significantly less smooth, and the mean transition length sensitivities are the least smooth.

Figure 5.6 presents the sensitivities for the unconstrained model. Plots are included for $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{13,x}$, $\bar{l}_{31,x}$, $\bar{l}_{11,y}$, $\bar{l}_{33,y}$, $\bar{l}_{13,y}$, and $\bar{l}_{31,y}$. As the symmetrically constrained model is used as the truth for the unconstrained sensitivity analysis, the sensitivity plots of $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{11,y}$, and $\bar{l}_{33,y}$ (mean facies length parameters) in Figure 5.6 are the same as in Figure 5.5 and are repeated in Figure 5.6 solely for reference. By inspecting the other plots (mean transition length parameters) in Figure 5.6, it is apparent that the mean transition lengths for the unconstrained model are relatively insensitive, especially $\bar{l}_{31,y}$.

The sensitivity analyses presented in Figures 5.4, 5.5, and 5.6 indicate that model inversions will benefit from constraints on the Markov-chain model. In cases where these constraints are not appropriate requiring the use of an unconstrained Markov-chain model, the mean facies lengths are still relatively sensitive, while the mean transition lengths exhibit low non-smooth sensitivities. In general the shape of sensitivity curves are convex with relatively well defined minima at the "true" values. This suggests that a gradient-based method for model inversion can be successfully applied in these cases. In the next section, inverse analyses demonstrate the affect of these sensitivities on the estimation of the mean lengths.

The sensitivity analyses were performed at half the horizontal grid spacing (25 m) over a range of 1000 m, resulting in 41 model calls. Therefore, a sensitivity analysis for a single parameter requires approximately 850 wall clock hours. As discussed above, the computation time can be greatly reduced by parallelization of the model calls.

5.5.3 Inverse analyses

Model inversion results for the proportionally random constrained model are presented in Figure 5.7. The initial values of the adjustable parameters were set to

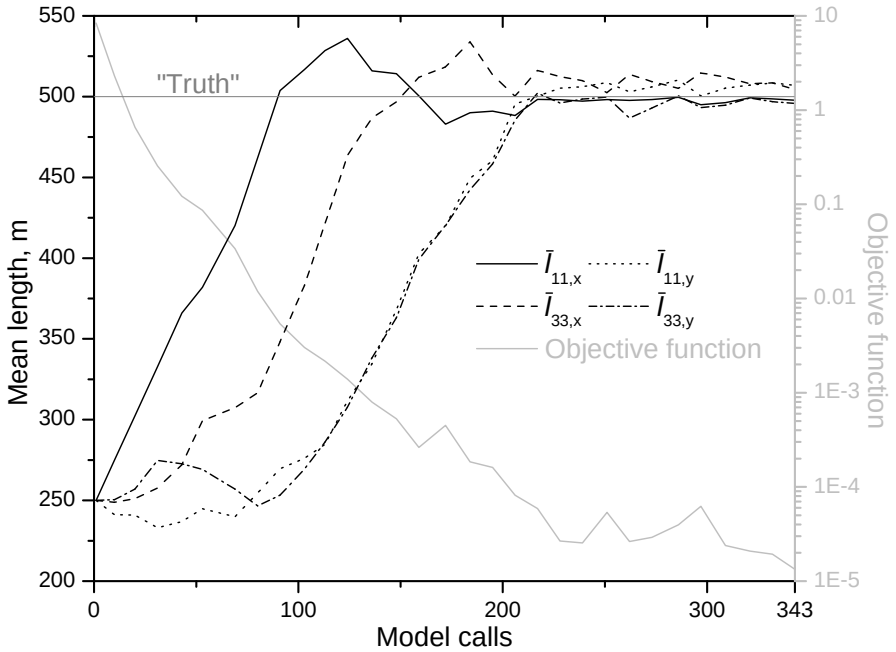


Figure 5.7: Proportionally random constrained model inversion results. Parameter values and objective function are plotted as a function of the number of model calls.

250 m from their "true" values of 500 m. However, the sensitivities in Figure 5.4 imply that the gradient-based optimization utilized here is expected to perform similarly if the initial values are defined within explored sensitivity ranges. Future work will address the effect of the initial values in the inverse process. As expected, the mean facies length of the high permeability category in the x-direction ($\bar{l}_{11,x}$) approaches the true value more quickly than the other mean facies lengths, followed by the mean facies length of the low permeability category in the x-direction ($\bar{l}_{33,x}$). This corresponds with the sensitivity analysis results presented in Figure 5.4, which indicate that the x-direction mean facies lengths are more sensitive than the y-direction facies mean lengths. To reiterate previous discussions, this is likely due to a combination of factors including the longer x-direction distance compared with the y-direction and the fact that the x-direction mean facies lengths are parallel to the direction of flow,

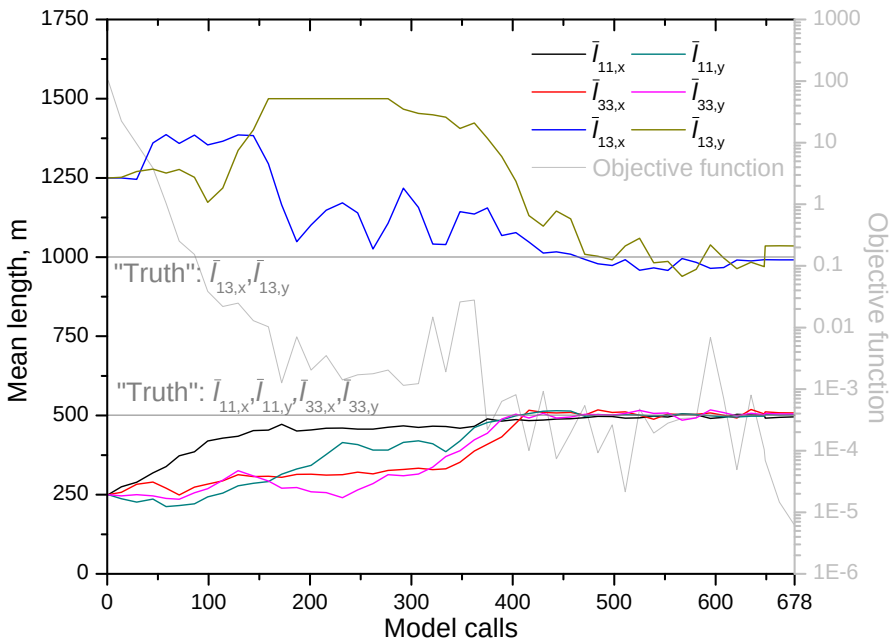


Figure 5.8: Symmetrically constrained model inversion results. Parameter values and objective function are plotted as a function of the number of model calls.

thereby affecting the hydraulic response of the aquifer by controlling the connectivity in the flow direction.

As indicated in Figure 5.7, the lowest objective function was achieved at 343 model calls. This corresponds to 31 Levenberg-Marquardt optimization iterations, which include multiple model calls for the calculation of derivatives and evaluation of Marquardt parameters. It was determined that no further progress would be achieved by 447 models calls, or 41 Levenberg-Marquardt optimization iterations.

Figure 5.8 presents the model inversion results for the symmetrically constrained model, where the mean transition lengths $\bar{l}_{13,x}$ and $\bar{l}_{13,y}$ are included. The corresponding diagonally-opposing mean transition lengths $\bar{l}_{31,x}$ and $\bar{l}_{31,y}$ are determined by the symmetry constraint. The initial values of the mean facies lengths were set

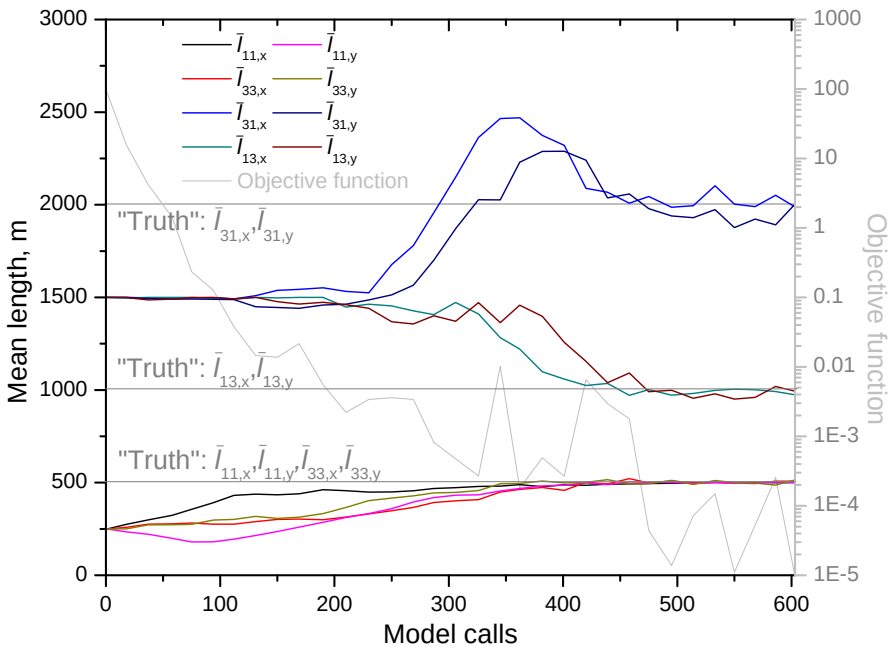


Figure 5.9: Unconstrained model inversion results. Parameter values and objective function are plotted as a function of the number of model calls.

to 250 m from their true value of 500 m, and the mean transition lengths were set to 1250 m from their true value of 1000 m. Based on Figure 5.5, it is apparent that mean facies length sensitivities (e.g. $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{11,y}$, and $\bar{l}_{33,y}$) are more smooth for values less than the truth in our synthetic example. This indicates that using a low estimate of the mean facies lengths will benefit the model inversion. Using this information, we have set the initial values of the mean facies lengths less than their truth. Future work will address the effect of the initial values in the inverse process.

