Computers & Geosciences 108 (2017) 8–20

Contents lists available at ScienceDirect

Computers & Geosciences

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

Research paper

iTOUGH2: A multiphysics simulation-optimization framework for analyzing subsurface systems

S. Finsterle [a,*,1], M. Commer [a], J.K. Edmiston [a,2], Y. Jung [a], M.B. Kowalsky [a], G.S.H. Pau [a], H.M. Wainwright [a], Y. Zhang [a]

[a] Lawrence Berkeley National Laboratory, Earth and Environmental Sciences, One Cyclotron Road, MS 74-316 C, Berkeley, CA, USA

ARTICLE INFO

Article history:
Received 15 June 2016
Received in revised form
25 August 2016
Accepted 12 September 2016
Available online 13 September 2016

Keywords:
Numerical simulation
Inverse modeling
Parameter estimation
Uncertainty quantification
Sensitivity analysis
Data-worth analysis

ABSTRACT

iTOUGH2 is a simulation-optimization framework for the TOUGH suite of nonisothermal multiphase flow models and related simulators of geophysical, geochemical, and geomechanical processes. After appropriate parameterization of subsurface structures and their properties, iTOUGH2 runs simulations for multiple parameter sets and analyzes the resulting output for parameter estimation through automatic model calibration, local and global sensitivity analyses, data-worth analyses, and uncertainty propagation analyses. Development of iTOUGH2 is driven by scientific challenges and user needs, with new capabilities continually added to both the forward simulator and the optimization framework. This review article provides a summary description of methods and features implemented in iTOUGH2, and discusses the usefulness and limitations of an integrated simulation-optimization workflow in support of the characterization and analysis of complex multiphysics subsurface systems.

1. Introduction

Advanced subsurface modeling software is capable of simulating how valuable resources—such as water, oil, gas, and heat—can be extracted from the pore space of soils and rocks, or how harmful materials—such as chemical and radioactive waste, or carbon dioxide—can be safely stored in the deep Earth or tracked as they migrate through the subsurface, or how fluids and energy could be temporarily stored underground for later retrieval.

Considerable effort is invested in the development and maintenance of such simulators. In particular, the mathematical representation of hydrological, thermal, geochemical, and geomechanical processes and their coupling is continuously refined, so is the numerical robustness and efficiency with which the resulting set of coupled governing equations is solved. Moreover, the simulators are enhanced to add interactions between the subsurface and other natural systems (e.g., surface water bodies, biosphere and atmosphere) as well as engineered components (e.g., wellbores, tunnels, containment systems, and sensors).

Powerful simulators are essential tools to help researchers and practitioners understand Earth systems. Moreover, numerical models are used for predictive purposes at specific sites. Direct human interventions in the subsurface (specifically injection, production, cleanup, and other reservoir management operations) or indirect interferences (e.g., through changes in climatic conditions or land use) can be predicted and analyzed, serving as the basis for optimizing or mitigating their impacts.

Assessing the behavior of actual systems requires that all relevant site-specific features be implemented and their properties characterized with a high enough accuracy so that the prediction uncertainty remains acceptable. This aspect of site-specific predictive simulation often leads to the most challenging tasks in the modeling process, such as the gathering, analysis, and pre-processing of diverse characterization data, conceptual and geologic model development, discretization of intricate geometric features, representation of complex forcing terms, model calibration, and predictive simulations with associated uncertainty quantification.

In this paper, we consider multiphysics problems in complex subsurface environments. In particular, we are interested in the flow of fluid mixtures, where each phase (i.e., gaseous, aqueous, and non-aqueous phase liquids) consists of multiple components (e.g., water, non-condensible gases, and volatile organic compounds), and where the coupling to strong thermal effects are of importance. Moreover, we may want to calculate geochemical reactions, geomechanical stresses and strains, or geophysical attributes as they are affected by temperature or changes in the distribution or composition of the fluids. Multiple simulators exist that are capable of modeling such coupled processes; a review of

* Corresponding author.
E-mail address: SAFinsterle@lbl.gov (S. Finsterle).
[1] Now at: Finsterle GeoConsulting, 315 Vassar Ave., Kensington, CA 94708, USA.
[2] Now at: Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94550, USA.

http://dx.doi.org/10.1016/j.cageo.2016.09.005
0098-3004/

these forward simulators is beyond the scope of this paper, but can be found in articles that describe code-comparison studies (Baca and Seth, 1996; O'Sullivan et al., 2001; Oldenburg et al., 2003; Pruess et al., 2004; MDH Engineered Solutions Corporation, 2005; Hudson et al., 2009; Birkle, 2011; Mukhopadhyay et al., 2015; Steefel et al., 2015; Fowler et al., 2016).

The iTOUGH2 simulation-optimization framework is built around software modules collectively known as the TOUGH suite of non-isothermal multiphase flow and transport simulators (Pruess et al., 2012; http://esd.lbl.gov/TOUGH). Developed at the Lawrence Berkeley National Laboratory in the early 1980s primarily for geothermal reservoir engineering, the suite of TOUGH simulators is now widely used at universities, government organizations, and private industry for applications related to geological carbon sequestration, nuclear waste disposal, energy production from geothermal, oil and gas reservoirs as well as methane hydrate deposits, environmental remediation, vadose zone hydrology, and other uses that involve coupled thermal, hydrological, geochemical, and geomechanical processes in permeable media. An overview and history of the TOUGH codes can be obtained from a series of review articles (Pruess, 2004; Finsterle et al., 2008, 2014).

iTOUGH2 (Finsterle, 2004; http://esd.lbl.gov/iTOUGH2) provides inverse modeling capabilities and (local and global) sensitivity and uncertainty propagation analyses for most TOUGH modules or any stand-alone simulation program (Finsterle and Zhang, 2011a). Such model-independent inverse modeling capabilities are also provided by other software packages, such as PEST (Doherty, 2010), UCODE (Poeter et al., 2014), and DAKOTA (Adams et al., 2016). Many specialized inversion codes exist for geophysical and other application areas. Finally, mathematical optimization algorithms are implemented in proprietary or open-source libraries and toolboxes, or are an integral part of computer languages. It is far beyond the scope of this paper to review and compare these approaches and related software packages. Instead, as part of this special issue on TOUGH, we present the capabilities of iTOUGH2 and how they can be used to develop a site-specific TOUGH model, enhance the understanding of input-output relationships, and quantify estimation and prediction uncertainties.

After presenting some general features of iTOUGH2 and providing an overview of iTOUGH2 case studies, we will discuss its various application modes, which include parameter estimation through automatic model calibration, sensitivity and data-worth analyses, and uncertainty quantification. We finally demonstrate some of iTOUGH2's features on a hydrogeomechanical problem.

2. iTOUGH2 overview

2.1. Description of simulation-optimization framework

iTOUGH2 has been referred to as a simulation-optimization framework, as it supports a workflow that aims at improving system understanding and decision making based on data and quantitative predictions using a numerical process simulator. Elements of this workflow include iterative improvement of a conceptual and numerical model based on prior knowledge and measured data, whereby data collection is driven by the need to determine influential model parameters with sufficient accuracy so that the subsequent predictions have an uncertainty level that is acceptable for the intended purpose of the model. This workflow and its support by the various iTOUGH2 toolsets are described in more detail in Finsterle et al. (2012) and Finsterle (2015a).

On its most basic level, iTOUGH2 generates various subsets of select input parameters, invokes the appropriate TOUGH module, and then extracts a specific subset of the simulator's output variables. The input-output relationship established by this simple scheme enables algorithms to (1) propose new parameter sets that likely lead to an improved fit of the model to measured data (parameter estimation), (2) examine the relative influence each parameter has on model predictions or related, quantifiable objectives (sensitivity analysis), (3) evaluate the information content of observable output variables (data-worth analysis), (4) assess the uncertainty and correlation structure of estimated parameters (error analysis), (5) examine the variability and uncertainty of model predictions (uncertainty propagation analysis), and (6) generate snapshots for the development of efficient surrogate models (reduced-order modeling). A vast array of methods and algorithms exist for each of these application modes, each with its distinct advantages and limitations.

We denote the vector of n adjustable input parameters as $\mathbf{p}$. The elements of $\mathbf{p}$ are potentially transformed parameters that can be mapped to one or multiple TOUGH input variables. They typically represent material properties (such as permeability, thermal conductivity, adsorption coefficient, or Young's modulus), but may also include initial and boundary conditions, forcing terms, and control parameters. Moreover, geostatistical properties as well as trends, shifts, and autocorrelation coefficients of the data can be considered input parameters to be varied, analyzed or estimated by iTOUGH2. In general, any aspect of the system that can be parameterized and included in TOUGH or its pre- and post-processors can be subjected to an iTOUGH2 analysis; an example is described in Wellmann et al. (2014).

The subset of discrete model output variables is compiled in vector $\mathbf{z}$ of length m. Elements of $\mathbf{z}$ are typically observable variables (such as pressures, temperatures, concentrations, flow rates, deformations) at select points in space and time. Prior information about the parameters in $\mathbf{p}$ can be considered the result of an observation; their initial values may thus be included in vector $\mathbf{z}$. Depending on the application, $\mathbf{z}$ may also contain target values that need to be minimized or maximized; examples include the total amount of contaminants in a system, costs, acceptable level of surface deformation, and energy production. In general, any TOUGH output variable or function thereof can be included in a performance measure to be analyzed by iTOUGH2. The elements of $\mathbf{z}$ are expected to depend on the parameter vector $\mathbf{p}$. In inverse modeling, each model output $z_i(\mathbf{p})$ corresponds to a measured data point, denoted by $z_i^*(i=1, \ldots, m)$.

In most applications, a priori uncertainty or variability is associated with each element of $\mathbf{p}$ and $\mathbf{z}$. These measures are used to weigh fitting errors, scale composite sensitivity measures, describe sampling distributions, or reflect acceptable estimation or prediction uncertainties. They are expressed as a variance and summarized in covariance matrix $\mathbf{C}_{zz}$ of dimensions $m \times m$.

Having the TOUGH simulator fully integrated into the iTOUGH2 framework, as indicated above, has considerable advantages in terms of user convenience, but also accuracy and efficiency. For example, iTOUGH2 has information about the spatial and temporal structure of the model. Moreover, it can control and interact with the simulator, should an individual forward run encounter convergence problems. To increase flexibility, iTOUGH2 supports the PEST protocol (Doherty, 2010; Finsterle and Zhang, 2011a), which links iTOUGH2's toolsets to external models. These external models can be pre- or postprocessors, or programs that are unrelated to the TOUGH suite of codes. The user may also conduct analyses in which a fully integrated TOUGH model is combined with an external model. PEST-style template and instruction files are used, respectively, to pass input parameters updated by the iTOUGH2 optimization routines to the model, and to retrieve the model-calculated values that correspond to observable variables.

