UQ for Complex Models

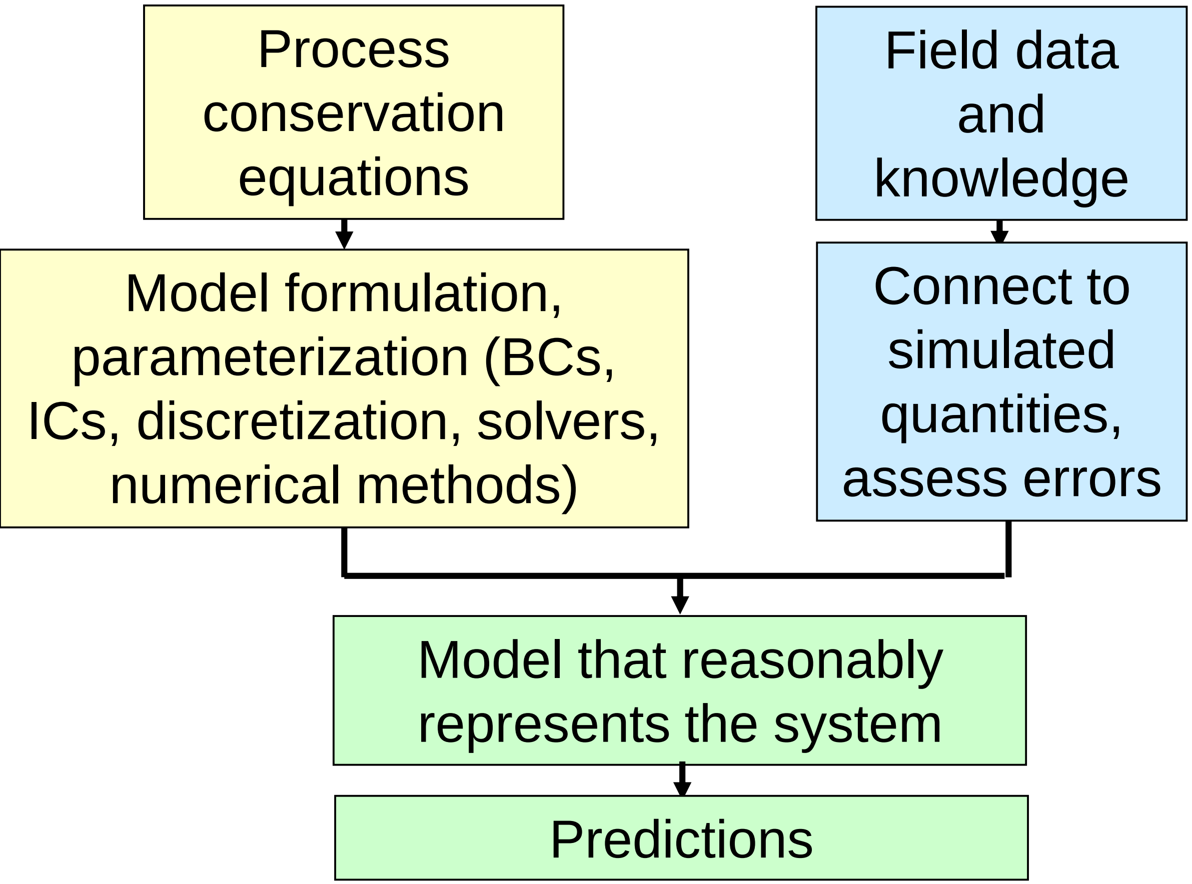
Professor Mary Hill

KU Department of Geology

Brian Klager

USGS, KU Department of Civil Engineering

Tools for Data-Model Integration



Issues	Tools
Scale Unknowns	-Hierarchical approaches -Reduce and UQ
Realism Compatibility Right results for right reasons? Predictive ability?	Model analysis methods -Compare obs, sim -Inversion -SA, UQ -Model Intercom-parisons

There are times to use semi-truck and a VW

There are "semi-trucks" and "volkswagons" to address any model analysis question

Methods largely based on work of U Minnesota statisticians Cook & Weisberg

	Volkswagons	Semitrucks
Common questions	**Frugal methods**	**Demanding methods**
Model Adequacy		
1. How can many data types with variable quality be included?	Error-based weighting and SOO or MAP	MOO, Pareto curve
2. Is model misfit/overfit a problem? Is the fit to prior knowledge and data subsets consistent? Are errors Gaussian?	RMSE, Nash-Sutcliffe, graphs, R^2_N, s_n^2, $s_{(n-p)}^2$ Compare fit to a priori error analysis using s_n^2, $s_{(n-p)}^2$	MOO, Pareto curve
3. How nonlinear is the problem?	Intrinsic nonlinearity, DELSA	DELSA, Explore objective function

From Hill Kavetski Clark Ye Arabi Lu Foglia Mehl 2015 Groundwater Journal

There are "semi-trucks" and "volkswagons" to address any model analysis question

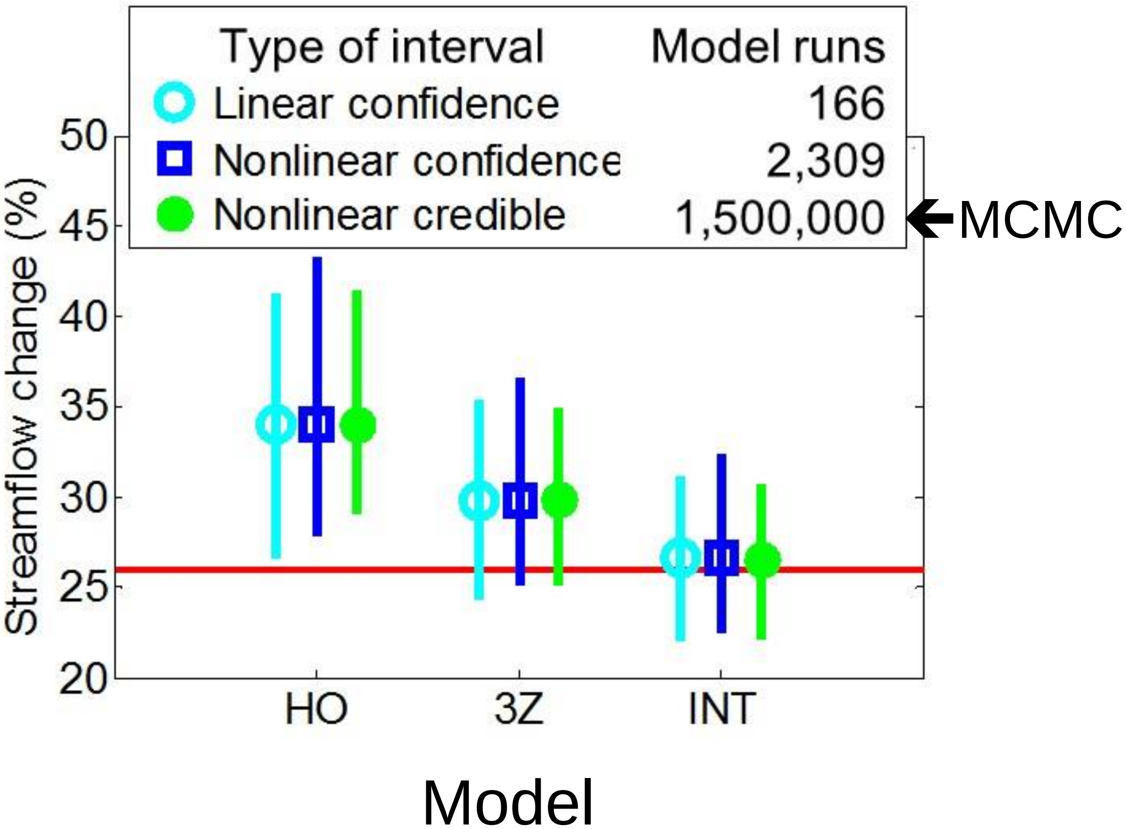
From Hill Kavetski Clark Ye Arabi Lu Foglia Mehl 2015 Groundwater Journal

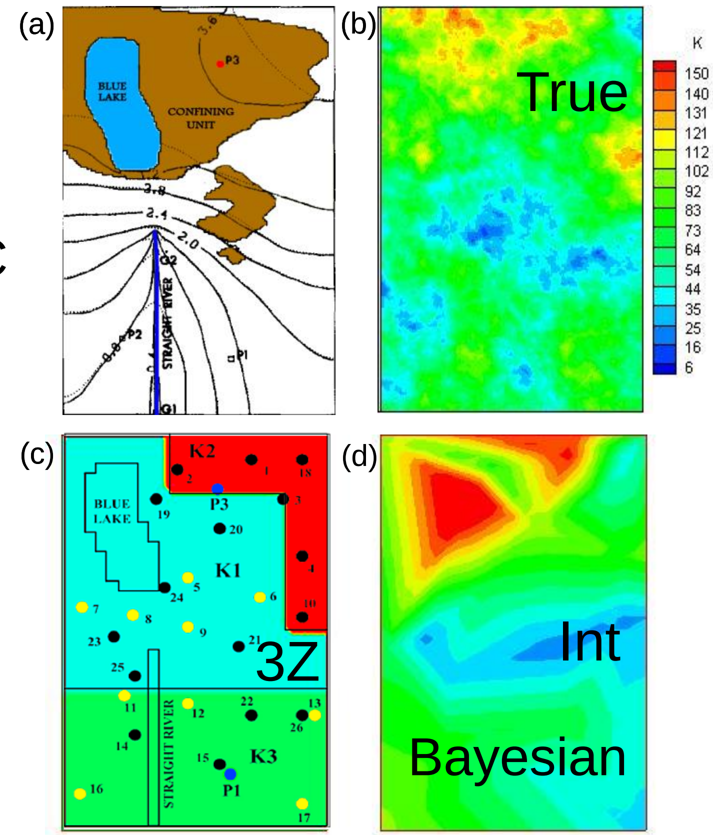
Common questions	Frugal methods	Demanding methods
Model Adequacy		
1. How can many data types with variable quality be included?	Error-based weighting and SOO or MAP	MOO, Pareto curve
2. Is model misfit/overfit a problem? Is the fit to prior knowledge and data subsets consistent? Are errors Gaussian?	RMSE, Nash-Sutcliffe, graphs, R^2_N, s_n^2, $s_{(n-p)}^2$ Compare fit to a priori error analysis using s_n^2, $s_{(n-p)}^2$	MOO, Pareto curve
3. How nonlinear is the problem?	Intrinsic nonlinearity, DELSA	DELSA, Explore objective function
Sensitivity and Uncertainty		
Observations (Obs) ←→ Parameters (Pars)		
4. What pars can and cannot be estimated with the obs?	Scaled local stats (CSS, ID, PCC, etc.), SVD, DoE, MoM(OAT, EE)	DoE, MoM(OAT, EE), eFAST, Sobol', RSA
5. Are any parts dominated by one obs and, thus, its error?	Scaled local stats (Leverage, DFBETAS)	Cross validation
6. How certain are the par values?	Par uncertainty intervals	Par uncertainty intervals
7. Which obs are important and unimportant to pars?	Scaled local stats (Leverage, Cook's D)	Cross validation
Parameters (Pars) ←→ Prediction (Preds)		
8. Which pars are important and unimportant to preds?	Scaled local stats (PSS, etc.), DELSA	DELSA, eFAST, Sobol'
9. How certain are the preds?	z/SD_z, Pred uncertainty intervals	Pred uncertainty intervals, multi-model analysis
10. Which pars contribute most and least to the pred uncertainty?	Scaled local stats (PPR VOII)	eFAST, Sobol'
Observations (Obs) ←→ Prediction (Preds)		
11. Which existing and potential obs are important to preds?	Scaled local stats (OPR VOII)	Cross validation
12. For multi-model analysis, which models are likely to produce accurate preds?	Analyze model fit and estimated parameters, AIC, AICc, BIC, KIC	Cross validation
Risk Assessment		
13. What risk is associated with a given decision strategy and set of scenarios?	Combine uncertainty analysis and scenario simulation. Smooth cost function	Combine uncertainty analysis and scenario simulation. Cost function need not be smooth.
14. What decisions are robust given a set of uncertain scenarios?	Evolutionary multiobjective optimization. Within this demanding method use frugal model analysis methods.	

