AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2014WR016653

Key Points:
- An innovative surrogate modeling approach to water management optimization
- Physically based hydrological modeling and optimization benefit each other
- A solution to the gap between complex modeling and real-world decision making

Supporting Information:
- Supporting Information S1

Correspondence to:
Y. Zheng,
yizheng@pku.edu.cn

Citation:
Wu, B., Y. Zheng, X. Wu, Y. Tian, F. Han, J. Liu, and C. Zheng (2015), Optimizing water resources management in large river basins with integrated surface water-groundwater modeling: A surrogate-based approach, *Water Resour. Res.*, *51*, 2153–2173, doi:10.1002/2014WR016653.

Received 10 NOV 2014
Accepted 5 MAR 2015
Accepted article online 10 MAR 2015
Published online 9 APR 2015

Optimizing water resources management in large river basins with integrated surface water-groundwater modeling: A surrogate-based approach

Bin Wu[1,2], Yi Zheng[1,2], Xin Wu[1,2], Yong Tian[1,2], Feng Han[1,2], Jie Liu[1,2], and Chunmiao Zheng[1,2,3]

[1]Department of Energy and Resources Engineering, College of Engineering, Peking University, Beijing, China, [2]Institute of Water Sciences, Peking University, Beijing, China, [3]Department of Geological Sciences, University of Alabama, Tuscaloosa, Alabama, USA

Abstract Integrated surface water-groundwater modeling can provide a comprehensive and coherent understanding on basin-scale water cycle, but its high computational cost has impeded its application in real-world management. This study developed a new surrogate-based approach, SOIM (Surrogate-based Optimization for Integrated surface water-groundwater Modeling), to incorporate the integrated modeling into water management optimization. Its applicability and advantages were evaluated and validated through an optimization research on the conjunctive use of surface water (SW) and groundwater (GW) for irrigation in a semiarid region in northwest China. GSFLOW, an integrated SW-GW model developed by USGS, was employed. The study results show that, due to the strong and complicated SW-GW interactions, basin-scale water saving could be achieved by spatially optimizing the ratios of groundwater use in different irrigation districts. The water-saving potential essentially stems from the reduction of nonbeneficial evapotranspiration from the aqueduct system and shallow groundwater, and its magnitude largely depends on both water management schemes and hydrological conditions. Important implications for water resources management in general include: first, environmental flow regulation needs to take into account interannual variation of hydrological conditions, as well as spatial complexity of SW-GW interactions; and second, to resolve water use conflicts between upper stream and lower stream, a system approach is highly desired to reflect ecological, economic, and social concerns in water management decisions. Overall, this study highlights that surrogate-based approaches like SOIM represent a promising solution to filling the gap between complex environmental modeling and real-world management decision-making.

1. Introduction

In many arid and semiarid areas, both surface water (SW) and groundwater (GW) resources are vital to sustainability of local society and ecosystem [*Schmidt et al*., 2013; *Shanafield et al*., 2012; *Simpson et al*., 2013]. To better manage the limited water resources in such areas, a systematic understanding on the integrated SW-GW system is highly desired. However, it would be very difficult to achieve the understanding solely based on field observations, especially for large river basins. Integrated SW-GW modeling can fuse theoretical knowledge with field observations, and provide a spatially and temporally detailed description of the hydrological cycle in a large basin [*Wu et al*., 2014]. A number of integrated SW-GW models have been developed, such as Hydrogeosphere [*Brunner and Simmons*, 2012], MIKE-SHE [*Graham and Refsgaard*, 2001], MODHMS [*Panday and Huyakorn*, 2004], ParFlow [*Kollet and Maxwell*, 2006], CATHY [*Weill et al*., 2011], PAWS+CLM [*Shen et al*., 2013], GEOtop [*Rigon et al*., 2006], and GSFLOW [*Doherty and Hunt*, 2009; *Markstrom et al*., 2008; *Surfleet and Tullos*, 2013; *Tian et al*., 2015]. The integrated models have been used to address a variety of water and environmental issues, such as impact of climate change [*Surfleet and Tullos*, 2013], water quality [*Fonseca et al*., 2014; *Wilby et al*., 2006], and basin-scale water resources management [*Hassanzadeh et al*., 2014; *Mazzega et al*., 2014].

Optimization has been widely practiced in developing water resources management plans [*Koech et al*., 2014; *Molinos-Senante et al*., 2014; *Sreekanth and Datta*, 2011]. A variety of optimization algorithms have been proposed, among which Genetic Algorithm (GA) [*Deb et al*., 2002], Shuffled Complex Evolution (SCE-UA) [*Duan et al*., 1992], PEST [*Doherty*, 1994], NSGA-II [*Deb et al*., 2002], Differential Evolution [*Price et al*.,

2005], Dynamically Dimensioned Search (DDS) [*Tolson and Shoemaker*, 2007], and stochastic Radial Basis Function (stochastic RBF) [*Regis and Shoemaker*, 2007] are frequently used in water resources research. More can be found in the review by *Maier et al*. [2014]. For complex optimization problems, heuristic approaches are less affected by local optima and more suitable than gradient-based approaches, but require much more evaluations of objective function. If an integrated SW-GW model (one model evaluation typically takes hours to days) is involved, the computational cost could be forbidding for a single heuristic search. How to perform an efficient and effective optimization in such cases has been addressed by few studies. *Condon and Maxwell* [2013] implemented a linear optimization water allocation algorithm into ParFlow. The objective was to optimize the pumping and diversion schedule to meet the irrigation water demand. In their specific case, the optimization needed no iterative evaluations of ParFlow, and therefore the computational cost was not a major concern. *Pisinaras et al*. [2013] coupled SWAT [*Arnold et al*., 1998] and MODFLOW [*Harbaugh*, 2005] to study seawater intrusion. The objective was to maximize the pumping while adequately preventing the aquifer from seawater intrusion. However, the model coupling is one-way, that is, SWAT provides the upper boundary condition for MODFLOW but receives no feedback from it. A general framework and innovative approaches are yet to be developed to optimize water recourses management based on a fully integrated SW-GW model.

Surrogate-based optimization approaches are promising for computationally expensive models [*Cai et al*., 2015; *Maier et al*., 2014; *Razavi et al*., 2012a, 2012b; *Wang et al*., 2014]. The core idea of the approaches is to replace the original complex model by a simple and computationally cheap surrogate model in the optimization process. A surrogate model can be regarded as a response surface between concerned model output variables and input variables, which emulates the original model in certain aspects. Some representative types of surrogate models include Probabilistic Collocation Method (PCM) [*Fen et al*., 2009; *Zheng et al*., 2011], Kriging [*Baú and Mayer*, 2006], Support Vector Machine (SVM) [*Zhang et al*., 2009], and RBF [*Regis and Shoemaker*, 2007]. There exist two typical strategies to perform a surrogate-based optimization. One is the batch approach which establishes a satisfactory surrogate model (i.e., response surface) before the optimization starts, and uses it throughout the optimization process [*Johnson and Rogers*, 2000; *Liong et al*., 2001]. It usually needs a large training set to guarantee a good substitution. The other is the adaptive approach which uses a small training set to build an initial surrogate model (most likely unsatisfactory) and iteratively updates it as the search for optima proceeds [*Ostfeld and Salomons*, 2005; *Regis and Shoemaker*, 2007]. The surrogate model guides the search, and each of its updates requires some additional training points. Surrogate-based optimization has been applied to hydrological models. For example, *Johnson and Rogers* [2000] used a linear regression model to replace MODFLOW in addressing a groundwater remediation problem; *Ostfeld and Salomons* [2005] calibrated the CE-QUAL-W2 water quality model using k-Nearest Neighbor algorithm (kNN), an adaptive surrogate-based approach. *Mugunthan and Shoemaker* [2006] implemented stochastic RBF, another surrogate-based approach, to perform automatic parameter calibration and uncertainty analysis for MODFLOW. However, to the authors' best knowledge, surrogate-based optimization has not been attempted for fully integrated SW-GW models in the context of water resources management.

This study developed a surrogate-based approach for optimizing water resources management based on the integrated modeling. It incorporates SVM [*Vapnik*, 1998] as the response surface and SCE-UA as the searching algorithm. Irrigation water management in Zhangye Basin (ZB), a semiarid to arid inland area in Gansu Province, China, was selected as the case study. ZB is the midstream section of the Heihe River Basin (HRB), China's second largest inland river basin. The irrigated agriculture in ZB's oases consumes a large amount of surface water diverted from the river, as well as groundwater pumped from the local aquifer [*Wu et al*., 2014]. The water diversion has significantly reduced the outflow from ZB to the lower HRB, which caused many environmental problems like desertification [*Ji et al*., 2006], salinization [*Wen et al*., 2005], and shrinkage of the terminal lake (i.e., East Juyan Lake) [*Guo et al*., 2009]. On the other hand, the excessive use of groundwater has led to the decline of groundwater level, posing a risk to the highly water-dependent society and ecosystem in ZB. Thus, it is imperative to improve the water management in this region, such that the growing conflict between agricultural productivity and environmental health could be possibly resolved.

The proposed approach was applied to optimize the partition of surface water and groundwater in the total irrigation water of ZB's irrigation districts, such that a better balance between the groundwater storage in the middle HRB and the environmental flow toward the lower HRB can be achieved. Through a series of

numerical experiments and benchmarking with another well-established approach, the applicability and advantages of our new approach were validated. The study results demonstrate that the integrated SW-GW modeling can provide great flexibility to formulation of water management optimization problems, and enhance the interpretability of the optimization results, since a model like GSFLOW can comprehensively simulate the hydrological variables in basin-scale water cycle with high spatial and temporal resolutions. The surrogate-based approach paves the path for embedding such complex modeling into the optimization. Management implications from this study include: first of all, environmental flow regulation needs to take into account the interannual variation of hydrologic conditions, as well as the spatial complexity of SW-GW interactions. Second, to resolve water use conflicts between upper stream and lower stream, a system approach is highly desired to adequately reflect ecological, economic, and social concerns in water management decisions.

The remaining part of this paper is organized as follows. Section 2 introduces the surrogate-based optimization approach. Section 3 introduces the study area and the integrated SW-GW model, formulates the optimization problems, and explains the numerical experiments. Section 4 presents the results and discussion, and section 5 concludes this study.

2. The Optimization Approach

2.1. Framework

Figure 1 is a flowchart of the proposed approach, which is referred to as SOIM (Surrogate-based Optimization for Integrated surface water-groundwater Modeling) hereafter. As demonstrated, once the optimization problem is defined and the integrated SW-GW model is established, the following steps can be taken in sequence.

1. Select *N* initial training points: *N* realizations of the decision variables (e.g., pumping ratios) are first sampled to train the surrogate model. To ensure a relatively even coverage over the decision space, Latin Hypercube Sampling (LHS) [*McKay et al*., 1979; *Stein*, 1987] is suggested to be the sampling scheme. *N* should be a relatively small number, such that the initial training would be computationally affordable.
2. Evaluate the SW-GW model: At the *N* points, the original model is run to simulate concerned model output variables (e.g., streamflow, groundwater level, etc.).
3. Build initial surrogate models: For each concerned output variable, a surrogate model is built to emulate its response to the decision variables. The mathematical nature of surrogate model is critical to the efficiency and effectiveness of the optimization, and support vector machine (SVM) is recommended here. As will be discussed later, SVM can generate good emulators for integrated SW-GW models with a small number of training points. The performance of the trained surrogate models can be further assessed with extra validating points that are also randomly picked. If the surrogates are not satisfactory, return to step (1) and increase *N*.
4. Perform the optimization based on the trained surrogate models: An effective heuristic searching algorithm is needed in addressing highly nonlinear problems, and we suggest SCE-UA whose applicability to hydrological modeling has been well proven [*Dakhlaoui et al*., 2012; *Duan et al*., 1992; *Khakbaz et al*., 2012]. In this step, SOIM operates *M* repeats of the surrogate-based optimization, each with a different random seed, and therefore produces *M* optimal points. Note that *M* should be a small number (e.g., several to a couple of tens) to keep the computational cost of the next step low.
5. Evaluate the original model at the *M* optimal points: A convergence check follows, which judges whether the current surrogate models need further improvement. Three criteria are considered. First, the objective function values, calculated with the origin model, of the *M* optimal points converge. Specifically, our approach requires that the distance between each of the *M* values and their centroid be less than a predefined small value e_1. Second, at each of the *M* optimal points, the difference between the objective function value calculated with the original model and that calculated with the surrogate models does not exceed a predefined small value e_2. Third, the repeats of step (6) reaches a predefined number T. If the first two criteria are both met, or the third one is met, proceed to step (7); otherwise, go to step (6). This convergence-check step guarantees that the surrogate models accurately approximate the original model at the *M* intermediate optimal points.

