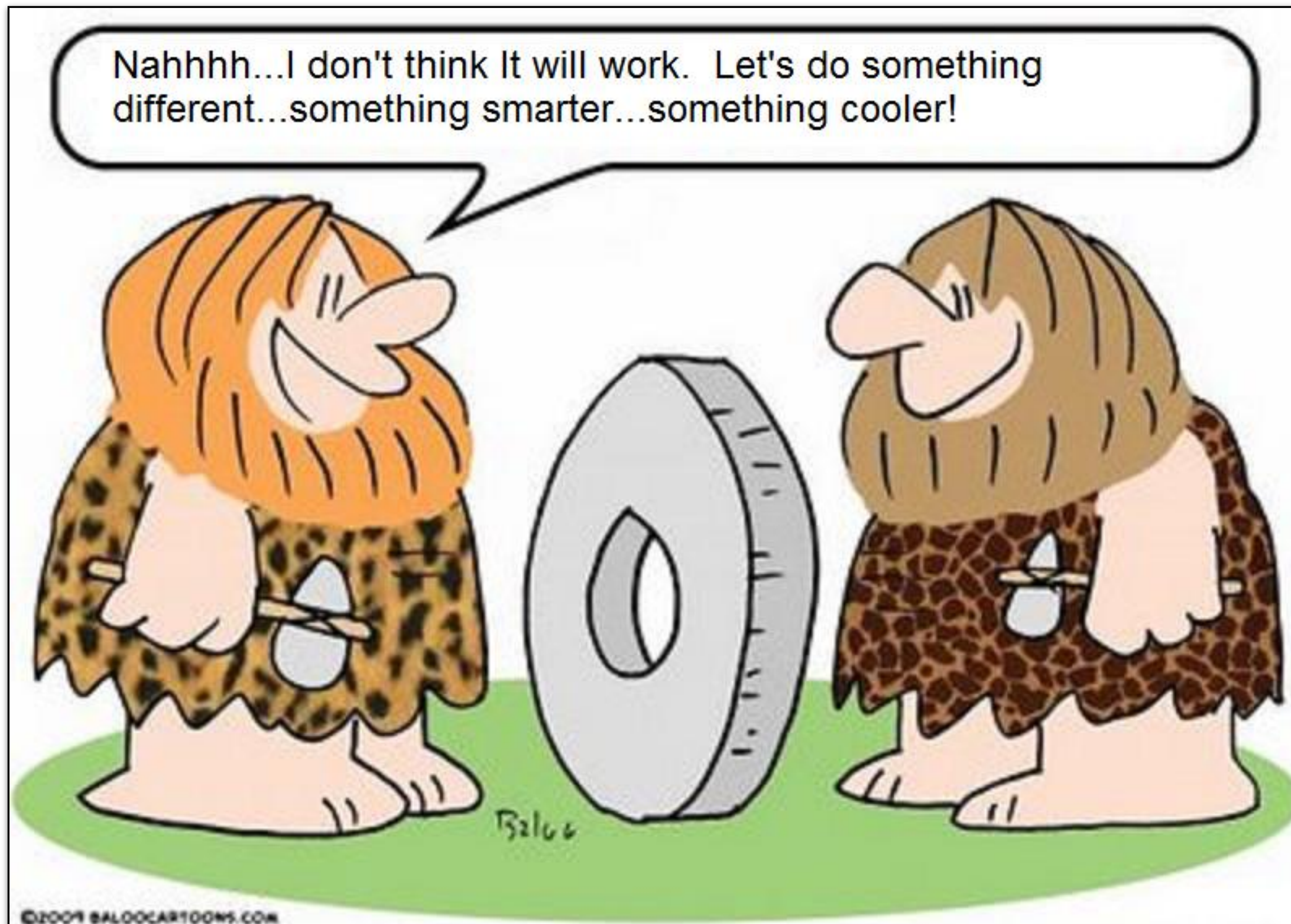
STEPS: Modeling for Problem Solving in ES

1. **Clearly define your goal (question you want to answer, hypothesis you want to test, prediction you want to make) - as precisely as possible**
2. Design or Select your model
3. Implement the model
4. Evaluate the model and quantify uncertainty
5. Apply the model to the goal
6. Communicate model results

Off-the shelf models

Useful because they usually have had multiple experts developing, testing and refining over time so you don’t have to re-invent the wheel



Model Selection

Is model structure APPROPRIATE

Is model performance GOOD ENOUGH

Then pick the simplest model that accounts for your processes of interest

Model Selection 1

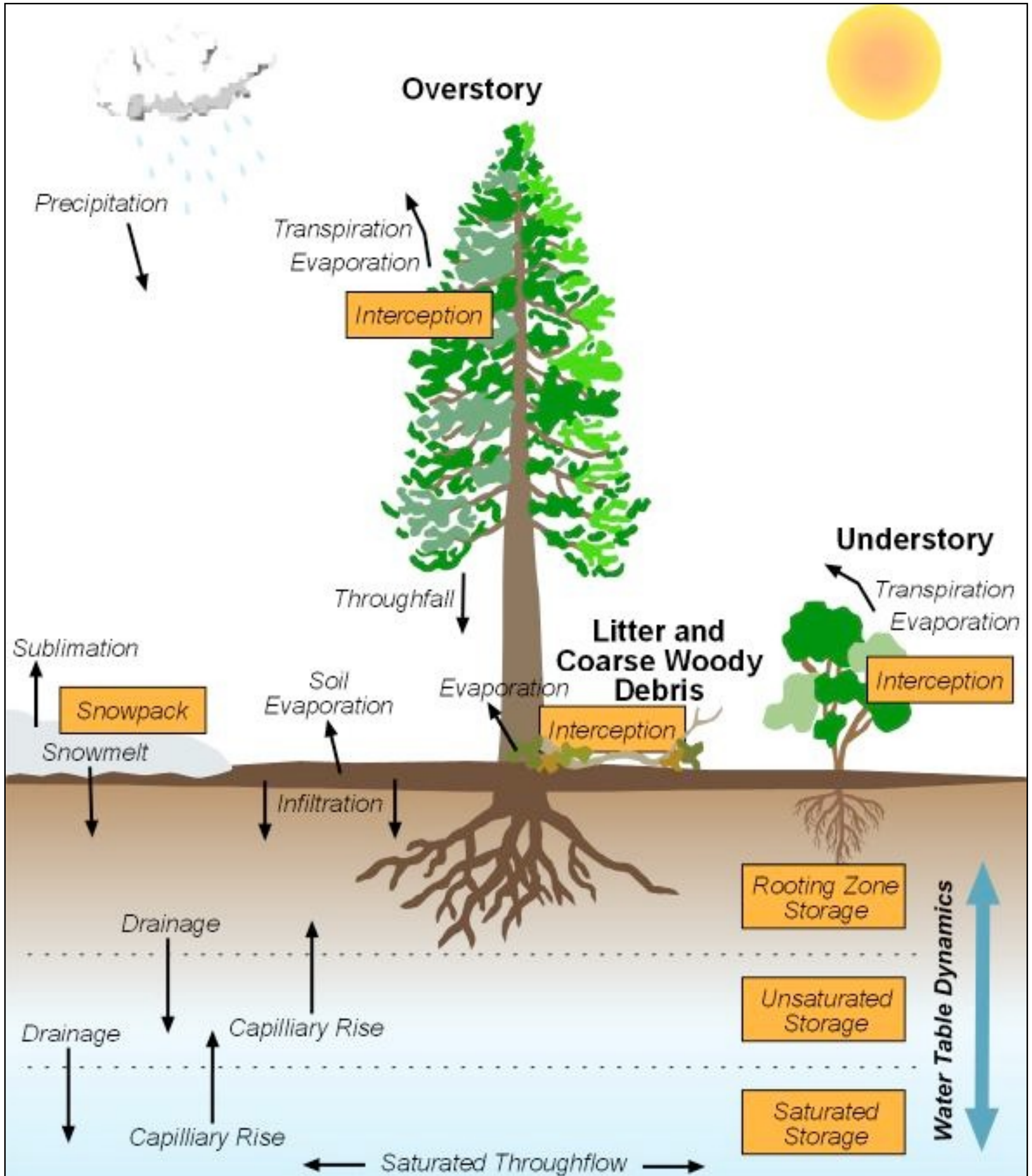
Does it represent the processes that you are interested in:

- outputs (e.g. streamflow, ET, N-export, forest water use)
- temporal and spatial resolution appropriate for the questions you are asking
- does it account for the mechanisms that are likely to be important in the questions you are asking - (conceptual model)

