Evaluation of alternative surface runoff accounting procedures using SWAT model

Haw Yen[1,2,3], Michael J. White[3], Jaehak Jeong[2], Mazdak Arabi[1], Jeffrey G. Arnold[3]

(1. *Department of Civil and Environmental Engineering, Colorado State University, Fort Collins, Colorado 80523, USA*;
2. *Blackland Research & Extension Center, Texas A&M Agrilife Research, Temple, Texas 76502, USA*;
3. *Grassland, Soil & Water Research Laboratory, USDA-ARS, Temple, Texas 76502, USA*)

Abstract: For surface runoff estimation in the Soil and Water Assessment Tool (SWAT) model, the curve number (CN) procedure is commonly adopted to calculate surface runoff by dynamically updating CN values based on antecedent soil moisture condition (SCSI) in field. From SWAT2005 and onward, an alternative approach has become available to apply the CN method by relating the runoff potential to daily evapotranspiration (SCSII). While improved runoff prediction with SCSII has been reported in several case studies, few investigations have been made on its influence to water quality output or on the model uncertainty associated with the SCSII method. The objectives of the research were: (1) to quantify the improvements in hydrologic and water quality predictions obtained through different surface runoff estimation techniques; and (2) to examine how model uncertainty is affected by combining different surface runoff estimation techniques within SWAT using Bayesian model averaging (BMA). Applications of BMA provide an alternative approach to investigate the nature of structural uncertainty associated with both CN methods. Results showed that SCSII and BMA associated approaches exhibit improved performance in both discharge and total NO_3 predictions compared to SCSI. In addition, the application of BMA has a positive effect on finding well performed solutions in the multi-dimensional parameter space, but the predictive uncertainty is not evidently reduced or enhanced. Therefore, we recommend additional future SWAT calibration/validation research with an emphasis on the impact of SCSII on the prediction of other pollutants.

Keywords: Soil and Water Assessment Tool (SWAT), curve number method, Bayesian model averaging, uncertainty analysis; hydrology, water quality

DOI: 10.3965/j.ijabe.20150803.833 Online first on [2015-03-03]

Citation: Yen H, White M J, Jeong J, Arabi M, Arnold J G. Evaluation of alternative surface runoff accounting procedures using SWAT model. Int J Agric & Biol Eng, 2015; 8(3): 54－68.

1 Introduction

Hydrological models are used extensively for water resources planning and management[1,2]. Among these models, the Soil and Water Assessment Tool (SWAT) is a widely used, semi-distributed river basin scale water quality and quantity model. The SWAT was developed to quantify the effects of land use change and other anthropogenic activities and has been applied to a variety of water resources investigations[3-5].

In the SWAT model, certain hydrologic and water quality processes can be represented using alternative

Received date: 2013-05-14 **Accepted date:** 2014-12-13

Biographies: Michael J. White, PhD, Agricultural Engineer, Research interests: agriculture, hydrology. Email: mike.white@ars.usda.gov. **Jaehak Jeong**, PhD, Assistant Professor, Research interests: hydrology, modeling. Email: jjeong@brc.tamus.edu. **Mazdak Arabi**, PhD, Assistant Professor, Research interests: Agriculture, Hydrology. Email: mazdak.arabi@colostate.edu. **Jeffrey G. Arnold**, PhD, Agricultural Engineer, Research interests: agriculture, hydrology. Email: jeff.arnold@ars.usda.gov.

***Corresponding author: Haw Yen**, PhD, Assistant Research Scientist, Research interests: modeling, optimization, uncertainty analysis. Department of Civil and Environmental Engineering, Colorado State University, Fort Collins, Colorado 80523, USA; Blackland Research & Extension Center, Texas A&M Agrilife Research, Temple, Texas 76502, USA; Grassland, Soil & Water Research Laboratory, USDA-ARS, Temple, Texas 76502, USA. Email: hyen@brc.tamus.edu; Phone: +1-(254) 774-6004.

algorithms which may be more representative of local conditions. For example, surface runoff may be estimated using either of two following methods: the NRCS curve number (CN) procedure[6] and the Green & Ampt infiltration method[7]. The Green & Ampt method is not as commonly applied as the CN method because it requires more detailed climate observations with shorter time intervals such as hourly or sub-hourly time step. The CN method accounts for short-term losses of precipitation such as canopy interception, depression storage, and infiltration, but may not properly reflect local hydrologic properties in long-term simulations (e.g., evapotranspiration)[8]. The retention parameter of CN method has been calculated using antecedent soil moisture conditions[9] in accordance with the original CN procedure[10]. This approach has been utilized in earlier versions of the SWAT before SWAT2005 and widely applied. However, this method tends to overestimate surface runoff in shallow soils when the storage condition is low[11-15].

An alternate approach that may provide improved output under such condition is to relate the CN retention parameter to daily evapotranspiration, which was first included in SWAT2005[3,4] and is an option available in all subsequent SWAT versions. By relating the daily CN value to plant evapotranspiration, the daily CN values tends not to excessively dependent on soil moisture and more dependent on antecedent climate conditions. To date, hundreds of SWAT studies have been conducted which used the CN antecedent soil moisture method[3,10]. However, relatively few SWAT studies have been performed using the CN approach based on accumulated plant evapotranspiration[9,11-12,16-17] and even fewer studies compared SWAT hydrologic performance[9,11-12] and water quality responses[12] between the two CN approaches. Water quality is largely driven by the quantity and timing of surface runoff and both of which are dependent upon the way CN values are estimated.

In addition, the new CN method may lead to increased uncertainty in the model output due to added complexity in the structure of the model, which is called structural uncertainty. Structural uncertainty is generally caused from the discrepancy between a mathematical model and the true system of a real life situation that the model describes. As most mathematical expressions almost always approximate the reality, structural uncertainty is intrinsic to all simulation models including SWAT. Standard statistical practices are generally insufficient to quantify the uncertainty in a model often leading to over-confident decisions. Bayesian model averaging (BMA)[18] is a recently developed technique that offers a systematic method for analyzing model uncertainty and checking the robustness of one model's results to alternative model specifications. Since a direct comparison between SCSI and SCSII may not be sufficient to identify the advantages and disadvantages of the methods, the BMA technique can be a sophisticated tool that provides an alternative approach or strategy to investigate the nature of structural uncertainty associated with these CN methods.

The overall goal of this study is to evaluate the role of structural uncertainty of the SWAT model attributable to these different CN methods on hydrologic and water quality predictions. Specifically, the following objectives are defined: (1) Characterization of the differences in SWAT performance on predicting runoff and water quality with the two different CN methods; and (2) Quantification of the improvement in the accuracy of SWAT output enhanced by utilizing a combination of the CN methods for differing climatic conditions. This study implements BMA to integrate hydrologic and water quality predictions by BMA weights and to characterize structural uncertainty associated with the different approaches on updating daily CN values.

2 Materials and methods

2.1 Site location and characteristics

The Eagle Creek watershed (ECW) is a 248 km^2 sub-watershed within the Upper White River Basin located in Boone, Hamilton, Hendricks and Marion counties in central Indiana (Figure 1). The average annual precipitation is between 960 and 1 020 mm; the lowest average temperature (-3°C) occurs in January and the highest average temperature is 24°C occurring in July[19]. The dominant land use is agricultural (59%), followed by rangeland (38%), forest (2%), and urban (1%).

Predominant soils include thin loamy deposits over loamy glacial till that are primarily classified as hydrologic group B (51%) and C (48%). In the headwater subbasins, soils are generally deep and poorly drained. Most soils are highly productive unless they are poorly drained.

Key data sources used for the study included a 30-m resolution Digital Elevation Model (DEM) obtained from the U.S. Geological Survey (USGS) National Elevation Dataset[20], the National Agriculture Statistics Service (NASS) Cropland Data Layer 2000-2003[21] for cropland inputs, and Soil Survey Geographic (SSURGO) soil data from National Resources Conservation Service (NRCS)[22]. These data were used in the delineation of the ECW in SWAT, which resulted in 35 subbasins and 446 hydrologic response units. Daily precipitation and the maximum and minimum temperature data for 12 years (1999-2010) at three locations (COOPID: 413943 in Harlingen, COOPID: 415836 in Mercedes, and COOPID: 419588 in Weslaco) were downloaded from the United States National Climatic Data Center[23]. Streamflow data are available at a single site (station 35, USGS Gauge #03353200) and periodic (approximately monthly) grab samples of nitrate concentration were also available from four additional monitoring stations 20, 22, 27, and 32 (Figure 1).

