Journal of Hydrology 561 (2018) 98–111

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

Adaptive surrogate model based multiobjective optimization for coastal aquifer management

Jian Song[a], Yun Yang[b], Jianfeng Wu[a,*], Jichun Wu[a], Xiaomin Sun[c], Jin Lin[c]

[a] Key Laboratory of Surficial Geochemistry, Ministry of Education, Department of Hydrosciences, School of Earth Sciences and Engineering, Nanjing University, Nanjing 210023, China
[b] School of Earth Sciences and Engineering, Hohai University, Nanjing 210098, China
[c] Nanjing Hydraulic Research Institute, Nanjing 210029, China

A R T I C L E I N F O

This manuscript was handled by G. Syme, Editor-in-Chief, with the assistance of Craig T. Simmons, Associate Editor

Keywords:
Seawater intrusion
Surrogate model
Multiobjective optimization
Memetic algorithm
Pumping optimization

A B S T R A C T

In this study, a novel surrogate model assisted multiobjective memetic algorithm (SMOMA) is developed for optimal pumping strategies of large-scale coastal groundwater problems. The proposed SMOMA integrates an efficient data-driven surrogate model with an improved non-dominated sorted genetic algorithm-II (NSGAII) that employs a local search operator to accelerate its convergence in optimization. The surrogate model based on Kernel Extreme Learning Machine (KELM) is developed and evaluated as an approximate simulator to generate the patterns of regional groundwater flow and salinity levels in coastal aquifers for reducing huge computational burden. The KELM model is adaptively trained during evolutionary search to satisfy desired fidelity level of surrogate so that it inhibits error accumulation of forecasting and results in correctly converging to true Pareto-optimal front. The proposed methodology is then applied to a large-scale coastal aquifer management in Baldwin County, Alabama. Objectives of minimizing the saltwater mass increase and maximizing the total pumping rate in the coastal aquifers are considered. The optimal solutions achieved by the proposed adaptive surrogate model are compared against those solutions obtained from one-shot surrogate model and original simulation model. The adaptive surrogate model does not only improve the prediction accuracy of Pareto-optimal solutions compared with those by the one-shot surrogate model, but also maintains the equivalent quality of Pareto-optimal solutions compared with those by NSGAII coupled with original simulation model, while retaining the advantage of surrogate models in reducing computational burden up to 94% of time-saving. This study shows that the proposed methodology is a computationally efficient and promising tool for multiobjective optimizations of coastal aquifer managements.

1. Introduction

The contradiction between intensive water demands and groundwater environment restrictions exists in planning and development of coastal regions. Groundwater in the coastal aquifers is susceptible to deterioration due to its proximity to saltwater in combination with pumping activity and climate change. Typically, irrigation behavior of groundwater resource utilization directly results in seawater intrusion (SI) and further reduces the amount of available freshwater resources, given that exceeding 1% of seawater (250 mg/L chloride concentration) indicates freshwater is unsuitable for drinking (Werner et al., 2013). Thus, pumping optimization of coastal aquifer is essential to alleviate SI caused by continuous and unplanned pumping activity while satisfying water demands and environmental constraints (Christelis and Mantoglou, 2016). The linked simulation-optimization (S/O) methodologies (Mantoglou et al., 2004; Mantoglou and Papantoniou, 2008; Qahman et al., 2005; Sedki and Ouazar, 2011; Abd-Elhamid and Javadi, 2011; Javadi et al., 2015) have been popularly used in the determination of optimal pumping strategies of coastal groundwater management. The S/O methodology has the advantages of applicability to non-linear groundwater system and universal optimum solution for policy-making. However, with the S/O methodology, thousands of evaluations of the simulation model need to be implemented before the optimal set is obtained. Thus, the precision and accuracy of model predictions and computational efficiency of optimization algorithm always determine the practicality and performance of linked S/O process (Bhattacharjya and Datta, 2005).

SI involves complex physical and chemical processes including

* Corresponding author at: Key Laboratory of Surficial Geochemistry, Ministry of Education, Department of Hydrosciences, School of Earth Sciences and Engineering, Nanjing University, Nanjing 210023, China.
E-mail addresses: jfwu@nju.edu.cn, jfwu.nju@gmail.com (J. Wu).

https://doi.org/10.1016/j.jhydrol.2018.03.063
Received 5 December 2017; Received in revised form 12 March 2018; Accepted 25 March 2018
Available online 27 March 2018

dispersive mixing, surface hydrology, and density effects caused unstable convection, anthropogenic influences and geological characteristics leading to multifaceted challenges for solving coastal groundwater management problems (Werner et al., 2013). Fortunately, many modular numerical codes, such as the commonly used finite difference program SEAWAT (Guo and Langevin, 2002), the finite element program FEMWATER (Lin et al., 1997), and the hybridization of finite-element and integrated-finite-difference program SUTRA (Voss and Provost, 2010) can be extensively used to solve the SI models for the coupled density-dependent flow and solute transport problems. However, for a regional coastal aquifer, the SI model is numerically complicated and computationally demanding. Beyond that, decision-makers always consider several conflicting objectives of optimization for coastal groundwater management (*e.g.*, maximization of total pumping from production well and minimization of the extent of SI), which generate a particularly intricate multiobjective optimization problem. Due to the complexity of coastal aquifer management, the S/O approaches by use of surrogate models assisted multiobjective evolutionary algorithms (SMOEAs) have been receiving increasing attention in recent SI optimization problems. Surrogate models, also known as *meta*-models, proxy models, lower fidelity models and response surfaces, are computationally efficient emulators designed to mimic key characteristics of simulation model and have great potential to overcome the computational limitations caused by the optimization management of complex models (Asher et al., 2015).