The computational burden of an iTOUGH2 analysis lies almost exclusively in the repeated computation of the TOUGH output for

Table 1
Overview of iTOUGH2 applications and case studies.

Area[a]	Description	Features used	References
NW	Estimation of two-phase flow parameters from mesoscale ventilation experiment	Scaled sensitivity coefficients; overall correlation measures; correction of estimation covariance matrix	Finsterle and Pruess (1995)
GT	Design and analysis of pressure-pulse decay experiments	Joint inversion of multiple experiments; analysis of correlation structure; robust estimators	Finsterle and Persoff (1997); Finsterle and Najita (1998); Hannon and Finsterle (in review)
VZ	Design and analysis of radial multistep outflow experiment	Model identification criteria	Finsterle and Faybishenko (1999)
NW VZ	Study of seepage into underground openings from fractured formation under unsaturated conditions	Inversions; sensitivity analysis; stochastic continuum model; Monte Carlo simulations	Wang et al. (1999); Finsterle (2000); Finsterle et al. (2003); Ghezzehei et al. (2004)
NW VZ	Analysis of barometric pressure fluctuations	Inversions followed by blind predictions for confirmation; use of prediction uncertainty for mitigation design; pilot point method	Ahlers et al. (1999); Unger et al. (2004); Jung et al. (2011)
ER	Optimization of remediation design	Gelation module; minimization of cost function with penalty term; global minimization algorithm	Moridis et al. (1999); Finsterle (2005)
NW	Analysis of heating experiments	Estimation of thermal and two-phase hydrologic parameters	Engelhardt et al. (2003); Engelhardt and Finsterle (2003); Kiryukhin et al. (2014)
NW	Analysis of above-boiling heating experiment on bentonite sample	Joint estimation of thermal and two-phase hydrologic parameters	Unger et al. (2004)
VZ	Analysis of data from bench-scale soil stabilization experiment	Gelation module; estimation of soil and fluid properties using concentration data	Gallagher and Finsterle (2004)
VZ	Hydrogeophysics	Joint estimation of hydrogeologic and petrophysical parameters using flow and geophysical data	Kowalsky et al. (2004, 2005, 2008, 2011, 2012); Linde et al. (2006); Finsterle and Kowalsky (2008); Doetsch et al. (2013); Commer et al. (2014)
NW	Analysis of temperature data	Estimation of various properties and boundary conditions from temperature data	Mukhopadhyay et al. (2009); Freifeld et al. (2008)
GT	History matching of geothermal data	Combined steady-state and transient data calibration; heat source estimation	White et al., (2004); Mannington et al. (2004); Porras et al. (2007); Kiryukhin et al. (2008); Kaya et al. (2014)
ER	Water table calibration	Drawdown calculation under unsaturated conditions; grid search	Zhang et al. (2011a)
ER	Tracer calibration	Pilot point method	Li et al. (2011)
CS	Well test design	Well test design based on joint inversion of multiple data sets	Zhang et al. (2011b); Rasmusson et al. (2014)
CS	Leakage pathway identification	Notional inversions, sensitivity and residual analyses	Lee et al. (2015)
GT	Reservoir analysis	Global sensitivity analysis	Finsterle et al. (2013); Wainwright et al. (2013)
CS	Carbon storage reservoir management	PEST	Birkholzer et al. (2012); Jung et al. (2013)
CS	Estimation of biogeochemical parameters	PEST	Hommel et al. (2015)

[a] Application area: CS: Carbon Storage; ER: Environmental Remediation; GT: Geothermal; NW: Nuclear Waste; VZ: Vadose Zone.

multiple parameter sets. In particular, the calculation of the elements of the Jacobian matrix, $J_{ij} = \partial z_i / \partial p_j$ by finite differences requires $(n+1)$ simulation runs. Fortunately, these forward simulations are independent of each other and can thus be run in parallel. iTOUGH2 also improves the efficiency of other methods (e.g., derivative-free global minimization algorithms and Monte Carlo simulations) to the extent that the algorithm can be "embarrassingly parallelized" (a technical term used to indicate that no data exchange between child processes is needed). As an alternative, the parallel versions of the TOUGH forward simulator (e.g., TOUGH3, Jung et al., 2016) can be linked to iTOUGH2 using the PEST protocol described above. Finally, Commer et al. (2014) developed MPiTOUGH2, a version of iTOUGH2 that uses a two-level parallelization approach. Independent TOUGH forward simulations are run in parallel, and at the same time, each TOUGH run is parallelized using domain decomposition.

2.2. iTOUGH2 case studies

Table 1 shows a partial list of iTOUGH2 case studies that were documented in peer-reviewed journal articles. The list demonstrates the diversity of questions addressed by iTOUGH2, which range from fundamental understanding of input-output relationships, to design of laboratory and field experiments, to the analysis of data collected on multiple scales, to the calibration of site-specific models. The list also indicates the usefulness of iTOUGH2 in a variety of application areas and different research and engineering environments.

3. Inverse modeling

3.1. General approach

As indicated by its name, the main purpose of iTOUGH2 is to solve the inverse problem based on the TOUGH forward simulator. Select parameters are estimated by minimizing some measure of the misfit between select simulation outputs and their corresponding measurement data. A single, best-estimate parameter set is obtained. This approach is referred to as a deterministic inversion (as opposed to a stochastic inversion; see, for example, Linde and Vrugt (2013)). Under certain conditions, the point estimate and its uncertainty can be related to the assumed distribution of the posterior residuals. The method is then referred to as maximum likelihood (ML) estimation, or—if a regularization term related to the prior distribution of the parameters is added—as maximum a posteriori (MAP) estimation.

The following subsections summarize the key elements of a formalized approach to parameter estimation by automatic history matching as implemented in iTOUGH2. After a short description of the joint inversion concept, we first discuss the objective function, followed by a list of local and global algorithms used to determine its minimum. Finally, we explain why a comprehensive residual

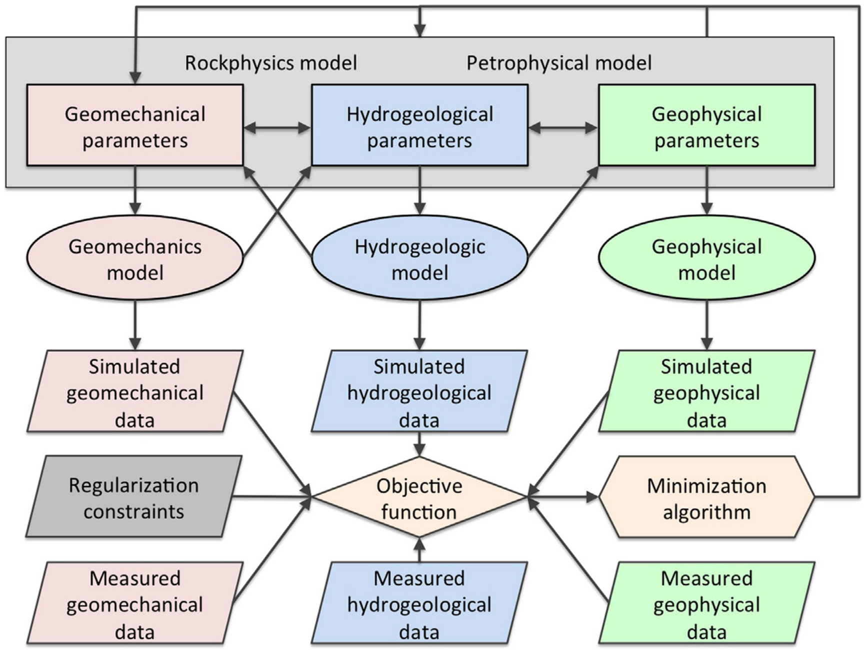

Fig. 1. Multiphysics joint inversion framework, showing two-way coupling or one-way linking of hydrogeological, geomechanical, and geophysical forward models. Parameters to be estimated are either shared among these model components, or are connected through petrophysical relationships. Observations of all types enter a single, weighted objective function, potentially augmented by regularization terms (prior information and Tikhonov regularization). The objective function is minimized by concurrently updating all influential parameters.

and error analysis is considered essential for a meaningful interpretation of the inverse modeling results.

3.2. Multiphysics joint inversion

The basic motivation for performing multiphysics joint inversions is that information contained in different data types is likely to be complementary to each other. For example, geophysical data generally contain information about the subsurface structure and spatial phase distribution, whereas hydrogeological data contain information about flow and transport properties. Combining the two data sets allows for the determination of the spatial distribution of flow properties. Geomechanical data also serve the purpose of better constraining a flow model, but may also be analyzed for the specific analysis of geomechanical properties and processes. Fig. 1 shows iTOUGH2's multiphysics joint inversion framework.

iTOUGH2's capabilities for jointly inverting hydrological and geophysical data for the estimation of hydrogeological, geophysical, and geostatistical parameters (Kowalsky et al., 2004, 2005, 2008, 2011, 2012; Finsterle and Kowalsky, 2008) have been recently extended to include time-domain electromagnetic and seismic data. These capabilities are implemented in MPiTOUGH2 (Commer et al., 2014). Time-lapse electrical DC resistivity data can also be simulated using the BERT finite-element code (Günther et al., 2006; Doetsch et al., 2013). Straight-ray, curved-ray, and full-waveform seismic simulators are integrated. Multiple petrophysical relations linking hydrological properties and states to various geophysical attributes are available. Geomechanical modeling is fully integrated in iTOUGH2, allowing the user to estimate geomechanical parameters based on displacement observations (and other sensitive data). An example of a joint inversion of hydrogeological and geomechanical data is shown in Section 8.

3.3. Objective function

The objective function is an integral measure of misfit between the observable model output and corresponding measurements. The topology of the objective function determines the solution of the inverse problem, including its uniqueness and stability. The objective function generally is a complex hypersurface in the n-dimensional parameter space, exhibiting multiple local minima, as well as ridges, saddle points, and long narrow valleys, making it difficult to identify the global minimum. Moreover, the shape, orientation, and convexity of the objective function near the minimum reflect, respectively, the non-linearity, correlation structure, and uncertainty of the estimated parameters. The topology of the objective function can be improved by proper parameterization of the problem, and by collecting data that are sensitive to and contain complementary information about the parameters to be estimated, which should be the declared purpose of numerically designing laboratory and field experiments through sensitivity analyses and notional inversions, as discussed in Finsterle (2015a).

