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

TEXAS A&M AGRILIFE RESEARCH

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

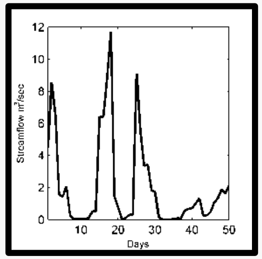
- Why and how do we calibrate?
  - Model parameters can be case sensitive
  - Before conducting model simulation for various scenarios
    - To ensure model responses are close to natural responses
    - To minimize the “differences” between observed/simulated data by adjusting values of model parameters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

Sources of Uncertainty in Watershed Modeling

Texas A&M AgriLife Research

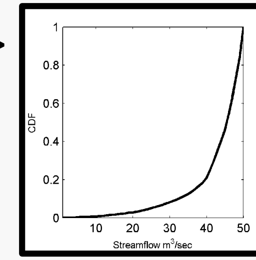
< Forcing Inputs >

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

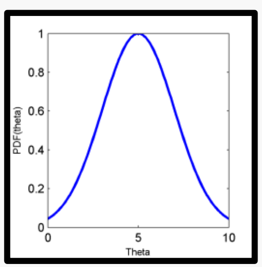


Inputs → **Watershed Model** → **Outputs**

$\theta = \{\theta_1, \theta_2, \ldots\}$ $\quad$ $M = \{M_1, M_2, \ldots\}$

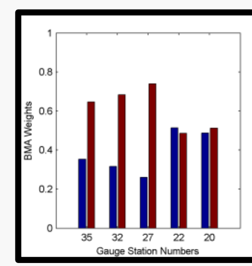
<Model Parameters>

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

TEXAS A&M AGRILIFE RESEARCH

- 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):

$$\overline{R}_t = \phi_t R_t ; \qquad \phi \sim N\left(m, \sigma_m^2\right)$$

$\overline{R}_t$: true rainfall depth at time *t*;

R_t : observed rainfall depth at time *t*;

ϕ_t : represents a random multiplier (noise) at time *t* with

mean *m*, $m \in [0.9, 1.1]$ and variance σ_m^2 , $\sigma_m^2 \in [10^{-5}, 10^{-3}]$

Incorporation of Uncertainty Sources

TEXAS A&M AGRILIFE RESEARCH

- 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:

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

$p(y)$: weighted output distribution based on M_k considered models;

$p(M_k \mid Y^T)$: 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.

$p(y_k \mid M_k)$: 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

$$\boxed{E_k = \frac{CF_k}{0.5}(Q_k^{obs} - Q_k^{sim})}$$

$$\boxed{CF_k = \begin{cases} normcdf(Q_k^{sim}, \mu, \sigma) - 0.5 & if\ \ Q_k^{sim} \geq Q_k^{obs} \\ 0.5 - normcdf(Q_k^{sim}, \mu, \sigma) & if\ \ Q_k^{sim} < Q_k^{obs} \end{cases}}$$

