The power of multi-objective calibration: Two case studies with SWAT

Sander Huisman

Thorsten Pohlert, Lutz Breuer & Hans-Georg Frede

SFB 299

Justus-Liebig-Universität Gießen, Germany

Institut für Landschaftsökologie und Ressourcenmanagement

The power of multi-objective calibration: A case study with SWAT

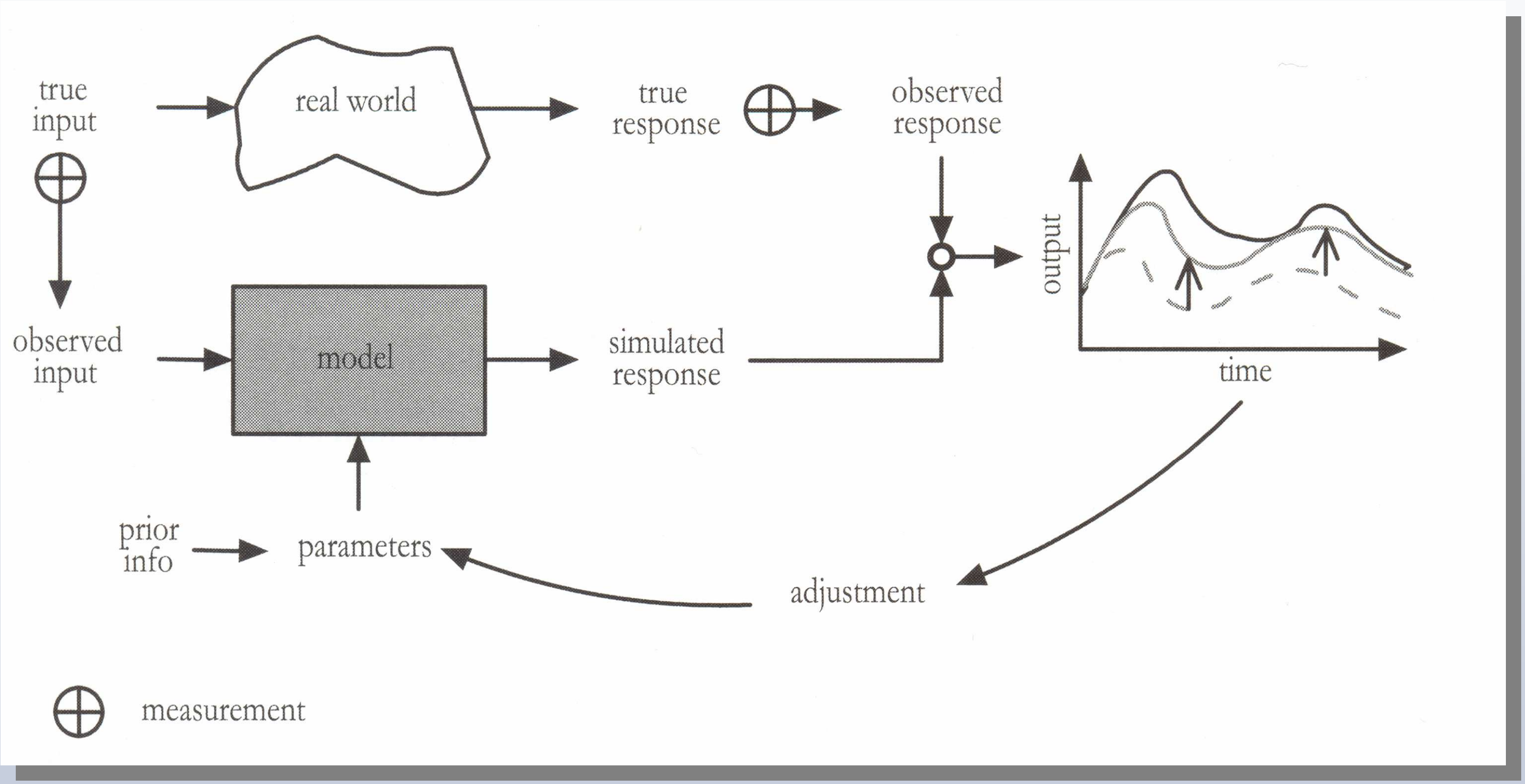
Sander Huisman

Thorsten Pohlert, Lutz Breuer & Hans-Georg Frede

Justus-Liebig-Universität Gießen, Germany

Institut für Landschaftsökologie und Ressourcenmanagement

Introduction to Calibration



Vrugt (2004)

Manual Calibration

- → Most widely used calibration method
- → Visual comparison of measured and simulated data
- → Semi-intiutive trial and error process for parameter adjustment
- → Closeness implicitly evaluated with several (>3) criteria
- → Excellent model calibrations, but manual calibration...
  - is highly labor-intensive (human resources)
  - is difficult to learn
  - procedures are model-dependent
  - results are user-dependent

Automatic Optimization

→ Algorithms that optimize an objective function by systematically searching the parameter space according to a fixed set of rules

- Local search algorithms
  - Nelder and Mead (Simplex) algorithm
  - Levenberg-Marquardt
  - Gauss-Newton
- Global search algorithms
  - Simulated Annealing
  - Genetic algorithms
  - Shuffled Complex Evolution

Objective Functions

→ Automation of calibration requires the formulation of „closeness“ measures (objective functions)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(d_i - o_i(\theta)\right)^2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|d_i - o_i(\theta)\right|$$