The weighted least-squares objective function is almost universally used in hydrogeological inverse problems and many other application areas, often without explicit justification. Many minimization algorithms (see Section 3.4) and analyses of the estimation errors (see Section 3.5) rely on the objective function being approximately quadratic, i.e.,

$$S(\mathbf{p})=\mathbf{r}^T\mathbf{C}_{zz}^{-1}\mathbf{r} \tag{1}$$

where $\mathbf{r}=(\mathbf{z}^* - \mathbf{z}(\mathbf{p}))$ is the residual vector of length m, and $\mathbf{C}_{zz}$ is the measurement covariance matrix, often assumed to be diagonal.

iTOUGH2 uses the following generalized likelihood function known as the skew exponential power (SEP) distribution (Schoups and Vrugt, 2010)

$$p(a_t\xi, \beta) = \frac{2\sigma_\xi}{\xi + \xi^{-1}}\omega_\beta \exp\left\{ -c_\beta \left|a_{\xi,t}\right|^{2/(1+\beta)} \right\} \tag{2}$$

where ξ is the skewness parameter and β is the kurtosis parameter. The coefficients $a_{\xi,t}$, μ_ξ, σ_ξ, c_β, and ω_β are computed as a function of ξ and β, as detailed in Schoups and Vrugt (2010). The density p is symmetric for $\xi=1$, and skewed to the left (right) for $\xi>1$ ($\xi<1$). For $\xi=1$, a uniform distribution is obtained for $\beta=-1$, a Gaussian for $\beta=0$, and a Laplacian for $\beta=1$. While the standard L_1 (least absolute value) and L_2 (least squares) estimators are included as special cases of the SEP, it is flexible enough to also handle non-symmetric, heavy-tailed distributions, as are often observed when analyzing residuals after a calibration. In addition to Eq. (2), iTOUGH2 allows the user to select from a set of predefined objective functions. These include functions that are based on a known probability density function (pdf), such as the Gaussian, and more heavy-tailed distributions such as the double-exponential and Cauchy distributions. Empirical functions are also included, such as the Huber and Andrews estimators, which do not correspond to a standard pdf (Finsterle and Najita, 1998). In particular, the Andrews estimator discards any residual larger than a user-specified threshold value, effectively eliminating the potentially strong bias introduced by outliers in the data. Non-normal and heteroscedastic errors can also be accounted for by employing a Box-Cox transformation to the calculated and simulated observables, and to represent autocorrelated residuals. Finally, the Nash-Sutcliffe and Kling-Gupta efficiency criteria (Gupta et al., 2009) may be selected as objective functions to be minimized.

Some of the objective functions mentioned above contain statistical parameters (such as skewness ξ, kurtosis β, Box-Cox parameter λ, or autocorrelation coefficient ρ) that may not be well known. They must be determined iteratively by an analysis of the final residuals after a preliminary inversion. Alternatively, through proper scaling of the objective functions in iTOUGH2, these statistical parameters may be jointly estimated along with the adjustable parameters of the physical system, as demonstrated in Finsterle and Zhang (2011b). It is recommended to perform a detailed residual analysis after an inversion to test whether the final residuals comply with the assumptions underlying the chosen objective function.

3.4. Minimization algorithm

Within the context of maximum likelihood estimation, the best-estimate parameter set is identified by finding the minimum of the objective function in the n-dimensional parameter space. Note that the well- or ill-posedness of an inverse problem is entirely determined by the topology of the objective function—not by the (in)ability of a given search algorithm to find its global minimum. It is noted, however, that an ill-posed inverse problem makes it difficult or impossible to detect a unique solution.

While many of the large number of available minimization algorithms perform well for a specific class of inverse problems, the difficulty for a modeler lies in navigating the tradeoff between generality and efficiency. This is particularly important for

Table 2
Overview of minimization algorithms implemented in iTOUGH2.

Algorithm	Advantages, Limitations	Comments
Derivative-based local minimization algorithms		
Gauss-Newton (e.g., Aster et al., 2013) $\Delta\mathbf{p}=\mathbf{F}^{-1}\mathbf{J}^T\mathbf{C}_{zz}^{-1}\mathbf{r}$	Very efficient for weakly non-linear least-squares problems; tends to overshoot for strongly non-linear problems	Most efficient method for linear least-squares models; rarely used for non-linear iTOUGH2 inversions; Levenberg-Marquardt method approaches Gauss-Newton method near minimum
Levenberg-Marquardt (Levenberg, 1944; Marquardt, 1963) $\Delta\mathbf{p}=(\mathbf{F}+\lambda\mathbf{D})^{-1}\mathbf{J}^T\mathbf{C}_{zz}^{-1}\mathbf{r}$	Efficient, flexible local algorithm for non-linear least-squares problems	Adaptively interpolates between robust gradient descent and efficient Gauss-Newton step; Tikhonov matrix **D** can be chosen as (1) **D**=**I**, (2) **D**=diag(**F**), or (3) **D**=eigenvalue(**F**), ordered according to parameter identifiability; initial Levenberg parameter value λ can be automatically determined; parameters can be adaptively selected based on composite sensitivity measure or independence measure or truncated SVD and superparameters
Derivative-free local minimization algorithm		
Downhill Simplex Algorithm (Nelder and Mead, 1965)	Minimizes continuous and discontinuous cost functions in relatively low-dimensional parameter spaces	Generates series of $(n+1)$-dimensional simplexes by (1) moving the point of the simplex with the highest cost function through the opposite face of the simplex, (2) reflecting, (3) expanding, or (4) contracting the simplex from the previous step. Initial simplex and contraction can be evaluated in parallel
Derivative-free, sampling-based global minimization algorithms		
Multistart Metric Stochastic Response Surface Method (Regis and Shoemaker, 2007)	Minimizes smooth, non-convex objective functions; uses local response surface approximation model to improve efficiency of global multistart sampling algorithm	Uses radial basis functions to approximate objective function; selects candidate point based on surrogate model estimate and distance from previously evaluated points; performs multistart local optimizations
Differential Evolutionary Algorithm (Storn and Price, 1997)	Sampling-based global algorithm; no restrictions on topology of objective function; robust but inefficient	Meta-heuristic minimization algorithm that is based on an analogy to evolutionary processes, including reproduction, cross-over, mutation, and fitness.
Harmony Search Algorithm (Geem et al., 2001)	Sampling-based global algorithm; no restrictions on topology of objective function; robust but inefficient	Meta-heuristic minimization algorithm that is based on an analogy to musical performance processes, such as improvisation and pitch adjustment
Simulated Annealing (Kirkpatrick et al., 1983)	Sampling-based global algorithm; no restrictions on topology of objective function; robust but inefficient	Continuous version of meta-heuristic minimization algorithm that is based on an analogy to the slow cooling of metals, with temperature and annealing schedule as control parameters
Monte Carlo Simulation, Latin Hypercube Sampling	No restrictions on topology of objective function; inefficient	Evaluates objective function at random points in n-dimensional parameter space; transient forward simulation automatically aborted at time t if $S(t)>S_{min}$; embarrassingly parallel
Derivative-free, global minimization algorithm		
Grid Search	No restrictions on topology of objective function; inefficient, only applicable for small n	Evaluate objective function on regular grid or at user-specified points in n-dimensional parameter space; transient forward simulation automatically aborted at time t if $S(t)>S_{min}$; embarrassingly parallel; unsorted output available for increased performance

computationally costly forward models, which prevent an exhaustive evaluation of the high-dimensional parameter space. While many of the global minimization algorithms are successful in estimating parameters and determining complex confidence regions for analytical test functions or highly simplified process models, their usefulness (and thus relevance) for solving real-life optimization problems is often limited. There exists a second tradeoff that needs to be addressed by the objectives of the study rather than the available algorithm. It is that between accurately determining the uncertainty of an oversimplified model versus obtaining an approximate measure of estimation uncertainty for an accurate, physics-based process model. Model-reduction techniques, if combined with a sound approximation error model, may provide a valid approach to address these issues.

Table 2 summarizes the minimization algorithms implemented in iTOUGH2. They are approximately ordered from efficient to computationally demanding methods. In Table 2, $\mathbf{F} = \left(\mathbf{J}^T\mathbf{C}_{zz}^{-1}\mathbf{J}\right)$ is the Fisher information matrix, an approximation of the Hessian. All algorithms are iterative, i.e., the parameter vector is updated as $\mathbf{p}_{k+1} = \mathbf{p}_k + \Delta\mathbf{p}_k$, where k is the iteration index (omitted in Table 2). For derivative-based local minimization algorithms, the Jacobian $\mathbf{J}_k$ and residual vector $\mathbf{r}_k$ need to be recalculated at each iteration. A Broyden rank-one update of the Jacobian matrix is available in iTOUGH2, whereby the need to perform a full recalculation of the Jacobian is automatically determined. The parameters to be updated by derivative-based algorithms can be adaptively selected based on a composite sensitivity measure (see Section 4), parameter independence measure (see Section 3.5), or parameter identifiability measure (Doherty and Hunt, 2009) in combination with a truncated singular value decomposition (SVD) of the Fisher matrix $\mathbf{F}$. Optimization can also be performed along so-called superparameters (Tonkin and Doherty, 2005). The truncation limit is user-specified and can be adaptively relaxed so that the number of parameters included in the optimization generally increases as the inversion proceeds.

Minimization of the objective function is stopped if one of several convergence criteria is met. Typically, the user specifies a maximum number of forward runs or parameter vector updates as a pragmatic criterion, but other measures (such as the value or gradient of the objective function, size of parameter update, or number of unsuccessful uphill steps) can be used.

The default minimization method in iTOUGH2 is the Levenberg-Marquardt algorithm with a Tikhonov matrix $\mathbf{D}$ that consists of the diagonal elements of the Fisher matrix $\mathbf{F}$, and a Jacobian $\mathbf{J}$ that is calculated (potentially in parallel) by forward finite differences. This method has been found to be both robust and efficient for many non-linear least-squares problems involving a computationally demanding TOUGH forward simulation. In addition, iTOUGH2 provides alternative local and global minimization algorithms to address a vast array of calibration and other optimization problems involving TOUGH, an external simulator, or a combination thereof.

3.5. Residual and error analysis

Identifying the global minimum of the objective function does not mean that the estimated parameters are optimal, or even meaningful, or that the model is an acceptable representation of the real system. Analyzing the final residuals and calculating the estimation errors gives the modeler a first indication about the quality of the inversion.

The estimated error variance is an overall goodness-of-fit measure:

$$s_0^2 = \frac{\mathbf{r}^T\mathbf{C}_{zz}^{-1}\mathbf{r}}{m-n} \tag{3}$$

The Fisher model test examines whether the overall fit to the data (as expressed by s_0^2) is consistent with the modeler's expectations, which were a priori expressed through matrix $\mathbf{C}_{zz}$. However, even if the overall fit is deemed acceptable, the residuals may be unacceptable for subsets of $\mathbf{r}$. A detailed residual analysis can shed light on an inconsistency between the data and the model, which becomes apparent if the residuals (1) do not follow the assumed distributional model, (2) show a systematic trend rather than being random, (3) exhibit outliers, or (4) are significantly larger or smaller than expected. The residual analysis (combined with a good physical understanding of the system) helps identify aspects of the model that need to be refined. A number of statistical tools supporting the residual analysis are implemented in iTOUGH2; they include scatter plots, a runs statistics, covariance matrices of the model prediction and residuals, statistical outlier identification, local reliability measures, and the relative contribution to the objective function. These residual statistics are discussed in Finsterle and Zhang (2011b).

