AGU PUBLICATIONS

Water Resources Research

REVIEW ARTICLE
10.1002/2015WR016967

Key Points:
- Three broad categories of surrogate model exist, each with their shortfalls
- Most methods common in water resources aren't applicable for large numbers of parameters
- A number of promising methods have not been applied to water resources models

Correspondence to:
M. J. Asher,
michael.asher@anu.edu.au

Citation:
Asher, M. J., B. F. W. Croke, A. J. Jakeman, and L. J. M. Peeters (2015), A review of surrogate models and their application to groundwater modeling, *Water Resour. Res.*, *51*, 5957–5973, doi:10.1002/2015WR016967.

Received 23 JAN 2015
Accepted 23 JUL 2015
Accepted article online 27 JUL 2015
Published online 21 AUG 2015

A review of surrogate models and their application to groundwater modeling

M. J. Asher[1,2], B. F. W. Croke[1,2], A. J. Jakeman[1], and L. J. M. Peeters[3]

[1]Integrated Catchment Assessment and Management, National Centre for Groundwater Research and Training, Australian National University, Canberra, Australian Capital Territory, Australia, [2]Mathematical Sciences Institute, Australian National University, Canberra, Australian Capital Territory, Australia, [3]CSIRO Land and Water, Adelaide, South Australia, Australia

Abstract The spatially and temporally variable parameters and inputs to complex groundwater models typically result in long runtimes which hinder comprehensive calibration, sensitivity, and uncertainty analysis. Surrogate modeling aims to provide a simpler, and hence faster, model which emulates the specified output of a more complex model in function of its inputs and parameters. In this review paper, we summarize surrogate modeling techniques in three categories: data-driven, projection, and hierarchical-based approaches. Data-driven surrogates approximate a groundwater model through an empirical model that captures the input-output mapping of the original model. Projection-based models reduce the dimensionality of the parameter space by projecting the governing equations onto a basis of orthonormal vectors. In hierarchical or multifidelity methods the surrogate is created by simplifying the representation of the physical system, such as by ignoring certain processes, or reducing the numerical resolution. In discussing the application to groundwater modeling of these methods, we note several imbalances in the existing literature: a large body of work on data-driven approaches seemingly ignores major drawbacks to the methods; only a fraction of the literature focuses on creating surrogates to reproduce outputs of fully distributed groundwater models, despite these being ubiquitous in practice; and a number of the more advanced surrogate modeling methods are yet to be fully applied in a groundwater modeling context.

1. Introduction

The physical properties and processes that determine groundwater flow are highly heterogeneous. To adequately capture such heterogeneity, many groundwater management problems require complex, fully distributed models that can accommodate fields for the hydraulic properties and boundary conditions that vary in time and space. There has been a tendency toward including more physical processes, increasing numerical resolution, and expanding the model domain in fully distributed groundwater models [*Leube et al.*, 2012; *Doherty and Simmons*, 2013], typically solved using a finite difference approximation such as implemented in MODFLOW [*Harbaugh*, 2005] or a finite element method, as used for example by FEFLOW [*Diersch*, 2005]. Greater conceptual model complexity, however, translates to a larger number of parameters and increased model runtimes.

Long runtimes inhibit the use of models in applications which require many model runs, such as integrated modeling (where groundwater flow models are coupled with models of different processes), uncertainty analysis, sensitivity analysis, and inverse modeling. Slow runtimes also prevent models being used in real time, necessary for applications such as decision support. Furthermore, the "curse of dimensionality" is encountered as the number of samples required to cover the parameter space in uncertainty analysis, sensitivity analysis, or calibration increases exponentially with the number of model parameters. An increase in model runtime means that numerical resolution needs to be reduced or physical processes ignored to decrease runtime and make many model runs computationally tractable. Surrogate models have the potential to speed up complex models without sacrificing accuracy or detail.

Also known as metamodels [e.g., *Blanning*, 1975], reduced models [e.g., *Willcox and Peraire*, 2002], model emulators [e.g., *O'Hagan*, 2006], proxy models [e.g., *Bieker et al.*, 2007], lower fidelity models [e.g., *Robinson et al.*, 2008], and response surfaces [e.g., *Regis and Shoemaker*, 2005], surrogate models are computationally cheaper models designed to approximate the dominant features of a complex model. While the main

Table 1. Taxonomy of Surrogate Models

Category	Also Known As	Examples
Data-driven surrogates involving empirical approximations of the complex model output calibrated on a set of inputs and outputs of the complex model (snapshots)	Response surface, statistical and black box methods	Polynomials [*Hussain et al*., 2002], neural networks [*Kourakos and Mantoglou*, 2009; *Yan and Minsker*, 2006], Gaussian processes [*Stone*, 2011; *Kennedy and O'Hagan*, 2001], kriging [*Baú and Mayer*, 2006; *Garcet et al*., 2006], radial basis functions [*Regis and Shoemaker*, 2005], support vector machines [*Yoon et al*., 2011], dynamic mode analysis [*Young and Ratto*, 2011], (generalized) polynomial chaos expansions [*Laloy et al*., 2013], genetic programming [*Fallah-Mehdipour et al*., 2013], Bayesian networks [*Fienen et al*., 2013]
Projection-based methods, where the governing equations are projected onto a reduced dimension subspace characterized by a basis of orthonormal vectors. Typically divided into SVD and Krylov-based methods	Reduced order, reduced basis and model reduction methods	Proper orthogonal decomposition (POD) [*McPhee and Yeh*, 2008; *Siade et al*., 2012; *Galbally et al*., 2010], Karhunen-Loève expansion [*Laloy et al*., 2013], proper generalized decomposition [*Chinesta et al*., 2011], Krylov subspace methods [*Dunbar and Woodbury*, 1989; *Woodbury et al*., 1990], dynamic mode decomposition [*Ghommem et al*., 2013], Fourier mode reduction [*Willcox and Megretski*, 2005] and (certified) reduced basis [*Lieberman et al*., 2010; *Chen et al*., 2010; *Knezevic and Peterson*, 2011; *Efendiev et al*., 2012]
Multifidelity based surrogates, built by simplifying the underlying physics or reducing numerical resolution	Multiscale, hierarchical and physically based methods	Multigrid method [*Ashby and Falgout*, 1996; *Saied and Mahinthakumar*, 1998; *Thum et al*., 2011], multiscale finite element method [*Shi et al*., 2012; *Hou and Wu*, 1997], heterogeneous multiscale method [*Weinan et al*., 2002], residual free bubbles [*Sangalli*, 2003], conservative subgrid [*Arbogast*, 2002] and variational multiscale method [*Ganapathysubramanian and Zabaras*, 2007; *Hughes et al*., 1998]

motivation for applying a surrogate model is achieving computational efficiency [*Razavi et al*., 2012a], other reasons do exist.

Simple surrogate models can reduce numerical instability which facilitates calibration and uncertainty analysis [*Doherty and Christensen*, 2011]. The process of building an emulator can reveal insensitive outputs and irrelevant parameters of a complex model [*Young and Ratto*, 2011]. Surrogates may serve as didactic tools for analyzing model simplification and the ways in which models simplify reality [*Watson et al*., 2013]. They can also be used to smooth an objective function surface, allowing the use of gradient based, nonlinear programming methods for optimization problems [*Hemker et al*., 2008; *Kavetski and Kuczera*, 2007], or to reduce ill conditioning of a conjugate gradient optimizer by using eigenvector approximations [*Vuik et al*., 1999]. The gain in computational efficiency opens the door for exploration of structural model uncertainty by simultaneous simulation and calibration of alternative model structures [*Matott and Rabideau*, 2008] or inclusion of data and physical processes at multiple scales [*Weinan and Engquist*, 2003]. Also, surrogates with sufficiently short runtimes have been used in interactive decision support environments [*Roach and Tidwell*, 2009]. Lastly, surrogates have been used for "complementary" modeling, where a simple model is fitted to the residual of a complex model to improve accuracy [*Demissie et al*., 2009; *Xu et al*., 2012].

This review is structured around a taxonomy of surrogate models, based on their mathematical structure. We follow *Robinson et al*. [2008] by classifying surrogate models into three categories as outlined in Table 1: data-driven surrogates involving empirical approximations of the complex model output calibrated on a set of inputs and outputs of the complex model (snapshots); projection-based methods, where the governing equations are projected onto a reduced dimensional subspace characterized by a basis of orthonormal vectors; and multifidelity methods, built by simplifying the underlying physics or reducing numerical resolution. The three categories are addressed in sections 2–4. Rather than describing each technique in detail, we give an illustrative example. We then discuss the other techniques in each category, how they relate to the example, and their historical and potential application to groundwater modeling. Since reducing runtime is a major motivation for using surrogate models, we discuss some alternative techniques for doing so in section 5.

Several recent publications compare a subset of surrogate modeling approaches [*Forrester and Keane*, 2009; *Castelletti et al*., 2012; *Frangos et al*., 2010; *Antoulas*, 2005; *Gugercin and Antoulas*, 2000]. *Razavi et al*. [2012a] published a review of surrogate models in the water resources literature. However, there appears a general imbalance in the literature toward data-driven methods. *Razavi et al*. [2012a], for example, devote 19 pages to data-driven methods, and only four to the section covering both multifidelity and projection-based

approaches. In particular, we note the sparse treatment of surrogates developed for spatially distributed models [*Pasetto et al*., 2011].

Two important motivations exist to correct this imbalance. First, *Razavi et al*. [2012b] have noted that despite the optimism about data-driven surrogate modeling in the literature, it can be an inefficient and unreliable approach to optimizing complex numerical models. Second, *Razavi et al*. [2012a] observe that "lower fidelity surrogates" (encompassing both projection-based and multifidelity methods reviewed here) overcome many of the limitations of data-driven approaches; namely, they can be applied to larger numbers of parameters, and perform better further from snapshots used in calibration of the surrogate. These advantages are particularly relevant to spatially distributed models which are ubiquitous in groundwater studies. Despite this, many sophisticated "lower fidelity" methods have not yet been applied to groundwater problems.

As well as summarizing the current state of the art in surrogate models, we aim to give a more comprehensive coverage of surrogate modeling approaches and discuss their historical, as well as their potential, application to distributed groundwater models. We aim to glean from the wider research community surrogate methods which show promise for groundwater applications.

