UQ for Complex Models

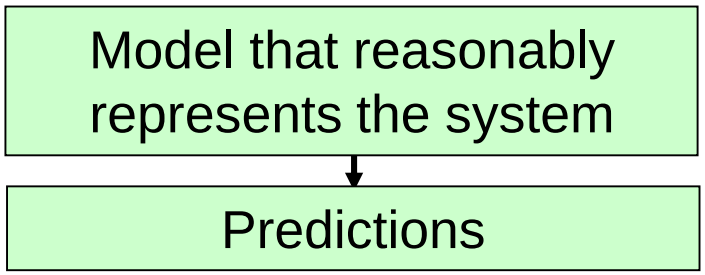
Professor Mary Hill

KU Department of Geology

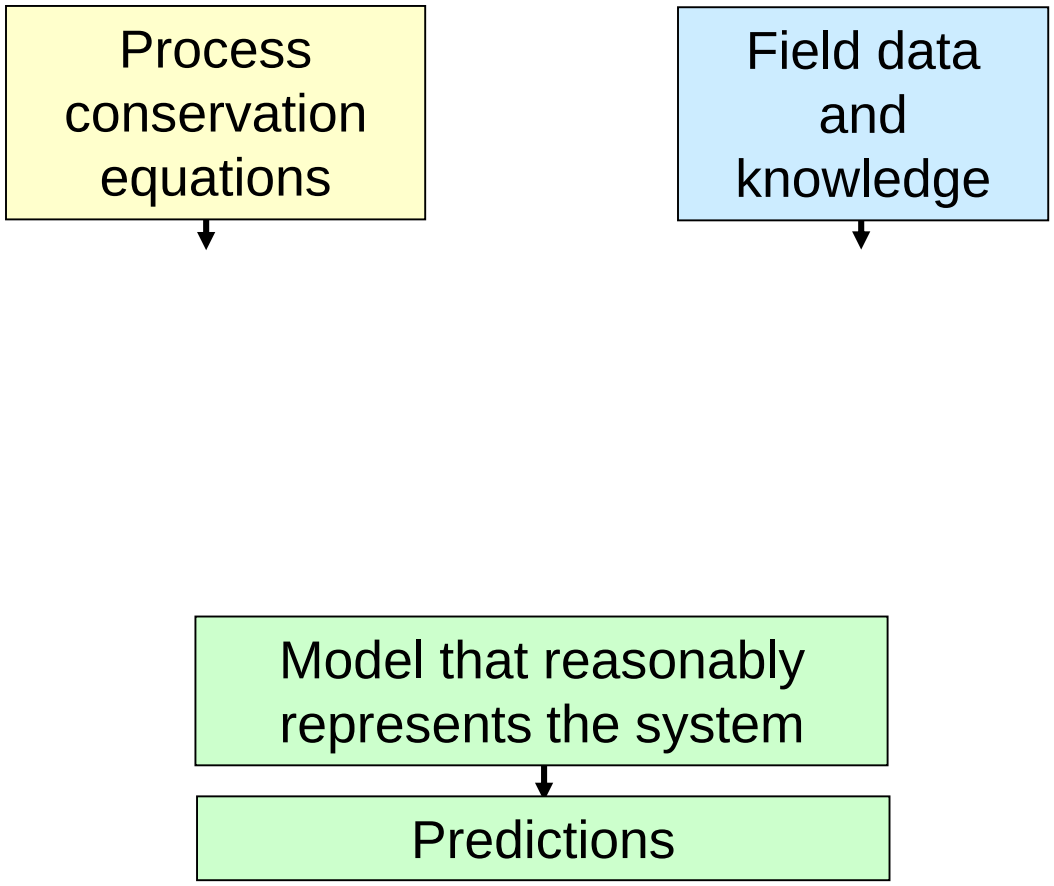
Brian Klager

USGS, KU Department of Civil Engineering

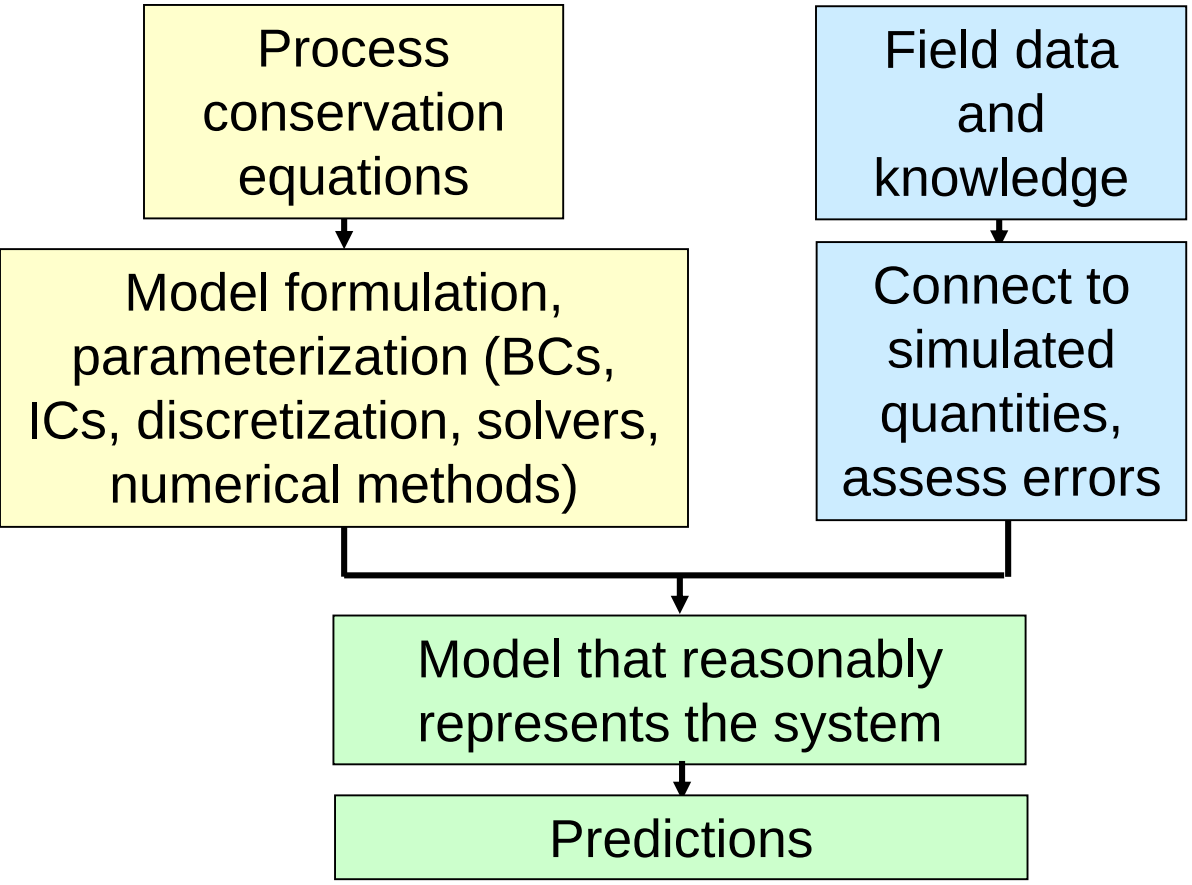
Tools for Data-Model Integration



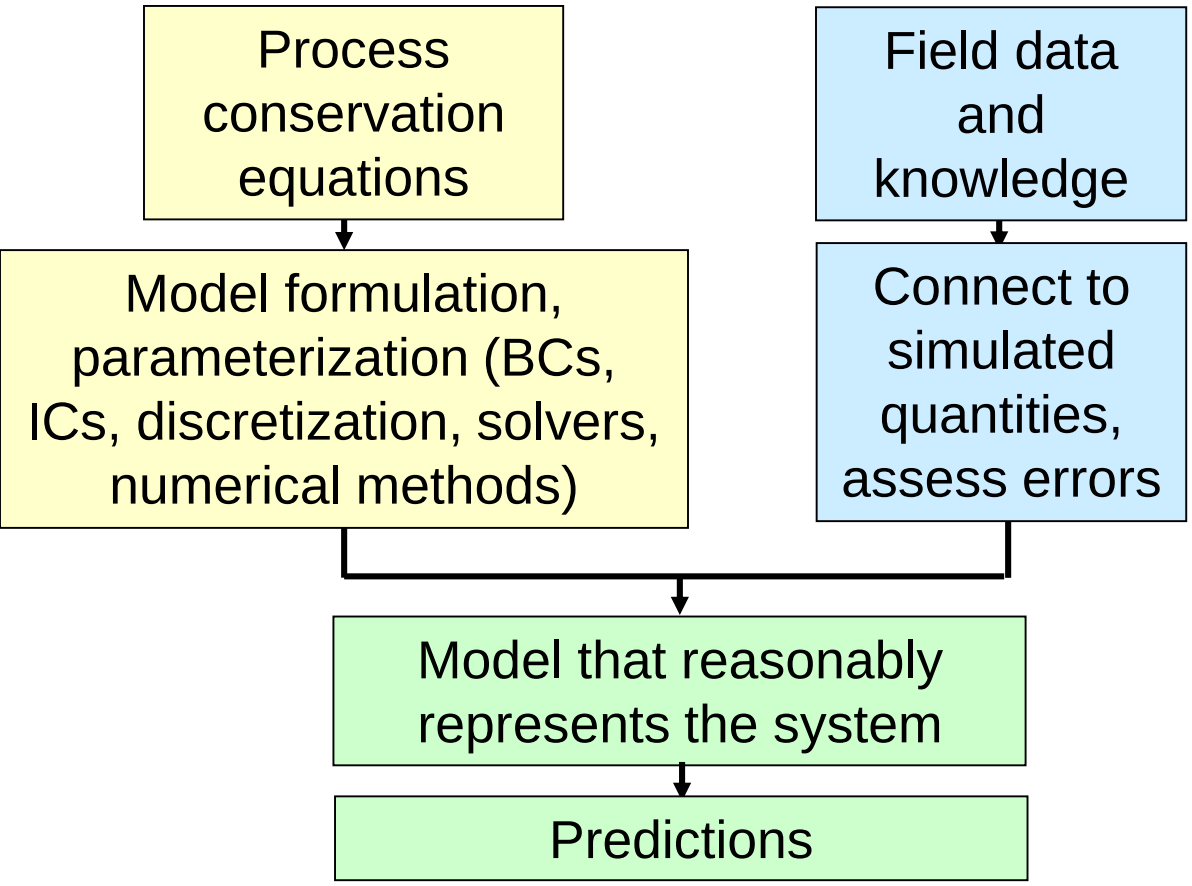