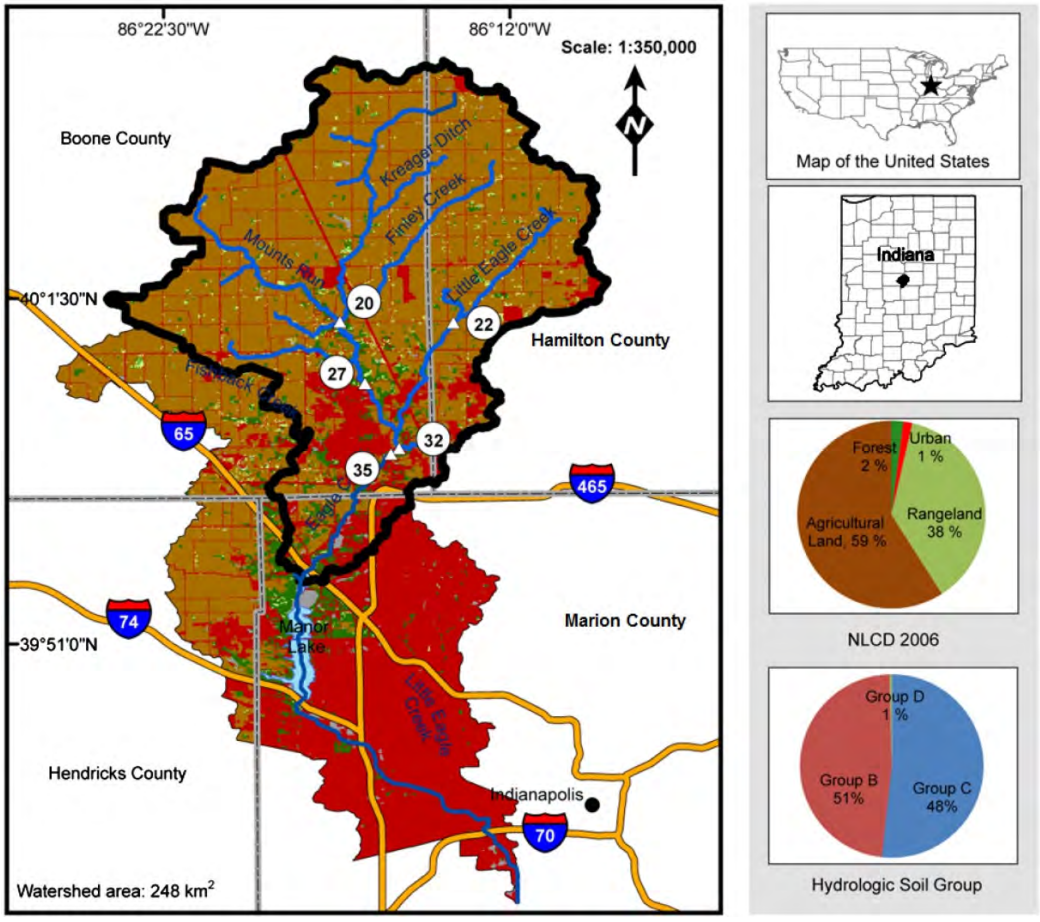


Figure 1 Case study area: Eagle Creek Watershed, Indiana

2.2 SWAT model

The SWAT model (version 2012; rev547)[1] is a continuous-time and semi-distributed watershed simulation model which is able to simulate hydrologic and water quality processes at the watershed level, and has been applied extensively in the field of watershed management[17,24-28] along with various modifications in structure[29]. A detailed description of the SWAT model is presented in the published study of Arnold, et al[30].

The water balance in SWAT is simulated based on the Equation (1):

$$S_t = S_0 + \sum_{i=1}^{t}(P_i - Q_i^{surf} - E_i - W_i^{seep} - Q_i^{gw}) \quad (1)$$

where, S_t is the final soil water content (H_2O, mm); S_0 is the initial soil water content at time step i (H_2O, mm); P_i

is the amount of precipitation at time step i (H_2O, mm); Q_i^{surf} is the amount of surface runoff at time step i (H_2O, mm); E_i is the amount of evapotranspiration at time step i (H_2O, mm); W_i^{seep} is the amount of water entering vadose zone from soil profile at time step i (H_2O, mm); Q_i^{gw} is the amount of return flow at time step i (H_2O, mm). In SWAT2005 and later versions of SWAT, the Green-Ampt method and the two other previously mentioned CN methods are available to calculate the amount of Q_i^{surf}.

2.3 SCS curve number procedure

The SCS curve number procedure[4] is designed to estimate surface runoff based on land use, hydrologic condition, and soil type[31]. Cumulative surface runoff is calculated by:

$$Q = \frac{(P - I_a)^2}{(P - I_a + S)} \tag{2}$$

where, Q is cumulative surface runoff (mm); P is cumulative precipitation (mm); I_a is cumulative initial abstraction, that is, canopy interception (mm) and depression storage (mm) and S is the retention parameter (mm). The initial abstraction is commonly approximated as 20% of the retention parameter ($0.2S$)[25]. Thus, Equation (2) can be re-written as:

$$Q = \frac{(P - 0.2S)^2}{(P + 0.8S)} \tag{3}$$

From Equation (3), the value of surface runoff is controlled by precipitation and the retention parameter. The retention parameter can be calculated as:

$$S = 25.4\left(\frac{1000}{CN} - 10\right) \tag{4}$$

where, CN is the empirically determined curve number associated with a given land use, slope, soil type, and cover condition, and in SWAT may be adjusted according to antecedent soil moisture condition or plant evapotranspiration. The optional calculation of the retention parameter from plant evaporation became available in SWAT2005[9].

2.3.1 Antecedent Soil Moisture Condition (SCSI)

The original SCS approach assumes an antecedent soil moisture condition (AMC) of II (average conditions). Curve numbers for antecedent soil moisture conditions I (dry) or III (wet) can be adjusted from reported AMC values as follows:

$$CN_1 = CN_2 - \frac{20(100 - CN_2)}{100 - CN_2 + \exp[2.533 - 0.0636(100 - CN^2)]} \tag{5}$$

$$CN_3 = CN_2 \exp[0.00673(100 - CN_2)] \tag{6}$$

where, CN_1 is the curve number of soil moisture condition I; CN_2 is the curve number of soil moisture condition II; and CN_3 is the curve number of soil moisture condition III. The definition of the dry condition is defined to be the wilting point; the average condition is defined to be the average moisture condition; and the wet condition is defined as the filed capacity which can be found from the SWAT manual of theoretical documentation[32].

Previous research indicates that the use of the AMC technique over predicts surface runoff results in shallow soils (<500 mm depth) with low total water storage[33]. Therefore, a new approach was proposed to address this problem by including a simple structure one-parameter module based on plant evaporation, which is discussed in the following section.

2.3.2 Plant Evapotranspiration (SCSII)

This method utilizes plant evapotranspiration to update the retention parameter (S_t) in place of antecedent soil moisture condition to adjust the CN value. By calculating the daily CN value as a function of plant evapotranspiration, this approach better reflects previous climatic conditions and less reliant on soil water storage, and has been shown to improve simulation of hydrologic processes in shallow soils[9,11] as follows:

$$S_t = S_{t-1} + PET_t \cdot \exp\left[\frac{-CNCOEF \cdot S_{t-1}}{S_{\max}}\right] - P_t + Q_t \tag{7}$$

where, S_t is the retention parameter at time; $S_{t\text{-}1}$ is the retention parameter at time; $S_{\max}$ is the maximum value of the retention parameter of all S_t; PET_t is the potential evapotranspiration for the given time t; $CNCOEF$ is the depletion coefficient; P_t s the precipitation depth for the given time and Q_t is the surface runoff[9]. The daily curve number adjusted for moisture content is calculated by rearranging Equation (4) and inserting the retention parameter, S_t, calculated for that moisture content. Obviously, the CN values for dry and wet soil moisture

conditions are not updated here because these parameters are only used in SCSI. To apply SCSII, users have to predefine or calibrate the value for the CNCOEF input parameter before implementing SCSII.

2.4 Parameter estimation procedure

Model performance during calibration and validation was assessed with statistically valid likelihood functions[34] used as objective functions and based on other related criteria established by Moriasi, et al[35]. The appropriate selection of likelihood function is critical for statistically sound validation of model performance[36] and the use of established performance criteria allows for a more robust assessment of model performance utilizing multiple objective functions such as Nash & Sutcliffe efficiency (NSE) and percent bias (PBIAS)[35].

