ISSN: 2583-7605 (Online)
www.ijsrmst.com

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

J. Singh[1], S. Maurya[2]

[1]Department of Physics, Shri L.B.S. Degree College, Gonda, Uttar Pradesh, India

[2]Researcher, DDU Gorakhpur University, Gorakhpur, Uttar Pradesh, India

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 a more catchment model to show what insights can be obtained using the Metropolis approach. Specifically, they explain how to evaluate correlations for split-sample tests, how to use prior data, and how to measure reliability for hydrological responses not involved in the calibration process. The Metropolis method outperforms more traditional assumptions as a first-order method of estimation when dealing with uncertainties in water samples.

Keywords: hydrologic models, Monte Carlo techniques, probability distribution, Metropolis sampling, parameter uncertainty.

INTRODUCTION:

The main focus of this work, which attempts to provide light on how to calibrate conceptual catchment models, is an accurate evaluation of parameters uncertainty and its influence on model prediction created using data other than the calibration data. 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 generalize 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 1990) or ARMA models do not take into account the physical properties of hydrological processes (Box and Jenkins, 1976). Conceptual models aim to avoid the scale concerns that hydrologists consider most essential and that plague reductionist models by adopting control quantities over which state variables and flow are temporarily and locally averaged. (Nash and Sutcliffe, 1970). Despite the fact that mass conservation

ISSN: 2583-7605 (Online)
www.ijsrmst.com

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. Several 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 x_t (by means of precipitation, evapotranspiration, and pollutant input) into fluxes, and also state variables. A vector of parameters may characterize 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 $f(x_t, \beta)$ and g to build the probability distribution f [ ]. For a given value of the error parameter vector g, we may assume that the residuals $[q_t\ , f(x_t, \beta)]$ 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 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 $f(x_t, \beta)$ 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 (γ, β) may be obtained, starting with some prior distribution $[q_t\ , f(x_t, \beta)\ ]$, by applying the Bayes rule to data D (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(γ, β) that should capture all the 'subjective' information about (,) prior to collecting sample D has every time been difficult for experts.

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 $[(b(\gamma, \beta) = 1]$ throughout the range of parameters. The integral may not equal 1 if the possible numbers are endlessly big. If the posterior

ISSN: 2583-7605 (Online)
www.ijsrmst.com

distribution truly represents a probability distribution, Bayesian inference can still be used even though the uniform prior is flawed. From a hydrological standpoint, the purpose of models calibrations to be determines the posterior probability distributions $p(\beta|D)$, which defines 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 essence, Duan et al. (1992) synthesise the literature and offer a reliable and effective probabilistic search methodology for determining the most likely value. Surprisingly little work has gone into creating techniques for accurate parameter uncertainty assessment in complicated hydrologic models. No credible hydrologist would assume that model parameters have a single, constant value since conceptual models are nothing more than the product of an empirical combination of mathematical operators defining the essential features of an idealised hydrologic cycle. For an analysis of first-order approximations in a hydrologic environment, See Kuczera (1988). The transfer function $f(x_t, \beta)$ is approximated at the first order in standard statistical theory, providing an approximate multinormal description of parameter uncertainty. The first order approximation produces appropriate results if the linearisation of $f(x_t, \beta)$ holds true everywhere in the domain of β for which there is sufficient uncertainty. However, hydrologic model seldom work with such a close approximation. These factors, among others, have led researchers to abandon traditional methods of statistical inference in favour of additional general Monte Carlo-based techniques (for examples, see van Straten and Keesman (1991), Hornberger and Spear (1980) 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 $p(\beta|D)$ with the aid of a random walk that discovers the real distribution. The approach may offer a more accurate representation of the distribution than significance sampling. When there is a substantial discrepancy between the predicted and actual significance distribution, Markov chain sampling is preferable. 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

ISSN: 2583-7605 (Online)
www.ijsrmst.com

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 $p(\beta|D)$ and can be effectively sampled. From the distribution $p(\beta|D)$, N weighted random samples are drawn according to the pattern b I, P I I = 1,..., N. Tanner (1992) and Gelman (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. Two such examples from 1991 are the GLUE approach by Beven and Binley and the Monte Carlo set membership method by van Straten and Keesman. 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 ($\beta|D$). In the case of parameter spaces with a lot of dimensions, this might need a lot of computational time and power. If each Monte Carlo sample were to cover the entire parameter space, it would take an absurd 1010 samples to generate a representative sample of a parameter in a hypercube with a resolution of one tenth of the parameter range, for example. 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. It is difficult to evaluate regional relationships between model parameters and catchment characteristics using first-order approximations and multinormal distributions, to evaluate the significance of deviations in split-sample tests, or to evaluate the usefulness of prediction/confidence limits on future hydrologic responses due to the nonlinear nature of hydrologic models. 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). Markov chain Monte Carlo sampling, which has gained a lot of attention in the literature on Bayesian statistics, is used in the second method. In contrast to significance sampling, this method 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

ISSN: 2583-7605 (Online)
www.ijsrmst.com

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

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

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

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

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

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

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

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

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

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

[10] Fiering, M.B., 1967. Streamflow Synthesis. Harvard UniversityPress, Cambridge, MA.

[11] Gelman, A., Carlin, J.B., G.E.P., Jenkins, G.M., 1976. Time Series Analysis: Forecast-ing and Control.

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

[13] 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.

[14] 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.

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

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