Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, No. 5, April : 2023

ANALYSIS OF PARAMETER UNCERTAINTY IN CONCEPTUAL CATCHMENT MODELS USING MONTE CARLO TECHNIQUES

G. Vincent Reddy, Research Scholar (student), Department of Mathematises, JS university- Shikohabad
Dr.Prabal Pratap sing, Professor and Head, Depart of Mathematises, JS university- Shikohabad

ABSTRACT
Two methods for assessing the uncertainty of parameters in complicated hydrologic models are discussed, both of which make use of the Monte Carlo method. To begin, the GLUE framework developed by Beven and Binley employs a technique called significance sampling, which is meant to mitigate bias when approximating unknown values. The Metropolis approach uses a random walk rather than significance sampling to account for parameter uncertainty, which has a non-normal probability distribution. Three examples are provided to illustrate the use of these Monte Carlo techniques. In the first, we take into account a straightforward water balance model for which we already know the solutions. The Metropolis sampling approach has been shown to be more effective than the importance sampling strategy. If insufficiently random samples are collected, results from significance sampling might be quite misleading. In the second and third examples, we use more complex catchment models to show what kind of insights may be achieved by using the Metropolis method. In particular, they illustrate how to assess the relevance of split-sample tests, make use of prior information, and assess confidence intervals for hydrologic responses that were not included in the calibration process. The Metropolis method is superior to the more commonplace first-order approximation-based inference when it comes to dealing with parameter uncertainty in hydrologic models.

INTRODUCTION
An accurate evaluation of parameter uncertainty and its impact on model predictions generated using data outside of the calibration data is the primary emphasis of this study, which aims to provide light on how to calibrate conceptual catchment models. A conceptual catchment model may be seen as an intermediate step between a physical reductionist model and a black box model.
Physically-based reductionist models attempt to generalise the physics of hydrologic processes from the lab scale to the watershed scale (Grayson et al., 1992). Black box models, such as neural networks (Chen et al., 1990) or ARMA models, do not take into consideration the physics of hydrologic processes (Box and Jenkins, 1976). Conceptual models aim to avoid the scale concerns that plague reductionist models by zeroing down on the processes that hydrologists believe most essential and by adopting control volumes over which state variables and fluxes are temporally and spatially averaged (Nash and Sutcliffe, 1970). Despite the fact that mass conservation requires the specification of all flows into and out of a control volume, conceptual, rather than physics-based, fluxequations are typically used. Therefore, conceptual models are less complicated to construct and need less information than reductionist models. Many of these states and fluxes, however, are too abstract to be identified by traditional measuring techniques. The conceptual model is characterised by the need to calibrate one or more model parameters against the catchment responses that may be seen in the real world.
The conceptual catchment model may be formalised by using an abstract statistical framework. The measured responses of the catchment at time t, t = 1,...,n, will be represented by an m-dimensional vector, qt. The onus is on the modeller to provide an explanation that makes room for. Just give me a moment to The catchment transfer function f(.) transforms inputs xt (such as precipitation, evapotranspiration, and pollutant input) into fluxes (such as streamflow, recharge, and pollutant loads), and also state variables (such as soil moisture and groundwater levels). A vector of parameters may

characterise mistakes in input (x), output (q), and model (uncertainty) states. The qt vector may be seen as a sample at random from a set of potential outcomes.

$$\mathbf{q}_t \leftarrow \phi[f(\mathbf{x}_t, \beta), \gamma], \quad t = 1, \ldots, n \qquad (1)$$

We may use f(xt,b) and g to build the probability distribution f[]. For a given value of the error parameter vector g, we may assume that the residuals [q f(xt,b)] behave like white noise.

Non-homogeneity of variance and serial dependence are examples of residual distortions that often arise during the model calibration process. Signs of a "poor" error structure include an overconfidence in the predicted results. Kuczera (1983) explains how to get around these restrictions using Box Cox transformations (Box and Tiao, 1973) in tandem with ARMA modelling of residuals (Box and Jenkins, 1976). The model's probability function takes on a multiplicative shape if the conditions of independence and homogeneity are accepted.

$$\prod_{t=1}^{n} L(\mathbf{q}_t, f(\mathbf{x}_t, \beta), \gamma) \qquad (2)$$

In the Bayesian statistical framework, the uncertainty of the parameters f(xt,) is measured using probability distributions. DeFinetti (1937) and Lindley (1996) establish the necessary axiomatic norms of logic for such a framework of statistical reasoning. A better understanding of (g,b) may be obtained, starting with some prior distribution [q- f(xt,)], by applying the Bayes rule to data [D = qt,xt,t = 1,...n] (Bayes, 1763).