Examples

1. Synthetic GW model – compare uncertainty measures (1 slide)
2. GW model in Death Valley region – observations important to predictions (1 slide)
3. Rainfall-runoff in the Netherlands – finding a model error (2 slides)

Example 1: GW model – Uncertainty Comparison

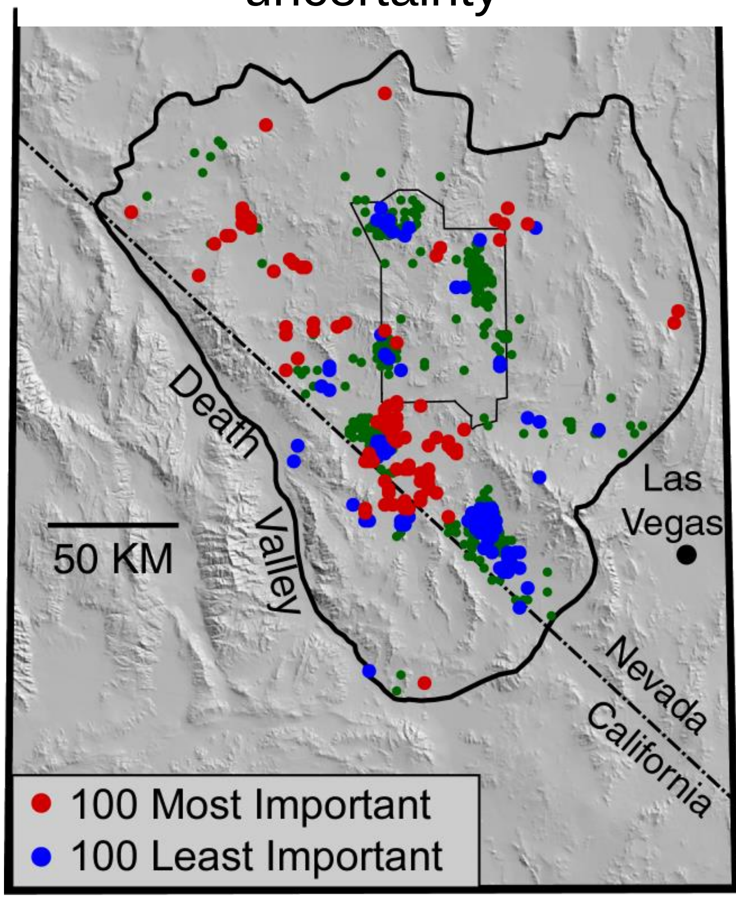




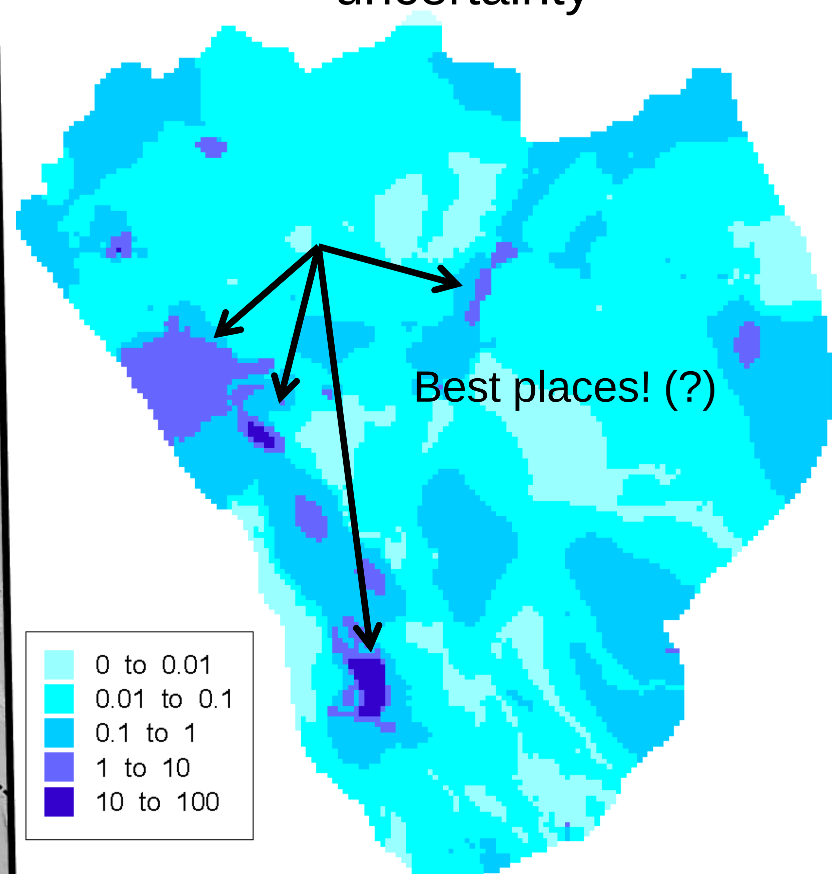
Lu + 2012 WRR

Example 2: Death Valley region GW

Existing obs that reduce uncertainty



New data that could reduce uncertainty



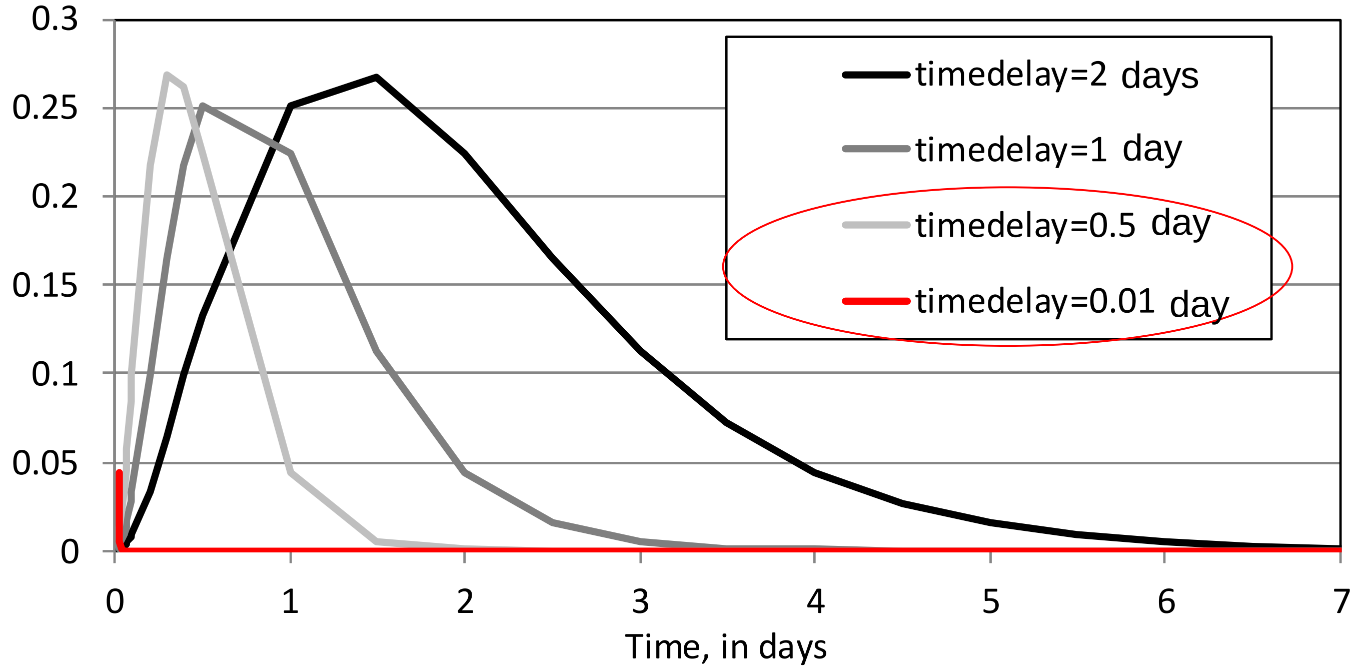
Consider calculated results in the context of model construction.
Odd results may indicate model construction problems.

Tiedeman + 2003, 2004 WRR

Documentation: Tonkin et al 2007 USGS

Example 3: Rainfall-Runoff in the Netherlands – Finding a model error

DELSA Rakovec Hill et al 2014 WRR +



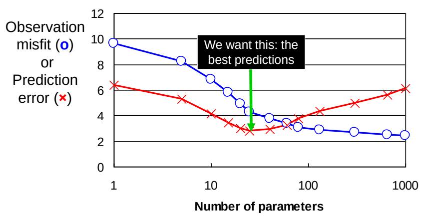
Gamma distributions for unit hydrographs.

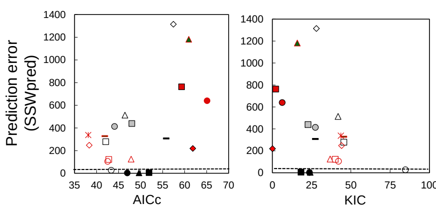
Model had a one day time step

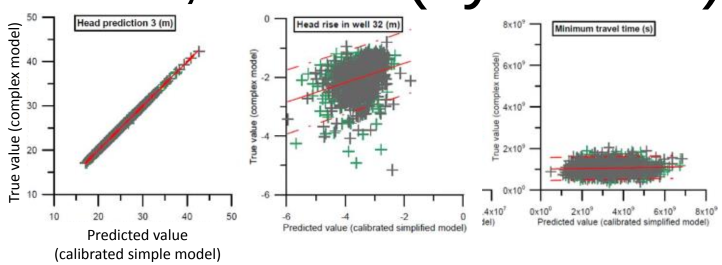
Found error because local sensitivity measures very large for small timedelay (and poor model fit).
Global methods like Method of Morris identify TIMEDELAY as most important parameter. They average results over parameter space, which hid the problem.

Model Intercomparisons

- Given data and model purpose, how do different methods of modeling and model analysis perform?
  - Exabyte scale computing plays critical role
- Selected published examples
  - In GW: Foglia + 2013 WRR (field case, cross-validation) Doherty Christensen 2011, WRR (synthetic)







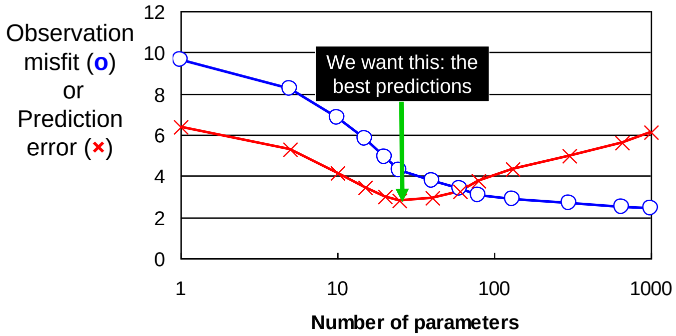
  - In SW: Initiatives such as DMIP

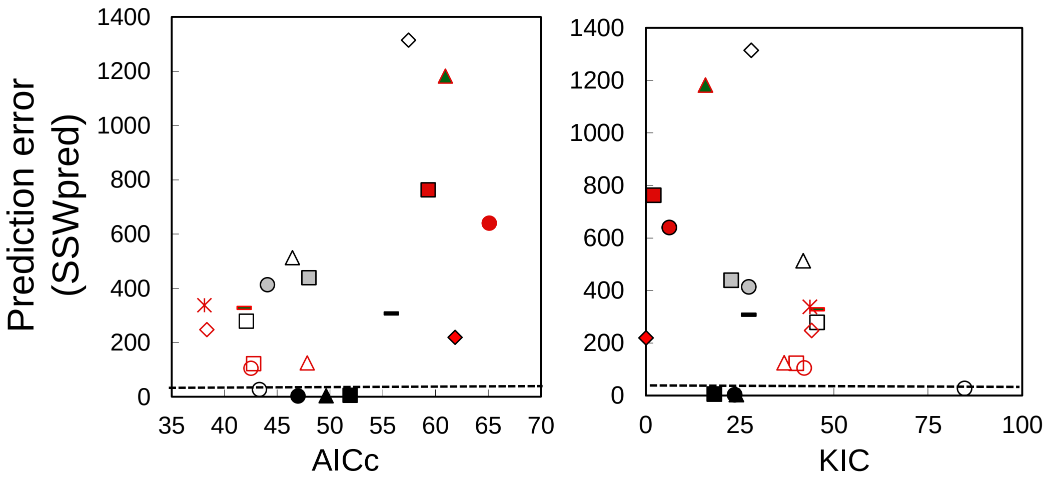
Maxwell + 2014 (test problems show similar performance of 7 complex integrated hydrologic models)

Best + 2015 (calibrated simple models better than uncalibrated process-based models)

