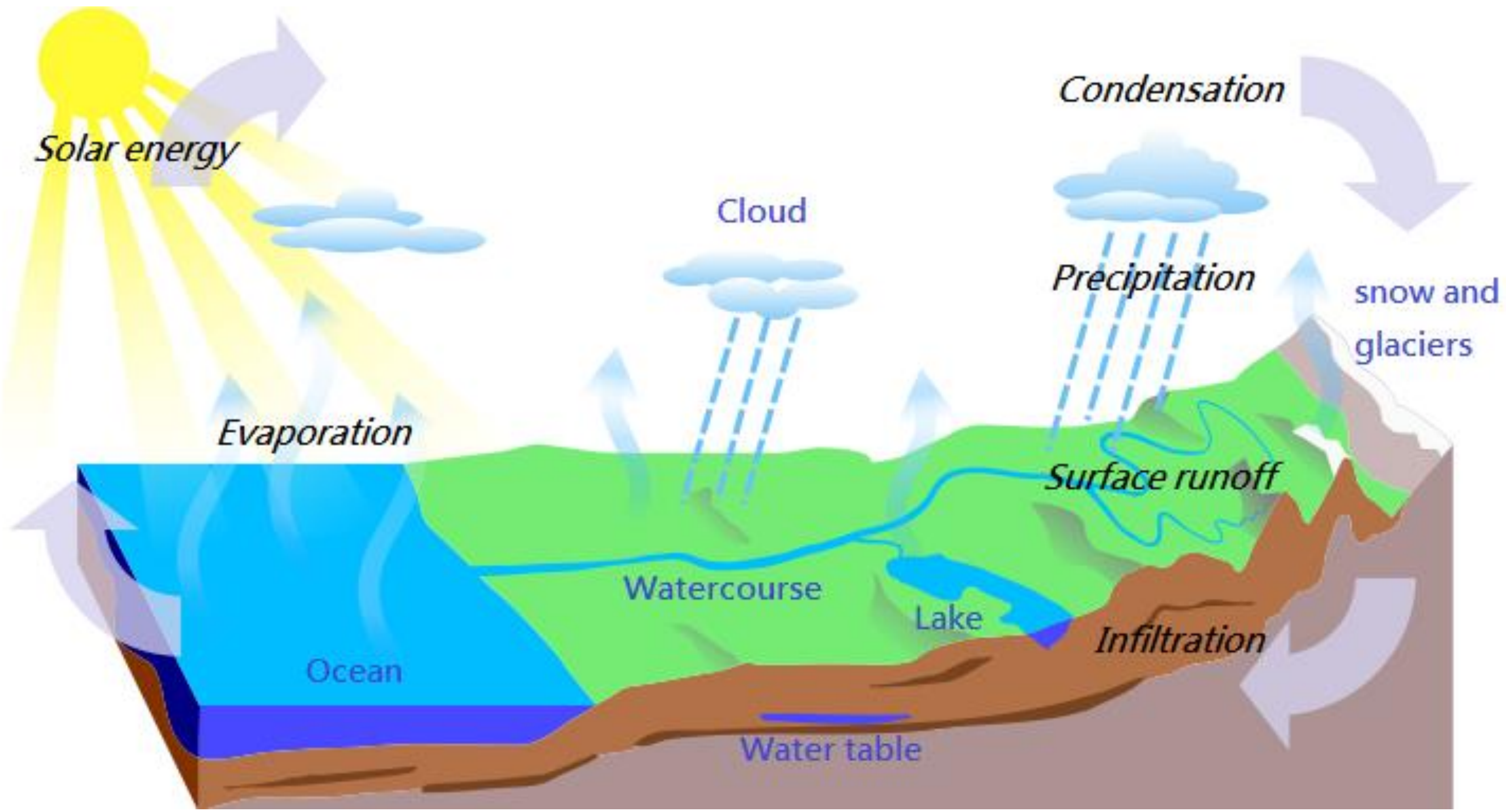
Hydrological modelling at the catchment scale:

Trusty Friend or Devious Foe?

Dmitri Kavetski, Martyn Clark, Mary Hill, Fabrizio Fenicia, Mark Thyer, Ben Renard and George Kuczera

Hydrological modeling and water cycle dynamics

- Routing of precipitation into streamflow, infiltration, evaporation
- Hillslope scale → catchment scale → continental scale
- Applications range from scientific to operational / management



GNU Copyright: Wikimedia: J de Bild and A. Chris

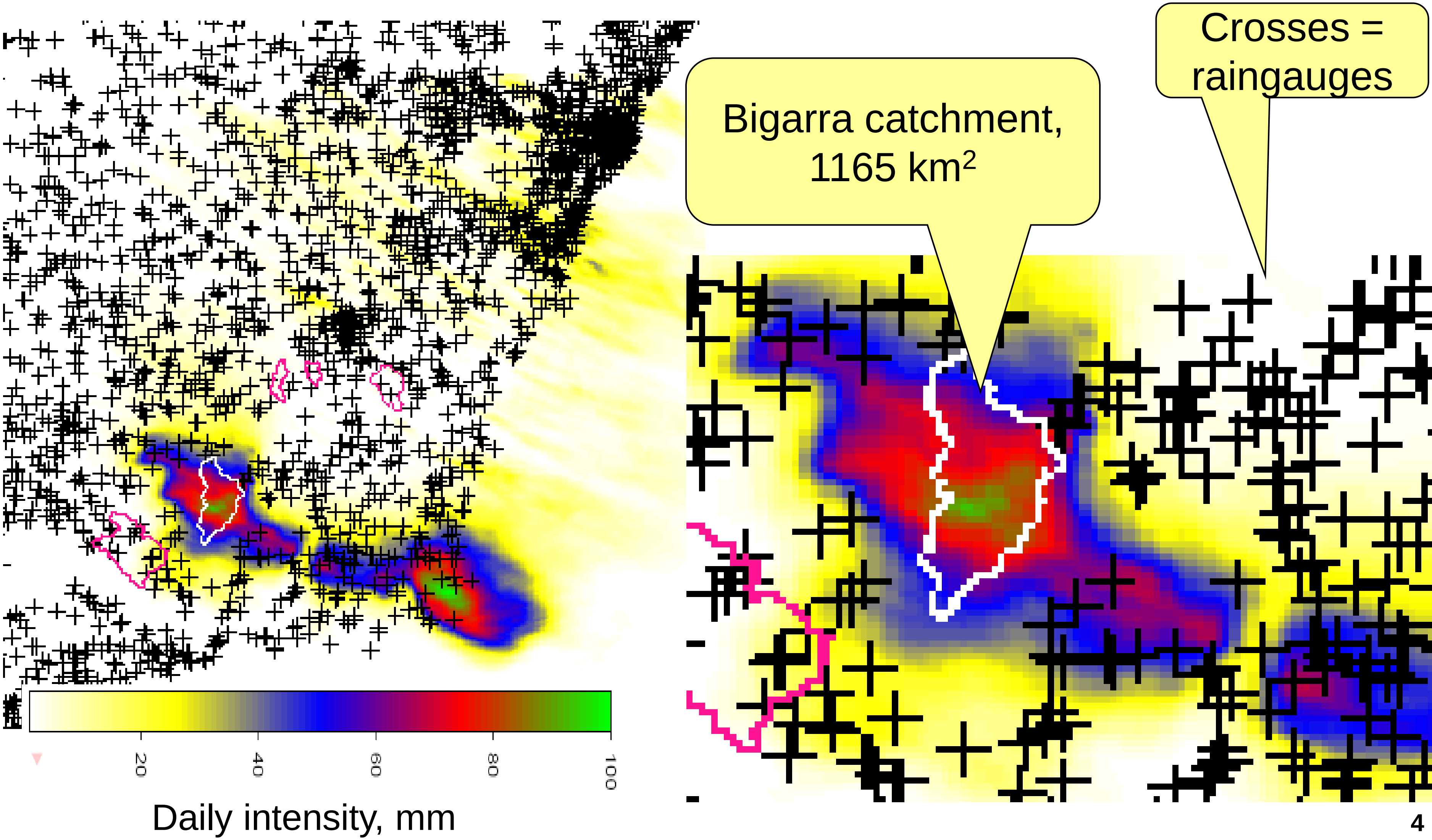
Hydrological prediction over a range of scales

- From "real time" (during floods) to seasonal (water management)
- In general, predictions / forecasts should be:
  - **Precise**
    - ±10% precision is better than ±50% precision
    - For this, we need good representation of catchment processes
      - » and need to extract maximum information from known data
  - **Reliable**
    - If we estimate ±10% errors, but routinely get 50% actual errors, we have a problem: misleading forecasts can undermine planning
    - Therefore, we need <u>reliable</u> uncertainty estimates
  - **Practical**
    - What good is a forecast if it takes forever to calculate or update?