Case Study Area

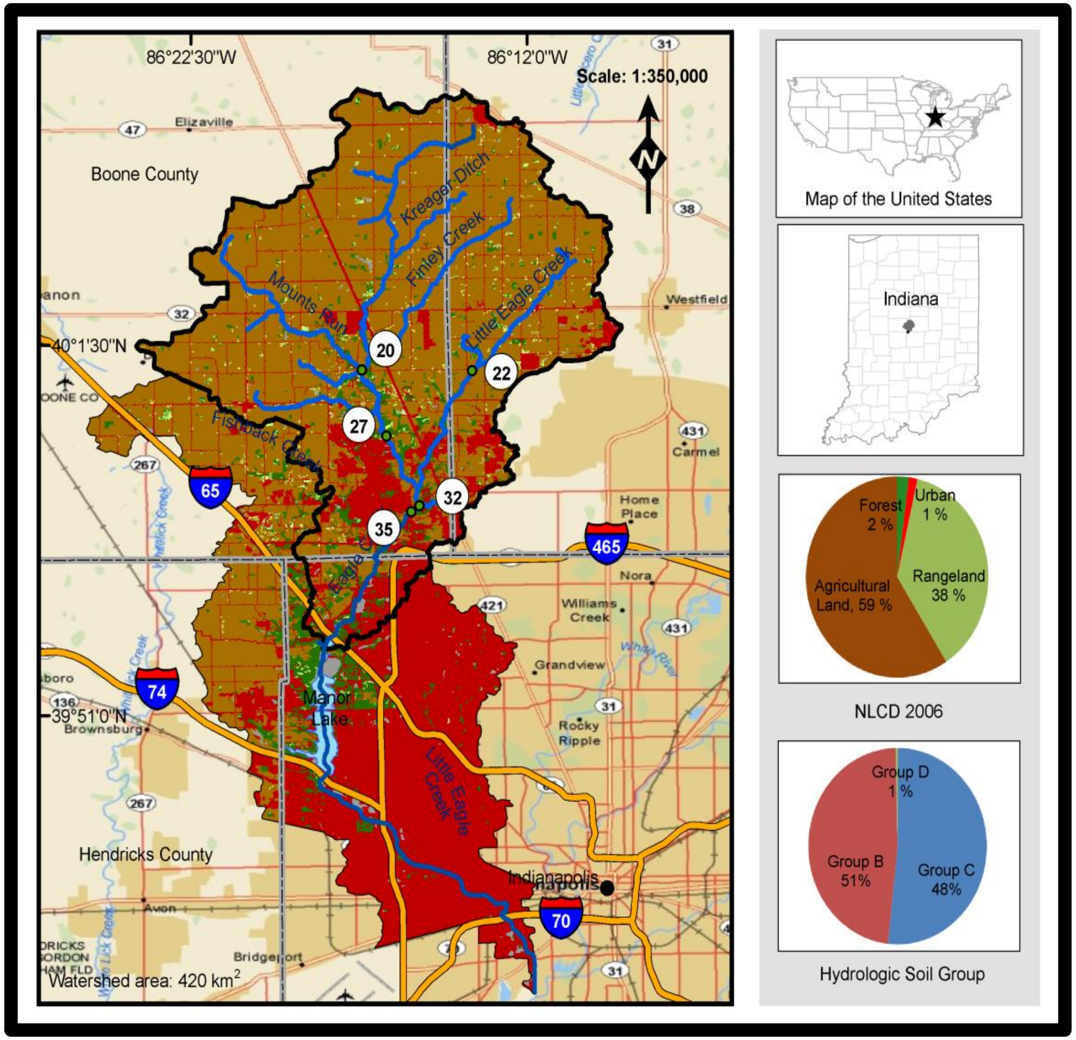
- Eagle Creek watershed
  - Central Indiana, USA
  - 248km²
  - Available data
    - 1997~2003
    - Streamflow (1 site)
    - NOX (4 sites)

Daily streamflow

Monthly Total Nitrate

Calibration (1997~2000)

Validation (2001~2003)



Case Study Settings

Case Scenarios	Scenario Setup
Scenario I	Calibration using SCSI
Scenario II	Calibration using SCSI + IU
Scenario III	Calibration using SCSII
Scenario IV	Calibration using SCSI + MU
Scenario V	Calibration using SCSI + IU + MU
Scenario VI	Calibration using SCSII + IU + MU
Scenario VII	Calibration using SCSI + IU + MU + Internal watershed behavior constraints
Scenario VIII	Calibration using SCSII + IU + MU + Internal watershed behavior constraints
Scenario IX	Apply BMA to Scenario V & VI
Scenario X	Apply BMA to Scenario VII & VIII

