SAND2008-6886C

A MODIFIED GENERALIZED LIKELIHOOD UNCERTAINTY ESTIMATION (GLUE) METHODOLOGY

Hui-Hai Liu, Yingqi Zhang and James Houseworth

Earth Sciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road Berkeley, California 94720, hhliu@lbl.gov

The Generalized Likelihood Uncertainty Estimation (GLUE) methodology has been widely used in many areas as an effective and general strategy for model calibration and uncertainty estimation associated with complex models. The application of the GLUE requires a formal definition of a likelihood function. However, it has been recognized that the choice of a likelihood function is inherently subjective, which, in turn, introduces a new kind of uncertainty–the epistemic uncertainty in the GLUE methodology. In this study, we propose a practical framework to address this uncertainty by using multiple likelihood functions. The final uncertainty estimate results are obtained by combining those results from individual likelihood functions based on probability theory. Through an analysis of the probabilities of four net infiltration maps at Yucca Mountain, Nevada, we demonstrate that (1) it is important to consider the uncertainty caused by the subjectivity of the likelihood selection in the GLUE application, and that (2) the proposed method can effectively address this epistemic uncertainty.

I. INTRODUCTION

Uncertainties associated with hydrological models may increase with the incorporation of more complex processes. Good modeling practice should quantify and reduce these uncertainties. Ref. 1 has provided a comprehensive review on understanding, quantifying, and reducing uncertainties in hydrological models. Among different uncertainty-analysis methodologies, the generalized likelihood uncertainty estimation (GLUE) methodology (Ref. 2) has been widely used in the hydrological literature, because it is relatively simple and easy to understand and implement. However, the application of GLUE requires a formal definition of a likelihood function. As a result of difficulties in rigorously defining a likelihood function for a given hydrological problem, the application of GLUE itself introduces epistemic uncertainty. (The epistemic uncertainty refers to the uncertainty related to incomplete or inadequate information in the model selection or parameter determination.) The associated subjectivity affects the calculation of model uncertainty, as shown previously in Ref. 3. Therefore, proper consideration of this epistemic uncertainty is essential for many practical applications. Although there is a large body of literature for the GLUE methodology, little attention has been given to addressing epistemic uncertainty associated with the selection of the likelihood function.

The main purpose of this study is to develop a practical framework for considering the epistemic uncertainty introduced by the subjectivity of likelihood function selection in the GLUE methodology.

II. A MODIFIED GLUE METHODOLOGY

II.A. Basic Concept of the GLUE Methodology

The GLUE methodology was first proposed in Ref. 2 as a framework to estimate uncertainty from equally acceptable models or parameter sets. It recognizes that equivalence of parameter sets or models in hydrological systems exists in many cases. Therefore, rather than using a single "optimal" parameter set or model, the GLUE approach assesses uncertainty in those parameter sets or models by weighting them according to how well a given model (or parameter set) has performed. The evaluation of the performance is based on a comparison between model results and measured data. In particular, Ref. 2 emphasized that the use of the term "likelihood" in the GLUE methodology is meant in a very general sense, as a possibilistic measure of how well the model conforms to the observed behavior of the hydrological system, and not in the restricted sense of the maximum likelihood theory. This flexibility is necessary because a seemingly mathematically rigorous likelihood function may not be satisfactory for complex practical problems, where the assumptions used to derive the likelihood function cannot be fully validated. Due to its simplicity and practicality, the GLUE methodology has been widely used in hydrological applications and other related areas. We refer readers to Ref. 2 for the details of the GLUE methodology.

As previously indicated, the choice of a likelihood function is inherently subjective in the GLUE, and this choice affects the calculation of model uncertainty [Ref.

3]. This means that the uncertainty in choosing a likelihood function needs to be appropriately addressed in the GLUE procedure.