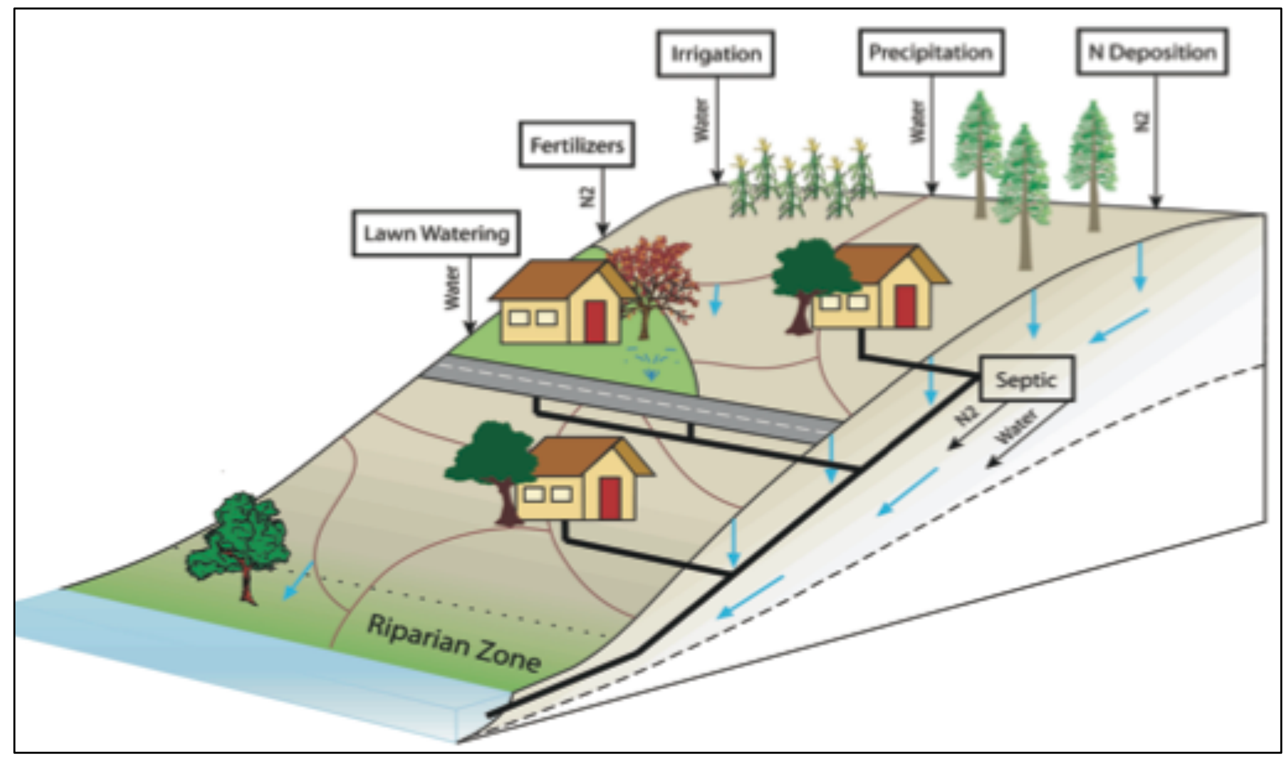
Is it APPROPRIATE

Then pick the simplest model that accounts for your processes of interest

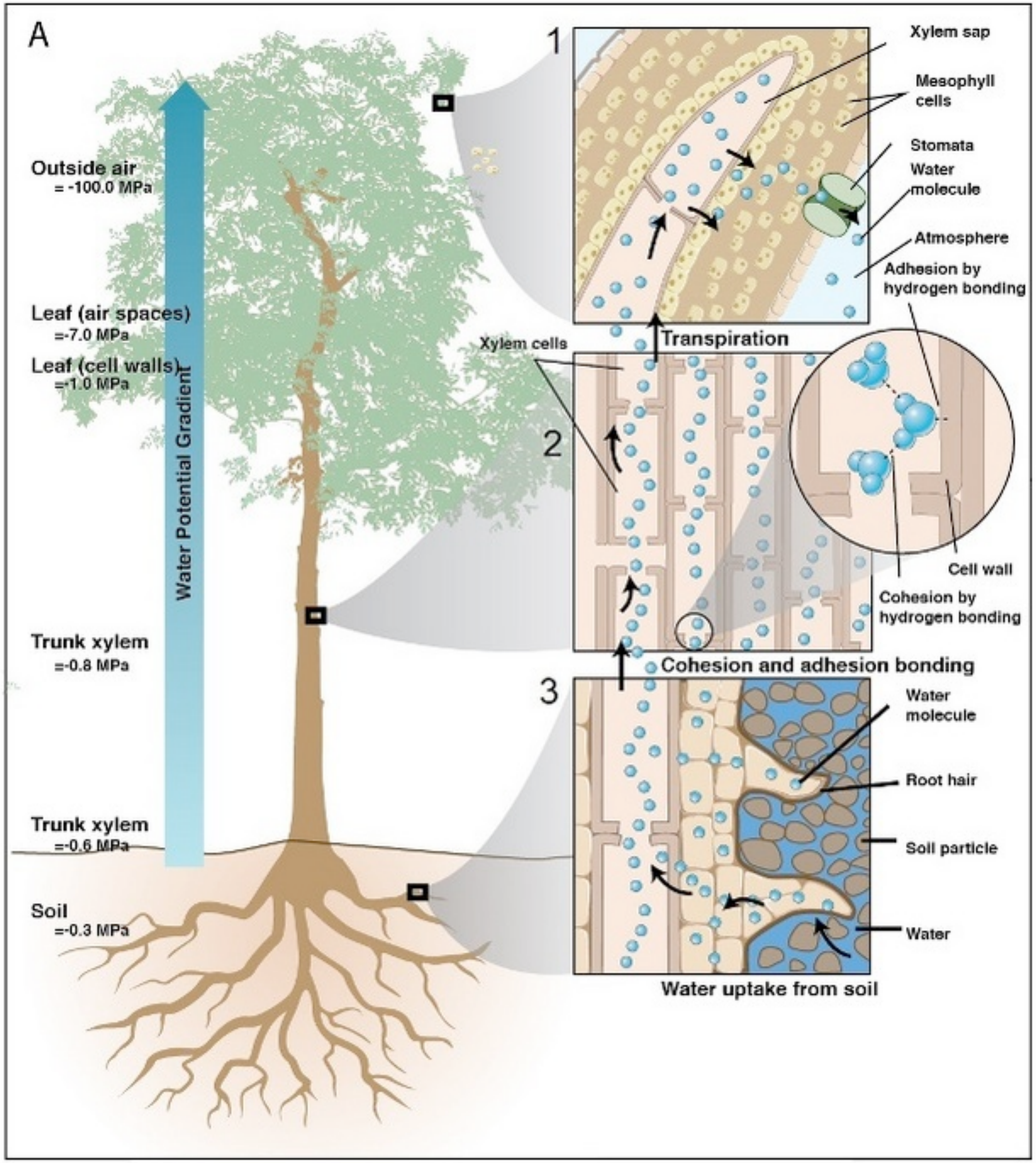
Hydrologic processes in RHESSys



GIS based model of water, carbon, and nitrogen cycling



Evapotranspiration



More complex models of plant hydrology

Citation: McElrone, A. J., Choat, B., Gambetta, G. A. & Brodersen, C. R. (2013) Water Uptake and Transport in Vascular Plants. Nature Education Knowledge 4(5):6

Model Selection 2

Does it capture the outputs of interest (and relationships between inputs/parameters and outputs) with sufficient *accuracy* to answer your research questions

IS IT GOOD ENOUGH?

Model Validation (Is it reasonable)

- Compare model results to simple thought experiments

- Similar to testing: conservation of mass, energy, behaviors under known conditions (zero rain = zero streamflow; zero CO_2 change = zero T change)?

- Are the values for outputs physically reasonable (e.g snowpack > 0, reservoir storage < reservoir)

Validation (is it accurate?)

Compare model results to observations
(either local or non-local)

How good is good enough?

Validation (is it accurate?)