2.4.1 Likelihood function

For a watershed model M with a vector of p parameters within the feasible parameter space (θ) that simulates the response vector of the watershed ($\hat{y}$), the discrete stochastic time-series vector of model residuals is:

$$\varepsilon(\theta)=\hat{y}-y=M(\theta)-y \quad \theta\in\Theta\subset\Re^n \tag{8}$$

The application of the first-order autoregressive (AR-1) transformation of the residuals can be used to account for correlated errors:

$$\varepsilon_i=\rho\varepsilon_{i-1}+\vartheta_i \quad i=1,\cdots,n \tag{9}$$

where, ρ is the lag-1 serial correlation coefficient for the residuals ε; $\vartheta \sim N(0,\sigma_\vartheta^2)$ is the innovation term with zero mean and constant variance σ_ϑ^2 and n is the total number of observed data. A proper likelihood function for multiple variables case is expressed as follows as proposed by Ahmadi[34]:

$$\ell(\varepsilon\,|\,\theta)=\sum_{j=1}^{m}\left\{-\frac{n_j}{2}\ln(2\pi)-\frac{1}{2}\ln\frac{\sigma_{\vartheta,j}^{2n_j}}{1-\rho_j^2}-\frac{1}{2}(1-\rho_j^2)\cdot\right.$$
$$\sigma_{\vartheta,j}^{-2}\cdot[\hat{y}_{1,j}(\theta)-y_{1,j}]^2-\frac{1}{2}\sigma_{\vartheta,j}^{-2}\cdot\sum_{i=2}^{n_j}\{(y_{i,j}-\rho_j y_{i-1,j})-$$
$$\left.[\hat{y}_{1,j}(\theta)-\rho_j\hat{y}_{i-1,j}(\theta)]\}^2\right\} \tag{10}$$

where, j is the number of variables; m is the total number of variables; n_j is the total number of observed data for variable j; $\sigma_{\vartheta,j}$ is the standard deviation of residuals for variable j; ρ_j is the lag-1 serial correlation coefficient for the residuals for variable j; The terms $\sigma_{\vartheta,j}$ and ρ_j can be estimated using the Bayesian approach[37] or can be assigned based on prior knowledge. The value of the likelihood function above will be minimized for all approaches in case studies.

2.4.2 Dynamically dimensioned search

Dynamically dimensioned search (DDS)[38] is an automatic calibration algorithm designed for the purpose of solving high dimensional problems. It has been shown that DDS has outstanding performance compared to other commonly used parameter estimation techniques[39] in different categories. In this study, DDS is adopted as the sampling technique to explore the role of structural uncertainty attributable to CN methodology in the SWAT model. Applications of DDS incorporated with SWAT can be found in literatures[5,15,40].

2.4.3 Model performance validation criteria

During calibration, proposed parameter sets (candidate parameter sets) and related model predictions are compared to observed data using standard statistical metrics to adjudge performance. In this study, the general performance ratings (GPR)[35], shown in Table 1, were used. These criteria are used to identify parameter sets with acceptable performances (e.g., Very good, Good, and Satisfactory) by statistical standards. However, methods which can generate more parameter sets with specific success rate do not guarantee better overall performance. The number of qualified parameter sets represents only the parameter sets with better manually assigned statistical (it can also be defined by many other different ways[41], thresholds and there is still chance that global optimal solution is located outside the region that satisfies additional validation criteria such as GPR. The purpose of conducting GPR is to take it as a supplementary guidance to have a better sense in evaluating output performance during model calibration.

Table 1 General performance ratings

Performance Rating	NSE	PBIAS/%	
		Streamflow	NOX
Very good	0.75 < NSE ≤ 1.00	PBIAS < ±10	PBIAS < ±25
Good	0.65 < NSE ≤ 0.75	±10 ≤ PBIAS < ±15	±25 ≤ PBIAS < ±40
Satisfactory	0.50 < NSE ≤ 0.65	±15 ≤ PBIAS < ±25	±40 ≤ PBIAS < ±70
Unsatisfactory	NSE ≤ 0.50	PBIAS ≥ ±25	PBIAS ≥ ±70

Note: NSE: Nash-Sutcliffe efficiency coefficient; PBIAS: Percent bias; NOX: Total nitrate; Statistical standards are taken from Moriasi, et al[35].

2.5 Bayesian model averaging (BMA)

The BMA[16,42] technique is a standard framework developed to combine models and predictive distributions[43-45]. According to the law of total probability, the posterior distribution of N different models with given data of observation Y can be written as[14]:

$$P(\Delta \mid Y) = \sum_{n=1}^{N} P(\Delta \mid M_n, Y) P(M_n \mid Y) \quad (11)$$

where, Δ is the quantity of prediction; M_n (n=1, 2, 3, …, N) is the ensemble of implemented model predictions; $P(\Delta|M_n, Y)$ is the posterior probability of M_n (assume it is correct for the training data[46] which reveals the manner how M_n fits the training data; $P(\Delta|M_n, Y)$ is the forecast posterior distribution of Δ given prediction quantities from model M_n and observation data Y. As shown in Equation (12), the term of posterior probability $P(M_n|Y)$ sums to one:

$$\sum_{n=1}^{N} P(M_n \mid Y) = \sum_{n=1}^{N} w_n = 1 \quad (12)$$

where, w_n is the posterior probability of prediction (the one with best solution). Therefore, the posterior probability of prediction can be regarded as weights which represent the contribution of each in favor of predictions. Recently, the BMA has been extended to ensembles of dynamic models where the forecast f_n is associated with a conditional probability distribution function (PDF)[44]. The BMA predictive model can be expressed as Equation (13) where $g_n(\Delta|f_n)$ is the conditional PDF of Δ given f_n.

$$P(\Delta \mid f_n, \cdots, f_N) = \sum_{n=1}^{N} w_n g_n(\Delta \mid f_n) \quad (13)$$

The assumption of the original form of BMA[44] suggests that the conditional PDF can be approximated by a linear function which is normal distribution centered. As shown in Equation (14), the mean of a normally distributed PDF is $a_n+b_nf_n$ with standard deviation.

$$\Delta \mid f_n \sim N(a_n + b_n f_n,\ \sigma^2) \quad (14)$$

From above, the BMA mean and variance can be written as follows[36]:

$$E(\Delta \mid f_n, \cdots, f_N) = \sum_{n=1}^{N} w_n (a_n + b_n f_n) \quad (15)$$

$$Var(\Delta \mid f_n, \cdots, f_N) = \sum_{n=1}^{N} w_n [(a_n + b_n f_n) - \sum_{i=1}^{N} w_i (a_i + b_i f_i)]^2 + \sum_{n=1}^{N} w_n \sigma_n^2 \quad (16)$$

In Equation (13), it is concerned in cases of discharge and water quality (both observation and simulation error are non-Gaussian) calibration that the BMA method is assumed to have conditional probability distribution to be Gaussian. Data transformation is required to properly perform the BMA procedure. Therefore, the log-likelihood function of Equation (13) is as follows (assume independence of forecast errors in time and space):

$$L(w_1, \cdots, w_N, \sigma^2) = \log[\sum_{n=1}^{N} w_n g_n(\Delta_{st} \mid f_{nst})] \quad (17)$$

The values of the weights and variance can be derived by applying maximum likelihood estimation through various optimization techniques such as the Expectation-Maximization (EM) algorithm or the Shuffle Complex Evolution Metropolis (SCEM-UA) algorithm[44-45,47]. In this study, Equation (17) will be solved by the EM algorithm.

2.6 Brier score

In this study, the brier score (BS) is implemented to compare performance across different models and scenarios[48,49]. The BS is a quantified scalar measure of model simulation/forecast and has been widely applied in multi-model topics. The BS is used for the evaluation of various models which generate same kind of outputs such as flow. For example, in the work done by Ajami, et al.[48], the BS was implemented to evaluate the performance of the Sacramento Soil Moisture Accounting Model (SAC-SMA)[50], Hydrologic model (HYMOD)[51], and Simple Water Balance Model (SWB)[52]. Users can take advantage of the BS as an index of reference to evaluate the performance among various models which generate same kind of outputs (e.g., flow). The base function of BS is written as follows[49]:

$$BS = 1 - \frac{1}{N} \sum_{i=1}^{N} [f(i) - o(i)]^2 \quad (18)$$

where, $f(i)$ is the frequency of the simulated target event at time step i estimated by the fraction of model simulations which satisfy (larger than) predefined threshold of Q^*; $o(i)$ equals to 1 if observed quantities at time step i are greater than Q^*, otherwise $o(i)$ is zero. Greater values of BS are preferred.

2.7 ECW case study scenarios

The ECW SWAT model is optimized for streamflow and NO_3 under five different calibration scenarios. In

scenario 1, the model is set to use the SCSI method while the SCSII method is used in scenario 2. In scenarios 3 to 5, the BMA technique is used for posterior output optimizations based on different configuration of weighting factors. Scenario 3 uses the best results obtained in scenarios 1 and 2 to combine an improved result using a single weighting factor for each output of the two scenarios. Scenario 4 expands the same concept to consider any seasonal variation in stream flow by splitting the year into 6 months of dry season and the other six months of wet season represented by two sets of weighting factors. In scenario 5, the division of wet and dry seasons is combined with cold and hot weather and dynamically updated every month based on 50-year measured weather data. An automated calibration technique (DDS) was employed in all scenarios to find an objective function that yielded the best performance criteria, or the optimal solutions.