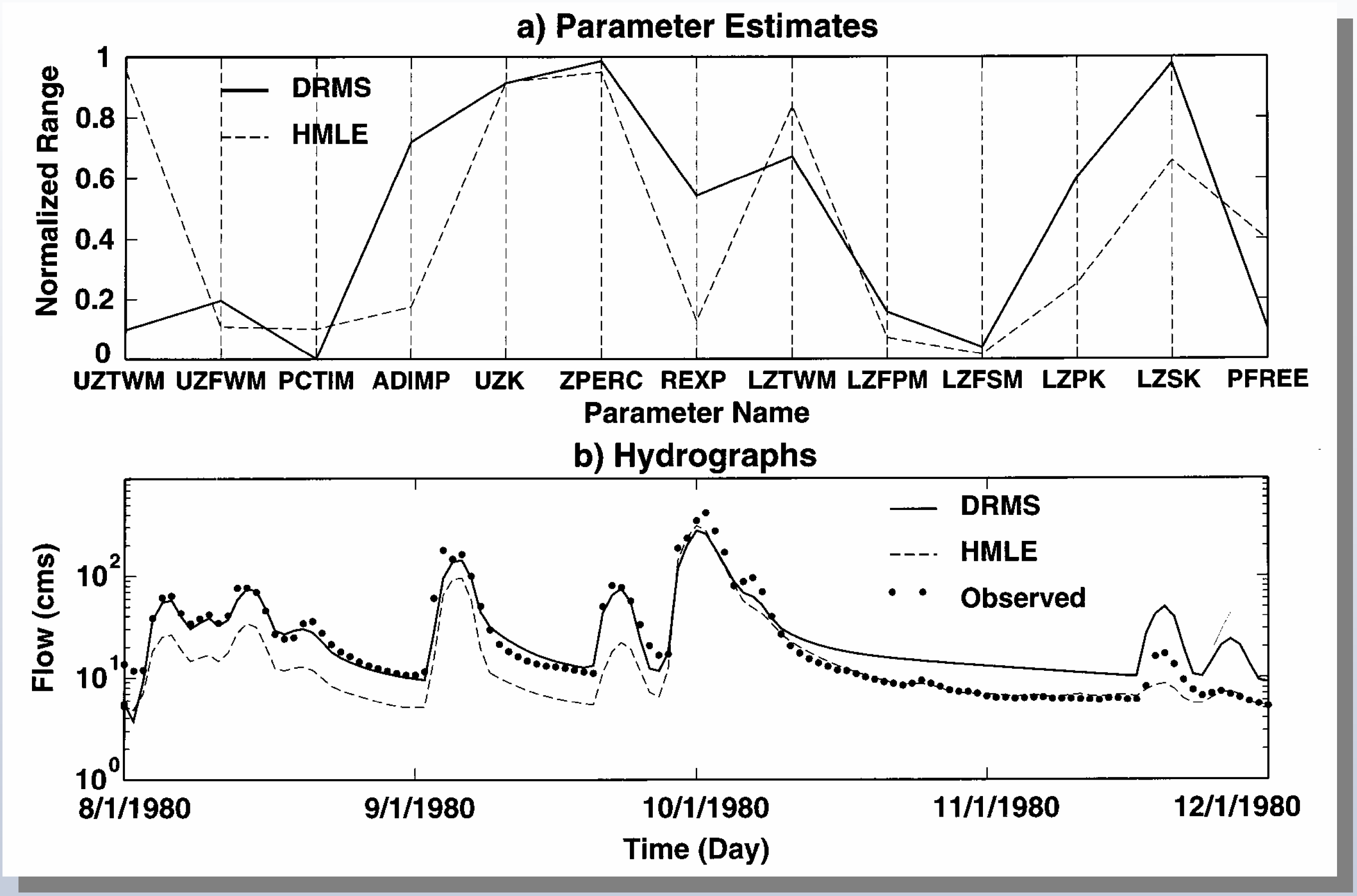
$$NS = 1 - \frac{\frac{1}{n}\sum_{i=1}^{n}\left(d_i - o_i(\theta)\right)^2}{\frac{1}{n}\sum_{i=1}^{n}\left(d_i - \bar{d}_i\right)^2}$$