II.B. A Modified GLUE Methodology

A modified GLUE methodology is developed herein to address the epistemic uncertainty associated with the selection of a likelihood function. It is difficult, if not impossible, to define a single best likelihood function for a given problem, while a number of acceptable likelihood functions may exist. (According to Ref. 2, a likelihood function in GLUE is considered to be acceptable when it equals to zero for all the simulations that are considered to exhibit behavior dissimilar to the system under study and increases monotonically as the similarity in behavior increases. These are not very restrictive requirements and could be satisfied by many likelihood functions.) Unlike the traditional GLUE methodology, in which a single likelihood function is used for a given problem, we propose to include all acceptable likelihood functions for the uncertainty analysis. The results from multiple likelihood functions can be combined using probability theory. Let $f_k(z)$ be the uncertainty estimation result using a single likelihood function L_k, and $P(L=L_k)$ be the probability for the likelihood function $L=L_k$ to be true (e.g., the kth likelihood function provides a true estimate). Then, the final estimation results are:

$$f(z) = \sum_k f_k(z) P(L = L_k) \quad (1)$$

In the case where $f_k(z)$ is represented in a discrete form, Eq. (1) can be written as:

$$P(Z = z_j) = \sum_k P(Z = z_j | L = L_k) P(L = L_k) \quad (2)$$

where $P(Z=z_j)$ is the probability for the event $Z = z_j$ (e.g., the jth model or values of the jth parameter set are true) to occur, and $P(Z = z_j | L = L_k)$ is the probability for $Z=z_j$ to occur when $L=L_k$ (e.g., the kth likelihood function is true). In the case study to be discussed in the next section, Eq. (2) is used to demonstrate the implementation procedure of our proposed method.

Our modified GLUE methodology can be considered a generalization of the original version of the GLUE. If the weight of an individual likelihood function L_k is assigned to be one, then our approach is reduced to the traditional GLUE. In other words, the traditional GLUE puts all confidence (or belief) into a single likelihood function, whereas our modified GLUE acknowledges the existence of multiple likelihood functions accepted for a given problem and considers all of them. Also, note that Ref. 2 suggested ways of combining likelihood functions to handle different types of observations. We use their suggested method to handle data at different locations and for different data types.

The modified GLUE methodology requires the selection of likelihood functions and the specification of weights for the different likelihood functions. In theory, all the *acceptable*, *independent* likelihood functions need to be considered. In practice, the selected likelihood functions should exhibit substantially different mathematical forms or have different theoretical bases such that a relatively large range of likelihood-function types are considered, and any overlaps between different selected likelihood functions are avoided. Furthermore, the selected likelihood functions should not be theoretically inconsistent with the data and physical processes considered in the model. Finally, the methodology also requires the independent specification of $P(L=L_k)$ – the probability that a likelihood function provides the true information. This probability can be thought of as likelihood function weights to be specified by the user. The specification of weights for different likelihood functions should be based on the confidence in (or reliability of) these functions for the problem under consideration. The weights are different if the confidence in these functions is different.

Note that the methodology that has been used for aggregating results of expert elicitation is similar to the modified GLUE. Expert elicitations (or judgments) enter several aspects of many scientific endeavors (Ref. 4). During an expert elicitation procedure, equally qualified experts from different related scientific areas may provide different probability density functions for a given parameter or analysis result. In most cases, a collective opinion (represented by a pdf) is needed from results of individual experts. An analogy exists between the expert elicitation and the modified GLUE methodology. In Eq (2), $P(Z=z_j)$ and $P(Z = z_j | L = L_k)$ correspond to the collective and individual uncertainty analysis results from experts, respectively. Each expert can be considered to be equivalent to a likelihood function L_k in this study. Similar probabilistic arguments were used in Ref. 4 for developing a mathematical formulation for aggregating results expert elicitations. Uniform weights were suggested in their work.

III. APPLICATION

This section demonstrates application of the modified GLUE methodology to determining probabilities for infiltration maps at Yucca Mountain, Nevada, the proposed site for disposal of high-level nuclear waste. The advantage of using the modified GLUE methodology for the problem under consideration is also discussed.

II.A. Infiltration Maps at Yucca Mountain and the Relevant Data

Net infiltration is a key hydrologic parameter for controlling percolation rate, groundwater recharge, potential seepage into waste emplacement drifts, and radionuclide transport. Ref. 5 developed net infiltration maps for the region immediately surrounding Yucca Mountain, based on present-day climate and two future climates. Forty infiltration maps were generated with the same probability of occurrence. From these forty maps, a cumulative distribution function (CDF) was obtained, and an annual mean net infiltration rate at different percentiles was found. Four infiltration maps, namely, 10^{th}, 30^{th}, 50^{th}, and 90^{th} percentile maps have been selected by the Yucca Mountain Project and will be used in a system performance analysis. However, during the process of generating net infiltration maps, only climate and shallow-soil-layer information has been taken into account because of computational limitations. Data from the deep unsaturated zone that provide additional information regarding infiltration/percolation processes at Yucca Mountain are not used. Using 10^{th}, 30^{th}, 50^{th}, and 90^{th} percentile net infiltration maps, we can simulate temperature and chloride concentration values at different locations in the unsaturated zone of Yucca Mountain (Ref. 6). Figure 1 shows a comparison between simulated and measured temperature data at a vertical borehole, and Figure 2 between simulated and measured chloride concentration data at a horizontal tunnel, with simulation results from the unsaturated zone flow model using these four net infiltration maps as input. In this study, both chloride concentration and temperature data observed in the unsaturated zone were used for uncertainty analysis of the four selected net infiltration maps.

II.B. Multiple Likelihood Functions

The modified GLUE methodology and four commonly used likelihood functions are used here for determining the probability for each of the four selected net infiltration maps based on Eq (2).

The traditional likelihood function (with the assumption that the residuals are independent and follow a Gaussian distribution) (Ref. 7) was chosen as our first likelihood function:

$$L_{1i} = \prod_{j=1}^{K} (2\pi\sigma_j^2)^{-0.5} \exp\{ -\frac{(x_{ij} - X_j)^2}{2\sigma_j^2} \} \qquad (3)$$

where subscript i (= 1, …, 4) is the index of the net infiltration map; j (=1, …, K) is the index of spatial location (with a total number of K), X is the measured value (temperature or chloride concentration), x is the corresponding simulated value, and σ is the standard deviation of residuals, which is the measurement error if the model were perfect. The same notations are applied in the following equations.

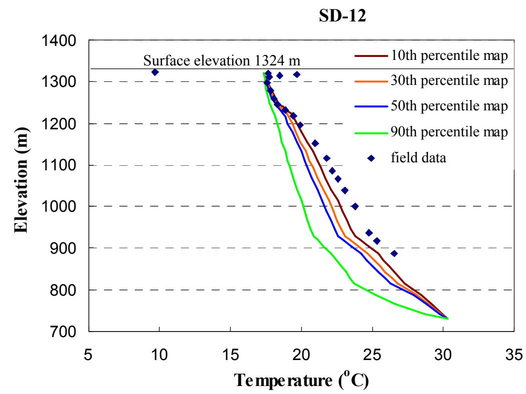

Fig. 1. A comparison between simulated and observed temperature distributions at borehole SD-12 (Refs. 6 and 8).

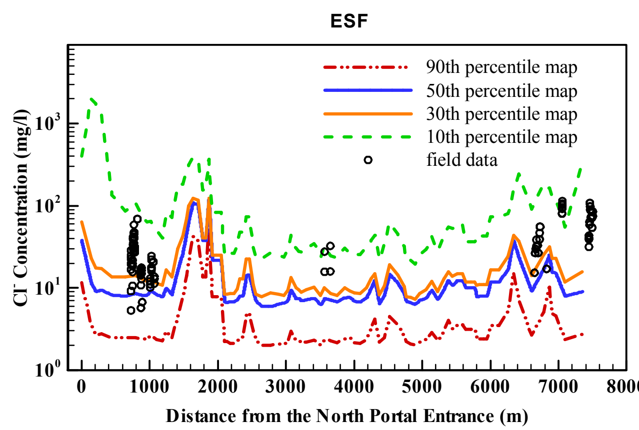

Fig. 2. A comparison between simulated and measured chloride concentration data at a horizontal tunnel (Refs. 6 and 8).

Ref. 2 provided a likelihood function for observations from multiple sites, as shown in Eq. (4).

$$L_{2i} = \{ \sum_{j=1}^{K} \frac{W^*_j}{(x_{ij} - X_j)^2} \}^N \qquad (4)$$

where W^*_j is the weight for observation j such that $\sum W^*_j = 1$. N is a parameter chosen by the user. Note 1) if $N = 0$, then all the net infiltration maps will have the same likelihood value, and if $N \to \infty$, the best simulation will be singled out, having a rescaled likelihood value of 1, with all others 0. The advantages of using this measure are that the parameter N gives us the flexibility of choosing to what degree the likelihood values of the maps are separated. In this study, the likelihood function defined in Eq. (4) (with $N = 0.5$ and $N = 1$) was chosen as the second likelihood function. The weights W_j* were chosen to be the same for all measurements.

The third likelihood function is directly taken from Ref. 2:

$$L_{3i} = 1/(\prod_{j=1}^{K} (x_{ij} - X_j)^2) \qquad (5)$$

In this function, the measurement error is not accounted for in the formulation.

The last likelihood function (equivalent to a simple fuzzy membership (Ref. 9)) is used to express a relative degree of belief that an infiltration scenario is a good estimate of the real infiltration history. Examples of this likelihood function include a triangular function, a trapezoidal membership function, and a beta distribution. For this analysis, the commonly used triangular function has been chosen, which can be expressed as (Ref. 9):

$$f_{ij} = 1 - \frac{|x_{ij} - X_j|}{\varepsilon} \qquad (6)$$

where ε defines acceptable error in an observation. In our case, ε was chosen as the maximum residual (absolute difference between simulated quantities and measured quantities, for temperature values and chloride concentrations) from all locations and all infiltration maps. Our last likelihood function (L_{4i},) is then defined as the arithmetic means of the likelihood function $f_{i,j}$ at multiple locations. This treatment is equivalent to one of the methods suggested in Ref. 11– summation of all the individual likelihood functions at multiple locations.

Although a relatively small number of likelihood functions are used here, they represent four different types of likelihood functions expected to give different measures of comparison between simulated and observed results, and have been commonly used in the GLUE framework or related applications.

II.C. Determination of probability values for infiltration maps

For a given likelihood function L_k , the probability of the ith infiltration map is

$$P_i^k = \frac{P_i^0 L_{ki}}{\sum_i P_i^0 L_{ki}} \qquad (7)$$

where P_i^0 is the prior probability for the ith infiltration map. Eq. (2) is used to determine the final probability value considering all the likelihood functions. In this study, we consider $P(L=L_k)$ to be the same for all the likelihood functions. When using both of temperature and chloride data sets, the final likelihood function is a product of likelihood functions for each individual data set.

III. RESULTS

We plot a box-and-whisker diagram (Fig. 3) for the probability values of the four infiltration maps, calculated using each of the likelihood functions and the combined two data sets. The diagram contains five summary statistics (the smallest value, lower quartile (25th percentile), median (50th percentile), upper quartile (75th percentile), and largest value). The final probability values for the four net infiltration maps considering all the likelihood functions from the boxplot are listed in Table 1. Results using each of the two data sets are very close to those from a combination of the two data sets. This supports the robustness of the modified GLUE and reasonableness of the estimated probability values (Table 1).

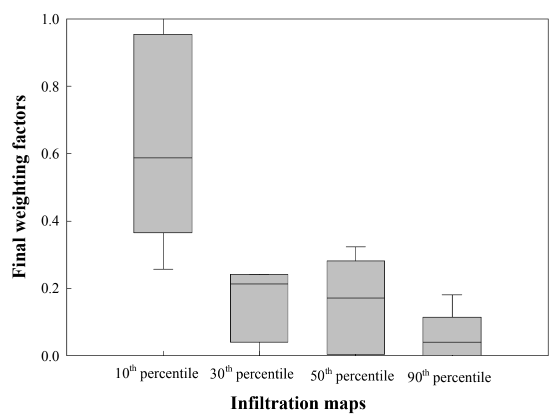

Fig.3. The box-and-whisker diagram for the probability values of the four net infiltration maps, calculated using each of the four likelihood functions (Ref. 8).

TABLE I. Probability Values for the Four Infiltration Maps Considering All the Likelihood Functions

	10th Percentile Map (%)	30th Percentile Map (%)	50th Percentile Map (%)	90th Percentile Map (%)
Probability	61.9	15.7	16.5	6.0

Fig. 3 shows that the probability values are found to be considerably sensitive to the selected likelihood function. For example, the probability value for the 10th percentile map varies from 25.7% to 100% for individual likelihood functions. Among the four likelihood functions, the two extreme likelihood functions are the first one L_1 and the last one L_4. L_1 puts all the weight on the infiltration map that gives the best matches to the observed data - with 100%, 0, 0, 0 for the 10th, 30th, 50th, and 90th percentile maps, respectively. L_4 tends to even out the weights on different infiltration maps - with 25.7%, 23.9%, 32.2%, 18.2% for the 10th, 30th, 50th, and 90th percentile maps, respectively. These results support the importance to account for the epistemic uncertainty caused by the subjective likelihood functions, because of the difficulty in defining a single best likelihood function and the large degree of variability in probability values calculated from different likelihood functions (Fig. 3). Also note that the probability value for the 30th percentile map (Table 1) is a little lower than the probability value for the 50th percentile map as a result of difference between the prior probabilities of the two maps.

The determined probability values (Table 1) are compared with results from an expert elicitation. A group of seven experts was assembled to participate in an expert elicitation panel to provide their judgments concerning key uncertainties associated with unsaturated flow at Yucca Mountain. The resulting assessments and probability distributions provide a reasonable aggregate representation of the knowledge and uncertainties concerning unsaturated zone flow at Yucca Mountain. A comparison is made between the cumulative probability distribution of the average infiltration rate over the repository footprint obtained using the modified GLUE methodology, and the expert elicitation probability distribution (Ref. 12) for percolation flux through the repository footprint. The expert elicitation concluded that net infiltration over the repository footprint and percolation through the repository footprint are quantitatively similar, because flow through the unsaturated zone above the repository was expected to be predominantly vertical. This behavior is consistent with the results of the unsaturated zone flow model.

As shown in Fig. 4, the probability distribution obtained using the modified GLUE methodology is consistent with the aggregate flux probability distribution obtained by expert elicitation. The aggregate probability distribution is the equally weighted combination of the individual probability distributions developed for percolation flux through the repository by each of the seven experts. The figure also shows the distribution for average net infiltration rate over the repository footprint as developed by the infiltration model, which provides the prior probabilities used in the modified GLUE methodology. The substantial difference between the distribution from the infiltration model and the integrated result using the GLUE methodology shows the importance of incorporating the temperature and chloride data into the evaluation of infiltration flux. The consistency between the probability distributions obtained from the two independent studies (expert elicitation and the modified GLUE) suggests that the probability values for net infiltration maps estimated with the modified GLUE methodology are reasonable, and that the use of multiple likelihood functions provides a more robust estimate than using a single likelihood function.

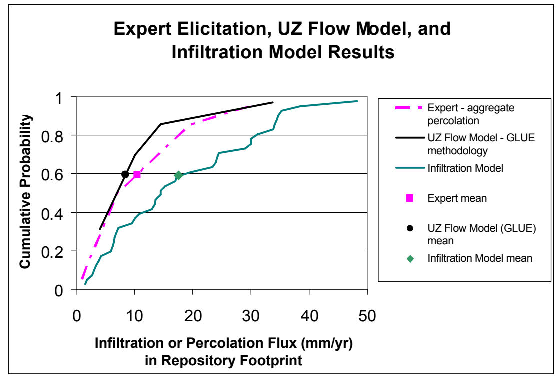

Fig. 4. Comparison among the water flux probability distributions over the repository footprint obtained from the modified GLUE, the expert elicitation (Ref. 11), and infiltration model (Ref. 5)