Few well-accepted criteria exist for assessing surrogate modeling approaches. *Razavi et al*. [2012a] note the need to validate the whole surrogate-enabled analysis framework rather than testing the surrogate in isolation. They also note the need for more developed metrics of computational efficiency gains, and recommend approaches which make use of estimates of surrogate-introduced uncertainty. Several standards of good modeling practice [*Jakeman et al*., 2006] are particularly relevant to surrogate modeling. Approaches should be as simple as possible to avoid coding errors. Justification should be given for the choice of surrogate technique. Insofar as possible, surrogate performance should be thoroughly analyzed. Finally, results and methods should be reported in sufficient detail as to allow informed criticism.

We analyze the surrogate approaches below with reference to the following criteria.

1. If purported to do so, the approach should significantly increase computational efficiency. Ideally this would be assessed based on:
   i. Average runtime of a single surrogate versus complex model run.
   ii. The number of complex model runs used to calibrate the surrogate, justified with reference to surrogate-introduced uncertainty.
   iii. Runtime for combining surrogate and complex models, typically this involves (iteratively) calibrating the surrogate on the output of complex model runs.
   iv. An indication of effort required to apply the method. Does the complex model code typically require modification?
2. The surrogate should allow more thorough analysis and testing of the original model.
3. An indication of surrogate-introduced uncertainty should be given.

2. Data-Driven Methods

In this section, we introduce data-driven methods by describing a particular example, Gaussian processes. We then discuss other techniques and their application to groundwater modeling.

Let us denote the output of a complex model F as $\mathbf{h}=F(\boldsymbol{\theta})$. Consider a vector, $\boldsymbol{\theta}$, whose values represent a groundwater model's inputs and parameters: the hydraulic properties, sources, sinks, and initial conditions. Data-driven surrogates attempt to emulate the mapping from $\boldsymbol{\theta}$ to $\mathbf{h}$ without considering the inner workings of F. The complex model is run on a set $\{\boldsymbol{\theta}^{(i)}\}$, known as the design of experiment, to produce a set of snapshots, $\{\mathbf{h}^{(i)}\}$. Depending on the data-driven method, a function

$$\hat{F}(\boldsymbol{\phi},\cdot), \tag{1}$$

is chosen, and the snapshots are used to fit the hyperparameters $\boldsymbol{\phi}$ such that

$$\hat{F}(\boldsymbol{\phi},\ \boldsymbol{\theta}^{(i)}) \approx F(\boldsymbol{\theta}^{(i)}) \qquad \text{for all } i. \tag{2}$$

Data-driven methods differ in how they select the snapshots and the functional form of $\hat{F}$, chosen to emulate the $\boldsymbol{\theta}$ to $\mathbf{h}$ relationship.

2.1. Gaussian Processes—An Example

Gaussian processes [e.g., *Kennedy and O'Hagan*, 2001; *Stone*, 2011] assume the relationship can be captured by a surrogate

$$\hat{F}(\boldsymbol{\phi},\ \boldsymbol{\theta}) = \mathbf{f}(\boldsymbol{\theta})^T\ \boldsymbol{\beta} + e(\boldsymbol{\theta}), \tag{3}$$

where $\mathbf{f}(\boldsymbol{\theta})$ are chosen regression functions, and the unknown hyperparameters, $\boldsymbol{\phi}$, include the regression coefficients, $\boldsymbol{\beta}$, and the parameters of the stochastic process, e. The stochastic process, e, has zero mean and a specified positive semidefinite covariance function. A common choice is a linear regression function

$$\mathbf{f}(\boldsymbol{\theta}) = (1,\ \boldsymbol{\theta}), \tag{4}$$

and the covariance function

$$Cov(e(\boldsymbol{\theta}^{(j)}), e(\boldsymbol{\theta}^{(k)})) = \sigma^2 \exp\left(-\sum_i \omega_i |\theta_i^{(j)} - \theta_i^{(k)}|^{\alpha_i}\right). \tag{5}$$

Two of the hyperparameters, $\boldsymbol{\beta}$ and σ, can be estimated by analytic expressions using a set of snapshots [*Sacks et al.*, 1989], but numerical optimization of ω_i and α_i is required.

2.2. Other Data-Driven Methods

Razavi et al. [2012a] thoroughly cover data-driven surrogates with only a few exceptions, which we mention briefly below.

2.2.1. Bayesian Networks

Bayesian networks are annotated acyclic graphs which give the joint probability distributions of a number of variables. Nodes are random variables $X_i, \ldots, X_n$ and edges (links) are probabilities $P(x_i|\pi_i)$ where x_i is a realization of X_i, and π_i its parents. *Fienen et al.* [2013] create a Bayesian network surrogate of a groundwater model. In order to reduce the number of parameters, surrogate parameters are taken as the maximum or mean of a property along a cross section of the complex model domain. As is a requirement for Bayesian networks, parameters are discretized into bins. The approach makes causal relationships easy to see and can be integrated simply with larger Bayesian networks. Being largely a static method, their main restriction is the difficulty in emulating the temporally varying outputs.

2.2.2. Transfer Functions

Transfer function models consist of a function

$$h_k = F(h_{k-1}, \cdots, h_{k-n}, u_k, \cdots, u_{k-m}), \tag{6}$$

which predicts an output time series h_k in terms of its history and possibly a driving input series u_k. *Young and Ratto* [2011] present a transfer function-based method which, as a dynamic model, purportedly exhibits greater predictive capability for a wider range of scenarios than static approaches. Immediate drawbacks are the limitation to linear models, and difficulty in creating a map from the transfer function to complex model parameters.

2.2.3. Response Matrices

Cheng et al. [2011] construct a response matrix surrogate using "one at a time" sensitivity analysis of a finite difference groundwater model's output heads to changes in pumping and recharge. The sensitivity analysis is used to compute influence coefficients, $\frac{\partial h_{i,T}}{\partial Q_{p,t}}$, which represent the change in head at location i and time T caused by a change in pumping rate of well p at time t. Using these derivatives, head can then be approximated by a truncated Taylor series around $Q_{p,0}$.

2.2.4. Genetic Programming

Genetic programming [e.g., *Fallah-Mehdipour et al.*, 2013] seeks to find a functional form that approximates the input-output relationship of a model. An evolutionary algorithm searches through relationships randomly created using inputs, outputs, random variables, and operators (functions, arithmetic operators, boolean operators, or logical expressions). Disadvantages of genetic programming surrogates include computationally demanding calibration and overfitting.

2.2.5. Polynomial Chaos

The popular polynomial chaos method involves expanding a random variable (parameter or output) in an orthogonal polynomial basis. Any stochastic process with finite variance can be expanded as a polynomial of random variables so long as the polynomials are orthogonal with respect to distribution of that random variable. *Xiu and Karniadakis* [2002] describes the method and the Askey scheme, which details which polynomial basis is to be used with which distribution. A common example is the use of Hermite polynomials of Gaussian

processes (normally distributed variables). The approach has recently been incorporated into a two-step MCMC (Markov Chain Monte Carlo) Bayesian inversion of a groundwater flow model [*Laloy et al*., 2013]. As is common in the wider literature [e.g., *Ghanem and Dham*, 1998] the polynomial chaos surrogate is coupled with a Karhunen-Loève parameterization of conductivity. The most common approaches of finding the coefficients of the polynomials are stochastic collocation and pseudo-spectral methods, which are nonintrusive as they do not involve editing the complex model code. However, it is possible to use an intrusive Galerkin method which would make the method a projection-based approach [*Herzog et al*., 2008].

2.3. Comparisons and Recommendations

There are notable similarities between a number of the data-driven methods listed in Table 1. *Forrester and Keane* [2009] note that certain forms of kriging, Guassian process models, radial basis functions, support vector machines, and single-layer neural networks with radial coordinate neurons are identical.

While many methods do share mathematical structures, such as Guassian function of distance, as in (5), it would be a mistake to consider all methods labeled as "kriging" or "radial basis functions" to be identical. Each of the terms listed in Table 1 represents a sizable volume of literature devoted to developing a variety of methods known by that name.

Furthering the confusion, recent efforts have combined several data-driven surrogate methods. Rather than selecting a data-driven surrogate method a priori, *Viana et al*. [2009] suggest a framework for using multiple surrogates simultaneously. *Matott and Rabideau* [2008] propose a method for the simultaneous calibration of multiple models. In their test case employing multiple analytic surrogates, they improve the optimized objective function and reduce runtime. R. Schöbi et al. (Polynomial-chaos-based kriging, submitted to *International Journal for Uncertainty Quantification*, 2015) combine polynomial chaos expansions with kriging.

Comparisons of data-driven techniques in the literature are numerous [e.g., *Garcet et al*., 2006; *Forrester and Keane*, 2009; *Razavi et al*., 2012b; *Villa-Vialaneix et al*., 2012; *Espinet and Shoemaker*, 2013]. However, as *Forrester and Keane* [2009] note, no method performs best universally. Results will depend on the application and factors such as the size of training set [*Breiman*, 2001]. There is some consensus [*Forrester and Keane*, 2009; *Villa-Vialaneix et al*., 2012; *Razavi et al*., 2012a] about the strength of kriging [*Jones et al*., 1998] and radial basis function-based [*Regis and Shoemaker*, 2007] optimization frameworks. Subsequent literature [*Espinet and Shoemaker*, 2013] indicates the viability of these methods for groundwater modeling.

For calibration and uncertainty quantification, we would, in addition to kriging and radial basis functions, advocate for the use of polynomial chaos expansions. The polynomial chaos method of *Marzouk and Najm* [2009] and related stochastic partial differential equation approaches form a very active area of research. These methods are worthy of consideration since they have only recently been applied to groundwater models and few comparisons with other data-driven approaches have been published other than the theoretical work of *O'Hagan* [2013]. Advantages of polynomial chaos methods include the ability to calculate sensitivity indices of parameters and their interaction analytically from the expansions [*Sudret*, 2008] and the depth of literature devoted to selecting samples at which to calibrate the surrogate [e.g., *Xiu and Hesthaven*, 2005]. While it has been noted [*Forrester and Keane*, 2009; *Razavi et al*., 2012a] that it is better to select snapshots iteratively, taking into account surrogate error, many surrogates do not have established methods for selecting snapshots of the complex model.

