Journal of Hydrology 544 (2017) 456–466

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

Calibration of an agricultural-hydrological model (RZWQM2) using surrogate global optimization

Maolong Xi [a], Dan Lu [b,*], Dongwei Gui [c,*], Zhiming Qi [d], Guannan Zhang [e]

[a] Department of Control Technology, Wuxi Institute of Technology, Wuxi, Jiangsu 214121, China
[b] Computer Science and Mathematics Division, Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
[c] Cele National Station of Observation & Research for Desert Grassland Ecosystem, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China
[d] Department of Bioresource Engineering, McGill University, Sainte-Anne-de-Bellevue, QC H9X3V9, Canada
[e] Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

ARTICLE INFO

Article history:
Received 8 September 2016
Received in revised form 22 November 2016
Accepted 23 November 2016
Available online 27 November 2016
This manuscript was handled by Corrado Corradini, Editor-in-Chief, with the assistance of Renato Morbidelli, Associate Editor

Keywords:
RZWQM2
Sparse grid
Surrogate model
QPSO algorithm
Global optimization

ABSTRACT

Robust calibration of an agricultural-hydrological model is critical for simulating crop yield and water quality and making reasonable agricultural management. However, calibration of the agricultural-hydrological system models is challenging because of model complexity, the existence of strong parameter correlation, and significant computational requirements. Therefore, only a limited number of simulations can be allowed in any attempt to find a near-optimal solution within an affordable time, which greatly restricts the successful application of the model. The goal of this study is to locate the optimal solution of the Root Zone Water Quality Model (RZWQM2) given a limited simulation time, so as to improve the model simulation and help make rational and effective agricultural-hydrological decisions. To this end, we propose a computationally efficient global optimization procedure using sparse-grid based surrogates. We first used advanced sparse grid (SG) interpolation to construct a surrogate system of the actual RZWQM2, and then we calibrate the surrogate model using the global optimization algorithm, Quantum-behaved Particle Swarm Optimization (QPSO). As the surrogate model is a polynomial with fast evaluation, it can be efficiently evaluated with a sufficiently large number of times during the optimization, which facilitates the global search. We calibrate seven model parameters against five years of yield, drain flow, and NO_3-N loss data from a subsurface-drained corn-soybean field in Iowa. Results indicate that an accurate surrogate model can be created for the RZWQM2 with a relatively small number of SG points (i.e., RZWQM2 runs). Compared to the conventional QPSO algorithm, our surrogate-based optimization method can achieve a smaller objective function value and better calibration performance using a fewer number of expensive RZWQM2 executions, which greatly improves computational efficiency.

Published by Elsevier B.V.

1. Introduction

Agricultural-hydrological models are important tools in quantifying and improving our understanding of hydrologic processes and evaluating the influence of agronomic management practices on the water cycle, water quality, and agricultural production. Several agricultural-hydrological models have been developed and each model has its own features. For example, the SWAT model evaluates effects of alternative management decisions on water resources and nonpoint-source pollution in large river basins (Arnold et al., 2012; Wang et al., 2015a); the HYDRUS model simulates water, heat, and solute movement in variably saturated media (Šimůnek et al., 2008); and a newly developed model in Alam and Dutta (2012) focuses on nutrient dynamics in river basin. In this study, we considered the Root Zone Water Quality Model (RZWQM2), which simulates plant growth and water, nutrient, and pesticide movement based on soil physical characteristics and hydraulic properties. The RZWQM2 was developed by the U. S. Department of Agriculture, and it has been widely used to investigate effects of different agricultural management practices on crop yield and water quality. The model performance has been evaluated in several U.S. states (e.g., Colorado, Georgia, Iowa, Minnesota, Missouri, Montana, Nebraska and Ohio—Ghidey et al., 1999; Wu et al., 1999; Landa et al., 1999; Abrahamson et al., 2006; Saseendran et al., 2008; Thorp et al., 2008; Qi et al., 2011, 2013; Ma et al., 2012), and in other countries (e.g., Canada, China and Portugal—Ahmed et al., 2007; Cameira et al., 2005; Fang

* Corresponding authors.
E-mail addresses: lud1@ornl.gov (D. Lu), guidwei@163.com (D. Gui).

http://dx.doi.org/10.1016/j.jhydrol.2016.11.051
0022-1694/Published by Elsevier B.V.

et al., 2014; Sun et al., 2016). Recently, RZWQM2 has been used to quantify climate change impacts on the hydrologic cycle and agro-environments such as cultivated fields (Ko et al., 2011; Islam et al., 2012; Wang et al., 2015b). In most cases, the model performed satisfactorily, but its performance is sensitive to particular and sometimes poorly identified parameter values. Thus, calibration of RZWQM2 is crucial to its successful application.

Traditionally, RZWQM2 has been calibrated by trial-and-error methods. While this straightforward and simple method continues to be widely used (Qi et al., 2011; Ma et al., 2012; Malonea et al., 2014; Shresthaa and Dattab, 2015), it is time-consuming and the termination process, heavily dependent on modelers' experience and expertise, suffers from subjectivity. Faced with the weakness inherent to manual methods, gradient-based automatic calibration tools, such as PEST (Doherty, 2004), have recently been employed to calibrate RZWQM2 (Malone et al., 2010; Fang et al., 2012). PEST adjusts model parameters to reduce the objective function value until the fit between model simulation results and measurement data is optimized. However, as a local optimizer, the use of PEST may result in different optimal parameter estimates when starting with different initial values, i.e., calibration results rely on the procedure initialization. In fact, for most gradient-based optimization methods, becoming trapped in a local objective function minimum is a common problem. To address this issue, Xi et al. (2015) applied a global search method, termed quantum-behaved particle swarm optimization (QPSO), to calibrate the RZWQM2 and obtained promising results.