Testing predictive skill using cross validation

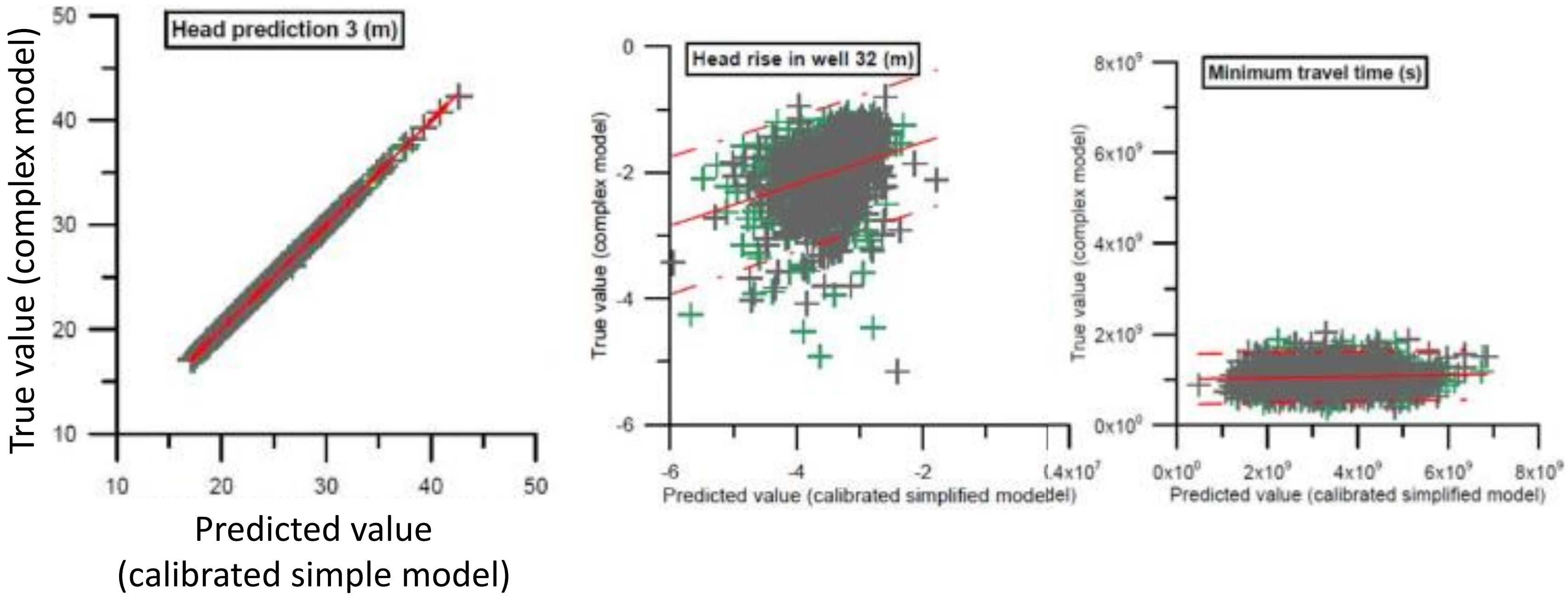
Foglia et al WRR 2013





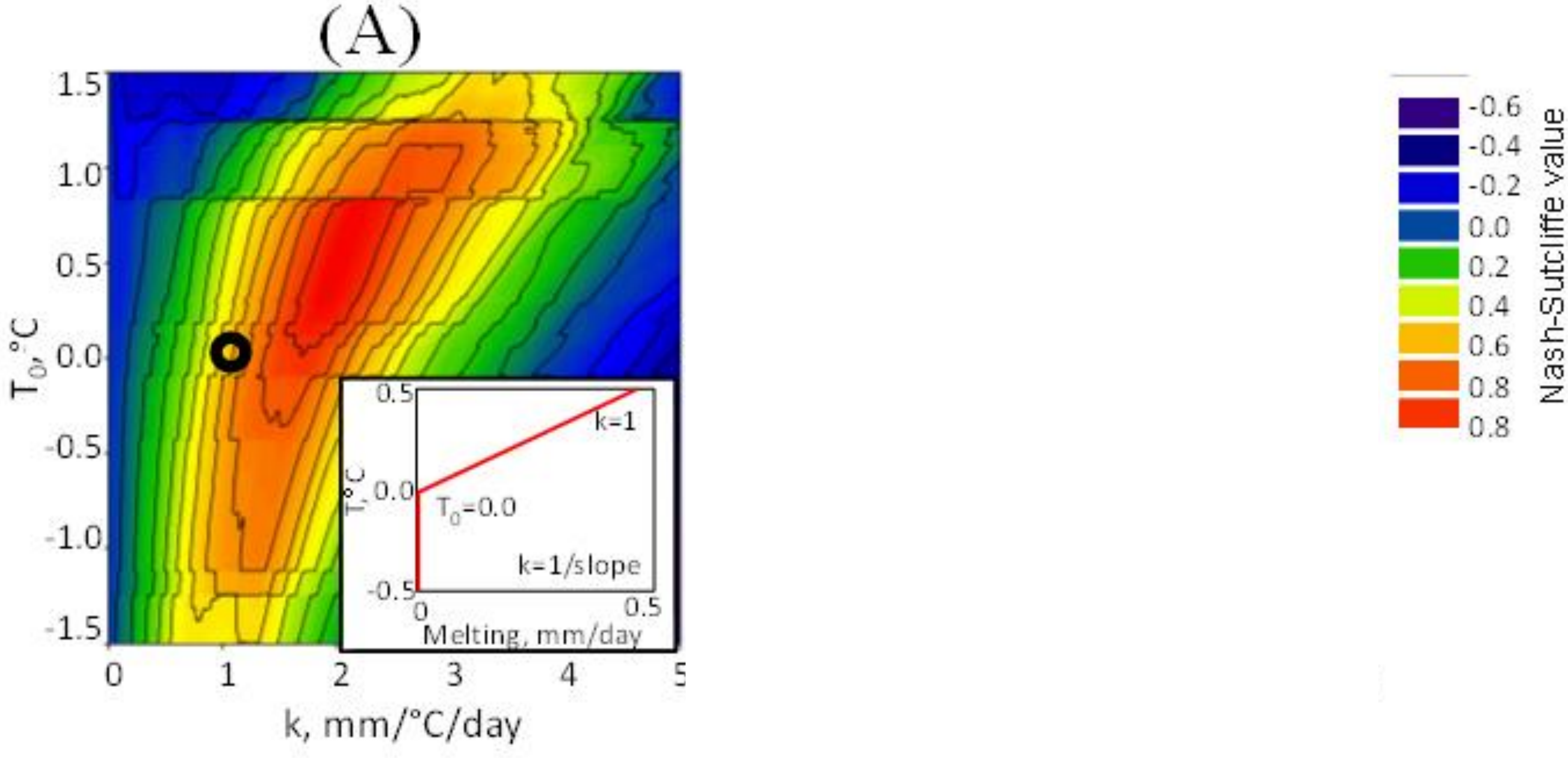
Synthetic model tests for prediction skill testing

Doherty and Christensen WRR 2011



How computationally frugal methods can fail – an example using snow melt

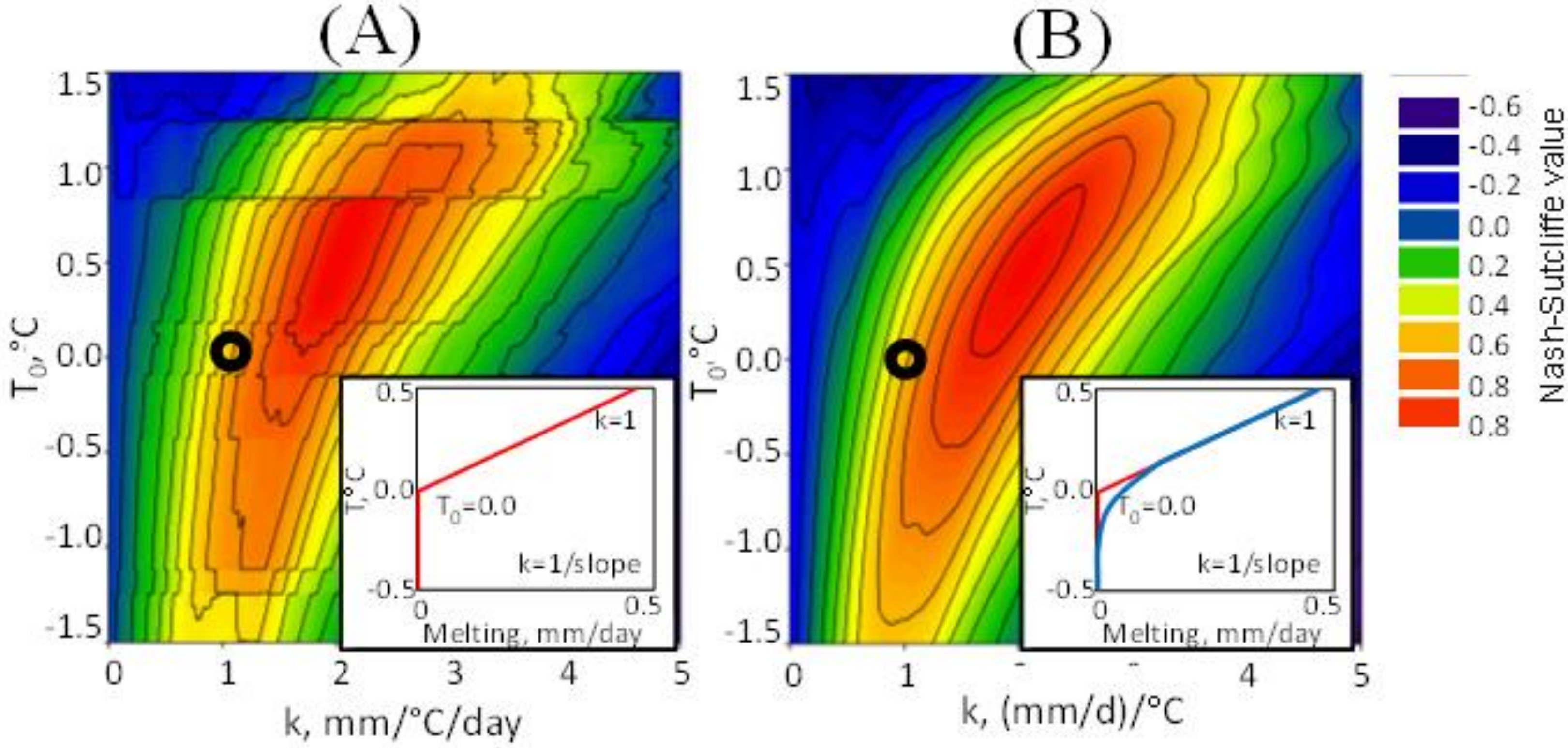
Called Numerical Daemons by Kavetski and Clark and Kavetski 2010



Kavetski and Kuczera 2007 WRR; Hill et al 2015 Groundwater

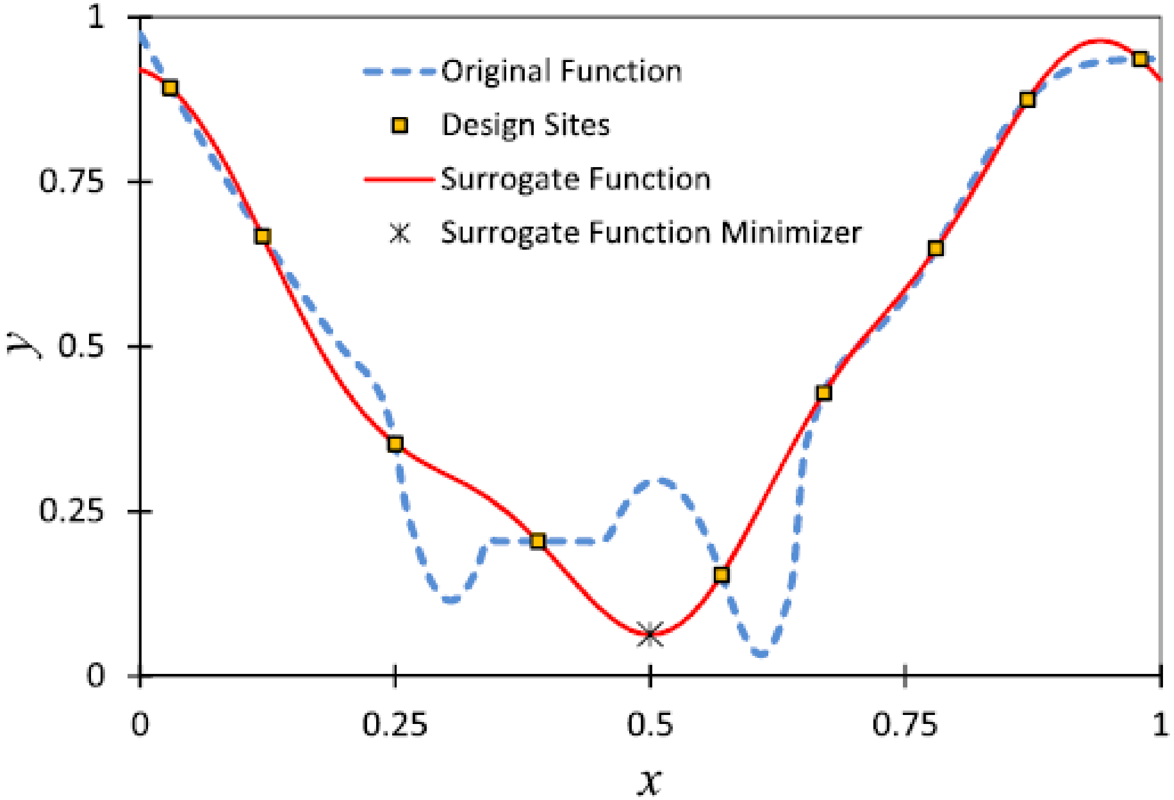
Fix the problem

Make the function more realistic



Kavetski and Kuczera 2007 WRR; Hill et al 2015 Groundwater

How surrogates can fail – a conceptual example

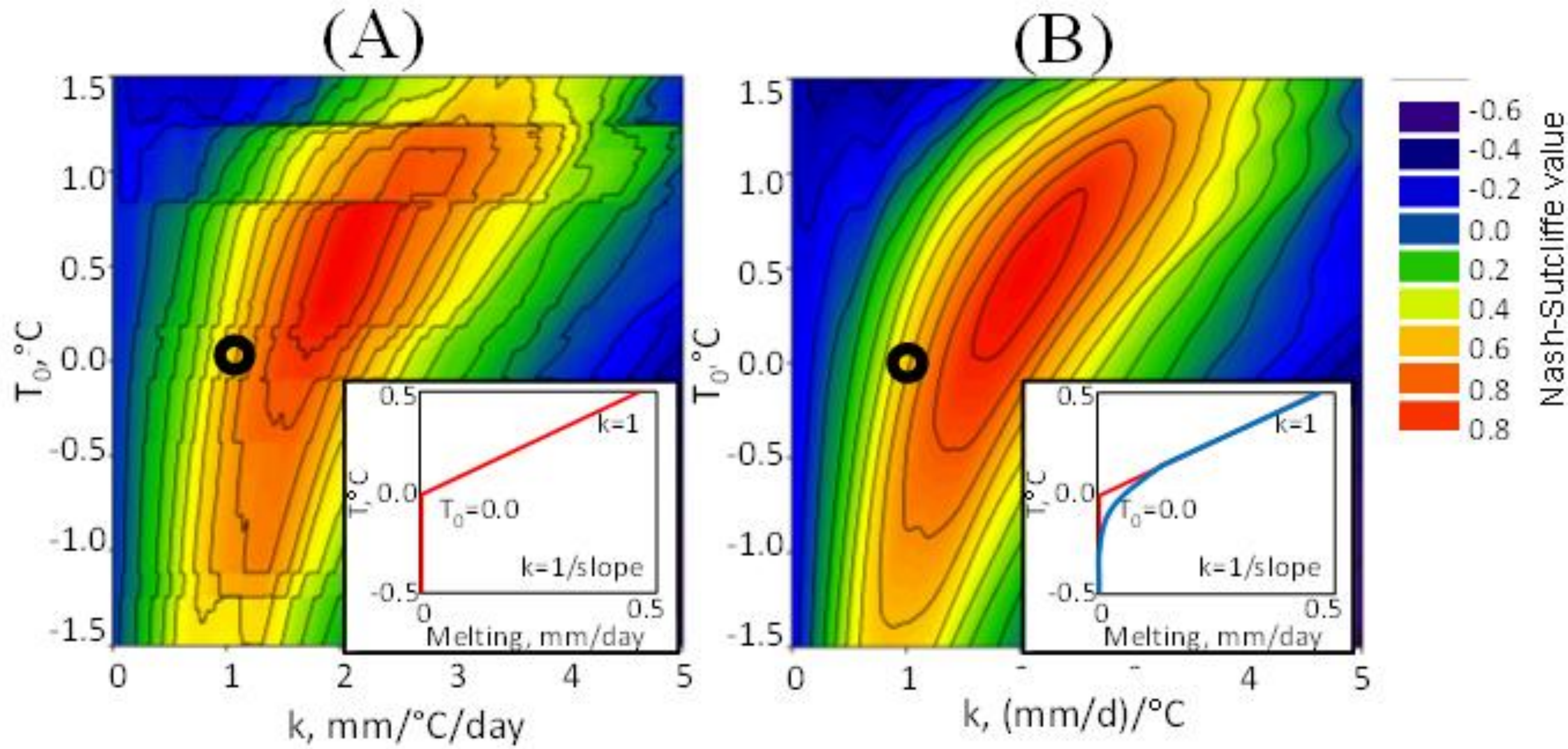


Smooth functions are also important to surrogate modeling

Razavi and Tolson 2012 WRR

Irregularities as in (A) can make surrogates inaccurate. Smooth functions as in (B) make surrogates easier to construct

