Journal of Hydrology 393 (2010) 245–256

Contents lists available at ScienceDirect

Journal of Hydrology

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

Multi-objective management of saltwater intrusion in coastal aquifers using genetic programming and modular neural network based surrogate models

J. Sreekanth *, Bithin Datta

Discipline of Civil and Environmental Engineering, School of Engineering and Physical Sciences, James Cook University, Townsville, QLD 4811, Australia
Cooperative Research Centre for Contamination Assessment and Remediation of the Environment (CRC-CARE), Mawson Lakes, SA 5095, Australia

A R T I C L E I N F O

Article history:
Received 29 December 2009
Received in revised form 3 July 2010
Accepted 25 August 2010

This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Luis E. Samaniego, Associate Editor

Keywords:
Salinity intrusion
Coastal aquifer
Pumping optimization
Surrogate model
Genetic programming
Modular neural network

S U M M A R Y

Surrogate model based methodologies are developed for evolving multi-objective management strategies for saltwater intrusion in coastal aquifers. Two different surrogate models based on genetic programming (GP) and modular neural network (MNN) are developed and linked to a multi-objective genetic algorithm (MOGA) to derive the optimal pumping strategies for coastal aquifer management, considering two objectives. Trained and tested surrogate models are used to predict the salinity concentrations at different locations resulting due to groundwater extraction. A two-stage training strategy is implemented for training the surrogate models. Surrogate models are initially trained with input patterns selected uniformly from the entire search space and optimal management strategies based on the model predictions are derived from the management model. A search space adaptation and model retraining is performed by identifying a modified search space near the initial optimal solutions based on the relative importance of the variables in salinity prediction. Retraining of the surrogate models is performed using input–output samples generated in the modified search space. Performance of the methodologies using GP and MNN based surrogate models are compared for an illustrative study area. The capability of GP to identify the impact of input variables and the resulting parsimony of the input variables helps in developing efficient surrogate models. The developed GP models have lesser uncertainty compared to MNN models as the number of parameters used in GP is much lesser than that in MNN models. Also GP based model was found to be better suited for optimization using adaptive search space.

1. Introduction

Optimization of pumping from coastal aquifers is a challenging groundwater management problem as excessive extraction of water from aquifers hydraulically connected to the sea often results in salinity intrusion. Salinity intrusion in coastal aquifers is a highly non-linear and complex process (Bear et al., 1999). Once salinity intrusion occurs, it involves long-term measures incurring huge costs to remediate these contaminated aquifers. Hence, carefully planned strategies of groundwater extraction are required to prevent the eventual contamination of the valuable resource.

Salinity intrusion management models are used to prescribe management strategies for the sustainable use of coastal aquifers by controlling salt water intrusion. Developing an optimal management model involves integrating a groundwater flow and transport simulation model within an optimization framework. Flow and transport equations for salinity intrusion are coupled together by the density variation occurring during the mixing process, requiring simultaneous solution of both the equations. The numerical model for the density dependent flow and transport simulation would be computationally expensive, especially when used in a simulation–optimization framework. Trained and tested surrogate models are capable of approximating the numerical simulation model for simulating flow and transport process in the aquifer. Such a surrogate model when linked to an optimal decision model can evolve multi-objective optimal management strategies for the aquifer with the least computational burden. The use of genetic programming (GP) and modular neural networks (MNN) as the surrogate models is presented in this study. Trained and tested GP and MNN models are linked with a multi-objective genetic algorithm to derive optimal management strategies for a coastal aquifer.

Simulation–optimization models have been extensively used in solving groundwater pumping management problems (Gorelick, 1983; Gorelick et al., 1984; Ahlfeld and Heidari,1994; Hallaji and Yazicigil, 1996; Emch and Yeh, 1998; Wang and Zheng, 1998; Das and Datta, 1999a,b; Cheng et al., 2000; Mantoglou, 2003; Mantoglou et al., 2004; Katsifarakis and Petala, 2006; Ayvaz and

* Corresponding author at: Discipline of Civil and Environmental Engineering, School of Engineering and Physical Sciences, James Cook University, Townsville, QLD 4811, Australia. Tel.: +61 0747815380; fax: +61 0747816788.

E-mail addresses: sreekanth.janardhanan@jcu.edu.au (J. Sreekanth), bithin.datta@jcu.edu.au (B. Datta).

0022-1694/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2010.08.023

Karahan, 2008). Sharp-interface models and variable density models are the two modelling approaches used for simulating salinity intrusion in coastal aquifers. A number of studies have used sharp interface salinity intrusion models in the simulation–optimization framework (Iribar et al., 1997; Dagan and Zeitoun, 1998; Mantoglou, 2003; Park and Aral, 2004; Mantoglou and Papantoniou, 2008). Sharp-interface models are relatively simple and are easier to be incorporated within optimization models. Using a 3D density dependent model within an optimization framework is constrained by the CPU time taken by the model. Different techniques like response matrix method (Gorelick, 1983), embedding technique (Das and Datta, 1999a,b), and externally linking the flow and transport simulation model to the optimization model (Dhar and Datta, 2009) have been used in the past studies. In an externally linked simulation–optimization framework the optimization model calls the simulation model each time a candidate solution is evaluated. Thus the simulation model is run thousands of times before the optimal solution are obtained, adding to the computational complexity of the management model. Dhar and Datta (2009) reported a 30-day run time, on a 2.4 GHZ Oeptron AMD machine with 4 GB RAM, for a linked simulation–optimization model applied to a small aquifer system to find optimal solution.