Proposed by Sun et al. (2004), QPSO is a population-based swarm intelligence algorithm theoretically guaranteed to find optimal solutions in search space. Given its strong global convergence, the QPSO algorithm has been applied to many studies in recent years (Davoodi et al., 2014; Xi et al., 2015; Hassani and Lee, 2016). A review on its application can be found in Fang et al. (2010). Many numerical results indicated that with sufficiently large iterations (typically thousands or even hundred thousand), the QPSO can successfully identify globally optimal parameters (Sun et al., 2004; Omkar et al., 2009). However, in practice, a large number of iterations are usually unaffordable, especially for complex system models such as the RZWQM2 where each model simulation typically takes several minutes or even hours. Consequently, application of the global QPSO algorithm may end up with providing a local optimum given the limited search resulting from the intensive expense of model runs.

In this study, the problem of the high computational cost of global optimization is resolved by combining surrogate modeling with the optimization operation to develop a surrogate-based optimization algorithm. The surrogate modeling involves constructing a cheap-to-evaluate surrogate model, which provides an accurate and efficient approximation to the input-output relation of the actual simulation model. Several surrogate methods have been developed including sparse-grid stochastic collocation that uses the sparse-grid interpolation (Nobile et al., 2008a, 2008b) and probabilistic collocation that uses the finite dimensional polynomial chaos expansion (Marzouk et al., 2007; Li and Zhang, 2007; Shi et al., 2009). Chang and Zhang (2009) provided a comprehensive comparison of the accuracy and efficiency of these two methods.

The idea of sparse-grid (SG) methods is to place a grid in the parameter space with sparse parameter samples (as opposed to a full tensor-product grid). Then the simulation model is solved only at these sparse parameter samples to construct an interpolant and this interpolant is the SG-based surrogate of the actual simulation model. The SG methods have been demonstrated to be efficient and effective in dealing with uncertainty quantification problems. For example, Shi and Yang (2009) and Lin and Tartakovsky (2009, 2010) used SG methods to estimate mean and covariance of groundwater state variables such as hydraulic head and solute concentrations. Ma and Zabaras (2009) and Zeng et al. (2012) used SG methods to build surrogate of geophysical models that were then employed to evaluate parameter distributions. Zhang et al. (2013) developed an adaptive SG method to accelerate Bayesian inference in groundwater reactive transport modeling. Similarly, Zeng et al. (2016) recently evaluated two SG surrogates for groundwater Bayesian uncertainty quantification. However, the SG based surrogate approach has rarely been employed in the context of parameter estimation and never in agricultural-hydrological models.

Considering the importance of the parameter optimization in the RZWQM2 and the challenging in the calibration of the model, this work explores a computationally efficient global optimization scheme by combining SG-based surrogate methods with the QPSO algorithm and applies it to the RZWQM2. The developed method takes advantage of the global convergence of the QPSO while overcoming its drawbacks in poor computational efficiency by using surrogate modeling. The key idea is to use SG interpolation to construct a surrogate of the RZWQM2 in a polynomial form, and then evaluate the cheap-to-run surrogate model in the global optimization process so as to improve overall computational efficiency. The main contribution of our work lies in exploring the combination of a state-of-the-art SG surrogate method with a global optimization technique, and successfully applying this efficient method to find better solutions of the RZWQM2 than previous study (Qi et al., 2011; Xi et al., 2015), which greatly improves model calibration performance and helps make rational and effective agricultural-hydrological decisions.

The paper is organized as follows. The RZWQM2 and the surrogate optimization method are described in Section 2. In Section 3, we apply the approach to calibrating the RZWQM2, and we examine the effectiveness and efficiency of the method in comparison with the conventional QPSO algorithm. In Section 4, we further discuss the optimization results and show some insight on future studies. Lastly, we conclude this paper in Section 5.

2. Materials and methods

2.1. Description of the RZWQM2 and calibration parameters

The Root Zone Water Quality Model (RZWQM2) is a one-dimensional system model that simulates major physical, chemical, and biological processes in an agricultural crop production system. It consists of components for hydrology, nutrition and pesticide transport and transformation, plant growth and crop production, and management activities (Ahuja et al., 2000; Ma et al., 2005, 2006). Each component is simulated by sub-models. For example, infiltration from rainfall, irrigation, or snowmelt is calculated using a modified Green-Ampt model. Water redistribution in the soil profile is simulated via the Richards equation by treating surface evaporation and plant root water uptake as sinks. Water stored in the soil profile, if exceeding the field capacity, is drained to build up a water table above the impermeable layer and the subsequent tile drainage flux is computed using the steady state Hooghoudt equation. In addition, lateral flow and seepage are quantified by user-defined parameters for constant bottom layer water flux rate and hydraulic gradient, respectively. For the upper boundary, soil evaporation and plant transpiration are estimated by the double layer model of Shuttleworth and Wallace (1985), which is an extension of the Penman-Monteith concept. Nutrient chemistry processes are simulated by OMNI (Shaffer et al., 2000), a state-of-the-art model for carbon and nitrogen cycling in soils. The DSSAT cropping system models (Jones et al., 2003) are incorporated to predict crop establishment and water and nutrient uptake.