Make the function more realistic



Kavetski and Kuczera 2007 WRR; Hill et al 2015 Groundwater

Programs developed to minimize numerical daemons – next generation models

- MODFLOW-OWHM (Hanson+ 2014 USGS)
  - Integrated GW-SW
  - Extensive land and water use support, including for
    - agriculture
    - diversions
    - demand-driven, supply-limited use of gw and sw
- Rainfall-runoff
  - TOPKAPI (Ciarapica & Todini 2002 Hyd Proc)
  - SUMMA (Clark + 2015a,b WRR; NCAR)

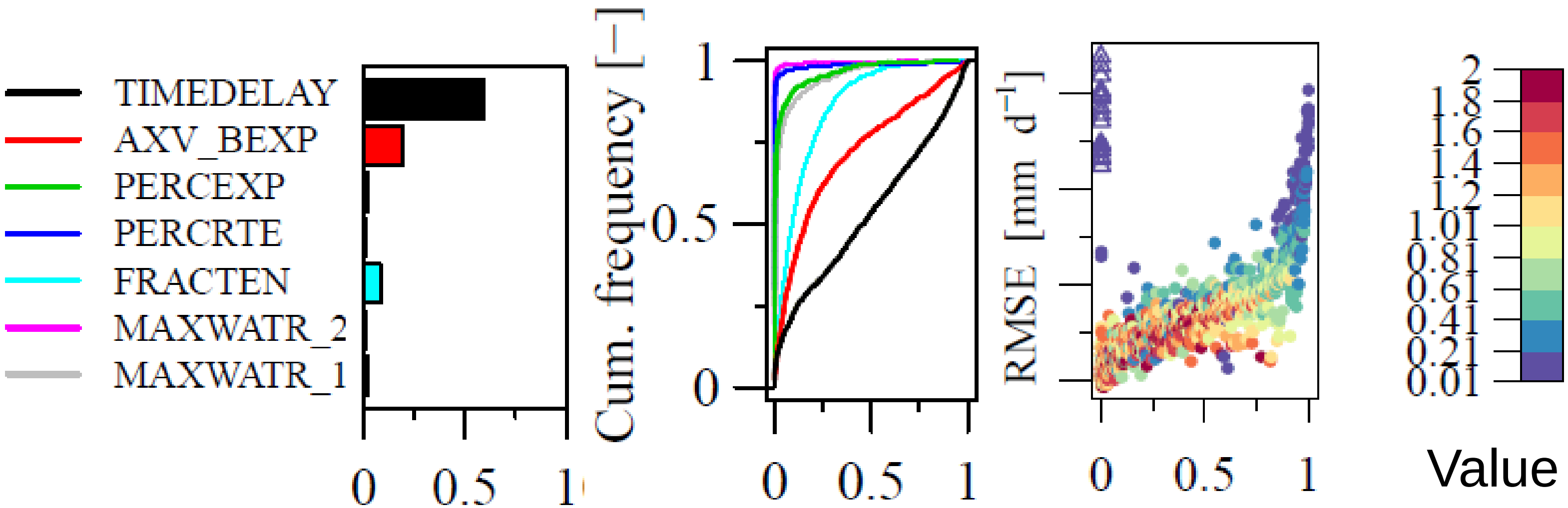
Conclusions

- This is a dynamic time in modeling
- Having a toolbox with a range or methods is important when doing UQ
- SA important to understanding system dynamics, explain the uncertainty, and understand what additional data would be most useful

Questions?

How the problem was obvious

Examine TIMEDELAY using DELSA results



Sobol’

One value “average” over parameter space

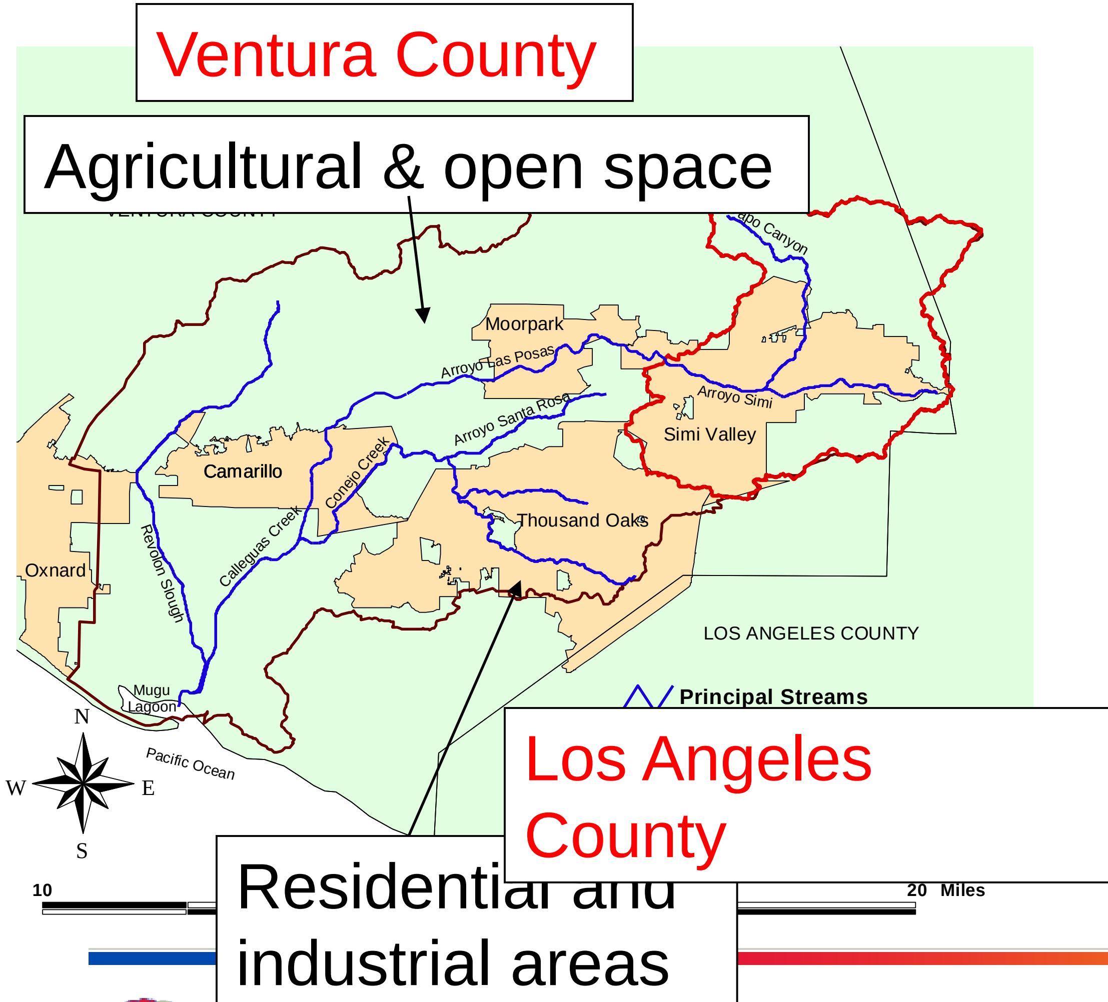
DELSA

Distribution of values over parameter space

DELSA for TIMEDELAY parameter

Value of TIMEDELAY (in days) parameter

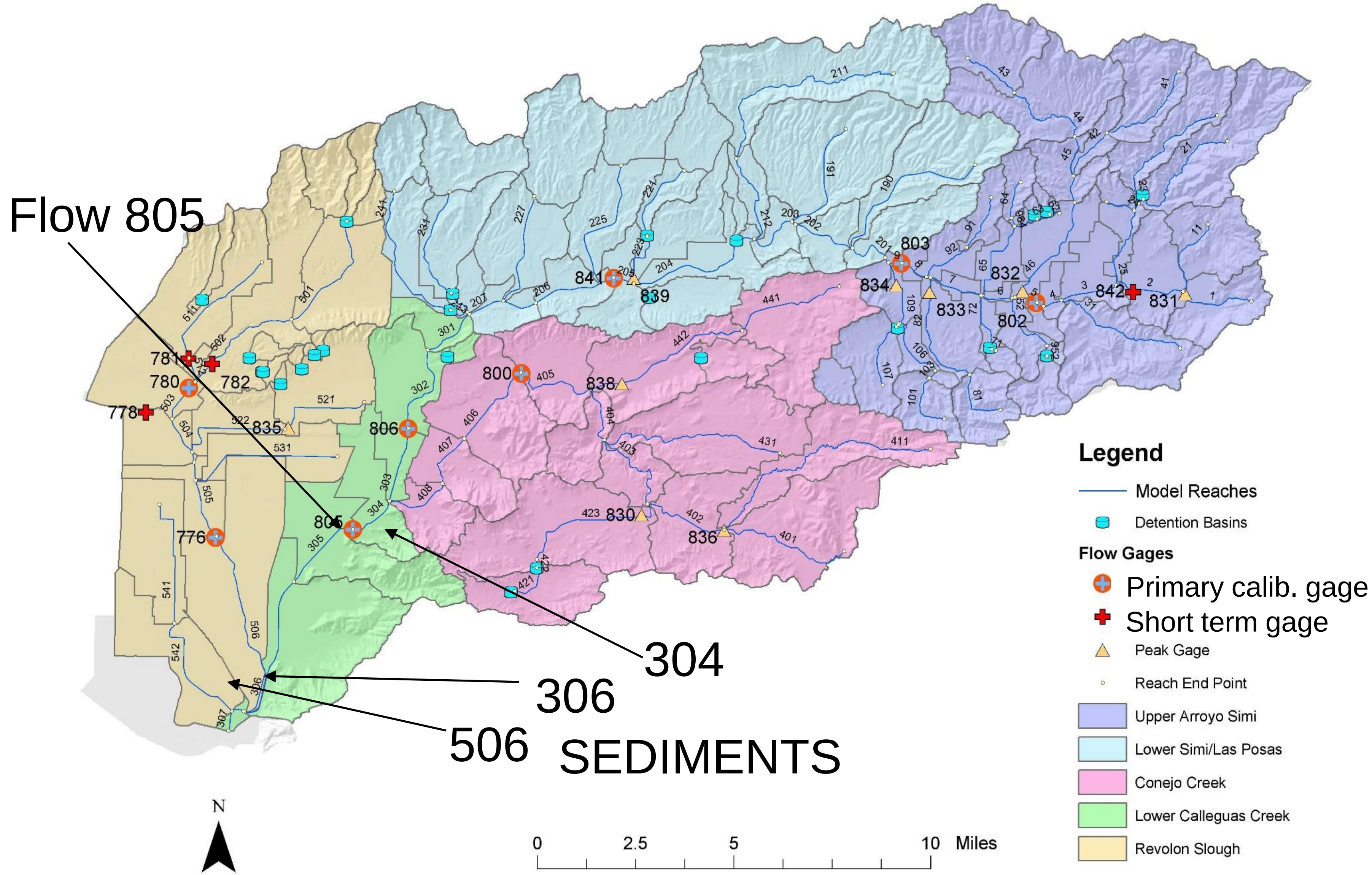
Example 1: The Calleguas Creek Watershed Model (CCWM)



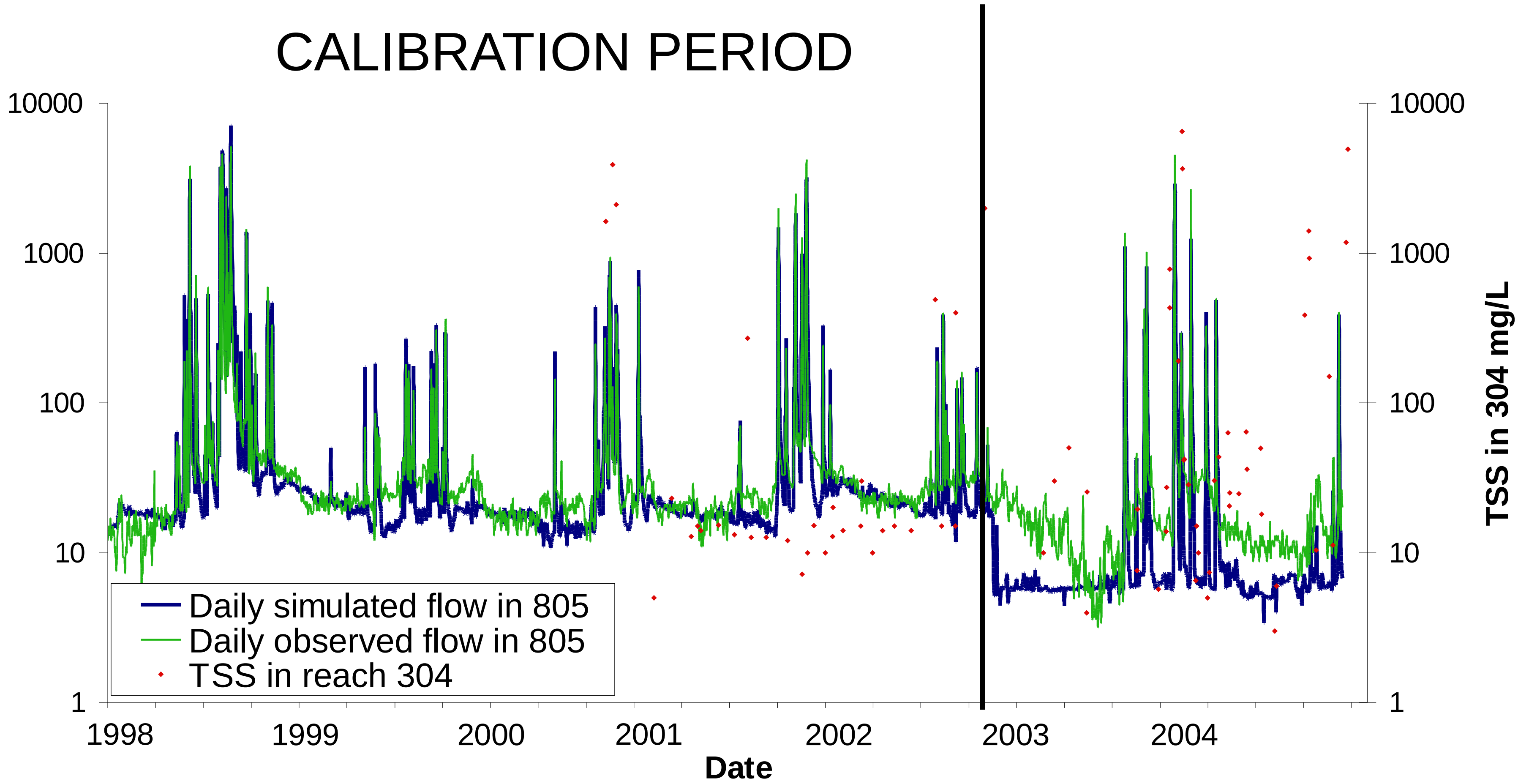
Main stem flows through flat valleys with urban and agricultural land.

