3B2v8.06a/w (Dec 5 2003).51c XML:ver:5.0.1 CAGEO : 1456 Prod.Type:FTP pp.1–16(col.fig.:NIL) ED:PradeepKumarR. PAGN:Uday SCAN:S.Geetha

ARTICLE IN PRESS

ELSEVIER

Computers & Geosciences ▮ (▮▮▮▮) ▮▮▮–▮▮▮

COMPUTERS & GEOSCIENCES

www.elsevier.com/locate/cageo

Assessment of a stochastic interpolation based parameter sampling scheme for efficient uncertainty analyses of hydrologic models

Faisal Hossain*, Emmanouil N. Anagnostou

Department of Civil and Environmental Engineering, Tennessee Technological University, Box 5015, 1020 Stadium Drive, Cookeville, TN 38505-0001, USA

Received 11 July 2004; received in revised form 10 October 2004; accepted 5 November 2004

Abstract

This study assesses a stochastic interpolation based parameter sampling scheme for efficient uncertainty analyses of stream flow prediction by hydrologic models. The sampling scheme is evaluated within the generalised likelihood uncertainty estimation (GLUE; Beven and Binley, 1992) methodology. A primary limitation in using the GLUE method as an uncertainty tool is the prohibitive computational burden imposed by uniform random sampling of the model's parameter distributions. Sampling is improved in the proposed scheme by stochastic modeling of the parameters' response surface that recognizes the inherent non-linear parameter interactions. Uncertainty in discharge prediction (model output) is approximated through a Hermite polynomial chaos approximation of normal random variables that represent the model's parameter (model input) uncertainty. The unknown coefficients of the approximated polynomial are calculated using limited number of model simulation runs. The calibrated Hermite polynomial is then used as a fast-running proxy to the slower-running hydrologic model to predict the degree of representativeness of a randomly sampled model parameter set. An evaluation of the scheme's improvement in sampling is made over a medium-sized watershed in Italy using the TOPMODEL (Beven and Kirkby, 1979). Even for a very high (8) dimensional parameter uncertainty domain the scheme was consistently able to reduce computational burden of uniform sampling for GLUE by at least 15–25%. It was also found to have significantly higher degree of consistency in sampling accuracy than the nearest neighborhood sampling method. The GLUE based on the proposed sampling scheme preserved the essential features of the uncertainty structure in discharge simulation. The scheme demonstrates the potential for increasing efficiency of GLUE uncertainty estimation for rainfall–runoff models as it does not impose any additional structural or distributional assumptions.

Keywords: Uncertainty estimation; Hydrologic models; Stochastic interpolation; Hermite polynomial chaos expansion; Parameter sampling; GLUE

1. Introduction

Due to ever-increasing computing power, the fully random Monte Carlo (MC) sampling is nowadays considered the preferred method for uncertainty analy-

*Corresponding author. Tel.: +1 931 372 3257; fax: +1 931 372 6239.
E-mail address: fhossain@tntech.edu (F. Hossain).

0098-3004/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cageo.2004.11.001

sis. Other reasons for the wide-spread preference of MC techniques are their lack of restrictive assumptions and completeness in sampling the error structure of the random variables (Beven and Freer, 2001; Beck, 1987; Kremer, 1983). MC sampling can also bypass several limitations of analytical techniques (such as first-order approximation methods; Bras and Rodriguez-Iturbe, 1993). An uncertainty estimation technique called Generalised Likelihood Uncertainty Estimation (GLUE) (Beven and Binley, 1992) is one such MC-based tool that can be employed to assess a hydrologic model's predictive uncertainty. This method evaluates the simulation results for each randomly sampled model parameter set against some observed data through a likelihood value. The method is originally founded on the principles of generalized sensitivity analysis (GSA) of Spear and Hornberger (1980). Because its structure is rooted in Bayesian theory, GLUE also allows blending of prior and current information for improved a posteriori inferences. While GLUE is not the only uncertainty assessment tool (Misirli et al., 2003; Thiemann et al., 2001; Tyagi and Haan, 2001; Krzysztofowicz, 2000; Young and Beven, 1994), it is one of the few convenient techniques currently available (Beven and Freer, 2001). GLUE has therefore found extensive application in the assessment of predictive uncertainty of many hydrologic variables like stream flow, flood inundation, ground water flow, land surface fluxes, etc. (Schulz and Beven, 2003; Christaens and Feyen, 2002; Beven and Freer, 2001; Schulz et al., 2001; Romanowicz and Beven, 1998; Franks et al., 1998; Franks and Beven, 1997; Freer et al., 1996; among many others). Recently, the GLUE technique has also proved to be a powerful tool in understanding the implications of remotely sensed rainfall error adjustment on flood prediction uncertainty (Hossain et al., 2004).

However, the GLUE method requires analysis of multiple simulation scenarios based on uniform random sampling of the model parameter hyperspace. This is considered a significant drawback of the scheme, as this requirement can be computationally prohibitive for physically complex hydrologic models that are distributed (Bates and Campbell, 2001; Beven and Binley, 1992). Beven and Binley (1992) have argued in detail that the assumption of uniform distribution is unlikely to prove critical for GLUE. Freer et al. (1996) have further justified uniform random sampling because it makes the GLUE procedure simple to implement and avoids the necessity to sample from some multivariate set of correlated distributions which is often very difficult to justify from observed data.

Nevertheless, the drawback of uniformity assumption in GLUE magnifies tremendously for hydrologic models when large number of parameters are involved. This is particularly evident if we consider the fact that, as computing power increases, the agenda for scientific inquiry correspondingly widens to take advantage of this increased power. Over the last decade, a review of the progression of literature reveals to us the following the realities: (1) more complex, physically-based and slow-running models are on the rise; (2) the time period and time step of scientific investigations are increasing and decreasing, respectively; (3) study regions are becoming larger (from small-sized basins to continental and global studies). For example, in an uncertainty assessment study involving an event-based distributed hydrologic model applied to a very small ($3.9\,km^2$) watershed with only four parameters, Beven and Binley (1992) reported the computing burden of GLUE to be 'significant' (with respect to the computing power that was available a decade ago). For 500 realizations of the model, 30–60 h of computing time were required by a large parallel computing system. With more increased computational power, GLUE has recently been applied to a fully-distributed and physically-based hydrologic model MIKE-SHE (Abbott et al., 1986; Christaens and Feyen, 2002). Yet, Christaens and Feyen (2002) reported therein a 50% loss in computing time due to model execution of unacceptable runs by uniform sampling.

In response to the computational burden imposed by MC-type uncertainty techniques (such as GLUE), researchers have strived to develop numerical schemes for efficient parameter sampling of hydrologic models. Kuczera and Parent (1998) and Bates and Campbell (2001) have explored the use of Markov Chain Monte Carlo (MCMC) methods for more efficient parameter uncertainty analyses. Bates and Campbell (2001) however reported that MCMC methods cannot be used as a blackbox—considerable care is required in its implementation when models have large number of parameters. A further criticism made by Beven and Freer (2001) was that MCMC methods can rarely be useful in making considerable savings in computing time when the model response surface with respect to parameters is not well defined and has the presence of multiple local maxima or plateux. Christaens and Feyen (2002) employed the Latin Hypercube Sampling (LHS) method to accelerate parameter sampling for MIKE-SHE model. However, LHS is based on the assumption of monotonicity of model output in terms of input parameters, in order to be unconditionally guaranteed of accuracy with an order of magnitude fewer runs than uniform random sampling (McKay et al., 1979; Iman et al., 1981). Hence, for hydrologic models, which are replete with multiple minima and maxima in the response surface (Duan et al., 1992), LHS can rarely be expected to perform to its full potential.

The present study is therefore motivated by the need to make the GLUE parameter sampling more efficient for hydrologic (i.e., rainfall–runoff) models. Such a technique should not impose additional structural or distributional assumptions that may otherwise compromise the inherent simplicity and validity of the GLUE method. We hypothesize that the presence of a complex parameter-output response surface is a manifestation of the inherent non-linear deterministic (chaotic) dynamics commonly observed in natural systems. Recently, much convincing evidence has been provided in this regard to promote this hypothesis (see Faybishenko, 2004; and Sivakumar 2004, for a review). In the current state of the art, GLUE would therefore require a stochastic and non-linear interpolator (hereafter called *interpolator*) for the model's complex parameter-output response surface. This *interpolator* could then act as a proxy to the slow running model and potentially identify the regions of high likelihood values of the parameter-output response surface. In this study we have chosen to develop a parameter sampling scheme that stochastically interpolates (non-linearly) the complex parameter-output response surface. Interpolation is based on 'Hermite Polynomial' (HP) chaos expansion that follows from the "*Theory of Homogeneous Chaos*" (Wiener, 1938). We do not demonstrate the presence or absence of chaotic behavior in this study. However, we are encouraged by the recent well-documented discovery of chaos in both streamflow and rainfall processes (Sivakumar et al., 2001a,b; Sivakumar 2000; Jayawardena and Lai, 1994). Basic concepts of our proposed scheme are derived from an uncertainty estimation tool originally developed by Isukapalli and Georgopoulos (1999). The evaluation of the *interpolator* within the GLUE framework is considered an unexplored topic in current literature on uncertainty estimation of rainfall–runoff models. An application is demonstrated on a medium-sized watershed in Northern Italy called Posina involving a 3-month-long hydrologic time series of rainfall and stream flow.