$$NSC = Number\ of\ sign\ changes$$

$$BIAS = \frac{1}{n}\sum_{i=1}^{n} d_i - o_i(\theta)$$

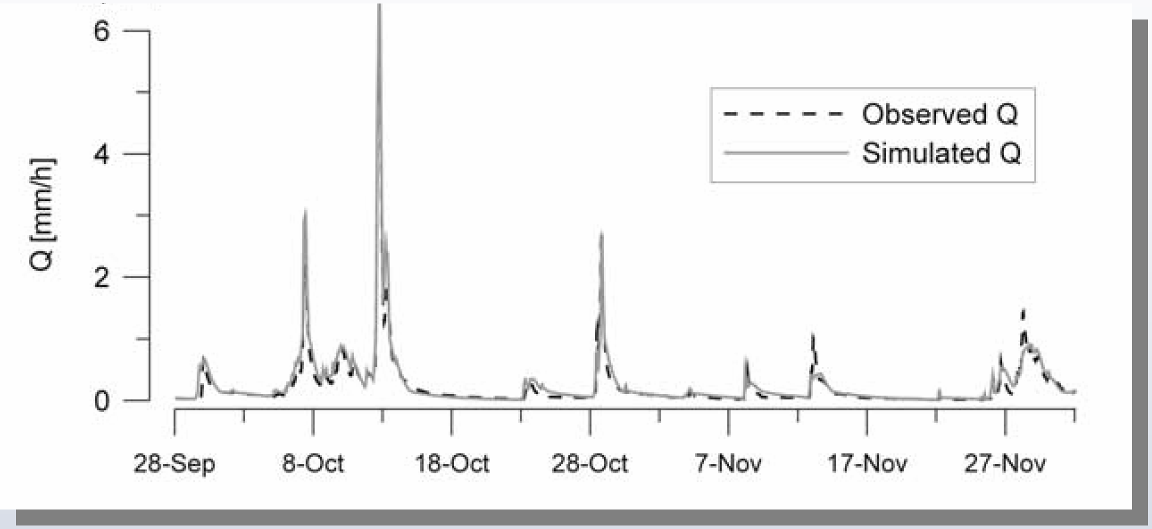
$$TMVOL = \sum_{j=1}^{n_{month}}\left(\frac{1}{n_{day}(i)}\frac{1}{n}\sum_{i=1}^{n_{day(i)}}\left(d_i - o_i(\theta)\right)\right)^2$$

Different evaluation criteria



Gupta et al. (1998)

Different output variables

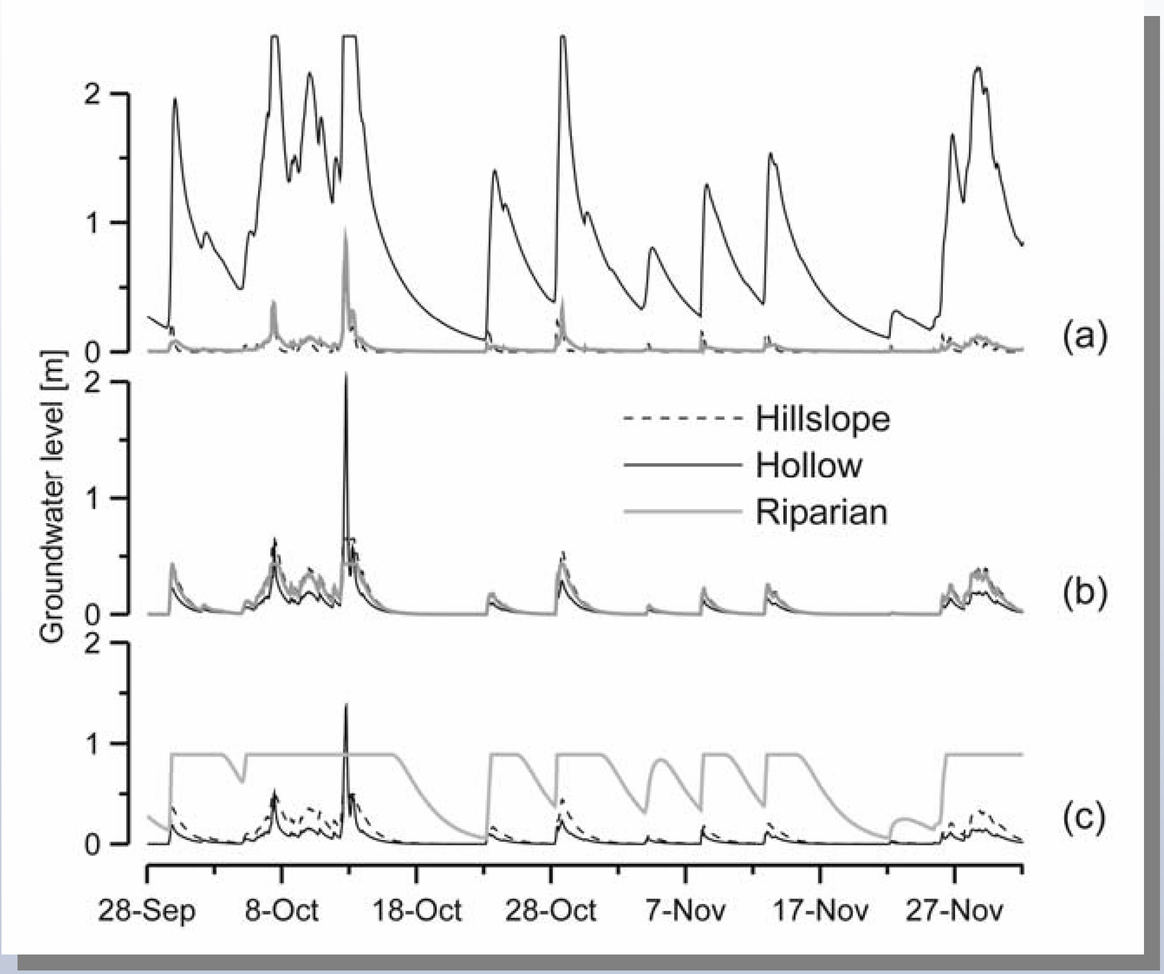


Seibert and McDonell (2002)

→ Calibration against discharge (Q) with a high efficiency (~0.93)