Watershed behavior constraints

- Denitrification

  Denitrification rate no more than 50 kg/ha/yr

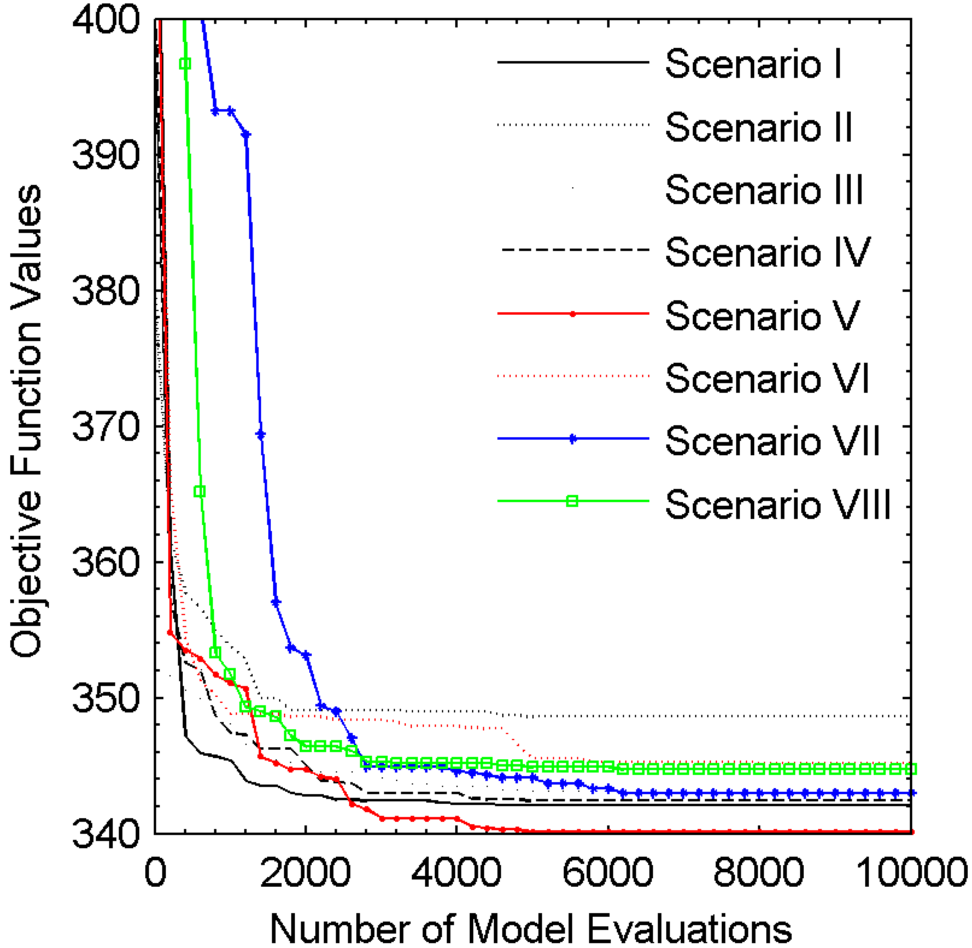
- Ratio of NO_3-N losses contributed from subsurface flow (SSQ) verse the total loss from SSQ and surface flow (SQ)