The study is organized in the following manner. In Section 2, a brief description of the watershed, data and the hydrologic model used in this study are discussed. Section 3 describes the GLUE method based on uniform parameter sampling. Section 4 provides the theoretical formulation of the *interpolator* and its method of employment with GLUE. Section 5 describes the simulation framework for assessing the *interpolator*. Section 6 provides comparisons of the *interpolator* based GLUE (hereafter called *interpolator-GLUE*) with traditional uniform sampling based GLUE (hereafter called *uniform-GLUE*). The *interpolator* sampling scheme is also compared with the nearest-neighborhood parameter sampling technique proposed earlier by Beven and Binley (1992) for computationally-challenged situations. Finally Section 7 presents the conclusions and further extensions that may extend capabilities of the *interpolator*.

2. Watershed, data and hydrologic model

The watershed chosen for this study (named Posina) is located in Northern Italy, close to Venice (Fig. 1, right panel). Posina has an area of 116 km^2 and altitudes ranging from 2230 to 390 m at the outlet (Fig. 1, left panel). Within a radius of 10 km from the center of the watershed there is a network of seven rain gauges providing representative estimates of the basin-averaged hourly rainfall. Posina is 68% forested thereby saturation-excess is the main rainfall–runoff generation mechanism of the basin.

The hydrologic data comprising rainfall and streamflow for Posina spanned a period from August 1, 1992 to October 31, 1992 totaling 2208 time steps at the hourly interval (Fig. 2). For estimation of potential evapotranspiration from the watershed, coincident meteorological data were available from a weather station located within 50 km of the watershed. A major storm

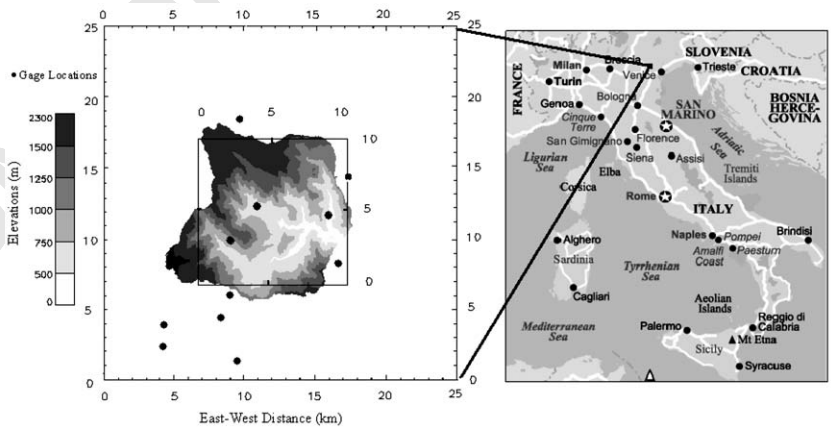

Fig. 1. Geographic location of Posina watershed (right panel), and watershed elevation map (left panel) overlaid by rain gauge network locations (in solid circles).

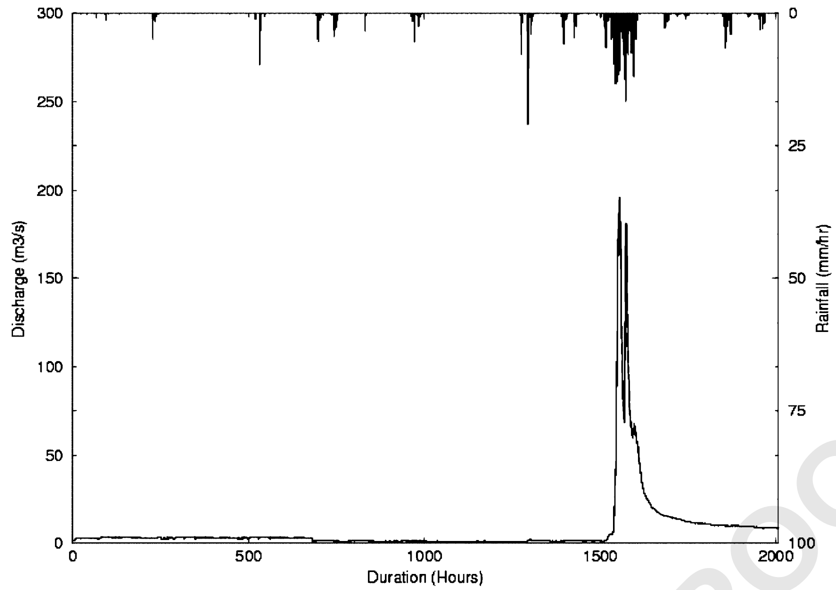

Fig. 2. Streamflow hydrograph (lower axis) and rainfall hyetograph (upper axis) for Posina from August 1 to October 31, 1992.

event took place from October 2 to October 7, 1992 and was associated with catastrophic flooding in the surrounding area (Fig. 2). The hydrologic data is considered particularly appropriate for the study of parameter sampling of hydrologic models because the period spans both dry (unsaturated) and wet (saturated) conditions of the watershed. Since baseflow (about 80% of timeseries) and surface runoff (about 20% of time-series) are adequately represented, the hydrologic data can be considered sufficiently long to characterize the complete structure of a model's parameter uncertainty for the watershed. The entire period of the hydrologic time-series was considered for rainfall–runoff simulation in this study.

The topographic index model (TOPMODEL) (Beven and Kirkby, 1979) was chosen to simulate the rainfall–runoff processes of the Posina watershed. This model makes a number of simplifying assumptions about the runoff generation processes that are thought to be reasonably valid in this wet, humid watershed. TOPMODEL is a semi-distributed watershed model that can simulate the saturation-excess mechanism of storm-runoff generation and incorporates the effect of topography on flow paths. The model is premised on the following two assumptions: (1) the dynamics of the saturated zone can be approximated by successive steady state representations; and (2) the hydraulic gradient of the saturated zone can be approximated by the local surface topographic slope. The generated runoff is routed to the main channel using an overland flow delay function. The main channel routing effects are considered using an approach based on an average flood wave velocity for the channel network (Beven and Kirkby, 1979; Beven et al., 1995). The major parameters of TOPMODEL are as follows: (1) *SZM*—the exponential decay rate of soil hydraulic properties with depth, (m); (2) *SR*0—the initial value of root zone deficit, (m); (3) *SRMAX*—the maximum storage capacity of the root zone, interpreted here as the soil moisture at field capacity, (m); (4) *XK*0—the vertical hydraulic conductivity, ($\mathrm{m\,h^{-1}}$); (5) T_0—the lateral transmissivity, interpreted here as the mean of $\ln(T_0)$, $\ln(\mathrm{m^2\,h^{-1}})$; (6) *TD*—the time delay parameter used to simulate the vertical unsaturated drainage flux, ($\mathrm{h\,m^{-1}}$); (7) *CHV*—the main channel flow velocity ($\mathrm{m\,h^{-1}}$); and (8) *RV*—the overland flow velocity ($\mathrm{m\,h^{-1}}$). The model was run at hourly intervals using basin-averaged rainfall input and considering homogeneous soils all over the watershed. We justify soil homogeneity considering the insignificant size of the watershed ($<500\,\mathrm{km^2}$) compared to the scale at which regional geology is expected to vary. TOPMODEL was initialized for the study period assuming that the first observed discharge is baseflow (see Fig. 2) and proportional to the initial subsurface storage deficit of the watershed (i.e., *SR*0). It should be noted that TOPMODEL, being a conceptual-type model, not all parameters are physically meaningful to be derived directly from in situ measurements. Hence the majority of the parameters were determined through calibration with rainfall-stream flow data, which is a common practice for hydrologic models today (Duan et al., 2003). Further information on the model can be found in (Beven et al., 1995) while previous TOPMO-

DEL applications on the Posina watershed are documented in Hossain et al. (2004).

