Generalized likelihood uncertainty estimation (GLUE) and approximate Bayesian computation: What's the connection?

David J. Nott,[1] Lucy Marshall,[2] and Jason Brown[2]

Received 5 July 2011; revised 11 September 2012; accepted 1 November 2012; published 20 December 2012.

[1] There has been a recent debate in the hydrological community about the relative merits of the informal generalized likelihood uncertainty estimation (GLUE) approach to uncertainty assessment in hydrological modeling versus formal probabilistic approaches. Some recent literature has suggested that the methods can give similar results in practice when properly applied. In this note, we show that the connection between formal Bayes and GLUE is not merely operational but goes deeper, with GLUE corresponding to a certain approximate Bayesian procedure even when the "generalized likelihood" is not a true likelihood. The connection we describe relates to recent approximate Bayesian computation (ABC) methods originating in genetics. ABC algorithms involve the use of a kernel function, and the generalized likelihood in GLUE can be thought of as relating to this kernel function rather than to the model likelihood. Two interpretations of GLUE emerge, one as a computational approximation to a Bayes procedure for a certain "error-free" model and the second as an exact Bayes procedure for a perturbation of that model in which the truncation of the generalized likelihood in GLUE plays a role. The intent of this study is to encourage cross-fertilization of ideas regarding GLUE and ABC in hydrologic applications. The connection we outline suggests the possibility of combining a formal likelihood with a kernel based on a generalized likelihood within the ABC framework and also allows advanced ABC computational methods to be used in GLUE applications. The model-based interpretation of GLUE may also be helpful in partially illuminating the implicit assumptions in different choices of generalized likelihood.

Citation: Nott, D. J., L. Marshall, and J. Brown (2012), Generalized likelihood uncertainty estimation (GLUE) and approximate Bayesian computation: What's the connection?, *Water Resour. Res.*, *48*, W12602, doi:10.1029/2011WR011128.

1. Introduction

[2] A common approach to uncertainty assessment in the hydrological literature is the generalized likelihood uncertainty estimation (GLUE) technique of *Beven and Binley* [1992]. There has been recent discussion in the hydrological literature concerning the relative merits of GLUE versus formal probabilistic approaches [see *Mantovan and Todini*, 2006; *Stedinger et al.*, 2008; *Beven et al.*, 2008; *Vrugt et al.*, 2009a; *Beven*, 2009, and references therein]. A recent comparison of the two methods concluded that "formal and informal Bayesian approaches have more common ground than the hydrologic literature and ongoing debate might suggest" [*Vrugt et al.*, 2009a, p. 1011]. In this note, we show that the connection between formal Bayes and GLUE goes beyond operational similarity by showing that it is possible to give a formal Bayesian interpretation of GLUE even when the corresponding generalized likelihood is not a true likelihood. In particular, recent developments in approximate Bayesian computation (ABC) methods (originating in genetics with related earlier precedents in the papers of *Diggle and Gratton* [1984] and *Gourieroux et al.* [1993]) provide an illuminating new way of looking at the GLUE procedure. In fact, while not originally motivated as an ABC algorithm, remarkably, GLUE predates the development of ABC methods: the first simple ABC algorithm was discussed in *Tavaré et al.* [1997].

[3] The ABC-GLUE connection is fruitful in a number of ways. First, even if the association is regarded as merely technical it allows advanced ABC computational methods and software to be brought to bear in GLUE applications. Second, in one of the ABC interpretations of GLUE, the generalized likelihood does not correspond to the likelihood but to a certain kernel function that appears in the algorithm; this naturally suggests a way to combine both a likelihood and generalized likelihood, with the probabilistic model allowing description of different sources of variability in a problem but the generalized likelihood reducing sensitivity to model misspecification in directions of interest to a modeler. Third, the model-based interpretation of GLUE may also be helpful in partially illuminating the implicit assumptions in different choices of generalized likelihood. The goal of this study is

[1]Department of Statistics and Applied Probability, National University of Singapore, Singapore.
[2]Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, Montana, USA.

Corresponding author: L. Marshall, Department of Land Resources and Environmental Sciences, Montana State University, 334 Leon Johnson Hall, PO Box 173120, Bozeman, MT 59717-3120, USA. (lmarshall@montana.edu)

0043-1397/12/2011WR011128

thus to promote cross-fertilization of ideas regarding GLUE and ABC, which we believe will motivate continuing interest in the underlying assumptions of GLUE and the numerical implementation of likelihood-free methods in hydrology.

[4] In the next section, we give a brief introduction to Bayesian inference, GLUE, and ABC methods. In section 3, we show how GLUE can be viewed as an ABC algorithm for a certain probabilistic model in which the corresponding generalized likelihood plays the role of the ABC kernel function. We also point out that the generalized likelihood (even when not a proper likelihood) can usually be considered to define a proper likelihood when the truncation in the GLUE algorithm is taken into account. We further include a case study that demonstrates how a representative ABC sampling algorithm and interpretation may be used for an established GLUE approach that attempts to address epistemic hydrologic model errors. Section 4 contains concluding discussion.