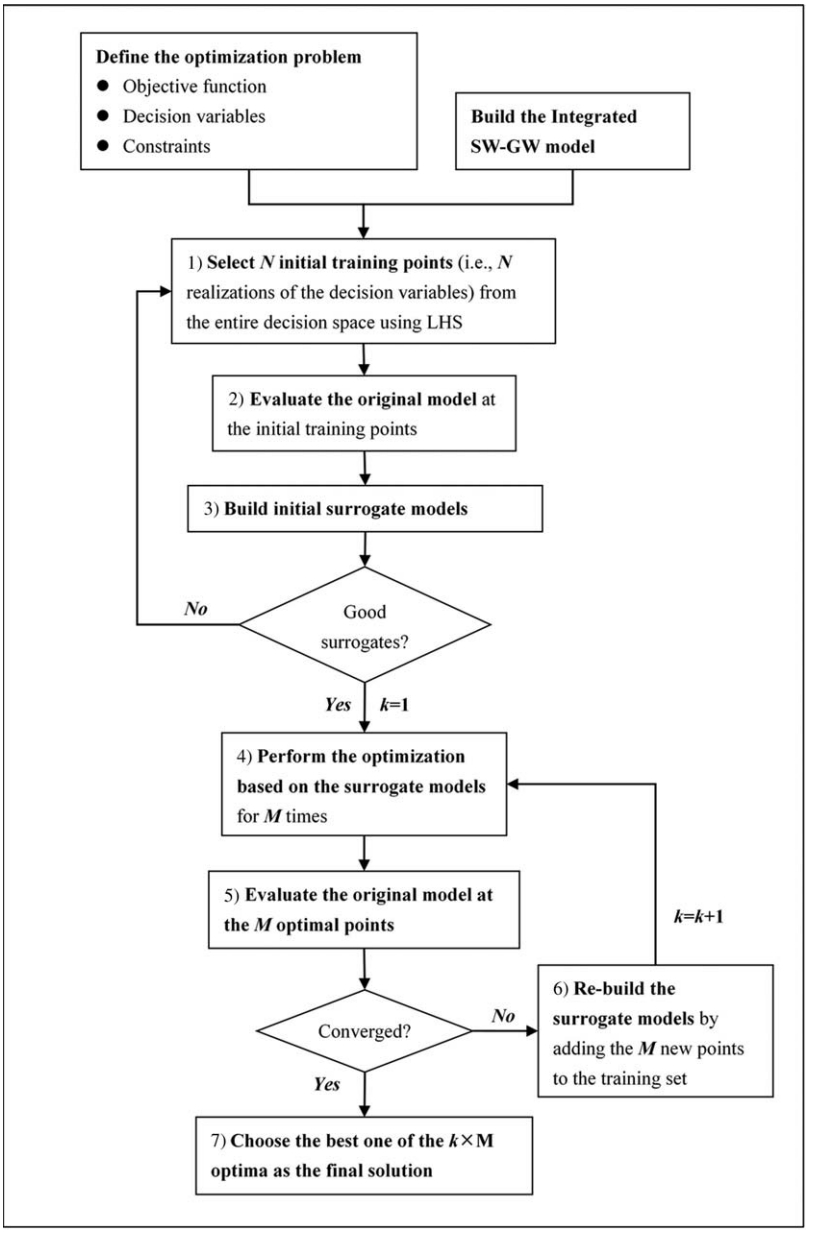

Figure 1. Framework of Surrogate-based Optimization for Integrated surface water-groundwater Modeling (SOIM).

6. Retrain the surrogate models: If the convergence is not reached, the M optima identified in step (5) are added to the training set, and the surrogate models are re-trained. This is a critical step in SOIM, whose rationale is as follows. Although the surrogate models in step (4) perform well at those randomly picked validating points, it is not guaranteed that they perform as well in the near-optimal regions of the decision space, because SOIM considers a relatively small N in steps (1)–(3). By introducing the optima as the additional training points, the performance of the surrogate models in the near-optimal regions should be significantly improved.
7. Choose the best one of the $k \times M$ optima as the final solution: SOIM innovatively incorporates SVM (an effective surrogate modeling technique) and SCE-UA (a well-established heuristic searching algorithm) to cope with integrated SW-GW modeling. It lies between typical batch approaches and typical adaptive approaches, because it requires acceptable surrogate models at the very beginning while still updates

the surrogate models iteratively. The *M*-repeat surrogate-based optimization in step (4) and convergence check in step (5) represent a unique quality assurance measure which existing surrogate-based approaches lack. This measure guarantees that the surrogate models can accurately approximate the original model at those intermediate optimal points, and therefore the search would not be misled by inaccurate surrogate models. A recent study by *Cai et al*. [2015] applied SVM to complex modeling-based water management optimization in the batch way. Compared to typical adaptive approaches and SOIM, the batch approach can generate more globally reliable surrogate models, since a much bigger training set is used. Its advantage is evident in cases when a large number of different optimization problems can utilize the same surrogate models. Nevertheless, our SOIM has its own advantages. First, if a limited number (say, several tens) of optimization problems are to be addressed, the total number of original model evaluations required by SOIM can be lower than the batch approach. Second, as discussed above, SOIM has an intrinsic quality assurance mechanism that the batch approach does not have. On the other hand, one major advantage of SOIM over typical adaptive approaches is that the initial surrogate models built in one optimization analysis could be reusable in others, such that the overall computational cost can be further reduced. It is also worth pointing out that the framework of SOIM (Figure 1) allows other suitable combinations of surrogate modeling techniques and searching algorithms. This study adopted SVM and SCE-UA because this combination was proved to be effective and efficient for the integrated SW-GW modeling.

2.2. Brief Introduction to SVM and SCE-UA

SVM [*Vapnik*, 1998] is a statistical learning method used for classification and regression. It has been applied to various environmental models [*Guo et al*., 2005; *Lin et al*., 2006; *Liong and Sivapragasam*, 2002; *Zhang et al*., 2009], but its applications to integrated SW-GW models have not been reported. In SOIM, we adopted the regression form of SVM to build surrogate models, which is:

$$f(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x} \rangle + b \tag{1}$$

where $\mathbf{x}$ is the input variable vector of training points; $f(\mathbf{x})$ represents the surrogate model; $\mathbf{w}$ and b are coefficients to be determined; and $< \cdot, \cdot >$ is the inner product operator. In SVM, $\mathbf{w}$ and b are determined based on the novel Structural Risk Minimization principle [*Shawe-Taylor et al*., 1998]:

$$\begin{aligned} &\min \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right) \\ &s.t. \\ &y_i - (\langle \mathbf{w}, x_i \rangle + b) \le \varepsilon + \xi_i \\ &(\langle \mathbf{w}, x_i \rangle + b) - y_i \le \varepsilon + \xi_i^* \\ &\xi_i^*, \xi_i \ge 0 \end{aligned} \tag{2}$$

C is a constant weight to balance the trade-off between the two items in the objective function. n is the number of training points; y_i is the original model's output for the ith training point; ε and ξ's are parameters that defined the feasible and infeasible regions. The constraints in equation (2) require the residuals, $y_i - (\langle \mathbf{w}, x_i \rangle + b)$, are within the range $[-\varepsilon - \xi_i^*, \varepsilon + \xi_i]$. Once the SVM surrogate is built up, for a target point x^*, it can be computed that:

$$f(x^*) = \langle \mathbf{w}(\mathbf{x}, \mathbf{y}), x^* \rangle + b \tag{3}$$

In nonlinear problems, equation (3) can be replaced by equation (4):

$$f(x^*) = \varphi(\mathbf{w}(\mathbf{x}, \mathbf{y}), x^*) + b \tag{4}$$

where φ is a kernel function which maps **w(x, y)** and x^* to a nonlinear space. More technical details about SVM can be found in *Burges* [1998].

SCE-UA [*Duan et al*., 1992] is a widely used global optimization technique developed in the field of hydrological modeling. It inherits the features of GA and Simplex downhill search, and incorporates the complex shuffling concept. In brief, SCE-UA divides the initial population into several complexes (i.e., subgroups), and each complex evolves separately, but is shuffled (i.e., remixed) with others after each iteration of

evolution. The evolutionary algorithm is called competitive complex evolution (CEE), which adopts reflection, contraction, and mutation operations to search for the next generation. More technical details about SCE-UA can be found in *Duan et al*. [1992].

3. Data and Methods

3.1. Study Area

The Zhangye Basin (ZB) is the midstream part of Heihe River Basin (HRB), the second largest inland river basin in China. ZB is a semiarid to arid area typical of Northwest China. Our modeling domain (Figure 2a)

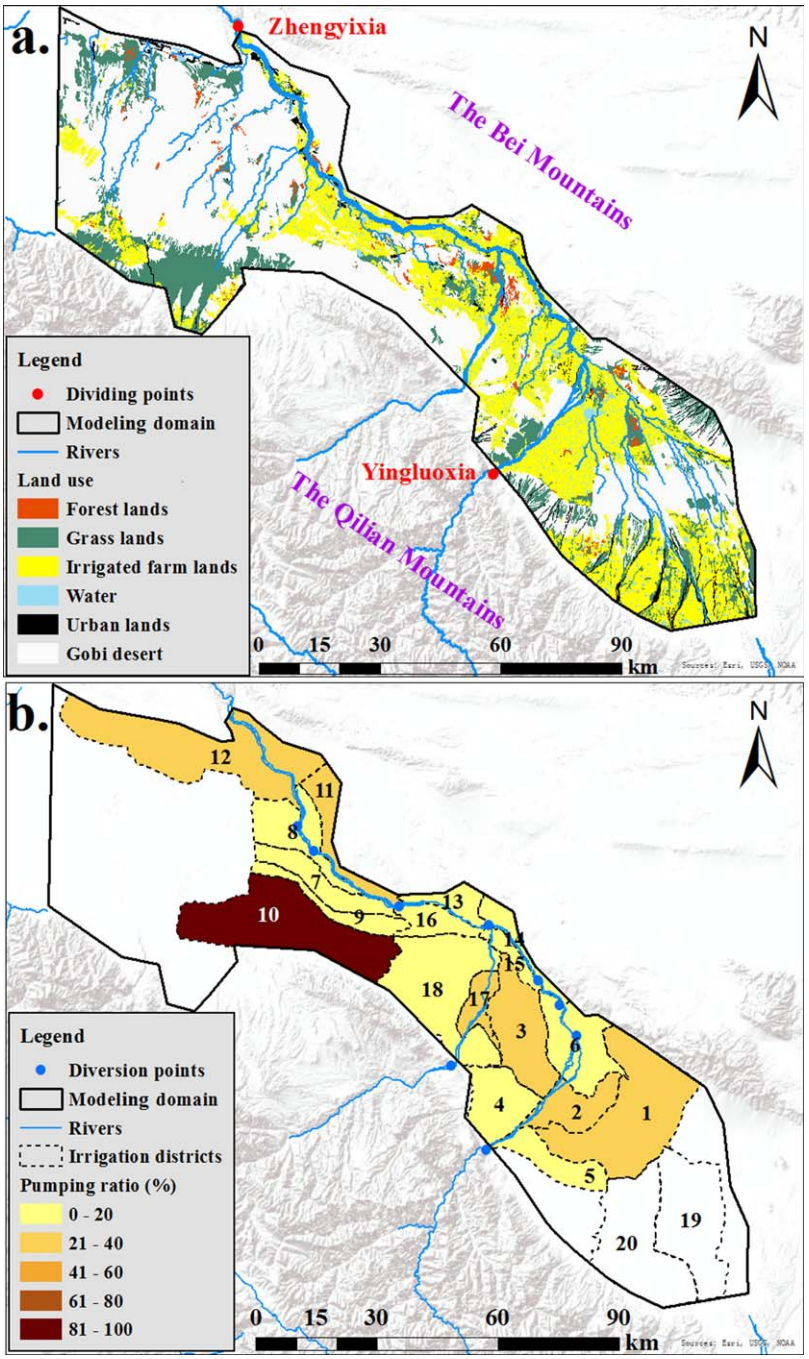

Figure 2. Study area. (a) The land use pattern within the modeling domain; (b): locations of the irrigation districts and their average pumping ratios (i.e., proportions of groundwater in total irrigation water) in 2000–2008.

has an area of 9097 km^2, and is bounded by the Qilian Mountains on the south, the Bei Mountains on the north, the Jiuquan-west Basin on the west, and the Maying Basin on the east [*Wu et al*., 2014]. The major land use types are irrigated farm lands in oases and Gobi desert with very poor vegetation. Corn and winter-wheat are the main crop types. The elevation of this area increases from 1290m in the northwest to 2200m in the southeast.

The average annual rainfall is about 190 mm, mainly in June–September, and the average annual potential evapotranspiration (ET) was estimated to be 1325 mm [*Wu et al*., 2010]. It has been estimated that, on average, 2.22 billion m^3/yr of surface water flows into ZB, mostly from the Qilian Mountains, and 1.72 billion m^3/yr is delivered by the Heihe River through Yingluoxia (the dividing point between the upper stream and midstream) [*Wu et al*., 2014]. The local precipitation mainly converts to ET and infiltration, generating limited surface runoff. The surface outflow is mainly through Zhengyixia (the dividing point between the midstream and lower stream) and averages 0.97 billion m^3/yr (2000–2008). The interactions between surface water and groundwater in this area are significant and complicated [*Hu et al*., 2007]. *Wu et al*. [2014] estimated that the GW to SW flux and SW to GW flux are 1.06 billion m^3/yr and 0.97 billion m^3/yr, respectively. The interactions (i.e., groundwater recharge and discharge) mainly occur in the river domain. In irrigated farm lands, groundwater is also recharged by the infiltrated irrigation water.

The agriculture in ZB takes about 1.89 billion m^3/yr (based on the data of 2000–2008), 82% of which is surface water diverted from the Heihe River, and the rest is pumped groundwater. There are 20 irrigation districts, as labeled in Figure 2b, and all except No.19 and No. 20 use both surface water and groundwater for irrigation. Before 2000, the luxurious agricultural water use in ZB had caused fast degradation of ecosystem in the lower HRB and shrinkage of the terminal lake. To protect the unique but fragile ecohydrological system of the lower HRB, a water allocation plan has been strictly enforced by the central government since 2000. The plan specified the amounts of environmental flow through Zhengyixia to be secured under different hydrologic conditions. In a normal year (i.e., 1.58 billion m^3 inflow at Yingluoxia), the outflow from Zhengyixia should be no less than 0.95 billion m^3. The regulation reduced the diversion and made more surface water available to the lower HRB. However, the reduction has led to a significant increase of groundwater pumping, as the pumping is largely unregulated in the study area [*Tian et al*., 2015]. The problem of groundwater overexploitation, as well as the companying ecological issues (e.g., wetland loss, vegetation degradation, etc.), is now looming in certain parts of ZB. Essentially, there may exist a trade-off between the recovery of the ecosystem in the lower HRB and the health of the ecosystem in the middle HRB.

3.2. The GSFLOW Model for the Zhangye Basin

GSFLOW is a typical integrated SW-GW model developed by USGS, which couples the 2-D surface hydrology model PRMS with the 3-D groundwater flow model MODFLOW [*Markstrom et al*., 2008]. In GSFLOW, between the soil zone and saturated zone, a vadose zone is conceptualized and handled by the Unsaturated Zone Flow package (UZF1). The surface domain of GSFLOW uses Hydrologic Response Units (HRUs), either regular grids or irregular polygons, as the basic computing units, and the subsurface domain is discretized into finite difference girds. GSFLOW defines "gravity reservoir", a storage in which an HRU exchanges water with the MODFLOW grid(s) it intersects. The exchange is two-way: when the groundwater head is below a threshold level, the "gravity reservoir" releases water to the vadose zone, and the water may eventually reach the saturated zone; and when the head exceeds the threshold level, the saturated zone discharges water to the "gravity reservoir", from which water can be further transformed to soil moisture, interflow, or overland flow. Streams are simulated using the Streamflow Routing package (SFR2). Stream-aquifer exchange is calculated using Darcy's law, based on the head difference between a reach and the subsurface grid(s) it intersects. More details about GSFLOW and its components can be found in *Markstrom et al*. [2008].

In our previous study [*Wu et al*., 2014], a GSFLOW model was established for the modeling domain (Figure 2a), which adequately simulates the hydrological processes and water budget in the study area. The simulation period is from 2000 to 2008, 9 years in total. As the present official version (v1.1.6) of GSFLOW does not directly simulate diversion and irrigation, an alternative way was used to represent them in the model simulation. More details about the model setup and calibration can be found in *Wu et al*. [2014].

3.3. Optimization Schemes

Because agriculture (mainly corn planting) is the major industry in this relatively undeveloped area and the local population keeps increasing, the demand for irrigation water is hard to be reduced in the short term. Thus, as the groundwater exploitation has not been well regulated, the flow conservation by reducing the diversion has to be at the cost of groundwater depletion. On the other hand, the irrigation efficiencies of surface water diversion and groundwater pumping are significantly different in this area. The river water is delivered by aqueducts, often poorly lined, over a long distance before it reaches the field, and a significant amount of water is lost to ET and aqueduct seepage. But a pumping well usually supplies water to its adjacent field, and the ET and seepage losses are much smaller. Therefore, it is well expected that the groundwater drawdown would be reduced if the partition of surface water and groundwater in the irrigation water could be spatially optimized. This study applied SOIM to address this optimization problem which is critical to the local water resources management. The problem specification is introduced below.

The decision variables (**X**) are proportions of groundwater in total irrigation water (i.e., pumping ratios). **X** is an 18-dimension vector, as the No. 19 and No.20 districts have ignorable pumping. Its ith element x_i ($i=1, 2,\ldots, 18$) corresponds to the ith irrigation district (Figure 2b). To simplify the optimization, x_i was assumed to be time-invariant. The groundwater pumped ($G_{i,t}$) and surface water diverted ($D_{i,t}$) for the ith district in the tth day can be expressed as:

$$G_{i,t}=F_{i,t}\times x_i \tag{5}$$

$$D_{i,t}=\frac{F_{i,t}\times(1-x_i)}{r_i} \tag{6}$$

where $F_{i,t}$ denotes the irrigation demand with the water losses in the aqueduct system excluded, and r_i stands for the time-invariant water efficiency of the aqueduct system, which is the ratio between the water volume entering the field and that taken from the river. Equation (5) assumes that the loss of pumped groundwater before it reaches the field is negligible, which adequately reflects the real situation. In the optimization, $G_{i,t}$ and $D_{i,t}$ vary with x_i, while $F_{i,t}$ and r_i are externally determined. r_i values were adopted from reports by the local water resources authority. $F_{i,t}$ was estimated as $D^*_{i,t}\cdot r_i+G^*_{i,t}$, where $D^*_{i,t}$ and $G^*_{i,t}$, respectively, represent the daily diversion and pumping rates that we complied as the inputs for the original GSFLOW model. This is equivalent to assuming that the demand was well satisfied historically and no water-saving strategies are to be implemented. Note that r_i is an important parameter and involves uncertainty. In this study, we used the best available data to set its values. A separate investigation may be necessary to explore how the uncertainty of r_i would impact the optimization results.

We considered two main model output variables, annual total outflow through Zhengyixia (denoted as R) and annual total change of regional groundwater storage (denoted as ΔS). GSFLOW computes the total groundwater storage in modeling domain at a daily time step, based on simulated groundwater heads and hydrogeological parameters (i.e., specific yield and/or specific storage), and ΔS is the difference in the storage between the last day and the first day of a year. A negative ΔS indicates a decrease of the storage. Two optimization schemes, scheme A and scheme B, were hypothesized. In scheme A, the objective is to maximize R, provided a constraint on ΔS. In scheme B, the objective is to maximize ΔS, given a constraint on R. Occasionally, the river flow may be insufficient for potential diversion during some low flow periods. Although GSFLOW can proceed in this situation anyway, numerical errors will be introduced to the water budget. To avoid significant numerical errors, in both schemes, we have additional constraints to ensure that the potential diversion at each of the nine diversion points (see Figure 2b) does not exceed the available river flow at a monthly scale. Scheme A is formulated as:

$$\begin{aligned}&\max_{\mathbf{X}} R(\mathbf{X})\\ &s.t.\\ &\Delta S \geq S_f\\ &Q_{j,m} \geq D_{j,m}\end{aligned} \tag{7}$$

where S_f represents a preset tolerance for the change of regional groundwater storage, usually a negative value. $Q_{j,m}$ represents flow rate at the jth ($j=1, 2,\ldots, 9$) diversion point (Figure 2b) in the mth month, before the potential diversion (denoted as $D_{j,m}$) occurs. Scheme B is formulated as:

$$\begin{aligned} &\max_{\mathbf{X}} \ \Delta S(\mathbf{X}) \\ &s.t. \\ &R \geq R_f \\ &Q_{j,m} \geq D_{j,m} \end{aligned} \tag{8}$$

where R_f represents the minimal outflow required at Zhengyixia. Scheme B mimics the present water regulation regime in the study area.

By introducing punishment functions, the optimization problems can be further transformed into a no-constraint form, $\max_{\mathbf{X}} h(\mathbf{X})$. For scheme A, now the objective function is

$$h(\mathbf{X}) = R(\mathbf{X}) + P(\Delta S) + \sum_j \sum_m P\left(D_{j,m}\right) \tag{9}$$

In equation (9), $P(\Delta S)$ and $P(D_{j,m})$ are punishment functions (both are nonpositive) in the form of equations (10) and (11), respectively.

$$P(\Delta S) = \begin{cases} 0 & \Delta S \geq S_f \\ (R_{\max} - R_{\min}) \cdot (\Delta S - S_f) & \Delta S < S_f \end{cases} \tag{10}$$

$$P\left(D_{j,m}\right) = \begin{cases} 0 & Q_{j,m} \geq D_{j,m} \\ (\Delta S_{\max} - \Delta S_{\min}) \cdot \left(Q_{j,m} - D_{j,m}\right) & Q_{j,m} < D_{j,m} \end{cases} \tag{11}$$

Here $R_{\max}$ and $R_{\min}$ are maximal and minimal simulated (with the original model) annual outflow of the initial training set, respectively; and $\Delta S_{\max}$ and $\Delta S_{\min}$ are maximal and minimal simulated annual storage changes, respectively. Similarly, for scheme B, the transformed objective function is

$$h(\mathbf{X}) = \Delta S(\mathbf{X}) + P(R) + \sum_j \sum_m P\left(D_{j,m}\right) \tag{12}$$

where $P(R)$ is another punishment function (nonpositive) in the following form

$$P(R) = \begin{cases} 0 & R \geq R_f \\ (\Delta S_{\max} - \Delta S_{\min}) \cdot (R - R_f) & R < R_f \end{cases} \tag{13}$$

Using $(R_{\max} - R_{\min})$ and $(\Delta S_{\max} - \Delta S_{\min})$ as the multipliers in the punishment functions ensures that three punishment functions (i.e., equations (10), (11), and (13)) are comparable in magnitude. In addition, if the constraints are not met, the punishment functions will be very large negative values, with regard to the magnitudes of R and ΔS. In this case, the optimization would give the priority to meeting the constraints, rather than improving the untransformed objective functions. Thus, scheme A favors the water resources conservation for the middle HRB, while scheme B favors that for the lower HRB.

3.4. Specification of SOIM and Benchmarking Analysis

Like any other optimization approaches, SOIM's speed and performance depend on its own parameters. In our case study, the number of initial training points was set to be $N=200$, and additional 100 points were selected as the validating points. Tests showed that this N value led to reliable optimization results while keeping the computational cost low. To build SVM surrogates, equation (14) was considered as the kernel function [*Burges*, 1998]

$$\varphi(x^*, x_i) = \exp\left(-\gamma \cdot |x^* - x_i|^2\right) \tag{14}$$

where x^* is a target point to be evaluated and x_i is the ith training point. For both of the schemes, surrogate models were needed for R, ΔS, and $Q_{j,m}$. We used the MATLAB code of SVM developed by *Fan et al.* [2005]. In this study, it has been found that the performance of SVM is sensitive to γ in equation (14), but not to other hyperparameters (e.g., C in equation (2)). Therefore, the value of γ was optimized within 0.01 to 1 (with a constant increment of 0.01) for individual SVM surrogates, while all other hyperparameters of SVM

Table 1. Design of the Numerical Experiments

Inflow Conditions		Actual Situations (Unit: Billion m^3)		Constraints (Unit: Billion m^3)				
				Scheme A (Objective: Max R)		Scheme B (Objective: Max ΔS)		
Year	Water Stress[a]	R	ΔS	Scenario A1 (No Storage Decline)	Scenario A2 (Moderate Storage Decline)	Scenario B1 (No Flow Decrease)	Scenario B2 (Normal Year)	Scenario B3 (Water Allocation Curve)
2000	35%	0.717	−0.143	$\Delta S > 0$	$\Delta S > -0.15$	$R > 0.717$	$R > 0.95$	$R > 0.818$
2001	47%	0.69	−0.391			$R > 0.69$		$R > 0.640$
2002	35%	0.967	−0.248			$R > 0.967$		$R > 0.992$
2003	27%	1.192	−0.024			$R > 1.192$		$R > 1.310$
2004	35%	0.819	−0.156			$R > 0.819$		$R > 0.871$
2005	32%	0.985	−0.104			$R > 0.985$		$R > 1.220$
2006	30%	1.104	−0.139			$R > 1.104$		$R > 1.216$
2007	25%	1.186	0.083			$R > 1.186$		$R > 1.531$
2008	29%	1.089	−0.028			$R > 1.089$		$R > 1.363$

[a]It is defined as the proportion of annual irrigation demand in the total water input. The total water input refers to the boundary surface water inflow from the mainstream and tributaries plus the local precipitation.

were set default. Compared to GSFLOW evaluations, the cost for determining and evaluating these surrogate models is trivial.

In step (4) of SOIM (Figure 1), SCE-UA was repeated 10 times (i.e., $M=10$). We used the MATLAB code developed by *Duan et al*. [1992]. The upper limit of objective function evaluation was set to 10,000 times. All other parameters of SEC-UA were set default.

The three criteria for convergence check after step (5) were specified as $e_1=1\times10^6$ m^3, $e_2=1\times10^6$ m^3, and $T=10$. Note that e_1 and e_2 represent two small quantities, since the magnitudes of simulated R and ΔS are in the order of 10^8 to10^9 m^3. Thus, in a single optimization calculation, at most 300 (i.e., $N+M\times T$) runs of GSFLOW were executed.

For the benchmarking purpose, we also applied an established approach, DYCORS (DYnamic COordinate search using Response Surface models) [*Regis and Shoemaker*, 2013], to the same optimization problems. DYCORS can be regarded as an upgraded version of stochastic RBF [*Regis and Shoemaker*, 2007], which couples dynamic coordinate search with stochastic RBF. DYCORS better balances global search and local search in generating new points, which would lead to better optimization results. In applying DYCORS, the maximum number of GSFLOW runs were set to 300, such that the comparison between DYCORS and SOIM can be fair. Note that using a small number (the default value is 37 that was also adopted by this study) of initial samples reflects the basic philosophy of DYCORS, that is, surrogate models are used to guide the search, rather than to fully replace the original model. In fact, using a larger number of initial samples may reduce the efficiency and effectiveness of DYCORS. A comparison experiment has clearly demonstrated that DYCORS with 37 initial samples actually performs better than DYCORS with 200 initial samples in scenario B1 (see supporting information Figure S8).

3.5. Numerical Experiments

In the two schemes, five optimization scenarios were designed (Table 1). In scenario A1, the constraint is $\Delta S > 0$ billion m^3, which means no storage decline would be allowed. Scenario A1 represents a very strict protection on the local groundwater resource and is referred to as "no-storage-decline scenario." In scenario A2, the constraint is $\Delta S > -0.15$ billion m^3. In 2000–2008, the annual ΔS ranged from −0.36 to 0.08 billion m^3. Hence, scenario A2 represents a "moderate-storage-decline scenario."

The three scenarios in scheme B differ in their constraints as well. The constraint in scenario B1 requires that R be no lower than the actual outflow at Zhengyixia, and the scenario is therefore referred to as "no-flow-decrease scenario." The constraints in scenarios B2 and B3 were both based on the existing "water allocation curve (WAC)" (Figure 3). WAC was designed to regulate the flow in the mainstream of Heihe River, which specifies the bottom line outflow at Zhengyixia under different inflow conditions at Yingluoxia [*Liu and Wang*, 2012]. WAC was endorsed by China's central government in 1997, but has been strictly enforced only since 2000. In fact, WAC is not a continuous curve, but consists of discrete points, as indicated by the circles in Figure 3. For example, according to WAC, in a normal year (i.e., average condition, with the inflow from Yingluoxia

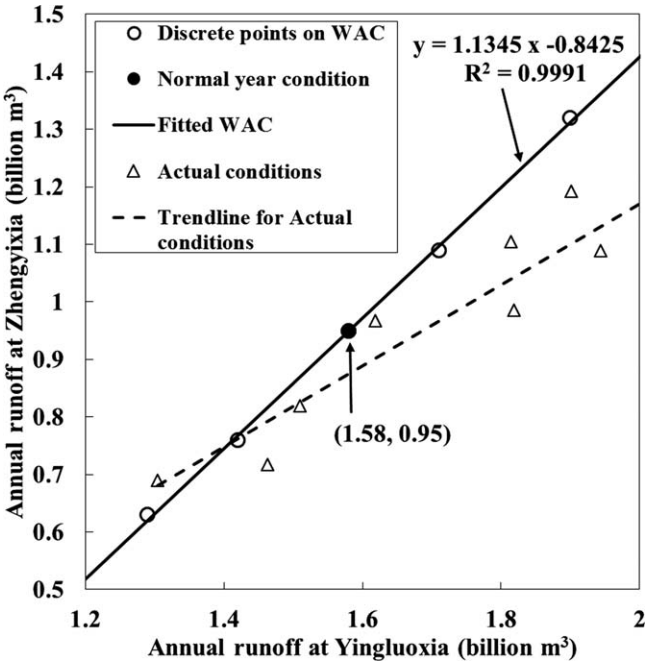

Figure 3. The water allocation curve (WAC) which represents the constraint of environmental flow at Zhengyixia depending on inflow conditions at Yingluoxia. The actual flow conditions are also plotted.

equals to 1.58 billion m^3), the outflow at Zhengyixia (i.e., *R*) should be no lower than 0.95 billion m^3, which forms the constraint in scenario B2. Scenario B2 can be referred to as "normal-year scenario." To extrapolate, we conducted linear fitting for the circles and obtained a fitted WAC, as shown in Figure 3. The constraint in scenario B3 is that *R* should be no lower than the flow limit interpolated or extrapolated from the fitted WAC. Scenario B3 is therefore referred to as "water-allocation-curve scenario."

For all the five scenarios, we performed the optimization, using both SOIM and DYCORS, for the individual years between 2000 and 2008 (Table 1), such that the impact of hydrological conditions on the optimization results could be explored. Thus, 90 numerical experiments in total were conducted, 45 using SOIM and 45 using DYCORS.

4. Results and Discussion

4.1. Effectiveness and Efficiency of SOIM

Figure 4 demonstrates the excellent performance of the SVM surrogates in replacing the GSFLOW model. Figures 4a–4c are for the 200 initial training points and Figures 4d–4f are for the 100 validating points. Note that, in the 45 SOIM experiments, a total number of 990 surrogate models were built for *R*, ΔS, and $Q_{j,m}$, and Figure 4 presents three examples of them. For those not shown, the R^2 values are all above 0.99 for the training, and all above 0.94 for the validation.

Figure 5 illustrates the optimization paths of SOIM and DYCORS, using scenario B1 as an example. As the effectiveness and efficiency of DYCORS for hydrological modeling has been proved by previous studies [*Espinet and Shoemaker*, 2013; *Regis and Shoemaker*, 2007, 2013], it was used as a benchmark for SOIM. First of all, both approaches resulted in a significant improvement of the objective function value, which indicates that they are both effective in solving the optimization problem. In most of the cases, SOIM led to a higher objective function value than DYCORS. Second, both of the two surrogate-based approaches are efficient. All the searches converged within the 300 GSFLOW run limit. Similar comparison results for the other four scenarios are illustrated by supporting information Figures S1–S4.

It is worth pointing out that the, for each inflow condition, the five scenarios share the same 200 initial training runs of GSFLOW in SOIM, as implied by the dashed portion of the SOIM curves in Figure 5. Thus, although DYCORS appears to converge faster than SOIM in individual searches, the computational cost of SOIM could be lower than DYCORS, if a variety of optimization problems are to be solved simultaneously based on a same complex model. Table 2 summarizes the numbers of GSFLOW evaluations required by the 45 SOIM experiments, taking into account the 100 validation runs in each inflow condition. GSFLOW evaluations (5120) were performed to finish all the 45 experiments, 114 for each on average. If more scenarios were considered, the average number can be further reduced, which highlights the computational advantage of SOIM. Note that an integrated SW-GW model for a large basin typically requires hours to days to finish a multiyear simulation on a PC. When such models are involved, heuristic optimization approaches are generally infeasible at a time scale allowed by decision-making. In our case, for example, even the 1 year

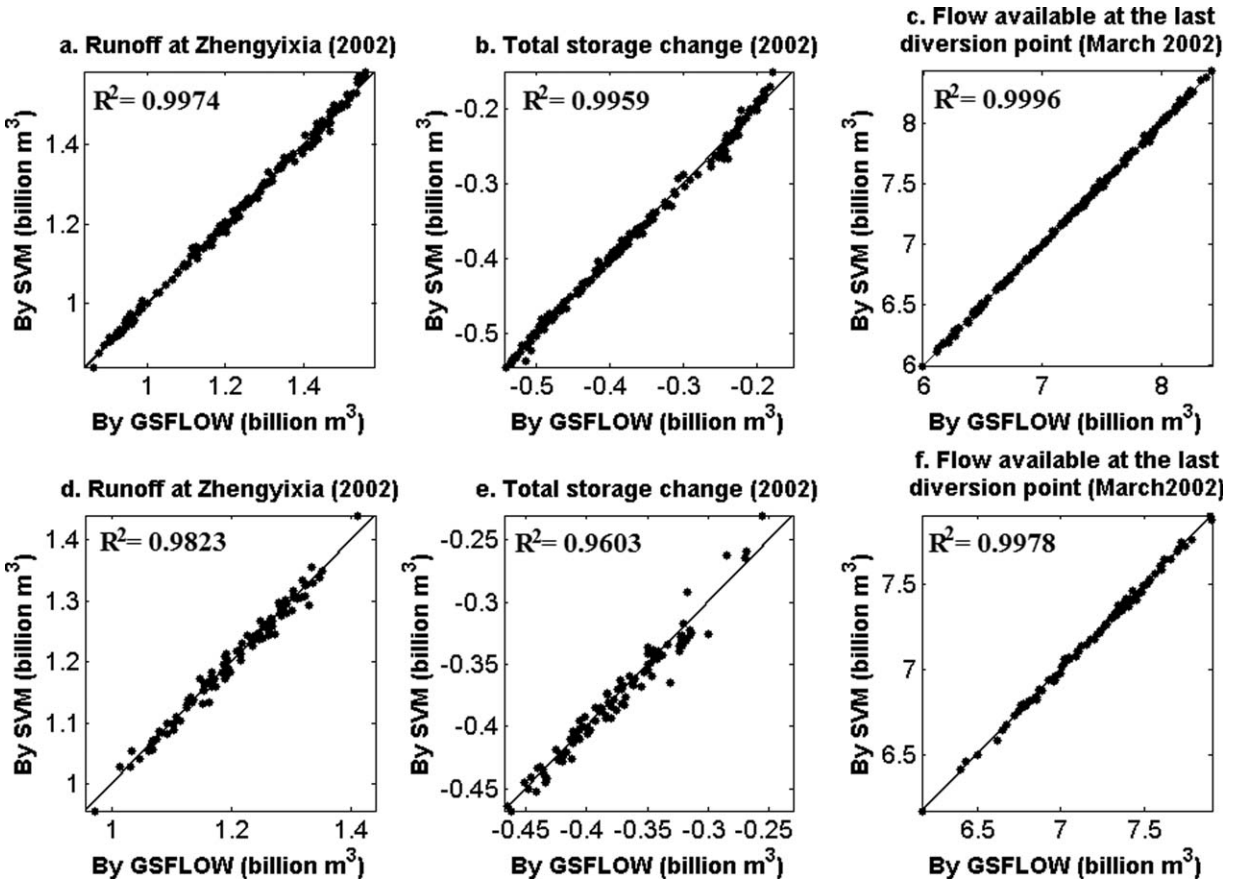

Figure 4. Comparison of SVM surrogate models against the original GSFLOW model. (a)–(c) Model simulation results at the 200 training points; (d)–(f) model simulation results at the 100 validating points. Runoff at Zhengyixia (Year 2002), total storage change (Year 2002), and flow available at the last diversion point (Year 2002) are the three examples.

GSFLOW simulation took 5 min to finish on a top-performance CPU (4.4 GHz), and typically the objective function needs to be evaluated for more than 5000 times in a single SCE-UA search. Without a surrogate-based approach like SOIM or DYCORS, the optimization would not be feasible.

4.2. Optimization Results by SOIM

Figure 6 illustrates the SOIM results for the nine experiments in scenario A1 (i.e., no-storage-decline scenario). In all cases, the final solution (solid circle) is far away from the actual condition (asterisk), which indicates the functioning of the optimization. The $k \times M$ additional training points (gray circles) are located in the region that is underrepresented by the initial training points (grey dots), which proves the necessity of the loop of steps (4)–(6) in SOIM (Figure 1). The arrows in Figure 6 indicate the optimization direction. Under the conditions of Years 2000–2002 and 2004–2006, in which the actual ΔS was far below zero, the direction is from lower right to upper left. As indicated by the column "Water stress" (defined as the proportion of annual irrigation demand in the total water input), these 6 years rank top 6 regarding the level of water stress. For these water-short years, there exists a sharp trade-off between streamflow decrease and groundwater storage decline. For Years 2000–2002 and 2004, the storage decline would be inevitable even if the pumping could approach zero, and therefore the optimization only made the status closer to $\Delta S=0$. On the other hand, under the conditions of Years 2003 and 2008, the optimization improves both R and ΔS, which is a desirable situation. Year 2007 represents a very wet year when the regional aquifer was significantly replenished, and therefore the optimization direction is from upper left to lower right. It is clear that the optimization has more space in wet conditions than in dry conditions.

Similarly, Figure 7 illustrates the optimization results in scenario B1 (i.e., the no-flow-decrease scenario). In all the nine cases, the flow constraint is met, and the optimization has led to an upward improvement. The

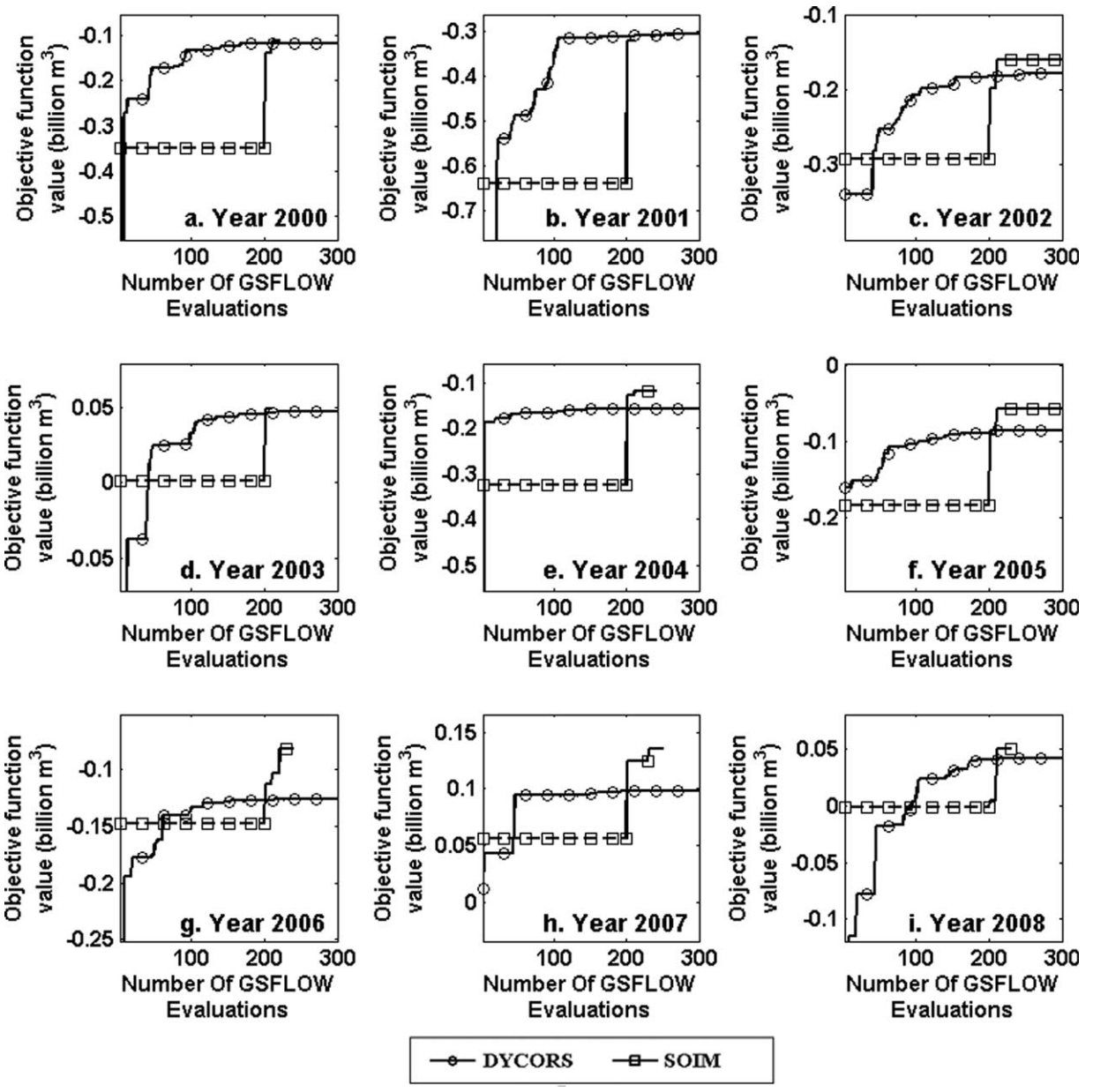

Figure 5. Evolution of the objective function value during the optimization by both SOIM and DYCORS. (a)–(i) The results for scenario B1 under the hydrologic conditions of Years 2000–2008, respectively. The dashed portion of the SOIM curves represents the 200 initial training runs of GSFLOW.

magnitude of the improvement varies in different years, and interestingly is not correlated with the water stress (see Table 1), which is because the flow constraint varies with the inflow condition as well.

To make the discussion succinct, the optimization results in other three scenarios are illustrated by supporting information Figures S5–S7.

4.3. Spatially Optimized Pumping Ratios

To ensure that the optimization results are physically reasonable, spatial patterns of the pumping ratio before and after the optimization were examined and compared. For example, Figure 8 shows the results for scenario B1 with the hydrological condition of Year 2002, which is closed to the normal year condition. Figures 8a and 8d together indicate that, in reality, districts in the recharge area (hereafter referred to as "recharge districts") depended heavily on groundwater, while those in the discharge

Table 2. Numbers of GSFLOW Evaluations in the 45 SOIM Experiments

			Added in the Iteration					
			Scenarios					
Inflow Conditions (Year)	For Training	For Validation	A1	A2	B1	B2	B3	Total
2000	200	100	100	10	20	10	10	450
2001	200	100	100	100	10	100	10	620
2002	200	100	100	70	100	100	100	770
2003	200	100	50	30	10	30	30	450
2004	200	100	100	30	40	70	20	560
2005	200	100	60	50	100	70	20	600
2006	200	100	20	10	40	50	10	430
2007	200	100	100	100	50	40	100	690
2008	200	100	70	100	30	20	30	550
Sum	1800	900	700	500	400	490	330	5120

area (hereafter referred to as "discharge districts") mostly used surface water. It is because the recharge districts are either of high elevation (e.g., No. 1, 2, 17) or far away from the river (e.g., No. 10), such that flow diversion is difficult to implement, while the surface water is much more accessible in the discharge districts. As Figures 8b and 8c demonstrate, the optimization has suggested reduction of the pumping ratio for most of the recharge districts, and increase for most of the discharge districts. It makes sense since the discharge districts have more shallow aquifers and are more adjacent to the river. The SW-GW exchanges are significant and quick in the discharge districts, such that the groundwater withdrawal can be quickly compensated by the infiltration water and river leakage.

Interestingly, it was also found that, although the areas A and B in Figure 8 were excluded from the optimization, they would still experience significant changes of ΔS according to the optimization. It implies that the connectivity of the regional groundwater system is good, such that local pumping activities may impact a much wider area. For example, the storage decrease in A is in fact due to the increase of groundwater pumping in the adjacent discharge districts, and the storage increase in B is caused by the reduction of pumping in the adjacent recharge districts.

According to Figure 8c, the average change of pumping ratio is +8% for the 18 districts, which means a significant increase of groundwater pumping. However, as Figure 7c shows, the groundwater storage would have an increase of around 0.1 billion m^3. Essentially, the storage increase is due to ET reduction. One apparent reason is that nonbeneficial ET from the aqueduct system can be reduced when part of the surface water diversion is replaced by groundwater pumping. Another important reason would be the reduced nonbeneficial ET from shallow groundwater. When the water table rises/declines (i.e., groundwater storage increases/decreases), ET from shallow groundwater should increase/decrease as well, but the response should be weaker in areas with larger depth to water table (DW). In scenario B1, 7 of the 18 districts would have an increase of ΔS (i.e., an increase of water table), and their average DWs (as of 2002) were all above 10 m (between 11.0 and 74.5 m). Among the other 11 districts with a decrease of ΔS (i.e., a decrease of water table), 7 had an average DW (as of 2002) smaller than 10 m (between 1.4 and 8.0 m). Therefore, the changed spatial pattern of groundwater storage (i.e., water table) due to the optimization would eventually lead to a decrease of ET from the regional aquifer.