Streamflow is calibrated on a daily basis at one gauge station (station 35, Figure 1) and a monthly basis for nitrate at four water quality monitoring locations (station 20, 22, 27, 32). Calibration and validation periods are from 1997-2000 and 2001-2003, respectively. Computational time for each model validation for a total of 10 000 iterations ranged from 450 to 500 hours on an Intel® Core™ 2 Duo CPU E8400@3.00 GHz, 32-bit system operated with Microsoft Windows XP. Model simulations are assessed using an objective function (statistically valid likelihood function) and multi-criteria model validation standards.

2.7.1 Scenario 1: SCSI calibration

In Scenario 1, the SWAT model utilizing SCSI was calibrated using 28 parameters (as shown in Table 2, total number of parameters under calibration: 28) where the CN is calculated by embracing information from antecedent soil moisture conditions. In the SWAT settings, the system parameter ICN (selection of daily curve number calculation method) should be set to the value of 0.

2.7.2 Scenario 2: SCSII calibration

In Scenario 2, the SWAT model calibration is using the approach of SCSII for surface runoff calculation where the plant evapotranspiration is applied (as shown in Table 2, total number of parameters under calibration: 29). In the SWAT settings, the system parameter ICN was set to a value of 1 (this method is available from the version of SWAT2005) with an extra model parameter CNCOEF included.

Table 2 Adjusted SWAT parameters for all case scenarios

Parameter	Description	Range
CDN	Denitrification exponential rate coefficient	0-1
CMN	Rate factor for humus mineralization of active organic nutrients	0.0003-0.03
NPERCO	Nitrogen percolation coefficient	0.01-1.00
SDNCO	Denitrification threshold water coefficient	0-1
SFTMP	Snowfall temperature/℃	±5
SMFMN	Melt factor for snow on 21 December/ (mm $H_2O \cdot d^{-1} \cdot ℃^{-1}$)	0-10
SMTMP	Snowmelt base temperature/℃	±5
SNOCOVMX	Minimum snow water content that corresponds to 100% snow cover/mm	0-650
SNO50COV	Snow volume fraction denoted by 50% snow cover of SNOCOVMX/mm	0.01-0.99
SURLAG	Surface runoff lag coefficient/d	0-12
SOL_ORGN	Initial organic nitrogen concentration in the soil layer/(mg $N \cdot kg^{-1}$)	1-10 000
ALPHA_BF	Baseflow alpha factor/d	0-1
GW_DELAY	Groundwater delay time/d	0-60
GW_REVAP	Groundwater "revap" coefficient	0.02-0.2
GW_SPYLD	Specific yield of the shallow aquifer/%	±50
GWQMN	Threshold water depth in the shallow aquifer required for return flow to occur/mm	0-5 000
ESCO	Evaporation compensation coefficient	0-1
CANMX	Evaporation compensation coefficient/mm	0-10
OV_N	Manning's "n" for overland flow	0.01-0.6
DEP_IMP	Depth to impervious layer in soil profile/mm	1 500-2 500
CN_F	SCS runoff curve number/%	±10
DDRAIN	Depth to subsurface drain/mm	500-1 500
CH_K2	Effective hydraulic conductivity in main channel alluvium/($mm \cdot h^{-1}$)	0-500
CH_N2	Manning's "n" value for the main channels	0.01-0.3
SOL_AWC	Soil available water capacity/%	±20
SOL_K	Saturated hydraulic conductivity /%	1-1 000
CH_K1	Effective hydraulic conductivity in tributary channel alluvium/($mm \cdot h^{-1}$)	0-300
CH_N1	Manning's "n" value for the tributary channels	0.01-0.3
CNCOEF*	Plant ET curve number coefficient	0-2

Note: *: CNCOEF is only applied for SCSII; **: Parameter values of GW_SPYLD, CN_F and SOL_AWC are the changes of fraction from default values.

2.7.3 Scenario 3: Application of universal BMA weights (BMAI)

The BMA will be applied in scenarios 3, 4 and 5 where the results of scenario 1 and 2 are aggregated by BMA weights using different posterior optimization schemes as follows. In scenario 3, the same set of BMA weights are assigned for all time series for each gauge station. As shown in Table 3, the BMA weights optimized for SCSI (i.e., scenario 1) and SCSII (i.e., scenario 2) appear to be

highly inclined to SCSII result in most scenarios and at most gage stations, which implies that the model output based on SCSII is preferred over the SCSI-based output by the BMA model.

Table 3 BMA model weights for scenario 3, 4, and 5 (daily streamflow: station 35; monthly NO_3: station 32, 27, 22, 20)

BMAI	ω @ st.35	ω @ st.32	ω @ st.27	ω @ st.22	ω @ st.20
SCSI	0.5237	0.0204	0.0001	0.0004	0.2086
SCSII	0.4763	0.9796	0.9999	0.9996	0.7914
BMAII Dry Season	ω @ st.35	ω @ st.32	ω @ st.27	ω @ st.22	ω @ st.20
SCSI	0.8636	<0.0001	0.0002	<0.0001	0.0007
SCSII	0.1364	>0.9999	0.9998	1.0000	0.9993
BMAII Wet Season	ω @ st.35	ω @ st.32	ω @ st.27	ω @ st.22	ω @ st.20
SCSI	0.2688	<0.0001	0.0006	0.0007	0.1879
SCSII	0.7312	>0.9999	0.9994	0.9993	0.8121
BMAIII Wet-Warm	ω @ st.35	ω @ st.32	ω @ st.27	ω @ st.22	ω @ st.20
SCSI	0.5204	0.0002	0.0004	0.0006	0.1029
SCSII	0.4796	0.9998	0.9996	0.9994	0.8971
BMAIII Dry-Warm	ω @ st.35	ω @ st.32	ω @ st.27	ω @ st.22	ω @ st.20
SCSI	0.5629	0.2415	0.3381	0.3174	0.3526
SCSII	0.4371	0.7585	0.6619	0.6826	0.6474

Note: ω: BMA weight (BMA weights in each gauge station is summed to one); st.: gauge station number shown in Figure 1.

2.7.4 Scenario 4: Application of wet/dry seasonal BMA weights (BMAII)

In scenario 4, the whole year is divided into a 6-month dry season and the other 6-month wet season based on the 55-year precipitation data (1950-2004) downloaded from the National Climatic Data Center[53]. Months with relatively more precipitation will be grouped as wet season (from March to August) and the rest of the year is defined as dry season (from September to the next February). Streamflow responses were assumed to show different characteristics during these dry and wet seasons, which are represented by two sets of BMA weights in scenario 4. This approach is designed to give the BMA model a flexibility to capture any seasonal variation in the streamflow characteristics while combining the SCSI and SCSII outputs.

2.7.5 Scenario 5: Application of warm/cool and wet/dry seasonal BMA weights (BMAIII)

Instead of using comparative quantity of precipitation to categorize different characters of dry/wet seasons, another approach is also implemented to classify weather conditions. The historical median of the temperature and precipitation data (1950-2004) is set as thresholds to evaluate if it is under the first classification of warm or cool conditions in temperature and the second classification of wet or dry conditions. In developing scenario 5, only wet-warm and dry-warm conditions were found to be significant because the temperature for the calibration period is above historical median constantly. The BMA weights for Scenario 5 are shown in Table 4.

Table 4 Success rate of parameter sets satisfy the performance criteria of all case studies in calibration and validation periods

Success rate/%	SCSI	SCSII	BMAI	BMAII	BMAIII
Calibration	67.19	75.18	75.3	75.93	76.37
Validation	0.30	3.01	2.56	1.31	1.04

Note: Success rate in Table 4 is applying the "Satisfactory" category in GPR.

3 Results

The effectiveness of the SCSI and SCSII methods as constructed in the five different scenarios was evaluated based on the following criteria: the convergence speed of objective functions, the percentage of model simulations meeting acceptable model performance criteria, and the BMA weights assigned to the CN calculation technique. Structural uncertainties in these CN methods were assessed using the fraction of observed data within the 95% confidence interval and the average width of uncertainty bands. Predictive uncertainty under the five approaches for calibration and validation periods was estimated by directly comparing the predicted and observed values in quantile-quantile (QQ) plots.

To characterize potential improvements in model predictions associated with the CN methods, we first tested model's ability to achieve better objective function values. Then, a GPR was implemented to compensate the disadvantage that the objective function may not be able to determine representative watershed behavior according to only one aggregated term. Third, the BMA technique was implemented through the calibration and validation periods to get corresponding statistical results. The BS was applied to evaluate the performance among different methods. Finally, an uncertainty analysis was performed to investigate the predictive uncertainties in all case scenarios.

3.1 Objective function validation

The advantage of the likelihood function as described in Equation (10) is that it provides a statistically proven objective function that incorporates multiple variables (in this study, they include streamflow and NO_3). The

objective function tends to decrease as results improve and asymptotically approaches to zero as the optimization proceeds and the result improves.