SMOEAs utilize surrogate models to partially substitute CPU-intensive simulation models to save computational cost and solve computationally intensive optimization problems (Sun et al., 2017). Surrogate models combined with coastal groundwater management in the linked S/O frameworks include a great diversity of available approaches such as radial basis functions (Christelis and Mantoglou, 2016), modular neural networks (Kourakos and Mantoglou, 2009; 2013), genetic programming (Sreekanth and Datta, 2010; 2011), multivariate adaptive regression splines (Roy and Datta, 2017) and evolutionary polynomial regressions (Hussain et al., 2015). Actually, these surrogate models can be divided into two types, *i.e.*, the on-line and off-line frameworks. At present, there are relatively limited on-line frameworks implemented in evolutionary algorithm (Sreekanth and Datta, 2015; Ketabchi and Ataie-Ashtiani, 2015) and major off-line frameworks presented in coastal aquifer managements (e.g., Hussain et al., 2015; Ataie-Ashtiani et al., 2014; Sreekanth and Datta, 2011; 2014; Roy and Datta, 2017; Nikolos et al., 2008; Bhattacharjya and Datta, 2009; Dhar and Datta, 2009). In general, the off-line framework involves three steps: (1) design of experiments which employ different space filling strategy to obtain key response characteristics of simulation model over design variables feasible space; (2) training and validation of surrogate model; (3) substitution of original simulation model throughout the optimization process. A potential disadvantage of off-line framework is that one-shot training surrogate model probably cannot accurately represent potential functional relationship between the original model and the interesting and important regions of design variable space, or the region of near Pareto optimal solutions (Razavi et al., 2012). For example, Sreekanth and Datta (2010) proposed genetic programming coupled with multiobjective genetic algorithm, which modified search space based on the relative importance of decision variables and further adaptively retrained surrogate model for achieving more accurate predictions in the vicinity of Pareto-optimal solutions, so as to prevent evolutionary search from falling into local optimum solutions (Jin, 2011). Kourakos and Mantoglou (2013) employed three criteria based on convergence and diversity of solutions for selecting the most promising offspring to implement adaptive training of modular neural network models and multiobjective optimization management in the central aquifer of Santorini Island. The results suggested that the Pareto solutions achieved by surrogate model were better than those solutions by NSGAII coupled with simulation model. However, the training process used for surrogate model was time-consuming even though Kourakos and Mantoglou (2013) reduced the number of decision variables to simplify the complexity of management model.

The construction of surrogate model combined with coastal groundwater management is a difficult and computationally demanding task especially in cases with high dimensional decision variables and multi outputs. To the best of our knowledge, there is no research about coastal groundwater optimization management that employs extreme learning machine (ELM) as surrogate model which has been proposed for training a single-hidden layer feedforward neural network (Huang et al., 2006). The essential advantages of ELM are high learning accuracy, very fast learning speed and least user intervention, which can explicitly tackle aforementioned challenges (Huang et al., 2015). The number of hidden layer nodes is a crucial parameter selected by time-consuming methods for the prediction performance of ELM. Kernel extreme learning machine (KELM) is developed, which replaces the hidden layer mapping by the kernel function mapping to avoid the hidden layer nodes selection problem (Wang and Han, 2014). KELM is more accurate and effective than ELM and has better generalization performance for regression and binary class classification applications (Chen et al., 2014).

The purpose of this study is to present a linked S/O framework that integrates SEAWAT with KELM assisted multiobjective memetic algorithm for pumping optimization of coastal groundwater management subject to environment constraints and water demands. The proposed algorithm is then applied to a real-world case study in the coastal aquifer underlying southern Baldwin County. Owning to increasing populations, economic development and tourism, groundwater pumping drastically increased and land use varies considerably in the study region (Murgulet and Tick, 2008; Lin et al., 2009). The multiobjective optimization management of coastal groundwater attempts to find out the optimal pumping design strategy between socioeconomic and environmental factors for simultaneously maximizing the total pumping rates and minimizing the extent of SI. The proposed S/O methodology is generally applicable to any highly parameterized SI models at the real-world field and large-scale coastal groundwater management problems.

This paper is organized in five sections. Following this introduction, Section 2 describes the surrogate model assisted multiobjective optimization framework for coastal aquifer management. Then Section 3 presents the application of the methodology to the real-world coastal groundwater management case study in the Baldwin County, Alabama. Next Section 4 illustrates the analysis of optimization results and discussions. Finally, Section 5 summarizes and concludes this paper.

2. Methodology

Surrogate model assisted multiobjective optimization framework developed in this study for pumping optimization of coastal aquifer comprises two components. The first one is the development of KELM surrogate model for substituting the physically based simulators of SEAWAT. The second component is an optimization model using surrogate model assisted multiobjective memetic algorithm integrated with kernel extreme learning machine (SMOMA-KELM) to optimize the pumping strategies satisfying the imposed environment and physical constraints.

2.1. Numerical model of variable density groundwater flow and solute transport

The MODFLOW (Harbaugh et al., 2000) and MT3DMS (Zheng and Wang, 1999) based program, SEAWAT, was designed to simulate density-dependent groundwater flow and multi-species solute transport (Guo and Langevin, 2002). The general mathematical equation describing variable-density groundwater flow and transport process can be stated as: