Environ Earth Sci (2015) 73:423–434
DOI 10.1007/s12665-014-3448-6

ORIGINAL ARTICLE

Development of multi-metamodels to support surface water quality management and decision making

Alexander Y. Sun · Roger M. Miranda · Xianli Xu

Received: 19 August 2013 / Accepted: 13 June 2014 / Published online: 4 July 2014

Abstract Watershed management and planning is a complex decision-making process, which not only involves deliberation using one or more watershed models, but also requires collaboration among multiple stakeholder groups with different ideologies, interests, and demographics. Web-based decision support tools have great potentials to enhance the transparency and participation of such decision making processes. Although physically based surface water quality models are well suited for offline water quality analyses, they are often too computationally demanding to be deployed in a web-based environment. In this work, three metamodels are developed to support decision-making activities related to surface water quality management at Arroyo Colorado Watershed, a coastal watershed located in Texas, US. All three metamodels are trained using an existing Soil and Water Assessment Tool (SWAT) model developed for the watershed. The main objectives of the metamodels are to support web-based decision support, including near-term nutrient load forecasting, online sensitivity study, and long-term load reduction planning. All metamodels either replicate or extend the capabilities of the original SWAT model and, thus, provide proxies for regulators and stakeholders to examine and discuss model results interactively. The novel, multi-metamodel methodology taken here is not only useful for supporting multigroup decision making and public education, but also provides a more effective way to leverage existing investment on watershed models.

Keywords Metamodeling · RBFN · PCM · Visual analytics · Collaborative decision making · Environmental decision support systems · SWAT model

A. Y. Sun (✉)
Bureau of Economic Geology, The University of Texas at Austin, Austin, TX, USA
e-mail: alex.sun@beg.utexas.edu

R. M. Miranda
Lyndon B. Johnson School of Public Affairs, The University of Texas at Austin, Austin, TX, USA

X. Xu
Institute of Subtropical Agriculture, Chinese Academy of Sciences, Changsha, China

Introduction

The nature of environmental decision making calls for a participatory approach, in which regulators, stakeholders, and researchers jointly examine management options through an iterative process and on a common platform. Web-based environmental decision-support systems (EDSS) have flourished since mid-1990s, in part to facilitate and sustain such multiparty decision-making processes. By definition, an EDSS consists of environmental models, databases, and assessment tools that are integrated under a common graphical user interface and realized by using spatial-data-management functionalities (Matthies et al. 2007). Thus, web-based EDSS are distinguished by their visual analytical capabilities, separating them from general geographic information systems that are mainly suitable for data display.

Implementation of analytical capabilities in EDSS is highly case dependent. While displaying static content (e.g., model outputs) may be sufficient in some cases, a growing number of projects now require web-based analytical tools that can assist end users to experiment with different scenarios and perform online simulation, forecasting, or even uncertainty analysis. Computational power has significantly grown in the last decade or so, but so has the computational demand of physically based, distributed watershed models (Castelletti et al. 2012a). As a result, serial execution of physically based models may still take minutes or even hours to finish, which make these models unsuitable for online use. Until distributed computing becomes more accessible, adapting computationally intensive models for online decision support will remain a practical challenge. The main objective of this paper is to demonstrate the application of model-reduction techniques in a web-based EDSS through a case study.

Model reduction refers broadly to techniques for developing and utilizing cheaper-to-run "surrogates" of a computer simulation model. The idea behind model reduction is not new. Indeed, a vast number of potential techniques exist in the literature and are often referred to as reduced-order modeling, surrogate modeling, response surface methods, metamodeling, or model emulation. For consistency, the term "metamodel" will be used throughout this paper. The original concept of metamodeling dates back to R. W. Blanning, who in the 1970s proposed to approximate a computer simulation model for expediting the solution of optimization problems (Blanning 1975). Since then, the subject has become an active research area in computer experiment design, uncertainty quantification, and risk assessment, as summarized in a number of monographs and literature surveys (Santner et al. 2003; Fang et al. 2006; Helton et al. 2006; Kleijnen 2008; Eldred 2009; Matott et al. 2009; Storlie et al. 2009; Young and Ratto 2009; Shan and Wang 2010; Castelletti et al. 2012a; Razavi et al. 2012).

Metamodeling mainly capitalizes on the idea that not all processes encapsulated by a process-level model are equally important and relevant to the main objectives of planning and management (Castelletti et al. 2012b). In other words, to be suitable for metamodeling, a problem needs to have both simulation space and decision space, with the latter being a subset of the former. Whilst the simulation space contains all information on relevant physical processes necessary to ensure model accuracy, the decision space only includes dominant system features that matter most to the specific objectives of decision making. Identification of an appropriate decision space is particularly important for web-based metamodel implementation.

Several classification schemes exist for metamodeling techniques. Depending on their parameterization processes, metamodeling techniques may be classified as either parametric or nonparametric. Parametric techniques attempt to replace the complex structure of a distributed model using simplified model/parameter structures, whereas nonparametric techniques train a "black-box" model to mimic the observed or simulated input-response relationships. The primary difference between the two is that coefficients from the latter usually do not have physical meanings. Some authors refer to the former techniques as "low-fidelity models" while reserving the term "metamodels" mainly for the latter techniques (e.g., Razavi et al. 2012). Some of the best-known nonparametric methods are spline-smoothing regression models (Friedman 1991), artificial neural networks (ANN) (Haykin 1999), and support vector machines (Vapnik 1998). In recent years, stochastic response surface methods have also received broad attention in reliability analyses and porous media modeling (Isukapalli et al. 1998; Eldred 2009; Xiu 2010). As in the spline-smoothing-based approaches, stochastic response surface methods seek to construct a response surface using results from a set of full-model runs for different parameter combinations; however, it does so in the stochastic space by treating uncertain model parameters as random variables. Of particular interest here is a stochastic collocation method called the probabilistic collocation method (PCM) (Li and Zhang 2007; Zheng et al. 2011; Sun et al. 2013). A major advantage of the PCM is that it does not require modification of an existing model (i.e., non-intrusive). Once constructed, the PCM metamodel can be used to generate results of a model run at virtually no additional cost. A disadvantage of the PCM is that it assumes the probability distributions of random variables are accessible, which may not always be the case. Recently, Zheng et al. (2011) applied PCM to approximate the Watershed Analysis Risk Management Framework (WARMF) model that is documented in Chen et al. (2005).

Depending on the modeling paradigm followed, metamodeling techniques can also be classified as either data-driven or objective-driven. Data-driven metamodels, such as ANN, are appropriate for predicting nonlinear input–output relationships without requiring knowledge of underlying physical processes. The success of such methods largely hinges on whether the training dataset sufficiently captures all expected system variation patterns. Extending data-driven models beyond the range of training data can lead to unpredictable results. On the other hand, objective-driven models often replace an existing complex model structure with a reduced-order model structure, which runs faster but still retains major features of the original model. Objective-driven metamodels may use a hybrid of parametric and nonparametric techniques; the modeling objectives are used not only to guide simplification of the original process-level model, but also to make appropriate performance metrics for gauging the level of model reduction. The development of objective-driven models is thus more problem specific, requiring significant domain knowledge to develop a simplified model structure (i.e., model abstraction) and then estimate the corresponding reduced set of model parameters.