Surrogate models are used to approximate the numerical simulation model in order to reduce the computational burden imposed by large scale numerical simulation models, especially within a linked simulation–optimization framework. Different methods like Artificial Neural Networks, radial-basis-function network, support vector machine etc. are used for constructing surrogate models. Extensive discussion of surrogate models can be seen in Jin (2005).

Artificial Neural Networks (ANN) have been widely used as surrogates for groundwater models (Ranjithan et al., 1993; Rogers et al., 1995; Aly and Peralta, 1999). Substantial research work has been done on using Artificial Neural Networks as surrogate models for simulation–optimization studies. Bhattacharjya (2003), Rao et al. (2004), Bhattacharjya and Datta (2005, 2009), Kourakos and Mantoglou (2006) and Dhar and Datta (2009) have used Neural Network surrogate models for developing salinity intrusion management models. Arndt et al. (2005) developed a neural network surrogate model implementing search interval adaptation. The adaptive neural network model was used as a surrogate for a finite element groundwater model and was used with an optimization algorithm to solve an optimal design problem. Yan and Minsker (2006) developed an Adaptive Neural Network Genetic Algorithm (ANGA) where the network was trained with search interval adaptation and genetic algorithm used to solve the optimization model. Behzadian et al. (2009) used adaptive neural networks in combination with multi-objective genetic algorithm NSGA-II to locate pressure loggers for a stochastic sampling design. Kourakos and Mantoglou (2009) developed a modular neural network (MNN) with a number of sub-networks replacing a global ANN. Salinity concentration in each monitoring well was predicted using a modular neural network and the intrusion is controlled by relatively few pumping wells falling within certain control distance from the monitoring wells. The networks were trained adaptively as optimization progresses. The computational time could be reduced considerably by using the modular neural networks.

A few studies in the broad area of hydrology and water resources have used GP models (Dorado et al., 2002; Makkeasorn et al., 2008; Parasuraman and Elshorbagy, 2008; Wang et al., 2009). GP has been used to develop prediction models run-off, river stage and real-time wave forecasting. (Babovic and keijzer, 2002; Sheta and Mahmoud, 2001; Gaur and Deo, 2008). Zechman et al. (2005) developed a GP based surrogate model for use in a groundwater pollutant source identification problem. The chemical signals at the observation wells were used to reconstruct the pollution loading scenario. The inverse problem was solved using a simulation–optimization approach using GA to conduct the search. The numerical model was replaced by a surrogate model developed using genetic programming to reduce the computational burden.

The present study uses two surrogate models, GP and MNN, linked with multi-objective genetic algorithm to solve the pumping optimization problem. Genetic programming (GP) models are developed as surrogates for the variable density flow and transport simulation model, FEMWATER, which is used to simulate the salinity concentration at each monitoring well location. The GP models are then coupled with a Multi-objective Genetic Algorithm, Non-dominated Sorted Genetic Algorithm- II (NSGA-II) (Deb, 2001) to derive optimal pumping strategies. Modular neural network models were also developed for predicting the salinity concentrations at these locations and linked with NSGA II to solve the same problem. Both GP and MNN models are trained with search space adaptation in two stages to increase the accuracy of prediction in a search space near the entire Pareto-optimal set of solutions.

The governing equations for the simulation of variable density flow and transport are described in Section 2. The framework of the management model and the optimization formulation are presented in Section 3. The GP-MOGA and MNN-MOGA models are presented in Sections 3.3 and 3.4 respectively. The methodology of search space adaptation and retraining the surrogate models for a multi-objective problem framework is described in Section 4. Section 5 presents the application of the developed methodologies to a small coastal aquifer and the relevant results and discussions. The conclusions are presented in Section 6.

2. Density dependent flow and transport simulation model

The three dimensional density dependent flow and transport simulation model FEMWATER (Lin et al., 1997) was chosen to simulate the coupled flow and transport process in the coastal aquifer system. The relevant equations for the density dependent flow and transport are as follows (Lin et al., 1997).

2.1. Flow equation

$$\frac{\rho}{\rho_o}F\frac{\partial h}{\partial t} = \nabla.\left[\mathbf{K.}\left(\nabla h + \frac{\rho}{\rho_o}\nabla z\right)\right] + \frac{\rho^*}{\rho_o}q \tag{1}$$

$$F = \alpha'\frac{\theta}{n} + \beta'\theta + n\frac{dS}{dh} \tag{2}$$

where F is the storage coefficient, h the pressure head, t the time, $\mathbf{K}$ the hydraulic conductivity tensor, z the potential head, q the source and/or sink, ρ the water density at the chemical concentration C, ρ_o the referenced water density at zero chemical concentration, ρ^* the density of either the injection fluid or the withdrawn water, θ the moisture content, α' the modified compressibility of water, n the porosity of the medium, S is the Saturation.

The hydraulic conductivity $\mathbf{K}$ is given by

$$\mathbf{K} = \frac{\rho g}{\mu}\mathbf{k} = \left(\frac{\frac{\rho}{\rho_o}}{\frac{\mu}{\mu o}}\right)\frac{\rho_o g}{\mu_o}\mathbf{k_s}k_r = \left(\frac{\frac{\rho}{\rho_o}}{\frac{\mu}{\mu_o}}\right)\mathbf{K_{so}}k_r \tag{3}$$

where μ is the dynamic viscosity of water at chemical concentration C, μ_o the referenced dynamic viscosity of water at zero chemical concentration, $\mathbf{k}$ the permeability tensor $\mathbf{k_s}$ the relative permeability or relative hydraulic conductivity, $\mathbf{K_{so}}$ is the referenced saturated hydraulic conductivity tensor.

The density dependence on concentration is given by,

$$\frac{\rho}{\rho_o} = a_1 + a_2C \tag{4}$$