The rate at which objective function values under SCSI and SCSII were reduced is shown in Figure 2. The SCSI method resulted in a slightly better overall model performance and converged more quickly, indicating slightly more computationally efficient calibration using SCSI. After 4 000 model validations no significant progress was made for either method. Final objective function values were similar for both methods (SCSI: 342.1; SCSII: 343.1).

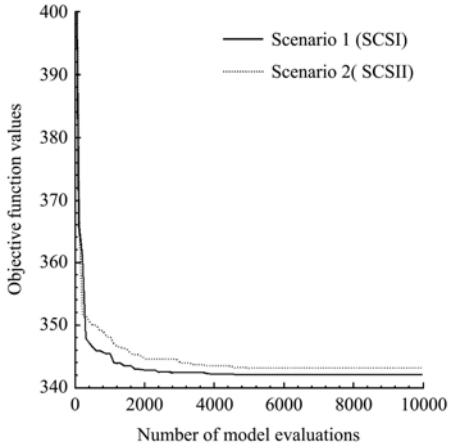


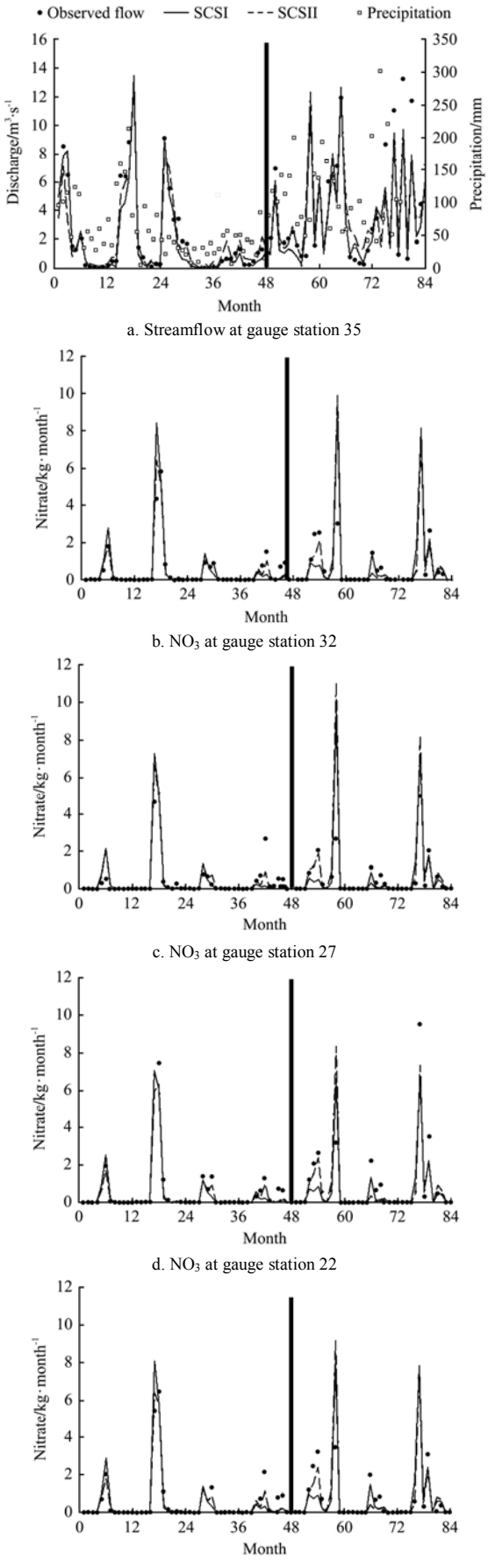
Figure 2 Overall performance of objective function versus model validations of scenario 1 (apply SCS method by antecedent soil moisture condition) and scenario 2 (apply SCS method by plant evapotranspiration)

3.2 Performance criteria

The percentage of parameter sets rated as "Satisfactory" or better is given in Table 4. All CN methods resulted in a relatively high success rate in the calibration period (67% to 76%) but not in the validation period (1% to 3%). This was primarily due to a severe drought (as shown in Figure 3a) in the last two years of calibration (1999 and 2000) where the validation period was not well predicted by SWAT using any CN methods.

3.3 BMA statistics

Monthly streamflow (streamflow is calibrated and validated in daily basis but shown monthly) and NO_3 are shown in Figure 3, and the statistical results are summarized in Table 5 based on NSE and PBIAS. In most cases as presented in Table 5, scenarios 2, 3, 4 and 5 generally perform better compared to scenario 1 (SCSI) in both calibration and validation periods. Therefore, the



a. Streamflow at gauge station 35

b. NO_3 at gauge station 32

c. NO_3 at gauge station 27

d. NO_3 at gauge station 22

e. NO_3 at gauge station 20

Figure 3 Observed and simulated (best results of 10 000 model runs) time series for streamflow discharge rate and NO_3 loads using SCSI and SCSII during the calibration (month 1 to 48) and validation (month 49 to 84) periods

BMA model assigned higher weights to scenario 2 to make better composite output of simulated data in scenarios 3, 4, and 5 (Table 3). The BMA posterior processor puts higher weights onto a reference output closer to the observation than onto an output that is farther off from the measured values. Since the SCSII method gave better output, all BMA scenarios tended to converge to similar statistics to that of SCSII.

Table 5 Results of statistics for calibration and validation period in all scenarios (daily streamflow: station 35; monthly NO_3: station 32, 27, 22, 20)

Period	Scenarios	Percent bias (PBIAS, %)				
		st.35	st.32	st.27	st.22	st.20
Calibration (1997-2000)	SCSI	11.94	-6.27	7.69	0.22	21.08
	SCSII	3.39	2.09	13.75	-5.37	21.43
	BMAI	8.16	2.87	14.6	-3.8	22.2
	BMAII	7.85	1.90	13.75	-5.37	21.20
	BMAIII	7.99	2.09	13.75	-5.37	21.20
Validation (2001-2003)	SCSI	14.18	-23.09	-1.60	-32.02	20.54
	SCSII	4.26	-23.73	-5.76	-55.94	5.91
	BMAI	9.69	-23.8	-5.58	-54.6	8.1
	BMAII	9.40	-23.69	-5.72	-55.94	8.92
	BMAIII	9.16	-23.73	-5.76	-55.94	6.77
Period	Scenarios	Nash-Sutcliffe Efficiency (NSE)				
		st.35	st.32	st.27	st.22	st.20
Calibration (1997-2000)	SCSI	0.92	0.54	0.78	0.62	0.93
	SCSII	0.91	0.85	0.93	0.76	0.95
	BMAI	0.94	0.85	0.93	0.76	0.95
	BMAII	0.94	0.85	0.93	0.76	0.95
	BMAIII	0.94	0.85	0.93	0.76	0.95
Validation (2001-2003)	SCSI	0.81	0.30	0.32	-1.96	0.56
	SCSII	0.84	0.26	0.42	-1.06	0.66
	BMAI	0.84	0.26	0.40	-1.09	0.56
	BMAII	0.84	0.27	0.42	-1.08	0.57
	BMAIII	0.84	0.26	0.42	-1.08	0.60

Note: st.: gauge station number shown in Figure 1.

3.4 Brier score model performance

The BS for SCSI, SCSII and BMA associated methods is shown in Figure 4. For scoring, the entire flow regime was divided into ten intervals with one representing low flow and ten representing high flow. A score value near one represents good matching between predicted and observed. Figure 4a shows the BS evaluated on streamflow, and Figure 4b depicts the score on NO_3 at the station 32. As noted earlier, either SCSI or SCSII performs unsatisfactorily in the low flow regime. The SCSI is demonstrated to better perform in predicting stream flow in the intermediate flow regime, but SCSII presents a better performance in the low flow regime. The BMAII scenario combined a significantly improved result on predicting streamflow, especially in the intermediate and high flow regimes. In Figure 4b, All BMA scenarios and the SCSII scenario results are almost identical making it difficult to identify the results in the figure. In case of NO_3 prediction, SCSII outperformed SCSI and the BMA scenarios successfully took the good result of SCSII output rather than from SCSI. The BS for other nitrate stations demonstrate similar results as that in Figure 4b which are not shown here (SCSII and BMA associated methods present superior forecast in simulating nitrate process).

