A Framework for Incorporating Uncertainty Sources in SWAT Modeling

Haw Yen
X. Wang, D. G. Fontane, R. D. Harmel, M. Arabi

July 30, 2014
2014 International SWAT Conference
Pernambuco, Brazil
Outline

- Overview
- Calibration of Watershed Simulation Models
- Sources of Uncertainty in Watershed Modeling
- Incorporation of Uncertainty Sources
- Case Study & Results
- Discussion & Conclusion
Overview

- Development of complex watershed models
  - Evaluate impact from climate changing, various human activities on issues such as:
    - Availability of water resources
    - Water quality
    - Watershed management

- Advanced technology in computer science
  - Complex watershed simulation models
    - Distributed in space & process-based
    - Long term simulations with large amount of data
Calibration of Watershed Models

- Why and how do we calibrate?
  - Model parameters can be case sensitive
  - Before conducting model simulation for various scenarios
    - To ensure model responses are close to natural responses
    - To minimize the “differences” between observed/simulated data by adjusting values of model parameters
  - “Differences” can be calculated as?
    - Error statistics (ex. RMSE, PBIAS, 1-NSE)
Sources of Uncertainty in Watershed Modeling

<Forcing Inputs>
- Climate: P, T, pressure, ..., etc
- Soils: types, texture, etc.
- Land use/land cover: type, etc
- Terrain and stream network

<Measured Fluxes>
- Streamflow
- Sediment
- Nutrients: N and P
- Chemicals: atrazine

Inputs → Watershed Model → Outputs

θ = {θ₁, θ₂, ...}
M = {M₁, M₂, ...}

<Model Parameters>
- Curve number
- Manning’s n
- Hydraulic K

<Model Structure>
- Surface, Subsurface Runoff processes
- Erosion and sedimentation
- Soil biogeochemical processes
- In-stream processes
Research Goal

- To incorporate the uncertainty from input, parameter, structural and measurement sources jointly during model calibration
- To understand the role and importance of four uncertainty sources during parameter estimation process
- To examine the effects of four uncertainty sources toward predictive uncertainty
Incorporation of Uncertainty Sources

- Parameter Uncertainty
  - Parameter estimation

**Dynamically dimensioned search (DDS)**

- no need for algorithm parameter tuning
- fast approximate stochastic global optimization
- search scaled to pre-specified max # of function evaluations
  (global search at the beginning and more localized in the end)
- perturbed variables are generated from a normal distribution
  centered on current best value.
Incorporation of Uncertainty Sources

Input Uncertainty

Input error model

The rainfall uncertainty was considered using an input error model which assumes a random Gaussian error as a multiplier for each input rainfall observation as proposed by Ajami et al. (2007) (WRR):

\[ \bar{R}_t = \phi_t R_t ; \quad \phi \sim N(m, \sigma_m^2) \]

\( \bar{R}_t \): true rainfall depth at time \( t \);
\( R_t \): observed rainfall depth at time \( t \);
\( \phi_t \): represents a random multiplier (noise) at time \( t \) with
mean \( m, m \in [0.9, 1.1] \) and variance \( \sigma_m^2, \sigma_m^2 \in [10^{-5}, 10^{-3}] \)
Incorporation of Uncertainty Sources

☯ Structural Uncertainty

**Bayesian model averaging (BMA)**

Bayesian theorem is applied over a set of considered models, $M_k$, to calculate a weighted probability distribution $p(y)$ for model output:

$$p(y) = \sum_{k=1}^{m} p(M_k | Y^T) p(y_k | M_k) = \sum_{k=1}^{m} w_k (y_k | M_k)$$

$p(y)$: weighted output distribution based on $M_k$ considered models;
$p(M_k | Y^T)$: posterior probability of model $M_k$ being correct model given the training data $Y^T$, and it reflects how well model $M_k$ fits the observed variable during the training period $T$, and it is also known as the BMA weight $w_k$.
$p(y_k | M_k)$: forecast pdf of output variable $y_k$ based on model $M_k$. 
Incorporation of Uncertainty Sources

- Measurement Uncertainty (2/2)
  - Incorporation of measurement uncertainty
    - Probability distribution (PD) method
      - Proposed by Harmel and Smith (2007)
      - Assign a correction factor on error between observation and simulation quantities

\[
E_k = \frac{CF_k}{0.5} (Q_{k}^{obs} - Q_{k}^{sim})
\]

\[
CF_k = \begin{cases} 
\text{normcdf} (Q_{k}^{sim}, \mu, \sigma) - 0.5 & \text{if } Q_{k}^{sim} \geq Q_{k}^{obs} \\
0.5 - \text{normcdf} (Q_{k}^{sim}, \mu, \sigma) & \text{if } Q_{k}^{sim} < Q_{k}^{obs}
\end{cases}
\]
Case Study Area