Regardless of the context, hydrological prediction is challenged by several major sources of predictive uncertainty

Data uncertainty: A sobering thought …

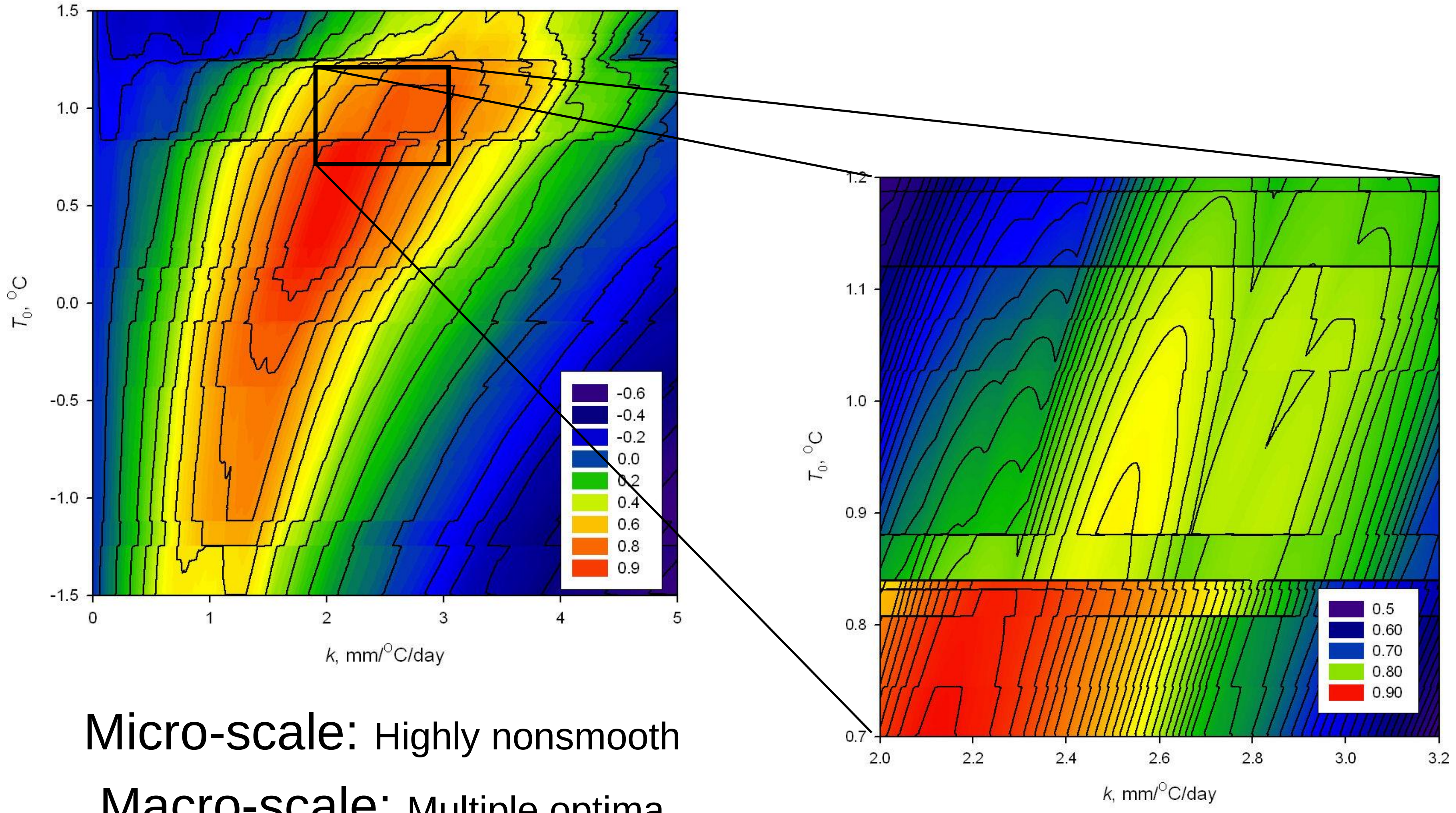
Radar rainfield image, 21 Nov 2009 AD, East coast of Australia, 512x512 km grid



A hydrological model for everywhere?

LITTLE SUCCESS IN BUILDING HYDROLOGICAL MODELS THAT WORK WELL “EVERYWHERE”, AT LEAST GIVEN CURRENT DATA

Nonlinear models, complex objective functions



Micro-scale: Highly nonsmooth

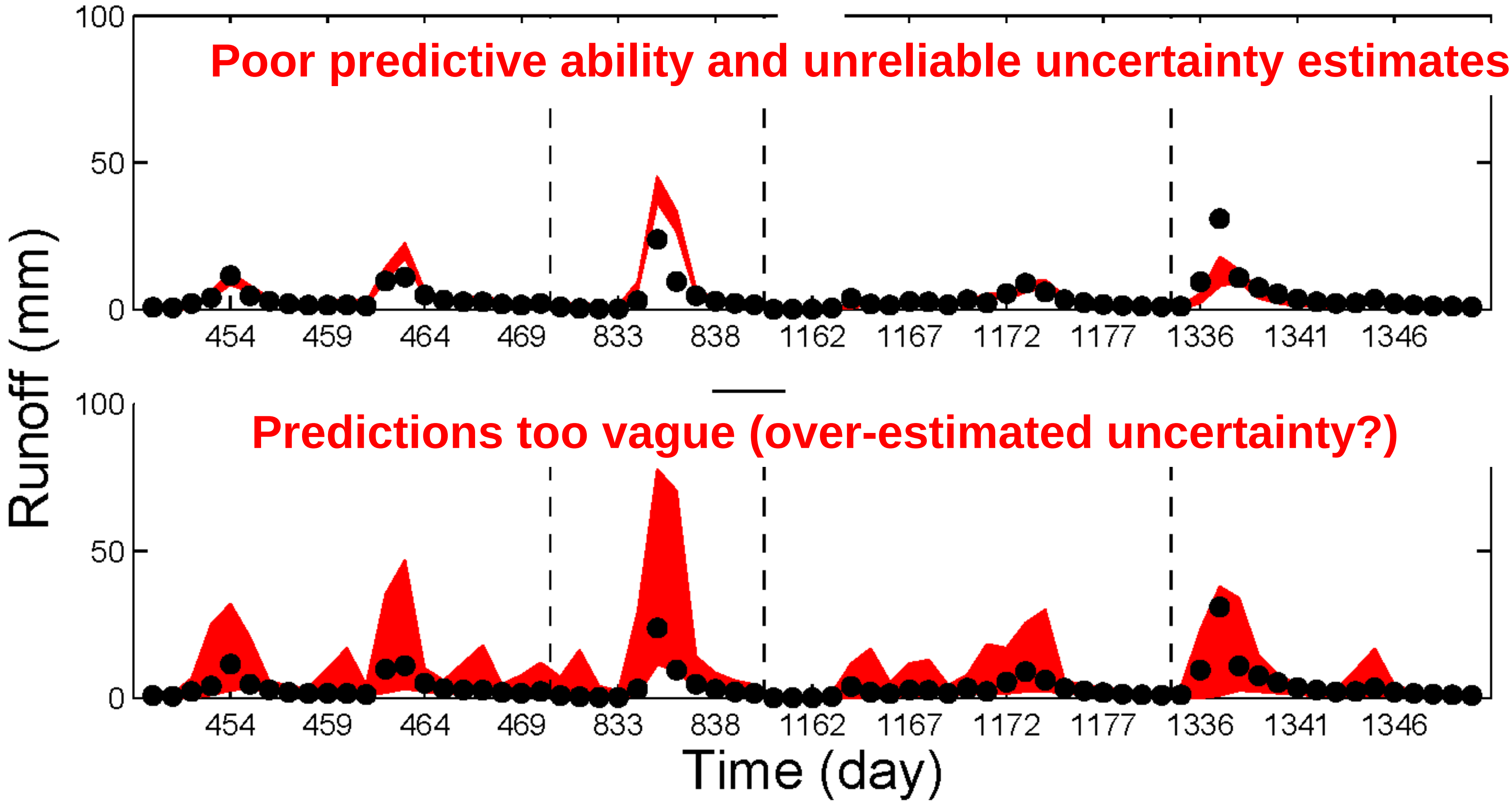
Macro-scale: Multiple optima

Duan et al, WRR1992; Kavetski and Kuczera, WRR2007

Hydrological "monsters" in the early XXI century

- Data monsters: Environmental data uncertainty
  - Significant errors in forcing-response data (eg, rainfall-runoff)
  - Non-Gaussian structure, time-dependencies, etc
- Physical monsters: Limited understanding of environmental dynamics
  - A real catchment vs a model representation?
- Mathematical monsters: Model nonlinearities
  - Require numerical approximations: more "physics" $\rightarrow$ more "CPU"
  - Require more complex optimization and statistical techniques