In a recent study, we developed a web-based EDSS, the Collaborative Geospatial Decision Support System (CGDSS), to support watershed management and decision making for the Arroyo Colorado Watershed (ACW), a

Texas coastal watershed located near the US–Mexico border. The main objective of the CGDSS project was to experiment with a co-management platform to support decision making and public outreach activities during the WPP implementation-tracking period. Previously, a SWAT model had been built for the ACW to support implementation of a Watershed Protection Plan (WPP) (Arroyo Colorado WPP 2007; Kannan et al. 2011, 2014). The Arroyo Colorado WPP is a "comprehensive watershed-based strategy" aimed to improve water quality and aquatic and riparian habitats in the watershed; it was designed to address impairments and concerns identified in the Texas Water Quality Inventory (Arroyo Colorado WPP 2007). The implementation period for Phase I of the Arroyo Colorado WPP is 2006–2015.

SWAT is a widely used, semi-distributed, watershed model for simulating the quality and quantity of surface and groundwater, and predicting the environmental impact of land use and land management practices (Gassman et al. 2007). A serial run of the Arroyo Colorado SWAT model takes more than 10 min to finish (see Sect. 4). In order to incorporate different aspects and use cases of the SWAT model in the CGDSS framework, the team developed three interactive metamodels:

- The first metamodel is a type of ANN model trained using the inputs/outputs of the SWAT model and is mainly used to predict pollutant loadings under user-specified forcing conditions.
- The second metamodel is developed using PCM and is used to examine sensitivity of loadings to different parameter values; it can also be used for uncertainty quantification.
- The third metamodel is a simple loading estimation model (i.e., PLOAD) used to examine the effect of land use changes on long-term pollutant loading behaviors.

The three metamodels are complementary and are designed to enhance public outreach and stakeholder engagement for a range of user levels. Although only total nitrogen (TN) loading is demonstrated in this work, metamodels for other parameters of interest can be implemented in a similar manner.

A note to keep in mind is that in the current study the SWAT model has already been adopted by regulators and local stakeholders. Thus, the main focus of metamodeling is to honor the original model behaviors as much as possible in the web-based decision space, rather than to improve the model quality in the simulation space. In other words, we deal with a sub-problem of the more complicated watershed quality management problem. Nevertheless, this sub-problem plays an important role in bridging the gap between scientific research and community-centered watershed management. To the best of our knowledge, few previous studies have attempted to implement multiple metamodels in a single web-based EDSS. Thus, experiences gained from this project will help to illuminate the technical challenges and issues associated with web-based implementation of metamodeling. The rest of paper is organized as follows. Section 2 provides detailed descriptions of the study area and background of study. Section 3 briefly introduces technical background for each type of metamodeling used. Section 4 demonstrates the three metamodels for TN loading and, finally, Sect. 5 discusses lessons learned.

Study area

The ACW covers an area of about 1,828 km^2 (Fig. 1) on the Gulf of Mexico coast. It is bounded on the west and south by the drainage divide to the Rio Grande, on the north by the drainage divide to the North Floodway, and on the east by the Lower Laguna Madre lagoon (Arroyo Colorado WPP 2007). The Arroyo Colorado River serves as the main drainage stream for this area of Texas. The lower third of the stream provides an inland waterway for commercial barge traffic and for recreational boating and fishing. Near the Gulf Coast, the Arroyo Colorado also serves as an important nursery and foraging area for numerous species of marine fish, shrimp, and crab (Arroyo Colorado WPP 2007). The Arroyo Colorado has two segments, the tidally influenced segment (Segment 2201) and the above-tidal segment (Segment 2202) (Fig. 1). Land within the ACW is intensively cultivated and irrigated. Dominant land-cover categories in the watershed are agriculture (54 %), range (18.5 %), urban (12.5 %), sugarcane (4 %), and other (Kannan et al. 2011) (see Fig. 2). Decades of human use have degraded habitat and water quality in the ACW and strained its ability to assimilate pollutants (Raines and Miranda 2002). Both Segments 2201 and 2202 are now listed as impaired due to high levels of bacteria. Segment 2201 is also listed as impaired due to low levels of dissolved oxygen. Nutrient concentrations (nitrogen and phosphorus compounds) are high in both segments (Arroyo Colorado WPP 2007). In the Arroyo Colorado WPP, Year 2000 was chosen as the starting benchmark for load calculations because many of the wastewater infrastructure improvement projects were completed in late 1990s. The 10-year load-reduction target for TN is 11 % below its Year 2000 level of 2,034 tons/year (Arroyo Colorado WPP 2007).

The ACW Partnership, which grew out of smaller groups of local stakeholders involved in the WPP development process, is now the leading stewardship organization in the watershed. The Partnership had been actively involved behind the development of the SWAT model. The ACW SWAT model was calibrated and validated for flow,

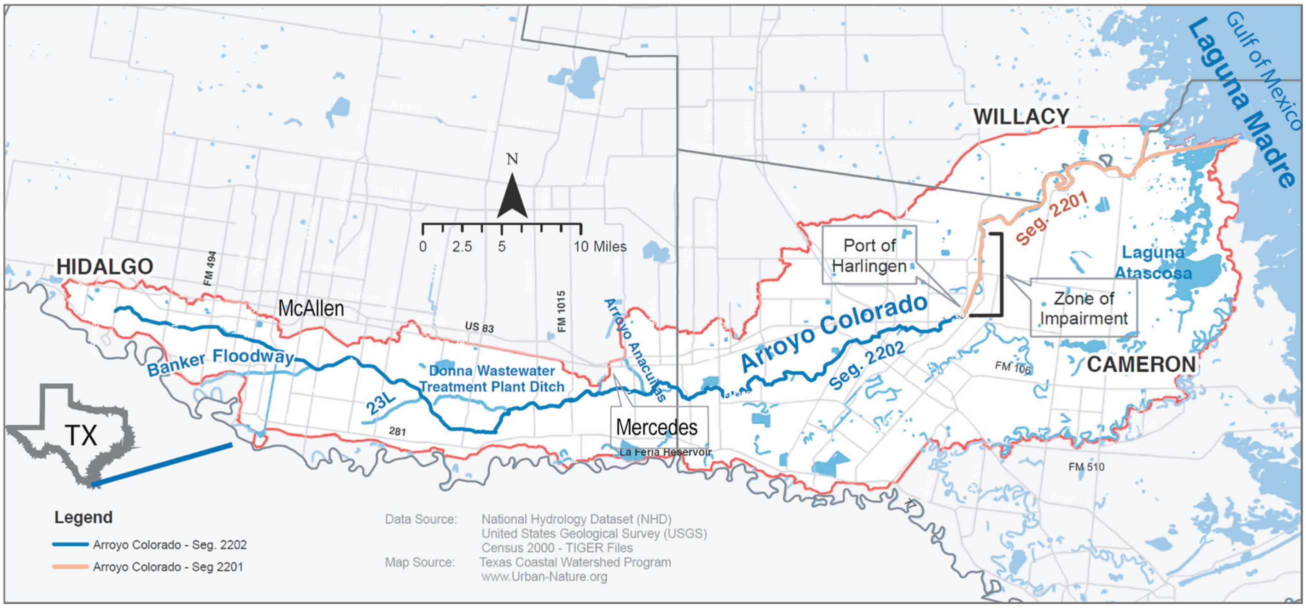

Fig. 1 ACW is located near US and Mexico border and has two impaired segments (Segment 2201 and 2202). Long-term hydrometeorological data are available at Harlingen, Mercedes, and McAllen gauges

