

Optimal groundwater management using state-space surrogate models: a case study for an arid coastal region

Tirthankar Roy, Niels Schütze, Jens Grundmann, Marco Brettschneider and Ashu Jain

ABSTRACT

A surrogate modeling framework is developed in this study to circumvent the computational burden of high-fidelity numerical groundwater models for arid coastal aquifers. Two different surrogate models, namely, artificial neural network (ANN) and Gaussian process model (GPM) are trained to replace the computationally expensive numerical flow and transport model OpenGeoSys. A novel time-dependent training scheme is introduced which helps the surrogates in tracking the discrete-time state-space trajectories of the high-fidelity model, thereby making them suitable for variable-time simulations. The surrogates are also tested in the extrapolation range corresponding to some extreme boundary conditions such as a very high rate of extraction. Both the surrogates show comparable accuracy in efficiently approximating the numerical model response; however, ANN is found to be much faster than GPM for the size of the data used. The trained surrogates are then used in developing a long-term planning and management framework for analyzing feasible management scenarios in the coastal aquifer of Oman.

Key words | artificial neural network, coastal aquifers, evolution strategies, Gaussian process model, optimal groundwater management, surrogate modeling

Tirthankar Roy
Niels Schütze
Jens Grundmann
Marco Brettschneider
Institute of Hydrology and Meteorology,
Technische Universität Dresden,
Bergstr. 66, 01069 Dresden,
Germany

Tirthankar Roy
Ashu Jain
Department of Civil Engineering,
Indian Institute of Technology Kanpur,
Kanpur 208016,
India

Tirthankar Roy (corresponding author)
Department of Hydrology and Water Resources,
The University of Arizona,
Tucson AZ 85721,
USA
E-mail: roy@email.arizona.edu

Marco Brettschneider
AXA Global P&C,
9 Avenue de Messine,
75018 Paris,
France

INTRODUCTION

Increasing water demand in different consumer sectors has contributed to higher water stresses in the aquifers in various parts of the world (WBCSD 2006). The consequences of this alteration vary from place to place depending on water consumption, aquifer characteristics, and climate conditions. Aquifers in coastal arid or semi-arid regions are often vulnerable to water table drawdown and seawater intrusion as a result of increased water pumping. As water is extracted from an aquifer, the water table drops at the pumping location, triggering drawdown. A sustainable development entails extraction to be equal to the 'capture' which is a balance between the change in virgin rate of recharge and/or the change in virgin rate of discharge (Bredenhoef 2002). Seawater intrudes into the aquifer subsequent to the negative pressure gradients created due to high extraction, and contaminates the fresh

water regions. This curtails drinking water availability and also jeopardizes agricultural production. Robust management strategies are therefore indispensable in this kind of situation for ensuring both profit from agriculture as well as aquifer sustainability.

In recent practice, management frameworks often couple a numerical groundwater model with an optimization algorithm to reach the optimal solutions of the control variables (Das & Datta 2001; Nicklow *et al.* 2010; Yeh 2015). This method is called simulation-optimization. The flow and transport processes can be simulated by various physically based high-fidelity (highly realistic) numerical models (e.g., Voss 1984; Lin *et al.* 1997; Guo & Langevin 2002; Liu *et al.* 2002; Hamidi *et al.* 2006; Delfs *et al.* 2009). For the optimization part, various nature-inspired metaheuristic algorithms, such as genetic algorithm

(Wang & Zheng 1998; Yan & Minsker 2006; Ataie-Ashtiani et al. 2014), evolution strategies (Bayer & Finkel 2004, 2007; Kollat & Reed 2006; Papadopoulou et al. 2010), particle-swarm-optimization (Mousavi & Shourian 2010; Gaur et al. 2011; Sedki & Ouazar 2011), ant-colony-optimization (Ataie-Ashtiani & Ketabchi 2011), differential evolution (de Paly et al. 2013), and honey-bee-mating-optimization (Haddad & Marino 2007, 2011) have gained vast popularity.

There are two main issues with a simulation-optimization framework that couples a high-fidelity numerical model with an optimization algorithm. First, the numerical models are often computationally expensive and second, the optimization algorithm requires the numerical model to run several thousands of times which makes the whole process overly time-consuming. In addition, the computational complexity can increase further if the considered optimal management problem has multiple criteria and/or a high number of decision variables (see Table 1). This computational burden can be reduced significantly by using different surrogate modeling techniques (Rao et al. 2004; Bhattacharjya & Datta 2009; Dhar & Datta 2009; Papadopoulou et al. 2010; Sreekanth

& Datta 2011, 2015; Ketabchi & Ataie-Ashtiani 2015). A surrogate model is an empirical or lower-fidelity alternative to either (A) the response of the high-fidelity numerical simulation model itself or (B) the objective function of the considered optimization problem or (C) the operating policy model in case of optimal control problems (Ratto et al. 2012). A surrogate model is simpler, considerably faster to run, and is generally derived through a training framework from the original simulation model (Razavi et al. 2012b).

A comprehensive review of the state-of-the-art of surrogate modeling of type (B) – the approximation of the objective function – can be found in Razavi et al. (2012b) where the authors analyzed 48 studies dealing with surrogate models directly from the water resources field and over 100 from a broader research community and summarized various concepts, methodological developments, and challenges involved in the field of surrogate modeling. They divided the surrogate model development techniques into four different categories: (1) basic ‘off-line’ sequential framework (e.g., Sreekanth & Datta 2011), (2) adaptive-recursive framework (e.g., Fen et al. 2009), (3)

Table 1 | Recent studies on groundwater management models for coastal groundwater systems

Reference	Model ^a	Problem ^b	Objective function ^c	#DV ^d	SM ^e	OA ^f	ST ^g	TD ^h	CE ⁱ
Bhattacharjya & Datta (2009)	3D DA-GW	GW (q)	$\max(Q_{out}), \min(C), \min(Q_{inj})$	33	MLP	NSGA-II	HC	TR3	$1,000 \times 2,500$
Bray & Yeh (2008)	FEMWATER	GW (q,x,y)	$\min(Q_{inj})$	215	–	GA	RC	TR5	Parallel
Dhar & Datta (2009)	FEMWATER	GW(q)	$\max(Q_{out}), \min(Q_{inj})$	33	MLP	NSGA-II	HC	TR3	$24 \times 1,800$
Haddad & Marino (2011)	3D DA-GW	GW(q)	$\max(C)$	26	–	HBMO	HC, RC	TI	6,000
Kourakos & Mantoglou (2010)	SEAWAT (2D)	GW(q)	$\min(C_{\delta k}), \min(C_{umw})$	4	–	NSGA-II	HC	TI	$? \times 400$
Papadopoulou et al. (2010)	PTC	GW(q)	$\max(Q_{out})$	5	RBF	DE	RC	TI	20×200
Rao et al. (2004)	SEAWAT	GW(q)	$\max(Q_{out}), \min(Q_{inj})$	27	MLP	SA	HC	TR3	10,000
Sreekanth & Datta (2011)	FEMWATER	GW(q)	$\max(Q_{out}), \min(Q_{inj})$	33	GP	NSGA-II	HC	TR3	200×750

^a3D DA-GW = 3D density driven groundwater model, FEMWATER, SEAWAT, PTC = see reference.

^bGW = groundwater management, q = abstraction (injection) rate, x = x-coordinate, y = y-coordinate of the well.

^c Q_{out} = pump volume, Q_{inj} = injection volume, C = costs.

^dNumber of decision variables.

^eSurrogate model: GP = genetic programming, MLP = multilayer perceptron, RBF = radial basis function network.

^fOptimization algorithm: GA = genetic algorithm, HBMO = honey-bee mating algorithm, NSGA-II = nondominated sorting genetic algorithm II, DE = differential evolution, SA = simulated annealing.

^gType of study: HC = hypothetical case, RC = real case.

^hTime dependence of decision variables: TI = time-invariant, TR <x> transient over x periods.

ⁱComputational effort: <population size > x <nr. of generations > .

metamodel-embedded evolution framework (e.g., Yan & Minsker 2011), and (4) approximation uncertainty based framework (e.g., Razavi et al. 2012a). In a basic sequential framework, the surrogates are globally fitted to the evaluated objective function values of some fixed design datasets. For the three other categories, the approximation framework is ‘on-line’; i.e., the design datasets and thus the surrogate model are updated during the course of the optimization. The approximation uncertainty based framework is the probabilistic version of the adaptive-recursive framework. Instead of considering the surrogate approximations as true deterministic values, this framework further incorporates the uncertainty associated with those approximations (Razavi et al. 2012a).

Table 1 shows some selected examples of simulation-optimization frameworks for the management and monitoring of coastal aquifers using surrogate models, and that are also reviewed by Ketabchi & Ataie-Ashtiani (2015) and Sreekanth & Datta (2015). All the reviewed studies employ surrogate modeling of type (B) – the approximation of the objective function – which thus require a fixed time horizon for training of the surrogates. Therefore, these surrogate models can be considered as time-independent surrogate models and an additional training is required if the given time horizon of the problem changes.

In this study, we introduce a new surrogate training framework which focuses on surrogate modeling of type (A); i.e., an approximation of the simulator, which makes the trained surrogate model time-dependent. These surrogates are able to track the discrete-time state-space trajectories of the original model and avoid an additional training when the simulation time is altered. This training scheme is ‘off-line’ in nature and exploits the information available in the variable space through a comprehensive set of feasible scenarios. We use two different surrogate models, namely, artificial neural network (ANN) and Gaussian process model (GPM). ANN is widely used in various hydrological applications (Hsu et al. 1995; Liriano & Day 2001; Jain et al. 2004; Shourian et al. 2008; Bhattacharjya & Datta 2009; Dhar & Datta 2009; Fu et al. 2010) whereas GPM is an emerging technique of non-parametric surrogate modeling (MacKay 1998; Williams 2002; Rasmussen & Williams 2006).

METHODS

The state-space surrogate modeling framework

Two different surrogate models are trained in this study to replace the computationally expensive numerical groundwater model OpenGeoSys (OGS) (Kolditz et al. 2012) in a simulation-optimization framework with a variable time horizon (see Figure 1). The inputs to the surrogates include water table heights and salinities at some specified locations in the problem domain. Since these inputs are a subset of the state variables of the numerical model, the surrogates can be seen as ‘state-space black-box’ models and formulated as:

$$\mathbf{y}(t+1) = f(\mathbf{y}(t), \mathbf{u}(t)) \text{ for } t = 0 \dots T-1 \quad (1)$$

where \mathbf{y} denotes the output variables; i.e., water table height \mathbf{h} and salinity \mathbf{S} at specific observation nodes, f is the numerical groundwater model, and \mathbf{u} represents the input variables; i.e., groundwater extraction by pumps and horizontal inflow into the aquifer (see Figure 1). The surrogate models are used to emulate the numerical model $f(\mathbf{y}(t), \mathbf{u}(t))$ by an approximation model $\tilde{f}(\mathbf{y}(t), \mathbf{u}(t), \theta)$ with parameters θ . Optimal values of the parameters are determined by some form of training or optimization which minimizes the error:

$$\theta^* = \arg \min_{\theta} f_{err} \left(f(\mathbf{y}(t), \mathbf{u}(t)), \tilde{f}(\mathbf{y}(t), \mathbf{u}(t), \theta) \right) \quad (2)$$

where f_{err} is a certain error function (see Figure 1).