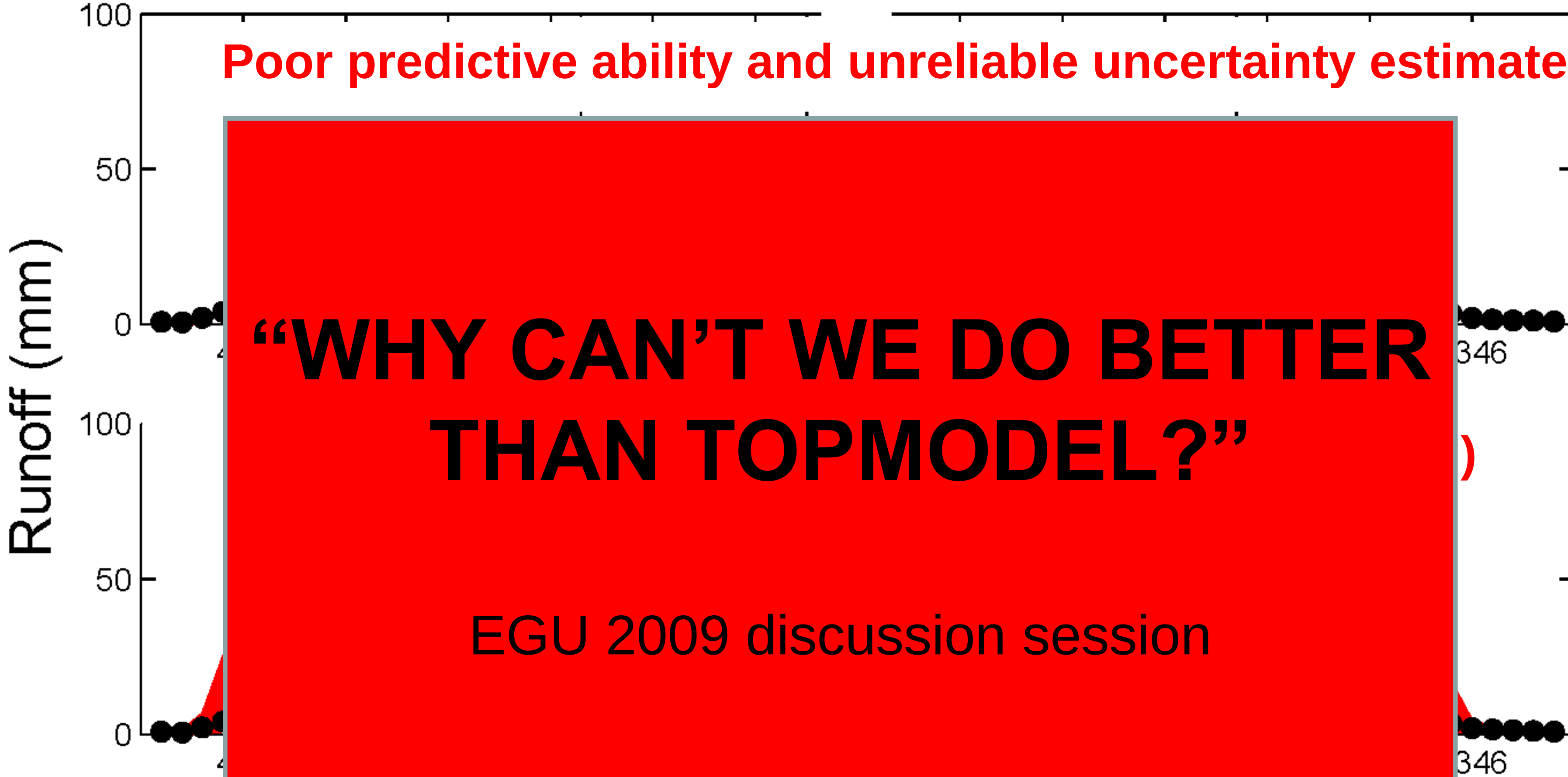
What have these monsters done to our models …



Many such exhibits at the *Monsters of Hydrology* workshop (Paris, 2008) and elsewhere in the published hydrological literature …

What have these monsters done to our models …

Poor predictive ability and unreliable uncertainty estimates



“WHY CAN’T WE DO BETTER THAN TOPMODEL?”

EGU 2009 discussion session

Many such exhibits at the *Monsters of Hydrology* workshop (Paris, 2008) and elsewhere in the published hydrological literature …

Hydrological "monsters" in the early XXI century

- Data monsters: Environmental data uncertainty
  - Significant errors in forcing-response data (eg, rainfall-runoff)
  - Non-Gaussian structure, time-dependencies, etc
- Physical monsters: Limited understanding of dynamics
  - A real catchment vs a highly simplified bucket model?
- Mathematical monsters: Model nonlinearities
  - Require numerical approximations: more "physics" $\rightarrow$ more "CPU"
  - Require more complex optimization and statistical techniques

What are some of the ways the community has approached these challenges?

How have we dealt with them so far …

- Data and Model Errors versus Calibration. Round *N*++
  - Many, many different objective functions
  - Multi-objective calibration
  - Global Optimization methods
  - Increasingly complex MCMC algorithms

- Entire new paradigms (eg, GLUE, MOCOM, “ABC”)
  - If we don’t want to construct a “formal” likelihood function, can we just sample from, for example, the Nash-Sutcliffe “distribution”?
  - Considerable debates in the last 20 years. Eg, Mantovan and Todini (2006): “incoherence of GLUE”; Freer et al (2006) “Just why would a modeller choose to be incoherent?”

- Computational brute force?
  - Multi-CPU clusters, code optimization, etc

How have we dealt with them so far …

- Data and Model Errors versus Calibration. Round *N*++
  - Many, many different objective functions
  - Mu
  - Gl
  - In
- Entir
  - If w … we jus … ?
  - Co … To … why wo
- Computational brute force?
  - Multi-CPU clusters, code optimization, etc

CALIBRATION FATIGUE?

Models dancing like “mathematical marionettes” to the tune of calibration schemes (Kirchner, 2006)

Outline of Presentation

1. Some current challenges in hydrologic modeling
   - Data and structural errors
   - Some historical ways we've dealt with these challenges
2. Eliminating unnecessary artefacts
   - Robust numerical formulation of hydrological models
   - Impact on hydrological calibration
3. Model development and hypothesis-testing
   - Too many models with too many differences?
   - Towards more systematic model comparison
4. Conclusions and a view to the future

Numerical solution/implementation aspects

- Analytical solution of water balances seldom possible
- Numerical "approximations" are (often tacitly) employed
  - Explicit Euler scheme (conceptual hydrology)

$$S_{n+1}^{(EE)} = S_n + inflow_n - outflow(S_n)$$

Flux at <u>start</u> of the time step

  - Implicit Euler scheme (engineering/groundwater)

$$S_{n+1}^{(IE)} = S_n + inflow_{n+1} - outflow(S_{n+1})$$

Flux at <u>end</u> of the time step

  - Adaptive numerical solutions (applied maths/engineering)