II. CONCLUSIONS

The GLUE methodology has been shown to be a powerful approach for conducting model calibrations and predictions for highly uncertain and complex hydrological systems. However, the choice of likelihood functions within the framework of GLUE is subjective, and different likelihood functions may give very different analysis results. To account for this uncertainty, we have proposed the modified GLUE methodology. Instead of seeking a single, best likelihood function for a given problem, we explicitly acknowledge that several acceptable likelihood functions may co-exist for a complex problem. Using the modified GLUE methodology, we performed an uncertainty analysis with each individual selected likelihood function, and then obtained the final result (based on a probabilistic framework) by weighting the results from individual likelihood functions using the probability for the corresponding individual likelihood function to be true. A similar methodology has been used in aggregating expert elicitation results in the literature.

To demonstrate the implementation procedure of the modified GLUE methodology, we presented a case study for determining probability values for four selected net infiltration maps at Yucca Mountain, Nevada. This determination is based on prior information of the probabilities and comparisons between simulated and observed chloride concentration and temperature distributions in the unsaturated zone. Four selected likelihood functions are used in the case study. As expected, significant variability in analysis results exists among different likelihood functions. The final analysis results obtained from the modified GLUE methodology are found to be very close to the independently obtained aggregation results of an expert elicitation of percolation flux distribution, supporting the practicality and robustness of the modified GLUE methodology. From this application of modified GLUE methodology, we have demonstrated that (1) it is important to consider the uncertainty caused by the subjectivity of the likelihood-function selection in the GLUE application, and that (2) the proposed method can effectively address this epistemic uncertainty.

ACKNOWLEDGMENTS

This work was supported by the Director, Office of Civilian Radioactive Waste Management, of the U.S. Department of Energy (DOE), under DOE Contract No. DE-AC02-05CH11231.

REFERENCES

1. Y. LIU and H.V. GUPTA, "Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework," *Water Resour. Res.*, **43,** W07401(2007).
2. K. BEVENB and A. BINLEY, "The future of distributed models: model calibration and uncertainty prediction," *Hydrological Processes*, **6**, 279(1992).
3. J. FREER, K. BEVENB and B. AMBROISE, "Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach," *Water Resources Research*, **32**, 2161(1996)
4. E. ZIOZ and G.E. APOSTOLAKIS, "Accounting for expert-to-expert variability: A potential source of bias in performance assessment of high-level radioactive waste repositories," *Ann. Nucl. Energy*, **24**, 751(1997)
5. J. S. STEIN, *Simulation of Net Infiltration for Present-Day and Potential Future Climates, MDL-NBS-HS-000023 REV 01*, Sandia National Laboratories (2007).
6. Y. Wu, *UZ Flow Models and Submodels MDL-NBS-HS-000006 REV 03*, Sandia National Laboratories (2007).
7. J. CARRERA and S.P. NEUMAN, "Estimation of Aquifer Parameters under Transient and Steady-State Conditions .1. Maximum-Likelihood Method Incorporating Prior Information," *Water Resources Research*, **22**, 199(1986).
8. Y.Q. ZHANG, H. H. LIU and J. HOUSEWORTH, "*A modified generalized likelihood uncertainty estimation (GLUE) methodology: Considering the Subjectivity of Likelihood Function Selection*," In review (2008).
9. K. BEVEN, "A manifesto for the equifinality thesis," *Journal of Hydrology*, **320**, 18(2006).
10. S. W. FRANKS and K.J. BEVEN, "Estimation of evapotranspiration at the landscape scale: A fuzzy disaggregation approach," *Water Resources Research*, **33**, 2929(1997).
11. S. K. ZAK, K. BEVEN and B. REYNOLDS, "Uncertainty in the estimation of critical loads: A practical methodology," *Water Air and Soil Pollution*, **98**, 297(1997).
12. CRWMS M&O, *Unsaturated Zone Flow Model Expert Elicitation Project*, Las Vegas, Nevada (1997).