Tools for Data-Model Integration



Tools for Data-Model Integration



Tools for Data-Model Integration



Issues

Scale

Unknowns

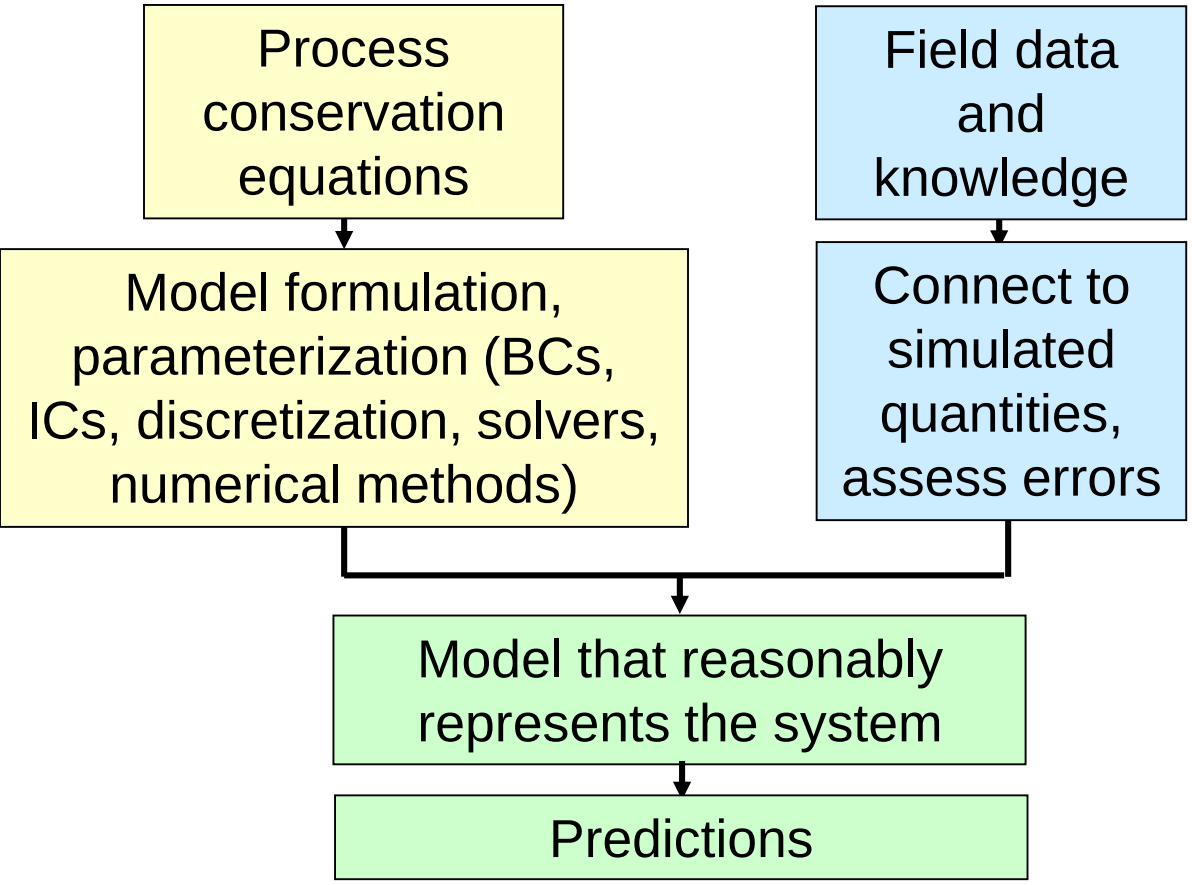
Realism

Compatibility

Right results for right reasons?

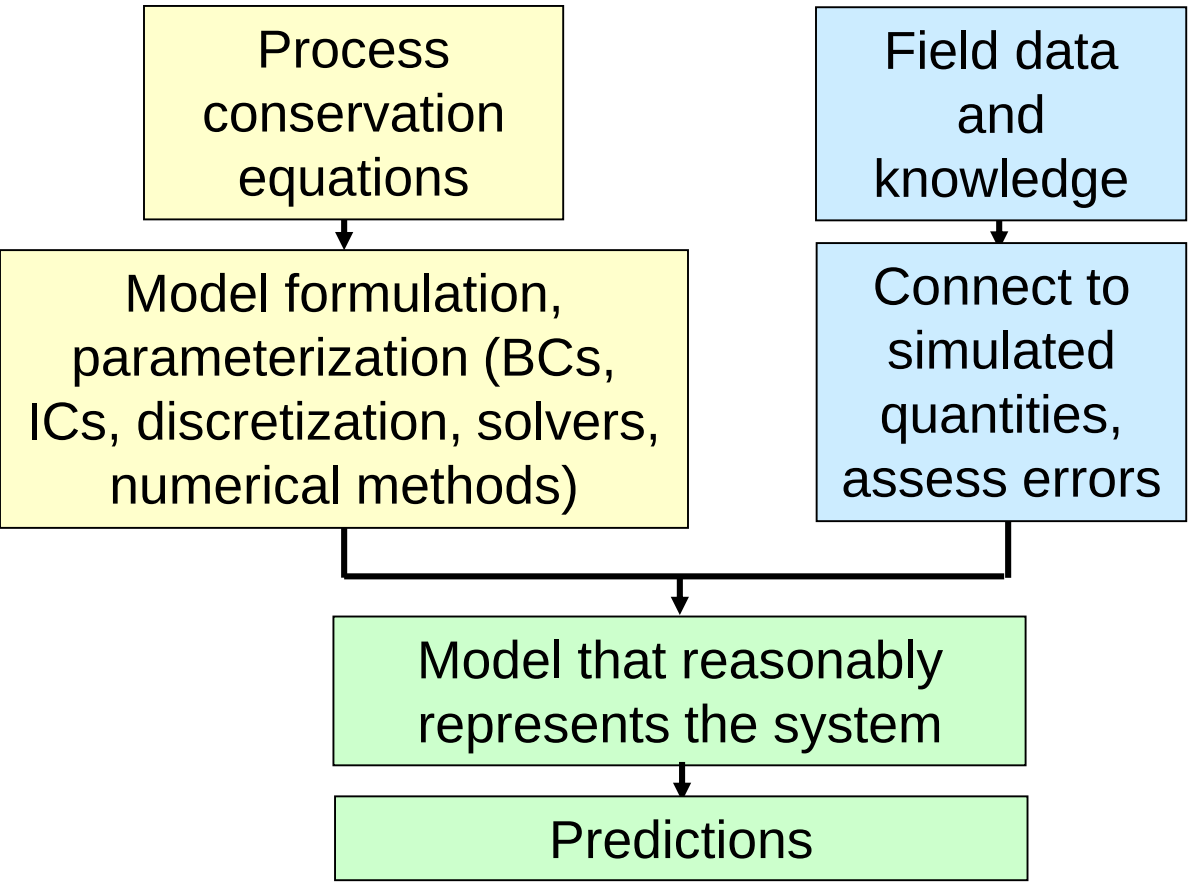
Predictive ability?