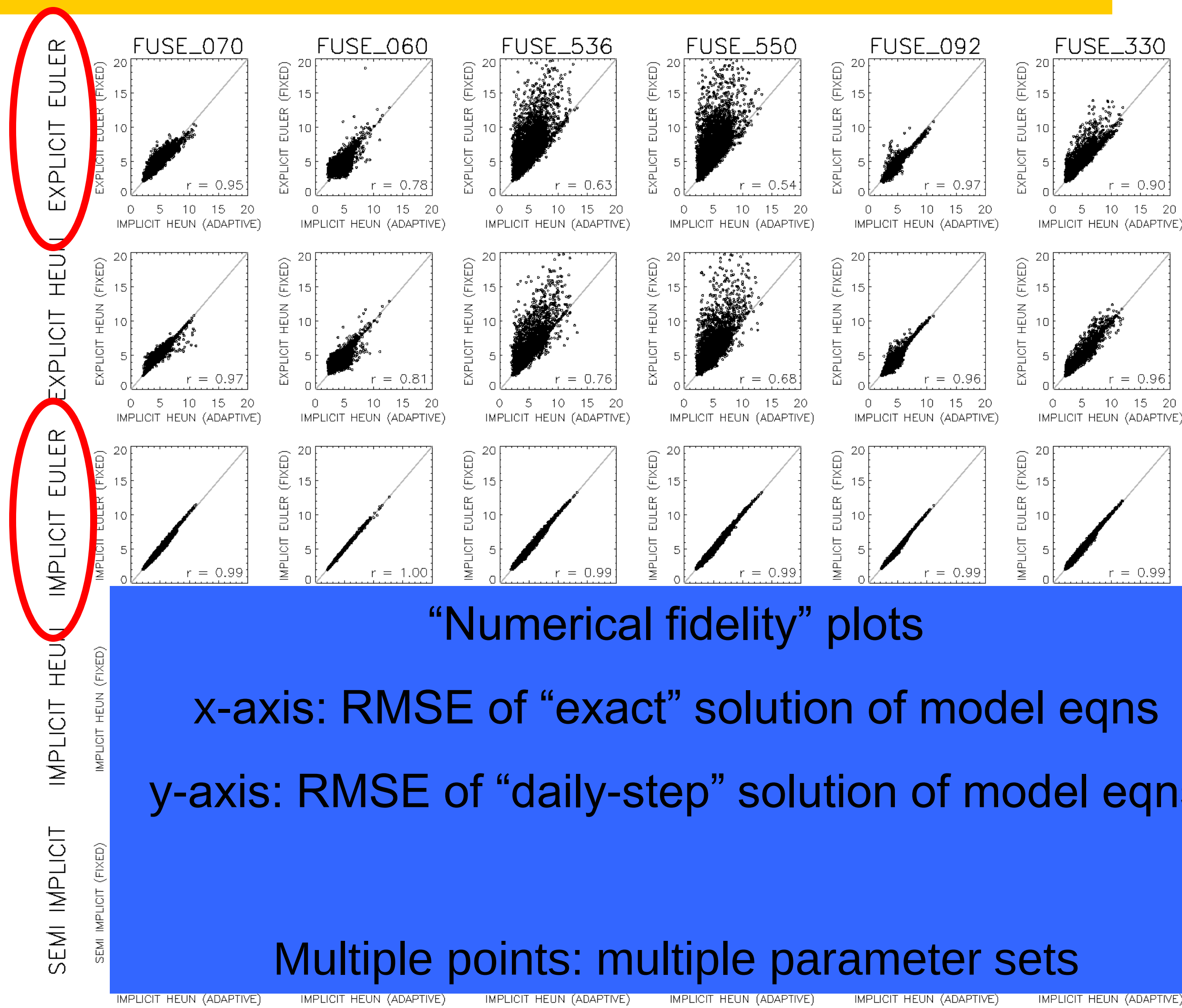
Though seemingly mundane, the numerical approximation technique has a profound impact on model behavior..

… yes, even when data is inexact and model is poor!

Numerical fidelity … Six models, 10,000 parameter sets

Explicit Euler scheme

Implicit Euler scheme



"Numerical fidelity" plots

x-axis: RMSE of "exact" solution of model eqns

y-axis: RMSE of "daily-step" solution of model eqns

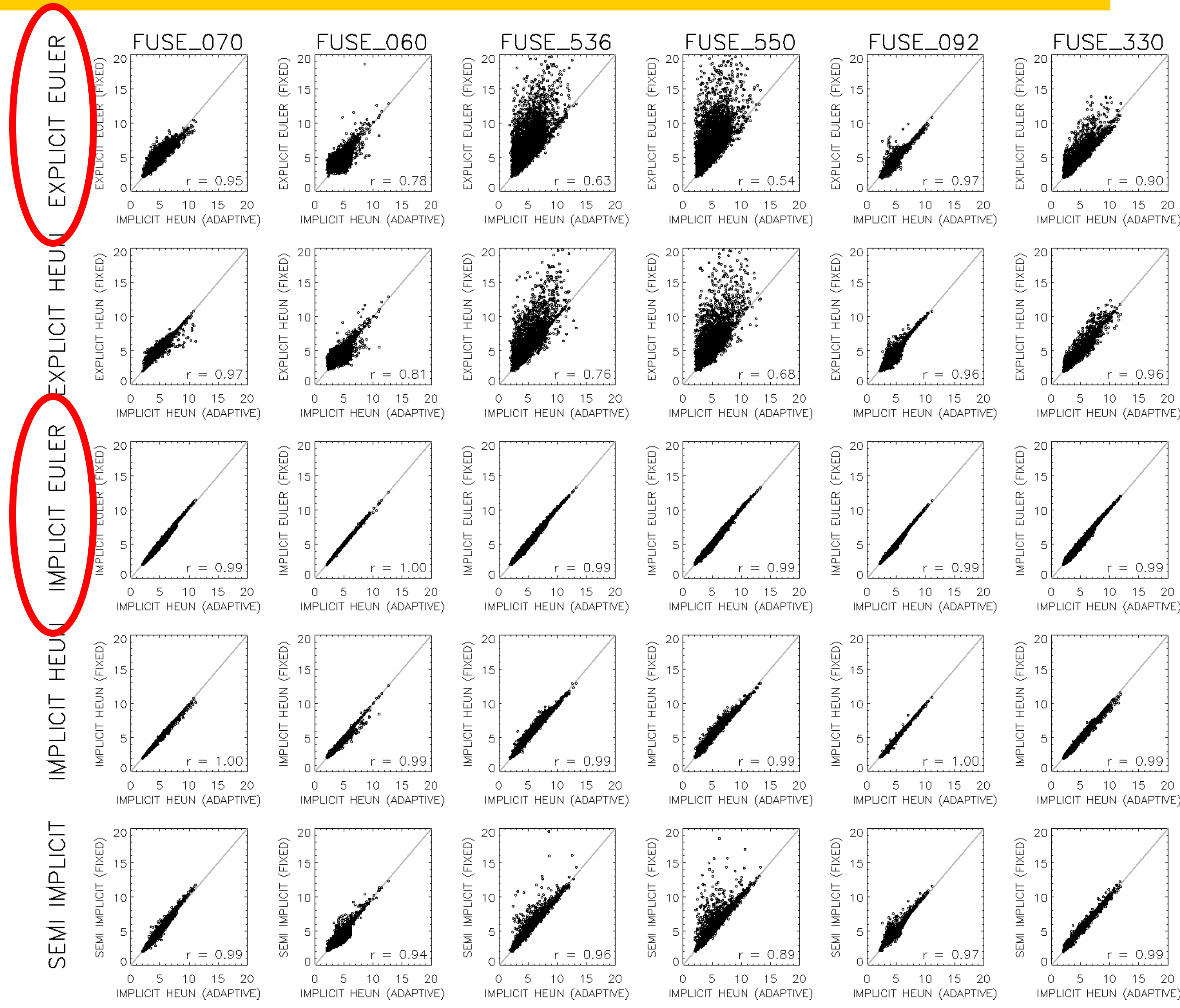
Multiple points: multiple parameter sets

Clark, M.P. and D. Kavetski (WRR 2010) The numerical dæmons of conceptual hydrological models – Parts 1&2