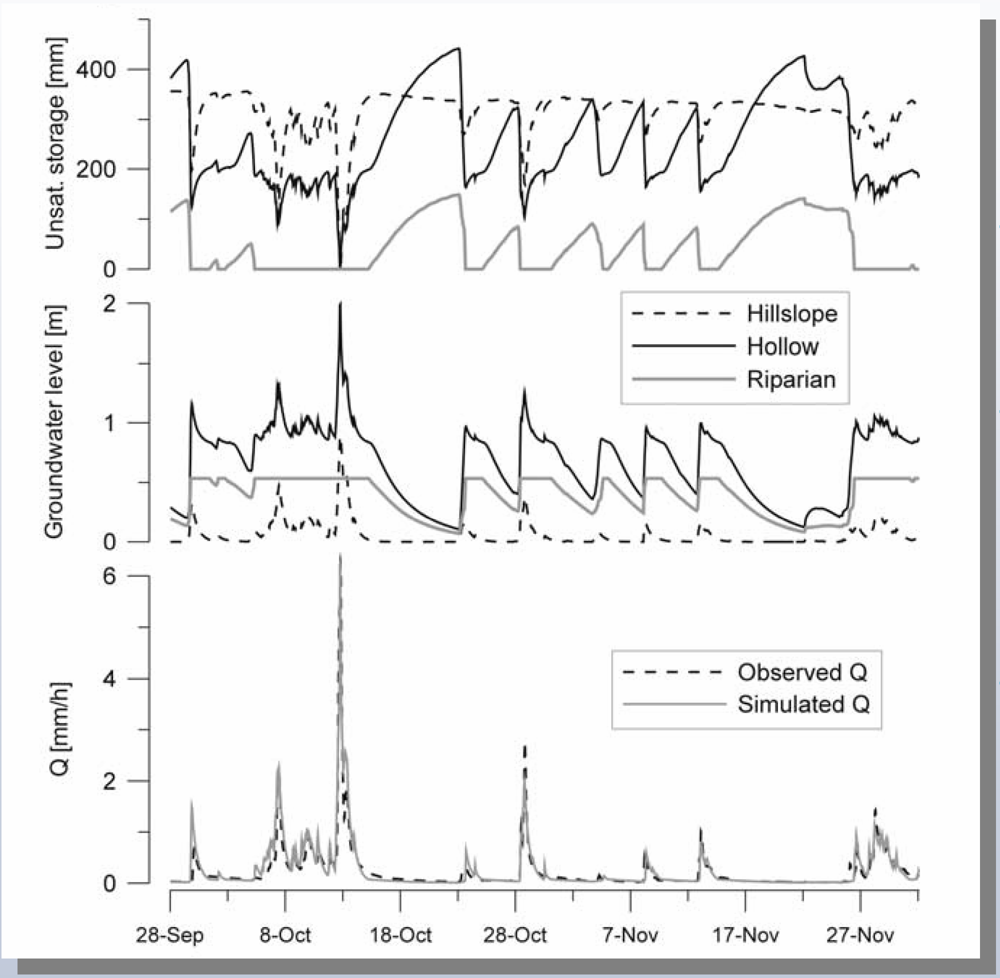
→ Does a good fit to measured discharge result in good predictions of other state variables?

Right for the Wrong Reasons



Seibert and McDonell (2002)

Less Right for the Right Reasons



Seibert and McDonell (2002)

- → Best overall agreement for GW and discharge
- → Model efficiency for discharge decreased from 0.93 to 0.84
- → Consistency with perceptual model strongly increased

Multi-Objective Calibration

→ Different objective functions result in different optimal parameters

- Is there an „optimal parameter set“ when using multiple objectives ?

→ The optimal parameter set for one signal might not be the best parameter set for another signal

- Can we find a comprimise that is satisfactory for all signals?

→ Three common approaches to multi-objective calibration

- Pareto-optimal parameter sets
- Aggregation of single objectives to a global objective criterion
- GLUE methodology