$$\pi(\gamma, \beta|\mathbf{D}) = \frac{\prod_{t=1}^{n} L[\mathbf{q}_t, f(\mathbf{x}_t, \beta), \gamma] b(\gamma, \beta)}{\int \prod_{t=1}^{n} L[\mathbf{q}_t, f(\mathbf{x}_t, \beta), \gamma] b(\gamma, \beta)\, d\beta\, d\gamma} \qquad (3)$$

Choosing the probability density b(,) that should capture all the'subjective' information about (,) prior to collecting sample D has always been difficult for practitioners.

Berger provides a total of eight different approaches to generating such distributions (1985). If we don't want to bias the model toward a beginning point, we may use a uniform prior [b(,) = 1] throughout the range of parameters. The integral may not equal 1 if the possible numbers are endlessly big. If the posterior distribution is a genuine probability distribution, then Bayesian inference may be conducted even with a defective uniform prior.

From a hydrological standpoint, the purpose of model calibration is to determine the posterior probability distribution p(l D), which describes the current degree of knowledge about the structural (or model) parameters given the data D and previous information. We get this value by doing a straight integration over the nuisance parameter. γ.

$$p(\beta|\mathbf{D}) = \int_{\text{all } \gamma} \pi(\gamma, \beta|\mathbf{D}) d\gamma \qquad (4)$$

Over the last several years, many scientists have laboured to estimate what the most probable value of is. In a nutshell, Duan et al. (1992) synthesise the research and provide a robust and efficient probabilistic search approach for identifying the most probable value. Surprisingly little effort has been put into developing methods for proper assessment of parameter uncertainty in complex hydrologic models. Conceptual models are nothing more than the result of an empirical combination of mathematical operators describing the key aspects of an idealised hydro-logic cycle, therefore no respectable hydrologist would presume that there is a single, invariable value for model parameters. it doesn't matter how skilled or imaginative the modeller is. Without a realistic assessment of parameter uncertainty, it has been impossible to perform tasks such as evaluating prediction/confidence limits on future hydrologic responses, assessing the significance of deviations in split-sample tests, and assessing the value of regional relationships between model parameters and catchment characteristics. See Kuczera for a discussion of first-order approximations in a hydrologic context (1988). The transfer function f(xt,) is approximated at the first order in standard statistical theory, providing an approximate multinormal description of parameter uncertainty. The first-order approximation yields appropriate

results if the linearization of f(xt,) holds true everywhere in the domain of b for which there is sufficient uncertainty. However, hydrologic models seldom work with such a close approximation. These factors, among others, have led researchers to abandon traditional methods of statistical inference in favour of more general MonteCarlo-based methods (for examples, see Hornberger and Spear (1980), van Straten and Keesman (1991), and Beven and Binley (2005)). (1992). One typical use of the Monte Carlo method is the estimation of prediction/confidence intervals. Despite its broad applicability, this approach has certain obvious restrictions in practise.

Sampling the posterior distribution using Monte Carlo

There are two common Monte Carlo techniques for drawing samples from the posterior distribution: significance sampling and Markov chain sampling.

Markov chain sampling

Markov chain methods may produce samples from the posterior distribution p(l D) using a random walk that learns the true distribution. The approach may offer a more accurate representation of the distribution than significance sampling. Markov chain sampling is preferred when there is a substantial gap between the expected and observed significance distribution. Gilks et al. provides various examples of the usage of Markov chain sampling, while Gelman et al. (1997) and Brooks (1998) give thorough discussions of the methodologies for Markov chain sampling (1996) . The Metropolis technique is often employed as a Markov chain sampler, despite the fact that it may not be the most efficient. Gibbs sampling offers the potential for faster convergence, but it requires conditional distribution sampling, which might be difficult for more complex models. Whereas, the Metropolis algorithm takes a more all-encompassing approach.

Importance sampling

An extensive section of Tanner (1992) is dedicated to the description of significance sampling, a frequent technique for choosing samples at random from a probability distribution. The strategy focuses on selecting a significance probability distribution I(b) that is near to the posterior probability p(b|D) and can be effectively sampled. From the distribution p (b l D), N weighted random samples are drawn according to the pattern b I, P I I = 1,..., N. Tanner (1992) and Gelman et al. (1997) both note that picking the right significance distribution is crucial to the effectiveness of this technique. It is possible for the algorithm to get the wrong result if one or more of the significance weights is set too high.