  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

TEXAS A&M AGRILIFE RESEARCH

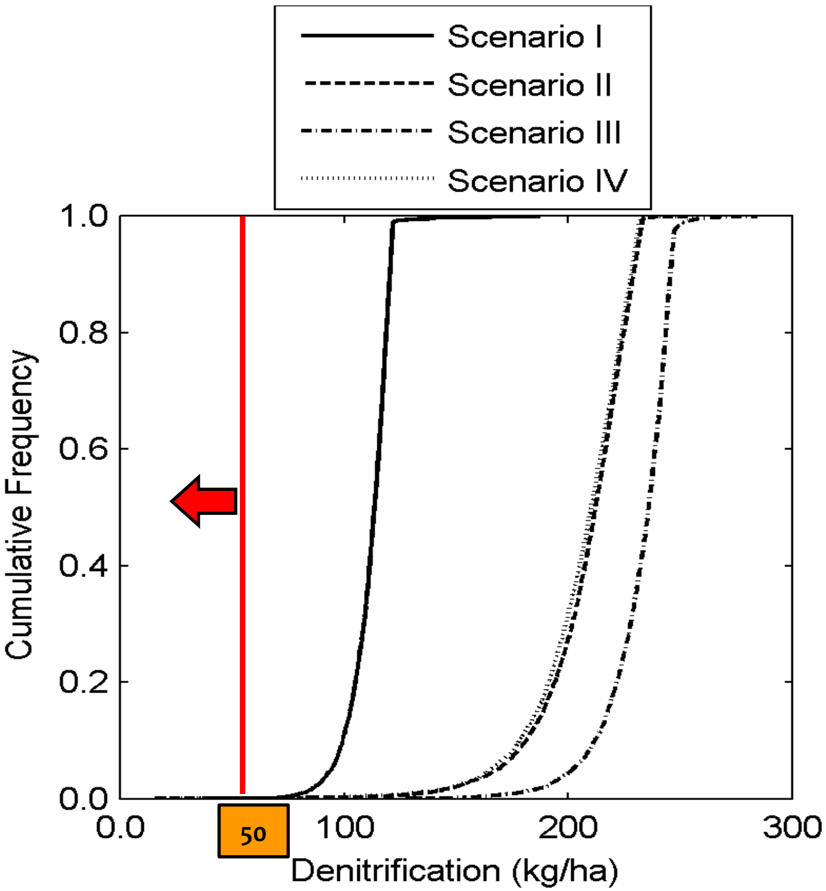
- Best objective function values and the corresponding outputs

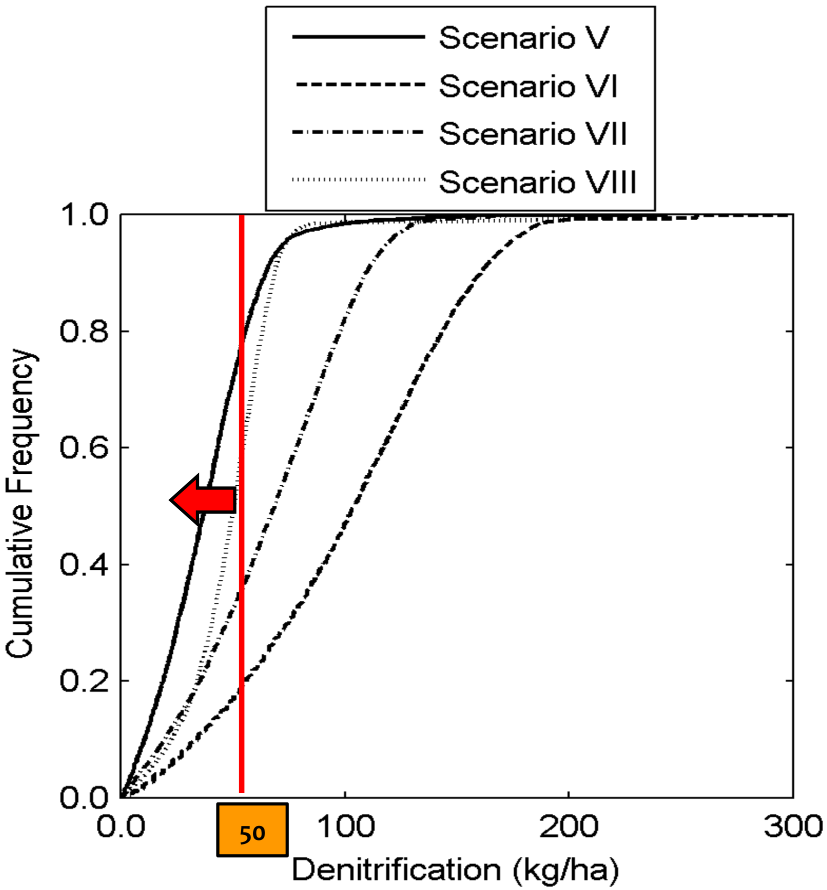
Scenarios	Objective Function	Denitrification (kg/ha)	NO3-N Loss Ratio
Scenario I	342.1 (399.6)	121.4 (16.9)	0.98 (0.94)
Scenario II	348.6 (379.4)	214.2 (30.4)	0.97 (0.99)
Scenario III	343.1 (373.1)	243.3 (7.1)	0.96 (0.98)
Scenario IV	342.4 (399.6)	211.4 (49.7)	0.98 (0.98)
Scenario V	340.1	36.1	0.96
Scenario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