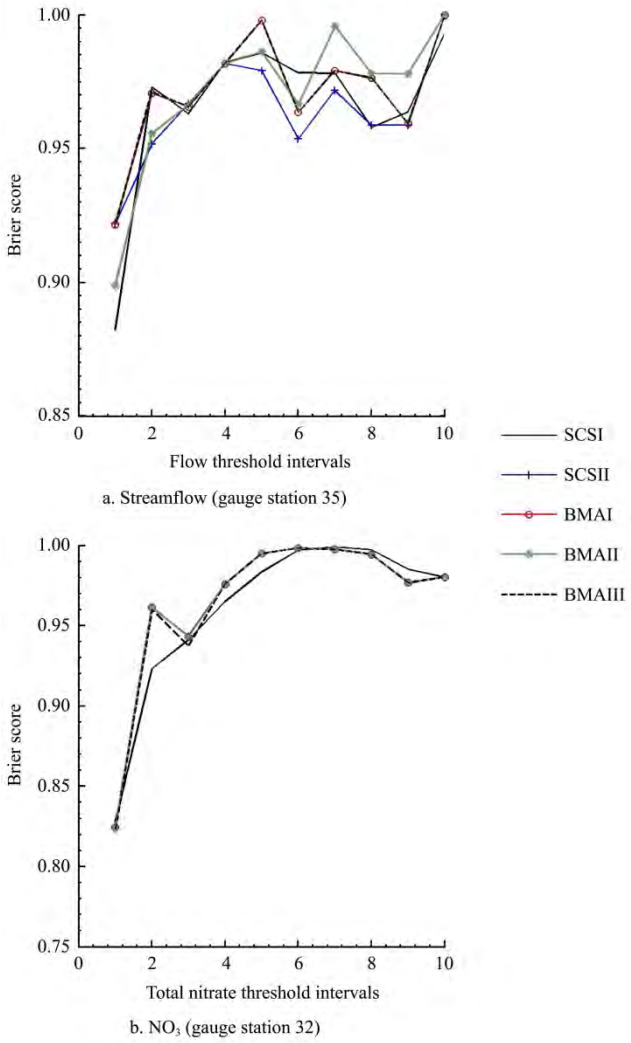


Figure 4 Brier score for SCSI, SCSII and BMA associated methods. The flow and the NO_3 threshold intervals are defined by ranking from the highest (interval = 10) to the lowest (interval = 1) quantities of observation data for specific gauge station

3.5 Structural uncertainty

Uncertainty may be embedded within model integrations and it can be added to model uncertainty as

new algorithms are developed in large models like SWAT. One way to quantify and compare such structural uncertainty is to construct uncertainty bands in the output and then to identify the differences of characteristics. The evaluation techniques may be supplemented by other techniques like QQ plots for validation purposes. The influence of different settings in the five scenarios on the streamflow and NO_3 outputs was estimated in each of the 10 000 iterations made during the calibration, then inclusion rates for each scenario and output variable were counted (Tables 6 and 7). Inclusion rate is defined as a number of iterations (or the outputs) that gave satisfactory statistical results among the 10 000 iterations. For uncertainty analysis, the central tendency of the inclusion rate was calculated by counting observation data points located within 95% of predictive uncertainty intervals and the spread was assumed to be the average width of the uncertainty band.

Table 6 Inclusion rates of observed streamflow/NO_3 data included in 95% confidence interval and the corresponding spread for the calibration period (1997-2000)

Calibration		st.35	st.32	st.27	st.22	st.20
Scenario 1	IR/%	33.33	18.52	33.33	14.81	33.33
	Spread	0.50	1.00	1.00	0.66	0.68
Scenario 2	IR/%	29.17	37.04	48.15	29.63	48.15
	Spread	0.48	0.82	0.82	0.86	0.80
Scenario 3	IR/%	41.67	44.44	51.85	40.74	55.56
	Spread	0.92	1.55	1.57	1.62	1.26
Scenario 4	IR/%	45.83	48.15	51.85	40.74	55.56
	Spread	0.94	1.57	1.57	1.62	1.36
Scenario 5	IR/%	41.67	44.44	40.74	29.63	48.15
	Spread	0.92	1.55	1.40	1.24	1.15

Note: IR: Inclusion rate; st.: gauge station number on Figure 1; Unit of spread for scenario 1 is m^3/s, and kg/month for scenario 2-5.

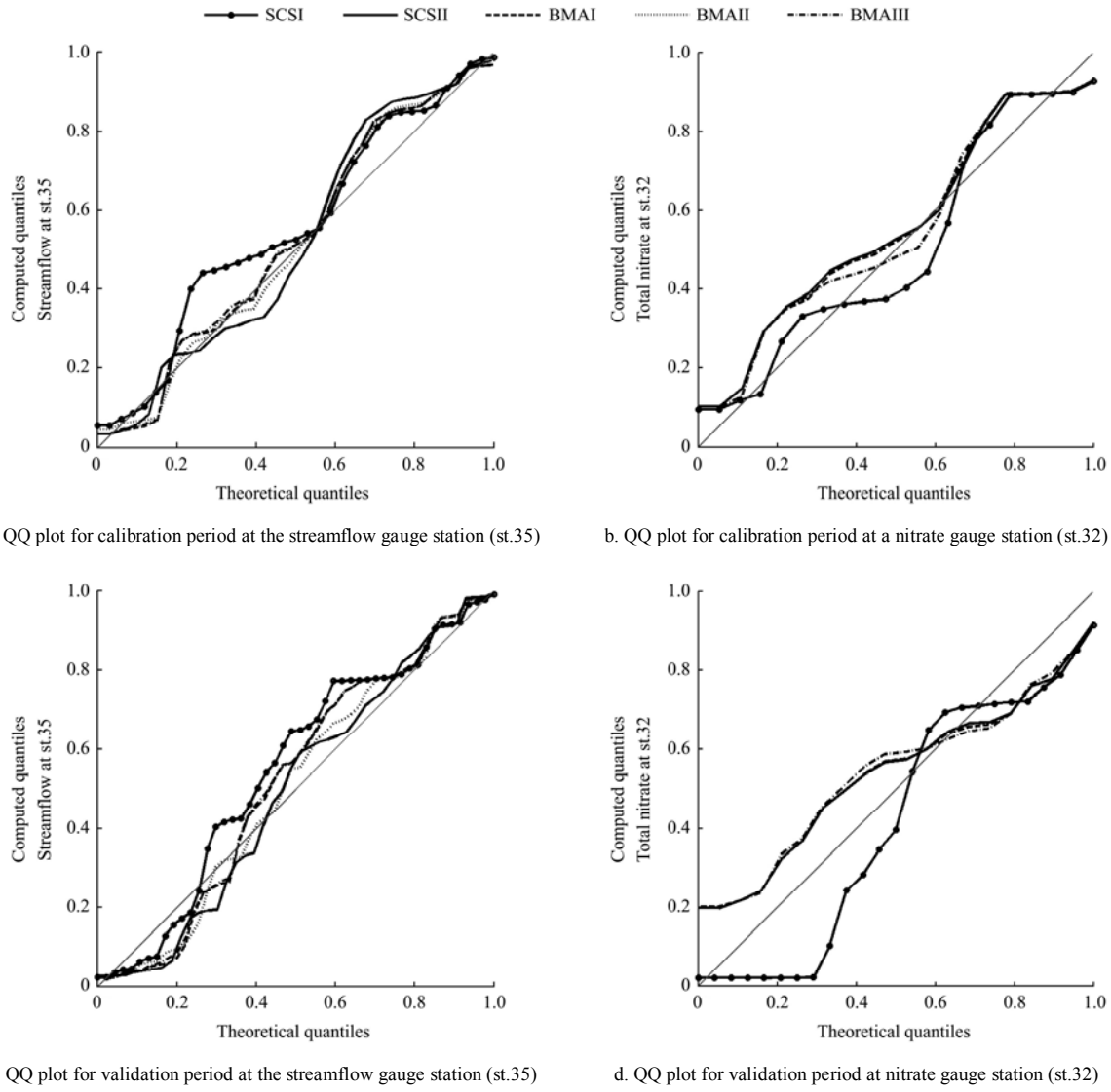
Table 7 Inclusion rates of observed streamflow/NO_3 data included in 95% confidence interval and the corresponding spread for the validation period (2001-2003)

Validation		st.35	st.32	st.27	st.22	st.20
Scenario 1	IR/%	38.89	33.33	27.78	27.78	38.89
	Spread	0.74	2.52	3.01	1.77	4.02
Scenario 2	IR/%	44.44	61.11	66.67	50.00	50.00
	Spread	1.30	1.80	1.85	1.68	1.78
Scenario 3	IR/%	44.44	61.11	66.67	50.00	55.56
	Spread	0.81	1.78	1.85	1.68	1.99
Scenario 4	IR/%	41.67	61.11	66.67	50.00	55.56
	Spread	0.83	1.80	1.85	1.68	1.62
Scenario 5	IR/%	41.67	55.56	55.56	44.44	61.11
	Spread	0.79	1.57	1.54	1.39	1.66

Note: IR: Inclusion Rate; st.: gauge station number on Figure 1; Unit of spread for scenario 1 is m^3/s, and kg/month for scenario 2-5.