2. Background

2.1. Bayesian Inference

[5] Suppose we observe real data $y = (y_1, \ldots, y_n)^T$ and that we have a model that generates an output corresponding to this data $\hat{y} = (\hat{y}_1, \ldots, \hat{y}_n)^T$. The model output depends on some parameters θ that are unknown, so that $\hat{y} = \hat{y}(\theta)$. θ could in principle be high-dimensional, including uncertain model inputs as well as parameters that relate to understanding of dominant hydrological processes. Through the model outputs and a specification for the distribution of model errors $y - \hat{y}$, we can specify a probabilistic model for the observable data y, $p(y|\theta)$ say. We call $p(y|\theta)$ the likelihood function. In a Bayesian approach to inference, we also specify in probabilistic form what we know about the unknown θ before observing data through a so-called prior distribution $p(\theta)$. Together, the prior and likelihood specify a full probabilistic model for (y, θ), $p(y, \theta) = p(\theta)p(y|\theta)$, and then once data are observed we can condition on y to obtain

$$p(\theta|y) \propto p(\theta)p(y|\theta). \tag{1}$$

$p(\theta|y)$ is called the posterior distribution, and it is a probabilistic representation of what we know about the unknown θ after observing data. Inference and prediction from the model are then based on the posterior distribution, at least after appropriate model checking and diagnostics.

[6] The above is a quick summary of the formal Bayesian approach to inference as it applies to hydrological modeling (see, for example, *Gelman et al.* [2003] for a complete discussion of Bayesian statistical inference). Although the application of formal Bayesian methods in hydrology has been increasing, some modelers have cautioned that insufficient attention is paid to the difficulties of specification of a full probabilistic model, with the possibility that inferences of interest might be sensitive to some arbitrary aspects of such a specification. While appropriate diagnostic checks can certainly guard against this sensitivity to some extent it could be argued that at least in some situations an alternative to a full probabilistic model specification is needed. On the other hand, it has also been suggested that one strength of the full Bayesian approach lies in the ability to understand and model contributions of different kinds of uncertainty (such as parameter uncertainty, structural uncertainty, and input uncertainty). *Vrugt et al.* [2009a] and *Beven* [2009] provide a recent discussion of these concerns.

[7] We say something also about how to generate Monte Carlo samples from the posterior distribution (1), as this relates to the interpretation of GLUE given later. Markov chain Monte Carlo methods are most commonly used for this task (see *Gelman et al.* [2003] for an introduction), but here we will discuss only importance sampling approaches. In an importance sampling approach, we start with a so-called importance density $h(\theta)$ on the parameter space, which is hopefully a good approximation to $p(\theta|y)$ and is easy to simulate from directly. Then we generate $\theta_i \sim h(\theta)$, $i = 1, \ldots, n$, and calculate the weights

$$w_i \propto \frac{p(\theta_i)p(y|\theta_i)}{h(\theta_i)}, \tag{2}$$

where the weights are normalized to sum to one, $\sum_{i=1}^{n} w_i = 1$. The points θ_i with weights w_i are a weighted sample from the posterior, in the sense that if $r(\theta)$ is some measurable function of θ, then,

$$\sum_{i=1}^{n} r(\theta_i)w_i$$

is a consistent estimate of $E(r(\theta)|y)$ as $n \to \infty$.

2.2. Generalized Likelihood Uncertainty Estimation

[8] The philosophy behind GLUE is to avoid a difficult full probabilistic model specification but to instead seek out regions in the parameter space where the model predictions are consistent with the observations. It is often the case that there may be widely separated regions of the parameter space consistent with the data (the so-called equifinality problem). The GLUE algorithm is:

Algorithm 1:

[9] 1. Generate points $\eta_1, \ldots, \eta_n$ from the prior $p(\theta)$.

[10] 2. For each η_i, calculate the generalized likelihood values $G(\eta_i; y)$, $i = 1, \ldots, n$.

[11] 3. For a selected threshold c, discard all η_i with $G(\eta_i; y) < c$ and retain the rest, resulting in a remaining "behavioral" set of solutions, which we denote by $\theta_1, \ldots, \theta_m$.

[12] 4. Compute weights

$$w_i = \frac{G(\theta_i; y)}{\sum_{j=1}^{m} G(\theta_j; y)}, i = 1, \ldots, m.$$

[13] 5. For a predictive quantity Δ, estimate the distribution of Δ by the discrete distribution with weights w_i on Δ_i, where Δ_i is the prediction using parameter θ_i, $i = 1, \ldots, m$.

[14] The generalized likelihood $G(\theta; y)$ can be defined in a number of different ways. It can be an actual likelihood function $p(y|\theta)$, in which case the GLUE algorithm corresponds to an importance sampling method for generating samples from the posterior distribution if the truncation at step 3 is ignored. The importance density here is the prior $p(\theta)$. More commonly, however, the generalized likelihood need not be a true likelihood but is merely a function that measures the degree of misfit between the observations and

model predictions in ways that seem relevant to the modeler. We note later, however, that the combination of the generalized likelihood together with the truncation of step 3 usually does define a proper likelihood [*Smith et al.*, 2008]. One common choice for the generalized likelihood is

$$G(\theta;y) = (s^2_{y-\hat{y}})^{-T},$$

where $s_{y-\hat{y}}$ is the standard deviation of the model errors $y-\hat{y}$, and $T > 0$ is a selected constant (termed the shaping factor).