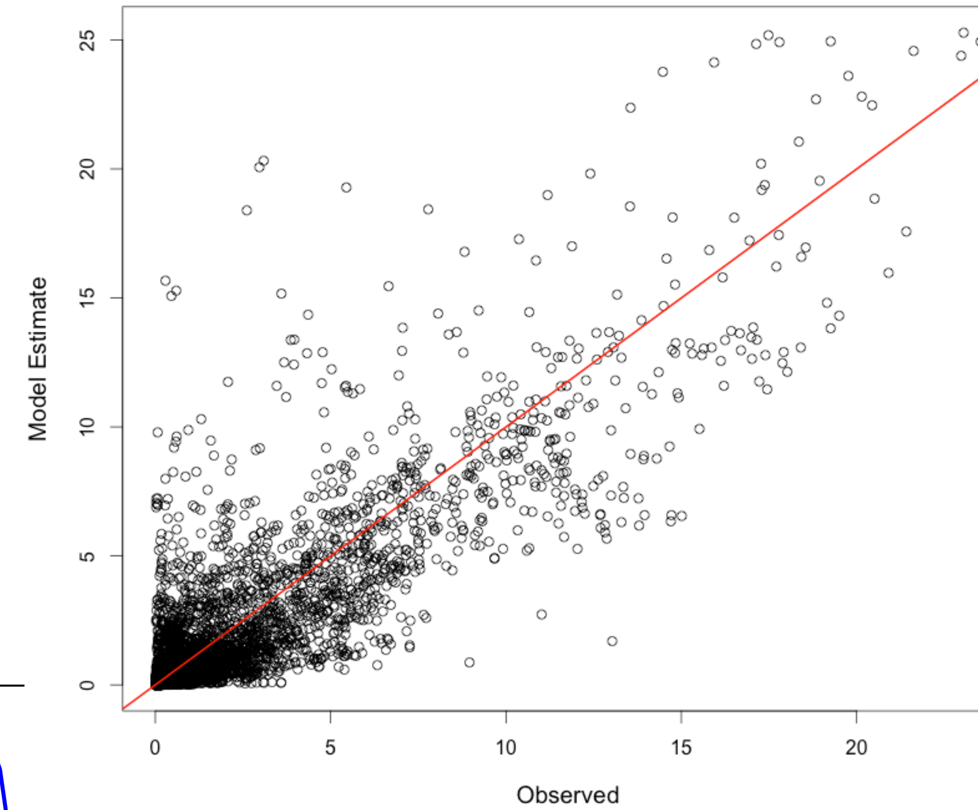
Compare model results to observations

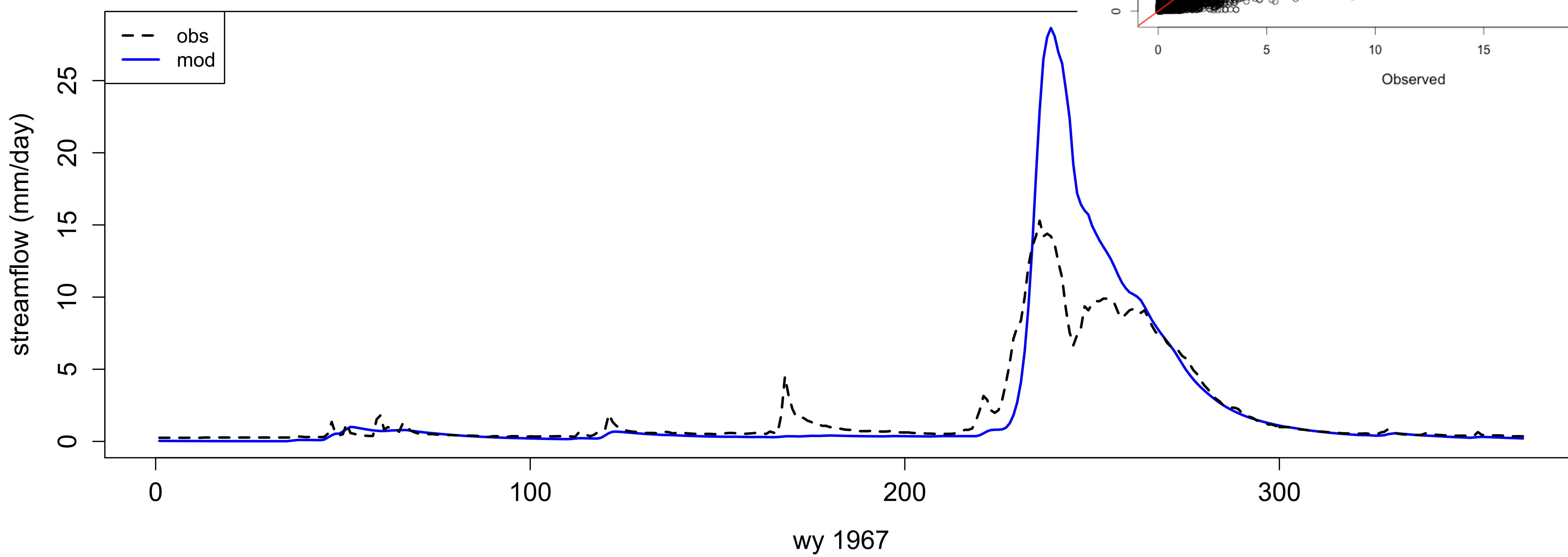
- Observations from the same site/scenario/ circumstance
  - streamflow from a rainfall-runoff model applied to a given watershed
  - estimates of population growth after an actual disturbance
- Observations of general patterns, relationships
  - ranges of precip/streamflow for watersheds in that region
  - ranges of population growth after similar disturbances

Model Performance

PLOT: First Step!!

Model and observed through time
Error through time
Model versus observed





Validation (is it accurate?)

IS IT GOOD ENOUGH?

MODEL + INPUTS/
PARAMETERS

Summary of good enough

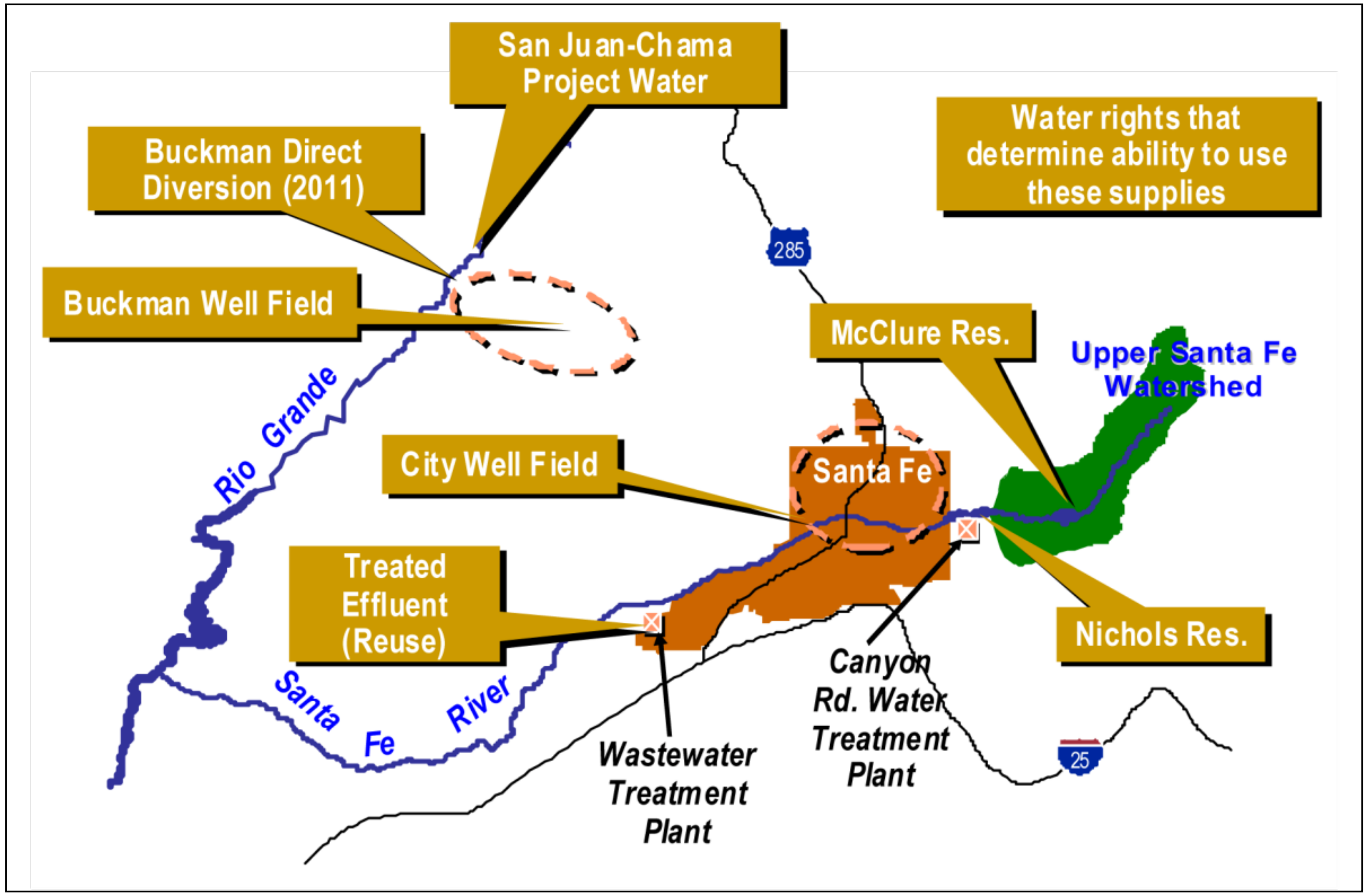
- Error/uncertainty in hydrologic variable of interest (or response of interest) is small relative to use of model results for decision making

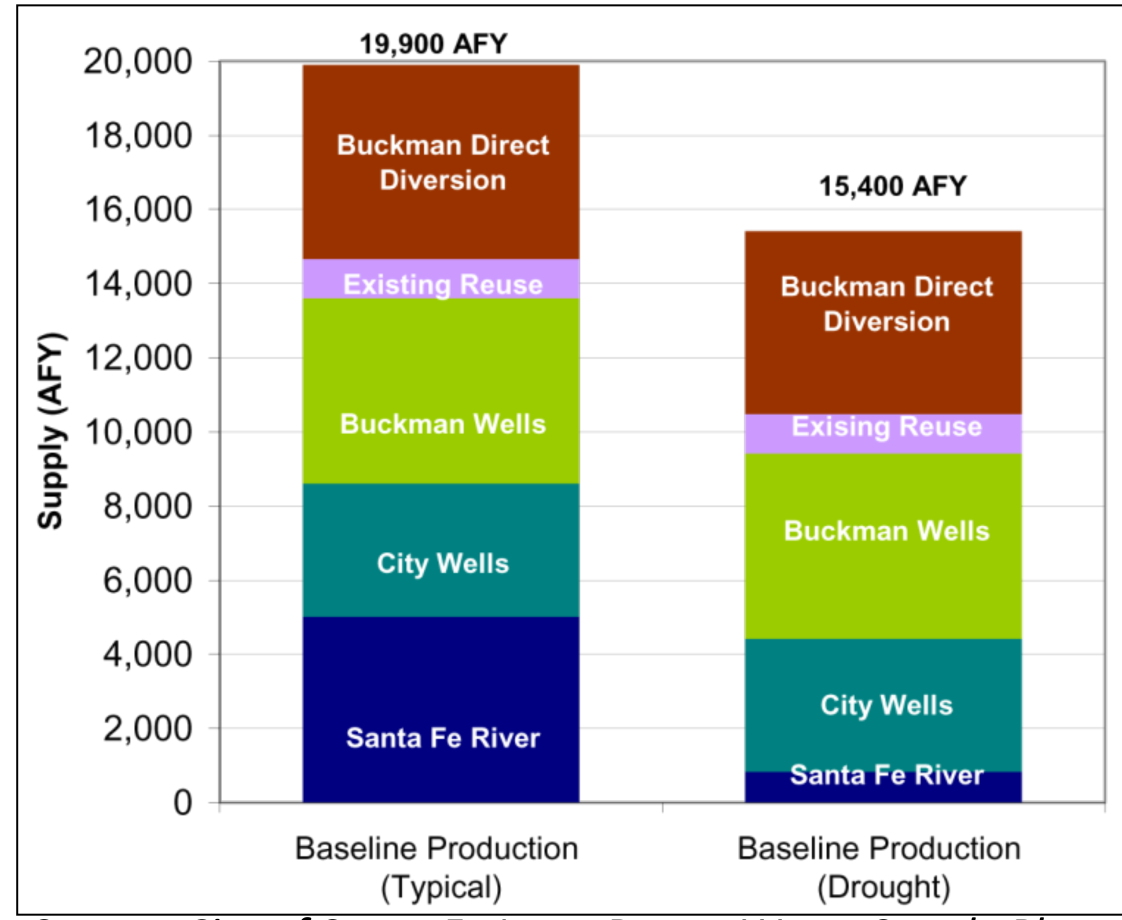
- Errors/uncertainty is small relative to simulated effect of change or relationships of interest

- Similar levels of performance by other models/studies reported in the literature - the “state of the art” argument

- Demonstration of improved performance

Change in flow at which Santa Fe water-managers would consider purchasing additional water rights or wells