Groundwater model OGS

OGS is a numerical simulation package for thermo-hydro-mechanical-chemical processes in porous and fractured media. It is based on the finite element method to solve the governing differential equations for different physical processes. Variable density flow simulation in porous media is highly dependent on the changes in solute concentration as well as temperature. OGS is used in this study to consider changes in density; however, the medium is

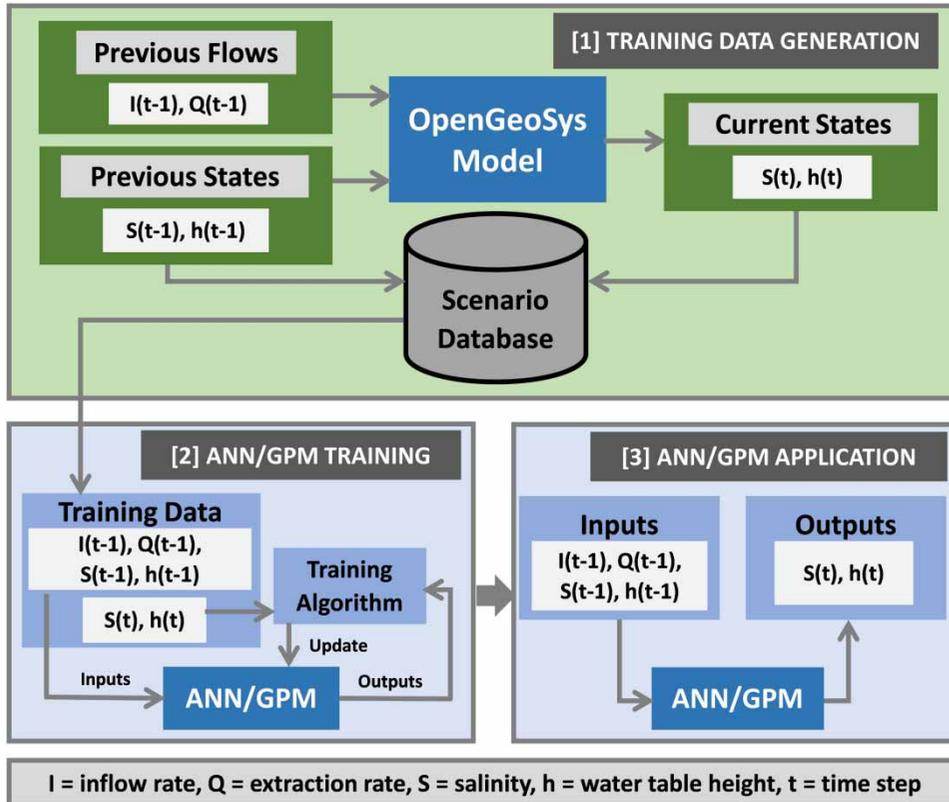


Figure 1 | Surrogate modeling framework for this study.

assumed to be isothermal. The three fundamental equations for variable density flow are (a) continuity equation, (b) momentum equation, and (c) contaminant transport equation, which are also linked to the bulk fluid density and hydrodynamic dispersion equations (Kolditz & Shao 2010).

Continuity equation

The flow equation for a variably saturated porous medium in terms of hydraulic head and mass concentration is given by:

$$\phi \frac{\partial S}{\partial t} + SS_0^h \frac{\partial h}{\partial t} + S\phi\lambda_C \frac{\partial C}{\partial t} + \nabla \cdot \vec{q} + \lambda_C \vec{q} \cdot \nabla C = Q_p \quad (3)$$

where ϕ is the porosity, S is the saturation ratio, t denotes time, S_0^h is the specific storativity, h is the hydraulic head, λ_C is the coefficient of expansivity, \vec{q} is the Darcy velocity vector, C is the relative concentration, and Q_p is the source term.

Momentum equation

The momentum equation for variable-density fluid flow in a porous medium is given by:

$$\vec{q} = -\frac{\hat{k}\rho_0\vec{g}}{\mu} \left(\nabla h + \left(\frac{\rho - \rho_0}{\rho_0} \right) \vec{e} \right) \quad (4)$$

where \hat{k} is the permeability tensor, ρ is the fluid density, ρ_0 is the reference fluid density, \vec{g} is the acceleration due to gravity, μ is the dynamic viscosity, and \vec{e} is the unit vector in the gravitational direction.

Contaminant transport equation

The advection-dispersion equation for solute transport is given by:

$$\frac{\partial(\phi C)}{\partial t} + \nabla \cdot (\phi \vec{v} C) - \nabla \cdot (\phi \hat{D} \cdot \nabla C) = Q_C \quad (5)$$

where \vec{v} is the fluid velocity vector and Q_C is the source term in terms of mass concentration. The dispersion tensor (\hat{D}) is given by:

$$\hat{D} = \gamma D_m \hat{\delta} + \alpha_T |v| \hat{\delta} + (\alpha_L - \alpha_T) \frac{\vec{v}_i \vec{v}_j}{|v|} \quad (6)$$

where γ is the tortuosity, D_m is the coefficient of molecular diffusion, $\hat{\delta}$ is the Kronecker-delta, α_T is the transverse dispersivity, and α_L is the longitudinal dispersivity.

The basic idea behind OGS dates back to the mid-1980s (Kolditz *et al.* 2012). It was formerly known as Rockflow/Geosys (Kolditz & Bauer 2004), and was first coupled with optimization by Bayer & Finkel (2007). Application of OGS can be found in some relatively new papers (e.g., Kolditz *et al.* 2008; Delfs *et al.* 2009; McDermott *et al.* 2009; Grundmann *et al.* 2012; Walther *et al.* 2012, 2014). The model was validated by Walther *et al.* (2012) against the Goswami–Clement benchmark problem (Goswami & Clement 2007) which is similar to the Henry's problem (Henry 1960) for salt water intrusion. The problem is posed as an initially freshwater environment with horizontal saltwater intrusion and withdrawal. The steady state density-dependent simulations of OGS were compared against the simulation results of the SEAWAT model (Langevin & Guo 2006) and the observations. It was found that OGS matched the observations more closely as compared to SEAWAT. The same experiment was extended to the transient state by Walther *et al.* (2014), and the ability of OGS to represent the transient density-dependent flow processes was also verified.

ANN

ANNs are biologically inspired mathematical models that imitate the basic structure and function of the human brain. Several interconnected artificial neurons form the neural network. A simple network consists of three different layers, namely, input layer, hidden layer, and output layer. Networks can be classified as feed-forward network and feed-back network based on the direction of the flow of information. Feed-forward network is a popular choice in the field of hydrologic

modeling where the flow of information is unidirectional. The connections between the neurons in a network are assigned with some weights which get updated during the training process based on the minimization of an error function. In this study, we apply the scaled conjugate gradient (SCG) algorithm (Møller 1993) for the training of ANNs.

GPM

In simple regression methods, a function is fitted between the input and the output variables. These methods are constrained by the choice of any particular fitting function (e.g., linear, higher order polynomial, sigmoidal, etc.). In an ideal regression framework, all possible function forms must be included, but there are infinite sets of feasible function forms. The concept of Gaussian process (GP) helps overcome this problem. A GP is defined as a distribution over functions, unlike a Gaussian distribution which is a distribution over vectors, or a Dirichlet process denoting distribution over distributions. GP can therefore be fully specified by a mean function and a covariance function, and once these two functions are specified, GP can be treated with the basic rules of probability applied to multivariate Gaussian distributions.

The joint Gaussian distribution for the training and testing datasets is given by:

$$\begin{bmatrix} f \\ f_* \end{bmatrix} \sim N\left(\begin{bmatrix} \mu \\ \mu_* \end{bmatrix}, \begin{bmatrix} K & K_* \\ K_*^T & K_{**} \end{bmatrix}\right) \quad (7)$$

where the training set input is denoted by X and the training set target by f . The testing set input and target are denoted by X_* and f_* , respectively. The conditional distribution of unknown test set outputs is given by:

$$f_* | f \sim N(\mu_* + K_*^T K^{-1} (f - \mu), K_{**} - K_*^T K^{-1} K_*) \quad (8)$$

where $(\mu_* + K_*^T K^{-1} (f - \mu))$ is the mean and $(K_{**} - K_*^T K^{-1} K_*)$ is the covariance, μ is the training input mean, μ_* is the testing input mean, K is the training set covariance matrix, K_* is the training-testing set covariance matrix, and K_{**} is the testing set covariance matrix.

The covariance function controls the collective variability in different function values for a given input to the functions. Learning/Training in GP implies calculating the optimal values of the covariance function parameters (also called *hyperparameters*) in light of the training data. Since GP is a non-parametric modeling approach, it requires the whole training dataset for prediction purpose. The main disadvantage of GP is its high computational time requirement for large datasets resulting from the inversion of the covariance matrix during training. To overcome this issue, various sparse pseudo-input GP techniques are proposed in which the actual input data are represented by a smaller subset (called *active set*) of the total input data selected based on some specific criteria (Smola & Bartlett 2000; Seeger et al. 2003; Quinonero-Candela & Rasmussen 2005; Snelson & Ghahramani 2006).

Model performance evaluation

The performance metrics used in this study for evaluating the surrogate models are the normalized root mean square error (NRMSE), correlation coefficient (R), Nash–Sutcliffe coefficient of efficiency (NSE; Nash & Sutcliffe 1970), normalized mean bias error (NMBE), threshold statistics (TSs), and normalized 95% quantile error (NQ95). Equations for these metrics are given in the following:

$$\text{NRMSE} = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (O_i - P_i)^2}}{(O_{\max} - O_{\min})} \quad (9)$$

$$\text{NQ95} = \frac{\sqrt{95\% \text{ Quantile of Square Errors}}}{(O_{\max} - O_{\min})} \quad (10)$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^N (O_i - P_i)^2}{\sum_{i=1}^N (O_i - \bar{O}_i)^2} \quad (11)$$

$$\text{R} = \frac{\sum_{i=1}^N (O_i - \bar{O}_i)(P_i - \bar{P}_i)}{\sqrt{\sum_{i=1}^N (O_i - \bar{O}_i)^2 \sum_{i=1}^N (P_i - \bar{P}_i)^2}} \quad (12)$$

$$\text{NMBE} = \frac{\frac{1}{N} \sum_{i=1}^N (P_i - O_i)}{(O_{\max} - O_{\min})} \times 100\% \quad (13)$$

$$\text{TSx} = \frac{n_x}{N} \times 100\% \quad (14)$$

where O is the observed value, P is the predicted value, N is the number of data points, O_{\max} is the maximum of the observed values, O_{\min} is the minimum of the observed values, \bar{O} is the average of the observed values, \bar{P} is the average of the predicted values, n_x is the number of points where the absolute relative error (ARE) is less than $x\%$.