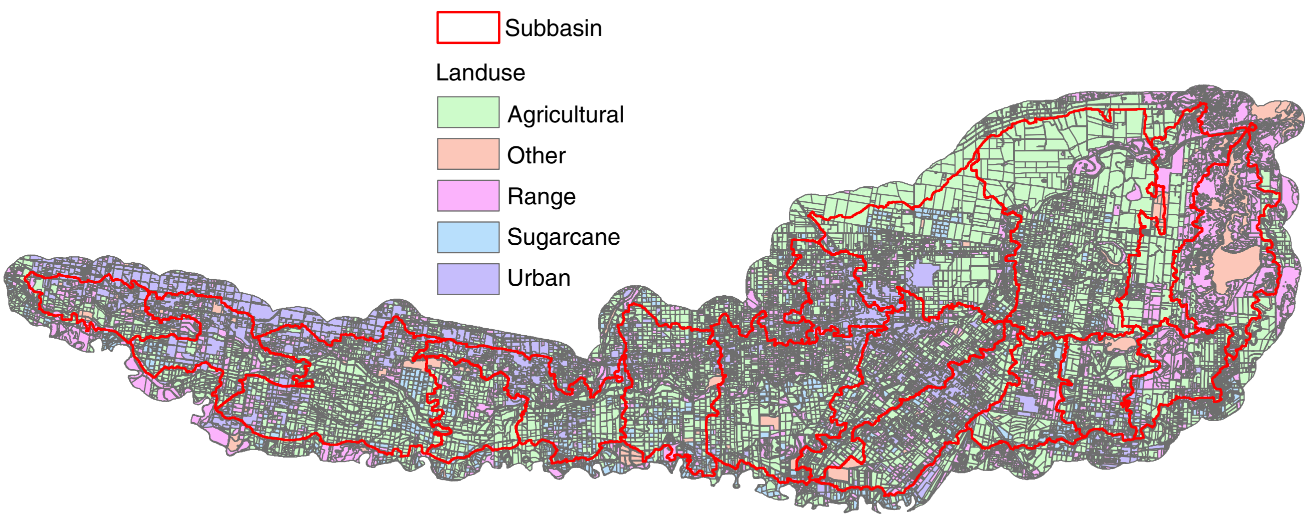

Fig. 2 Land use/land cover map of the ACW. Dominant land use/land cover categories in the watershed are agriculture (54 %), range (18.5 %), urban (12.5 %), other (open waters) (6 %), and sugarcane (4 %). The watershed is divided into 17 subbasins as shown by the *thin solid lines*

sediment, nitrogen, phosphorous using seven years (2000–2006) of data. Data from 1999 was used for spin-up run. Most of the water quality data used for model calibration and testing came from the gauge near Harlingen (see Fig. 1). Figure 3 shows a comparison of observed and predicted daily streamflow and TN loading values at the Harlingen gauge. The daily TN loading was estimated by using daily streamflow and observed TN concentrations. Note that no TN observations were collected in December 2002 when the high-flow period occurred (Fig. 3a). Overall, the fit is deemed satisfactory. A detailed documentation of SWAT calibration and validation methodologies is given in a recent report by Kannan (2012). In the following, the technical background of the three metamodels will be described for completeness.

Methodologies

The ANN metamodel

ANNs have been extensively used in streamflow forecasting and water resources management (e.g., Hsu et al. 1995; ASCE 2000a, b; Moradkhani et al. 2004; Chang and Chang

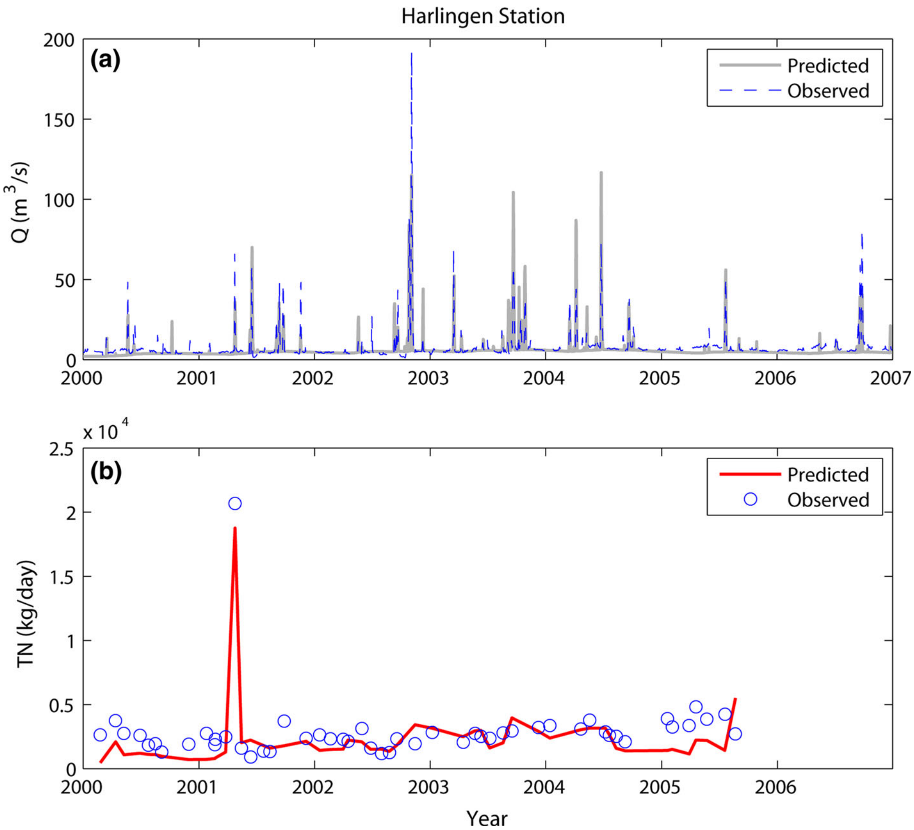

Fig. 3 Predicted and observed **a** streamflow and **b** TN loading during calibration (2000–2003) and validation (2004–2006) periods at Harlingen gauge

2006; Maier et al. 2010). A major strength of neural networks lies in their universal approximation property—an ANN with a single hidden layer can be trained to approximate the causal relationship of any nonlinear dynamic system without a priori assumptions about the underlying physical system (Haykin 1999). In this work, the radial basis function network (RBFN), a type of feedforward neural network, is used to learn the nonlinear input–output dynamics simulated by the SWAT model. As mentioned before, the main focus of metamodeling is replicating model responses in the decision space; thus, the model outputs instead of actual observations are used for training.