Calibration Lysimeter Brandis

SFB 299

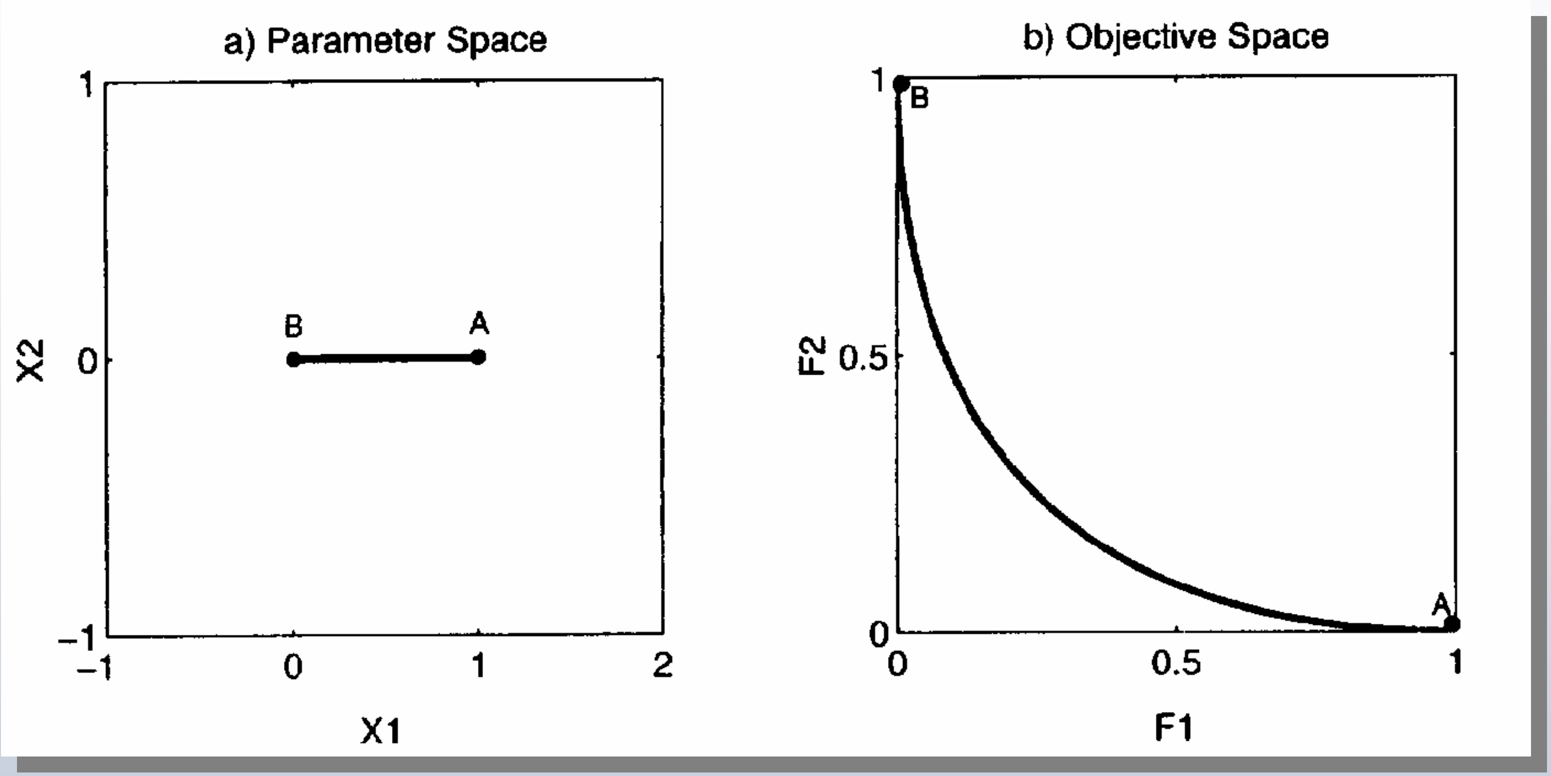
Table 1. Prior parameter ranges used in the multi-objective calibration for the lysimeter data.

Model Parameter	Lower limit	Upper limit
*Bulk density (BD, g cm^{-3})	1.45	1.55
*Available water content (AWC, -)	0.16	0.19
*Saturated hydraulic conductivity (K_{sat}, mm hr^{-1})	0	750
*Moist soil albedo (ALB, -)	0.20	0.30
Maximum rooting depth (RD_{max}, mm)	500	2000
Soil evaporation compensation factor (ESCO, -)	0.1	1.0
Plant uptake compensation factor (EPCO, -)	0.0	1.0
Rate factor for humus mineralization (CMN, -)	0.00	0.01
Residue decomposition factor (RSDCO, -)	0.00	0.10
Maximum daily denitrification rate (MAX_WDN, kg ha^{-1} d^{-1})	0.0	0.3
Maximum daily nitrate uptake (MAX_NUP, kg ha^{-1} d^{-1})	0.0	10.0

→ Calibration with 12 years of data for mean monthly percolation, evapotranspiration and nitrate leaching

→ Calibration against two evaluation criteria (bias and NS-efficiency)