Reproducing the observed system response at the calibration points is a necessary but not sufficient condition for a successful inversion. The parameter set may still be of little or no practical use because of large estimation uncertainty. Under the assumption that the forward model is linear and the residuals are Gaussian, the covariance matrix of the estimated parameters is given by

$$\mathbf{C}_{pp} = s_0^2\left(\mathbf{J}^T\mathbf{C}_{zz}^{-1}\mathbf{J}\right)^{-1} \tag{4}$$

Eq. (4) is only an approximation of the actual uncertainty region if the model is nonlinear. Nevertheless, it provides useful information about the estimation uncertainty and parameter correlations. Estimates may be highly uncertain if (1) the model does not fit the data well, i.e., s_0^2 is large, (2) the observations are not sufficiently sensitive to the parameters of interest, i.e., the values in a column of the Jacobian matrix $\mathbf{J}$ are too small, or (3) the parameters exhibit strong statistical correlations, i.e., columns of $\mathbf{J}$ are nearly linearly dependent. Strong parameter correlation is the main culprit of large estimation uncertainty, and is usually an indication of overparameterization, a poor choice of the parameter set, or a lack of calibration data with complementary information content. Again, physical insight is needed to remedy the situation, as demonstrated in Finsterle and Persoff (1997).

A normalized measure of a parameter's overall independence can be defined as the ratio between its marginal standard deviation (the square-root of the diagonal element of $\mathbf{C}_{pp}$) and the standard deviation calculated under the assumption that all the other parameters are uncorrelated (or perfectly known). An ill-posed inverse problem can be improved by better constraining a parameter with a low rank in parameter independence and a high rank in its composite sensitivity measure. Finally, iTOUGH2 evaluates various model identification criteria, which can be used to measure the relative performance of competing alternative conceptual models with different parameterizations and different calibration data (Carrera and Neuman, 1986; Ye et al., 2008).

4. Sensitivity and data-worth analysis

Sensitivity analysis is a prime application mode of a numerical model to evaluate which parameters have a high impact on the system response. The usefulness of sensitivity analysis and its relation to uncertainty quantification is amply described in the literature (see, e.g., Saltelli et al., 2008). iTOUGH2 supports both local and global sensitivity analyses. The local sensitivity coefficients are the partial derivatives of select output variables z_i with respect to select input parameters p_j, evaluated at the reference parameter set $\mathbf{p}^*$. Because sensitivity coefficients have units of the model output over the units of the parameter, these sensitivity coefficients cannot be readily compared to each other if inputs and outputs of different types are involved. We therefore introduce a scaled, dimensionless local sensitivity coefficient

$$\bar{S}_{ij} = \left.\frac{\partial z_i}{\partial p_j}\right|_{\mathbf{p}*} \cdot \frac{\sigma_{p_j}}{\sigma_{z_i}} \tag{5}$$

which can be used to calculate composite sensitivity measures (e.g., the sum of the absolute values of the scaled sensitivity coefficients of each row and column of $\bar{\mathbf{S}}$, or separated for each data set and data type), which contain information about the overall influence of a parameter or the overall sensitivity of a subset of the data. In Eq. (5), σ_p is the parameter scaling factor, reflecting the parameter's expected variability or acceptable estimation uncertainty, and σ_z is the output scaling factor, reflecting measurement error, acceptable average residual, or acceptable prediction uncertainty. The interpretation and determination of these factors in various modeling contexts are discussed in detail in Finsterle (2015a).

The highly nonlinear character of most TOUGH models may warrant performing a global sensitivity analysis. The methods of Morris (1991) and Saltelli et al. (2006) are implemented in iTOUGH2. In the elementary effects method of Morris (1991), a point in the parameter space is selected, each parameter is perturbed—one at a time—and the impact on the output (elementary effect) is evaluated. The procedure is repeated for n_p random starting points (paths). The mean of the n_p elementary effects assesses the overall influence of the respective parameter on the output; the standard deviation indicates whether the effects are linear and additive or nonlinear, or whether interactions among the parameters are involved. The method provides valuable information about a parameter's relative influence for nonlinear models even for a relatively small number of paths, which makes it attractive for the analysis of computationally expensive TOUGH models.

The variance-based method of Saltelli et al. (2006) computes the output variance based on randomly perturbed parameters, and evaluates how much the output variance can be reduced by fixing each parameter. This provides first-order sensitivity indices for each parameter. In addition, the total sensitivity index can be computed by varying one parameter while keeping all other parameters fixed. The total sensitivity index includes the total effect of each parameter, accounting for interaction effects with other parameters. The two global sensitivity methods are compared and related to each other in a TOUGH modeling study performed by Wainwright et al. (2014).

A data-worth analysis (Dausman et al., 2010) is a byproduct of a local sensitivity analysis. It identifies the contribution each (potential or existing) data point makes to the solution of an inverse problem and a subsequent predictive simulation. It examines how the addition of potential data (or removal of existing data) reduces (or increases) the uncertainty in predictions made by a model that will be calibrated against these data. iTOUGH2 supports a workflow in which calibration and prediction phases are combined in a single data-worth analysis, which thus automatically identifies data that contain information about those parameters that are most influential on the predictions of interest. The general approach is visualized in Fig. 2. A set of parameters that potentially influence key predictions is shared between two numerical models: one simulating actual or potential calibration data, and the other simulating the ultimate system response of interest. A notional inversion is performed using the calibration model, and the covariance matrix of the estimated parameters is calculated. Uncertainties in these parameters are then propagated through the prediction model to arrive at a covariance matrix of the target model output. This process is repeated by removing existing or adding potential calibration data points (or entire data sets). The relative change in the prediction uncertainty is calculated and used as a measure of data worth. The process is explained and demonstrated in detail in Finsterle (2015a).

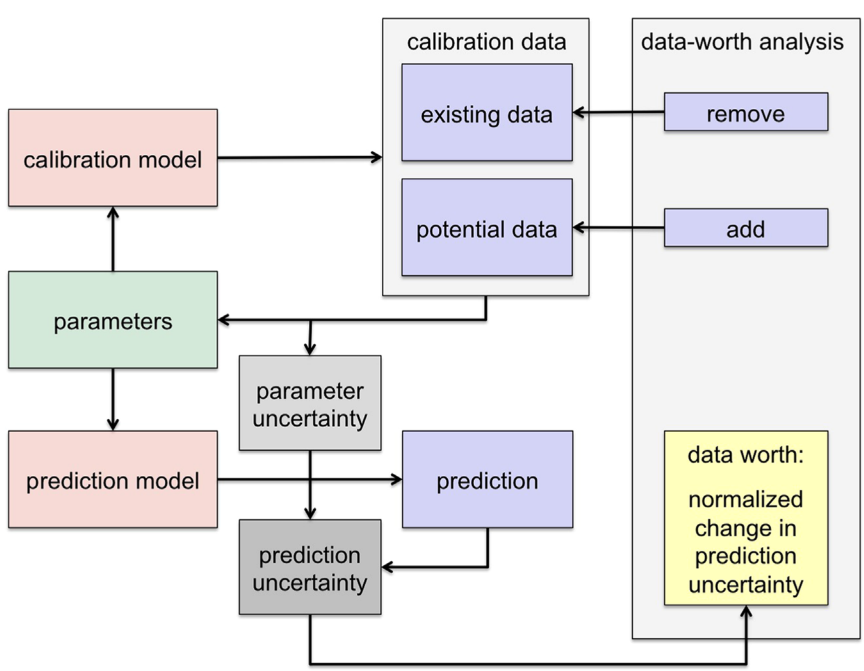

Fig. 2. Elements of data-worth analysis. Influential parameters to be estimated from calibration data are shared with a prediction model to determine how the removal of existing data or addition of potential data impacts prediction uncertainty.

5. Uncertainty propagation analysis

Uncertainty in a model prediction consists of multiple components. Specifically, errors or uncertainties in the conceptual model tend to have a dominant effect not only on the numerical results, but also on the interpretation of these results. Nevertheless, we limit the discussion here on the impact of uncertainties in parameters (which includes any aspect of the conceptual model that can be parameterized) on model predictions.

Assuming normality, the covariance matrix of model predictions $\hat{\mathbf{z}}$ is given by

$$\mathbf{C}_{\hat{z}\hat{z}} = \hat{\mathbf{J}}\mathbf{C}_{pp}\hat{\mathbf{J}}^{T} \tag{6}$$

where the Jacobian $\hat{\mathbf{J}}$ holds sensitivity coefficients of the model predictions with respect to the parameters $\mathbf{p}$, whose uncertainties are described by $\mathbf{C}_{pp}$. Eq. (6) is not only efficiently evaluated in parallel, but also readily accounts for statistical correlations among the parameters. Moreover, iTOUGH2 provides a simple procedure to correct for mild nonlinearities, due to Carrera (1984) and demonstrated in Finsterle and Pruess (1995).

Monte Carlo sampling of the parameter space is a way of examining prediction uncertainty without making any assumptions about the linearity or probability distributions. To partly address the fact that sufficient coverage of the entire parameter space requires a very large number of function evaluations, iTOUGH2 uses Latin Hypercube sampling and executes the TOUGH simulations in parallel. A simple response surface approximation approach is also available (see below). Finally, accounting for statistical correlations among the uncertain input parameters is essential. Neglecting such correlations has a far greater impact on the resulting prediction uncertainty distributions than an insufficient number of samples. For standard Monte Carlos simulations, iTOUGH2 uses empirical orthogonal functions (Kitterød and Finsterle, 2004); to account for correlations in Latin Hypercube sampling, the method of Zhang and Pinder (2003) is employed.

6. Reduced-order modeling

Sampling-based methods (such as Monte Carlo methods, global sensitivity analyses, and global minimization algorithms) are generally unfeasible because of their computational demands. Reduced-order modeling approaches attempt to approximate either the objective function or the high-fidelity model itself with fast surrogate models; see Ravazi et al. (2012) for an overview. iTOUGH2 supports reduced-order modeling in that so-called snapshot simulations can be generated (or externally calculated results can be read in), and a simple inverse-distance weighted interpolation function is used to evaluate the objective function at unsampled points in the parameter space. Examples of other model-reduction approaches using TOUGH as the high-fidelity model are described in Lehikoinen et al. (2009, 2010), Pau et al. (2013, 2014, 2016), Zhang et al. (2016) and Liu et al. (2016).