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

Results

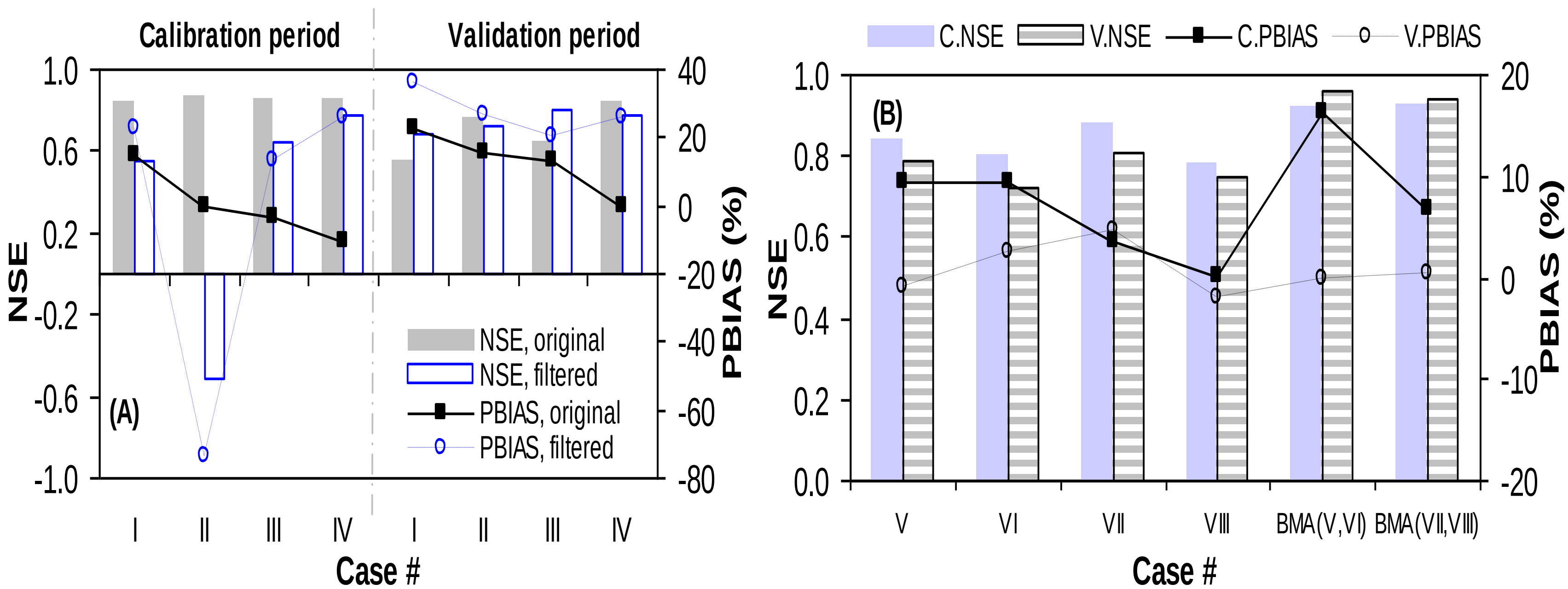
- Applications of internal watershed behavior constraints during calibration