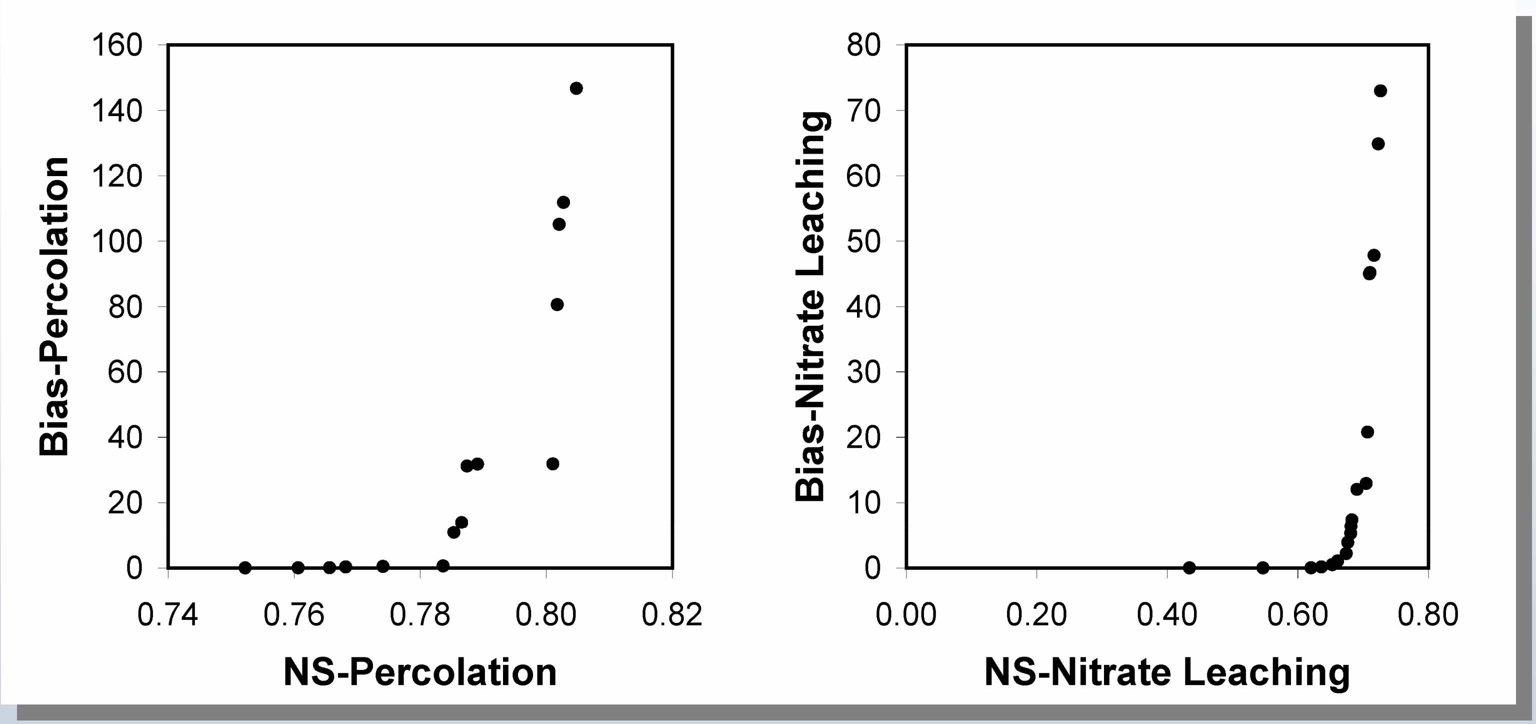
Pareto-Optimal Parameter Sets



Yapo et al. (1998)

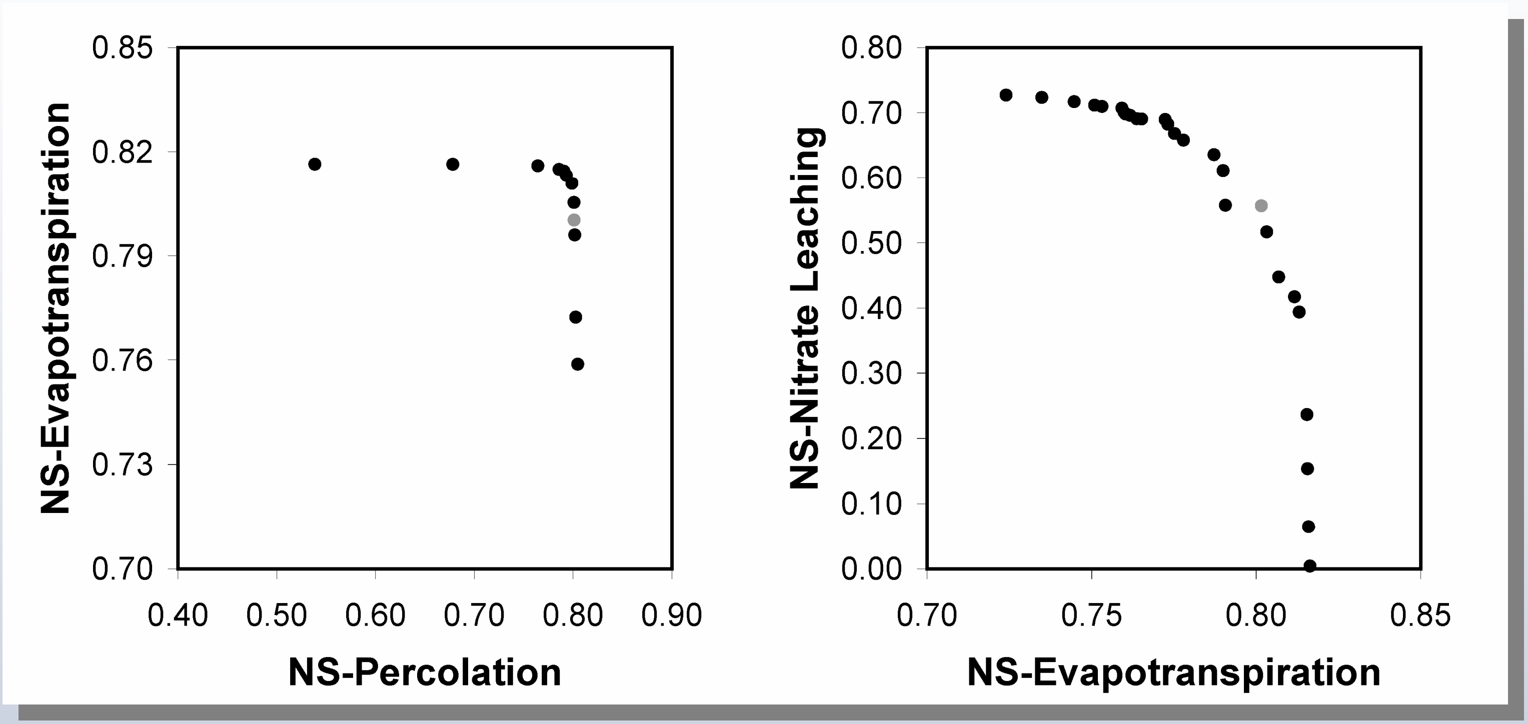
→ Algorithms to find Pareto front are MOCOM-UA, MOSCEM-UA and the approximation of Madsen (2000).

