Title
Model Reduction via Proper Orthogonal Decomposition of Transient Confined and Unconfined Groundwater-Flow

Author
Boyce, Scott Elliott

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Model Reduction via Proper Orthogonal Decomposition of Transient Confined and Unconfined Groundwater-Flow

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Civil Engineering

by

Scott Elliott Boyce

2015

ABSTRACT OF THE DISSERTATION

Model Reduction via Proper Orthogonal Decomposition

of Transient Confined and Unconfined Groundwater-Flow

By

Scott Elliott Boyce

Doctor of Philosophy in Civil Engineering

University of California, Los Angeles, 2015

Professor William W. Yeh, Chair

Understanding groundwater resources is enhanced through the application of mathematical models that simulate the dynamics of an aquifer system. Conducting advanced analyses such as inverse problems for parameter estimation or optimization of pumping schedules under different scenarios requires a large number of simulations. Such analyses are intractable for complex, highly-discretized models with large computational requirements. Reducing the computational burden associated with these simulation models provides the opportunity to perform more advanced analyses on a wider spectrum of groundwater management problems. Projection based model reduction via Proper Orthogonal Decomposition (POD) has been shown to reduce the state space dimension by several orders of magnitude and thus reduces the computational burden. Two new POD techniques have been developed that improve the computation of high dimensional groundwater modeled systems. The first method provides a framework for

developing parameter independent reduced models for solving inverse problems of confined groundwater models. This methodology is validated using synthetic test cases to solve a traditional inverse problem and Bayesian inverse problem. The second method presents a novel technique that allows for model reduction of unconfined groundwater flow, a nonlinear system of equations, using the Newton formulation of MODFLOW. This method extends POD to nonlinear equations and reduces the computational burden of solving the inverse of the Jacobian required by the Newton formulation. Multiple test cases are presented to illustrate how a POD model is constructed and applied to different groundwater models. These two techniques result in several orders of magnitude of reduction in the state dimension and reduce to the total CPU time. For the case of the Bayesian inverse problem, the synthetic example's parameter posterior distributions that are described with the Metropolis-Hastings Markov chain Monte Carlo method results in a time savings of 48 days when using the reduced model.

The dissertation of Scott Elliott Boyce is approved.

Steven A. Margulis

Michael K. Stenstrom

Lieven Vandenberghe

William W. Yeh, Committee Chair

University of California, Los Angeles

2015

DEDICATION

This dissertation is dedicated unto itself.

May it live on forever as a series of structured electrons within the interconnected world computing network.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 2 Figures

Chapter 3 Figures

LIST OF TABLES

Chapter 2 Tables

ACKNOWLEDGMENTS

As a child there was one thing that I would always hear from one grandparent growing up and that was “You should be a doctor, the family needs to have doctor in it. It does not matter what kind, but you should keep focused in school and be a doctor.” Those were Grandmother’s kind words and sometimes not so kind words to me throughout my childhood. I like to first acknowledge Evelyn Siamon for being the earliest believer in my abilities.

I would like to thank my parents, Gordon and Linda Boyce, for their emotional and financial support throughout my entire life. They gave my sister and me a perfect childhood; the emphasis that my family placed on academics and intellectual pursuits has made me who I am today.

I want to thank and acknowledge my sister and her husband, Nicole and Jeff Parker, whom both were more excited than I when I decided to pursue my Ph.D. They also generously donated their time with editing and applying their strong rhetoric skills to ensure that this document was well written.

During the course of my Ph.D. studies I made many close friends and acquaintances. Two close friends stand out and deserve recognition here. Wei-Chen Cheng, who took me under his wing and kept me informed about what to expect at each stage of my Ph.D. Without his guidance, I surely would not have made it to this point in my academic career. Throughout the years Sami Maalouf provided great advice and emotional support during the worst of times during my Ph.D. Sami also assisted me with understanding the requirements of the final dissertation. Both Wei-Chen and Sami are lifelong and dear friends as a result of this Ph.D. and I thank them dearly for all their help through the years.

I am indebted to my advisor and mentor, Professor William W-G. Yeh, whose patience and support during my Ph.D. has made me into the professional hydrologist I am today. He introduced me to many topics of research from hydropower systems, to biostatistics, to optimization theory, to the concepts and research that I discuss in this dissertation. I surely would not have ever completed my Ph.D. had he not guided me through it.

Lastly I like acknowledge my longest known friend, Lawrence Wong. We first met in 1998 during the engineering orientation at the UCSB before our academic careers even started. He and I, while not always able to see each other frequently, have always remained close friends. It always been nice to have someone throughout the years that actually thinks I am smart.

And finally the more formal acknowledgements.

This material is based on work supported by NSF under award EAR-1314422.

Chapter 2 contains the following reprinted journal article:

Boyce, S. E., & Yeh, W. W.G. (2014). Parameter-independent model reduction of transient groundwater flow models: Application to inverse problems. *Advances in Water Resources*, *69*, 168-180, http://dx.doi.org/10.1016/j.advwatres.2014.04.009.

Chapter 3 contains the following reprinted journal article:

Boyce, S. E., Nishikawa, T., & Yeh, W. W.G. (2015). Reduced order modeling of the Newton formulation of MODFLOW to solve unconfined groundwater flow. *Advances in Water Resources, 83,* 250-262, http://dx.doi.org/10.1016/j.advwatres.2015.06.005

and retains the following copyright:

VITA

1998	Salutatorian, Paso Robles High School Paso Robles, California
2000-2002	University of California's Leadership Excellence through Advanced Degrees (UC LEADS) Scholar University of California, Santa Barbara
2001	UC Berkeley Chemical Engineering Plasma Deposition and Glow Discharge Process Research Internship
2001-2002	Reserve Officers' Training Corps (ROTC) Cadet University of California, Santa Barbara
2003	B.S., Chemical Engineering University of California, Santa Barbara
2003-2004	Chemical Engineer CerOx Corp. Santa Maria, California
2004-2007	Environmental Resources Engineering graduate coursework, Humboldt State University GPA: 3.43
2009	M.S. Civil Engineering University of California, Los Angeles GPA: 3.74
2009-2015	Teaching Assistant School of Engineering and Applied Science University of California, Los Angeles
2010-present	Hydrologist, United States Geological Survey: California Water Science Center San Diego, CA
2013-2014	California Water Science Center Research Program Awardee, United States Geological Survey
2015	United States Geological Survey, Office of Groundwater Research award for development of a new groundwater subsidence software package for MODFLOW and MODFLOW-OWHM

SELECTED PUBLICATIONS AND PRESENTATIONS

Boyce, S.E., Nishikawa, T., and Yeh, W. W.G. (2015). Reduced order modeling of the Newton formulation of MODFLOW to solve unconfined groundwater flow. Advances in Water Resources, 83, 250-262. http://dx.doi.org/10.1016/j.advwatres.2015.06.005

Boyce, S.E. and Hanson, R.T. (2015). An Integrated Approach to Conjunctive-Use Analysis with the One-Water Hydrologic Flow Model, MODFLOW-OWHM. MODFLOW and More 2015: Modeling A Complex World - Integrated GroundWater Modeling Center, Golden, Colorado, May 31–June 3, 2015, p. 6-10. http://igwmc.mines.edu/conference/Mod2015/MM15_Proceedings.pdf

Nava, A.P., Villanueva, C.C., Villarreal, F.C., Hanson, R.T., and **Boyce**, S.E. (2015). A New Integrated Hydrologic Model for Mexico Valley, Mexico City, Mexico. MODFLOW and More 2015: Modeling A Complex World - Integrated GroundWater Modeling Center, Golden, Colorado, May 31–June 3, 2015, p. 148. http://igwmc.mines.edu/conference/Mod2015/MM15_Proceedings.pdf

Hanson, R.T., Traum J., **Boyce**, S.E., Schmid, W., Hughes, J.D, W. W. G. (2015), Examples of Deformation-Dependent Flow Simulations of Conjunctive Use with MF-OWHM. Ninth International Symposium on Land Subsidence (NISOLS), Nagoya, Japan.

Boyce, S.E. and Yeh, W. W.G. (2014). Parameter-independent model reduction of transient groundwater flow models: Application to inverse problems. Advances in Water Resources, 69, 168-180. http://dx.doi.org/10.1016/j.advwatres.2014.04.009

Boyce, S.E. and Yeh, W. W.G. (2014), POD Model Reduction: A Method for Reducing the Computational Burden of Solving Systems of Equations from Numerical Models. California Water & Environmental Modeling Forum, Folsom, California.

Hanson, R.T., **Boyce**, S.E., Schmid, W., Hughes, J.D., Mehl, S.M., Leake, S.A., Maddock, Thomas, III, and Niswonger, R.G. (2014). One-Water Hydrologic Flow Model (MODFLOW-OWHM): U.S. Geological Survey TM6-A51, 120 p., http://dx.doi.org/10.3133/tm6A51

Ferguson, I.M., **Boyce**, S.E., Hanson, R.T., and Llewellyn, D. (2014), Fully-Integrated Simulation of Conjunctive Use from Field to Basin Scales: Development of a Surface Water Operations Module for MODFLOW-OWHM. American Geophysical Union Fall Meeting, San Francisco, California.

Boyce, S.E., Nishikawa, T., and Yeh, W. W.G. (2013), Projection-based Model Reduction of Unconfined Groundwater Systems. American Geophysical Union Fall Meeting, San Francisco, California.

Boyce, S.E. and Yeh, W. W.G. (2012), Parameter Independent Model Reduction of a 2-Dimensional Groundwater-Flow Model for Bayesian Inverse Problems. American Geophysical Union Fall Meeting, San Francisco, California.

Chapter 1
Introduction

Water, the elixir of life, is an essential part of humanity's existence. Throughout time civilizations flourished when there were nearby stable and clean water supplies. Until recently water was viewed to be an unlimited resource from both the surface and subsurface. As populations have grown, demand for clean water has increased significantly [1,2]. These increases necessitate a better understanding of groundwater resources which is done through the application of mathematical models that simulate the dynamics of an aquifer system. Complex, highly-discretized groundwater simulation models often have a large computational requirement that prevents advanced analysis, such as parameter estimation and optimization of pump schedules. Reducing the computational burden associated with these simulation models provides the opportunity to perform more advanced analyses on a wider spectrum of groundwater problems. This dissertation presents two new model reduction techniques that significantly lower the computational burden by reducing the state dimensionality of a groundwater simulation model.

This chapter introduces key concepts related to this dissertation and its new contributions to the scientific field. This chapter is subdivided into seven sections. Section 1.1 covers basic of concepts of groundwater. Section 1.2 extends these concepts into how saturated groundwater flow is modeled. Section 1.3 covers the basics of parameter estimation with a focus on hydraulic conductivity. Section 1.4 discusses Bayesian inverse problems. Section 1.5 provides the Bayesian mathematical development of the Metropolis-Hastings Markov Chain Monte Carlo method for solving the posterior distribution of hydraulic conductivity and a simulation model’s

error. Section 1.6 provides an overview of the groundwater simulation software used in this dissertation. Section 1.7 introduces the concept of Model Reduction and its basic implementation.

This dissertation itself is broken into five total chapters. The first being the chapter you are reading now, the Introduction. Chapter 2 is a reprint of an *Advances in Water Resources* journal paper, Boyce and Yeh [3], that describes a novel algorithm on how to construct a parameter independent reduced model. The parameter independent reduced models are then evaluated with both the deterministic and Bayesian inverse problems. Chapter 3 is a reprint of another *Advances in Water Resources* journal paper, Boyce et.al. [4], that derives a new method for applying model reduction to the nonlinear problem of unconfined groundwater flow. The author of this dissertation was the main contributor and author of both reprinted articles. Chapter 4 concludes the dissertation with final remarks on the two reprinted journal papers. Finally, the dissertation closes on Chapter 5 with suggestions for future research and ideas.

1.1 Groundwater

Beneath the Earth's surface the interstitial spaces within rocks and soils have the capability of storing water. If this space contains water and is capable of transport, it is called an aquifer or groundwater system. Aquifers serve as a stable source of potable water for agricultural, industrial, and domestic use. Aquifer flow occurs from either hydraulic pressure differences from within or gravitational pull to the sea. An undisturbed aquifer system reaches a hydraulic pressure equilibrium called steady state, when no (or a constant) external forcing is applied to it. The most common human external forcing is the result of an extraction pumping well [5].

Understanding groundwater resources is enhanced through the application of mathematical models that simulate the dynamics of an aquifer system. Groundwater flow is modeled in one,

two or three dimensions. In 1856, Henry Darcy investigated the flow of water in vertical homogeneous sand filters in connection with the fountains of the city of Dijon. Darcy stated that the specific discharge, *q*, is equal to the hydraulic conductivity, *K*, times gradient of the hydraulic head, *h*. This relationship, $q = -K\nabla h$, is referred to as Darcy's Law [5,6]. The parameter hydraulic conductivity describes the nature of the porous medium that makes up the groundwater system. Hydraulic conductivity is a function of the mean grain diameter, a shape constant of the soils called specific permeability, dynamic viscosity, and specific gravity [8].

1.2 Groundwater Modeling

Combining Darcy's Law with the continuity equation derives the confined, saturated groundwater flow. The following is the governing equation for confined groundwater flow:

$$\nabla \cdot \left(K \nabla h \right) \pm Q = S_s \frac{\partial h}{\partial t} \tag{1.1}$$

where $\nabla \cdot$ is the divergence operator, ∇ is the gradient operator, K is the hydraulic conductivity tensor [L/T], h is the hydraulic head [L], Q is a volumetric flux per unit volume in or out of the system [T^{-1}], S_S is the specific storage [L^{-1}], t is the time [T]. If there is no confining layer above the groundwater flow the flow is referred to as unconfined, water table, or phreatic groundwater flow. This upper free surface boundary changes the confined groundwater equation to following nonlinear, saturated unconfined groundwater flow equation:

$$\nabla \cdot \left(Kh \nabla h \right) \pm Q = S_y \frac{\partial h}{\partial t} \tag{1.2}$$

where Q is a volumetric flux per unit area in or out of the system [L/T] and S_y is the specific yield [-].

Conducting advanced analyses for groundwater systems with Equation 1.1 and 1.2 requires a large number of simulations. Examples include optimization of pumping schedules under different scenarios or evaluating climate projection influence over water resources. Such analyses are intractable for complex, highly-discretized in time and space models with large computational requirements.

The confined (Eqn. 1.1) and unconfined groundwater (Eqn. 1.2) flow equations are typically discretized then solved through Finite Differences or Finite Elements. Embedded in this discretization are all the parameters, forcing, and boundary conditions that describe the groundwater system. To simplify the groundwater system, parameters are lumped into zonation patterns from which regions of the discretization have common parameter values. Choices in the parameter zonation design depend on scientific interpretation of the geologic characteristics of the aquifer [7]. For hydraulic conductivity the zonation can be as large as a single zone encompassing the entire model discretization or as refined as a zone for each individual finite difference block or finite element [5].

An issue with the representation hydraulic conductivity within a groundwater simulation model is that it cannot be directly measured. Soil samples can directly measure hydraulic conductivity, but the results rarely are applicable to a discretized groundwater model due to the spatial lumping of the parameter. In addition, hydraulic conductive stochastically varies within the subsurface. This stochastic variation can be described by the log-normal probability distribution [9] or gamma/log-gamma distribution [10]. Due to the limitation of not being able to directly measure hydraulic conductivity, it is inferred by solving inverse problems such as a pump test or optimization techniques called parameter estimation. If hydraulic conductivity is treated as stochastic, then it is solved with a Bayesian inverse problem.

1.3 Overview of Parameter Estimation and Deterministic Inverse Problems

Parameter estimation calibrates a groundwater simulation model's hydraulic conductivity to a set of field observations through an inverse problem. Yeh [11] does a comprehensive review of the various strategies used in solving an inverse problem. Common strategies involve an optimization approach that finds the best fit to observation data by choice of hydraulic conductivity. This approach henceforth will be referred as the deterministic inverse problem.

A common objective for the deterministic groundwater inverse problem is identifying a vector of hydraulic conductivity values that minimizes the Sum of Squared Errors (SSE) between a set of observed hydraulic head data and the solution to Equation 2 [11]. This problem is mathematically formulated as:

$$\arg\min_{K \in \mathcal{K}} \sum_{j=1}^{N_t} \sum_{i=1}^{N_l} \left(D_{i,j} - h_{\mathrm{op}(i),j} \right)^2 \tag{1.3}$$

where the decision variable, K, is the set of zoned hydraulic conductivity values used to solve Equation 2; $\mathcal{K}$ is the search space of K, which is usually a continuous interval from a $K_{\min}$ to $K_{\max}$; i and j are the observation location and time indices, respectively; N_l and N_t are the total number of observation locations and times, respectively; $D_{i,j}$ and $h_{\mathrm{op}(i),j}$ are the observed and model simulated hydraulic head; and $\mathrm{op}(i)$ maps the location index to the model nodal index.

The inverse problem is usually solved with a gradient-based search method. Gradient-based search methods rely on derivatives to determine a search direction and reach optimization when the derivative is zero. A problem with gradient-based search methods is that the inverse problem is non-unique because observation errors, model structure errors and the difficulty in

calculating derivatives for search directions. Another issue is that there may exist multiple optimal solutions, called local solutions, and one global. These issues make a gradient search method highly dependent on the initial guess [5].

Genetic Algorithms (GAs) are a set of heuristic methods that follow along the concepts of survival of the fittest to solve optimization problems that are capable of overcoming some of the gradient-based search method's shortcomings. Genetic Algorithms apply a random sample of decision variables, called chromosomes, to evaluate the objective, called the fitness function, to determine which decision variables influence the resampling, the next generation, of the decision variables. The resampling can undergo a series of emulated symbolic biological features such as chromosome cross over or mutation of values. The advantages of GAs are that they are insensitive to the initial guess of the decision variable and do not require the computation of derivatives [12,13]. They are also capably of exploring the entire solution space to identify the global optimum solution.

One problem with applying GAs to an inverse problem is the number of groundwater simulation model calls required to have a sufficiently large chromosome population, and the number of successive generations needed to have the population converge to a solution. This requires repeatedly running the discretized simulation model until the GA finds the global optimal solution. Improvement on a single model simulation time of the model can drastically improve upon the total time for GA to find the optimal solution.

Despite this problem GA is well suited for overcoming the pitfalls of gradient-based searched methods, and it has been applied to many optimization problems involving groundwater. An example of applying GA to solving the inverse problem of a pump test for a

confined aquifer is discussed in Samuel and Jha [14]. GA was also used by Tsai [7] to determine an optimal zonation structure of an aquifer that is subdivided by a Voronoi tesselation.

1.4 Bayesian Inverse Problems

A disadvantage of the deterministic inverse problem is its inability to quantify the uncertainty of the optimal parameters or derive a posterior distribution of them. By approaching the inverse problem from a statistical point of view, the resulting solution is a posterior probability distribution of the parameters conditioned on the model structure and observations [15]. A Bayesian inverse problem is a stochastic method that incorporates prior information to construct a posterior probability distribution of the parameter. Yeh [11], Beck [16] and Freeze et al. [15] cover comprehensive reviews that include uncertainty quantification in simulation models.

Markov Chain Monte Carlo (MCMC) are a set of methods for solving a Bayesian inverse problem. The MCMC simulation uses a proposal distribution to generate a sample of parameter values from which Markov chain is constructed. As samples are drawn from the proposal, the Markov chain converges to a stationary distribution, which is the hydraulic conductivity's posterior probability distribution [17,18]. One algorithm for constructing the MCMC chain is the Metropolis-Hastings (M-H) algorithm [19,20]. In the case of groundwater this would be the hydraulic conductivity's posterior probability distribution. This distribution can be used to quantify the uncertainty and assess its effects on model predictions [21].

The Bayesian inverse problem through MCMC requires many repeated model simulations to characterize an unknown parameter’s posterior probability distribution. It is computationally infeasible to solve a Bayesian inverse problem of a discretized groundwater flow model with a high dimension state space. Model reduction has been shown to reduce the

dimension of a groundwater model by several orders of magnitude, which makes it well suited for Bayesian inverse problems.

1.5 Markov Chain Monte Carlo Development for Hydraulic Conductivity

Bayesian methods link prior knowledge of a model's parameters to observed information and the models output (Eqn. 1.4). This linkage produces a posterior distribution of the model parameters given the observed data. When there is not an analytical solution available, the posterior distribution can be derived from a Markov Chain Monte Carlo (MCMC) simulation that explores Bayes' Theorem without the proportionality constant as presented in Equation 1.5 [18].

$$P\left(A\middle|B\right)=\frac{P\left(B\middle|A\right)\Pi\left(A\right)}{P(B)}=\frac{P\left(B\middle|A\right)\Pi\left(A\right)}{\int P\left(B\middle|A\right)\Pi\left(A\right)\mathrm{dA}} \tag{1.4}$$

$$P\left(A\middle|B\right)\propto P\left(B\middle|A\right)\Pi\left(A\right) \tag{1.5}$$

For a groundwater model, observed data is the hydraulic head at observation or pumping well location. This observed data, $D(x,t)$, is assumed to be related to the model results, $M(x,t;K)$, plus a model error term, $\varepsilon(x,t)$. The model error term is assumed to be normally distributed (Eqn. 1.7) with zero mean and an unknown random variance σ_{ϵ}^{2}. Since σ_{ϵ}^{2} is unknown, its posterior distribution has to be part of the MCMC simulation and will be referred to as the likelihood variance. This is done by using the normal distribution's variance conjugate prior, the inverse-gamma distribution (Eqn. 1.8), to derive a posterior distribution [22,23],

$$D(x,t) = M(x,t;K) + \varepsilon(x,t) \tag{1.6}$$

$$D(x,t) - M(x,t;K) = \varepsilon(x,t) \sim \mathrm{N}\left(0, \sigma_{\epsilon}^2\right) \tag{1.7}$$

$$\Pi\left(\sigma_{\varepsilon}^2\right) \sim \mathrm{IG}(\alpha,\beta) \propto \left(\frac{1}{\sigma_{\varepsilon}^2}\right)^{\alpha-1} \exp\left[-\left(\frac{1}{\sigma_{\varepsilon}^2}\right)\beta\right] \tag{1.8}$$

The posterior distribution of the hydraulic conductivity can be obtained through application of Bayes' Theorem (Eqn. 1.9). Since the value of P(D) is a proportionality factor and is difficult to analytically derive, it is removed from the equation. The remaining portion of Equation 1.9 is composed of a likelihood function and a prior distribution. Equation 1.10, is the likelihood of the collected data given the model parameters are chosen correctly. The prior distribution for the hydraulic conductivity (Eqn. 1.11) is assumed to be log-normally distributed [9]. The hydraulic conductivity's mean and variance are chosen based on the prior knowledge of the hydraulic conductivity.

$$P\left(K \middle| D, \sigma_{\varepsilon}^2\right) = \frac{P\left(D \middle| K, \sigma_{\varepsilon}^2\right)\Pi(K)}{P(D)} \propto P\left(D \middle| K, \sigma_{\varepsilon}^2\right)\Pi(K) \tag{1.9}$$

$$P\left(D \middle| K, \sigma_{\varepsilon}^2\right) = \left(2\pi\sigma_{\varepsilon}^2\right)^{-N/2} \exp\left[-\frac{1}{2\sigma_{\varepsilon}^2}\sum_{i=1}^{N}\left(D(x,t) - M(x,t;K)\right)^2\right] \tag{1.10}$$

$$\Pi(K) = \frac{1}{K\sqrt{2\pi\sigma_{\pi}^2}} \exp\left[-\frac{\left(\ln K - \mu_{\pi}\right)^2}{2\sigma_{\pi}^2}\right] \tag{1.11}$$

The Metropolis-Hastings method is applied to Equation 1.9 to solve for the posterior distribution of the hydraulic conductivity. The MCMC chain is generated by creating a series of proposed values based on an initial guess. A proposed value, K^*, is generated from a proposal density that is conditional on the current value in the Markov chain. The proposal (Eqn. 1.12) is a

lognormal distribution with a mode equal to the current value in the chain, *K*, and variance, σ_p^2, set to a user specified value. To delineate the proposal density function from the other distributions the symbol Pr( ) is used. The variance of the proposal density function affects the rate of convergence of the final Markov chain and is chosen by doing several short simulations to see how the Markov chain responds. When a candidate is generated from Equation 1.12, its acceptance probability, $\hat{\alpha}$, is calculated by taking the ratio presented in Equation 1.13.

$$\Pr\left(K^* \middle| K\right) = \mathrm{P}\left(K \to K^*\right) = \frac{1}{K^*\sqrt{2\pi\sigma_p^2}} \exp\left[-\frac{\left(\ln(K^*) - \ln(K) - \sigma_p^2\right)^2}{2\sigma_p^2}\right] \tag{1.12}$$

$$\hat{\alpha} = \min\left\{ \frac{P\left(D\middle|\underline{K}^*, \sigma_\varepsilon^2\right)}{P\left(D\middle|\underline{K}, \sigma_\varepsilon^2\right)} \prod_{i=1}^{\mathrm{N}}\left[\frac{\Pi\left(K_i^*\right)}{\Pi\left(K_i\right)}\right] \prod_{i=1}^{N} \frac{\Pr\left(K_i\middle|K_i^*\right)}{\Pr\left(K_i^*\middle|K_i\right)},\ 1 \right\} \tag{1.13}$$

Applying the given likelihood (Eqn. 1.10) and prior for the hydraulic conductivity (Eqn. 1.11) to the natural log of the acceptance probability yields the Equation 1.14. If the prior's

$$\begin{aligned} \ln\hat{\alpha} = & \frac{1}{2\sigma_\varepsilon^2}\sum_{i=1}^{N}\left(D(x,t) - M(x,t;K)\right)^2 \\ & - \frac{1}{2\sigma_\varepsilon^2}\sum_{i=1}^{N}\left(D(x,t) - M(x,t;K^*)\right)^2 \\ & + \frac{1}{2\sigma_\pi^2}\sum_{i=1}^{N}\left[\left(\ln K_i - \mu_\pi\right)^2 - \left(\ln K_i^* - \mu_\pi\right)^2\right] \end{aligned} \tag{1.14}$$

variance is not known, then it is set to an arbitrarily large value to lower its influence on the posterior. The natural log is taken to simplify the equations for faster computation on a computer. This natural log of the acceptance probability is then compared to the natural log of a random uniform (zero to one) draw and if it is larger, the value of the hydraulic conductivity is

accepted as a new value in the chain [18]. The formal algorithm for the Metropolis-Hastings MCMC is as follows:

$$K^t := \left(K_1^t, \ \ldots \ , K_{\mathrm{N}}^t\right)$$

Given $K^0 := \left(K_1^0, \ \ldots \ , K_{\mathrm{N}}^0\right)$, $K^{t-1} := K$, and $K^t := K^*$; iterate algorithm for $t = 1, \ 2, \ \ldots$

0. Establish K^0, set $t = 1$

1. Generate a Candidate from a proposal distribution $K^t \sim \Pr\left(K^t | K^{t-1}\right)$

2. Generate from the Uniform distribution $\alpha \sim \mathrm{U}[0,1]$

3. Using the target density solve:

$$\hat{\alpha} = \min\left\{ \frac{P\left(D\middle|\underline{K}^t, \sigma_\varepsilon^2\right)}{P\left(D\middle|\underline{K}^{t-1}, \sigma_\varepsilon^2\right)} \prod_{i=1}^{\mathrm{N}}\left[\frac{\Pi\left(K_i^t\right)}{\Pi\left(K_i^{t-1}\right)}\right] \prod_{i=1}^{\mathrm{N}} \frac{\Pr\left(K_i^{t-1}\middle|K_i^t\right)}{\Pr\left(K_i^t\middle|K_i^{t-1}\right)}, \ 1 \right\}$$

4. If $\alpha < \hat{\alpha}$ then keep K^t else set $K^t = K^{t-1}$.

5. Set $t = t + 1$ and go to Step 1.

The posterior distribution of σ_ϵ^2, Equation 1.15, is a combination of a likelihood function and its prior distribution. Since its prior is conjugate, the posterior is an inverse gamma represented in Equation 1.16. Typical values of α, β for the prior in Equation 1.11 are equal to one or less. For this model a value of one was used for both α and β. These values have little impact on the posterior because the sum of squared errors tends to be very large ($\sum_{i=1}^{N}\left(D(x,t) - M(x,t;K)\right)^2 \gg \beta$) and the observed data sets are significantly larger than one ($\mathrm{N} \gg \alpha$).

$$P\left(\sigma_\varepsilon^2 \middle| K, D\right) \propto P\left(D \middle| K, \sigma_\varepsilon^2\right)\Pi\left(\sigma_\varepsilon^2\right) \tag{1.15}$$

$$\begin{aligned} P\left(\sigma_\varepsilon^2 \middle| K, D\right) &\propto \frac{1}{\left(\sigma_\varepsilon^2\right)^{N/2}} \exp\left[-\frac{\sum_{i=1}^{N}\left(D(x,t)-M(x,t;K)\right)^2}{2\sigma_\varepsilon^2}\right]\left(\frac{1}{\sigma_\varepsilon^2}\right)^{\alpha-1}\exp\left[-\left(\frac{1}{\sigma_\varepsilon^2}\right)\beta\right] \\ &\propto \left(\frac{1}{\sigma_\varepsilon^2}\right)^{N/2+\alpha-1}\exp\left[-\frac{1}{\sigma_\varepsilon^2}\left(\frac{\sum_{i=1}^{N}\left(D(x,t)-M(x,t;K)\right)^2}{2}+\beta\right)\right] \end{aligned} \tag{1.16}$$

$$\sim \mathrm{IG}\left(N/2+\alpha, \frac{\sum_{i=1}^{N}\left(D(x,t)-M(x,t;K)\right)^2}{2}+\beta\right)$$

Solving for the two posterior distributions described in Equation 1.9 and 1.15 involves a twostep process. The first step is to use the Metropolis-Hastings algorithm to get a hydraulic conductivity sample from Equation 1.9 for a given likelihood variance, σ_ε^2. This sample will either be the previous value in the Markov chain or a new accepted value. This new sample for the hydraulic conductivity, *K*, is then used in the second step to sample for the likelihood variance from Equation 1.16. With the new value for the likelihood variance, the method proceeds with sampling the next hydraulic conductivity sample. This process continues until the two Markov chains reach stationarity.