[15] The behavioral threshold, c, may be based on what is referred to as the *limits of acceptability* [*Beven*, 2006]. This refers to an estimate of the effective model error that attempts to take into account multiple sources of uncertainty in the modeling process (such as observational errors and model structure errors, e.g., *Liu et al.* [2009] and *Blazkova and Beven* [2009]).

2.3. Approximate Bayesian Computation

[16] ABC methods originated in genetics with *Tavaré et al.* [1997] and is commonly used in situations, where we have a model that is easy to simulate from, given the model parameters, but where the likelihood is difficult or impossible to calculate. ABC methods are able to simulate from the posterior distribution in Bayesian inference, producing summaries of the posterior useful for inference and prediction but without requiring that the likelihood be calculated. A recent survey of ABC methods is given by *Marin et al.* [2011].

[17] To describe the most basic form of ABC, suppose that the data y are discrete. Consider the following algorithm:

Algorithm 2:

[18] 1. Sample $(\theta, z) \sim p(\theta)p(z|\theta)$ (that is, first sample θ from the prior $p(\theta)$ and then for the sampled θ sample z from $p(z|\theta)$).

[19] 2. If $z = y$ (i.e. if the simulated data from $p(z|\theta)$ matches the observed data y) then output the value of θ simulated at step 1 and stop. Otherwise, return to step 1.

[20] It is not hard to see that the above algorithm is simply a rejection algorithm for simulating from $p(\theta|y)$. Furthermore, provided we can simulate from the model there is no need to evaluate the model likelihood in this algorithm.

[21] The rejection rate of the above algorithm might be very high if the number of possible values for y is large, and if y is continuous the algorithm in the form stated above is not applicable at all. This objection is usually addressed in a number of ways. First, often the data are summarized through summary statistics s that might be of much lower dimension than y (if the statistics s are sufficient statistics there is no loss of information in such a reduction). Second, in step 2, we in general do not require an exact match of z and y but instead require only that $d(y,z) < \epsilon$, where d is some distance function and $\epsilon > 0$ is a tolerance parameter. Third, in addition to the first two innovations there are many advanced computational methods that can be employed to make the methods work well in problems of moderate dimension (see the review of *Marin et al.* [2011], for an entry point into the literature). It is not our intention to discuss advanced ABC methods here. In what follows and with a slight abuse of notation, we continue to write y for the data that we condition on for inference, even if this is a summary statistic obtained from the full data set.

[22] Note that in Algorithm 2, if we accept the sampled θ in step 2 when $\mathrm{d}(y,z) < \epsilon$ we are in effect sampling from the distribution

$$p^*(z,\theta) = \frac{p(\theta)p(z|\theta)I(\mathrm{d}(y,z) < \epsilon)}{\int p(\theta)p(z|\theta)I(\mathrm{d}(y,z) < \epsilon)\mathrm{d}\theta\mathrm{d}z}, \tag{3}$$

i.e., we are sampling from the joint distribution of $p(\theta)p(z|\theta)$ conditioned by the indicator $I(\mathrm{d}(y,z) < \epsilon)$, such that we accept samples for θ which satisfy $\mathrm{d}(y,z) < \epsilon$. The resulting draws for θ then come from the corresponding marginal distribution, which is obtained by integrating equation (3)

$$p^*(\theta|y) \propto p(\theta)\int p(z|\theta)I(\mathrm{d}(y,z) < \epsilon)\mathrm{d}z, \tag{4}$$

which effectively corresponds to replacing the likelihood in equation (1) with an approximate likelihood obtained by averaging in a neighborhood of y. In equation (4), we have shown dependence on y explicitly in the notation to emphasize that this is an approximation to the posterior distribution $p(\theta|y)$ with the approximation increasingly good as $\epsilon \to 0$ (provided that y is a sufficient statistic).

[23] One can also replace the term $I(\mathrm{d}(y,z) < \epsilon)$ in equation (4) with a kernel function $K_\epsilon(y-z)$, where ϵ now denotes a bandwidth parameter, with $K_\epsilon(h)$ concentrating on zero as $\epsilon \to 0$. This results in approximations

$$p^*(z,\theta) \propto p(\theta)p(z|\theta)K_\epsilon(y-z) \tag{5}$$

and

$$p^*(\theta|y) \propto p(\theta)\int p(z|\theta)K_\epsilon(y-z)\mathrm{d}z \tag{6}$$

corresponding to equations (3) and (4), respectively. R. D. Wilkenson (Approximate Bayesian computation (ABC) gives exact results under the assumption of model error, 2008, available at http://arxiv.org/abs/0811.3355, hereinafter referred to as Wilkenson, 2008) has pointed out that if $K_\epsilon(h)$ is a probability density, then we can think of $p^*(\theta|y)$ as an exact posterior distribution in which the data are considered to be generated as $y = z + \delta$ with $z \sim p(z|\theta)$ and $\delta \sim K_\epsilon(\delta)$. Hence, equation (6) can be regarded as an approximation to $p(\theta|y)$, or as an exact posterior for a different model to the one embodied by the likelihood $p(y|\theta)$ where there is an additional additive random error. Wilkinson (2008) has discussed different interpretations that might be given to this error.