The lowest objective function for the symmetric model inversion was achieved at 678 model calls, corresponding to 49 Levenberg-Marquardt optimization iterations. It was determined that no further progress would be achieved at 820 model calls, or 59 Levenberg-Marquardt optimization iterations.

Figure 5.9 presents the model inversion results for the unconstrained model. As indicated by equation (5.30), in this case it is necessary to specify 4 mean lengths for the spatial correlation model in each direction. Therefore, there are eight mean length parameters plotted in Figure 5.9. The initial values of the mean facies lengths were set to 250 m from their true value of 500 m, while the mean transition lengths are set to 1500 m from their true values of 2000 m ($\bar{l}_{31,x}$ and $\bar{l}_{31,y}$) and 1000 m ($\bar{l}_{13,x}$ and $\bar{l}_{13,y}$). These model inversion results confirm the sensitivity analysis presented in Figure 5.6. It is apparent that $\bar{l}_{11,x}$ converges towards its true value more quickly than the other parameters, as in the proportionally random and symmetrically constrained model inversions. It is also apparent that the mean transition lengths from the low conductivity to the high conductivity facies ($\bar{l}_{31,x}$ and $\bar{l}_{31,y}$) converge to their true values slower than the other parameters. This behavior is expected considering the relatively low, erratic sensitivities for these parameters in Figure 5.9. It is important to note that despite having poor sensitivities, these parameters do converge towards the "truth" during the inversion.

The lowest objective function for the unconstrained model inversion was achieved at 603 model calls, corresponding to 33 Levenberg-Marquardt optimization iterations. It was determined that no further progress would be achieved at 780 model calls, or 43 Levenberg-Marquardt optimization iterations.

5.6 Conclusions

While much effort has been directed towards geostatistical stochastic inversion methods utilizing covariance-based geostatistics, little effort has been spent on similar approaches employing Markov-chain based geostatistics. This paper demonstrates the feasibility of a stochastic inversion utilizing a Markov-chain model of spatial variability. The approach hinges on the use of inverse transition rates as adjustable

parameters, where inverse auto-transition rates are defined as mean facies lengths and inverse cross-transition rates are defined as mean transition lengths.

The proposed approach has been developed using established software applications (e.g. PEST (*Doherty*, 1994), T-PROGS *Carle* (1999)), thereby, hopefully providing an easily interpretable approach for utilizing stochastic theory in hydrogeology as suggested by *Dagan* (2004), *Neuman et al.* (2004), *Sudicky* (2004), and others.

Simplifying constraints to the Markov-chain model allowing reductions in the number of adjustable parameters are explored within the stochastic inversion. These simplifying constraints include proportionally random spatial correlations and symmetric correlations in non-background categories. In addition to these constrained models of spatial correlation, an unconstrained Markov-chain model was evaluated as well. Both constrained Markov-chain model inversions performed well, indicating that these models can be useful in cases where these simplifications are expected to be valid, or as references to be used in conjunction with an unconstrained analysis to determine the level of deviation from these simplified spatial correlation models. The unconstrained analysis demonstrates that it is not necessary to constrain the Markov-chain model to obtain a successful inverse analyses. The lack of constraints on the geostatistical model in the inverse process allows for characterization of wider ranges of aquifer heterogeneities.

Using the Chebyshev inequality to provide an overestimate of the length of the confidence limits of the sample mean of the drawdown predictions, we determine that 1000 realizations provide adequate convergence for the drawdown predictions at the scale of our synthetic models. In this way, we ensure that limitations are due to measurement imprecision, not sample statistic inference imprecision. It is important to note that a larger number of realizations is expected to produce smoother sensitivities, however the cost will be computational efficiency.

Sensitivity analyses indicate that the proportionally random model provides parameters with well-behaved sensitivities. The symmetrically constrained model produces decent sensitivities for all adjustable parameters, with the poorest sensitivities in the mean transition lengths. The mean facies length sensitivities are also decent for the unconstrained model, while significantly poorer sensitivities are observed for the mean transition length sensitivities. Nevertheless in general the shape of sensitivity curves are convex with relatively well defined minima at the "true" values. This suggested that gradient-based method for model inversion could be successfully applied in these cases. In general, the x-direction mean length parameters are more sensitive than their corresponding y-direction mean length parameters. This is likely the result that the x-direction is parallel to the flow direction and that the model domain is longer in the x-direction than the y-direction. Both factors are expected to provide better characterization of properties of the aquifer heterogeneities in the x-direction.

The example inversion runs demonstrate that despite some low sensitivity adjustable parameters, all parameter do converge towards their true values for all models. This demonstrates the feasibility of a Levenberg-Marquardt optimization on a Markov-chain based stochastic inversion. In contrast, previous analyses (*Harp et al.*, 2008) were performed using a stochastic global search method called AMALGAM (*Vrugt and Robinson*, 2007) which allowed for more detailed search for optimal solution, but at the expense of requiring a significant number of model calls. The successful application of a gradient-based method is important for practical applications where computational efficiency is critical. Further research is needed to extend this framework to allow variable and/or non-uniform hydraulic conductivity, greater number of stratigraphic units, and transport information.

CHAPTER APPENDIX

5.A Derivation of mean facies length and mean transition lengths

As demonstrated by *Carle and Fogg* (1996), mean facies length can be derived considering two categories. In the following, we review their derivation and present the derivation of the mean transition length for a 1-D Markov-chain model.

Considering the definition of an auto-transition rate, presented in equation (5.10), and the definition of the mean facies length, presented in equation (5.25), the auto-transition probability can be related to the mean facies length as

$$\frac{\partial t_{ii}(0)}{\partial h} = r_{ii} = -\frac{1}{\bar{l}_{ii}}. \tag{5.45}$$

Using the definition of a derivative, equation (5.45) can be expressed as

$$\lim_{h \to 0} \frac{t_{ii}(h) - t_{ii}(0)}{h} = -\frac{1}{\bar{l}_{ii}}. \tag{5.46}$$

Recognizing that $t_{ii}(0) = 1$ (i.e. probability of an auto-transition at lag 0 is 1) and assuming a two-category Markov-chain model, $t_{ij}(h)$ to be expressed as

$$t_{ij}(h) = 1 - t_{ii}(h) = t_{ii}(0) - t_{ii}(h). \tag{5.47}$$

To consider cases with greater than two categories, the jth category can represent

the union of all categories other than the ith category. Substituting equation (5.47) into equation (5.46) produces

$$\lim_{h\to 0} \frac{t_{ij}(h)}{h} = -\frac{1}{\bar{l}_{ii}}. \tag{5.48}$$

The two-location joint probability $p_{ij}(h)$ introduced in equation (5.34) can be defined as

$$p_{ij}(h) = \frac{T_{ij}(h)}{N(h)} \tag{5.49}$$

where $T_{ij}(h)$ is the number of transitions encountered from the ith to the jth category at lag h in a particular direction and $N(h)$ is the total number of lag intervals. Substituting equation (5.49) into equation (5.34) and dividing by h produces

$$\frac{t_{ij}(h)}{h} = \frac{T_{ij}(h)}{p_i N(h) h} \tag{5.50}$$

As $N(h) * h$ is the total length, $p_i N(h) h$ is the total length of the ith category. Since categories i and j are assumed to be mutually-exclusive exhaustively-defined categories, $T_{ij}(h)$ is equivalent to the number of embedded occurrences of the ith category. Therefore, substituting equation (5.50) into equation (5.48) produces

$$\lim_{h\to 0} \frac{T_{ij}(h)}{p_i N(h) h} = \frac{\text{Number of embedded occurences of the } i\text{th unit}}{\text{Total length of the } i\text{th unit}}. \tag{5.51}$$

Which leads to a linguistic definition of the mean facies length as

$$\bar{l}_{ii} = \frac{\text{Total length of the } i\text{th unit}}{\text{Number of embedded occurences of the } i\text{th unit}}. \tag{5.52}$$

The mean transition length can be derived similarly starting from the cross-transition rate definition presented in equation (5.10) and the definition of a mean transition length presented in equation (5.26) as

$$\frac{\partial t_{ij}(0)}{\partial h} = r_{ij} = \frac{1}{\bar{l}_{ij}}. \tag{5.53}$$

Using the definition of a derivative, equation (5.53) can be expressed as

$$\lim_{h \to 0} \frac{t_{ij}(h) - t_{ij}(0)}{h} = \frac{1}{\bar{l}_{ij}}. \tag{5.54}$$

Recognizing that $t_{ij}(0) = 0$ (i.e. probability of a cross-transition at lag 0 is 0) equation (5.54) can be expressed as

$$\lim_{h \to 0} \frac{t_{ij}(h)}{h} = \frac{1}{\bar{l}_{ij}}, \tag{5.55}$$

where $t_{ij}(h)/h$ has been defined in equation (5.50). This demonstrates that in a 2-category Markov-chain model, $\bar{l}_{ii} = \bar{l}_{ij}$, as presented in equation (5.29). However, in the derivation of the mean transition length, we are not restricted to 2-category Markov-chain models. Therefore, $\bar{l}_{ij}$ can be defined linguistically as

$$\bar{l}_{ij} = \frac{\text{Total length of the } i\text{th unit}}{\text{Number of transitions from the } i\text{th to the } j\text{th unit}}. \tag{5.56}$$

This indicates that $\bar{l}_{ij}$ is the mean length of the ith category between transitions from the ith to the jth category. This should not be confused with the total length between transitions from the ith to the jth category considering lengths of intervening

segments (i.e. segments from units other than the ith category) as the lengths of intervening segments are not considered in the mean transition length.