Data-driven surrogates have a number of limitations. *Razavi et al*. [2012b] warn they may be subject to computationally demanding calibration, subjective structure, and overfitting. They are of the opinion that the optimism in the literature around data-driven surrogates is in some areas ill founded, pointing out that in some cases optimization methods without surrogates were more effective than those employing surrogates. It is worth noting that there are well-established methods for addressing issues such as overfitting [*Hastie et al*., 2009]. Other disadvantages include the possibility of getting trapped in local minima [*Demissie et al*., 2009] and the limitation of only being able to handle a relatively small number of parameters. This latter limitation often results in aquifer parameters being assumed homogeneous [*Mugunthan et al*., 2005] or known a priori [*Yan and Minsker*, 2006]. However, novel methods of calibrating polynomial chaos surrogates using adaptive sparse grids [*Jakeman and Roberts*, 2013] have been applied to increasingly large numbers of parameters.

Data-driven surrogates are not expected to perform well away from design sites. A known limitation of all global surrogates, where a single surrogate is used for the full parameter range of interest, is their inability

to adequately capture heterogeneity [*Najm*, 2009]. Solutions for polynomial chaos surrogates have been proposed involving piecewise polynomial bases [*Wan and Karniadakis*, 2005], multivariate wavelet bases [*Le Maitre et al*., 2004], or sparse grid collocation with local interpolants [*Matthies and Keese*, 2005; *Xiu and Hesthaven*, 2005]. *Marzouk and Najm* [2009] propose adaptive polynomial degree, sparse truncation of the basis, or partitioning the prior support as other possible improvements.

Despite their drawbacks, well used data-driven approaches remain a valuable tool in applications such as decision support and integrated modeling, where it may be necessary to limit both the number of parameters and the ranges which they take. Quick runtimes once calibrated and their nonintrusive nature make data-driven methods particularly useful for these applications.

3. Projection-Based Methods

The essence of projection-based surrogates is to replace a vector space, $\mathbf{h}$, by a linear combination of (orthogonal) basis vectors

$$\mathbf{h} \approx \mathbf{\Phi h}_r. \tag{7}$$

Complexity is reduced because the number of basis vectors (columns of $\mathbf{\Phi}$), and therefore elements in $\mathbf{h}_r$, needed to approximate $\mathbf{h}$ is often much smaller than the dimension of $\mathbf{h}$. Projection-based methods vary according to which vector space is approximated and how the basis vectors are found. The process invariably involves substituting $\mathbf{\Phi h}_r$ for $\mathbf{h}$ in the governing equations, and using an orthogonality condition to simplify the result.

3.1. Proper Orthogonal Decomposition—An Example

The proper orthogonal decomposition (POD) method (also referred to as singular value decomposition (SVD), principal component analysis, and empirical orthogonal functions) is a common projection-based surrogate. It proceeds as follows [*McPhee and Yeh*, 2008]. For a complex model F, we obtain a set of snapshots $\{\mathbf{y}^{(i)}\}$ from a set of samples of the input space $\{\boldsymbol{\theta}^{(i)}\}$ where $\mathbf{y}^{(i)}=F(\boldsymbol{\theta}^{(i)})$. A set of normalized snapshots is then combined in a matrix

$$\mathbf{Y}=\left[\frac{\mathbf{y}^{(1)}}{||\mathbf{y}^{(1)}||} \quad \frac{\mathbf{y}^{(2)}}{||\mathbf{y}^{(2)}||} \quad \cdots \quad \frac{\mathbf{y}^{(n)}}{||\mathbf{y}^{(n)}||}\right]. \tag{8}$$

The method requires the eigenvalues and eigenvectors of the covariance matrix $\mathbf{C}=\mathbf{YY}^T$. However, in practice the smaller $\mathbf{C}_s=\mathbf{Y}^T\mathbf{Y}$ is used to solve the eigenproblem

$$\mathbf{C}_s\mathbf{g}^{(i)}=\mathbf{g}^{(i)}\lambda^{(i)} \qquad i \in \{1,\cdots,n\}. \tag{9}$$

The eigenvectors $\mathbf{\Phi}$ of $\mathbf{C}$ are then computed as

$$\mathbf{\Phi}=\mathbf{Y}\left[\mathbf{g}^{(1)} \quad \mathbf{g}^{(2)} \quad \cdots \quad \mathbf{g}^{(n)}\right] \mathbf{\Lambda}^{-\frac{1}{2}}, \tag{10}$$

where

$$\mathbf{\Lambda}^{-\frac{1}{2}}=\begin{bmatrix}\sqrt{\lambda^{(1)}} & & & \\ & \sqrt{\lambda^{(2)}} & & \\ & & \ddots & \\ & & & \sqrt{\lambda^{(n)}}\end{bmatrix}^{-1}. \tag{11}$$

The dimensions of $\mathbf{\Phi}$ are the same as $\mathbf{Y}$. However, the proportion of the variance in $\mathbf{Y}$ explained by each eigenvector (column of $\mathbf{\Phi}$) is given by the normalized value of its corresponding eigenvalue

$$\psi^{(i)}=\frac{\lambda^{(i)}}{\sum_{j=1}^{n}\lambda^{(j)}}. \tag{12}$$

In practice, a minimum acceptable explained variance $\theta_{\min}$ can be chosen, and m identified such that $\sum_{i=1}^{m}\psi^{(i)} \geq \theta_{\min}$, where $\psi^{(i)}$ are arranged in descending order. A useful surrogate is created when m is orders of magnitude smaller than n.

In the case of confined aquifers (linear flow) where transmissivity does not vary in time, the discretized groundwater equation may be written [*Harbaugh*, 2005]

$$\mathbf{M}\frac{d\mathbf{h}}{dt}+\mathbf{A}\mathbf{h}=\mathbf{q}, \tag{13}$$

where $\mathbf{h}$ is head, $\mathbf{q}$ a source-sink term, and $\mathbf{M}$ and $\mathbf{A}$ are derived from hydraulic conductivity, storativity, and head-dependent boundary conditions.

Substituting $\mathbf{\Phi h}_r$ for $\mathbf{h}$ in (13), left multiplying with $\mathbf{\Phi}^T$, and noting $\mathbf{h}_r$ is time dependent but not $\mathbf{\Phi}$, we obtain

$$\mathbf{\Phi}^T\mathbf{M}\ \mathbf{\Phi}\frac{d\mathbf{h}_r}{dt}+\ \mathbf{\Phi}^T\mathbf{A}\ \mathbf{\Phi h}_r=\ \mathbf{\Phi}^T\mathbf{q}. \tag{14}$$

Although this reduced order equation can be solved using similar methods as those used for (13), this typically involves editing the complex model solvers.

3.2. The Lanczos Method—Another Example

As with the POD method described above, the Lanczos [*Dunbar and Woodbury*, 1989] approach is based on the reduced order equation (14). However, the basis vectors making up $\mathbf{\Phi}$ are computed differently. Rather than using the eigenvectors of the snapshot covariance as basis vectors, the Lanczos method uses the solution to the generalized eigenvalue problem

$$\mathbf{A}\ \mathbf{\Phi}=\mathbf{M}\ \mathbf{\Phi}\ \mathbf{\Lambda}, \tag{15}$$

where $\mathbf{A}$ and $\mathbf{M}$ are from (13), $\mathbf{\Phi}$ are the eigenvectors to be found, and $\mathbf{\Lambda}$ is the diagonal matrix of corresponding eigenvalues. Again, we only need find the m smallest eigenvalues and corresponding vectors. Using the orthogonality properties of the eigenvectors, $\mathbf{\Phi}^T\mathbf{M}\ \mathbf{\Phi}=\mathbf{I}$ and $\mathbf{\Phi}^T\mathbf{A}\ \mathbf{\Phi}=\ \mathbf{\Lambda}$, (14) can be reduced to

$$\frac{d\mathbf{h}_r}{dt}+\ \mathbf{\Lambda h}_r=\ \mathbf{\Phi}^T\mathbf{q}. \tag{16}$$

If $\boldsymbol{g}=\mathbf{\Phi}^T\mathbf{q}$ is time independent, this has the analytic solution

$$(h_r)_i(t)=(h_r)_i^0e^{-\lambda_i t}+\frac{g_i}{\lambda_i}(1-e^{-\lambda_i t}). \tag{17}$$

A time stepping algorithm can be applied for calculating time dependent $\boldsymbol{g}$ as done by *Sahuquillo* [1983]. However, finding the m smallest eigenvalues of $\mathbf{A}\ \mathbf{\Phi}=\mathbf{M}\ \mathbf{\Phi}\ \mathbf{\Lambda}$ is a computationally challenging problem.

Note that we can rewrite (15) as

$$\mathbf{\Phi}\ \mathbf{\Lambda}^{-1}=\mathbf{A}^{-1}\mathbf{M}\ \mathbf{\Phi}. \tag{18}$$

The Lanczos algorithm seeks the solution to the tridiagonal system

$$\mathbf{LT}=\mathbf{A}^{-1}\mathbf{ML}, \tag{19}$$

by finding the tridiagonal matrix

$$\mathbf{T}=\begin{bmatrix}\alpha_1 & \beta_2 & & & \\ \beta_2 & \alpha_2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_m \\ & & & \beta_m & \alpha_m\end{bmatrix}, \tag{20}$$

and the so-called Lanczos vectors $\mathbf{L}=(L_1, L_2, \cdots, L_m)$.

The eigenvalues of $\mathbf{T}$ give good approximations to the smaller eigenvalues of $\mathbf{A}\ \mathbf{\Phi}=\mathbf{M}\ \mathbf{\Phi}\ \mathbf{\Lambda}$. The strength of the approach lies in the efficient algorithm for computing L_i, α_i and β_i using a standard tridiagonal solution algorithm. Similarly to above, substituting $\mathbf{Lh}_r$ for $\mathbf{h}$ in (13) and left multiplying with $\mathbf{L}^T\mathbf{MA}^{-1}$, gives

Table 2. Methods for Computing a Reduced Basis

Method	Basis	Reference
POD	SVD of snapshot covariance	*McPhee and Yeh* [2008]
Proper generalized decomposition	Separation of variables	*Chinesta et al.* [2011]
Dynamic mode decomposition	SVD of forward operator	*Ghommem et al.* [2013]
Lanczos and Arnoldi	Krylov subspace	*Dunbar and Woodbury* [1989] and *Woodbury et al.* [1990]
Fourier model reduction	Fourier expansion of the discrete-frequency transfer function	*Willcox and Megretski* [2005] and *Gugercin and Willcox* [2008]
(Certified) reduced basis	Greedy algorithm	*Lieberman et al.* [2010], *Chen et al.* [2010], and *Knezevic and Peterson* [2011]

$$\mathbf{L}^T\mathbf{M}\mathbf{A}^{-1}\mathbf{M}\mathbf{L}\frac{d\mathbf{h}_r}{dt}+\mathbf{L}^T\mathbf{M}\mathbf{L}\mathbf{h}_r=\mathbf{L}^T\mathbf{M}\mathbf{A}^{-1}\mathbf{q}, \tag{21}$$

which can be reduced to

$$\mathbf{T}\frac{d\mathbf{h}_r}{dt}+\mathbf{h}_r=\mathbf{L}^T\mathbf{M}\mathbf{A}^{-1}\mathbf{q}, \tag{22}$$

using the orthogonality properties $\mathbf{L}^T\mathbf{M}\mathbf{A}^{-1}\mathbf{M}\mathbf{L}=\mathbf{T}$ and $\mathbf{L}^T\mathbf{M}\mathbf{L}=\mathbf{I}$. A solution for $\mathbf{h}_r$ can then be found using standard time integration techniques.