For a model to be called well performing, the NRMSE, NMBE, and NQ95 values should be as low as possible. Values of R lie within a range of -1 to 1 where a value close to 1 denotes strong positive correlation and a value close to -1 indicates strong negative correlation. Values of NSE vary from negative infinity to 1 where any value close to 1 is considered to be good. A negative NSE value indicates that the performance of the model is worse than a simple persistence model. TSs give an idea about the overall ARE (or any other error) distribution. For example, a TS value of 90% at 25% ARE level (TS25) would imply that for 90% of the total points, ARE is less than 25% . Hence, larger value of TS at lower ARE level is always desired.

Water management system for coastal regions

The surrogates are applied in the frame of a simulation-based integrated water management system published by Grundmann et al. (2012, 2013). This system aims at generating optimal strategies for the management of water and agricultural resources considering water quantity and quality in the coastal regions that are infected by saltwater intrusion. A simulation-optimization approach is used where an evolution strategy is linked with an integrated simulation model that describes the coupled groundwater-agriculture hydrosystem. The trained surrogate models are used within the integrated simulation model to simulate the aquifer behavior including the dynamics of the sea-water interface. Behavior of the highly productive

agricultural farms is simulated by a database of generated stochastic crop water production functions (SCWPF) which describe the relationship between crop yield, salinity of irrigation water, and water amount for irrigation assuming an optimal water application. The SCWPF are derived by employing soil-vegetation-atmosphere-transfer models together with optimal irrigation scheduling and control algorithms (Schütze & Schmitz 2010; Grundmann et al. 2012). The objective function of the management problem combines two contradicting objectives of agricultural profit of the farmers and sustainability of the groundwater system.

The sustainability of the groundwater system is quantified in this study by the sustainability index (*SI*). *SI*, as formulated in this study, measures the stability of the groundwater system by assessing its shift from the steady state condition consequent to pumping effects:

$$SI(t_n) = 0.5 \left(\frac{1}{m_1} \sum_{k=1}^{m_1} \frac{(h(t_n, x_k) - h(t_0, x_k))}{h_{max} - h(t_0, x_k)} + \frac{1}{m_2} \sum_{k=1}^{m_2} \frac{(S(t_n, x_k) - S(t_0, x_k))}{S_{max} - S(t_0, x_k)} \right) \quad (15)$$

where k is the location index, m_1 and m_2 indicate the number of locations considered for water table height and salinity, $h(t_0, x_k)$ and $S(t_0, x_k)$ are the steady state water table height and salinity at the k^{th} location, $h(t_n, x_k)$ and $S(t_n, x_k)$ are the water table height and salinity at the k^{th} location at the end of the time period n . A lower value of *SI* indicates higher stability of the groundwater system.

The agricultural profit is calculated by a *Profit* function considering the revenues and costs of the agricultural production based on the salinity of irrigation water, cropping pattern, extraction rates, and cultivated area.

$$Profit = \sum_{i=1}^n \left[\left(\sum_{j=1}^m P_j Y_j(t_i) L(t_i) \right) - CI_j(t_i) \right] - CP(t_i) \quad (16)$$

where i is the index for cultivation period, n is the total number of time steps, j is the index for the number of the cultivated crops [$j = 1 \dots m$], m is the total number of cultivated crops, P_j is the current market prices of j^{th} crop [\$/t], Y_j is the yield of the j^{th} crop [t/ha], $L(t_i)$ is the total cultivated area at time step t_i [ha], $CI_j(t_i)$ is the cost of irrigation for

j^{th} crop at time step t_i including the fixed and the variable costs [\$], $CP(t_i)$ is the cost of pumping at time step t_i including fixed and variable costs [\$].

Both the objectives are contradicting in nature; because if the farmers aim at earning high *Profit*, they tend to pump more water, which has a negative impact on the sustainability, since excessive pumping worsens the condition of the aquifer due to a declining water table and an enforced salt water intrusion. The multi-objective optimization problem is solved in this study using a single objective evolution strategy CMA-ES (Covariance Matrix Adaptation Evolution Strategy) (Hansen & Ostermeier 2001; Hansen et al. 2003; Hansen 2006) where the contradicting objectives are coupled using varying weights (Equation (17)). This allows the design and evaluation of different management scenarios in accordance with the preferences of the stakeholders and decision-makers.

$$\max(OF) = \max(w_1 Profit + w_2 SI); \quad w_1 + w_2 = 1 \quad (17)$$

$$(A_1(t_1) \dots A_1(t_n) \dots A_m(t_1) \dots A_m(t_n), L(t_1) \dots L(t_n), \dots \\ Q(t_1) \dots Q(t_n)) = \operatorname{argmax}(OF)$$

where *OF* is the objective function that needs to be maximized, w_1 and w_2 are the weights on profit and sustainability, respectively, n is the total number of time steps, m is the total number of cultivated crops, $L(t_i)$ is the total cultivated area at time step t_i [ha], $A_j(t_i)$ is the fraction of cultivated area for j^{th} crop at time step t_i and $Q(t_i)$ is the extraction rate for the time step t_i [m³/m/s].

DEVELOPMENT OF SURROGATE MODELS FOR COASTAL AQUIFER MANAGEMENT

Study area and problem description

This study is carried out in the Barka region of Al-Batinah coastal plain situated at the north of the Sultanate of Oman, as shown in Figure 2. The plain is bordered by the Sea of Oman in the north and the Hadjar Mountain range in the south and has a width of almost 20 km from the sea to the interior mountains. Agriculture plays an important

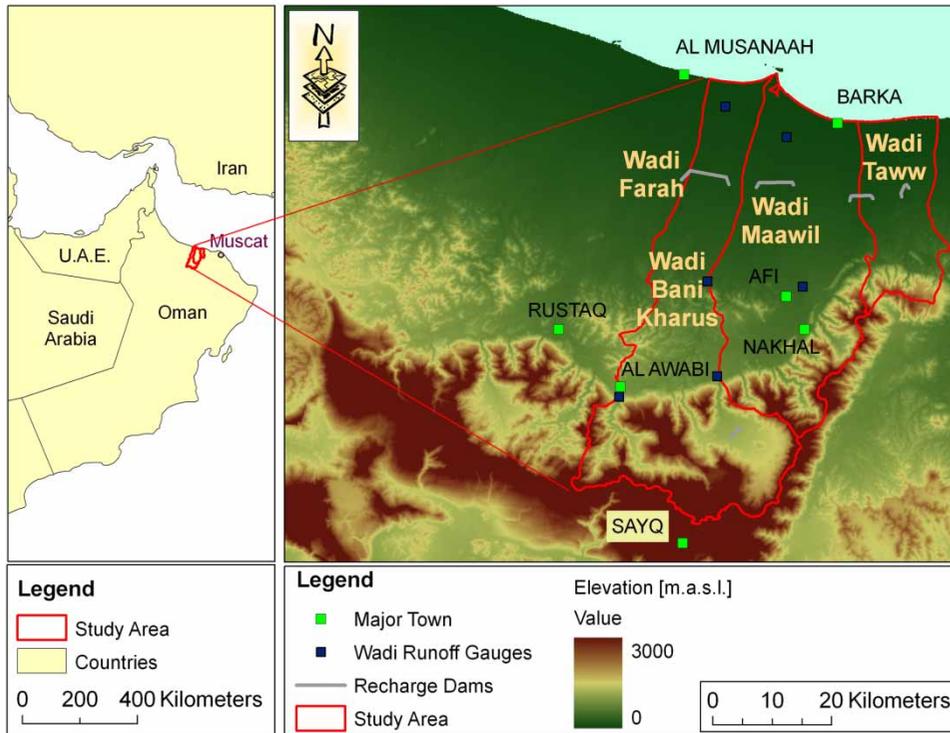


Figure 2 | Study area: Barka region at Al-Batinah plain of Oman.

role in the region since almost one-third of the economically active Omani population works in this sector (FAO 2008). A huge variety of crops is grown on numerous small-scaled farms which are situated within a distance of 6 km from the coastline. Irrigation is mandatory in the region due to the arid climate. Groundwater is pumped by many uncontrolled wells from the coastal aquifer system which is characterized by a complex heterogeneous structure of fluvial, aeolian, and marine deposits, consisting of locally cemented gravels, differently sized sand, silt, loam, and clay (Walther *et al.* 2012).

An increasing trend of groundwater extraction and a large drawdown of the water table as well as saltwater intrusion have been evident in the region since the 1960s, although the amount of the total extraction was lower at the beginning (Stranger 1985). The extraction rate in the Barka region was reported to be 30 Mm³/year in 1970 which increased to about 120 Mm³/year in the recent past (Al-Shoukri 2008). The coastal aquifer system is mainly replenished by rainfall in the southern hard rock mountains. Rainfall characteristics are highly variable

between the years. The mean annual precipitation is about 250 mm/year in the upper mountains and 50 mm/year along the coastline. Therefore, direct recharge of the coastal aquifer from rainfall is negligible. An average lateral recharge rate from the mountain range into the coastal aquifer was estimated to be 68 Mm³/year (Walther *et al.* 2012). Since the groundwater extraction exceeded the recharge rate, an inversion of the groundwater's natural flow gradient occurred together with an ongoing process of saltwater intrusion into the aquifer. As a result, the water for irrigation became saline which led to economic losses for the agricultural sector and many abandoned farms near the coastline.

Numerical groundwater model setup

Based on the work of Walther *et al.* (2012), who set up a heterogeneous 3D density-dependent groundwater flow model of the alluvial coastal aquifer for the study site, a 2D vertical slice of the unconfined 3D alluvial coastal aquifer model was constructed to demonstrate the functionality

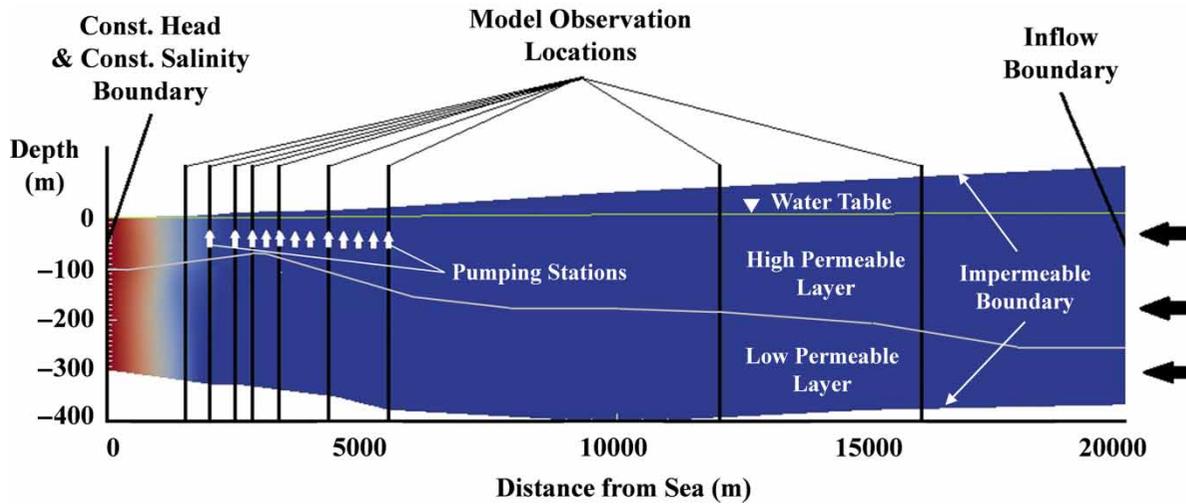


Figure 3 | Problem domain for OGS model simulations including problem geometry, boundary conditions, pumping station locations, and model observation locations.