Results

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

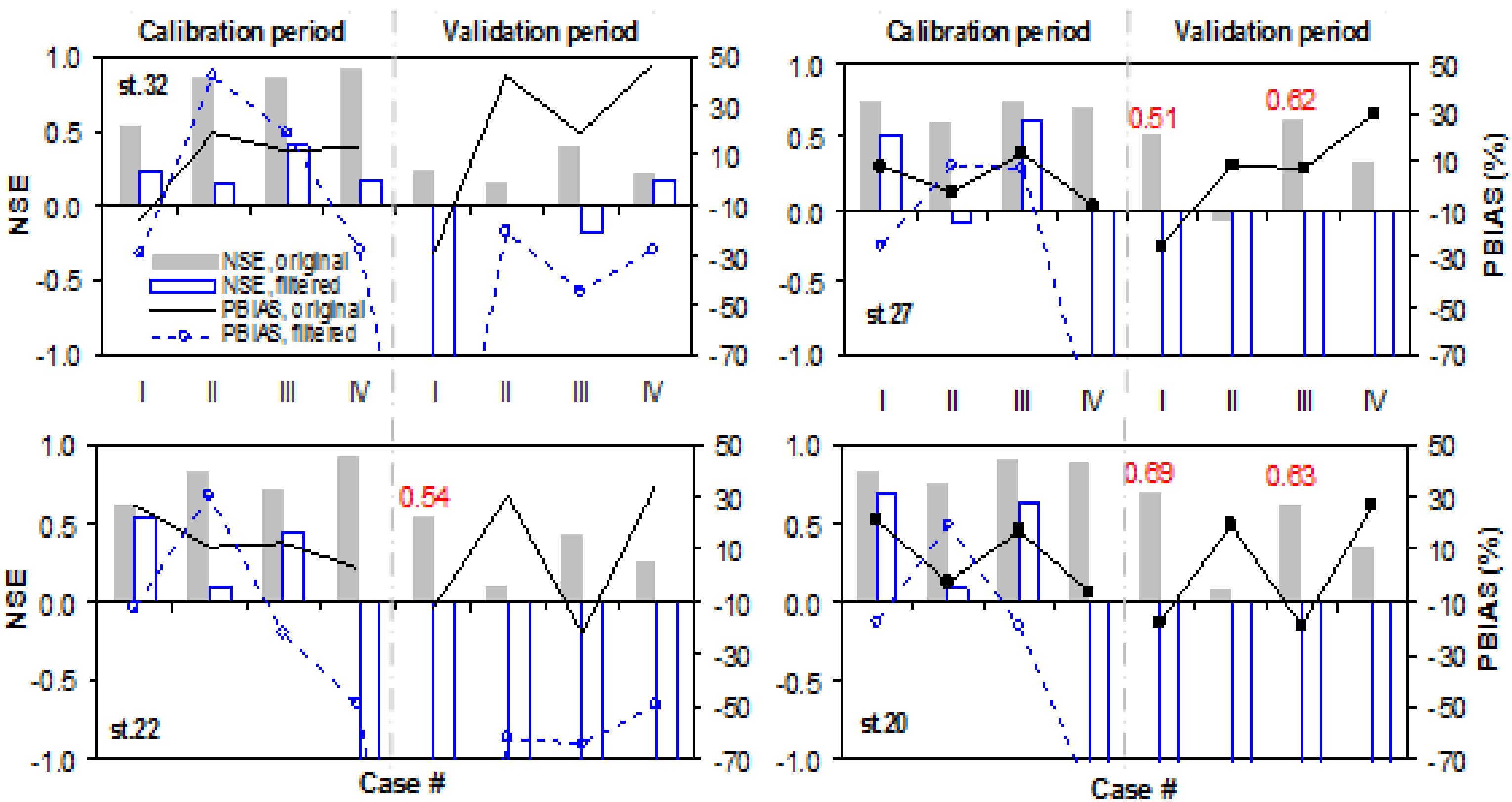


"original": original calibration results;

"filtered": post-processed results after removing runs violated behavior constraints