3. Generalised likelihood uncertainty estimation (GLUE)

GLUE is based on Monte Carlo simulation: a large number of model runs are made, each with random parameter values selected from probability distributions for each parameter. GLUE assumes uniform probability distribution of all model parameters for reasons already alluded in Section 1. The acceptability of each run is assessed by comparing predicted to observed hydrologic measurement through some chosen likelihood measure. Runs that achieve a likelihood below a certain threshold may then be rejected as '*non-behavioral*' (accepted runs are referred to as '*behavioral*'). The likelihoods of these *non-behavioral* parameters are set to zero and are thereby removed from the subsequent analysis. Following the rejection of *non-behavioral* runs, the likelihood weights of the retained (*behavioral*) runs are rescaled so that their cumulative total is one (Freer et al., 1996). In this study the GLUE method was applied to uncertainty estimation of discharge (streamflow) prediction by TOPMODEL at the basin outlet. Thus at each time step the predicted discharge from the retained runs are likelihood weighted and ranked to form a likelihood-weighted cumulative distribution function of discharge from which chosen quantiles can be selected to represent model uncertainty. While GLUE is based on a Bayesian conditioning approach, the likelihood measure is achieved through a goodness of fit criterion as a substitute for a more traditional likelihood function. The likelihood associated with a particular parameter value may therefore be expected to vary depending on the values of the other parameters, and there may be no clear optimum parameter set.

Because GLUE allows the choice to be subjective, two likelihood measures were employed in this study for evaluating the proposed *interpolator* sampling scheme. These are (1) the classical index of efficiency (Nash and Sutcliffe, 1970), hereafter referred to as *Efficiency index*; and (2) a weighted peak runoff–runoff volume index (hereafter referred to as *PR–RV* Index). We define the *Efficiency* Index as follows:

$$\mathit{Efficiency\ \ index} = \left[1 - \frac{\sigma_e^2}{\sigma_o^2}\right], \tag{1}$$

where, σ_e is the variance of errors and σ_o, the variance of observations. The *PR–RV* Index is defined as the weighted average of percentage error in Peak Runoff (*PR*) and Runoff Volume (*RV*) where 60% weight is given to *PR* error and 40% to *RV* error. Because the discharge data had only one major storm event spanning 20% of the total timeseries, we observed the error in Time to Peak (*TP*) to be relatively less sensitive to the goodness of fit (i.e., root mean square of error) of simulations. Hence error in *TP* was not considered herein. The error in the hydrologic parameters (*PR* and *RV*) is defined as follows:

$$PR\ \text{error}(\%) = \left|\frac{\text{Peak runoff}_{\text{obs}} - \text{Peak runoff}_{\text{sim}}}{\text{Peak runoff}_{\text{obs}}}\right| \times 100, \tag{2a}$$

$$RV\ \text{error}(\%) = \left|\frac{\text{Runoff volume}_{\text{obs}} - \text{Runoff volume}_{\text{sim}}}{\text{Runoff volume}_{\text{obs}}}\right| \times 100, \tag{2b}$$

Subscripts 'obs' and 'sim' imply the observed and simulated hydrologic parameters, respectively. The *PR–RV* Index is now defined as,

$$PR{-}RV\ \text{Index}(\%) = 0.6(PR\ \text{Error}) + 0.4(RV\ \text{Error}). \tag{3}$$

Both likelihood measures (Eqs. (1) and (3)) are consistent with the requirements of the GLUE, as they change monotonically with increasing similarity of behavior in discharge simulation. Note that, the *Efficiency* Index *increases* while the *PR–RV* Index *decreases* monotonically with more accurate simulations. Hence, we considered the reciprocal (inverse) of the *PR–RV* Index as the GLUE-required likelihood measure in the rescaling of likelihood weights. It is appropriate to note, at this stage, that the choice of relative weights assigned to *PR* and *RV* was arbitrary. The purpose of having a *PR–RV* index was to assess the performance of the proposed sampling across two widely different likelihood measures. Hence, this study does not address how the assignment of relative weights to *PR* and *RV* would affect the performance of the sampling scheme. Using the hydrologic parameter calibration algorithm of Duan et al. (1992) we found the highest *Efficiency* index to be 0.975 and the lowest *PR–RV* index as 1.9%. Due to unknown complexities in the parameter-response surface and limitations of current non-linear optimization algorithms (Duan et al., 1992) the two optimized parameter sets (for each index) however did not match.

To implement the GLUE methodology, each parameter of TOPMODEL was specified a range of possible values. Table 1 lists the ranges assigned to all eight TOPMODEL parameters used for GLUE. For a rigorous assessment of the *interpolator*, we considered it important to assume all eight parameters potentially sensitive and having highly non-linear interactions in simulation of discharge.

Table 1
Parameter value ranges used for GLUE sampling

	Minimum value (p)	Maximum value (q)
1 SZM (m)	0.0001	0.2
2 SR0 (m)	0.0001	1.0
3 SRMAX (m)	0.0001	1.0
4 XK0 ($\mathrm{m\,h^{-1}}$)	0.0001	10.0
5 T0 ln($\mathrm{m^2\,h^{-1}}$)	0.0001	15.0
6 TD ($\mathrm{h\,m^{-1}}$)	0.0001	5.0
7 CHV ($\mathrm{m\,h^{-1}}$)	500.0	5000.0
8 RV ($\mathrm{m\,h^{-1}}$)	50.0	2500.0

4. Formulation of the stochastic interpolator

The principle of the *interpolator* is founded on the *Theory of Homogeneous Chaos* (Wiener, 1938). Wiener (1938) had shown that if a deterministic dynamical model (where model output is random) bears a highly non-linear relationship with model inputs (and with a tendency to exhibit chaotic behavior), then it is possible to approximate both inputs and outputs (treated here as random processes) of the uncertain model through series expansion of standard random variables using Hermite polynomials (HP). Although the presence of chaotic behavior in the hydrologic system under study is not addressed herein, recent literature supports the wisdom of choosing the *Theory of Homogeneous Chaos* as a basis for formulation of the *interpolator*. We cite a few examples from literature as follows: (1) Both rainfall and streamflow have been observed to exhibit chaotic behavior over long-time scales (Jayawardena and Lai, 1994; Sivakumar et al., 2001a,b); (2) Sivakumar et al. (2001a) have demonstrated the presence of chaos in the rainfall–runoff transformation process (also see Sivakumar (2004) for a general overview). It is however worthwhile to mention that the sampling interval (hourly) chosen for this study may have unknown effects on the outcome of the proposed sampling method as most studies (cited herein) have investigated chaos in data at much coarser scales (> hourly).

There are three major steps involved in the algorithm formulation of the *interpolator*. We describe these steps below. For more details on the mathematical theory, one is referred to Isukapalli and Georgopolous (1999) and Ghanem and Spanos (1991).

Step 1: Transformation of parameter distributions. Our TOPMODEL model input parameter uncertainty domain is represented by an 8-D hypercube (Table 1) with the distribution of each parameter being uniform (the norm for GLUE). It is defined as follows:

$$X_i \sim U(p_i, q_i), \quad i = 1, \ldots\ldots, 8, \tag{4}$$

where p and q form the lower and upper parameter ranges (columns 2 and 3 of Table 1). Subscript 'i' refers to the specific parameter type (from 1 to 8 as listed in Table 1). 'X' represents the parameter value. These uniformly distributed parameters are then expressed as a series of a standard normal random variable (*srv*) as,

$$x_{i,j} = p_i + (q_i - p_i)\left(\frac{1}{2} + \frac{1}{2} erf(\varepsilon_{i,j}/\sqrt{2})\right),$$
$$i = 1, \ldots\ldots, 8, \tag{5}$$

where ε is a *srv* $\sim N(0,1)$ and 'j' denotes the index for a random realization. $erf(xx)$ is the error function defined by the following integral:

$$erf(xx) = \frac{2}{\sqrt{\pi}} \int_0^{xx} \mathrm{e}^{-w^2}\, \mathrm{d}w. \tag{6}$$

In Eq. (6), xx is the *srv* and ww an intrinsic independent variable of the error function.

We have now expressed the random inputs (uniformly distributed model parameters) via *srvs* as $\{\varepsilon\}_{i=1}^{n}$ (where, $n = 8$). The choice of transforming the model parameters to the normal *srvs* is justified by mathematical tractability of functions of these *srvs* (Devroye, 1986). For example, other common univariate distributions such as gamma, exponential, Weibull, log-normal can also be transformed explicitly to normal *srvs*.