of the surrogates for simulating groundwater flow and sea-water intrusion. The 2D vertical model domain for OGS simulations is shown in Figure 3. The left-hand side of the 20 km long domain is the lower boundary (sea-side in Figure 2) and the right-hand side is the upper boundary (mountain-side in Figure 2). The lower and the upper boundaries are 300 m and 365 m deep below the mean sea level (MSL), respectively. The lower boundary is characterized by a zero pressure boundary condition at sea level whereas the upper boundary has horizontal inflow. Increase in pressure per meter depth in the lower boundary is about 10,065.06 Pa assuming the density of the sea water to be $1,026 \text{ kg/m}^3$ and the gravitational acceleration to be 9.81 m/s^2 ($1,026 \times 9.81 = 10,065.06$). The salinity at the lower boundary is set to its maximum; i.e., the salinity of the sea water. Upper and lower boundaries are considered to be impermeable. The assumption of impermeable upper boundary is reasonable for the study area since groundwater recharge subsequent to rainfall is negligible in the coastal plain due to the arid climate conditions with significantly low rainfall (see the section ‘Study area and problem description’). Two distinct layers are created to approximately match the actual hydrogeology of the aquifer (Figure 3). The agricultural pumping region is represented by 12 pumping stations located within 2 to 5.5 km from the coastline (Figure 3). In order to roughly imitate the actual field conditions, all the pumps are kept equidistant with an interval of 300 m,

except for the first two where the interval is taken to be 500 m. Individual pumps are represented by vertical lines of 20 m length situated at a depth of 35 to 55 m below MSL. A grid with variable node densities is adapted in this study for the numerical modeling of the flow and transport processes. The grid size varies from 10 m near the pumping region to 50 m in the inner part of the domain. Near the sea front, the grid size is kept at 15 m and near the inflow boundary it is 50 m. The generated mesh has 15,373 nodes and 29,352 triangular elements. The node density values are also checked with the Peclet and Courant conditions for numerical stability.

Table 2 shows the parameter values of the OGS model used in this study. For variable density flow simulation, the

Table 2 | Final parameter set for OGS model simulations

Parameters	Upper layer	Lower layer
Porosity	0.20	0.15
Tortuosity	1	1
Isotropic permeability	$1 \times 10^{-10} \text{ m}^2$	$1 \times 10^{-12} \text{ m}^2$
Longitudinal dispersion length	100 m	100 m
Transverse dispersion length	50 m	50 m
Diffusion coefficient	$5 \times 10^{-9} \text{ m}^2/\text{s}$	$5 \times 10^{-9} \text{ m}^2/\text{s}$
Solid density	$2 \times 10^3 \text{ kg/m}^3$	$2 \times 10^3 \text{ kg/m}^3$
Viscosity	$1 \times 10^{-3} \text{ kg/m/s}$	$1 \times 10^{-3} \text{ kg/m/s}$

difference between the density of saltwater and the fresh water is assumed to be 26 kg/m^3 and the permeability reduction factor (Goswami & Clement 2007) is taken to be 0.001.

The transient model with the aforementioned settings was run in three steps; i.e., (1) steady state simulation, (2) initial state simulation, and (3) future state simulation. In order to achieve the steady state condition, OGS was first run for 300 years with an inflow rate of $68 \text{ Mm}^3/\text{year}$ and an extraction rate of $36.9 \text{ Mm}^3/\text{year}$, as reported by Walther *et al.* (2012). The upper permeable layer reached the steady state condition within the first 50 years but the lower layer took almost 200 years. The initial state was derived by further running the model for another 40 years with a higher extraction rate and the same inflow rate as before. The extraction rate had to be kept high in order to ensure reasonable water table drawdown and progress of the seawater front that match the current state of the aquifer. The end point of the initial state simulation (i.e., water table heights and salinity after 40 years) was considered to be the starting point of the future state simulations, which were carried out to generate a scenario database for developing the surrogates. Water table heights and salinity values were extracted from the model simulation results at nine different model observation locations (MOLs) at 1.5, 2, 2.5, 2.8, 3.4, 4.3, 5.5, 12, and 16 km from the coastline (Figure 3). All inflow and extraction rates were divided by the third dimension of the model domain (30 km, see Walther *et al.* (2012)) to make them representative of the 2D model.

Generation of scenario database

The scenario database for developing the surrogate models was created by running future state simulations (see the section 'Numerical groundwater model setup') with OGS for 60-year periods with different combinations of inflow and extraction rates. In total, four different datasets were created for training, testing, validation, and extrapolation. ANNs were trained using the training set and some superior models were selected based on their performance in testing. The final ANN model was selected from the validation results. GPM was trained using the training set and was also run with the validation set. Finally, both the surrogates

were run with the extrapolation set to investigate the model consistency for extreme values. Note that the term 'extrapolation' throughout this paper corresponds to any inflow or extraction rate that is outside its range used in training. It is not a temporal extrapolation.

For the training set preparation, five different extraction rates and five different inflow rates (in order to account for the variability in groundwater replenishment) were selected such that they include values from low to high ranges. Additionally, an adaptive pumping scheme was applied during the higher extraction rates by deactivating the seaside pumps in order to reproduce farmers' adaptation on saline water intrusion. If five extraction rates were to be distributed over the 60-year period considering yearly time step, the total number of unique patterns would be 5^{60} . The same number of patterns could be generated for five inflow rates as well. If we consider both inflow and extraction patterns together, the number of total patterns would become 25^{60} . Clearly, there was an obvious need to reduce the number of patterns. For that purpose, inflow and extraction rates were treated as fixed pairs, meaning, the first extraction pattern would always concur with the first inflow pattern, the second extraction pattern with the second inflow pattern, and so on. This reduced the total number of patterns from 25^{60} to 5^{60} . These patterns were then applied to the first 12 years of the 60-year period. In order to further reduce the size of the data, we considered a 3-year time step for the management problem, which essentially implies that the inflow and the extraction rates are constant for 3 successive years. This was a valid assumption in our case mainly for three reasons: (1) we were interested in long-term planning and management, (2) constant pumping rates can arise from water quota systems, (3) since inflow rates depend on the longer time scales, the 3-year duration follows the dynamics of the groundwater recharge more closely. These 12-year patterns with 3-year time step were then repeated four more times to extend the entire management period of 60 years. This technique reduced the number of patterns drastically to $5^4 = 625$. Data from OGS were extracted at a fixed interval of 3 years, thereby creating 20 points for each 60-year pattern. Thus, the final size of the training data became $625 \times 20 = 12,500$. Similar methodology was applied for generating the other three datasets as well. For testing and validation sets, four different inflow and four different

extraction rates were chosen, and the size of each one of the datasets was $4^4 \times 20 = 5,120$. Furthermore, the inflow and the extraction rates for testing and validation were chosen such that they lie within their range used in training. The extrapolation set was created with two values of inflow and two values of extraction rates such that one value is lower than the lowest and the other value is higher than the highest used in training. Thereby, the size of the extrapolation set was $2^4 \times 20 = 320$. Different inflow and extraction rates used for the data preparation in this study are summarized in Table 3.

Table 3 | Different inflow and extraction rates (Mm³/year) used for data preparation

	Inflow rates	Extraction rates
Training	50, 60, 70, 80, 90	50, 150, 300, 400, 600
Testing	55, 65, 75, 85	100, 200, 350, 500
Validation	52, 68, 72, 88	100, 250, 450, 500
Extrapolation	40, 100	40, 700

There were still repetitions in the datasets that needed to be filtered out. For example, the first extraction pattern for the 60-year management period in the training dataset would look like:

E ₁																				
----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------

where each of the 20 square boxes represents a time period of 3 years and E₁ denotes the first extraction rate out of the five rates selected during training. The second pattern looks like:

E ₁	E ₁	E ₁	E ₂	E ₁	E ₁	E ₁	E ₂	E ₁	E ₁	E ₁	E ₂	E ₁	E ₁	E ₁	E ₂	E ₁	E ₁	E ₁	E ₁	E ₂
----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------

This time it has also included the second extraction rate E₂. As can be seen, the first three values of the second pattern are redundant once the first pattern is already fed to the model. Removing the first three values would create a modified second pattern as:

E ₂	E ₁	E ₁	E ₁	E ₂	E ₁	E ₁	E ₁	E ₂	E ₁	E ₁	E ₁	E ₂	E ₁	E ₁	E ₁	E ₁	E ₂
----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------

This technique was applied to all the four datasets which in addition to reducing the number of data points, also helped in avoiding overfitting (Table 4).

OGS had to be run once for each individual pattern. The training data had 625 patterns, each of the testing and validation data had 256 patterns, and the extrapolation data had 16 patterns, which produced a total number of $625 + 256 + 256 + 16 = 1,153$ patterns. Therefore, there were 1,153 OGS runs which were distributed over 14 dual core computers, each running two OGS programs simultaneously. The entire data preparation process took around 16 hours.

Table 4 | Size of different datasets before and after removing the repetitions

	With repetitions	Without repetitions
Training	12,500	10,780
Testing	5,120	4,436
Validation	5,120	4,436
Extrapolation	320	286

Datasets for the surrogates were generated from the OGS model based on different inflow–extraction scenarios for a management period of 60 years. Water table heights were extracted from all nine MOLs whereas only the first

seven MOLs were considered for the salinity because of the fact that salinity was negligible at the last two MOLs. One inflow rate, one extraction rate, nine water table heights, and seven salinity values from any given time step

were considered to be the inputs for that time step whereas nine water table heights and seven salinity values from the next time step were used as outputs. Table 5 shows the first three rows of any dataset with superscripts indicating the time steps. As is seen, for each time step there are 18

Table 5 | First three rows of any dataset

Time step	Inputs		Targets			
1	Inflow ¹	Extraction ¹	9 WT ¹	7 Sal ¹	9 WT ²	7 Sal ²
2	Inflow ²	Extraction ²	9 WT ²	7 Sal ²	9 WT ³	7 Sal ³
3	Inflow ³	Extraction ³	9 WT ³	7 Sal ³	9 WT ⁴	7 Sal ⁴

inputs from that time step and 16 outputs from the next time step. Training based on the pattern of the data shown in Table 5 makes the surrogates suitable for variable-time simulations.