Tools for Data-Model Integration



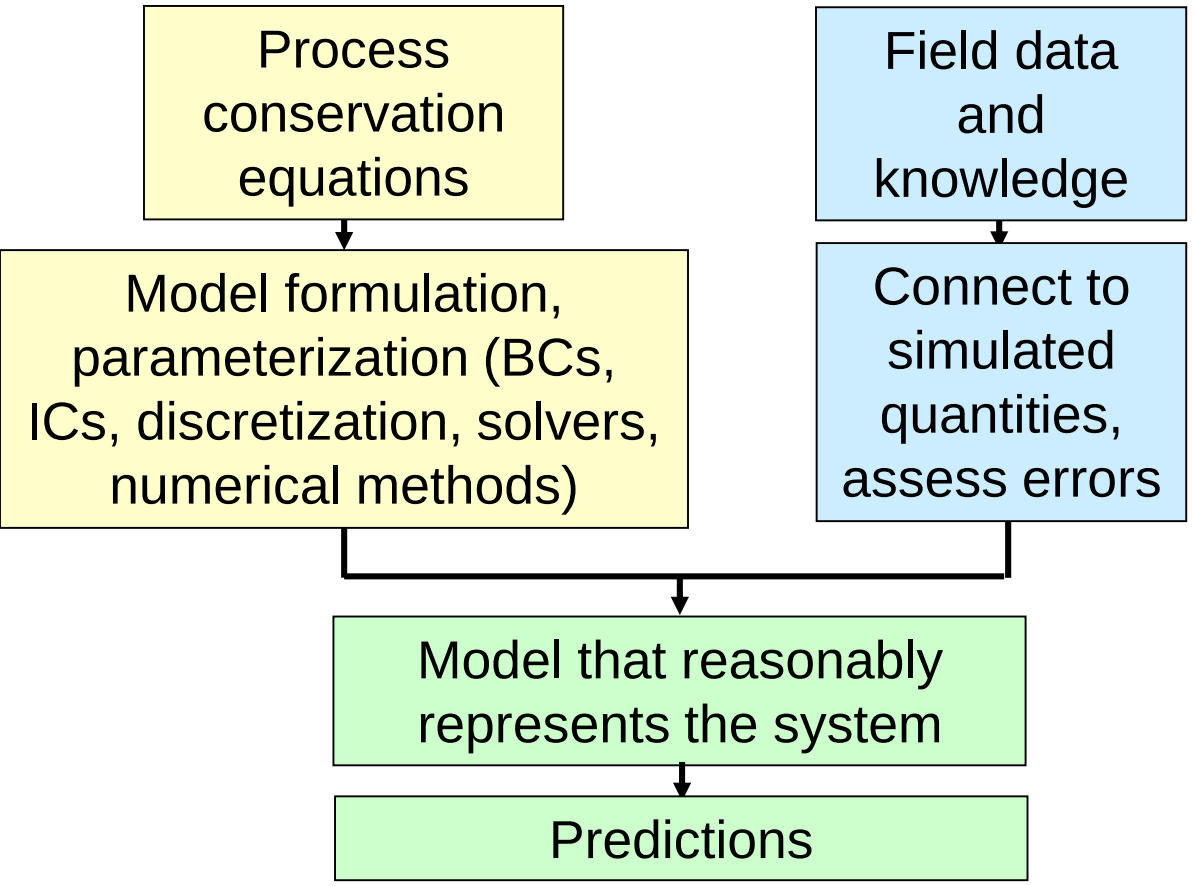
Issues	Tools
Scale	-Hierarchical approaches
Unknowns	-Reduce and UQ

Tools for Data-Model Integration



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

Tools for Data-Model Integration



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

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?

It’s a problem to have to use a semi truck every time you want to go to the store

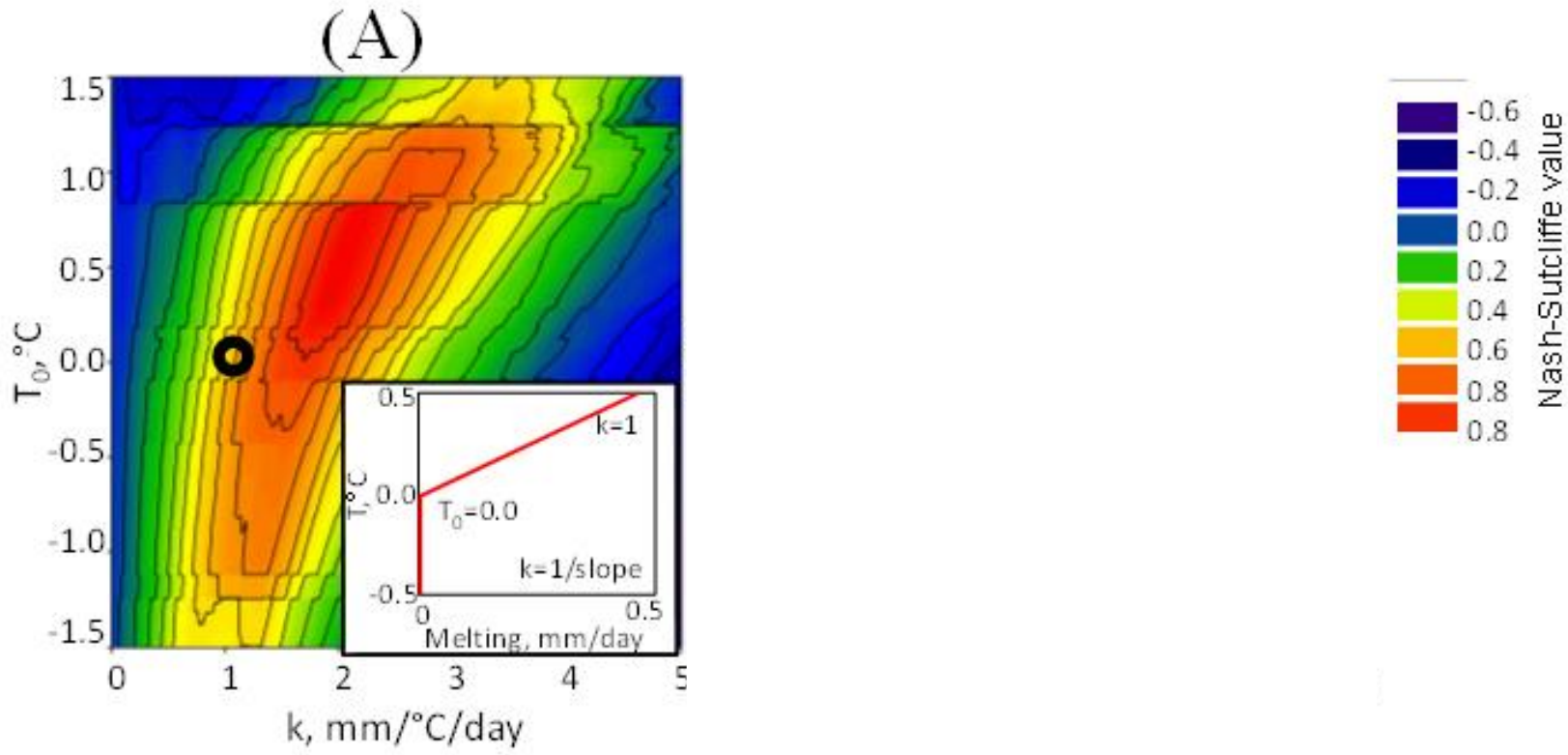
There convenient and demanding ways to address any model analysis question

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.	

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

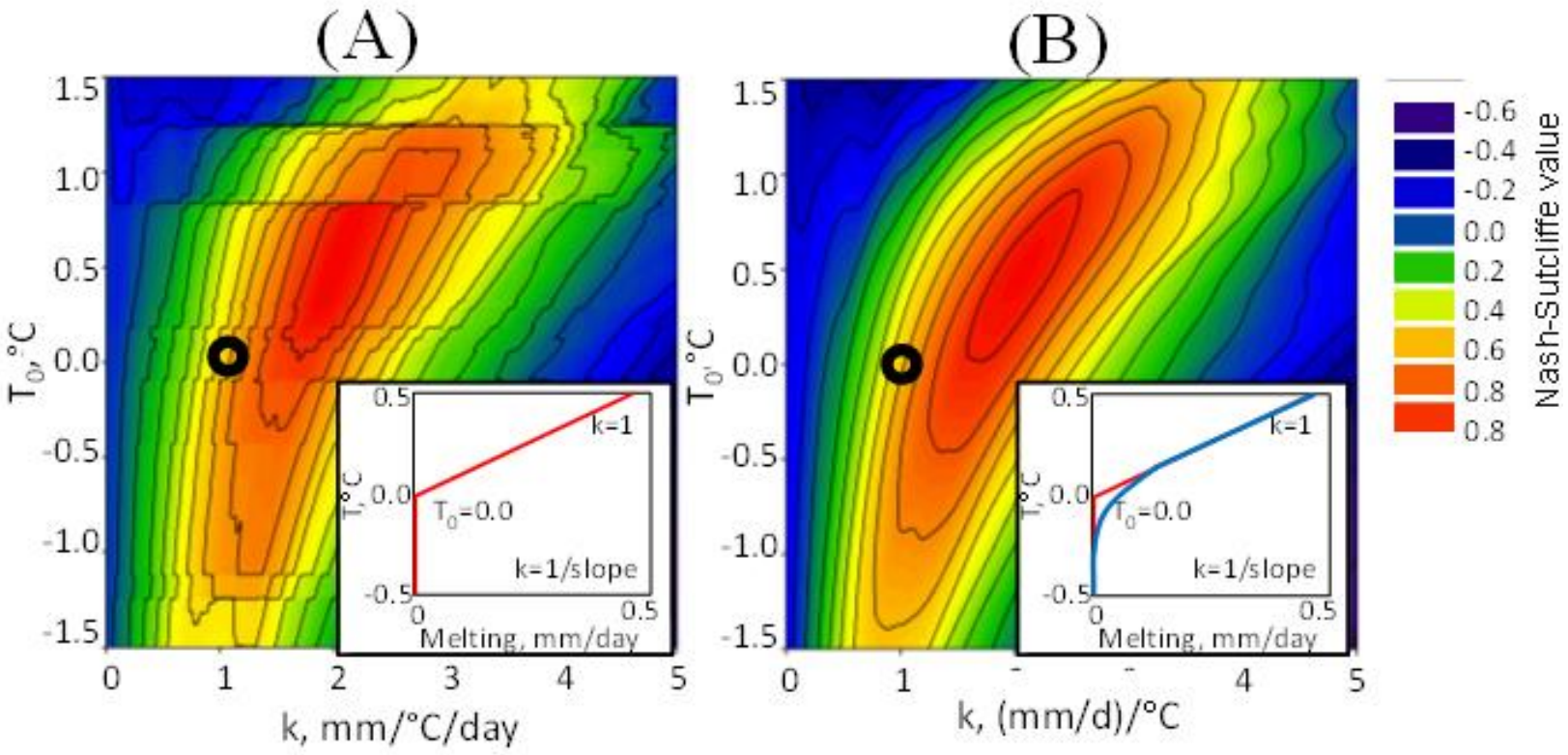
Called Numerical Daemons by Kavetski and Clark and Kavetski 2010