7. Forward model enhancements

iTOUGH2 is wrapped around standard TOUGH2 (Pruess et al., 2012). However, many modifications to the TOUGH simulator have been made. Some of these features are motivated by the fact that —if used within the iTOUGH2 optimization framework—the simulation problem has to be solved in a single run without user intervention. This requirement has led to a number of useful features, such as the ability to connect steady-state and transient simulations, to change mesh geometry, to change primary variables and certain material properties and flags at specified restart times, and to select alternative convergence criteria.

Other enhancements were driven by specific user needs, such as the incorporation of non-Darcy flow based on the Forchheimer equation and choked flow in gas wells, coupled overland-groundwater flow (Akhavan et al., 2012), internal generation of spatially correlated, random property fields using geostatistics, anisotropic permeability modifiers, time-dependent Dirichlet and free-drainage boundary conditions, scaling of capillary strength parameter based on permeability and temperature, inclusion of the active fracture concept (Liu et al., 1998), material-related sinks and sources, coupling to geomechanics (see Section 8), and vapor-pressure reduction to prevent disappearance of the liquid phase.

A third group of enhancements includes features that simply increase user convenience, such as the signal handler (which allows the user to request printout or to terminate a TOUGH run at any point during the simulation), free-format and tabular reading, improved time-stepping and printout control, five- to nine-character element names, and intermediate saving of restart files. Most of these options are described in a dedicated report (Finsterle, 2015b). They are useful even if iTOUGH2 is only run in forward mode.

8. Joint inversion of production and deformation data

The coupling of TOUGH's multiphase flow and transport processes with rock mechanics is an active area of research and code development (Finsterle et al., 2014). These developments are partly motivated by the need to stimulate reservoirs through fracturing (e.g., for enhanced geothermal systems or tight shale gas formations), to understand the risk that confining layers are compromised (e.g., at carbon storage sites or in heat-generating nuclear waste repositories), to address the issue of induced seismicity (either as a risk or monitoring option during reservoir stimulation), and many other applications (e.g., uplift and subsidence calculations caused by injection and production operations).

Observable deformation data do not only contain information about geomechanical processes, but may help identify magnitude and direction of fluid and heat flows that cause stress changes in the reservoir. To be able (1) to determine geomechanical properties through inverse modeling, and (2) to make use of deformation data to constrain other properties and processes within a multiphysics joint inversion framework, non-isothermal fluid flow and geomechanics must be coupled. We integrated the ROCMECH simulator (Kim et al., 2012) into iTOUGH2. ROCMECH calculates elastoplastic deformations caused by thermal and mechanical stresses using a sequential approach. Geomechanical properties (such as Young's and shear moduli, Poisson's ratio, Biot coefficient, yield stress, hardening parameter, friction and dilation angles, and thermal dilation coefficient) are part of the list of parameters that can be estimated, and deformation observations (magnitude and orientation) can be used as calibration data.

iTOUGH2's multiphysics joint inversion capabilities are demonstrated here for a synthetic enhanced oil recovery (EOR) operation, in which steam with a temperature of 300 °C is injected to mobilize and displace heavy oil towards a production well through a highly heterogeneous reservoir. Injectors and producers are arranged in a five-spot configuration, and all wells were hydraulically stimulated, assuming to have created large ellipsoidal, vertical fractures. Sequential Gaussian simulations are used to generate a spatially correlated, anisotropic, heterogeneous

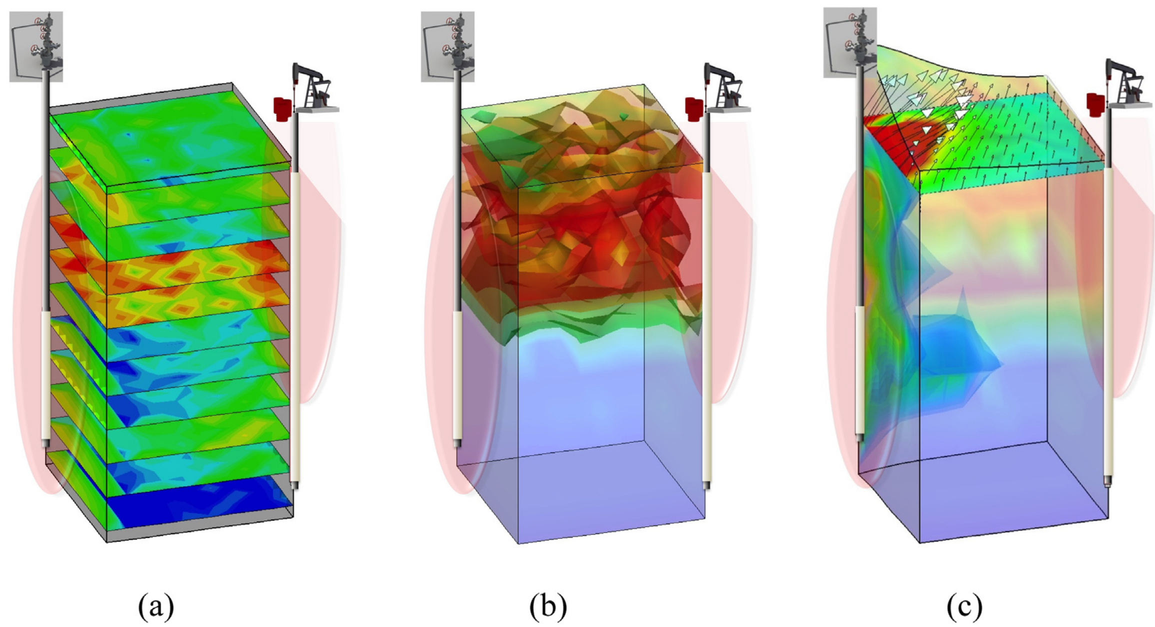

Fig. 3. Simulation of coupled nonisothermal fluid flow and geomechanics; (a) heterogeneous, anisotropic permeability field, (b) initial oil saturation, and (c) steam saturation isosurfaces and deformations at the top of the reservoir (vectors) after 50 days of steam injection and oil production.

permeability field (Fig. 3(a)). Initial pressure, temperature, and oil saturation distributions are calculated by running the system to near-steady conditions (Fig. 3(b)). Fifty days of steam injection followed by ten days of recovery are simulated using the non-isothermal three-phase module TMVOC (Pruess and Battistelli, 2002), which is integrated with ROCMECH for stress/strain calculations, GSLIB (Deutsch and Journel, 1992) for geostatistical simulations, PVM for parallelization, and iTOUGH2 for inverse modeling. The main outputs of interest from this simulation are the oil production rate and reservoir deformations.

Fig. 3(c) shows simulated deformations at the top of the reservoir in response to steam injection as well as oil and water production. Pressure and temperature changes lead to effective and thermal stresses that result in considerable uplift and subsidence, which may be observed at the land surface using InSAR or tiltmeters. Such deformation data can then be used (in combination with injection, production and thermal data) to calibrate the reservoir model using iTOUGH2's multiphysics joint inversion capabilities.

Table 3 shows the unknown or uncertain parameters that are analyzed by iTOUGH2. They include geometrical, hydrogeological, geomechanical, thermal, and geostatistical parameters, as well as autocorrelation coefficients for individual data sets. In addition, estimates of log-permeability modifiers at certain points are used to condition the geostatistically generated property field. This approach, referred to as the pilot point method (de Marsily, 1978; RamaRao et al., 1995), improves data fits by adjusting the reservoir structure during the inversion. The 61 parameters listed in Table 3 are determined by calibrating the model against data collected in the injection and production wells and at the top of the model (Table 3).

The calibration model is different from the model used to generate the synthetic data in that the permeabilities in each grid block are simply interpolated from the pilot point values using kriging, i.e., the resulting property field is much smoother than that obtained by geostatistical simulations. This means that the calibration model is fundamentally incapable of accurately reproducing the data, thus mimicking the conditions of matching actual data and avoiding the so-called “inverse crime” (Kaipio and Somersalo, 2004). The presence of a conceptual error is indicated by the follow diagnostic measures:

- Assuming that the prior assumptions about the final residuals (reflected in $\mathbf{C}_{zz}$) were solely based on measurement errors (i.e., without considering modeling errors), the overall match is declared poor, as reflected by an estimated error variance that is significantly larger than 1.0.
- The error in the conceptual model leads to systematic differences between the calculated system response and the corresponding data, as confirmed by the runs statistic performed by iTOUGH2. Consequently, the residuals are non-normal and autocorrelated, as reflected by non-zero estimates of the AR (1) parameters (see Table 3).
- The estimated parameter values are different from the “true” values used to generate the data. Specifically, parameters that are not very influential and that are strongly correlated to other parameters (see Table 3) tend to be adjusted such that they compensate for the systematic errors in the conceptual model. However, it is essential to realize that the parameters are—by definition—related to the conceptual model, so the notion of a “true” parameter set does not strictly apply, neither in reality nor in a synthetic study that avoids committing an inverse crime.
- Estimation uncertainty is strongly related to a parameter's overall independence, i.e., it is not sufficient to have a high composite influence measure. This is specifically true for absolute permeability, as shown in Table 4.
- Similarly, high composite sensitivity does not necessarily imply that the corresponding data set is of high value, as it may contain redundant information. This is specifically true for densely sampled tiltmeter data, as shown in Table 4.

It is not the purpose of this review article to examine this particular inversion in detail; the goal is simply to demonstrate the capabilities of iTOUGH2 to analyze disparate data and variety of parameter types using a joint inversion approach that incorporates a coupled process simulator.

Table 3
Parameters estimated by multiphysics joint inversion.