ANN development

A two-layer feed-forward neural network was developed in this study with linear output activation function and mean square error function in training using the NETLAB Toolbox (Nabney 2002). The ANN was trained using the fully automated SCG algorithm with maximum iteration being the termination criterion. The training scheme was applied in two steps: (1) systematic analysis and (2) refinement. The main purpose of the systematic analysis step was to gain some preliminary knowledge about a reasonable range of the number of hidden neurons and maximum iterations. It is always desirable to have a simple (less hidden neurons) model that is able to demonstrate almost similar performance as compared to its more complex counterparts. Moreover, the number of iterations must not be so high that the model tends to overfit the training data. In the systematic analysis step, the hidden neurons were varied from 10 to 190 with an increment of 20, and the iterations were varied from 50 to 500 with an increment of 50. As can be seen from Figure 4(a), the NRMSE shows a decreasing trend even at 500th iteration. Therefore in the refinement step, the iterations were increased to a maximum of 5,000 with an increment of 500 (Figure 4(b)). Furthermore, networks that showed poor performance during systematic analysis step (e.g., 10, 30) and those with a very high (>100) number of hidden neurons were also removed before executing the refinement step. Some superior models were selected based on their testing performance whereas the final model was selected from the validation results. The model selection process was based on evaluating the performance

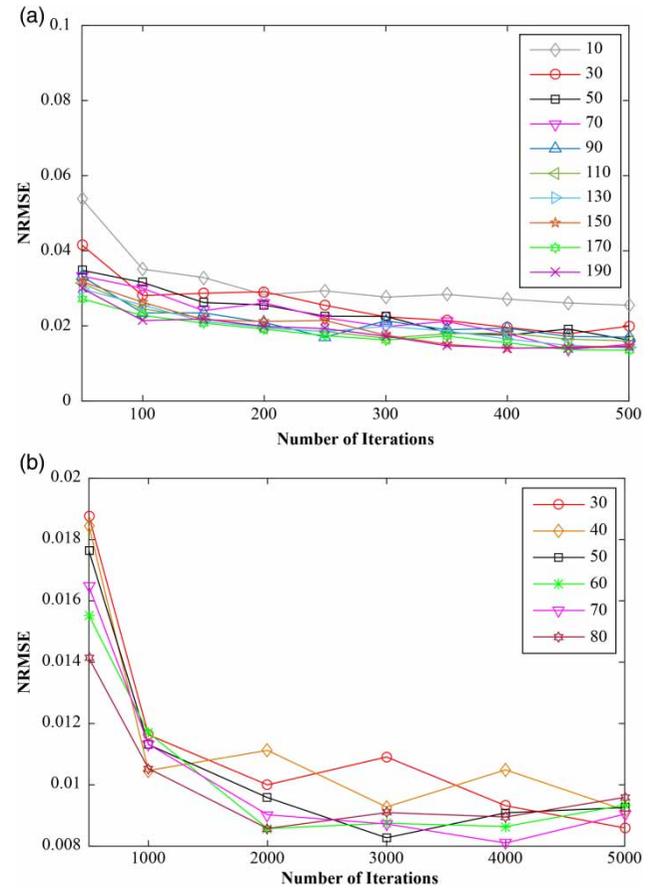


Figure 4 | Performance evaluation of ANNs using NRMSE for different numbers of hidden neurons and iterations during the two step training procedure: (a) systematic analysis step; (b) refinement step.

metrics described in the section ‘Model performance evaluation’. The final model had 18 inputs, 30 hidden neurons, and 16 outputs. The input to hidden layer connection had $18 \times 30 = 540$ weights and 30 biases, and the hidden to output connection had $30 \times 16 = 480$ weights and 16 biases. Therefore, the total number of weights and biases became $540 + 480 + 30 + 16 = 1,066$.

GPM development

A total of 16 single-output GPMs were trained in this study for 16 different outputs and they were combined together to build a unified modeling framework. As for the GPM configuration, linear mean function and composite covariance function (linear and squared exponential function forms) were used. The fully independent training

conditional (Naish-Guzman & Holden 2007) approximation scheme with 500 pseudo inputs was applied as a sparse GP technique. The GPML (Gaussian Processes for Machine Learning) Toolbox (Rasmussen & Williams 2006) was used for the implementation of GPM.

Setup of water management system application

For demonstrating the ability of the surrogate models to emulate the behavior of a physically based numerical groundwater model, an application of the surrogate models within the frame of a water management system for coastal regions is presented as an example (see Figure 1, Step 3). The long-term evaluation of the environmental and economic development over 60 years in a hypothetical farm located in the vicinity of the sea at MOL 3, was investigated for a more profit oriented management scenario with $w_1=0.8$ and $w_2=0.2$ (see Equation (17)). For the sake of simplicity, cultivation of two different crops growing in two seasons per year was considered: (1) maize, a salt-sensitive cash crop with a market price of \$175/t, and (2) sorghum, a lower priced but more salt-resistant crop with a market price of \$100/t. Additionally, regional specific cost models for the irrigated agriculture and the groundwater abstraction were incorporated (Malik & Al-Zubeidi 2006). The three decision variables considered in this study were the cropping pattern, groundwater extraction rate, and cultivated area, which would produce a total number of $3 \times 60 = 180$ variables if they were to be applied to a 60-year period with yearly time stepping. However, since a 3-year time step

was established throughout the course of this study, the total number of variables reduced to $3 \times 20 = 60$. In order to keep the optimization problem manageable for this example application, it was assumed that the variables were constant for each set of non-overlapping 12 consecutive years (i.e., four time steps). Thus, the final number of decision variables reduced to $60/4 = 15$. The CMA-ES algorithm needed 2,914 iterations and a population size of 12 to solve the optimization problem.

RESULTS

Generation of surrogate models

The average validation error statistics for all 16 outputs of the final model are given in Table 6. The performance of ANN was also investigated for the extrapolation dataset in order to check the model consistency for extreme values. As seen in Table 6, the model performance was quite satisfactory even in the extrapolation range.

The average validation and extrapolation error statistics of the combined GPM framework are given in Table 7. Similar to ANN, the performance of GPM in the extrapolation also turns out to be quite satisfactory. Both the surrogates show comparable performance as captured by the performance metrics. They both have very low values of NRMSE, NQ95, and NMBE, and very high values of NSE, R, and threshold statistics. Performance of GPM is relatively better, particularly in the extrapolation range.

Table 6 | Validation and extrapolation error statistics of the final ANN model

	NRMSE	NQ95	NSE	R	NMBE	TS25	TS100
Validation	0.0159	0.0258	0.9784	0.9974	-0.7772	65.34	85.23
Extrapolation	0.0762	0.1516	0.6941	0.9993	0.5339	38.83	67.19

Table 7 | Average validation and extrapolation error statistics of the combined GPMs

	NRMSE	NQ95	NSE	R	NMBE	TS25	TS100
Validation	0.0136	0.0244	0.9865	0.9969	-0.1254	74.18	90.19
Extrapolation	0.0064	0.0110	0.9977	0.9993	-0.3288	79.31	90.78

Behavior of surrogate models for interpolation

The term ‘interpolation’ within the context of this paper corresponds to any scenario which is a trajectory over the states of the system that are within the range of the states used in training. This means that a scenario can have a time period which is shorter, equal, or longer than the 60-year reference time period used in training. To demonstrate the interpolation behavior of the surrogates, a long-term simulation was performed over 84 years. The values of inflow and extraction rates selected for this purpose had never been used in any of the datasets prepared earlier. Figure 5 shows the resulting water levels and salinity values for selected MOLs at different distances from the sea in comparison to the OGS model outputs. Although the surrogates are trained using 60-year scenario database, they perform reasonably well beyond that range. As can be seen, the results are quite similar only with a few exceptions.

GPM tends to follow OGS simulation more closely as compared to ANN in case of salinity at MOL 1. For the water table height at MOL 4, the performance of ANN is quite satisfactory within the first 60-year time period; however, it tends to underestimate the heights beyond that range. GPM, on the other hand, consistently performs well even while temporally extrapolated. Salinity values are overall quite low at MOL 4 since that location is quite far away from the sea.

Behavior of surrogate models at their boundaries

In order to analyze the behavior of the surrogate models at the boundaries of their trained space, 60-year simulation runs were started from the initial conditions of the aquifer with constant low ($55 \text{ Mm}^3/\text{year}$) and high ($595 \text{ Mm}^3/\text{year}$) extraction rates. The resulting temporal variations of water table and salinity profiles over the distance from the

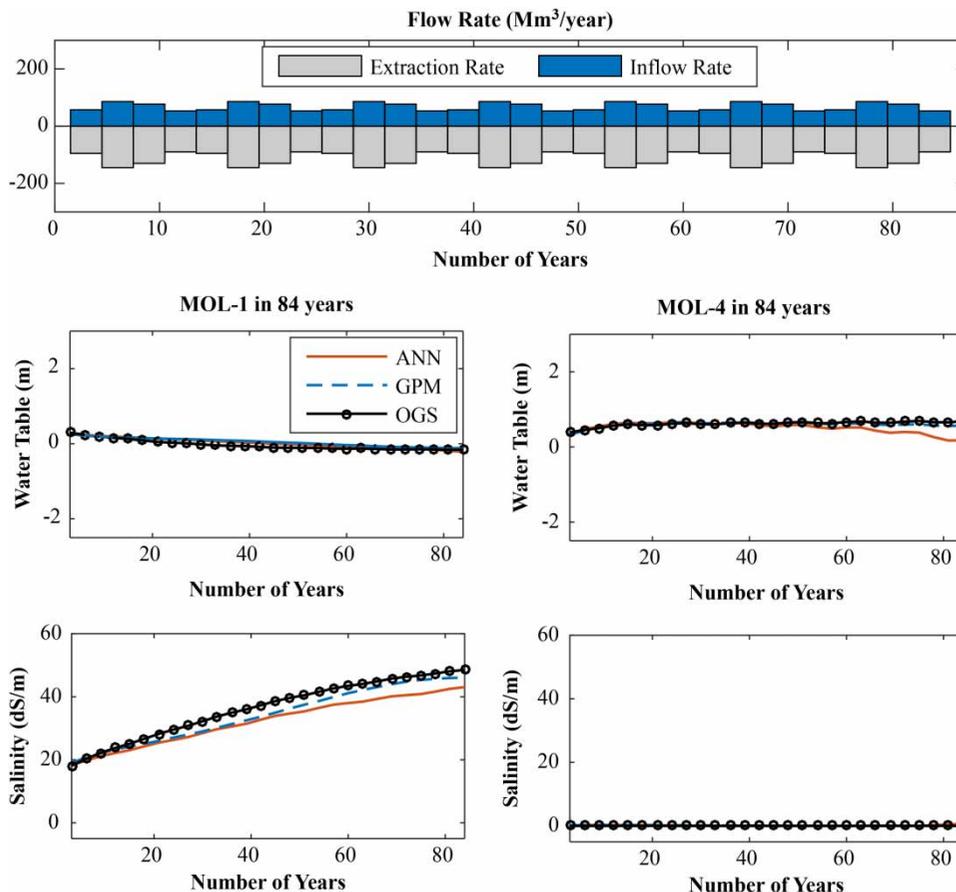


Figure 5 | Comparison of model outputs by water levels and salinity at MOL 1 and MOL 4 for 84-year simulation period with changing inflow and extraction rates.

