Environmental Modelling & Software 100 (2018) 213–221

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

A tool for efficient, model-independent management optimization under uncertainty

Jeremy T. White [a, *], Michael N. Fienen [b], Paul M. Barlow [c], Dave E. Welter [d]

[a] GNS Science, Wairakei, New Zealand
[b] U.S. Geological Survey, Middleton, WI, United States
[c] U.S. Geological Survey, Reston, VA, United States
[d] Computational Water Resource Engineering, United States

ARTICLE INFO

Article history:
Received 22 June 2017
Received in revised form
26 September 2017
Accepted 9 November 2017

Keywords:
Uncertainty quantification
Model independent
Parameter estimation
Optimization under uncertainty

ABSTRACT

To fill a need for risk-based environmental management optimization, we have developed PESTPP-OPT, a model-independent tool for resource management optimization under uncertainty. PESTPP-OPT solves a sequential linear programming (SLP) problem and also implements (optional) efficient, "on-the-fly" (without user intervention) first-order, second-moment (FOSM) uncertainty techniques to estimate model-derived constraint uncertainty. Combined with a user-specified risk value, the constraint uncertainty estimates are used to form chance-constraints for the SLP solution process, so that any optimal solution includes contributions from model input and observation uncertainty. In this way, a "single answer" that includes uncertainty is yielded from the modeling analysis. PESTPP-OPT uses the familiar PEST/PEST++ model interface protocols, which makes it widely applicable to many modeling analyses. The use of PESTPP-OPT is demonstrated with a synthetic, integrated surface-water/groundwater model. The function and implications of chance constraints for this synthetic model are discussed.

Software availability

The source code for PESTPP-OPT is available as part of the PEST++ software suite (Welter et al., 2015):

https://github.com/dwelter/pestpp.

In addition to the source code, the git repository includes statically-linked OSX and PC executables, as well as three example problems adapted from GWM (Ahlfeld et al., 2005) (e.g., the "dewater" problem, the "seawater" problem and the "supply2" problem), including the example problem presented herein.

1. Introduction

Environmental modeling analyses are frequently undertaken with the focus of providing a decision-support tool for resource managers, a critical role for modeling (Gorelick and Zheng, 2015; Horne et al., 2016). Rigorous resource management optimization is a widely-recognized approach for providing optimal and unbiased answers to resource management questions. Readers are referred to Singh (2012); Yeh (2015); Gorelick and Zheng (2015); Horne et al. (2016) among others, for recent reviews of the importance and application of environmental resource management optimization.

However, for environmental models to be used appropriately in a risk-based decision-making context, these models should include estimates of uncertainty in important model outcomes. This uncertainty conveys a clear understanding of the reliability of the simulation results and the management solutions that depend on these simulation results, providing a margin of safety to account for imperfections in the model and data supplying it (Anderson et al., 2015). Unfortunately, in practice, providing resource managers with a range of possible model outcomes can raise more questions than the modeling analysis answers. Phrases such as "How can we manage to a range of outcomes?" or "We need a single number" are common responses to modeling analyses presented in the context of uncertainty. One possible solution to this conundrum is the use of optimization under uncertainty techniques (Sahinidis, 2004). This type of management optimization problem seeks an optimal solution to a resource-management problem, but includes recognized sources of uncertainty. In this way, the optimal solution provides a "single answer", but that answer includes

* Corresponding author.
E-mail address: j.white@gns.cri.nz (J.T. White).

https://doi.org/10.1016/j.envsoft.2017.11.019
1364-8152/

uncertainty—uncertainty arising from uncertain model inputs (and optionally observation noise) is propagated to constraints, which in turn affects the optimal solution.

Unfortunately, while many approaches and techniques to optimization under uncertainty have been proposed in the literature (Joodavi et al., 2015; Zekri et al., 2015; Nouiri et al., 2015; Tsoukalas and Makropoulos, 2015; Sreekanth et al., 2016; Beh et al., 2017), few, if any, generalized (model-independent/non-intrusive) tools exist for practitioners to apply these techniques to environmental models (Gorelick and Zheng, 2015; Horne et al., 2016). Compounding this lack of tools is the large computational burden and high dimensionality associated with many types of environmental models—especially groundwater or integrated surface-water/groundwater models—that can preclude the application of many approaches to optimization and optimization under uncertainty.

Herein, we present PESTPP-OPT, an efficient, model-independent (non-intrusive) tool for optimization under uncertainty. PESTPP-OPT implements the simplex algorithm (Dantzig et al., 1955) to solve the linear programming problem and also implements sequential linear programming problem (SLP: Ahlfeld and Mulligan, 2000) to resolve mild nonlinearities in the relation between decision variables and constraints arising from the numerical model. We extend the work of Wagner and Gorelick (1987) to use a Bayesian formulation of first-order, second-moment (FOSM) uncertainty techniques to efficiently and seamlessly (without user intervention) estimate prior or posterior uncertainty in the constraints derived from model output, thereby propagating model parameter and observation uncertainty to the constraints used in the optimization process. This FOSM-based constraint uncertainty estimation happens “on-the-fly” (programmatically and without user intervention) and, depending on the selected settings, requires no additional model runs.