Step 2: Polynomial chaos expansion. Next, we represent our uncertain model output, L—the likelihood measure (left-hand side of Eq. (1) or (3)), as an nth order expansion of a Hermite Polynomial of *srvs*. This step, called "Polynomial Chaos Expansion", follows from Ghanem and Spanos (1991). In this study we have considered 2nd and 3rd order expansions which are defined as follows:

$$L_2 = a_{0,2} + \sum_{i=1}^{n} a_{i,2}\varepsilon_i + \sum_{i=1}^{n} a_{ii,2}(\varepsilon_i^2 - 1) + \sum_{i=1}^{n-1}\sum_{j>1}^{n} a_{ij,2}\varepsilon_i\varepsilon_j, \tag{7}$$

$$\begin{aligned} L_3 = {} & a_{0,3} + \sum_{i=1}^{n} a_{i,3}\varepsilon_i + \sum_{i=1}^{n} a_{ii,3}(\varepsilon_i^2 - 1) \\ & + \sum_{i=1}^{n} a_{iii,3}(\varepsilon_i^3 - 3\varepsilon_i) \\ & + \sum_{i=1}^{n-1}\sum_{j>1}^{n} a_{ij,3}\varepsilon_i\varepsilon_j, + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ijj,3}(\varepsilon_i\varepsilon_j^2 - \varepsilon_i) \\ & + \sum_{i=1}^{n-2}\sum_{j>1}^{n-1}\sum_{k>j}^{n} a_{ijk,3}\varepsilon_i\varepsilon_j\varepsilon_k, \end{aligned} \tag{8}$$

where the subscript after L represents the order of the expansion.

Step 3: Calibration of coefficients of the *interpolator*. From the above equations (7 and 8), it can be seen that the number of unknown coefficients (the 'a's in the right-hand side) to be determined for second and third order polynomial chaos expansions are 45 and 153, respectively. These unknown coefficients are now identified by generating the same number of model data points and solving the system of linear algebraic equations. Isukapalli and Georgopoulos (1999) provide guidelines on choosing model points for robust calibration of coefficients. The choice of the model points in this study is, however, left open. We investigated this issue herein and observed that the model points for calibration is best chosen as scattered uniformly in the entire domain of possible likelihood values. However, we did not find the *interpolator*'s performance to be overly sensitive to the choice of model points. For calibration of polynomial coefficients we used the singular value decomposition (SVD) method (Press et al., 1999) because of its ability to handle ill-conditioned matrices (Press et al., 1999). This is important for higher order expansions or when the likelihood measures and coefficients suffer from scaling problems.

In Fig. 3 we summarize the algorithm for the *interpolator*. First, we generate a set of uniformly distributed model parameter sets from *srvs* (using Eq. (5) and Table 1). 45 and 153 points on the TOPMODEL's parameter-output (L) response surface are then chosen for the 2nd and 3rd order *interpolators*, respectively. The *interpolator* is then calibrated for its coefficient values by solving the system of linear algebraic equations by the SVD method. Once calibrated for TOPMODEL and the watershed using data available, we evaluate the efficiency of the *interpolator* in parameter sampling in the following 4 steps: (i) sample N (0,1) *srvs*; (ii) generate the corresponding family of uniformly distributed TOPMODEL parameters from Eq. (5); (iii) compute the *interpolator*-predicted likelihood value—L values from Eq. (7) or Eq. (8); and (iv) if the *interpolator* predicts a sampled parameter set to be *behavioral,* then test its accuracy by actual execution of TOPMODEL for that parameter set. Note that the use of the *interpolator* within the GLUE framework does not violate the requirement that parameters be sampled from their marginal uniform distributions (discussed further in the following sections). It only helps to make an informed decision on sampling by providing an indication on whether the sampled parameter set is likely to be *behavioral* or *non-behavioral* before making the actual time-consuming TOPMODEL model run.

5. Simulation framework

The *interpolator* is potentially a few (at least 2–3) orders faster in computation than TOPMODEL itself and can therefore serve as a fast-running proxy for making Bayesian decisions on the degree of representa-

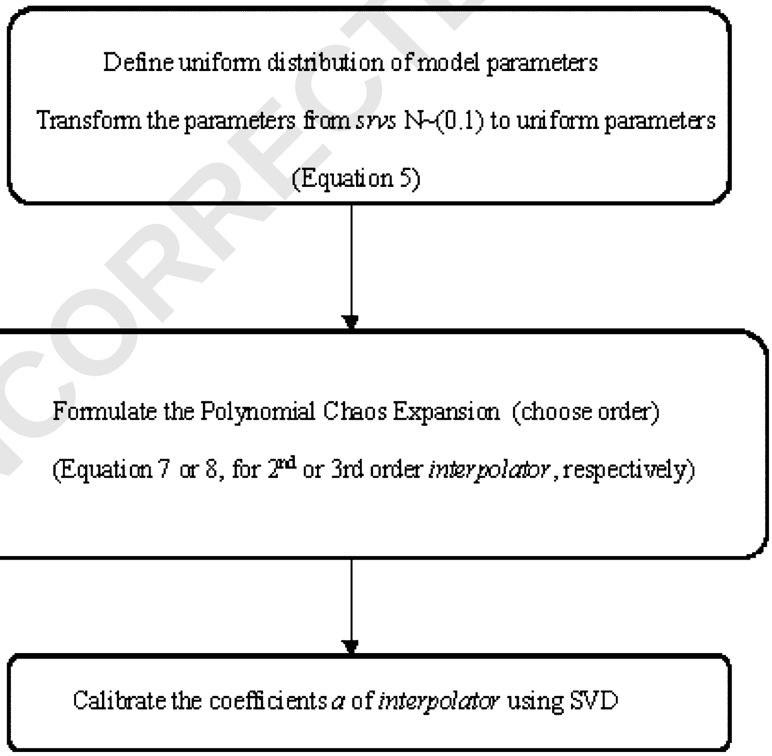

Fig. 3. Flow-chart for algorithm formulation of *interpolator.*

tiveness of sampled parameter sets. In almost all previous GLUE applications reported in literature, *behavioral* and *non-behavioral* parameter sets were identified through the actual time-consuming execution of the hydrologic model. This often resulted in a high wastage of computational time where a large majority of the runs were found to be *non-behavioral* (see Christaens and Feyen, 2002, for example). In the simulation framework we tested the accuracy of the *interpolator* in modeling the parameter-output response surface for GLUE and assessed its potential in reducing the computational time due to the *non-behavioral* runs (that are not detected a priori by uniform sampling in GLUE).

From the specified parameter ranges (Table 1), a total of 200,000 TOPMODEL parameter sets were sampled and the respective hydrographs simulated. All sets had an *Efficiency* index greater than 0.0 or a *PR–RV* Index less than 100%. This large set of parameters now formed the reference database for evaluation of the *interpolator*. This ensemble was further divided into 50 sub-divisions each containing 4000 parameter sets. Each set within the sub-division had its corresponding '*true*' model response in terms of likelihood measures L (*Efficiency* Index and *PR–RV* Index from Eqs. (4) and (5), respectively). These true values were archived from actual execution of TOPMODEL. We then evaluated the accuracy of the *interpolator* within each of these 50 sub-divisions to make generalizations on the mean and variability of its performance of the *interpolator* as a fast-running proxy to the model. We first present a confusion matrix (i.e., a matrix where observed and simulated vectors are presented in a matrix format) for sampled parameter sets below to define the performance measures whose description follows next (*Note*: '*N*' in each quadrant represents the number of samples; *Behavioral* (*Non-behavioral*) refer to sets greater(less) than a threshold performance measure.

Prediction (from *interpolator*)	Truth (from TOPMODEL execution): *Behavioral*	Truth (from TOPMODEL execution): *Non-behavioral*
Behavioral	N_A	N_B
Non-behavioral	N_C	N_D

To define the probability of *interpolator* to successfully predict whether a sampled parameter set is *behavioral* or *non-behavioral* (based on a given threshold for likelihood measure L) we define *success ratio* (SR) as,

$$SR = \frac{N_A}{N_A + N_B}. \tag{9}$$

The SR indicates only a partial assessment of sampling efficiency. There can be instances where the *interpolator* is overly conservative in predicting a set as *behavioral* and thereby achieve a spuriously very high or very low SR over very small samples of model executions. Specific instances where the SR may not be a reliable indicator of efficiency is when the parameter uncertainty domain is significantly under-represented. Thus, another measure, *Bias Score* (BS, Eq. (10)) was also defined. BS quantifies the propensity of the *interpolator to* predict unsuccessfully the *behavioral* sets as *non-behavioral* or missing regions of potential high likelihood values of the response surface.

$$BS = \frac{N_A + N_B}{N_A + N_C}. \tag{10}$$

A BS value of less than 1 would indicate that the *interpolator* has a tendency to be conservative in predicting correctly a sampled parameter set's likelihood value. A BS value greater than 1 would indicate the *interpolator's* propensity to predict samples as *behavioral*. An ideal *interpolator* should therefore have a BS of near 1.0 and *Success Ratio* that is higher than that for uniform sampling.