Mathematically, radial basis functions represent a family of basis functions expressed as

$$\phi_i = \phi(\|\mathbf{x} - \mathbf{c}_i\|), \tag{1}$$

where $\|\cdot\|$ is typically the Euclidean norm and is used as a measure of distance between any input data sample $\mathbf{x}$ and data center $\mathbf{c}_i$. The data centers play the role of connection weights between the input and hidden layers. One of the most commonly used radial basis functions is the Gaussian kernel,

$$\phi_i(\|\mathbf{x} - \mathbf{c}_i\|) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{c}_i\|^2}{\sigma_i^2}\right), \tag{2}$$

which is completely specified by two parameters, the data center $\mathbf{c}_i$, and spread σ_i. In practice, a global spread parameter σ is often used and is sufficient for most purposes (Park and Sandberg 1991; Demuth et al. 2008). Training of RBFNs typically proceeds in two steps. In the first step, the basis functions connecting the input and hidden layers are built. To avoid overfitting, the hidden neurons and, thus, data centers, are added adaptively until either a predefined training error is reached on the training set or the maximum number of neurons is exceeded. In the second step, weights connecting the hidden and output layers are determined by solving a linear system of equations. The hyperparameters of the RBFN model are the global spread and training error, both are adjusted as part of the training process. We used the Neural Network Toolbox from MATLAB R2010 to build and train the RBFN model offline (Demuth et al. 2008). The hyperparameters were optimized by using the genetic algorithm, a global optimization program that is also available in MATLAB.

Inputs to the RBFN consist of monthly total precipitation (P), minimum ($T_{\min}$) and maximum ($T_{\max}$) temperatures, and average streamflow (Q), which were aggregated/averaged based on daily records collected from the Harlingen, McAllen, and Mercedes gauges (Fig. 1) from 2000 to 2006. Of the entire dataset, 70 % of the data was used for training and the rest for testing. All forcing data have been used in calibrating and validating the original SWAT model. A 1-month-head RBFN was trained for TN loading (i.e., the target variable) using SWAT outputs. In addition to aforementioned hydrometeorological variables, the SWAT-predicted TN loading values from the antecedent months were also used for one-month-ahead training (see Sect. 4.1). ANN is a deterministic method and may be combined with an ensemble method to improve its performance while providing estimates of predictive uncertainty (Sun 2013b). However, the number of required ANNs can be large as the number of uncertain parameters increases. Therefore, the more efficient PCM is used instead for sensitivity studies.

The PCM metamodel

Sensitivity analysis and uncertainty quantification are critical components in all watershed decision-making activities. In a co-learning and co-management environment, in which syntheses are performed jointly and stakeholders participate directly in the decision-making process (Berkes 2009), hands-on experience with the watershed model can be extremely valuable. PCM was developed to support such analytical capability in CGDSS.

As mentioned in the Introduction, PCM is a type of stochastic response surface methods that represent parametric uncertainties as an expansion of orthogonal polynomials of independent random variables. The starting point of PCM is to represent the model output of interest, y, in terms of input variables, such as $\mathbf{X} = \{X_i\}_{i=1}^{n}$, using polynomial chaos expansion (Ghanem and Spanos 1991; Isukapalli et al. 1998). The input variables are transformed into standard random variables, $\boldsymbol{\xi} = \{\xi_i\}_{i=1}^{k}$. The generalized polynomial chaos expansion is then given in the following summation form (Xiu 2010)

$$y = \sum_{i=0}^{\infty} a_i \Psi_i(\boldsymbol{\xi}), \tag{3}$$

in which a_i are deterministic coefficients and $\Psi_i(\boldsymbol{\xi})$ are products of one-dimensional basis polynomials. A one-to-one mapping exists between polynomial families and many commonly used continuous and discrete probability distributions (Xiu 2010). For example, for Gaussian random variables, the optimal family of orthogonal basis polynomials is Hermite polynomial, for which the one-dimensional form is given by

$$H_p(\xi) = (-1)^p e^{\xi^2/2} \frac{d^p}{d\xi^p} e^{-\xi^2}, \tag{4}$$

where p is the degree of Hermite polynomial. For uniformly distributed random variables, the optimal family of orthogonal basis polynomials is Legendre, for which the one-dimensional form is

$$L_p(\xi) = \frac{1}{2^p p!} \frac{d^p}{d\xi^p} \left[\left(\xi^2 - 1\right)^p \right]. \tag{5}$$

The significance of Eq. (3) is that it provides a way to approximate complex functions having uncertain inputs. In practice, the infinite series in Eq. (3) is truncated at finite terms

$$y \approx \sum_{i=0}^{N_p} a_i \Psi_i(\boldsymbol{\xi}), \tag{6}$$

in which N_p is given by

$$N_p = \frac{(n+p)!}{n!p!}, \tag{7}$$

where p is the order of polynomial expansion and n is the number of random inputs. To determine a_i, two methodologies can be used, the least-squares approach and the PCM. Both methods are non-intrusive and use an existing forward model to generate collocation points [i.e., the left-hand side of Eq. (6)]. The least squares approach (Le Maıtre et al. 2002; Eldred 2009) generally performs stratified sampling in the parameter space to form an over-determined system of equations and the recommended number of model runs is at least two times of N_{p}. In contrast, the PCM (Tatang et al. 1997; Li and Zhang 2007) uses a selected subset of polynomial roots to generate exactly N_{p} collocation points. The points are selected so that those having higher probabilities on the corresponding distributions are selected first (Li and Zhang 2007). Once the deterministic coefficients are obtained, the resulting coefficient matrix can be used to estimate the first moments of model output, interpolate model output for any parameter combination in the defined parametric ranges, or to quickly conduct sensitivity analysis (Zheng et al. 2011). If the dimension of parameter space is low (<15), PCM offers a more efficient alternative to Monte Carlo simulation using the original model.

Extensive studies have been performed to evaluate the effect of parametric uncertainty on SWAT output (van Griensven et al. 2006; Rouholahnejad et al. 2012). In general, hydrologic parameters are dominant in controlling water-quality predictions (van Griensven et al. 2006). Kannan et al. (2011) considered the sensitivity of simulated streamflow in the ACW to 15 hydrologic parameters and concluded that the five most influential SWAT parameters are: available water capacity (AWC), soil evaporation compensation factor (ESCO), plant evaporation compensation factor (EPCO), groundwater re-evaporation coefficient (GW_REVAP), and surface runoff lag factor (SURLAG). They found that streamflow is not sensitive to the runoff curve number (CN2), probably due to the short duration of the monsoon/rainy season in the ACW. Table 1 gives the list of five parameters and their ranges used for demonstration of PCM. In accordance with van Griensven

Table 1 List of parameters used in PCM sensitivity study

Name	Definition	Process	Min	Max
ALPHA-BF	Baseflow alpha factor (day)	Groundwater	0.001	1
EPCO	Plant evaporation compensation factor	Evaporation	0.001	1
ESCO	Soil evaporation compensation factor	Evaporation	0.001	1
SOIL_AWC[a]	Available water capacity of the soil layer (mm/mm soil)	Soil	0	1
GW_REVAP	Groundwater revap coefficient	Groundwater	0.02	0.2

[a] These parameters are varied from −50 to 50 % to maintain the spatial relationship (van Griensven et al. 2006)

et al. (2006), we assumed that all parameters vary uniformly in their ranges and thus the optimal orthogonal polynomial family is Legendre. The order of polynomials was set to four on the basis of our previous experience (Sun et al. 2013). The total number of SWAT models runs required for training the PCM with $n = 5$ and $p = 4$ is 126, as indicated by Eq. (7).