1.6 Groundwater Simulation Software

Two different groundwater flow simulation softwares are used in this dissertation. Their source code was modified to apply model reduction to the groundwater flow equations (Eqn. 1.1 and 1.2). This was necessary to validate and provide a real word test case for the new methods described in Chapters 2 and 3. All modifications were thoroughly tested to ensure the highest accuracy of the code in representing the developed mathematical concepts.

The first model, which is used in Chapter 2, is Saturated 2-Dimensional Flow (SAT2D) [24]. SAT2D is a finite element-based software that simulates two-dimensional, saturated groundwater flow. It applies finite elements to only Equation 1.1 and then solves the resulting system of linear equations with a preconditioned conjugate gradient (PCG) solver. The software uses a generic file input system that specifies the finite element discretization scheme and properties of the aquifer system. This allows the same software to be reused for multiple modeling exercises. The model reduction is applied by intercepting the system matrices produced by SAT2D before they are solved by PCG and instead uses LAPACK's LU decomposition solver [25].

The second model, which is used in Chapter 3, is the most recent release of the MODFLOW family. This version of MODFLOW is called the One-Water Hydrologic Flow Model, MODFLOW-OWHM [26,27]. MODFLOW-OWHM uses a three-dimensional finite difference scheme to solve both the unconfined and confined groundwater flow equations (Eqn. 1.1 and 1.2). Its design allows for addition features to be easily added that are called packages and processes. Due to this design there are numerous options and features that make it flexible for modeling different water resource scenarios and evaluating sustainable groundwater usage. Boyce and Hanson [27] state that MODFLOW-OWHM represents a complete integrated hydrologic model that fully links the movement and use of groundwater, surface water, and imported water for consumption by agriculture and natural vegetation on the landscape, and for potable water. This version of MODFLOW is unique compared to previous releases because it retains and keeps track of all the water during simulation of the hydrosphere, MODFLOW-OWHM thus accounts for "all of the water everywhere and all of the time".

1.7 Model Reduction via Proper Orthogonal Decomposition and the Galerkin Projection

Projection-based model reduction proceeds by projecting a discretized groundwater-flow model, $h \in \mathbb{R}^{n}$, onto a reduced model subspace of $h_r \in \mathbb{R}^{r}$, where $r \ll n$. This process is done through the substitution of h_r to replace h and premultiplying the groundwater system matrices by the transpose of the Galerkin projection operator, which is composed of orthonormal basis vectors. Proper Orthogonal Decomposition (POD) is a mathematical operation that transforms a basis to an orthonormal basis and ranks each vector in terms of quality of information. In the literature the quality of information is referred to as percent variance or percent energy [31,4].

Model reduction proceeds by assuming that the hydraulic head, h, can be mapped linearly onto the reduced space, h_r, by some orthonormal matrix, $P \in \mathbb{R}^{n \times r}$, which contains information about the discretized groundwater-flow model's solutions in time [28,3 ,4]:

$$h = Ph_r \tag{1.17}$$

The vectors in the P matrix are sometimes referred to as the POD basis, spatial basis functions, empirical orthogonal functions, or principal vectors [4,29,30,31].

The matrix P is derived from taking multiple sets, k, of model solutions at different times, h^{t_k}. These solutions are called “snapshots” [4] and are collected from the groundwater-flow model under a constant/reference forcing (e.g. pumping well set at a specific rate). These snapshots provide the necessary information about how the groundwater model responds to external forces enabling the reduce model to replicate the response in a lower dimensional space at any forcing value (e.g. pumping well at different rates in time).

A set of snapshots must be taken for each, time invariant, linearly independent forcing term in the groundwater-flow model. Time invariant forcings do not change during the course of the simulation, such as a constant boundary condition, and are included along with all the time variant snaphots. For example, if a model is surrounded by my a constant head, Dirichlet, boundary condition that never changes the head value and has two extraction wells that have their rates change throughout a simulation and there is no linear relation between them, then two sets of snapshots must be taken. The first set would include the constant head bound and use one of the wells set at a constant rate (the other being set to a rate of zero). The value of the rate does not matter, but does affect the optimal choice of snapshot times. The second set of snapshots would then include the same constant head boundary and the second well set to a constant rate. If there are any constant in time linear relationships between forcings (e.g. $\text{Well A} = 2 \times \text{Well B} \ \ \forall t$), then a single set of snapshots can be taken at that relationship. If the linear dependence is not known or there is any potential for linear independence, then separate snapshots should be taken for each forcing as any linear dependence will later be removed mathematically from the snapshot set. All the sets of snapshots are then combined to form the final snapshot set:

$$\begin{aligned} \mathrm{S}^{\mathrm{j}} &= \left[h^{t_1},\ h^{t_2},\ h^{t_3},\ \ldots,\ h^{t_k} \right] \quad \forall \mathrm{j} = 1,\ 2,\ \ldots,\ \mathrm{NSET} \\ \mathrm{S} &= \left[\mathrm{S}^{1},\ \mathrm{S}^{2},\ \ldots,\ \mathrm{S}^{\mathrm{NSET}} \right] \in \mathbb{R}^{\mathrm{n} \times \mathrm{m}} \end{aligned} \tag{1.18}$$

where S^{j} is a set of snapshots collected for the j^{th} linearly independent forcing and is composed of k snapshots, NSET is the total number of linearly independent forcings, and S is the final snapshot set containing m snapshots. Note that the length of each vector in S is the same as the model discretization and that row 1 of S should correspond with node 1 of the model (i.e. first row of *h*).

For each snapshot set, S^j, the number of snapshots, k, and their specific solution time, t_k, are problem-dependent. Both the number of snapshots and specific solution times do not have to be the same for each snapshot set, S^j. This is a limitation of model reduction in that the quality of the reduce model is highly dependent on the snapshot selection. This both includes the number of snapshots, the simulated times that they are taken at, and, for nonlinear models, the value of the forcing itself. Snapshot selection is an area of research beyond the scope of this dissertation, which is focused on the model reduction itself. The method discussed in Boyce and Yeh [3] does include a systematic approach to selecting the optimal number of snapshots and their corresponding times. This is an area of research that needs further investigation.

From empirical tests in literature [28,3,4,31] snapshot set, S^j, sizes should range between 10 to 50 snapshots that are exponentially distributed in time, t_k, from slightly beyond the initial condition to quasi-steady state for one constant forcing value. If the model is highly nonlinear, then more snapshots maybe necessary with a higher density at simulation times and responses where the reduce model performs unsatisfactorily. Boyce et.al. [4] discovered that nonlinear responses were better replicated by the reduced model by taking an additional 10 to 50 exponentially distributed in time snapshots that capture the aquifer recovery from pumping (compared to only taking snapshots with pumping). For this case it is important to pick a pumping rate that dewaters the aquifer to its deepest observed drawdown point, collecting snapshots at that rate, and then collect the additional snapshots of the aquifer recovery from pumping. If this fails to improve the reduced model, then additional snapshot sets should be taken at different pumping rates at the same location.

The snapshots from Equation 1.18 form a basis that, when made orthonormal, derives the Galerkin projection operator [3,4,28]. This operator is the same matrix as the linear map, P, defined in Equation 1.17. The orthonormalization of the snapshots, S, is performed by Proper Orthogonal Decomposition (POD), which, for the discrete case, is the Singular Value Decomposition (SVD) [4,32,33]:

$$\mathrm{S} = \mathrm{U}\Sigma\mathrm{V}^{\mathrm{T}} = \sum_{i=1}^{\mathrm{rank(S)}} \sigma_i u_i v_i^{\mathrm{T}} \quad \Rightarrow \quad \mathrm{P} = \mathrm{U} \tag{1.19}$$

where $\mathrm{rank}(\mathrm{S})$ is the minimum between the number of linearly independent columns or rows in S, σ_i are the singular values that are ordered from largest to smallest, Σ is a diagonal matrix with singular values along the diagonal, $\mathrm{U} = [u_1, u_2, \ldots]$ are the left singular vectors that comprise the POD basis set equal to P, and $\mathrm{V} = [v_1, v_2, \ldots]$ are the right singular vectors that represent the proper orthogonal modes.

There are two other methods for deriving the Galerkin projection operator by transforming the snapshot sets, S, to an orthonormal basis. The first is a direct orthonormalization of the snapshots through the Gram-Schmidt method or QR decomposition. These methods produce an orthonormal basis, but do not provide any information about the quality of each basis vector (viz. no singular values). The second exploits the fact that Eigenvalue decomposition (EVD) is identical to SVD for symmetric matrices. The snapshot set is then made symmetric and less computationally demanding by solving for the EVD of $\mathrm{S}^{\mathrm{T}}\mathrm{S}$. A transformation is then applied to get the orthonormal basis that represents the range space of S. These two methods of building the projection matrix have been used extensively in literature [34,35, 28, 31]. The following is how EVD is applied and then transformed to obtain the Galerkin projection operator.

$$S^T S = V \Sigma V^T$$
$$P = S V \Sigma^{-1/2} \tag{1.20}$$

A note of caution is that Equation 1.20 is considered numerically unstable when solved on a computer due to accumulation of round off errors from the matrix multiplication, $S^T S$. There are extensive examples and literatures as to why this should not be done [36,37,38,33]. Gram-Schmidt is considered numerically unstable as well. There is a Modified Gram-Schmidt that improves the numerical stability, but only when solving a linear least squares problem and not for constructing a basis and it fails if there is any linear dependence amongst the snapshots [39,40,41].

Without knowledge of the optimal choice for snapshots, the singular values can serve as weights to distinguish the left singular vectors, subsequently referred to as POD basis vectors, that contribute the most amount of information to the projection operator P. This is accomplished by solving for r in the following relationship:

$$\min \; r$$
$$\text{Subject to: } \frac{\sum_{i=1}^{r} \sigma_i}{\sum_{i=1}^{\text{rank(S)}} \sigma_i} \times 100\% \geq \text{PE} \tag{1.21}$$

where PE is a user-specified minimum required percent energy contribution of the sum of the r largest singular values. Typical values for PE range from 99% to 99.99%, where smaller values lead to fewer basis vectors at a cost of less accurate reduced model results. If Equation 1.20 is used to construct the Galerkin projection operator, then Equation 1.21 is modified to use Eigenvalues instead of the singular values. The user specified number, PE, is then referred as the percent variance due to the similarity of $S^T S$ to a covariance matrix [35].

Once r has been established, the final truncated POD basis is $\mathrm{P} = \left[u_1, u_2, \ \ldots, u_r\right]$, which defines the dimension of the reduced model, h_r (Eqn. 1.17). This process of removing insignificant basis vectors is called singular-value truncation; if S is symmetric, then it is also called principal component analysis.

Model reduction begins by constructing the orthonormal matrix, P (full snapshots or truncated) with POD. Once P is established the reduced dimension state variable, h_r, is substituted into the groundwater system matrices with Equation 1.17 to remove the full dimension variable h. After this the Galerkin projection is applied by premultiplying the groundwater system matrices by P^{T} making them dimensionally consistent. The vector h_r is then solved for in the place of h. This reduces the number of equations from n equations to r, drastically simplifying the calculation. To obtain h from h_r the same linear map defined in Equation 1.17 is used. This then solves for $h \in \mathbb{R}^{\mathrm{n}}$ with r equations instead of n.

1.8 References

[1] Fetter, C. W. (2001): *Applied hydrogeology*. 4th ed. Upper Saddle River, N.J: Prentice Hall.

[2] Loucks, D. P., Van Beek, E., Stedinger, J. R., Dijkman, J. P., & Villars, M. T. (2005). *Water resources systems planning and management: an introduction to methods, models and applications*. Paris: UNESCO.

[3] Boyce, S. E., & Yeh, W. W.G. (2014). Parameter-independent model reduction of transient groundwater flow models: Application to inverse problems. *Advances in Water Resources, 69,* 168-180, http://dx.doi.org/10.1016/j.advwatres.2014.04.009.

[4] Boyce, S. E., Nishikawa, T., & Yeh, W. W.G. (2015). Reduced order modeling of the Newton formulation of MODFLOW to solve unconfined groundwater flow. *Advances in Water Resources*, *83*, 250-262, http://dx.doi.org/10.1016/j.advwatres.2015.06.005

[5] Willis, R. & Yeh, W. W.G. (1987). *Groundwater systems planning and management*. Englewood Cliffs, NJ: Prentice-Hall.

[6] Bear, J. (1988): *Dynamics of fluids in porous media*. New York, N.Y: Dover Publications.

[7] Tsai, F. T. C., & Yeh, W. W. G. (2004). Characterization and identification of aquifer heterogeneity with generalized parameterization and Bayesian estimation. *Water Resources Research*, *40*(10).

[8] Viessman, W. & Lewis, G.L. (2003): *Introduction to hydrology*. 5th ed. Upper Saddle River, NJ: Prentice Hall.

[9] Freeze, R. A. (1975). A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media. *Water Resources Research*, *11*(5), 725-741.

[10] Loáiciga, H. A., Yeh, W. W. G., & Ortega-Guerrero, M. A. (2006). Probability density functions in the analysis of hydraulic conductivity data. *Journal of Hydrologic Engineering*, *11*(5), 442-450.

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

[12] Davis, L. (1991), *Handbook of genetic algorithms*, XII, 385, Van Nostrand Reinhold, New York.

[13] Goldberg, D. E. (1989), *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley Publishing Company, Reading, Mass.

[14] Samuel, M. P., & Jha, M. K. (2003). Estimation of aquifer parameters from pumping test data by genetic algorithm optimization technique. *Journal of irrigation and drainage engineering*, *129*(5), 348-359.

[15] Freeze, R. A., Massmann, J., Smith, L., Sperling, T., & James, B. (1990). Hydrogeological decision analysis: 1. A framework. *Groundwater*, *28*(5), 738-766.

[16] Beck, M. B. (1987). Water quality modeling: a review of the analysis of uncertainty. *Water Resources Research*, *23*(8), 1393-1442.

[17] Geyer, C. J. (1992). Practical Markov chain monte carlo. *Statistical Science*, 473-483.

[18] Robert, C., & Casella, G. (2013). *Monte Carlo statistical methods*. Springer Science & Business Media.

[19] Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*(1), 97-109.

[20] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, *21*(6), 1087-1092.

[21] Hassan, A. E., Bekhit, H. M., & Chapman, J. B. (2009). Using Markov Chain Monte Carlo to quantify parameter uncertainty and its effect on predictions of a groundwater flow model. *Environmental Modelling & Software*, *24*(6), 749-763.

[22] Daniels, M. J. (1999). A prior for the variance in hierarchical models. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*, 567-578.

[23] Lynch, S. M. (2007). *Introduction to applied Bayesian statistics and estimation for social scientists*. Springer Science & Business Media.

[24] Gambolati, G., M. Putti, and C. Paniconi (1999). Three-dimensional model of coupled density-dependent flow and miscible salt transport in groundwater, in *Seawater Intrusion in Coastal Aquifers: Concepts, Methods, and Practices*, edited by J. Bear et al., pp. 315-362, Kluwer Acad., Dordrecht, Netherlands.

[25] E. Anderson , Z. Bai , C. Bischof , L. S. Blackford , J. Demmel , Jack J. Dongarra , J. Du Croz , S. Hammarling , A. Greenbaum , A. McKenney , D. Sorensen (1999). LAPACK Users' guide (third ed.), xxi, 404, Society for Industrial and Applied Mathematics, Philadelphia. http://dx.doi.org/10.1137/1.9780898719604

[26] Hanson, R.T., Boyce, S.E., Schmid, Wolfgang, Hughes, J.D., Mehl, S.M., Leake, S.A., Maddock, Thomas, III, and Niswonger, R.G. (2014), One-Water Hydrologic Flow Model (MODFLOW-OWHM): *U.S. Geological Survey Techniques and Methods*, 6–A51, 120 p., http://dx.doi.org/10.3133/tm6A51.

[27] Boyce**,** S.E. & Hanson, R.T. (2015). An Integrated Approach to Conjunctive-Use Analysis with the One-Water Hydrologic Flow Model, MODFLOW-OWHM. *MODFLOW and More,* Integrated GroundWater Modeling Center, Golden, Colorado, p. 6-10. http://igwmc.mines.edu/conference/Mod2015/MM15_Proceedings.pdf

[28] Vermeulen, P. T. M., Heemink, A. W., & Te Stroet, C. B. M. (2004). Low-dimensional modelling of numerical groundwater flow. *Hydrological processes*, *18*(8), 1487-1504.

[29] Lang, Y. D., Malacina, A., Biegler, L. T., Munteanu, S., Madsen, J. I., & Zitney, S. E. (2009). Reduced Order Model Based on Principal Component Analysis for Process Simulation and Optimization†. *Energy & Fuels*, *23*(3), 1695-1706.

[30] Park, H. M., & Cho, D. H. (1996). Low dimensional modeling of flow reactors. *International Journal of Heat and Mass Transfer*, *39*(16), 3311-3323.

[31] Vermeulen, P. T. M., Heemink, A. W., & Te Stroet, C. B. M. (2004). Reduced models for linear groundwater flow models using empirical orthogonal functions. *Advances in water resources*, *27*(1), 57-69.

[32] Laub, A. J. (2005). *Matrix analysis for scientists & engineers*, Society for Industrial and Applied Mathematics, Philadelphia, 157 p.

[33] Laub, A. J. (2012). *Computational matrix analysis*, Society for Industrial and Applied Mathematics, Philadelphia, 154 p.

[34] McPhee, J., & Yeh, W. W. G. (2008). Groundwater management using model reduction via empirical orthogonal functions. *Journal of Water Resources Planning and Management*, *134*(2), 161-170.

[35] Siade, A. J., Putti, M., & Yeh, W. W. G. (2010). Snapshot selection for groundwater model reduction using proper orthogonal decomposition. *Water Resources Research*, *46*(8).

[36] Chen, S., Billings, S. A., & Luo, W. (1989). Orthogonal least squares methods and their application to non-linear system identification. *International Journal of control*, *50*(5), 1873-1896.

[37] Golub, G. (1965). Numerical methods for solving linear least squares problems. *Numerische Mathematik*, *7*(3), 206-216.

[38] Golub, G. H., & Reinsch, C. (1970). Singular value decomposition and least squares solutions. *Numerische mathematik, 14*(5), 403-420.

[39] Gentle, J. E. (2007). *Matrix algebra: theory, computations, and applications in statistics*. Springer Science & Business Media.

[40] Golub, G. H., & Van Loan, C. F. (1996). *Matrix computations*. Johns Hopkins University, Press, Baltimore, MD, USA, 374-426.

[41] Stewart, G. W. (1973), *Introduction to matrix computations*, Academic Press, New York, 441 p.

Chapter 2

Advances in Water Resources 69 (2014) 168–180

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Parameter-independent model reduction of transient groundwater flow models: Application to inverse problems

Scott E. Boyce, William W.-G. Yeh *

Department of Civil and Environmental Engineering, University of California, Los Angeles, CA 90095, USA

ARTICLE INFO

Article history:
Received 8 January 2014
Received in revised form 6 April 2014
Accepted 8 April 2014
Available online 18 April 2014

Keywords:
Model reduction
Proper orthogonal decomposition
Inverse problem
Markov Chain Monte Carlo
Greedy algorithm
Snapshot selection

ABSTRACT

A new methodology is proposed for the development of parameter-independent reduced models for transient groundwater flow models. The model reduction technique is based on Galerkin projection of a highly discretized model onto a subspace spanned by a small number of optimally chosen basis functions. We propose two greedy algorithms that iteratively select optimal parameter sets and snapshot times between the parameter space and the time domain in order to generate snapshots. The snapshots are used to build the Galerkin projection matrix, which covers the entire parameter space in the full model. We then apply the reduced subspace model to solve two inverse problems: a deterministic inverse problem and a Bayesian inverse problem with a Markov Chain Monte Carlo (MCMC) method. The proposed methodology is validated with a conceptual one-dimensional groundwater flow model. We then apply the methodology to a basin-scale, conceptual aquifer in the Oristano plain of Sardinia, Italy. Using the methodology, the full model governed by 29,197 ordinary differential equations is reduced by two to three orders of magnitude, resulting in a drastic reduction in computational requirements.

1. Introduction

Traditional inverse problems involve solving a weighted history matching optimization problem that yields a set of optimally selected parameters. The Bayesian inverse problem reformulates the solution by treating the parameters as random variables that are described by a posterior probability distribution. Markov Chain Monte Carlo (MCMC) methods are a powerful set of algorithms capable of exploring the probability space of the random variables used in the formulation of the Bayesian inverse problem. The MCMC simulation constructs a posterior distribution from samples generated by a proposal distribution. The samples are accepted or rejected to a Markov chain through the posterior density function and proposal density function evaluated at both the current step (previously accepted sample) and the proposed step (new sample). If the Markov chain is constructed correctly it should converge to a stationary distribution that represents the posterior distribution [1,2]. Shi et al. [3] evaluated for vadose zone modeling the confidence interval predictive performance of MCMC and compared with nonlinear regression. Their results indicated that MCMC produced better results and for small parameter dimensions was more computationally efficient. One method to lower the dimensionality of the parameter space is to parameterize it further with the Karhunen–Loeve expansion (KLE). Das et al. [4] applied MCMC with a KLE parameterization of saturated hydraulic conductivity fields for soil moisture problems.

One of the most commonly used MCMC methods for determining the acceptance of parameter samples is the Metropolis–Hastings (M–H) algorithm [5,6]. In the case of groundwater, and the focus of this paper's Bayesian inverse problem, the parameter of interest is the posterior probability distribution of hydraulic conductivity given historical water level data. This distribution can be used to quantify the uncertainty and assess its effects on model predictions [7]. A problem with solving the Bayesian inverse problem through M–H MCMC is that it requires a large number of sequential model simulations to characterize an unknown parameter's posterior probability distribution. There are parallel versions of MCMC, but they still require many sequential model simulations to construct the chains. As a result, Bayesian inversion for parameter estimation of a highly discretized groundwater simulation model can be computationally infeasible. Another alternative is to use a two-stage MCMC framework that relies on a surrogate model, composed of a coarser grid or simplified flow process, to first evaluate the acceptance of a proposed value before

* Corresponding author. Tel.: +1 310 825 2300; fax: +1 310 825 7581.
E-mail addresses: Boyce@engineer.com (S.E. Boyce), williamy@seas.ucla.edu (W.W.-G. Yeh).

http://dx.doi.org/10.1016/j.advwatres.2014.04.009
0309-1708/

its acceptance is evaluated by the full model [8]. This can further be enhanced by using adaptive sparse-grids [9].

Model reduction based on the Galerkin projection is a technique that projects a high dimensional model characterized by ordinary differential equations (ODEs) onto a low dimensional subspace, spanned by a small number of optimally chosen basis functions (principal components) [10]. The application of model reduction to a confined, groundwater model has been shown to reduce the dimensionality by several orders of magnitude. A variety of papers have been written on model reduction techniques. Vermeulen et al. [11] applied proper orthogonal decomposition (POD) for model reduction to groundwater equations by collecting an ensemble of hydraulic head solutions, called snapshots, at specific times from the simulation and at a constant, reference pumping rate. Snapshots have to be taken for each reference extraction/injection well. Vermeulen et al. [11] then applied principal component analysis (PCA) to this ensemble to form a projection matrix that reduces the groundwater model. McPhee and Yeh [12] followed this methodology and demonstrated that a POD reduced model maintains its sensitivity of head with respect to pumping, enabling it to embed in a management optimization problem. Baú [13] increased the utility of POD by deriving a reduced model for each Monte Carlo realization of hydraulic conductivity to solve a stochastic, multi-objective, confined groundwater management problem.

In principle, the model reduction technique applies to linear systems, such as confined aquifers, because it uses the principle of superposition. Application to nonlinear systems is possible, but the reduced model error would be greater and may require significantly more basis functions to characterize the model [14]. Robinson et al. [15] and Li and Hu [16] applied POD model reduction to several synthetic one- and two-dimensional mass transport models without chemical reactions. Buchan et al. [17] solved for the population growth of free moving neutrons, an eigenvalue problem, in a nuclear reactor system. The eigenvalue problem was reformulated to create pseudo-time dependence that describes the snapshots used in their projection basis.

In general, the reduced model depends on the data used to construct the projection matrix. The data consists of snapshots generated from the original full model for a given set of model parameter values. Thus, the reduced model may be sensitive to changes in parameters. This causes problems when the reduced model is used for solving the inverse problem of parameter estimation.

Developing parameter-independent reduced models is a new area of active research. Vermeulen [18] applied a reduced model to an inverse problem by taking snapshot sets over a specific range of parameter combinations. The drawback of this procedure is that with a large number of parameters the combinations can get very large. Additionally, if the parameters move away from the specified range, the accuracy of the reduced model drops and a new set of snapshots is required. Lieberman et al. [19] proposed a greedy algorithm for the construction of a projection-based reduced model that reduces the parameter and state spaces for a steady state statistical inverse problem. A greedy algorithm solves a multi-stage optimization problem by combining the optimal solution obtained from each stage. At each stage, the algorithm selects the local optimum and moves on to the next stage. A solution to the original multi-stage optimization problem is built up stage by stage. In general, this greedy strategy does not guarantee global optimum, but in many instances yields a good approximation to the optimal solution. The advantages of the algorithm are its easy implementation and fast execution. The objective function proposed by Lieberman et al. [19] for the selection of the optimal parameter set maximized the error at steady state between the original full model and the reduced model. The parameter set that resulted from the optimization and its corresponding steady state solution were added to their respective projection matrices. The procedure was repeated until the specified error criterion was satisfied. Pasetto et al. [20] proposed an algorithm to reduce the computational burden associated with combinatorial search. The algorithm applied a greedy algorithm that searched over sets of parameter combinations and a snapshot selection strategy proposed by Siade et al. [21]. The greedy objective was evaluated using a scaled residual derived from the reduced model. This reduces the number of full model evaluations required for the determination of the principal components to be included in the reduced model.

In this study, we develop a new methodology for building the projection matrix for transient groundwater flow. Our proposed methodology is intended to work for linear, regional groundwater models where the zonation structure already has been determined; that is, the aquifer has been divided into a finite number of zones and each zone is characterized by a constant parameter (or parameters). The challenge of determining the optimum zonation structure of a random field by parameterization lies outside the scope of this paper.

This paper is organized into six sections. Section 1 is an introduction. Section 2 presents the governing equation for the confined aquifer and defines the notations. Section 3 discusses the deterministic inverse problem and the Bayesian inverse problem. Section 4 reviews the concept of projection-based model reduction and develops a parameter independent model reduction methodology. Section 5 applies the model reduction methodology to one- and two-dimensional test cases. Section 6 concludes the findings and discusses the results.