3. The ABC-GLUE Connection

3.1. A Theoretical Analysis

[24] Consider the model

$$y = \hat{y}(\theta) + \delta, \tag{7}$$

where $\delta = (\delta_1, \ldots, \delta_n)^T$ is specified as a zero mean random vector. We recognize that the assumption of additive stochastic errors holds only for idealized cases in hydrologic modeling (see, for example, *Beven and Westerberg* [2011]), but note that this assumption is useful for illustrating the ABC-GLUE connection. We leave the exact distribution of δ unspecified, but consider the limiting case where simulating data from the model with parameter θ just results in the deterministic output $\hat{y}(\theta)$. For a generalized likelihood $G(\theta;y)$, suppose it is a function of $y-\hat{y}(\theta)$ only, $G(\theta;y)=g(y-\hat{y})$ say, with g having a global maximum at the origin. Commonly used generalized likelihoods that we know of satisfy these requirements. Now consider the ABC target equation (5) for this model where the kernel $K_\epsilon(h)$ is chosen to be

$$K_\epsilon(h) \propto \begin{cases} g(h), & \text{if } g(h) > c, \\ 0, & \text{otherwise,} \end{cases}$$

where $K_\epsilon(h)$ is normalized to integrate to one, so that $K_\epsilon(h)$ is a probability density function (because of the truncation this is usually possible, regardless of whether $G(\theta;y)$ is a real likelihood, and is certainly possible if g is bounded, say if g is unbounded at the origin and not integrable, then one may consider a truncation of g at some high level since usually $y-\hat{y}\approx 0$ is not attainable anyway in situations where GLUE is applied). Here the kernel parameter ϵ can be considered some decreasing function of c, $\epsilon=\epsilon(c)$ say. If we sample from equation (5) by importance sampling, using the importance density $p(\theta)p(z|\theta)$, we obtain exactly the GLUE algorithm, since if we generate $(\theta_i,z_i)\sim p(\theta)p(z|\theta)$, $i=1,\ldots,n$, the importance weights equation (2) are of the form

$$w_i \propto \frac{p(\theta_i)p(z_i|\theta_i)K_\epsilon(y-z_i)}{p(\theta_i)p(z|\theta_i)} = K_\epsilon(y-z_i)$$

and since $K_\epsilon(y-z_i)$ is just $g(y-z_i)$ if this is bigger than c and zero otherwise, the importance sampling weights are zero for the nonbehavioral solutions of GLUE, and the behavioral solutions have the GLUE weights.

[25] So GLUE corresponds to an importance sampling ABC algorithm where the kernel function is defined from the generalized likelihood. Furthermore, as we pointed out in section 2.3 and following an observation by Wilkinson (2008), the above ABC algorithm can be considered as sampling from an exact posterior rather than an approximation, but where $\delta\sim K_\epsilon(\delta)$. Hence, the GLUE algorithm does correspond to a certain exact Bayesian analysis, where the error distribution depends on the generalized likelihood (even if not a proper likelihood) since the kernel defined by the truncation of the generalized likelihood generally does result in a well-defined model.

[26] If it is said that GLUE is incoherent, what does this mean if GLUE actually is a Bayesian analysis for a certain model? Actually, there are a number of serious points here. The model we have discussed is defined implicitly through the generalized likelihood $G(\theta;y)$ and c, and if we consider models for $y_1,\ldots,y_n$ and for $y_1,\ldots,y_m$ with $m<n$ given θ defined in this way, then the latter model is not obtained from the former in general by marginalization. So if one requires temporal consistency between models that might be considered at different times then GLUE does not have this kind of consistency. Often it is pointed out that with GLUE and certain choices of generalized likelihood there is no increase in information as the sample size increases (but it should be noted that if we were to allow a sample size-dependent choice of c, this would not be the case). Proponents of GLUE may counteract that the purpose of this is to use parameter uncertainty to guard against false confidence resulting from model misspecification and the formal Bayesian likelihood. To this in turn critics might respond that it is unlikely to be the case that parameter uncertainty can adequately represent predictive uncertainty or compensate for model misspecification, and that the uncertainty intervals produced by GLUE are often not well calibrated in the usual statistical sense. It has been argued that this is not always the purpose of such intervals anyway (see, for example, remarks in *Beven* [2006]). In addition, it is suggested that for real problems formal statistical measures are inappropriate (or incoherent in the qualitative sense, *Beven et al.* [2008]). We believe that consideration of the model-based interpretation of GLUE above may in any case be helpful in identifying the implicit assumptions in different choices of generalized likelihood function and might shed some light on how GLUE is best implemented.