Kavetski and Kuczera 2007 WRR; Hill et al 2015 Groundwater

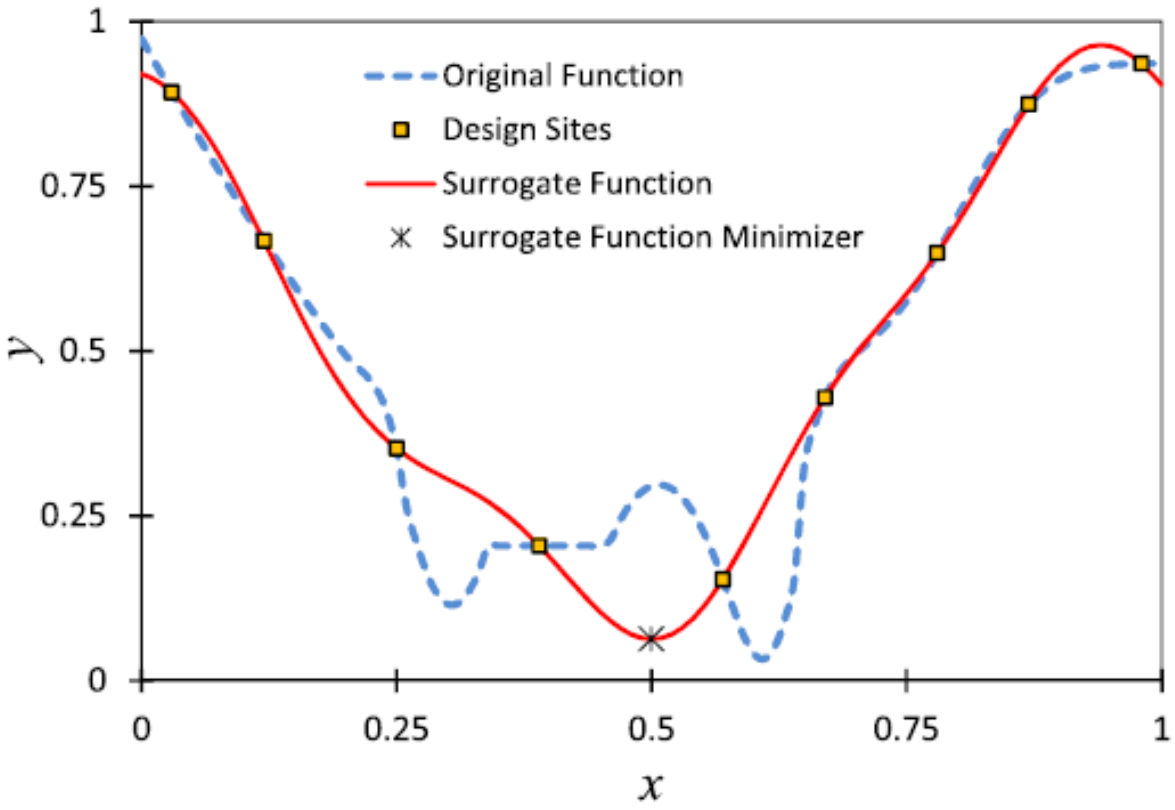
Fix the problem

Make the function more realistic



Kavetski and Kuczera 2007 WRR; Hill et al 2015 Groundwater

How surrogates can fail – a conceptual example



Razavi and Tolson 2012 WRR

Using DELSA to Detect Numerical Daemons

If they are there -- Fix them or use computationally demanding methods

DELSA: Distributed Evaluation of Local Sensitivity Analysis

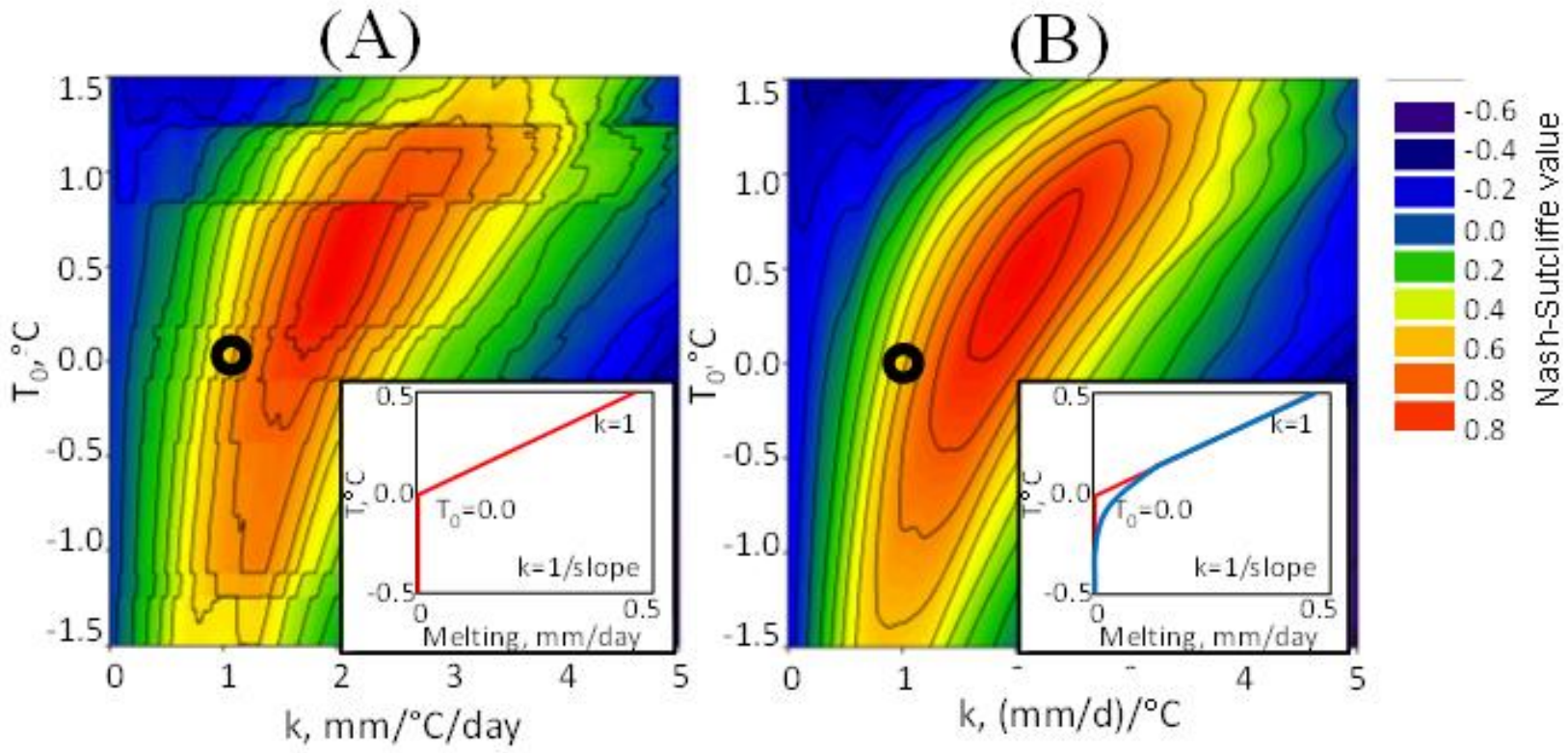
Rakovec et al 2014 Water Resources Research (+ some new plots): Compares DELSA with demanding method of method (Sobol') for sensitivity analysis

Sobol' is

- global
- widely used in hydrology and
- often takes 1,000s to 10,000s of model runs

DELSA would produce erratic, useless UQ or SA results for (A), and useful results for (B)