The proposed methodology constructs a projection matrix that covers the entire parameter space in the original full model and does not require taking new snapshots while solving the inverse problem. The projection matrix is assembled from snapshots generated iteratively by two greedy algorithms that select optimal parameter sets and snapshot times between the parameter space and the time domain. The proposed methodology is validated using a conceptual one-dimensional model that compares the result of a deterministic inverse problem with the empirical statistics from a Bayesian inverse problem. We then apply the methodology to a basin-scale, conceptual aquifer in the Oristano plain of Sardinia, Italy. Using the methodology, the full model is reduced by two to three orders of magnitude, resulting in a drastic reduction in computational requirements.

2. Confined groundwater modeling

The governing equation for confined, anisotropic, saturated groundwater flow can be expressed by the following parabolic partial differential equation [22,23]:

$$\frac{\partial}{\partial x}\left(K_x\frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(K_y\frac{\partial h}{\partial y}\right)+\frac{\partial}{\partial z}\left(K_z\frac{\partial h}{\partial z}\right)\pm Q=S_S\frac{\partial h}{\partial t} \tag{1}$$

where h is the hydraulic head (L); K_x, K_y, K_z are the hydraulic conductivities (L/T) in the x, y, and z directions; S_S is the specific storage (L^{-1}); Q is a volumetric flux per unit volume in or out of the system (T^{-1}); and t is the time (T). Eq. (1) is subject to the following initial and boundary conditions:

$$h(x,y,z,t)=h_I(x,y,z),\quad (x,y,z)\in\Gamma_F,\ t=0$$
$$h(x,y,z,t)=h_D(x,y,z,t),\quad (x,y,z)\in\Gamma_D$$
$$K\frac{\partial h(x,y,z,t)}{\partial n}=q_N(x,y,z,t),\quad (x,y,z)\in\Gamma_N$$
$$\Gamma_D\cup\Gamma_N=\Gamma_B$$

where h_I is the initial condition, h_D is a specified Dirichlet boundary condition, q_N is a specified Neumann boundary condition, $\frac{\partial}{\partial n}$ is the normal derivative, Γ_F is the flow region, and Γ_B is the boundary

of the aquifer composed of Dirichlet (Γ_D) and Neumann (Γ_N) conditions.

Applying finite-difference or finite-element approximations to the spatial derivatives of Eq. (1) yields a system of ordinary differential equations (ODEs) of the form

$$Ah + f = B\frac{dh}{dt} \tag{2}$$

where $h \in \mathbb{R}^n$ is a vector of hydraulic heads at each nodal location, $A \in \mathbb{R}^{n\times n}$ is the stiffness matrix that contains spatial information about hydraulic conductivity, $B \in \mathbb{R}^{n\times n}$ is the mass matrix, and f is a vector containing all boundary conditions and external forcing (e.g., pumping). Typical construction of A and B yields large, sparse, symmetric and positive definite matrices. Eq. (2) subsequently will be referred to as the full model.

3. The inverse problem

3.1. Deterministic inverse problem

A common objective for the groundwater inverse problem is identifying a vector of hydraulic conductivity values that minimizes the Sum of Squared Errors (SSE) between a set of observed hydraulic head data and the solution to Eq. (2) [24]. This problem is mathematically formulated as:

$$\arg\min_{K\in\mathcal{K}} \sum_{j=1}^{N_t}\sum_{i=1}^{N_l}(D_{i,j} - h_{\mathrm{op}(i),j})^2 \tag{3}$$

where the decision variable, K, is the set of zoned hydraulic conductivity values used to solve Eq. (2); $\mathcal{K}$ is the search space of K, which is usually a continuous interval from a $K_{\min}$ to $K_{\max}$; i and j are the observation location and time indices, respectively; N_l and N_t are the total number of observation locations and times, respectively; $D_{i,j}$ and $h_{\mathrm{op}(i),j}$ are the observed and model simulated hydraulic head; and op(i) maps the location index to the model nodal index.

3.2. Bayesian inversion problem through Markov Chain Monte Carlo

Bayesian methods link prior knowledge of a model's parameters to observed information and the model's output. A posterior distribution of the parameters can be derived from a Markov Chain Monte Carlo (MCMC) simulation [2]. The residual between an observation, $D_{i,j}$, and its corresponding groundwater simulation model result, $h_{\mathrm{op}(i),j}$, is assumed to be normally distributed, $D_{i,j} - h_{\mathrm{op}(i),j} = \varepsilon_{i,j} \sim N(0, \sigma_\varepsilon^2)$, with zero mean and an unknown variance, σ_ε^2, that has an Inverse Gamma prior distribution, $\Pi(\sigma_\varepsilon^2) \sim IG(\alpha, \beta)$ [25,26]. The variable, ε, then lumps together a representation of all sources of error such as model structure, parameter error, and observation error [3]. Using Bayes' Rule with the above-mentioned assumption the following relationships can be derived:

$$P(K|D,\sigma_\varepsilon^2) = \frac{P(D|K,\sigma_\varepsilon^2)\Pi(K)}{P(D)} \propto P(D|K,\sigma_\varepsilon^2)\Pi(K) \tag{4}$$

where:

$$P(D|K,\sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-(N_t+N_l)/2}\exp\left[-\frac{1}{2\sigma_\varepsilon^2}\sum_{j=1}^{N_t}\sum_{i=1}^{N_l}(D_{i,j} - h_{\mathrm{op}(i),j})^2\right] \tag{5}$$

$$\Pi(K) = \frac{1}{K\sqrt{2\pi\sigma_\pi^2}}\exp\left[-\frac{(\ln K - \mu_\pi)^2}{2\sigma_\pi^2}\right] \tag{6}$$

Eq. (4) is simplified to a proportional relationship consisting of a likelihood function, Eq. (5), and a known prior distribution of hydraulic conductivity, Eq. (6). The prior distribution for the hydraulic conductivity is assumed to be log-normally distributed with shape parameters μ_π and σ_π^2 and parameters are uncorrelated [27]. Its mean and variance are chosen based on the prior knowledge of the hydraulic conductivity and are used to compute the shape parameters.

The Metropolis–Hastings algorithm is an MCMC method that generates a sequence of samples that approximate the posterior distribution of the hydraulic conductivity. The Markov chain is constructed from a set of proposed values, K^*, that are sampled from a proposal density function and are either accepted or rejected based on the likelihood function and the prior distribution.

We apply the following proposal density function (Pr( )) for generating hydraulic conductivity samples:

$$\Pr(K^*|K) = P(K \to K^*) = \frac{1}{K^*\sqrt{2\pi\sigma_p^2}}\exp\left[-\frac{(\ln(K^*) - \ln(K) - \sigma_p^2)^2}{2\sigma_p^2}\right] \tag{7}$$

The proposal function is log-normally distributed with a mode equal to the current value in the chain, K, and shape parameter, σ_p^2, set to a user-specified value. When a candidate, composed of N hydraulic conductivity values, is generated from Eq. (7), its acceptance probability, $\hat{\alpha}$, is calculated by taking the ratio presented in the following equation:

$$\hat{\alpha} = \min\left\{\frac{P(D|\underline{K}^*,\sigma_\varepsilon^2)}{P(D|\underline{K},\sigma_\varepsilon^2)}\prod_{i=1}^{N}\left[\frac{\Pi(K_i^*)}{\Pi(K_i)}\right]\prod_{i=1}^{N}\frac{\Pr(K_i|K_i^*)}{\Pr(K_i^*|K_i)}, 1\right\} \quad \text{where } \underline{K}, \underline{K}^* \in \mathbb{R}^N \tag{8}$$

Applying the given likelihood, Eq. (5), and prior for the hydraulic conductivity, Eq. (6), to the log of the acceptance probability yields the following equation:

$$\ln\hat{\alpha} = \min\left\{\frac{1}{2\sigma_\varepsilon^2}\sum_{j=1}^{N_t}\sum_{i=1}^{N_l}(D_{i,j} - h_{\mathrm{op}(i),j})^2 - \frac{1}{2\sigma_\varepsilon^2}\sum_{j=1}^{N_t}\sum_{i=1}^{N_l}(D_{i,j} - h^*_{\mathrm{op}(i),j})^2 + \frac{1}{2\sigma_\pi^2}\sum_{i=1}^{N}[(\ln K_i - \mu_\pi)^2 - (\ln K_i^* - \mu_\pi)^2], 0\right\} \tag{9}$$

If the prior's variance is not known, then it is set to an arbitrarily large value in order to lower its influence on the posterior. The log is taken to simplify the equations for faster computation. This log of the acceptance probability is then compared to the log of a random uniform (0 to 1) draw. If it is larger, the value of the hydraulic conductivity is accepted as a new value in the chain [2]. The following algorithm summarizes the steps that involve the generation of a Markov chain based on the Metropolis–Hastings method:

Algorithm 1. Metropolis–Hastings Method

Given $K^{(t)} := K^*$ and $K^{(0)} := (K_1^{(0)}, \ldots, K_N^{(0)})$; iterate algorithm for $t = 1, 2, \ldots$

1. Generate a candidate from a proposal distribution $K^t \sim \Pr(K^t|K^{t-1})$.
2. Generate from the uniform distribution $\alpha \sim U[0,1]$.
3. Using the target density to solve $\hat{\alpha} = \min\left\{\frac{P(D|\underline{K}^t,\sigma_\varepsilon^2)}{P(D|\underline{K}^{t-1},\sigma_\varepsilon^2)}\prod_{i=1}^{N}\left[\frac{\Pi(K_i^t)}{\Pi(K_i^{t-1})}\right]\prod_{i=1}^{N}\frac{\Pr(K_i^{t-1}|K_i^t)}{\Pr(K_i^t|K_i^{t-1})}, 1\right\}$.
4. If $\alpha < \hat{\alpha}$, then keep K^t, else set $K^t = K^{t-1}$.
5. Set $t = t + 1$ and go to Step 1.

The posterior distribution of σ_ε^2 is a combination of a likelihood function and its prior distribution, $\Pi(\sigma_\varepsilon^2)$:

$$P(\sigma_\varepsilon^2|K,D) \propto P(D|K,\sigma_\varepsilon^2)\Pi(\sigma_\varepsilon^2) \sim \mathrm{IG}\left(\frac{N_t+N_l}{2}+\alpha, \frac{\sum_{j=1}^{N_t}\sum_{i=1}^{N_l}(D_{i,j}-h_{\mathrm{op}(i),j})^2}{2}+\beta\right) \quad (10)$$

Since its prior distribution is a conjugate distribution, its posterior can be derived analytically as an inverse gamma and directly sampled from Eq. (10) [26]. The choice for parameters α and β have a small impact on the posterior because the groundwater model SSE tends to be very large, $\sum_{j=1}^{N_t}\sum_{i=1}^{N_l}(D_{i,j}-h_{\mathrm{op}(i),j})^2 \gg \beta$, and the observed data count is likely to be significantly larger than one, $N_t+N_l \gg \alpha$.

Solving for the two posterior distributions described in Eqs. (4) and (10) involves an initial starting point and then iterating through a three step process. The Metropolis–Hastings algorithm is independent of the initial value; however, it can affect the length of time it takes for the chain to converge. To improve the speed of convergence the initial values, $K^{(0)}$, are the solution to the deterministic inverse problem in Eq. (3). With the initial hydraulic conductivity values given, the first step is to generate a sample of the likelihood variance, σ_2, from Eq. (10). The second step is to sample a new hydraulic conductivity set, (K^*), from Eq. (7), run the simulation model with the sample, and then evaluate the acceptance probability, $\hat{\alpha}$. If the new values are accepted they are used to generate a new likelihood variance, σ_ϵ^2, and the process is repeated. If they are rejected, then the process is repeated with the original values. This process continues until the Markov chain converges to the stationary distribution that represents a sample from the posterior distribution.

4. Model reduction of a confined groundwater system

4.1. Projection-based model reduction

Projection-based model reduction proceeds by projecting a linear, discretized groundwater model, $h \in \mathbb{R}^n$, onto a reduced model subspace of $h_r \in \mathbb{R}^r$, where $r \ll n$. For this to occur, it is assumed that the hydraulic head, h, can be mapped linearly onto the reduced space, h_r, by some orthonormal matrix, $P \in \mathbb{R}^{n\times r}$, which contains information about the discretized groundwater model's solutions in time [11]:

$$h = Ph_r \quad (11)$$

The vectors in the P matrix are sometimes referred to as spatial basis functions, empirical orthogonal functions, or principal vectors.

Let S be a set of snapshots:

$$S = [h^{t_1}, h^{t_2}, h^{t_3}, \ldots, h^{t_r}] \quad (12)$$

Snapshots are solutions to the groundwater model, h^{t_i}, under a constant/reference forcing (e.g. pumping well) taken at specific simulation times, t_i. In principle, a set of snapshots has to be taken for each forcing term in the groundwater model. For example, if a model has two extraction wells, then two sets of snapshots must be taken. The first set is taken with one extraction well pumping at a constant rate, while the other well's pumping is set to zero. The second set is taken by reversing the pumping rate between the two wells [28]. The snapshots form a basis that has to be made orthonormal to use as a projection matrix, P, that is the same matrix as the linear map in Eq. (11). The orthonormalization of the snapshots, S, is performed by either QR decomposition [29,30],

$$S = QR \rightarrow P = Q \quad (13)$$

or Singular Value Decomposition (SVD),

$$S = U\Sigma V^T \rightarrow P = U \quad (14)$$

It is important to recognize that Q may not necessarily be equal to U, but they do have the same range space that is spanned by the snapshots. We choose to use QR decomposition because of its efficiency and numerical accuracy. Further reduction of the projection matrix may be possible by applying principal component analysis (PCA) on snapshots [11,12,18,21,28,31,32]. However, this paper's methodology finds that further removing basis vectors from snapshots through PCA results in an insignificant change in the reduced space dimension, r, and may cause the reduced model to lose its parameter independence. The reason PCA has a minimal impact is that the proposed methodology, presented in the next section, systematically includes the worst fitting snapshot until there is a high fidelity between the full and reduced models.

The discretized groundwater model presented in Eq. (2) is projected onto a reduced space by substituting in $h = Ph_r$ and applying the Galerkin projection to form the following reduced model:

$$P^TAPh_r + P^Tf = P^TBP\frac{dh_r}{dt} \quad (15)$$

Let $A_r = P^TAP \in \mathbb{R}^{r\times r}$, $B_r = P^TBP \in \mathbb{R}^{r\times r}$, and $f_r = P^Tf \in \mathbb{R}^r$, such that Eq. (15) becomes

$$A_rh_r + f_r = B_r\frac{dh_r}{dt} \quad (16)$$

Eq. (16) subsequently will be referred to as the reduced model. As we can see, the reduced model has the identical form of the full model in Eq. (2). The reduced system now only requires the solution of r ODEs compared to the full model that has n ODEs. For the reduced model to be efficient, it is necessary that $r \ll n$. A special property of this substitution and projection is that if A is positive definite and P is full column rank, then A_r is positive definite.

4.2. Parameter-independent projection matrix construction

Snapshots collected to form the projection matrix, P, are predicated on the parameters that are used to generate the snapshots. Since both deterministic and Bayesian inverse problems explore a wide range of parameter values, the validity of a projection matrix that is based on one set of parameters is subject to question. This problem is overcome by combining sets of snapshots from different parameter sets, $S(K_j) = [h^{t_1}(K_j), h^{t_2}(K_j), h^{t_3}(K_j), \ldots, h^{t_r}(K_j)]$ where S and h^{t_i} are now a functions of K_j, into one large snapshot set that, when made orthonormal, forms a parameter-independent projection matrix as follows:

$$\tilde{S} = [S(K_1), S(K_2), \ldots, S(K_j)] = QR \rightarrow P_j = Q \quad (17)$$

The questions that arise are which parameters are best to take snapshots from and what are the best simulation times, t_i, to use. The questions are answered through the use of two greedy algorithms that iteratively build the projection matrix, P, until it is sufficiently robust for all parameter values within a specified range and their corresponding snapshot times. Given a projection matrix, P_j, the first greedy algorithm, called *Parameter Greedy Step*, solves the following equation:

$$\arg\max_{K_{j+1}\in\mathcal{K}} \mathcal{J}_{K_{j+1}} = \left\| \frac{h^{SS}(K_{j+1})}{\|h^{SS}(K_{j+1})\|} - \frac{P_jh_r^{SS}(K_{j+1})}{\|P_jh_r^{SS}(K_{j+1})\|} \right\|_2^2 \quad (18)$$

Eq. (18) searches for a hydraulic conductivity set, K_{j+1}, within a range of reasonable values, $\mathcal{K}$, that maximizes the error between a normalized full model's steady state head, $h^{SS}(K_{j+1})$, and the normalized reduced model's steady state head, $P_jh_r^{SS}(K_{j+1})$. This optimization approach can be viewed as searching for a parameter set, K_{j+1}, where the reduced model constructed from P_j fails to represent the shape of the original full model. The reason for using the

steady state solution is that it represents the final shape of the model's response to pumping. This final response should always be included as a snapshot for constructing an accurate reduced model. Snapshots are then taken from the optimized parameter set to improve the robustness of the reduced model. Once the snapshots for K_{j+1} have been included into $\tilde{S}$, the projection matrix P_{j+1} is formed and Eq. (18) determines the next parameter set. This process continues until the following equation is satisfied:

$$\mathcal{J}_{K_{j+1}} < \alpha_K \mathcal{J}_{K_2} \quad \forall\, j > 2 \tag{19}$$

Eq. (19) is satisfied when $\mathcal{J}_{K_{j+1}}$ is reduced by α_K orders of magnitude from the initial $\mathcal{J}_{K_2}$ value, where $\mathcal{J}_{K_2}$ is obtained from the initial projection matrix, P_1, used to initiate the algorithm. From numerical experiments it has been found that values of α_K that work the best range from 10^{-3} to 10^{-5} (less accurate to more accurate, respectively).

The steady state heads in Eq. (18) are normalized ($h^{SS}/\|h^{SS}\|$) to improve the solution stability by comparing their shapes rather than the magnitudes of drawdown. Without the normalization, the optimal parameter set would approach zero to maximize drawdown. The shape is of greater importance because, ultimately, the final projection matrix is made orthonormal which removes the magnitudes of the snapshots.

Given the jth parameter set selected from the *Parameter Greedy Step*, a second greedy algorithm, called *Time Greedy Step*, searches for the optimal snapshot times, t_i, by solving the following maximization problem:

$$\arg\max_{t_i \in \mathcal{T}} \mathcal{J}_{t_i} = \left\| h^{t_i}(K_j) - P_j h_r^{t_i}(K_j) \right\|_2^2 \tag{20}$$

where $\mathcal{T}$ is simulation time frame. The snapshots are added sequentially to $S(K_j)$, updating $\tilde{S}$, which is then made orthonormal to reform P_j. Each subsequently added snapshot from the *Time Greedy Step* (i) increases the column dimension of P_j by one, and as a consequence renders the reduced model more accurate for the parameter set K_j. The reduced model is considered sufficiently accurate when one of the following criteria is satisfied:

$$\mathcal{J}_{t_i} < \alpha_t \mathcal{J}_{t_1} \tag{21}$$

$$\|h^{t_i}(K_j) - P_j h_r^{t_i}(K_j)\|_\infty < \epsilon \tag{22}$$

Eq. (21) is the main convergence criterion, which ensures that the reduced model is sufficiently accurate for K_j when the objective value, $\mathcal{J}_{t_i}$, has been reduced as a fraction of the first objective value, $\mathcal{J}_{t_1}$. A reduction in three orders of magnitude, $\alpha_t = 10^{-3}$, provides the best balance between snapshot count and reduced model accuracy. The second convergence criterion, Eq. (22), sets a maximum absolute error bound; for example, $\epsilon = 1$ mm.

For this methodology to begin it needs a starting parameter set and an initial projection matrix to build upon. The first hydraulic conductivity values are set equal to one for all zones, $K_1 = \mathbf{1}$ (a vector of 1's), and the corresponding steady state solution is used to compute the first projection matrix, P_1. To start the algorithm, an equal weight is applied to all zones. If prior information on parameter values is available, for example, the mean hydraulic conductivity values (μ_K), those values should be used to start the algorithm ($K_1 = \mu_K$). Again, we use the steady state solution as the starting vector as it represents the final shape of the model's response to pumping and should always be included as a snapshot. Consequently, the snapshot time selected to start the algorithm corresponds to the time when a steady state solution is reached. The *Time Greedy Step* builds upon the starting projection matrix P_1 until the reduced model is sufficiently accurate for K_1. Then the algorithm begins to iterate between the *Parameter Greedy Step* and the *Time Greedy Step* until their convergence criteria are satisfied (Eqs. (19), (21) and (22)). The convergence bounds α_K, α_t and ϵ can be changed depending on the level of accuracy required for the final reduced model. The overall algorithm for constructing the parameter-independent projection matrix, P_j, is summarized below and illustrated in Fig. 1.

Algorithm 2.

1. Set indices $i = 1, j = 1$, and the initial hydraulic conductivity as constant for all zones, $K_1 = \mathbf{1}$ (a vector of 1's). Initialize $\tilde{S}$ as an empty set.
2. Solve for the steady state snapshot using K_j, add it to $\tilde{S}$ and make it orthonormal to get P_j.
3. Solve the ith *Time Greedy Step* using K_j.
4. If $i = 2$, then establish $\mathcal{J}_{t_2}$, else if the convergence criteria is satisfied go to Step 7.
5. Add the new snapshot at t_i to $S(K_j)$, which is added to $\tilde{S}$.
6. Make $\tilde{S}$ orthonormal via QR decomposition to reform P_j, set $i = i + 1$, and go to Step 3.
7. Solve the $j + 1$ *Parameter Greedy Step* to find K_{j+1} using P_j.
8. If $j = 2$, then establish $\mathcal{J}_{K_2}$, else if the convergence criteria is satisfied, exit with P_j as the final parameter-independent projection matrix.
9. Set $j = j + 1$, $i = 1$ and go to Step 2.

5. Application of parameter-independent model reduction to a confined groundwater model

5.1. One-dimensional test case

Saturated 2-dimensional flow (SAT2D) is a finite-element model for a two-dimensional, confined, groundwater flow [33]. Using SAT2D, we construct a one-dimensional model to demonstrate and validate the proposed methodology. To simulate a one-dimensional groundwater flow in SAT2D, all parameters are distributed appropriately across the width of the model. The model dimensions are 100 m in length by 10 m wide by 1 m deep. The lithology is partitioned into five hydraulic conductivity zones, $K_j = [K_{j,1}, K_{j,2}, K_{j,3}, K_{j,4}, K_{j,5}]$, with a specific storage, S_s, of 1 m^{-1} (an unusually large value is used here to shorten the time required to reach a steady state solution). An initial condition of 0 m is assumed. The two fixed-head Dirichlet boundary conditions are set to 0 m. There is an extraction well located at the center with a pumping rate of 20 m/d (per unit area). Fig. 2 shows the one-dimensional test case. The finite element grid consists of three rows of 101 nodes (because SAT2D requires three rows of 101 nodes to create a finite element mesh to simulate the aquifer), which requires solving a system of 303 linear ODEs for each advancement in time (Δt). The ODEs are solved through an implicit finite difference scheme.

The methodology presented in Algorithm 2 is applied to the one-dimensional model to construct a parameter-independent projection matrix. As stated in step 1, the starting parameters, K_1, are set to 1 m/d for all five zones. Eqs. (18) and (20), from Steps 7 and 3, respectively, are solved as bounded fitness functions through a Genetic Algorithm program developed by Carroll (FORTRAN genetic algorithm driver, version 1.7a, 2001, available at http://cuaerospace.com/carroll/ga.html). The GA code solves for each hydraulic conductivity set in Eq. (18) using a population of 100 chromosomes with jump mutation of 0.01, creep mutation of 0.02, uniform crossover, elitism, and niching. The lower and upper bounds used by GA for solving the *Parameter Greedy Step*, Eq. (18),

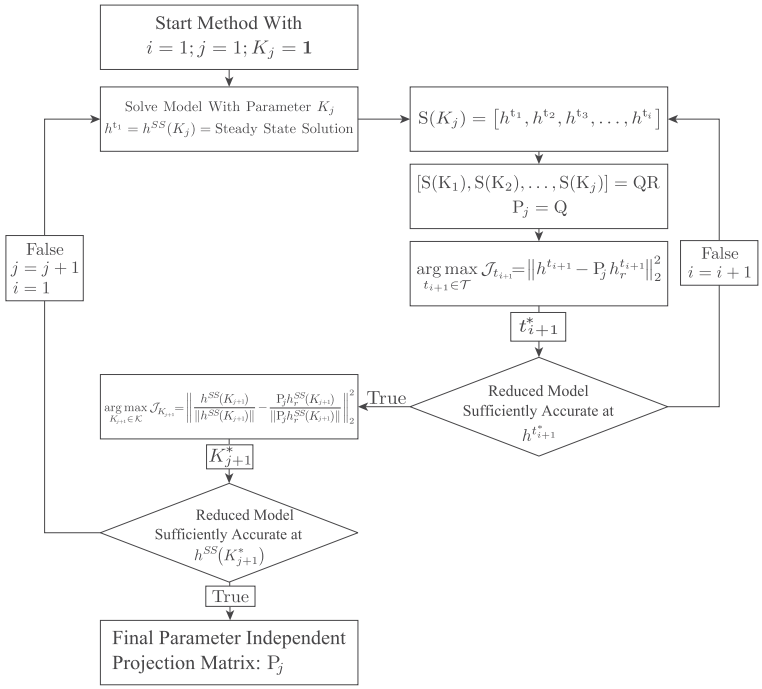

Fig. 1. Flowchart of the methodology for construction a parameter-independent projection matrix. Note that K*, t* are the solutions to their respective argmax optimization problems.

are 0.1 m/d and 20 m/d, respectively, for each of the hydraulic conductivity zones. The advantage of using GA instead of a gradient-based search method is that it is derivative-free and, with multiple starting values, it is more likely to find the global optimum for a nonlinear and non-convex optimization problem. Additionally, it is an "embarrassingly parallel" workload for solving each generation. Once the hydraulic conductivity set is selected the GA code again solves the *Time Greedy Step*, Eq. (20). The lower and upper bounds for the snapshot times are 0.1 d and 1000 d, respectively. Another advantage of using GA is that the *Time Greedy Step* only requires a single model call per generation (i.e. step towards optimal solution) to evaluate a set of snapshot times.

Each *Parameter Greedy Step* progressively reduces the objective function value until the convergence criterion of Eq. (19) is satisfied. Fig. 3 shows the rate of convergence for each of the optimally selected parameter sets. The final parameter-independent projection matrix from Algorithm 2 contains 33 snapshots from five different hydraulic conductivity sets. When this projection matrix is applied as shown in Eq. (15), it is capable of reducing the system of 303 ODEs required at each time step to 33 ODEs for a wide range of hydraulic conductivity values. Fig. 4 shows the hydraulic conductivity sets selected at each *Parameter Greedy Step* and their corresponding snapshot times selected by the *Time Greedy Step*.

To evaluate the reduced model for solving the inverse problems a synthetic set of 41 head observations ($t_{obs} = 1, 5, 10, 15, \ldots, 195, 200$ d) are generated from the full model. Observations are taken at the center of each hydraulic conductivity zone at an assumed "true value", $K_{True} = [1, 5, 10, 8, 5]$ m/d. The generated observations are corrupted by adding Gaussian noise with zero mean and 0.1 variance. Using the corrupted observations, hydraulic conductivity is estimated by Eq. (3) using Carroll's FORTRAN genetic algorithm driver without parallelization. The upper and lower bounds are assumed to be 0.01 and 50 m/d, respectively. The full and reduced models' solutions to the deterministic inverse problem are presented in Table 1. These values are used as $K^{(0)}$ in Algorithm 1 to minimize the burn-in period and construct a Markov chain of 100,000 samples. The chain is subsequently thinned by keeping every tenth value for a final chain length of 10,000. The chains from the full model and the reduced model are constructed with a prior mean equal to the solution of the deterministic inverse problem and variance of 100. The proposal distribution has a mode equal to the current value in the chain and shape parameters hand calibrated to generate candidates with a 40% acceptance rate. Using the full and reduced Markov chains the posterior distributions are reconstructed using Python's (SciPy) kernel density estimation function that uses the mathematics defined in [34]. Fig. 5 shows the comparison of the posterior density functions

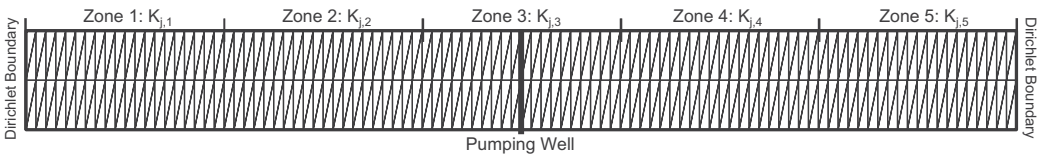

Fig. 2. One-dimensional groundwater flow model.

between the full model and the reduced model. Visually the two distributions are indistinguishable. The descriptive statistics in Table 1 show that the results obtained from the full model and the reduced model are within a relative error of 1% to 5% of each other.