3.2. A Case Study

[27] The practical connection between GLUE and ABC methods in hydrology can further be demonstrated by a simple case study. A recent GLUE-based study by *Westerberg et al.* [2011] considered hydrologic model calibration using a catchment's flow duration curve (FDC), arguing the method appeared less sensitive to epistemic model errors than applications using traditional model performance measures. The approach is particularly useful for modeling scenarios where observation time periods for streamflow and model input data do not overlap such that a likelihood cannot be formulated.

[28] We implement here a conceptual rainfall-runoff model used frequently in hydrologic model applications. The probability distributed model (PDM) [*Moore*, 2007] simulates catchment streamflow based on saturation excess runoff processes by conceptualizing the spatial distribution of soil moisture storage as a probability density function. Our version of the model is adapted from *Smith and Marshall* [2010] and contains six unknown parameters (Table 1): *cmax*, the maximum soil moisture storage capacity; *b*, the spatial variability of storage capacity; *kb*, the groundwater recharge constant; *tr*1, the rate of outflow of a surface storage; *tr*2, the rate of outflow of a subsurface storage; and *cf*, a storage threshold factor. The model is applied to the Stringer Creek subcatchment of Tenderfoot Creek Experimental Forest (TCEF) located in Montana. Inputs to the model are time-series data of observed precipitation and potential evapotranspiration on a 6-h time step for the period of study extending from October 2005 to September 2008. Snowmelt and precipitation data are calculated from meteorological observations at the Onion Park SNOTEL station located at TCEF [see *Nippgen et al.*, 2011]. TCEF is of particular interest to this application as it has been previously demonstrated that there is a strong connection between the FDC and shallow groundwater hydrologic connectivity [*Jencso et al.*, 2009; *Jencso and McGlynn*, 2011]. This suggests the FDC will be useful here for identifying hydrologic model parameters and for model diagnostics.

Table 1. PDM Parameters and Sampling Ranges

Parameter	Description (Units)	Sampling Range
cmax	Maximum storage capacity (mm)	[0–400]
b	Exponent of Pareto distribution controlling spatial distribution of storage (none)	[0–1.5]
kb	Groundwater recharge constant (none)	[0–0.05]
tr1	Rate of outflow from surface storage (none)	[0–0.2]
tr2	Rate of outflow from groundwater storage (none)	[0.2–1.0]
cf	Minimum storage threshold (mm)	[0–30]

[29] We calibrate the model using two methods: a GLUE approach using simple Monte Carlo sampling and specifying a generalized likelihood based on the catchment's FDC; and an ABC approach that uses sequential Monte Carlo (SMC) sampling and a summary statistic of the same form as the GLUE generalized likelihood. We set this summary statistic (or generalized likelihood) as the mean absolute error (MAE) of the observed and simulated FDC. Adapting the method used by *Westerberg et al.* [2011], we specify 30 evaluation points (EPs) along the FDC based on equal intervals of observed discharge and use these to calculate MAE. In both GLUE and ABC approach, we specify a model error tolerance, ϵ, (or limit of acceptability) equal to 5% of the FDC exceedance percentage. It should be noted that this tolerance may be based on discharge measurement error as in *Westerberg et al.* [2011], but in the absence of stage-duration data, we select a value such that we can readily demonstrate the connection between GLUE and ABC.

[30] For our GLUE algorithm, we randomly sample candidate parameter sets from specified parameter ranges (Table 1) and simulate the catchment's FDC. Using GLUE terminology, parameter sets that provide an FDC within the identified limit of acceptability are deemed "behavioral" and retained to indicate the parameter uncertainty. For this application, we continue sampling until 500 behavioral parameters are obtained.

[31] Our ABC-SMC approach uses the algorithm developed by *Sisson et al.* [2007]. In SMC-based simulation, a population of parameters (referred to as particles) is evolved from an initial prior distribution, through a sequence of intermediary distributions, until it ultimately represents a sample from the target parameter distribution. In the ABC framework, particles are sampled from preceding populations, and each sampled particle is moved according to a Markov kernel to improve particle dispersion. A decreasing sequence of tolerances $\epsilon(t)$ is defined such that as $\epsilon(t) \to 0$ we obtain the target distribution of interest. Sampled particles that fall outside the prespecified tolerance are rejected. Repeated resampling (and moving) of particles is undertaken until a fixed number of particles are obtained for each population. For our case study, we define the sequence of error tolerances as an exponential decay function for the MAE of the FDC. Particles are initially sampled from the specified parameter ranges (Table 1). We sample 10 sequential populations with the final error tolerance set at 5% of the FDC. The population size is specified to be 500 (the same as our GLUE-based study).