Watershed subject to flooding and erosion. Sediment deposition downstream in Mugu Lagoon.

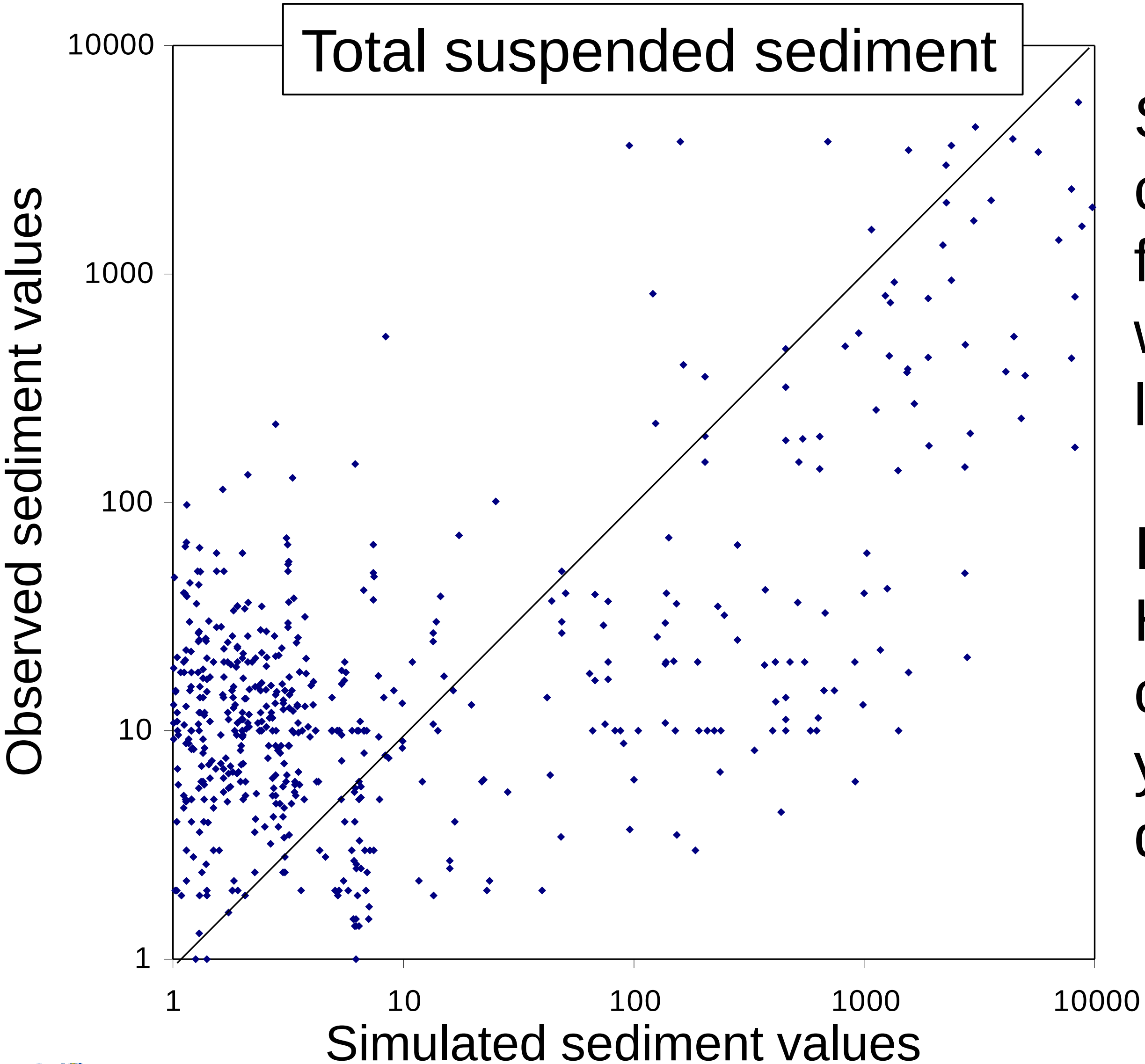
Study conducted with HSPF by
Laura Foglia, UC Davis, Larry Walker and Associates


Flow 805
304
306
506
SEDIMENTS
Legend
Model Reaches
Detention Basins
Flow Gages
Primary calib. gage
Short term gage
Peak Gage
Reach End Point
Upper Arroyo Simi
Lower Simi/Las Posas
Conejo Creek
Lower Calleguas Creek
Revolon Slough
N
0
2.5
5
10
Miles

Observed and simulated flows and observed sediment in downstream location



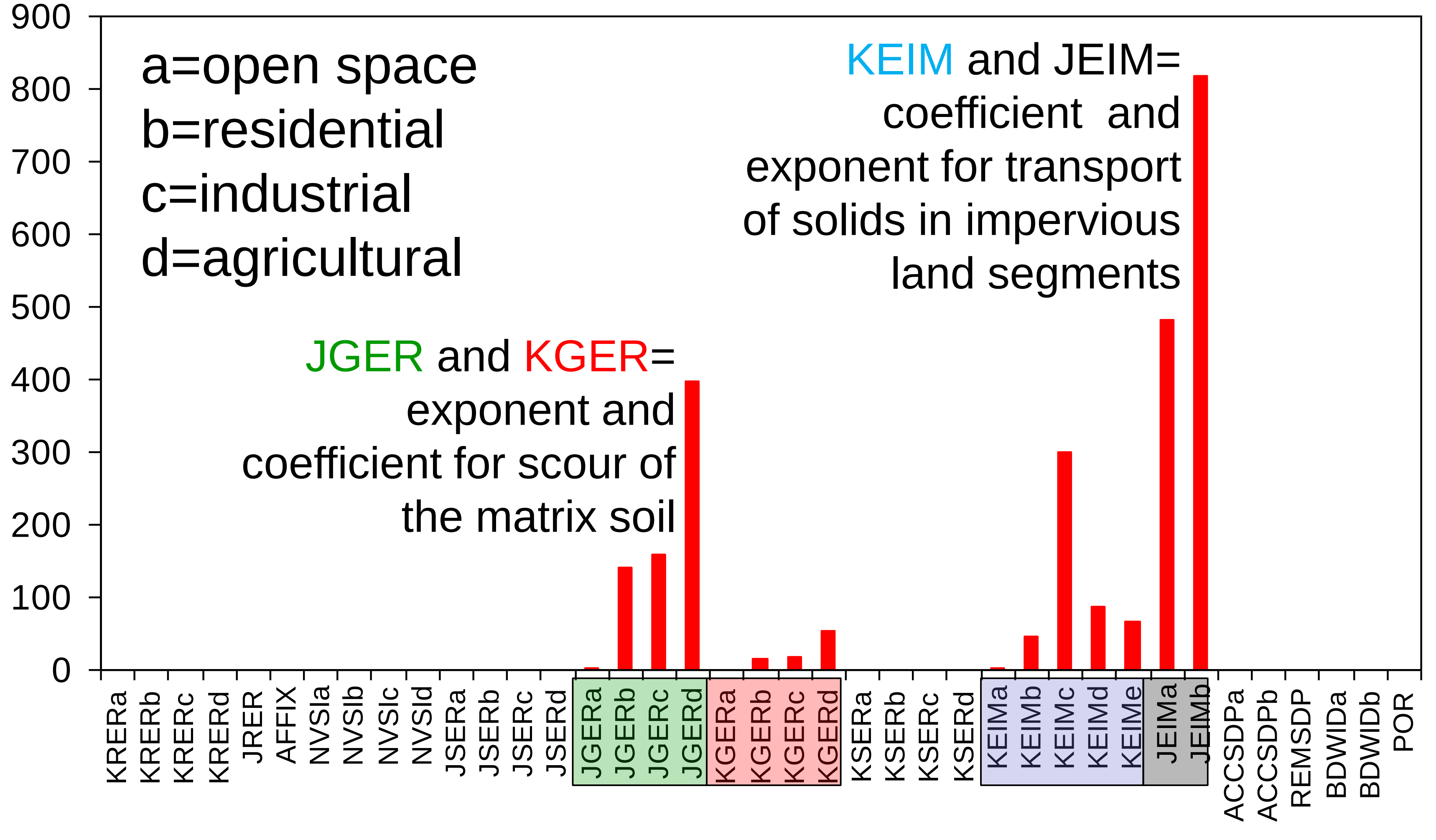
Sediment calibration: preliminary results



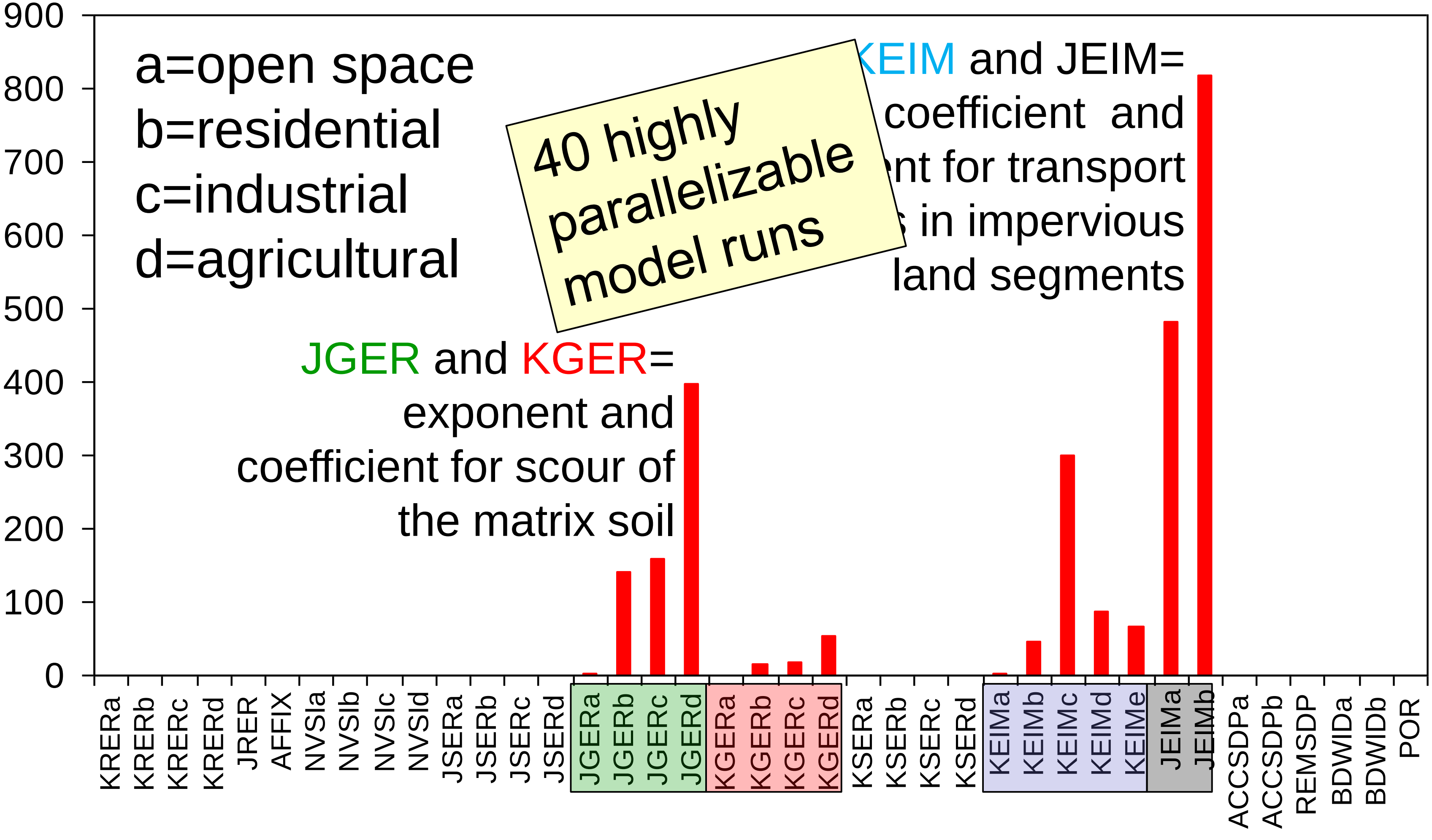
Sediment calibration:
few sparse data with
low quality

In contrast:
Hydrology calibration has years of continuous data, high quality

Consider sediment observations and parameters:
Composite scaled sensitivity
parameter correlation coefficients mostly small