Four statistical tests are performed to validate that the Markov chains produced from the full model and the reduced model are equivalent. The first two tests are the Levene and the Fligner–Killeen Equal Variance Tests, both contain the null hypothesis that the two sample populations (i.e. the full and the reduced Markov chains) have equal variance. The third test is the one-way analysis of variance (ANOVA), which employs a null hypothesis that the two sample populations have the same population mean. The fourth test is the Kolmogorov–Smirnov 2 Sample Equal Distribution Test, which features a null hypothesis that the two sample populations originated from the same probability distribution. All four statistical tests failed to reject their corresponding null *hypotheses* at a significance level of 0.05 (the smallest p-value is 0.18). This provides strong evidence that the full and reduced Markov chains are statistically equivalent.

To evaluate the robustness of the reduced model, we perform a series of error analysis tests using the following set of equations:

$$\begin{aligned} \mathrm{MAE} &= (1/\mathrm{N})\sum_{i=1}^{\mathrm{N}} |h_i - (\mathrm{P}h_r)_i| \\ \mathrm{RMSE} &= \sqrt{(1/\mathrm{N})\sum_{i=1}^{\mathrm{N}} (h_i - (\mathrm{P}h_r)_i)^2} \\ \mathrm{MARE} &= 100\% \times (1/\mathrm{N})\sum_{i=1}^{\mathrm{N}} |(h_i - (\mathrm{P}h_r)_i)/h_i| \;\forall\; h_i \neq 0 \\ \mathrm{NRMSE} &= 100\% \times \mathrm{RMSE}/(h_{\max} - h_{\min}) \end{aligned} \tag{23}$$

where MAE is the mean absolute error, RMSE is the root mean square error, MARE is the mean absolute relative error, NRMSE is the normalized root mean square error, N is the number of model nodes evaluated, i is the node index, and $h_{\max}$ and $h_{\min}$ are the maximum and minimum head value across all nodes, respectively. It is assumed that errors propagate in the reduced model with each time step, so the greatest error will occur at the last time step. To evaluate the error analysis tests, Eq. (23), we use 5000 randomly generated hydraulic conductivity sets (zone 1 through zone 5) that are sampled from the prior, Eq. (6), with the mean and variance obtained from the deterministic inverse solution. The solutions from the full and reduced models for each of the sampled parameter sets are stored for all nodes at 200 days (last time step). These values are used to compute the error statistics described in Eq. (23).

Fig. 6 shows exceedance curves from the error statistics described in Eq. (23) for the 5000 randomly generated hydraulic conductivity sets. The error analysis results demonstrate that the reduced model accurately represents the full model over the entire parameter space. The few parameter sets that result in large errors occur for hydraulic conductivity samples near zero in Zone 3 (well location).

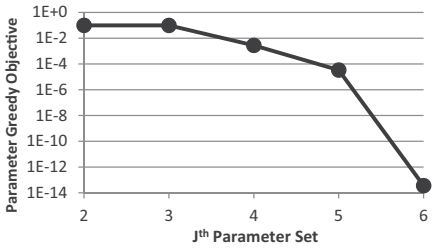

Fig. 3. Semi-log plot of the one-dimensional, five zone, *Parameter Greedy Step* objective for each of the optimally selected parameter sets.

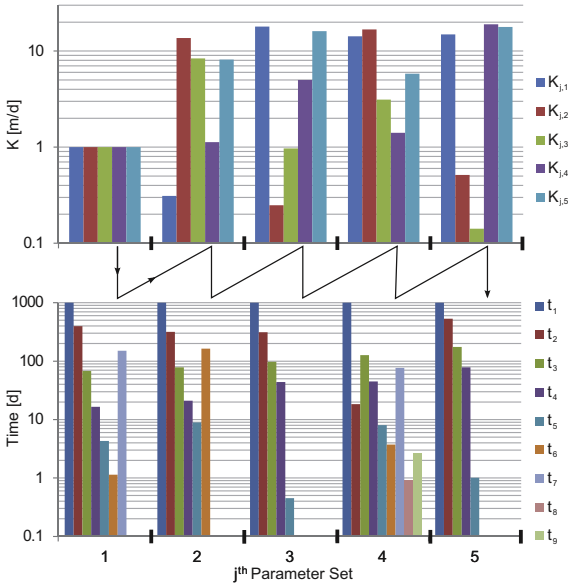

Fig. 4. Hydraulic conductivity sets selected from the one-dimensional model at each *Parameter Greedy Step* (top) and their corresponding snapshot times selected by the *Time Greedy Step* (bottom). Note that for all the parameter sets the first snapshot time corresponds the time when a steady state is reached and is graphically represented as 1000 days.

To benchmark the time savings from the reduced model, we calculate the average runtime of 1,000 SAT2D simulations using the "true values" for the hydraulic conductivities. The simulation model is compiled with default optimizations on the Intel Fortran Compiler and linked to the Intel Math Kernel Library on a single core of an Intel 2.67 GHz Xeon X5550 running CentOS 6. The full model uses the default PCG solver that comes with SAT2D while the reduced model uses LAPACK's Cholesky decomposition (DPOSV) as its solver. For a simulation time of 200 days, the average runtime of the full model is 0.3454 s while the average runtime of the reduced model is 0.2649 s. The application of model reduction results in a computational time savings of 23.3% for the one-dimensional case. Due to the simplicity of the one-dimensional model the time saving is insignificant. The purpose of using a simple, one-dimensional model is to verify the proposed methodology. As will be seen later, for a highly discretized model, the reduction in dimension is drastic and time-saving is significant.

5.2. Two-dimensional test case: Oristano plain conceptual model

We test the proposed parameter-independent model reduction methodology using synthetically generated data on a conceptual, large, confined aquifer located in the Oristano plain of Sardinia, Italy [20,32,35,36]. Synthetic data is used to avoid unknown sources of error, such as model structure error, and keep the focus on the validation of the proposed model reduction methodology. The model, referred to as the Oristano model, is solved with SAT2D [33]. The two-dimensional finite-element model consists of 57,888 triangular elements and 29,197 nodes. The confined aquifer has a constant depth of 100 m and a constant specific storage of 10^{-5} m^{-1}. The flow domain is bounded by a zero-Dirichlet boundary condition. There are six pumping wells in the aquifer with a constant pumping rate of 1000 m/d. Fig. 7 shows the Oristano model boundary, finite element grids and pumping well locations. Fig. 8 shows the three different zonation patterns and the

Table 1
Comparison of the deterministic inverse problem (DIP), 5% confidence interval (CL), 95% confidence interval, and median from the full and reduced one-dimensional models. Note that $K_1,\ldots,K_5$ refers to hydraulic conductivity for zone 1 to zone 5.

	DIP		5% CL		95% CL		50% CL (Median)	
	Full	Reduced	Full	Reduced	Full	Reduced	Full	Reduced
K_1	1.005	0.972	0.829	0.830	1.305	1.298	1.000	0.996
K_2	5.099	5.198	4.594	4.590	5.774	5.759	5.124	5.122
K_3	9.602	9.275	8.488	8.488	11.010	11.007	9.574	9.582
K_4	8.344	8.810	7.219	7.218	9.861	9.876	8.377	8.372
K_5	4.900	4.878	4.450	4.454	5.291	5.292	4.861	4.863
σ_ϵ^2	–	–	0.0792	0.0791	0.1718	0.1722	0.1138	0.1139

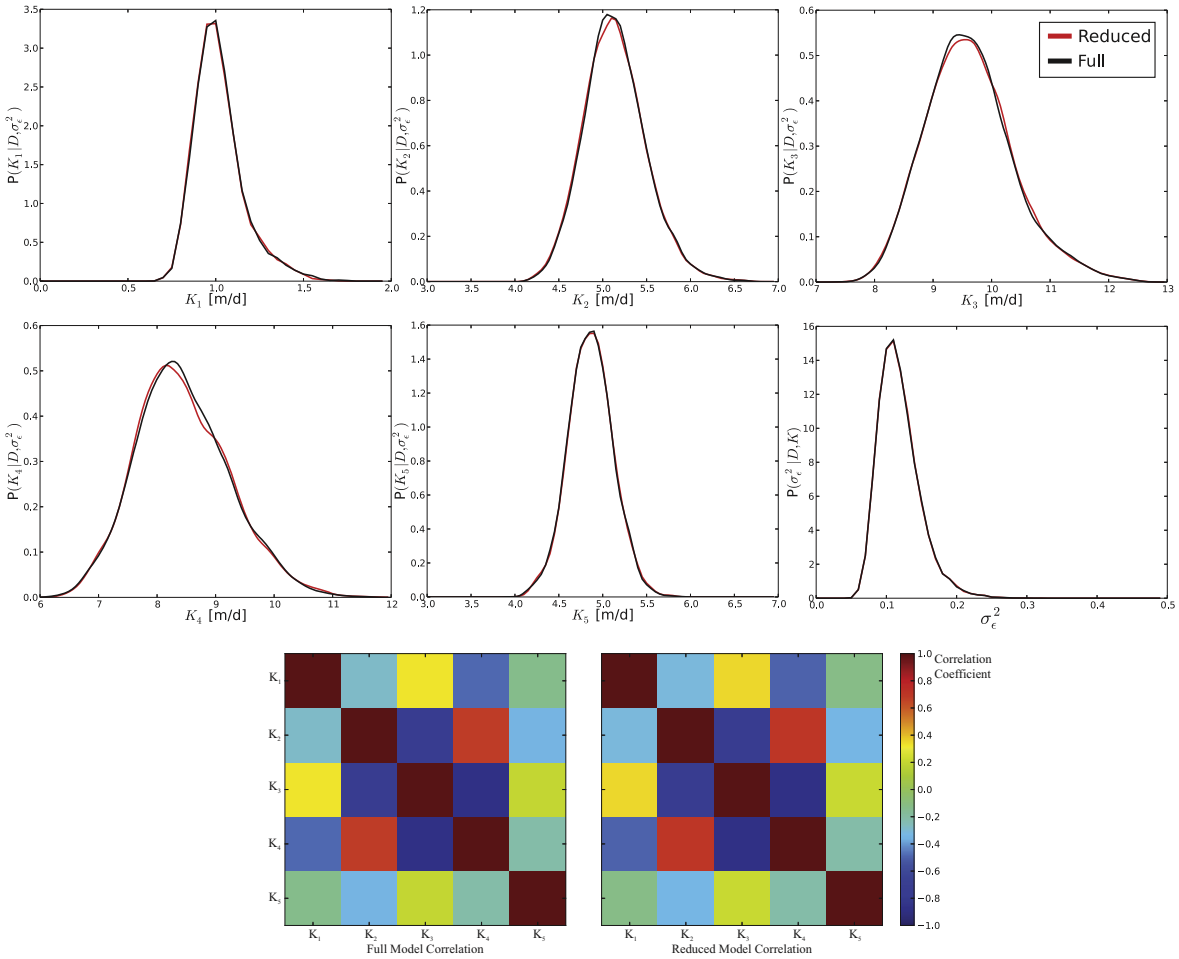

Fig. 5. Reconstruction of the posterior distributions from the one-dimensional, five zone's full and reduced models' Markov chains. From left to right and top to bottom the distributions are for the hydraulic conductivity of zone 1 through 5, then the variance, σ_ϵ^2, and then correllation matricies from the full and reduced models. Note that $K_1,\ldots,K_5$, refers to hydraulic conductivity for zone 1,…,zone 5, respectively.

observation well locations used for demonstration (Z3, Z7 and Z15) [20,32,35]. The number of zones starts at 3, increases to 7, and finally increases to 15. It also is assumed that the hydraulic conductivity in each zone is constant.

The parameter-independent projection matrices for each of the three Oristano zonation patterns are constructed by solving Eqs. (18) and (20), as shown in Fig. 1, with the same convergence criteria and parameter bounds as the one-dimensional case (Section 5.1). The lower and upper bounds for the snapshot times are 10^{-5} d and 10 d, respectively, because the solution reaches a quasi-steady state after five days. The starting hydraulic conductivity is set to 1 m/d for each of the three zonation patterns. Algorithm

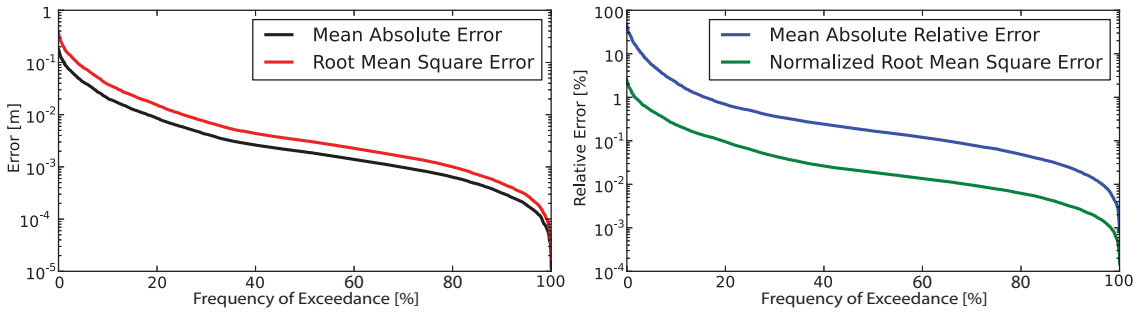

Fig. 6. Exceedance curve of the error statistics, Eq. (23), at 200 days derived from 5000 randomly sampled hydraulic conductivity sets.

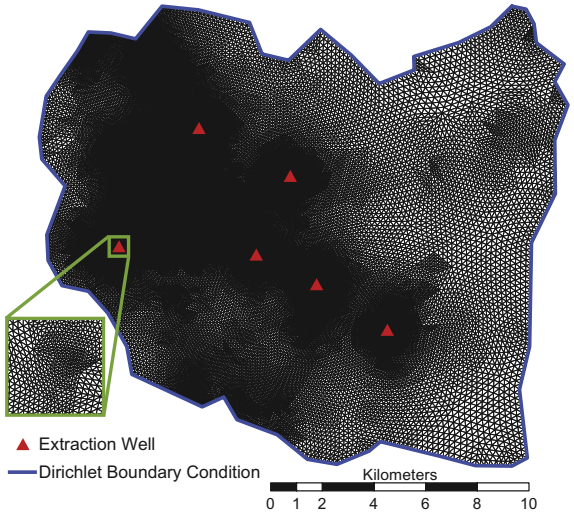

Fig. 7. Oristano model boundary, finite element grid and well locations used in the SAT2D simulation.

2 progressively reduces the objective function value, as seen in Fig. 9, for each *Parameter Greedy Step*, until the convergence criterion is satisfied.

At convergence the 3, 7, and 15 zone models require 69, 127, and 273 snapshots from 8, 15, and 34 parameter sets, respectively. Each parameter set requires between 7 and 12 snapshots for the *Time Greedy Step* to converge. Even though the Z3 Oristano model contains fewer zones than the one-dimensional model (Section 5.1), it requires more parameter sets to achieve convergence due to the added complexity of a two-dimensional model. The actual parameter sets and their corresponding snapshot times for the Z7 model are presented in Fig. 10. The greedy optimal snapshot times tend to start by selecting a quasi-steady state time; then subsequent times decrease until near zero and then select another large simulation time.

Using the same benchmark criteria for the one-dimensional case (Section 5.1) but with a model simulation time of five days, the average runtimes of the full Z3, Z7, and Z15 models are 32.16 s, 33.41 s, 33.15 s, respectively. The average runtimes of the reduced models are 1.82 s, 2.92 s, 5.17 s, respectively, resulting in a reduction in computational time of 94.3%, 91.3%, and 84.4%, respectively. This time savings is very significant given that it would take an estimated 38 days to construct 100,000 sequential samples from the full model to build the Markov chain. The reduced model, while less accurate, is the only computationally reasonable option when many sequential model calls are required.

To demonstrate the applicability of the three Oristano reduced models, we compare the solutions from the reduced models with the solutions obtained from the full model. The "true parameter values" are listed in Table 2. Using these values, a synthetic set of head observations is generated at 24 observation wells, marked on Fig. 8, for 40 exponentially spaced simulation times ranging from 0.04 d to 5 d. The generated observations subsequently are corrupted by adding Gaussian noise with zero mean and 0.1 variance. Using the corrupted observations, we first solve the

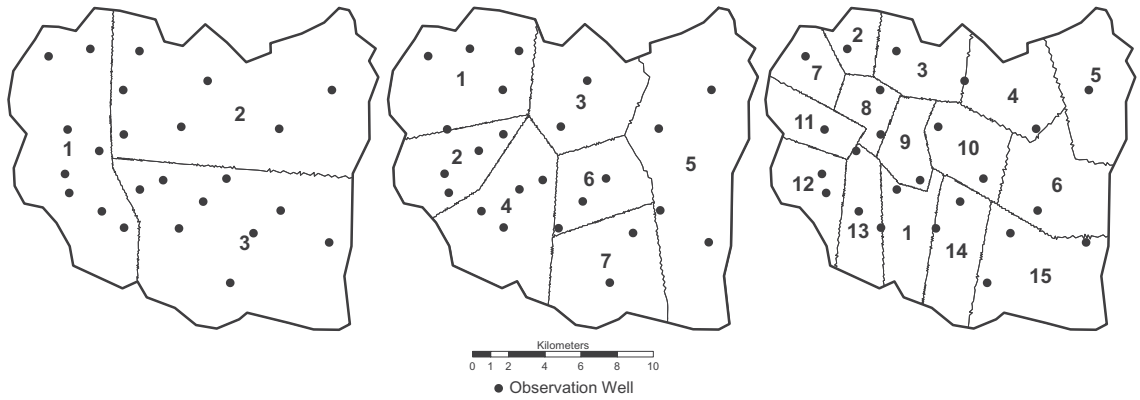

Fig. 8. The three Oristano zonation patterns, (Left) Z3; (Middle) Z7; (Right) Z15. The black dots represent the observation well locations.

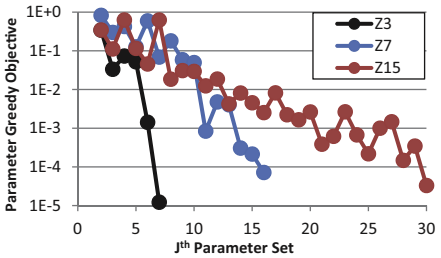

Fig. 9. Semi-log plot of the Oristano *Parameter Greedy Step* objective values for the Z3, Z7, and Z15 model's selected parameter sets.

deterministic inverse problem with Carroll's FORTRAN genetic algorithm driver as described in the one-dimensional case (Section 5.1). The optimized parameter values, shown in Table 2, are used as the starting values for the Bayesian inverse problem. Nearly identical results are obtained with the full model and the reduced model for the Z3 and Z7 cases. The results for the Z15 case show small discrepancies between the reduced model and the full model. Given the large number of zones considered and the limited number of observation locations, this is not unexpected. Results could be improved by increasing the pumping rate to increase drawdown, adding more observation locations or increasing the number observation times, as in Siade et al. [32]. To solve the deterministic inverse problem, the Z3, Z7, Z15 full models, running Carroll's GA in serial, require approximately 200 h to complete. The reduced models required 15, 28 and 57 h, respectively. Note that the reduced models are constructed without any knowledge of the true parameter values other than the upper and lower bounds specified in Eq. (18).

Table 2
Comparison of the solution to the deterministic inverse problem between the full and reduced Oristano models.

Zone number	K_{True}	Full 3 zone	Reduced 3 zone	Full 7 zone	Reduced 7 zone	Full 15 zone	Reduced 15 zone
1	15	15.026	15.025	15.081	15.194	15.574	16.424
2	5	5.075	5.075	5.081	5.077	7.253	7.071
3	7	6.960	6.959	6.932	6.922	6.997	6.899
4	12			12.170	12.090	12.067	13.186
5	3			2.736	2.705	3.464	2.725
6	16			15.585	15.618	14.002	13.055
7	10			10.031	10.041	10.943	12.978
8	2					1.992	1.980
9	9					8.989	8.753
10	18					17.378	17.352
11	17.5					14.115	14.503
12	4.3					4.379	4.377
13	0.5					0.706	0.833
14	16.1					16.154	15.641
15	1					1.001	1.005

The optimized parameters from the deterministic inverse problem are the first value, $K^{(0)}$, in the Metropolis–Hastings algorithm, Algorithm 1, to generate Markov chains from the full and reduced models. The reduce model easily constructs a Markov chain of 100,000 values, but due to the difficulty of a long run time the full Z3 model's chain contains 10,000 samples and the Z7 and Z15 models' chains contain 20,000 samples. All the Markov chains are thinned by keeping every tenth value and comparisons are made using the same length of chains (i.e. first 10,000 samples from the Z3 reduced model are compared to the full version). The posterior variances of the chains, $P(\sigma_\varepsilon^2|K,D)$, from the Z3, Z7, Z15 full and reduced models pass the Levene Equal Variance Test, the Fligner–Killeen Equal Variance Test, ANOVA, and the Kolmogorov–Smirnov 2 Sample Equal Distribution Test at a significance level

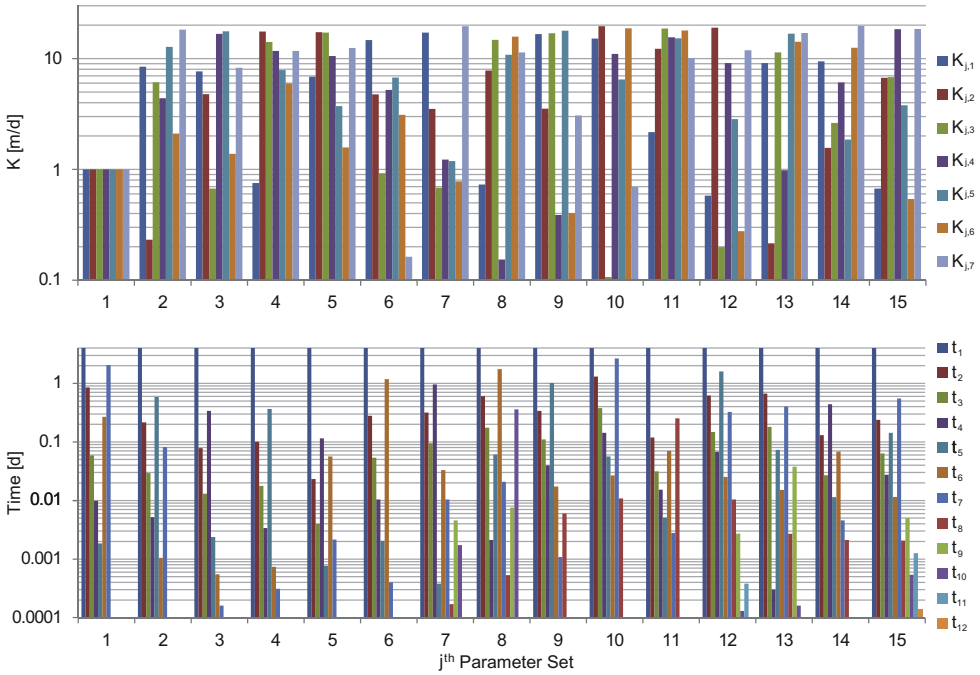

Fig. 10. For the seven zone Oristano model the hydraulic conductivity sets selected at each *Parameter Greedy Step* (top) and their corresponding snapshot times selected by the *Time Greedy Step* (bottom). Note that for all the parameter sets the first snapshot time corresponds to the time when a steady state is reached and is graphically represented as 4 days.

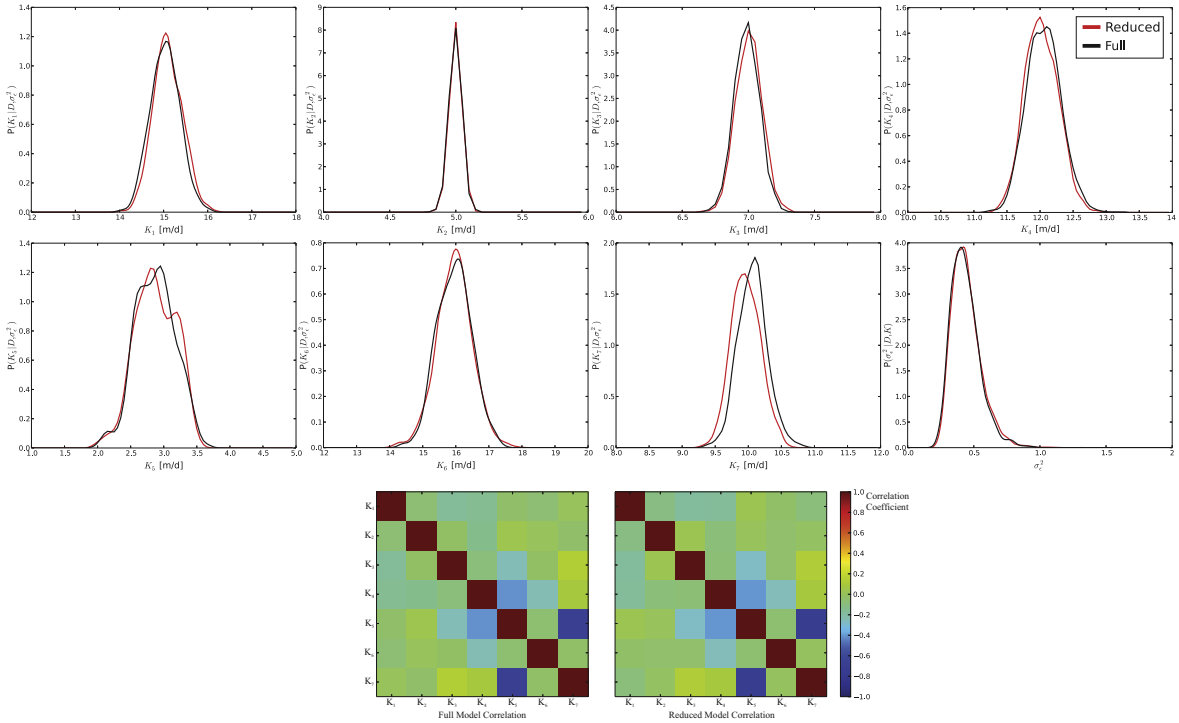

Fig. 11. Reconstruction of the posterior distributions from the Z7 full model's Markov chain and the first 20,000 values of the Z7 reduced model's chain. From left to right and top to bottom the distributions are for the hydraulic conductivity of zone 1 through 7, then the variance, σ_ϵ^2, and then correllation matricies from the full and reduced models. Note that $K_1,\ldots,K_7$, refers to hydraulic conductivity for zone 1,...,7, respectively.

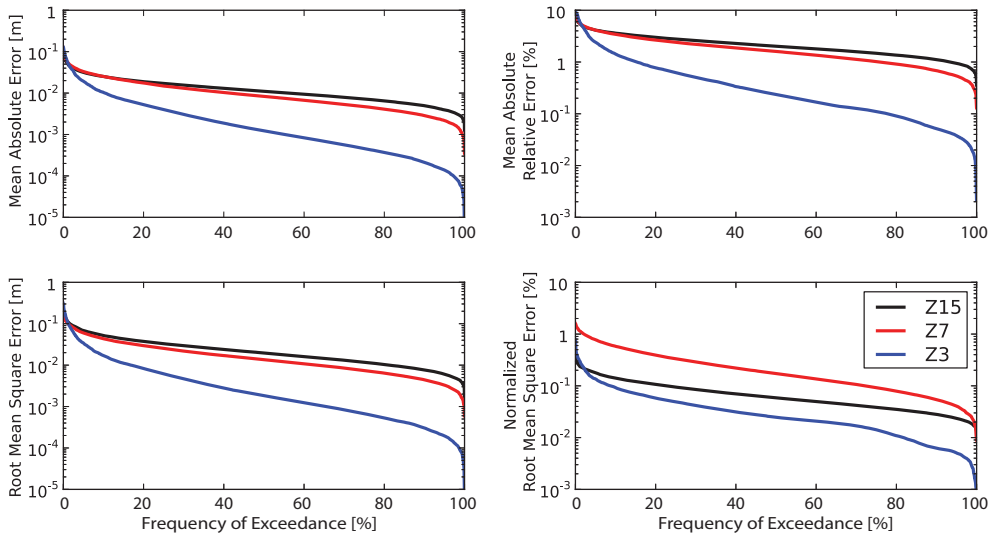

Fig. 12. Exceedance curve of the error statistics, Eq. (23), at 5 days derived from 5000 randomly sampled hydraulic conductivity sets.

of 0.05. The full and reduced Z3 model Markov chains produce a posterior of hydraulic conductivity, $P(K|D,\sigma_\epsilon^2)$, that passes the four aforementioned statistical tests at a significance of 0.05. The Z7 model has three zones that pass the statistical tests at a significance of 0.01 and none of the Z15 pass. This rejection occurs because more complex models require more samples to fully explore their posterior distributions. This is impossible to achieve since it requires 13 days for the Z7 and Z15 full models to generate 20,000 samples. In contrast, it takes 7 and 12 days, respectively, for the reduced models to generate 100,000 samples. Using the Markov chains from the full and reduced models the posterior distributions for the Z7 model are reconstructed, shown in Fig. 11, using

Python's (SciPy) kernel density estimation function. Visually the two distributions compare favorably even though they fail some of the statistical tests.