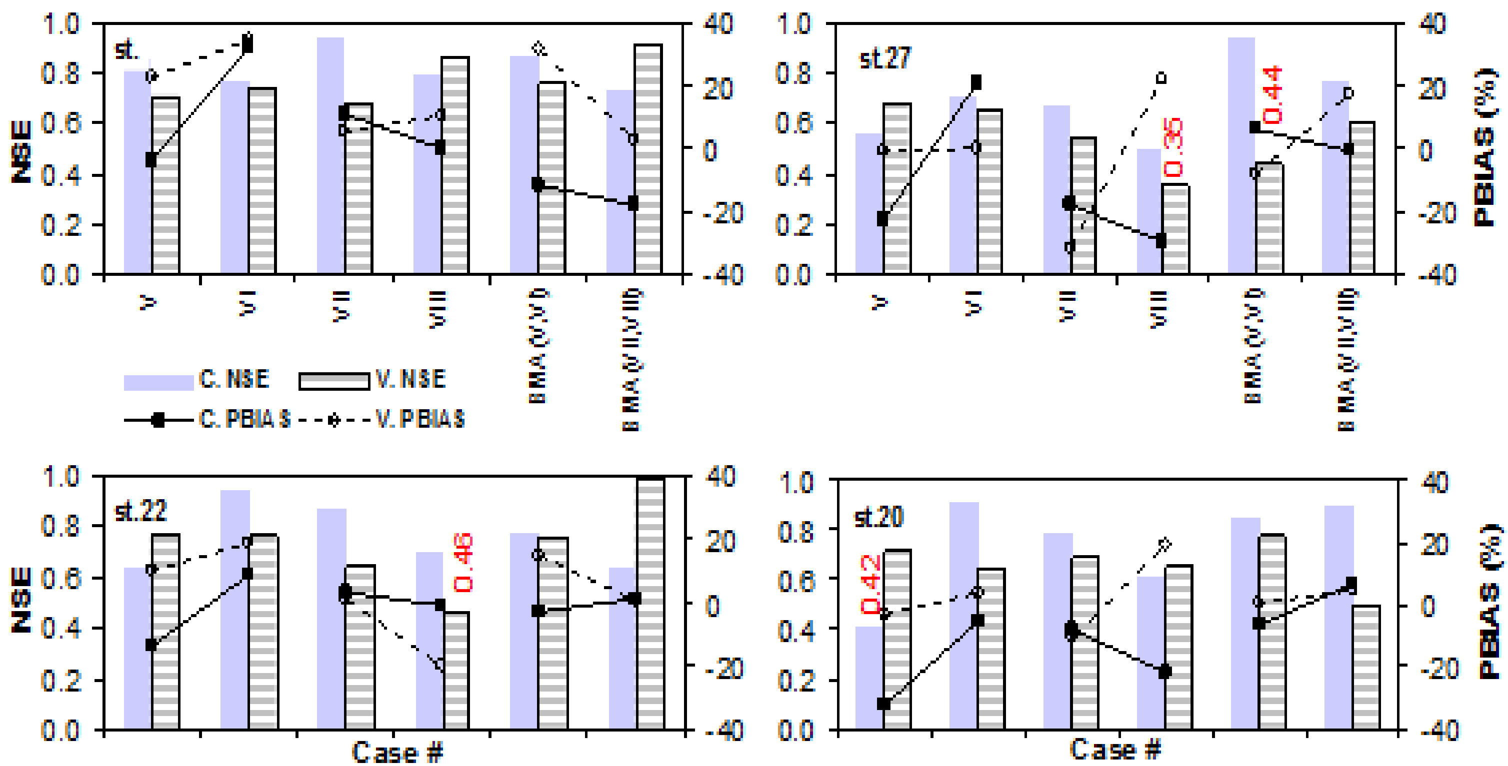
Results

NSE and PBIAS for calibration/validation periods for calibration cases I~IV at the 4 USGS stations (st.) for NO_3-N loss.