4.4. Potential Water Saving

Figure 9 illustrates the relationship between water saving and change of ET, taking into account all the 45 SOIM solutions. Here the saving refers to the potential change of $(R+\Delta S)$ due to the optimization. The two variables exhibit an excellent 1:1 relationship ($R^2=0.9084$), which means one unit reduction of ET would roughly cause one unit of water saving. This is consistent with what have been found in scenario B1 (see section 4.3), that is, the potential storage increase after the optimization is due to the reduction of nonbeneficial ET. The small deviation of the points from the trend line should be due to the minor storage change in the soil and vadose zone. Note that some points of A1 and

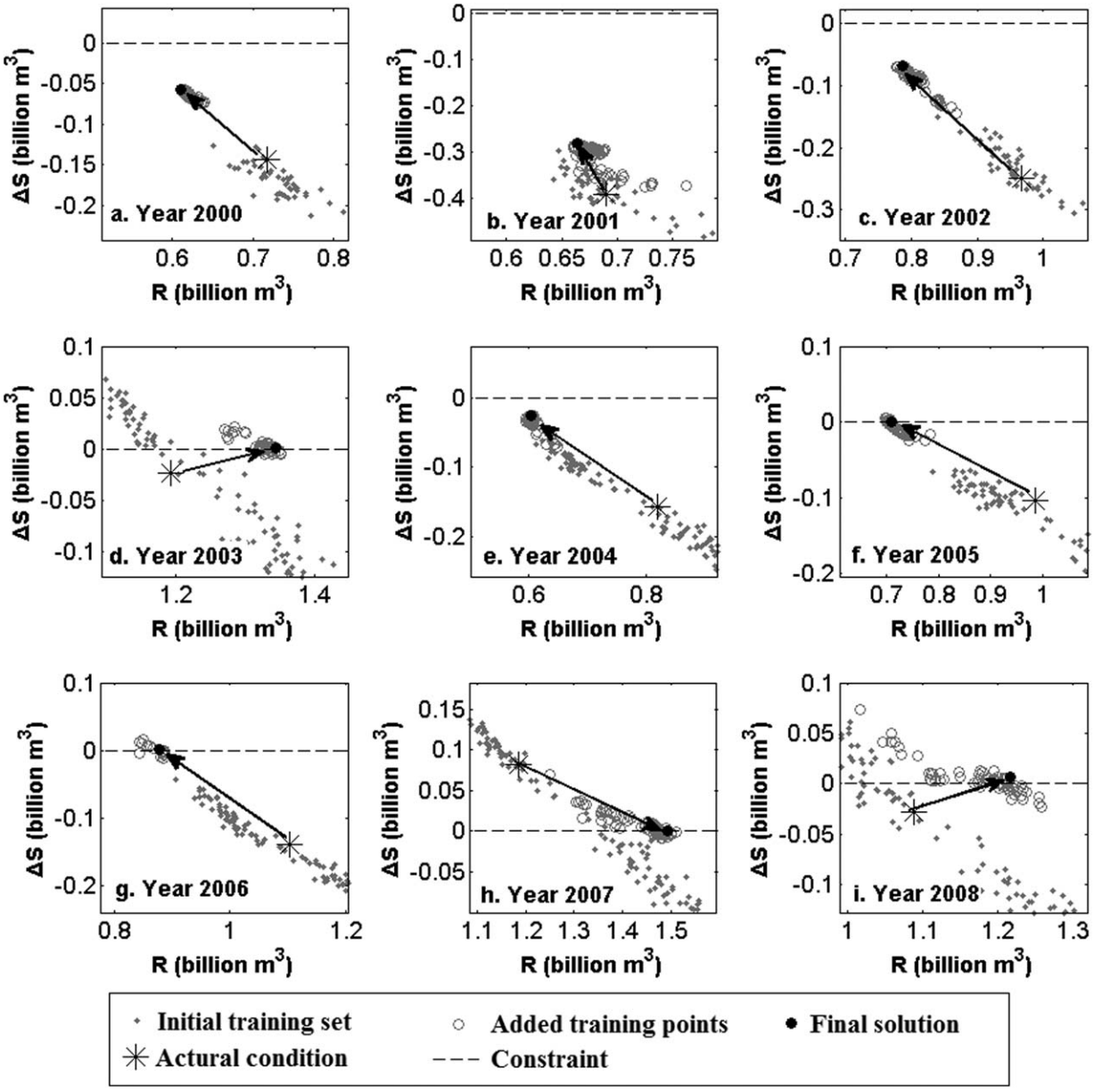

Figure 6. SOIM's optimization results in scenario A1. (a)–(i) The results under the hydrologic conditions of Years 2000–2008, respectively.

B2 are on the left side of the x=0 line, indicating a negative water saving. In scenario A1, as already discussed in section 4.2, these negative-saving points correspond to very dry years in which the constraint $\Delta S>0$ would be unachievable given the irrigation demand, even if the pumping were fully stopped. In such cases, moving toward the constraint would result in negative water saving. On the contrary, in scenario B2, the negative-saving points correspond to extremely wet years, in which cases 0.95 billion m^3 represents a trivial constraint for the outflow at Zhengyixia. In this case, the optimization would favor the surface water diversion rather than groundwater pumping, such that the nonbeneficial ET would be increased.

Figure 10 further compares the relationships between potential water saving and total water input in scenarios B1, B2, and B3. Here the total water input refers to the boundary surface water inflow from the mainstream and tributaries plus the local precipitation. The saving in scenario B1 has no

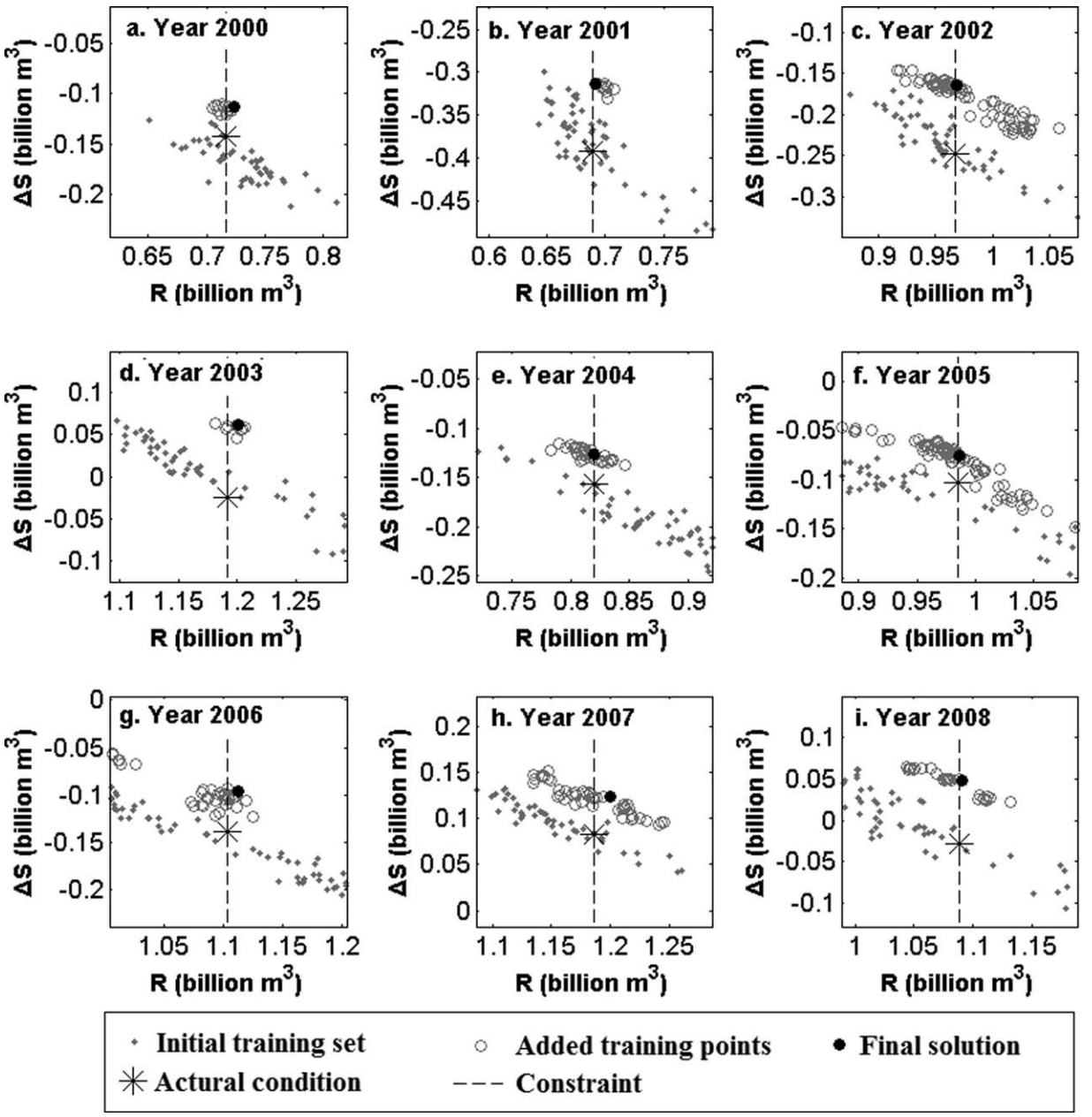

Figure 7. SOIM's optimization results in scenario B1. (a)–(i) The results under the hydrologic conditions of Years 2000–2008, respectively.

dependence on the water input, since the flow constraint varies with the water input condition, as already mentioned in section 4.2. The reason for the negative relationship in scenario B2 has also been mentioned before. Given the fixed constraint of 0.95 billion m^3, surface water diversion with more ET loss is favored in wetter years. Interestingly, the relationship is positive in scenario B3, which in fact can be explained by Figure 3. As the figure shows, the trend line of the actual conditions is mostly below WAC and has a smaller slope than WAC. In wetter years, the discrepancy between the actual outflow and the outflow limit imposed by WAC is larger, and therefore groundwater pumping with less ET loss would be further enhanced. Figure 10 sheds light on the importance of setting flexible environmental flow objective based on hydrological conditions. Further discussions on WAC are provided in the next section.

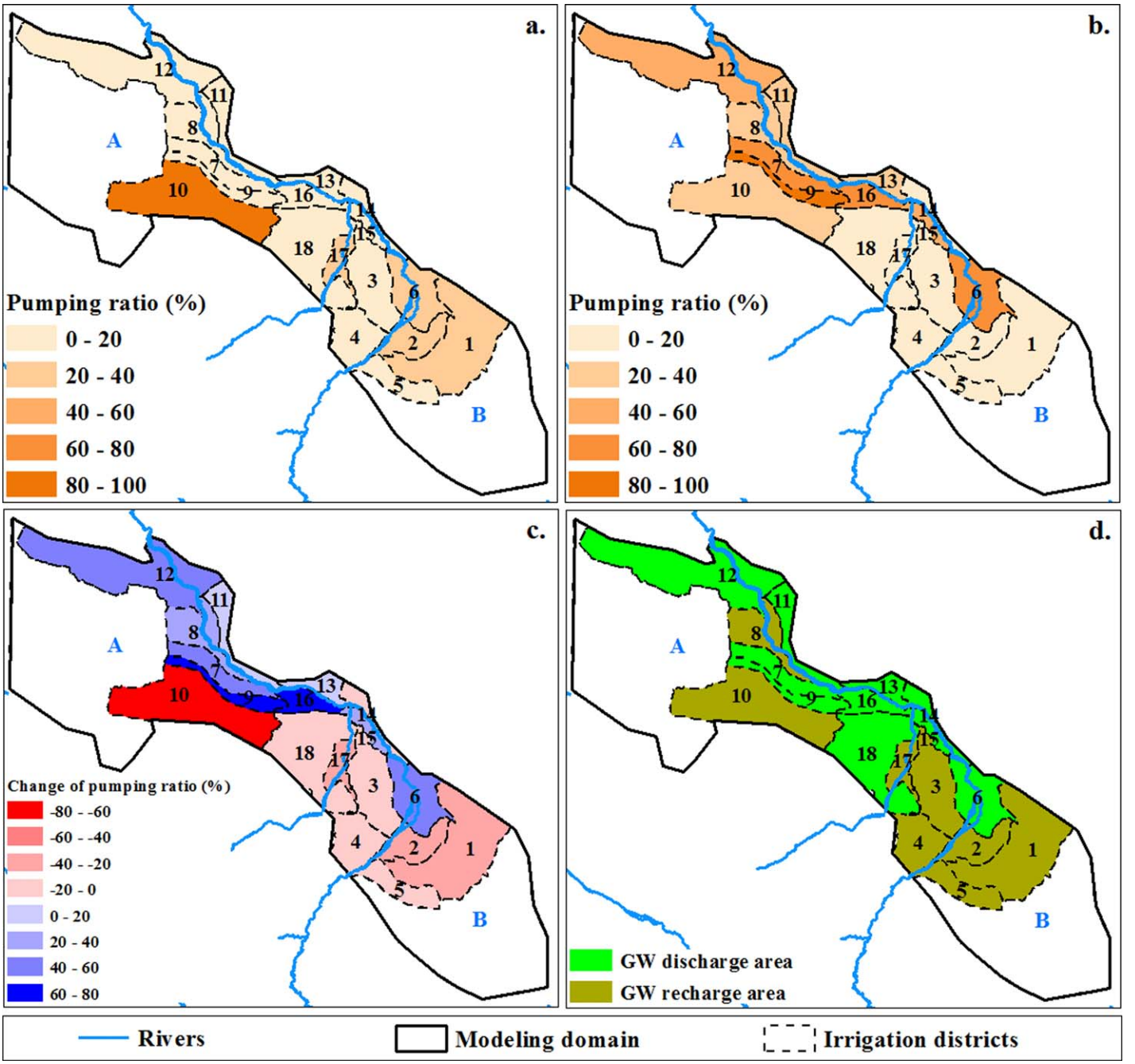

Figure 8. Spatial patterns of the pumping ratios at the 18 irrigation districts. (a) Actual pumping ratios in Year 2002; (b): optimized pumping ratios for Year 2002 in scenario B1; (c) change of the ratios after the optimization; and (d) distribution of groundwater discharge and recharge areas.