Chapter 6

Dissertation Conclusions

Chapters 2 through 5 present hydrogeological developments that attempt to reconcile our current limitations to adequately model hydrogeological phenomena with the operational demands of modeling decision support. These approaches present hydrogeological engineering solutions based on current hydrogeological theory and available data designed to provide technical understanding to support water management decisions. In the dissertation introduction, the approaches presented in Chapters 2 through 5 are discussed within the context of the current hydrogeological perspective. This section highlights the conclusions that can be attained from Chapters 2 through 5 with respect to the issues discussed in the dissertation introduction. Detailed conclusions for Chapters 2, 3, 4, and 5 are found in Sections 2.6, 3.6 ,4.6, and 5.6, respectively, and therefore will not be repeated here.

The conclusions from this research include:

1. Chapters 2 and 3 demonstrate hydrogeological investigation approaches utilizing an existing dataset from a water-supply/monitoring well network. This type of dataset is often neglected in hydrogeological investigations due to its highly complex nature. Given the measurement limitations and expense associated

with hydrogeological data collection, the use of methods to extract information from available data presents a cost effective alternative to traditional methods of pump tests.

2. Chapters 2 and 3 demonstrate inverse modeling approaches employing minimally parameterized models in the form of an analytical solution to the groundwater flow equation. Therefore, the inversions have restricted degrees of freedom compared to distributed-parameter models, enabling them to fail and provide incisive conclusions on the validity of their assumptions.

3. The parameter estimation method described in Chapter 2 provide cross-hole hydraulic characteristics consistent with conventional cross-hole pumping test estimates. As distinct property values are allowed for individual pumping/monitoring well pairs, inconsistent with the assumption of homogeneity implicit in the use of the Theis solution, these estimates do not estimate properties within the theoretical context of the Theis solution. Therefore, this approach presents a hydrogeological engineering solution, providing the decomposition of pumping influences at a monitoring location.

4. The parameter estimates obtained in Chapter 3 provide convergent, late-time estimates of effective aquifer properties (i.e. effective transmissivity and connectivity indicators) at inter-well support scale within the theoretical context of the Theis solution.

5. Chapters 4 and 5 demonstrate approaches to reduce the ill-posedness of inversions utilizing distributed-parameter models.

6. Chapters 4/5 introduce the use of Markov-chain-based geostatistics to identify/characterize hydrostratigraphies in an automated inverse framework utilizing hydraulic and geologic data.

7. Chapter 4 demonstrates the use of a single stochastic realization to represent the Markov-chain spatial correlation model in a geostatistical inversion. While the single realization approach provides a constraint on the inversion, and was successfully demonstrated in a 2-D, 2-stratigraphic unit model, issues concerning the inability to adequately characterize the hydraulic response of the geostatistical models or fully explore alternative, equally probable hydrostratigraphies within a geostatistical model are conceivable.

8. Chapter 5 demonstrates the ability to infer hydraulic characteristics of a Markov-chain geostatistical model to evaluate the plausibility of statistical hydrostratigraphic characteristics within an inverse framework. The approach is demonstrated on a 3-D, three-stratigraphic unit model. This approach recognizes information content limitations of current datasets, attempting to identify statistical characteristics of a hydrostratigraphy, as opposed to assuming that the information content of typical datasets can constrain an inversion to identify a single hydrostratigraphy.

Appendices

Appendix A

A Genetic-Fuzzy Approach for Modeling Complex Systems with an Example Application in Masonry Bond Strength Prediction[1]

Abstract

A genetic fuzzy learning from examples (GFLFE) approach is presented for determining fuzzy rule-bases generated from input/output datasets. The method is less computationally intensive than existing fuzzy rule base learning algorithms as the

[1]An edited version of this paper was published by ASCE. Copyright (2009) American Society of Civil Engineers: Harp, D.R., M. Reda Taha, and T.J. Ross (2009), A Genetic-Fuzzy Approach for Modeling Complex Systems with an Example Application in Masonry Bond Strength Prediction, *Journal of Computing in Civil Engineering, 23*(3), 193-199.

optimization variables are limited to the membership function widths of a single rule, which is equal to the number of input variables to the fuzzy rule base. This is accomplished by primary width optimization of a fuzzy learning from examples (FLFE) algorithm. The approach is demonstrated by a case study in masonry bond strength prediction. This example is appropriate as theoretical models to predict masonry bond strength are not available. The GFLFE method is compared to a similar learning method using constrained nonlinear optimization. Our results indicate that the use of a genetic optimization strategy as opposed to constrained nonlinear optimization provides significant improvement in the fuzzy rule base as indicated by a reduced fitness (objective) function and reduced root-mean-squared-error of an evaluation data set.

A.1 Introduction

Civil engineers are often confronted with systems of input/output data where the relationships between inputs and output is not well understood. Often, the physical basis of these relations are unknown or ambiguous. In other cases, the data may be too imprecise, or the complexity might be too large, to extract the relationships in a deterministic way. All too often, the situation is an unknown combination of these factors. In these situations, fuzzy set theory can be used to handle the uncertainty due to modeling ambiguity and/or data imprecision. Fuzzy set theory can be used where information in the form of complex relationships between inputs and output are extracted into a fuzzy rule-base. In order to extract a useful model in the form of a fuzzy rule-base from a numerical dataset, a method is needed to learn the fuzzy rule-base.

Genetic algorithms have been used to develop fuzzy rule-bases as well as to tune existing fuzzy rule-bases (*Cordón et al.*, 2004). The use of genetic algorithms to

fully develop a fuzzy rule-base, generally called *learning*, is a challenging task (*Ross*, 2004). Common approaches for general rule-base learning have been the Pittsburgh approach (*Smith*, 1980), the Michigan approach (*Holland and Reitman*, 1978), and iterative rule learning (*Venturini*, 1993). The Michigan approach and iterative rule learning methods encode each rule as a chromosome, while in the Pittsburgh approach, an entire set of rules is encoded in a single chromosome (*Castro and Camargo*, 2004). In this way, the Michigan approach and iterative rule learning are more computationally efficient, however they have the disadvantage that the individual rules are pitted against each other for survival to the next generation (*Castro and Camargo*, 2004). In the Pittsburgh approach, a population of rule-bases compete for survival, thereby allowing the collective synergism of the rules within a rule-base to influence the fitness of a particular rule-base (*Castro and Camargo*, 2004). The drawback of the Pittsburgh approach is the computational effort required to handle the complexity of chromosome evolution.

Tuning an existing fuzzy rule-base entails optimizing the membership function parameters of the rule-base. But tuning may also refer to adjusting the functions used to scale the input and output variables (*Cordón et al.*, 2004; *Jang et al.*, 1997). Although the process of tuning an existing rule-base is in general less complicated than learning a rule-base, the optimization can be complicated by the number of membership function parameters in the rule-base. The disadvantage of tuning an existing rule-base is that the search will be limited to solutions possessing the pre-specified structure. Learning algorithms, such as the one presented here, do not have this limitation as the structure is not pre-specified, but derived from the training data.

A methodology is presented here that allows the automatic extraction of the relationships between the inputs and output into a fuzzy rule-base. Our proposed method employs a genetic algorithm to optimize a modified learning from example

(MLFE) approach after *Passino and Yurkovich* (1998) (referred to in general as fuzzy learning from examples (FLFE) hereafter). The novelty of the proposed method lies in its ability to significantly enhance the learning process by solely optimizing the first rule developed by the FLFE algorithm. Existing learning algorithms that optimize an entire fuzzy rule base will be computationally more expensive.

The membership function widths for the first rule developed by the FLFE approach are differentiated from the subsequent membership function widths by the terms primary widths and secondary widths, respectively. This categorization is useful as the determination of the primary and secondary widths differ during the development of the fuzzy rule-base using FLFE. The primary widths must be specified to initiate the FLFE method, while the secondary widths are determined by the method itself to allow overlap of adjacent MFs using a pre-specified weighting factor (*Ross*, 2004). Varying the primary widths supplied to the FLFE method will alter the resulting rule base. Primary width optimization exploits this property by determining the most appropriate primary widths for the FLFE method.

The initialized rule base can be fine tuned by adjusting the output membership function centers using a recursive least squares method (*Ross*, 2004). A more complete fine tuning of the algorithm is possible using the gradient method (*Ross*, 2004), where all membership function centers and widths are tuned. Genetic FLFE (GFLFE) couples the computational efficiency of primary width optimization of the FLFE method and recursive least squares method with a global search strategy able to traverse the complicated fitness landscape of this optimization.

To illustrate the proposed methodology, an example civil engineering application is presented. The application extracts a fuzzy rule-base to model the flexural bond strength of masonry. While experimental research has indicated correlations between several input variables and flexural bond strength of masonry, the complex interactions of these various input variables are not well known (*Sugo*, 2000). We

utilize the results of a sensitivity analysis by *Reda Taha et al.* (2005) to select the most influential parameters for modeling flexural bond strength of masonry such as brick strength, absorption, sorptivity index, cement content and relative humidity. This application presents a case where the lack of knowledge about the complex interaction between inputs and output, or modeling ambiguity, can be dealt with by automatically extracting the relationships from the data into a fuzzy rule-base.

A.2 Methods

The following sections describe the rule base structure, discuss the fuzzy learning from examples (FLFE) method used for learning physical phenomenon, present the GFLFE approach to primary width optimization, and discuss the constrained nonlinear optimization method.

A.2.1 Rule base structure

The knowledge rule-base is comprised of a collection of if-then rules where the rules are in the form of a premise clause and an associated consequence. This can be considered a Mamdami system of deductive inference where the fuzzy linguistic terms have been replaced with fuzzy values (*Ross*, 2004). Here, we assume Gaussian MFs for the inputs described as

$$\mu(x) = exp\left[-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2\right] \tag{A.1}$$

where x is the input variable, c is the center of the membership function, and σ is the relative width of the membership function (*Ross*, 2004). It is important to note that

the choice of Gaussian MFs does not imply the use of probabilistic distributions, but were chosen as these MFs can be described by a single continuous equation. Delta functions have been used for the output MFs. Other MF types can be used for input and output MFs without greatly influencing the proposed method. The proposed method implements the center-average defuzzification method.