Because PESTPP-OPT uses the familiar PEST (Doherty, 2015)/PEST++ (Welter et al., 2015) model-independent (non-intrusive) framework, the existing PEST and PEST++ user base will now be able to easily apply these sophisticated management optimization techniques with little additional user effort beyond typical parameter estimation.

The following sections briefly describe the theory of SLP and FOSM, and present an example application of PESTPP-OPT to solve a chance-constrained integrated groundwater surface-water management problem.

2. Theory

2.1. Terminology

Parameter estimation and management optimization share many common elements, but, in some cases, employ different terminology, or worse, have similar terminology with differing definitions. Therefore, we now explicitly define how we use these terms in this paper.

In parameter estimation (PE) and uncertainty quantification (UQ) parlance, “parameters” are uncertain model inputs (any numeric quantity used both the simulator) that are nominated for adjustment during history matching and/or uncertainty analysis; “observations” in PE and UQ analyses are measured data, typically collected from the environmental system being modeled. In management optimization parlance, following Ahlfeld et al. (2009), “decision variables” are model inputs whose values are to be determined by the optimization process. That is, decision variables are some model inputs that can be controlled. Constraints are conditions that must be satisfied by any optimal solution, including maximum and minimum allowable values for decision variables, or, alternatively, constraints may be derived from model output, or both. Constraints based on model output may include simulated groundwater levels, stream flows, and/or streamflow depletions (Barlow and Leake, 2012; Fienen et al., 2017). Constraints may also be derived from model output. For example, simulated differences in model simulated states in space or time (e.g., drawdowns or water level difference across a confining unit)—most simulators only output simulated states, but users can post-process these states to calculate differences.

In PE and UQ parlance, “forecasts” are unobserved quantities of interest derived, at least in part, from model outputs. Because model inputs (e.g., parameters) are uncertain, so too are forecasts in as much as a given forecast depend on the parameters. Similarly, constraints that are derived from model output are unobserved quantities that, because of parameter uncertainty, are also uncertain. Note that observations of system states may be available at constraint locations (e.g. water levels from an existing well). However, the simulated response to a given management scenario remains uncertain, and is therefore, similar to a “forecast” in a PE and UQ analysis.

In PE and UQ parlance, the “objective function” is the functional typically composed of differences between observations and model-simulated equivalents. Typically, this functional is based on the $\mathscr{L}_2$ norm (sum of squares) of the differences. The focus of PE is to minimize this functional. In management optimization, the objective function is composed of combinations of decision variables and may either be minimized or maximized, depending on the specific application, subject to constraints. Moreover, herein we use linear programming, so that the objective function is a weighted linear combination of decision variables.

2.2. Linear programming

Linear programming (LP) is a solution to the optimization problem that relies on an assumption of a linear relationship between decision variables and constraints as well as an objective function that is a linear combination of decision variables (Nocedal and Wright, 2006). Using vector notation, LP can be summarized as

$$\begin{aligned} &\text{minimize}: \mathbf{c^T x} \\ &\text{subject to}: \ \mathbf{Ax} \le \mathbf{b} \\ &\mathbf{x} \ge 0 \end{aligned} \tag{1}$$

where $\mathbf{x}$ is a vector of m decision variables, $\mathbf{c}$ is vector of m objective function coefficients, $\mathbf{A}$ is an $n \times m$ matrix of constraint coefficients, and $\mathbf{b}$ is a vector of n specified constraint values. Following Ahlfeld and Mulligan (2000), herein the matrix $\mathbf{A}$ is called a response matrix and is calculated, in part, by evaluating the model with the perturbed decision variables and recording the change in constraints. Specifically,

$$\frac{\partial \mathbf{A_i}}{\partial \mathbf{x_j}} \approx \frac{\mathbf{A}_i(\mathbf{x}_j + \delta\mathbf{x}_j) - \mathbf{A}_i(\mathbf{x}_j)}{\delta\mathbf{x}_j} \tag{2}$$

where $\mathbf{A_i}$ is a vector of simulated constraint values (row of $\mathbf{A}$), $\mathbf{x}_j$ is the j^{th} decision variable, and $\delta\mathbf{x}_j$ is a small perturbation of the j^{th} decision variable. This formulation is similar to the finite-difference approximation used to fill the Jacobian matrix ($\mathbf{J}$) of sensitivities in many PE algorithms (e.g. Doherty, 2015; Welter et al., 2015).

Constraints in the LP process can be formed directly from