Performance of the *interpolator* was compared with the fully uniform sampling of parameter sets using the above two measures (Eqs. (9) and (10)). The Nearest-Neighborhood (*NN*) search for interpolating parameter sets's likelihood value was also compared herein (hereafter called *NN* method). This type of sampling method was first proposed by Beven and Binley (1992) to address the computational concerns of the GLUE method. In the *NN* method, a sampled point in parameter hyperspace is searched for the '*n*' nearest neighboring points in a model's response surface that is constructed from a finite number of sample points (= 1000 points in this study). The probable likelihood value is then interpolated by the inverse squared distance technique. We have considered 6 and 12 neighbors for the *NN* method. A point to note is that the *NN* method requires a computationally intensive sorting algorithm to rank all the distances from a sampled point. The computing time for sorting increases as N^2 where N is the size of the pre-constructed model points (Press et al., 1999). Hence a compromise is needed with the size of the pre-constructed model points when the dimension of the parameter hyperspace is high.

6. Results and discussion

In Fig. 4 we show a comparison of *SRs* for the various sampling schemes—*interpolator*, *NN* method and the uniform sampling. The *SR* shown is the mean of the 50 subdivisions represented with one standard deviation of variability in performance. The inverse of the standard deviation is a measure of how consistent the sampling scheme is in predicting correctly. In Tables 2a and b, we also present the mean values (of 50 sub divisions) for *BS* and the total confusion matrix values—N_A, N_B, N_C and N_D, as a function of *behavioral* threshold for *Efficiency* Index and *PR–RV* Index, respectively. These values are presented for the *interpolator* and *NN* method only. Joint assessment of *SR* with *BS* statistics leads us to the following observations on the relative merits and limitations of the *interpolator* sampling scheme with respect to the *NN* method:

(1) The *interpolator* sampling scheme appears to sample more efficiently for *Efficiency* Index likelihood measure than the *PR–RV* Index likelihood measure (Fig. 4). This may hint at the importance of careful formulation of the likelihood measure for GLUE sampling and potentially indicate a structural weakness in the *PR–RV* Index to serve as a reliable likelihood measure. However, the *interpolator* generally samples more efficiently than the uniform sampling scheme (note: some rare exceptions using the 2nd order *interpolator*).

(2) For *Efficiency* Index, the 2nd order *interpolator* is found to be more accurate in sampling than the 3rd order *interpolator* (upper panels of Fig. 4 and Table 2a). For *PR–RV* Index, it appears that the 3rd Order *interpolator* is more accurate in sampling than the 2nd order *interpolator* (lower panels of Fig. 4 and Table 2b). At this stage, it is difficult to identify possible reasons behind such an observation and detailed investigation is necessary. Recent work by Field and Grigoriu (2004) indicated that the order of the Hermite Polynomial approximation bears a complex relationship to the nature of the system being modeled. The *Efficiency* Index based *interpolator* potentially reduces the total computing time by uniform sampling for *behavioral* parameter sampling by about 15—25% for the 8-dimensional parameter hyperspace.

(3) Although the *NN* sampling method has the highest *Success Ratio* (*SR*) of the three sampling methods, it also has the highest variability (Fig. 4). This variability (standard deviation), which is about 10–15 times higher

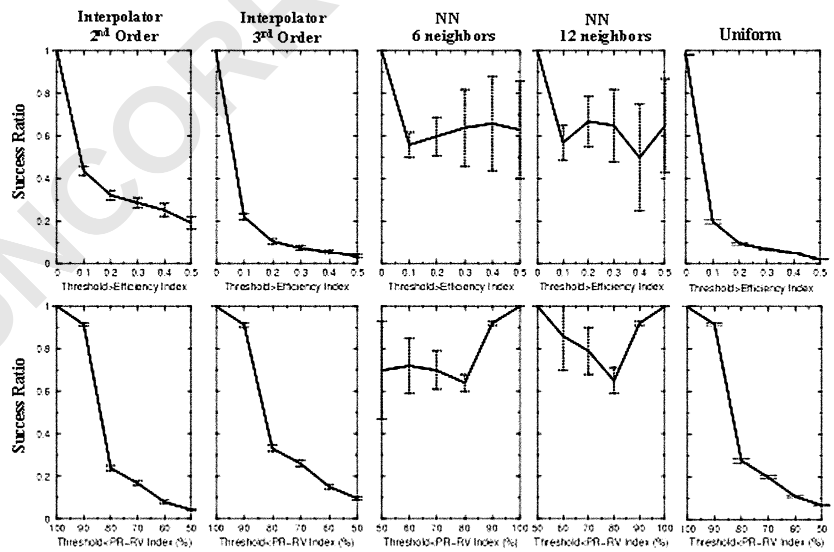

Fig. 4. *Success ratios* (*SRs*) of sampling methods. Upper panels—*Efficiency* Index; Lower panels—*PR–RV* Index.

Table 2
(a) Mean Bias Scores (*BS*) and total confusion matrix numbers for *Efficiency* Index likelihood measure

Behavioral threshold >efficiency index	Bias score *BS*	N_A	N_B	N_C	N_D
Interpolator-2nd Order					
0.00	0.367	66147	0	113943	0
0.10	1.274	222624	28624	17724	111456
0.20	2.150	12616	26313	5548	135613
0.30	2.297	8582	21237	4448	145823
0.40	2.304	5742	16890	4131	153327
0.50	2.558	3346	13907	3458	159379
Interpolator-3rd Order					
0.00	0.506	91050	0	89040	0
0.10	2.065	18428	65155	21636	74879
0.20	4.107	7969	67389	10173	94554
0.30	5.089	4955	62128	8055	104956
0.40	5.911	3206	55766	6362	114469
0.50	7.461	1830	49482	5264	123814
Nearest neighborhood (*NN*)-6 neighbors					
0.00	1.00	180090	0	0	0
0.10	0.76	16227	12151	20326	131386
0.20	0.23	2670	1290	14701	161429
0.30	0.07	251	133	12425	167281
0.40	0.03	10	10	9578	170492
0.50	0.01	3	4	9584	170499
Nearest neighborhood (*NN*)-12 neighbors					
0.00	1.00	180090	0	0	0
0.10	0.80	16087	13325	20466	130212
0.20	0.45	4438	3233	12933	159486
0.30	0.16	1007	742	11669	166672
0.40	0.09	172	160	9416	170342
0.50	0.07	21	21	6554	1703485
(b) Mean Bias Scores (*BS*) and total confusion matrix numbers for *PR–RV* Index likelihood measure					
Interpolator-2nd Order					
100	0.75	135382	0	44708	0
90	0.80	121884	11503	44165	2538
80	2.62	32284	99048	17969	30789
70	2.56	22110	106923	14246	36811
60	6.34	10957	115726	9108	44299
50	10.64	6076	118203	5673	50138
Interpolator-3rd Order					
100	0.54	96756	0	83334	0
90	0.58	87494	8210	78493	5893
80	1.89	31474	63187	18770	66659
70	2.58	24569	69066	11788	74667
60	4.64	14249	78365	5813	81663
50	7.83	8944	82545	2808	85793
Nearest neighborhood (*NN*)-6 neighbors					
100	1.00	180090	0	0	0
90	1.08	166105	13983	1	1
80	0.98	31503	17541	18712	112334
70	0.19	4979	1730	31366	142015

Table 2 (*continued*)

Behavioral threshold >efficiency index	Bias score *BS*	N_A	N_B	N_C	N_D
60	0.05	416	109	19640	159925
50	0.01	10	6	11744	168330
Nearest neighborhood (*NN*)-12 neighbors					
100	1.00	180090	0	0	0
90	1.08	166035	13913	71	71
80	0.99	31461	18174	18754	11701
70	0.34	8387	3981	27958	139764
60	0.13	1551	775	18505	159259
50	0.06	223	149	11531	168187

than the interpolator's *SR*, increases as a function of *behavioral* threshold.

(4) *NN* sampling method has very low *Bias Scores*, which decreases as the *behavioral* threshold criterion increases (Tables 2a and b). This indicates the *NN* method has a higher tendency to miss regions of high likelihood values in the sampling than the *interpolator*. The *NN* sampling scheme formulated herein is found to be an ineffective global sampling tool. Another major drawback is that the sorting algorithm in the *NN* scheme increases the computational burden of sampling. For example, after a total of 200,000 executions by the *NN* sampling method, only 21 behavioral sets exceeding *Efficiency Index* > 0.5 (Table 2a) were yielded. For the 2nd order *interpolator* the total number of *behavioral* sets yielded was much larger (3346 sets, Table 2a) and took insignificant computing time.