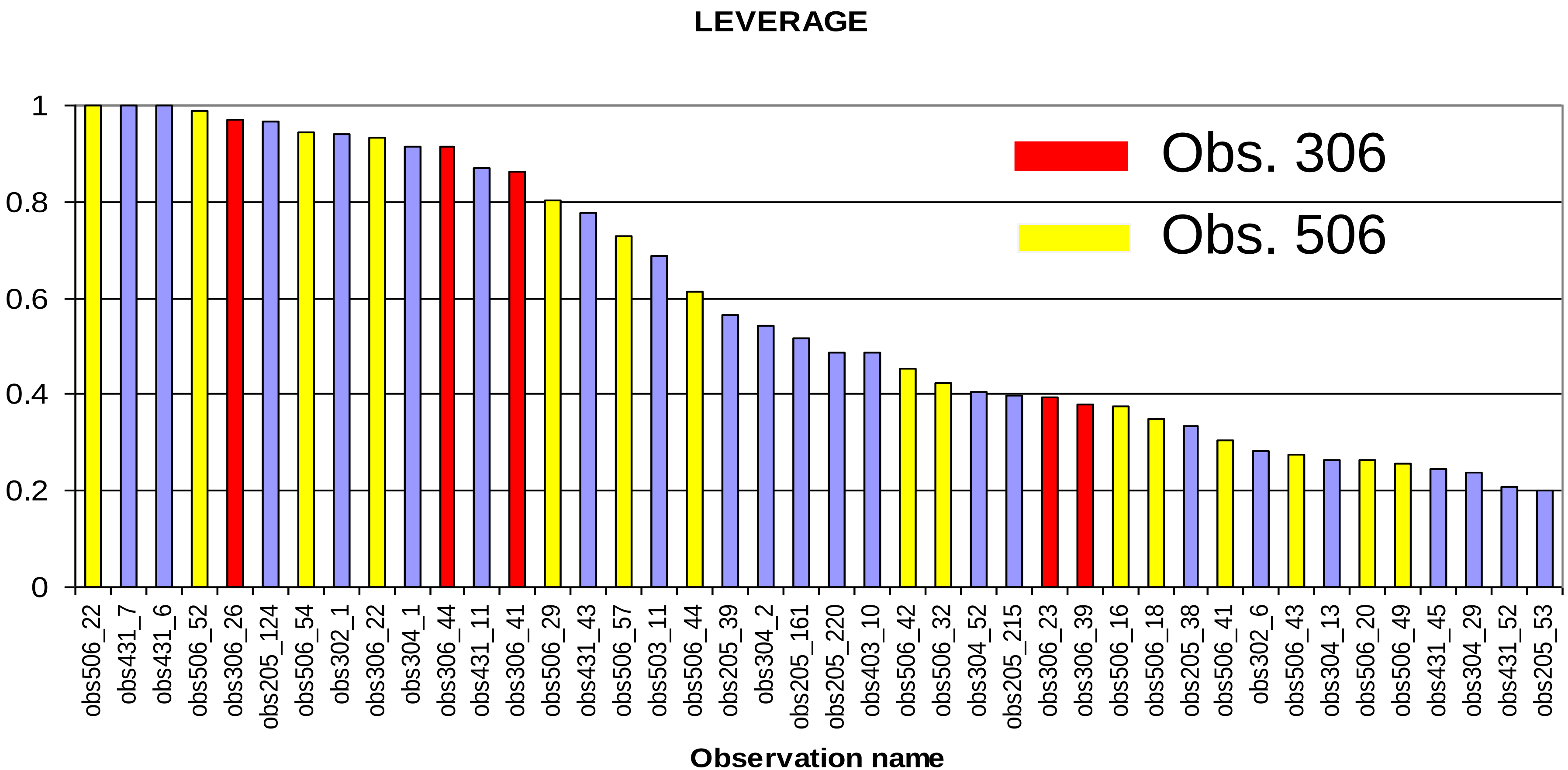
Consider sediment observations and parameters: Composite scaled sensitivity parameter correlation coefficients mostly small



Leverage statistics: which are the reaches with more influential observations?

None of the observations from the upstream reaches is included in the most influential.

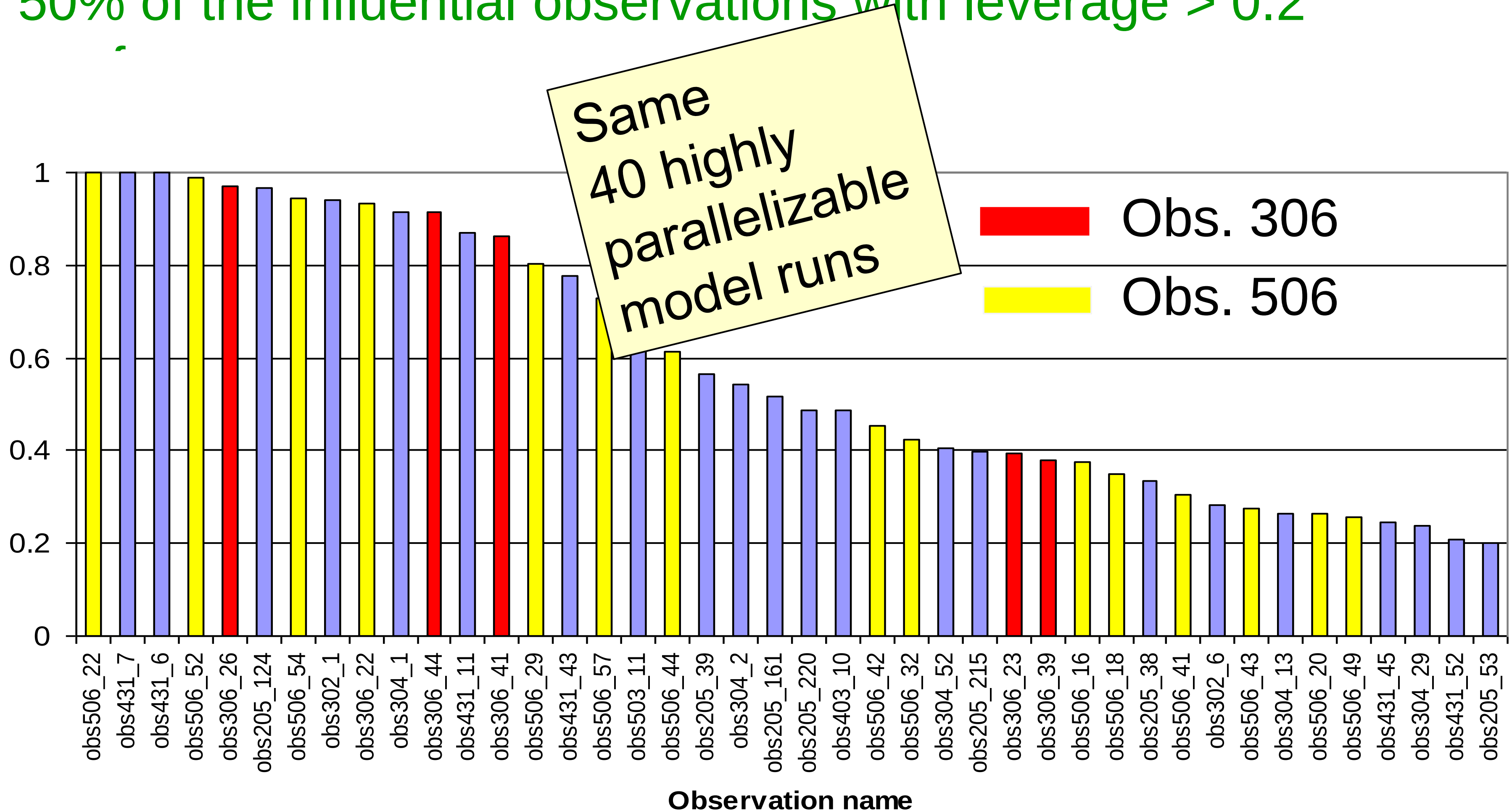
50% of the influential observations with leverage > 0.2



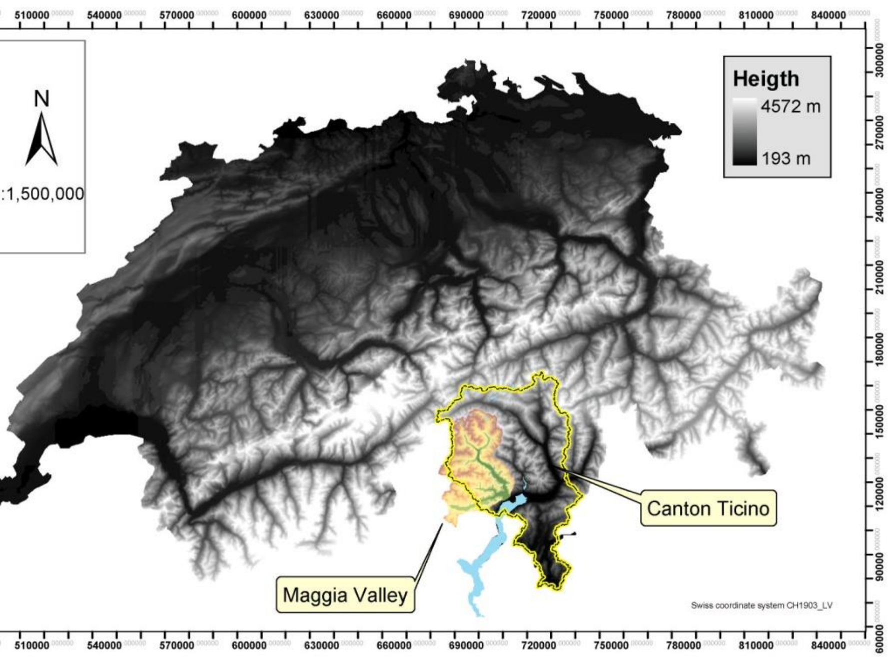
Leverage statistics: which are the reaches with more influential observations?

None of the observations from the upstream reaches is included in the most influential.

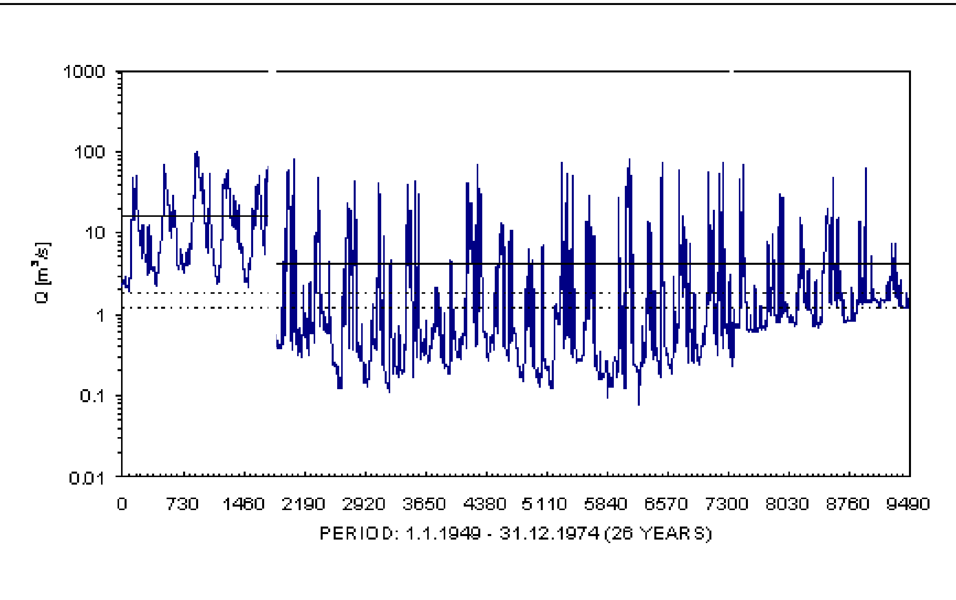
50% of the influential observations with leverage > 0.2



Example 2: Maggia Valley, southern Switzerland



- Goal: Integrated hydrologic model to help manage the ecology of this altered hydrologic system.



Maggia Valley, southern Switzerland

Series of studies to identify and test a convenient (fast to calculate) set of model analysis methods for use with the eventual computationally demanding integrated hydrologic model.

Use the computationally fast component models (GW & SW).

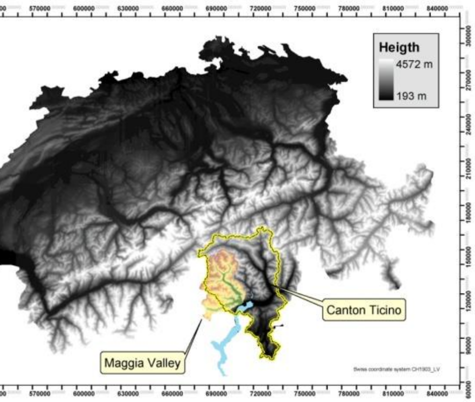
1. Test frugal sensitivity analysis (SA) using cross-validation