Both PCM and RBFN are metamodels that aim to honor the SWAT model outputs as much as possible. However, the downside is that their "parameters" offer little physical interpretation. When it comes to long-term load reduction planning, practitioners may prefer to use models that are physically-based and yet simple to manage. This leads to the PLOAD model, which falls in the category of "low-fidelity models" mentioned in the Introduction.

The PLOAD metamodel

PLOAD, one of the models in US EPA's BASINS suite (http://water.epa.gov/scitech/datait/models/basins/index.cfm), offers a simple, first-order analytical tool for calculating nonpoint-source pollutant loadings (EPA 2001). In our study, PLOAD was selected as a reduced-order model to approximate the SWAT process-level model because of its simplicity, rational nature, and familiarity to most regulators. The key concept in PLOAD is export coefficients that relate long-term (typically annual) pollutant loading rates to land use types

$$L_i = \sum_{j=1}^{N} R_{ij} A_j, \tag{8}$$

where L_i is loadings due to the ith pollutant (i.e., TN in the current case), N is the total number of land use types in a watershed, and R_{ij} and A_j are the pollutant export coefficient and the total area of the j-th land use type, respectively. The latter information was derived from the land use land cover map used by the ACW SWAT model, which divides the watershed into 17 subbasins (Fig. 2). Export coefficients are typically determined by monitoring pollutant loadings from small catchments with a predominant land use or by using field plots to isolate individual land uses; however, the resulting export coefficients may not be applicable to the entire watershed because they do not necessarily represent average land conditions and land use practices over the watershed. Because of the lack of continuous water quality gauges at most subbasins, the SWAT model outputs were used to "train" export coefficients using a linear regression method (Shrestha et al. 2008). In this case, the semi-distributed SWAT model is essentially used as a spatiotemporal interpolator to generate loadings at ungauged subbasin outlets.

Using the SWAT-simulated, average annual loadings at the outlet of the subbasin (i.e., L_i) and the areas of different land uses within that subbasin (i.e., A_j), we obtain 17 equations using Eq. (8), one per subbasin. This leads to a linear system, from which the unknown export coefficients can be estimated using least squares. Recall that ACW has five land use types, including range, urban and industrials, sugarcane, agricultural, and other. We excluded the category other because it mostly represents open waters; thus four unknowns were solved from an over-determined system. We emphasize that the linear regression procedure described here is mainly for extracting information from the Arroyo Colorado SWAT model as initial guesses. End users can easily alter export coefficients from the web user interface, as described in details in the next section.

Results and discussion

RBFN

The RBFN was trained offline using SWAT inputs and outputs. As part of the RBFN model's development, a stepwise selection procedure was used to choose inputs (or predictors). Starting with a complex input structure that includes two antecedent lags (i.e., $t-1$ and $t-2$) for all variables, we removed the input variables one at time to see their relative contribution to the reduction (or increase) of root-mean square errors (RMSE). The selected set of input variables for the one-month-ahead prediction consists of

$$\begin{aligned}&P_{\mathrm{ha}}(t-1), P_{\mathrm{md}}(t-2),\\&T_{\mathrm{min,ha}}(t-1), T_{\mathrm{max,ha}}(t-1), T_{\mathrm{min,mc}}(t-1), T_{\mathrm{max,mc}}(t-1),\\&Q_{\mathrm{ha}}(t-1), Q_{\mathrm{ha}}(t-2),\\&\dot{M}(t-1), \dot{M}(t-2)\end{aligned} \tag{9}$$

where the subscripts ha, md, and mc denote the Harlingen, Mercedes, and McAllen gauges, respectively (see Fig. 1); $\dot{M}$ represents monthly TN loading from the watershed, and

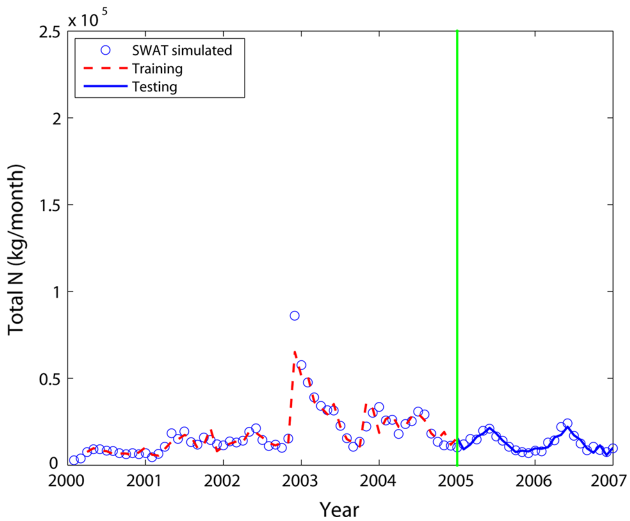

Fig. 4 Training and testing of RBFN metamodel for TN output using SWAT results (*circles*). Training and testing periods separated by *vertical line*

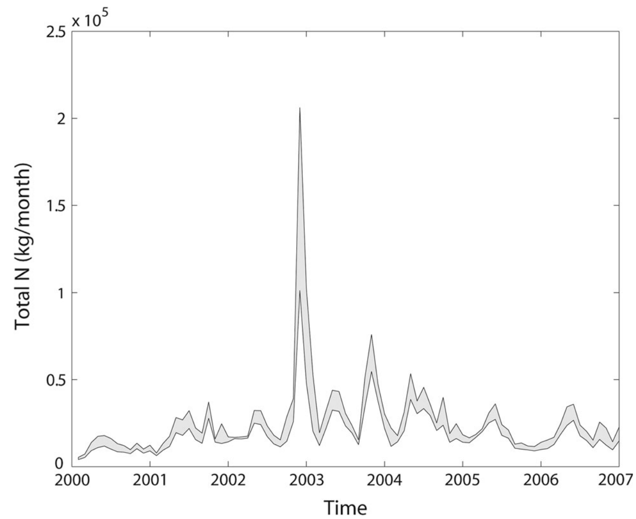

Fig. 5 Confidence bounds (95 %) obtained by running 2,000 Monte Carlo simulations using the PCM metamodel

all other symbols are defined under Sect. 3.1. Note that the precipitation input for the upstream Mercedes gauge is lagged behind the Harlingen gauge by 1 month.