As shown in Table 6, the BMA scenarios (scenarios 3, 4 and 5) resulted in increased uncertainty and a higher inclusion rate in the calibration period (including all streamflow and NO_3 stations) than the other SCSI and SCSII scenarios. In the validation period (Table 7), the SCSII and BMA associated approaches (scenario 2) resulted in higher spread and inclusion rate for streamflow and NO_3 (scenario 1, 3, 4 and 5 have similar results). In addition, the average spread of NO_3 stations in scenarios 2, 3, 4 and 5 is narrower than that in scenario 1 but with higher inclusion rate. In general, the inclusion rate for NO_3 increases with the application of BMA (scenarios 3 and 4) in both calibration and validation periods. However, the dynamic monthly updates made to the BMA weights in scenario 5 do not improve the results of either SCSI or SCSII.

The QQ plots are a graphical tool to assess model predictive uncertainty. A QQ plot compares the quantiles of two variables such as predicted and observed flows. If the variables come from the same type of distribution, or identical to each other, the QQ plot is a straight line. Instructions as guidance in interpreting QQ plot can be found in literature[54]. The QQ plots for calibration and validation periods are shown in Figure 5. As shown in Figure 5a, 5b for calibration period, most scenarios do not have major differences in matching with observation data (theoretical quantiles) but slightly overestimating streamflow and NO_3. The SCSI is the only case with distinct performance compared to other four cases. According to a visual inspection of the QQ plots for validation period, the SCSII and BMA associated scenarios perform better than SCSI in predicting both streamflow and NO_3 (Figures 5c and 5d). In addition, the SCSII and BMA associated methods are generating QQ plots with smoother shape which means more consistent results (i.e., less fluctuation in cumulative distribution) can be expected from these approaches and the state of over- or under-estimated predictive uncertainty is not as noticeable as SCSI. The QQ plots for other nitrate gauge stations also showed similar patterns which are not presented in this paper.



a. QQ plot for calibration period at the streamflow gauge station (st.35)

b. QQ plot for calibration period at a nitrate gauge station (st.32)

c. QQ plot for validation period at the streamflow gauge station (st.35)

d. QQ plot for validation period at nitrate gauge station (st.32)

Figure 5 Quantile-Quantile (QQ) plots for calibration and validation periods for all cases

4 Discussion and conclusions

From the viewpoint of finding better solutions and the speed of convergence of the objective function, SCSI and SCSII show compatible results although the SCSII converges a little slower. The application of a formal likelihood function sustains high inclusion rate in the "Satisfactory" level for calibration but it decreases dramatically in the validation period. The SCSII and BMA associated methods derive more solutions which satisfy additional model performance validation criteria in both calibration and validation periods. As mentioned previously, SCSI tends to overestimate surface runoff in shallow soils when storage condition is low[10,12]. During the severe drought in the first year of the validation period, SCSII and BMA associated methods show better quality solutions. In other words, SCSI is less capable of properly simulating low flow or drought conditions[10,12]. Moreover, SCSII and BMA associated approaches consistently generate solutions with better statistics in both calibration and validation periods.

The higher inclusion rate for NO_3 with the incorporation of the BMA technique offers another major benefit of aggregating SCSI and SCSII during optimization processes by improving the likelihood of confining the observed pollutograph within the uncertainty bands of predicted output. As demonstrated in Table 3, the SCSII output had a dominant impact to the BMA scenarios with significant weights for NO_3 gauge stations compared to SCSI, which means that SCSI is not able to simulate NO_3 processes as well as SCSII in practice. Therefore, in this study, SCSII (and also BMA

applications along with SCSII) performed relatively better on predicting nutrient loads than the SCSI method. The SCSII method also performed better than the SCSI method in predicting streamflow. However, it resulted in higher PBIAS at two stream gages among the five gages.

In this study, the impact of structural uncertainty on hydrologic and water quality predictions which may be intrinsic in the two SWAT CN methods was investigated with an application of the BMA technique. Several conclusions are made based on the analyses of the five scenarios.

(1) SCSI and SCSII show similar performance in convergence speed or optimization of the objective function. However, SCSII and BMA methods give superior model predictions with statistically better solutions and higher success rate compare to SCSI.

(2) SCSI has relatively higher BMA weights on streamflow outputs over SCSII (except for the wet season in scenario 4) but its predicted streamflow is not perceptibly superior to SCSII.

(3) In simulating water quality variables, SCSII outperforms SCSI in simulating NO_3 in the case study of ECW. Evidently, higher BMA weights are assigned to SCSII for NO_3 outputs. Error statistics, predictive uncertainty and NO_3 profiles of the BMA associated scenarios were particularly close to that in SCSII because dominant weights are assigned to SCSII on NO_3 outputs.

(4) The seasonal classification approach does not show substantial improvements in either statistical metrics or predictive uncertainty. Applications of BMA have positive effects in increasing inclusion rate but the predictive uncertainty is not evidently reduced.

For a complex large-scale watershed simulation model as SWAT, optional alternatives can be applied on different functions for specific hydrologic/nutrient processes (e.g., surface runoff in this study). Advantages and disadvantages ensue with the application of each algorithm can be demonstrably affected especially when methods are designed to catch dissimilar watershed characters or various engineering purposes. In this study, the use of the plant evapotranspiration CN method in SWAT demonstrated outstanding performance in the prediction of both discharge and NO_3. SCSII was originally designed to improve runoff simulation, and it also improves SWAT's ability to predict pollutants which result in water quality impairment as demonstrated in this research by nitrate. We recommend additional SWAT calibration/validation research with an emphasis on the impact of SCSII on the prediction of other pollutants (e.g., pesticide, phosphorus). In addition, we also recommend testing with additional watersheds with differing climatic conditions to establish recommendations on the best CN method to utilize.

Acknowledgements

This study was supported in part by the USDA - National Institute of Food and Agriculture grants 2007-51130-03876, 2009-51130-06038, the Research Program for Agricultural Science & Technology Development (Project No. PJ008566), National Academy of Agricultural Science, Rural Development Administration, Republic of Korea, and the USDA-NRCS Conservation Effects Assessment Project (CEAP) – Wildlife and Cropland components. Awesome comments by the associate editor and reviewers greatly improved the quality of the manuscript. Thanks a lot! In addition, please remember that USDA is an equal opportunity employer and provider.

[References]

[1] Daniel E B, Camp J V, LeBoeuf E J, Penrod J R, Dobbins J P, Abkowitz, M D. Watershed modeling and its applications: A state-of-the-art review. The Open Hydrology Journal, 2011; 5: 26–50.

[2] Borah D K, Yagow G, Saleh A, Barnes P L, Rosenthal W, Krug E C, Hauck L M. Sediment and nutrient modeling for TMDL development and implementation. Transactions of the ASABE, 2006; 49(4): 967–986.

[3] Gassman P W, Reyes M R, Green C H, Arnold J G. The Soil and Water Assessment Tool: Historical development, applications, and future research directions. Transactions of The ASABE, 2007; 50(4): 1211–1250.

[4] Williams J R, Arnold J G, Kiniry J R, Gassman P W, Green C H. History of model development at Temple, Texas. Hydrological Sciences Journal, 2008; 53(5): 948–960.

[5] Yen H, Bailey R T, Arabi M, Ahmadi M, White M J, Arnold J G. The role of interior watershed processes in improving parameter estimation and performance of watershed models. Journal of Environmental Quality, Published online. doi:

10.2134/jeq2013.03.0110.

[6] USDA-NRCS, 2004. Chapter 10: Estimation of direct runoff from storm rainfall. In Part 630: Hydrology: NRCS National Engineering Handbook, USDA National Resources Conservation Service, Washington, DC. Available at: http://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/mi/technical/?cid=stelprdb1043063. Accessed on [2008-06-16].

[7] Green W H, Ampt G A. Studies on soil physics, part I, the flow of air and water through soils. Journal of Agricultural Sciences, 1911; 4(1): 1–24.

[8] Ponce V M, Hawkins R H. Runoff curve number: Has it reached maturity? Journal of Hydrologic Engineering, 1996; 1: 11–19.

[9] Kannan N, Santhi C, Williams J R, Arnold J G. Development of a continuous soil moisture accounting procedure for curve number methodology and its behaviour with different evapotranspiration methods. Hydrological Processes, 2008; 22(13): 2114–2121, doi: 10.1002/ hyp.6811.

[10] Green C H, Tomer M D, Di Luzio M, Arnold J G. Hydrologic evaluation of the soil and water assessment tool for a large tile-drained watershed in Iowa. Transactions of the ASABE, 2006; 49(2): 413–422.

[11] Amatya D M, Jha M K. Evaluating the SWAT model for a low-gradient forested watershed in coastal South Carolina. Transactions of the ASABE, 2011; 54(6): 2151–2163.

[12] Gassman P W. A simulation assessment of the Boone River watershed: Baseline calibration/validation results and issues, and future research needs. Doctoral Dissertation, 2008, Iowa State University, Ames, Iowa.

