On the Auto-calibration of Watershed Models: Multisite Many-Objective Measures of Information

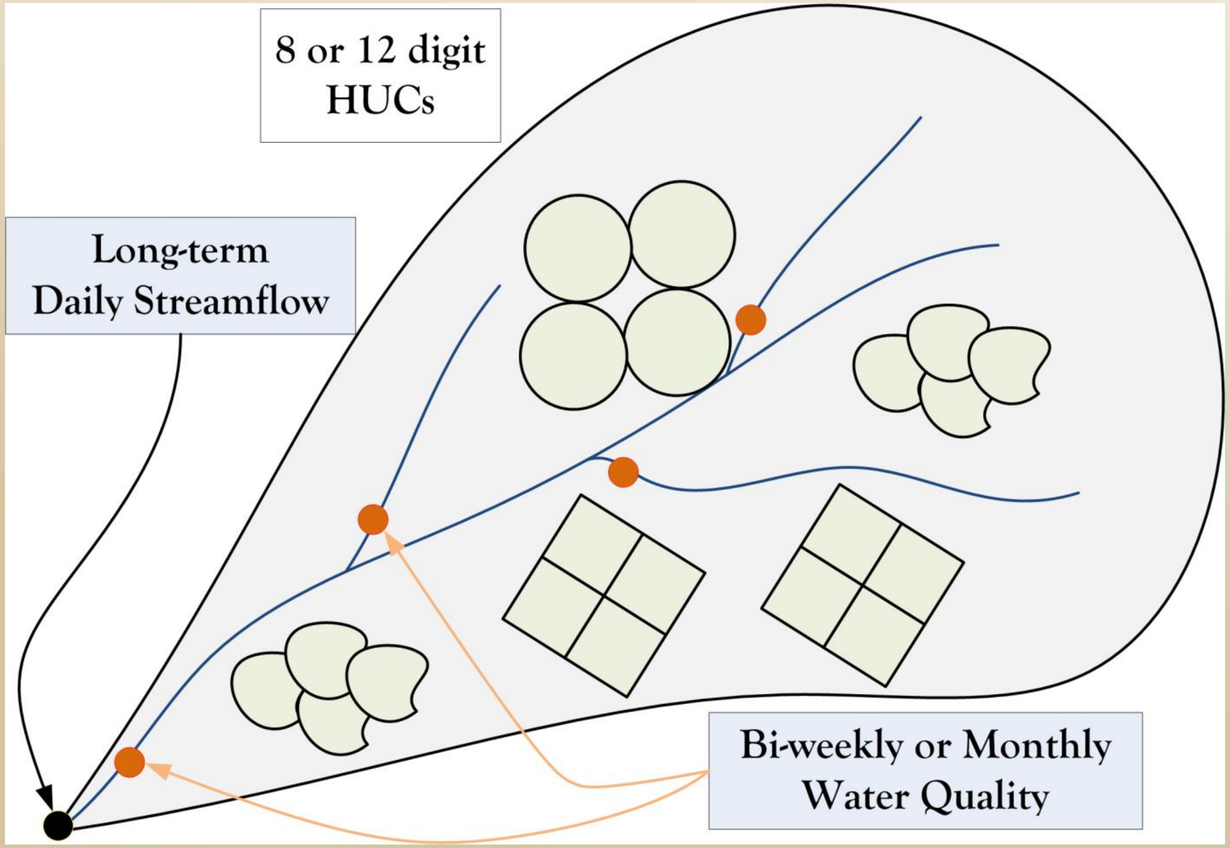
Mehdi Ahmadi
Mazdak Arabi
2013 SWAT Conference
Toulouse, France

Colorado State University

Model Calibration

With increased complexity of watershed models, efficient and effective use of observed data is vital for calibration of complex spatially distributed process-based models.

- ❑ Daily streamflow
- ❑ Weekly/monthly water quality



Multisite Many-Objective Calibration

- Aggregation of information for response variables at multiple sites into a single objective function of model errors.

- Hydrologic and water quality observations are characterized by varying measurement errors, varying sample size, and are typically noncommensurable.

- These considerations must be taken into account when using data in construction of the objective function.