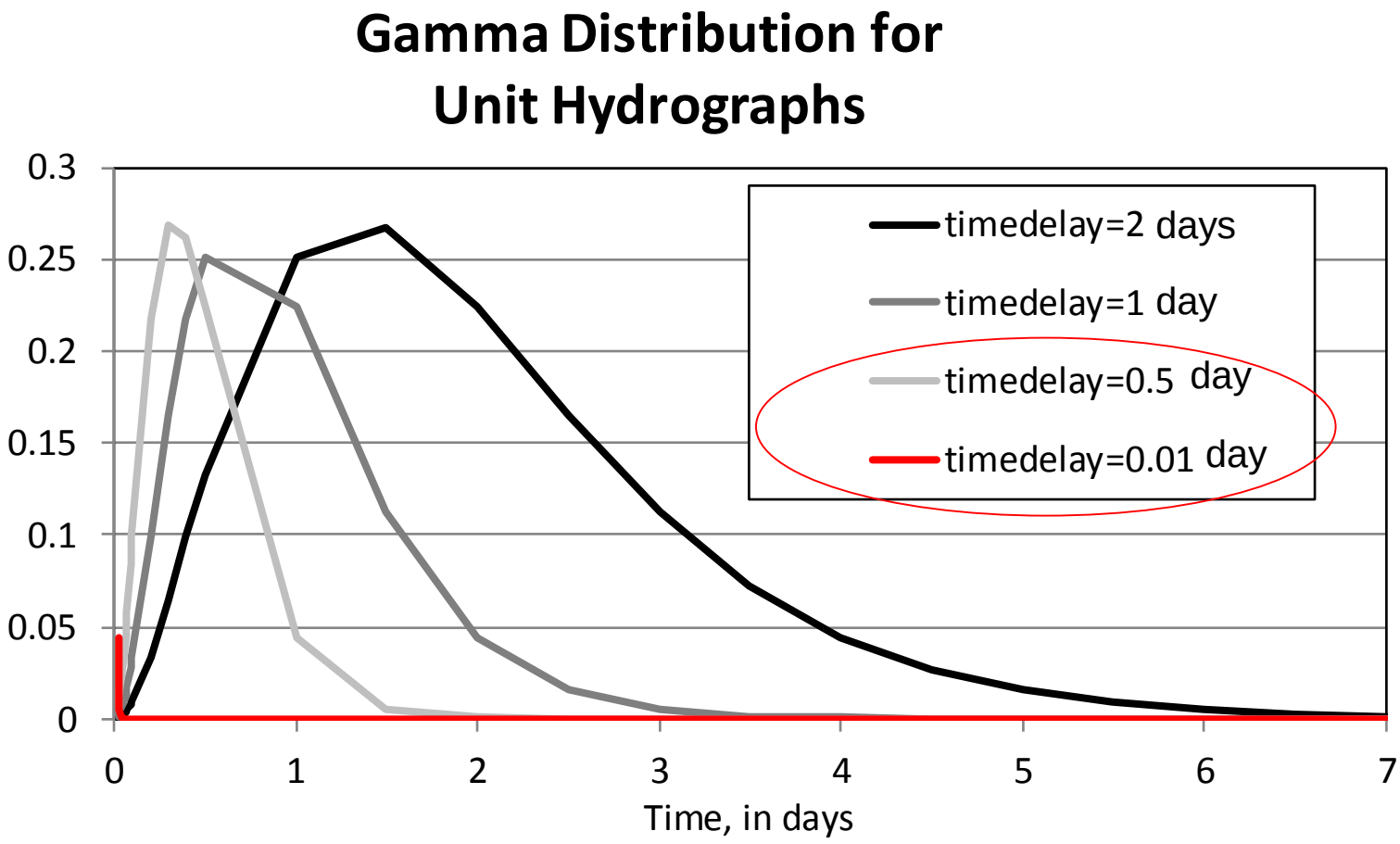
Make the function more realistic



Kavetski and Kuczera 2007 WRR; Hill et al 2015 Groundwater

Use DELSA to Find Numerical Daemon

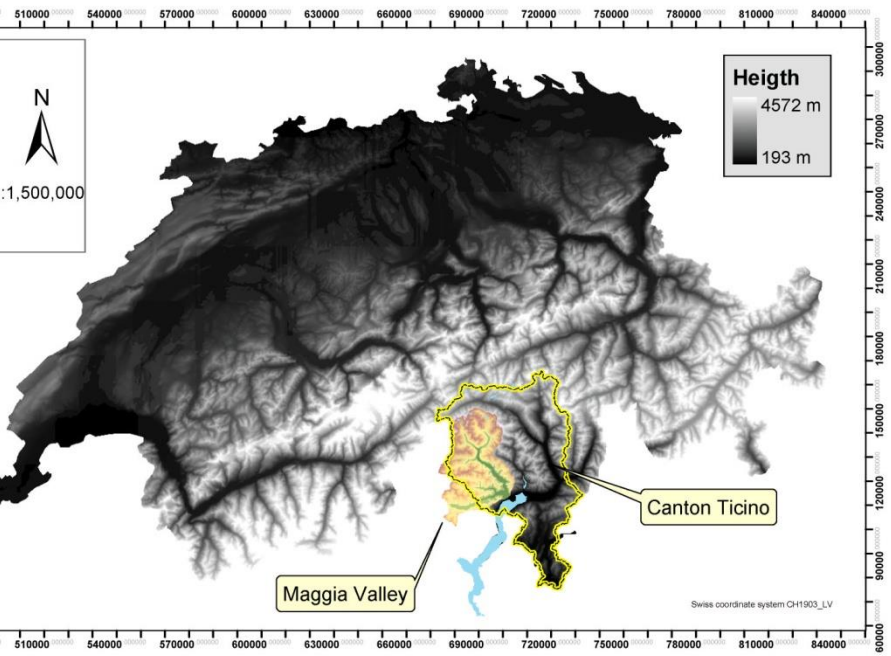
DELSA was introduced by Rakovec Hill Clark + 2014 WRR



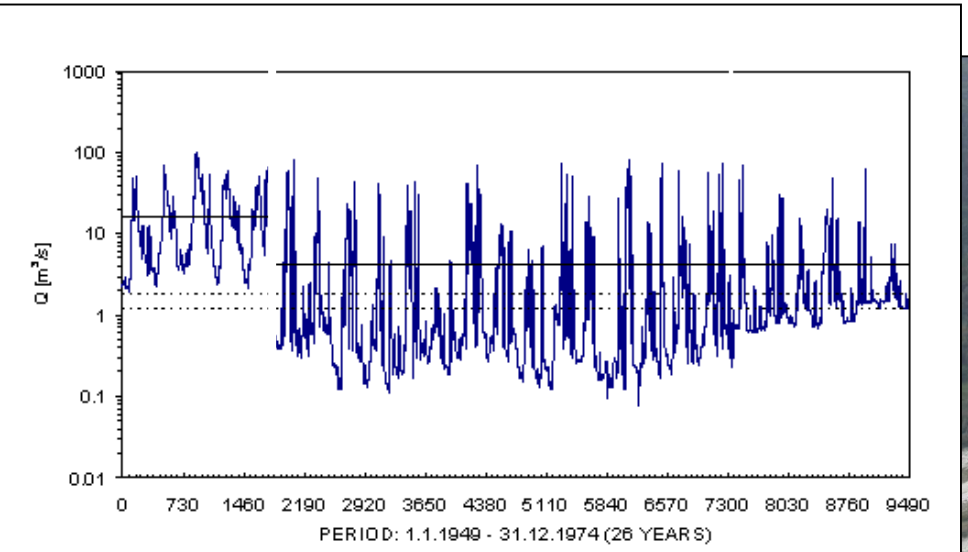
Model had a one day time step

Sensitivity very large for very small timedelay. Global methods like Sobol' and Method of Morris identify TIMEDELAY as most important parameter. Their averaged results hide the problem

Example: 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 of how well AIC, AICc, BIC, KIC identify models with good predictive ability using cross validation

- Use **SA** and **calibration methods** (Foglia + 2013 WRR) GW

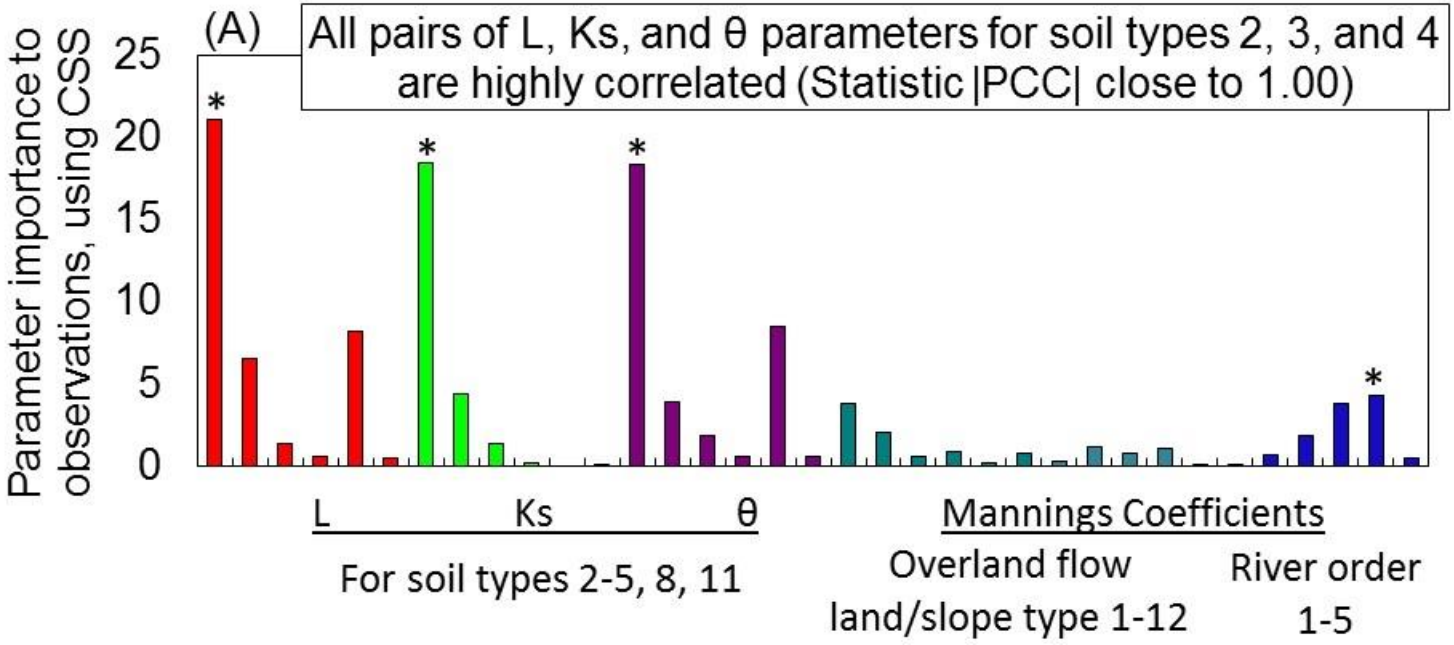
Methods largely based on work of U Minnesota statisticians Cook & Weisberg