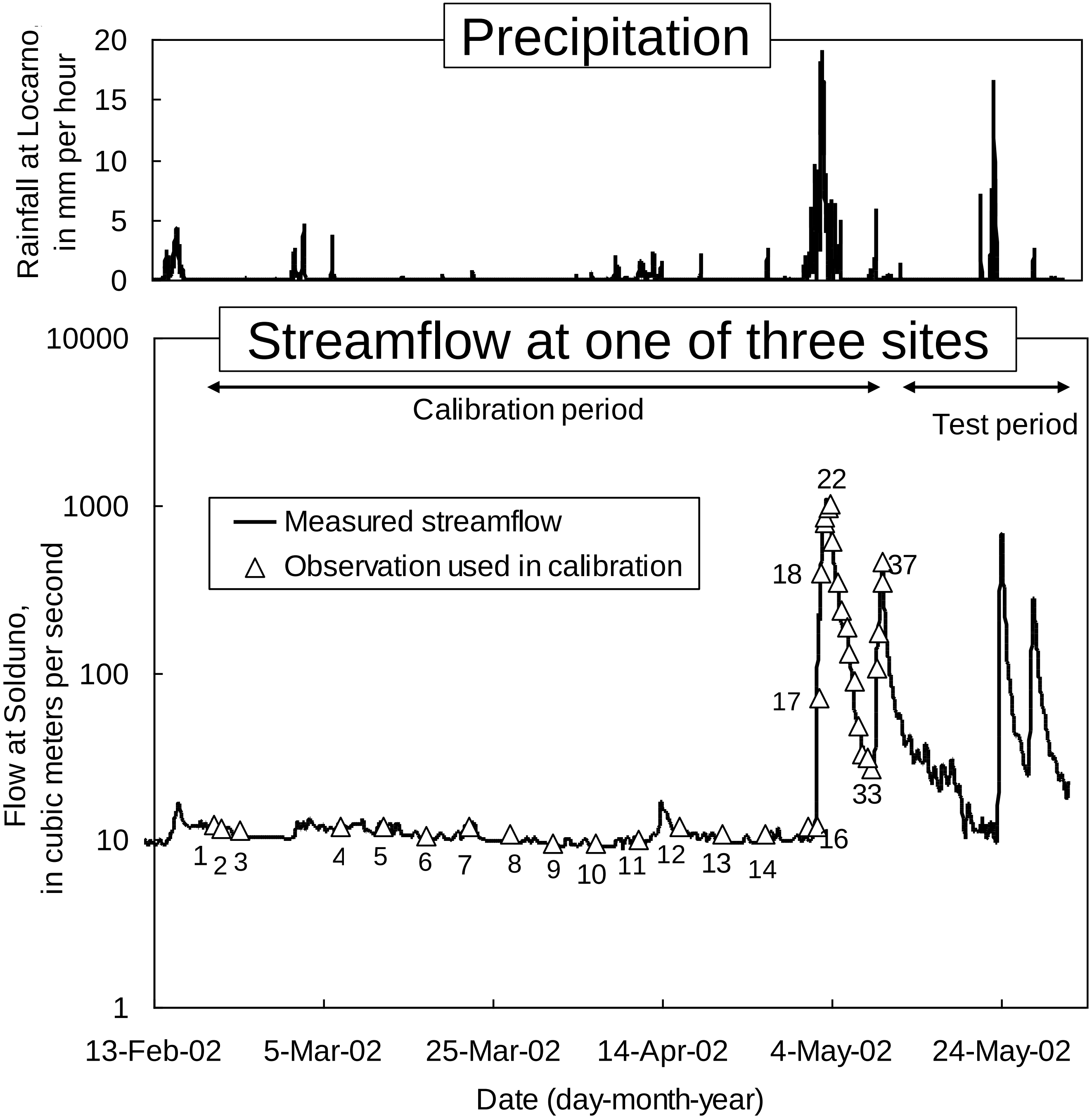
–Foglia + 2007 GW

2. Demonstrate frugal optimal calibration method

–(Foglia + 2009 WRR) TOPKAPI SW

3. Test multiple model UQ methods
   - Use **SA** and **calibration methods** (Foglia + 2013 WRR) GW



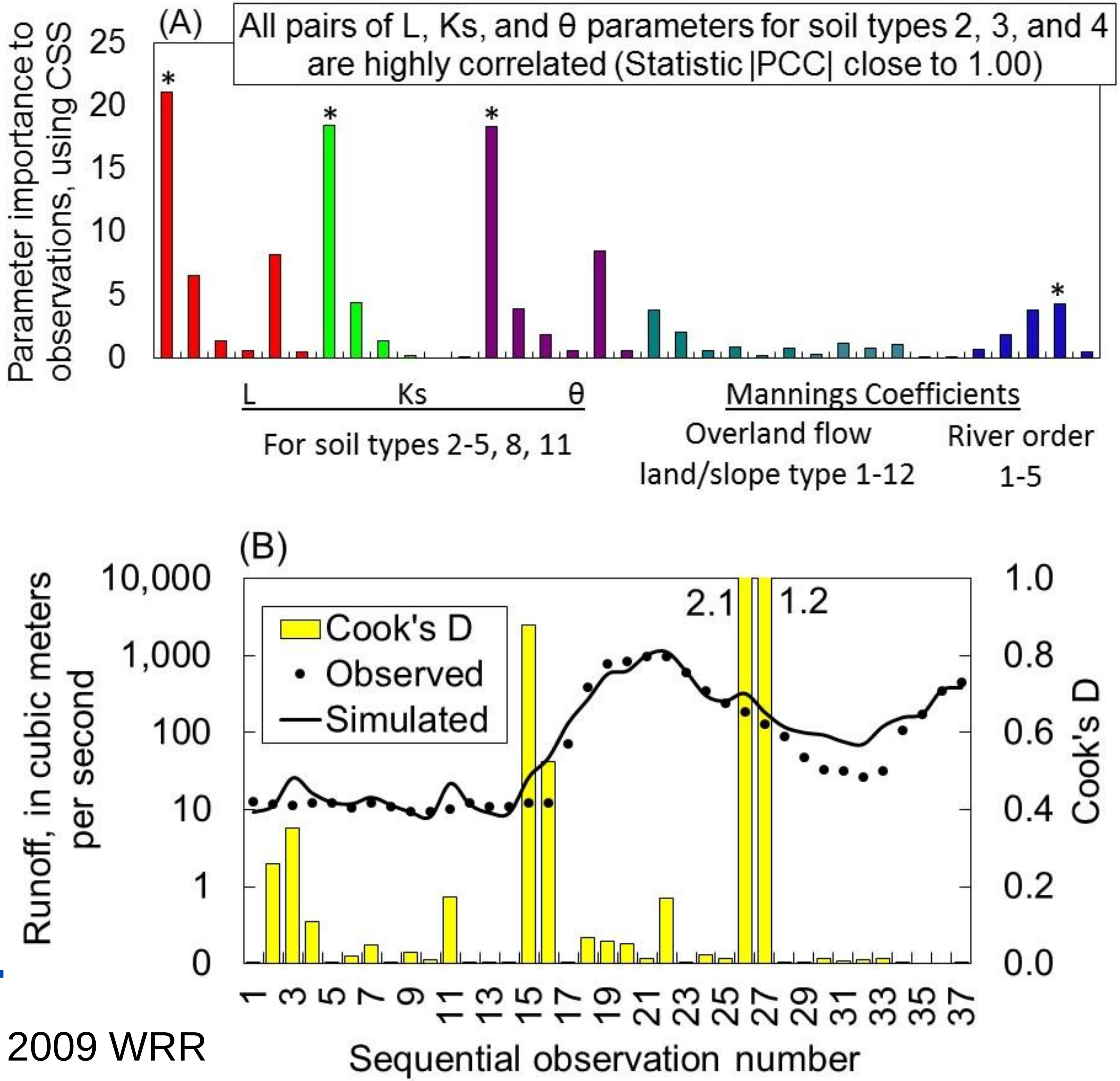


Error-based weighting.
Errors:
10% for most flows.
40% for flows corrected for dam operation and for all flows at one site with poor measurements

Foglia + 2009 WRR

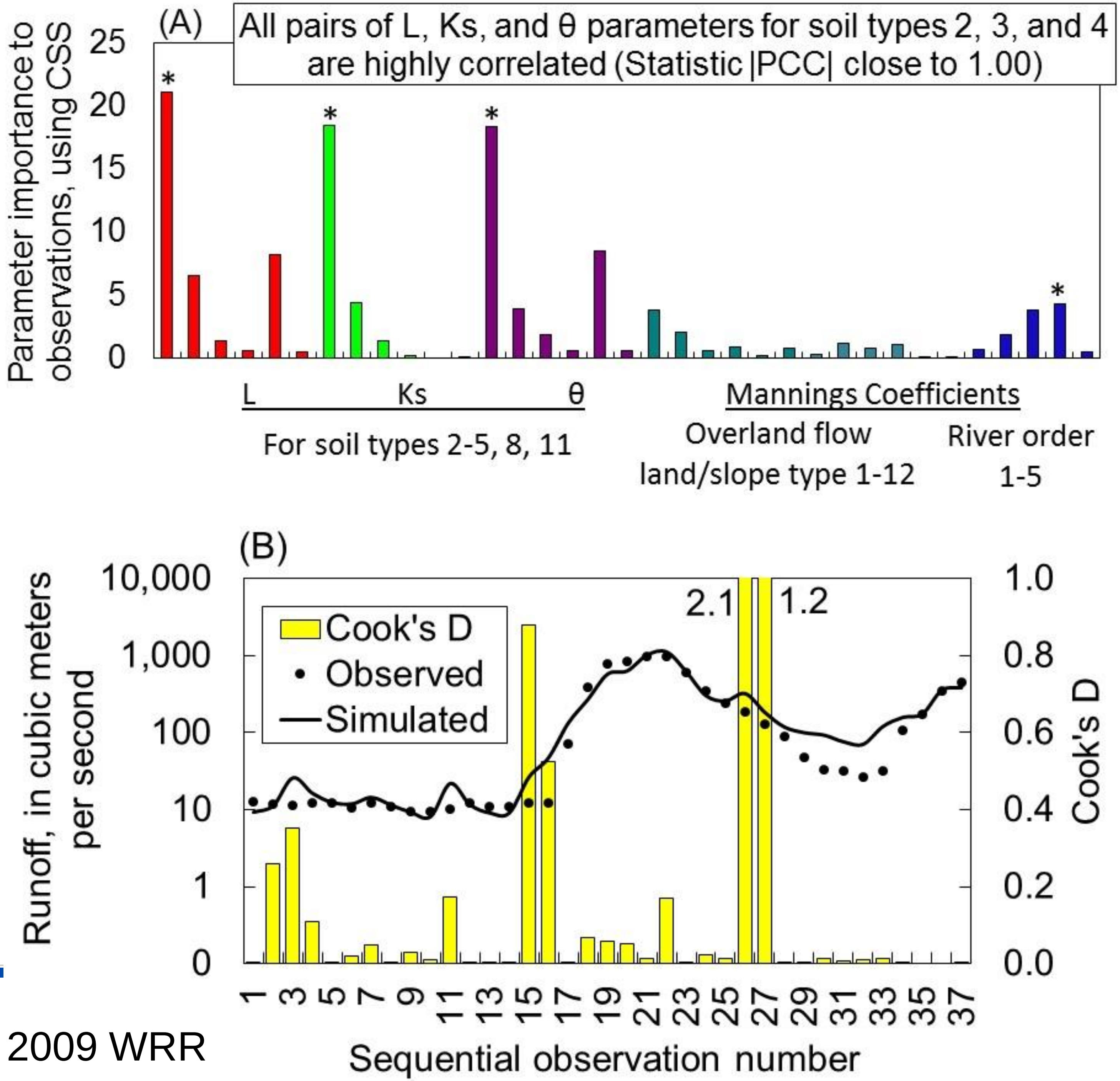
Parameter analysis: Observations and prior information are sufficient to estimate 3 parameters

Observation analysis: Low flows more important that expected. Resample low flows.