Parameter	Units	Estimated value	Uncertainty	Composite influence[a]	Independence[b]
Hydrogeological Parameters[c]					
k_{rh} hori. perm. reservoir	log (m^2)	-13	3.2	95,690	0.001
k_{rv} vert. perm. reservoir	log (m^2)	-14	3.2	95,700	0.001
k_{f1} perm. fracture 1	log (m^2)	-9.2	0.2	44,330	0.14
k_{f2} perm. fracture 2	log (m^2)	-12	0.4	38,970	0.15
f_k perm. modifiers	log (–)	$-2 < f_k < 2$	$0.01 < \sigma < 4$	~5000	~0.3
ϕ_r porosity reservoir	(–)	0.30	0.03	10,190	0.19
ϕ_{f1} porosity fracture 1	(–)	0.25	0.22	4820	0.13
ϕ_{f2} porosity fracture 2	(–)	0.12	>0.12	6250	0.10
S_{wr} residual water sat.	(–)	0.14	0.02	27,470	0.12
S_{or} residual oil sat.	(–)	0.05	0.01	21,250	0.48
n_{RP} exponent	(–)	2.05	0.01	28,000	0.18
n_{PC} exponent	(–)	1.73	0.02	13,550	0.21
Geometrical Parameters[d]					
Z_1 elevation fracture 1	(m)	-86	2.2	7000	0.58
Z_2 elevation fracture 2	(m)	-27	0.9	4810	0.16
dX_1 X extent fracture 1	(m)	65.0	>65.0	1230	0.001
dX_2 X extent fracture 2	(m)	20.6	0.2	8200	0.46
dZ_1 Z extent fracture 1	(m)	39.4	0.1	13,590	0.68
dZ_2 Z extent fracture 2	(m)	43.5	0.5	4700	0.38
Thermal Parameters					
λ heat conductivity	(W m^{-1} °C^{-1})	3.0	0.6	4710	0.22
C specific heat	(J kg^{-1} °C^{-1})	1200[g]	–	2310	0.39
ρ rock grain density	(kg m^{-3})	2500[g]	–	2070	0.01
α_T thermal dilation	log (°C^{-1})	-5.08	0.02	49,930	0.29
Geomechanical Parameters					
E Young's modulus	log (Pa)	9.6	3.0	37,290	0.004
G Shear modulus	log (Pa)	8.3	3.0	37,280	0.004
ν Poisson's ratio	(–)	0.3[g]	–	1410	0.37
α Biot coefficient	(–)	0.86	0.05	4870	0.03
Geostatistical Parameters[e]					
c_0 nugget effect	log (–^{2})	0.265	0.004	18,580	0.19
c sill value	log (–^{2})	0.200	0.005	19,810	0.18
a correlation length	(m)	55.5	1.0	2760	0.22
β_h hori. anisotropy	log (m)	0.8	>10.0	36,860	0.00
β_v vert. anisotropy	log (m)	-0.7	>10.0	36,860	0.00
Statistical Parameters[f]					
ρ_1 AR(1) deformation	(–)	0.2	>0.2	7450	0.003
ρ_1 AR(1) tilt	(–)	-0.1	>0.1	120	0.97
ρ_1 AR(1) pressure	(–)	0.4	>0.4	3820	0.003
ρ_1 AR(1) steam rate	(–)	0.8	0.1	1490	0.71
ρ_1 AR(1) oil rate	(–)	0.9	0.1	2000	0.06
ρ_1 AR(1) water rate	(–)	0.4	0.1	890	0.63

[a] Sum of the absolute values of the scaled sensitivity coefficients (Eq. (5)) for a given parameter.
[b] Ratio of the conditional to the marginal estimation uncertainty; a value of 1 indicates a perfectly independent parameter; a value of 0 refers to a fully correlated parameter.
[c] k is absolute permeability; f_k are log-permeability modifiers at 24 pilot points (not listed individually); ϕ is porosity; S_{wr}, S_{or}, and n_{RP} are parameters of the Stone (1970) three-phase relative permeability functions; n_{CP} is the exponent in the Parker et al. (1987) capillary pressure functions.
[d] Z is the elevation of the midpoint of the ellipsoidal hydrofractures; dX and dZ are the lengths of the semiaxes of the ellipsoidal hydrofractures.
[e] Geostatistical parameters refer to a spherical semivariogram (Deutsch and Journel, 1992).
[f] ρ is the autocorrelation coefficient of a first-order AR(1) autoregressive model.
[g] Parameter is at its lower or upper bound.

9. Concluding remarks

Since its first public release in 1997, iTOUGH2 has evolved from an automatic history matching tool for TOUGH models to a more comprehensive simulation-optimization platform that allows scientists and reservoir engineers to better integrate data into their models, design experiments, field tests, and monitoring campaigns, assess parameter and prediction uncertainty, and optimize operations for the recovery or storage of resources and contaminants. While the optimization component of iTOUGH2 greatly enhances the analysis of simulation results, it is our opinion that an accurate, physically based process simulator needs to be at the core of modeling studies that aim to improve understanding of the subsurface and reliably predict its response to natural changes or man-made interventions. Multiphysics forward and joint inverse modeling supports these aims.

We will continue to update iTOUGH2 and add new features and analysis methods to both its forward model and inversion framework in response to user requests, and to address scientific challenges.

The source code of iTOUGH2 is licensed by the University of California and distributed through the Berkeley Lab Software Center at http://esd1.lbl.gov/research/projects/tough/licensing/index.html#agreements.

Table 4
Observations used for multiphysics joint inversion.

Observation	Units	σ_p[a]	R[b]	Composite sensitivity[c]	Data Worth[d]
Δz uplift/ subsidence	(m)	0.01	0.91	10,420	4.5
θ_x tilt X	(deg)	0.001	0.88	55,860	5.7
θ_y tilt Y	(deg)	0.001	0.94	114,110	2.2
P_{rec} recovery pressure	(bar)	1.0	0.99	390	0.8
T_{pro} temperature	(°C)	0.1	0.91	10,783	0.1
q_{steam} steam injection	(kg s^{-1})	5.0	0.99	402,780	40.1
q_{oil} oil production	(kg s^{-1})	1.0	0.99	12,370	19.6
q_{wat} water production	(kg s^{-1})	5.0	0.99	219,320	26.9

[a] Square-root of diagonal element of a priori observation covariance matrix $\mathbf{C}_{zz}$; this standard deviation is also used to generate normally distributed measurement noise, added to the synthetic data.
[b] Pearson correlation coefficient of regression observed vs. calculated.
[c] Sum of the absolute values of the scaled sensitivity coefficients (Eq. (5)) for a given data set.
[d] Data-worth Metric 1 of Finsterle (2015a).

Acknowledgments

We appreciate the anonymous reviewers' constructive comments. iTOUGH2 has been developed over many years with contributions from the co-authors and the following colleagues: R. Ahlers, M. Akhavan, A. Arnaldsson, A. Battistelli, G. Björnsson, L. Blanco Martin, D. Bullivant, J. Doetsch, C. Doughty, T. Ghezzehei, F. Hale, T. Hiratsuka, M. Jonsson, N. Keen, J. Kim, N.-O. Kitterød, H.-H. Liu, L. Magnusdottir, J. O'Sullivan, M. O'Sullivan, R. Senger, T. Sonnenborg, S. Webb, S. White, Y.-S. Wu, M. Takeda, K. Zhang, and others. iTOUGH2 turning out to be a useful tool is mainly due to Karsten Pruess, the developer of its powerful engine: The TOUGH simulator. This work was supported, in part, by the U.S. Dept. of Energy under Contract no. DE-AC02-05CH11231. Enhancements to iTOUGH2 were sponsored by numerous DOE offices, non-DOE sponsors, as well as collaborating institutions and individuals. Some developments were funded through the TOUGH Development Grant, which is supplied by the TOUGH developers, who donate their portion of the royalty income from TOUGH licenses to the Grant.