The final RBFN for TN prediction consists of 30 hidden neurons. Figure 4 plots the TN loadings simulated by the RBFN in both the training and testing periods. The training period includes a sufficient number of variation patterns, including a high loading peak near the end of 2002. Although the trained RBFN does not fully capture the high peak in 2003, overall it gives satisfactory performance. The Nash–Sutcliffe Efficiency (NSE) calculated over the testing period is 0.78, which is acceptable at the monthly scale.

PCM

The PCM coefficients a_i in Eq. (6) are calculated offline. As mentioned in Sect. 3.2, results from a total of 126 simulation runs were used to formulate a linear system of equations for determining the PCM coefficients. Each SWAT run takes about 13.8 min to complete on a Dell Precision T3500 Workstation equipped with an Intel® Xeon processor, in addition to 3.5 min of pre-processing time required to generate SWAT input files. The computational demand required by the sensitivity or uncertainty analysis highlights the need for metamodels. Using PCM, we performed Monte Carlo simulation of TN loading from the entire watershed with 2000 realizations, and the total running time was under 2 s, making the technology well suited for online deployment. Figure 5 shows the confidence bounds derived from the Monte Carlo simulation using PCM metamodel.

PLOAD

PLOAD export coefficients were estimated by using linear regression, as described in Sect. 3.3. The results for TN are 3.3, 1.5, 5.6, 3.2 (in kg/ha/year) for range, urban, sugarcane, and agricultural land use types, respectively. We compared the results to published values for TN (e.g., Beaulac and Reckhow 1982; EPA 2001; Shrestha et al. 2008) and found they are generally within the range of published values. The average monthly TN loading from the entire watershed was estimated to be 3.9×10^4 kg/month (or 4.7×10^5 kg/year) using the export coefficients. As an independent check, we turned to the online version of the US Geological Survey's SPARROW (SPAtially Referenced Regression on Watershed attributes) model (http://cida.usgs.gov/sparrow/map.jsp?model=35), which is another popular long-term pollutant loading model. The ACW long-term TN loading estimated by the SPARROW model is 2.7×10^5 kg/year. The particular SPARROW results cited here were based on water quality measurements around 2002. The PLOAD estimate is more conservative than the average TN loading simulated by SWAT (see Fig. 4) and SPARROW. Given its simple model structure, the PLOAD model is still valuable. Moreover, we emphasize that the export coefficients reported here only provide a starting point for multiple groups of people to perform high-level planning and scenario analysis. During the watershed management and decision making process, both export coefficients and land use types can be modified in a collaborative manner to reflect more up-to-date information.

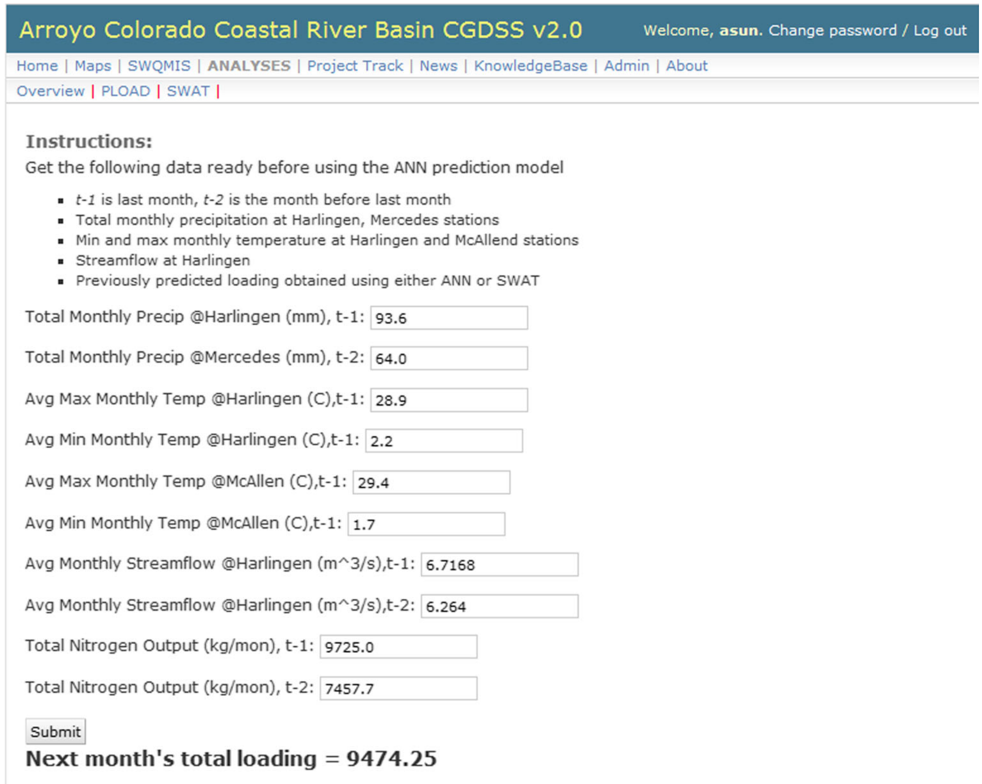

Fig. 6 Screenshot of the web interface for 1-month-ahead TN loading prediction using RBFN

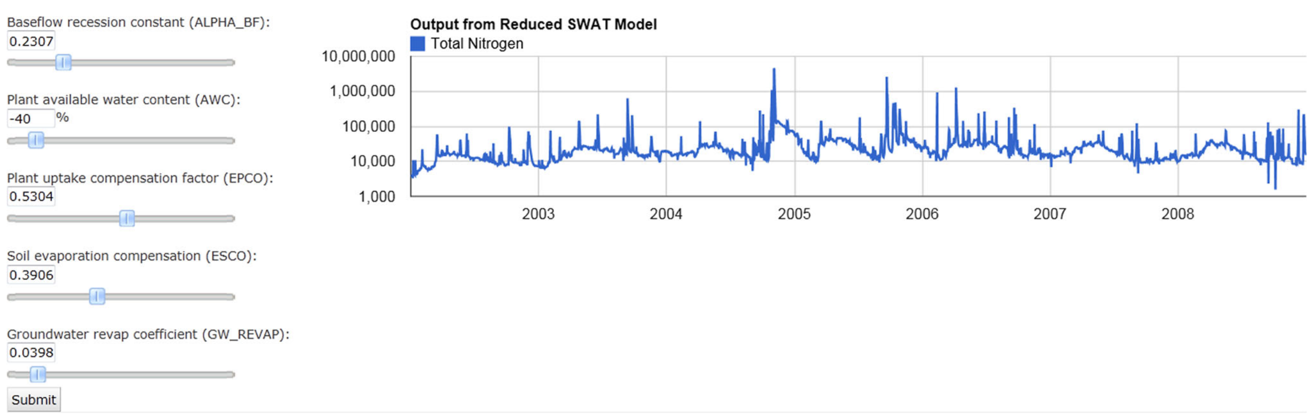

Fig. 7 Screenshot of the web interface used for performing sensitivity study with PCM

Web implementation

This subsection summarizes implementation of the web-based metamodels. In this work, an open-source content management system Django (https://www.djangoproject.com) is adopted. Django provides a number of tools for quick implementation and deployment of web applications. Additional server-side (Python) and client-side (Java) scripts were used to enable web-GIS functionalities and to implement the metamodels (Sun 2013a).