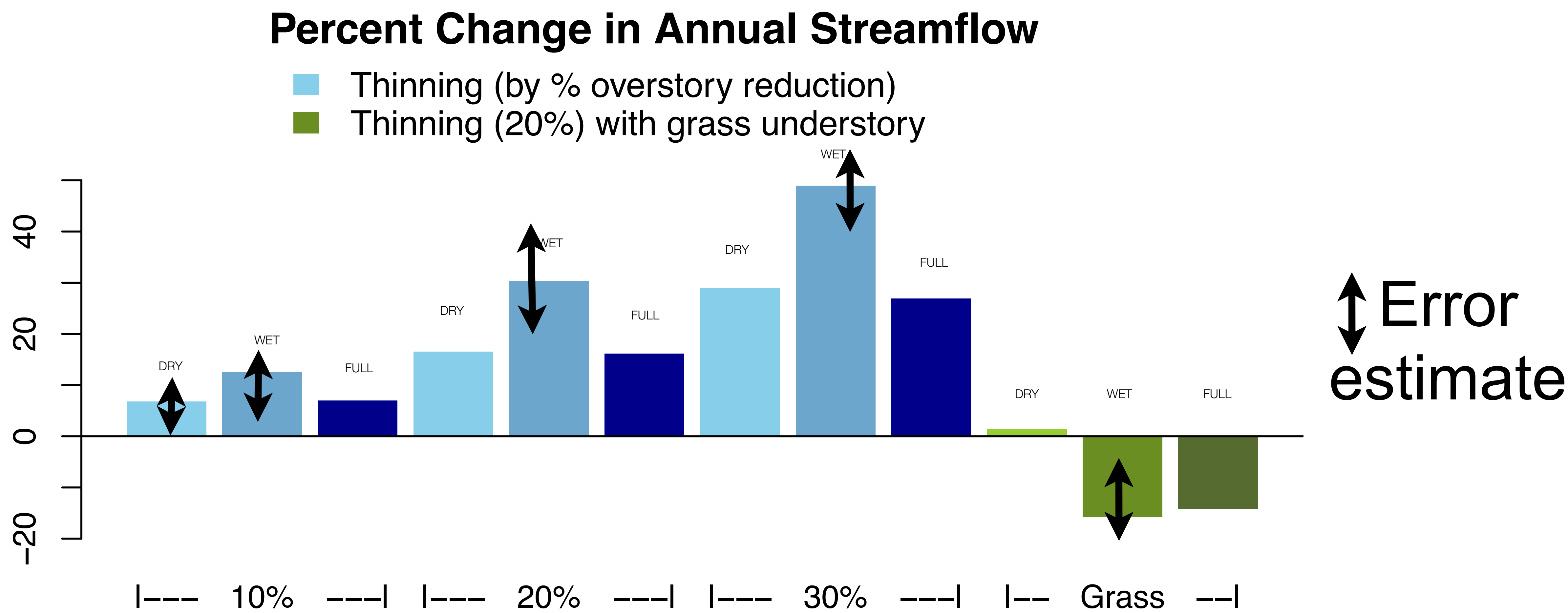


Source: City of Santa Fe Long-Range Water Supply Plan, 2008.

Supply < 2000 AFY, costs X dollars to “buy” additional water..., what is the cost effective decision given climate change estimates for the next decade?

What is good enough?

Errors/uncertainty is small relative to simulated effect of change or process of interest

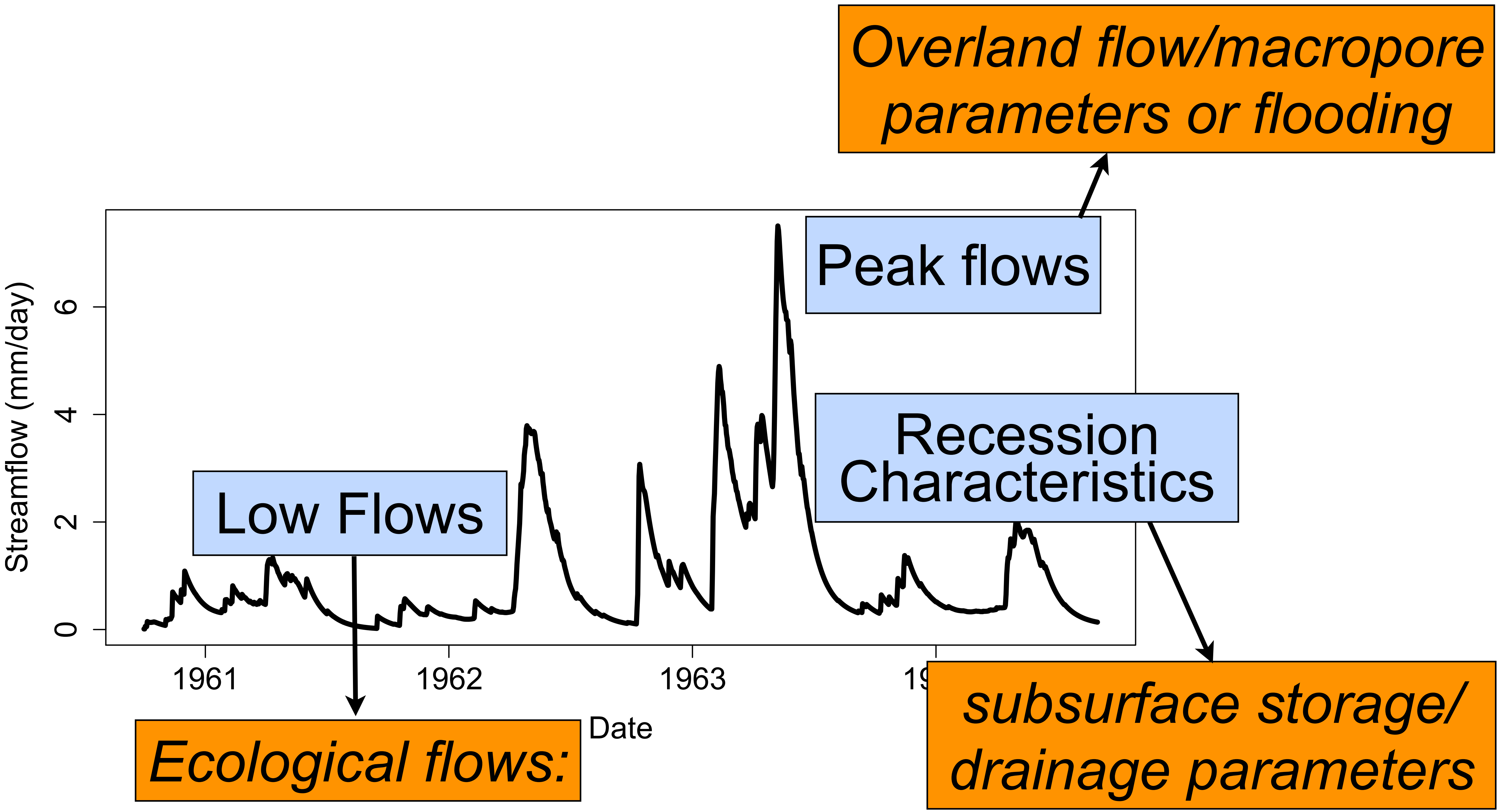


adapted from Dugger et al., 2013

Could a grass understory impact how streamflow changes with overstory reduction (thinning)?

Information content of observations:

What you try to get "right" depends on the processes that the parameters influence and your use of the model



What is good enough?

2b. Demonstration of improved performance

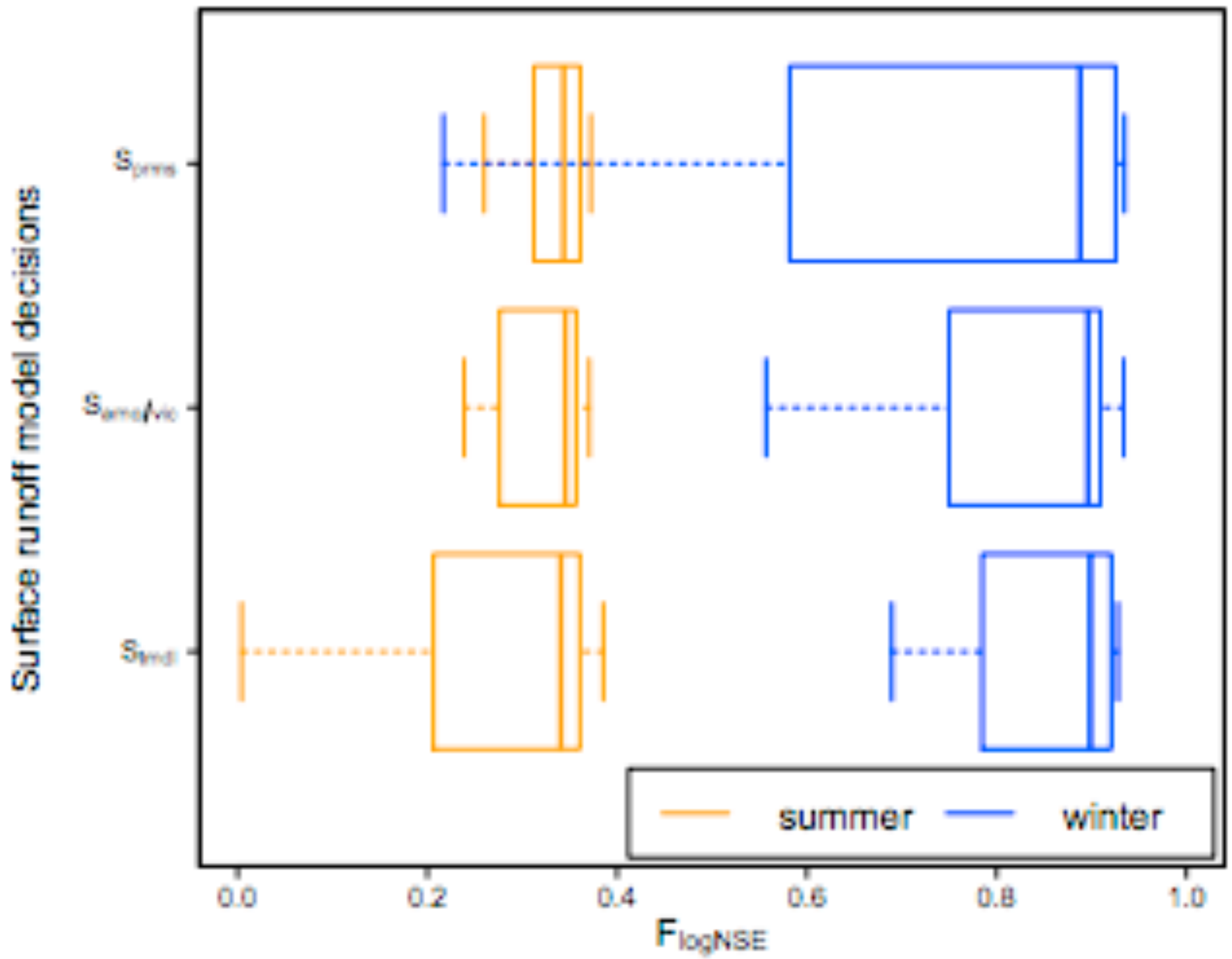


Fig. 8. Boxplots of model performance for summer and winter streamflow simulations for the three surface runoff decision options.