The fuzzy knowledge rule-base implements a product t-norm for the premise as

$$\mu_i(\boldsymbol{x}) = \prod_{j=1}^{n} \left[-\frac{1}{2} \exp \left(\frac{x_j - c_j^i}{\sigma_j^i} \right)^2 \right] \tag{A.2}$$

where $\mu_i(\boldsymbol{x})$ is the membership value of the input data-tuple $\boldsymbol{x}$ in the ith rule, c_j^i and σ_j^i are the center and width for the jth input of the ith rule, respectively, and n is the number of inputs (*Ross*, 2004). Along with center-average defuzzification, product implication is used for the output as

$$f(\boldsymbol{x}|\theta) = \frac{\sum_{i=1}^{R} b_i \mu_i(\boldsymbol{x})}{\sum_{i=1}^{R} \mu_i(\boldsymbol{x})} \tag{A.3}$$

where θ is a matrix that contains the rule-base parameters c_i, σ_i, and b_i, where b_i defines the delta function for the output of the ith rule, and R is the number of rules in the rule-base. By substituting Equation (A.2) into Equation (A.3), the output of

the rule-base can be explicitly described by the rule-base parameters; c, σ, and b, as

$$f(\boldsymbol{x}|\theta) = \frac{\sum_{i=1}^{R} b_i \prod_{j=1}^{n} \left[-\frac{1}{2} \exp\left(\frac{x_j - c_j^i}{\sigma_j^i} \right)^2 \right]}{\sum_{i=1}^{R} \prod_{j=1}^{n} \left[-\frac{1}{2} \exp\left(\frac{x_j - c_j^i}{\sigma_j^i} \right)^2 \right]}. \tag{A.4}$$

Equation (A.4) is the numerical equivalent of a collection of R fuzzy rules which employ Gaussian input (antecedent) MFs and delta function output (consequent) MFs with center average defuzzification.

A.2.2 Automated fuzzy learning

Similar to most fuzzy rule-base development methods, a collection of training data-tuples are required. The training data are separated into initialization and optimization data. The initialization data is processed into a first approximation of the rule-base using the FLFE method.

We start by assigning the values of the inputs for the first data-tuple in the initialization dataset to the centers of the input MFs and the associated output value to the delta function describing the output membership function for the first rule. This leaves the designation of the input membership function widths to fully describe the first rule. As there is no way to extract the values of these widths directly from the data, optimization can be used to appropriately determine these values. This is the basis of primary width optimization, and these are the MF widths introduced previously as primary widths.

The proposed method then iterates through the remaining data-tuples in the initialization dataset, determining that a new rule is necessary when the following

inequality is violated:

$$\epsilon_f \leq |f(\boldsymbol{x}|\theta) - y| \tag{A.5}$$

where ϵ_f is a specified test factor, $f(\boldsymbol{x}|\theta)$ is the fuzzy rule-base output described by Equation (A.4), and y is the measured output for the current data-tuple. If inequality (A.5) is violated, a new rule is created by assigning the inputs of the current data-tuple to the centers of the input MFs and the output of the current data-tuple to the delta function for the output membership function for the new rule. The membership function widths for each new rule, introduced as secondary widths previously, are calculated to allow for the overlap of adjacent MFs specified by the weighting factor w as

$$\sigma_s = \frac{1}{w}|c - c_{min}| \tag{A.6}$$

where c is the center of the membership function in question, and c_{min} is the nearest existing center to c (*Ross*, 2004). It is in this way that the FLFE method is able to derive all MF parameters, except for the primary widths, directly from the training data. Once the algorithm has iterated through the initialization dataset, the initialized rule-base is passed to a recursive least squares method in order to fine tune the output membership functions (*Passino and Yurkovich*, 1998).

A.2.3 Genetic fuzzy learning from examples (GFLFE)

Fig. A.1 presents a diagram of the primary width optimization method for fuzzy rule-base development described below. The fitness function (referred to as the objective

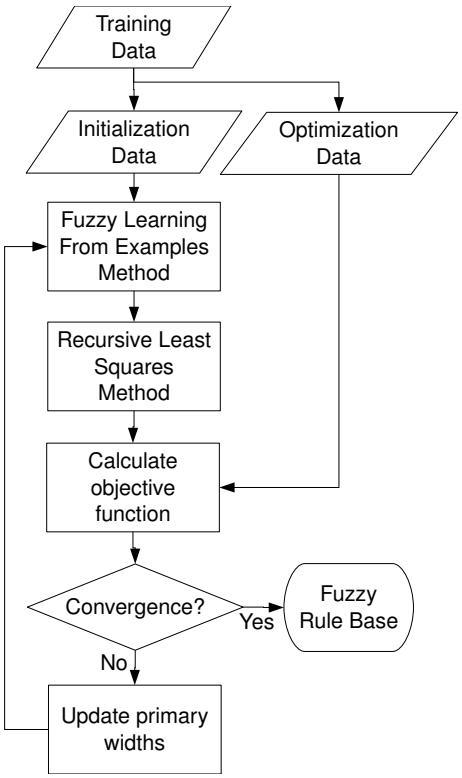


Figure A.1: Primary width optimization flow diagram.

function in the constrained nonlinear case) is defined as

$$\min_{\beta \in \mathbf{B}} \left[\frac{1}{m} \sum_{k=1}^{m} (f(\boldsymbol{x}_k|\theta) - y_k)^2 \right]^{0.5} \tag{A.7}$$

where β is a vector containing the primary widths constrained by $\mathbf{B}$, m is the number of data-tuples in the optimization dataset, $f(\boldsymbol{x}_k|\theta)$ is the fuzzy rule-base output for the kth data-tuple $\boldsymbol{x}_k$ of the optimization dataset, and y_k is the kth measured output in the optimization dataset. The use of a designated optimization dataset as opposed to reusing the initialization dataset ensures that the fuzzy rule-base is not over-trained. Over-training or over-fitting tends to produce systems with limited

ability to generalize the phenomenon being learned (*Haykin*, 1998). By substituting Equation (A.4) into Equation (A.7), the primary widths ($\sigma_j^1, j = 1, \ldots, n$) can be expressed explicitly in the fitness function as

$$\min_{\beta \in \mathbf{B}} \left[\frac{1}{m} \sum_{k=1}^{m} \left(\left(\frac{\sum_{i=1}^{R} b_i \prod_{j=1}^{n} \left[-\frac{1}{2} \exp\left(\frac{x_{k,j} - c_j^i}{\sigma_j^i} \right)^2 \right]}{\sum_{i=1}^{R} \prod_{j=1}^{n} \left[-\frac{1}{2} \exp\left(\frac{x_{k,j} - c_j^i}{\sigma_j^i} \right)^2 \right]} \right) - y_k \right)^2 \right]^{0.5} \tag{A.8}$$

where $x_{k,j}$ is the jth input of the kth data-tuple of the optimization dataset. It is important to note that primary width optimization is not simply the minimization of Equation (A.8) by varying the primary widths. The optimization also includes the FLFE learning algorithm, which, given different values of primary widths, will alter the rest of the rule base further by influencing the selection of rules and altering the secondary widths ($\sigma_j^i,\ j = 1, \ldots, n\ i = 2, \ldots, R$) and output membership function centers ($b_i,\ i = 1, \ldots, R$), as described in section A.2.2.

A binary genetic algorithm is used in this research, where a population of possible solutions is converted to binary form for implementation of the genetic algorithm (*Haupt and Haupt*, 2004). The population of possible solutions corresponds to a collection of alternative sets of primary widths. The fitness of these solutions within the population is determined using the fitness function (Equation (A.8)). A stochastic uniform selection method is employed, where parents are laid out along a line where their length is proportional to their scaled length based on their fitness. The algorithm moves along the line in equal steps, selecting the parent it lands on at each step. Crossover is performed using a random binary vector, where a value of zero indicates the gene from one parent is passed to the offspring, while a one indicates the gene from the other parent is passed to the offspring. The number of offspring derived from crossover at each generation is set to 80%. Mutation is performed by

selecting a random number from a Gaussian distribution with an initial variance equal to two times the length of the range of the variable, specified by the upper and lower bounds. The variance was decreased linearly until it reached zero at the last generation. A single optimal solution, known as an elite, is guaranteed to survive each generation. These methods emulate natural selection and are designed to allow optimal solutions to evolve (*Haupt and Haupt*, 2004).

The initial population used in the optimization was created using a uniform distribution of values, where the values were restricted to provide realistic membership function widths and numerical stability. A parametric study to identify GA population size was performed. This parametric analysis is presented in Figure A.2, where the number of function evaluations necessary to reduce the fitness function below a designated value is evaluated 20 times for each population size from 20 to 200 at increments of 20. By inspecting Figure A.2, it is apparent that a reduced number of function evaluations and standard deviation are achieved at a population size of 120, without significant improvements at larger population sizes. Therefore, a population size of 120 was selected for the GA optimization process. Similar results were obtained from a parametric study of population size for a bare-soil evaporation case study (not presented here), however, it is not verified that these results will apply to all case studies.

Primary widths outside of the prescribed range are discouraged in successive generations by penalizing the fitness function (Equation (A.8)). The number of generations was set to 150 as this appeared to enable the optimization process to converge to a stable solution.