Despite not passing the statistical tests, the confidence levels produced by the full model and the reduced model are very similar, with relative percent errors between 1% and 3% for most of the 5%, 50%, and 95% confidence interval values. However, the errors for seven of the confidence levels for the Z15 reduced model varied from 6% to 13%. These slightly large errors occur near the lower bound of Eq. (18) or are located far from a pumping well with only one observation location. The reduced models are sensitive to samples drawn beyond the lower bound of $\mathcal{K}$, but can handle any value greater than the upper bound.

To analyze and evaluate the robustness of the Oristano reduced models, we perform the same error analysis tests as described in the one-dimensional case. The last time step (5 d) is assumed to have the greatest error. The full and reduced models are run using 5000 hydraulic conductivity sets (zone 1 through 15) that are randomly sampled from the prior, Eq. (6), using the deterministic inverse problem's solution as the mean and a variance of 100. The error exceedance curves for the Z3, Z7, and Z15 models are presented in Fig. 12. The Z7 and Z15 models have similar MAE, RMSE, and MARE, with the Z15 model slightly less accurate. This is expected since the Z15 reduced model has to reproduce a more complex full model.

6. Conclusion

The key contribution of this paper is its development of a new methodology for assembling snapshots from different sets of hydraulic conductivity that form the basis of a parameter-independent projection matrix. This projection matrix, when applied through the Galerkin projection, can significantly reduce the dimensionality of a transient, confined groundwater model. Most model reduction methods are predicated on the parameter values selected to generate the snapshots. This limits the range and validity of the reduced model when it is used for parameter estimation through either a deterministic or Bayesian inverse solution. The proposed methodology determines the best parameter sets to use to generate snapshots that form a parameter-independent projection matrix. This final projection matrix, composed of snapshots from different parameter sets, is used to reduce the full model such that the reduced model reproduces accurate solutions for the entire parameter space. Thus, the reduced model can be used to solve the Bayesian inverse problems of highly discretized groundwater models (e.g., a model with 57,888 elements and 29,197 nodes). Such problems were not possible to solve in the past because of excessive computational requirements.

The proposed methodology constructs the parameter-independent projection matrix by iteratively selecting parameter sets and their corresponding snapshot times through the use of two greedy algorithms. The parameter sets are selected by solving a bounded optimization problem, Eq. (18), at each greedy stage, called the *Parameter Greedy Step*. In between each Parameter Greedy Step a set of corresponding greedy optimal snapshot times are selected, Eq. (20). At each greedy stage, called the *Time Greedy Step*, a single snapshot time is selected and added to the parameter-independent projection matrix. The bounds of Eq. (18) determine the extent to which the parameter independence works sufficiently. A general rule of thumb is to pick a parameter range based on the prior distribution, using values within three to four standard deviations around its mean.

We validated the proposed methodology on a conceptual one-dimensional, five zone confined aquifer. The constructed projection matrix reduced the dimensionality from 303 to 33 ODEs, resulting in an average savings of 23% (0.08 s) in CPU time. We then tested the reduced model on a synthetic observation set to solve the deterministic inverse and Bayesian inverse problems. The results obtained from the reduced model were compared with the full model solutions. The deterministic inverse solution of the reduced model differed from the full model solution with relative errors from 1% to 5%, while the resulting Markov chains for the Bayesian inverse solutions were found to be statistically equivalent between the reduced model and the full model.

The range and validity of the proposed parameter-independent model reduction methodology was demonstrated by exploring three different large-scale, highly discretized models of a two-dimensional, conceptual, confined aquifer in the Oristano plain of Sardinia, Italy. The three versions (Z3, Z7, and Z15) differed only in their number of hydraulic conductivity zones and all required solving 29,197 ODEs per simulated time step. The proposed methodology reduced the Z3, Z7, and Z15 models to 69, 127, and 273 ODEs, respectively, resulting in an average CPU time savings of 94%, 91%, and 84%, respectively. The reduced model maintains a high level of accuracy for a wide range of parameter values, with the majority of the mean absolute errors 3 cm or less and the mean absolute relative errors 4% or less, as shown in Fig. 12. The time savings is substantial considering that it would take an estimated 38 to 60 days to construct a Markov chain of 100,000 samples from the full Oristano model. In contrast, the Z3, Z7, and Z15 reduced models require 4, 7, and 12 days, respectively, to complete.

Acknowledgments

This material is based on work supported by NSF under awards EAR-0910507 and EAR-1314422, by ARO under award W911NF-10-1-0124, and by an AECOM endowment. This research also was supported in part by the Institute of Digital Research and Education at the University of California, Los Angeles through the Hoffman2 shared computing resources. The authors would like to thank five anonymous reviewers for their in-depth and constructive reviews.

References

[1] Geyer CJ. Practical Markov chain Monte Carlo. Stat Sci 1992;7(4):473–83.
[2] Robert CP, Casella G. Monte Carlo statistical methods. 2nd ed. New York: Springer; 2004. 645p.
[3] Shi X, Ye M, Finsterle S, Wu J. Comparing nonlinear regression and Markov chain Monte Carlo methods for assessment of prediction uncertainty in vadose zone modeling. Vadose Zone J 2012;11(4). http://dx.doi.org/10.2136/vzj2011.0147.
[4] Das NN, Mohanty BP, Efendiev Y. Characterization of effective saturated hydraulic conductivity in an agricultural field using Karhunen–Loève expansion with the Markov chain Monte Carlo technique. Water Resour Res 2010;46(6). http://dx.doi.org/10.1029/2008WR007100.
[5] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. J Chem Phys 1953;21(6):1087.
[6] Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 1970;57(1):97–109.
[7] Hassan AE, Bekhit HM, Chapman JB. Using Markov chain Monte Carlo to quantify parameter uncertainty and its effect on predictions of a groundwater flow model. Environ Model Soft 2009;24(6):749–63. http://dx.doi.org/10.1016/j.envsoft.2008.11.002.
[8] Laloy E, Rogiers B, Vrugt JA, Mallants D, Jacques D. Efficient posterior exploration of a high-dimensional groundwater model from two-stage Markov chain Monte Carlo simulation and polynomial chaos expansion. Water Resour Res 2013;49:2664–82. http://dx.doi.org/10.1002/wrcr.20226.
[9] Zhang G, Lu D, Ye M, Gunzburger M, Webster C. An adaptive sparse-grid high-order stochastic collocation method for Bayesian inference in groundwater reactive transport modeling. Water Resour Res 2013;49. http://dx.doi.org/10.1002/wrcr.20467.
[10] Kunisch K, Volkwein S. Galerkin proper orthogonal decomposition methods for parabolic problems. Numer Math 2001;90(1):117–48. http://dx.doi.org/10.1007/s002110100282.
[11] Vermeulen PTM, Heemink AW, te Stroet CBM. Low-dimensional modeling of numerical groundwater flow. Hydrol Process 2004;18(8):1487–504. http://dx.doi.org/10.1002/hyp.1424.

[12] McPhee J, Yeh WW-G. Groundwater management using model reduction via empirical orthogonal functions. J Water Resour Planning Manage 2008;134(2):161–70. http://dx.doi.org/10.1061/(ASCE)0733-9496(2008)134:2(161).
[13] Baú DA. Planning of groundwater supply systems subject to uncertainty using stochastic flow reduced models and multi-objective evolutionary optimization. Water Resour Manage 2012;26(9):2513–36. http://dx.doi.org/10.1007/s11269-012-0030-4.
[14] Cardoso MA, Durlofsky LJ, Sarma P. Development and application of reduced-order modeling procedures for subsurface flow simulation. Int J Numer Methods Eng 2009;77(9):1322–50. http://dx.doi.org/10.1002/nme.2453.
[15] Robinson BA, Lu Z, Pasqualini D. Simulating solute transport in porous media using model reduction techniques. Appl Math 2012;3(10). http://dx.doi.org/10.4236/am.2012.310170.
[16] Li X, Hu BX. Proper orthogonal decomposition reduced model for mass transport in heterogenous media. Stoch Environ Res Risk Assess 2012:1–11. http://dx.doi.org/10.1007/s00477-012-0653-2.
[17] Buchan AG, Pain CC, Fang F, Navon IM. A POD reduced-order model for eigenvalue problems with application to reactor physics. Int J Numer Methods Eng 2013;95(12):1011–32. http://dx.doi.org/10.1002/nme.4533.
[18] Vermeulen PTM. Inverse modeling of groundwater flow using model reduction. Water Resour. Res. 2005;41(6). http://dx.doi.org/10.1029/2004WR003698.
[19] Lieberman C, Willcox K, Ghattas O. Parameter and state model reduction for large-scale statistical inverse problems. SIAM J. Sci. Comput. 2010;32(5):2523. http://dx.doi.org/10.1137/090775622.
[20] Pasetto D, Putti M, Yeh W-G. A reduced-order model for groundwater flow with random hydraulic conductivity: application to Monte Carlo. Water Resour Res 2013;49:3215–28. http://dx.doi.org/10.1002/wrcr.20136.
[21] Siade AJ, Putti M, Yeh WW-G. Snapshot selection for groundwater model reduction using proper orthogonal decomposition. Water Resour Res 2010;46(8). http://dx.doi.org/10.1029/2009WR008792.
[22] Bear J. Dynamics of fluids in porous media. New York, NY: Dover Publications; 1988. 764p.
[23] Willis R, Yeh WW-G. Groundwater systems planning and management. Englewood Cliffs, NJ: Prentice-Hall; 1987. 416p.
[24] Yeh WW-G. Review of parameter identification procedures in groundwater hydrology: the inverse problem. Water Resour Res 1986;22(2):95. http://dx.doi.org/10.1029/WR022i002p00095.
[25] Daniels MJ. A prior for the variance in hierarchical models. Can J Stat 1999;27(3):567–78. http://dx.doi.org/10.2307/3316112.
[26] Lynch SM. Introduction to applied Bayesian statistics and estimation for social scientists. New York: Springer; 2007. 357p.
[27] Freeze RA. A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media. Water Resour Res 1975;11(5):725. http://dx.doi.org/10.1029/WR011i005p00725.
[28] Vermeulen PTM, Heemink A, Te Stroet C. Reduced models for linear groundwater flow models using empirical orthogonal functions. Adv Water Resour 2004;27(1):57–69. http://dx.doi.org/10.1016/j.advwatres.2003.09.008.
[29] Laub AJ. Matrix analysis for scientists & engineers. Philadelphia: SIAM; 2005. 157p.
[30] Laub AJ. Computational matrix analysis, vol. xiii. Philadelphia: Society for Industrial and Applied Mathematics; 2012. 154p.
[31] Pasetto D, Guadagnini A, Putti M. POD-based Monte Carlo approach for the solution of regional scale groundwater flow driven by randomly distributed recharge. Adv Water Resour 2011;34(11):1450–63. http://dx.doi.org/10.1016/j.advwatres.2011.07.003.
[32] Siade AJ, Putti M, Yeh WW-G. Reduced order parameter estimation using quasilinearization and quadratic programming. Water Resour Res 2012;48(6). http://dx.doi.org/10.1029/2011WR011471.
[33] Gambolati G, Putti M, Paniconi C. Three-dimensional model of coupled density-dependent flow and miscible salt transport in groundwater. In: Bear J et al., editors. Seawater intrusion in coastal aquifers: concepts, methods, and practices. Dordrecht, Netherlands: Kluwer Acad.; 1999. p. 315–62.
[34] Scott DW. Multivariate density estimation: theory, practice, and visualization, vol. 383. John Wiley & Sons; 2009.
[35] Cau PL, Lecca G, Putti M, Paniconi C. The influence of a confining layer on saltwater intrusion and surface recharge and groundwater extraction conditions. In: Hassanizadeh SM et al., editors. Computational methods in water resources. Developments in water resources, vol. 1. Amsterdam, Netherlands: Elsevier; 2002. p. 493–500.
[36] Ushijima TT, Yeh WWG. Experimental design for estimating unknown groundwater pumping using genetic algorithm and reduced order model. Water Resour Res 2013;49(10):6688–99. http://dx.doi.org/10.1002/wrcr.20513.

Chapter 3

Advances in Water Resources 83 (2015) 250–262

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Reduced order modeling of the Newton formulation of MODFLOW to solve unconfined groundwater flow

Scott E. Boyce [a,b], Tracy Nishikawa [a], William W-G. Yeh [b,*]

[a] U.S. Geological Survey, California Water Science Center, 4165 Spruance Rd., Suite 200, San Diego, CA 92101-0812, USA
[b] Department of Civil and Environmental Engineering, University of California, 5732B Boelter Hall, Los Angeles, California, 90095, USA

ARTICLE INFO

Article history:
Received 27 February 2015
Revised 12 June 2015
Accepted 13 June 2015
Available online 17 June 2015

Keywords:
Model reduction
Nonlinear model reduction
Proper orthogonal decomposition
Unconfined groundwater flow
MODFLOW

ABSTRACT

Projection-based model reduction techniques have been shown to be very effective for reducing the computational burden of high-dimensional groundwater simulations, but only applied to confined groundwater flow. A new methodology is proposed that reduces the dimension of a discretized, transient, unconfined groundwater-flow model. This unconfined model reduction technique is based on Galerkin projection and the Newton formulation of MODFLOW. The method is implemented following the standard package design and code structure that MODFLOW employs for all its features. When the package is invoked within MODFLOW it can collect snapshots, produce a basis, construct the reduced model and propagate the reduced model forward in time. The new formulation accurately represents the water-table surface under a variety of nonlinear settings, such as intraborehole flow from a Multi-Node Well. The unconfined model reduction is applied to four test cases to illustrate its flexibility in handling nonlinear features. Several test cases are discussed to demonstrate the unconfined model reduction applicability. The final test case applies the new model reduction methodology to a scoping MODFLOW model of Santa Barbara, CA composed of 113,578 cells, which requires solving 113,578 equations per time step, and reduces it to 127 equations.

1. Introduction

Understanding of groundwater resources is enhanced through the application of mathematical models that simulate the dynamics of an aquifer system. Conducting advanced analyses such as optimization of pumping schedules under different scenarios requires a large number of simulations. Such analyses are intractable for complex, highly-discretized models with large computational requirements. Reducing the computational burden associated with these simulation models provides the opportunity to perform more advanced analyses on a wider spectrum of groundwater management problems.

Projection-based model reduction techniques via the Galerkin method have been shown to be very effective for reducing the computational burden of large-scale, linear, groundwater simulations [1–4]. This type of model reduction involves construction of a basis spanned by solutions, referred to as snapshots, from the original simulation model. To form the basis, selected snapshots are collected as a snapshot set and analyzed with Proper Orthogonal Decomposition (POD). POD is a statistical pattern analysis technique that extracts a low-dimensional basis from a high-dimensional process, such as the snapshot set [5]. The discrete form of POD is identical to singular value decomposition where the POD basis, which is the Galerkin projection matrix, are the left singular vectors and the proper orthogonal modes are the right singular vectors [6,7]. POD has also been referred to as Principal Component Analysis [8], Karhunen–Loève Expansion [9], and Empirical Orthogonal Functions [10,2]. The selection of the snapshots that compose the snapshot set is an area of active research, but not the focus of this paper. For more detailed information regarding snapshot selection see Boyce and Yeh [4] and Siade et al. [11].

Currently, most model reduction applications are applied to a linear, confined, groundwater model and demonstrate a reduction in dimensionality by several orders of magnitude. Vermeulen et al. [1] applied model reduction to the confined, groundwater-flow equation by collecting from a simulation model a snapshot set of simulated hydraulic heads at a constant, reference pumping rate. Vermeulen et al. [1] then applied principal component analysis (PCA) to the snapshot set to generate the basis used by the Galerkin method to derive the reduced model. McPhee and Yeh [2] followed this methodology and demonstrated that a reduced model maintains its sensitivity of hydraulic head with respect to pumping, enabling it to embed in a management optimization problem. Baú [12] increased the utility of POD model reduction by deriving a set of reduced models for each Monte Carlo realization of hydraulic conductivity to solve a stochastic, multi-objective, confined groundwater management problem.

* Corresponding author. Tel.: 3108252300; fax: 3108257581.
E-mail addresses: seboyce@usgs.gov (S.E. Boyce), tnish@usgs.gov (T. Nishikawa), williamy@seas.ucla.edu (W.W-G. Yeh).

http://dx.doi.org/10.1016/j.advwatres.2015.06.005
0309-1708/

In theory, the Galerkin projection model reduction technique only applies to linear systems, such as confined aquifers. Application to nonlinear systems is possible, but the reduced model's error is greater and may require significantly more basis functions (snapshots) to characterize the model [13]. The basis functions derived from POD can be thought of as tangent spaces that approximate the nonlinear system [5]. Robinson et al. [14] and Li and Hu [15] applied POD model reduction to several synthetic one- and two-dimensional mass-transport models without chemical reactions. Buchan et al. [16] solved for the population growth of free moving neutrons, an Eigenvalue problem, in a nuclear reactor system. The Eigenvalue problem was reformulated to create pseudo-time dependence that describes the snapshots used in their projection basis. With these nonlinear examples there has been limited application of model reduction to the nonlinear, unconfined groundwater flow equations.

Simulating unconfined groundwater flow is an essential part of groundwater modeling projects. MODFLOW [17] and its variants are one of the most popular groundwater modeling software packages available that solve the confined and unconfined equations. MODFLOW is designed as a series of independent modular pieces called packages. To solve the governing equations, MODFLOW uses a three-dimensional, block-centered finite difference scheme to discretize the model domain and implicit finite differencing to discretize time. The nonlinear system of equations is solved using the Picard method. To improve stability of solving the MODFLOW unconfined flow equations the Newton method has been shown to be a useful alternative to the Picard method [18,19]. This led to the development of the Newton formulation of MODFLOW, MODFLOW-NWT [20], which expands upon previous work with upstream weighting for determining horizontal intercell conductance, conductance smoothing functions, and a Newton solver. The two main packages for the Newton formulation of MODFLOW are the flow package Upstream Weighting (UPW) and solver package Newton Solver (NWT). Conductance is MODFLOW's finite difference representation of hydraulic conductivity between two modeled grid cells and is derived from Darcy's Law and finite differences. While the Newton solver is more stable than the Picard method, it requires solving for the inverse of the nonsymmetric, Jacobian matrix for each Newton step (inner iteration).

Solving unconfined models with a high discretization through the Newton method can become computationally intractable due the difficulty of solving for the inverse of the nonsymmetric Jacobian. To reduce the computational burden, and thus the simulation runtime, we propose an approach that combines the Newton formulation of MODFLOW [20] with projection-based model reduction. This approach is implemented following the standard package design and code structure that MODFLOW employs for all its features. The source code modifications are implemented within a recent release of MODFLOW called MODFLOW One-Water Hydrologic Flow Model (MF-OWHM) [21]. The model-reduction package initially collects snapshots from a MF-OWHM simulation and converts them to a basis. Once the basis is constructed, the package performs a reduced model simulation by intercepting the NWT package where it solves for the inverse of the Jacobian. The reduced model uses Galerkin projection to form a reduced dimension Jacobian, solves for the inverse of the reduced Jacobian, and then calculates the Newton step.

2. Groundwater-flow modeling

2.1. Discretized groundwater-flow modeling

Darcy's Law is combined with the continuity equation to derive the governing equation for anisotropic, saturated groundwater flow. The resulting confined and unconfined groundwater flow equations are the following [19,22,23]:

$$\nabla \cdot (K h \nabla h) \pm Q = S_s \frac{\partial h}{\partial t} \tag{1}$$

$$\nabla \cdot (K h \nabla h) \pm Q = S_y \frac{\partial h}{\partial t} \tag{2}$$

where $\nabla \cdot$ is the divergence operator, ∇ is the gradient operator, K is the hydraulic conductivity tensor [L/T], h is the hydraulic head [L], Q is a volumetric flux per unit volume in or out of the system [T^{-1} for Eq. 1 and L/T for Eq. 2], S_S is the specific storage [L^{-1}], t is the time [T], and S_y is the specific yield [–]. Within a three-dimensional groundwater system, Eq. (1) applies to regions under confined conditions (fully saturated model layers) and Eq. (2) applies to regions under unconfined conditions (not fully saturated model layers). During the course of a model simulation a particular model cell can switch from confined (Eq. 1) to unconfined (Eq. 2) conditions and vice-versa. For unconfined conditions the additional h variable, representing the saturated thickness, makes the equation nonlinear and acts as an upper boundary condition created by the water table.

Applying a spatial discretization to Eqs. (1) and (2) through a differencing scheme (e.g. finite differences or finite elements) yields a system of nonlinear ordinary differential equations (ODEs) of the form:

$$\mathrm{A}h + \mathrm{f} = \mathrm{B}\frac{dh}{dt} \tag{3}$$

where $h \in \mathbb{R}^n$ is a vector of hydraulic heads at each nodal location; $\mathrm{A} \in \mathbb{R}^{n \times n}$ is a nonlinear stiffness matrix, which contains spatial information about hydraulic conductivity and is a function of h (i.e. the saturated thickness) for model cells that are under unconfined conditions; $\mathrm{B} \in \mathbb{R}^{n \times n}$ is the mass matrix; and $\mathrm{f} \in \mathbb{R}^n$ is a vector containing all Dirichlet, Neumann and part of the Cauchy boundary conditions. To solve for the time derivative an implicit difference scheme is applied to the time derivative to form the following system of nonlinear equations:

$$(\mathrm{B} - \Delta t \mathrm{A}) h^{t+1} = \mathrm{B}h^t + \Delta t \mathrm{f} \tag{4}$$

where t is the time step index and $\Delta t = t^{t+1} - t^t$ is the time step that propagates the model forward to a new time. Nonlinearity arises because matrix A is a function of h for layers that are not fully saturated.

2.2. Newton formulation of the groundwater-flow equation

The Newton formulation, as described in Niswonger et al. [20] and Painter et al. [18], rearranges Eq. (4) into the following:

$$(\mathrm{B} - \Delta t \mathrm{A}^i) h^{t+1,i} - \mathrm{B}h^t - \Delta t \mathrm{f} = \mathrm{R}^i \tag{5}$$

$$J^i \Delta h = \mathrm{R}^i \tag{6}$$

$$J^i h^{t+1,i+1} = \mathrm{R}^i + J^i h^{t+1,i} \tag{7}$$

where i is the Newton step index, A^i is the stiffness matrix evaluated at $h^{t+1,i}$, $h^{t+1,i}$ is the head at the ith Newton step for the $t+1$ th time step, $\mathrm{R}^i \in \mathbb{R}^n$is the residual that results from an estimate of $h^{t+1,i}$, J^i is the Jacobian matrix defined as $J^i = \partial \mathrm{R}^i / \partial h^{t+1,i} \in \mathbb{R}^{n \times n}$, and $\Delta h = h^{t+1,i+1} - h^{t+1,i}$. The Newton procedure then iteratively solves for $h^{t+1,i+1}$ until $\mathrm{R}^i \to 0$.

The original MODFLOW approach removes dewatered cells from the finite difference equations and treats them as a no-flow boundary. A cell is dewatered when its hydraulic head (h) fall below its bottom elevation. Unlike the original version of MODFLOW, dewatered model cells are not removed in the Newton formulation of MODFLOW. Due to this design, the reduced model must reproduce the results from all the model cells, including the dewatered ones.

3. Model reduction via POD and the Galerkin projection for the Newton formulation of groundwater-flow

3.1. The Galerkin projection matrix

Projection-based model reduction proceeds by projecting a discretized groundwater-flow model, $h \in \mathbb{R}^n$, onto a reduced model subspace of $h_r \in \mathbb{R}^r$, where $r \ll n$. For this to occur, it is assumed that the

hydraulic head, h, can be mapped linearly onto the reduced space, h_r, by some orthonormal matrix,$\mathrm{P} \in \mathbb{R}^{n\times r}$, which contains information about the discretized groundwater-flow model's solutions in time [1,4]:

$$h = \mathrm{P}h_r \tag{8}$$

The vectors in the P matrix are sometimes referred to as the POD basis, spatial basis functions, empirical orthogonal functions, or principal vectors [4,8–10].

The matrix P is derived from a time series of solutions, h^{t_k}, called snapshots, that are collected from the groundwater-flow model under a constant/reference forcing (e.g. pumping well) taken at specific simulation times, t_k. A set of snapshots must be taken for each linearly independent forcing term in the groundwater-flow model. For example, if a model has two extraction wells that have no linear relation between them, then two sets of snapshots must be taken. If there are any linear relationships (e.g. Well A = 2 × Well B), then a single set of snapshots can be taken at that relationship. All the sets of snapshots are then combined to form the final snapshot set:

$$\begin{aligned} &S^j = [h^{t_1}, h^{t_2}, h^{t_3}, \ldots, h^{t_k}] \qquad \forall j = 1, 2, \ldots, \mathrm{NSET} \\ &S = [S^1, S^2, \ldots, S^{\mathrm{NSET}}] \in \mathbb{R}^{n\times m} \end{aligned} \tag{9}$$

where S^j is a set of snapshots collected for the jth linearly independent forcing and is composed of k snapshots, NSET is the total number of linearly independent forcings, and S is the final snapshot set containing m snapshots.

For each snapshot set, S^j, the number of snapshots, k, and their specific solution time, t_k, are problem-dependent. Both the number of snapshots and specific solution times do not have to be the same for each snapshot set, S^j. From the literature and empirical tests using this study's methodology, each snapshot set, S^j, should be composed of snapshots that have 10–50 exponentially distributed values of t_k from the initial condition to quasi-steady state with the forcing set to a constant value and then another 10–50 snapshots with no forcing to adequately capture the model's nonlinear response and recovery. For simplicity, a set of exponentially distributed snapshots times are selected by visually examining the model output. For a more systematic approach to estimate the optimal number of snapshots, see Boyce and Yeh [4].

The snapshots from Eq. (9) form a basis that, when made orthonormal, derives the Galerkin projection operator. This operator is the same matrix as the linear map, P, defined in Eq. (8). The orthonormalization of the snapshots, S, is performed by POD, which for the discrete case is the Singular Value Decomposition (SVD) [24,25]:

$$S = \mathrm{U}\Sigma\mathrm{V}^{\mathrm{T}} = \sum_{i=1}^{\mathrm{rank}(S)} \sigma_i u_i v_i^{\mathrm{T}} \Rightarrow \mathrm{P} = \mathrm{U} \tag{10}$$

where rank(S) is the minimum between the number of linearly independent columns or rows in S, σ_i are the singular values that are ordered from largest to smallest, Σ is a diagonal matrix with singular values along the diagonal, $\mathrm{U} = [u_1, u_2, \ldots]$ are the left singular vectors that comprise the POD basis set equal to P, and $\mathrm{V} = [v_1, v_2, \ldots]$ are the right singular vectors that represent the proper orthogonal modes. For the test cases presented in Sections 4 and 5 the SVD is performed with the LAPACK library [26].

Without knowledge of the optimal choice for snapshots, the singular values can serve as weights to distinguish the left singular vectors, subsequently referred to as POD basis vectors that contribute the most amount of information to the projection operator P. This is accomplished by solving for r in the following relationship:

$$\begin{aligned} &\min \quad r \\ &\text{Subject to:} \quad \frac{\sum_{i=1}^{r}\sigma_i}{\sum_{i=1}^{\mathrm{rank}(S)}\sigma_i} \times 100\% \geq \mathrm{PE} \end{aligned} \tag{11}$$

where PE is a user-specified minimum required percent energy contribution of the sum of the r largest singular values. Typical values for PE range from 99% to 99.99%, where smaller values lead to fewer basis vectors at a cost of less accurate reduced model results. Once r has been established, the final truncated POD basis is $\mathrm{P} = [u_1, u_2, \ldots, u_r]$, which defines the dimension of the reduced model, h_r (Eq. 8). This process of removing insignificant basis vectors is called singular-value truncation; if S is symmetric, then it is also called principal component analysis.