Plot demonstrates the effect of different model structures (surface runoff assumptions adapted from VIC, PRMS, TOPMODEL) on performance

Staudinger, M., Stahl, K., Seibert, J., Clark, M. P., and Tallaksen, L. M.: Comparison of hydrological model structures based on recession and low flow simulations, Hydrol. Earth Syst. Sci. Discuss., 8, 6833-6866, doi:10.5194/hessd-8-6833-2011, 2011.

Performance Measures

Root Mean Square Error (RMSE)

$$SSE = \frac{1}{n}\sum_{i=1}^{n}\left(m_i - o_i\right)^2$$

$$RMSE = \sqrt{SSE}$$

Nash Sutcliffe Efficiency (NSE)

Nash and Sutcliffe, 1970, J. of Hydrology

Widely used in hydrology

Range – infinity to +1.0

Overly sensitive to extreme values

$$NSE = \frac{\sum_{i=1}^{n}\left(o_i - m_i\right)^2}{\sum_{i=1}^{n}\left(o_i - \bar{o}\right)^2}$$

BIAS or Percent Error (Err)

Useful for determining if there is a long term flow over or under estimation

$$Err = \frac{\left(\bar{m} - \bar{o}\right)}{\bar{o}} * 100$$

*Others: Cor, R^2

Model Performance

```
#' nse
#'
#' Compute NSE between observation and model
#' @param  m  model estimates
#' @param  o  observations
#' @return nse


nse = function(m,o) {

err = m-o
meanobs = mean(o)
mse = sum(err*err)
ovar = sum((o-meanobs)*(o-meanobs))
nse = 1.0-mse/ovar

return(nse)
}
```

Model Performance

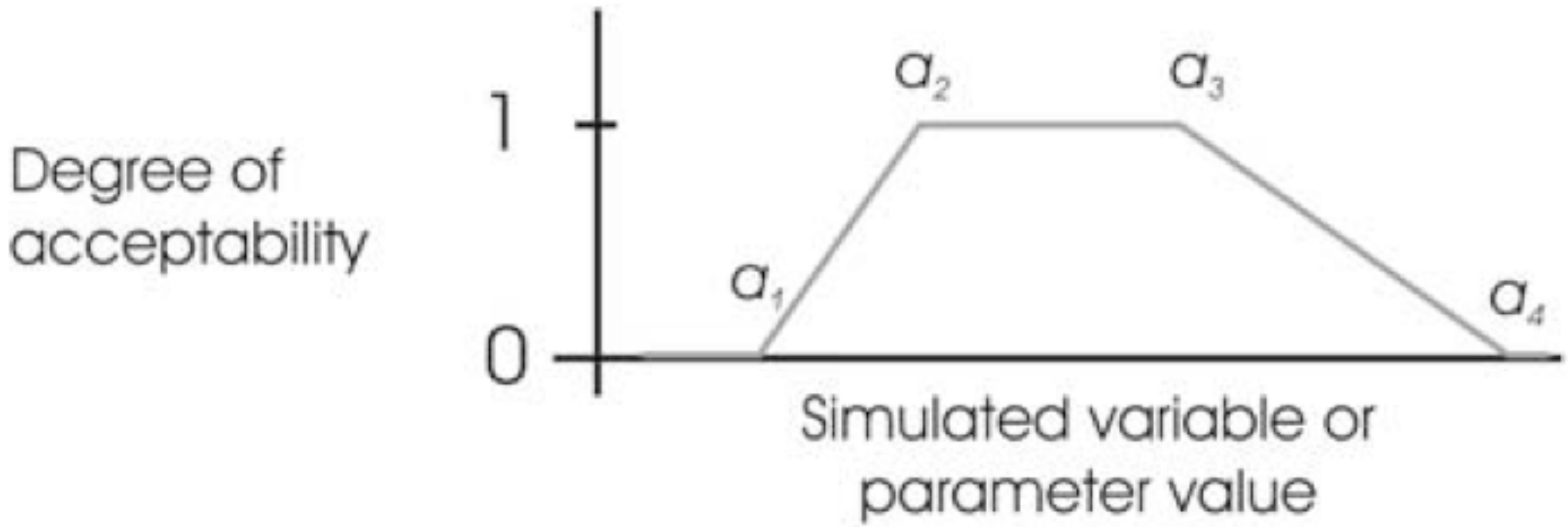
```
#' relerr
#'
#' Compute percent error between observation and model
#' @param  m  model estimates
#' @param  o  observations
#' @return relerr


relerr = function(m,o) {

err = m-o
meanobs = mean(o)
meanerr = mean(err)

res = meanerr/meanobs
return(res)
}
```

Soft metrics - Fuzzy-Evaluation



$$\mu(x) = \begin{cases} 0 & \textit{if } x \le a_1 \\ \dfrac{x - a_1}{a_2 - a_1} & \textit{if } a_1 \le x < a_2 \\ 1 & \textit{if } a_2 \le x < a_3 \\ \dfrac{a_4 - x}{a_4 - a_3} & \textit{if } a_3 \le x < a_4 \\ 0 & \textit{if } x > a_4 \end{cases}$$

For data where there is a lot of uncertainty in observed values (imprecise measurements)

Seibert, J., and J. J. McDonnell, On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, Water Resour. Res., 38(11), 1241, doi:10.1029/2001WR000978, 2002.

Performance: Metrics

R also has built in functions that can be helpful

- cor(x,y) - correlation coefficient
- help(cor)

Model Performance

Combining objective functions

Multiplicative approach

Metric A * Metric B …

Metric A * weighting A + Metric B * weighting B…

Requires metrics to be normalized (same range)

- 0-1
- divide by maximum value

Metrics must all work in the same direction

increase = better OR decrease = better

Performance Metrics

- If you are combining = need to increase with "better" models

- Transform metrics like RMSE that work in reverse

$$SSE = \frac{1}{n}\sum_{i=1}^{n}(m_i - o_i)^2$$

$L = (SSE)^{-n}$ where n is a shaping parameter
(Freer et al., (1997)

$L = \exp(-nSSE)$

$L = (\max(RMSE) - RMSE) / (\max(RMSE) - \min(RMSE))$

$$\text{relErr} = \frac{(\bar{m} - \bar{o})}{\bar{o}}$$

Transform to 0-1, and positive

mErr = 1.0 - min(1.0, abs(relErr))

mErr 1.0 - min(1.0,abs(relErr)/max(abs(relErr))

Combining

cperf = mErr * max(NSE, 0)
cperf = 0.75*mErr * 0.25*max(NSE,0)

Model Performance

```
#' cper
#'
#' Compute a performance measure (0-1) between observation and model
#' based on both NSE and relative error
#' @param  m  model estimates
#' @param  o  observations
#' @param  weight.nse weighting to give NSE metric
#' @param  weight.relerr weighting to give relative error metric
#' @return  combined 0-1 performance measure


cper = function(m,o,weight.nse=0.5, weight.relerr=0.5) {

nse = nse(m,o)
mnse = max(nse,0)

rel.err = relerr(m,o)
merr = 1.0-min(1.0, abs(rel.err)/max(abs(rel.err)))

combined = weight.nse*mnse + weight.relerr*merr

return(combined)

}
```

Calibration-Optimization

- choosing parameter sets to use based on comparison with observed data
- calibration is very similar to sensitivity analysis
  - we could use LHS or SOBEL function to generate parameter sets and model runs
  - compute performance metrics for each run
  - graph and decide on a ‘cut off point’ of ‘acceptable parameters
- optimization
  - a way to to calibration - search procedure

Calibration in R

```
#' compute annual yield'
#'
#' Function to compute yeild of different fruits as a function of annual temperature and precipitatio
#' @param T annual temperature (C)
#' @param P annual precipitation (mm)
#’ @param crop.pars - list that contains the following
#' @param Topt optimal temperature (C)
#' @param max.water  maximum water requirement (mm)
#' @param ts slope on temperature
#' @param tp slope on precipitation
#' @param base.yield baseline yield  (kg)
#' @param irr  irrigation in (mm)
#' @return yield in kg


compute_yield = function(T, P, irr, crop.pars) {

with(as.list(crop.pars), {

nyears=length(T)
irr.peryear = rep(irr, times=nyears)
water.input = P+irr.peryear;
yield = ifelse(water.input < max.water,
        tp*water.input - ts*abs(T-Topt) + base.yield,
        tp*max.water - ts*abs(T-Topt) + base.yield )
yield=pmax(yield,0)
return(yield)
})

}
```