3.3. Other Projection-Based Methods

Other reduced basis methods employ different techniques to find bases which reduce the dimensionality of the parameter or output space. Table 2 lists a number of different approaches.

We note here that projection-based surrogates, despite employing many of the same techniques, differ from parameterization methods. In a discussion of parameterization methods for calibrating reservoir models, *Oliver and Chen* [2011] mention zonation, pilot points, splines, Karhunen-Loève approximations, wavelets, and singular vectors of the model sensitivities. While these might be considered surrogate parameterizations, they do not emulate the model. The distinguishing feature of the work listed in Table 2 is that each uses an intrusive (model driven) method to derive an orthogonal basis.

3.4. Comparisons and Recommendations

Antoulas [2005] and *Gugercin and Antoulas* [2000] divide "reduced models" into two categories: SVD based (e.g., POD) and Krylov subspace based. While SVD-based methods have an error bound they cannot be applied to highly complex systems. *Antoulas* [2005] argue that Krylov methods can be implemented iteratively and so are more appropriate for systems of high complexity. He describes a method combining the two categories designed to overcome the limitations of both. However, *Frangos et al.* [2010] claim that Krylov methods are limited to linear cases. There seems to be very limited application of Krylov methods to groundwater, some being the Lanczos algorithm by *Dunbar and Woodbury* [1989] and the Arnoldi method by *Woodbury et al.* [1990].

There has been more attention paid to the POD method which, for a given dimension, minimizes the least squares error of the surrogate on the snapshots [*Frangos et al.*, 2010]. *Hay et al.* [2012] note that variation in parameters reduces the effectiveness of a POD surrogate. They propose replacing each mode with a first-order Taylor expansion. A finite difference method is used to compute the sensitivity of the mode to variations in each parameter. *Ghommem et al.* [2013] compare POD and dynamic mode decomposition methods and find the latter basis makes more accurate predictions for a variety of parameter and boundary condition values. *Chinesta et al.* [2011] advocate for proper generalized decomposition based on empirical studies, but note the lack of rigorous mathematical foundations. *Nouy* [2007] note that the method allows greater online computational savings compared to POD.

For a linear system, *Willcox and Megretski* [2005] compare their approach to POD and Arnoldi methods, and demonstrate superior computational efficiency and error bounds over a wide range of frequencies. *Chen et al.* [2010] points out that POD-based methods have only the largest ignored singular value as an error approximation, and Krylov methods have no estimate at all. He advocates the certified reduced basis method employing the "greedy" algorithm [e.g., *Knezevic and Peterson*, 2011]), so called because it iteratively adds basis vectors in the direction of the residual. *Buffa et al.* [2012] provide the proof underlying the

greedy algorithm's error estimate. *Lieberman et al.* [2010] employ the reduced basis method on the parameter and state spaces of a groundwater inverse problem.

Rewieński and White [2006] propose a method which combines a reduced basis and a quasi piecewise linear approximation of the state function for a nonlinear differential equation. It is claimed the approach outperforms POD on their case studies.

The projection-based approach of *Sahuquillo* [1983] has been applied by *Andreu et al.* [1996] to provide real-time spatially distributed groundwater flow modeling for decision support.

Projection-based methods have two principal drawbacks: the basis vectors depend on the snapshots used to compute them, making inverse modeling and uncertainty analysis difficult, and solving the reduced model typically involves editing the model code. To address the former issue, recent work aims at selecting optimal snapshots which cover both the parameter space and the time domain. *Lieberman et al.* [2010] develop an approach for the reduced basis method, *Siade et al.* [2010] and *Baú* [2012] for POD, and *Pasetto et al.* [2013, 2014] and *Boyce and Yeh* [2014] for syntheses of the two methods. For confined aquifers, reduced models have been demonstrated to run orders of magnitude faster than the full order equivalent. However, the approach does not yield similar results in nonlinear problems as many more basis vectors need to be included [*Cardoso et al.*, 2009].

4. Multifidelity Methods

Multifidelity surrogates refer to those constructed from the complex model by reducing numerical resolution, increasing tolerances, or removing processes. Perhaps the simplest surrogate possible is created by reducing the numerical discretization of the complicated model. Previous research has focused on how to upscale properties from the scale of measurements or a fine grid to a coarse grid for rapid computation. However, *Farthing et al.* [2012] note that underresolved models can produce inaccurate objective function values and may produce "false" solutions.

The disadvantages of simply reducing resolution and the requirement for finer detail has led to multifidelity methods that attempt to combine models at multiple levels of complexity to attain the detail of the complex, at the speed of the simple. This typically involves solving the global problem on a coarse grid, along with multiple local problems on a finer grid. Methods often differ in how they relate the results; in particular, in how they set the boundary conditions for the local problems.

Weinan and Engquist [2003] differentiate between homogeneous multiscale methods, which use identical models at different scales and heterogeneous methods, which allow for different processes at each scale; for example, Darcy's law at one scale, and lattice Boltzmann pore-scale effects at another. In this review we consider only homogeneous multiscale methods, which we refer to as multifidelity methods as they can be more readily classified as surrogates. While they do not have the advantage of incorporating multiple physical processes, they do still simplify the inclusion of data at multiple scales, possibly negating the need for upscaling.

4.1. The Multiscale Finite Volume Method—An Example

Consider the groundwater flow equation (the continuous form of (13))

$$\nabla \cdot (K \nabla h) = S_s \frac{\partial h}{\partial t} + q. \tag{23}$$

For the multiscale finite volume method (MsFV) developed by *Jenny et al.* [2003] (for a detailed explanation, see *Hajibeygi et al.* [2008]), the model domain Ω is divided into a coarse grid with M control volumes (grid cells) $\bar{\Omega}_m$ and a dual coarse grid, formed by joining the midpoints of the control volumes of the coarse grid. The dual grid has N control volumes $\tilde{\Omega}_n$. Head on the fine grid is computed as

$$h(\mathbf{x}) \approx \sum_{n=1}^{N} \left(\sum_{m=1}^{M} \phi_{m,n}(\mathbf{x}) \bar{h}_m + \phi_n^*(\mathbf{x}) \right), \tag{24}$$

where basis functions $\phi_{m,n}$ and correction functions ϕ_n^* are computed by solving the following systems of equations for each dual grid cell [*Lunati and Jenny*, 2008]:

$$\begin{aligned} &\nabla \cdot (K\nabla\phi_{m,n}) = 0 && x \in \tilde{\Omega}_n \\ &(\mathbf{n}\cdot\nabla)((K\nabla\phi_{m,n})\cdot\mathbf{n}) = 0 && x \in \partial\tilde{\Omega}_n, \\ &\phi_{m,n}(x_i) = \delta_{ni} && \text{at dual grid point } x_i \end{aligned} \tag{25}$$

$$\begin{aligned} &\nabla \cdot (K\nabla\phi_n^*) = R_n && x \in \tilde{\Omega}_n \\ &(\mathbf{n}\cdot\nabla)((K\nabla\phi_n^*)\cdot\mathbf{n}) = R_n^* && x \in \partial\tilde{\Omega}_n, \\ &\phi_n^*(x_i) = 0 && \text{at dual grid point } x_i \end{aligned} \tag{26}$$

where R_n and R_n^* are some specified functions. In practice, good results are achieved with $R_n = 0$, but more advanced techniques exist [e.g., *Hajibeygi et al*., 2008]. In a typical finite volume fashion, (24) can then be substituted into (23), the resulting equation integrated over $\tilde{\Omega}_m$ and Gauss' theorem applied to find the coarse grid heads $\bar{h}_m$. The MsFV method has potential for wide-scale application where it is to be included in the recently released finite volume package MODFLOW-USG [*Panday et al*., 2013].

4.2. Other Multifidelity Methods

In this section we discuss the application of several multifidelity methods to groundwater modeling. The multiscale finite element method (MsFEM) [*Efendiev and Hou*, 2007] proceeds similarly to the above MsFV approach, but without the correction functions. The approach has recently been applied to groundwater flow by *He et al*. [2013]. Recent work [*Efendiev et al*., 2012, 2013] has combined MsFEM with projection-based approaches to reduce online runtime. *Sun* [2008] reviews a number of multiscale methods for groundwater modeling including the multiscale finite element method (MsFEM), the mixed MsFEM, subgrid upscaling, mixed mimitec multiscale methods for corner-point grids, the stochastic variation multiscale method, the MsFV method, and ghost node local grid refinement. All but the latter two are mixed finite element methods. Although MsFEM and MsFV methods are not typically more computationally efficient than the corresponding fine grid solution, they allow greater parallelization [*Sun*, 2008].

Multigrid [e.g., *Bastian and Reichenberger*, 2000] and adaptive mesh refinement methods [e.g., *Mansour and Spink*, 2013] allow the problem to be solved on a hierarchy of resolutions, providing methods for interpolating between multiple scales on the same computational domain. Similar to traditional multigrid methods, the ghost node local grid refinement method involves an iterative interpolation of head from the coarse-grid solution to local fine-grid boundaries, and flux solutions in the opposite direction. The approach is widely available as the MODFLOW-LGR package [*Mehl and Hill*, 2005]. *Vilhelmsen et al*. [2012] conclude that the method is more efficient than a uniform fine grid only when the area of refinement covers less than 10–15% of the total model.