Pareto Front: Multiple Criteria



→ Excessive calibration on one evaluation criteria leads to a detoriation of other criteria

Pareto Front: Multiple Variables



→ Excessive calibration on one output variable leads to a detoriation of other output variables

Aggregation to Global Objective Criteria

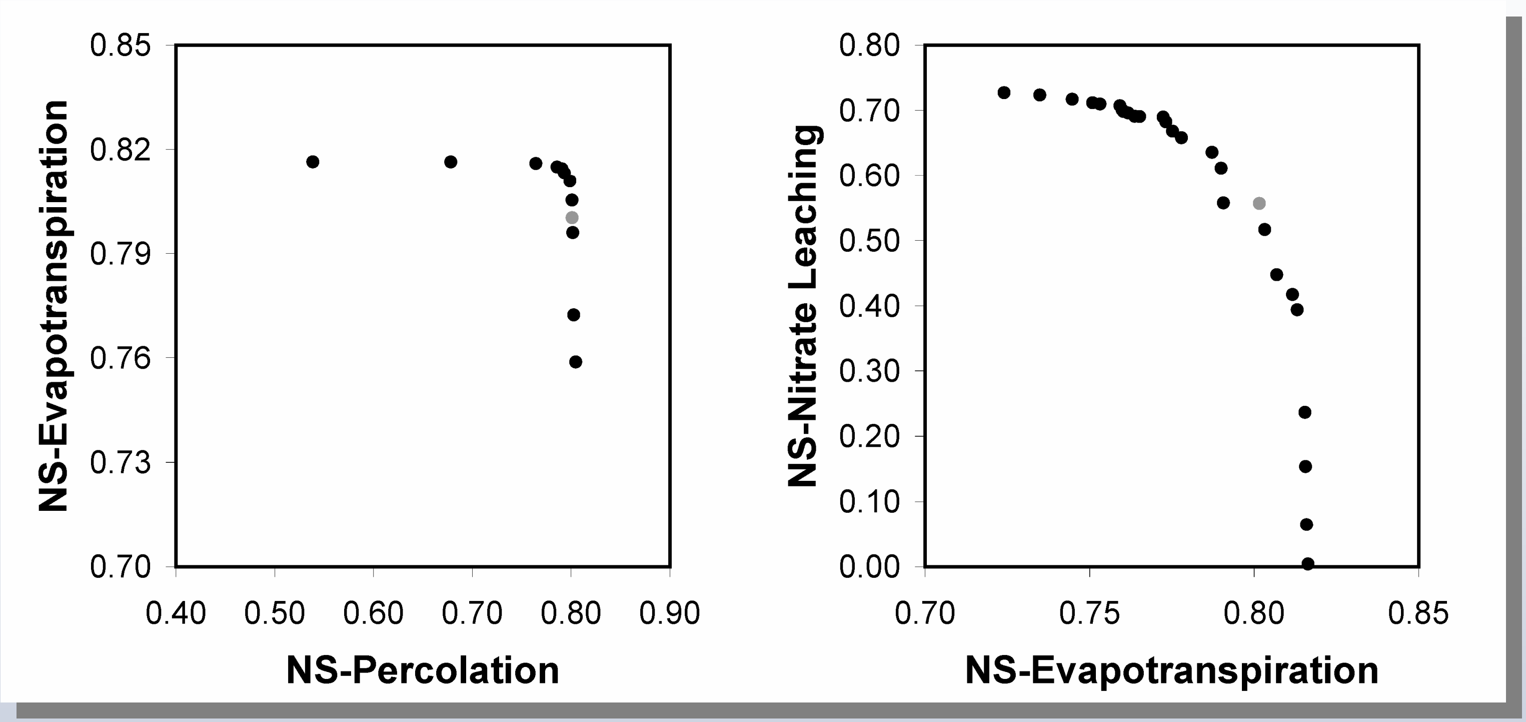
→ Pareto front esitmation more computationally expensive than single objective calibration (10000 vs. 3000)

→ Aggregation to save computation time:

$$GOC_j = \sum_{i=1,m} f_i(OF_{i,j})$$

→ Weights are based on mean and standard deviation of random sample from parameter space

Aggregation to Global Objective Criteria



→ Aggregation often leads to a good trade-off solution, but provides no information on how the strong the compromise is.