There are two methods that can improve the selection of basis vectors before applying SVD to construct the POD basis. The first method involves normalizing each column in S (i.e., each snapshot) by dividing it by its respective L^2 norm. Normalization produces a better spread of singular values by removing any issues of scale from the snapshots. The second method is to mean center and redefine S as follows:

$$\begin{aligned} &\bar{h} = \frac{1}{m}[S\mathbf{1}_m] \\ &S := S - \bar{h}\mathbf{1}_m^{\mathrm{T}} \end{aligned} \tag{12}$$

where $\mathbf{1}_m$ is a column vector of m ones, $\bar{h} \in \mathbb{R}^n$ contains the mean for each row in S. If the snapshots are mean centered, then Eq. (8) must be adjusted to include the missing information from the POD basis (P), i.e.,

$$h = \mathrm{P}h_r + \bar{h} \tag{13}$$

The advantage of mean centering is that it removes all time-invariant effects from the snapshots and can stabilize problems that may arise from boundary conditions. If both normalization and mean centering are required, it is important to first apply the mean centering and then normalize the redefined S to preserve the relationship defined in Eq. (13). For most situations the reduced form of the Newton formulation of MODFLOW (Section 3.2) produces the most accurate reduced models with both mean centering and normalization. Of the two improvements, our test cases indicate that mean centering has the greatest effect on improving accuracy and normalization only results in a minor improvement.

3.2. The Galerkin projection and the Newton formulation of MODFLOW

The Newton formulation of MODFLOW [20] is projected onto a reduced space via the Galerkin projection and a substitution that makes the system of equations dimensionally correct. The Galerkin projection involves pre-multiplying Eq. (7) by the transpose of the POD basis (P^{T}). The resulting system of equations now requires one of two possible substitutions to make it dimensionally correct. If mean centering is not applied, then substituting in $h^{t+1,i+1} = \mathrm{P}h_r^{t+1,i+1}$ is required to form Eq. (14). If there is mean centering, then substitute $h^{t+1,i+1} = \mathrm{P}h_r^{t+1,i+1} + \bar{h}$ to form Eq. (15).

$$\mathrm{P}^{\mathrm{T}}J^i\mathrm{P}h_r^{t+1,i+1} = \mathrm{P}^{\mathrm{T}}[\mathrm{R}^i + J^i h^{t+1,i}] \tag{14}$$

$$\mathrm{P}^{\mathrm{T}}J^i\mathrm{P}h_r^{t+1,i+1} = \mathrm{P}^{\mathrm{T}}[\mathrm{R}^i + J^i h^{t+1,i} - J^i\bar{h}] \tag{15}$$

Let $J_r^i = \mathrm{P}^{\mathrm{T}}J^i\mathrm{P} \in \mathbb{R}^{r\times r}$ and $b_r^i = \mathrm{P}^{\mathrm{T}}[\mathrm{R}^i + J^i h^{t+1,i}] \in \mathbb{R}^r$, or if there is mean centering, $b_r^i = \mathrm{P}^{\mathrm{T}}[\mathrm{R}^i + J^i h^{t+1,i} - J^i\bar{h}] \in \mathbb{R}^r$, then Eqs. (14) and (15) are then transformed to:

$$J_r^i h_r^{t+1,i+1} = b_r^i \tag{16}$$

Eq. (16) is subsequently referred to as the reduced model and Eq. (7) as the full model. The reduced model only requires the solution of r equations for each Newton step compared to the full model's n equations. After solving for $h_r^{t+1,i+1}$ for one Newton step either Eq. (8) (no mean centering) or Eq. (13) (mean centering) provide a means of obtaining $h^{t+1,i+1}$. For the reduced model to be efficient, it is necessary that $r \ll n$.

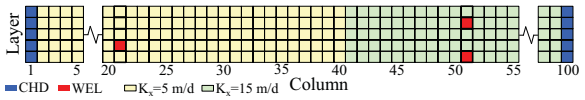

Fig. 1. Two-dimensional, vertical test case, TC1, with three wells (WEL) and time-variant specified-head (CHD) boundary conditions.

4. Two-dimensional synthetic, vertical test cases

A synthetic, unconfined, two-dimensional (2D), vertical groundwater-flow model is constructed in MF-OWHM [21] to demonstrate the application of model reduction to the Newton formulation of MODFLOW. The modeled domain contains 500 simulated cells with a total length (column direction) of 10 km and depth of 50 m. The model's top and bottom elevations are set to 0 m and –50 m, respectively. The initial condition is set to a uniform water level of –2 m to force an unconfined condition. The model discretization consists of five, 10 m-thick, convertible layers, one 100 m-wide row, and one-hundred, 100 m-long columns. A convertible layer in MF-OWHM uses, where appropriate, the confined (Eq. (1)) and unconfined (Eq. (2)) equations. The domain is divided into two zones with columns 1 through 40 and columns 41 through 100 having horizontal hydraulic conductivity values of 5 m/d and 15 m/d, respectively. The vertical hydraulic conductivity is set to be one-tenth that of the horizontal. The entire model domain has a specific yield (S_y) of 0.3 and specific storage (S_s) of 1E–5 m^{-1}. Using the aforementioned properties we evaluate two different test cases using different stresses and boundary conditions.

4.1. Three wells and fixed head boundary condition

The first test case, TC1, applies model reduction to a 2D groundwater-flow model that uses MODFLOW's well (WEL) and time-variant specified-head (CHD) packages (Fig. 1). The WEL package simulates pumping into or out of the model and CHD applies a time varying fixed-head boundary condition [23]. The three WEL locations are at layer 2, column 51; layer 4, column 21; and layer 5, column 51 (Fig. 1) and are subsequently identified only by their layer. All model cells that are specified with the CHD package are set to a constant value of –2 m to maintain unconfined conditions throughout the simulation.

Three sets of 90 snapshots are collected (NSET = 3) to construct the reduced model. Each snapshot set, j, is taken with one of the wells set to a pumping rate of 400 m^3/d for 50,000 days, the quasi-steady state time, followed by 50,000 days of recovery (no pumping). Taking snapshots for each well allows their values to vary independently within the reduced model. The distribution of snapshot times, t_k, are shown in Fig. 2, where the 46th snapshot is the first one taken without pumping to record the model's recovery. The three sets of 90 snapshots are combined to form $S \in \mathbb{R}^{500\times270}$, which is subsequently mean centered and normalized. Using a PE of 99.99% results in a final POD basis composed of 61 vectors. This basis then reduces the Newton method's system of equations (Eq. (7)) from 500 equations to 61 (Eq. (16)). Further reduction could have been achieved if it was assumed that two or more of the wells pumped at the same rate. Then their snapshots could have been combined, resulting in a smaller final basis.

The simulation time frame used to compare the full and reduced models is 3000 days with pumping followed by 3000 days of recovery with a time step of 10 days. The pumping rate is set to 1000 m^3/d for each well to illustrate that the reduced model is independent of the pumping rate that was used for the snapshots (i.e., 400 m^3/d).

To evaluate the reduced model's accuracy the following equations are defined:

$$\begin{aligned} \mathrm{MAE} &= (1/\mathrm{N}) \textstyle\sum_{i=1}^{\mathrm{N}} |h_i - (\mathrm{P}h_r)_i| \\ \mathrm{RMSE} &= \sqrt{(1/\mathrm{N}) \textstyle\sum_{i=1}^{\mathrm{N}} (h_i - (\mathrm{P}h_r)_i)^2} \\ \mathrm{NRMSE} &= 100\% \times \mathrm{RMSE}/(h_{\max} - h_{\min}) \end{aligned} \tag{17}$$

where MAE is the mean absolute error, RMSE is the root mean squared error, NRMSE is the normalized root mean squared error, N is total the number of model cells, i is the model cell index (e.g., row 1, column 1, layer 1 has index of 1), and $h_{\max}$ and $h_{\min}$ are the maximum and minimum head values across all model cells, respectively. The NRMSE weights the RMSE by the spread of information to provide a method of comparing accuracy of different models. The largest value of NRMSE from all the time steps is 0.075%. Fig. 3 presents a histogram of the MAE calculated at each time step and an error exceedance curve of the RMSE for TC1. The MAE errors are clustered around 0.007 m (Fig. 3, left) with the largest values being 0.015 m. The reduced model replicates the full model with all the errors less than 0.02 m. The error exceedance curve demonstrates that the RMSE is very small with a 5% chance of being 0.016 m or greater. In addition, the full and reduced models' cumulative percent mass balance errors at the end of the simulation are 0% and 0.39%, respectively.

The water table is calculated as the hydraulic head of the upper most model cells that are not dewatered. Fig. 4 presents the simulated water-table elevation for the full and reduced models at selected times during the 6000 day simulation. In the figure the model cells that are above the water table have hydraulic heads below their bottom cell elevation; conversely cells below the water table have hydraulic heads above their top elevation. During the first 3000 days

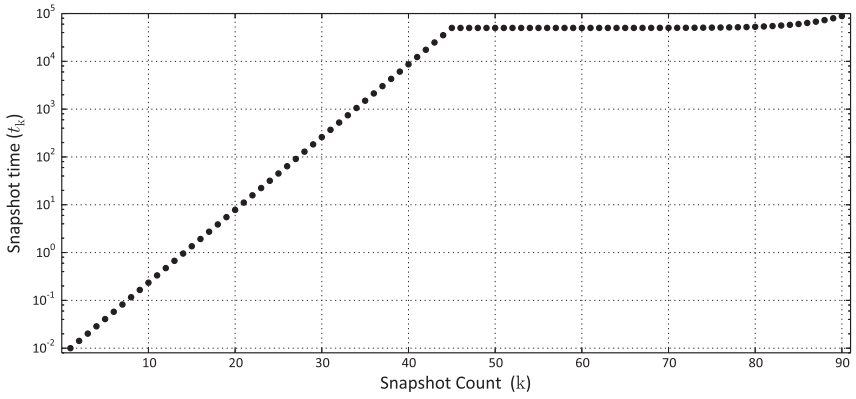

Fig. 2. Distribution of snapshot times. The first 45 are taken with pumping and the remainder without.

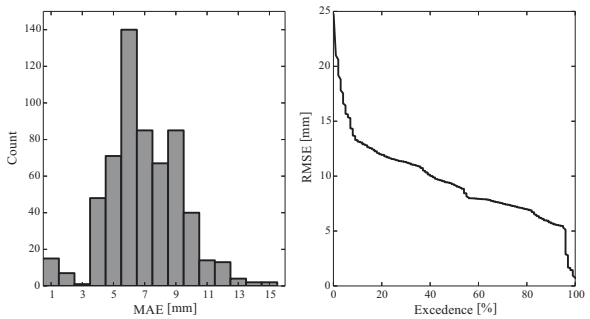

Fig. 3. TC1 (left) histogram of the mean absolute errors from all time steps and (right) exceedance curve of the RMSE. Note that the units are in mm.

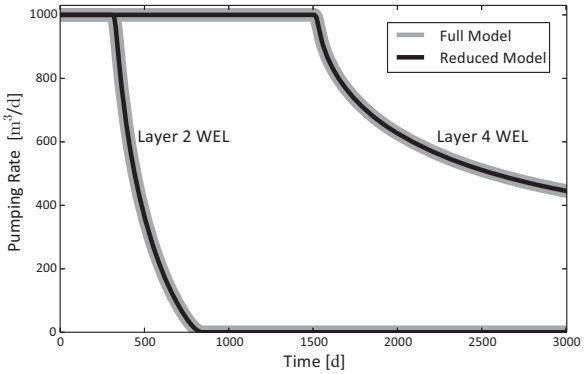

Fig. 5. Two-dimensional test case, TC1, full and reduced model's loss of pump capacity due to drawdown.

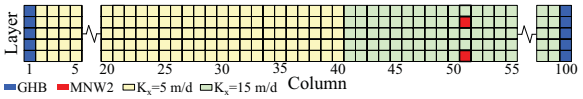

Fig. 6. Two-dimensional test case, TC2, with one multi-node well with two screened intervals (MNW2) and general head boundary conditions (GHB).

of simulation (Fig. 4, left) the pumping well removes water causing the water table to decline below layers 1, 2, and 3 within most of the model domain. This illustrates that despite multiple layers being dewatered, the reduced model accurately represents the full model. During the next 3000 days (Fig. 4, right) there is no pumping and the reduced model still accurately reproduces the nonlinear recovery of the water table.

Unique to the Newton formulation of MODFLOW is that a WEL model cell simulates loss of production due to dewatering by reducing the pumping rate as a function of saturation in the pumping cell. The pumping wells located in layers 2 and 4 (Fig. 1) undergo this reduction in rate during the first 3000 days of pumping (Fig. 5). Due to dewatering of the cell, the well located at layer 2 completely stops production after 850 days. The reduced model follows the same trend and visually matches the full model's reduction in pumping as illustrated in Fig. 5.

4.2. One multi-node well and a general head boundary condition

The second test case, TC2, applies model reduction to a two-dimensional model that uses MODFLOW's revised Multi-Node Well (MNW2) [27,28] and the General-Head Boundary (GHB) packages, shown in Fig. 6. The MNW2 package [28] simulates wells that extend over more than one model cell and is a more realistic representation of a groundwater pumping well than the WEL package. The MNW2 package calculates the head level within the well; its possible production rate is based on a desired flow rate and allows for intraborehole flow (i.e., flow within the well itself). When there is no desired flow rate specified, an MNW2 well can transmit water between layers via intraborehole flow. The location of the MNW2 well is at column 51 with a screened interval from –10 m to –20 m and –40 m to –50 m (layers 2 and 5, respectively, in Fig. 6). The MNW2 well in TC2 has the losstype set to "skin" [28], a radius of 0.15 m (inner radius), a skin radius of 0.2 m (outer radius). The hydraulic conductivity of the skin material is 1.5 m/d.

The GHB package simulates head-dependent flux boundaries by allowing water to flow into or out of the system based on the water level and a user-specified conductance. For TC2, the GHB boundary head (B_{head}) is set to –2 m and the conductance assigned to columns 1 and 100 is calculated such that it represents the model grid's hydraulic conductivity.

To construct the reduced model, one set of 90 snapshots is collected (NSET = 1). The first 45 snapshots are taken with the MNW2 well set to a desired pumping rate (Qdes) of 400 m^3/d for 50,000 days—the quasi-steady state time, to capture the response to MNW2 pumping. After the quasi-steady state is reached, 45 more snapshots are taken over 50,000 days with the MNW2 well's desired pumping rate set to 0 m^3/d to capture the recovery from pumping and any associated intraborehole flow. The distribution of snapshot times, t_k, are

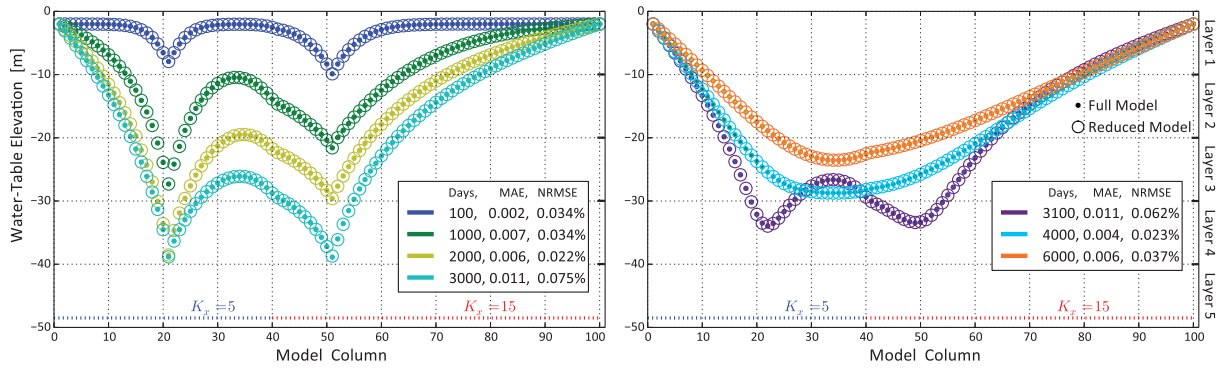

Fig. 4. Two-dimensional test case, TC1, simulated water-table elevation during the first 3000 days of pumping (left) and the next 3000 days of recovery (right) at selected times for the full and reduced models. Note that MAE and NRMSE are calculated from the simulated head at all 500 model cells.

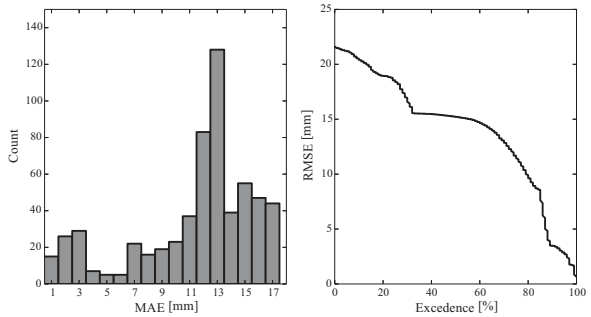

Fig. 7. TC2 (left) histogram of the mean absolute errors from all time steps and (right) exceedance curve of the RMSE. Note that the units are in mm.

the same as used in TC1 (Fig. 2). The TC2 snapshots from the MNW2 well are then mean centered and normalized. Using a PE of 99.99% results in a final POD basis composed of 22 vectors for reducing the Jacobian. The TC2 basis is smaller than the TC1 basis because the former is able to exploit the fact that the nodes within the MNW2 well are related while the latter assumes no relationship between the three wells.

Like TC1, the simulation timeframe used to compare the full and reduced models is progressed with a 10-day time step and composed of 3000 days of pumping at a desired rate (Qdes) of 1000 m^3/d followed by 3000 days of recovery at a desired rate of 0 m^3/d. At the end of this timeframe the full and reduced models both have a 0% cumulative percent mass balance errors. The MAE and RMSE results from the full and reduced models are summarized in Fig. 7. The reduced model accurately represents the full model with all MAE's being less than 0.018 m and the largest NRMSE across all time steps is 0.135%. The exceedance curve increases slightly compared to TC1 with a 5% chance of being 0.021 m or greater. The errors are slightly greater than TC1 because there are more nonlinear features that the reduced model must represent and the reduced model dimension is smaller. This increase is nominal considering the overall accuracy of the reduced model and that the full model dimension is reduced by 96%.

The simulated water-table elevation for the full and reduced models at selected times is presented in Fig. 8. The water table is accurately represented by the reduced model despite dewatering layers 1, 2, and 3 and an MNW2 well extracting water from layers 2 and 5 (Fig. 8). MNW2 determines the flow rate at each screen location based on a calculated well water level (hwell) relative to the head in layers 2 and 5 as it attempts to meet the desired pumping rate. Fig. 9 shows the full and reduced model's simulated well water level and screen flow rates. During the first 3000 days, the pumping rate of 1000 m^3/d lowers the well water level (Fig. 9, top). When layer 2 dewaters at the well site (Time: ~1500 days; Fig. 9, left) the screen's flow rate lowers to zero and the layer 5 screen increases to meet the desired rate. This nonlinear response is visually identical between the full and reduced models. For the 3000 days of recovery (Time: 3000–6000 days; Fig. 9, right) the desired flow rate is set to zero; however, intraborehole flow may still occur. Water from layer 5 flows into the well bore and travels into layer 2 (Fig. 9, right). The reduced model does not visually match this nonlinear response, but the differences are negligible considering the scale of the flow rates, with the largest deviations no more than 0.06 m^3/d.

5. Three-dimensional test cases

Two synthetic, unconfined, three-dimensional (3D) groundwater-flow models are evaluated with the proposed model reduction methodology. The first 3D test case (TC3) is similar to TC1 and TC2, except a third dimension is added as well as other MODFLOW packages. The second 3D test case (TC4) is a complex-refinement of a scoping groundwater-flow simulation of Santa Barbara, CA [29]. This model is partially calibrated and only constructed to illustrate model reduction for the presented scenario.

5.1. Test case 3: synthetic model with multiple packages

The third test case, TC3, is a synthetic, three-dimensional, unconfined groundwater-flow model that incorporates multiple MODFLOW packages. The model is composed of 5 layers, 100 rows and 100 columns for a total of 50,000 simulated cells. Each layer is 10 m thick with the model's top and bottom elevations set to 0 m and –50 m, respectively. The length of each row and column is 100 m. The horizontal domain of TC3 is presented in Fig. 10. The initial condition is set to –2 m for the entire model domain. The domain is divided into two zones with columns 1 through 40 (Zone 1) and columns 41 through 100 (Zone 2) having a horizontal hydraulic of 5 m/d and 15 m/d, respectively, with no horizontal anisotropy. The vertical hydraulic conductivity is set to one-tenth of the horizontal. The entire domain has a $S_y = 0.3$ and $S_s = 1E{-}5\ m^{-1}$.

TC3 is bounded by two GHB conditions located along columns 1 and 100, across all rows and layers, and has a boundary head of –2 m and conductance calculated to represent a hydraulic conductivity of 5 m/d and 15 m/d, respectively. Within the domain there are two recharge fields and a drain field simulated using the recharge (RCH) and drain (DRN) packages, respectively (Fig. 10). The recharge fields distribute a specified flux over the top of the model and are located

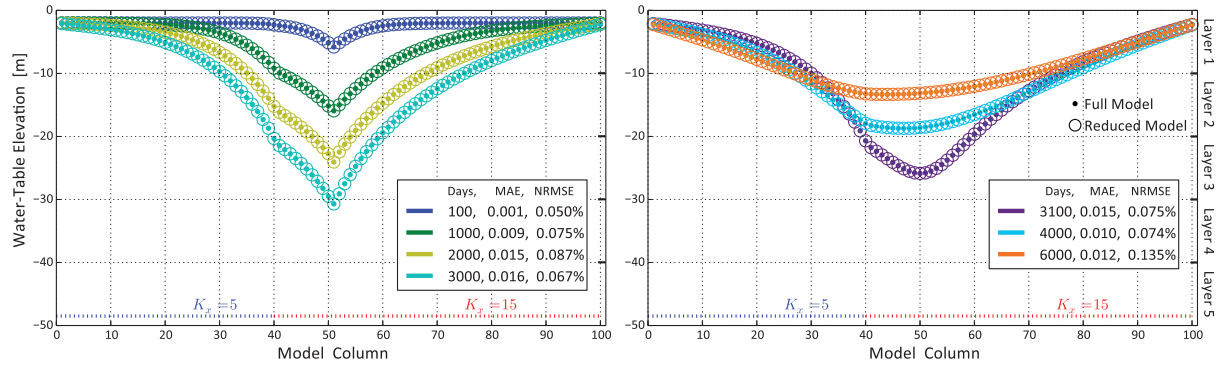

Fig. 8. TC2 simulated water-table elevation during the first 3000 days of pumping (left) and the next 3000 days of recovery (right) at select times for the full and reduced models. Note that MAE and NRMSE are calculated from the simulated head at all 500 model cells.

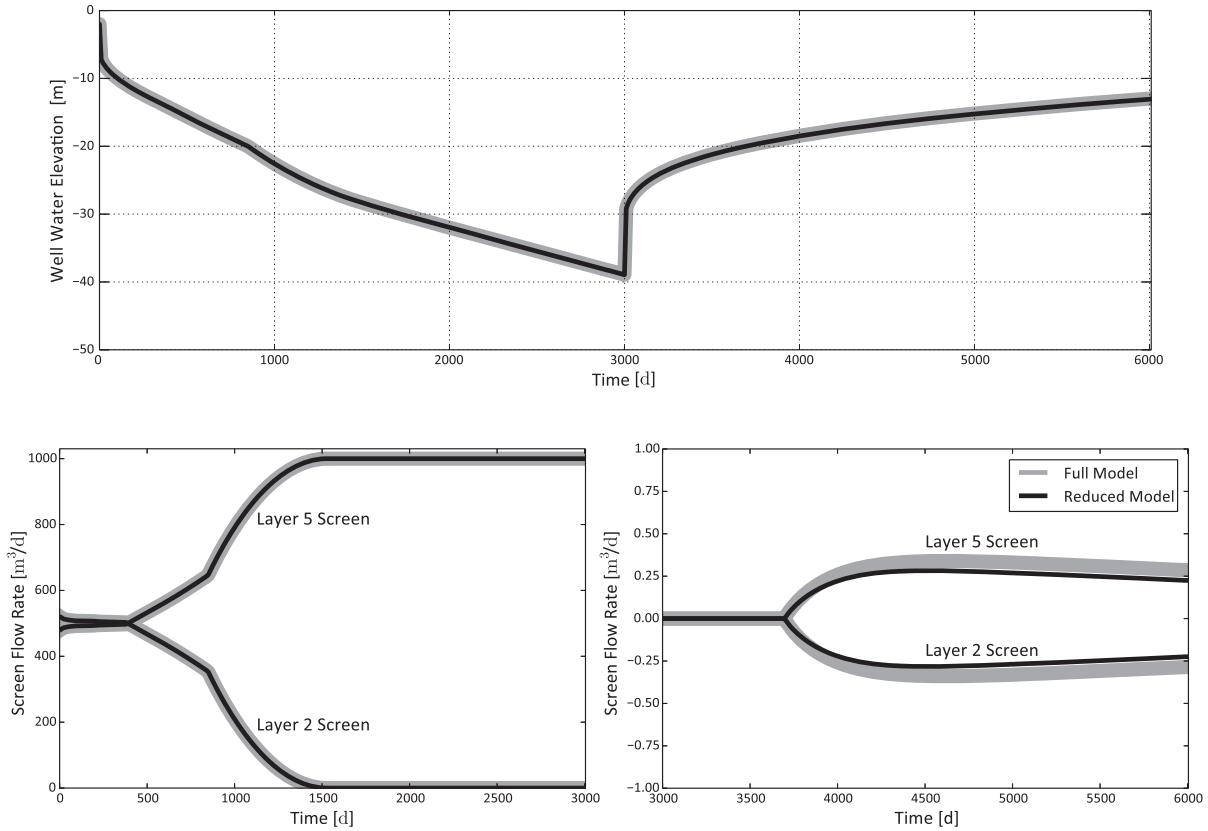

Fig. 9. TC2 full and reduced model's simulated well water level (top) and screen flow rate with (left) and without pumping (right). The MNW2 screen flow rate is positive when flow is out of the layer and into well. Note the center and right parts are on different scales.

in the upper left (columns 11–30) and upper right (columns 11–30) of the model and have a recharge flux of 0.1 mm/d. Note that since the Newton formulation does not allow cells to deactivate when dewatered, a dewatered upper layer cell will receive the recharge and then transmit it downward toward the upper-most saturated layer. The drain field is located between columns 21 and 30, has a drain elevation set to –5 m, and the hydraulic conductance of the interface between the aquifer and the drain is 100 m^2/d. Drains only remove water from the model when the water table elevation is greater than –5 m.

A fault line is simulated using a horizontal flow barrier (HFB) [30] that is between two columns. There is a break and shift in the barrier near a river to allow groundwater flow through and to simulate a shifted and washed-out barrier. The barrier is located across all layers and is either between columns 37 and 38 or between columns 39 and 40. The barrier has a thickness of 5 m and a hydraulic conductivity of 0.001 m/d. Running parallel to the barrier is a river (RIV package) that crosses it near its break and then runs parallel again. The riverbed bottom elevation (Rbot) is –2 m and the head of the river (Stage) is set to 1 m. The riverbed hydraulic conductivity is 0.25 m/d and 0.5 m/d for Zones 1 and 2, respectively.

Water is extracted from the groundwater system by seven wells (WEL) and one multi-node well with two screens (MNW2). The well field represents four wells that always pump at the same rate and are located at layer 4 and along column 80. The left-most well is located at layer 4, row 26 and column 21, and the remaining two wells in Fig. 10 pump at the same rate and are located at layers 2 and 5 at row 71 and column 51. The location of the MNW2 well is at row 26 and column 65 with a screened interval from –20 m to –30 m and –40 m to –50 m (layers 3 and 5, respectively). The MNW2 well uses the "skin" losstype, has a radius of 0.1 m (inner radius), and a skin radius of 0.25 m (outer radius). The hydraulic conductivity of the skin material is 0.5 m/d.