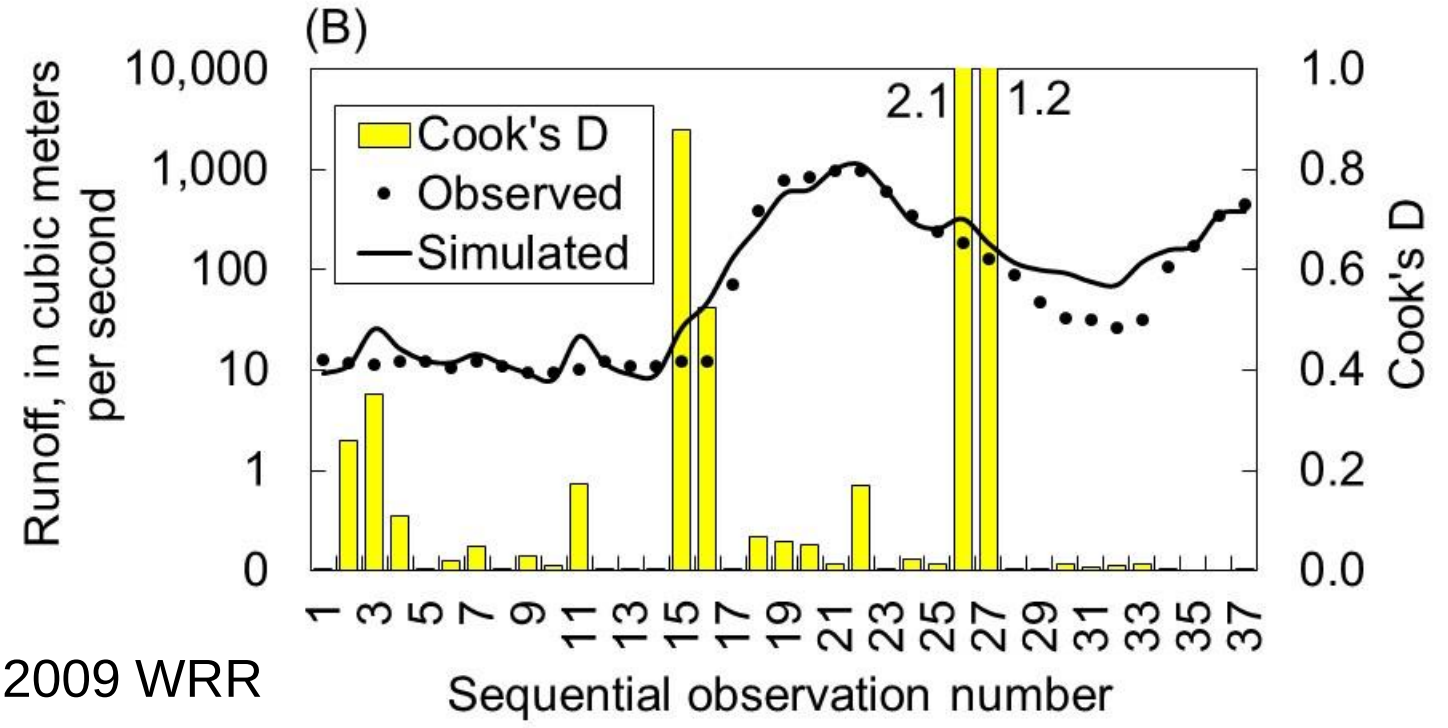
Parameter analysis:

Observations and prior information are sufficient to estimate 3 parameters

Observation analysis:

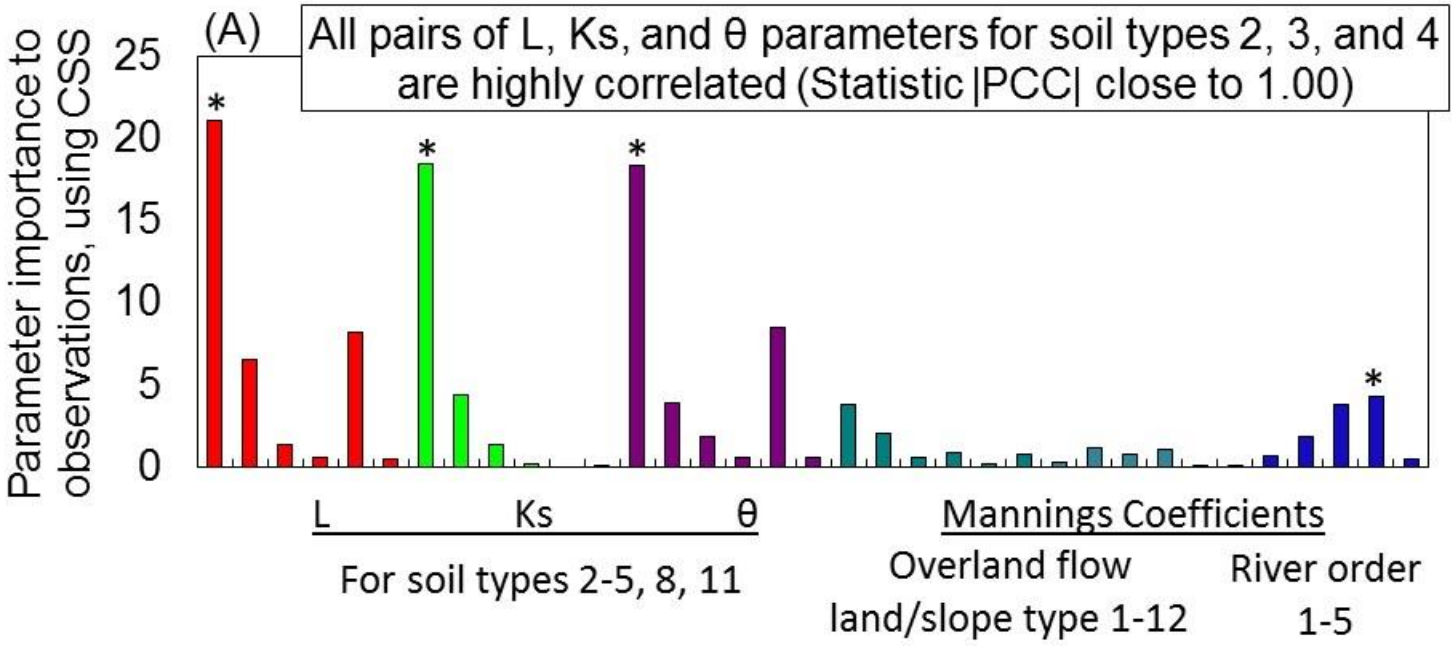
Low flows more important that expected. Resample low flows.