[32] Weighted histograms of the resulting GLUE and ABC sampled parameters for the PDM are given in Figure 1. It should be noted that the GLUE-derived parameter distributions adhere reasonably well to the ABC-SMC parameters. Differences in the distributions may be attributed to sampling density, parameter insensitivity, or differences in the robustness of the simple Monte Carlo GLUE approach and the SMC algorithm. It is obvious that the GLUE algorithm applied here is simply an ABC rejection algorithm as described in Algorithm 2. However, these results also demonstrate the practical similarity between the GLUE and ABC algorithms for hydrologic applications. In the case presented here, the summary statistic is selected based on the lack of streamflow observations during the period of model simulation. In related cases (especially those regarding predictions in ungauged basins) summary statistics may be derived based on certain hydrologic signatures [*Wagener and Montanari*, 2011]. Following the connection outlined in the preceding connection, the use of generalized likelihoods based on such signatures combined with probabilistic error models may provide a way to explicitly model different sources of error but also address epistemic model errors. Recent discussions regarding the sufficiency of certain types of ABC summary statistics [*Blum et al.*, 2012] may thus prove useful to the hydrologic community for understanding the implications of certain types of objective functions in model inference.

4. Discussion

[33] By making a link between the GLUE algorithm and ABC, we have further clarified the relationship between GLUE and formal Bayesian methods. It is remarkable that GLUE is an example of an ABC method that predates the development of ABC [*Tavaré et al.*, 1997]. It might be argued that the link we have outlined here is merely technical. However, even if this is the case, there are some immediate and concrete benefits. The frenetic recent activity on developing more efficient ABC approximations and software [*Marjoram et al.*, 2003; *Sisson et al.*, 2007, 2009; *Beaumont et al.*, 2009] can be brought to bear for GLUE-based applications. A common objection to the use of GLUE revolves essentially around the computational inefficiency of importance sampling. For example, *Vrugt et al.* [2009b], in discussing the Differential evolution adaptive metropolis (DREAM) Markov chain Monte Carlo (MCMC) scheme of *Vrugt et al.* [2008, 2009a], state that "However, DREAM will have a much better efficiency in finding 'acceptable' models as it uses adaptive proposal updating to search for high quality solutions. Use of a simple uniform sampling distribution... as typically done in GLUE, can result in an algorithm that, even after billions of model evaluations, may only have generated a handful of good solutions...". Recent work has then shown how likelihood-based MCMC algorithms may be modified to address GLUE problems [e.g., *Blasone et al.*, 2008; *McMillan and Clark*, 2009]. We believe that the connection outlined here could further encourage adoption of new numerical methods for GLUE, and the availability of ABC software will also help in this regard. Once the distribution targeted by GLUE is understood, it becomes possible to advance the use of sophisticated SMC algorithms, Markov chain Monte Carlo algorithms, and so on with GLUE.

[34] The continuing interest in characterizing input uncertainty in hydrologic modeling (such as the Bayesian total error analysis [BATEA] method of *Kavetski et al.* [2006a, 2006b]) may also benefit from ABC methods. The quantification of hydrologic input error often requires parameter