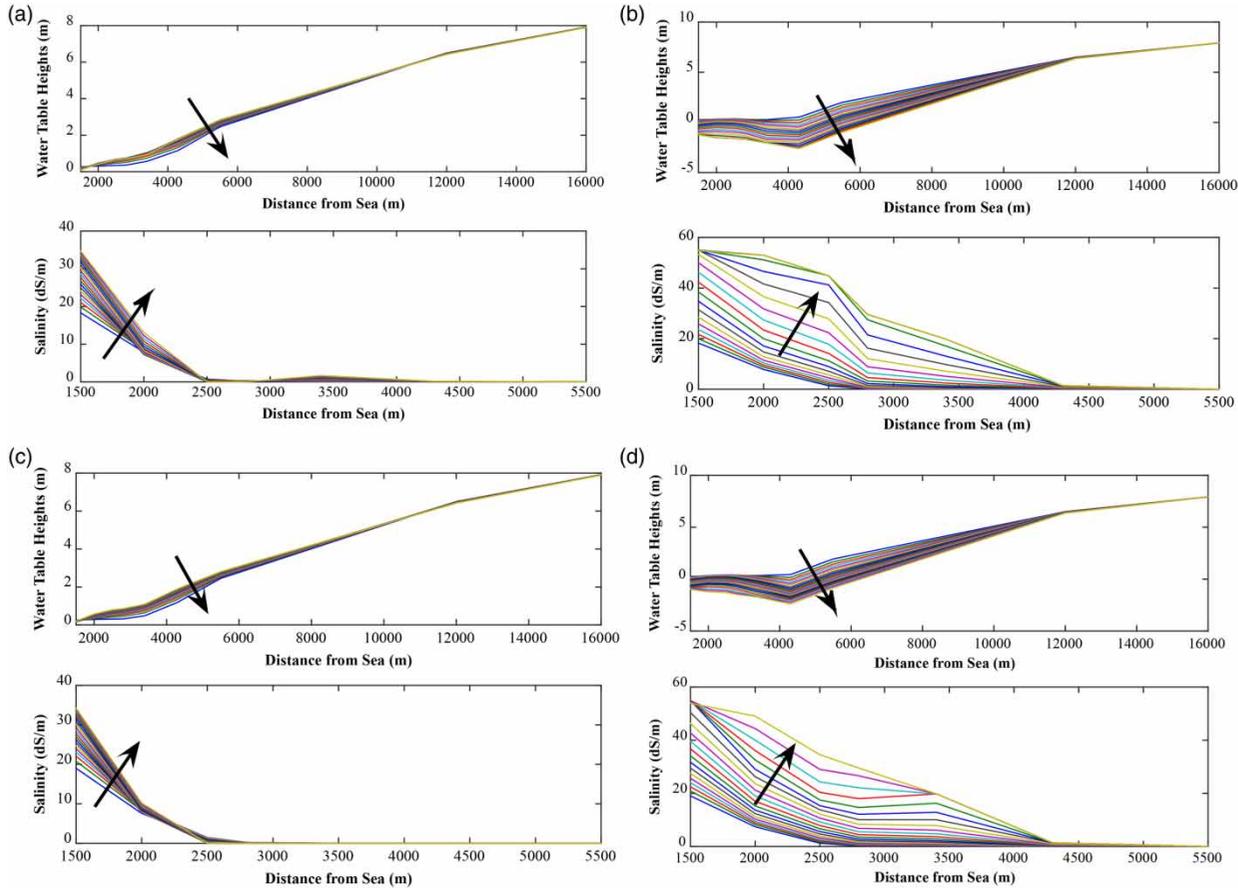


Figure 6 | Temporal variations of water table and salinity profiles over distance from the sea for ANN (a, b) and GPM (c, d) at lower (a, c) and higher (b, d) extraction rates.

sea are shown in Figure 6. The arrows in the figure indicate the directions of change over time.

As seen in Figure 6, both ANN and GPM perform quite satisfactorily and there is a good agreement in their outputs. They follow almost a similar temporal pattern while applied to the model domain. For the low extraction case, the water table drawdown is quite small and it always stays above the MSL whereas the salinity goes up to almost 35 dS/m. Water table drawdown is predominant for high extraction rate where it goes down to 2.5 m below MSL. The salinity value in this case also reaches its maximum at the end of the simulation period.

Surrogate model results are then compared against the simulation results of OGS in low ($55 \text{ Mm}^3/\text{year}$) and high ($595 \text{ Mm}^3/\text{year}$) extraction cases at the end of the simulation period, as shown in Figure 7. Water table heights are well approximated by both the surrogates for low as well as high extraction rates. The performance differs mainly for salt

concentration simulation. As is seen, GPM follows the salt concentration profile simulated by OGS more closely at the low extraction case. In the case of high extraction, GPM underestimates salt concentration at the sea-side pumps.

Although GPM showed comparable accuracy as compared to ANN, the computational time was not improved much. GPM was around five to six times faster than OGS for this particular application whereas ANN was more than 10,000 times faster. While these results substantiate the successful use of ANN as a surrogate model, they also indicate that GPM has the potential to become a prominent candidate for surrogate modeling if the computational time can be reduced through future research.

Application of the water management system

Selected results from an application of ANN within the frame of a water management system over 60 years are

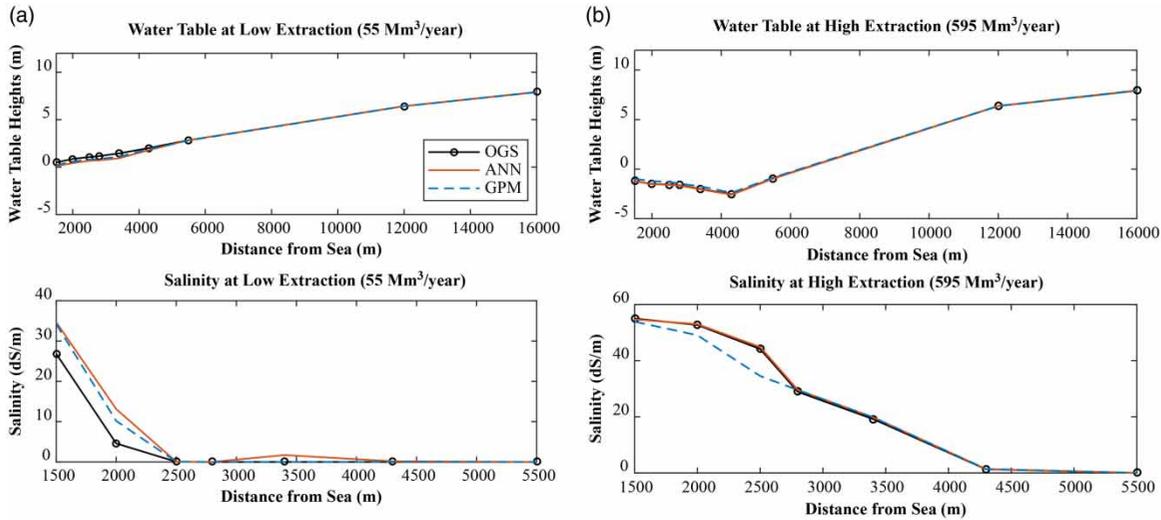


Figure 7 | Comparison of OGS, ANN, and GPM outputs for water table heights (upper) and salinity (lower) over distance from the sea for lower (left) and higher (right) extraction rates at the end of the simulation period. (a) Low extraction (55 Mm³/year). (b) High extraction (595 Mm³/year).

presented in Figure 8. Optimal values of the decision variables for a more profit oriented management scenario have been estimated in order to manage the agricultural production in the coastal arid region which is infected by saltwater intrusion. Considering the trade-off between the

two contradicting objectives of profit and sustainability, an overall 60-year profit was estimated to be \$158,118 and the value of *SI* was 0.1714. The estimated profit is quite significant for the particular application concerned and the value of *SI* also indicates that the groundwater system will

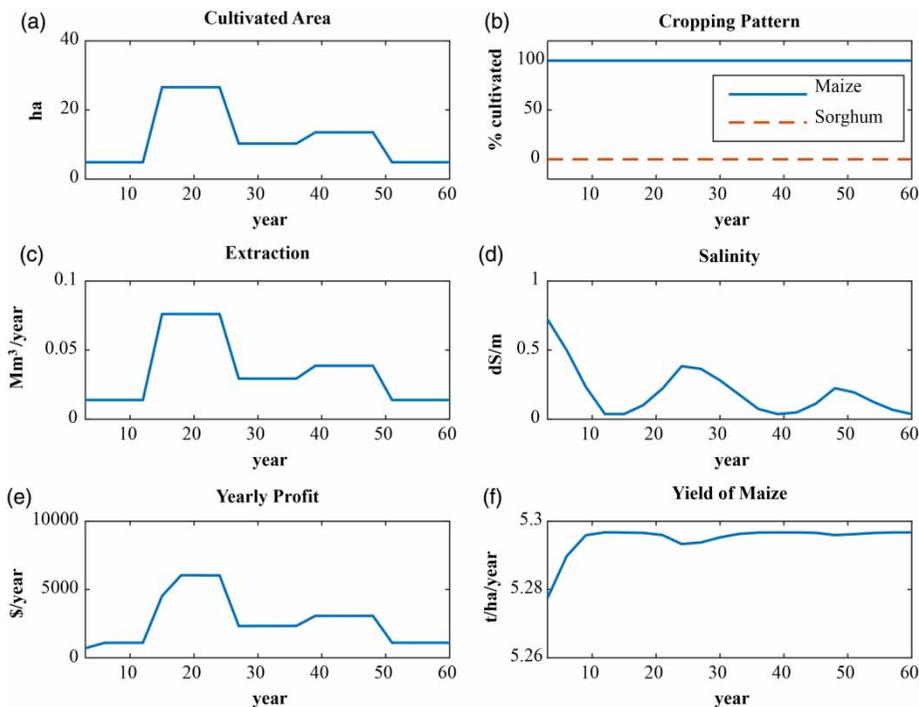


Figure 8 | Selected results from a long-term (60 years) optimization run for a profit-oriented management scenario at a specific farm location.

remain in an almost sustainable state at the end of the management period. As is seen, cropping of maize is suggested throughout the whole management period (Figure 8(b)). The salinity increases twice in 60 years but with small peaks (Figure 8(d)). At the end of the management period, it achieves a very low value which means that the groundwater system is sustainable. The algorithm increases the extraction rates from the 12th year to almost the 27th year and then again from the 36th year to the 47th year (Figure 8(c)). The increased extraction rates also correspond to the increased cultivated area (Figure 8(a)) and the increased profit (Figure 8(e)). The yield of maize (salt-sensitive crop) increases with the decrease in salinity and vice versa (Figure 8(f)). Table 8 shows the optimal parameter set for this scenario where each row corresponds to a time period of 12 consecutive years. This scenario can be adapted to a real life problem since it has reasonable trade-off between the two contradicting objectives of profit and sustainability.

The optimization results presented so far in this paper are derived by using ANN as the surrogate model because of its fast performance. Figure 9 shows the cumulative profit over time which is derived by running the

management framework in simulation mode with all the three models (ANN, GPM, and OGS) and the optimal values of the decision variables given in Table 8. As can be seen, GPM slightly underestimates the cumulative profit whereas ANN shows very small overestimation. Overall, the graphs look quite similar to each other which substantiates the successful application of the surrogates to replace the high-fidelity numerical model. However, there were noticeable differences in the computation time requirements. To run the same simulation, OGS and GPM took 456 seconds and 148 seconds, respectively, whereas ANN only took 0.8 seconds.