Numerical fidelity … Six models, 10,000 parameter sets

Explicit Euler: numerically fragile

Implicit Euler numerically reliable

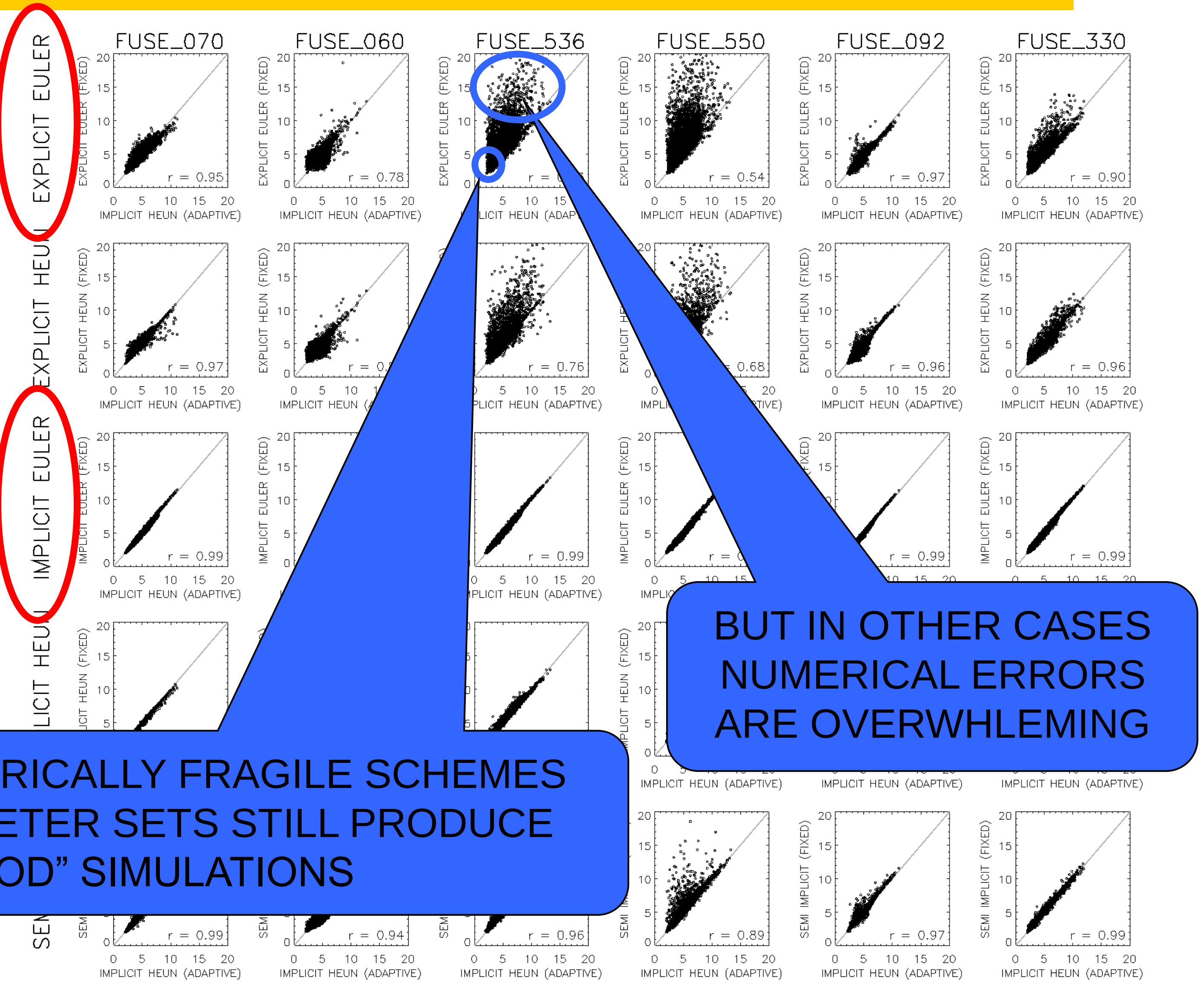


Clark, M.P. and D. Kavetski (WRR 2010) The numerical dæmons of conceptual hydrological models – Parts 1&2

Numerical fidelity … Six models, 10,000 parameter sets

Explicit Euler: numerically fragile

Implicit Euler numerically reliable



Clark, M.P. and D. Kavetski (WRR 2010) The numerical dæmons of conceptual hydrological models – Parts 1&2

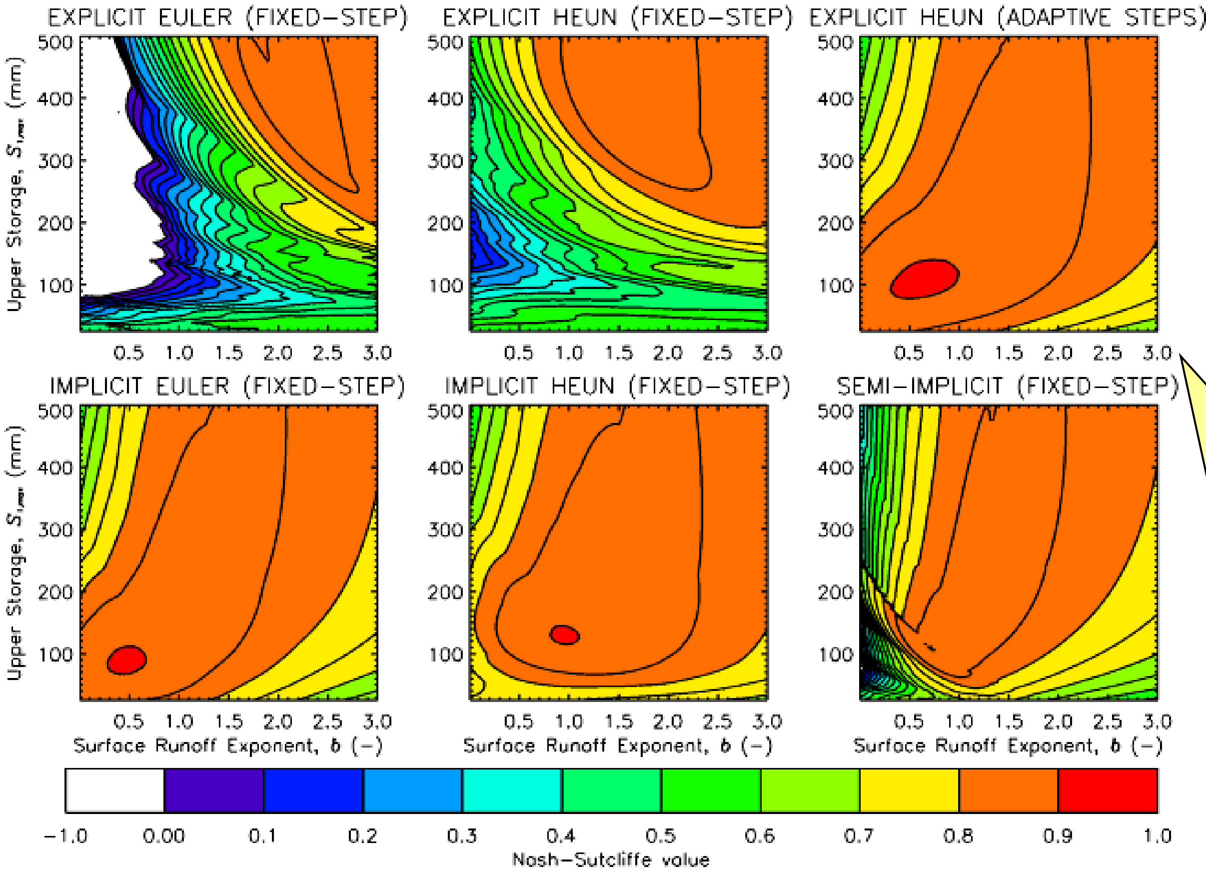
Impact of time stepping scheme (same model)