(5) Efficiency of the *NN* sampling method does not appear very sensitive to the number of neighbors used in the parameter search (Fig. 4 and Tables 2a, b). This is expected as *NN* method samples on the principle of inverse-squared distance interpolation which fails to recognize the greater non-linearity in the parameter-output response surface.

The assessment of the *interpolator* using *SR* and *BS* is not a complete test of its eligibility to accelerate the uniform parameter sampling for GLUE parameter. The question as—*does the interpolator alter the structural properties of the GLUE uncertainty analyses?*—requires investigation. For this, we have chosen to examine the dotty plots of parameters sampled by the *interpolator* and compare them to the reference dotty plots by uniform sampling. Dotty plots were first proposed by Beven and Binley (1992) as a simple way to demonstrate the parameter equifinality (non-uniqueness) of a model. Against the likelihood value presented along the *y*-axis, the scatter of the parameters along the *x*-axis is accepted as a qualitative measure of parameter equifinality. If the dotty plots derived from uniform sampling are assumed as the reference, then the parameters sampled as *behavioral* by the pre-screening of the *interpolator* should show similar scatter to represent consistent equifinality. This is an important aspect to assess for any parameter sampling scheme, which otherwise may render itself unsuitable for GLUE analysis. Note that a parameter set was always deemed *behavioral* only after an actual TOPMODEL run. The sole purpose of the *interpolator* is to filter out the potentially *non-behavioral* sets that could otherwise increase computational time of model execution. We show herein dotty plots pertaining to 5000 sampled parameter sets determined as *behavioral* with the *Efficiency* Index likelihood measure >0.3 (Figs. 5a–c) and *PR–RV* Index $<100\%$ (Figs. 6a–c). By comparing among the figures ('a' with 'b' and 'c'), we observe that the *behavioral* parameters sampled via the *interpolator* represent, at least qualitatively, the same degree of equifinality (non-uniqueness) as the reference uniformly sampled dotty plots (Figs. 5a and 6a). The *interpolator* imposes no specific regions of local attraction that causes a sampling pattern incompatible with that by purely uniform (non-*interpolator*) random sampling.

A more definitive test for preservation of equifinality however, would be to consider all 28 (i.e., ${}^{8}C_{2}$) combinations of parameter covariations in lieu of the one-to-one parameter dotty comparisons. Since this is a large number of comparisons, we adopted an alternative, yet a definitive way nevertheless in our opinion, of answering if the *interpolator* altered the uncertainty structure of the model or not. In Fig. 7, we show a GLUE analysis with 90% quantiles (confidence limits) in discharge simulation uncertainty obtained from the aforementioned 5000 *behavioral* parameter sets (Figs. 5a–c, 6a–c). The prediction quantiles produced by uniform random sampling (leftmost panels, Fig. 7) are assumed as the reference for comparison here. For both likelihood measures (*Efficiency* Index—upper panel, Fig. 7; *PR–RV* Index—lower panel, Fig. 7) we observe negligible difference in the uncertainty estimation at the 90% confidence limits. A subsequently more rigorous test for the preservation of the uncertainty structure in simulation is then provided in Fig. 8. Here we compare the *exceedance probability* (*EP*) against the width of confidence limits from 10% quantile width (45% upper and 55% lower) to 90% quantile width (5% upper and 95% lower). *EP* is defined as the number of times the observed discharge is not enveloped by the confidence limits normalized by the total number of time-step in simulation. *EP* would typically decrease monotonically with decreasing quantile width. A very close similarity of the monotonic decrease in EP with increasing quantile width is observed between the *interpolator*-GLUE (middle and rightmost panels—Fig. 8) and *uniform-GLUE* (leftmost panels—Fig. 8).

7. Conclusion

A stochastic and non-linear interpolation based parameter sampling scheme for uncertainty analyses of hydrologic models was presented. The scheme was based on the principles of the 'Theory of Homogeneous Chaos'. The sampling scheme was evaluated within the generalised likelihood uncertainty estimation (GLUE; Beven and Binley, 1992) methodology for uncertainty analysis. Uncertainty in discharge prediction (model output) was modeled through a Hermite polynomial chaos approximation of normal random variables that represented the model's parameter (model input) uncertainty. The unknown coefficients of the polynomial were then calculated using limited number of model simulation runs. The calibrated Hermite polynomial (*interpolator*) was then used as a fast-running proxy to the slower-running hydrologic model to predict the degree of representativeness of a randomly sampled model parameter set. An evaluation of the scheme's improvement in sampling was then made through comparison with the fully uniform sampling (the norm for GLUE) and the nearest-neighborhood sampling technique using TOPMODEL over a medium-sized watershed in Italy. A notable reduction of computational burden in the ranges of 15–25% was observed even for a high dimensional parameter uncertainty. The GLUE based on the proposed stochastic interpolation sampling scheme preserved the essential features of the uncertainty structure in discharge simulation. The stochastic *interpolator* demonstrates potential to make GLUE uncertainty estimation more efficient for models where large number of parameters (>4) are involved, although further investigation is necessary to explore this issue in detail. An additional advantage is that the *interpolator* does not impose any additional structural or distributional assumptions upon GLUE.

It is appropriate to note at this stage the limitations of the Hermite polynomial approximation—which is the basis for formulation of our proposed *interpolator* scheme. Errors are inherent when the Hermite Polynomial Chaos is approximated as a 2nd, 3rd or higher order approximation (depending on the order of approximation). These errors may or may not be significant, depending on the application. In this study, we have observed a complex relationship among the efficiency of sampling, the order of approximation and the formulation of the likelihood function. In any case, it is wise to understand further and quantify the consequences of the approximations before using the scheme for other applications involving GLUE method (see Field and Grigoriu, 2004 for a detailed assessment on the limitations of the Hermite polynomial approximations).

Some of the natural extensions of this stochastic interpolation based sampling scheme include: (i) appli-

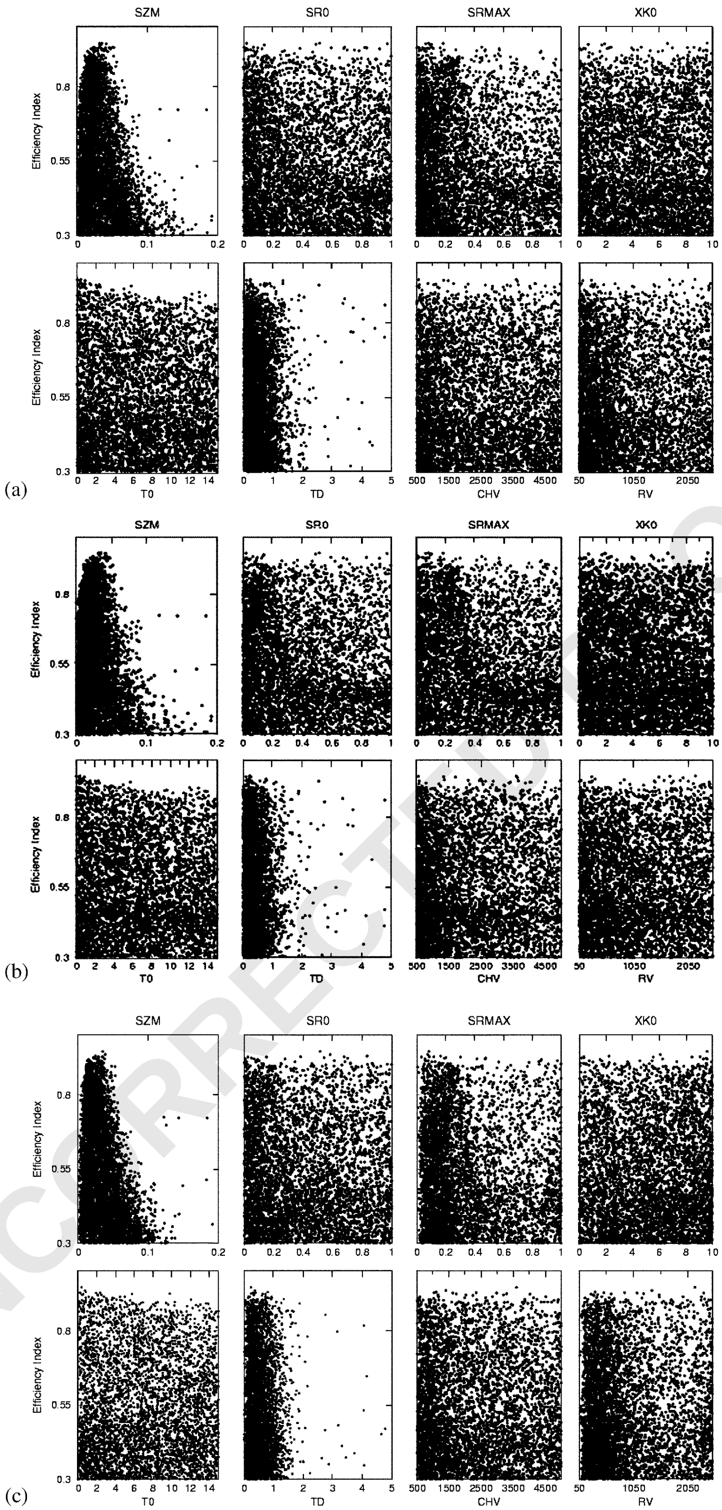

Fig. 5. (a) Dotty plots from uniform sampling with *Efficiency* Index as the likelihood measure. (b) Dotty plots from 2nd order *interpolator* with *Efficiency* Index as likelihood measure. (c) Dotty plots from 3rd order *interpolator* with *Efficiency* Index as likelihood measure.