4.5. Water Allocation Curve Revisited

The water allocation curve (WAC) represents the doctrine for the water resources management in HRB, and has been enforced for more than a decade. Nevertheless, criticisms on WAC have never been stopped. One major concern is that WAC is too stringent for the midstream irrigation districts, and therefore the ecosystem recovery in the lower HRB may be at the cost of ecosystem degradation in the middle HRB [*Liu and Wang*, 2012; *Tian et al.*, 2015]. To maintain the doctrine, refraining the irrigation water use in the midstream districts seems to be inevitable. However, as the lower HRB has witnessed a fast growth of water-intensive agriculture (e.g., watermelon planting) in recent years, a

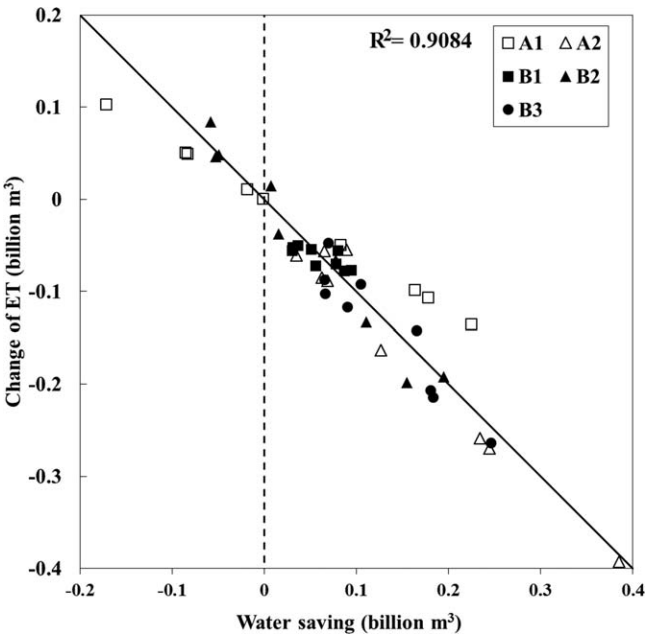

Figure 9. Water saving and change of ET at the 45 optimal solutions.

serious equity problem is now looming in HRB. Equity is also a key issue for the human society to address in adapting to environmental changes [e.g., *Field et al.*, 2014].

Figure 11 compared the optimization results in scenarios A1 and A2 against WAC. As mentioned before, scheme A favors the water resources conservation for the middle HRB. Thus, Figure 11 provides a way to evaluate WAC from the perspective of the middle HRB. As illustrated, the fitted optimal curve in scenario A1 sits far below WAC and only intersects WAC at its two ends. This implies that, in this strict groundwater-protection scenario, the conflict between the regulation and the water demand by the districts would be keen, except in extraordinarily dry or extraordinarily wet years. In extraordinarily dry years, the no-storage-decline constraint is impossible to meet, but the flow limit by WAC is trivial as well. In extraordinarily wet years, the total water input is more than enough to satisfy everyone. Scenario A2 adopts a no-accelerated-decline constraint instead. In this scenario, the optimal points and the fitted curve are very close to WAC in the low and medium flow conditions, and are above WAC in the high flow condition. Figure 11 suggests that, if both the irrigation demand and WAC keep unchanged, the decline of the groundwater storage is unstoppable. Even in the optimal situation, the decline would continue in the present pace.

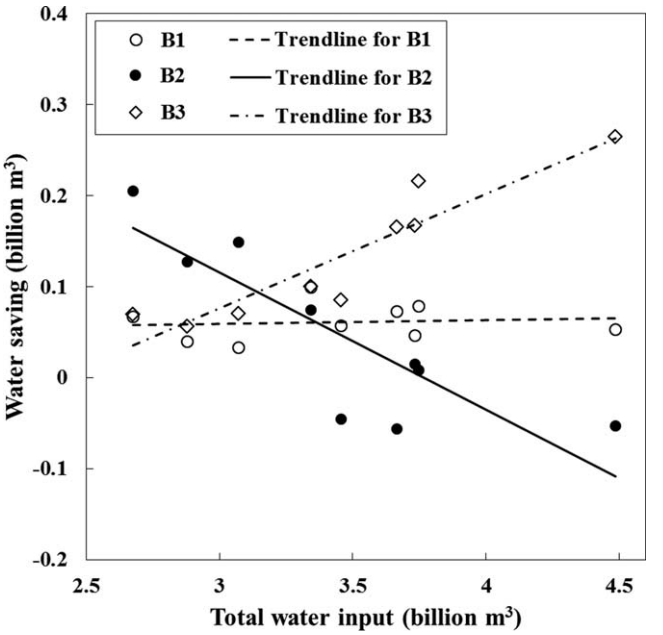

Figure 10. The relationship between water saving and total water input in scenarios B1, B2, and B3.

5. Conclusions

This study developed a new surrogate-based approach, abbreviated as SOIM, to optimizing water resources management in large river basins, which allows a complex SW-GW model into the optimization. SOIM's applicability and advantages were validated through a case study in the Zhangye Basin (ZB), for which a GSFLOW model was built. Five optimization scenarios and nine hydrological conditions were investigated in optimizing the partition of surface water and groundwater in the total irrigation water. As the 45 numerical experiments revealed, due to the strong and complicated SW-GW interactions, basin-scale water saving could be achieved by

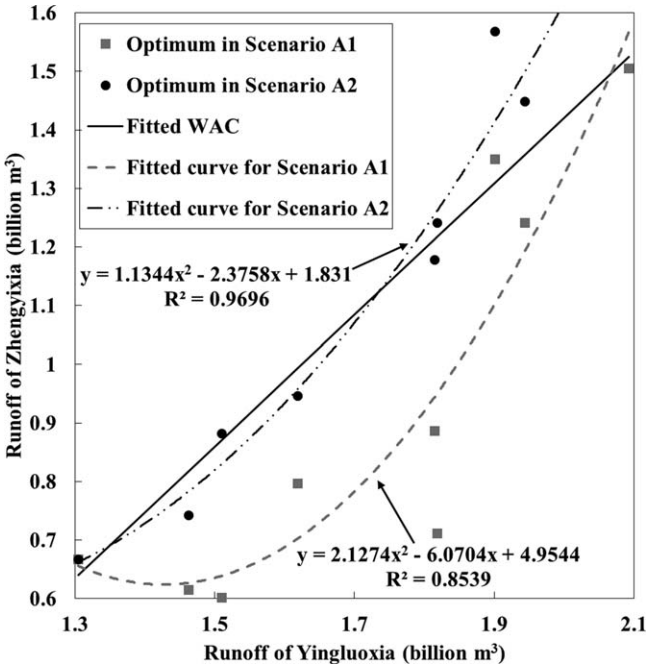

Figure 11. Optimization results in scenarios A1 and A2 compared against the water allocation curve (WAC).

spatially optimizing the ratios of groundwater use in different irrigation districts. The water-saving potential essentially stems from the reduction of nonbeneficial evapotranspiration from the aqueduct system and shallow groundwater, and its magnitude largely depends on both water management schemes and hydrological conditions. Important implications for water resources management in general include: first, environmental flow regulation needs to take into account interannual variation of hydrological conditions, as well as spatial complexity of SW-GW interactions; and second, to resolve water use conflicts between upper stream and lower stream, a system approach is highly desired to reflect ecological, economic, and social concerns in water management decisions.

On the other hand, this study clearly demonstrated the benefits of using a physically based integrated SW-GW model, instead of a more conceptual one, in addressing complex real-world management issues. An integrated model like GSFLOW comprehensively simulates the state variables and fluxes in basin-scale water cycle, with high spatial and temporal resolutions. Thus, it can provide great flexibility to formulating proper objective functions and constraints for different optimization problems. The model also allows one to scrutinize how various hydrological variables would be impacted by optimized water management strategies, and therefore makes the optimization results more physically interpretable and verifiable.

Overall, this study highlights the value of incorporating an integrated SW-GW model in optimizing water resources management, and stresses the point that surrogate-based approaches like SOIM and DYCORS pave the path for such incorporation. These approaches represent a promising solution to filling the gap between complex environmental modeling and real-world management decision-making. Although SOIM was proposed for integrated SW-GW models, it is potentially applicable to other complex models such as land surface, climate, and water quality models.

Nevertheless, as an early effort to fuse water management optimization, physically based hydrological modeling and surrogate modeling, this study has some limitations to be addressed by future work. First, although SVM works nicely for the integrated model in this study, it may not be treated as a general choice for physically based hydrological modeling. It is also not clear whether the performance of SVM would be seriously degraded when more complicated optimization problems (e.g., multiple objectives, a much higher dimension of decision space, etc.) are considered. Future studies may compare multiple surrogate techniques (e.g., SVM, Kriging, RBF, etc.) within the SOIM framework for different complexity levels of optimization problems. Second, this study formulated a relatively simple optimization problem in which water quantity is the only concern. A more realistic optimization problem may incorporate other considerations (e.g., water quality, ecological health, engineering feasibility, economic cost, etc.) and essentially be multiobjective. Exploring the Pareto front in multiobjective optimization involving a complex model is a very challenging task, even with surrogate modeling. Whether and how the existing approaches including SOIM could effectively tackle this task, or, if a ground breaking approach needs to be explored, deserves future research efforts.

Acknowledgments
This work was supported by the National Natural Science Foundation of China (NSFC) (91125021; 91225301; 41371473). The data set was provided by Cold and Arid Regions Science Data Center in Lanzhou, China (http://westdc.westgis.ac.cn). We thank Juliane Mueller and Christine A. Shoemaker for providing us the MATLAB code of DYCORS (Version 1.3).

References

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams (1998), Large area hydrologic modeling and assessment: Part I. Model development, *J. Am. Water Resour. Assoc.*, *34*(1), 73–89.

Baú, D. A., and A. S. Mayer (2006), Stochastic management of pump-and-treat strategies using surrogate functions, *Adv. Water Resour.*, *29*(12), 1901–1917.

Brunner, P., and C. T. Simmons (2012), HydroGeoSphere: A fully integrated, physically based hydrological model, *Groundwater*, *50*(2), 170–176.

Burges, C. J. C. (1998), A tutorial on Support Vector Machines for pattern recognition, *Data Min. Knowl. Discov.*, *2*(2), 121–167.

Cai, X., R. Zeng, W. H. Kang, J. Song, and A. J. Valocchi (2015), Strategic planning for drought mitigation under climate change, *J. Water Resour. Plann. Manage.*, 04015004, doi:10.1061/(asce)wr.1943–5452.0000510.

Condon, L. E., and R. M. Maxwell (2013), Implementation of a linear optimization water allocation algorithm into a fully integrated physical hydrology model, *Adv. Water Resour.*, *60*, 135–147.