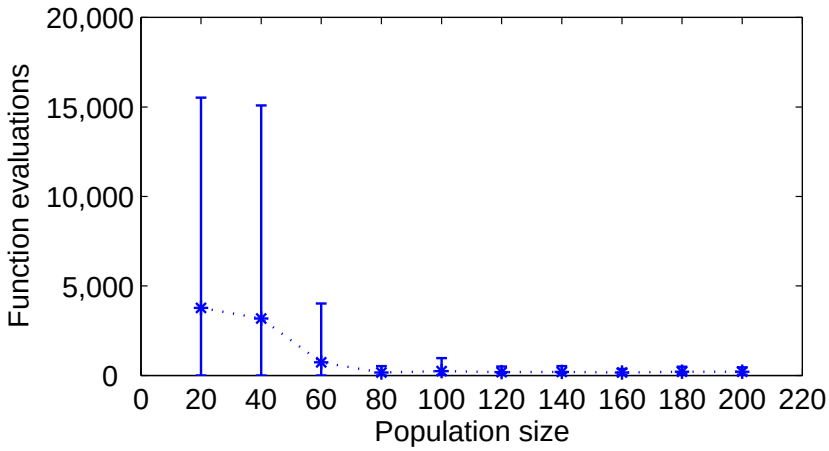


Figure A.2: Function (learning algorithm) evaluations versus genetic algorithm population size. The number of function evaluations necessary to reduce the fitness function below 0.13 MPa is determined 20 times at each population size from 20 to 200 at increments of 20. The stars connected by a dotted line represent the mean number of function evaluations, while the error bars indicate the standard deviation about the mean at each population size.

A.2.4 Constrained nonlinear optimization

The GFLFE method is compared to a similar learning approach using constrained nonlinear optimization (CNO) proposed by the authors previously (*Harp et al.*, 2007). The two methods are similar in that both learning methods attempt to develop optimal knowledge rule bases by primary width optimization. This comparison demonstrates the necessity of using a global optimization strategy as indicated by the improved performance of the GFLFE method. Primary width optimization by constrained nonlinear optimization is performed using Sequential Quadratic Programming (SQP) here. The SQP method closely mimics the Newton method for constrained optimization. As the calculation of exact derivatives is not possible given the nature of the optimization, quasi-Newton methods must be used to approximate their values. A quasi-Newton updating method is used at each major iteration to

approximate the Hessian, or second derivatives, of the Lagrangian function ($\mathcal{L}(x, \lambda)$) (*Nocedal and Wright*, 1999) defined as

$$\mathcal{L}(x, \lambda) = f(x) - \lambda_i c_i(x) - \ldots - \lambda_n c_n(x) \tag{A.9}$$

where $f(x)$ is the objective function, c_i is the ith constraint placed on the optimization, λ_i is the ith Lagrange multiplier, and n is the number of constraints. This approximation is used to generate a Quadratic Programming (QP) subproblem that is used to form a search direction for a line search procedure.

A.3 Case Study

The case study involves the evaluation of masonry bond strength by testing masonry prisms made of four types of mortar and four types of brick units under two curing regimes. The masonry prisms were tested at three ages up to a year creating an experimental database of 96 data tuples. Table A.1 presents the mix proportions of four masonry mortars used in the experiments, while Table A.2 presents the properties of the brick units. Twenty-four, five-high stack bonded prisms were constructed from each mortar type and brick unit. Each mix was tested at dry (20% relative humidity (RH)) and moist (100% RH) airing conditions at 20°C. Four prisms were tested from each curing condition at 28, 180, and 360 days of age. The masonry bond strength was examined using a bond wrench test apparatus as designed by *Shrive and Tillemann* (1992). The top half of the brick is gripped between two neoprene pads in a clamp and torque wrench is attached to the clamp such that the centerline of the torque wrench arm is centered over the brick. The method complies with the requirements of ASTM C1072-94 1994. Table A.3 presents the bond strength database. Six input parameters were selected based on a ranking of significance

Table A.1: Mix proportions by volume of the four masonry mortar used in the experimental program.

Mortar Group	Portland Cement	Hydrated Lime	Fly Ash Type (F)	Sand
A	1	0.5	0	4.5
B	1	1	0	6
C	0.8	0.4	0.3	4.5
D	0.8	0.8	0.4	6

Table A.2: Brick types and properties.

Group Designation	Compressive Strength (MPa)	IRA[a] ($kg/m^2/min$)	Total Absorption%	Sorptivity Index
1	43.9	3.67	8.11	2347
2	38.5	6.55	7.98	5781
3	59.2	2.40	8.32	1465
4	72.0	2.28	6.70	1216

[a]IRA (Initial rate of absorption)

performed by *Reda Taha et al.* (2005). These include: Portland cement volume, masonry age, curing index, compressive strength, total absorption, and sorptivity.

A.4 Results and Discussion

A comparison of the performance of constrained nonlinear and genetic optimization strategies in primary width optimization is presented. The comparison includes the results of 20 sequential optimization runs using each strategy starting from random initial values. A training data set of 72 data-tuples is separated into an initialization data set (48 data-tuples) and an optimization data set (24 data-tuples). The resulting

Table A.3: Bondstrength (MPa) for all brick and mortar types.

Mortar Mix	Age, days	Brick Type							
		1		2		3		4	
		Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry
A	28	0.59	0.47	0.37	0.35	0.39	0.46	0.48	0.45
	180	1.42	1.16	0.72	0.6	.071	0.59	1.06	0.68
	360	1.19	0.94	0.72	0.6	0.71	0.59	0.97	0.67
B	28	0.67	0.86	0.43	0.43	0.39	0.34	0.6	0.5
	180	1.16	0.94	0.89	0.49	0.72	0.4	0.97	0.64
	360	1.2	0.95	1.01	0.55	0.78	0.38	0.99	0.69
C	28	0.63	0.62	0.42	0.36	0.44	0.33	0.59	0.51
	180	0.88	0.66	0.7	0.39	0.76	0.38	0.78	0.7
	360	0.96	0.75	0.81	0.42	0.78	0.45	0.82	0.68
D	28	0.48	0.59	0.34	0.26	0.38	0.32	0.58	0.47
	180	0.87	0.65	0.42	0.38	0.57	0.51	0.86	0.58
	360	0.88	0.67	0.53	0.36	0.61	0.47	0.88	0.58

objective function values (although fitness value would be more appropriate for the genetic algorithm case, objective function value will be used here for both the CNO approach and GFLFE method, as the fitness and objective functions are equivalent here) are presented in box and whisker plots in Figure A.3, where the horizontal lines indicate the lower quartile, median, and upper quartile of the objective function values and the whiskers indicate the extent of the remaining values. It is apparent in this figure that the GFLFE method produces lower objective function values. The GFLFE method runs result in a mean fitness value of 0.121 MPa with a standard of deviation of 0.0072, while the CNO approach result in a mean of 0.138 with a standard of deviation of 0.0073. This indicates a decrease of over 12% in the fitness value using the GFLFE method with nearly equivalent standard deviation. The average run time on a 2.80 GHz processor for the GFLFE method is 9.88 minutes with a standard deviation of 0.86 minutes, while the average for the constrained nonlinear optimization is 0.40 minutes with a standard of deviation of 0.28 minutes. While

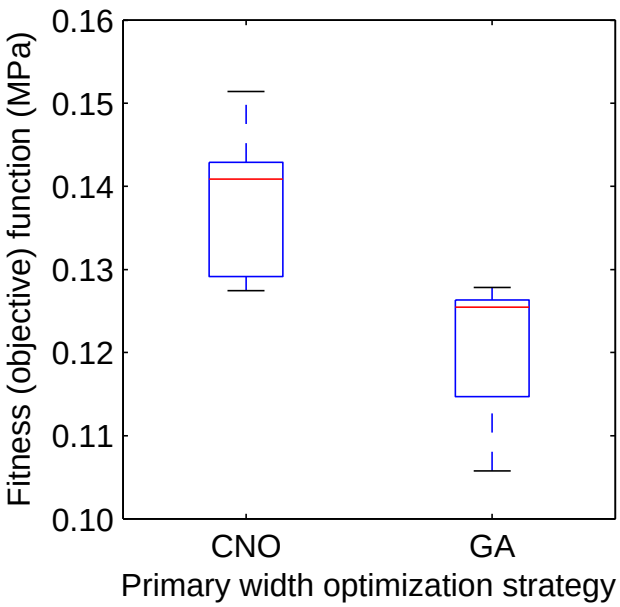


Figure A.3: Box and whisker plots comparing the ability of a constrained nonlinear optimization (CNO) and a genetic algorithm (GA) to reduce the fitness (objective) function in primary width optimization. The plot presents results for 20 runs using each optimization strategy. The horizontal lines indicate the lower quartile, median, and upper quartile values, while the whiskers indicate the extent of the remainder of the results. In both cases, no outliers were identified.

this indicates that GFLFE run time was longer than that of the CNO approach, it is important to realize that the CNO approach clearly converged to suboptimal solutions. The relatively long run time for the GFLFE approach is expected as the method relies on the evolution of optimal solutions, as opposed to a gradient-based search utilized by the CNO approach. The advantage of the evolutionary approach is the ability to escape local minima providing a strategy for locating globally optimal solutions for complex objective functions.

Table A.4 presents the optimized primary widths for the run with the lowest objective function value for both optimization strategies. It is apparent that there are

significant differences in the optimized primary widths for the two cases. This is an indication of the difficulty of primary width optimization, providing support for the use of a global search strategy by GA as opposed to a local search strategy by CNO. The input and output MFs corresponding to the GFLFE primary widths presented in Table A.4 are presented in Figure A.5. Thirteen rules have been developed during the learning algorithm resulting in an efficiency quotient (No. training data-tuples/No. of rules) of greater than 5.5.

Finally, an evaluation data set, comprised of 24 data-tuples that are excluded from the training data, is used to compare the ability of the developed fuzzy rule bases to predict bondstrength. In both cases, the fuzzy rule base resulting from the optimization run with the lowest objective function value is used to make the predictions. Figure A.4 presents these results in predicted versus measured plots. The root-mean-squared-error (RMSE) for the constrained nonlinear optimization fuzzy rule base (Figure A.4(a)) is 0.637 MPa and for GFLFE (Figure A.4(b)) 0.610 MPa. This indicates a decrease of over 4% in RMSE using the GFLFE method.