For the RBFN metamodel, the connection weights between input and hidden layers (i.e., centers of radial basis functions) and between hidden and output layers are stored on the web server. CGDSS provides a web form for the user to specify antecedent forcing variable information. On submission of the form, the backend Python scripts calculate the total loadings using the saved RBFN parameters. The web interface for RBFN is shown in Fig. 6.

The implementation of the PCM model is similar to that of the RBFN model—the PCM coefficient matrix is stored on the server. A Python script is used to calculate basis function values based on user inputs and the PCM output is obtained through a matrix–vector multiplication; both tasks can be executed in real time. Figure 7 shows the web

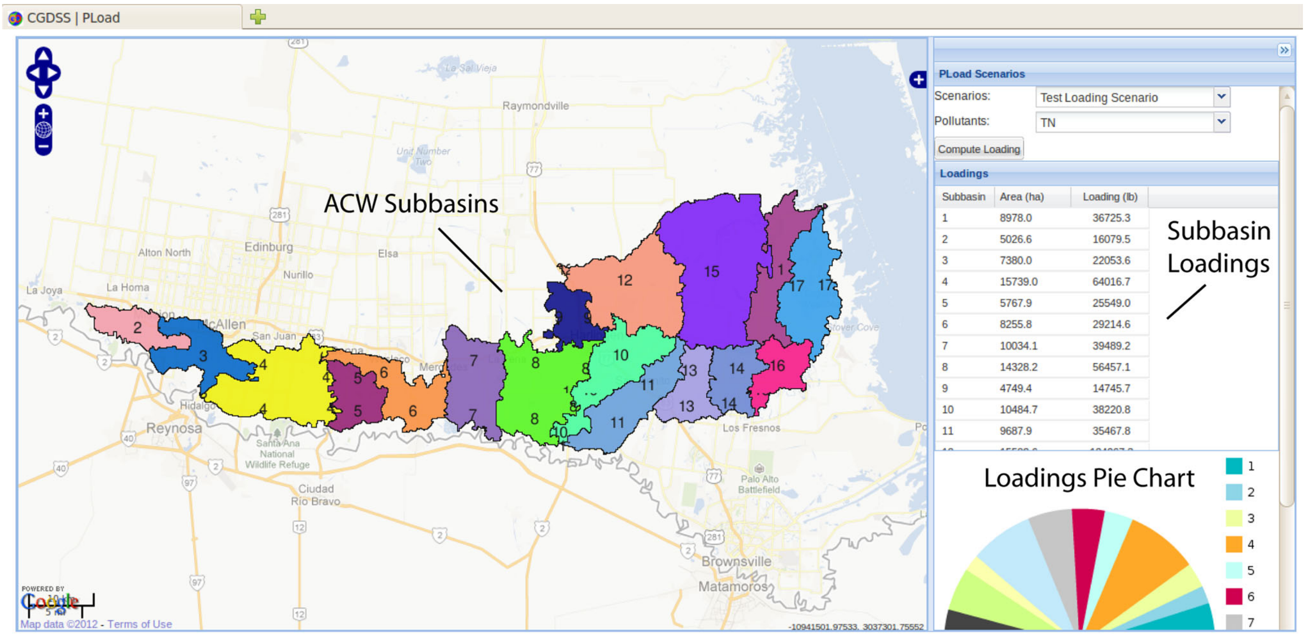

Fig. 8 Screenshot of the PLOAD module, which shows TN loadings by subbasins

interface designed for performing sensitivity analysis using the PCM. A user may change any of the five uncertain parameters to examine the sensitivity of loadings. The RBFN and PCM web interfaces are mainly used for rapid prototyping in this work. In the future, more user friendly features will be added to improve usability. Although not done here, both RBFN and PCM can be configured to generate outputs at the subbasin level.

Implementation of the PLOAD metamodel is more involved because of the extra geospatial processing. A script was developed for the user to upload a land use/land cover shapefile and subsequently store in an object-oriented spatial database. However, this only needs to be done once for each land use/land cover map. The user can then create different loading scenarios using the web interface, and the land use types in each scenario may be assigned using different export coefficient values. The input fields are initially populated using values obtained through the linear regression procedure as discussed in Sect. 4.3. As part of the joint decision-making process, users can choose to set the export coefficients to different values to see the impact on load reduction. A full documentation of the PLOAD web interface implementation is provided in Sun (2013a). Figure 8 shows a screen capture of the PLOAD metamodel interface that displays TN outputs by subbasins.

Summary and conclusions

Metamodeling represents a fast expanding field of research in system modeling, uncertainty quantification, and risk analysis. Metamodeling can be especially suitable to support visual analytics and such aspect has not been extensively explored for web-based EDSS. In this study, we developed three metamodels to support analytical functionalities of a web-based EDSS. Like many other environmental decision-making processes, the watershed management and decision-making process is often rife with controversies and challenges. Our web application aims to mitigate some of these challenges by making the watershed decision making process more collaborative, transparent, and educational.

The nature of watershed decision making process makes it difficult to develop a one-size-fit-all metamodel. Our experience through this project indicates that it is more fruitful to develop a set of metamodels to fulfill different decision-making needs than to develop a single model for all use cases. If a process-level model is already adopted by regulators and stakeholders, then metamodeling provides a decision space in which end users can experiment with different scenarios and observe the consequences. If the process-level model development is still ongoing, all metamodels can be easily re-trained after a new version of process-level model becomes available. Thus, it is crucial for a metamodel to replicate the process-level model as much as possible. On the other hand, modelers and developers should not lose sight of the overarching objective of the decision-making activity and the practical needs of end users in the process of fulfilling scientific research needs. Arguably the more technical details embedded in a model the more difficult it is for laypeople to grasp. Therefore, the watershed modeling is hierarchical:

at the top-level, abstraction models are used for long-term planning (e.g., PLOAD); at the middle level, more specialized metamodels are used for near-term, data-driven forecasting (e.g., ANN); finally, at the lowest level, process-level models (e.g., SWAT) are used to support all higher-level models. The level of users can be classified similarly when they register with the EDSS. As a result, different metamodels may be exposed to different end user groups, depending on their role and interests in the decision making process.

Acknowledgments Funding for this research was partially provided by NASA (Grant NNX09AR63G) through a subcontract from Southwest Research Institute. Mr. Roger M. Miranda is a Ph.D. student in Environmental Policy at the University of Texas Austin. The authors are grateful to the two anonymous reviewers for their constructive comments. We thank Drs. Narayanan Kannan (currently at Texas Tarleton State University) and Jaehak Jeong at Texas Blackland Research and Extension Center for providing the SWAT model.