GLUE methodology

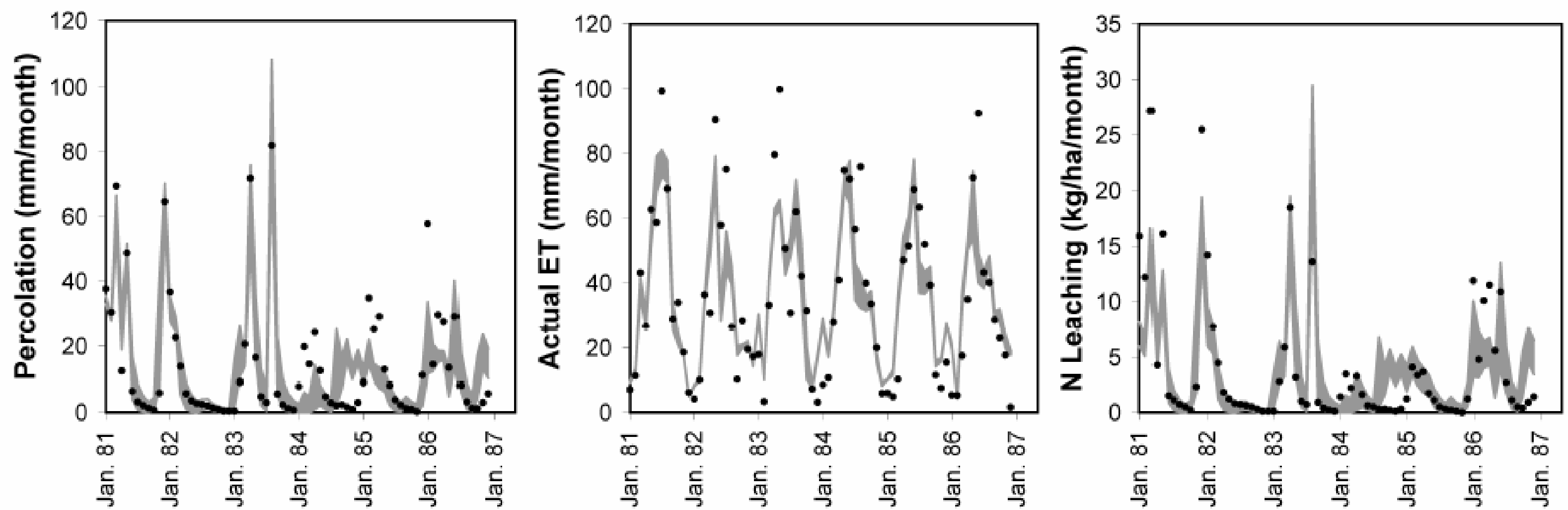
→ Many parameter combinations that lead to acceptable simulations (equifinality)

- Step 1: define ranges for each model parameter
- Step 2: define acceptable
- Step 3: run model
- Step 4: reject unacceptable
- Step 5: iterate until enough acceptable runs are obtained

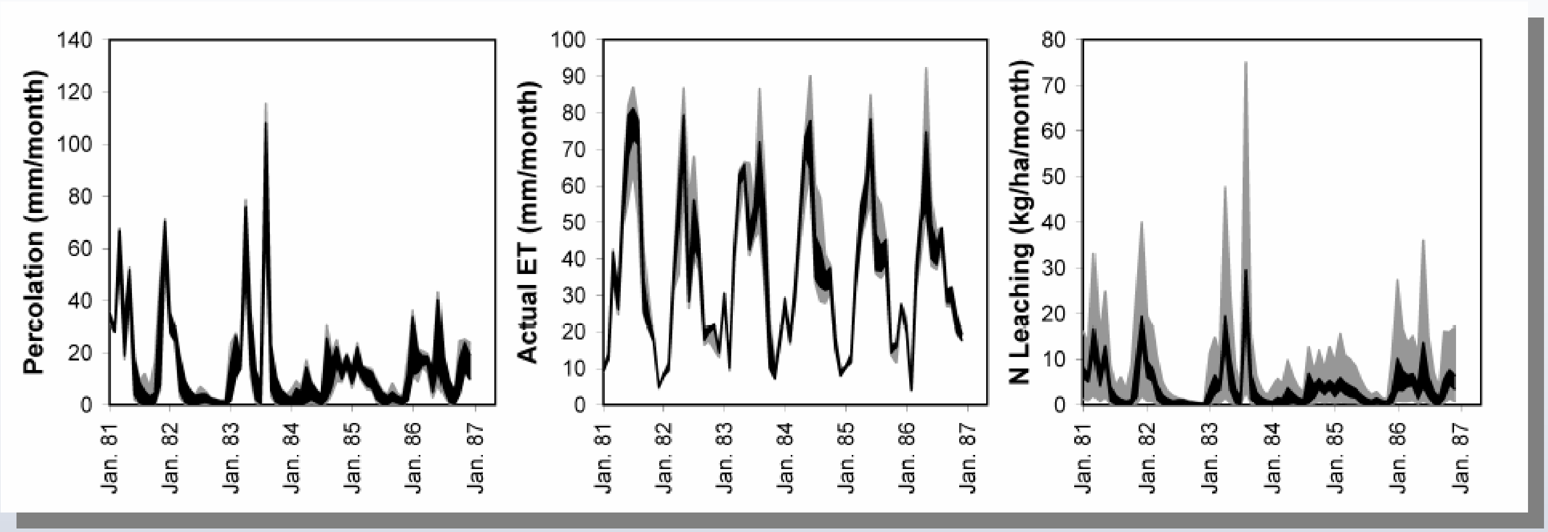
Results of GLUE Methodology I

Table 3. Number of behavioral model runs in the GLUE analysis and the percentage of model runs retained after considering six successive performance criteria.

Total	Percolation (NS > 0.7)	Percolation (Bias < 50)	Actual ET (NS > 0.7)	Actual ET (Bias < 50)	N-leaching (NS > 0.7)	N-leaching (Bias < 50)
157502	78568	3145	2545	2048	157	97
	49.9%	4.0%	80.9%	80.5%	7.7%	61.8%



Results of GLUE Methodology I



→ Percolation and ET can be simulated correctly despite inadequate representation of N (remains of the modular structure of SWAT).

Conclusions

→ Use of multiple objectives makes you realize the deficiencies of single objective calibration

→ Aggregation of objective functions is a relatively computationally inexpensive method to find a decent compromise solution.

→ More advanced methods can indicate model deficiencies, model weaknesses, etc.

[redacted]