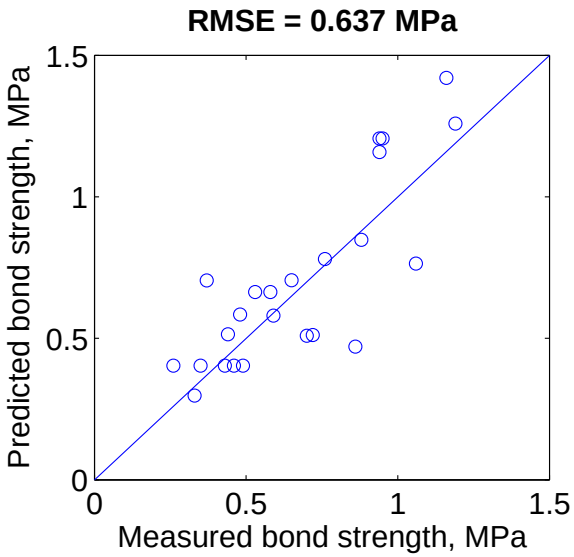
A.5 Conclusions

A genetic algorithm-based approach for primary width optimization has been presented for creating fuzzy rule-bases from input/output datasets providing a consistent modeling approach. The fuzzy rule-bases are used to describe linguistically the knowledge rule-base of the input/output process. Such information cannot be deduced from other methods such as neural nets. This approach can be used to extract inferences describing the relationships between inputs and an output for complex systems. The most appropriate use for the GFLFE method is in the case where modeling ambiguity and/or data imprecision exist, as is often the case in practice.

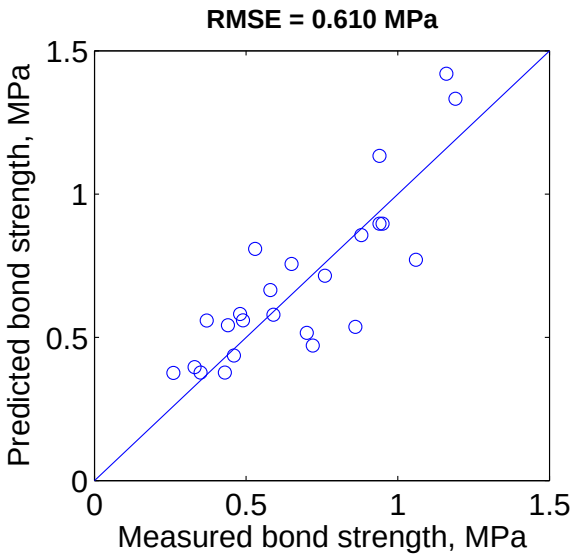
Existing methods for learning fuzzy rule-bases using genetic algorithms require

significant computational effort, while existing tuning algorithms are limited in their ability to find optimal solutions. We suggest that a genetic fuzzy learning from examples method that targets primary width optimization can significantly reduce the complexity of the optimization process.

A comparison of modeling masonry bond strength using the GFLFE approach and a similar approach using CNO (*Harp et al.*, 2007) is presented. This comparison demonstrates the ability of the GFLFE method to reduce the objective function by over 12% and enhance learning. GFLFE demonstrates good learning efficiency for the case study with an efficiency quotient of over 5.5. While the increase in run time for the GFLFE method is significant, we propose that the additional computational effort is justified in most modeling scenarios given the improvements in the rule base resulting in increased prediction accuracy.



(a) CNO-FLFE



(b) GFLFE

Figure A.4: Plots of measured versus predicted bond strength for an evaluation data set using models developed by (a) nonlinear constrained primary width optimization (CNO-FLFE) and (b) GFLFE.

Table A.4: Optimized primary widths.

Optimization Strategy	Cement volume	Masonry age, days	Curing index	Compressive strength, MPa	Total absorption%	Sorptivity Index
Constrained nonlinear	0.0249	67.15	0.320	13.35	0.562	1826
GFLFE	0.0611	114.6	0.096	13.00	0.582	1678

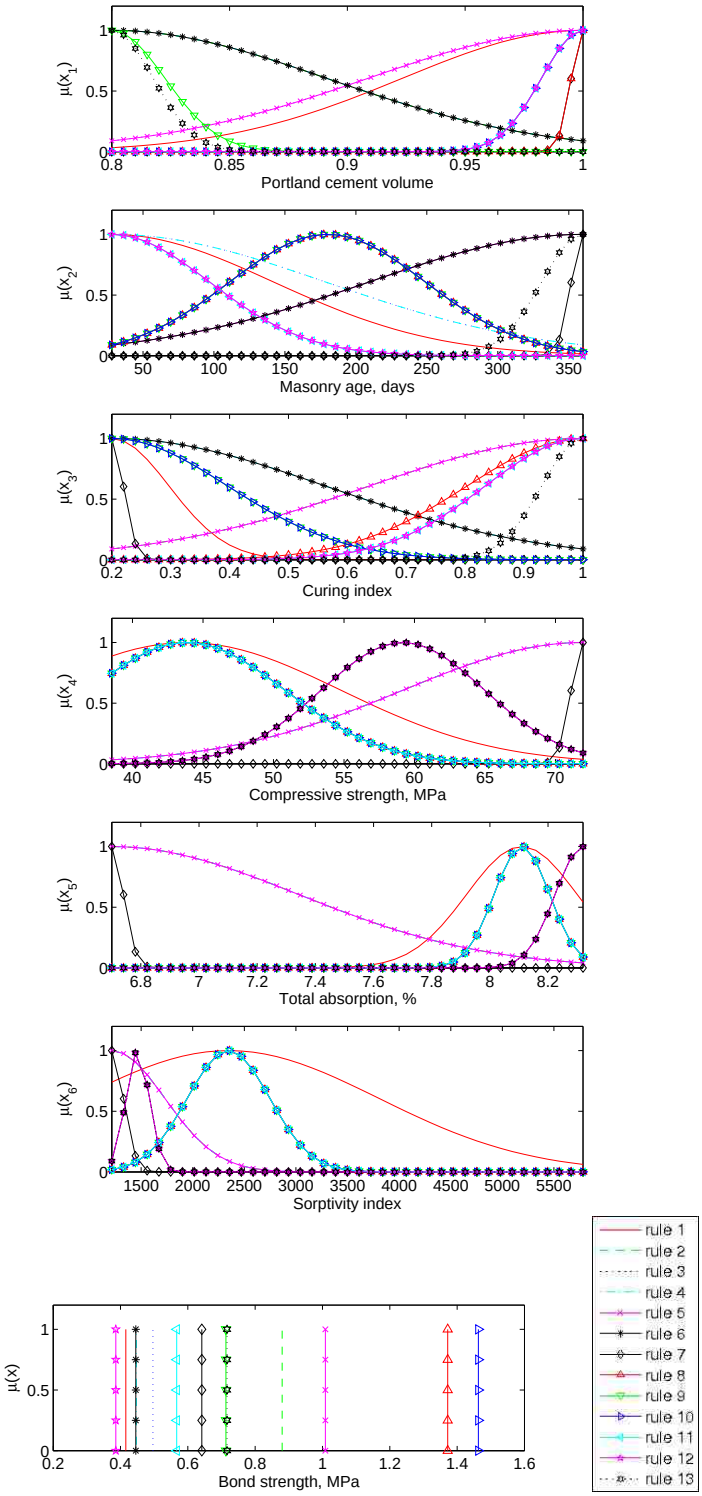


Figure A.5: Input and output MFs developed by GFLFE

References

Agterberg, F. (1974), *Geomathematics*, 596 pp., Elsevier Sci., New York.

Allen, S., and R. Koch (2008), Groundwater level status report for fiscal year 2007, *Progress Report LA-14358-PR*, Los Alamos National Laboratory.

ASTM (1994), Standard specifications for standard test method for measurement of masonry flexural bond strength, in *C1072-94*, Philadelphia, Pa.

Ballio, F., and A. Guadagnini (2004), Convergence assessment of numerical Monte Carlo simulations in groundwater hydrology, *Water Resources Research*, *40*, W04603, doi:10.1029/2003WR002876.

Barnett, D., J. Rieger, and E. Thornton (2003), Results of tritium tracking and groundwater monitoring at the Hanford Site 200 area state-approved land disposal site–fiscal year 2003, *Tech. Rep. PNNL-14449*, Pacific Northwest National Laboratory.

Bear, J. (1972), *Dynamics of Fluids in Porous Media*, American Elsevier.

Ben Ameur, H., G. Chavent, and J. Jaffre (2002), Refinement and coarsening indicators for adaptive parametrization: Application to the estimation of hydraulic transmissivities, *Inverse Problems*, *18*, 775–794.

Beven, K. (1993), Prophecy, reality, and uncertainty in distributed hydrological modelling, *Advances in Water Resources*, *16*, 41–51.

Beven, K. (2000), Uniqueness of place and process representations in hydrological modelling, *Hydrology and Earth Systems Sciences*, *4*(2), 203–213.

Beven, K. (2002), Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system, *Hydrological Processes*, *16*, 189–206.

Beven, K. (2006), A manifesto for the equifinality thesis, *Journal of Hydrology*, *320*, 18–36.

Bureau de Recherches Géologiques et Miniéres (1990), Interpretation of hydraulic tests (pulse tests, slug tests and pumping tests) from the third testing campaign at the cabril site (Spain) (in French), technical report for Spanish Nuclear Waste Management Company.

Butler, J. (1988), Pumping tests non-uniform aquifers—the radially symmetric case, *Journal of Hydrology*, *101*(1/4), 15–30.

Butler, J. (1990), The role of pumping tests in site characterization: Some theoretical considerations, *Ground Water*, *28*(3), 394–402.

Carle, S. (1999), *T-PROGS: Transition Probability Geostatistical Software, Version 2.1*, Hydrological Sciences Graduate Group, University of California, Davis.

Carle, S., and G. Fogg (1996), Transition probability-based indicator geostatistics, *Mathematical Geology*, *28*(4), 453–476.