** Can represent the dominant source of error (dwarfs data + structure)

** Can massively distort parametric dependencies of the model

** Degradation of performance in predictive ("validation") mode

** See "dæmonic papers" (Clark & Kavetski, WRR2010)



Same model assumptions,

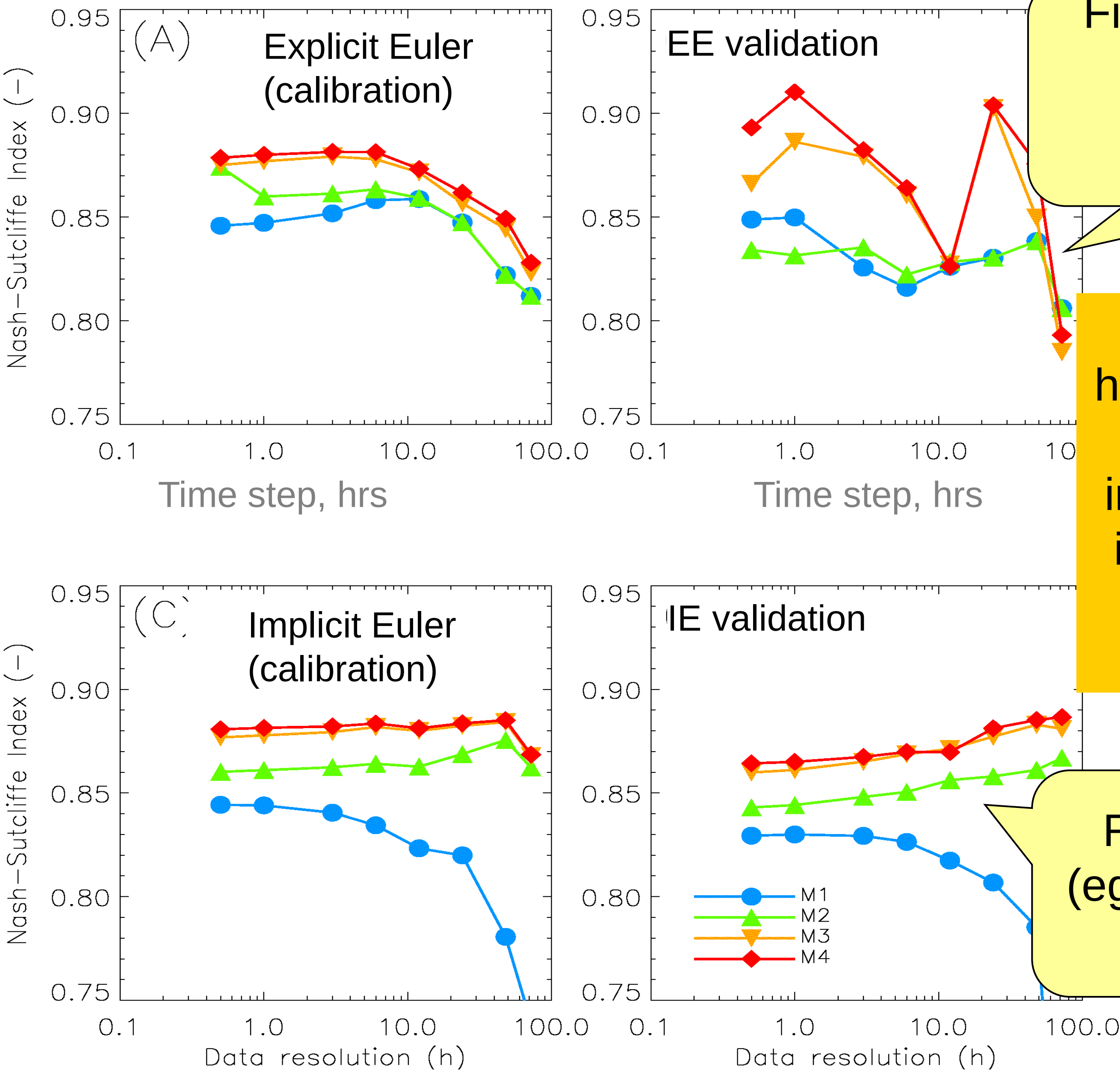
Same uncertain data,

Same objective function

but …

Different time-stepping scheme

Hypothesis-testing at a range of time scales

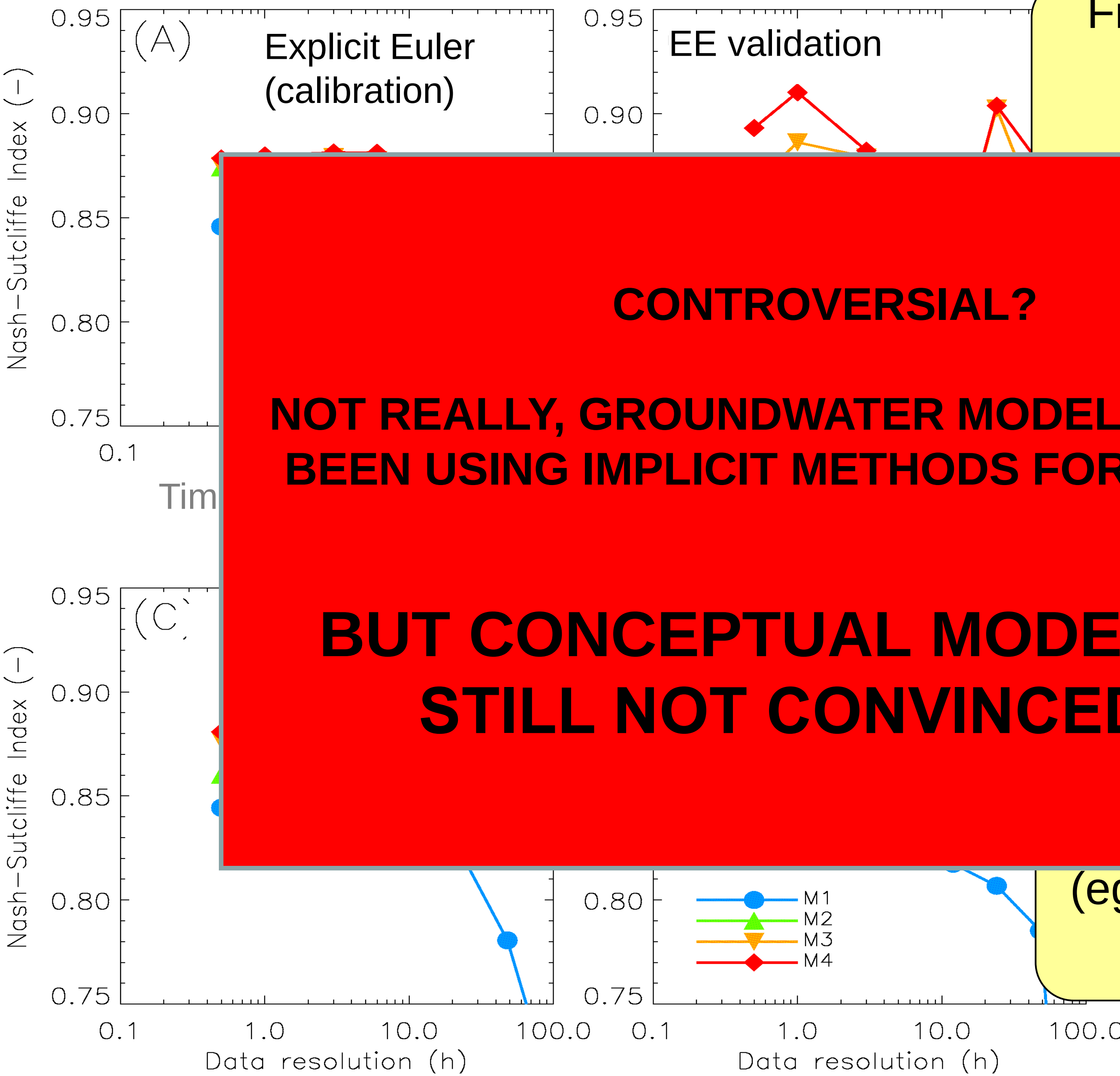


Fixed-step explicit time stepping behaves inconsistently in predictive mode

Q: How would a hydrologist unaware of numerical issues interpret these results in the context of time scale behavior and model selection?

Robust time stepping (eg, fixed-step implicit) is enviably fidelious

Hypothesis-testing at a range of time scales



Fixed-step explicit time stepping behaves inconsistently in … de

… d a … /are of … ues … esults … f time … and … on?

CONTROVERSIAL?

NOT REALLY, GROUNDWATER MODELLERS HAVE BEEN USING IMPLICIT METHODS FOR DECADES

BUT CONCEPTUAL MODELLERS STILL NOT CONVINCED …

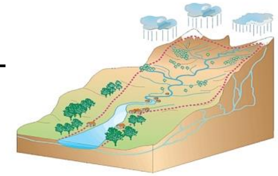
… oping (eg, fixed-step implicit) is enviably fidelious

Outline of Presentation

1. Some current challenges in hydrologic modeling
   - Data and structural errors
   - Some historical ways we've dealt with these challenges
2. Eliminating unnecessary artefacts
   - Robust numerical formulation of hydrological models
   - Impact on hydrological calibration
3. Model development and hypothesis-testing
   - Too many models with too many differences?
   - Towards more systematic model comparison
4. Conclusions and a view to the future

Errors in Environmental Modelling

Reality



True Input(s) → **True processes** → **True response(s)**

Input errors

e.g. rainfall sampling errors

Highly variable in time/space

Low gauge density

Structural errors

e.g. lumping processes

Ubiquitous but

Poorly understood

Output errors

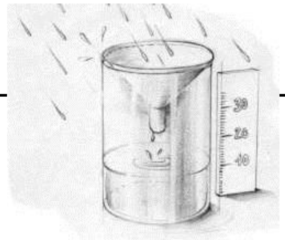
e.g. rating curve errors

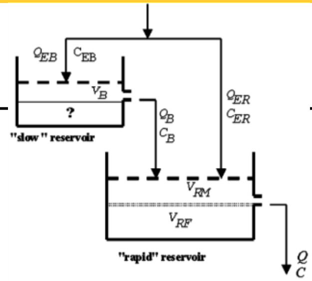
General accuracy ~10-20%

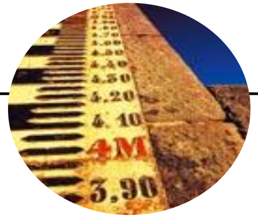
larger errors in larger floods?