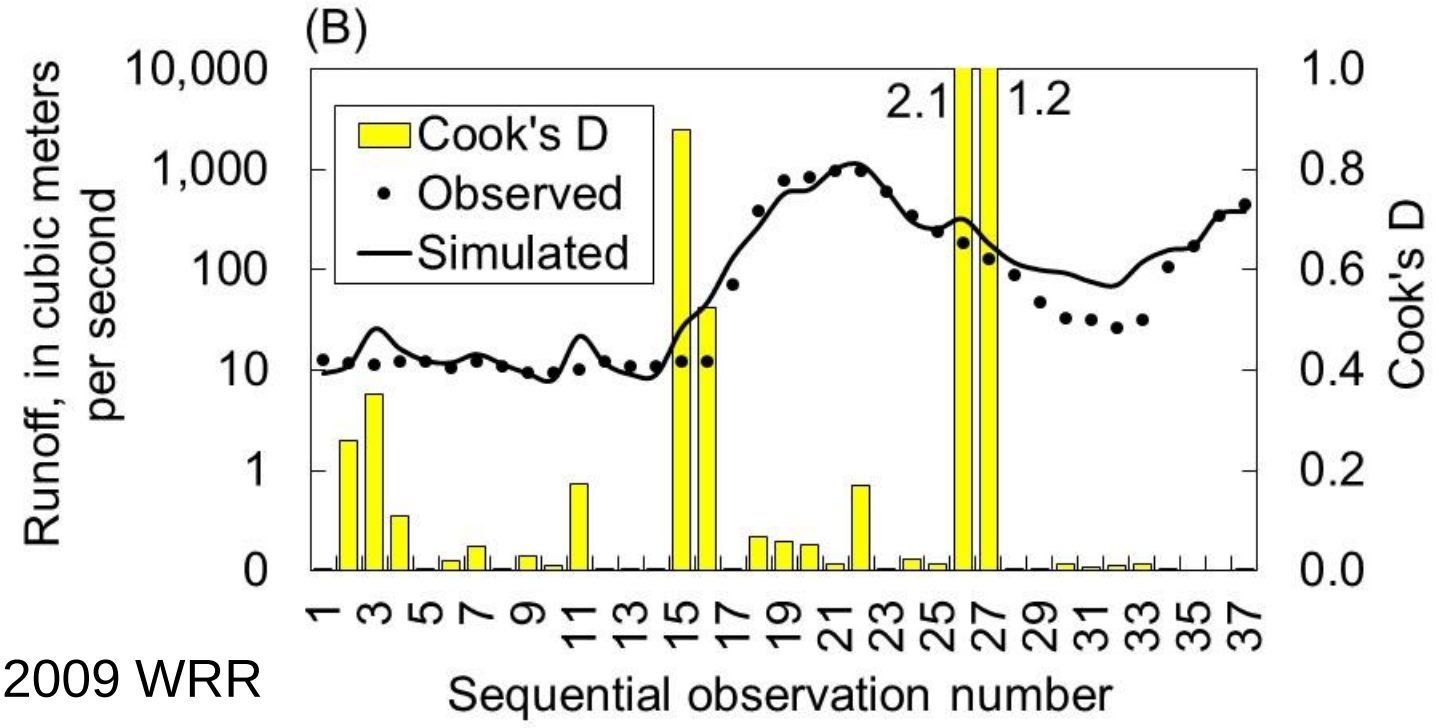




Foglia + 2009 WRR

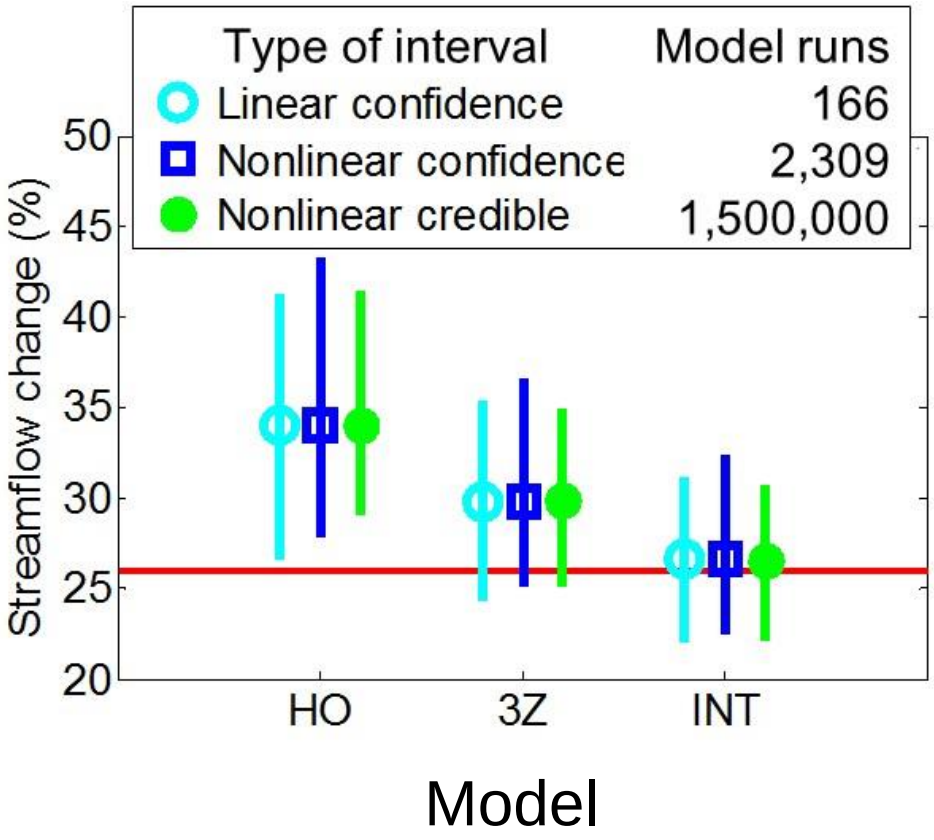
71 highly parallelizable model runs

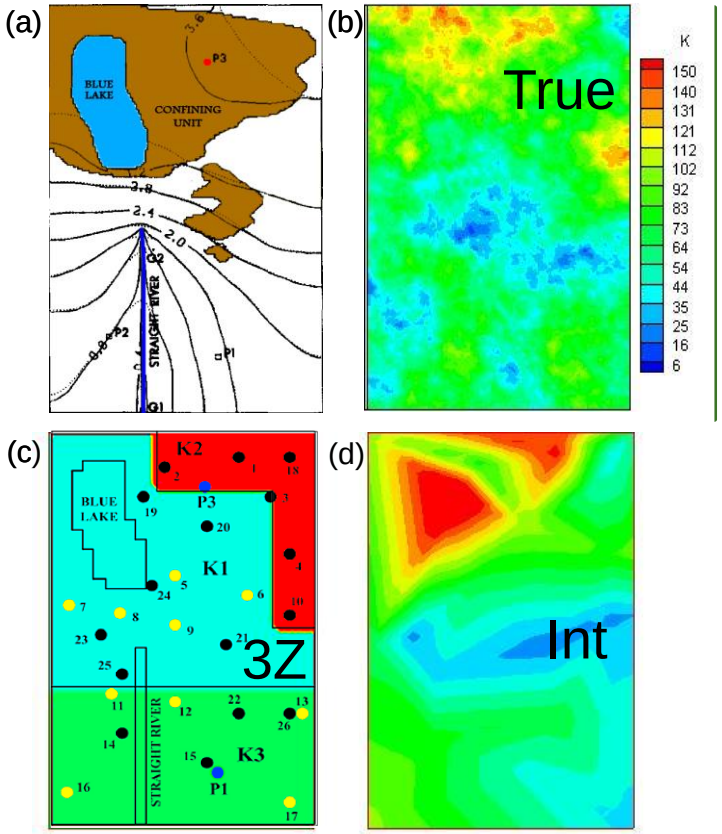




Foglia + 2009 WRR

Uncertainty Quantification using GW model

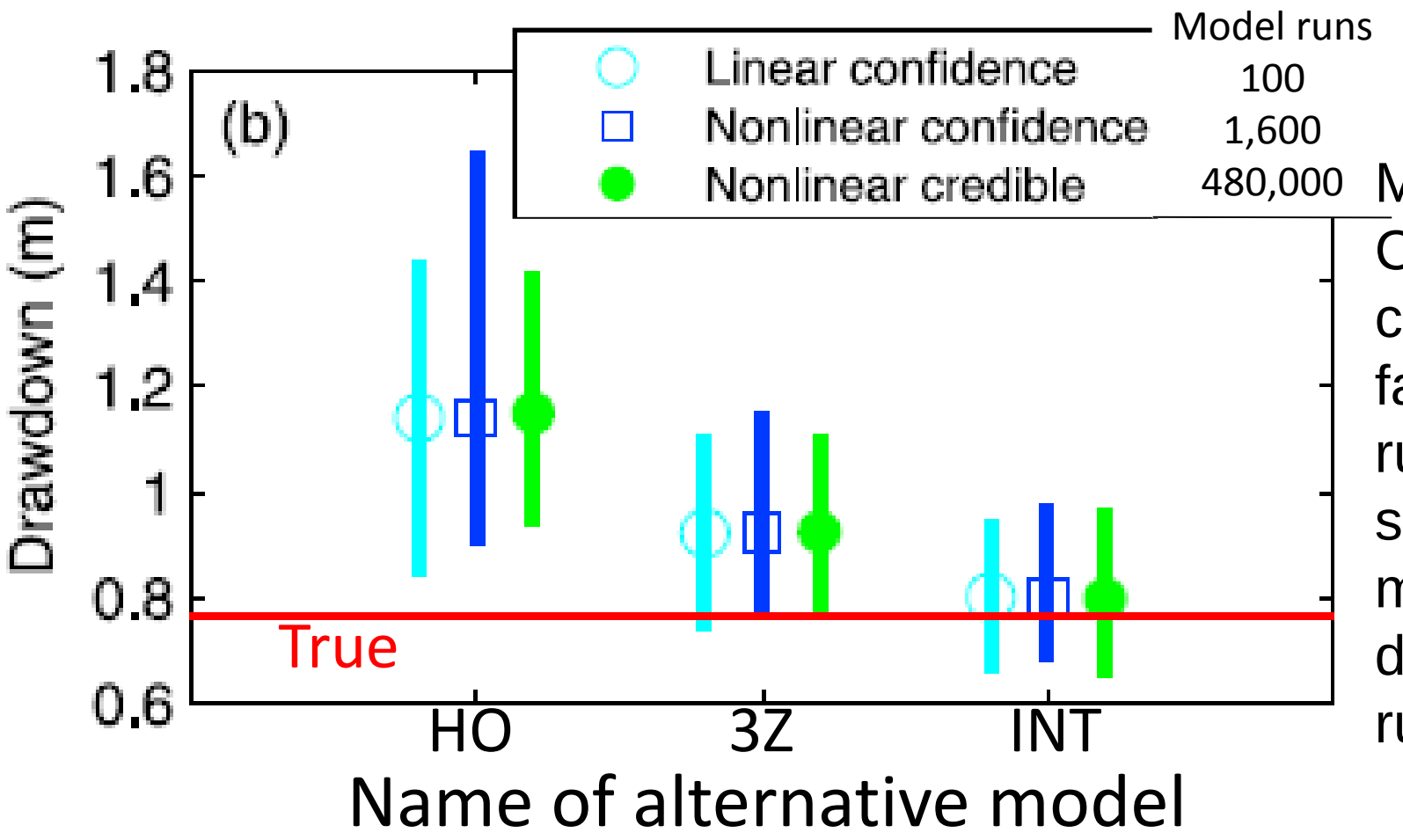




Lu + 2012 WRR

THE UNIVERSITY OF KANSAS

Uncertainty Quantification

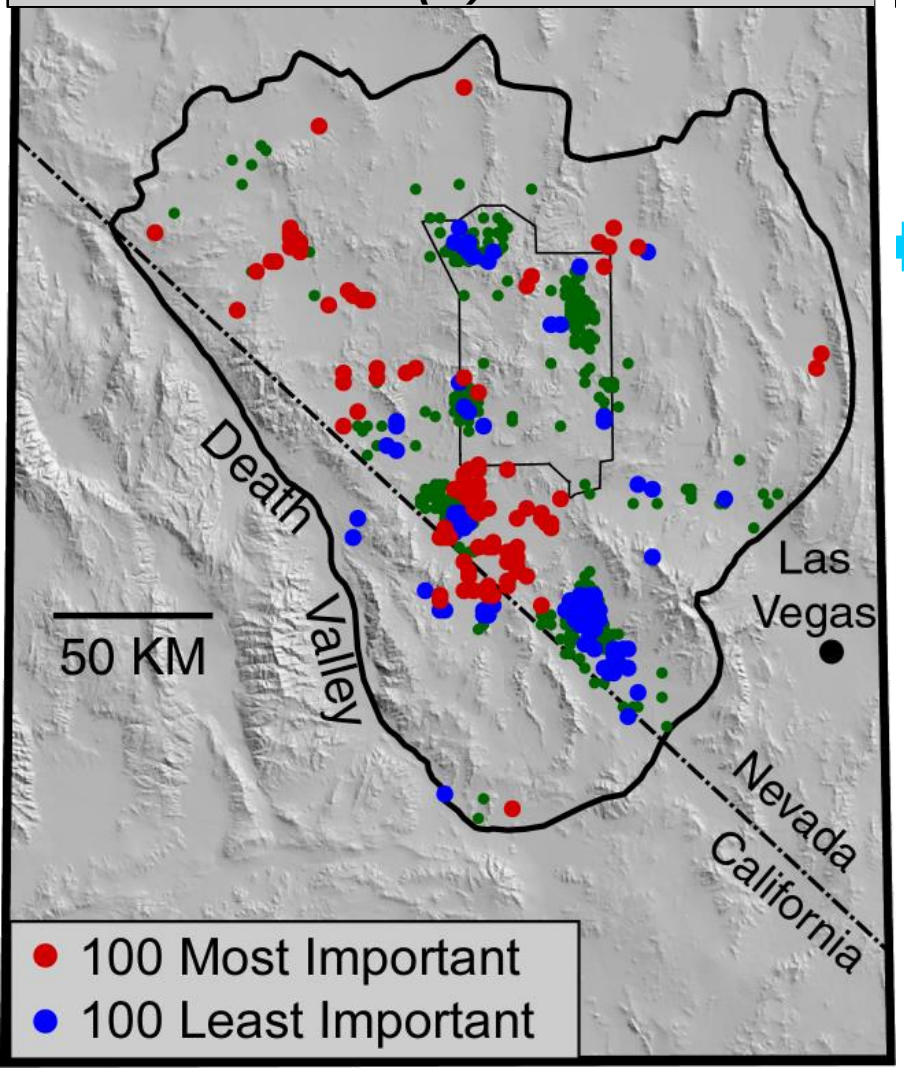


MCMC Can create a fast running surrogate model to do these runs

Value of Improved Information

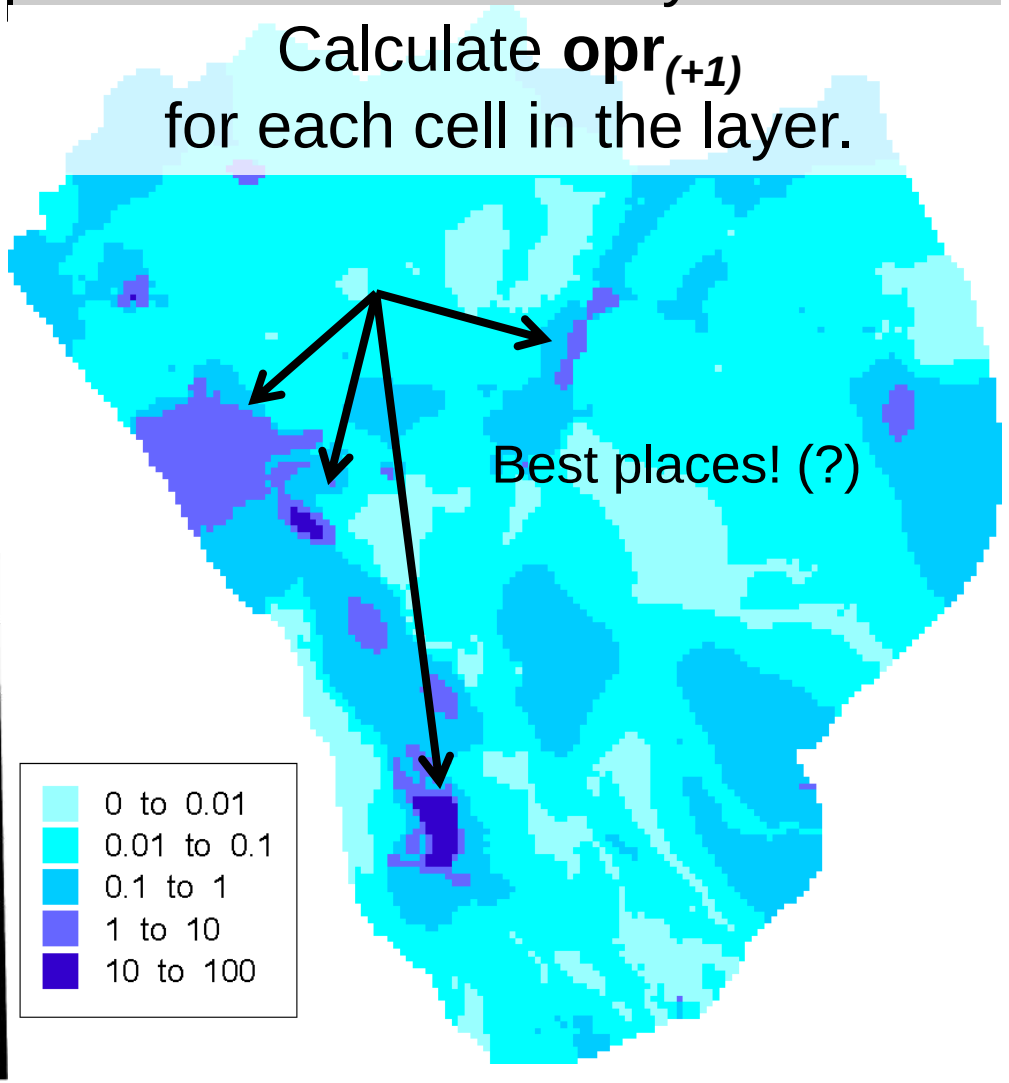
Value of information statistics,
Tiedeman + 2003, 2004 WRR
Program documentation: Tonkin et al 2007 USGS

Important existing obs?
Calculate **opr**$_{(-1)}$ for each obs



Consider one potential new observation in layer 1.