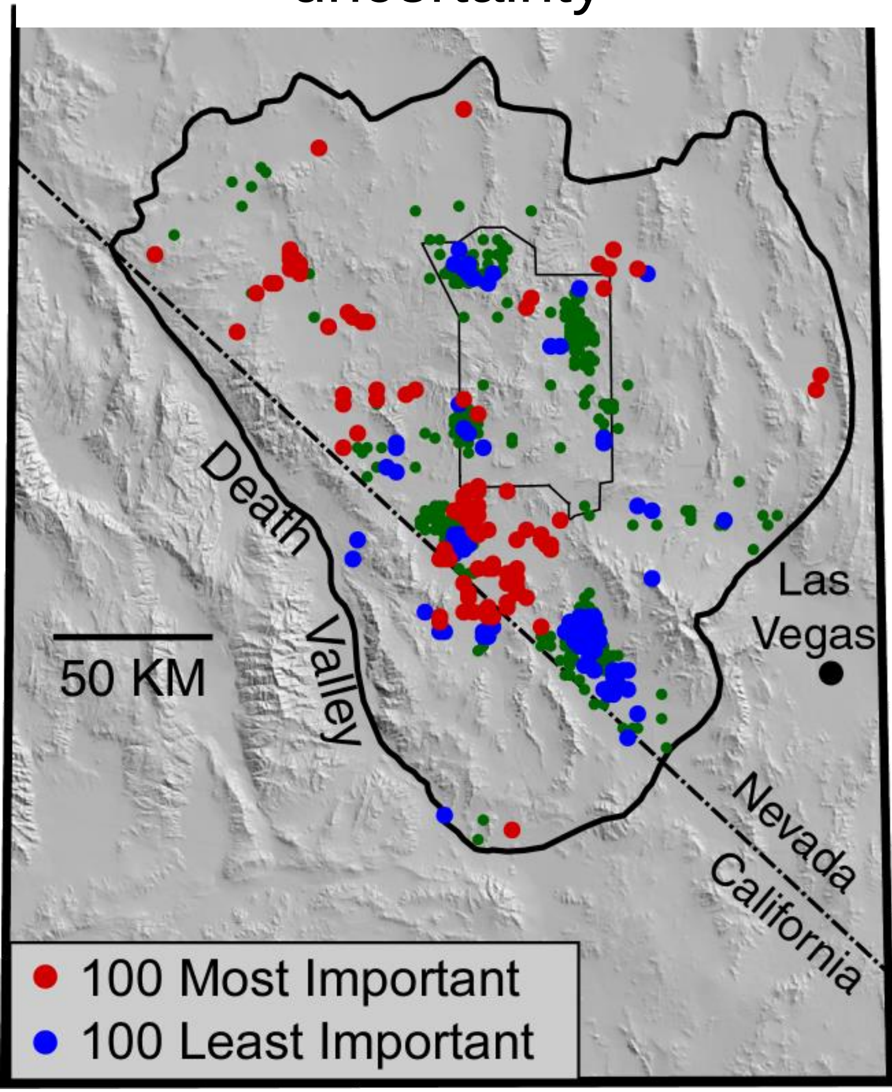
Foglia + 2009 WRR

71 highly parallelizable model runs

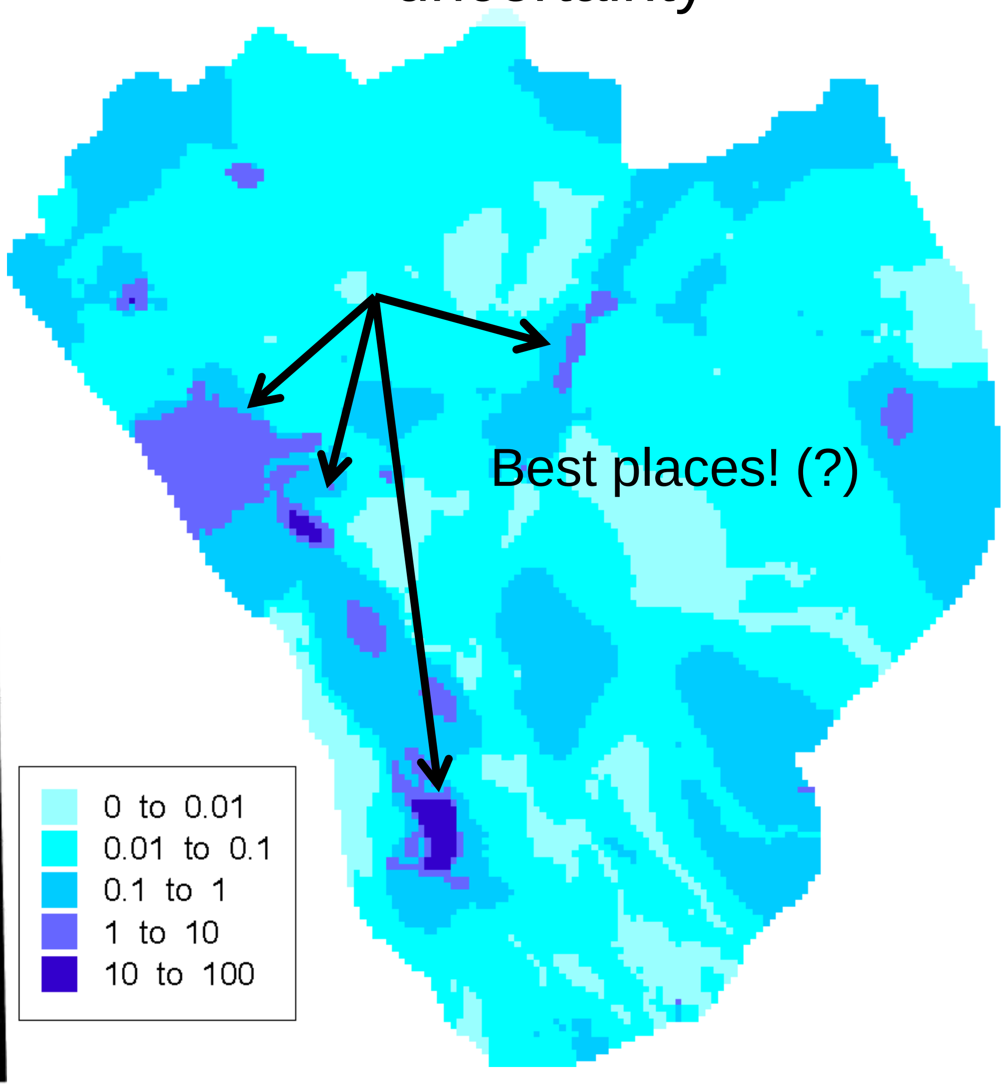


Foglia + 2009 WRR

Existing obs that reduce uncertainty



New data that could reduce uncertainty



Consider calculated results in the context of model construction.
Odd results may indicate model construction problems.

Tools for Data-Model Integration

Model that reasonably represents the system

↓

Predictions

Tools for Data-Model Integration

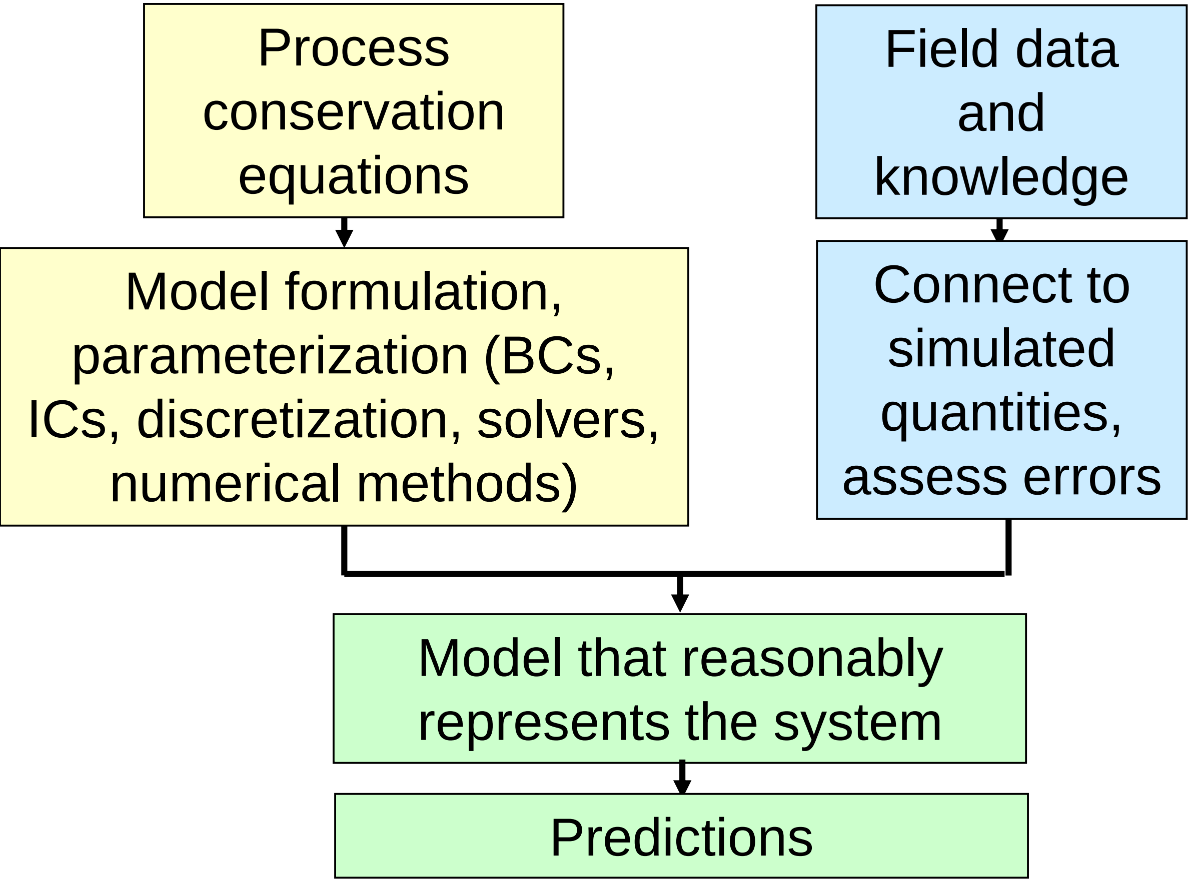
Process conservation equations

Field data and knowledge

Model that reasonably represents the system

Predictions

Tools for Data-Model Integration



Tools for Data-Model Integration

Process conservation equations

Model formulation, parameterization (BCs, ICs, discretization, solvers, numerical methods)

Field data and knowledge

Connect to simulated quantities, assess errors

Model that reasonably represents the system

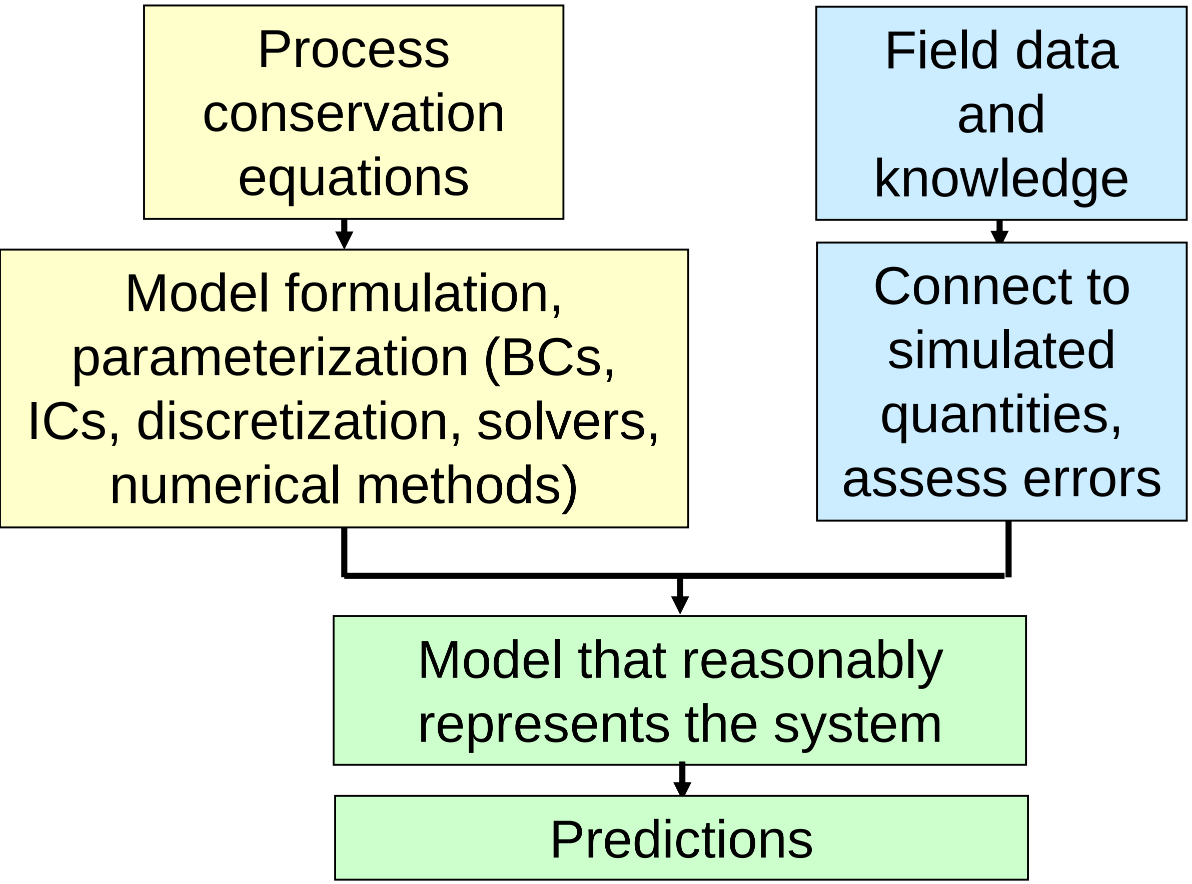
Predictions

Issues

Scale

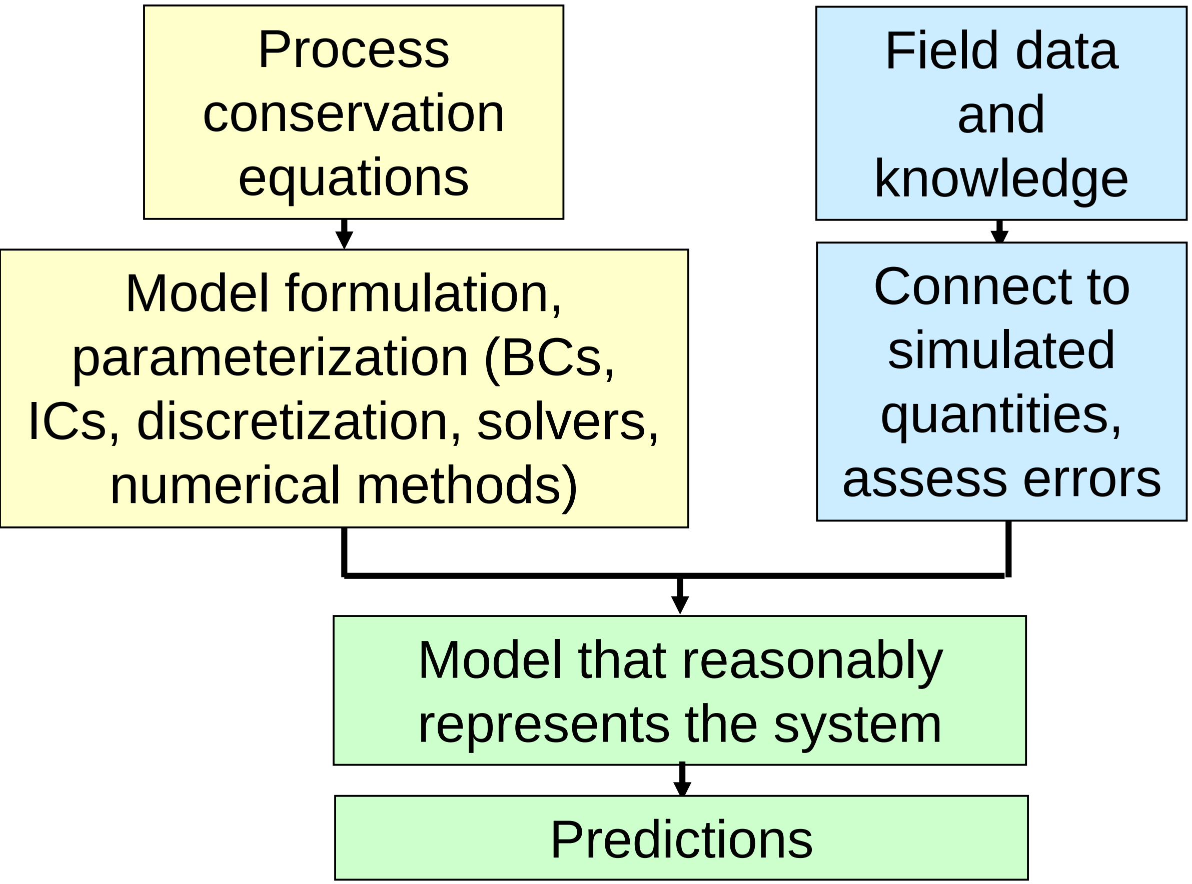
Unknowns

Tools for Data-Model Integration



Issues	Tools
Scale	-Hierarchical approaches
Unknowns	-Reduce and UQ

Tools for Data-Model Integration



Issues	Tools
Scale	-Hierarchical approaches
Unknowns	-Reduce and UQ
Realism	
Compatibility	
Right results for right reasons?	
Predictive ability?	

How to conduct SA and UQ?

It would be really nice to have a convenient (fast to calculate) set of model analysis methods

What questions will be addressed by the model analysis?

- What parameters are important to fitting observations?
- What observations are important to parameters?
- What observations and parameters are important to predictions and UQ?

Why is convenience important?

- Answer questions liked “Did the change I just made change the important parameters and observations, or UQ results?