Observed Input(s) → **Conceptualized processes** → **Observed response(s)**

Model



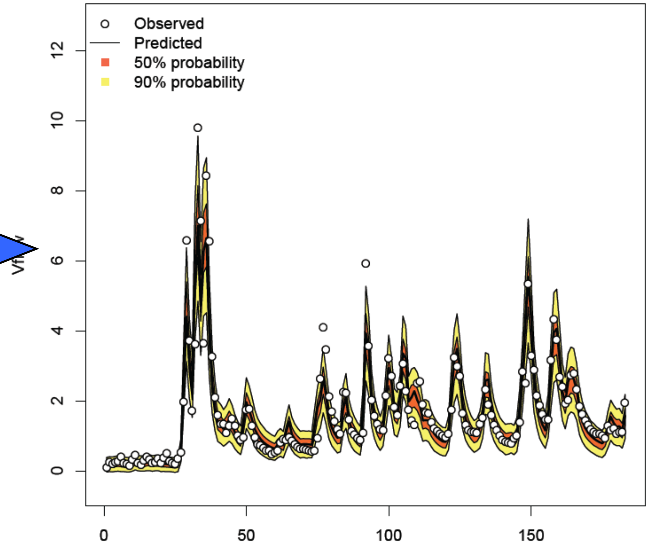


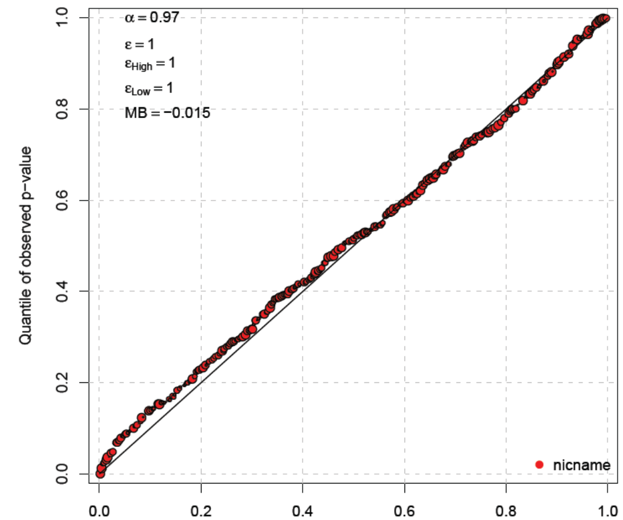


Aggregate vs Decompositional inference/prediction

1. Aggregate:

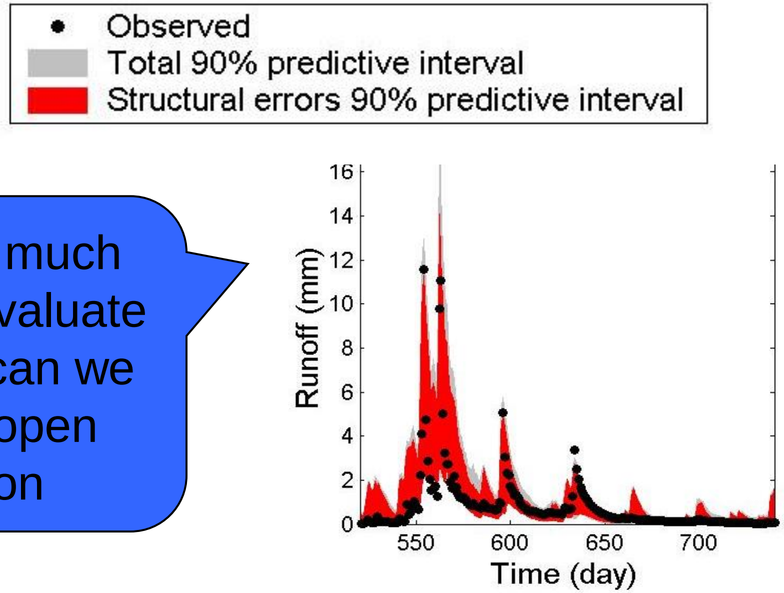
Reliability "relatively" easier to achieve and test – unless extrapolating!

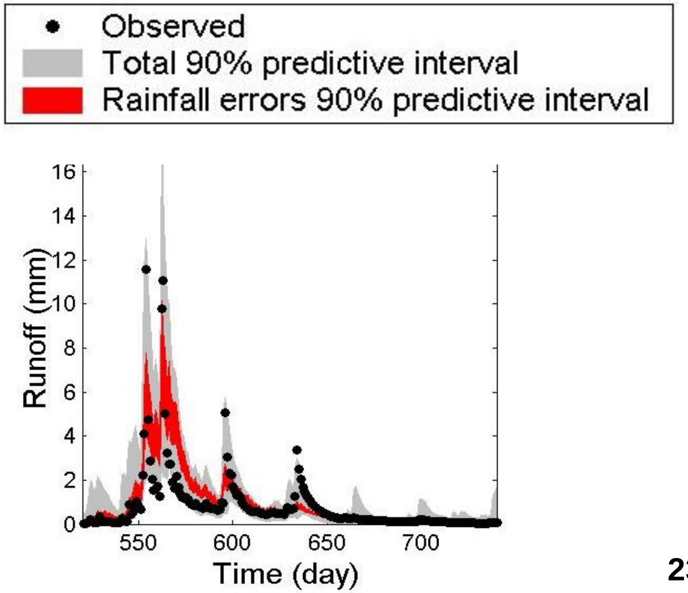




2. Decompositional: requires far more information

Reliability much harder to evaluate – how far can we go is an open question





Outline of Presentation

1. Some current challenges in hydrologic modeling
   - Data and structural errors
   - Some historical ways we've dealt with these challenges
2. Eliminating unnecessary artefacts
   - Robust numerical formulation of hydrological models
   - Impact on hydrological calibration
3. Model development and hypothesis-testing
   - Too many models with too many differences?
   - Towards more systematic model comparison
4. Conclusions and a view to the future

Hydrological and environmental modeling as a "decision-making" scientific process

- Some modeling decisions can be based on "well-understood" physics
  - Use Richards equation for the unsaturated soil zone
  - Explicitly simulate snow surface energy exchanges

- Other modeling decisions more ambiguous
  - Preferential / macropore flow – is it significant or even dominant, and, if so, how should it be represented?
  - How to characterize (unknown) bedrock topography/permeability

- Other modeling decisions are more pragmatic, based on the computer budget and other considerations
  - What is the best way to represent the spatial variability of snow depth across a hierarchy of scales?
  - Is a lumped model sufficient, or a distributed model required?

Hydrological and environmental modeling as a "decision-making" scientific process

- Some modeling decisions can be based on "well-understood" physics
  - U
  - E
- Othe
  - P ... ant,
    a
  - H ... ility
- Othe ... puter
  budg
  - W ... w
    d
  - Is a lumped model sufficient, or a distributed model required?

…CURRENTLY, LITTLE AGREEMENT REGARDING A "CORRECT" MODEL STRUCTURE

… "EVERY MODELLER HAS THEIR OWN MODEL"!

SIGNIFIES A "PROBLEM" FOR THE DISCIPLINE OF HYDROLOGY

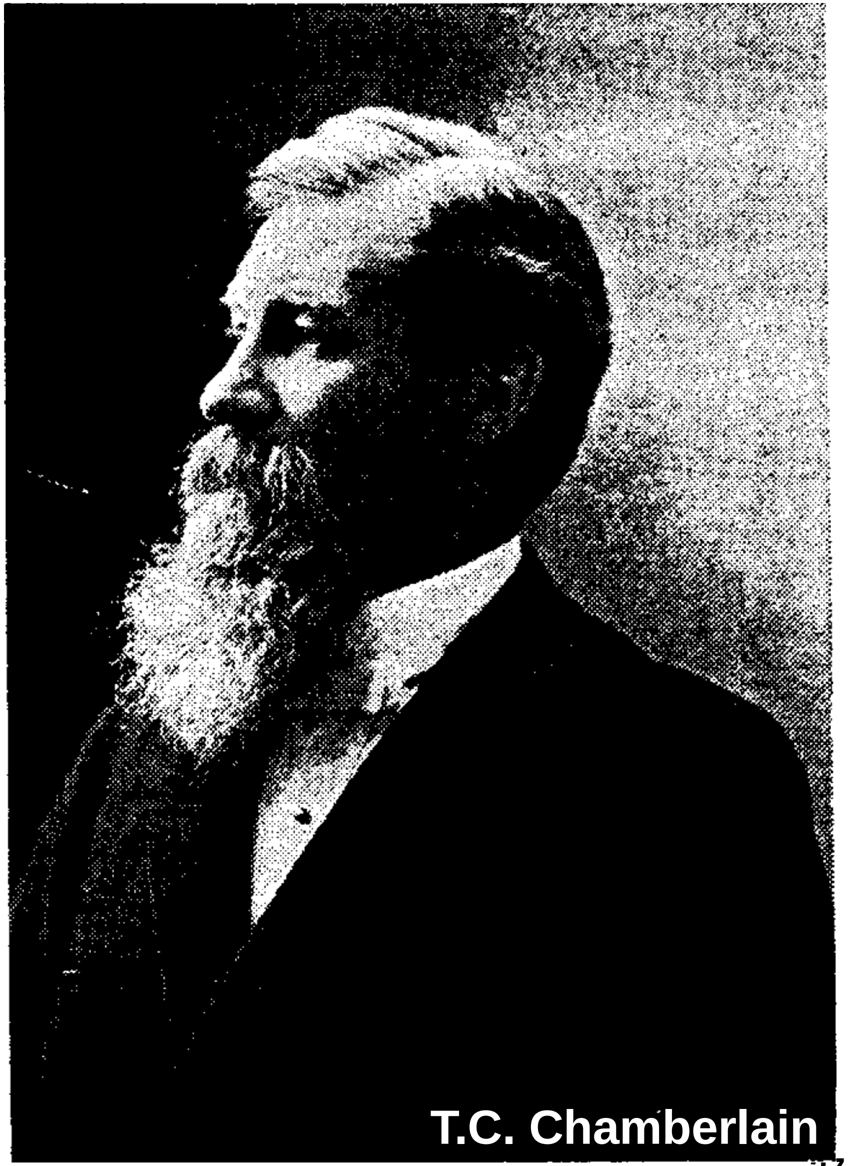
Pursuing the method of multiple working hypotheses in catchment-scale hydrological modelling

- Scientists often develop "parental affection" for their theories

- Chamberlin's method of multiple working hypotheses

"…the effort is to bring up into view every rational explanation of new phenomena… the investigator then becomes parent of a family of hypotheses: and, by his parental relation to all, he is forbidden to fasten his affections unduly upon any one"

Chamberlin (1890)



Wait, aren’t we doing that already?