References

Adams, B.M., Bauman, L.E., Bohnhoff, W.J., Dalbey, K.R., Ebeida, M.S., Eddy, J.P., Eldred, M.S., Hough, P.D., Hu, K.T., Jakeman, J.D., Stephens, J.A., Swiler, L.P., Vigil, D. M., and Wildey, T.M., 2016. Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 6.4 User's Manual, Sandia Technical Report SAND2014-4633.
Ahlers, C.F., Finsterle, S., Bodvarsson, G.S., 1999. Characterization of subsurface pneumatic response at Yucca Mountain. J. Contam. Hydrol. 38 (1–3), 47–68.
Akhavan, M., Imhoff, P.T., Finsterle, S., Andres, A.S., 2012. Application of a coupled overland flow-vadose zone model to rapid infiltration basin systems. Vadose Zone J. 11 (2). http://dx.doi.org/10.2136/vzj2011.0140.
Aster, R.C., Borchers, B., Thurber, C.H., 2013. Parameter Estimation and Inverse Problems, 2nd ed. Academic Press, Oxford, UK.
Baca, R.G., Seth, M.S., 1996. Benchmark testing of thermohydrologic computer codes, Report CNWRA 96-003, Center for Nuclear Waste Regulatory Analyses, San Antonia, Texas, 45 pp.
Birkholzer, J.T., Cihan, A., Zhou, Q., 2012. Impact-driven pressure management via targeted brine extraction-Conceptual studies of CO_2 storage in saline formations. Int. J. Greenh. Gas Control 7, 168–180.
Birkle, P., 2011. Advances in geochemical modeling for geothermal applications. In: Bundschuh, J., Zilberbrand, M. (Eds.), Geochemical modeling of groundwater, vadose and geothermal systems. ISBN 978-0-415-66810-1CRC Press, Leiden, The Netherlands, p. 332.
Carrera, J., Neuman, S.P., 1986. Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior information. Water Resour. Res. 22 (2), 199–210.
Carrera, J., 1984. Estimation of aquifer parameters under transient and steady state conditions (Ph.D. dissertation), Dept. of Hydrol. and Water Resour., Univ. of Ariz., Tucson, AZ.
Commer, M., Kowalsky, M.B., Doetsch, J., Newman, G.A., Finsterle, S., 2014. MPiTOUGH2: a parallel parameter estimation framework for hydrological and hydrogeophysical applications. Comput. Geosci. 65, 127–135. http://dx.doi.org/10.1016/j.cageo.2013.06.011.
Dausman, A.M., Doherty, J., Langevin, C.D., Sukop, M.C., 2010. Quantifying data worth toward reducing predictive uncertainty. Ground Water 48 (5), 729–740.
de Marsily, G., 1978. De l'identification des systèmes hydrologiques (Ph.D. thesis), Univ. Paris VI, Paris.
Deutsch, C.V., Journel, A.G., 1992. GSLIB, Geostatistical Software Library and User's Guide. Oxford University Press, New York, p. 1992.
Doetsch, J., Kowalsky, M.B., Doughty, C., Finsterle, S., Ajo-Franklin, J.B., Carrigan, C. R., Yang, X., Hovorka, S.D., Daley, T.M., 2013. Constraining CO_2 migration by coupled modeling and inversion of electrical resistance and gas composition data. Int. J. Greenh. Gas Control 18, 510–522. http://dx.doi.org/10.1016/j.ijggc.2013.04.011.
Doherty, J., Hunt, R.J., 2009. Two statistics for evaluating parameter identifiability and error reduction. J. Hydrol. 366, 119–127.
Doherty, J., 2010. PEST: Model-independent Parameter Estimation, User manual: 5th edition, Watermark Numerical Computing, Brisbane, Australia, 336 pp.
Engelhardt, I., Finsterle, S., 2003. Thermal-hydrologic experiments with bentonite/crushed rock mixtures and estimation of effective parameters by inverse modeling. Appl. Clay Sci. 23, 111–120.
Engelhardt, I., Finsterle, S., Hofstee, C., 2003. Experimental and numerical investigation of flow phenomena in nonisothermal, variably saturated bentonite/crushed rock mixtures. Vadose Zone J. 2, 239–246.
Finsterle, S., 2000. Using the continuum approach to model unsaturated flow in fractured rock. Water Resour. Res. 36 (8), 2055–2066. http://dx.doi.org/10.1029/2000WR900122.
Finsterle, S., 2004. Multiphase inverse modeling: review and iTOUGH2 applications. Vadose Zone J. 3, 747–762.
Finsterle, S., 2005. Demonstration of optimization techniques for groundwater plume remediation using iTOUGH2. Environ. Model. Softw. 21 (5), 665–680. http://dx.doi.org/10.1016/j.envsoft.2004.11.012.
Finsterle, S., 2015a. Practical notes on local data-worth analysis. Water Resour. Res. 51 (12), 9904–9924. http://dx.doi.org/10.1002/2015WR017445.
Finsterle, S., Pruess, K., 1995. Solving the estimation-identification problem in two-phase flow modeling. Water Resour. Res. 31 (4), 913–924. http://dx.doi.org/10.1029/94WR03038.
Finsterle, S., Persoff, P., 1997. Determining permeability of tight rock samples using inverse modeling. Water Resour. Res. 33 (8), 1803–1811. http://dx.doi.org/10.1029/97WR01200.
Finsterle, S., Najita, J., 1998. Robust estimation of hydrogeologic model parameters. Water Resour. Res. 34 (11), 2939–2947. http://dx.doi.org/10.1029/98WR02174.
Finsterle, S., Faybishenko, B., 1999. Inverse modeling of a radial multistep outflow experiment for determining unsaturated hydraulic properties. Adv. Water Resour. 22 (5), 431–444. http://dx.doi.org/10.1016/S0309-1708(98)00030-X.
Finsterle, S., Kowalsky, M.B., 2008. Joint hydrological-geophysical inversion for soil structure identification. Vadose Zone J. 7, 287–293. http://dx.doi.org/10.2136/vzj2006.0078.
Finsterle, S., Zhang, Y., 2011a. Solving iTOUGH2 simulation and optimization problems using the PEST protocol. Environ. Model. Softw. 26, 959–968. http://dx.doi.org/10.1016/j.envsoft.2011.02.008.
Finsterle, S., Zhang, Y., 2011b. Error handling strategies in multiphase inverse modeling. Comput. Geosci. 37, 724–730. http://dx.doi.org/10.1016/j.cageo.2010.11.009.
Finsterle, S., Kowalsky, M.B., Pruess, K., 2012. TOUGH: model use, calibration and validation. Trans. ASABE 55 (4), 1275, 129.
Finsterle, S., Sonnenthal, E.L., Spycher, N., 2014. Advances in subsurface modeling: the TOUGH suite of simulators. Comput. Geosci. 65, 2–12. http://dx.doi.org/10.1016/j.cageo.2013.06.009.
Finsterle, S., Ahlers, C.F., Trautz, R.C., Cook, P.J., 2003. Inverse and predictive modeling of seepage into underground openings. J. Contam. Hydrol. 62–63, 89–109. http://dx.doi.org/10.1016/S0169-7722(02)00174-2.
Finsterle, S., Zhang, Y., Pan, L., Dobson, P., Oglesby, K., 2013. Microhole arrays for improved heat mining from enhanced geothermal systems. Geothermics 47, 104–115. http://dx.doi.org/10.1016/j.geothermics.2013.03.001.
Finsterle, S., Doughty, C., Kowalsky, M.B., Moridis, G.J., Pan, L., Xu, T., Zhang, Y., Pruess, K., 2008. Advanced vadose zone simulations using TOUGH. Vadose Zone J. 7, 601–609. http://dx.doi.org/10.2136/vzj2007.0059.
Finsterle, S., 2015b. Enhancements to the TOUGH2 Simulator Implemented in iTOUGH2, Report LBNL-7016E, Lawrence Berkeley National Laboratory, Berkeley, California.
Fowler, S.J., Kosakowski, G., Driesner, T., Kulik, D.A., Wilhelm, S., Masset, O., 2016. Numerical simulation of reactive fluid flow on unstructured meshes. Transp. Porous Media 112, 283–312. http://dx.doi.org/10.1007/s11242-016-0645-7.
Freifeld, B.M., Finsterle, S., Onstott, T.C., Toole, T., Pratt, L.M., 2008. Ground surface temperature reconstructions: using in situ estimates for thermal conductivity acquired with a fiber-optic distributed thermal perturbation sensor. Geophys. Res. Lett. 35, L14309. http://dx.doi.org/10.1029/2008GL034762.
Gallagher, P.M., Finsterle, S., 2004. Physical and numerical model of colloidal silica injection for passive site stabilization. Vadose Zone J. 3 (3), 917–925.
Geem, Z.W., Kim, J.-H., Loganathan, G.V., 2001. A new heuristic optimization