Weinan et al. [2007] refer to a number of other classical and recent multiscale methods. Domain decomposition aims to allow the problem to be solved independently on a number of subdomains to facilitate parallel computation. Wavelet methods decompose a model into its components of different frequencies, allowing each component to be computed using the appropriate resolution.

In certain cases a simpler model which has a different structure or ignores certain physical processes may be available. For example, *Keating et al*. [2010] build an ad hoc analytic surrogate to predict influences on groundwater head of nuclear tests. Although potentially useful, the approach is not generic enough to be classified and contrasted against more general methods.

4.3. Comparisons and Recommendations

The multifidelity surrogates discussed above aim to solve the forward model, and can therefore be substituted for the complex model in any application. While they are intrusive, the MsFEM, MsFV, and multigrid methods maintain the detail and accuracy of the complex model. This makes them ideal for those implementing groundwater modeling codes, but less relevant to practitioners.

Related to such forward model surrogates, several approaches have been developed to use multifidelity models in optimization and uncertainty analysis. *Robinson et al*. [2008] aim to develop a general framework for optimization using multiple models with different sets of parameters. A space mapping, linking low to high-fidelity parameters is created by varying low-fidelity parameters to match high-fidelity output. A common uncertainty analysis approach is to use a coarse model to increase acceptance rates of MCMC during

Bayesian inversion of a complex model [*Efendiev et al*., 2005]. Samples are evaluated by the complex model only if accepted by the surrogate. *Cliffe et al*. [2011] apply the multilevel Monte Carlo method to groundwater flow. *Narayan et al*. [2014] develop an uncertainty quantification method based on multifidelity stochastic collocation which uses low-fidelity results to inform sampling locations for the high-fidelity model.

Doherty and Christensen [2011] advocate the use of both simple and complex models. Complex models allow the use of expert information in prior distributions, since their parameters correspond more readily to physical quantities. Simple models allow numerical stability and efficient calibration. They develop a method to compare the results of simple and complex models to infer whether errors are due to dependency on the complex model null-space, measurement noise in the calibration data, or surrogate parameters compensating for the structural simplification. Results allow the correction of predictions of a simple model calibrated on field data. A case study is presented using two finite difference models of different grid resolutions.

If different fidelity models (e.g., a complex model and a data-driven surrogate or the same model at different resolutions) exist, these multifidelity inverse modeling methods are particularly attractive since they can be applied nonintrusively. As noted by *Robinson et al*. [2008], such approaches require a mapping from surrogate to high-fidelity parameters. For groundwater models, even upscaling parameters from a fine to a coarse grid is no simple matter [*Wen and Gómez-Hernández*, 1996; *Vermeulen et al*., 2006; *Mehl and Hill*, 2010]. To what extent multifidelity models can be used in the aforementioned inverse modeling frameworks, using simple and practical mappings, remains an open question.

5. Decreasing Runtime

Decreasing model runtime can be a major motivation for employing one of the surrogates mentioned in sections 2–4. In this section we note some of the techniques available for reducing model runtime other than employing a surrogate.

5.1. A Simpler Model

Determining appropriate model complexity is a complex issue. At one extreme, proponents of analytic models [*Matott et al*., 2006; *Craig and Read*, 2010; *Estabragh et al*., 2013] espouse not just their fast runtimes, but other advantages such as numerical stability. Others, such as *Doherty and Christensen* [2011] and *Miller et al*. [2013], argue for the many advantages of increasing model complexity. In any case, computational expense should certainly feature as a consideration when determining model complexity [*Hill*, 2006].

5.2. Parameterization

Rather than simplifying the model itself, a common approach is to reduce the number of parameters, thus reducing the number of samples necessary for uncertainty analysis or calibration. A number of parameterization methods are common in practice; for example, zones, pilot points, splines, Karhunen-Loève, wavelets, and singular vectors of the model sensitivities [*Oliver and Chen*, 2011]. Such approaches are often coupled with surrogate modeling techniques. *Laloy et al*. [2013], for example, combine a polynomial chaos surrogate with a Karhunen-Loève parameterization of conductivity.

5.3. Uncertainty Analysis Algorithm

Surrogates are commonly used to accelerate uncertainty quantification and the choice of uncertainty analysis algorithm may also have a large effect on total runtime by reducing the number of model runs required. A large number of uncertainty analysis methods exist. *Matott et al*. [2009] for example, evaluate 65 software packages for evaluating model uncertainty. *Mariethoz et al*. [2010] propose a method, iterative spatial resampling, that requires fewer model runs than MCMC but yields "reasonably" similar posterior distributions. *Franssen and Kinzelbach* [2009] compare ensemble Kalman filtering to a Monte Carlo-based inversion algorithm and note a significant speed up. The field of stochastic partial differential equations provides many alternatives to sampling. For example, moment differential equations [*Winter et al*., 2003] involve solving the flow equation for as many probability moments as one is interested in. Perturbation-based solutions exist for the first and second moments of head and flux [*Guadagnini and Neuman*, 1999]. The wide-scale application of the approach is limited by its ability to account for highly heterogeneous media. *Vrugt et al*. [2003] combine the Metropolis algorithm with an evolutionary approach to reduce the number of model runs required for Bayesian inversion compared to traditional MCMC. *Shafii et al*. [2014] suggest that simply

relaxing the convergence criterion of a MCMC sampler may yield sufficiently accurate uncertainty estimates at a fraction of the computational cost.

5.4. Optimization Algorithm

As is the case with uncertainty analysis, an apt choice of algorithm may reduce the number of forward runs necessary to calibrate a model. The SVD-assist [*Tonkin and Doherty*, 2005], adjoint sensitivity [*LaVenue and Pickens*, 1992], and Principal Component Geostatistical [*Kitanidis and Lee*, 2014] approaches all claim to reduce the computational burden of the inverse problem. *Franssen et al*. [2009], *Oliver and Chen* [2011], and *Zhou et al*. [2014] review some of the many optimization algorithms available to modelers.

5.5. Computational Techniques

If the aim of a surrogate model is simply to reduce runtime, improving computational techniques may be a viable alternative. Relevant topics, including choice of solvers, programming language, and parallelization techniques are covered by *Miller et al*. [2013]. Seemingly banal choices, such as that of compiler, may have significant implementations. For example, Intel compiled MODFLOW can be almost eight times faster than the gfortran version according to *Dong and Li* [2009].

Of particular note, parallelizing the forward model can reduce runtimes for a variety of applications. HYDROLAB [*Erhel et al*., 2009], ParFlow [*Ashby and Falgout*, 1996], and PFLOTRAN [*Mills et al*., 2009] are examples of softwares developed explicitly for highly parallelized groundwater modeling. Linear reduction of parallel walltime versus the number of processors, up to 27,580 cores, has been achieved using PFLOTRAN. By comparison, the most significant result in the data-driven surrogate approaches reviewed by *Razavi et al*. [2012a] is a 97% reduction in complex model evaluations reported by *Regis and Shoemaker* [2012]. Ignoring time to develop and run the surrogate, this time saving could be achieved by parallelization of the original model on a 30 core machine. Parallel implementations of the algebraic multigrid solver [*Thum et al*., 2011] exist for both MODFLOW and FEFLOW. *Fienen and Hunt* [2015] outline approaches for further parallelizing high throughput applications, where it is unnecessary for computations to interact when running, such as uncertainty analysis and calibration.

6. Further Research

Based on our review of the methods above, we consider that a number of areas warrant further research. The limitations of global surrogates have been identified. In particular there is a need to deal with complex fields with discontinuities. Localization techniques attempt to overcome these difficulties by dividing parameter space into a number of subdomains, and employing a different surrogate on each. Such methods are still in their infancy, but are receiving increasing attention (see polynomial chaos references in section 2.3). While localized parameterizations have been developed for groundwater models [e.g., *Nan and Wu*, 2011], localized surrogates, to our knowledge, have not.

Quantification of the uncertainty introduced by the surrogate model is another nascent field. Many approaches have not yielded rigorous bounds on surrogate computed posteriors. Application of, and comparison between, approaches is made difficult by the lack of an established measure of surrogate-induced uncertainty [*Chen et al*., 2010], and surrogate-enabled runtime reduction [*Razavi et al*., 2012b].

Multiscale methods which preserve the accuracy and detail of the complex forward model hold promise for implementation in industry groundwater modeling codes. For example, as mentioned in section 4, there is the possibility of a MsFV or MsFEM version of the numerous finite volume and finite element codes in widespread use. Further investigation is required into the possible increases in computational efficiency of each multiscale method.

Many data-driven methods rely on ad hoc approaches to select snapshots on which to calibrate the surrogate. Further research is warranted into the application to groundwater model surrogates of the innovative snapshot selection methods from the multifidelity [e.g., *Narayan et al*., 2014] and projection-based [e.g., *Pasetto et al*., 2014] literature. Although these approaches were developed in an uncertainty analysis framework, they could as easily be applied to snapshot selection for integrated modeling or decision support.

In addition to the above areas which require further research, a number of promising techniques appear as active areas of research in the literature. For a small number of parameters, radial basis function and

kriging-based optimizers [*Regis and Shoemaker*, 2012] have been shown to compare favorably to industry standard methods such as gradient-based PEST [*Espinet and Shoemaker*, 2013]. Tensor-based sparse grid collocation methods for polynomial chaos-based uncertainty analysis [*Espig et al*., 2013; *Jakeman and Roberts*, 2013] deserve attention, as it is a rapidly improving field of research with the ability to handle increasing numbers of parameters. Parameter-independent projection-based methods [*Boyce and Yeh*, 2014; *Pasetto et al*., 2014] have been shown capable of significant runtime reductions while maintaining spatially distributed parameters and outputs, making them applicable to inverse modeling along with other applications.

7. Conclusion

The purpose of this review is to summarize approaches to surrogate modeling which are applicable to groundwater modeling. As has been already indicated, no surrogate method is universally superior. We conclude by summarizing our findings on surrogate models appropriate to different use cases.

Multiscale and parameter independent projection-based methods have potential to replace groundwater models in any context, since they can emulate the full output of a complex model. However, their likely application is by developers of groundwater model codes, rather than everyday users, because they are intrusive. If the aim of a surrogate model is simply to reduce runtime, we add that computational techniques may be a viable alternative in any context.

For decision support, where very short runtimes are required, data-driven approaches are the obvious choice. Since decision support typically involves a small number of parameters, many of the drawbacks of data-driven methods are irrelevant. The only alternatives are certain projection-based approaches which only require the once off computation of basis vectors.