References

Arroyo Colorado WPP (2007) Arroyo Colorado Watershed Protection Plan. http://www.arroyocolorado.org/watershed-protection-plan. Accessed 12 May 2014

ASCE (2000a) Committee on the application of ANNs in hydrology artificial neural networks in hydrology I: preliminary concepts. J Hydrol Eng 5:115–123

ASCE (2000b) Committee on the application of ANNs in hydrology artificial neural networks in hydrology, II: hydrologic application. J Hydrol Eng 5:124–137

Beaulac MN, Reckhow KH (1982) An examination of landuse-nutrient export relationships. J Am Water Resour As 18:1013–1024

Berkes F (2009) Evolution of co-management: role of knowledge generation, bridging organizations and social learning. J Environ Manage 90:1692–1702

Blanning RW (1975) Construction and implementation of metamodels. Simulation 24:177–184

Castelletti A, Galelli S, Ratto M, Soncini-Sessa R, Young PC (2012a) A general framework for dynamic emulation modelling in environmental problems. Environ Model Softw 34:5–18

Castelletti A, Galelli S, Restelli M, Soncini-Sessa R (2012b) Data-driven dynamic emulation modelling for the optimal management of environmental systems. Environ Model Softw 34:30–43

Chang FJ, Chang YT (2006) Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. Adv Water Resour 29:1–10

Chen CW, Herr JW, Goldstein RA, Ice G, Cundy T (2005) Retrospective comparison of watershed analysis risk management framework and hydrologic simulation program FORTRAN applications to mica creek watershed. J Environ Eng Asce 131:1277–1284

Demuth H, Beale M, Hagan M (2008) Neural network toolbox™ user's guide. The MathWorks Inc., Natick

Eldred M (2009) Recent advances in non-intrusive polynomial chaos and stochastic collocation methods for uncertainty analysis and design. 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Palm Springs, California

EPA (2001) PLOAD version 3.0: An arcview GIS tool to calculate nonpoint sources of pollution in watershed and stormwater projects. User's Manual. EPA, Washington, DC

Fang K, Li R, Sudjianto A (2006) Design and modeling for computer experiments. Chapman & Hall/CRC, Boca Raton

Friedman JH (1991) Multivariate adaptive regression splines. Ann Stat 19:1–67

Gassman PW, Reyes MR, Green CH, Arnold JG (2007) The soil and water assessment tool: historical development, applications, and future research directions. T Asabe 50:1211–1250

Ghanem R, Spanos PD (1991) Stochastic finite elements: a spectral approach. Springer, New York

Haykin SS (1999) Neural networks: a comprehensive foundation. Prentice Hall, Upper Saddle River

Helton JC, Johnson JD, Sallaberry CJ, Storlie CB (2006) Survey of sampling-based methods for uncertainty and sensitivity analysis. Reliab Eng Syst Safe 91:1175–1209

Hsu KL, Gupta HV, Sorooshian S (1995) Artificial neural-network modeling of the rainfall-runoff process. Water Resour Res 31:2517–2530

Isukapalli SS, Roy A, Georgopoulos PG (1998) Stochastic response surface methods (SRSMs) for uncertainty propagation: application to environmental and biological systems. Risk Anal 18:351–363

Kannan N (2012) SWAT modeling of the Arroyo Colorado Watershed, Technical Report No. 426. Texas Water Resources Institute, Temple, TX

Kannan N, Jeong J, Srinivasan N (2011) Hydrologic modeling of a canal-irrigated agricultural watershed with irrigation best management practices: case study. J Hydrol Eng 16:746–757

Kannan N, Omani N, Miranda R (2014) Water quality modeling of an agricultural watershed with best management practices. Int J Res Eng Technol 3:553–564

Kleijnen JPC (2008) Design and analysis of simulation experiments. Springer, New York

Le Maıtre OP, Reagan MT, Najm HN, Ghanem RG, Knio OM (2002) A stochastic projection method for fluid flow: II. Random process. J Comput Phys 181:9–44

Li H, Zhang D (2007) Probabilistic collocation method for flow in porous media: comparisons with other stochastic methods. Water Resour Res 43:W09409

Maier HR, Jain A, Dandy GC, Sudheer KP (2010) Methods used for the development of neural networks for the prediction of water resource variables in river systems: current status and future directions. Environ Model Softw 25:891–909

Matott LS, Babendreier JE, Purucker ST (2009) Evaluating uncertainty in integrated environmental models: a review of concepts and tools. Water Resour Res 45:W06421. doi:10.1029/2008WR007301

Matthies M, Giupponi C, Ostendorf B (2007) Environmental decision support systems: current issues, methods and tools. Environ Model Softw 22:123–127

Moradkhani H, Hsu K-L, Gupta HV, Sorooshian S (2004) Improved streamflow forecasting using self-organizing radial basis function artificial neural networks. J Hydrol 295:246–262

Park J, Sandberg IW (1991) Universal approximation using radial-basis-function networks. Neural Comput 3:246–257

Raines TH, Miranda RM (2002). Simulation of flow and water quality of the Arroyo Colorado, Texas, 1989–99. United States Geological Survey—Water Resources Investigations Report, No: 02-4110. USGS

Razavi S, Tolson BA, Burn DH (2012) Review of surrogate modeling in water resources. Water Resour Res 48:W07401. doi:10.1029/2011WR011527

Rouholahnejad E, Abbaspour KC, Vejdani M, Srinivasan R, Schulin R, Lehmann A (2012) A parallelization framework for calibration of hydrological models. Environ Model Softw 31:28–36

Santner TJ, Williams BJ, Notz W (2003) The design and analysis of computer experiments. Springer, New York

Shan S, Wang GG (2010) Survey of modeling and optimization strategies to solve high-dimensional design problems with computationally-expensive black-box functions. Struct Multidiscip Optim 41:219–241

Shrestha S, Kazama F, Newham LTH (2008) A framework for estimating pollutant export coefficients from long-term in-stream water quality monitoring data. Environ Model Softw 23:182–194

Storlie CB, Swiler LP, Helton JC, Sallaberry CJ (2009) Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. Reliab Eng Syst Safe 94:1735–1763

Sun AY (2013a) Enabling collaborative decision making in watershed management using cloud computing services. Environ Model Softw 41:93–97

Sun AY (2013b) Predicting groundwater level changes using GRACE data. Water Resour Res 49:1–13

Sun AY, Zeidouni M, Nicot JP, Lu Z, Zhang D (2013) Assessing leakage detectability at geologic CO_2 sequestration sites using the probabilistic collocation method. Submitted. Adv Water Resour, 49–60

Tatang MA, Pan W, Prinn RG, McRae GJ (1997) An efficient method for parametric uncertainty analysis of numerical geophysical models. J Geophys Res 102:21925–21932

van Griensven A, Meixner T, Grunwald S, Bishop T, Di Luzio M, Srinivasan N (2006) A global sensitivity analysis tool for the parameters of multivariable catchment models. J Hydrol 324:10–23

Vapnik V (1998) The support vector method of function estimation. In: Nonlinear Modeling, pp 55–85

Xiu D (2010) Numerical methods for stochastic computations: a spectral method approach. Princeton University Press, Princeton

Young PC, Ratto M (2009) A unified approach to environmental systems modeling. Stoch Env Res Risk A 23:1037–1057

Zheng Y, Wang WM, Han F, Ping J (2011) Uncertainty assessment for watershed water quality modeling: a probabilistic collocation method based approach. Adv Water Resour 34:887–898