algorithm: harmony search. Simulation 76 (2), 60–68.
Ghezzehei, T.A., Trautz, R.C., Finsterle, S., Cook, P.J., Ahlers, C.F., 2004. Modeling coupled evaporation and seepage in ventilated tunnels. Vadose Zone J. 3 (3), 806–818.
Günther, T., Rücker, C., Spitzer, K., 2006. Three-dimensional modelling and inversion of DC resistivity data incorporating topography – II. Inversion. Geophys. J. Int. 166, 506–517.
Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. J. Hydrol. 377 (1–2), 80–91.
Hannon, M.J., Finsterle, S., 2016. Considering anisotropy in multidimensional pressure-pulse-decay experiments. Transp. Porous Media, (submitted).
Hommel, J., Lauchnor, E., Phillips, A., Gerlach, R., Cunningham, A.B., Helmig, R., Ebigbo, A., Class, H., 2015. A revised model for microbially induced calcite precipitation: improvements and new insights based on recent experiments. Water Resour. Res. 51, 3695–3715.
Hudson, J.A., Bäckström, A., Rutqvist, J., Jing, L., Backers, T., Chijimatsu, M., Christiansson, R., Feng, X.-T., Kobayashi, A., Koyama, T., Lee, H.-S., Neretnieks, I., Pan, P.Z., Rinne, M., Shen, B.T., 2009. Characterising and modelling the excavation damaged zone (EDZ) in crystalline rock in the context of radioactive waste disposal. Environmental Geology 57, 1275–1291.
Jung, Y., Imhoff, P., Finsterle, S., 2011. Estimation of landfill gas generation rate and gas permeability field of refuse using inverse modeling. Transp. Porous Media 90 (1), 41–58. http://dx.doi.org/10.1007/s11242-010-9659-8.
Jung, Y., Zhou, Q., Birkholzer, J.T., 2013. Early detection of brine and CO_2 leakage through abandoned wells using pressure and surface-deformation monitoring data: concept and demonstration. Adv. Water Resour. 62, 555–569.
Jung, Y., Pau, G.S.H., Finsterle, S., 2016. TOUGH3: a new base version of the TOUGH suite of codes, Comput. Geosci., (this issue).
Kaipio, J., Somersalo, E., 2004. Statistical and Computational Inverse Problems. Springer, New York, NY.
Kaya, E., O'Sullivan, M.J., Hochstein, M.P., 2014. A three dimensional numerical model of the Waiotapu, Waikite and Reporoa geothermal areas, New Zealand. J. Volcanol. Geotherm. Res. 283, 127–142.
Kim, J., Sonnenthal, E., Rutqvist, J., 2012. Formulation and sequential numerical algorithms of coupled fluid/heat flow and geomechanics for multiple porosity materials. Int. J. Numer. Methods Eng. 92 (5), 425–456.
Kirkpatrick, S., Gelatt Jr, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220 (4598), 671–680.
Kiryukhin, A.V., Asaulova, N.P., Finsterle, S., 2008. Inverse modeling and forecasting for the exploitation of the Pauzhetsky geothermal field, Kamchatka, Russia. Geothermics 37, 540–562.
Kiryukhin, A.V., Polyakov, A.Y., Mushinskii, A.V., 2014. Measurements of thermal conductivity and specific heat in volcanogenic rocks. J. Volcanol. Seismol. 8 (5), 283–293.
Kitterød, N.-O., Finsterle, S., 2004. Simulating unsaturated flow fields based on saturation measurements. J. Hydraul. Res. 42, 121–129.
Kowalsky, M.B., Finsterle, S., Rubin, Y., 2004. Estimating flow parameter distributions using ground-penetrating radar and hydrological measurements during transient flow in the vadose zone. Adv. Water Resour. 27 (6), 583–599. http://dx.doi.org/10.1016/j.advwatres.2004.03.003.
Kowalsky, M.B., Gasperikova, E., Finsterle, S., Watson, D., Hubbard, S.S., 2011. Coupled modeling of hydrogeochemical and electrical resistivity data for exploring the impact of recharge on subsurface contamination. Water Resour. Res. 47, W02509. http://dx.doi.org/10.1029/2009WR008947.
Kowalsky, M.B., Birkholzer, J., Peterson, J., Finsterle, S., Mukhopadhyay, S., Tsang, Y., 2008. Sensitivity analysis for joint inversion of ground-penetrating radar and thermal-hydrological data from a large-scale underground heater test. Nucl. Technol. 164 (2), 169–179.
Kowalsky, M.B., Finsterle, S., Peterson, J., Hubbard, S., Rubin, Y., Majer, E., Ward, A., Gee, G., 2005. Estimation of field-scale soil hydraulic parameters and dielectric parameters through joint inversion of GPR and hydrological data. Water Resour. Res. 41, W11425. http://dx.doi.org/10.1029/2005WR004237.
Kowalsky, M.B., Finsterle, S., Williams, K.H., Murray, C., Commer, M., Newcomer, D., Englert, A., Steefel, C.I., Hubbard, S.S., 2012. On parameterization of the inverse problem for estimating aquifer properties using tracer data. Water Resour. Res. 48, W06535. http://dx.doi.org/10.1029/2011WR011203.
Lee, S.J., McPherson, B.J., Vasquez, F.G., 2015. Leakage pathway estimation using iTOUGH2 in a multiphase flow system for geologic CO_2 storage. Environ. Earth Sci. 74 (6), 5111–5128.
Lehikoinen, A., Finsterle, S., Voutilainen, A., Kowalsky, M.B., Kaipio, J.P., 2009. Dynamical inversion of geophysical ERT data: state estimation in the vadose zone. Inverse Probl. Sci. Eng. 17 (6), 715–736. http://dx.doi.org/10.1080/17415970802475951.
Lehikoinen, A., Huttunen, J.M.J., Finsterle, S., Kowalsky, M.B., Kaipio, J.P., 2010. Dynamic inversion for hydrological process monitoring with electrical resistance tomography under model uncertainties. Water Resour. Res. 46, W04513. http://dx.doi.org/10.1029/2009WR008470.
Levenberg, K., 1944. A method for the solution of certain nonlinear problems in least squares. Q. Appl. Math. 2, 164–168.
Li, L., Gawande, N., Kowalsky, M.B., Steefel, C., Hubbard, S.S., 2011. Physicochemical heterogeneity controls on uranium bioreduction rates at the field scale. Environ. Sci. Technol. 45 (23), 9959–9966.
Linde, N., Vrugt, J.A., 2013. Distributed soil moisture from crosshole ground-penetrating radar travel times using stochastic inversion. Vadose Zone J. 12 (1). http://dx.doi.org/10.2136/vzj2012.0101.
Linde, N., Finsterle, Hubbard, S., 2006. Inversion of tracer test data using tomographic constraints. Water Resour. Res. 42 (4), W04410. http://dx.doi.org/10.1029/2004WR00380.
Liu, H.-H., Doughty, C., Bodvarsson, G.S., 1998. An active fracture model for unsaturated flow and transport in fractured rocks. Water Resour. Res. 34 (10), 2633–2646.
Liu, Y., Pau, G.S.H., Finsterle, S., 2016. Implicit sampling combined with reduced order modeling for the inversion of vadose zone hydrological data, Comput. Geosci., (this issue).
Mannington, W., O'Sullivan, M., Bullivant, D., 2004. Computer modelling of the Wairakei-Tauhara geothermal system, New Zealand. Geothermics 33 (4), 401–419.
Marquardt, D.W., 1963. An algorithm for least squares estimation of nonlinear parameters. SIAM J. Appl. Math. 11, 431–441.
MDH Engineered Solutions Corporation, 2005. Evaluation of Computer Models for Predicting the Fate and Transport of Hydrocarbons in Soil and Groundwater, Pub No. 808, ISBN No. 0-7785-4040-5, p. 61.
Moridis, G.J., Finsterle, S., Heiser, J., 1999. Evaluation of alternative designs for an injectable barrier at the Brookhaven National Laboratory Site, Long Island, New York. Water Resour. Res. 35 (10), 2937–2953. http://dx.doi.org/10.1029/1999WR900184.
Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics 33 (2), 161–174.
Mukhopadhyay, S., Tsang, Y.W., Finsterle, S., 2009. Parameter estimation from flowing fluid temperature logging data in unsaturated fractured rock using multiphase inverse modeling. Water Resour. Res. 45, W04414. http://dx.doi.org/10.1029/2008WR006869.
Mukhopadhyay, S., Doughty, C., Bacon, D., Li, J., Wei, L., Yamamoto, H., Gasda, S., Hosseini, S.A., Nicot, J.-P., Birkholzer, J.T., 2015. The Sim-SEQ project: comparison of selected flow models for the S-3 site. Transp. Porous Media 108 (1), 207–231.
Nelder, J.A., Mead, R., 1965. A simplex method for function minimization. Comput. J. 7 (4), 308–313.
Oldenburg, C.M., Law, D.H.-S., LeGallo, Y., White, S.P., 2003. Mixing of CO_2 and CH_4 in gas reservoirs: code comparison studies. Greenhouse Gas Control Technologies 1, 443–448.
O'Sullivan, M.J., Prues, K., Lippmann, M.J., 2001. State of the art of geothermal reservoir simulation. Geothermics 30, 395–429.
Parker, J.C., Lenhard, R.J., Kuppusamy, T., 1987. A parametric model for constitutive properties governing multiphase flow in porous media. Water Resour. Res. 12 (1), 618–624.
Pau, G., Zhang, Y., Finsterle, S., 2013. Reduced order models for many-query subsurface flow applications. Comput. Geosci. 7 (4), 705–721. http://dx.doi.org/10.1007/s10596-013-9349-z.
Pau, G.S.H., Finsterle, S., Zhang, Y., 2016. Fast high-resolution prediction of multiphase flow in fractured formations. Adv. Water Resour. 88, 80–85. http://dx.doi.org/10.1016/j.advwatres.2015.12.008.
Pau, G.S.H., Zhang, Y., Finsterle, S., Wainwright, H., Birkholzer, J., 2014. Reduced order modeling in iTOUGH2. Comput. Geosci. 65, 118–126. http://dx.doi.org/10.1016/j.cageo.2013.08.008.
Poeter, E.P., Hill, M.C., Lu, D., Tiedeman, C.R., Mehl, S., 2014. UCODE_2014, with New Capabilities to Define Parameters Unique to Predictions, Calculate Weights Using Simulated Values, Estimate Parameters with SVD, Evaluate Uncertainty with MCMC, and More, Integrated Groundwater Modeling Center, Report Number GWMI 2014-02, 172 pp.
Porras, E.A., Tanaka, T., Fujii, H., Itoi, R., 2007. Numerical modeling of the Momotombo geothermal system, Nicaragua. Geothermics 36 (4), 304–329.
Pruess, K., 2004. The TOUGH codes — a family of simulation tools for multiphase flow and transport processes in permeable media. Vadose Zone J. 3, 738–746.
Pruess, K., Battistelli. A., 2002. TMVOC, a Numerical Simulator for Three-phase Non-isothermal Flows of Multicomponent Hydrocarbon Mixtures in Saturated-unsaturated Heterogeneous Media, Report LBNL-49375, Lawrence Berkeley National Laboratory, Berkeley, California, 192 pp.
Pruess, K., Oldenburg, C., Moridis, G. 2012. TOUGH2 User's Guide, Version 2.1, Report LBNL-43134, Lawrence Berkeley National Laboratory, Berkeley, California, 204 pp.
RamaRao, B.S., LaVenue, A.M., de Marsily, G., Marietta, M.G., 1995. Pilot-point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields—1, theory and computational experiments. Water Resour. Res. 31 (3), 475–494.
Rasmusson, K., Rasmusson, M., Fagerlund, F., Bensabat, J., Tsang, Y., Niemi, A., 2014. Analysis of alternative push-pull-test-designs for determining in situ residual trapping of carbon dioxide. Int. J. Greenh. Gas Control 27, 155–168.
Ravazi, S., Tolson, B.A., Burn, D.H., 2012. Review of surrogate modeling in water resources. Water Resour. Res. 48, W07401.
Regis, R.G., Shoemaker, C.A., 2007. A stochastic radial basis function method for the global optimization of expensive functions. INFORMS J. Comput. 19 (4), 497–509.
Saltelli, A., Ratto, M., Tarantola, S., Campolongo, F., 2006. European Commission, Joint Research Centre of Ispra (I), Sensitivity analysis practices: Strategies for model-based inference. Reliab. Eng. Syst. Safe. 91 (10-11), 1109–1125.
Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. Global Sensitivity Analysis, the Primer. John Wiley & Sons Ltd, West Sussex, England, p. 292.
Schoups, G., Vrugt, J.A., 2010. A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heterscedastic, and

non-Gaussian errors. Water Resour. Res 46, W10531.
Steefel, C.I., Appelo, C.A.J., Arora, B., Jacques, D., Kalbacher, T., Kolditz, O., Lagneau, V., Lichtner, P.C., Mayer, K.U., Meeussen, J.C.L., Molins, S., Moulton, D., Shao, H., Simunek, J., Spycher, N., Yabusaki, S.B., Yeh, G.T., 2015. Reactive transport codes for subsurface environmental simulation. Comput. Geosci. 19 (3), 445–478.
Stone, H.L., 1970. Probability model for estimating three-phase relative permeability. J. Pet. Technol. 22 (2), 214–218.
Storn, R., Price, K., 1997. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. J. Glob. Optim. 11, 341–359.
Tonkin, M., Doherty, J., 2005. A hybrid regularized inversion methodology for highly parameterized models. Water Resour. Res. 41, W10412.
Unger, A., Finsterle, S., Bodvarsson, G.S., 2004. Transport of radon gas into a tunnel at Yucca Mountain—estimating large-scale fractured tuff hydraulic properties and implications for the ventilation system. J. Contam. Hydrol. 70 (3–4), 152–171. http://dx.doi.org/10.1016/j.jconhyd.2003.07.001.
Wainwright, H.M., Finsterle, S., Zhou, Q., Birkholzer, J.Y., 2013. Modeling the performance of large-scale CO_2 storage systems: a comparison of different sensitivity analysis methods. Int. J. Greenh. Gas Control 17, 189–205. http://dx.doi.org/10.1016/j.ijggc.2013.05.007.
Wainwright, H.M., Finsterle, S., Jung, Y., Zhou, Q., Birkholzer, J.T., 2014. Making sense of global sensitivity analyses. Comput. Geosci. 65, 84–94. http://dx.doi.org/10.1016/j.cageo.2013.06.006.
Wang, J.S.Y., Trautz, R.C., Cook, P.J., Finsterle, S., James, A.L., Birkholzer, J., 1999. Field tests and model analyses of seepage into drift. J. Contam. Hydrol. 38 (1–3), 323–347, doi:0.1016/S0169-7722(99)00019-4.
Wellmann, F., Croucher, A., Finsterle, S., 2014. Adding geology to the equation: towards integrating structural geological data into inverse modeling with iTOUGH2. Comput. Geosci. 65, 95–109. http://dx.doi.org/10.1016/j.cageo.2013.10.014.
White, S.P., Creighton, A.L., Bixley, P.F., Kissling, W.M., 2004. Modeling the dewatering and depressurization of the Lihir open-pit gold mine, Papua New Guinea. Geothermics 33 (4), 443–456.
Ye, M., Meyer, P.D., Neuman, S.P., 2008. On model selection criteria in multimodel analysis. Water Resour. Res. 44, W03428.
Zhang, Y., Hubbard, S., Finsterle, S., 2011a. Factors governing sustainable groundwater pumping near a river. Ground Water 49 (3), 432–444. http://dx.doi.org/10.1111/j.1745-6584.2010.00743.x.
Zhang, Y., Liu, Y., Pau, G.S.H., Oladyshkin, S., Finsterle, S., 2016. Evaluation of multiple reduced-order models to enhance confidence in global sensitivity analyses. Int. J. Greenh. Gas Control 49, 217–226. http://dx.doi.org/10.1016/j.ijggc.2016.03.003.
Zhang, Y., Freifeld, B., Finsterle, S., Leahy, M., Ennis-King, J., Paterson, L., Dance, T., 2011b. Single-well experimental design for studying residual trapping of supercritical carbon dioxide. Int. J. Greenh. Gas Control 5, 88–98. http://dx.doi.org/10.1016/j.ijggc.2010.06.011.
Zhang, Y.Q., Pinder, G., 2003. Latin hypercube lattice sample selection strategy for correlated random hydraulic conductivity fields. Water Resour. Res. 39 (8), 1226. http://dx.doi.org/10.1029/2002WR001822.