To construct the reduced model, six sets of snapshots are collected (NSET = 6) from MNW2, WEL, RIV, and RCH for a total of 460 snapshots. For MNW2 and WEL there are 45 snapshots, exponentially distributed in time, taken while pumping is active for 50,000 days (quasi-steady state time), and then 45 more snapshots are taken for 22,500 days of recovery (no pumping). Since the well field always pumps at the same rate, it only requires one set of snapshots to represent the four wells. For both RCH and RIV there are 50 snapshots, exponentially distributed in time, taken for over 50,000 days. The snapshot set (S) is mean centered and normalized; SVD is then applied to obtain the singular values and POD basis. The magnitude of each singular value and its corresponding percent energy is presented in Fig. 11. This illustrates that there is a good spread of information among the chosen snapshots. Even with a good spread of singular values the PE indicates a rapid increase to 99.9% around the 60th singular value. To demonstrate an extreme reduction, a PE of 99% is selected, which results in a POD basis composed of 28 vectors. The lower PE still produces acceptable errors with the extreme reduction of the Jacobian dimension by 99.9% and Eq. (16) from 50,000 equations to 28 equations.

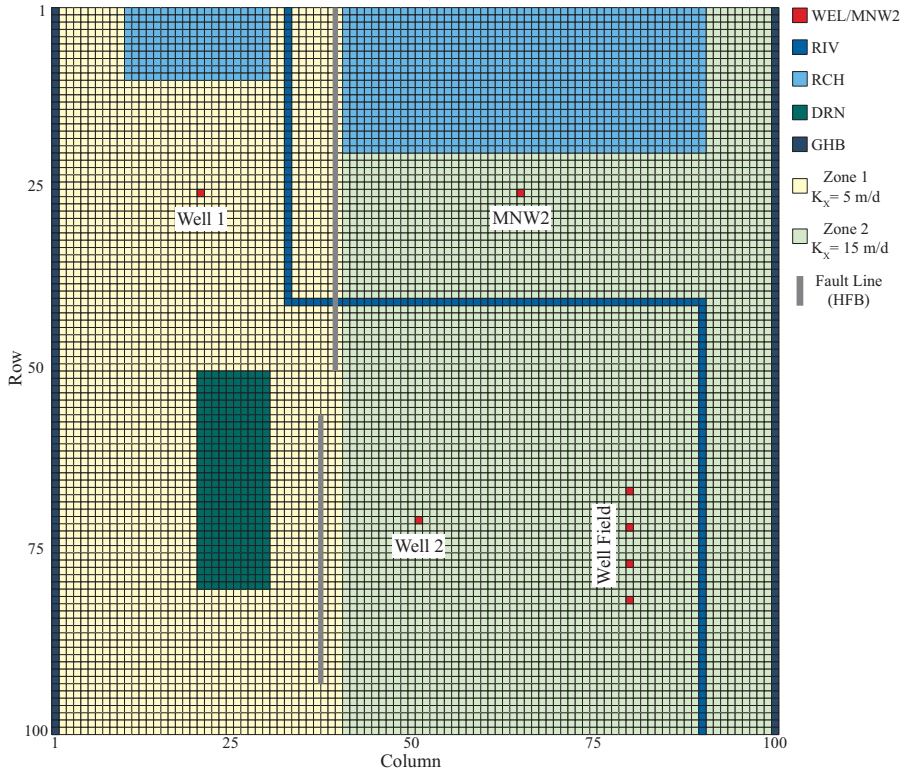

Fig. 10. Plan overview of test case 3 (TC3).

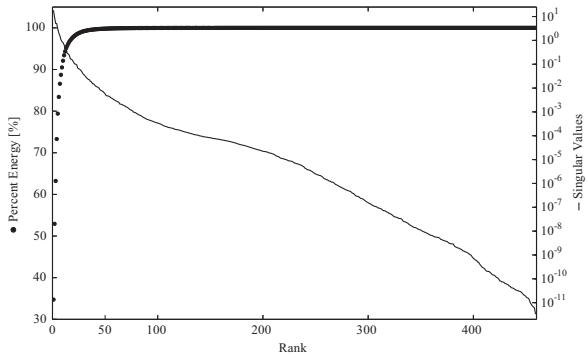

Fig. 11. Percent energy (left *y*-axis) and singular values (right *y*-axis) from applying SVD to the 460 snapshots taken from TC3.

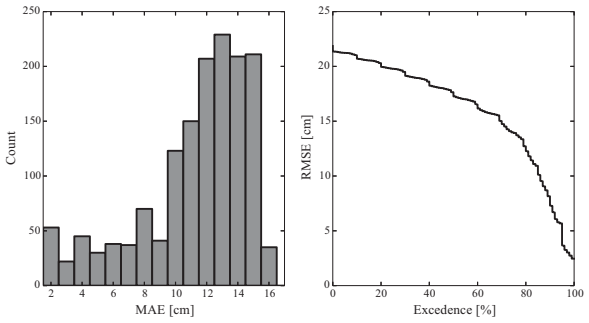

Fig. 12. TC3 (left) histogram of the mean absolute errors from all time steps and (right) exceedance curve of the RMSE.

To compare the full and reduced models, a simulation timeframe of 3000 days is divided into 20 150-day, stress periods with a time step of 2 days. Water is extracted from the domain from the wells during the odd-numbered stress periods (e.g. 1, 3, etc.) and there is no pumping during the even-numbered stress periods (MNW2 may still have intraborehole flow). For all stress periods the GHB, DRN, RCH, and RIV remain at non-zero values. At the end of the simulation the full and reduced models' cumulative percent mass balance errors are 0% and 0.6%, respectively. The largest NRMSE across all time steps is 3.5%. The error results from the full and reduced models are summarized in Fig. 12. The histogram and RMSE exceedance curve is compiled by computing the MAE and RMSE from the reduced and full models for each time step of the simulation. The mean average errors are acceptable considering the level of reduction demonstrated in TC3. The exceedance curve has a 5% chance of being 0.212 m or greater, which is about an order of magnitude greater than the 5% from TC1 and TC2. This increase is expected since the number of cells in TC3 is an order of magnitude greater than TC1 and TC2.

Assuming that errors propagate from previous time steps, the results of the final time step should be the least accurate. Fig. 13 is a contour plot of the water-table elevation at the last pumping stress period (2850 days) and the last recovery stress period (3000 days), which occur at times 1425 and 1500 days, respectively. The reduced model matches the full model with the contours in general

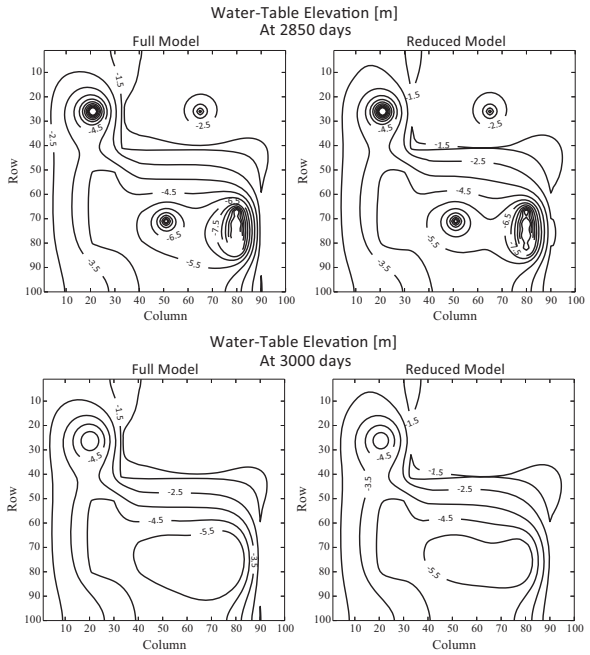

Fig. 13. TC3 comparison between the full (left) and reduced (right) model's water table elevation at the last pumping stress period at 2850 days (top) and the last recovery stress period at 3000 days (bottom).

agreement of each other (Fig. 13). By splitting the calculation the MAE and NRMSE for only the time steps with pumping (Fig. 13, top) and those with recovery (Fig. 13, bottom) shows that the reduced model is slightly more accurate during pumping. These associated errors for pumping and recovery are a MAE of 0.108 m and 0.116 m, respectively, and NRMSE of 0.62% and 2.4%, respectively. The larger error metrics during the recovery time steps is most likely the result of selected snapshots not adequately representing the recovery of the model or information that was related to the recovery is removed by using a PE of 99% (a higher PE would include more information from the snapshots).

Fig. 14 presents the pumping from the MNW2 well in TC3 for the full and reduced models. Since the desired pumping rate (Qdes) is non-zero for the odd stress periods, the head in the well rapidly drops due to pumping. During the even stress periods there is no pumping and the well water level in the MNW2 well (Fig. 14, left) recovers as water seeps in from the two screened layers. During the course of an odd stress period (pumping) the net flow rate from the layer 3 screen starts to drop as the head in the well lowers. Consequently the layer 5 screen's net rate increases to meet the desired pumping rate. The reduced model does not exactly reproduce these nonlinear responses; however, the differences are negligible and the overall trend matches.

5.2. Three-dimensional test case (TC4) of Santa Barbara groundwater-flow model

In order to present our method's applicability to a "real-world" problem, we developed a scoping groundwater-flow model based on the published model of the Santa Barbara groundwater basin [29], located in Santa Barbara, CA, Fig. 15. The original model input files are designed for MODFLOW-88 [31]. To apply model reduction the flow and solver input files are converted to the UPW/NWT input files, and the remainder of the packages updated for simulation with MF-OWHM. This model has been partially calibrated and is constructed to illustrate the applicability of model reduction for the presented test case.

The Santa Barbara model, now referred to as TC4, is composed of 56 layers, 58 rows, 152 columns and has 113,578 active model cells. The first layer follows the land topography with a maximum and minimum elevation of 179 m mean sea level (msl) and –4 m msl, respectively. Each model layer has an average thickness of 6 m and the deepest model cell bottom has an elevation of –340 m msl. The row and column lengths are a constant value of 76 m. The TC4 domain (Fig. 15) contains six mapped fault lines that are represented as horizontal flow barriers (HFB). Mountain-front recharge along the northeast boundary is represented by the recharge package (RCH). In the southeast region there is a drain network (DRN) that removes water from the groundwater system. The rivers located in the active model domain are simulated using the RIV package. The ocean boundary is represented as a freshwater equivalent head using the GHB package and the remainder of the model boundary is no-flow. The cell-wise hydraulic conductivity data is developed by compiling drillers' log data, electrical-resistivity data, and geologic maps to derive a semi-continuous hydraulic conductivity value for every model cell. Seven production wells are simulated with the WEL package and have screening intervals that span across multiple layers. Unlike MNW2, the well package cannot accommodate pumping spread across multiple model cells directly. Instead, multiple WEL cells are required to represent a well's screened intervals. The total pumping from a well is distributed across the multiple WEL cells according to the cell's hydraulic-conductivity value. Due to this limitation the seven production well sites (Fig. 15) are represented by 114 WEL cells.

To construct the reduced model eight sets of snapshots, exponentially distributed in time, are collected (NSET = 8). To capture the model's unpumped dynamics, one set of 28 snapshots is collected, without pumping but with RCH, RIV, GHB, and DRN, spanning from 0.1 days to 5000 days (near the initial condition to quasi-steady state time). The next seven sets of snapshots contain 19 exponentially distributed snapshots from 0.03 days to 1323 days (near the pumping quasi-steady state) with one of the seven wells set to a constant

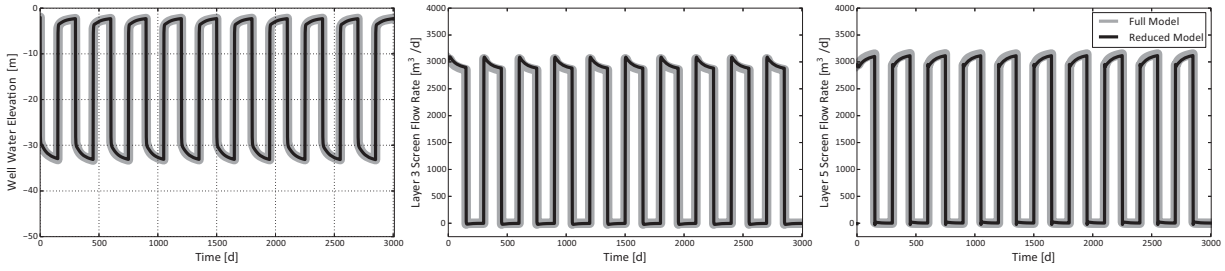

Fig. 14. TC3 full and reduced model's (left) MNW2 well water level and (center and right) MNW2 screen flow rate where positive flow is out of the layer and into the well.

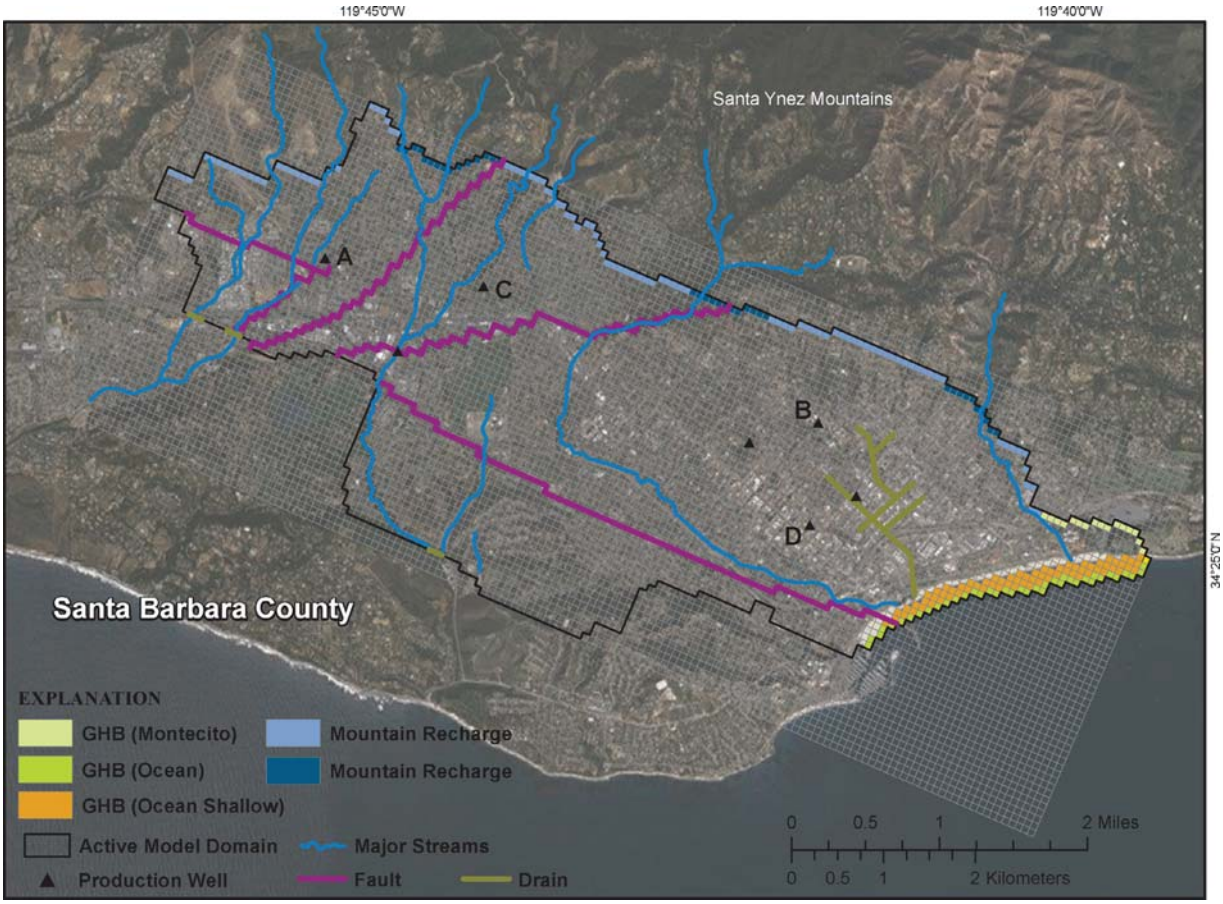

Fig. 15. Overview of the Santa Barbara Model, TC4.

pumping rate. The combined snapshot sets are mean centered, normalized and then SVD is applied to obtain the POD basis. The associated errors from different POD basis dimensions are then evaluated using the TC4 scenario.

The accuracy of the reduced model is compared with the full model during a 2-year simulation with monthly pumping changes (24 stress periods). Each stress period is equal to the days in a month (e.g. 31, 28, 31 days) with four time steps per stress period and beginning in January. During the first year, the pumping rate for the seven wells increases from a small rate to a maximum rate from stress period 1 to 9 and then decreases to zero from 10 to 12. The pumping rates are different across the seven wells with the smallest stress period 1 rate being 178 m^3/d and largest stress period 9 rate being 4823 m^3/d. The second year repeats the same pumping schedule that is applied to the first.

For this synthetic scenario, the full and reduced models are evaluated at 99%, 99.9%, and 99.99% PE, which results in a POD basis composed of 86, 127, and 147 vectors, respectively. The three different bases derive reduced models with different levels of accuracy. At the end of the simulation the full and 99%, 99.9%, and 99.99% PE reduced models have a cumulative percent mass balance errors of 0.1%, 2.9%, 1.9%, and 1.88%, respectively. Fig. 16 presents the snapshot set's singular values and their corresponding PE values along with the associated errors for the three aforementioned PE values. The largest NRMSE across all time steps for the three PE values is 0.16%, 0.13%, and 0.13%, respectively. There is a direct correlation with the accuracy loss by lowering the PE below 99.99%. As would be expected, smaller values of PE have the benefit of a smaller POD basis, but at a loss in accuracy. This is not necessarily true for larger values of PE (>99.99%) due to the fact that eventually basis vectors will be included that contribute nearly zero information to the reduced model. This is demonstrated by (Fig. 16) the large improvement in the reduced model's MAE histogram and exceedance curves between PE values of 99% and 99.9%, but a negligible improvement from PE values of 99.9–99.99%. In fact, the RMSE for PE values of 99.9% and 99.99% are nearly identical (Fig. 16).

Since there is little improvement beyond a PE of 99.9% the subsequent discussion and analysis are from this PE value, which results in a POD basis composed of 127 basis vectors. This POD basis then reduces the model derived from TC4 from 113,578 equations (Eq. 7) to 127 (Eq. 16), reducing the dimension by 99.9%. Fig. 17 presents four hydrographs from the full and reduced models at well locations specified as A, B, C and D in Fig. 15. Only the well sites' upper-most (top WEL cell) and lower-most (bottom WEL cell) WEL cells are presented in the hydrographs. Due to the three-dimensional nature of this highly heterogeneous model, the Top WEL Cell does not necessarily have a higher hydraulic head. Many factors can cause this effect, such as having the top WEL Cell being surrounded by lower hydraulic conductivity cells than those that surround the bottom WEL Cell. The reduced model precisely represents the full

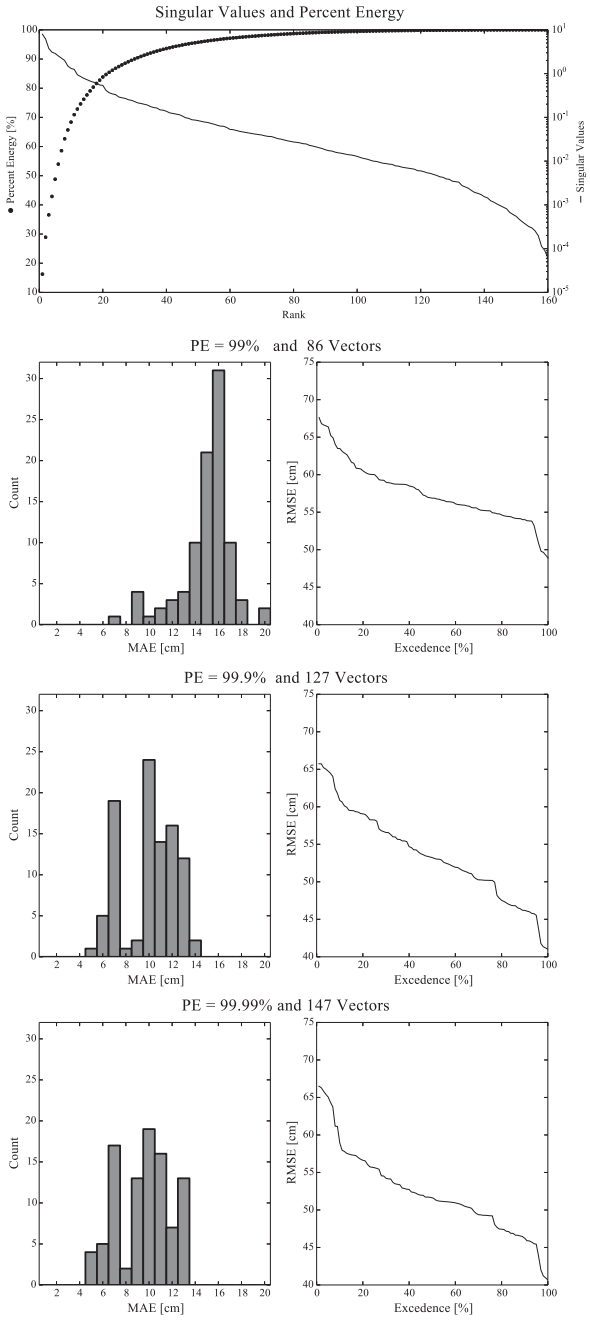

Fig. 16. TC4 snapshot set's singular values and corresponding percent energy (top) and the error between the full and reduced models at different percent energies (PE).

model at each location through both the pumping and recovery of TC4.

The simulated water table at the end of the 2-year scenario is presented in Fig. 18. The simulated water-table elevation varies from −1.3 m msl to 133.5 m msl (Fig. 18). The effect of the fault lines (Fig. 18) is easily identified by the sharp contrasts in water-table elevation (Fig. 18). The drain network (Fig. 18) is also apparent from the lowered water-table elevation. At the scale of Fig. 18, visually, the simulated water tables for the full and reduced models are indistinguishable. This is quite significant considering the original model requires the solution of 113,578 equations for each Newton step compared to the reduced model's 127 equations.

5.3. Quantification of CPU time savings

It is difficult to quantify improvement in CPU time due to many factors that affect overall simulation time. For example, the reduced model's system of equations are solved by using the LAPACK's General Matrix LU decomposition with partial pivoting and row interchanges (DGESV) [26], while the NWT package includes a choice of two highly optimized solvers specifically designed for its matrix structure. The current implementation is also undertaken in a way to ensure that the computer code performs as expected with minimal code refactoring for speed purposes. What is common between the full and reduced models is that they are compiled with the Intel Visual Fortran Composer XE 2013 SP1 for single thread execution and run on a Dell Precision M6700 with Windows 7,32 GB RAM, a Samsung 840PRO SSD hard drive, and Intel Core i7-3840QM processor at 2.8 GHz. With the aforementioned computer the CPU time required for completion of the full and reduced models for TC3 are 24.3 min and 7.6 min, respectively. For TC4 the CPU times are 33.5 min and 6 min, respectively. For this initial investigation of unconfined model reduction, there is a significant savings in CPU time. Further improvement is possible through better identification of snapshots and optimized solvers.

6. Discussions and conclusion

This paper describes a novel method for applying projection-based model reduction to unconfined groundwater-flow systems and applies it to four test cases. The POD basis is derived from taking snapshots from the original, full model and applying singular value decomposition to identify the dominant proper orthogonal modes. This POD basis is applied through the Galerkin projection to the Jacobian in the Newton formulation of MODFLOW. This results in a significant reduction in the dimensionality of the Jacobian for transient, unconfined groundwater-flow models. This method is not limited to the unconfined groundwater flow and can be applied to a groundwater system that is strictly confined (Eq. 1). In fact, the model reduction of the four test cases were originally validated using only the confined equation (Eq. 1), before their evaluation with the unconfined equations (Eqs. 1 and 2).

The four test case examples illustrate that the reduced model can replicate the full model under a variety of scenarios and maintain a small mass balance error. TC1 and TC2 demonstrate that the reduced model replicates the original model when multiple MODFLOW layers are dewatered. TC2 and TC3 incorporate an MNW2 well to increase the complexity. The reduced model verifies that not only for the simulated water table; it also closely matches the simulated well head level and intraborehole flows. TC3 and TC4 produce reduced models that simulate three-dimensional groundwater flow in a heterogeneous system. TC4 is a more complicated case composed of 113,578 model cells with a spatially varying horizontal hydraulic conductivity, vertical hydraulic conductivity, and specific storage. In each of the test cases, the dimensionality reduction is at least an order of magnitude, resulting in faster simulation runs.

Snapshot selection does affect the reduced model's fidelity to the full model. Poorly selected snapshots result in a poor basis and consequently a reduced model that does not accurately reflect the full one. In this case, the best reduced models are developed from exponentially distributed snapshot times starting just after the initial condition to steady state under a constant forcing. For certain nonlinear problems the reduced model is further improved by continuing to take snapshots with the forcing turned off in order to capture the

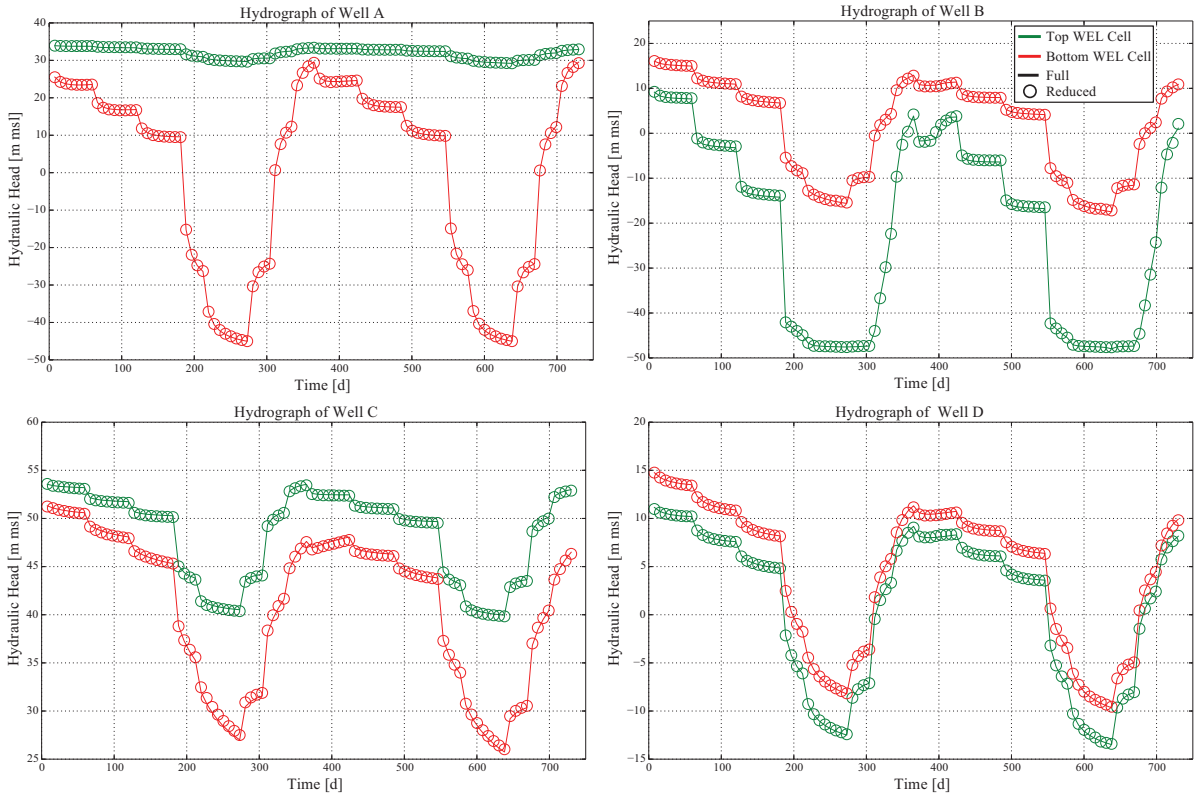

Fig. 17. TC4's full and reduced models' simulated hydrographs at the four production wells identified as A, B, C and D in Fig. 15.