Carle, S., and G. Fogg (1997), Modeling spatial variability with one and multidimensional continuous-lag markov chains, *Mathematical Geology*, *29*(7), 891–918.

Carle, S., E. LaBolle, G. Weissmann, D. VanBrocklin, and G. Fogg (1998), Conditional simulation of hydrofacies architecture: A transition probability/markov approach, in *Hydrogeological Models of Sedimentary Aquifers SEPM Concepts in*

Hydrogeol. Environ. Geol., edited by O. Fraser and J. Davis, pp. 147–170, Soc. for Sediment. Geol., Tulsa, OK.

Carrera, J., and S. Neuman (1986), Estimation of aquifer parameters under steady-state and transient conditions: 2. uniqueness, stability, and solution algorithms, *Water Resour. Res.*, *22*(2), 211–227.

Carrera, J., A. Alcolea, A. Medina, J. Hidalgo, and L. Slooten (2005), Inverse problem in hydrogeology, *Hydrogeology Journal*, *13*, 206–222.

Castro, P. A., and H. A. Camargo (2004), Learning and optimization of fuzzy rule base by means of self-adaptive genetic algorithm, in *Proceedings of the 13th IEEE International Conference on Fuzzy Systems*, vol. 2, pp. 1037–1042, Budapest, Hungary.

Chen, J., and Y. Rubin (2003), An effective bayesian model for lithofacies estimation using geophysical data, *Water Resour. Res.*, *39*(5), 1118, doi:10.1029/2002WR001666.

Cooley, R. (1983), Incorporation of prior information on parameters into nonlinear regression groundwater flow models: 2. applications, *Water Resour. Res.*, *19*(3), 662–676.

Cooper, H., and C. Jacob (1946), A generalized graphical method for evaluating formation constants and summarizing well-field history, *Eos Trans. AGU*, *27*(4), 526–534.

Cordón, O., F. Gomide, F. Herrera, F. Hoffman, and L. Magdalena (2004), Ten years of genetic fuzzy systems: Current framework and new trends, *Fuzzy Sets and Systems*, *141*, 5–31.

Dagan, G. (2004), On application of stochastic modeling of groundwater flow and

transport, *Stochastic Environmental Research and Risk Assessment*, *18*, 266–267, doi:10.1007/s04477-004-0191-7.

Dai, Z., and J. Samper (2006), Inverse modeling of water flow and multicomponent reactive transport in coastal aquifer systems, *J. of Hydrology*, *327*, 447–461.

Dai, Z., R. Ritzi, and D. Dominic (2005), Improving permeability semivariograms with transition probability models of hierarchical sedimentary architecture derived from outcrop-analog studies, *Water Resour. Res.*, *41*, W07032, doi:10.1029/2004WR003515.

Dai, Z., A. Wolfsberg, Z. Lu, and R. Ritzi Jr. (2007), Representing aquifer architecture in macrodispersivity models with an analytical solution of the transition probability matrix, *Geophysical Research Letters*, *34*, L20406, doi:10.1029/2007GL031608.

Dawson, K., and J. Istok (1991), *Aquifer Testing: Design and Analysis of Pumping and Slug Tests*, Lewis Publishers.

de Marsily, G., G. Lavedan, M. Boucher, and G. Fasanino (1984), Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model, in *Proc Geostatistics for Natural Resources Characterization. Part 2*, edited by Verly et al, D. Reidel Pub. Co.

de Marsily, G., F. Delay, J. Gonçalvés, P. Renard, V. Teles, and S. Violette (2005), Dealing with spatial heterogeneity, *Hydrogeology Journal*, *13*(1), 161–183.

Deutsch, C., and A. Journel (1992), *GSLIB Geostatistical Software Library and User's Guide*, Oxford University Press, New York.

Doherty, J. (1994), *PEST Model-Indpendent Parameter Estimation*, Watermark Computing, Corinda, Australia.

Doherty, J. (2003), Groundwater model calibration using pilot points and regularization, *Ground Water*, *41*(2), 170–177.

Doherty, J. (2004), *PEST Model-Independent Parameter Estimation*, Watermark Numerical Computing, Corinda, Australia.

Eppstein, M., and D. Dougherty (1996), Simultaneous estimation of transmissivity values and zonation, *Water Resour. Res.*, *32*(11), 3321–3336.

Freeze, R., and J. Cherry (1979), *Groundwater*, Prentice Hall, Englewood Cliffs, NJ.

Frick, U. (1992), Grimsel test site, the Radionuclide Migration Experiment—overview of investigations 1985-1990, *Tech. Rep. Nagra Tech. Rep. 91-04*, Nagra, Wettingen, Switzerland.

Gégo, E., G. Johnson, and M. Hankins (2001), An evaluation of methodologies for the generation of stochastic hydraulic conductivity fields in highly heterogeneous aquifers, *Stochastic Environmental Research and Risk Assessment*, *15*, 47–64.

Grayson, R., I. Moore, and T. McMahon (1992), Physically based hydrologic modeling 2. is the concept realistic?, *Water Resources Research*, *26*(10), 2659–2666.

Gross, L. (2007), Use of geostatistics to evaluate the monitoring well network of a TCE plume, *International Journal of Environment and Pollution*, *29*(4), 370–382.

Harp, D., and V. Vesselinov (2009), Identification and analysis of long-term water-level fluctuations due to spatially and temporally variable water-supply pumping, *Ground Water*, in review.

Harp, D., M. Reda Taha, J. Stormont, E. Farfan, and J. Coonrod (2007), An evaporation estimation model using optimized fuzzy learning from example algorithm with an application to the riparian zone of the Middle Rio Grande in New Mexico, U.S.A., *Ecological Modelling*, *208*, 119–128.

Harp, D., Z. Dai, A. Wolfsberg, J. Vrugt, B. Robinson, and V. Vesselinov (2008), Aquifer structure identification using stochastic inversion, *Geophysical Research Letters*, L08404, doi:10.1029/2008GL033585, l08404, doi:10.1029/2008GL033585.

Haupt, R., and S. Haupt (2004), *Practical Genetic Algorithms*, second ed., John Wiley & Sons Inc., Hoboken, NJ.

Haykin, S. (1998), *Neural Networks: A Comprehensive Foundation*, Prentice Hall, Englewood Cliffs, N.J., USA.

Heredia, J., A. Medina, and J. Carrera (2000), Estimation of parameter geometry, in *Computational Methods for Flow and Transport in Porous Media*, edited by J. Crolet, Kluwer, Dordrecht.

Herweijer, J. (1996), Constraining uncertainty of groundwater flow and tranport models using pumping tests, in *Calibration and Reliability in Groundwater Modelling*, pp. 473–482, IAHS Publ., 237.

Herweijer, J., and S. Young (1991), Use of detailed sedimentological information for the assessment of aquifer tests and tracer tests in a shallow fluvial aquifer, in *Proceedings of the 5th Annual Canadian/American Conference on Hydrogeology: Parameter Identification and Estimation for Aquifer and Reservoir Characterization*, Natl. Water Well Assoc., Dublin, Ohio.

Hix, G. (2007), Ground-water monitoring in the Tucson Basin, Pima County, Arizona, *Ground Water Monitoring and Remediation*, *1*, 36–38.

Holland, J., and J. Reitman (1978), *Cognitive Systems Based on Adaptive Algorithms, in: D.A. Waterman, F. Hayes-Roth (Eds.), Pattern-Directed Inference Systems*, 313-329 pp., Academic Press, New York.

Jacob, C. (1940), On the flow of water in an elastic artesian aquifer, *Trans. Amer. Geophys. Union*, *21*, 574–586.

Jang, J., C. Sun, and E. Mizutani (1997), *Neuro-Fuzzy and Soft Computing, A Computational Approach to Learning and Machine Intelligence*, Prentice Hall, NJ, USA.

Journel, A. (1983), Nonparametric estimation of spatial distribution, *Mathematical Geology*, *15*(3), 445–468.

Kirchner, J. (2006), Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, *Water Resources Research*, *42*, W03S04, doi:10.1029/2005WR004362.

Kitanidis, P. (1996), On the geostatistical approach to the inverse problem, *Adv. Water Resour.*, *19*(6), 333–342.

Klemes, V. (1988), A hydrological perspective, *Journal of Hydrology*, *100*, 3–28.

Knudby, C., and J. Carrera (2006), On the use of apparent hydraulic diffusivity as an indicator of connectivity, *J. of Hydrology*, *329*, 377–389.

Koch, R., and S. Schmeer (2009), Groundwater level status report for 2008, Los Alamos National Laboratory, *Tech. Rep. LA-14397-PR*, Los Alamos National Laboratory.

LaBolle, E., and G. Fogg (2001), Role of molecular diffusion in contaminant migration and recovery in an alluvial aquifer system, *Transport in Porous Media*, *42*(1-2), 155–179.

Levenberg, K. (1944), A method for the solution of certain nonlinear problems in least squares, *Q. Appl. Math.*, *2*, 164–168.

LeVenue, A., B. RamaRao, G. de Marsily, and M. Marietta (1995), Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields 2. application, *Water Resources Research*, *31*(3), 475–493.

Los Alamos National Laboratory (2008a), Pajarito canyon investigation report, *Tech. Rep. LA-UR-085852*, Los Alamos National Laboratory.

Los Alamos National Laboratory (2008b), Fate and transport investigations update for chromium contamination from Sandia Canyon, *Tech. rep.*, Environmental Programs Directorate, LANL, Los Alamos, New Mexico, lA-UR-08-4702.

Lu, Z., and B. Robinson (2006), Parameter identification using the level set method, *Geophysical Research Letters*, *33*, L06404, doi:10.1029/2005GL025541.

Lu, Z., and D. Zhang (2002), On stochastic modeling of flow in multimodal heterogeneous formations, *Water Resources Research*, *38*(10), 1190, doi:10.1029/2001WR001026.

Marquardt, D. (1963), An algorithm for least-squares estimation of nonlinear parameters, *J. Soc. Ind. Appl. Math.*, *11*, 431–441.