Dakhlaoui, H., Z. Bargaoui, and A. Bárdossy (2012), Toward a more efficient Calibration Schema for HBV rainfall-runoff model, *J. Hydrol.*, *444–445*, 161–179.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan (2002), A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Trans. Evol. Comput.*, *6*(2), 182–197.

Doherty, J. (1994), PEST: A unique computer program for model-independent parameter optimisation, *Water Down Under 94: Groundwater/Surface Hydrol. Common Interest Pap.*, 551, Barton, ACT: Institution of Engineers, Australia.

Doherty, J., and R. J. Hunt (2009), Two statistics for evaluating parameter identifiability and error reduction, *J. Hydrol.*, *366*(1–4), 119–127.

Duan, Q., S. Sorooshian, and V. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, *28*(4), 1015–1031.

Espinet, A. J., and C. A. Shoemaker (2013), Comparison of optimization algorithms for parameter estimation of multi-phase flow models with application to geological carbon sequestration, *Adv. Water Resour.*, *54*, 133–148.

Fan, R., P. Chen, and C. Lin (2005), Working set selection using the second order information for training SVM, *J. Mach. Learn. Res.*, *6*, 1889–1918.

Fen, C.-S., C. Chan, and H.-C. Cheng (2009), Assessing a response surface-based optimization approach for soil vapor extraction system design, *J. Water Resour. Plann. Manage.*, *135*(3), 198–207.

Fonseca, A., C. Botelho, R. A. Boaventura, and V. J. Vilar (2014), Integrated hydrological and water quality model for river management: A case study on Lena River, *Sci. Total Environ.*, *485–486*, 474–489.

Field, C. B., V. R. Barros, K. J. Mach, and M. D. Mastrandrea (2014), Technical summary, in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge Univ. Press, Cambridge, U. K.

Graham, N., and A. Refsgaard (2001), *MIKE SHE: A Distributed, Physically Based Modelling System for Surface Water/Groundwater Interactions*, in MODFLOW 2001 and Other Modeling Odysseys, Conference Proceedings, Colorado School of Mines, Golden, Colo.

Guo, Q., M. Kelly, and C. H. Graham (2005), Support vector machines for predicting distribution of sudden Oak death in California, *Ecol. Modell.*, *182*(1), 75–90.

Guo, Q., Q. Feng, and J. Li (2009), Environmental changes after ecological water conveyance in the lower reaches of Heihe River, northwest China, *Environ. Geol.*, *58*(7), 1387–1396.

Harbaugh, A. W. (2005), MODFLOW-2005, The U.S. Geological Survey modular ground-water model: The ground-water flow process, *U.S. Geol. Surv. Tech. Methods, 6-A16*.

Hassanzadeh, E., A. Elshorbagy, H. Wheater, and P. Gober (2014), Managing water in complex systems: An integrated water resources model for Saskatchewan, Canada, *Environ. Modell. Software*, *58*, 12–26.

Hu, L. T., C. X. Chen, J. J. Jiao, and Z. J. Wang (2007), Simulated groundwater interaction with rivers and springs in the Heihe river basin, *Hydrol. Processes*, *21*(20), 2794–2806.

Ji, X., E. Kang, R. Chen, W. Zhao, Z. Zhang, and B. Jin (2006), The impact of the development of water resources on environment in arid inland river basins of Hexi region, Northwestern China, *Environ. Geol.*, *50*(6), 793–801.

Johnson, V. M., and L. L. Rogers (2000), Accuracy of neural network approximators in simulation-optimization, *J. Water Resour. Plann. Manage.*, *126*(2), 48–56.

Khakbaz, B., B. Imam, K. Hsu, and S. Sorooshian (2012), From lumped to distributed via semi-distributed: Calibration strategies for semi-distributed hydrologic models, *J. Hydrol.*, *418–419*, 61–77.

Koech, R. K., R. J. Smith, and M. H. Gillies (2014), Evaluating the performance of a real-time optimisation system for furrow irrigation, *Agric. Water Manage.*, *142*, 77–87.

Kollet, S. J., and R. M. Maxwell (2006), Integrated surface-groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model, *Adv. Water Resour.*, *26*(7), 945–958.

Lin, J.-Y., C.-T. Cheng, and K.-W. Chau (2006), Using support vector machines for long-term discharge prediction, *Hydrol. Sci. J.*, *51*(4), 599–612.

Liong, S.-Y., and C. Sivapragasam (2002), Flood stage forecasting with support vector machines, *J. Am. Water Resour. Assoc.*, *38*(1), 173–186.

Liong, S.-Y., S.-T. Khu, and W.-T. Chan (2001), Derivation of Pareto front with genetic algorithm and neural network, *J. Hydrol. Eng.*, *6*(1), 52–61.

Liu, X., and L. Wang (2012), Discussion on the scheduling scheme and water allocation curve of Heihe River [in Chinese], *Gansu Water Resour. Hydropower Technol.*, *48*(10), 16–18.

Maier, H. R., et al. (2014), Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions, *Environ. Modell. Software*, *62*, 271–299.

Markstrom, S. L., R. G. Niswonger, R. S. Regan, D. E. Prudic, and P. M. Barlow (2008), GSFLOW: Coupled ground-water and surface-water flow model based on the integration of the Precipitation-Runoff Modeling System (PRMS) and the Modular Ground-Water Flow Model (MODFLOW-2005), *U.S. Geol. Surv. Tech. Methods*, 6-D1, 240 pp.

Mazzega, P., O. Therond, T. Debril, H. March, C. Sibertin-Blanc, R. Lardy, and D. Sant'ana (2014), Critical multi-level governance issues of integrated modelling: An example of low-water management in the Adour-Garonne basin (France), *J. Hydrol.*, *519*, 2515–2526, doi:10.1016/j.jhydrol.2014.1009.1043.

McKay, M. D., R. J. Beckman, and W. J. Conover (1979), A Comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics*, *21*(2), 239–245.

Molinos-Senante, M., F. Hernández-Sancho, M. Mocholí-Arce, and R. Sala-Garrido (2014), A management and optimisation model for water supply planning in water deficit areas, *J. Hydrol.*, *515*, 139–146.

Mugunthan, P., and C. A. Shoemaker (2006), Assessing the impacts of parameter uncertainty for computationally expensive groundwater models, *Water Resour. Res.*, *42*, W10428, doi:10.1029/2005WR004640.

Ostfeld, A., and S. Salomons (2005), A hybrid genetic-instance based learning algorithm for CE-QUAL-W2 calibration, *J. Hydrol.*, *310*(1–4), 122–142.
Panday, S., and P. S. Huyakorn (2004), A fully-coupled physically-based spatially distributed model for evaluating surface/subsurface flow, *Adv. Water Resour.*, *27*(4), 361–382.
Pisinaras, V., C. Petalas, V. A. Tsihrintzis, and G. P. Karatzas (2013), Integrated modeling as a decision-aiding tool for groundwater management in a Mediterranean agricultural watershed, *Hydrol. Processes*, *27*(14), 1973–1987.
Price, K. V., R. M. Storn, and J. A. Lampinen (2005), *Differential Evolution: A Practical Approach to Global Optimization*, Springer, Berlin.
Razavi, S., B. A. Tolson, and D. H. Burn (2012a), Numerical assessment of metamodelling strategies in computationally intensive optimization, *Environ. Modell. Software*, *34*, 67–86.
Razavi, S., B. A. Tolson, and D. H. Burn (2012b), Review of surrogate modeling in water resources, *Water Resour. Res.*, *48*, W07401, doi: 10.1029/2011WR011527.
Regis, R. G., and C. A. Shoemaker (2007), A stochastic radial basis function method for the global optimization of expensive functions, *INFORMS J. Comput.*, *19*(4), 497–509.
Regis, R. G., and C. A. Shoemaker (2013), Combining radial basis function surrogates and dynamic coordinate search in high-dimensional expensive black-box optimization, *Eng. Optim.*, *45*(5), 529–555.
Rigon, R., G. Bertoldi, and T. M. Over (2006), GEOtop: A distributed hydrological model with coupled water and energy budgets, *J. Hydrometeorol.*, *7*(3), 371–388.
Schmidt, S., T. Geyer, A. Marei, J. Guttman, and M. Sauter (2013), Quantification of long-term wastewater impacts on karst groundwater resources in a semi-arid environment by chloride mass balance methods, *J. Hydrol.*, *502*, 177–190.
Shanafield, M., P. G. Cook, P. Brunner, J. McCallum, and C. T. Simmons (2012), Aquifer response to surface water transience in disconnected streams, *Water Resour. Res.*, *48*, W11510, doi:10.1029/2012WR012103.
Shawe-Taylor, J., P. L. Bartlett, R. C. Williamson, and M. Anthony (1998), Structural risk minimization over data-dependent hierarchies, *IEEE Trans. Inform. Theory*, *44*(5), 1926–1940.
Shen, C., J. Niu, and M. S. Phanikumar (2013), Evaluating controls on coupled hydrologic and vegetation dynamics in a humid continental climate watershed using a subsurface-land surface processes model, *Water Resour. Res.*, *49*, 2552–2572, doi:10.1002/wrcr.20189.
Simpson, S. C., T. Meixner, and J. F. Hogan (2013), The role of flood size and duration on streamflow and riparian groundwater composition in a semi-arid basin, *J. Hydrol.*, *488*, 126–135.
Sreekanth, J., and B. Datta (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, doi:10.1029/2010WR009683.
Stein, M. (1987), Large sample properties of simulations using Latin hypercube sampling, *Technometrics*, *29*(2), 143–151.
Surfleet, C. G., and D. Tullos (2013), Variability in effect of climate change on rain-on-snow peak flow events in a temperate climate, *J. Hydrol.*, *479*, 24–34.
Tian, Y., Y. Zheng, B. Wu, X. Wu, J. Liu, and C. Zheng (2015), Modeling surface water-groundwater interaction in arid and semi-arid regions with intensive agriculture, *Environ. Modell. Software*, *63*, 170–184.
Tolson, B. A., and C. A. Shoemaker (2007), Dynamically dimensioned search algorithm for computationally efficient watershed model calibration, *Water Resour. Res.*, *43*, W01413, doi:10.1029/2005WR004723.
Vapnik, V. (1998), *Statistical Learning Theory*, 768 pp., John Wiley, N. Y.
Wang, C., Q. Duan, W. Gong, A. Ye, Z. Di, and C. Miao (2014), An evaluation of adaptive surrogate modeling based optimization with two benchmark problems, *Environ. Modell. Software*, *60*, 167–179.
Weill, S., A. Mazzia, M. Putti, and C. Paniconi (2011), Coupling water flow and solute transport into a physically-based surface-subsurface hydrological model, *Adv. Water Resour.*, *34*(1), 128–136.
Wen, X., Y. Wu, J. Su, Y. Zhang, and F. Liu (2005), Hydrochemical characteristics and salinity of groundwater in the Ejina Basin, Northwestern China, *Environ. Geol.*, *48*(6), 665–675.
Wilby, R. L., P. G. Whitehead, A. J. Wade, D. Butterfield, R. J. Davis, and G. Watts (2006), Integrated modelling of climate change impacts on water resources and quality in a lowland catchment: River Kennet, UK, *J. Hydrol.*, *330*(1-2), 204–220.
Wu, B., Y. Zheng, Y. Tian, X. Wu, Y. Yao, F. Han, J. Liu, and C. Zheng (2014), Systematic assessment of the uncertainty in integrated surface water-groundwater modeling based on the probabilistic collocation method, *Water Resour. Res.*, *50*, 5848–5865, doi:10.1002/2014WR015366.
Wu, Y. Q., Y. H. Zhang, X. H. Wen, and J. P. Su (2010), *Hydrologic cycle and water resource modeling for the Heihe River Basin in northwestern China [in Chinese]*, 199 pp., Science Press, Beijing.
Zhang, X., R. Srinivasan, and M. Van Liew (2009), Approximating SWAT model using artificial neural network and support vector machine, *J. Am. Water Resour. Assoc.*, *45*(2), 460–474.
Zheng, Y., W. M. Wang, F. Han, and J. Ping (2011), Uncertainty assessment for watershed water quality modeling: A probablilistic collocation method based approach, *Adv. Water Resour.*, *34*(7), 887–898.