Calculate **opr**$_{(+1)}$
for each cell in the layer.



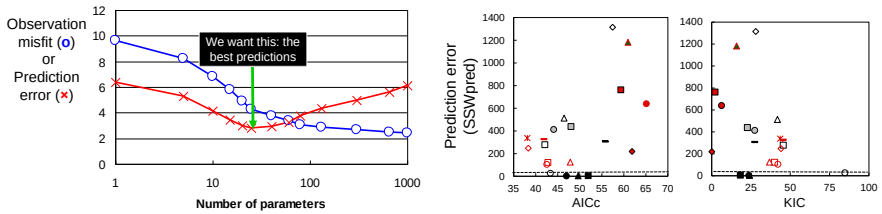
$OPR = 100x(s_{z(i)}-s_z)/s_z$ $\quad s_z=[(\partial z/\partial \mathbf{b})^T[s^2(\mathbf{X}^T\omega\mathbf{X})^{-1}]\ (\partial z/\partial \mathbf{b})]^{1/2}$

Programs developed to minimize numerical daemons – next generation models

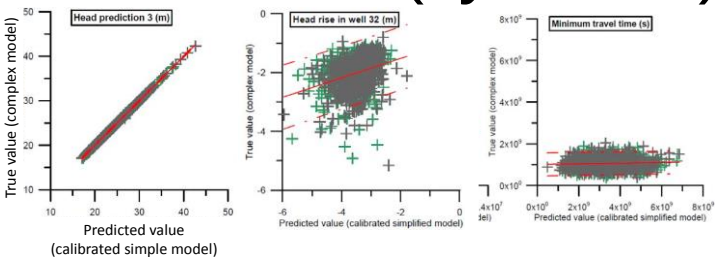
- MODFLOW-OWHM (Hanson+ 2014 USGS)
  - Integrated GW-SW
  - Extensive land and water use support, including for
    - agriculture
    - diversions
    - demand-driven, supply-limited use of gw and sw
- Rainfall-runoff
  - TOPKAPI (Ciarapica & Todini 2002 Hyd Proc)
  - SUMMA (Clark + 2015a,b WRR; NCAR)

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)

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?