DISCUSSION

The training scheme applied in this study produces time-dependent surrogate models which do not require any additional training if the simulation time is altered (see Figure 5). Optimization with varying planning horizon is rather very common in several groundwater management applications such as remediation or restoration of coastal aquifers. A variable period affects the pumping rates and the removal efficiency of the remediation systems (Singh & Chakrabarty 2011; Peralta & Kalwij 2012). Thus, time is an important decision variable, or it can be included as an additional objective when trade-offs between costs and time are of interest. In all those applications, surrogate modeling approaches assuming fixed time horizons are not that useful because for each model run, a new training of the surrogate becomes compulsory. Moreover, in cases where the objective function is not fully defined at the beginning and needs to be refined during the investigation for a satisfying solution, the usefulness of variable-time surrogate modeling and decoupling of the model and the objective function is quite obvious. The modeler can use a single trained model to carry out this interactive exploration and also several other tasks such as sensitivity analyses, uncertainty analyses at different time scales, comparing long- and short-term management decisions, etc., without carrying out a new training procedure.

Therefore, our goals in this investigation are to understand: (1) how closely the state-space surrogates follow the results calculated by the numerical groundwater model; (2)

Table 8 | Optimal values of the decision variables in the profit-oriented long-term management scenario

Cropping pattern (fraction of maize)	Extraction rates (Mm ³ /year)	Cultivated area (hectare)
1	50	4.85
1	274	26.57
1	106	10.25
1	139	13.51
1	50	4.85

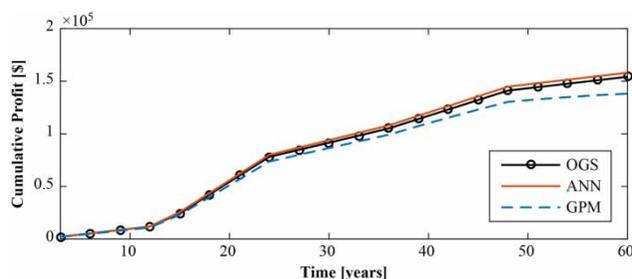


Figure 9 | Comparison of the surrogates and OGS in terms of predicted cumulative profit in the long-term water management scenario.

how well the surrogates behave in extrapolation which corresponds to some extreme boundary conditions not considered during training; (3) how similar the optimal management decisions are when trained surrogates are used within the optimization routine instead of the actual numerical groundwater model; and (4) how computationally efficient the surrogates are as compared to the numerical groundwater model. Both the surrogates show comparable accuracy in validation and extrapolation (Tables 6 and 7). The NRMSE, NQ95, and NMBE statistics are all quite low whereas the values of NSE, R, and threshold statistics are on the higher range for both the models, which collectively indicate superior model performance. GPM also tends to perform better in the extrapolation range. During interpolation (Figure 5), GPM follows OGS more closely for salinity at MOL 1. For water table heights at MOL 4, the performance of ANN is quite satisfactory for the first 60 years, after which it shows some underestimations. Overall, the results are quite similar and they also substantiate a successful application of the variable-time simulation framework since the surrogates in this case were trained for 60 years but were run for 84 years. Comparison of the surrogates at the boundaries of their training space (Figure 7) also shows good agreement in the outcomes. Both the surrogates approximate water table heights closely for the low as well as high extraction rates. In the case of low extraction, GPM imitates the salinity profile more closely, but when the extraction rate is kept high, it underestimates the salt concentration near the sea-side pumps.

Although GPM has shown high accuracy in approximating the OGS model, it still suffers from its relatively higher computational time requirement. It did perform faster than OGS, however, it was considerably slower as compared to ANN. Improving the GPM speed should be regarded as an avenue for future research. Attempts should be made through a more rigorous training with increased number of inflow and extraction rates to solve the issue with ANN in Figure 5. It is worth mentioning that ANN in this case runs out of the valid state-space of the training dataset, which could be one of the reasons for such an abnormal behavior. One solution approach that might work would be to train the ANNs for the longest time period and then carry out variable-time simulations within that temporal range. Another probable reason

behind the behavior of ANN in Figure 5 could be the implementation of the adaptive pumping scheme, because there is a chance that it might have introduced discontinuities in the ANN. We address this particular issue more elaborately in our ongoing research.

SUMMARY AND CONCLUSIONS

In this study, we present a simulation-optimization framework for solving coupled groundwater-agriculture management problems in the arid coastal aquifer of Oman which is infected with saline water as a result of uncontrolled pumping over the last few decades. The framework couples a groundwater model with an optimization routine to estimate the optimal values of the decision variables of the management problem. This process usually involves thousands of runs of the computationally expensive groundwater model which eventually leads to a significantly high overall computational time requirement. To overcome this issue, faster surrogate models are used as alternatives to the numerical groundwater model. In this study, we replace the high-fidelity numerical groundwater model OGS with two state-space surrogate models (ANN and GPM) in a water management framework for the coastal aquifer of Oman. The surrogates are trained using the scenario database generated from the physically based OGS model considering 2D transient density-dependent groundwater flow and salt water transport processes in a two-layer aquifer system. Keeping ANN as the main candidate for the surrogate modeling, we further analyze the comparative benefits and drawbacks of GPM. We then demonstrate an example of a long-term (60-year) management problem and analyze the agricultural profit as well as aquifer sustainability through the application of our methodology.

Although both ANN and GPM showed comparable accuracy in approximating the response of OGS, the main drawback of GPM was its relatively higher computational time requirement for the particular problem presented in this study. It did perform faster than the high-fidelity numerical model; however, the computational speed needs to be improved further, which should be regarded as an avenue for future research. ANN, on the other hand, approximated the numerical model quite closely and was also significantly

fast. The training scheme applied in this study made the surrogates suitable for variable-time simulations. This kind of a training scheme is highly beneficial for various applications where simulation period needs to be altered frequently; i.e., for exploring optimal water management strategies on different time horizons (Grundmann et al. 2012). The applied multi-objective water management framework facilitates the evaluation of different management scenarios by changing the weights on profit and sustainability terms in the objective function. Furthermore, it also allows the exploration of the entire trade-off between the contradicting objectives within a multi-criteria optimization (Grundmann et al. 2013) in a reasonable computational time.

ACKNOWLEDGEMENTS

The first author of this manuscript was supported by the DAAD Master's Scholarship. The manuscript was prepared within the research project IWAS funded by the German Federal Ministry of Education and Research (BMBF) under grant no. 02WM1028. We would like to thank the Ministry of Regional Municipalities and Water Resources of the Sultanate of Oman for supporting the IWAS-IWRM project. We thank Marc Walther (TU Dresden/UFZ Leipzig) for his help with the OGS model.

REFERENCES

- Al-Shoukri, S. S. 2008 Modeling of Groundwater Flow in Wadi Ma'awil Catchment, Barka in Sultanate of Oman. PhD Thesis, Arabian Gulf University, Bahrain.
- Ataie-Ashtiani, B. & Ketabchi, H. 2011 Elitist continuous ant colony optimization algorithm for optimal management of coastal aquifers. *Water Resour. Manage.* **25** (1), 165–190.
- Ataie-Ashtiani, B., Ketabchi, H. & Rajabi, M. M. 2014 Optimal management of a freshwater lens in a small island using surrogate models and evolutionary algorithms. *J. Hydrol. Eng.* **19** (2), 339–354.
- Bayer, P. & Finkel, M. 2004 Evolutionary algorithms for the optimization of advective control of contaminated aquifer zones. *Water Resour. Res.* **40**, W06506.
- Bayer, P. & Finkel, M. 2007 Optimization of concentration control by evolution strategies: formulation, application, and assessment of remedial solutions. *Water Resour. Res.* **43** (2), W02410.
- Bhattacharjya, R. K. & Datta, B. 2009 ANN-GA-based model for multiple objective management of coastal aquifers. *J. Water Resour. Plann. Manage.* **135** (5), 314–322.
- Bray, B. S. & Yeh, W. W. G. 2008 Improving seawater barrier operation with simulation optimization in Southern California. *J. Water Resour. Plann. Manage.* **134** (2), 171–180.
- Bredheoef, J. D. 2002 The water budget myth revisited: why hydrogeologists model. *Groundwater* **40** (4), 340–345.
- Das, A. & Datta, B. 2001 Application of optimisation techniques in groundwater quantity and quality management. *Sadhana – Academy Proceedings in Engineering Sciences* **26** (Part 4), 293–316.
- Delfs, J. O., Park, C.-H. & Kolditz, O. 2009 A sensitivity analysis of Hortonian flow. *Adv. Water Resour.* **32** (9), 1386–1395.
- de Paly, M., Bürger, C. M. & Bayer, P. 2013 Optimization under worst case constraints – a new global multimodel search procedure. *Struct. Multidis. Optimiz.* **48** (6), 1153–1172.
- Dhar, A. & Datta, B. 2009 Saltwater intrusion management of coastal aquifers. I: linked simulation-optimization. *J. Hydrol. Eng.* **14** (12), 1263–1272.
- FAO 2008 Geography, climate, population. FAO's Information System on Water and Agriculture AQUASTAT. www.fao.org/nr/water/aquastat/main/index.stm.
- Fen, C. S., Chan, C. C. & Cheng, H. C. 2009 Assessing a response surface-based optimization approach for soil vapor extraction system design. *J. Water Resour. Plann. Manage.* **135** (3), 198–207.
- Fu, G., Makropoulos, C. & Butler, D. 2010 Simulation of urban wastewater systems using artificial neural networks: embedding urban areas in integrated catchment modelling. *J. Hydroinform.* **12** (2), 140–149.
- Gaur, S., Chahar, B. R. & Graillot, D. 2011 Analytic elements method and particle swarm optimization based simulation-optimization model for groundwater management. *J. Hydrol.* **402** (3–4), 217–227.
- Goswami, R. R. & Clement, T. P. 2007 Laboratory-scale investigation of saltwater intrusion dynamics. *Water Resour. Res.* **43** (4), 1–11.
- Grundmann, J., Schütze, N., Schmitz, G. H. & Al-Shaqsi, S. 2012 Towards an integrated arid zone water management using simulation-based optimisation. *Environ. Earth Sci.* **65** (5), 1381–1394.
- Grundmann, J., Schütze, N. & Lennatz, F. 2013 Sustainable management of a coupled groundwater-agriculture hydrosystem using multi-criteria simulation based optimization. *Water Sci. Technol.* **67** (3), 689–698.
- Guo, W. & Langevin, C. D. 2002 *User guide to SEAWAT: A Computer Program for Simulation of Three-Dimensional Variable-Density Groundwater Flow*. Report of the US Geological Survey. USGS, Reston, VA, USA.
- Haddad, O. B. & Marino, M. A. 2007 Dynamic penalty function as a strategy in solving water resources combinatorial optimization problems with honey-bee mating optimization (HBMO) algorithm. *J. Hydroinform.* **9** (3), 233–250.