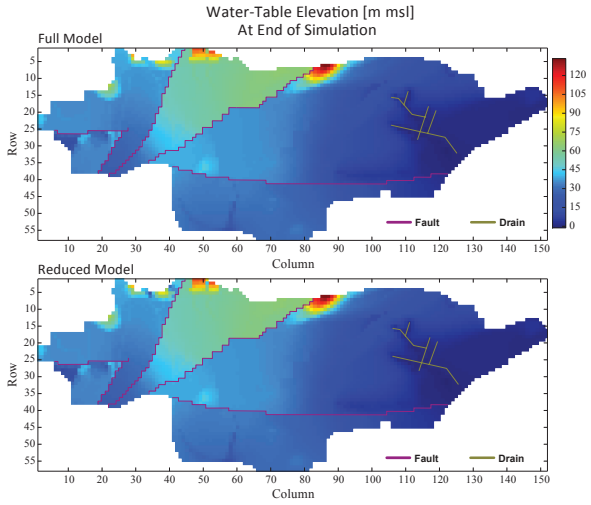

Fig. 18. TC4's full (top) and reduced (bottom) models' water table elevation at the end of the 2 year scenario.

dynamics of the nonlinear recovery. This is especially important for multi-node wells in order to capture the intraborehole flows under no pumping. Another improvement, which was not needed for the presented test cases, is to collect snapshots at different constant rates (e.g. two snapshot sets for one well, but at different rates). Taking two sets at different rates makes the set less sensitive to the actual snapshot times, though generally this results in larger-dimensioned reduced models. After the snapshots are collected, applying mean centering to the final set greatly improves the reduced model and is recommended for the presented model reduction formulation. An important area for future research and expansion is for an automatic or optimal decision tree for the selection of snapshots.

Acknowledgment

Funding was provided by the US Geological Survey California Water Science Center research program and the NSF under award EAR-1314422. Partial support also was provided by an AECOM endowment. The authors would like to thank five anonymous reviewers and two USGS reviewers for their in-depth and constructive reviews.

References

[1] Vermeulen PTM, Heemink AW, te Stroet CBM. Low-dimensional modeling of numerical groundwater flow. Hydrol Process. 2004;18(8):1487–504.
[2] McPhee J, Yeh WW-G. Groundwater management using model reduction via empirical orthogonal functions. J Water Resour Plan Manag. 2008;134(2):161–70.
[3] Siade AJ, Putti M, Yeh WW-G. Reduced order parameter estimation using quasi-linearization and quadratic programming. Water Resour Res. 2012;48(6):W06502 http://dx.doi.org/10.1029/2011WR011471.
[4] Boyce S, Yeh WW-G. Parameter-independent model reduction of transient groundwater flow models: application to inverse problems. Adv Water Resour. 2014;69::168–80 ISSN 0309-1708 http://dx.doi.org/10.1016/j.advwatres.2014.04.009.
[5] Kerschen G, Golinval JC. Physical interpretation of the proper orthogonal modes using the singular value decomposition. J Sound Vibrat. 2002;249(5):849–65.
[6] Chatterjee A. An introduction to the proper orthogonal decomposition. Curr Sci. 2000;78(7):808–17.

[7] Hinze M, Volkwein S. Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: error estimates and suboptimal control. Dimension reduction of large-scale systems. Berlin Heidelberg: Springer; 2005. p. 261–306.
[8] Lang YD, Malacina A, Biegler LT, Munteanu S, Madsen JI, Zitney SE. Reduced order model based on principal component analysis for process simulation and optimization. Energy and Fuels 2009;23(3):1695–706.
[9] Park HM, Cho DH. Low dimensional modeling of flow reactors. Int J Heat Mass Transf. 1996;39(16):3311–23.
[10] Vermeulen PTM, Heemink A, Te Stroet C. Reduced models for linear groundwater flow models using empirical orthogonal functions. Adv Water Resour. 2004;27(1):57–69.
[11] Siade AJ, Putti M, Yeh WW-G. Snapshot selection for groundwater model reduction using proper orthogonal decomposition. Water Resour Res. 2010;46(8):W08539 http://dx.doi.org/10.1029/2009WR008792.
[12] Baú DA. Planning of groundwater supply systems subject to uncertainty using stochastic flow reduced models and multi-objective evolutionary optimization. Water Resour Manag 2012;26(9):2513–36.
[13] Cardoso MA, Durlofsky LJ, Sarma P. Development and application of reduced-order modeling procedures for subsurface flow simulation. Int J Numer Meth Eng. 2009;77(9):1322–50. http://dx.doi.org/10.1002/nme.2453.
[14] Robinson BA, Lu Z, Pasqualini D. Simulating solute transport in porous media using model reduction techniques. Appl Math. 2012;3(10):1161–9 http://dx.doi.org/10.4236/am.2012.310170.
[15] Li X, Hu BX. Proper orthogonal decomposition reduced model for mass transport in heterogenous media. Stochast Environ Res Risk Asses 2013;27(5):1181–91 http://dx.doi.org/10.1007/s00477-012-0653-2.
[16] Buchan AG, Pain CC, Fang F, Navon IM. A POD reduced-order model for eigenvalue problems with application to reactor physics. Int J Numer Meth Eng. 2013;95(12):1011–32.
[17] Harbaugh AW. MODFLOW-2005, the US geological survey modular ground-water model: the ground-water flow process. US Dep Inter, US Geo Surv Tech Meth 2005:6–A16.
[18] Painter S, Başağaoğlu H, Liu A. Robust representation of dry cells in single-layer MODFLOW Models. Ground Water 2008;46(6):873–81.
[19] Keating E, Zyvoloski G. A stable and efficient numerical algorithm for unconfined aquifer analysis. Ground Water 2009;47(4):569–79.
[20] Niswonger RG, Panday S, Ibaraki M. MODFLOW-NWT, a Newton formulation for MODFLOW-2005. US Geo Surv Tech Meth 2011:6–A37.
[21] Hanson RT, Boyce SE, Schmid W, Hughes JD, Mehl SM, Leake SA, et al. One-water hydrologic flow model (MODFLOW-OWHM). US Geo Surv Tech Meth. 2014;120:6–A51 http://dx.doi.org/10.3133/tm6A51.
[22] Willis R, Yeh WW-G. Groundwater systems planning and management. Englewood Cliffs, NJ: Prentice-Hall; 1987. p. 416.
[23] Harbaugh AW, Banta ER, Hill MC, McDonald MG. MODFLOW-2000, the U.S. geological survey modular ground-water model – User guide to modularization concepts and the ground-water flow process. US Geo Surv Open-File Rep. 2000;121:00–92.
[24] Laub AJ. Matrix analysis for scientists & engineers. Philadelphia: Society for Industrial and Applied Mathematics; 2005. p. 157.
[25] Laub AJ. Computational matrix analysis, xiii. Philadelphia: Society for Industrial and Applied Mathematics; 2012. p. 154.
[26] Anderson E, Bai Z, Bischof C, Blackford LS, Demmel J, Dongarra JackJ, Du Croz J, Hammarling S, Greenbaum A, McKenney A, Sorensen D. LAPACK Users' guide. 3rd ed., xxi. Philadelphia: Society for Industrial and Applied Mathematics; 1999. http://dx.doi.org/10.1137/1.9780898719604 p. 404.
[27] Halford KJ, Hanson RT. User guide for the Drawdown-Limited, multi-node well (MNW) package for the U.S. geological survey's modular three-dimensional finite-difference ground-water flow model, Versions MODFLOW-96 and MODFLOW-2000. US Geo Surv Open-File Rep. 2002;33:02–293 http://pubs.usgs.gov/of/2002/ofr02293/text.html.
[28] Konikow LF, Hornberger GZ, Halford KJ, Hanson RT. Revised multi-node well (MNW2) package for MODFLOW ground-water flow model. US Geo Surv Tech Meth 2009;67:6–A30.
[29] Freckleton JR, Martin P, Nishikawa T. Geohydrology of storage unit III and a combined flow model of the Santa Barbara and Foothill ground-water basins, Santa Barbara County, California. Sacramento, California. US Geo Surv Water-Res Investig Rep. 1998;80:97–4121.
[30] Hsieh PA, Freckleton JR. Documentation of a computer program to simulate horizontal-flow barriers using the modular three-dimensional finite-difference ground-water flow model. US Geo Surv Open-File Rep. 1993;59:92–477.
[31] McDonald M.G., Harbaugh A.W. A modular three-dimensional finite-difference ground-water flow model: Techniques of water-resources investigations of the United States geological survey, Book 6, Chapter A1; 1988. p. 586.

Chapter 4
Conclusion

Groundwater resources provide a stable, potable water supply when exploited appropriately. Assessing groundwater resources through the application of mathematical models that simulate the dynamics of an aquifer system provides a better understanding of them. Mathematical models provide information on how to minimize negative environmental impacts and aid in quantifying and analyzing groundwater sustainable usage. Complex, highly-discretized groundwater simulation models often have a large computational requirement that can prevent more advanced understanding and analysis. This dissertation presented two new model reduction techniques that significantly lower the computational burden by reducing the state dimension, hydraulic head, of a groundwater simulation model. This chapter concludes and summarizes the two papers presented in Chapters 2 and 3.

4.1 Parameter-Independent Model Reduction

Groundwater simulation models require the specification of hydraulic conductivity values that are assigned to a predefined zonation pattern. The zonal hydraulic conductivity cannot be directly measured requiring its inference from an inverse problem. Solving both a deterministic and Bayesian inverse problem require numerous simulation runs to complete their evaluation of hydraulic conductivity. Model reduction can significantly lower simulation time, but only works for zonal values of hydraulic conductivity near the zonal values used to construct reduced models Galerkin projection operator. This limits the applicability of model reduction to inverse problems because as hydraulic conductivity deviates from the value used to build the Galerkin projection operator would necessitate recollecting new snapshots and rebuilding the operator. This is

especially problematic for Bayesian inverse problems because sampled hydraulic conductivity varies lognormally, necessitating a new Galerkin projection operator for each sample. Presented in Chapter 2 is a new methodology that constructs a Galerkin projection operator that does not have to be rebuilt for any sampled hydraulic conductivity. The reduced model has a high reusability because it can also work for solving a deterministic inverse problem and pump schedule management problems.

Most model reduction methods are predicated on the parameter values selected to generate the snapshots. Chapter 2's methodology automatically determines parameter sets that should be represented in the reduced model and then automatically selects the appropriate snapshots for that parameter set. The proposed methodology constructs a reduced model that accepts any reasonable combination of hydraulic conductivity as input. The reduction is done via a Galerkin projection operator, which is an orthonormal basis, and called, in Chapter 2, the parameter-independent projection matrix. This orthonormal basis is constructed by iteratively selecting parameter sets and their corresponding snapshot times through the use of two greedy algorithms. A greedy algorithm solves a multi-stage optimization problem by combining the optimal solution obtained from each stage. In general, a greedy strategy does not guarantee global optimum, but, in many instances, yields a good approximation to the optimal solution. The advantages of Greedy optimization are its easy implementation, fast execution, and often it yields a solution when multi-objective optimization problems fail due to nonuniqueness.

The parameter sets are selected by solving a bounded optimization problem at each greedy stage, called the Parameter Greedy Step. In between each Parameter Greedy Step a set of corresponding greedy optimal snapshot times are selected. The process of selecting a new snapshot time is called the Time Greedy Step and yields a single snapshot for each Greedy stage that is added to the parameter-independent projection matrix. These snapshots from different groups of parameters are collected into one large basis that when made orthonormal becomes a parameter independent Galerkin projection operator. This operator reduces by one to three orders of magnitude the dimensionality of the original, full model and reproduces accurate solutions for

any reasonable combination of parameter values.

The proposed methodology is validated on a conceptual one-dimensional, five zone confined aquifer, Chapter 2-Figure 2. The constructed Galerkin projection operator reduces the groundwater flow system equations from 303 to 33, resulting in an average time savings of 23% (0.08 s) using the Chapter 2's described computer architecture. A synthetic observation set, that is corrupted with noise, is derived to validate the reduced models ability to reproduced the same deterministic inverse and Bayesian inverse problems solutions as the original, full dimensional model. The deterministic inverse solution of the reduced model differed from the full model solution with relative errors from 1% to 5%, while the resulting Markov chains for the Bayesian inverse solutions were found to be statistically equivalent between the reduced model and the full model.

The range and validity of the proposed parameter-independent model reduction methodology was demonstrated by exploring three different large-scale, highly discretized models of a two-dimensional, conceptual, confined aquifer in the Oristano plain of Sardinia, Italy. The three versions contain the same well locations and boundary conditions and differ only in their number of hydraulic conductivity zones. The number of zones between the different versions are: 3, 7 and 15 zones and referred to as Z3, Z7, and Z15, respectively. All three versions required the solution of 29,197 equations per simulated time step. The proposed methodology reduced the Z3, Z7, and Z15 models to 69, 127, and 273 equations, respectively, resulting in an average CPU time savings of 94%, 91%, and 84%, respectively. The reduced model maintains a high level of accuracy for a wide range of parameter values, with the majority of the mean absolute errors being 3 cm or less and the mean absolute relative errors 4% or less. The time savings is substantial considering that it would take an estimated 38 to 60 days to construct a Markov chain of 100,000 samples from the full Oristano model for all three versions.

In contrast, the Z3, Z7, and Z15 reduced models require 4, 7, and 12 days, respectively, to complete.

4.2 Model Reduction of Unconfined Groundwater Flow

When the piezometric surface of a confined aquifer lowers below its upper confining unit or there exists no upper confining unit, then the aquifer dynamics change from a linear response to nonlinear. This change is the result of how water is released from storage under confined conditions compared to unconfined. The unconfined condition produces a free surface upper boundary called the water table or phreatic surface. Since most shallow well systems extract water from unconfined aquifers, and excessive pumping of confined aquifers can dewater them to unconfined conditions, the simulation of unconfined flow is an essential part of regional groundwater modeling projects.

Complex, highly discretized unconfined models can have a substantial computational requirement to simulate groundwater flow. Using MODFLOW-OWHM and the Newton Formulation of MODFLOW, which is a variant of the Newton-Raphson method, improves the stability of the solution and the convergence speed of the simulation, but for many scenarios is computationally intractable to allow advanced analysis. The issue with the Newton formulation is it involves solving the inverse of the nonsymmetric Jacobian. To reduce the computational burden, and thus the simulation runtime, Chapter 3 uses the Galerkin projection to form a reduced dimension Jacobian, solves for the inverse of the reduced Jacobian, and then calculates the Newton step. This drastically reduces the computational demands of solving the inverse of the Jacobian allowing more advanced unconfined analysis.

The projection-based model reduction uses a POD basis that is derived from taking snapshots from the original, full model. The basis removes any linear dependence and identifies

the most dominant basis vectors, the proper orthogonal modes, through singular value decomposition. This POD basis is applied through the Galerkin projection to the Jacobian in the Newton formulation of MODFLOW. This results in a significant reduction in the dimensionality of the Jacobian for transient, unconfined groundwater-flow models. This method is not limited to the unconfined groundwater flow and can be applied to a groundwater system that is strictly confined with significant improvements on accuracy and time savings because the groundwater flow equations become linear.

Chapter 3 presents four test cases that illustrate that the reduced model can replicate the full model under a variety of scenarios and maintain a small mass balance error. The most significant reduction occurs in the fourth test case. This case contained a real world simulation model of the Santa Barbara and Foothill groundwater basins and requires solving for the inverse of a Jacobian matrix composed of 113,578 rows for each Newton-Raphson iteration. Applying the Galerkin projection results in a reduced dimension Jacobian is composed of 127 rows. That is a three order of magnitude of reduction (~99.9% reduction) in the Jacobian dimension.

Snapshot selection does affect the reduced model's fidelity to the full model. Poorly selected snapshots result in a poor basis and consequently a reduced model that does not accurately reflect the full one. This limitation of POD includes the number of snapshots, the simulated times that they are taken at, and, for nonlinear models, the value of the forcing itself. From empirical tests, each independent forcing should have between 10 to 50 snapshots that are exponentially distributed in time. These snapshots are collected under a constant reference forcing and range from slightly beyond the initial condition to quasi-steady state. If the model is highly nonlinear, then more snapshots maybe necessary increasing the number of snapshots at times where the reduced model performs unsatisfactorily.

The reduced model for unconfined flow tends to be more accurate while under stress (i.e. pumping) compared to its recovery from that stress. To increase the accuracy during recovery, it is recommended to take an additional 10 to 50 exponentially distributed in time snapshots. These snapshots begin slightly beyond the termination of the forcing and end when the groundwater system returns to its no forcing, steady state condition.

Selection of the pumping rate does influence the accuracy of the reduced model to replicate nonlinear responses. A good rule of thumb is to select a rate that dewaters the aquifer layer or produces a drawdown equivalent to the largest observed drawdown at that well location. This ensures that the snapshots that are collected contain all drawdown and the recovery information that the original, full model could potentially calculate as output. If this is problematic or the reduced model fails to reproduce the original, full model, then another alternative is to take sets of snapshots at different pumping rates for the same pumping well. These two sets of snapshots, taken at different pumping rates for the same well, would start just beyond the initial condition, end at their respective steady states, and then capture their respective recovery.

After the snapshots sets have been collected and combined, the size of the final projection basis is determined by the Percent Energy (PE). Typical values for PE range from 99% to 99.99%, where smaller values lead to fewer basis vectors at a cost of less accurate reduced model results. The fourth test case explored the accuracy of the reduce model at different values of PE. This test case at 99%, 99.9%, and 99.99% PE results in a POD basis composed of 86, 127, and 147 vectors, respectively. PE less than 99% yielded reduced models that produced unacceptable results. Increasing PE increases the reduced model accuracy with diminishing returns. Eventually

the gain in accuracy is nominal and the reduced model dimension approaches the original, full model providing no significant time savings.

Chapter 5

Future Research

Two new model reduction techniques are presented in this dissertation. These techniques open new aspects of research and potential investigations that were not previously available. This final chapter discusses some potential future research ideas.

5.1 Parameter-Independent Model Reduction and Sedimentary Texture Groundwater Flow Models

A common practice for numerical modeling of groundwater flow is to distribute hydraulic conductivity into an aggregate of model grid cells, called zones. The zonal hydraulic conductivity values are calibrated by an inverse procedure using water level observations. It has been shown in the literature that a highly discretized model can be reduced by three orders of magnitude through methods developed for model reduction. The most popular method for model reduction is based on the Galerkin projection of the high dimensional model equations onto a subspace, approximated by a small number of optimally chosen basis functions. Chapter 2 demonstrates that for a small number of zones, it is possible to develop a parameter-independent reduced model that will cover the entire parameter space in the original full-scale model. This is done by using basis functions from different combinations of parameter values. However, for a model with numerous zones it becomes infeasible to search for all parameter combinations.

To reduce the number of zones, current groundwater modeling efforts have been making the assumption that horizontal (K_h) and vertical (K_v) hydraulic conductivities are correlated to sediment texture. Sediment texture is defined as the fraction of coarse-grained and fine-grained

sediment. An example location is a texture defined as 45% coarse and 55% fine (must sum to 100). This assumption is based on the spatial correlation between saturated hydraulic conductivity and pore-size distributions in geologic media [1]. A method for estimating hydraulic conductivity based on this assumption has been applied successfully in previous groundwater-flow models of the central western San Joaquin Valley [2,3] and northeastern San Joaquin Valley [4,3]. The method uses the estimated sediment texture assigned to every model cell as a fraction of coarse and fine along with horizontal and vertical hydraulic conductivity estimates for each textural end member. The end members represent 100% coarse (K_c) and 100% fine (K_f) sediment.

Faunt and others [5] identify the power mean as a useful means for estimating hydraulic conductivity values. In addition, their work includes a review of the literature that describes the use of the power mean for calculating hydraulic conductivity. A power mean is a mean (M) of the form:

$$\mathrm{M}^{p}\left(x\right)=\left(\frac{1}{n}\sum_{k=1}^{n}x_{k}^{p}\right)^{1/p} \qquad (5.1)$$

where

p is the averaging power-mean exponent from range -1 to 1,

n is the number of elements being averaged, and

x_k is the k^{th} element in the list.

The power mean can take any number set and derive different statistical means based on the value of *p*. The most common values of *p* are -1, 0, 1 to produce the harmonic mean, geometric mean, and arithmetic mean, respectively. A partial mean can be derived by using a fractional value of *p*, such as 0.8 would produce a mean between an arithmetic and geometric.

The horizontal hydraulic conductivity ($K_{h,i}$) can be calculated as the weighted arithmetic mean (p = 1) of the hydraulic conductivities for each cell (i) of the coarse-grained (K_c) and fine-grained (K_f) lithologic end members and the distribution of sediment texture:

$$K_{h,i} = \left[K_c F_{c,i} + K_f F_{f,i} \right] \tag{5.2}$$

where:

$F_{c,i}$ is the fraction of coarse-grained sediment in a cell, estimated from sediment texture data as described in the previous section, and

$F_{f,i}$ is the fraction of fine-grained sediment in a cell ($F_{f,i} = 1 - F_{c,i}$).

Because K_f is much smaller that K_c, the arithmetic mean largely is influenced by the K and fraction of the coarse-grained end member.

Vertical hydraulic conductivity between layers ($K_{v,k+1/2}$) was calculated as the p^{th} weighted power mean of the hydraulic conductivities of the coarse-and fine-grained lithologic end members [5]:

$$K_{v,k+\frac{1}{2}} = \left[F_{c,k+\frac{1}{2}} K_c^p + F_{f,k+\frac{1}{2}} K_f^p \right]^{\frac{1}{p}} \tag{5.3}$$

where

$F_{c,k+\frac{1}{2}}$ is the fraction of coarse-grained sediment between layer midpoints, and

$F_{f,k+\frac{1}{2}}$ is the fraction of fine-grained sediment between layer midpoints.

Phillips and Belitz [2] determined that vertical conductivities could be calculated using either weighted harmonic or weighted geometric means. Belitz and others [3] represented the

vertical conductivities with the weighted harmonic mean. Faunt and others [5] calculated the vertical conductivities as power means in which *p* varied between –1.0 (the harmonic mean) and 0.0 (the geometric mean).

The relationship between hydraulic conductivity and the percentage coarse-grained deposits that are combined with the hydraulic conductivity end members and the exponent of the power mean is nonlinear. K_f is sensitive to the averaging method used. Both the harmonic and geometric means more heavily weight the fine-grained end member and as a result, the calculated vertical hydraulic conductivity is much lower than the horizontal.

The sedimentary texture formulation results in a semi-continuous representation of hydraulic conductivity that only requires calibrating the two end-member K values (K_c and K_f). This formulation changes the search space for the construction of the parameter-independent reduced model from the number of zones to just the two end-member K values. This would change the *Parameter Greedy Step* to the following:

$$K_{j+1} = [\mathrm{K}_{c,j+1}, \mathrm{K}_{\mathrm{f},j+1}] \in \mathbb{R}^2$$

$$\arg\max_{K_{j+1} \in \mathcal{K}} \mathcal{J}_{K_{j+1}} = \left\| \frac{h^{SS}\left(K_{j+1}\right)}{\left\|h^{SS}\left(K_{j+1}\right)\right\|} - \frac{\mathrm{P}_j h_r^{SS}\left(K_{j+1}\right)}{\left\|\mathrm{P}_j h_r^{SS}\left(K_{j+1}\right)\right\|} \right\|_2^2 . \tag{5.4}$$

Equation 5.4 searches for the two end members, K_c and K_f, within a range of reasonable values, $\mathcal{K}$, that maximizes the error between a normalized full model's steady state head, $h^{SS}(K_{j+1})$, and the normalized reduced model's steady state head, $\mathrm{P}_j h_r^{SS}(K_{j+1})$. This drastically reduces the search space in the construction of the parameter-impendent reduced model.

The proposed methodology can be tested on a sedimentary texture version of the Chapter 2 test case of the confined aquifer in Oristano, Italy. The Oristano model can be altered from its original zonation pattern by defining the 57,888 elements with a fraction of coarse- and fine-

grained sediment. This highly-heterogeneous model is reduced by selecting combinations of the end member coarse (K_c) and fine (K_f) hydraulic conductivity at a specific/constant power, *p*, of the power mean function. Using this parameter independent reduced model, the deterministic inverse problem and Bayesian inverse problem can be evaluated to test the accuracy of the reduced model and time savings.

5.2 Unconfined Model Reduction Using the Picard Method

Chapter 3 presents a new methodology that reduces the dimension of the Jacobian required to solve discretized, transient, unconfined groundwater flow via the Newton-Raphson method. Previously, unconfined groundwater has not been successfully combined with model reduction. The code discussed in Chapter 3 was implemented as a standalone package within the groundwater simulation software, MODFLOW-OWHM [6,7]. Another method of solving unconfined flow is to use Picard iterations until the solved heads match the heads used to satisfy the nonlinearity. In practice the Picard method is less stable than the Newton when modelled layers dewater (i.e. the simulated head is below the model cell bottom). However in practice if model layers do not dewater, then the Picard method is significantly faster than the Newton. Using the existing framework described in Chapter 3, it should be easy to modify the code to solve unconfined flow with the Picard solution and model reduction.

5.3 Optimal Snapshot Selection and Unconfined Flow

One issue found in Chapter 3 is that the snapshot selection is critical for producing accurate reduced models. The *Time Greedy Step* described in Chapter 2 presents a unique way of identifying greedy optimal snapshots. This method could be applied for selecting snapshots for the unconfined groundwater reduced ground model. Also since the original publication of Boyce

and Yeh [8], Chapter 2, there has been many new and more advanced releases of Genetic Algorithm (GA) codes that can handle more complex multi-objective optimizations. Given these GA codes the *Time Greedy Step* can be recast as follows:

$$\underset{Q_i \in \mathcal{Q},\, t_k \in \mathcal{T}}{\arg\max}\, \mathcal{J}_{Q_i} = \sum_{k=1}^{n} \left\| h^{t_k}\left(K_j, Q_i\right) - \mathrm{P}_j h_r^{t_k}\left(K_j, Q_i\right) \right\|_2^2 \qquad (5.5)$$

where each greedy stage, *i*, would select the greedy optimal pumping rate, *Q*, within the pumping range $\mathcal{Q}$, to take snapshots a user specified count of snapshots, *n*, within the time frame $\mathcal{T}$. Note that now the resulting hydraulic head, $h^{t_k}\left(K_j, Q_i\right)$, is a function of both the current hydraulic conductivity set, K_j, and the pumping rate. Since snapshot times are not added sequentially, this method will lose the greedy optimal minimum snapshot count. However, instead of using QR to evaluate P_j at each Greedy stage, this method would use singular value decomposition and percent energy to automatically trim the insignificant basis vectors. This will lead to excess snapshots being taken, but will find the greedy optimal pumping rate to collect snapshots from and provide an automatic framework for determining if multiple pumping rates should be employed to collect snapshots.

5.4 References

[1] Russo, D., & Bouton, M. (1992). Statistical analysis of spatial variability in unsaturated flow parameters. *Water Resources Research*, *28*(7), 1911-1925.

[2] Phillips, S. P., & Belitz, K. (1991). Calibration of a Texture-Based Model of a Ground-Water Flow System, Western San Joaquin Valley, Californiaa. *Groundwater*, *29*(5), 702-715.

[3] Phillips, S.P., Green, C.T., Burow, K.R., Shelton, J.L., & Rewis, D.L., 2007, *Simulation of multistage ground-water flow in part of the Northeastern San Joaquin Valley, California*: U.S. Geological Survey Scientific-Investigations Report 2007–5009, 43 p.

[4] Burow, K.R., Shelton, J.L., Hevesi, J.A., & Weissmann, G.S. (2004), *Hydrogeologic characterization of the Modesto area, San Joaquin Valley, California*: U.S. Geological Survey Scientific Investigations Report 2004–5232, 54 p

[5] Faunt, C.C., ed., 2009, *Groundwater Availability of the Central Valley Aquifer, California*: U.S. Geological Survey Professional Paper 1766, 225 p.

[6] Hanson, R.T., Boyce, S.E., Schmid, Wolfgang, Hughes, J.D., Mehl, S.M., Leake, S.A., Maddock, Thomas, III, and Niswonger, R.G. (2014), One-Water Hydrologic Flow Model (MODFLOW-OWHM): *U.S. Geological Survey Techniques and Methods*, 6–A51, 120 p., http://dx.doi.org/10.3133/tm6A51.

[7] Boyce, S. E., Nishikawa, T., & Yeh, W. W.G. (2015). Reduced order modeling of the Newton formulation of MODFLOW to solve unconfined groundwater flow. *Advances in Water Resources, 83*, 250-262, http://dx.doi.org/10.1016/j.advwatres.2015.06.005

[8] Boyce, S. E., & Yeh, W. W.G. (2014). Parameter-independent model reduction of transient groundwater flow models: Application to inverse problems. *Advances in Water Resources, 69*, 168-180, http://dx.doi.org/10.1016/j.advwatres.2014.04.009.