There are only two possible approaches to prioritisation when dealing with continuous multivariate situations. The evenly sampled hypercube is the first example. Beven and Binley's GLUE method and van Straten and Keesman's Monte Carlo set membership technique are two such examples (1991). If significant parameter interaction produces narrow-curving ridges on the posterior surface p, then it may be required to sample the hypercube extensively to prevent dominating weights (b l D). In the case of parameter spaces with a lot of dimensions, this might need a lot of computational time and power. To produce a representative sample of a parameter in a hypercube with a resolution of one tenth of the parameter range, for instance, would need an insane 1010 Monte Carlo samples, assuming that each sample covers the whole parameter space. Undersampling crucial portions of the parameter space might lead to a handful of dominant significance weights if the sampling density isn't maintained high enough.

CONCLUSION

In the scientific literature, parameter uncertainty assessment in sophisticated hydrologic models is seldom discussed. The nonlinearity of hydro-logic models makes it difficult to evaluate prediction/confidence limits on future hydrologic responses, evaluate the significance of deviations in split-sample tests, or evaluate the value of regional relationships between model parameters and

catchment characteristics using traditional statistical theory based on first-order approximations and multinormal distributions. To evaluate the parameter uncertainty in complex hydrologic models, this research compares two Monte Carlo-based methodologies. The first, significance sampling, has seen extensive use in hydrology, the most recent example being the GLUE method introduced by Beven and Binley (1992). The second method is the use of Markov chain Monte Carlo sampling, which has been extensively discussed in the Bayesian statistics literature. This method, in contrast to significance sampling, employs a random walk that adjusts to the underlying probability distribution. The Metropolis algorithm was the first of its type to use Markovchain sampling. In spite of the fact that it is not the most effective option, it was selected for this investigation due to its scalability. In this article, we do not investigate the problem of convergence in the Metropolis approach. Several practical concerns must be addressed, since posterior distributions might cover a large region of parameter space and have several local optimums. Among them include settling on a manageable sample size for both keeping and throwing out and determining the best method for random seeding .

References

A., Hornberger, G., 1993. How much complexity is war-ranted in a rainfall-runoff model, Water Resources Res., 29 (8),2637–2649.

Bayes, T., 1763. An essay towards solving a problem in thedoctrine of chances. Phil. Trans. Roy. Soc. 53.

Berger, J.O., 1985. Statistical Decision Theory and BayesianAnalysis, Springer Verlag, New York.Box,

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

Brooks, S.P., 1998. Markov chain Monte Carlo and its application,The Statistician. J. Royal Statist. Soc. Series

D.Chen, S., Billings, S.A., Grant, P.M., 1990. Nonlinear system iden-tification using neural networks, Int. J. Control, 51 (6), 1215–1228.

DeFinetti, B., 1937. La pre´vision, ses lois logiques, ses sourcessubjectives. Herman, Paris.

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

Edijatno, E., Michel, C., 1989. A three parameter daily rainfall-runoff model (in French), Houille Blanche, 2, 113–121.

Fiering, M.B., 1967. Streamflow Synthesis. Harvard UniversityPress, Cambridge, MA.Gelman, A., Carlin, J.B.,

G.E.P., Jenkins, G.M., 1976. Time Series Analysis: Forecast-ing and Control. Holden-Day, San Fransisco.Box,

G.E.P., Tiao G.C., 1973. Bayesian inference in statisticalanalysis. Addison-Wesley, Reading, MA.

G.M., Spear, R.C., 1980. Eutrophication in Peel inlet, I:the problem defining behavior and a mathematical model for thephosphorus scenario, Water Resources Res., 14, 29–42.Jakeman,

Grayson, R.B., Moore, I.D., McMahon, T.A., 1992. Physicallybased hydrologic modelling 2. Is the concept realistic?, WaterResources Res., 26 (10), 2659–2666.Hornberger,

S., Spiegelhalter, D. (Eds.), 1996. PracticalMarkov Chain Monte Carlo, Chapman-Hall, New York.

Stren, H.S., Rubin, D.B., 1997. BayesianData Analysis. Chapman and Hall, London, p. 526.Gilks, W.R., Richardson,