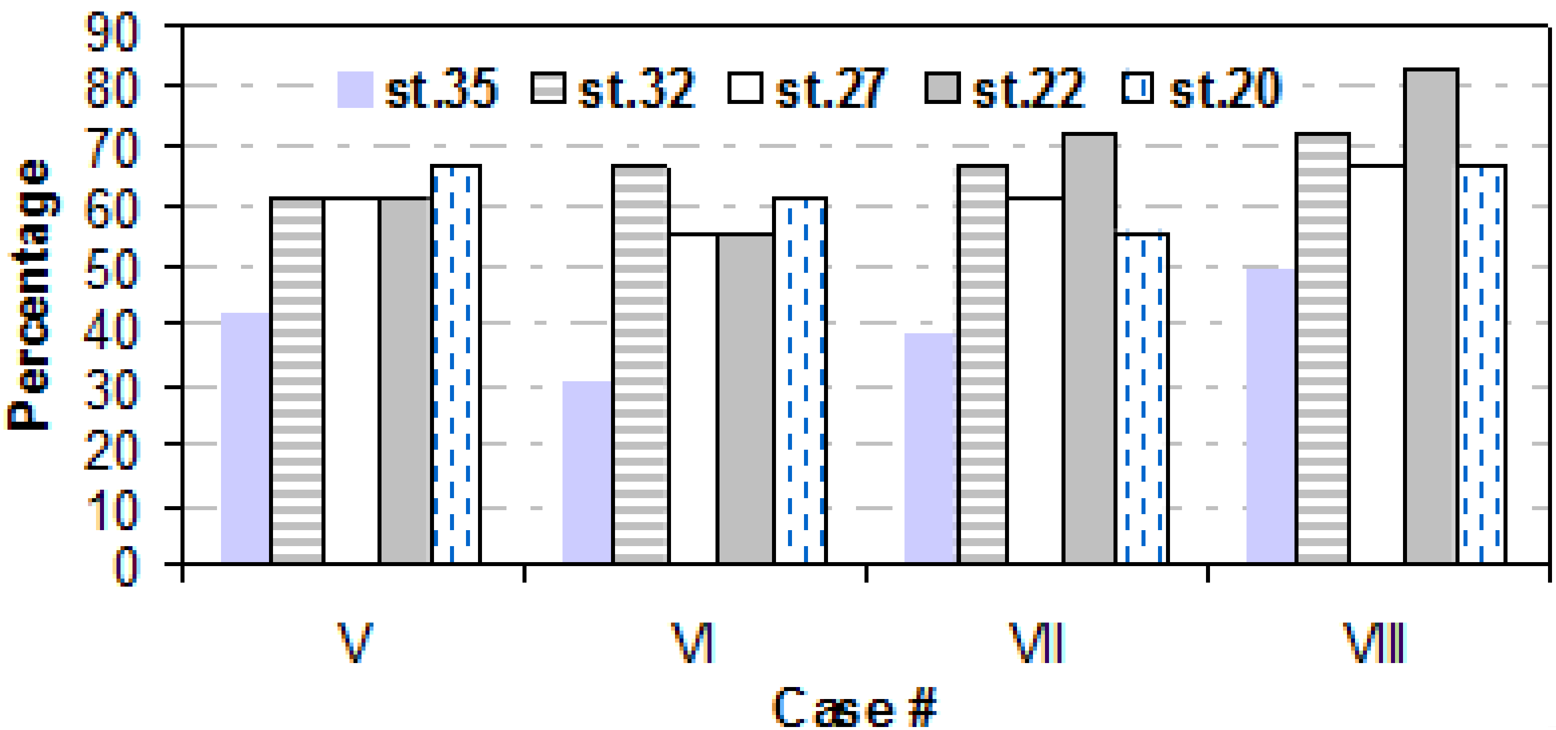
Results

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

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

- **Yen, H.**, Bailey, R. T., Arabi, M., Ahmadi, M., White, M. J., Arnold, J. G. **(2014b)**. "Evaluation of watershed model performance using general watershed Information: Beyond Typical Accuracy." ***Journal of Environmental Quality***, doi:10.2134/jeq2013.03.0110 (In Press).

- **Yen, H.**, M. J. White, M. Arabi, J. G. Arnold **(2014c)** "Evaluation of alternative surface runoff accounting procedures using the SWAT model." ***International Journal of Agricultural and Biological Engineering*** (In Press).

- Wang, X., **H. Yen**, J. Liu, Q. Liu **(2014)** "An auto-calibration tool for the Agricultural Policy Environmental eXtender (APEX) model." ***Transactions of the ASABE***, 54(7), pp. 1-13, doi: 10.13031/trans.57.10601.

Thanks for your attention!

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

Texas A&M AgriLife Research