- Eagle Creek watershed
  - Central Indiana, USA
  - 248km²
- Available data
  - 1997~2003
  - Streamflow (1 site)
  - NOX (4 sites)

Daily streamflow
Monthly Total Nitrate
Calibration (1997~2000)
Validation (2001~2003)
## Case Study Settings

<table>
<thead>
<tr>
<th>Case Scenarios</th>
<th>Scenario Setup</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario I</td>
<td>Calibration using SCSI</td>
</tr>
<tr>
<td>Scenario II</td>
<td>Calibration using SCSI + IU</td>
</tr>
<tr>
<td>Scenario III</td>
<td>Calibration using SCSII</td>
</tr>
<tr>
<td>Scenario IV</td>
<td>Calibration using SCSI + MU</td>
</tr>
<tr>
<td>Scenario V</td>
<td>Calibration using SCSI + IU + MU</td>
</tr>
<tr>
<td>Scenario VI</td>
<td>Calibration using SCSII + IU + MU</td>
</tr>
<tr>
<td>Scenario VII</td>
<td>Calibration using SCSI + IU + MU + Internal watershed behavior constraints</td>
</tr>
<tr>
<td>Scenario VIII</td>
<td>Calibration using SCSII + IU + MU + Internal watershed behavior constraints</td>
</tr>
<tr>
<td>Scenario IX</td>
<td>Apply BMA to Scenario V &amp; VI</td>
</tr>
<tr>
<td>Scenario X</td>
<td>Apply BMA to Scenario VII &amp; VIII</td>
</tr>
</tbody>
</table>
Watershed behavior constraints

- Denitrification
  Denitrification rate no more than 50 kg/ha/yr

- Ratio of NO$_3$-N losses contributed from subsurface flow (SSQ) verse the total loss from SSQ and surface flow (SQ)
  Greater than 0.6

If results violate these constraints, then the corresponding model run is penalized by assigning an extreme value to the objective function used in DDS minimization procedure. Therefore, new search can avoid poor local optima.
Results

- Results of objective function values
## Results

- **Best objective function values and the corresponding outputs**

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Objective Function</th>
<th>Denitrification (kg/ha)</th>
<th>NO3-N Loss Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario I</td>
<td>342.1 (399.6)</td>
<td>121.4 (16.9)</td>
<td>0.98 (0.94)</td>
</tr>
<tr>
<td>Scenario II</td>
<td>348.6 (379.4)</td>
<td>214.2 (30.4)</td>
<td>0.97 (0.99)</td>
</tr>
<tr>
<td>Scenario III</td>
<td>343.1 (373.1)</td>
<td>243.3 (7.1)</td>
<td>0.96 (0.98)</td>
</tr>
<tr>
<td>Scenario IV</td>
<td>342.4 (399.6)</td>
<td>211.4 (49.7)</td>
<td>0.98 (0.98)</td>
</tr>
<tr>
<td>Scenario V</td>
<td>340.1</td>
<td>36.1</td>
<td>0.96</td>
</tr>
<tr>
<td>Scenario VI</td>
<td>345.2</td>
<td>14.5</td>
<td>0.63</td>
</tr>
<tr>
<td>Scenario VII</td>
<td>343.0</td>
<td>36.1</td>
<td>0.92</td>
</tr>
<tr>
<td>Scenario VIII</td>
<td>344.7</td>
<td>49.8</td>
<td>0.79</td>
</tr>
</tbody>
</table>

( * ): **Behavior Definitions** applied (General Performance Ratings by Moriasi et al. 2007)
Results

- Applications of internal watershed behavior constraints during calibration
Results

Nash-Sutcliffe efficiency (NSE) and percent error (PBIAS) for calibration/validation periods at station #35 for streamflow. C.: calibration and V: validation.

“original”: original calibration results;
“filtered”: post-processed results after removing runs violated behavior constraints
Results

NSE and PBIAS for calibration/validation periods for calibration cases I–IV at the 4 USGS stations (st.) for NO$_3$-N loss.
Results

NSE and PBIAS for calibration/validation periods for calibration cases V-BMA(VII-VIII) at the 4 USGS stations (st.) for NO$_3$-N loss.
Results

Percentage of observations within prediction bounds during validation for cases considered uncertainty in parameter, input data and calibration/validation data.
Conclusions

- Watershed behavior was more realistically represented when three or four major sources of uncertainty were considered without having to embed watershed behavior constraints in auto-calibration procedure;
- Inclusion of four uncertainty sources improved model simulations for both the calibration period and validation period;
- Application of watershed behavior constraints improved the quality of calibration results.
Reference


Thanks for your attention!

Haw Yen, Ph.D.
hyen@brc.tamus.edu