Proposed Framework

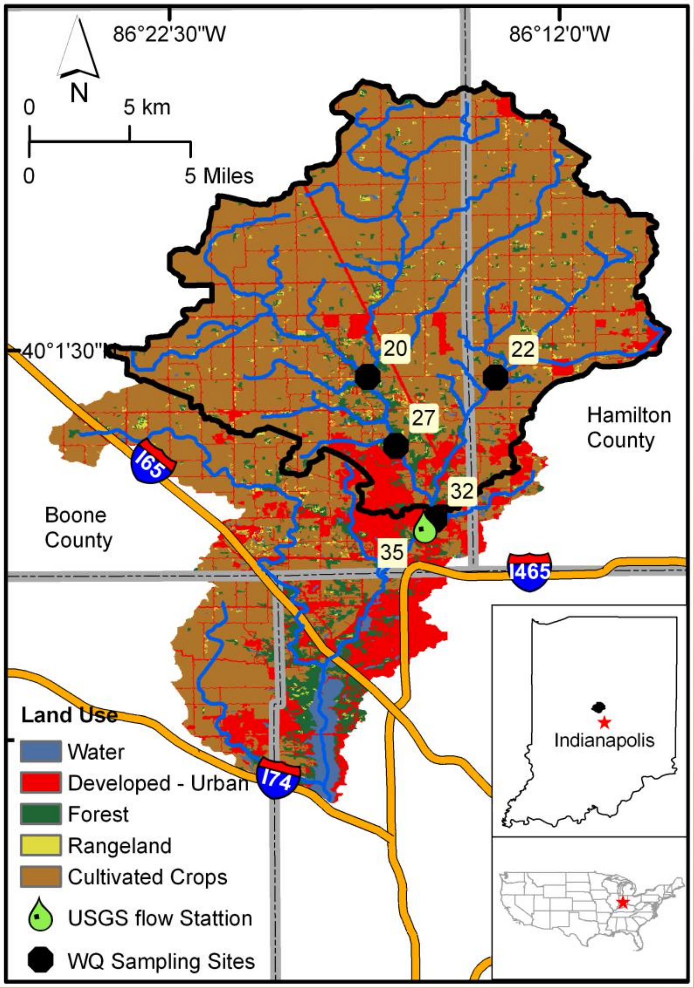
We propose a framework that consists of four major components to be used for calibration and evaluation of hydrologic and water quality models:

1. An a-priori characterization of system behavior;
2. A formal and statistically correct formulation of objective function(s) of model errors;
3. An efficient optimization method;
4. A multi criteria decision analysis (MCDA)

Study Area

Eagle Creek Watershed (ECW), Indiana

- Drainage area: 248.1 km^2
- Land use
  - 52% cropland
  - 27% pasture
  - 12% urban
  - 9% forest
- Observed data
  - Daily streamflow data at the outlet
  - Instantaneous WQ samples at multiple locations
  - WQ loads estimated using LOADEST



Characterization of system behavior

Behavioral solutions of a model comprise a subset of conceptually plausible responses that are judged by the analyst to be satisfactory according to past observations of the system under study:

- $0.65 < \text{GW Nitrate/Toal Nitrate} < 0.95$
- $5.0 < \textit{denitrification rate}\left(\frac{kg}{ha.yr}\right) < 50.0$

Performance Rating	Percent Bias (PBIAS, %)		Nash−Sutcliffe Efficiency (NSE)	
	Daily Streamflow	Monthly NOx	Daily Streamflow	Monthly NOx
Satisfactory	< 20	< 20	> 0.60	> 0.45
Good	< 15	< 15	> 0.65	> 0.50
Very Good	< 10	< 10	> 0.70	> 0.55

Objective Function of Model Errors

$$\hat{y} = M(\boldsymbol{\theta}) \quad \boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^n$$

$$\varepsilon(\boldsymbol{\theta}) = y - \hat{y} = y - M(\boldsymbol{\theta})$$

Bayesian Statistics:

$$P(\boldsymbol{\theta}|y) \propto P(\boldsymbol{\theta}) \cdot P(y|\boldsymbol{\theta})$$