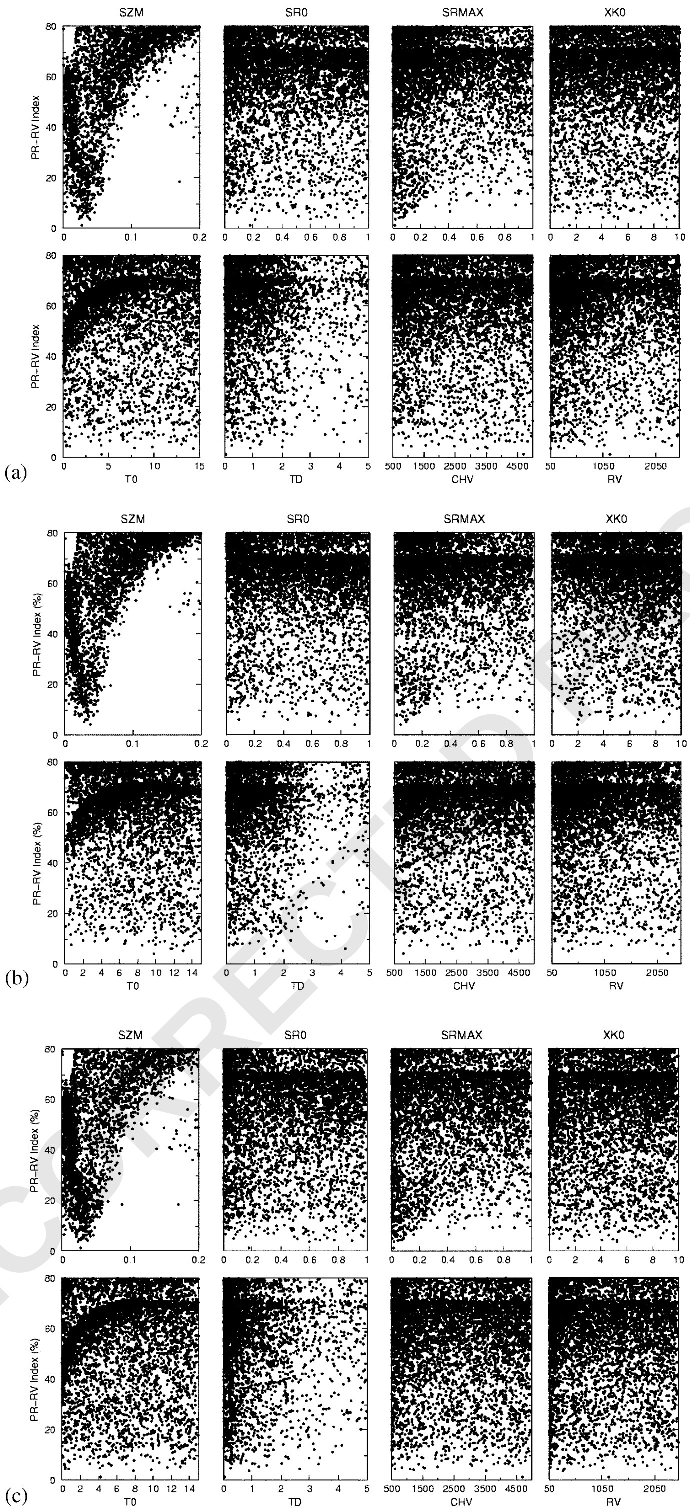

Fig. 6. (a) Dotty plots from uniform sampling with *PR–RV* Index as likelihood measure. (b) Dotty plots from 2nd order *interpolator* with *PR–RV* Index as likelihood measure. (c) Dotty plots from 3rd order *interpolator* with *PR–RV* Index as likelihood measure.

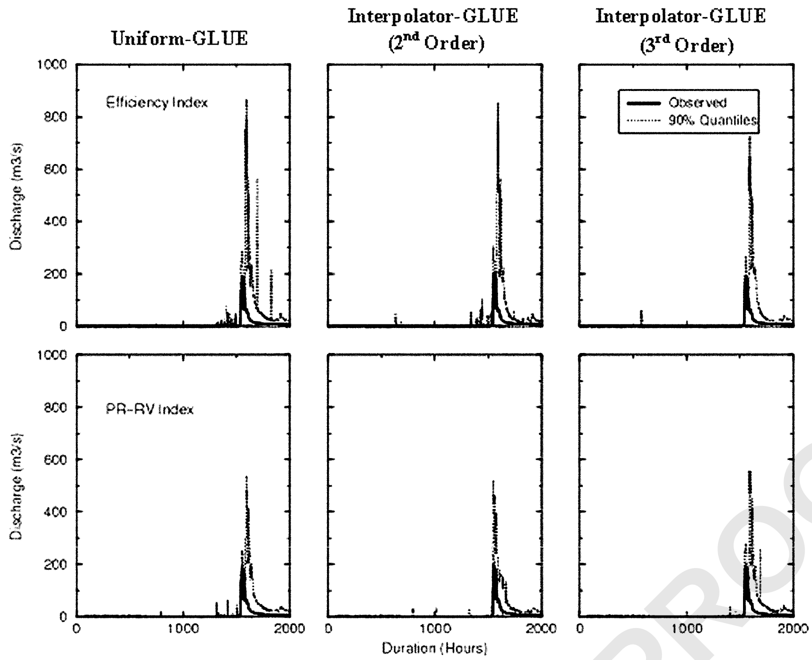

Fig. 7. The GLUE uncertainty estimation of discharge simulation at 90% quantile widths (confidence limits) for uniform random sampling (leftmost panels) and *interpolator* (middle and right most panels). Upper panels represent GLUE for *behavioral Efficiency* Index (>0.3) while lower panels are from *behavioral PR–RV* index (<80%). Uncertainty estimation for each scenario was conducted from corresponding set of 5000 sampled sets shown as dottyplots in Figs. 5a, b, c, 6a, b and c.

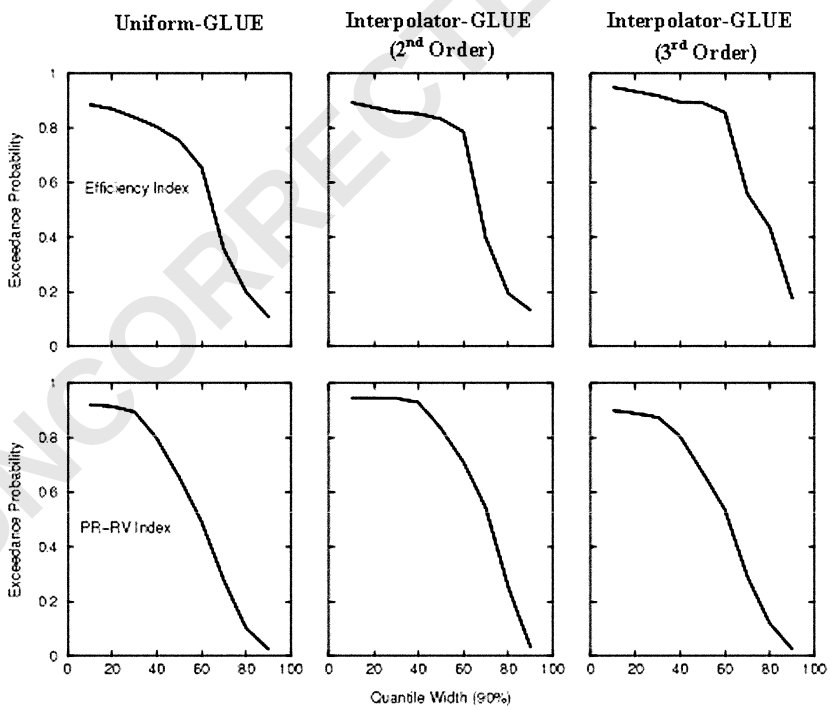

Fig. 8. Exceedance probability (*EP*) as a function of quantile width. Leftmost panels—*uniform*-GLUE; middle panels—*interpolator*-GLUE (2nd order); rightmost panel—*interpolator* -GLUE (3rd order); upper panels—*Efficiency* Index (>0.3); lower panels—*PR–RV* Index as likelihood measure (<80%).