Note how different metrics help to define different parameters

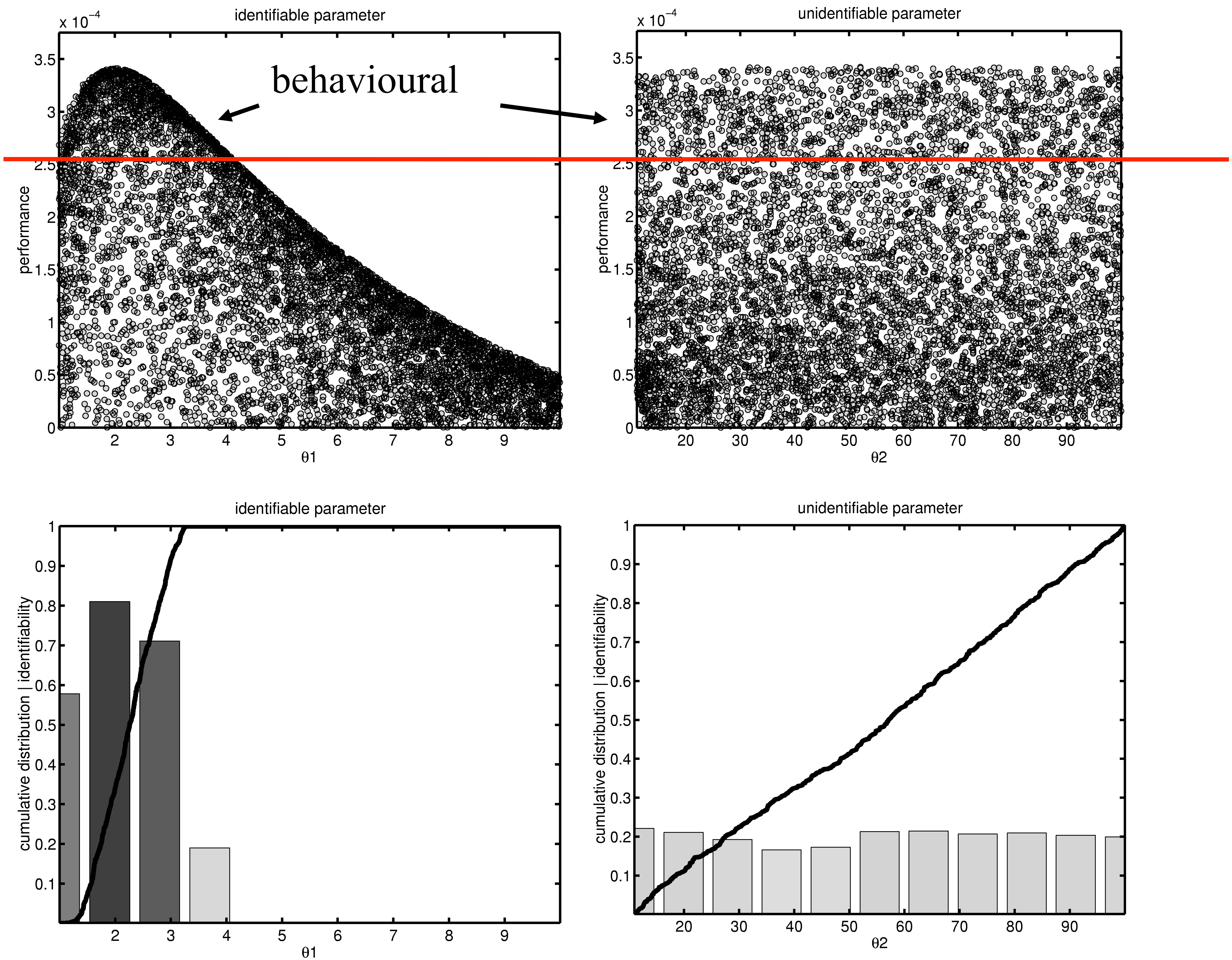
Calibration

- Choose parameter values that give you the best performance
- Sometimes this is obvious, but often things are much more messy

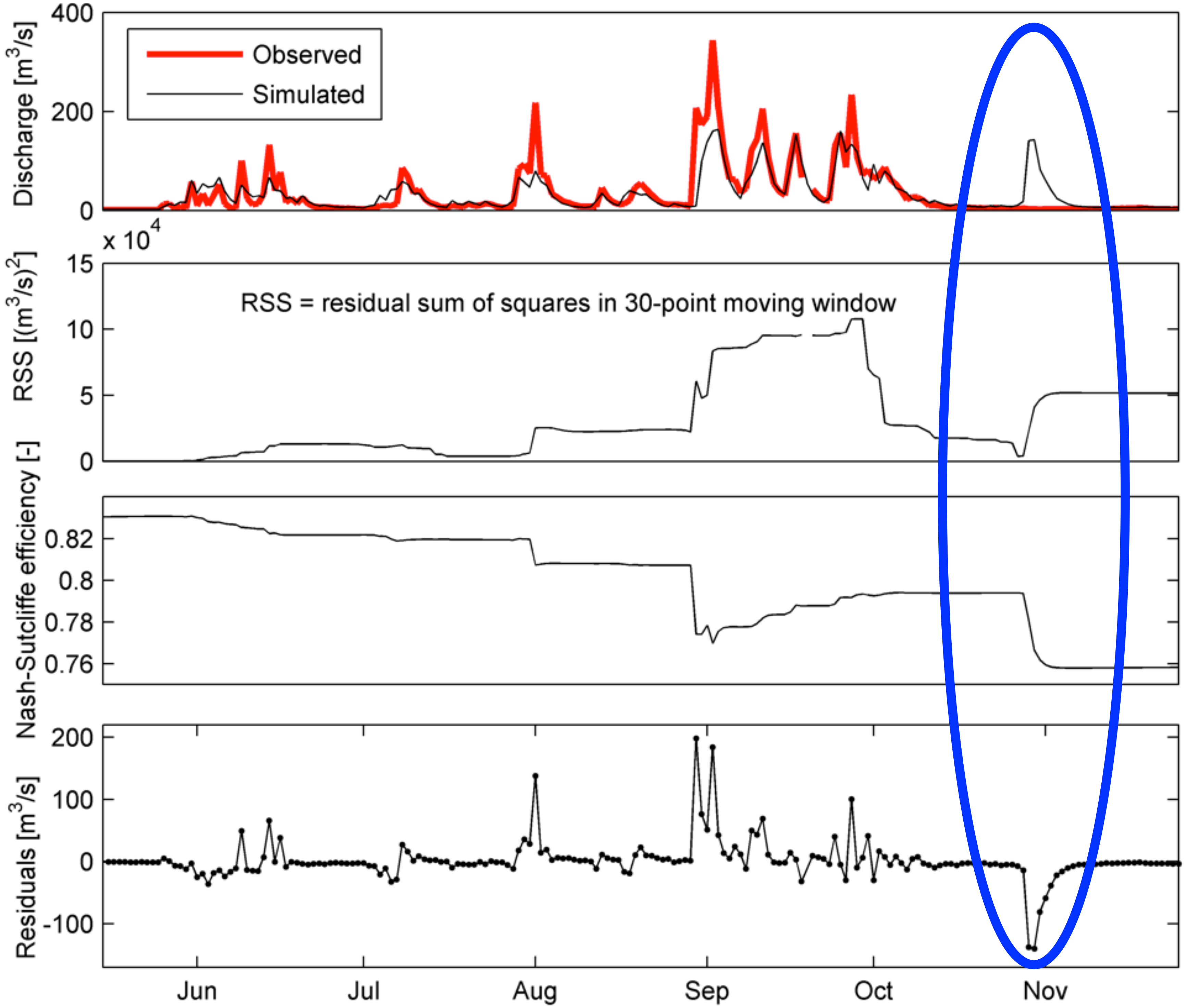
Equifinality

- Many parameter values can give equal performance
- And comparison with observed data is limited
    - errors in observed data
    - over-fitting to the particular time/place/set of observations available