Assuming $\varepsilon \sim N(0, \sigma_e)$

$$L(\boldsymbol{\theta}|y) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_e^2}} \exp\left[-\frac{(y_i - \hat{y}_i(\boldsymbol{\theta}))^2}{2\sigma_e^2}\right]$$

Copy with auto-correlated and non-Guassian errors

Using successive log and AR(1) transformations:

$$\ell(\boldsymbol{\theta}|y) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln\frac{\sigma_{\vartheta}^{2n}}{1-\rho^2} - \frac{1}{2}(1-\rho^2).\sigma_{\vartheta}^{-2}[\hat{y}_1(\boldsymbol{\theta}) - y_1]^2 - \frac{1}{2}\sigma_{\vartheta}^{-2}.\sum_{i=2}^{n}\{(y_i - \rho y_{i-1}) - [\hat{y}_i(\boldsymbol{\theta}) - \rho\hat{y}_{i-1}(\boldsymbol{\theta})]\}^2$$

Box-Cox or other transformation of responses may also help with the issue of heteroscedasticity.

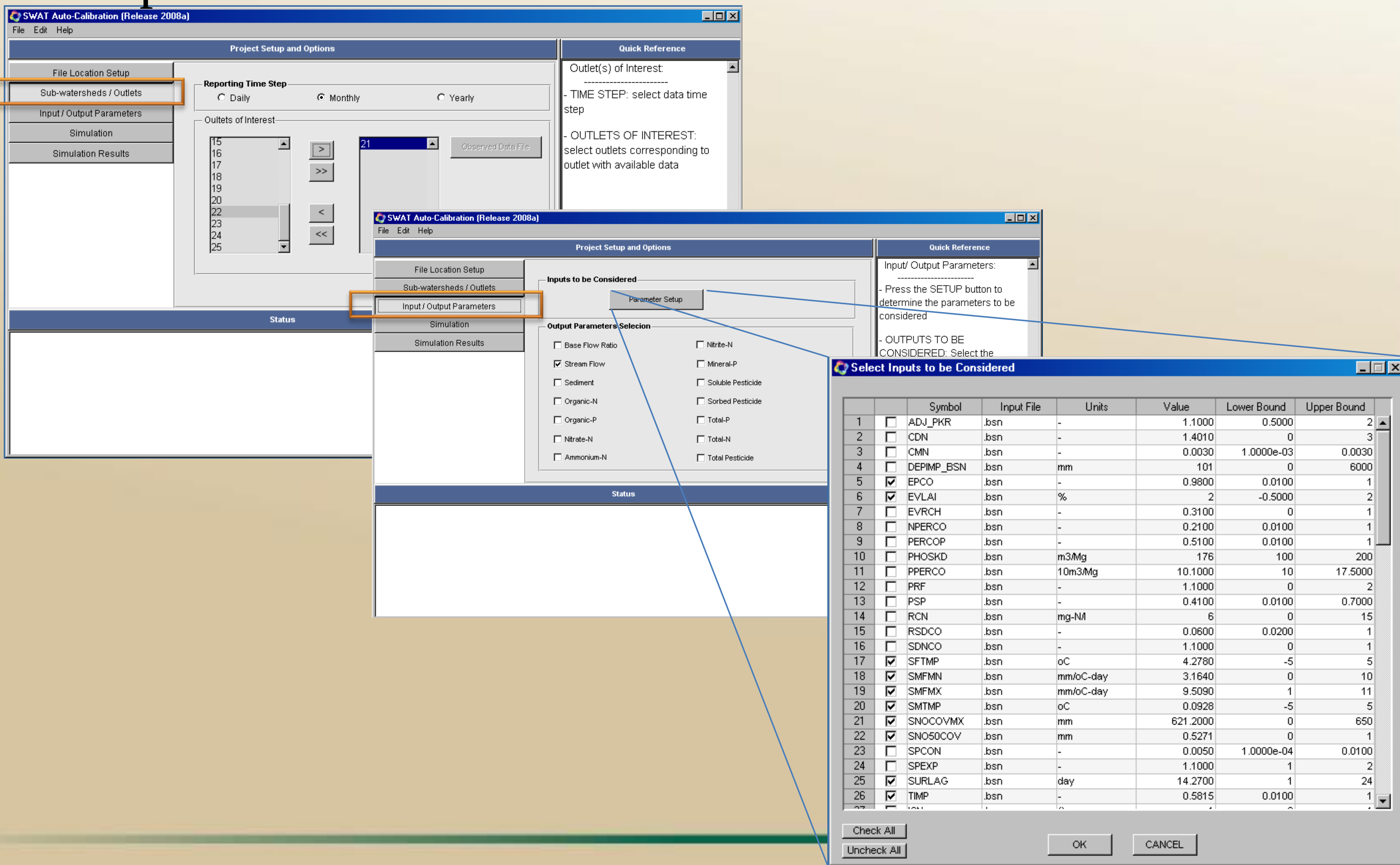
Parameter Estimation Technique

- Single objective methods (all information are aggregated)
  - Shuffled Complex Evolutionary (SCE)
  - Dynamically Dimensioned Search (DDS)
  - Differential Evolution Adaptive Metropolis (DREAM)

- Multi-objective methods: Nondominated Sorted Genetic Algorithm II (NSGA-II)
  - Two-objective (2OF NSGA-II): Streamflow responses at the outlet and Nitrate data are aggregated
  - Five-objective (5OF NSGA-II): Streamflow responses at the outlet and nitrate responses at 4 stations