cation of the *interpolator* to other physically-complex models and hydrologic variables within the GLUE framework; (ii) investigating the conditions or assumptions that give rise to a chaotic and non-chaotic behavior in the hydrologic system and thereby attempt to connect the relationship of the hydrologic variable to the polynomial chaos expansions; and (iii) investigating the effect of the dimensional size of the parameter hyperspace on the sampling efficiency of the *interpolator*. It has also been suggested that when the gradient information of the parameters with respect to model output is assimilated in the polynomial chaos expansion, an increase in the prediction accuracy of the *interpolator* can be expected (Isukapalli and Georgopoulos, 1999). Another potential use of the stochastic interpolation sampling scheme would be in applications to large-scale land surface simulations where model parameters are distributed as a matrix (2-D spatial domain) over synoptic scales (in this study the parameters were a vector). For such applications, further study is needed to explore ways to mathematically reformulate the *interpolator* to handle such distributed parameters in spatial format. Work is on-going on some of the above aspects and we hope to report them in future.

Acknowledgements

The first author wishes to dedicate this work to ‘Mano’—his best friend at all times at the University of Connecticut. Dr. Marco Borga of the University of Padua, Legarno, Italy, provided the watershed and storm data for the study. This research effort initiated two years ago as part of the authors’ appreciation of the GLUE method as a relatively easier ‘engineering’ tool for deriving answers to scientific questions on the error propagation of precipitation remote sensing. The constructive comments received from two anonymous reviewers are also gratefully acknowledged.

References

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O’Connell, P.E., Rasmussen, J., 1986. An introduction to the European Hydrological System—Systeme Hydrologique European, ‘SHE’, 2. Structure of a physically based, distributed modeling system. Journal of Hydrology 87, 61–77.

Bates, B.C., Campbell, E.P., 2001. A Markov Chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall–runoff modeling. Water Resources Research 37 (4), 937–947.

Beck, M.B., 1987. Water quality modeling: a review of the analysis of uncertainty. Water Resources Research 23 (8), 1393–1442.

Beven, K.J., Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. Hydrological Processes 6, 279–298.

Beven, K.J., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology. Journal of Hydrology 249, 11–29.

Beven, K.J., Kirkby, M.J., 1979. A physically-based variable contributing area model of basin hydrology. Hydrological Sciences Journal 24 (1), 43–69.

Beven, K.J., Lamb, R., Quinn, P., Romanowicz, R., Freer, J., 1995. TOPMODEL. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water Resource Publication, Fort Collins, Colorado, pp. 627–668.

Bras, R.L., Rodriguez-Iturbe, I., 1993. Random Functions and Hydrology, Dover Publications, New York, 559pp.

Christaens, K., Feyen, J., 2002. Constraining soil hydraulic parameter and output uncertainty of the distributed hydrological MIKE SHE model using the GLUE framework. Hydrological Processes 16 (2), 373.

Devroye, L., 1986. Non-uniform random variate generation. Springer, New York.

Duan, Q., Sorooshian, S., Gupta, V.K., 1992. Effective and efficient global optimization for conceptual rainfall–runoff models. Water Resources Research 28, 1015–1031.

Duan, Q., Gupta, H.V., Sorooshian, S., Rousseau, A.N., Turcotte, R., 2003. Calibration of watershed models, Water Science and Application 6. AGU Publications, Washington, DC.

Faybishenko, B., 2004. Non-linear dynamics in flow through fractured porous media: status and perspectives. Reviews of Geophysics 42, RG2003.

Field Jr., R.V., Grigoriu, M., 2004. On the accuracy of the polynomial chaos approximation. Probabilistic Engineering Mechanics 19, 65–80.

Franks, S.W., Beven, K.J., 1997. Bayesian estimation of uncertainty in land surface-atmosphere flux predictions. Journal of Geophysical Research 102 (D20), 23,991–23,999.

Franks, S.W., Gineste, P., Beven, K.J., Merot, P., 1998. On constraining the predictions of a distributed model: the incorporation of fuzzy estimates of saturated areas into calibration process. Water Resources Research 34 (4), 787–797.

Freer, J., Beven, K.J., Ambroise, B., 1996. Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach. Water Resources Research 32 (7), 2161–2173.

Ghanem, R., Spanos, P.D., 1991. Stochastic Finite Elements: A Spectral Approach. Springer, New York.

Hossain, F., Anagnostou, E.N., Borga, M., Dinku, T., 2004. Hydrological Model Sensitivity to Parameter and Radar Rainfall Estimation Uncertainty. Hydrological Processes 18.

Iman, R.L., Helton, J.C., Campbell, J.C., 1981. An approach to sensitivity analysis of computer models: Part I—Introduction, input variable selection and preliminary variable assessment. Journal of Quality Technology 13 (3), 174–183.

Isukapalli, S.S., Georgopolous, P.G., 1999. Computational methods for efficient sensitivity and uncertainty analysis of models for environmental and biological systems. Technical Report CCL/EDMAS-03, Rutgers University.

Jayawardena, A.W., Lai, F., 1994. Analysis and prediction of chaos in rainfall and stream flow time series. Journal of Hydrology 153, 23–52.

Kremer, J.N., 1983. Ecological implications of parameter uncertainty in stochastic simulation. Ecological Modelling 18, 187–207.

Krzysztofowicz, R., 2000. Hydrologic uncertainty processor for probabilistic river stage forecasting. Water Resources Research 36 (11), 3265–3277.

Kuczera, G., Parent, E., 1998. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. Journal of Hydrology 211, 69–85.

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

Misirli, F., Gupta, H.V., Sorooshian, S., Thiemann, M., 2003. Bayesian recursive estimation of parameter and output uncertainty for watershed models. In: Duan, Q.J., Gupta, H.V., Sorooshian, S., Rousseau, A.N., Turcotte, R. (Eds.), Calibration of Watershed Models. Water Science Application No. 6, AGU, Washington DC, pp. 125–1132.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models, 1, A discussion of principles. Journal of Hydrology 10, 282–290.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1999. Numerical Recipes in Fortran 77, second ed. Cambridge University Press, Cambridge, UK.

Romanowicz, R., Beven, K.J., 1998. Dynamic real-time prediction of flood inundation probabilities. Hydrological Sciences 43 (2), 181–196.

Schulz, K., Beven, K.J., 2003. Data-supported robust parameterizations in land surface–atmosphere flux predictions: towards a top-down approach. Hydrological Processes 17, 2259–2277.

Schulz, K., Jarvis, A., Beven, K.J., 2001. The predictive uncertainty of land surface fluxes in response to increasing ambient carbon dioxide. Journal of Climate 14, 2551–2562.

Sivakumar, B., 2000. Chaos theory in hydrology: important issues and interpretations. Journal of Hydrology 227, 1–20.

Sivakumar, B., 2004. Chaos theory in geophysics: past, present and future. Chaos, Solitons Fractals 19 (2), 441–462.

Sivakumar, B., Berndtsson, R., Olsson, J., Jinno, K., 2001a. Evidence of chaos in the rainfall–runoff process. Hydrological Sciences Journal 46 (1), 131–146.

Sivakumar, B., Sorooshian, S., Gupta, V.J., Gao, X., 2001b. A chaotic approach to rainfall disaggregation. Water Resources Research 37 (1), 61–72.

Spear, R.C., Hornberger, G.M., 1980. Eutrophication in Peel Inlet, II, Identification of critical uncertainties via generalised sensitivity analysis. Water Resources Research 4, 43–49.

Thiemann, M., Trosset, M., Gupta, H., Sorooshian, S., 2001. Bayesian recursive parameter estimation for hydrologic models. Water Resources Research 37 (10), 2521–2535.

Tyagi, A., Haan, C.T., 2001. Uncertainty analysis using first-order approximation method. Water Resources Research 37 (6), 1847–1858.

Young, P.C., Beven, K.J., 1994. Database mechanistic modeling and rainfall-flow non-linearity. Environmentrics 5 (3), 335–363.

Wiener, N., 1938. The homogeneous chaos. American Journal of Mathematics 60, 897–936.