Dotty Plots and Identifiability Analysis

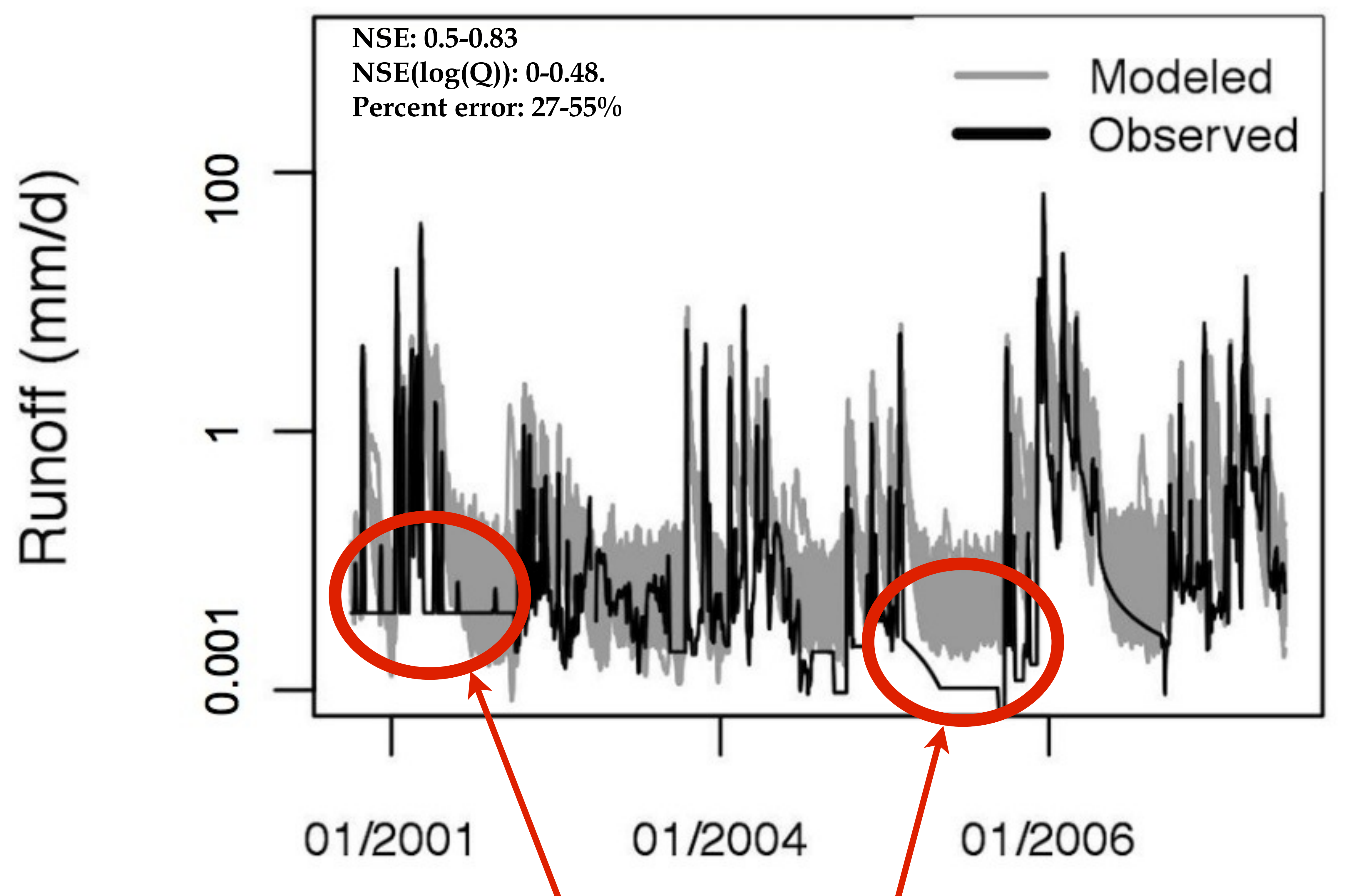


Disinformation in calibration data



Application of WASMOD to Pasa La Ceiba, Honduras
(from Ida Westerberg, Uppsala)

Other Issues:Observed Error



Error in observed streamflow can bias calibration:
Remove those periods from calibration

Problem:

model selection based on performance AND calibration

- Parameter optimization/evaluation: will not be robust
- calibration period
- performance measure
- input/measurement errors
- concept of equifinality

Beven, JH, 2006, Manifesto for the Equifinality Thesis

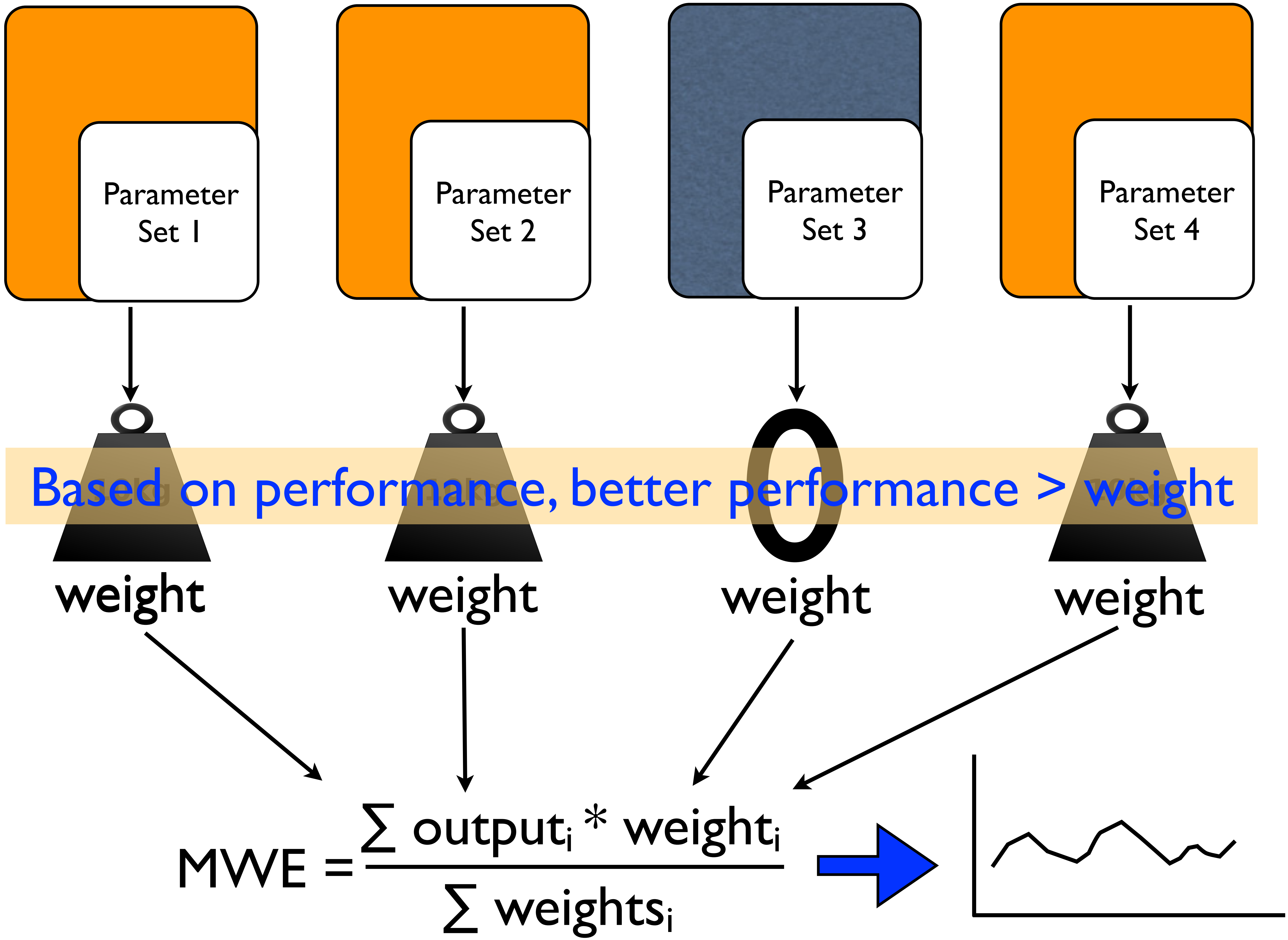
A possible solution:

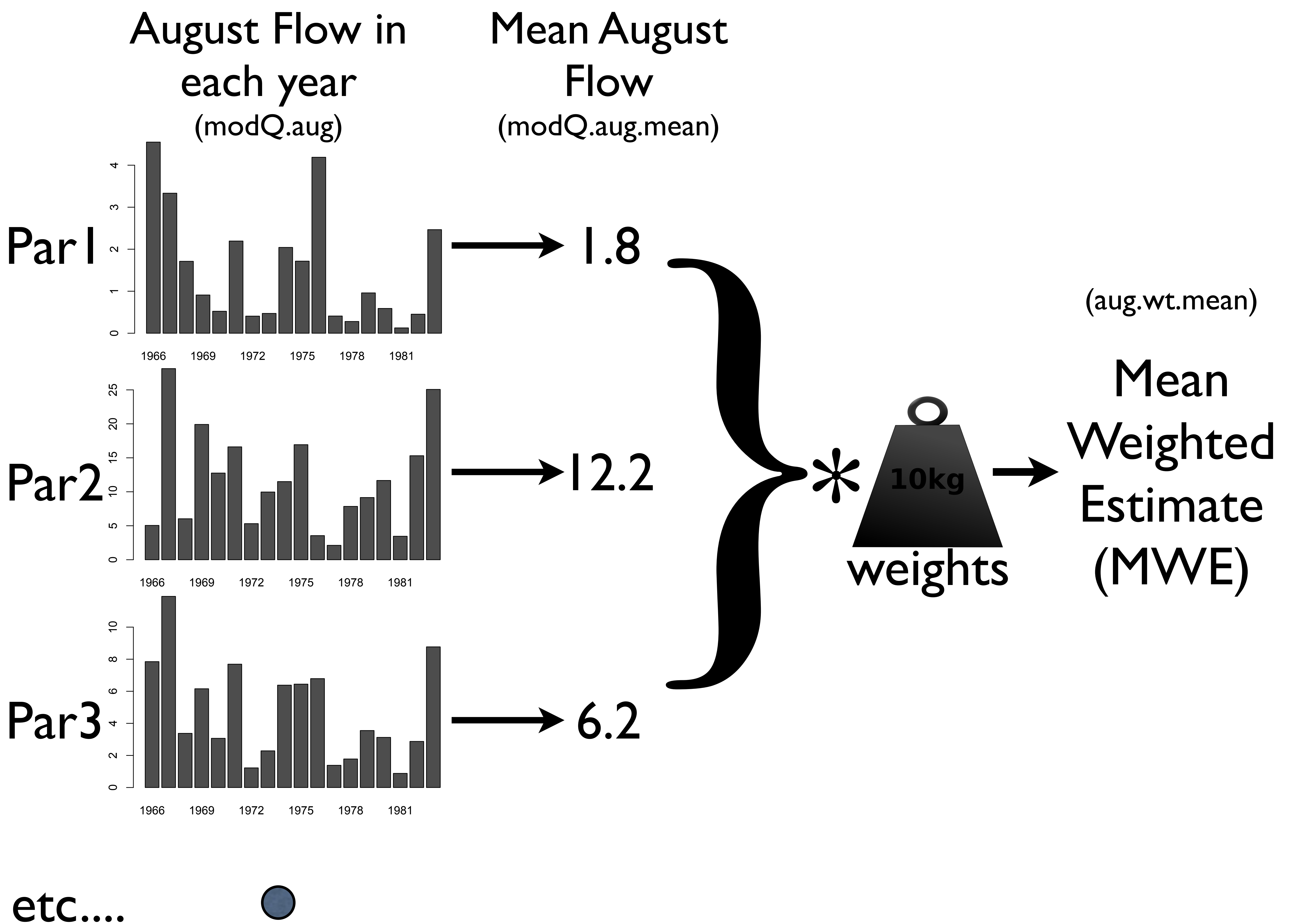
Generalized Likelihood Uncertainty Estimation (GLUE)

- assess the likelihood of different models + parameters being good predictors of the system of interest
- reject (give zero likelihood) those models that are clearly not good predictors of calibration data
- Can be done with different model structures as well as different parameter sets

What to do about Equifinality

- Keep all parameter sets that are acceptable
  - acceptable: above some threshold of performance
  - always run the model for those parameter sets and use range of model output to define uncertainty bound
  - if you need a single model estimate:
    - combine results from all acceptable parameters
      - » average
      - » weight by performance


Parameter Set 1
Parameter Set 2
Parameter Set 3
Parameter Set 4
Based on performance, better performance > weight
0
weight
weight
weight
weight
MWE = ∑ output_i * weight_i / ∑ weights_i


August Flow in each year
(modQ.aug)
Mean August Flow
(modQ.aug.mean)
Par1
0
1
2
3
4
1966
1969
1972
1975
1978
1981
1.8
Par2
0
5
10
15
20
25
1966
1969
1972
1975
1978
1981
12.2
Par3
0
2
4
6
8
10
1966
1969
1972
1975
1978
1981
6.2
*
10kg
weights
(aug.wt.mean)
Mean Weighted Estimate (MWE)
etc....