- Haddad, O. B. & Marino, M. A. 2011 Optimum operation of wells in coastal aquifers. *Proc. Inst. Civil Eng. – Water Manage.* **164** (3), 135–146.
- Hamidi, M., Reza, S. & Yazdi, S. 2006 Numerical modeling of seawater intrusion in coastal aquifer using finite volume unstructured mesh method. *WSEAS Trans. Math.* **5** (6), 648.
- Hansen, N. 2006 The CMA evolution strategy: a comparing review. In: *Towards a New Evolutionary Computation. Advances in the Estimation of Distribution Algorithms* (J. A. Lozano, P. Larranaga, I. Inza & E. Bengoexea, eds), Springer, Berlin, pp. 75–100.
- Hansen, N. & Ostermeier, A. 2001 Completely derandomized self-adaptation in evolution strategies. *Evol. Comput.* **9**, 159–195.
- Hansen, N., Müller, S. D. & Koumoutsakos, P. 2003 Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evol. Comput.* **11**, 1–18.
- Henry, H. R. 1960 Salt Intrusion Into Coastal Aquifers. PhD Thesis, Columbia University, New York, NY, USA.
- Hsu, K. L., Gupta, H. V. & Sorooshian, S. 1995 Artificial neural network modeling of the rainfall-runoff process. *Water Resour. Res.* **31** (10), 2517–2530.
- Jain, A., Sudheer, K. P. & Srinivasulu, S. 2004 Identification of physical processes inherent in artificial neural network rainfall runoff models. *Hydrol. Process.* **18** (3), 571–581.
- Ketabchi, H. & Ataie-Ashtiani, B. 2015 Review: coastal groundwater optimization-advances, challenges, and practical solutions. *Hydrogeol. J.* **23** (6), 1129–1154.
- Kolditz, O. & Bauer, S. 2004 A process-orientated approach to compute multi-field problems in porous media. *J. Hydroinform.* **6**, 225–244.
- Kolditz, O. & Shao, H. (Eds) 2010 *Opengeosys: Developer-Benchmark-Book 5.04*. UFZ Publisher, Leipzig, Germany.
- Kolditz, O., Delfs, J.-O., Bürger, C., Beinhorn, M. & Park, C.-H. 2008 Numerical analysis of coupled hydrosystems based on an object-oriented compartment approach. *J. Hydroinform.* **10** (3), 227–244.
- Kolditz, O., Bauer, S., Bilke, L., Böttcher, N., Delfs, J. O., Fischer, T., Görke, U. J., Kalbacher, T., Kosakowski, G., McDermott, C. I., Park, C. H., Radu, F., Rink, K., Shao, H., Shao, H. B., Sun, F., Sun, Y. Y., Singh, A. K., Taron, J., Walther, M., Wang, W., Watanabe, N., Wu, Y., Xie, M., Xu, W. & Zehner, B. 2012 Opengeosys: an open-source initiative for numerical simulation of thermo-hydro-mechanical/chemical (THM/C) processes in porous media. *Environ. Earth Sci.* **67** (2), 589–599.
- Kollat, J. B. & Reed, P. M. 2006 Comparing state-of-the-art evolutionary multi-objective algorithms for long-term groundwater monitoring design. *Adv. Water Resour.* **29** (6), 792–807.
- Kourakos, G. & Mantoglou, A. 2010 Simulation and multi-objective management of coastal aquifers in semi-arid regions. *Water Resour. Manage.* **25** (4), 1063–1074.
- Langevin, C. D. & Guo, W. 2006 MODFLOW/MT3DMS-based simulation of variable-density ground water flow and transport. *Ground Water* **44** (3), 339–351.
- Lin, H.-C. J., Richards, D. & Talbot, C. 1997 *FEMWATER: A Three-Dimensional Finite Element Computer Model for Simulating Density-Dependent Flow and Transport in Variably Saturated Media*. U.S. Army Corps of Engineers, Washington, DC, USA.
- Liriano, S. & Day, R. 2001 Prediction of scour depth at culvert outlets using neural networks. *J. Hydroinform.* **3**, 231–238.
- Liu, F., Turner, I. & Anh, V. 2002 An unstructured mesh finite volume method for modelling saltwater intrusion into coastal aquifers. *Korean J. Comput. Appl. Math.* **9** (2), 391–407.
- MacKay, D. J. C. 1998 Introduction to Gaussian process. *NATO ASI Series F Comp. Syst. Sci.* **168**, 133–166.
- Malik, A. S. & Al-Zubeidi, S. 2006 Electricity tariffs based on long-run marginal costs for central grid system of Oman. *Energy* **31** (12), 1703–1714.
- McDermott, C. I., Walsh, R., Mettler, R., Kosakowski, G. & Kolditz, O. 2009 Hybrid analytical and finite element numerical modeling of mass and heat transport in fractured rocks with matrix diffusion. *Comput. Geosci.* **13** (3), 349–361.
- Möller, M. F. 1993 A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks* **6** (4), 525–533.
- Mousavi, S. J. & Shourian, M. 2010 Adaptive sequentially space-filling metamodeling applied in optimal water quantity allocation at basin scale. *Water Resour. Res.* **46**, W03520.
- Nabney, I. 2002 *NETLAB: Algorithms for Pattern Recognition*. Springer Science & Business Media, London.
- Naish-Guzman, A. & Holden, S. 2007 The generalized FITC approximation. *Advances in Neural Information Processing Systems 20*, NIPS, pp. 1057–1064.
- Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models. Part I – a discussion of principles. *J. Hydrol.* **10** (3), 282–290.
- Nicklow, J., Reed, P., Savic, D., Dessalegne, T., Harrell, L., Chan-Hilton, A., Karamouz, M., Minsker, B., Ostfeld, A., Singh, A., Zechman, E. & ASCE Task Committee on Evolutionary Computation in Environmental and Water Resources Engineering 2010 State of the art for genetic algorithms and beyond in water resources planning and management. *J. Water Resour. Plann. Manage.* **136** (4), 412–432.
- Papadopolou, M. P., Nikolos, I. K. & Karatzas, G. P. 2010 Computational benefits using artificial intelligent methodologies for the solution of an environmental design problem: saltwater intrusion. *Water Sci. Technol.* **62** (7), 1479–1490.
- Peralta, R. C. & Kalwij, I. 2012 *Groundwater Optimization Handbook*. CRC Press, Boca Raton, FL, USA.
- Quinero-Candela, J. & Rasmussen, C. E. 2005 A unifying view of sparse approximate Gaussian process regression. *J. Machine Learn. Res.* **6**, 1939–1959.
- Rao, S., Sreenivasulu, V., Bhallamudi, S., Thandaveswara, B. & Sudheer, K. 2004 Planning groundwater development in coastal aquifers. *Hydrol. Sci. J.* **49** (1), 155–170.
- Rasmussen, C. E. & Williams, C. K. I. 2006 *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, MA, USA.
- Ratto, M., Castelletti, A. & Pagano, A. 2012 Emulation techniques for the reduction and sensitivity analysis of complex environmental models. *Environ. Modell. Softw.* **34**, 1–4.

- Razavi, S., Tolson, B. A. & Burn, D. H. 2012a Numerical assessment of metamodelling strategies in computationally intensive optimization. *Environ. Modell.Softw.* **34**, 67–86.
- Razavi, S., Tolson, B. A. & Burn, D. H. 2012b Review of surrogate modeling in water resources. *Water Resour. Res.* **48**, W07401.
- Schütze, N. & Schmitz, G. H. 2010 OCCASION: new planning tool for optimal climate change adaption strategies in irrigation. *J. Irrig. Drain. Eng.* **136** (12), 836–846.
- Sedki, A. & Ouazar, D. 2011 Swarm intelligence for groundwater management optimization. *J. Hydroinform.* **13** (3), 520–532.
- Seeger, M., Williams, C. & Lawrence, N. 2003 Fast forward selection to speed up sparse Gaussian process regression. In: 9th International Workshop on Artificial Intelligence and Statistics. Society for Artificial Intelligence and Statistics. EPFL-CONF-161318.
- Shourian, M., Mousavi, S., Menhaj, M. & Jabbari, E. 2008 Neural-network-based simulation-optimization model for water allocation planning at basin scale. *J. Hydroinform.* **10** (4), 331–343.
- Singh, T. S. & Chakrabarty, D. 2011 Multiobjective optimization of pump-and-treat-based optimal multilayer aquifer remediation design with flexible remediation time. *J. Hydrol. Eng.* **16** (5), 413–420.
- Smola, A. J. & Bartlett, P. 2000 Sparse greedy Gaussian process regression. Advances in Neural Information Processing Systems 13, NIPS.
- Snelson, E. & Ghahramani, Z. 2006 *Sparse Gaussian Processes Using Pseudo-Inputs*. In: *Advances in Neural Information Processing Systems*. MIT Press, Cambridge, MA, USA, pp. 1257–1264.
- Sreekanth, J. & Datta, B. 2011 Coupled simulation-optimization model for coastal aquifer management using genetic programming-based ensemble surrogate models and multiple-realization optimization. *Water Resour. Res.* **47**, W04516, p. 17.
- Sreekanth, J. & Datta, B. 2015 Review: simulation-optimization models for the management and monitoring of coastal aquifers. *Hydrogeol. J.* **23** (6), 1155–1166.
- Stranger, G. 1985 Coastal salinization: a case history from Oman. *Agric. Water Manage.* **9**, 269–286.
- Voss, C. I. 1984 *A Finite-Element Simulation Model for Saturated-Unsaturated, Fluid-Density-Dependent Groundwater Flow with Energy Transport or Chemically-Reactive Single-Species Solute Transport*. Water Resources Investigation Report 84-4369, USGS, Reston, VA, USA.
- Walther, M., Delfs, J. O., Grundmann, J., Kolditz, O. & Liedl, R. 2012 Saltwater intrusion modeling: verification and application to an agricultural coastal arid region in Oman. *J. Comput. Appl. Math.* **236** (18), 4798–4809.
- Walther, M., Bilke, L., Delfs, J. O., Graf, T., Grundmann, J., Kolditz, O. & Liedl, R. 2014 Assessing the saltwater remediation potential of a three-dimensional, heterogeneous, coastal aquifer system. *Environ. Earth Sci.* **72** (10), 3827–3837.
- Wang, M. & Zheng, C. 1998 Ground water management optimization using genetic algorithms and simulated annealing: formulation and comparison. *J. Am. Water Resour. Assoc.* **34** (3), 519–530.
- WBCSD 2006 *Facts and Trends: Water. Report of the World Business Council for Sustainable Development*, Washington, DC.
- Williams, C. K. I. 2002 Gaussian processes. In: *The Handbook of Brain Theory and Neural Networks*. 2nd edn (M. A. Arbib, ed.). The MIT Press, Cambridge, MA, USA.
- Yan, S. & Minsker, B. 2006 Optimal groundwater remediation design using an adaptive neural network genetic algorithm. *Water Resour. Res.* **42** (5), W05407, p. 14.
- Yan, S. & Minsker, B. 2011 Applying dynamic surrogate models in noisy genetic algorithms to optimize groundwater remediation designs. *J. Water Resour. Plann. Manage.* **137** (3), 284–292.
- Yeh, W. W. G. 2015 Review: optimization methods for groundwater modeling and management. *Hydrogeol. J.* **23** (6), 1051–1065.

First received 27 April 2015; accepted in revised form 12 January 2016. Available online 18 February 2016