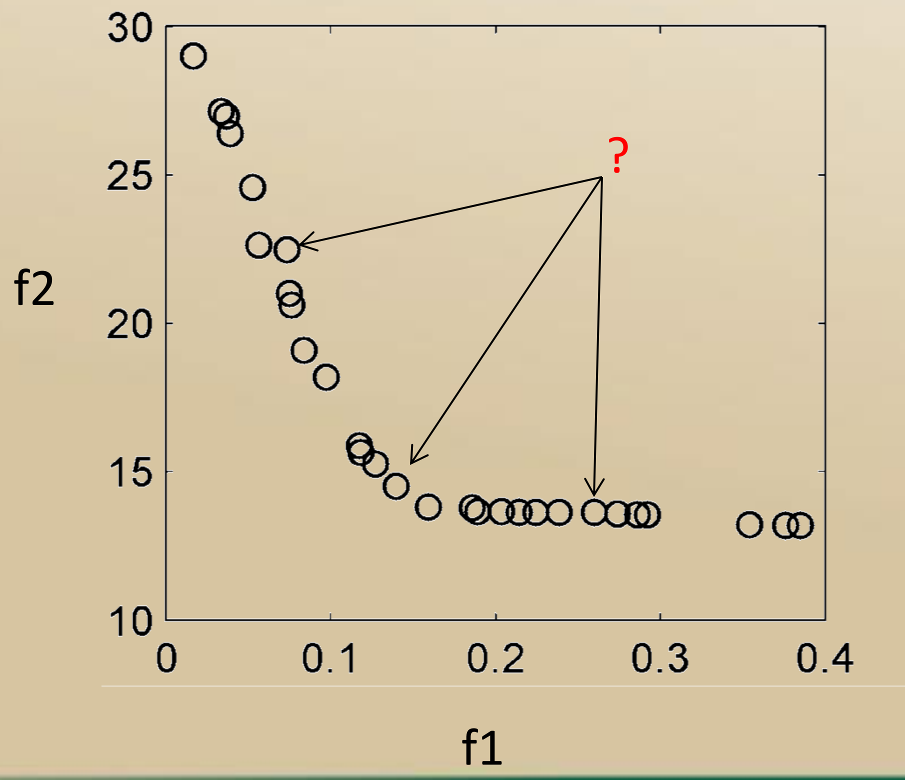
The Calibration Tool in MATLAB

3. An optimization method: Auto-calibration tool

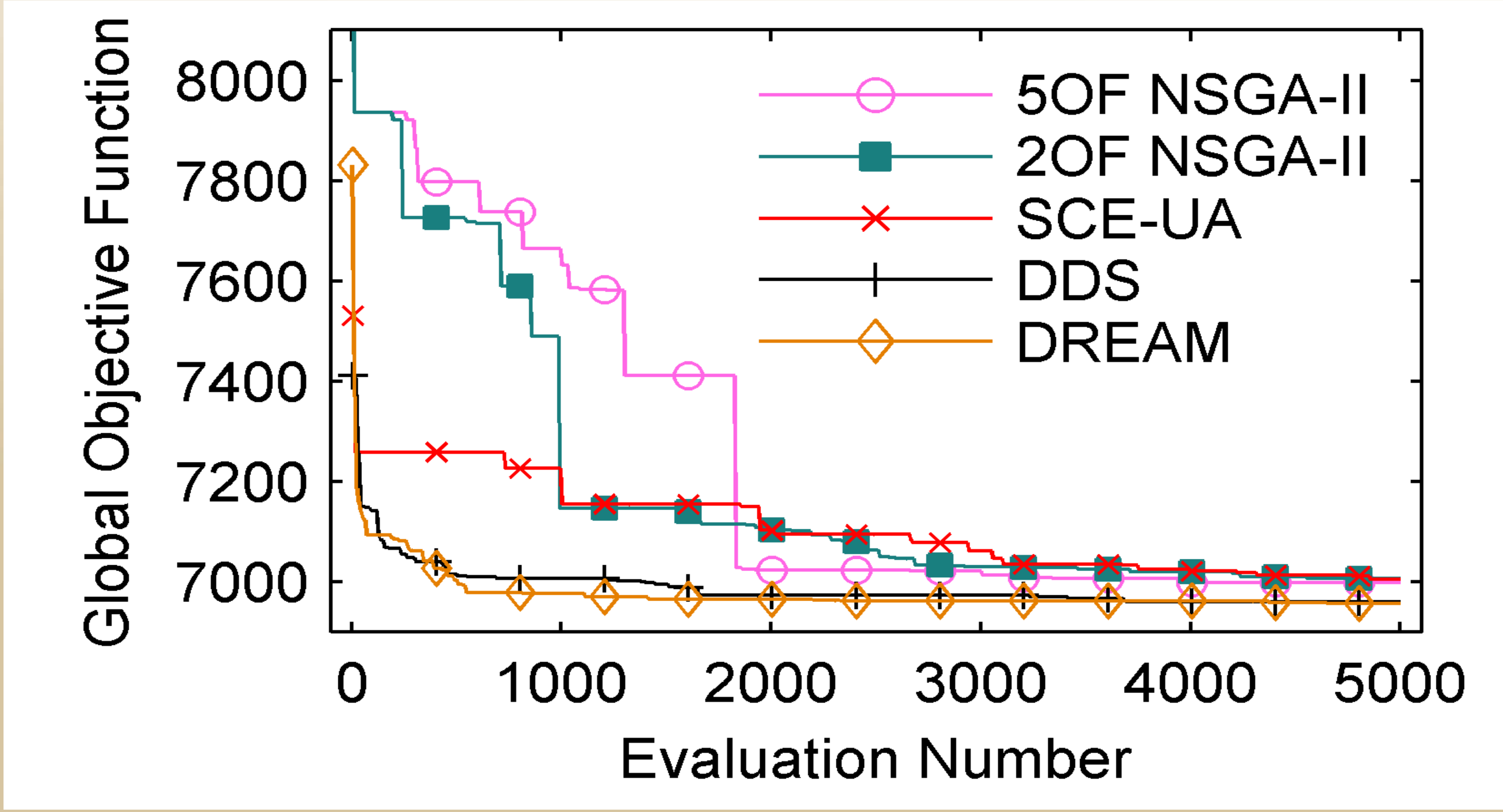


Multi Criteria Decision Analysis

MCDA provides an objective approach for selection of non-dominated solutions from the Pareto-optimal front that are most consistent with the goals of the modeling study.