Calibration - Beyond GLUE: Sampling the parameter space -Optimization and Formal Bayesian analysis

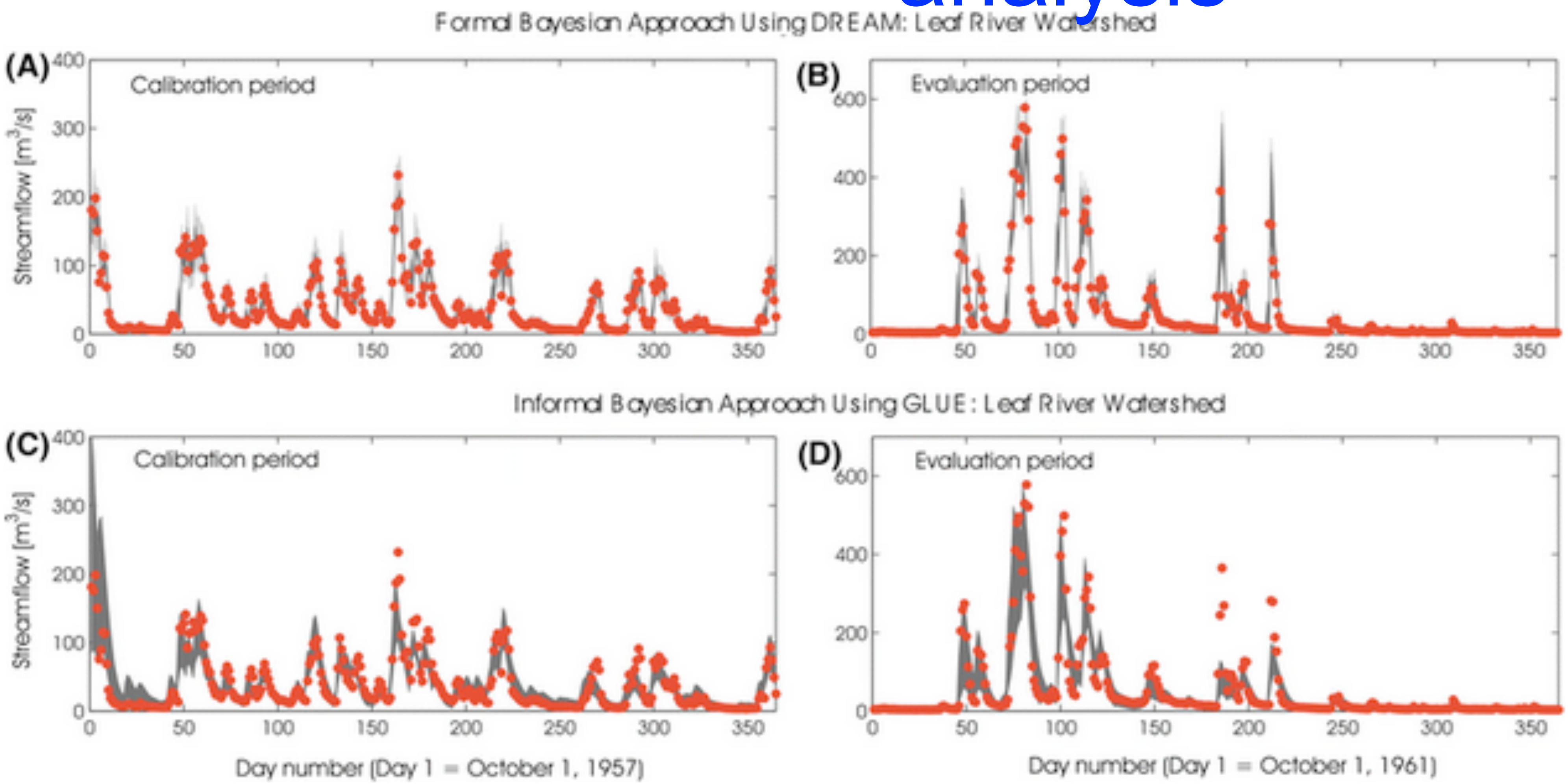


Fig. 5

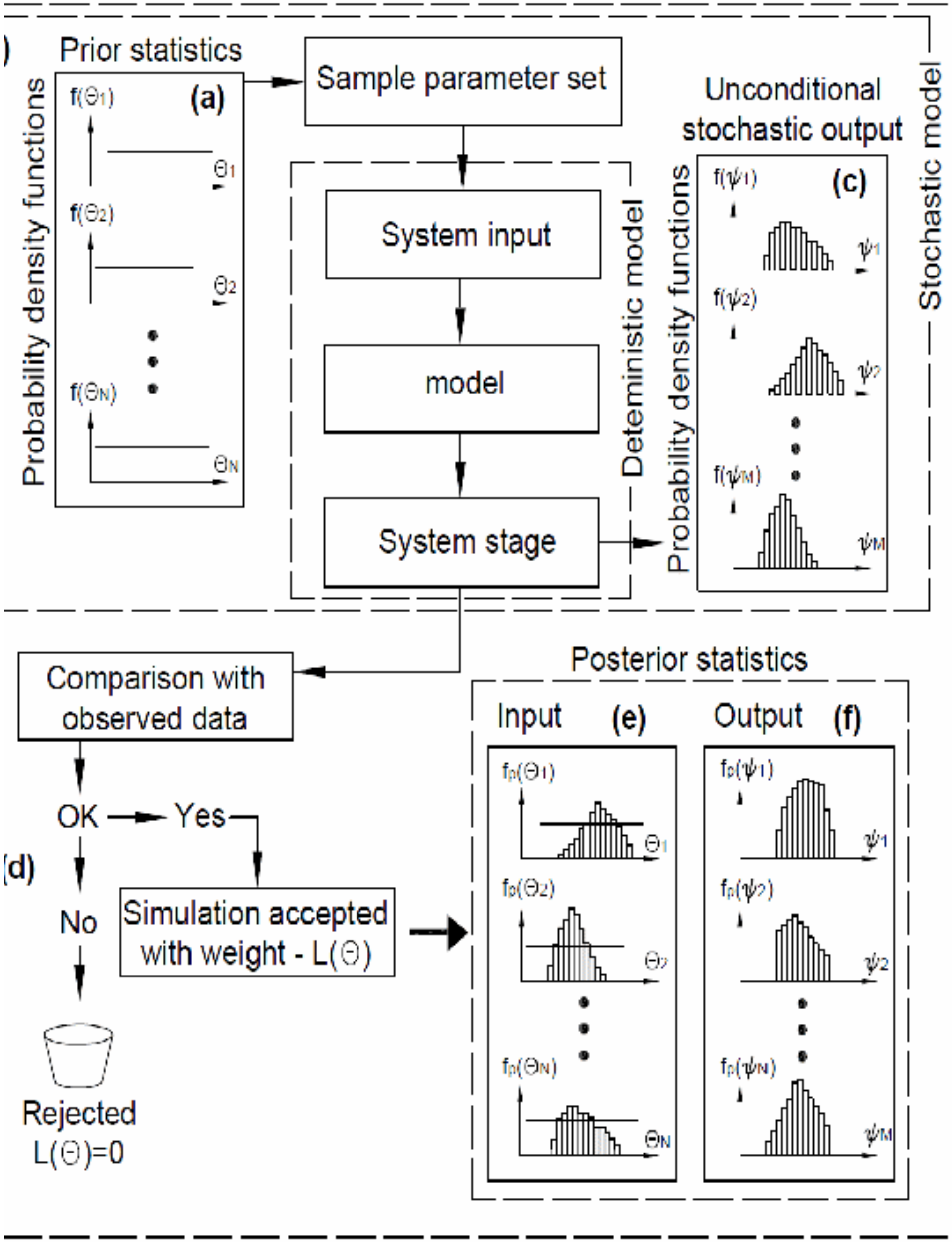
Streamflow prediction uncertainty ranges derived with DREAM (*top panels*) and GLUE (*bottom panels*) for a representative portion of the calibration (*left column*) and evaluation period (*right column*) for the Leaf River watershed. In each DREAM graph, the *dark gray* region represents the 95% confidence intervals of the output prediction due to parameter uncertainty, whereas the *light gray* region represents the additional 95% ranges of the prediction uncertainty. For GLUE the 95% prediction quantiles are presented. The *solid circles* denote the streamflow observations

J.A. Vrugt, C.J.F. ter Braak, H.V. Gupta, B.A. Robinson Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stochastic Environmental Research and Risk Assessment, 44 (2008)http://dx.doi.org.proxy.library.ucsb.edu:2048/10.1007/s00477-008-0274-

Generally found similar results using GLUE and formal MCMC calibration and uncertainty estimation

GLUE:
Generalized Likelihood Uncertainty Estimation

K. Beven and A. Binley, “The future of distributed models: model calibration and uncertainty prediction,” Hydrological Processes, vol. 6, no. 3, pp. 279–298, 1992. View at Scopus



http://dc434.4shared.com/doc/R33bPROP/preview_html_m537a3256.gif