In the case of inverse modeling, we echo previous concerns that despite the ease of use and popularity of data-driven methods, they have well-established limitations and should be used with care. Preference should be given to approaches which have been validated in the context in which they are to be used. For example, the kriging and radial basis functions enabled optimization, or polynomial chaos expansion uncertainty analysis methods mentioned in section 6. Projection-based approaches have only recently been applied to the inverse modeling of groundwater models, with promising results for runtime reduction. Doubtless, the benefits and disadvantages of these methods will be exposed more clearly as they are further explored. Multifidelity inverse modeling methods (discussed in section 4.3) offer potential for both direct application and future development. Again, they have only recently been applied to groundwater models, but hold great promise. Developing maps from low to high-fidelity models remains the biggest hurdle to their widespread application.

Acknowledgments
Many thanks go to the three reviewers who provided many insightful comments, and greatly improved the quality of the manuscript. The first author would like to thank the Department of Industry and CSIRO Land and Water for their financial support. While this manuscript does not contain any data, if the reader has any data-related inquiries please contact the corresponding author Michael Asher (michael.asher@anu.edu.au).

References

Andreu, J., J. Capilla, and E. Sanchís (1996), Aquatool, a generalized decision-support system for water-resources planning and operational management, *J. Hydrol.*, *177*(3), 269–291.
Antoulas, A. (2005), An overview of approximation methods for large-scale dynamical systems, *Annu. Rev. Control*, *29*(2), 181–190, doi: 10.1016/j.arcontrol.2005.08.002.
Arbogast, T. (2002), Implementation of a locally conservative numerical subgrid upscaling scheme for two-phase Darcy flow, *Comput. Geosci.*, *6*(3–4), 453–481.
Ashby, S. F., and R. D. Falgout (1996), A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow simulations, *Nucl. Sci. Eng.*, *124*(1), 145–159.
Bastian, P., and V. Reichenberger (2000), Multigrid for higher order discontinuous Galerkin finite elements applied to groundwater flow, technical report 2000–37, Univ. Heidelberg, Heidelberg.
Baú, D. A. (2012), Planning of groundwater supply systems subject to uncertainty using stochastic flow reduced models and multiobjective evolutionary optimization, *Water Resour. Manage.*, *26*(9), 2513–2536.
Baú, D. A., and A. S. Mayer (2006), Stochastic management of pump-and-treat strategies using surrogate functions, *Adv. Water Resour.*, *29*(12), 1901–1917.
Bieker, H. P., et al. (2007), Real-time production optimization of oil and gas production systems: A technology survey, *SPE Prod. Oper.*, *22*(04), 382–391.
Blanning, R. W. (1975), The construction and implementation of metamodels, *Simulation*, *24*(6), 177–184.
Boyce, S. E., and W. W.-G. Yeh (2014), Parameter-independent model reduction of transient groundwater flow models: Application to inverse problems, *Adv. Water Resour.*, *69*, 168–180.
Breiman, L. (2001), Random forests, *Mach. Learning*, *45*(1), 5–32.
Buffa, A., Y. Maday, A. T. Patera, C. Prudhomme, and G. Turinici (2012), A priori convergence of the greedy algorithm for the parametrized reduced basis method, *ESAIM: Math. Modell. Numer. Anal.*, *46*(03), 595–603.

Cardoso, M., L. Durlofsky, and P. Sarma (2009), Development and application of reduced-order modeling procedures for subsurface flow simulation, *Int. J. Numer. Methods Eng.*, *77*(9), 1322–1350.
Castelletti, A., S. Galelli, M. Ratto, R. Soncini-Sessa, and P. Young (2012), A general framework for dynamic emulation modelling in environmental problems, *Environ. Modell. Software*, *34*, 5–18.
Chen, Y., J. S. Hesthaven, Y. Maday, and J. Rodriguez (2010), Certified reduced basis methods and output bounds for the harmonic Maxwell's equations, *SIAM J. Sci. Comput.*, *32*(2), 970–996, doi:10.1137/09075250X.
Cheng, W.-C., M. Putti, D. R. Kendall, and W. W.-G. Yeh (2011), A real-time groundwater management model using data assimilation, *Water Resour. Res.*, *47*, W06528, doi:10.1029/2010WR009770.
Chinesta, F., P. Ladeveze, and E. Cueto (2011), A short review on model order reduction based on proper generalized decomposition, *Arch. Comput. Methods Eng.*, *18*(4), 395–404.
Cliffe, K., M. Giles, R. Scheichl, and A. L. Teckentrup (2011), Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients, *Comput. Visualization Sci.*, *14*(1), 3–15.
Craig, J. R., and W. W. Read (2010), The future of analytical solution methods for groundwater flow and transport simulation, in *XVIII International Conference on Computational Methods in Water Resources*, edited by J. Carrera, pp. 1–8, Int. Cent. for Numer. Methods in Eng., Barcelona.
Demissie, Y. K., A. J. Valocchi, B. S. Minsker, and B. A. Bailey (2009), Integrating a calibrated groundwater flow model with error-correcting data-driven models to improve predictions, *J. Hydrol.*, *364*(3), 257–271.
Diersch, H. (2005), FEFLOW finite element subsurface flow and transport simulation system, in *Reference Manual*, 292 pp., WASY GmbH Inst. Water Resour. Plann. Syst. Res., Berlin.
Doherty, J., and S. Christensen (2011), Use of paired simple and complex models to reduce predictive bias and quantify uncertainty, *Water Resour. Res.*, *47*, W12534, doi:10.1029/2011WR010763.
Doherty, J., and C. Simmons (2013), Groundwater modelling in decision support: Reflections on a unified conceptual framework, *Hydrogeol. J.*, *21*(7), 1531–1537, doi:10.1007/s10040-013-1027-7.
Dong, Y., and G. Li (2009), A parallel PCG solver for MODFLOW, *Groundwater*, *47*(6), 845–850.
Dunbar, W. S., and A. D. Woodbury (1989), Application of the Lanczos algorithm to the solution of the groundwater flow equation, *Water Resour. Res.*, *25*(3), 551–558.
Efendiev, Y., and T. Hou (2007), Multiscale finite element methods for porous media flows and their applications, *Appl. Numer. Math.*, *57*(5), 577–596.
Efendiev, Y., A. Datta-Gupta, V. Ginting, X. Ma, and B. Mallick (2005), An efficient two-stage Markov chain Monte Carlo method for dynamic data integration, *Water Resour. Res.*, *41*, W12423, doi:10.1029/2004WR003764.
Efendiev, Y., J. Galvis, and F. Thomines (2012), A systematic coarse-scale model reduction technique for parameter-dependent flows in highly heterogeneous media and its applications, *Multiscale Modell. Simul.*, *10*(4), 1317–1343.
Efendiev, Y., J. Galvis, R. Lazarov, M. Moon, and M. Sarkis (2013), Generalized multiscale finite element method. Symmetric interior penalty coupling, *J. Comput. Phys.*, *255*, 1–15.
Erhel, J., J.-R. de Dreuzy, A. Beaudoin, E. Bresciani, and D. Tromeur-Dervout (2009), A parallel scientific software for heterogeneous hydrogeology, *Parallel Comput. Fluid Dyn.*, *67*, 39–48.
Espig, M., W. Hackbusch, A. Litvinenko, H. G. Matthies, and E. Zander (2013), Efficient analysis of high dimensional data in tensor formats, in *Sparse Grids and Applications*, edited by J. Garcke and M. Griebel, vol. 88, pp. 31–56, Springer, Heidelberg.
Espinet, A. J., and C. A. Shoemaker (2013), Comparison of optimization algorithms for parameter estimation of multi-phase flow models with application to geological carbon sequestration, *Adv. Water Resour.*, *54*, 133–148.
Estabragh, A., M. Pereshkafti, and A. Javadi (2013), Comparison between analytical and numerical methods in evaluating the pollution transport in porous media, *Geotech. Geol. Eng.*, *31*(1), 93–101.
Fallah-Mehdipour, E., O. B. Haddad, and M. Mariño (2013), Prediction and simulation of monthly groundwater levels by genetic programming, *J. Hydro-environ. Res.*, *7*(4), 253–260.
Farthing, M. W., K. Fowler, X. Fu, A. Davis, and C. Miller (2012), Effects of model resolution on optimal design of subsurface flow and transport problems, *Adv. Water Resour.*, *38*, 27–37.
Fienen, M. N., and R. J. Hunt (2015), High-throughput computing versus high-performance computing for groundwater applications, *Groundwater*, *53*(2), 180–184.
Fienen, M. N., J. P. Masterson, N. G. Plant, B. T. Gutierrez, and E. R. Thieler (2013), Bridging groundwater models and decision support with a Bayesian network, *Water Resour. Res.*, *49*, 6459–6473, doi:10.1002/wrcr.20496.
Forrester, A. I., and A. J. Keane (2009), Recent advances in surrogate-based optimization, *Prog. Aerospace Sci.*, *45*, 50–79, doi:10.1016/j.paerosci.2008.11.001.
Frangos, M., Y. Marzouk, K. Willcox, and B. van Bloemen Waanders (2010), Surrogate and reduced-order modeling: A comparison of approaches for large-scale statistical inverse problems, in *Large-Scale Inverse Problems and Quantification of Uncertainty*, edited by L. Biegler et al., pp. 123–149, John Wiley, Chichester, U. K.
Franssen, H. H., and W. Kinzelbach (2009), Ensemble Kalman filtering versus sequential self-calibration for inverse modelling of dynamic groundwater flow systems, *J. Hydrol.*, *365*(34), 261–274, doi:10.1016/j.jhydrol.2008.11.033.
Franssen, H. H., A. Alcolea, M. Riva, M. Bakr, N. Van der Wiel, F. Stauffer, and A. Guadagnini (2009), A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments, *Adv. Water Resour.*, *32*(6), 851–872.
Galbally, D., K. Fidkowski, K. Willcox, and O. Ghattas (2010), Non-linear model reduction for uncertainty quantification in large-scale inverse problems, *Int. J. Numer. Methods Eng.*, *81*(12), 1581–1608.
Ganapathysubramanian, B., and N. Zabaras (2007), Modeling diffusion in random heterogeneous media: Data-driven models, stochastic collocation and the variational multiscale method, *J. Comput. Phys.*, *226*(1), 326–353.
Garcet, J. P., A. Ordonez, J. Roosen, and M. Vanclooster (2006), Metamodelling: Theory, concepts and application to nitrate leaching modelling, *Ecol. Modell.*, *193*(3), 629–644.
Ghanem, R., and S. Dham (1998), Stochastic finite element analysis for multiphase flow in heterogeneous porous media, *Transp. Porous Media*, *32*(3), 239–262.
Ghommem, M., M. Presho, V. M. Calo, and Y. Efendiev (2013), Mode decomposition methods for flows in high-contrast porous media. Global-local approach, *J. Comput. Phys.*, *253*, 226–238, doi:10.1016/j.jcp.2013.06.033.
Guadagnini, A., and S. P. Neuman (1999), Nonlocal and localized analyses of conditional mean steady state flow in bounded, randomly nonuniform domains: 1. Theory and computational approach, *Water Resour. Res.*, *35*(10), 2999–3018, doi:10.1029/1999WR900160.
Gugercin, S., and A. Antoulas (2000), A comparative study of 7 algorithms for model reduction, in *Proceedings of the 39th IEEE Conference on Decision and Control*, vol. 3, edited by A. Antoulas and A. Varga, pp. 2367–2372, IEEE, Sydney.