[13] Setegn S G, Srinivasan R, Melesse A M, Dargahi B. SWAT model application and prediction uncertainty analysis in the Lake Tana Basin, Ethiopia. Hydrological Processes, 2010; 24: 357–367, doi: 10.1002/hyp.7457.

[14] Yen H. Confronting Input, Parameter, Structural, and Measurement Uncertainty in Multi-site Multiple Responses Watershed Modeling using Bayesian Inferences. Doctoral Dissertation, Colorado State University, 2012.

[15] Yen H, Wang X, Fontane D G, Harmel R D, Arabi M. A framework for propagation of uncertainty contributed by input data, parameterization, model structure, and calibration/validation data in watershed modeling. Environmental Modelling and Software, 2014; 54: 211–221. doi: 10.1016/j.envsoft.2014.01.004

[16] Jajarmizadeh M, Harun S, Gharaman B, Mokhtari M H. Modeling daily stream flow using plant evapotranspiration method. International Journal of Water Resources and Environmental Engineering, 2012: 4(6): 218–226. doi: 10.5897/IJWREE12.019.

[17] Srinivasan R X, Zhang J A. Swat ungagged: Hydrological budget and crop yield predictions in the upper Mississippi River Basin. Transactions of the ASABE, 2010; 53(5): 1533–1546.

[18] Hoeting J A, Madigan D, Raftery A E, Volinsky C T. Bayesian model averaging : A tutorial. Statistical Science, 1999; 14(4): 382–417.

[19] Newman J E. The Natural Heritage of Indiana: Our Changing Climate, Edited by M. T. Jackson, 1997.

[20] USGS NED. United States Geological Survey National Elevation Database, 1 arc Second Digital Elevation Model, 2010.

[21] USDA NASS. United States Department of Agriculture - National Agricultural Statistics Service, Cropland Data Layer, 2003.

[22] USDA NRCS. United States Department of Agriculture - Soil Data Mart, 2010.

[23] United States National Climatic Data Center (2013), http://www.ncdc.noaa.gov/; Accessed on [2013-12-03].

[24] Arnold J. G, Allen P M, Volk M, Williams J R, Bosch D D. Assessment of different representations of spatial variability on SWAT model performance. Transactions of the ASABE, 2010; 53(5): 1433–1443.

[25] Daggupati P, Douglas-Mankin K R, Sheshukov A Y, Barnes P L, Devlin D L. Field-level targeting using SWAT: Mapping output from HRUs to fields and assessing limitations of GIS input data. Transactions of the ASABE, 2011; 54(2): 501–514.

[26] Douglas-Mankin K R, Srinivasan R, Arnold J G. Soil and Water Assessment Tool (SWAT) model: current developments and applications. Transactions of the ASABE, 2010; 53(5): 1423–1431.

[27] Hoque Y M, Hantush M M, Govindaraju R S. On the scaling behavior of reliability-resilience-vulnerability indices in agricultural watersheds. Ecological Indicators, 2014; 40: 136–146.

[28] Osorio J, Jeong J, Arnold J G, Beiger B. Influence of potential evapotranspiration on the water balance of sugarcane fields in Maui, Hawaii-special issue: evapotranspiration. J. Water Res. 2014; 6(9): 852–868.

[29] Yen H, Ahmadi M White M J, Wang X, Arnold J G. C-SWAT: The Soil and Water Assessment Tool with Consolidated Input Files in Alleviating Computational Burden of Recursive Simulations. Computers & Geosciences. 2014; 72: 221–232. doi: 10.1016/j.cageo.2014.07.017.

[30] Arnold J G, Srinivasan R, Muttiah R S, Williams J R. Large area hydrologic modeling and assessment. I. Model development. Journal of the American Water Resources Association, 998; 34(1): 73–89.

[31] Rallison R E, Miller N. Past, present, and future SCS runoff procedure. In Rainfall Runoff Relationship. ed. V. P. Singh. Littleton, Colo.: Water Resources Publication, 1981; 353–364.

[32] Neitsch S L, Arnold J G, Kiniry J R, Williams J R. Soil and Water Assessment Tool: Theoretical Documentation Version 2009, Texas Water Resources Institute TR-406, College Station, Texas, 2011.

[33] Williams J, Kannan N, Wang X, Santhi C, Arnold J. Evolution of the SCS runoff curve number method and its application to continuous runoff simulation. Journal of Hydrologic Engineering, 2012; 17(11): 1221–1229.

[34] Ahmadi M. A Multi Criteria Decision Support System for Watershed Management under Certain Conditions. Doctoral Dissertation, Colorado State University, 2012.

[35] Moriasi D N, Arnold J G, Liew M W V, Bingner R L, Harmel R D, Veith T L. Model validation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE, 2007; 50(3): 885–900.

[36] Stedinger J R, Vogel R M, Lee S U, Batchelder R. Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. Water Resources Research, 2008; 44: 1–17. doi: 10.1029/2008WR006822.

[37] Vrugt J A, Braak C J F, Gupta V K, Robinson B A. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stochastic Environmental Research and Risk Assessment, 2009; 23(7): 1011–1026. doi: 10.1007/s00477-008-0274-y.

[38] Tolson B A, Shoemaker C A. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. Water Resources Research, 2007; 43(1): 1–16, doi: 10.1029/2005WR004723.

[39] Yen H, Jeong J, Tseng W-H, Kim M-K, Records R M, Arabi M. Computational procedure in evaluating sampling techniques for parameter estimation and uncertainty analysis in watershed modeling. Journal of Hydrologic Engineering. 2014; doi: 10.1061/(ASCE)HE.1943- 5584.0001095.

[40] Seo M-J, Yen H, Jeong J. Transferability of Input Parameters between SWAT 2009 and SWAT 2012. Journal of Environmental Quality. 2014; 43(3): 869–880. doi: 10.2134/jeq2013.11.0450

[41] Beck M B, Fath B D, Parker A K, Osidele O O, Cowie G M, Rasmussen T C, et al. Developing a concept of adaptive community learning: Case study of a rapidly urbanizing watershed. Integrated Assessment, 2002; 3(4): 299–307, doi: 10. 1076/iaij.3.4.299.13583.

[42] Kass R E, Raftery A E. Bayes factors. Journal of the American Statistical Association, 1995; 90: 773–795.

[43] Leamer E E. Specification Search. Wiley, 1978; pp 370.

[44] Raftery A E, Gneiting T, Balabdaoui F, Polakowski M. Using Bayesian Model Averaging to Calibrate Forecast Ensembles. American Meteorological Society, 2005; 133: 1155–1174.

[45] Wöhling T, Vrugt J A. Combining multiobjective optimization and Bayesian model averaging to calibrate forecast ensembles of soil hydraulic models. Water Resources Research, 2008; 44(12): 1–18. doi: 10.1029/2008WR 007154.

[46] Raftery A E, Balabdaoui F, Gneiting T, Polakowski M. Using Bayesian model averaging to calibrate forecast ensembles, Tech. Rep. 440, Dep. of Stat., Univ. of Wash., Seattle, 2003.

[47] Duan Q, Ajami N K, Gao X, Sorooshian S. Multi-model ensemble hydrologic prediction using Bayesian model averaging. Advances in Water Resources, 2007; 30(5): 1371–1386. doi: 10.1016/j.advwatres.2006.11.014.

[48] Ajami N K, Duan Q, Sorooshian S. An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. Water Resources Research, 2007; 43(1): 1–19. doi:10.1029/2005WR 004745.

[49] Georgakakos K, Seo D, Gupta H, Schaake J, Butts M. Towards the characterization of streamflow simulation uncertainty through multimodel ensembles. Journal of Hydrology, 2004; 298(1-4): 222–241, doi: 10.1016/ j.jhydrol.2004.03.037.

[50] Burnash R J, Ferral R L, McGuire R A. A generalized streamflow simulation system conceptual modeling for digital computers. Report, U. S. Dep. of Commer. Natl. Weather Serv. and State of Calif. Dep. of Water Resour, 1973.

[51] Boyle D. Multicriteria calibration of hydrological models. Ph.D. dissertation, Univ. of Ariz., Tucson, 2001.

[52] Schaake J C, Koren V, Duan Q Y, Chen F. Simple water balance model for estimating runoff at different spatial and temporal scales. Journal of Geophysical Research, 1996; 101(D3): 7461–7475.

[53] NOAA (2012), Local weather observation station record, NOAA Satellite and Information Service and National Climatic Data Center, Available at: http://www.ncdc. noaa.gov/oa/ncdc. html. Accessed on [2013-05].

[54] Thyer M, Renard B, Kavetski D, Kuczera G, Franks S W, Srikanthan S. Critical validation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. Water Resources Research, 2009; 45(3): 1–22. doi: 10.1029/2008 WR006825.