Mason, J., S. Sebree, and T. Quinn (2005), Monitoring-well network and sampling design for ground-water quality, Wind River Indian Reservation, Wyoming, *Scientific Investigations Report 2005-5027*, USGS.

Matheron, G. (1967), Eléments pour une théorie des milieux poreux [elements for a theory of porous media], Masson, Paris.

McDonnell, J., M. Sivapalan, K. Vaché, S. Dunn, G. Grant, R. Haggerty, C. Hinz, R. Hooper, J. Kirchner, M. Roderick, J. Selker, and M. Weiler (2007), Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology, *Water Resources Research*, *43*, W07301, doi:10.1029/2006WR005467.

McLin, S. (2005), Analyses of the PM-2 aquifer test using multiple observation wells, *Tech. Rep. LA-14225-MS*, Los Alamos National Laboratory.

McLin, S. (2006a), Analyses of sequential aquifer tests from the guaje well field, *Tech. Rep. LA-UR-06-2494*, Los Alamos National Laboratory.

McLin, S. (2006b), Analyses of the PM-4 aquifer test using multiple observation wells, *Tech. Rep. LA-14252-MS*, Los Alamos National Laboratory.

Meier, P., J. Carrera, and X. Sanchez-Vila (1998), An evaluation of Jacob's method for the interpretation of pumping tests in heterogeneous formations, *Water Resources Research*, *34*(5), 1011–1025.

Neuman, S. (1973), Calibration of distributed parameter groundwater flow models viewed as a multiple-objective decision process under uncertainty, *Water Resources Research*, *9*(4), 1006–1021.

Neuman, S. (1987), Stochastic continuum representations of fractured rock permeability as an alternative to the REV and fracture network concepts, in *Rock Mechanics: Proceedings of the 28th U.S. Symposium*, edited by I.W. Farmer et al., pp. 533–561, Springfield, Vt.

Neuman, S. (1990), Universal scaling of hydraulic conductivities and dispersivities in geologic media, *Water Resources Research*, *26*(8), 1749–1758.

Neuman, S., and V. Di Federico (2003), Multifaceted nature of hydrogeologic scaling and its interpretation, *Reviews of Geophysics*, *41*(3), doi:10.1029/2003RG000130.

Neuman, S., and P. Witherspoon (1972), Field determination of the hydraulic properties of leaky multiple aquifer systems, *Water Resour. Res.*, *8*(5), 1284–1298.

Neuman, S., A. Guadagnini, and M. Riva (2004), Type-curve estimation of statistical heterogeneity, *Water Resources Research*, *40*, W04201, doi:10.1029/2003WR002405, w04201, doi:10.1029/2003WR002405.

Nocedal, J., and S. Wright (1999), *Numerical Optimization*, Springer-Verlag, New York.

Passino, K., and S. Yurkovich (1998), *Fuzzy control*, Addison Wesley Longman, Menlo Park, CA.

R Development Core Team (2005), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Ramarao, B., A. Lavenue, G. de Marsily, and M. Marietta (1995), Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields, 1. theory and computational experiments, *Water Resources Research*, *31*(3), 475–493.

Ramey, H. (1982), Well loss function and the skin effect: A review, in *Recent Trends in Hydrogeology*, edited by T. Narasiman, pp. 265–271, GSA, Boulder, CO, GSA Spec. Paper 189.

Rasmussen, T., and L. Crawford (1997), Identifying and removing barometric pressure effects in confined and unconfined aquifers, *Ground Water*, *35*(3), 502–511.

Reda Taha, M., J. Lucero, and T. Ross (2005), Examining the significance of mortar and brick unit properties on masonry bond strength using bayesian model screening, in *10th Canadian Masonry Symposium*, pp. 742–751, Banff, Alberta.

Ritzi, R., Z. Dai, D. Dominic, and Y. Rubin (2004), Spatial correlation of permeability in cross-stratified sediment with hierarchical architecture, *Water Resour. Res.*, *40*(3), W03513, doi:10.1029/2003WR002420.

Ross, S. (1993), *Introduction to Probability Models*, 5 ed., 556 pp., Academic Press, San Diego, California.

Ross, T. (2004), *Fuzzy Logic with Engineering Applications*, second ed., John Wiley and Sons, Ltd., West Sussex, England.

Sanchez-Vila, X., J. Carrera, and J. Girardi (1996), Scale effects in transmissity, *Journal of Hydrology*, *183*, 1–22.

Sanchez-Vila, X., P. Meier, and J. Carrera (1999), Pumping tests in heterogeneous aquifers: An analytical study of what can be obtained from their interpretation using Jacob's method, *Water Resources Research*, *35*(4), 943–952.

Schad, H., and G. Teutsch (1994), Effects of investigation scale on pumping test results in heterogeneous porous media, *J. Hydrol.*, *159*, 61–77.

Shrive, N., and D. Tillemann (1992), A simple apparatus and method for measuring on-site flexural bond strength, in *Proceedings of the 6th Canadian Masonry Symposium*, pp. 283–294, Saskatoon, Canada.

Smith, S. (1980), A learning system based on genetic adaptive systems, Phd dissertation, University of Pittsburgh.

Straface, S., T.-C. J. Yeh, J. Zhu, S. Troisi, and C. Lee (2007), Sequential aquifer tests at a well field, Montalto, Uffugo Scalo, Italy, *Water Resources Research*, *43*, W07432, doi:10.1029/2006WR005287.

Sudicky, E. (2004), On certain stochastic hydrology issues, *Stochastic Environmental Research and Risk Assessment*, *18*, 285, doi:10.1007/s00477-004-0196-2.

Sugo, H. (2000), Strength and microstructural characteristics of brick/mortar bond, Phd dissertation, University of Newcastle, Newcastle, Australia.

Sun, N.-Z. (2005), Structure reduction and robust experimental design for distributed parameter identification, *Inverse Problems*, *21*, 739–758.

Sun, N.-Z., and W.-G. Yeh (1985), Identification of parameter structure in groundwater inverse problem, *Water Resour. Res.*, *21*(6), 869–883.

Theis, C. (1935), The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using ground-water storage, *Eos Trans. AGU*, *16*, 519–524.

Trinchero, P., X. Sánchez-Vila, and D. Feràndez-Garcia (2008), Point-to-point connectivity, an abstract concept or a key issue for risk assessment studies?, *Advances in Water Resources*, *31*, 1742–1753.

Tsai, F.-C., N.-Z. Sun, and W.-G. Yeh (2005), Geophysical parameterization and parameter structure identification using natural neighbors in groundwater inverse problems, *Journal of Hydrology*, *308*, 269–283, doi:10.1016/j.jhydrol.2004.11.004.

Turk, G. (1982), Markov chains: Letters to the editor, *Mathematical Geology*, *14*(5), 539–542.

Venturini, G. (1993), Sia: A supervised inductive algorithm with genetic search for learning attribute based concepts, in *European Conf. on Machine Learning*, pp. 280–296, Vienna.

Vesselinov, V. (2004a), An alternative conceptual model of groundwater flow and transport in saturated zone beneath the pajarito plateau, *Tech. Rep. LA-UR-05-6741*, Los Alamos National Laboratory.

Vesselinov, V. (2004b), Logical framework for development and discrimination of alternative conceptual models of saturated groundwater flow beneath the pajarito plateau, *Tech. Rep. LA-UR-05-6876*, Los Alamos National Laboratory.

Vesselinov, V., S. Neuman, and W. Illman (2001), Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff 2. equivalent parameters, high-resolution stochastic imaging and scale effects, *Water Resources Research*, *37*(12), 3019–3041.

Vistelius, A. (1949), On the question of the mechanism of formation of strata, *Dokl. Akad. Nauk SSSR*, *65*(2), 191–194.

Vrugt, J., and B. Robinson (2007), Improved evolutionary optimization from genetically adaptive multimethod search, *PNAS*, *104*(3), 708–711, doi:10.1073/pnas.0610471104.

Vrugt, J., B. Robinson, and J. Hyman (2008), A universal multimethod search strategy for computationally efficient global optimization, in Press.

Walthall, S., and J. Ingram (1984), The investigation of aquifer parameters using multiple piezometers, *Ground Water*, *22*(1).

Weissmann, G., and G. Fogg (1999), Multi-scale alluvial fan heterogeneity modeled with transition probability geostatistics in a sequence stratigraphic framework, *Journal of Hydrology*, *226*, 48–65.

Wu, C.-M., T.-C. Yeh, J. Zhu, T. Lee, N. Hsu, C.-H. Chen, and A. Sancho (2005), Traditional analysis of aquifer tests: Comparing apples to oranges?, *Water Resources Research*, *41*, W09402, doi:10.1029/2004WR003717.

Ye, M., S. Neuman, and P. Meyer (2004), Maximum likelihood bayesian averaging of spatial variability models in unsaturated fractured tuff, *Water Resour. Res.*, *40*, W05113, doi:10.1029/2003WR002557.

Yeh, T.-C. J., and C.-H. Lee (2007), Time to change the way we collect and analyze data for aquifer characterization, *Ground Water*, *45*(2), technical Commentary.

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

Zhang, C., and W. Li (2008), A comparative study of nonlinear Markov chain models for conditional simulation of multinomial classes from regular samples, *Stochastic Environmental Research and Risk Assessment*, *22*, 217–230, doi:10.1007/s00477-007-0109-2.

Zyvoloski, G. (2007), FEHM: A control volume finite element code for simulating subsurface multi-phase multi-fluid heat and mass transfer, *Tech. Rep. LAUR-07-3359*, Los Alamos National Laboratory.

Zyvoloski, G., B. Robinson, Z. Dash, and L. Trease (1997), Summary of the models and methods for the fehm application—a finite-element heat- and mass-transfer code, *Tech. Rep. LA-13307-MS*, Los Alamos Natl. Lab., Los Alamos, NM.