Gugercin, S., and K. Willcox (2008), Krylov projection framework for Fourier model reduction, *Automatica*, *44*(1), 209–215, doi:10.1016/j.automatica.2007.05.007.
Hajibeygi, H., G. Bonfigli, M. A. Hesse, and P. Jenny (2008), Iterative multiscale finite-volume method, *J. Comput. Phys.*, *227*(19), 8604–8621.
Harbaugh, A. W. (2005), MODFLOW-2005, the US geological survey modular ground-water model: The ground-water flow process, *U.S. Geol. Surv. Tech. Methods, 6-A16*, 1–99.
Hastie, T., R. Tibshirani, J. Friedman, T. Hastie, J. Friedman, and R. Tibshirani (2009), *The Elements of Statistical Learning*, vol. 2, Springer, N. Y.
Hay, A., I. Akhtar, and J. T. Borggaard (2012), On the use of sensitivity analysis in model reduction to predict flows for varying inflow conditions, *Int. J. Numer. Methods Fluids*, *68*(1), 122–134, doi:10.1002/fld.2512.
He, X., L. Jiang, and J. David Moulton (2013), A stochastic dimension reduction multiscale finite element method for groundwater flow problems in heterogeneous random porous media, *J. Hydrol.*, *478*, 77–88.
Hemker, T., K. R. Fowler, M. W. Farthing, and O. von Stryk (2008), A mixed-integer simulation-based optimization approach with surrogate functions in water resources management, *Opt. Eng.*, *9*(4), 341–360.
Herzog, M., A. Gilg, M. Paffrath, P. Rentrop, and U. Wever (2008), Intrusive versus non-intrusive methods for stochastic finite elements, in *From Nano to Space*, edited by M. H. Breitner, G. Denk, and P. Rentrop, pp. 161–174, Springer, Berlin.
Hill, M. C. (2006), The practical use of simplicity in developing ground water models, *Groundwater*, *44*(6), 775–781.
Hou, T. Y., and X.-H. Wu (1997), A multiscale finite element method for elliptic problems in composite materials and porous media, *J. Comput. Phys.*, *134*(1), 169–189.
Hughes, T. J., G. R. Feijóo, L. Mazzei, and J.-B. Quincy (1998), The variational multiscale method: A paradigm for computational mechanics, *Comput. Methods Appl. Mech. Eng.*, *166*(1), 3–24.
Hussain, M. F., R. R. Barton, and S. B. Joshi (2002), Metamodeling: Radial basis functions versus polynomials, *Eur. J. Oper. Res.*, *138*(1), 142–154.
Jakeman, A. J., R. A. Letcher, and J. P. Norton (2006), Ten iterative steps in development and evaluation of environmental models, *Environ. Modell. Software*, *21*(5), 602–614.
Jakeman, J. D., and S. G. Roberts (2013), Local and dimension adaptive stochastic collocation for uncertainty quantification, in *Sparse Grids and Applications*, edited by J. Garcke and M. Griebel, pp. 181–203, Springer, Heidelberg.
Jenny, P., S. Lee, and H. Tchelepi (2003), Multi-scale finite-volume method for elliptic problems in subsurface flow simulation, *J. Comput. Phys.*, *187*(1), 47–67.
Jones, D. R., M. Schonlau, and W. J. Welch (1998), Efficient global optimization of expensive black-box functions, *J. Global Opt.*, *13*(4), 455–492.
Kavetski, D., and G. Kuczera (2007), Model smoothing strategies to remove microscale discontinuities and spurious secondary optima in objective functions in hydrological calibration, *Water Resour. Res.*, *43*, W03411, doi:10.1029/2006WR005195.
Keating, E. H., J. Doherty, J. A. Vrugt, and Q. Kang (2010), Optimization and uncertainty assessment of strongly nonlinear groundwater models with high parameter dimensionality, *Water Resour. Res.*, *46*, W10517, doi:10.1029/2009WR008584.
Kennedy, M. C., and A. O'Hagan (2001), Bayesian calibration of computer models, *J. R. Stat. Soc., Ser. B*, *63*(3), 425–464.
Kitanidis, P. K., and J. Lee (2014), Principal component geostatistical approach for large-dimensional inverse problems, *Water Resour. Res.*, *50*, 5428–5443, doi:10.1002/2013WR014630.
Knezevic, D. J., and J. W. Peterson (2011), A high-performance parallel implementation of the certified reduced basis method, *Comput. Methods Appl. Mech. Eng.*, *200*(13), 1455–1466.
Kourakos, G., and A. Mantoglou (2009), Pumping optimization of coastal aquifers based on evolutionary algorithms and surrogate modular neural network models, *Adv. Water Resour.*, *32*(4), 507–521.
Laloy, E., B. Rogiers, J. A. Vrugt, D. Mallants, and D. Jacques (2013), Efficient posterior exploration of a high-dimensional groundwater model from two-stage Markov chain Monte Carlo simulation and polynomial chaos expansion, *Water Resour. Res.*, *49*, 2664–2682, doi:10.1002/wrcr.20226.
LaVenue, A. M., and J. F. Pickens (1992), Application of a coupled adjoint sensitivity and kriging approach to calibrate a groundwater flow model, *Water Resour. Res.*, *28*(6), 1543–1569.
Le Maitre, O., O. Knio, H. Najm, and R. Ghanem (2004), Uncertainty propagation using Wiener–Haar expansions, *J. Comput. Phys.*, *197*(1), 28–57.
Leube, P. C., W. Nowak, and G. Schneider (2012), Temporal moments revisited: Why there is no better way for physically based model reduction in time, *Water Resour. Res.*, *48*, W11527, doi:10.1029/2012WR011973.
Lieberman, C., K. Willcox, and O. Ghattas (2010), Parameter and state model reduction for large-scale statistical inverse problems, *SIAM J. Sci. Comput.*, *32*(5), 2523–2542.
Lunati, I., and P. Jenny (2008), Multiscale finite-volume method for density-driven flow in porous media, *Comput. Geosci.*, *12*(3), 337–350.
Mansour, M., and A. Spink (2013), Grid refinement in Cartesian coordinates for groundwater flow models using the divergence theorem and Taylor's series, *Groundwater*, *51*(1), 66–75.
Mariethoz, G., P. Renard, and J. Caers (2010), Bayesian inverse problems and optimization with iterative spatial resampling, *Water Resour. Res.*, *46*, W11530, doi:10.1029/2010WR009274.
Marzouk, Y. M., and H. N. Najm (2009), Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems, *J. Comput. Phys.*, *228*(6), 1862–1902.
Matott, L. S., and A. J. Rabideau (2008), Calibration of complex subsurface reaction models using a surrogate-model approach, *Adv. Water Resour.*, *31*(12), 1697–1707.
Matott, L. S., A. J. Rabideau, and J. R. Craig (2006), Pump-and-treat optimization using analytic element method flow models, *Adv. Water Resour.*, *29*(5), 760–775.
Matott, L. S., J. E. Babendreier, and S. T. Purucker (2009), Evaluating uncertainty in integrated environmental models: A review of concepts and tools, *Water Resour. Res.*, *45*, W06421, doi:10.1029/2008WR007301.
Matthies, H. G., and A. Keese (2005), Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations, *Comput. Methods Appl. Mech. Eng.*, *194*(12), 1295–1331.
McPhee, J., and W. W.-G. Yeh (2008), Groundwater management using model reduction via empirical orthogonal functions, *J. Water Resour. Plann. Manage.*, *134*(2), 161–170.
Mehl, S., and M. C. Hill (2010), Grid-size dependence of Cauchy boundary conditions used to simulate stream–aquifer interactions, *Adv. Water Resour.*, *33*(4), 430–442.
Mehl, S. W., and M. C. Hill (2005), MODFLOW-2005, the US geological survey modular ground-water model: Documentation of shared node local grid refinement (LGR) and the boundary flow and head (BFH) package, *U.S. Geol. Surv. Tech. Methods, 6-A21*, 13 pp.