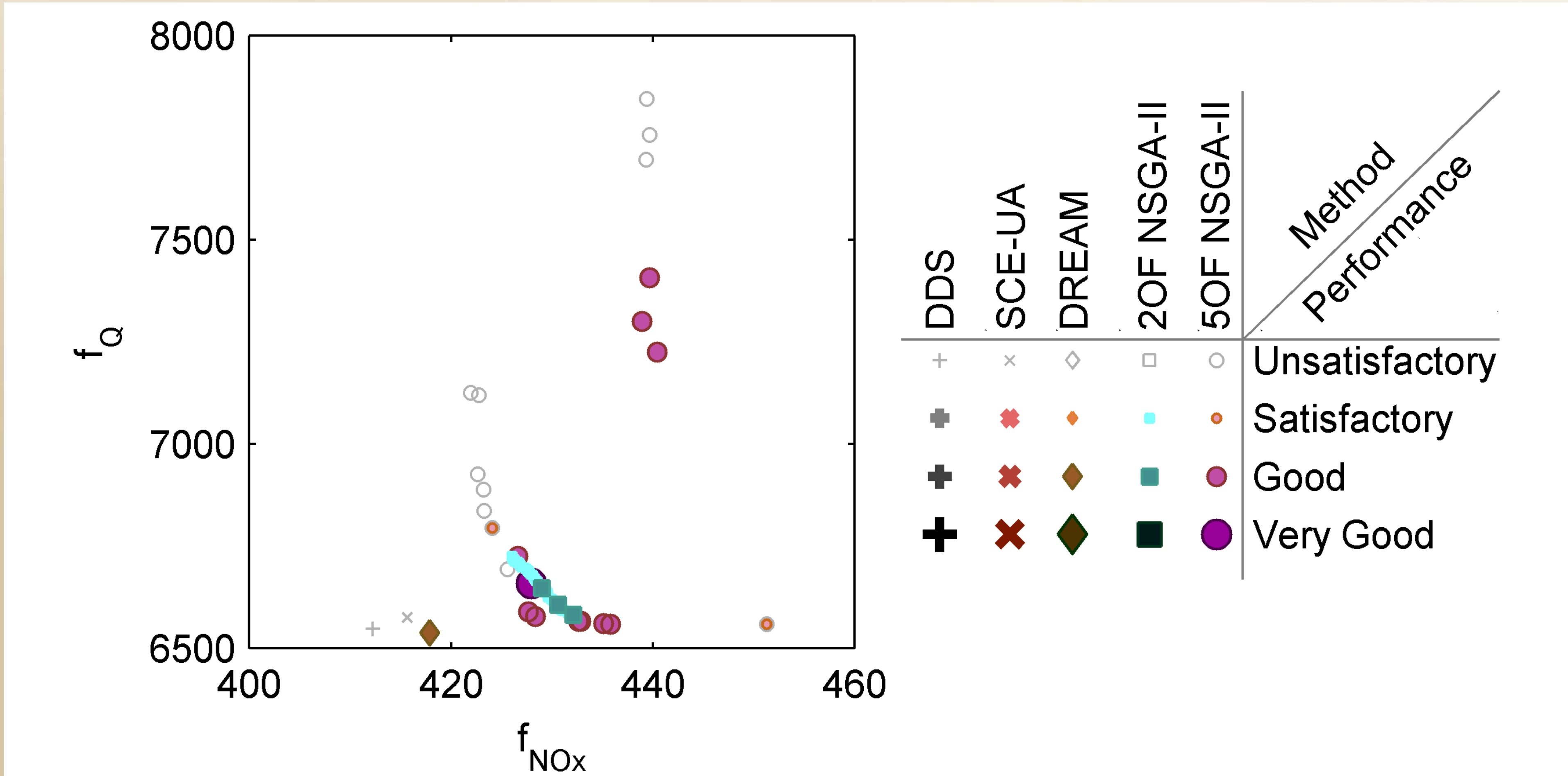


Results: Effectiveness vs. Efficiency

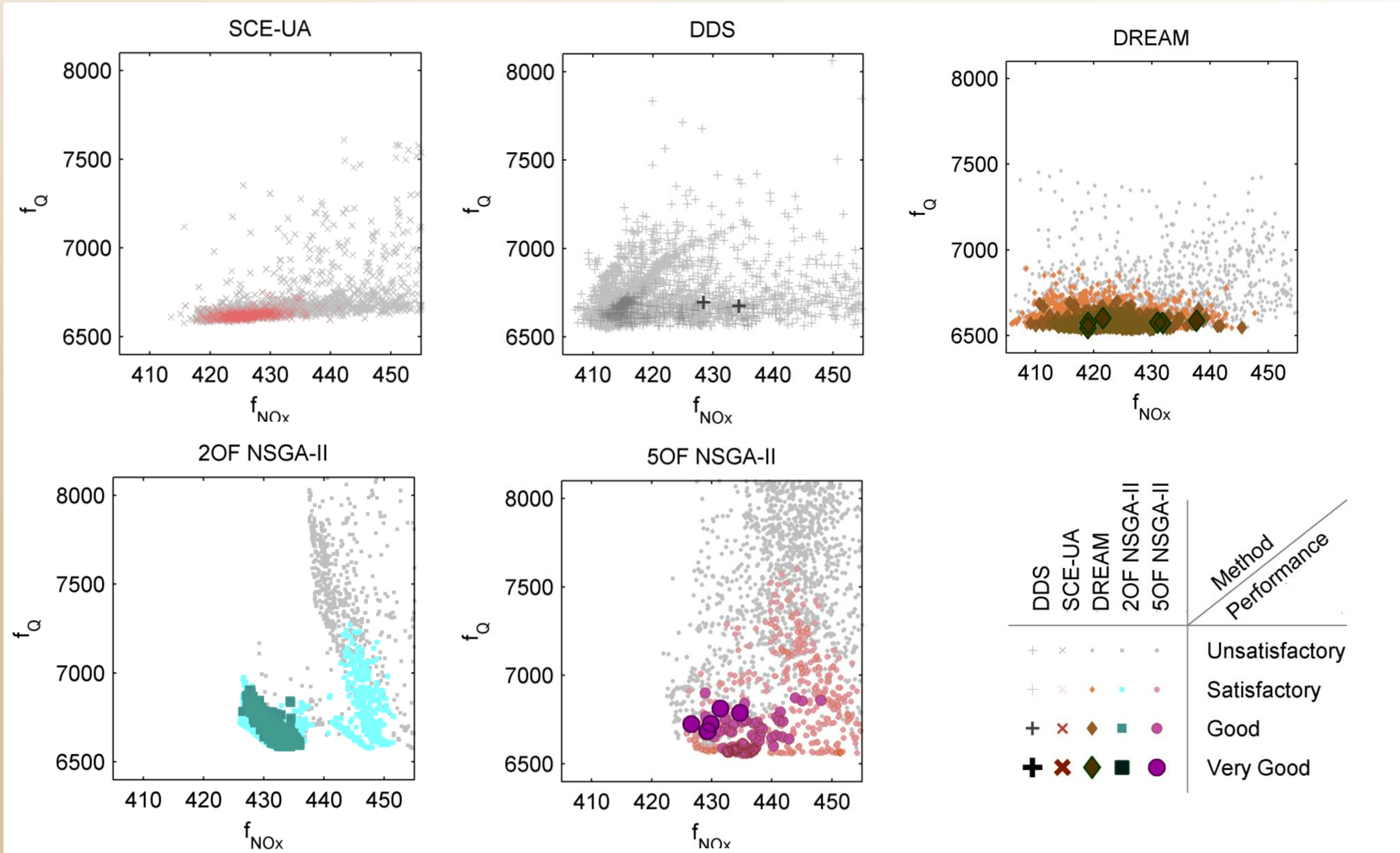


environmental Risk Assessment & Management System (eRAMS)

Results: Effectiveness vs. Efficiency



Behavioral Suboptimal Solutions



Conclusion

- For multisite-multiobjective automatic calibration of a watershed model, both a **formal likelihood function** considering the structure of residuals and a **multiobjective optimization approach** are essential

- This is particularly required when a strict definition of **system behavior** is considered.

- The use of the solutions from the single objective techniques was limited because the simulations did not mimic the observed behavior of the system for all objectives at all sites

Conclusion

- Results of 2OF and 5OF NSGA-II suggest that the aggregation of information for the same response variable (nitrate in this study) at different observational sites using the proposed likelihood function appeared as a pragmatic approach for enhancing the speed of convergence to the Pareto-optimal front.