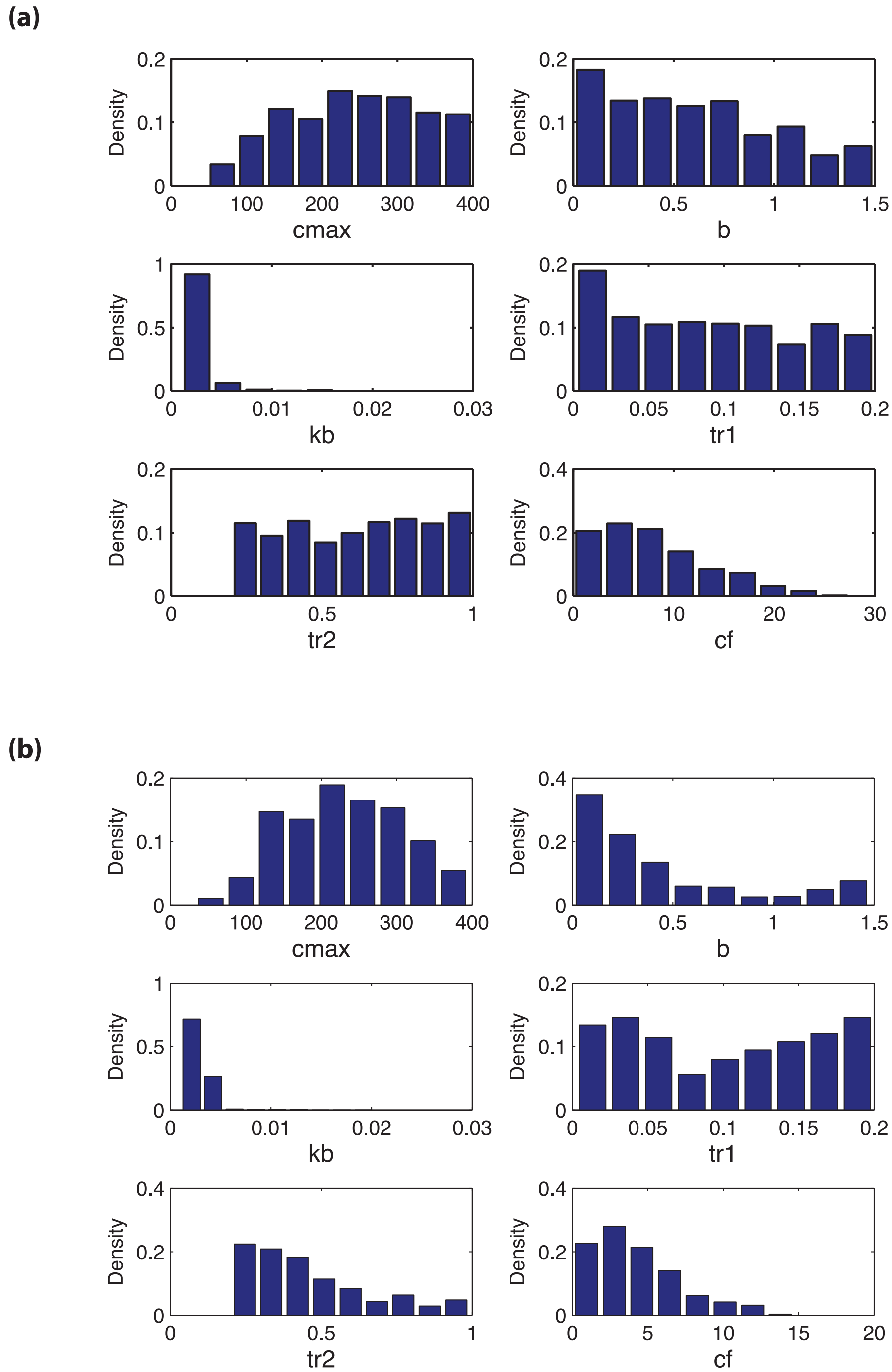

Figure 1. Weighted histograms of (a) sampled GLUE parameters and (b)ABC-SMC particles. A description of the model parameters may be found in Table 1.

sampling from very high-dimensional models due to the introduction of probabilistic models representing rainfall uncertainty. The ABC approach can provide a way of integrating out input error parameters so that model predictive uncertainty may be more easily addressed (see *Nott et al.* [2012] for initial work in this regard).

[35] The connection with ABC also hints at ways to extend GLUE where we want to specify both a probabilistic model as well as a generalized likelihood (since the generalized likelihood is identified with the ABC kernel and not the likelihood we are free to specify a likelihood as well). This allows us to model different sources of uncertainty but to also mitigate the effects of misspecification in directions of interest through choosing the kernel function based on a generalized likelihood. We are currently pursuing work based on these extensions.

[36] **Acknowledgments.** The authors thank Keith Beven, Jasper Vrugt, and a third anonymous reviewer for their thoughtful comments provided on this article. David Nott gratefully acknowledges the support and contributions of the Singapore-Delft Water Alliance (SDWA). The research presented in this work was carried out as part of the SDWAs multireservoir research program (R-264-001-001-272).

References

Beaumont, M., J.-M. Cornuet, J.-M. Marin, and C. Robert (2009), Adaptive approximate Bayesian computation, *Biometrika*, *96*, 983–990.

Beven, K. J. (2006), A manifesto for the equifinality thesis, *J. Hydrol.*, *320*, 18–36.

Beven, K. J. (2009), Comment on "Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?" by Jasper A. Vrugt, Cajo J. F. ter Braak, Hoshin V. Gupta and Bruce A. Robinson, *Stoch. Environ. Res. Risk Assess.*, *23*, 1059–1060.

Beven, K. J., and A. M. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Process.*, *6*, 279–298.

Beven, K. J., and I. Westerberg (2011), On red herrings and real herrings: Disinformation and information in hydrological inference, *Hydrol. Process.*, *25*(10), 1676–1680.

Beven, K. J., P. J. Smith, and J. E. Freer (2008), So just why would a modeller choose to be incoherent?, *J. Hydrol.*, *354*, 15–32.

Blasone, R.-S., J. A. Vrugt, H. Madsen, D. Rosbjerg, M. A. Robinson, and G. A. Zyvoloski (2008), Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling, *Adv. Water Res.*, *31*, 630–648.

Blazkova, S., and K. Beven (2009), A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic, *Water Resour. Res.*, *45*, W00B16, doi:10.1029/2007WR006726.

Blum, M. G., M. A. Nunes, D. Prangle, and S. A. Sisson (2012), A comparative review of dimension reduction methods in approximate Bayesian computation, *Stat. Sci.*, in press.

Diggle, P. J., and R. J. Gratton (1984), Monte Carlo methods of inference for implicit statistical models, *J. R. Stat. Soc. Ser. B*, *46*, 193–227.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2003), *Bayesian Data Analysis*, 2nd ed., CRC Press, Boca Raton, Fla.

Gourieroux, C., A. Monfort, and E. Renault (1993), Indirect inference, *J. Appl. Econ.*, *8*, 85–118.

Jencso, K. G., B. L. McGlynn, M. N. Gooseff, S. M. Wondzell, K. E. Bencala, and L. A. Marshall (2009), Hydrologic connectivity between landscapes and streams: Transferring reach and plot scale understanding to the catchment scale, *Water Resour. Res.*, *45*, W04428, doi:10.1029/2008WR007225.