Miller, C. T., C. N. Dawson, M. W. Farthing, T. Y. Hou, J. Huang, C. E. Kees, C. Kelley, and H. P. Langtangen (2013), Numerical simulation of water resources problems: Models, methods, and trends, *Adv. Water Resour.*, *51*, 405–437.
Mills, R. T., G. E. Hammond, P. C. Lichtner, V. Sripathi, G. K. Mahinthakumar, and B. F. Smith (2009), Modeling subsurface reactive flows using leadership-class computing, *J. Phys. Conf. Ser.*, *180*, 12–62.
Mugunthan, P., C. A. Shoemaker, and R. G. Regis (2005), Comparison of function approximation, heuristic, and derivative-based methods for automatic calibration of computationally expensive groundwater bioremediation models, *Water Resour. Res.*, *41*, W11427, doi: 10.1029/2005WR004134.
Najm, H. N. (2009), Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics, *Annu. Rev. Fluid Mech.*, *41*, 35–52.
Nan, T., and J. Wu (2011), Groundwater parameter estimation using the ensemble Kalman filter with localization, *Hydrogeol. J.*, *19*(3), 547–561.
Narayan, A., C. Gittelson, and D. Xiu (2014), A stochastic collocation algorithm with multifidelity models, *SIAM J. Sci. Comput.*, *36*(2), A495–A521.
Nouy, A. (2007), A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations, *Comput. Methods Appl. Mech. Eng.*, *196*(4548), 4521–4537, doi:10.1016/j.cma.2007.05.016.
O'Hagan, A. (2006), Bayesian analysis of computer code outputs: A tutorial, *Reliab. Eng. Syst. Safety*, *91*(10), 1290–1300.
O'Hagan, A. (2013), Polynomial chaos: A tutorial and critique from a statistician's perspective, *SIAM/ASA J. Uncertainty Quantification*, *20*, 1–20.
Oliver, D. S., and Y. Chen (2011), Recent progress on reservoir history matching: A review, *Comput. Geosci.*, *15*(1), 185–221.
Panday, S., P. Huyakorn, R. Therrien, and R. Nichols (2013), MODFLOW–USG version 1: An unstructured grid version of MODFLOW for simulating groundwater flow, *Adv. Water Resour.*, *25*(5), 497–511.
Pasetto, D., A. Guadagnini, and M. Putti (2011), POD-based Monte Carlo approach for the solution of regional scale groundwater flow driven by randomly distributed recharge, *Adv. Water Resour.*, *34*(11), 1450–1463.
Pasetto, D., M. Putti, and W. W.-G. Yeh (2013), A reduced-order model for groundwater flow equation with random hydraulic conductivity: Application to Monte Carlo methods, *Water Resour. Res.*, *49*, 3215–3228, doi:10.1002/wrcr.20136.
Pasetto, D., A. Guadagnini, and M. Putti (2014), A reduced-order model for Monte Carlo simulations of stochastic groundwater flow, *Comput. Geosci.*, *18*(2), 157–169.
Razavi, S., B. A. Tolson, and D. H. Burn (2012a), Review of surrogate modeling in water resources, *Water Resour. Res.*, *48*, W07401, doi: 10.1029/2011WR011527.
Razavi, S., B. A. Tolson, and D. H. Burn (2012b), Numerical assessment of metamodelling strategies in computationally intensive optimization, *Environ. Modell. Software*, *34*, 67–86.
Regis, R. G., and C. A. Shoemaker (2005), Constrained global optimization of expensive black box functions using radial basis functions, *J. Global Opt.*, *31*(1), 153–171.
Regis, R. G., and C. A. Shoemaker (2007), A stochastic radial basis function method for the global optimization of expensive functions, *INFORMS J. Comput.*, *19*(4), 497–509.
Regis, R. G., and C. A. Shoemaker (2012), A quasi-multistart framework for global optimization of expensive functions using response surface models, *J. Global Opt.*, *56*, 1719–1753.
Rewieński, M., and J. White (2006), Model order reduction for nonlinear dynamical systems based on trajectory piecewise-linear approximations, *Linear Algebra Appl.*, *415*(2), 426–454.
Roach, J., and V. Tidwell (2009), A compartmental-spatial system dynamics approach to ground water modeling, *Groundwater*, *47*(5), 686–698, doi:10.1111/j.1745-6584.2009.00580.x.
Robinson, T., M. Eldred, K. Willcox, and R. Haimes (2008), Surrogate-based optimization using multifidelity models with variable parameterization and corrected space mapping, *AIAA J.*, *46*(11), 2814–2822.
Sacks, J., et al. (1989), Design and analysis of computer experiments, *Stat. Sci.*, *4*(4), 409–423.
Sahuquillo, A. (1983), An eigenvalue numerical technique for solving unsteady linear groundwater models continuously in time, *Water Resour. Res.*, *19*(1), 87–93.
Saied, F., and G. Mahinthakumar (1998), Efficient parallel multigrid based solvers for large scale groundwater flow simulations, *Comput. Math. Appl.*, *35*(7), 45–54.
Sangalli, G. (2003), Capturing small scales in elliptic problems using a residual-free bubbles finite element method, *Multiscale Modell. Simul.*, *1*(3), 485–503.
Shafii, M., B. Tolson, and L. Matott (2014), Uncertainty-based multi-criteria calibration of rainfall-runoff models: A comparative study, *Stoch. Environ. Res. Risk Assess.*, *28*(6), 1493–1510, doi:10.1007/s00477-014-0855-x.
Shi, L., L. Zeng, D. Zhang, and J. Yang (2012), Multiscale-finite-element-based ensemble Kalman filter for large-scale groundwater flow, *J. Hydrol.*, *468–469*, 22–34.
Siade, A. J., M. Putti, and W. W.-G. Yeh (2010), Snapshot selection for groundwater model reduction using proper orthogonal decomposition, *Water Resour. Res.*, *46*, W08539, doi:10.1029/2009WR008792.
Siade, A. J., M. Putti, and W. W.-G. Yeh (2012), Reduced order parameter estimation using quasilinearization and quadratic programming, *Water Resour. Res.*, *48*, W06502, doi:10.1029/2011WR011471.
Stone, N. (2011), Gaussian process emulators for uncertainty analysis in groundwater flow, PhD thesis, Univ. of Nottingham, Nottingham.
Sudret, B. (2008), Global sensitivity analysis using polynomial chaos expansions, *Reliab. Eng. Syst. Safety*, *93*(7), 964–979.
Sun, A. Y. (2008), State-of-the-art multiscale approaches for flow and transport modeling: A literature review, *Tech. Rep. 2008-1*, U.S. Nucl. Regul. Commiss., Cent. for Nucl. Waste Regul. Anal., San Antonio, Tex.
Thum, P., W. Hesch, and K. Stüben (2011), LMG2: Accelerating the SAMG multigrid-solver in MODFLOW, paper presented at MODFLOW and MORE Conference, Colo., Colorado School of Mines, Golden, Colo.
Tonkin, M. J., and J. Doherty (2005), A hybrid regularized inversion methodology for highly parameterized environmental models, *Water Resour. Res.*, *41*, W10412, doi:10.1029/2005WR003995.
Vermeulen, P., C. Te Stroet, and A. Heemink (2006), Limitations to upscaling of groundwater flow models dominated by surface water interaction, *Water Resour. Res.*, *42*, W10406, doi:10.1029/2005WR004620.
Viana, F. A., R. T. Haftka, and V. Steffen Jr. (2009), Multiple surrogates: How cross-validation errors can help us to obtain the best predictor, *Struct. Multidisciplinary Optim.*, *39*(4), 439–457.
Vilhelmsen, T. N., S. Christensen, and S. W. Mehl (2012), Evaluation of MODFLOW-LGR in connection with a synthetic regional-scale model, *Groundwater*, *50*(1), 118–132.

Villa-Vialaneix, N., M. Follador, M. Ratto, and A. Leip (2012), A comparison of eight metamodeling techniques for the simulation of N_2O fluxes and N leaching from corn crops, *Environ. Modell. Software*, *34*, 51–66, doi:10.1016/j.envsoft.2011.05.003.
Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003), A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, *39*(8), 1201, doi:10.1029/2002WR001642.
Vuik, C., A. Segal, and J. Meijerink (1999), An efficient preconditioned cg method for the solution of a class of layered problems with extreme contrasts in the coefficients, *J. Comput. Phys.*, *152*(1), 385–403.
Wan, X., and G. E. Karniadakis (2005), An adaptive multi-element generalized polynomial chaos method for stochastic differential equations, *J. Comput. Phys.*, *209*(2), 617–642.
Watson, T. A., J. E. Doherty, and S. Christensen (2013), Parameter and predictive outcomes of model simplification, *Water Resour. Res.*, *49*, 3952–3977, doi:10.1002/wrcr.20145.
Weinan, E., and B. Engquist (2003), Multiscale modeling and computation, *Notices Am. Math. Soc.*, *50*(9), 1062–1070.
Weinan, E., et al. (2002), The heterogeneous multi-scale method, *Commun. Math. Sci.*, *1*, 87–132.
Weinan, E., B. Engquist, X. Li, W. Ren, and E. Vanden-Eijnden (2007), Heterogeneous multiscale methods: A review, *Commun. Comput. Phys.*, *2*(3), 367–450.
Wen, X.-H., and J. J. Gómez-Hernández (1996), Upscaling hydraulic conductivities in heterogeneous media: An overview, *J. Hydrol.*, *183*(1), ix–xxxii.
Willcox, K., and A. Megretski (2005), Fourier series for accurate, stable, reduced-order models in large-scale linear applications, *SIAM J. Sci. Comput.*, *26*(3), 944–962.
Willcox, K., and J. Peraire (2002), Balanced model reduction via the proper orthogonal decomposition, *AIAA J.*, *40*(11), 2323–2330.
Winter, C., D. Tartakovsky, and A. Guadagnini (2003), Moment differential equations for flow in highly heterogeneous porous media, *Surv. Geophys.*, *24*(1), 81–106, doi:10.1023/A:1022277418570.
Woodbury, A. D., W. S. Dunbar, and B. Nour-Omid (1990), Application of the Arnoldi algorithm to the solution of the advection-dispersion equation, *Water Resour. Res.*, *26*(10), 2579–2590, doi:10.1029/WR026i010p02579.
Xiu, D., and J. S. Hesthaven (2005), High-order collocation methods for differential equations with random inputs, *SIAM J. Sci. Comput.*, *27*(3), 1118–1139.
Xiu, D., and G. E. Karniadakis (2002), The Wiener–Askey polynomial chaos for stochastic differential equations, *SIAM J. Sci. Comput.*, *24*(2), 619–644.
Xu, T., A. J. Valocchi, J. Choi, and E. Amir (2012), Improving groundwater flow model prediction using complementary data-driven models, paper presented at 19th International Conference on Water Resources, Univ. of Ill., Urbana-Champaign.
Yan, S., and B. Minsker (2006), Optimal groundwater remediation design using an adaptive neural network genetic algorithm, *Water Resour. Res.*, *42*, W05407, doi:10.1029/2005WR004303.
Yoon, H., S.-C. Jun, Y. Hyun, G.-O. Bae, and K.-K. Lee (2011), A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer, *J. Hydrol.*, *396*(1), 128–138.
Young, P. C., and M. Ratto (2011), Statistical emulation of large linear dynamic models, *Technometrics*, *53*(1), 29–43, doi:10.1198/TECH.2010.07151.
Zhou, H., J. J. Gómez-Hernández, and L. Li (2014), Inverse methods in hydrogeology: Evolution and recent trends, *Adv. Water Resour.*, *63*, 22–37.