공학석사 학위논문

A Comparative Study on the GLUE Likelihood Definitions for a Hydrologic Model

수문 모델을 위한 GLUE 우도함수 정의에 관한 비교 연구

2019년 8월

서울대학교 대학원

건설환경공학부

한 기 돈

Abstract

The uncertainty analysis is an important subject to hydrologic models and to assess climate change. For hydrologic model, GLUE is frequently adopted methodology. However, the method is still in discussion on the definitions of likelihood measure; the informal and formal likelihood definitions. Informal likelihood definition usually results in unreliable uncertainty interval with respect to its significance level. Due to this limitation, some researchers have suggested a formal likelihood definition based on a Bayesian approach. In the course of its definition, the explicit form of model error should be considered.

For real application, the model error should be considered as non-normal, correlated, heteroscedastic. The diverse strategy was adopted for this task. Different from the previous strategy, three different approaches were used for generalization of error structure in this study. The stable distribution was used for the non-normal attribute, ARFIMA(0,d,0) was adopted for identifying correlation structure, division of period into subperiods was applied for heteroscedasticity. Prior to this analysis of error structure, identifiable epistemic uncertainty was reduced by the regression method.

After all, uncertainty estimation results from informal and formal likelihood definitions were compared with each other. The result shows us that a formal likelihood definition gives us a more statistically reliable result but have large uncertainty intervals due to consideration of model error terms. In contrast to

this, informal likelihood gives consistent result for the calibration period and validation period, but it lost statistical meaning when the uncertainty interval should cover some extreme values. Also, the facet of the heavy-tail property of stable distribution as error distribution was excavated, and it gives fruitful insight in the context of climate change.

Keywords : Uncertainty analysis, GLUE methodology, formal likelihood, informal likelihood, error structure

Student ID : 2017-24832

Table of Contents

Lists of Figures

List of Tables

Chapter 1. Introduction

1.1 Background and Necessity of Study

1.1.1 Watershed model

The water is an indispensable component of life on the earth. The fact that famous cradles of civilization were located around the river has implication about the important position of water in the advancement of human civilization. After the industrial revolution, the contamination of water increases a lot because of the advent of many kinds of anthropogenic activities. Moreover, the conflicts between nations increase due to the scarcity of the water by the impact of climate change in many regions (Barlow & Clark, 2017). Now, the water became important resources to be managed carefully as a cohesive whole.

This integrated water management could be accomplished within watershed modeling. There has been a proliferation of watershed models since the development of the Stanford Watershed Model in 1955 by Crawford & Linsley. The watershed model can be distinguished as a lumped or distributed model according to their process description. A distributed model uses equations considering spatial variability of processes, input, boundary conditions, system characteristics, but lumped model doesn't take account of it (SinghVijay, 2012). For example, the well-known hydrological model named SWAT (Soil & Water Assessment Tool) is a semi-distributed parameter model

because it uses the concept of HRU (hydrologic response unit) to consider the spatial variability but not same as the fully distributed model. Likewise, the model which will be discussed in this thesis is one of physically combined lumped/distributed model, Modified Generalized Watershed Loading Function (MGWLF). The GWLF model simulates runoff, sediment, nutrient loadings for source areas within a watershed. It uses the SCS curve number (CN) approach to calculate runoff, and uses the universal soil loss equation (USLE) algorithm to compute sediment yield. It also calculates nutrients yield by assigning N and P coefficients to surface runoff and to sediment from agricultural source areas (Haith Douglas, Mandel Ross, Wu Ray Shyan, 1992).

In recent works, Qi et al. (2017) compared SWAT and GWLF simulation performance in humid south and semi-arid north of China using evaluation statistics (R^2, NSE, RSR, PBIAS) (Qi, et al., 2017) . They reported that both models reproduce almost the same accuracy in simulating monthly streamflow, sediment, total nitrogen loadings during the simulation period. While SWAT performed better at detailed representation, GWLF could produce much better average values. In summary, they suggested that GWLF is recommended when there are not enough data to apply SWAT because GWLF can reproduce similar result with SWAT even if we need fewer input data to run it due to its simpler structure.

Seoul National University Watershed Simulation (SNU-WS) is a GIS-based watershed simulation tool for assessment of water balance and nutrient and sediment loads caused by point and non-point sources in watershed and

stream systems. All the component of it is connected and customized into a non-proprietary, open source GIS MapWindow software (www.mapwindow.org) as plug-ins. The modeling tool is programmed in VB.NET, and designed to complement and interoperate with full-featured MapWindow GIS functions. Minimum data input requirements, ease of applications and providing reasonable outcomes are the appealing points of this modeling tool. This tool has been developed based on a modified version of GWLF which can be applied for sparsely and ungauged regions as mentioned above. However, to support the calibration and validation processes, an uncertainty analysis is still on development.

Because uncertainty is a nature of hydrological and climatological data, we cannot produce a meaningful result without performing uncertainty analysis. In addition, if climate change is taken into account for future study on the responses of hydrological processes to its impacts, the importance of uncertainty analysis will increase due to an accumulation of more uncertainty characteristics. Upon the above needs, the consideration of uncertainty analysis in the hydrologic model cannot be neglected.

1.1.2 Uncertainty analysis

To deal with uncertainty analysis, we need to figure out what the uncertainty is. This calls for some philosophical considerations. Usually, Uncertainty can be classified as two types even though there is controversy about it. One is Aleatory Uncertainty and the other is Epistemic Uncertainty

(Roy & Oberkampf, 2011). To be brief, Aleatory Uncertainty is implicit uncertainty in nature which we cannot eliminate even if we get more information about the system. So, it can be interpreted as variability of nature. This kind of uncertainty is mostly analyzed by describing it as probability density function or probability mass function. On the other hand, Epistemic uncertainty is the one which can be figured out by getting more information about the system (Roy & Oberkampf, 2011). This one usually can be described using some interval which has possible values in it. For example, when we set the probability density function for some random variable, this description of variability can be regarded as aleatory uncertainty. But if we don't know the exact values of some parameters which determine the form of a probability density function, this uncertainty can be distinguished as epistemic uncertainty. In most engineering situation, it is noted that the epistemic uncertainty is more difficult to be modelled appropriately than aleatory uncertainty.

In Roy and Overkampf (2011), they suggested the framework for uncertainty quantification in scientific computing. First of all, they categorized the uncertainties in the system as uncertainty in model inputs, the one in numerical approximation, the one in model form. Then, they regard the uncertainty of inputs as a combination of aleatory uncertainty and epistemic uncertainty and both the one of numerical approximation and model form as epistemic uncertainty. Through this classification, they quantified the uncertainty of system responses. However, the uncertainty cannot be classified simply as above in most case of hydrological engineering application. There are a variety of sources of uncertainty in hydrological modeling such as

measurement uncertainty. This cannot be differentiated and analyzed separately simply as in the case of the above.

In this context, there is popular methodology, so-called GLUE (Beven & Binley, The future of distributed models: Model calibration and uncertainty prediction, 1992), which is frequently used in hydrological modeling application. The most beneficial point of this method is that it analyzes the uncertainty in the system as a cohesive whole, not separately. Actually, this approach is more persuasive in most real application because many uncertainties in the real world are connected with each other. Through accepting this concept, we can get a philosophy of equifinality. Equifinality means that there can be many equally probable settings of model which behaves well in same level (Beven & Binley, 1992 ; Beven & Binley, 2014). The procedures are mainly composed of three steps (Mirzaei , Huang, Shatirah, & Ahmed, 2015). The first step is sampling the parameter of interest with the prior distribution of it. The next step is to evaluate the simulation result for each parameter through assigning likelihood measure. The final one is to make empirical cumulative distribution function using simulation result weighted by its likelihood measure and denote specified quantile value to represent uncertainty interval. The more sophisticated description will be given in chapter 2, theoretical background.

After the first application of this method by beven (Beven & Binley, The future of distributed models: Model calibration and uncertainty prediction, 1992), there was a great debate about objectivity of this method due to the

subject choice of likelihood function dependent on individuals (Mantovan & Todini, 2006; Stedinger, Vogel, Lee, & Batchelder, 2008). Numerous researches (Stedinger, Vogel, Lee, & Batchelder, 2008) which argue about the conventional GLUE methodology recommended the use of formal likelihood by this respect of the method. The formal likelihood definition is based on formal Bayesian approach and this always incorporates the explicit representation of model error term. The first research for this approach (Stedinger, Vogel, Lee, & Batchelder, 2008) assumes model error as normal, independent, homoscedastic. After this study, many researchers started to generalize those assumptions to a more general case. In Schoups & Vrugt (2010), they used skew exponential power distribution as a descriptor of non-normal property and used a linear relationship between error standard deviation and mean flow to consider heteroscedasticity. In Smith et al. (2010), zero-inflated, independent, heteroscedastic errors transformed by Box-Cox transformation was considered. As a reaction to formal likelihood approach, Beven (Beven, Smith, & Freer, 2008 ; Beven & Binley, 2014) also reputes these kinds of criticism by saying that in formal statistical analysis, modeling error is approximated by aleatory structural model, but in the real world the epistemic uncertainty exists, and the effect of this uncertainty cannot be omitted. The controversy about the objectivity of GLUE method is now ongoing and need to be studied more.

1.1.3 Necessity of study

After the first introduction of GLUE methodology, it was applied to diverse engineering research due to its ease of implementation. The subjectivity concerned in likelihood definition usually makes the user easily apply it to any real situation. However, due to its subjectivity, it also is argued by many researchers. In these studies, the likelihood was formally defined based on formal Bayesian framework. Afterwards, the hot debate about the appropriate likelihood definition keeps going on.

Considering the previous controversy, a comparative study of both approaches is needed. Uncertainty estimation based on both approaches should be compared with each other quantitively and also in a qualitative manner. This will clarify the difference between both approaches and their effects on the uncertainty analysis results. In addition to this, the more general framework for the description of error structure in formal likelihood definition is needed. Following previous studies, the more precise description of correlation, heteroscedasticity, non-normality should be tried. As Beven (Beven, Smith, & Freer, 2008 ; Beven & Binley, 2014) highlights the importance of epistemic uncertainty dissolved in model error, the procedure for reducing identifiable epsitemic error before statistical analysis of error structure is also on demand.

1.2 Objectives

The main objective is the comparative study of informal and formal likelihood definitions. In a way to do this, a new approach for generalization of

error structure will be applied. ARFIMA (0,d,0) will be used to identify correlation structure; and the stable distribution will be used as a descriptor of the non-normal attribute of error structure. For heteroscedasticity, the division of period to its subperiods will be performed for daily outcomes. And as the procedure for reduction of epistemic uncertainty dissolved in residuals before all these error structure analysis, linear regression method will be used.

For comparison, a real application to monthly and daily streamflow for Springcreek watershed in Pennsylvainia, USA will be carried out. SNU-WS will be used as a hydrologic simulation tool, and codes for the implementation of basic GLUE methodology into SNU-WS will be accomplished. MATLAB code will be combined in VB.NET framework for code development. This can be accomplished by using SDK complier in MATLAB.

Eventually, the uncertainty estimation given by both definitions will be compared with each other quantitively and in a qualitative manner. The comparison will be based on two aspects. The first one is the parameter uncertainty, and the second one is predictive uncertainty. Parameter uncertainty will be investigated through parameter posterior probability distribution. As formal likelihood approach explicitly defines model error term, model and parameter uncertainties can be differentiated. As opposed to this, an informal likelihood approach considers all kinds of uncertainties in a single subjectively chosen likelihood measure. So, parameter uncertainty turns to be higher in the informal likelihood approach. The predictive uncertainty will be investigated through drawing the uncertainty interval for a certain significance level.

Because formal likelihood approach accumulates model error term to the original simulation result, it will be more statistically reliable with respect to informal likelihood approach. Also, the facet of the new approach for generalization of error structure will be highlighted through analyzing their effect on uncertainty interval.

Chapter 2. Theoretical background

2.1 Modified Generalized Watershed Loading Function (MGWLF)

For mathematical modeling of the estimates of nonpoint sources of nitrogen and phosphorus in streamflow, export coefficient, chemical simulation models can be considered to be applied. The method of the export coefficient is based on the assumption that, for a given climatological regime, specific land use types will yield characteristic quantities of nutrients on the annual cycle (Winter & Duthie, 2000). So, it is of limited value for determining monthly loads or capturing seasonal variation. Compared to this, chemical simulation model deals with a mechanistic description of the transport and chemical behavior of nutrient. It is the most sophisticated description of nutrient loads, but they require too much data for use in practical purpose (Haith Douglas, Mandel Ross, Wu Ray Shyan, 1992).

GWLF compromises these two approaches. Mechanistic modeling such as water balance will be applied to the water system and sediment movement. Various chemical behaviors of the nutrients will be simplified through empirical formula or ignored like the case of the export coefficient. Through this compromise, it can be an appropriate model for application with acceptable accuracy.

The general structure of GWLF can be described through the streamflow nutrient flux. The nutrient in water can have two types of phases. One is the dissolved phase, and the other is a solid phase. In this model, each phase of the nutrient have their own contributors. It can be represented as follows.

$$LD_t = DP_t + DR_t + DG_t + DS_t$$

$$LS_t = SP_t + SR_t + SU_t$$

where LD denotes dissolved nutrient load, LS denotes solid-phase nutrient load, DP, DR, DG, DS are dissolved nutrient load from point sources, rural runoff, groundwater, septic system, SP, SR, SU are solid-phase nutrient load from point source, rural runoff, urban runoff on the day t. Through this classification, nutrient loads from each component will be treated separately. Before introducing procedures for calculation of these loads, the implementation of water balance principle will be figured out.

GWLF used in SNU-WS was modified to produce daily output (Nguyen & Nguyen, 2014). And the improved method was used in the procedure to estimate DR (Schneiderman, Pierson, Lounsbury, & Zion, 2002). Because it had been modified, the model used in SNU-WS will be called Modified Generalized Watershed Loading Function (MGWLF).

2.1.1 Water Balance

In the context of hydrological cycle, daily water balances for unsaturated zone and shallow saturated zone will be analyzed (HaithDouglas, MandelRoss,

WuRay shyan, 1992).

$$\mathbf{U_{t+1} = U_t + R_t + M_t - Q_t - E_t - PC_t} \quad \textbf{(cm)}$$

$$\mathbf{S_{t+1} = S_t + PC_t - G_t - D_t} \quad \textbf{(cm)}$$

U_t and S_t are the unsaturated and shallow saturated zone soil moistures at the beginning of day t, and R_t, M_t, Q_t, E_t, PC_t, G_t, D_t are rainfall, snowmelt, watershed runoff, evapotranspiration, percolation into the shallow saturated zone, groundwater discharge to the stream/channel, seepage flow to the deep saturated zone. Each term in this equation has significant meaning in the hydrological cycle. So, their theoretical background will be reviewed.

2.1.1.1 Runoff calculation

Runoff is computed from daily weather data by the U.S. Soil Conservation Service's Curve Number (SCS CN number) Equation (Chow, 1964).

$$Q_t = \frac{\left(R_t + M_t - 0.2DS_t\right)^2}{R_t + M_t + 0.8DS_t} \quad \textbf{(cm)}$$

The unit of all terms in here is centimeter. Rainfall and snowmelt can be given by daily precipitation and temperature data. When daily mean air temperature T_t (in celsius) is lower than 0, all precipitation will be regarded as snow, so it will be added to snow storage. Otherwise, it will become rainfall (R_t). Snowmelt will be computed when T_t is lower than 0.

$$M_t = 0.45 T_t \qquad T_t > 0 \qquad \textbf{(cm)}$$

The detention parameter DS_t is determined by curve number CN of the day t by the relation shown below.

$$DS_t = \frac{2540}{CN_t} - 25.4 \qquad \textbf{(cm)}$$

Here, curve number becomes the function of time because it can be affected by antecedent moisture content. It will be a function of 5-day antecedent precipitation (A_t) in this model. The exact functional form of the CN value according to A_t is given by the figure shown below.

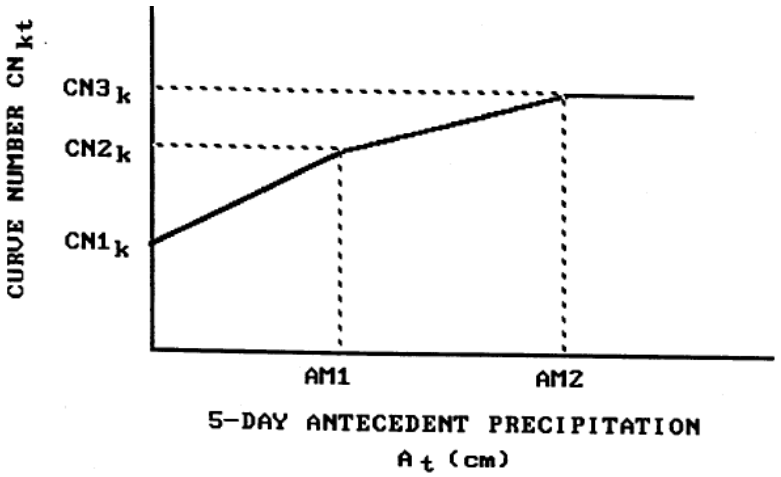

Figure 1. CN value as a function of A_t

Subscription k means corresponding land use for CN value. Recommended values for AM1 and AM2 are 1.3 cm, 2.8 cm for dormant seasons and 3.6 cm, 5.3 cm for growing seasons. CN1, CN2, CN3 denote the curve number in the state of driest, average, wettest moisture condition. CN2

should be given by the user of model, and CN1, CN3 will be computed through the following equation.

$$CN1 = \frac{CN2}{2.334 - 0.01334\ CN2} \qquad CN3 = \frac{CN2}{0.4036 + 0.0059\ CN2}$$

Recently, Lal Mohan et al. (2019) reverified the above equations (Lal , Mishra, & Kumar, 2019). They evaluated the performance of 5 existing equations with suggested 3 equations of their owns through utilizing the data from agricultural field located at Roorkee, Uttarakhand, India and available published data around the globe. In this model, we adopted most traditional one suggested as above. Because CN value is most renown effective parameter in hydrological modeling, its impact will be assessed by uncertainty analysis.

2.1.1.2 Evapotranspiration

The evapotranspiration will be computed through following equation.

$$E_t = Min\left(CV_t\ PE_t\ ,\ U_t + R_t + M_t - Q_t\right) \qquad (\mathbf{cm})$$

$$PE_t = \frac{0.021 H_t^2\ e_t}{T_t + 273} \qquad (\mathbf{cm})$$

for which CV_t is a cover coefficient, and PE_t is a potential evapotranspiration, H_t is the number of daylight hours per day during the month containing day t, e_t is the saturated water vapor pressure (mbar) on day t, and T_t is the temperature on day t by Celsius. When T_t is below 0 or the same with

it, PE_t will be set to 0. PEt equation above was given by (Hamon, 1960).

Saturated vapor pressure can be approximated as by the equation given below (Bosen, 1960).

$$e_t = 33.8639\left[\left(0.00738T_t + 0.8072\right)^8 - 0.000019\left(1.8T_t + 48\right) + 0.001316\right]$$
$$T_t \geq 0 \ (\text{mbar})$$

Important mechanism in here is that evapotranspiration should be limited by available moisture in the unsaturated zone.

2.1.1.3 Percolation

Percolation is the mechanism which the water in the unsaturated zone move to shallow saturated zone when its amount exceeds available soil water capacity U*.

So, it can be represented by the following expression.

$$PC_t = Max(0, U_t + R_t + M_t - Q_t - E_t - U^*) \qquad (\mathbf{cm})$$

It means that percolation will be calculated after runoff and evaporation are considered.

2.1.1.4 Groundwater discharge and deep seepage

In this model, the shallow unsaturated zone is modeled as a simple linear reservoir which can be characterized by the following expression where

S denotes the water content in the zone.

$$\frac{dS_t}{dt} = -rS_t$$

Here, r is the groundwater recession coefficient (day $^{-1}$). Then, Groundwater discharge and deep seepage are calculated by the relation shown below.

$$G_t = r\,S_t \qquad (\mathbf{cm})$$

$$D_t = s\,S_t \qquad (\mathbf{cm})$$

Seepage constant s (day $^{-1}$) is usually set to 0 by conservative assumption. When s becomes 0, it means that all precipitation exits the watershed in evapotranspiration or streamflow. Otherwise, it should be estimated through parameter calibration (HaithDouglas, MandelRoss, WuRay shyan, 1992). Groundwater recession coefficient has a value of usual range which is from 0 to 0.2.

2.1.1.5 Streamflow

As a result of the water balance equation, streamflow can be calculated. After summing up the runoff from all land uses, the groundwater discharge will be added on to it. Finally, tile drainage and water withdrawal terms will be also added, resulting in the streamflow in target watershed. Water withdrawal information will be given by the user, and tile drainage could be calculated

through the following equation.

$$T_t = T_t^{RO} + T_t^{GW} \qquad (\mathbf{cm})$$

$$T_t^{RO} = Q_t^{AG} \times TD \qquad T_t^{GW} = G_t \times \frac{AG}{AT} \times TD$$

where T_t denotes tile drainage, T_t^{RO} denotes tile drainage contributed by runoff process from agricultural land, T_t^{GW} denotes tile drainage contributed by groundwater discharge from agricultural land, Q_t^{AG} means runoff from agricultural land, AG means agricultural land use area, AT means total watershed area, and finally TD denotes tile drainage density which should be given by the user.

2.1.2 Dissolved nutrient load

Nutrient load of dissolved phase will be computed through water balance equation given above. Nutrient load from rural runoff, groundwater, a septic system will be investigated one by one.

2.1.2.1 Rural runoff load (DR)

Dissolved nutrient loads from rural runoff will be calculated by multiplying runoff by dissolved nutrient concentration and source area.

$$DR_t = 0.1 \sum_k Cd_k Q_{kt} A_k \qquad (\mathbf{Mg})$$

where Cd_k means nutrient concentration in runoff from source area k

(mg/l), Q_{kt} means runoff from source area (cm) in day t, and A_k denotes the area of source area k (ha). 0.1 in the equation is adopted due to unit conversion.

2.1.2.2 Groundwater load (DG)

Dissolved nutrient load from groundwater can be calculated through multiplying groundwater discharge by nutrient concentration in groundwater with the total watershed area.

$$DG_t = 0.1\, C_g \times G_t \times AT \qquad (\mathbf{Mg})$$

where C_g denotes the concentration of nutrient in groundwater (mg/l), G_t means groundwater discharge on day t, and AT represents total watershed area (ha). Like as rural runoff case, 0.1 is adopted again due to unit conversion. Different from rural runoff case, we don't need to consider different source area k in here. We use a lumped model for modeling groundwater, so we just need to use groundwater discharge from the total watershed area.

2.1.2.3 Septic system load (DS)

Septic system loads can be classified by four components.

$$DS_t = DS_{1t} + DS_{2t} + DS_{3t} + DS_{4t} \qquad (\mathbf{kg})$$

where DS_{1t}, DS_{2t}, DS_{3t}, DS_{4t} are the dissolved nutrient load from normal systems, short-circuited systems, ponded systems, and direct systems

discharge systems in day t.

1. Normal systems

When some septic system is constructed and operated by recommended procedures such as those suggested by the EPA design manual for on-site wastewater disposal systems, we call it as normal septic system. Effluents from such systems infiltrate into the soil and go down to the shallow saturated zone. Nitrogen in the effluent will be converted to nitrate and will pass through plant uptake process, then it will flow into the stream by groundwater discharge. On the contrary to this, phosphates in the effluent are adsorbed and remain in the soil. Hence there are no phosphorous loads from normal septic systems which flow into streamflow. The nitrogen load from a normal septic system can be determined through the equation given below.

$$DS_{1t} = 0.001 A_f A_s n_1 (e - u) \qquad \textit{Growing Season} \qquad (\mathbf{kg})$$

$$DS_{1t} = 0.001 A_f A_s n_1 e \qquad \textit{Dormant Season} \qquad (\mathbf{kg})$$

where DS_{1t} is nitrogen load from normal septic system in day t, e is daily nutrient load in septic tank effluent (g/day) , u is daily nutrient uptake by plants (g/day), and n_1 is the number of normal septic systems, A_f is the attenuation factor for subsurface flow loss, A_s is the attenuation factor for soil loss. u is considered only in the growing season.

2. Short circuited systems

On the contrary to normal septic systems, these systems are located close to a surface waterbody, which means that we cannot ignore the phosphorus load. The plant uptake could happen, so nitrogen load and phosphorus load will be calculated by the following equation.

$$DS_{2t} = 0.001 n_2 (e - u) \qquad Growing\ Season \qquad (\mathbf{kg})$$
$$DS_{2t} = 0.001 n_2 e \qquad Dormant\ Season \qquad (\mathbf{kg})$$

The form of the equation is similar to the case of normal systems. n_2 means the number of short-circuited systems.

3. Ponded systems

These systems result in surfacing of the effluent. In principle, when the daily mean temperature is below 0 or there exists snowpack, the effluent will freeze in a thin layer at the ground surface. Otherwise, it will be delivered to surface waters through overland flow in the corresponding month.

$$EF_t = 0.001 n_3 (e - u) \qquad Growing\ Season \qquad (\mathbf{kg})$$
$$EF_t = 0.001 n_3 e \qquad Dormant\ Season \qquad (\mathbf{kg})$$

$$FN_{t+1} = FN_t + EF_t \qquad SN_t > 0 \ or \ T_t \le 0 \qquad (\mathbf{kg})$$
$$FN_{t+1} = 0 \qquad otherwise \qquad (\mathbf{kg})$$

where EF_t denotes effluent discharge of the day t, FN_t means the amount of accumulated frozen effluent, SN_t is the amount of snowpack in the start of day t, and n_3 means the number of ponded

systems.

The accumulated effluent will melt when the two conditions are accomplished. The first one is that snowpack should be eliminated, and the second one is the daily mean temperature should be above 0. Using these conditions, we can get the nutrient loads from ponded systems in overland flow on the day t.

$$DS_{3t} = FN_t + EF_t \qquad SN_t = 0 \ and \ T_t > 0 \qquad (\mathbf{kg})$$

$$DS_{3t} = 0 \qquad otherwise \qquad (\mathbf{kg})$$

4. Direct discharge systems

The septic tank effluent from illegal systems will directly go into surface waters. Hence,

$$DS_{4m} = 0.001 n_4 \, e \qquad (\mathbf{kg})$$

where n_4 denotes the number of direct discharge systems.

2.1.3 Solid-phase nutrient load

Solid-phase loads in rural runoff and urban runoff will be considered in sequence. The runoff calculated by water balance equation will be used in this calculation.

2.1.3.1 Rural runoff load (SR)

Solid-phase rural nutrient loads can be estimated by the product of

daily watershed sediment yield Y_t (Mg) and average sediment nutrient concentration C_s (mg/kg).

$$SR_t = 0.001 C_s Y_t \qquad (\mathbf{Mg})$$

To calculate daily sediment yield, we need to utilize the universal soil loss equation (USLE). The procedure will be given as the same manner in Schneiderman et al. (2002). Schneiderman et al. (2002) modified the original GWLF model through revision of sediment yield timing. In the original model, an annual sediment yield was calculated for sediment year which starts from April and ends in March. And the timing of sediment release was assumed to be proportional to transport capacity of runoff. In these procedures, no carryover of sediment release from one year to next year was assumed. Due to this assumption, there was a discontinuity in sediment yield boundary which will not happen in a real situation. To solve this problem, they changed the yearly sediment yield equation to the one of long term average. This can be resolved by the following expression.

$$\overline{Y} = \overline{E} \times SDR \qquad (\mathbf{Mg})$$

in which, SDR is sediment delivery ratio, and it is empirically determined from annual export studies. $\overline{Y}$ (Mg), $\overline{E}$ (Mg) are a long term average of annual sediment yield, annual sediment erosion. $\overline{E}$ will be calculated by the following equation.

$$\overline{E} = \sum_{k} \left(\frac{\sum_{t=1}^{n} X_{kt}}{n} \times 365.25\, days / year \right) \qquad (\mathbf{Mg})$$

where X_{kt} expressed sediment erosion from source area k on the day t. At this point, we need to resolve USLE for computation of X_{kt}. USLE is given by the following expression (Wischmeier & Smith, 1978).

$$X_{kt} = 0.132\, RE_t\, K_k\, LS_k\, C_k\, P_k\, A_k \qquad (\mathbf{Mg})$$

where the unit of all terms except RE_t is U.S. customary units due to its first development by U.S. Department of Agriculture (USDA) (Foster, McCool, Renard, & Moldenhauer, 1981). The physical meaning of each term in this equation will be investigated in sequence.

The first term RE_t is the rainfall erosivity index. It represents the kinetic energy of rainfall which causes potential erosion risk. As expected, this term varies geographically. Richardson et al. (1983) suggested following formula to get this term (Richardson, Foster, & Wright, 1983).

$$RE_t = 64.6\, a_t\, R_t^{1.81} \qquad (\mathbf{MJ \bullet mm / ha \bullet h})$$

where coefficient a_t varies with season and geographical location. Because RE_t in this equation is in unit MJ-mm/ha-h, so unit conversion factor 0.132 was added in USLE. The suggested formula will be used in our application.

The second term K_k is called the soil erodibility factor in source area k. The soil erodibility is an inherent soil characteristic which resists transport caused by raindrops or overland flow. It is cohesiveness or bonding characteristic of soil type.

The third term LS_k reflects topographic effects of source area k. The Steeper and longer slope will reproduce higher overland flow velocities which will affect sediment erosion. This effect is considered in this term.

The fourth term C_k represents the effect of crop management of source area k. This is defined as the ratio of soil loss of land cropped under the specified condition to corresponding loss under tilled, continuous fallow condition.

The fifth term P_k considers some kinds of practice in source area k which usually applied in actual situation. This includes contouring, strip cropping and terracing.

The last term A_k means the area of source area k.

After getting $\overline{E}$ from USLE, we can associate it with the following basic empirical relationship.

$$Y_t = k\,TC_t, \qquad (\mathbf{Mg}) \qquad\qquad TC_t = Q_t^{1.67}$$

TC_t means the transport capacity of runoff.

Taking average in each side, the formula for k will be elucidated by equating the term $\overline{Y}$ with the one in the previous equation.

$$\overline{Y} = k\ \overline{TC} \qquad \overline{Y} = SDR \times \overline{E}$$

$$k\ \overline{TC} = SDR \times \overline{E}$$

$$k = \overline{E}\, SDR \frac{1}{\overline{TC}}$$

$$where \qquad \overline{TC} = \frac{\sum_{t=1}^{n} TC_t}{n} \times 365.25\, days / year$$

By substituting k in an original basic empirical relationship, we get the final equation describing daily sediment yields.

$$Y_t = \overline{E}\, SDR\, \frac{TC_t}{\overline{TC}} \qquad (\mathbf{Mg})$$

Through Y_t with C_s, we get SR_t by the first equation given in this section, which is solid-phase nutrient load in the rural runoff.

2.1.3.2 Urban runoff load (SU)

Nutrient loads from urban runoff are considered to be entirely in the solid phase. The mathematical derivation presented here is attributed to Overton & Meadows (1976) (Overton & Meadows, 1976). Nutrients accumulate on urban surfaces along time passes and are washed off by runoff mechanism. The accumulated nutrient load on source area k on day t in dry period can be modeled by following ordinary differential equation.

$$\frac{dN_k(t)}{dt} = n_k - \beta N_k(t)$$

where $N_k(t)$ is the accumulated nutrient load on source area k on day t (kg/ha), n_k is a constant accumulation rate (kg/ha-day), β is a depletion rate constant (day^{-1}). Solving above differential equation, we get the general solution of the form,

$$N_k(t) = N_k(0)e^{-\beta t} + \frac{n_k}{\beta}(1-e^{-\beta t}) \qquad (\mathbf{kg/ha})$$

This result implies that the nutrient loads will increase but not above certain asymptotic value. If we conservatively assume that $N_k(t)$ reaches 90% of asymptotic value in 20 days, then with the initial value of 0, we get the depletion rate constant.

$$0.9\frac{n_k}{\beta} = \frac{n_k}{\beta}\left(1-e^{-20\beta}\right) \qquad \rightarrow \qquad \beta = 0.12$$

By setting initial time as a day t, the $N_k(t+1)$ can be identified through the previous equation. Moreover, when the effect of wash off is considered, it can be represented as the following expression.

$$N_k(t+1) = N_k(t)e^{-0.12} + \frac{n_k}{0.12}(1-e^{-0.12}) - W_k(t) \qquad (\mathbf{kg/ha})$$

in which, $W_k(t)$ is runoff nutrient load from source area k on day t (kg/ha). Finally, this runoff load will be defined as,

$$W_k(t) = w_k(t)\left[N_k(t)\,e^{-0.12} + \frac{n_k}{0.12}(1-e^{-0.12})\right] \qquad (\mathbf{kg/ha})$$

$$w_k(t) = 1-e^{-1.81Q_k(t)}$$

$w_k(t)$ in here is called the first-order wash off function suggested by Amy (1974) (Amy, 1974).

Afterwards, solid-phase nutrient loads in urban runoff will be computed by multiplying $W_k(t)$ by each source area and summing all of it.

$$SU_t = 0.001 \sum_k W_k(t) A_k \qquad (\mathbf{Mg})$$

Eventually, dissolved nutrient loads of various types and solid-phase nutrient loads from each contributor are discussed and identified. From this theoretical background, SNU-WS uses MGWLF for calculation of daily nutrient loads in the watershed. In the next step, the theoretical background for GLUE will be investigated.

2.2 Generalized Likelihood Uncertainty Estimation (GLUE)

The original GLUE method was suggested by Beven & Binley (1992) (Beven & Binley, The future of distributed models: Model calibration and uncertainty prediction, 1992). Before starting the discussion of the GLUE methodology, the philosophical aspect of it should be introduced.

In the usual sense, calibration accompanies the search for the most optimal set of parameters whose outcomes are alingn with the observed one. The global optimum parameter set is found, and parameter and predictive uncertainty are assessed with respect to that global optimum. However, there have been many studies which illustrated the difficulties of finding such an optimal parameter

set in the high dimensional parameter space associated with hydrological models. Parameters in many hydrologic models are usually intercorrelated with each other, and residuals frequently possess autocorrelation and heteroscedascity. These attributes of parameters result in local minima, valleys and other complicated structure in parameter response surface, and it will make it hard to apply classical calibration method like trial and error, automatic search, random search. Moreover, the hydrologic model usually includes numerous hydrological processes. And these processes will be represented by a particular combination of parameters. So, the optimal parameter set should also be considered in the light of qualitative sense. In the philosophy of classical optimization, the optimal parameter set is defined as the one which produces outcomes matching the observed one in a quantitative sense. It means that when we determine optimized parameter set as a global optimum one, there can be the possibility of excluding sets of parameters that give a qualitatively more correct simulation of the response mechanism in the catchment. This is a difficulty of differentiating multiple optima in the response surface by many perspectives. In addition, it might also exclude parameter sets far from the global optimum one, while giving closer results with respect to a different period of observations.

Due to a number of drawbacks of previous calibration philosophy, Beven & Binley (1992) asserts that there can be many parameter sets which can be identified as a good simulator of the system. As noted earlier, due to the error in model structure, observation, measurement, it is undesirable to think that only one true parameter set exists in a given model structure. Rather, it is natural

to set the possibility of parameter set being a good simulator of the system by assigning likelihood values for each of them. This concept is called equifinality. The original GLUE method is based on this philosophy.

It consists of 4 major steps.

1. Sampling parameters based on an appropriate definition of the prior distribution
2. Definition of likelihood measure
3. Estimation of parameter uncertainty and predictive uncertainty using calculated likelihood
4. Model validation and prediction within the limits of uncertainty

Uncertainty intervals given by these procedures are dependent on the model's performance. Because given uncertainty intervals are entirely dependent on model simulation results, poorer model performance implies poorer coverage ratio of observation values in those uncertainty intervals. Beven & Binley (2014) also states that "… There are, however, applications where it is clear that the range of models tried cannot match particular observations in either calibration or validation. This could be because of model structural error, or, as noted earlier, it could be because of epistemic error in the inputs." In this respect, it is more desirable if we can figure out epistemic uncertainty purely introduced by our model structural error and reduce it. Therefore, in this study, statistical treatment was performed to accomplish this after parameter sampling procedures. After the first introduction of GLUE, there was a huge controversy about its usefulness due to the subjectivity involved in the selection of likelihood. Following Stedinger et al. (2008),

there were a number of researches about formal likelihood definitions and its outcomes (Smith, Sharma, Marshall, Mehrotra, & Sisson, 2010; Schoups & Vrugt, 2010). Formal likelihood definition is determined by a formal statistical approach based on Bayes theorem. When we follow this approach, the identification of the error structure should be performed, and many assumptions are adopted in this step. Usually, it incorporates the assumptions about correlation, heteroscedasticity, nonnormality. In this study, ARFIMA(0,d,0) model was used to model the correlation structure of error, and stable distribution was considered to consider nonnormal error structure. Also, for the daily flow, heteroscedasticity was considered as dividing the period into several subdivisions. Every step of GLUE methodology with the added procedures introduced above will be covered one by one in the following sections.

2.2.1 Parameter sampling

There are numerous parameters in the typical hydrologic model. First of all, the parameters of interest should be designated. Then, appropriate parameter ranges will be determined, and a prior distribution of parameter set will be assumed based on the model structure. Of course, there may be subjectivity at this step. Without useful information about parameters, uniform distribution is usually chosen. Because it can be updated after evaluating likelihood values of it, this assumption is not critical at all. After setting prior distribution, the parameter value will be randomly sampled. The drawback of typical GLUE application is a considerable computational burden due to the random sampling technique. To improve this aspect, practitioners of the GLUE

methodology generally adopt a Latin hypercube sampling strategy (LHS). A Latin square is a square grid which has the only one sample in each row and column, and Latin hypercube is a generalization of this concept in a higher dimension. Compared to random sampling, LHS can economize sampling times and produce greater computing efficiency. In this study, LHS was adopted to capture the general state of the response function of parameter space in an efficient way.

2.2.2 Identification of an epistemic error

Epistemic uncertainty is usually introduced by an imperfect model description of physical processes in the real world. According to its level, it can incorporate overestimation or underestimation of the model outputs compared to real observations. To figure it out, we can apply statistical analysis such as regression analysis. For example, the daily streamflow variable in previous MGWLF model can be considered. As covered by earlier sections about MGWLF, the daily streamflow is composed of runoff, groundwater flow and other minor components. As will be clear in section 3.3.2, inexact modeling of runoff calculation usually produce large peaks which deviate from the observed value. Also, the inexact modeling of baseflow produces constant error along a low flow period. By comparing each simulation results given by sampled parameter set with observation values, the model result which gives the smallest squared error sum could be chosen. This simulation result will be considered as the one which gives the smallest epistemic error among the other simulation

results given by numerous sampled parameter set. So, it will be assumed that the other simulation results include at least the level of epistemic error of this selected simulation result. It means that by considering the epistemic error given by selected simulation result in every other simulation results, we can get new corrected simulations whose epistemic error or epistemic uncertainty have been reduced apparently. For the example case, we can find an empirical relation between error specified by the selected simulation result and runoff, groundwater flow component through regression analysis. Through applying this empirical relation equally to simulation results given by other sampled parameter set, we can figure out the epistemic error and subtract it from each error to get the simulation results corrected for epistemic error. As noted earlier, these corrected results will give us better uncertainty intervals in both formal and informal likelihood cases.

2.2.3 Likelihood measure

2.2.3.1 Informal likelihood

The informal likelihood measure is the likelihood measure originally suggested by Beven & Binley (1992). It can be defined as a measure of how much model predictions coincide with observation values. It should increase monotonically as the similarity increases. Also, the threshold value can be adopted to set the likelihood measure as 0 under some designated value. This measure can be regarded as a fuzzy measure. The fuzzy measure represents the degree certain element belongs to a specific set. Many goodness of fit indices

used frequently could be used for definition in this procedure. Typical likelihood measures are given below.

$$L = 1 - \frac{\sigma_e^2}{\sigma_o^2}$$

where σ_e^2 is the variance of residual, and σ_o^2 is the variance of observations. This is the Nash–Sutcliffe model efficiency coefficient (NSCE) which is frequently used in the hydrologic application. In usual hydrologic time series, as time scale goes down, sudden enormous peak values appear frequently. If the model cannot capture the physics behind it, the variance of residuals will be influenced enormously. To mitigate this effect, the new likelihood definition based on L-moment can be developed. It will be given as follows.

$$L = 1 - \frac{\lambda_{2,\varepsilon}}{\lambda_{2,o}}$$

where $\lambda_{2,\varepsilon}$ is the L-scale of residual, and $\lambda_{2,o}$ is the L-scale of observed values. Another example is the inverse error variance which is suggested by Beven & Binley (1992).

$$L = \frac{1}{\left(\frac{\sigma_e^2}{\sigma_o^2}\right)^N}$$

where σ_e^2 is the variance of residual, σ_o^2 is the variance of observations, and N is the shape factor determined by the user. Here, the variance of residual

was nondimensionalized through dividing it by variance of observations. When $N = 0$, each simulation will have an equal likelihood. When $N \to \infty$, the single best simulation will have a renormalized likelihood value of 1 with other simulation having values of 0.

In the most hydrologic model, there are numerous types of outputs including discharge, sediment yield, nutrient loads etc. And there can be observation from multiple sites. In this case, we should combine the individual likelihood measures to make the overall likelihood measure. There can be some ways of doing this. One is to make a weighted sum of individuals. In the case of likelihood measures suggested by Beven & Binley (1992) defined above, the combined result can be written by the one below.

$$L_{combined} = \left(\sum_{j=1}^{m} \frac{W_j}{\sigma_{ej}^2} \right)^N$$

where m is the number of observed results, W_j is the corresponding weight, σ_{ej}^2 is the variance of residuals of the j-th observed one, and again N is the parameter given by the user. Ruark et al.(2011) used weights based on the first order sensitivity of the output to assess parameter uncertainty in sediment transport modeling applications (Ruark, Niemann, Greimann, & Arabi, 2011). They performed global sensitivity analysis through fourier amplitude sensitivity testing (FAST), and used the result from it to make a combined likelihood measure for 3 distinguished response variables. Alternatively, the individual likelihood measure could be regarded as a fuzzy measure, so we can

apply a number of generalized fuzzy set operations. Two possible suggestions are set union and set intersection in fuzzy set theory. It can be written by the expression shown below.

$$\bigcup_{1\le j\le m} L_j = \max\left[L_1, L_2, \ldots L_m\right] \qquad \bigcap_{1\le j\le m} L_j = \min\left[L_1, L_2, \ldots L_m\right]$$

where m is the number of observed results. Because there are multiple outputs in our model (discharge, nutrient loads etc), the illustrated method can be applied.

Subjectivity associated with the choice of appropriate informal likelihood measure includes the consideration of purposes of modeling. According to the goal of modeling, the emphasis in simulation results and observation values can be different, which affects the definition of likelihood measure. One more thing which should be underlined is that the likelihood measure is allotted to a set of parameters. Due to this fact, intercorrelation between parameters could be resolved implicitly in the process of uncertainty estimation.

When we get this informal likelihood measure, we can find posterior distribution through multiplying prior distribution by calculated likelihood measure. This value will be the final likelihood measure which will be used in section 2.2.4, 2.2.5.

2.2.3.2 Formal likelihood

Previous informal likelihood definition has a lot of subjectivity in the choice of likelihood definition. It assumes that this subjective choice of likelihood measure will handle implicitly all the uncertain issues. However, Stedinger et al. (2008) argues that " How is it possible that a simple subjective likelihood measure can understand and represent all these issues (model nonlinearity, errors in model structure etc.)?" Moreover, it ignores the model error and does not consider it when making uncertainty intervals. So, we cannot expect that those uncertainty intervals cover observations with a specified significance level. This apparent drawback of informal likelihood definition makes the researchers investigate about formal likelihood definition. Among these researches, Stedinger et al. (2008) could be considered as a landmark because it gives us insight, which shows how much informal likelihood and formal likelihood definition results differently even at a simple linear regession case. Although it had been generalized more, this paper is the main reference for the formal likelihood method adopted in this study. The theoretical background of formal likelihood method of this study will be introduced step by step now.

The first starting point of formal likelihood method is Bayes theorem which can be represented as follows.

$$P(\theta_i | Q_{Obs}) \propto P(Q_{Obs} | \theta_i) P(\theta_i)$$

where θ_i is the ith parameter set, Q_{Obs} is observed quantities, $P(\theta_i)$ is

prior pdf of θ_i, $P(\theta_i|Q_{Obs})$ is posterior pdf of θ_i, and $P(Q_{Obs}|\theta_i)$ is likelihood function in the context of Bayesian approach. Here, $P(\theta_i)$ corresponds to prior distribution of parameter set given in section 2.2.1, and $P(\theta_i|Q_{Obs})$ corresponds to the final likelihood measure which will be used in section 2.2.3, 2.2.4, 2.2.5. To get this value, the only task is to calculate $P(Q_{Obs}|\theta_i)$.

The model can be conceptualized as the following expression.

$$Q = \hat{Q}(\theta_i) + \varepsilon$$

where Q denotes true values, $\hat{Q}(\theta_i)$ means model simulation results derived by θ_i, and ε denotes a model error. Now, if we define pdf of ε, as $P(\varepsilon)$, this becomes $P(Q - \hat{Q}(\theta_i))$, and if we substitute Q by observed values, this will be interpreted as likelihood function given in Bayes theorem. So, to find $P(Q_{Obs}|\theta_i)$, we just need to analyze the error structure. Stedinger et al. (2008) considered the simplest case whose error structure is independent, normal and homoscedastic. Afterward, many researchers extend this condition to more general case whose error structure is correlated, nonnormal, heteroscedastic (Smith, Sharma, Marshall, Mehrotra, & Sisson, 2010; Schoups & Vrugt, 2010). In this study, different methods from other reseaches were

adopted for this generalization.

First of all, correlation embedded in the error structure was modelled by ARFIMA model. The previous researches usually adopted the ARMA model for consideration of correlation structure. However, in terms of correlation, it is very well- known fact that hydrologic time series contains the property of long range dependence. Hurst phenomenon strongly supports this feature as the usual hurst exponent of hydrologic series is higher than 0.5. The hurst exponent (H) is a measure of the extent of long-range dependence in a time series. Its value is between 0 and 1. A value of 0.5 means the absence of long-range dependence. The value lower than 0.5 denotes anti-persistency, which indicates strong negative correlation signifying largely fluctuating process. On the other hand, the value higher than 0.5 implies long-range dependence. The closer the value to 1, the larger the degree of persistence which corresponds to long-range dependence. In time series modeling, the long range dependence could be treated using ARFIMA model. It extends the integration order of the ARIMA model from integer to real number. The simplest form of the model, ARFIMA(0,d,0), can be represented by the following expression.

$$(1-B)^d X_t = \varepsilon_t \qquad d = H - \frac{1}{2}$$

Theoretically, $(1-B)^d$ could be expanded by following way.

$$(1-B)^d X_t = \sum_{m=0}^{\infty} (-1)^m \frac{\Gamma(d+1)}{\Gamma(m+1)\Gamma(d-m+1)} X_{t-m}$$

By the way, this expression cannot be coded completely because it includes infinity. Therefore, the finite expression will be coded in a real application. This is given in the following manner.

$$(1-B)^d X_t = \sum_{m=0}^{l} (-1)^m \frac{\Gamma(d+1)}{\Gamma(m+1)\Gamma(d-m+1)} X_{t-m} \quad , \quad t-l=c$$

where c is the first time step in time series. This definition corresponds to the definition of Riemann-Liouville fractional differentiation. By generalizing differentiation order in usual calculus to a real number, it defines derivative as

$${}_c D_x^{\alpha} f(x) = \frac{1}{\Gamma(n-\alpha)} \left(\frac{d}{dt}\right)^n \int_c^x \frac{f(y)}{(x-y)^{\alpha-n+1}} dy \qquad (n-1) \le \alpha < n$$

c can be set as any value, and this is the starting point of integration in the definition. Theoretically, the meaning of this definition is equivalent to the meaning of the definition used in a real application (Oldham & Spanier, 1974). Also, c in both of definitions has the same meaning of starting point. Therefore, we can notice that even if we change a definition of $(1-B)^d$ as introduced, it is in accordance with theoretical meaning. Furthermore, for real application, it was enough to use the simplest model, ARFIMA(0,d,0). It means that p,q in ARFIMA(p,d,q) could be ignored. Section 3.3.4 shows this fact clearly. Finally, we will adopt this method for real application.

Secondly, after identifying appropriate difference order in ARFIMA model, the stable distribution was introduced for an explanation of nonnormality of

errors. The stable distribution can be defined by the following the statement.

Let us consider X_1 and X_2 be independent copies of random variable X, if for any constants a > 0 and b > 0 the distribution of random variable aX_1+bX_2 is same with the one of cX+d of some constants c > 0 and d, the random variable X is called stable distribution.

The normal distribution surely satisfies this property, so we can say the stable distirubution is the generalization of normal distribution. In this respect, the central limit theorem can also be generalized to embrace stable distribution as the asymptotic result distribution in the theorem. It asserts the following argument.

The sum of a number of random variables with symmetric distributions having power-law tails, decreasing as $|x|^{-\alpha-1}$ where $0<\alpha\leq 2$, tend to a stable distribution $f(x;\alpha,0,c,0)$ as the number of variables grows.

The power-law tails are also called heavy tails because it gives us a higher portion in the tail than exponentail decrease case. This kind of phenomena frequently appears in natural phenomena. Usually, natural phenomena have a nonlinear attribute, which shows chaotic characteristic. And this makes us hard to predict future observation because it makes large epistemic uncertainty. This exotic behavior of nature produces a heavy-tailed distribution of error. In the context of climate change, this trend will be amplified and cannot be ignored for uncertainty analysis. Taking into account this aspect of nature, it is more desirable if we can adopt heavy-tailed distribution when modeling error

structure. Since the asymptotic distribution of linear combination of random variables of heavy-tailed distribtuion is stable distribution, using stable distribution as a model for general error distribution is a very good choice. Janicki & Weron (1993) also supports this idea by showing that stable stochastic processes are appropriate for simulating chaotic behavior which is prevalent in nature. The stable distribution does not have an explicit representation of its form in pdf, but has explicit form by characteristic function. It is four parameteric distribution given below.

$$\varphi(t;\alpha,\beta,c,\mu)=\exp(it\mu-|ct|^{\alpha}(1-i\beta\,\mathrm{sgn}(t)\Phi))$$

$$\Phi=\tan(\frac{\pi\alpha}{2})\ \ for\ \alpha\neq 1 \qquad \Phi=-\frac{2}{\pi}\log|t|\ \ for\ \alpha=1$$

where α gives us the information about the level of the heaviness of its tail, β is in charge of asymmetry of distribution, c shows us a measure of the width of the distribution, and μ is shift parameter.

Except in the case of the normal distribution, stable distribution does not have finite variance. Even when $\alpha\leq 1$, it does not have a finite mean. For this reason, in a real application, truncated stable distribution (TSD) will be used. This can be defined by setting maximum and minimum values of a stable random variable.

Lastly, the heteroscedasticity was considered by dividing the whole period by

its subperiod. Previous research such as Schoups & Vrugt (2010) uses the assumed linear relation between error standard deviation and mean flow to account for heteroscedasticity. However, in this research, the epistemic error of snowmelt process in daily outcome was too large in its amount and time of occurrence, so heteroscedasticity cannot be treated using a linear relationship determined by mean flow. Since snowmelt process usually has its own characteristic time period (It does not occur all the days in 1 year), it is reasonable to divide 1 year to 12 months and analyze error distribution of each month seperately to consider snowmelt process.

Through these three innovations, the error structure could be identified, and $P(Q_{Obs} - \hat{Q}(\theta_i))$ could be calculated by the following manner.

$$P(Q_{Obs} - \hat{Q}(\theta_i)) = \prod_{t=1}^{N} S_{m(t)}(\nu_t) \qquad N : Total\ time\ step$$

$$\nu_t = (1-B)^d P(Q_{Obs,t} - \hat{Q}_t(\theta_i))$$

$$S_j(\nu_t) : TSD\ of\ j^{th}\ month \quad m(t) : Month\ of\ time\ step\ t$$

Eventually, the final formal likelihood value could be calculated by Bayes theorem using this result. These procedures will be presented in section 3.3.4 in a more detail by real application to Springcreek watershed.

2.2.4 Parameter posterior probability distribution

After finishing the calculation of likelihood measure through definition given in the previous section, parameter posterior probability distribution could be found. This is accomplished through regarding each likelihood measure as a

weight for each parameter set. Because likelihood measure given in the previous step could not satisfy the normalization condition, which means the sum of likelihood measure may not be 1. To satisfy condition we can divide each likelihood measure by a sum of total measure.

After this step, we can use likelihood measure as a weight for each parameter set and get the corresponding empirical pdf of the parameter. For each parameter of interest, this procedure could be performed. Especially for informal likelihood case, the response function of parameter set space usually has a very irregular and fractal property. However, our goal is not to see the exact form of response function but to have an insight into the general tendency or trend. It is based on the fact that there is no certain response function which has absolute consequence. We can compare response functions one another as changing their informal likelihood definition. This gives us insight that it is more important to see the general trend according to changing likelihood definition than investigating each response function delicately. Accepting the above argument, the kernel function will be used to see the general tendency or trend when producing parameter posterior probability distribution. As other researches did, empirical cumulative distribution also could be generated using the calculated likelihood measure.

2.2.5 Uncertainty interval

Now, our interest moves onto model outcome which denotes variable of interest. At this time, the calculated likelihood measure can be considered as a weight for each simulation result. When this simulation result is given as time

series, we can fix the time step and get empirical cumulative distribution by the same procedure given in the previous section. Through estimating quantiles of a certain significance level for each time step, the uncertainty interval for corresponding significance level could be calculated. Finally, we can investigate if each observation value resides in this interval or not.

When we make the interval, we need one more step for a formal likelihood case. As explained in the previous section, formal likelihood method is based on the model description including the error term. Therefore, when we make the uncertainty interval, we should consider the error term. So, after adding a certain number of sampled error terms to each simulation results, we use these renewed values to get the uncertainty intervals. Likelihood measure will be assigned to new values according to the corresponding parameter set.

To illustrate this situation, the concept of the confidence interval and the prediction interval can be thought. Usually, confidence interval informs us about parametric uncertainty meanwhile prediction interval indicates both of model uncertainty and parameter uncertainty. The confidence interval only captures average values when it is assumed that the average of model error is zero. Compared to this, the prediction interval captures observation values. Likewise, when we draw uncertainty interval using only simulation results, it corresponds to the confidence interval. When we draw it by considering error terms, it will correspond to the prediction interval. Our goal is to cover the real observation values, so it is more reasonable to use the interval corresponding to the prediction interval.

2.2.6 Model validation and prediction

Uncertainty interval given in the previous step was applied in only calibration period because we got likelihood measure from the period where we have observation values. However, we can use likelihood measure calculated from the calibration period in a period where we do not have observation values and draw uncertainty interval through the way suggested in the previous section. After importing real observed values in this period, we can check if this interval cover observation value well or not. Then, this period can be called as validation period. If the validation succeeds, we can repeat the previous procedure to get another uncertainty interval in the future period for prediction. This could be called a prediction period. In this study, the results from calibration and validation periods were compared with each other. This will be covered in section 4. Through this analysis, we can confirm if our method for definition of likelihood was successful or not.

Chapter 3. Methodology

3.1 SNU-WS

Before introducing the specific methodology of GLUE application, the general structure of SNU-WS should be investigated. The operation of SNU-WS can be divided into two major parts, namely Pre-processing and MGWLF component. Pre-processing part is the procedure for preparing input data for main MGWLF model. It will be completed with the GIS function. After that, the output will be put into MGWLF part, resulting in the final output. The uncertainty analysis step will be added in the process of the main MGWLF model.

The pre-processing part will be done with Mapwindow GIS which is an open source. It should be noted that SNU-WS is built in the .NET framework and is written by visual basic.NET language. It is released by Microsoft in 2002 as a successor to the original Visual Basic computer programming language (Frenze, 2002). In the next part, the pre-processing part will be firstly figured out, and execution of main MGWLF will be introduced by the second step. All the procedures will be introduced with input layers for Springcreek watershed in Pennsylvania, USA. The data was downloaded from Mapshed Demo data (http://www.mapshed.psu.edu/download.htm).

3.1.1 Pre-processing

Pre-processing procedure is conducted with MapWindow GIS by building a plug-in merged into Mapwindow. Necessary information for MGWLF will be generated through performing clipping approach based on attributes of GIS layers. As a first step, data will be uploaded after setting plug-in. Figure 2 shows an interactive data upload window.

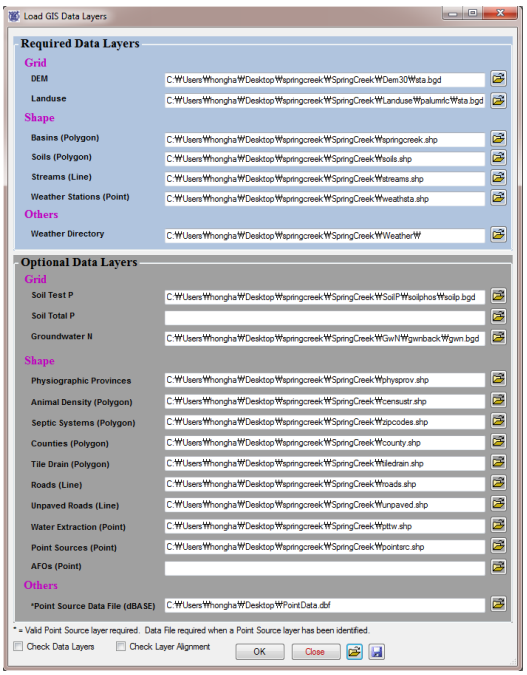

Figure 2. GIS data layers uploaded for pre-processing

Data layers in this step can be identified as two types, one is required data layers, and the other is optional data layers. The required data layers consist of geographic data, weather data, soil data, land use data and the optional data layers consist of point sources data, groundwater nitrogen data, soil phosphorus data, water extraction data, tile drains data, unpaved roads data and roads data,

county boundaries data, septic systems data, animal density data, physiographic province data. The required data layers represent essential information for application of MGWLF model. Geographic data layers include Digital Elevation Model (DEM) which is the grid file as a 3D representation of a terrain's surface, basins, streams. Those determine geometric features of the watershed and weather data will determine the input value of temperature, precipitation as input for MGWLF model. Land use and soil data will be used to get parameters like CN value as a representative example. The optional data layers include the information of concentration of nitrogen and phosphorus, point sources, septic systems and some other supplementary information such as tile drains, animal density etc. Figure 3 shows the result after importing input layers.

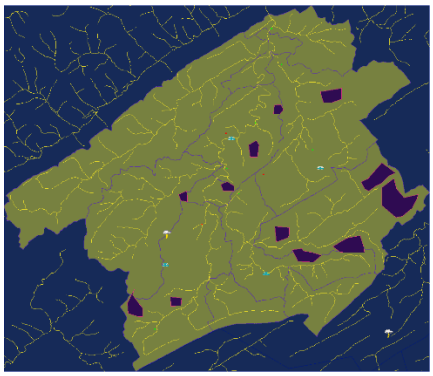

Figure 3. Uploaded layers

By the next step, subbasins of interest will be selected by GIS function in basin layer and input data for MGWLF will be generated for corresponding subbasins. During this step, there is an option to choose if the selected basins are aggregated or not. If aggregation is chosen, selected basins will be considered as a cohesive whole.

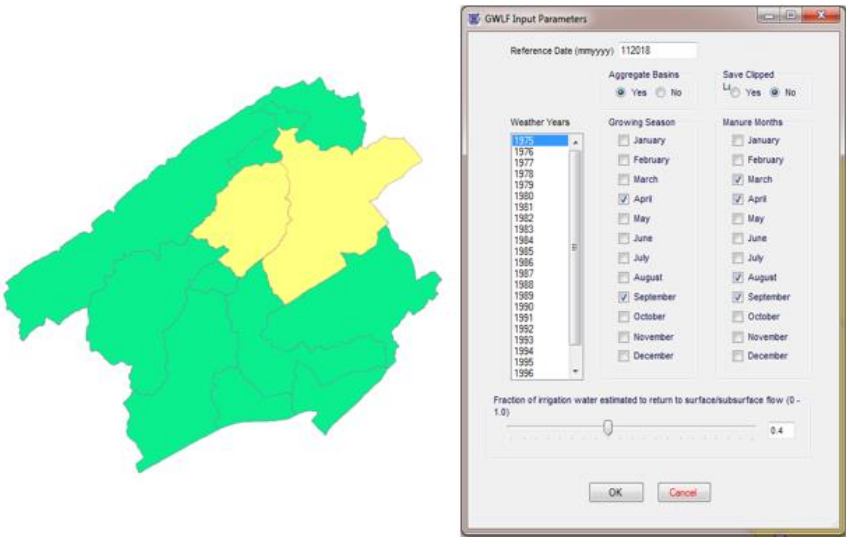

Figure 4. Creating SNU-WS input

Through these procedures, transport.dat, nutrient.dat, weather.dat files are generated. Transport.data file has information about transport properties with corresponding parameters such as CN numbers for each source area, USLE factors for each source area, recession coefficients etc. Nutrient.dat file has information of input for nutrient loads calculation such as nitrogen and phosphorus concentration in groundwater. Weather.dat file has information about weather stations with their data such as precipitation and Min/Max temperature. In the following section, the acquisition of results from MGWLF model using these input files will be discussed.

3.1.2 MGWLF component

Input files prepared by the previous step will be entered into MGWLF model as a second major procedure. We can also manually edit the input files. Figure 5 shows the interface for this procedure.

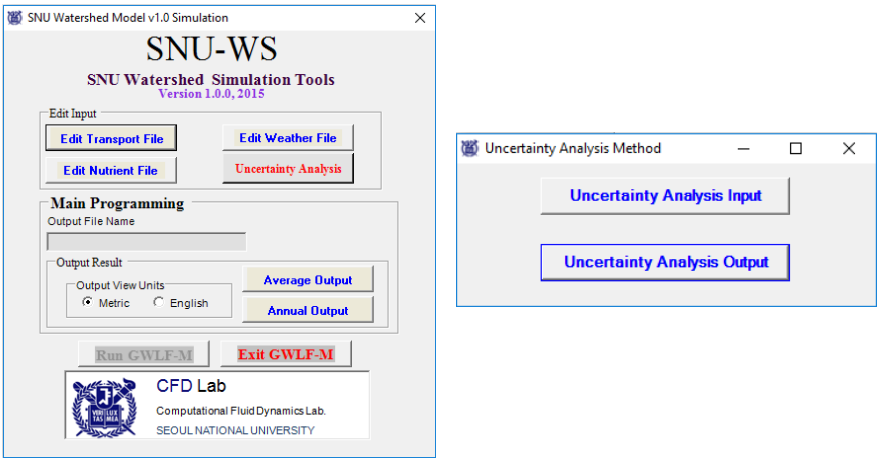

Figure 5. Interface for MGWLF component operation

After writing the output file name, the button for Run GWLF-M will be activated. Through clicking it, prepared input transport file should be selected, then the analysis will begin. Afterward, we can see output results as two types, average output and annual output. Figure 6 shows two types of outputs.

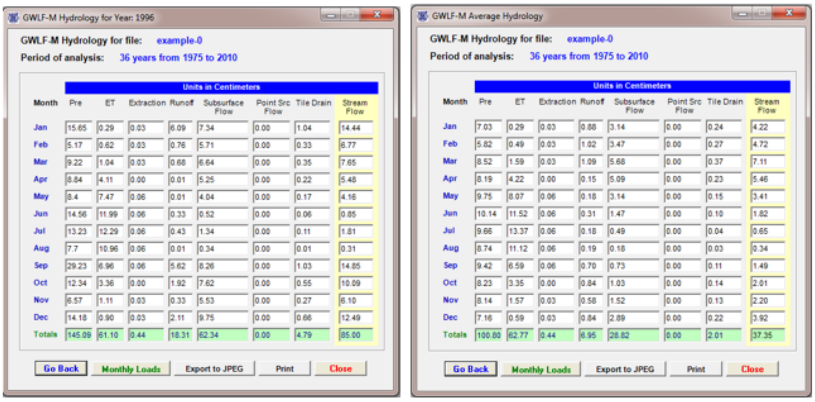

Figure 6. Average output (left) and annual output for 1996 (right)

The results can be analyzed by the user by comparing them with real observation data. Also, they can change the parameter values in the editing interface to see the most appropriate result. In the next section, the strategy for the development of uncertainty analysis based on GLUE methodology will be

excavated.

3.2 Uncertainty Analysis tool

3.2.1 Development of uncertainty analysis input interface

As noted in the earlier section, uncertainty input interface was designed in the second part, namely MGWLF part. When uncertainty analysis input button is clicked, the interface in figure 7 will be shown.

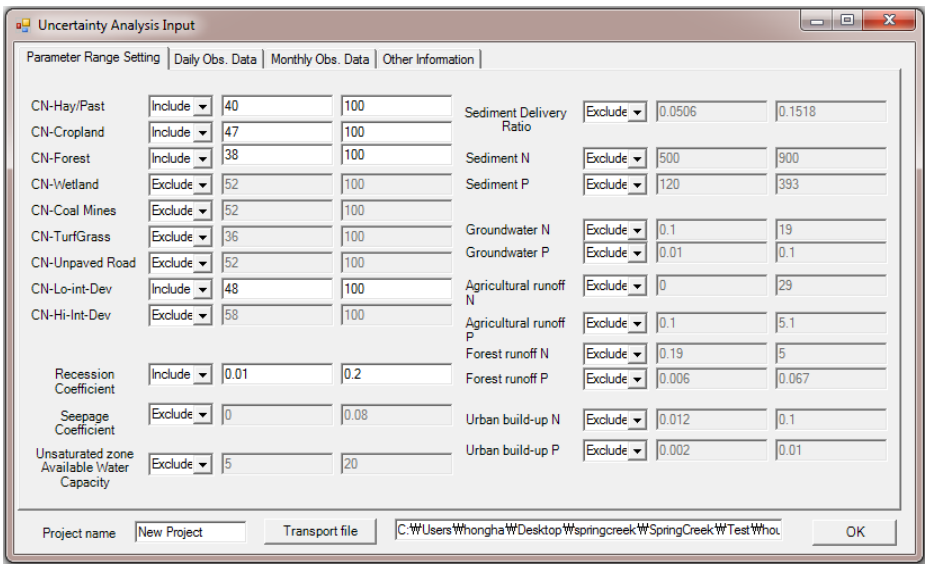

Figure 7. Uncertainty analysis input interface

The first tab of this interface is the parameter range setting page. There are a number of parameters which are engaged in MGWLF model. However, previous research suggested parameters which would be considered as effective parameters for output variable (Li, Weller, & Jordan, 2010). Moreover, it presents an appropriate parameter range for each parameter. In this study, these values were set as default values for parameter range except CN number,

sediment delivery ratio. For CN number, imported input CN number was used to get upper and lower bounds. The lower bound was set as a larger value between 30 and imported value – 35. The upper bound was set as a smaller value between 100 and imported value + 35. Likewise, in the case of sediment delivery ratio, half of the imported value was subtracted and added from imported value, and these values were set as lower bound and upper bound. The parameters in the left column have an effect on streamflow, and parameters in the right column only affect water quality output (nitrogen concentration, phosphorous concentration). By selecting "include" item in combobox for each parameter, we can set the parameters of interest.

For the next step, the observed values should be imported to do uncertainty analysis. The variables of interest could be classified as daily output and monthly output. Two interfaces for importing observation values are given in the second and the third tabs. Those are presented in figure 8. First of all, the user can select the variables of interest and also select periods of interest. Then, according to its selection, we can import the observation data file for each variable in each period by clicking the select file button. Alongside this, the time period for periods of interest should be defined in this step. The default value for these periods will be set as maximum and minimum date or month defined by transport.dat file.

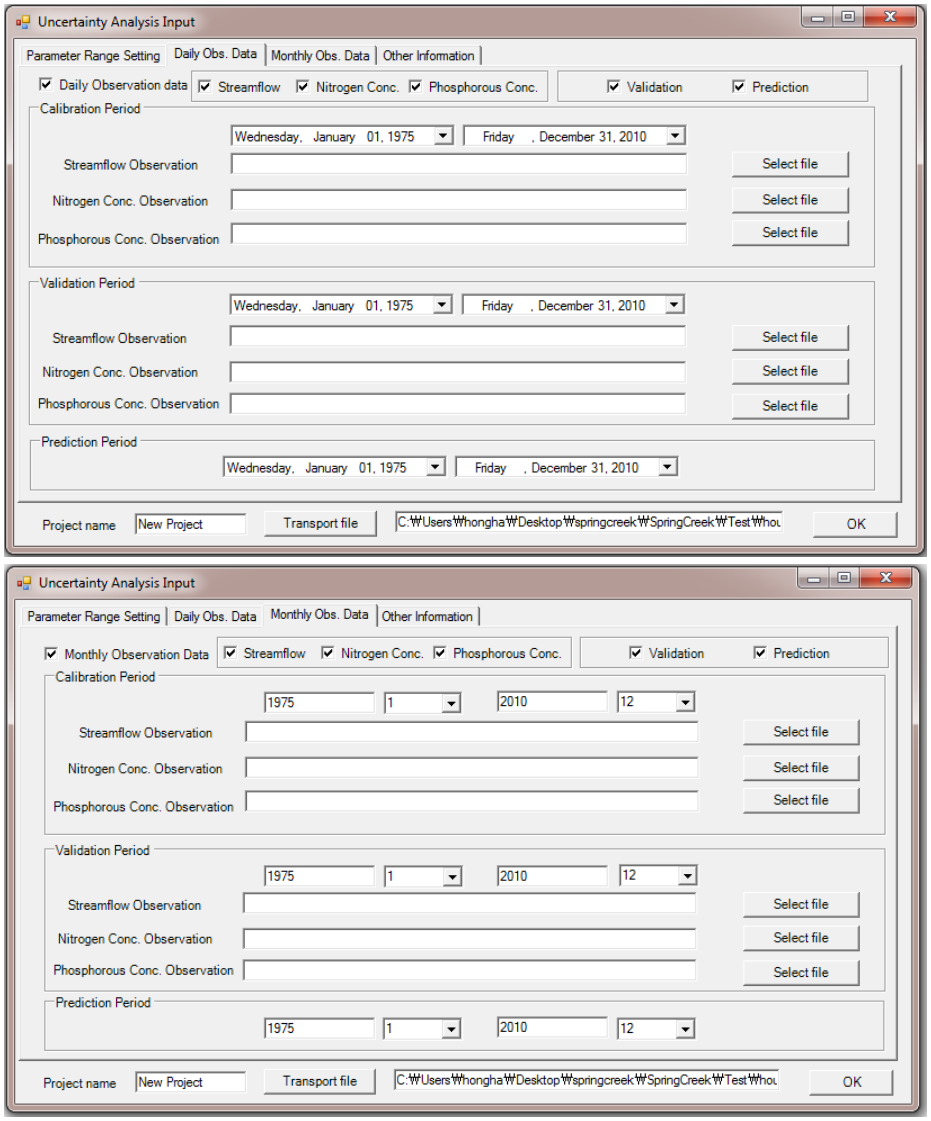

Figure 8. Daily observation data (upper), Monthly observation data (lower) import interface

At the final tab page, other important information should be defined by the user. Likelihood measure definition must be defined in this step. For the ease of application, only informal likelihood definition was implemented by code. Also associated with this, likelihood combination method should be determined if multiple variables of interest were chosen. There are two options for sampling

methods, which are random sampling and LHS. Finally, the simulation number should be given, and the significance level for uncertainty interval should be defined. The default value of the simulation number and significance level is 1000 and 2.5%, 97.5%. After defining this required information, the OK button in right-downside could be clicked to perform uncertainty analysis.

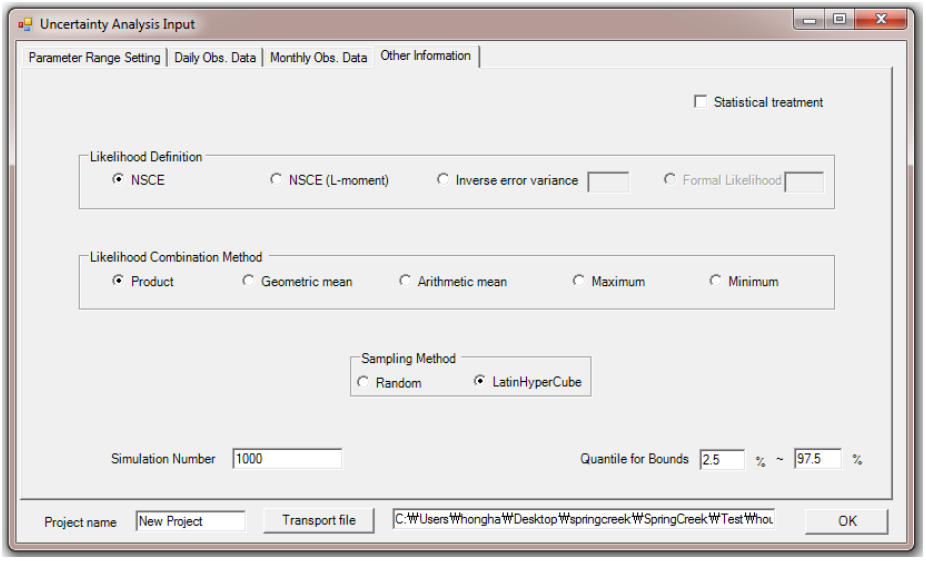

Figure 9. Other information interface

3.2.2 Development of uncertainty analysis output interface

After finishing the uncertainty analysis, the results will be saved as a form of a text file. The output files for the daily and monthly variable will be collected separately, and the name of the text file will contain the information about likelihood definition, variable of interest, project name, etc. In the case of uncertainty interval file, the corresponding time period is written in the first line. In the case of parameter empirical cumulative distribution function file,

simulation number, the information about parameters of interest and each of their names are recorded. When uncertainty analysis output button in figure 5 is clicked, we can see the graph of result information by choosing the appropriate file. After clicking the button, two options will be given (Figure 10).

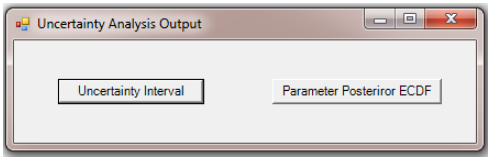

Figure 10. Uncertainty analysis output interface

One is the option for uncertainty interval, and the other is for parameter posterior ECDF. If the uncertainty interval button is clicked, we can select the uncertainty interval result file and can see the result by the graph in matlab function. In the same way, if the parameter posterior ECDF button is clicked, again parameter empirical cumulative distribution function file could be selected, and the results will be shown as a matlab figure. For the example case, sample results for Springcreek watershed were given in figure 11. However, it should be emphasized that statistical treatment for reducing epistemic uncertainty was not performed in this implementation. So, it is the result of original GLUE methodology which was suggested by Beven & Binley (1992).

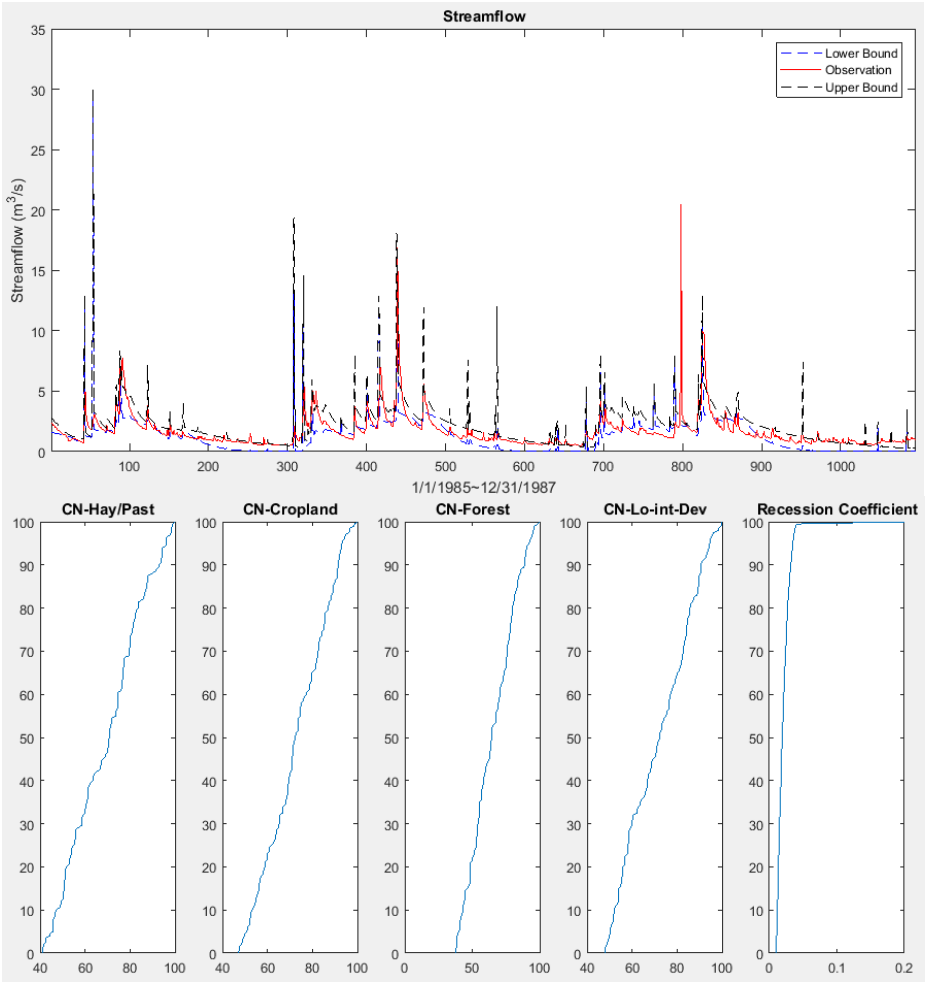

Figure 11. Uncertainty interval (Upper) result and parameter empirical cumulative distribution function (Lower)

3.3 Application

To compare the result from informal likelihood with the result from formal likelihood, the real application of glue methodology suggested in section 2.2 was performed. The input data for simulation is the data of Springcreek watershed in Pennsylvania, USA which was used in section 3.1, section 3.2.

The corresponding observation data is imported from a national water information system, USGS (https://waterdata.usgs.gov/nwis/uv?01546400). Full simulation period is from 1975 to 2010. The variable of interest was streamflow. The information of period of observation data is given in Table 1.

Table 1. Period of observation data

	Daily flow	Monthly flow
Calibration period	1985 ~ 1992	1985 ~ 1996
Validation period	1993 ~ 2000	1997 ~ 2008

Because there is one leap year per 4 years, multiples of 4 years were chosen as appropriate periods for calibration and validation for daily flow and monthly flow. Sampling method was chosen as LHS, and the simulation number was set as 1000. Comparison of the uncertainty interval result for simulation number 1000 with uncertainty interval from 500 shows us the convergence, which confirms the appropriateness of simulation number 1000.

3.3.1 Parameter setting

According to Qi, et al., 2017, recession coefficient and CN2 are identified as important parameters for streamflow. The usual range of recession

coefficient is from 0.01 to 0.2 (Qi, et al., 2017). Formally, 9 land uses were presented in this watershed, namely Hay/Past, Cropland, Forest, Wetland, Coal mines, Turf grass, Unpaved road, Low intensity development area, High intensity development area. However, 4 land uses occupy 96.4% of total area. So, CN2s for these 4 land uses out of 9 land uses will be considered in this application. Corresponding parameter range was chosen as default value which was introduced in 3.2.1. Table 2 shows the portion of these chosen land use areas compared to total watershed area with their parameter ranges.

Table 2. Information about CN2 for selected land uses

	Hay/pasture	Cropland	Forest	Low intensity development area
Range	40 ~ 100	47 ~ 100	38 ~ 100	48 ~ 100
Portion	16.6%	21.1%	40.3%	18.4%

Different from the application of original GLUE methodology in section 3.2.2, the simulation results from SNU-WS were exported as a text file to apply the method introduced in section 2.2. To help the understanding of the readers, the case for daily streamflow was presented step by step following the consecutive procedures.

3.3.2 Reducing epistemic uncertainty

To identify the epistemic uncertainty, simulation result which gives the smallest sum of squared residual was found. Then, the corresponding residual was plotted with runoff, baseflow components. Figure 12 shows this. Here, a

strong relationship between peak residual and runoff, constant residual in the autumn period and baseflow is apparent. Also, when we plot residual against each component including Runoff, baseflow (groundwater input), the other components, the linear relationship appears in each of the figures. The left column of figure 13 shows this fact. It means that we can find the linear relationship between residual and these components using linear regression model. This can be expressed by the following equation.

$$\varepsilon = \beta_0 + \beta_1 RO + \beta_2 BS + \beta_3 Etc + \varepsilon^*$$
$$RO : Runoff \quad BS : Baseflow \quad Etc : Other\ components$$

After completing linear regression analysis, each coefficient (β) can be determined, and corrected residual (ε^*) will be left. The right column of figure 13 shows the relationship between this corrected residual with each component.

As can be seen in the figure, the linear trend disappeared. This implies the epistemic uncertainty related to each streamflow component has been identified. In figure 14, this effect is proven to be indisputable. After adding deterministic part of the regression equation to the original simulation result, this corrected result was compared with the original simulation result. By the comparison, we can notice that peaks occurring by the inexact modeling of runoff process disappeared, and the constant baseflow in the autumn period (part surrounded by the red rectangle) appeared by correction. Comparison of corresponding residuals was also given in the same figure. Using the given equation, all the simulation result was converted to corrected results. These results will be used in subsequent sections.

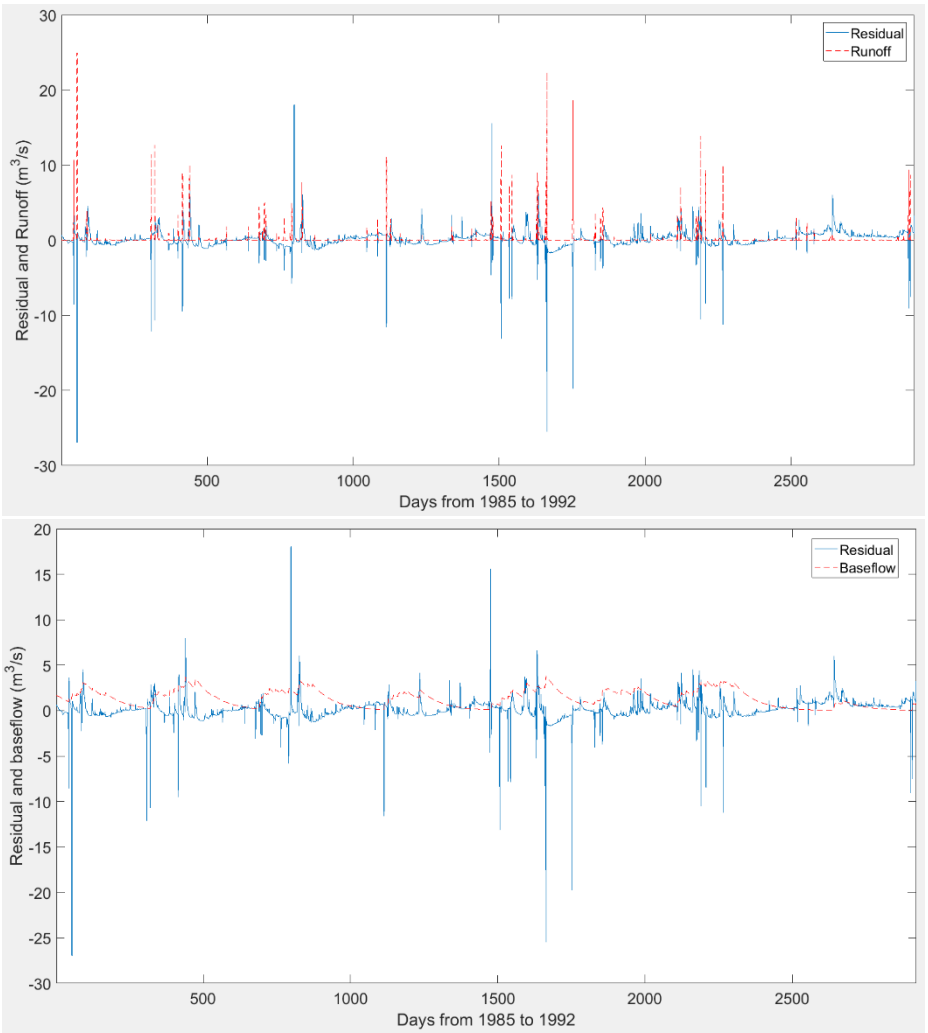

Figure 12. Residual and runoff (Upper), Residual and baseflow (Lower)

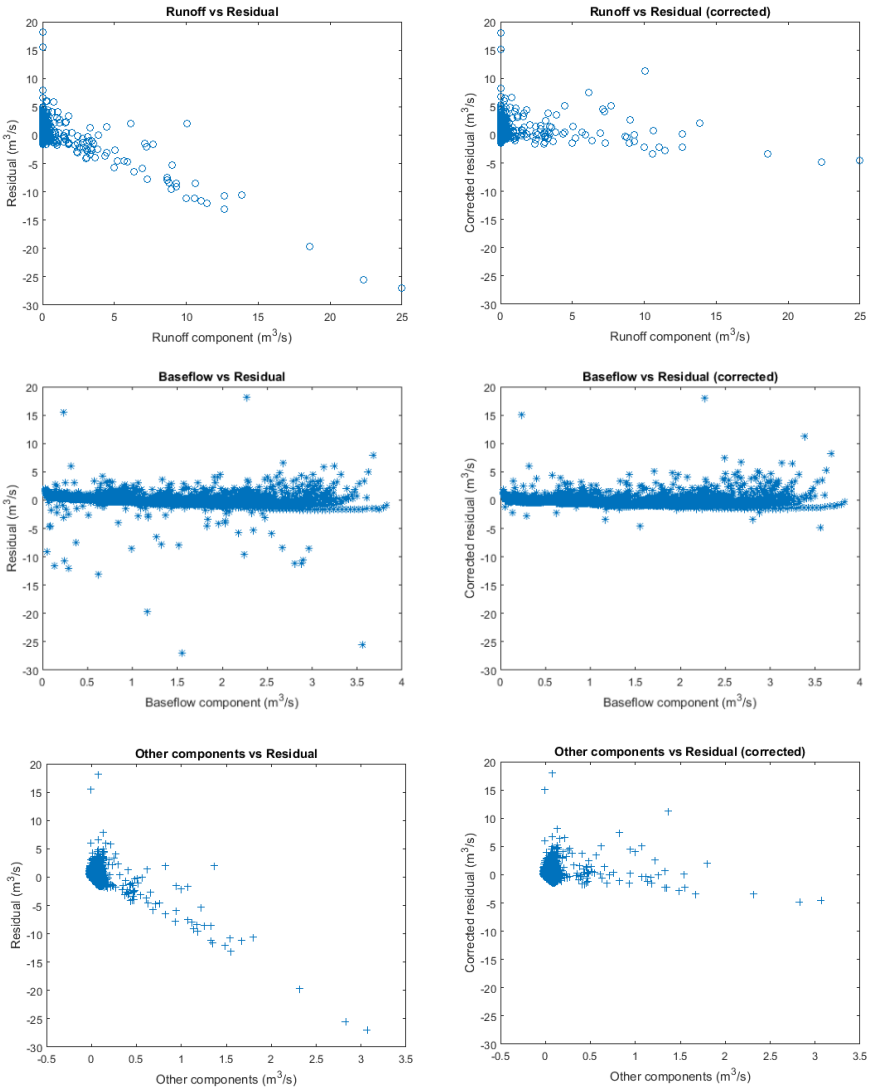

Figure 13. Relationship of each component against residual (left column)

Relationship of each component against corrected residual (right column)

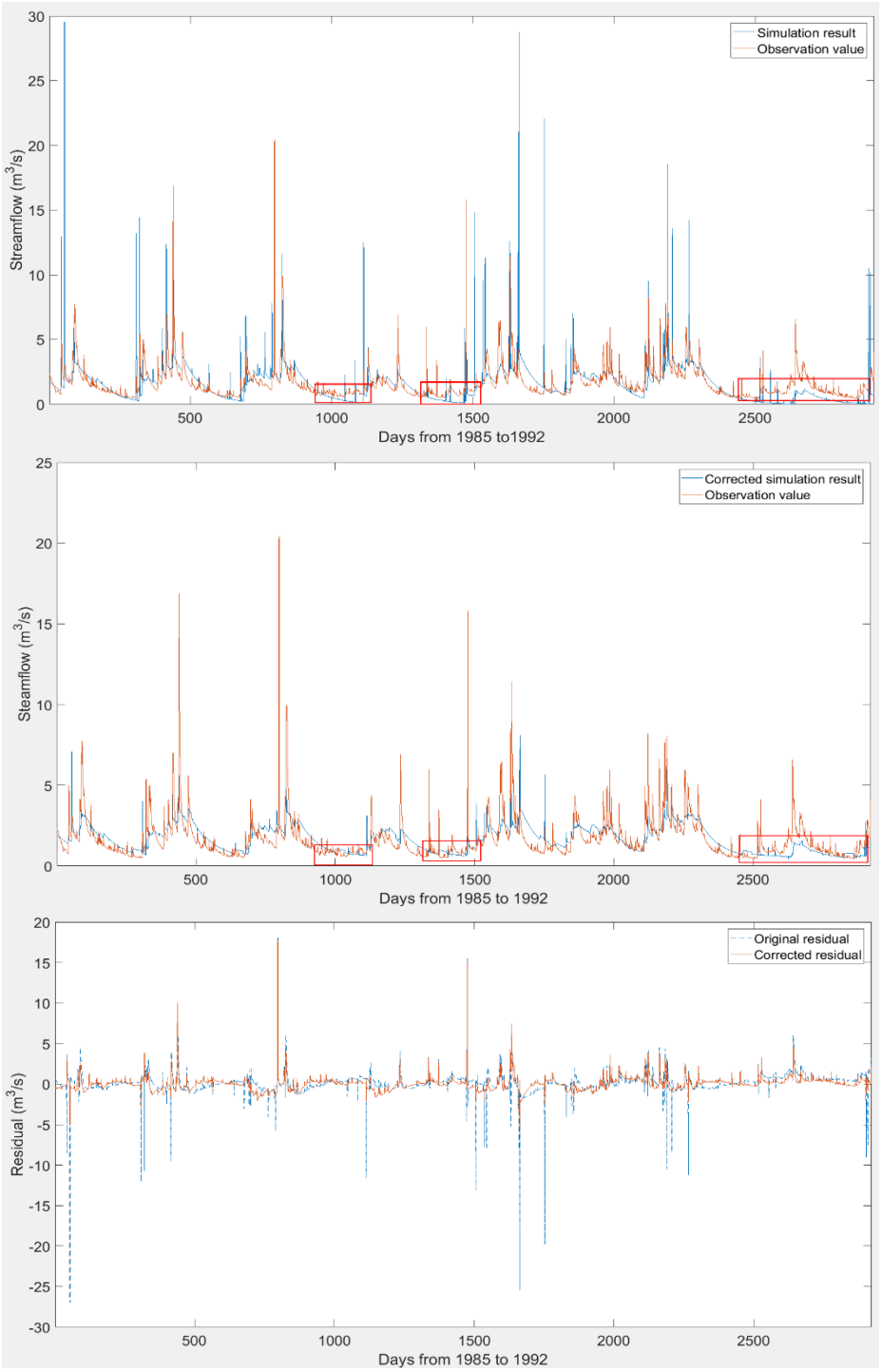

Figure 14. Corrected simulation result and the original one (the second and the first figure), Comparison of corresponding residuals (The third figure)

3.3.3 Informal likelihood definition

For the application of informal likelihood, the corrected simulation results could be used to generate uncertainty interval. This task is quite straightforward because we just apply the method in 2.2.3.1 to get the likelihood measure and follow other procedures in section 2.2 sequentially. The result of the application will be presented in section 4.

3.3.4 Formal likelihood definition

For the application of formal likelihood, the corrected residual should be used to identify the error structure. As noted earlier, the main features of error structure can be considered as the following concepts.

1. Correlation structure
2. Distribution
3. Heteroscedasticity

The treatment of these three features in the real application will be presented in order.

Stedinger et al. (2008) estimated the varianve of error (parameter of error structure) through model results from maximum likelihood estimates of the parameter set. As an approximation, Ruark et al. (2011) assumed that the maximum likelihood estimates of the parameter set as the one which gives the smallest sum of squared residual. Likewise, maximum likelihood estimate of parameter set which gives the smallest sum of squared residual was found, and

the subsequent procedure was applied on corresponding residual. First of all, the time series analysis should be performed. Figure 15 shows the autocorrelation function of the corrected residual of MLE parameter set ($\varepsilon_t^{'}$).

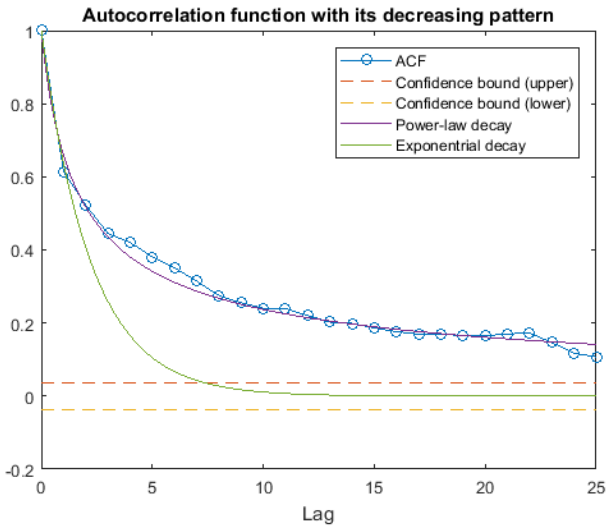

Figure 15. Autocorrelation function of $\varepsilon_t^{'}$ and its decreasing pattern

The formula for power-law decay is $\frac{1}{(1+Lag)^{0.6}}$ and $e^{-0.45\times Lag}$ for the exponential decay. As can be seen apparently in the graph, the decreasing pattern is completely close to power-law like. By finding appropriate d in ARFIMA (0,d,0), this residual could be made independent by fractionally differencing it. This residual after treatment will be denoted as $\nu_t = (1-B)^d \varepsilon_t^{'}$.

The autocorrelation function and partial autocorrelation function of ν_t and $\varepsilon_t^{'}$ are shown in figure 16. It is clear that ν_t is independent. The appropriate

order d was found as 0.45, which shows high long-range dependence of time series. Afterwards, ν_t will be considered to find its statistical distribution.

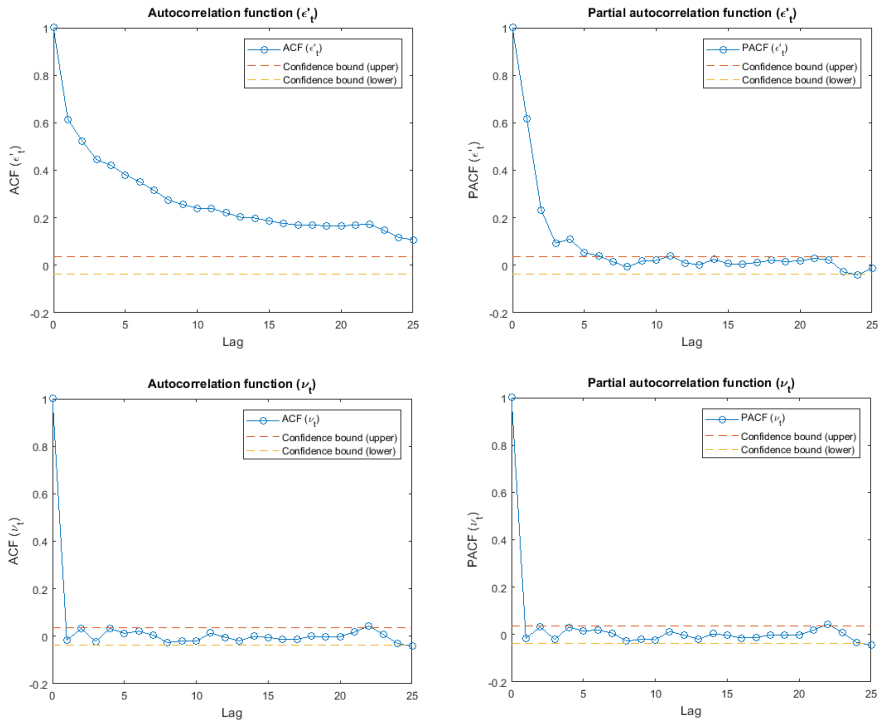

Figure 16. The ACF and PACF of ε_t' (the first row) and ν_t (the second row)

ν_t will be used to find appropriate parameters of stable distribution to get statistical distribution. However, in the course of finding distribution, heteroscedasticity was also considered in daily streamflow. In figure 17, the time series of ν_t and smoothed trend of standard deviation were drawn. Standard deviation was calculated for time series whose span is determined as 31 days centered at each time step. The figure shows clear heteroscedasticity of

ν_t. Also, there are some peak values which occur in a quasiperiodic way. This may attribute to the effect of the snowmelt process. This process usually makes very big sudden peaks in daily streamflow. If only one distribution is found along all time period, this large peak will be dissolved into an estimated one distribution. If this distribution is used to make an uncertainty interval, it will produce a large interval in the period where it is for sure that there will not be a snowmelt process. This large uncertainty interval could be minimized by dividing 1 year to 12 months. By physical consideration, it is reasonable to divide the period because snowmelt occurs at a certain time period in 1 year although it can be changed slightly year by year. In this sense, ν_t was divided according to their corresponding month. Then, the parameters of stable distribution were estimated for each of month using divided ν_t. After estimation, the largest and the smallest value of ν_t in each month was set as the maximum and minimum value of the truncated stable distribution. This is the final distribution which shows error structure. The result is shown in figure 18, 19 with their estimated parameters in Table 3.

α in stable distribution indicates fatness of tail. When we compare α of each month one another, we can see that the value of May, June, July, October, November, December is quite smaller than the one from other months. This may due to the effect of the snowmelt process. Because the snowmelt process introduces an abrupt increase in streamflow, it will make the tail of estimated distribution heavier. This inspection also supports the statement that the stable distribution estimated for each month is a good choice to model the snowmelt

process. Considering a more general case where there are many unknown physical processes happening, this argument can be extended to more general modeling framework. It also should be noted that heteroscedasticity was not considered in the monthly case because the effect of the snowmelt process was not severe in monthly case.

Now, all the information about the error structure is determined. So, we could compute a formal likelihood for each parameter set and get uncertainty interval. First of all, calculate the difference between observation and each simulated result. Then, through fractionally differencing it by order of 0.45, ν_t could be found. ν_t can be classified along each month. Finally, the PDF value of each day will be multiplied for all time steps in the period of interest. At this step, it should be noted that the estimated stable distribution for the corresponding month was considered for PDF value of each day. Although TSD should be considered rigorously, it was too restrictive to produce a reasonable outcome.

In the case of multiplication, the result will be totally zero if one multiplicand becomes zero. Because the PDF of TSD results in zero when the ν_t is out of certain range, the multiplication always results in zero when TSD was adopted. As an alternative, estimated stable distribution was used because PDF of stable distribution never produce zero value for its entire domain.

In contrast to informal likelihood, there is one more step when making uncertainty interval in formal likelihood application. This is the step for adding several realizations of error term to each corrected model simulation result. The

number of model error terms was chosen as the number of time series period Then, using new generated results, uncertainty interval was made. For illustration, several realizations of model error terms are shown in figure 20.

Table 3. Estimated parameters of truncated stable distribution for each month

Parameters / Months	Alpha	Beta	Gamma	Delta	Maximum	Minimum
1	1.207	-0.497	0.107	-0.244	-5.952	14.731
2	1.051	0.173	0.111	0.147	-4.718	3.521
3	1.046	0.557	0.149	1.007	-7.085	18.172
4	1.334	0.095	0.149	-0.128	-1.098	2.960
5	0.879	0.224	0.113	-0.164	-0.980	2.900
6	0.865	0.344	0.064	-0.125	-0.960	5.796
7	0.808	0.096	0.058	0.000	-5.187	1.860
8	0.919	0.529	0.043	-0.212	-1.092	2.911
9	0.927	0.801	0.043	-0.372	-0.598	1.398
10	0.821	0.289	0.034	-0.108	-1.908	3.495
11	0.864	0.063	0.083	-0.029	-2.819	3.561
12	0.763	-0.478	0.073	0.045	-1.801	3.737

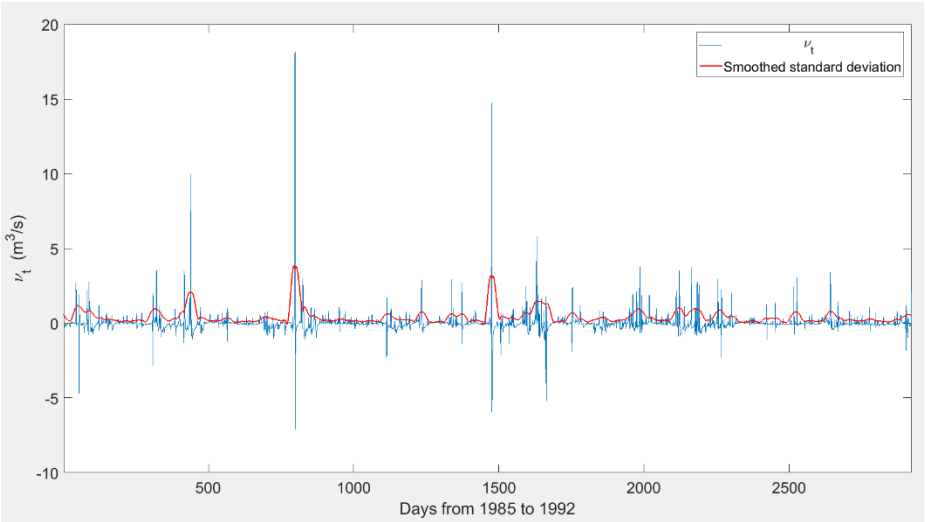

Figure 17. ν_t with a smoothed trend of standard deviation

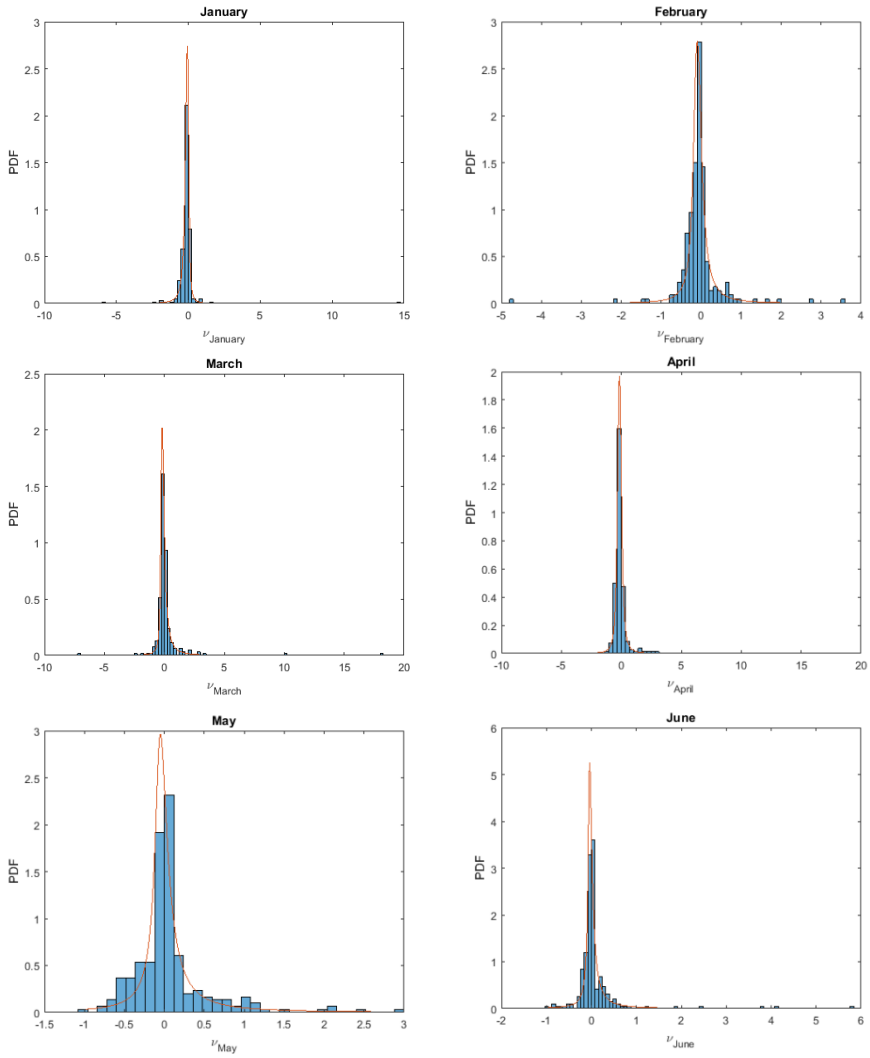

Figure 18. PDF derived from ν_t with their estimated stable distribution

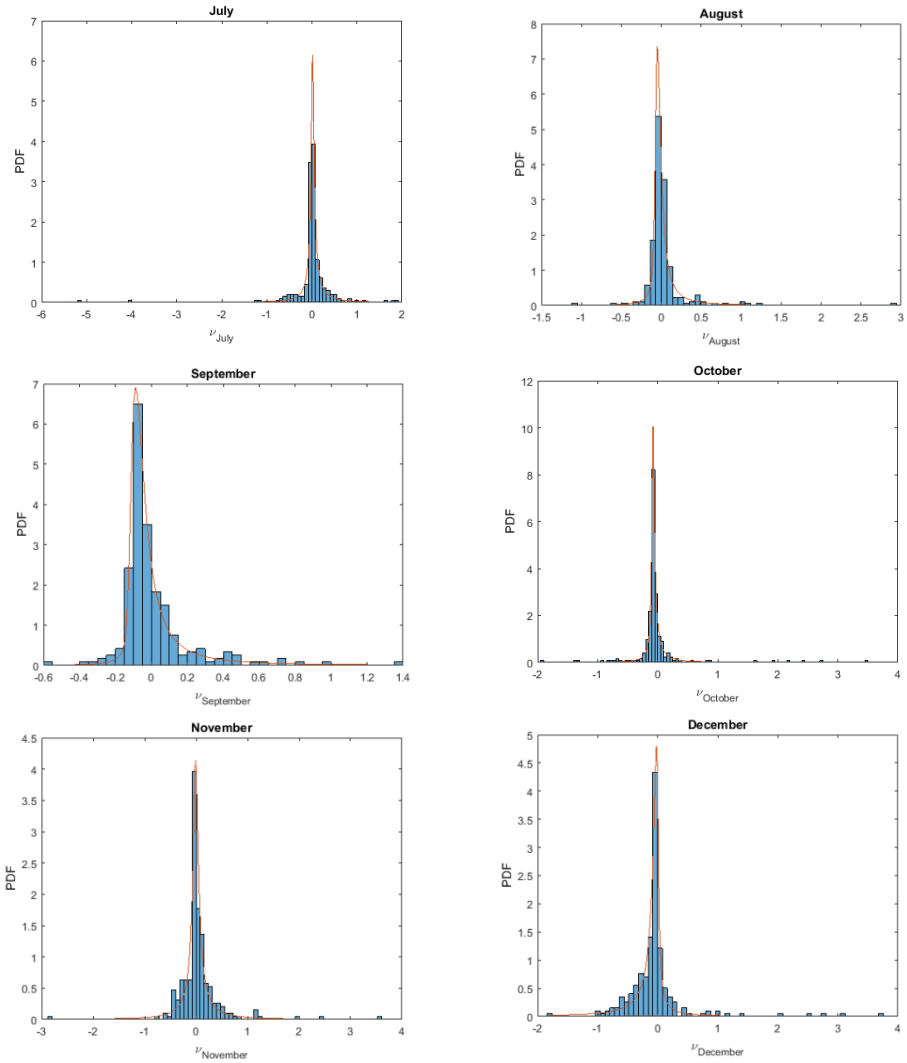

Figure 19. PDF derived from ν_t with their estimated stable distribution

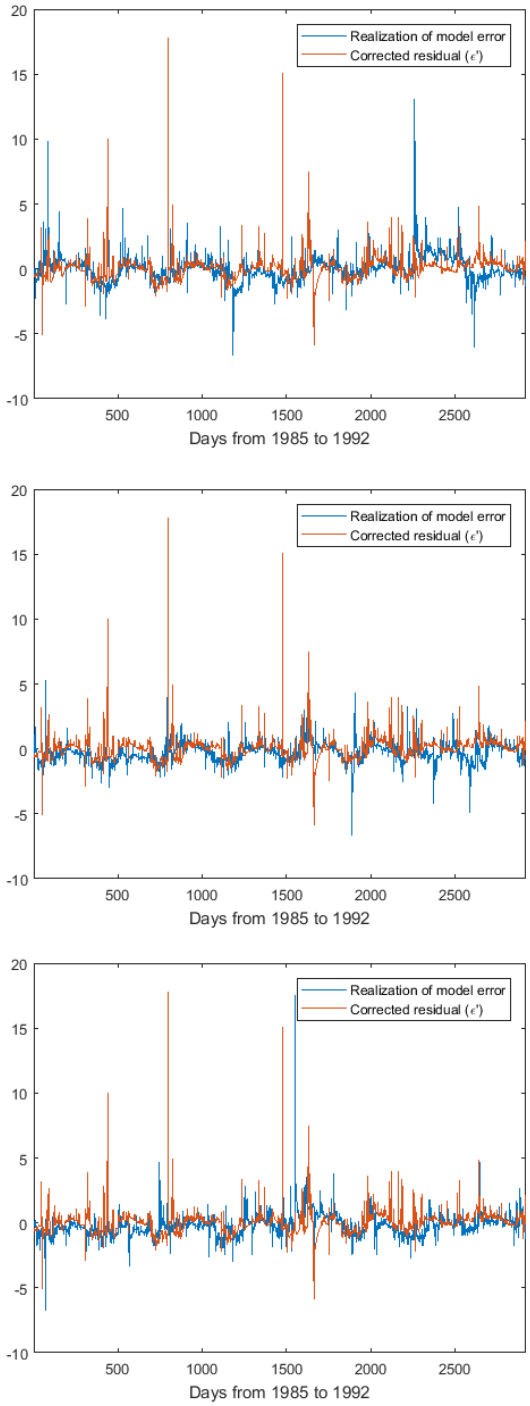

Figure 20. Several realizations of model error with corrected residual (ε')

Chapter 4. Result

4.1 Monthly flow

4.1.1Informal likelihood

There are numerous definitions of informal likelihood definition because of its subjectivity. In this application, three cases are considered. NSE, inverse error variance of shaping factor 1 (IEV1) and 10 (IEV10). To investigate parameter uncertainty, parameter posterior probability distribution was explored. Figure 21 shows this result. The results from NSE and IEV1 are similar for all parameters and quite uniform in CN2 but have some higher probability on low values in recession coefficient. However, the result from IEV10 shows a clear trend in each function. This is because IEV10 gives a more rapid change of likelihood measures. The most obvious one is recession coefficient. It is for sure that some low recession coefficient gives the highest pdf value. As an explanation, this watershed has low recession coefficient, which means a small portion of groundwater moves to streamflow.

For the investigation of predictive uncertainty, the uncertainty interval of calibration and validation period for a 95% significance level was given in figure 22 and figure 23. As an interpretation of the statistical meaning of significance level, uncertainty interval should cover observation values in the proportion of its significance level, but the actual coverage ratio was found to

be significantly less than theoretical one. These results are summarized in table 4. It does not coincide with the expected values. Therefore, we can conclude that the uncertainty interval of a certain significance level is not statistically meaningful in informal likelihood case. Also, IEV10 case gives a short interval in both periods, which implies a rapid change of its likelihood values.

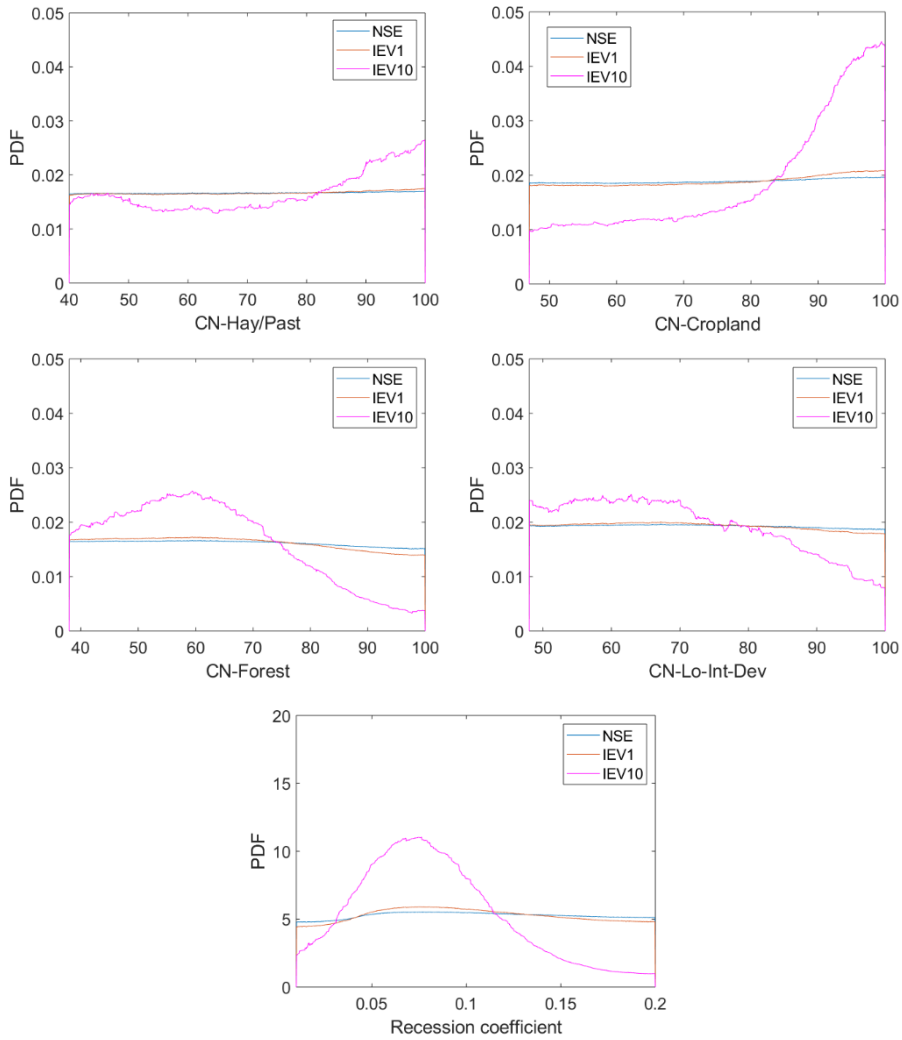

Figure 21. Parameter posterior probability distribution according to informal likelihood definition (Monthly streamflow)

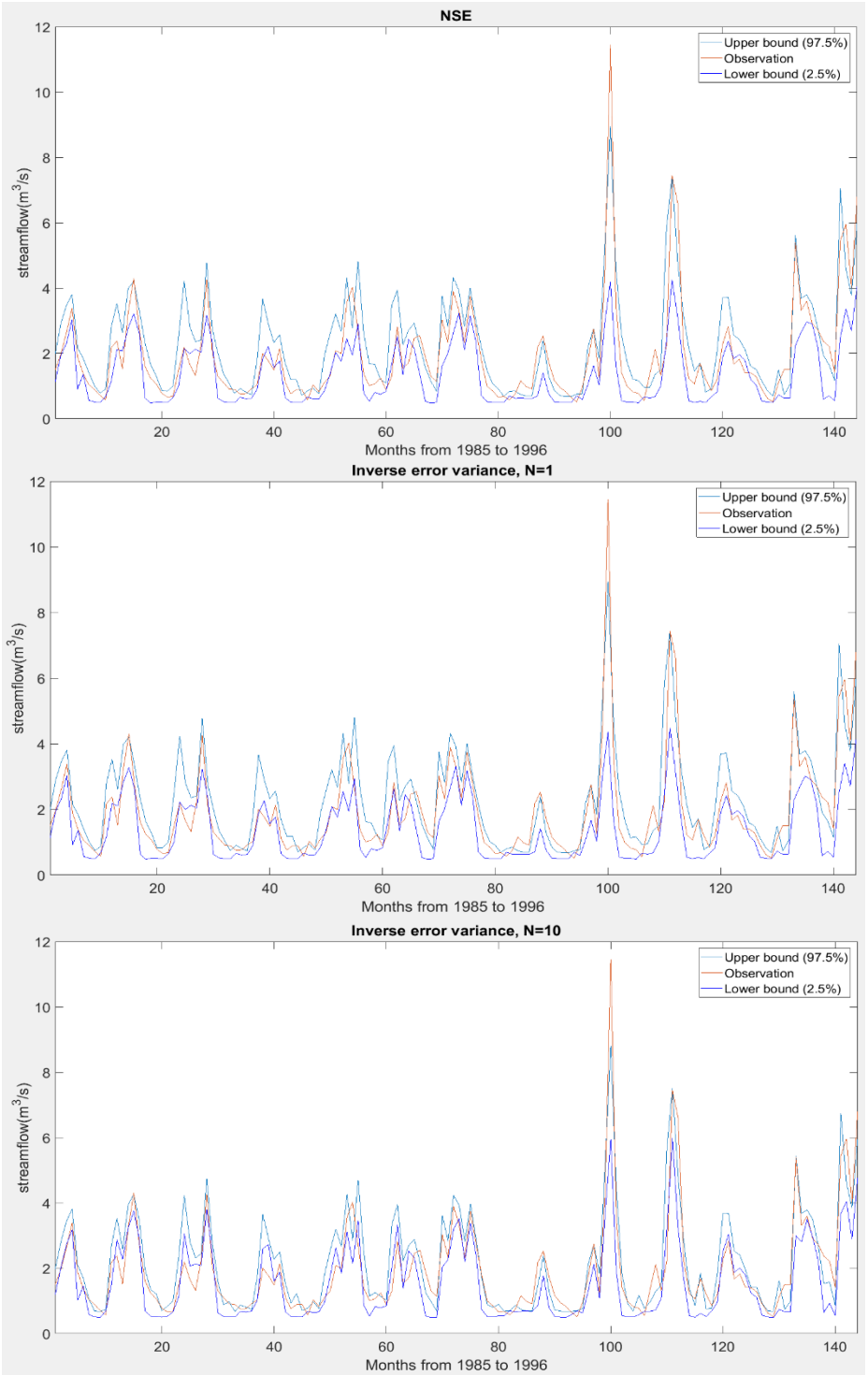

Figure 22. Uncertainty interval for 95% significance level (calibration)

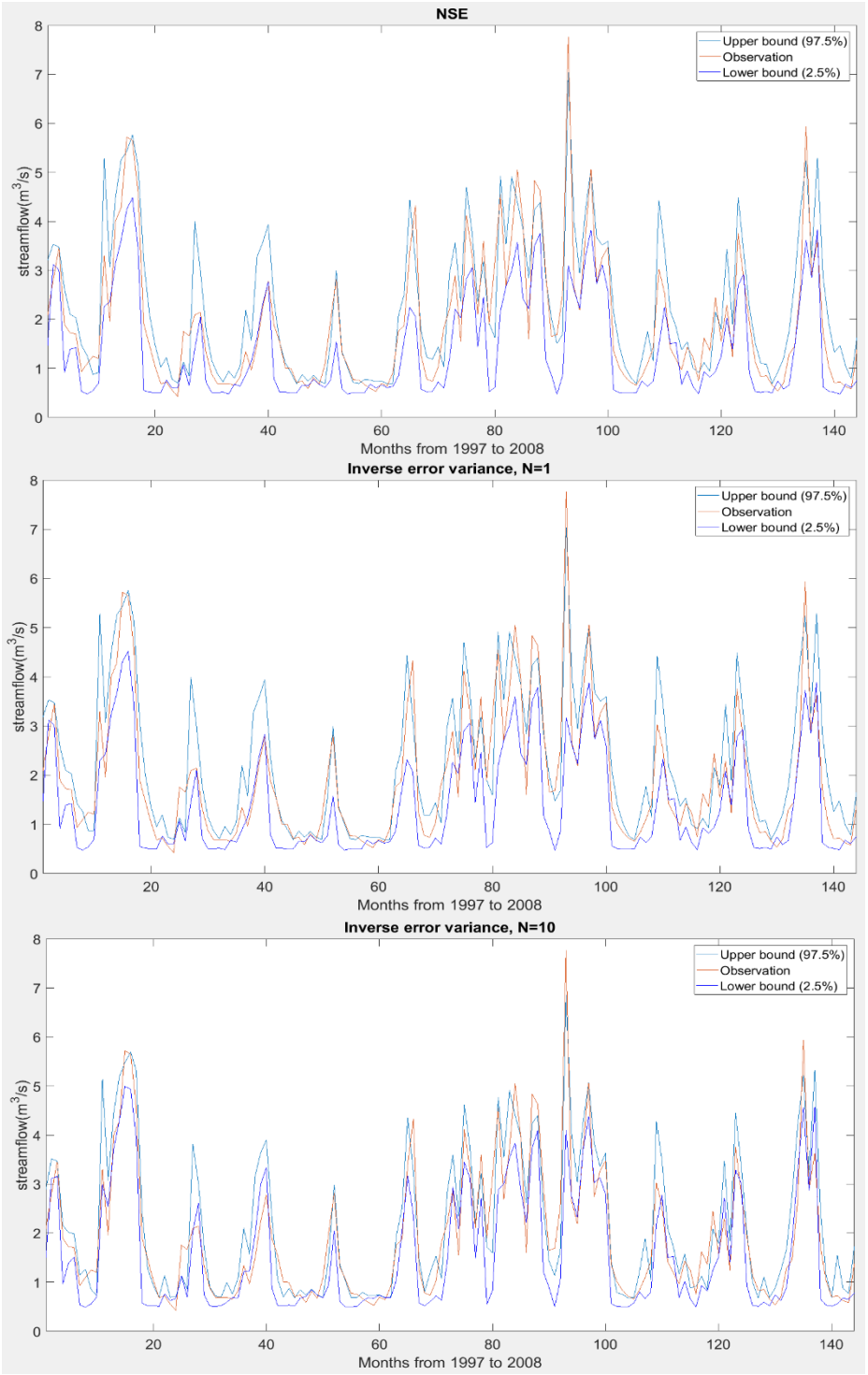

Figure 23. Uncertainty interval for 95% significance level (validation)

4.1.2Formal likelihood

Parameter posterior probability distribution was shown in figure 24. It shows a clear trend in most function similar to IEV10. This is due to the restrictive property of formal likelihood definition. Formal likelihood definition incorporates numerous multiplications when calculating the likelihood measure. Because multiplication gives restrictive property to its outcomes (If some multiplicand is zero, the result will be zero.), following results show a rapid change in its values. When the number of multiplicands becomes higher, the result will be more restrictive. This will be seen again in the daily streamflow case. Although steep attribute of function is the sameas the result of IEV10, the pattern is quite different. However, the result of recession coefficient is similar. For both formal and informal likelihood cases, recession coefficient has higher pdf values in some low values. This assures the interpretation given in section 4.1.1.

For an exploration of predictive uncertainty, the uncertainty intervals were drawn as before (figure 25, figure 26). Now, the coverage ratio is near its significance level (Table 4). It means that the uncertainty interval made by formal likelihood definition is reliable in its statistical meaning in both periods.

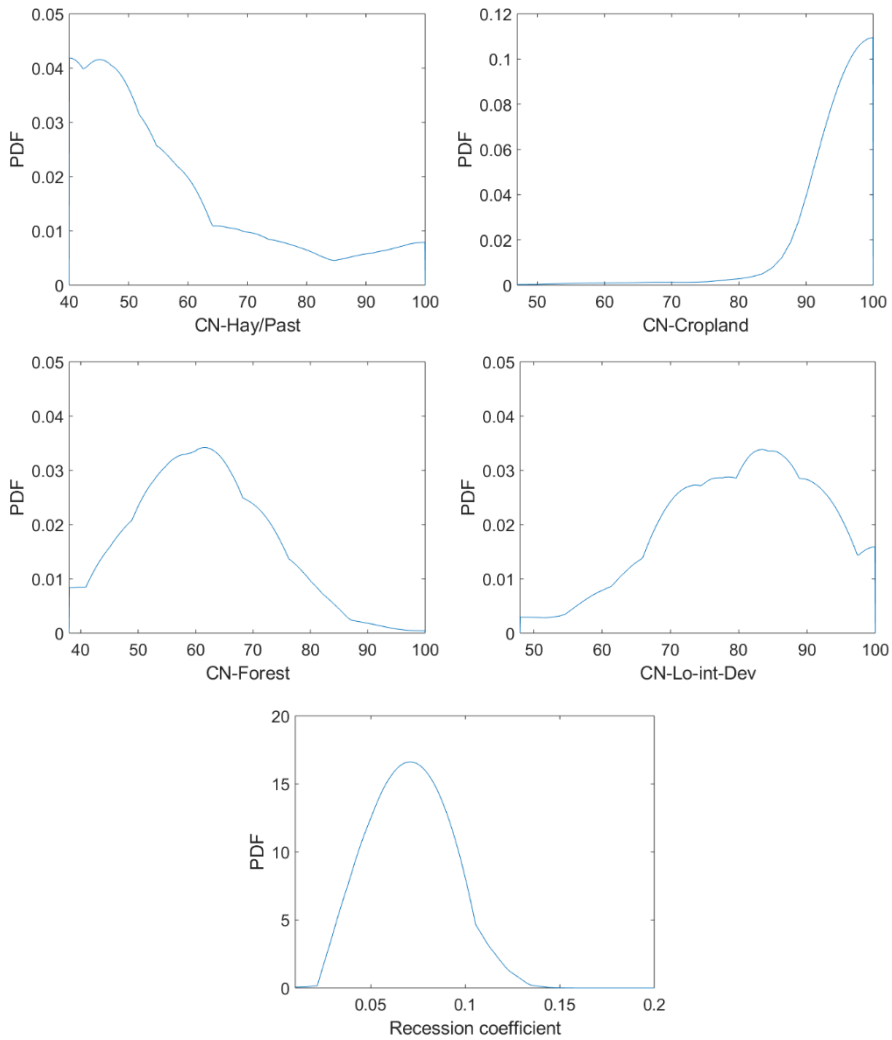

Figure 24. Parameter posterior probability distribution according to formal likelihood definition (Monthly streamflow)

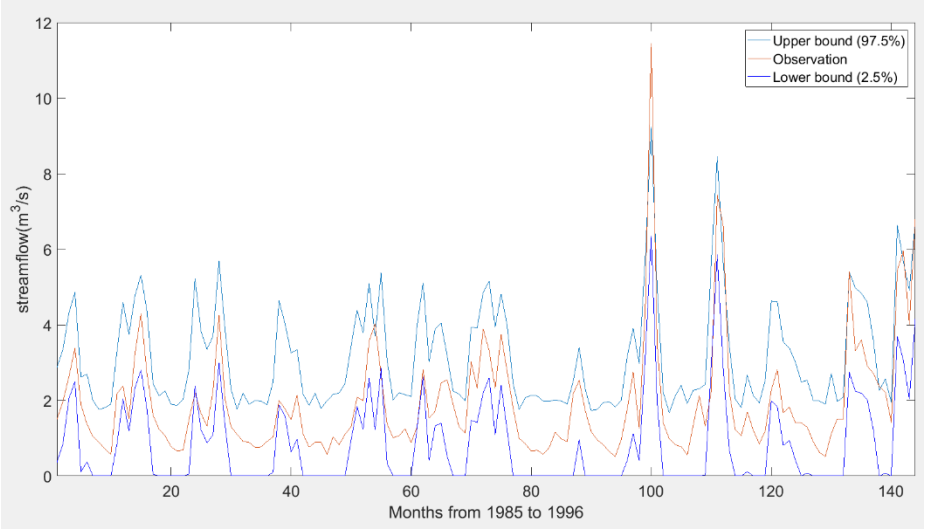

Figure 25. Uncertainty interval for 95% significance level (calibration)

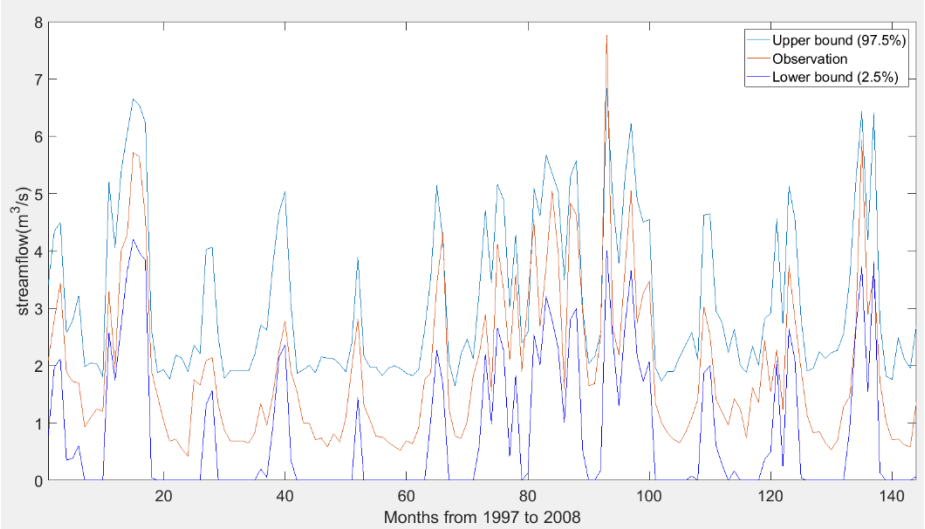

Figure 26. Uncertainty interval for 95% significance level (validation)

Table 4. The coverage ratio of uncertainty interval in both periods according to likelihood definitions

	Informal likelihood			Formal likelihood
	NSE	IEV (N=1)	IEV (N=10)	
Calibration	0.6181	0.6042	0.4444	0.9306
Validation	0.6042	0.5764	0.4375	0.9722

4.1.3 Comparison

For parameter uncertainty, formal likelihood gives steeper functional shape than informal likelihood case. It means formal likelihood results in less parameter uncertainty. This fact also appears in daily flow case, and the effect is more obvious. The reason will be explained in section 4.2.3 in detail.

By comparing the predictive uncertainty of two cases, we can observe that coverage ratio fits well in its significance level for formal likelihood case meanwhile it does not match well in informal likelihood case. This is due to the absence of modeling of model error term in informal likelihood case. However, the interval size is quite bigger in formal likelihood case than informal likelihood case. This also can be understood by the fact that the uncertainty interval in a formal likelihood case is made after adding sampled model error terms into original simulation results.

Finally, we can conclude that a formal likelihood gives more statistical reliability but gives larger intervals and represents less parameter uncertainty. Informal likelihood case is opposite to this.

4.2 Daily flow

4.2.1 Informal likelihood

NSE and inverse error variance definition were used for comparison. Two shape factors (N) were examined in inverse error variance, 1 (IEV1), 50 (IEV50). Parameter uncertainty was assessed through parameter posterior probability distribution (figure 27). Similar to the monthly case, the results from NSE and IEV1 are not quite different, but the result from IEV50 shows again a very clear trend. In the case of NSE and IEV1, almost uniform distribution was found for CN2 parameters, and the same but clearer trend compared to the monthly case was found in recession coefficient due to scale effect. In the case of IEV50, the trend in CN2 is quite different from the result from the monthly case, but the trend in recession coefficient is more obvious. Some low value gives high pdf values. The difference between the results from two different time scales could be understood as a scale effect (because time scale had been changed). Nevertheless, all the results from both time scales support the fact that a certain low recession coefficient is a good descriptor of this watershed. Again, we could say that this watershed is less likely to allow the movement of groundwater to channel.

For predictive uncertainty, uncertainty intervals of 95% significance level for the calibration period were presented in figure 28, 29, 30. The corresponding one for validation period were also presented in figure 31, 32, 33. When we see these outcomes, it can be noticed that the outcome from inverse error variance of shaping factor 50 has a smaller interval compared to the others. This shows the effect of shaping factor on likelihood, the higher the

shaping factor, the steeper likelihood function. To see how much the interval captures the observations, the coverage ratio versus corresponding significance level were drawn (Figure 34). Here, the coverage ratio means how much uncertainty interval cover the observation values when uncertainty intervals are defined by two quantile lines , $\left(100-\frac{100-p}{2}\right)\%$ and $\left(\frac{100-p}{2}\right)\%$, where p is the significance level. By inspection, the coverage ratio shows limited efficiency and high deficiency to capture observation values when the significance level becomes higher. This is because the uncertainty interval made by informal likelihood does not consider model error. It could not capture the values which are out of the range where the model can describe (Usually, these values are extreme values). This is the limitation of informal likelihood. One notable feature also should be mentioned. It is the fact that the graph of coverage ratio does not change much according to its period classification. This can be understood by the consideration that efficiency of informal likelihood is limited to model efficiency which is not changed along the time.

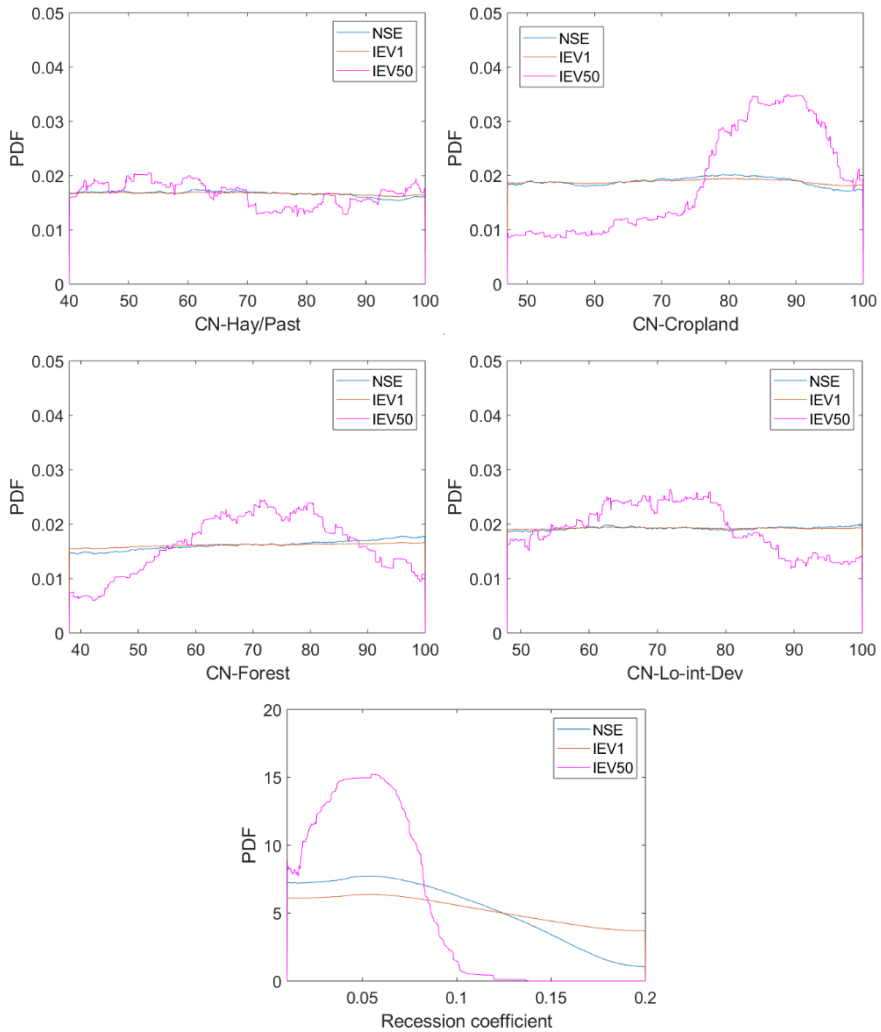

Figure 27. Parameter posterior probability distribution according to informal likelihood definition (Daily streamflow)

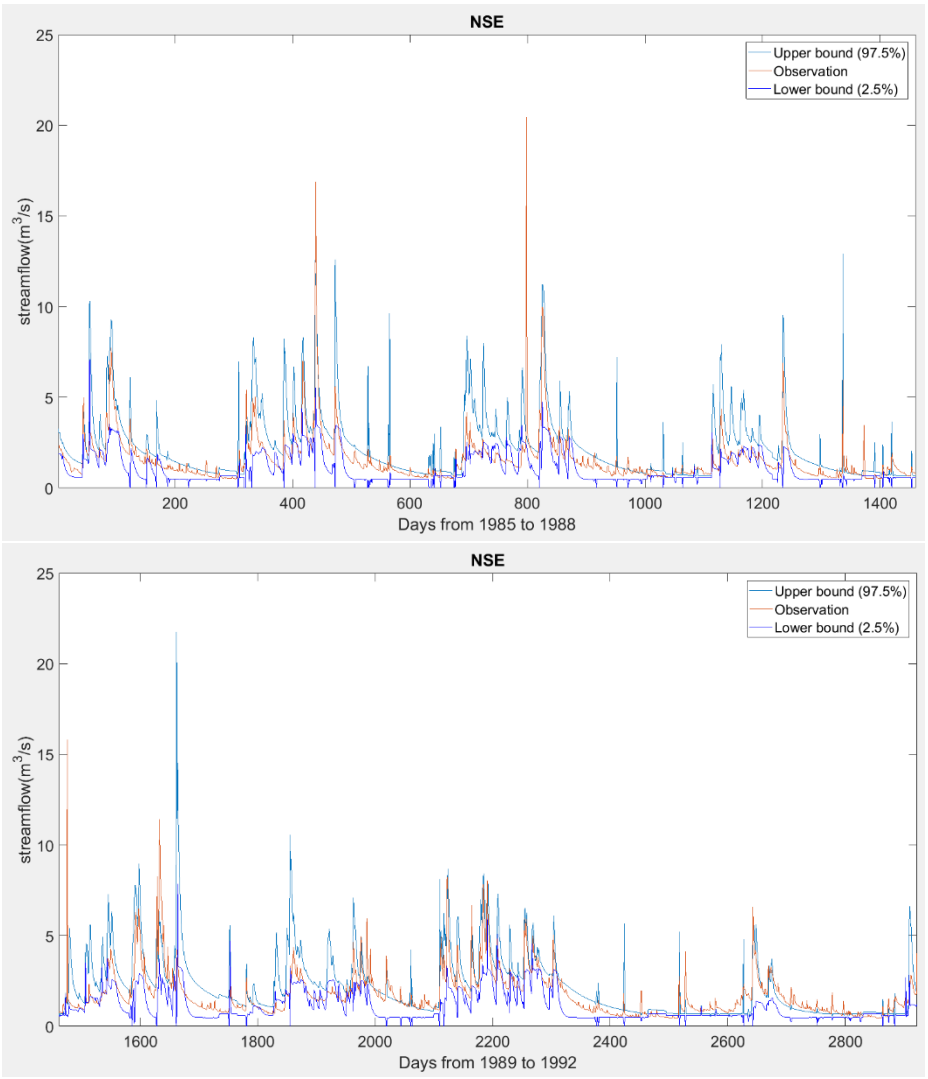

Figure 28. Uncertainty interval for 95% significance level in calibration period (NSE)

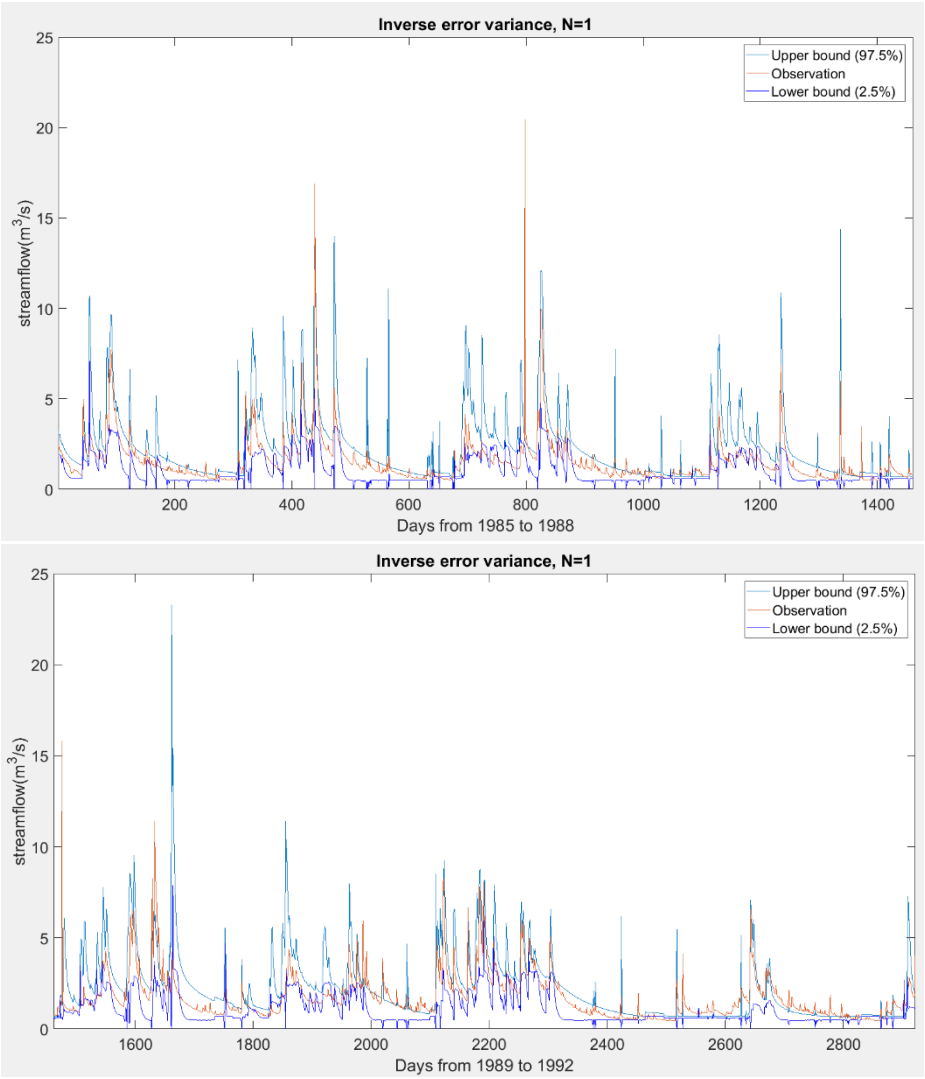

Figure 29. Uncertainty interval for 95% significance level in calibration period (Inverse error variance, N=1)

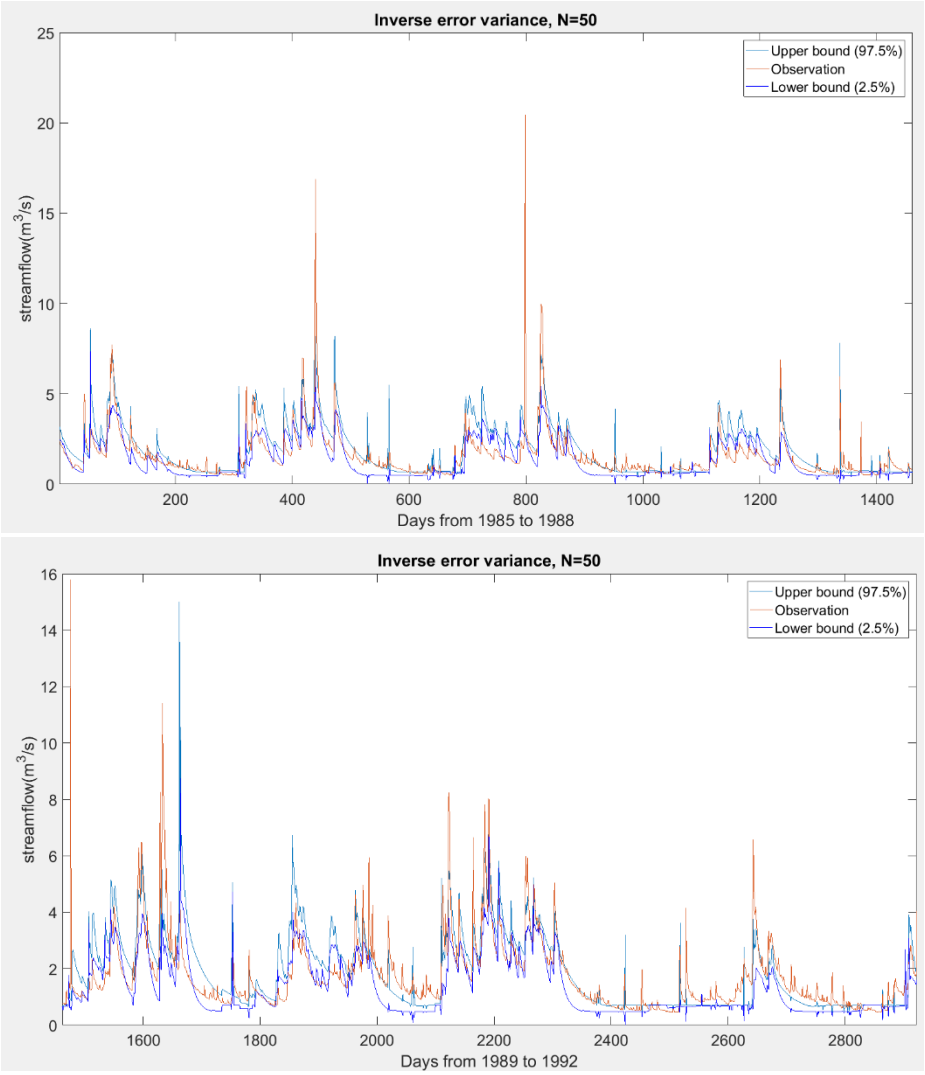

Figure 30. Uncertainty interval for 95% significance level in calibration period (Inverse error variance, N=50)

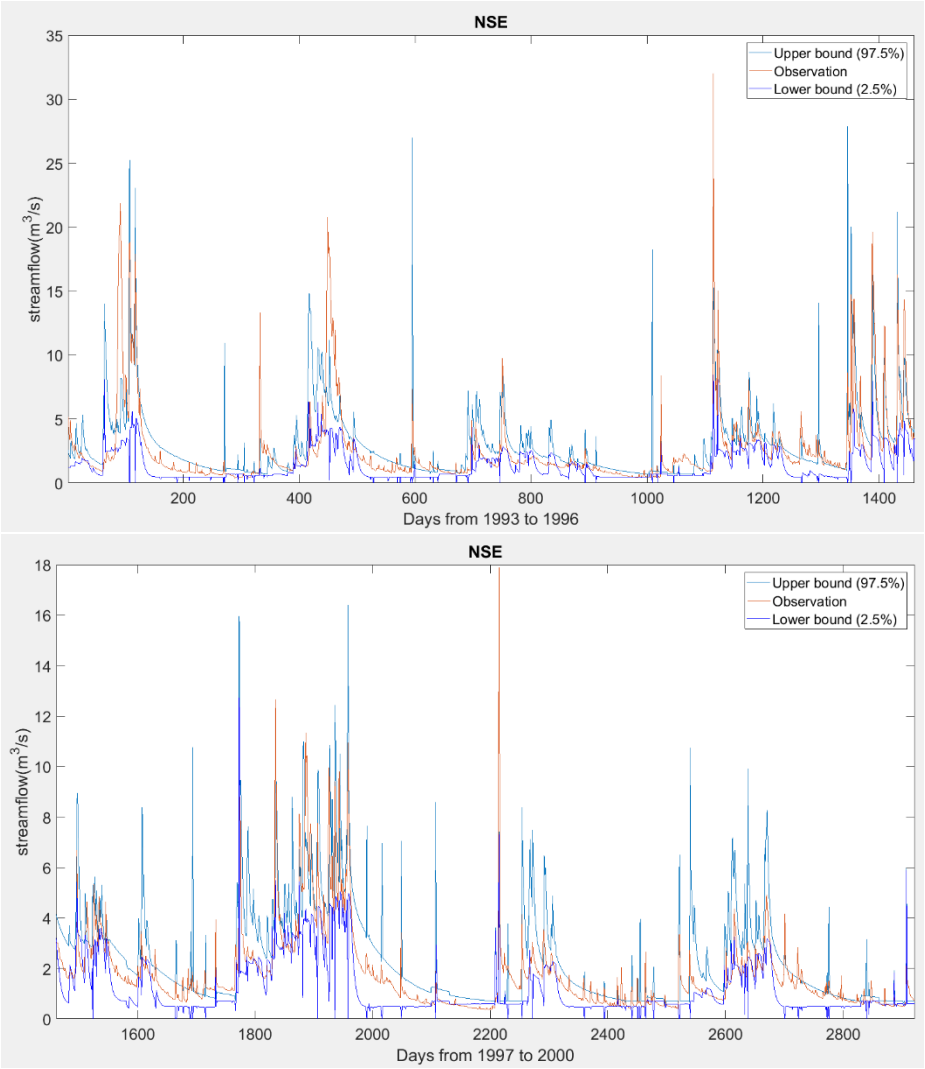

Figure 31. Uncertainty interval for 95% significance level in validation period (NSE)

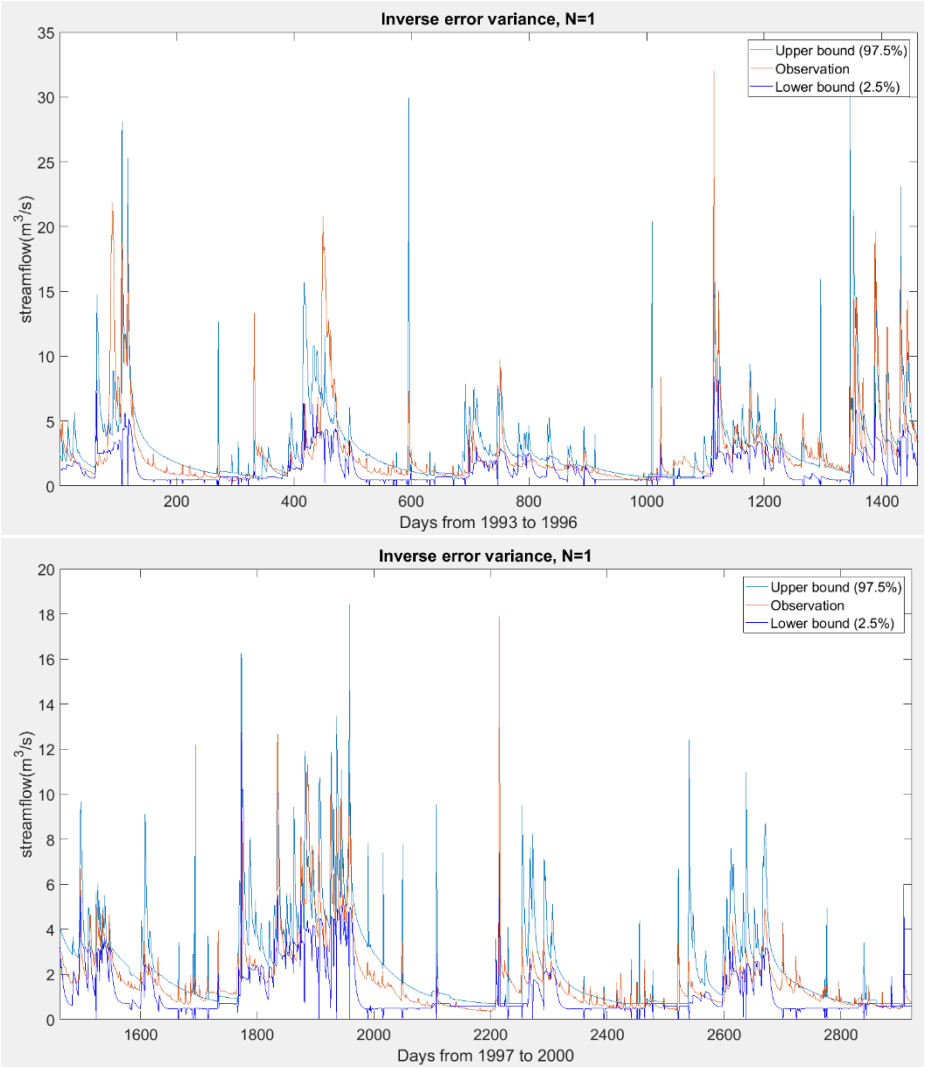

Figure 32. Uncertainty interval for 95% significance level in validation period (Inverse error variance, N=1)

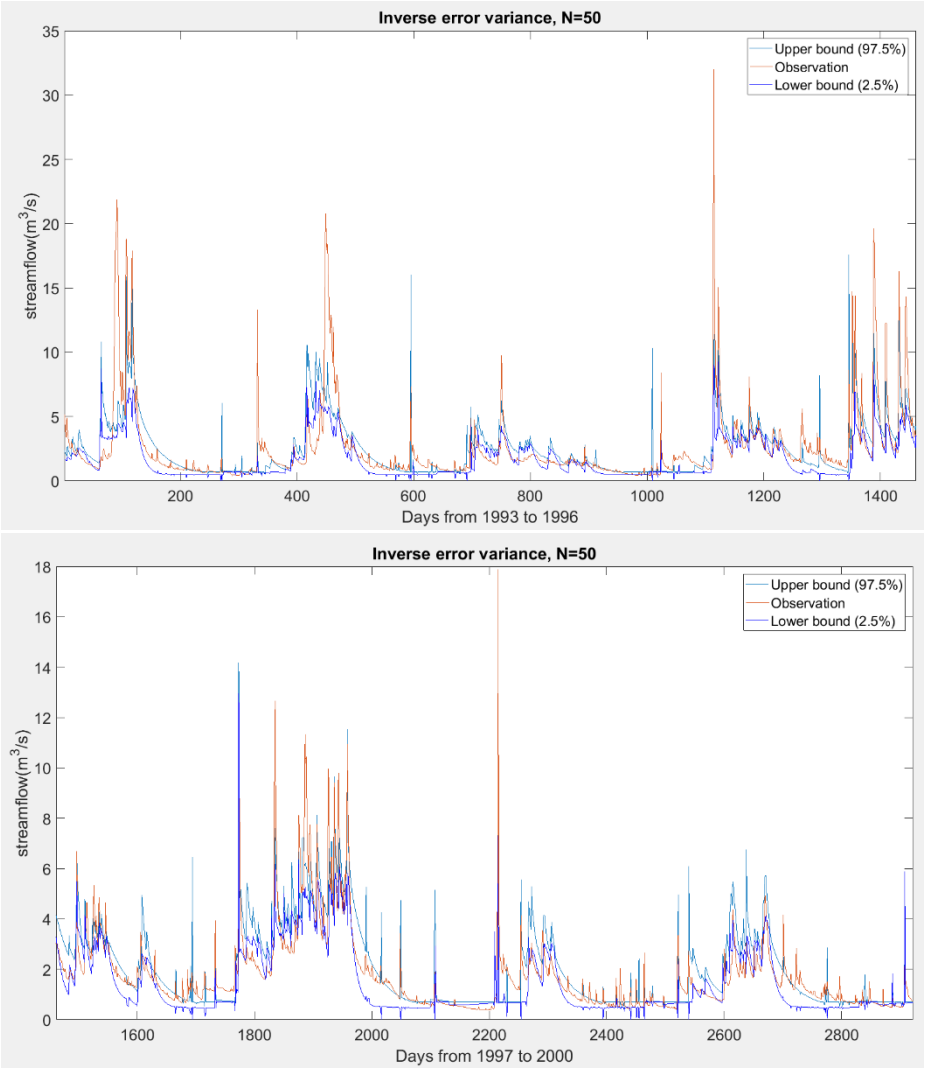

Figure 33. Uncertainty interval for 95% significance level in validation period (Inverse error variance, N=50)

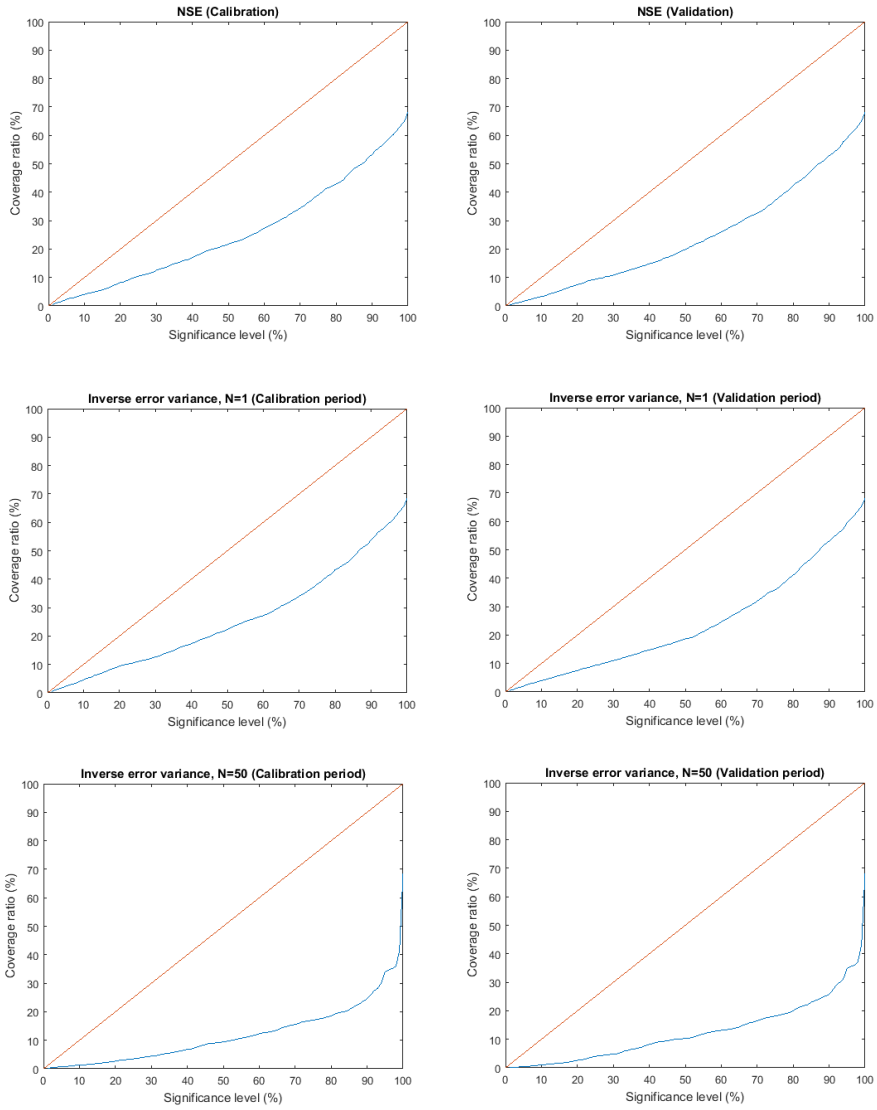

Figure 34. Coverage ratio and significance level of each period for various likelihood definitions

4.2.2 Formal likelihood

First of all, parameter uncertainty was assessed through calculated likelihoods. As covered in section 4.1.2, formal likelihood results in steeper parameter posterior probability distribution than the informal case. This attributes to the fact that the calculation of formal likelihood accompanies a lot of multiplication. Now, this trend became more apparent in the daily streamflow case. Because the number of multiplicands in daily streamflow is 2922, most calculated likelihoods converge to zero. Therefore, parameter uncertainty is highly reduced, resulting in Dirac-delta function shape in the specific parameter value. These parameter values are the values which give the best simulation result considering model error. The results are shown in figure 35.

Secondly, predictive uncertainty was assessed through uncertainty intervals. In figure 36, the uncertainty interval of calibration period was represented for a 95% significance level. Corresponding uncertainty interval of the validation period is shown in figure 37. For both cases, the corresponding coverage ratio was found as 0.9391, 0.8638. For the calibration period, the value is quite closer to the theoretical one as opposed to the validation period. This result could be understood well when we plot coverage ratio according to many significance levels. Figure 38 shows these plots for each period. As can be seen in the graph, the coverage ratio is mostly consistent with the theoretical one in the calibration period. It gives even higher coverage ratio on most significance levels. The significance level when the theoretical one matches exactly with coverage ratio was found as 86%. Contrary to this, the coverage ratio in the validation period does not match well in the high significance level

region. Accordingly, the significance level when the theoretical one matches exactly with coverage ratio was found to be around 60%. The mismatch between theoretical one and computed one in higher significance level region could be understood as the failure to capture the exact timing of physical processes which had not been modelled by SNU-WS. This situation could be understood by the graph in figure 39. In figure 39, several high peaks which the intervals could not capture were shown. As can be noticed, the uncertainty interval was successful for capturing the amount of peak value in their range but failed to capture the exact timing of their occurrence. This means that the more sophisticated treatment of heteroscedasticity is needed because heteroscedasticity involves the concept of temporal variation of the statistical attribute. Finally, some notable characteristics of the uncertainty interval introduced by the use of stable distribution should be emphasized here. To present characteristics of heavy-tailed distribution, the gaussian distribution and stable distribution of the alpha value of 0.7627 were drawn in figure 40. Because heavy-tailed distribution gives more probability on extreme values, the decreasing pattern of quantiles in the tail part is quite rapid. This fact could be recognized by the inspection of the pdf of each distribution. Considering this argument, we can expect that the uncertainty interval will be changed largely in higher significance levels. In figure 41, the quantiles for 99%, 99.5%, 100% in the calibration period were shown. As expected before, the upper bound for uncertainty interval changes largely according to various significance level. The implication of this phenomena is that we need a larger uncertainty interval to capture the more extreme event, and stable distribution gives the appropriate

size of uncertainty interval to cover those events. Moreover, if we did not use stable distribution and used the normal distribution, the uncertainty interval could be made lower but may not capture the extreme events consistently. Therefore, considering the consequence of extreme events in the usual design experience, the adoption of stable distribution can be thought of as a good choice.

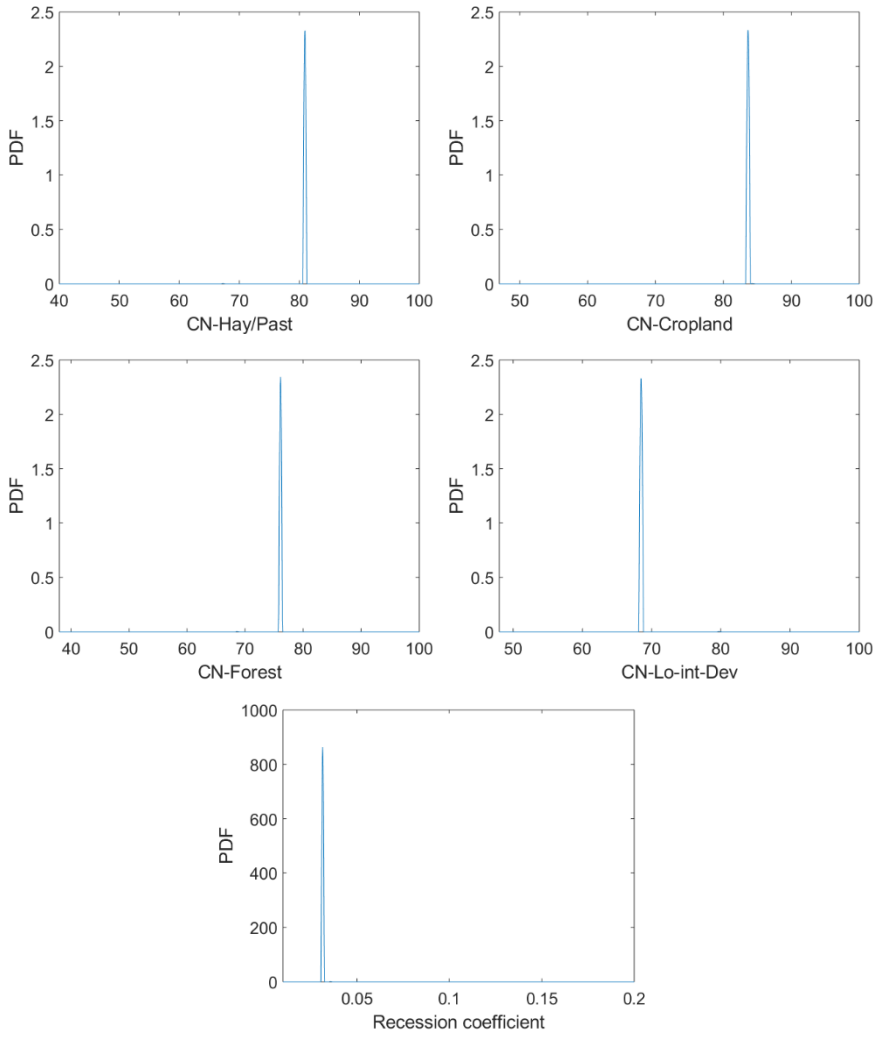

Figure 35. Parameter posterior probability distribution of 5 parameters

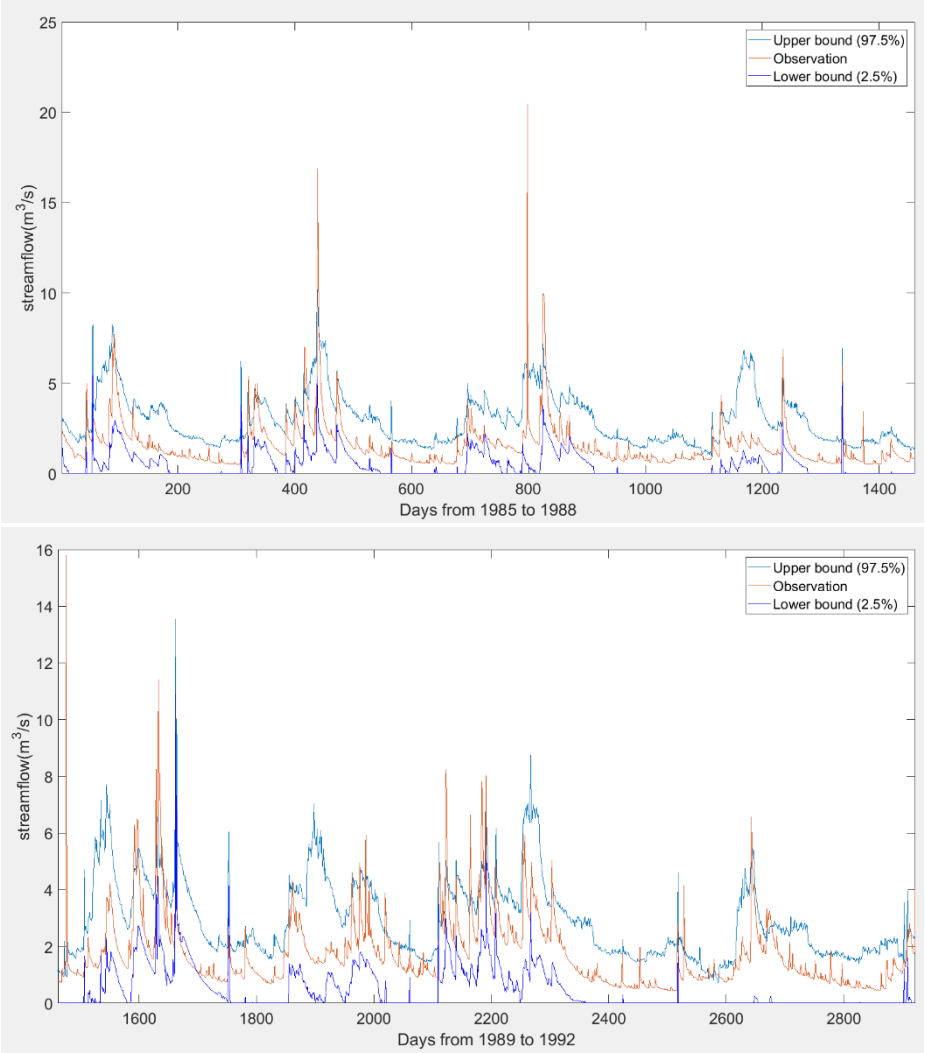

Figure 36. Uncertainty interval for 95% significance level in calibration period

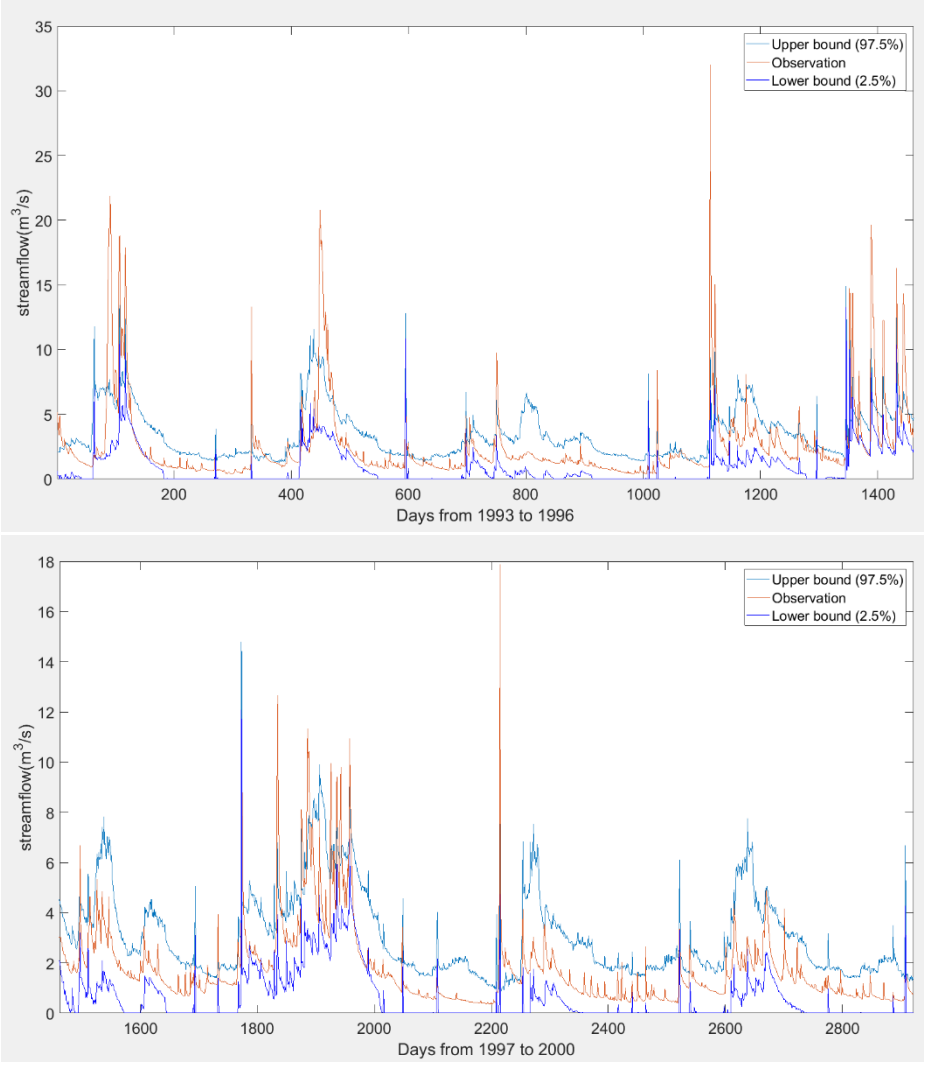

Figure 37. Uncertainty interval for 95% significance level in validation period

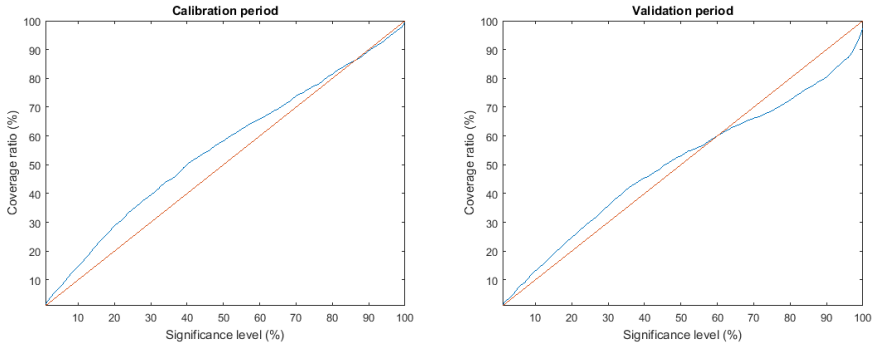

Figure 38. Coverage ratio according to its significance level

(blue line is computed one and the red line is theoretical one)

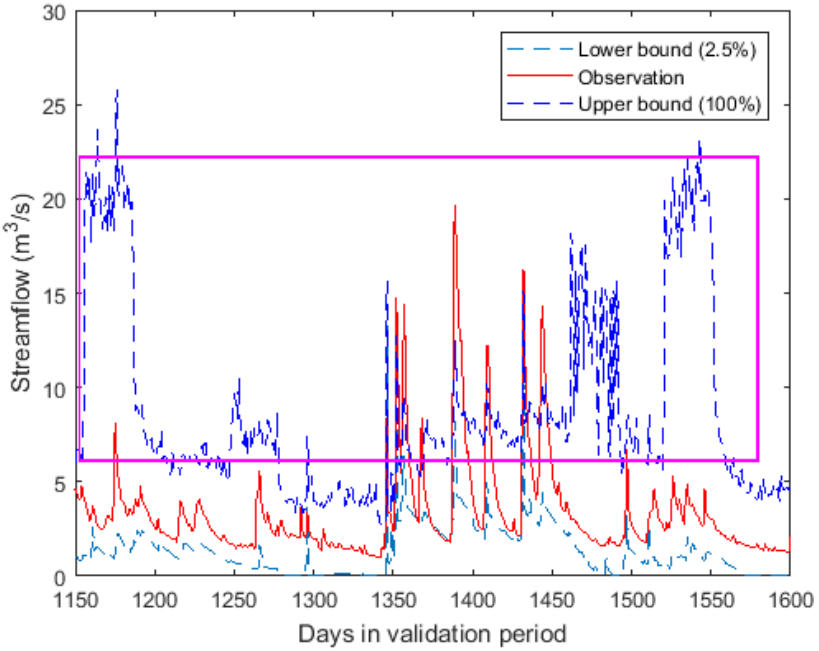

Figure 39. Example peaks which the uncertainty interval could not cover

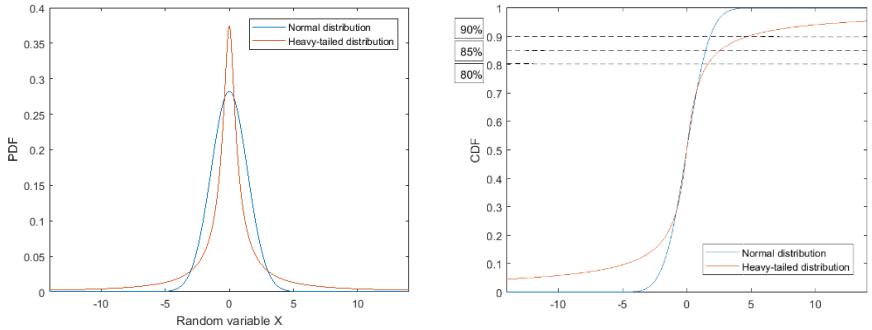

Figure 40. PDF and CDF of normal and a heavy-tailed distribution

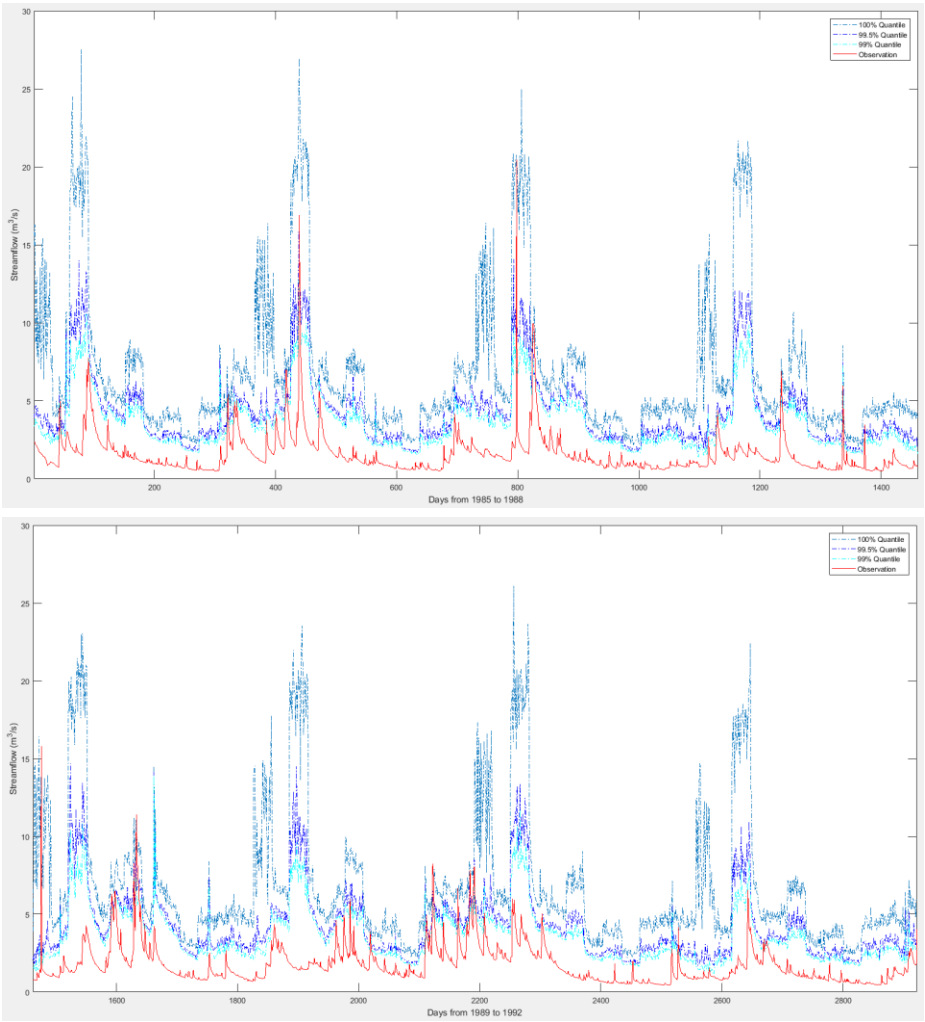

Figure 41. Various quantiles to capture extreme phenomena in calibration period

4.2.3 Comparison

In respect to parameter uncertainty, the informal likelihood gives us a more uncertain picture for all parameters as opposed to results from formal likelihood. This feature can be explained by the attribute of each methodology. Because the definition of formal likelihood involves the explicit definition of error structure, the parameter uncertainty and model uncertainty could be analyzed by a separated one. On the other hand, all kinds of uncertainty are analyzed together by single subjectively chosen likelihood measure in informal likelihood definition. Hence, parameter uncertainty in informal likelihood should be higher than formal likelihood case because all uncertainties are mingled in parameter uncertainty.

In terms of predictive uncertainty, we should compare the coverage ratio plot of the two cases. This could be accomplished by comparing figure 34 and figure 38. In the case of informal likelihood, the plot is not quite different between calibration period and validation period. However, for the case of formal likelihood, the difference is apparent, and most difference occurs at the higher significance region. Informal likelihood definition does not consider model error, hence coverage ratio derived by it should be entirely dependent on model performance. It makes the coverage ratio not quite different in both periods because model performance does not change along time. On the other hand, a formal likelihood incorporates error modelling. It implies the possibility that the error structure determined by analysis in the calibration period could not be accurate, which means that it could be inappropriate in another period. This inexactness of error analysis can introduce discrepancy of coverage ratio

in both periods. Finally, it is a certain fact that the uncertainty interval from formal likelihood definition is more consistent with a statistical significance level than the case of informal likelihood. Even though the coverage ratio in the validation period of formal likelihood definition does not match well with theoretical one compared to the calibration period, it is more consistent with theoretical one than the outcomes of informal likelihood definition. Eventually, we can conclude that the formal likelihood definition gives use more meaningful results in a statistical sense than informal likelihood case. However, in terms of interval size, formal likelihood definition reproduces a larger one than informal likelihood definition. This trend increases as the significance level becomes higher. This also attributes to the fact that formal likelihood definition incorporates the consideration of model error terms. Considering model error terms, interval will be larger than the case without considering it. This effect will increase as its significance level increases. Figure 42 shows this trend.

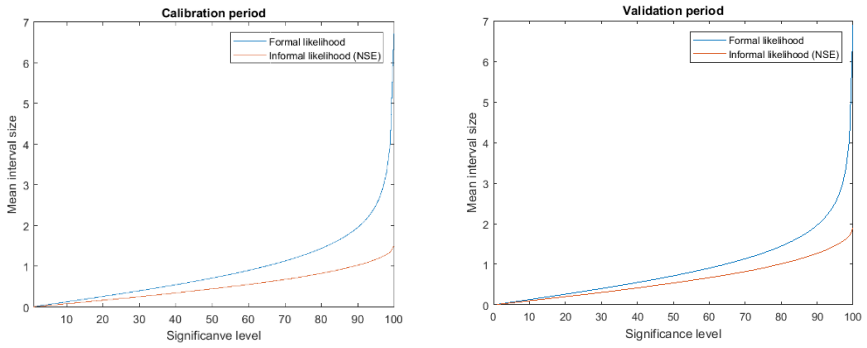

Figure 42. Comparison of mean interval size of two methods according to the corresponding significance level

Chapter 5. Summary and Conclusion

In this study, two likelihood definition approaches were adopted for investigation of GLUE methodology, namely informal and formal likelihood definitions. First of all, epistemic uncertainty was identified by finding a relationship between residuals and subcomponents of the variable of interest. After subtracting this epistemic uncertainty from the original residual, the new corrected residual was used for each approach. For informal likelihood definition, NSE and IEV were explored. For formal likelihood definition, the Bayesian approach was adopted, and the attribute of model error term was generalized to the non-normal, correlated, heteroscedastic (daily) situation. Different from other researches, non-normality of error term was defined through stable distribution , and correlation structure was modeled by ARFIMA(0,d,0) model, and heteroscedasticity was considered through the division of time period into its sub-periods.

For real application, SNU-WS was used for the hydrologic model, and target watershed was set as Springcreek watershed in Pennsylvania, USA. The code was implemented to SNU-WS for basic application of informal likelihood. Also, different from conventional GLUE methodology, the step for reducing identifiable epistemic uncertainty was added after the parameter sampling procedure. This was done by finding the statistical relationship between residual and subcomponents of the target variable (streamflow). Finally, an

uncertainty analysis was performed for monthly streamflow and daily streamflow considering 5 significant parameters.

The results from both time scales show that formal likelihood definition gives us more reliable uncertainty interval in its significance level than the other one. This fact is more obvious in high significance level than in low significance level and daily flow than monthly flow. This is because informal likelihood definition does not consider model error term. As time scale becomes reduced, and the significance level becomes higher, there may exist many extreme events to be covered in its uncertainty interval. Usually, an informal likelihood fails to cover these values because the coverage ratio of informal likelihood definition is limited to model efficiency. However, in cases of formal likelihood, the inability of the model to capture those values could be complemented by adding more estimated error terms to original model simulation results. So, it gives a reliable result for the corresponding significance level.

Furthermore, the benefit of adopting stable distribution was explained. Thinking about many nonlinear chaotic natural phenomena reproducing extreme events which cannot be fully described by the usual hydrologic model, the normal distribution is not an appropriate choice to model error. It should be generalized to stable distribution which includes various degrees of heavy-tailed distributions (Heaviness can be determined by parameter α). The uncertainty interval considering stable distribution shows high upper bounds for large significance level due to its description of more extreme events. It

means that we need a very large uncertainty interval if we want to cover most of the observation values, and stable distribution gives us a way to compute that interval quantitively according to its significance level. As compared to this, if we adopt normal distribution, we will get a smaller interval, but it does not have any meaning for covering extreme events because normal distribution is not heavy-tailed distribution.

There is also some disadvantage in formal likelihood definition caused by considering model error terms. That is the interval size problem. As explained before, to get a statistically consistent result, the uncertainty interval in a formal likelihood definition was larger than the one from informal likelihood. This reminds us of the idiom, “There ain’t no such thing as a free lunch”. We should give up small interval size to get a statistically consistent result.

Now, the choice is up to the user of GLUE methodology. We can choose informal likelihood definition to consider the values except for extreme ones and get smaller interval size. Or we can apply a formal likelihood definition to cover extreme values but have large intervals. However, in spite of these issues, if we consider that extreme events will occur more frequently in the future based on climate change context, it is more desirable to apply formal likelihood definition in real application of engineering purpose.

REFERENCES

Abbaspour, K. C. (2015). SWAT-CUP: SWAT Calibration and Uncertainty Programs - A User Manual.

Amy, G. (1974). *Water quality management planning for urban runoff.* US Environmental Protection Agency.

Barlow, M., & Clark, T. (2017). Blue gold: the battle against corporate theft of the world's water. Routledge.

Beven, K., & Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrological processes*, 6(3), 279-298.

Beven, K., & Binley, A. (2014). GLUE: 20 years on. *Hydrological processes*, 28(24), 5897-5918.

Beven, K., Smith, P. J., & Freer, J. E. (2008). So just why would a modeller choose to be incoherent? *Journal of hydrology*, 354(1-4), 15-32.

Bosen, J. F. (1960, 8 1). A FORMULA FOR APPROXIMATION OF THE SATURATION VAPOR PRESSURE OVER WATER. Washington D.C, Washington, U.S.

Chow, V. T. (1964). Handbook of applied hydrology.

Foster, G. R., McCool, D. K., Renard, K. G., & Moldenhauer, W. C. (1981). Conversion of the universal soil loss equation to SI metric units. *Journal of Soil and Water Conservation*, 36(6), 355-359.

Frenze, c. M. (2002). Visual Basic and Visual Basic.NET for scientists and Engineers.

Haith, D. A., Mandel, R., & Wu, R. (1992). Generalized Watershed Loading Functions Version 2.0 User's Manual. Department of Agricultural & Biological Engineering Cornell University.

Hamon, R. W. (1960). Estimating potential evapotranspiration. *Doctoral dissertation, Massachusetts Institute of Technology*.

Janicki, A., & Weron, A. (1993). *Simulation and chaotic behavior of alpha-stable stochastic processes (Vol. 178)*. CRC Press.

Lal , M., Mishra, S. K., & Kumar, M. (2019). Reverification of antecedent moisture condition dependent runoff curve number formulae using experimental data of Indian watersheds. *CATENA*, 173, 48-58.

Li, X., Weller, D. E., & Jordan, T. E. (2010). Watershed model calibration using multi-objective optimization and multi-site averaging. *Journal of Hydrology*, 380(3-4), 277-288.

Mantovan, P., & Todini, E. (2006). Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. *Journal of hydrology*, 330(1-2), 368-381.

Mirzaei , M., Huang, Y., Shatirah, A., & Ahmed, E.-S. (2015). Application of the generalized likelihood uncertainty estimation (GLUE) approach for assessing uncertainty in hydrological models: a review. *Stochastic environmental research and risk assessment*, 29(5), 1265-1273.

Nguyen, H., & Nguyen, V. (2014). Development of a GIS-Based Watershed Modeling Tool for Water Balance and Nutrient Loads. *Doctoral dissertation Seoul National University*.

Nystrom, E. A., & Burns, D. A. (2011). *TOPMODEL Simulations of Streamflow and Depth to Water Table in Fishing Brook Watershed, New York, 2007–09*. US Department of the Interior, US Geological Survey.

Oldham, K. B., & Spanier, J. (1974). *The fractional calculus theory and applications of differentiation and integration to arbitrary order (Vol. 111)*. Elsevier.

Overton, D. E., & Meadows, M. E. (1976). Simulating pollutographs and loadographs, stormwater modeling.

Qi, Z., Kang, G., Chu, C., Qiu , Y., Xu , Z., & Wang, Y. (2017). Comparison of SWAT and GWLF Model Simulation Performance in Humid South and Semi-Arid North of China. *Water*, 9(8), 567.

Richardson, C. W., Foster, G. R., & Wright, D. A. (1983). Estimation of erosion index from daily rainfall amount. *Transactions of the ASAE*, 26(1), 153-0156.

Roy , C. J., & Oberkampf, W. L. (2011). A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. *Computer methods in applied mechanics and engineering*, 200(25-28), 2131-2144.

Ruark, M. D., Niemann, J. D., Greimann, B. P., & Arabi, M. (2011). Method for assessing impacts of parameter uncertainty in sediment transport modeling applications. *Journal of hydraulic engineering*, 137(6), 623-636.

Schneiderman, E. M., Pierson, D. C., Lounsbury, D. G., & Zion, M. S. (2002). MODELING THE HYDROCHEMISTRY OF THE CANNONSVILLE WATERSHED WITH GENERALIZED WATERSHED LOADING FUNCTIONS (GWLF). *JAWRA Journal of the American Water Resources Association*, 38(5), 1323-1347.

Schoups, G., & Vrugt, J. A. (2010). A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. *Water Resources Research*, 46(10).

Singh, V. P. (2012). Computer Models of Watershed Hydrology. Water Resources Publications.

Smith, t., Sharma, A., Marshall, L., Mehrotra, R., & Sisson, S. (2010). Development of a formal likelihood function for improved Bayesian inference of ephemeral catchments. *Water Resources Research*, 46(12).

Stedinger, J. R., Vogel, R. M., Lee, S., & Batchelder, R. (2008). Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. *Water resources research*, 44(12).

Vallam, P., Qin, X. S., & Yu, J. J. (2014). Uncertainty quantification of hydrologic model. *APCBEE procedia*, 10, 219-223.

Winter, J. G., & Duthie, H. C. (2000). EXPORT COEFFICIENT MODELING TO ASSESS PHOSPHORUS LOADING IN AN URBAN WATERSHED1. *JAWRA Journal of the American Water Resources Association*, 36(5), 1053-1061.

Wischmeier, W. H., & Smith, D. D. (1978). Predicting rainfall erosion losses-a guide to conservation planning. *Predicting rainfall erosion losses-a guide to conservation planning*.

초록

수문 모델링과 기후 변화의 연구에 있어 불확실성 분석은 빼놓을 수 없는 중요한 주제이다. 일반적으로 수문 모형에 적용되는 불확실성 분석 방법 중에 GLUE라는 방법론이 있다. 하지만, 처음 제안되었던 GLUE 방법론은 우도 함수의 정의에 있어 사용자의 주관성이 크게 작용하여 통계적 신뢰성이 떨어진다는 점에서 다양한 비판을 받아왔다. 이러한 단점 때문에, 베이즈 이론에 근거한 형식적 우도 함수 정의에 관한 연구가 많이 이루어져왔다. 모델의 부정확성으로 인해 생기는 오차를 명시적으로 밝혀내는 과정은 이러한 연구들의 요체를 이루고 있다.

실제 적용에 있어서, 이러한 모델 오차는 일반적으로 비정규 분포를 따르고, 시간에 대해 독립적이지 않으며 이분산성을 보인다. 이러한 특성을 표현하기 위해 다양한 모델들이 차용되어져 왔다. 본 연구에서는 이전의 연구와 다른 방법을 차용하여 모델 오차를 묘사하고자 하였다. 안정 분포를 통해 비정규분포를 표현하였고 오차 시계열의 상관구조를 모의하기 위해 ARFIMA(0,d,0) 모델을 사용하였다. 또한, 일유량에 대해서는 기간을 더 작은 기간들로 나눔으로써 이분산성을 고려하였다. 또한, 이러한 오차구조의 분석 전에 회귀분석법을 사용하여 계산 가능한 인식론적 불확실성을 최대한 줄이는 과정을 거쳤다.

이러한 과정을 거쳐 계산된 형식적 우도함수와 전통적인 GLUE 방법론에서 사용되는 비형식적 우도함수의 정의에 따른 불확실성 분석 결과가 서로 비교되었다. 결과적으로, 형식적 우도함수를 통하여 구하여진 불확실성 분석 결과가 더욱 통계적으로 유의미한 것을 확인할 수 있었다. 그러나, 모델 오차의 고려로 인해 불확실성 구간이 일반적으로 크게 측정되는 단점이 있었다. 이에 비해, 비형식적 우도함수의 경우 높은 유의 수준을 가진 불확실성 구간일수록 통계적으로 무의미한 결과를 나타내었다. 이는 높은 유의 수준에 도달하기 위해 구간 안에 포함되어져야 할 다양한 극값들을 비형식적 우도함수로는 포함할 수 없기 때문인 것으로 분석되었다. 이에 반해, 형식적 우도 함수는 모델 오차를 고려하여 커진 구간을 통해 이러한 극값들을 잘 포함하였고 그에 상응하는 통계적으로 유의미한 결과를 보였다. 또한, 모델 오차를 묘사하기 위해 사용된 안정 분포가 어떻게 불확실성 구간이 극값들을 잘 포함하는데 기여하는지, 그리고 그 것이 기후변화 상황에서 어떤 의미를 가지는지 논의되었다.

주요어 : 불확실성 분석, GLUE 방법론, 형식적 우도함수, 비형식적 우도함수, 오차 구조

학번 : 2017-24832