Jencso, K. G., and B. L. McGlynn (2011), Hierarchical controls on runoff generation: Topographically driven hydrologic connectivity, geology, and vegetation, *Water Resour. Res.*, *47*, W11527, doi:10.1029/2011WR010666.

Kavetski, D., G. Kuczera, and S. W. Franks (2006a), Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory, *Water Resour. Res.*, *42*, W03407, doi:10.1029/2005WR004368.

Kavetski, D., G. Kuczera, and S. W. Franks (2006b), Bayesian analysis of input uncertainty in hydrological modeling: 2. Application, *Water Resour. Res.*, *42*, W03408, doi:10.1029/2005WR004376.

Liu, Y., J. Freer, K. Beven, and P. Matgen (2009), Towards a limits of acceptability approach to the calibration of hydrological models: Extending observation error, *J. Hydrol.*, *367*, 93–103.

Mantovan, P., and E. Todini (2006), Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology, *J. Hydrol.*, *330*, 368–381.

Marin, J.-M., P. Pudlo, C. P. Robert, and R. Ryder (2011), Approximate Bayesian computational methods, *Stat. Comput.*, *22*(6), 1167–1180, doi:10.1007/s11222-011-9288-2.

Marjoram, P., J. Molitor, V. Plagnol, and S. Tavaré (2003), Markov chain Monte Carlo without likelihoods, *Proc. Natl. Acad. Sci. U. S. A.*, *100*, 15,324–15,328.

McMillan, H., and M. Clark (2009), Rainfall-runoff model calibration using informal likelihood measures within a Markov chain Monte Carlo sampling scheme, *Water Resour. Res.*, *45*, W04418, doi:10.1029/2008WR007288.

Moore, R. J. (2007), The PDM rainfall runoff model, *Hydrol. Earth Syst. Sci.*, *11*, 483–499, doi:10.5194/hess-11-483-2007.

Nippgen, F., B. L. McGlynn, L. A. Marshall, and R. E. Emanuel (2011), Landscape structure and climate influences on hydrologic response, *Water Resour. Res.*, *47*, W12528, doi:10.1029/2011WR011161.

Nott, D. J., Y. Fan, L. Marshall, and S. A. Sisson (2012), Approximate Bayesian computation and Bayes linear analysis: Towards high-dimensional ABC, *J. Comp. Graph. Stat.*, in press.

Sisson, S. A., Y. Fan, and M. Tanaka (2007), Sequential Monte Carlo without likelihoods, *Proc. Natl. Acad. Sci. U. S. A.*, *104*, 1760–1765.

Sisson, S. A., Y. Fan, and M. Tanaka (2009), Sequential Monte Carlo without likelihoods: Errata, *Proc. Natl. Acad. Sci. U. S. A.*, *106*, p. 16889.

Smith, T. J., and L. A. Marshall (2010), Exploring uncertainty and model predictive performance concepts via a modular snowmelt-runoff modelling framework, *Environ. Modell. Software*, *25*(6), 691–701.

Smith, P. J., K. J. Beven, and J. Tawn (2008), Informal likelihood measures in model assessment: Theoretic development and investigation, *Adv. Water Res.*, *31*, 1087–1100.

Stedinger, J. R., R. M. Vogel, R. Batchelder, and S. U. Lee (2008), Appraisal of the generalized likelihood uncertainty estimation (GLUE) method, *Water Resour. Res.*, *44*, W00B06, doi:10.1029/2008WR006822.

Tavaré, S., D. J. Balding, R. C. Griffiths, and P. Donnelly (1997), Inferring coalescence times from DNA sequence data, *Genetics*, *145*, 505–518.

Vrugt, J. A., C. J. F. ter Braak, M. P. J. M. Clark Hyman, and B. A. Robinson (2008), Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation, *Water Resour. Res.*, *44*, W00B09, doi:10.1029/2007WR006720.

Vrugt, J. A., C. J. F. ter Braak, H. V. Gupta, and B. A. Robinson (2009a), Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?, *Stoch. Environ. Res. Risk Assess.*, *23*, 1011–1026.

Vrugt, J. A., C. J. F. ter Braak, C. G. H. Diks, B. A. Robinson, J. M. Hyman, and D. Higdon (2009b), Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling, *Int. J. Nonlin. Sci. Numer. Simul.*, *10*, 273–290.

Wagener, T., and A. Montanari (2011), Convergence of approaches toward reducing uncertainty in predictions in ungauged basins, *Water Resour. Res.*, *47*, W06301, doi:10.1029/2010WR009469.

Westerberg, I. K., J.-L. Guerrero, P. M. Younger, K. J. Beven, J. Seibert, S. Halldin, J. E. Freer, and C.-Y. Xu (2011), Calibration of hydrological models using flow-duration curves, *Hydrol. Earth Syst. Sci.*, *15*, 2205–2227, doi:10.5194/hess-15-2205-2011.