- What about multi-model comparisons?
  - Eg, the MOPEX experiment: 12 catchments, 8-10 models
- What about multi-model ensemble methods?
  - Haven‘t we been already combining models in various statistically rigorous ways?

Wait, aren’t we doing that already?

- Sure, but ... the models in most of these experiments differed in many <u>uncontrolled</u> ways

- eg, GR4J versus Sacramento model:
  - Different quickflow representation,
  - Different groundwater representation
  - Different evaporation/transpiration representation
  - Different time stepping scheme

- Uncontrolled differences make it hard for hydrologists to understand model performance and how to improve it

- Plus … the original MOPEX experiment reported detailed model comparisons … for “anonymous models”!
  - The results/figures did not identify the models “by name”!

Wait, aren't we doing that already?

- Sure, but ... the models in most of these experiments differed in many <u>uncontrolled</u> ways

- eg, C
  - Di
  - Di
  - Di
  - Di

… and are our models really "different" from each other, or are they all wrong in the same way?

How would this affect model ensemble methods?

- Unc… to
  unde

- Plus … the original MOPEX experiment reported detailed model comparisons … for "anonymous models"!
  - The results/figures did not identify the models "by name"!

Towards multi-hypothesis frameworks

(Clark, Kavetski and Fenicia, WRR, 2011):

Accommodate different options regarding process selection and representation

- *Choice of state variables*
- *Choice of processes to include/exclude*
- *Choice of parameterizations for individual processes*

- For example, a possible state equation for the unsaturated zone is

$$\frac{\mathrm{d}S_1}{\mathrm{d}t} = p - e - q_d - q_{sx}$$

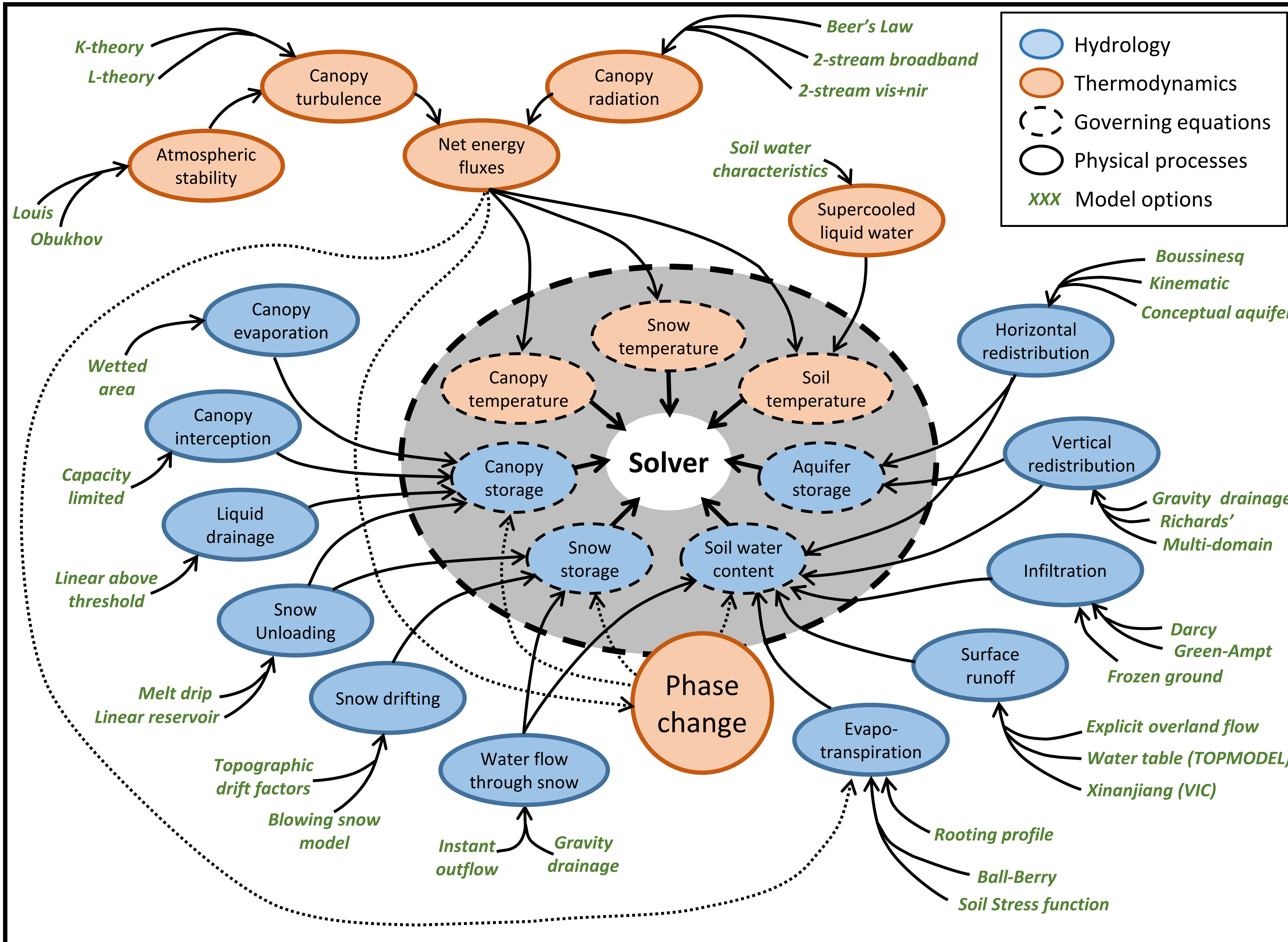
precipitation — *evaporation* — *vertical percolation* — *surface runoff*

many options

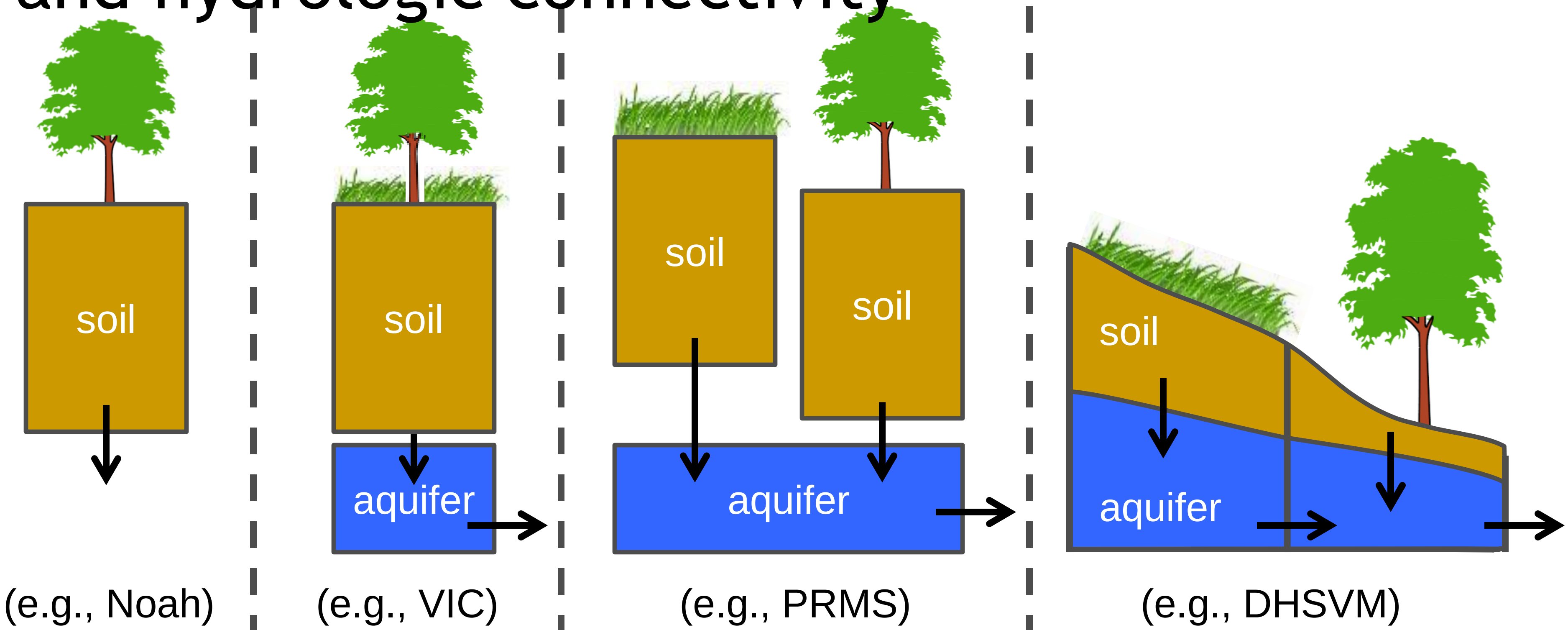
- Two popular representations of surface runoff:

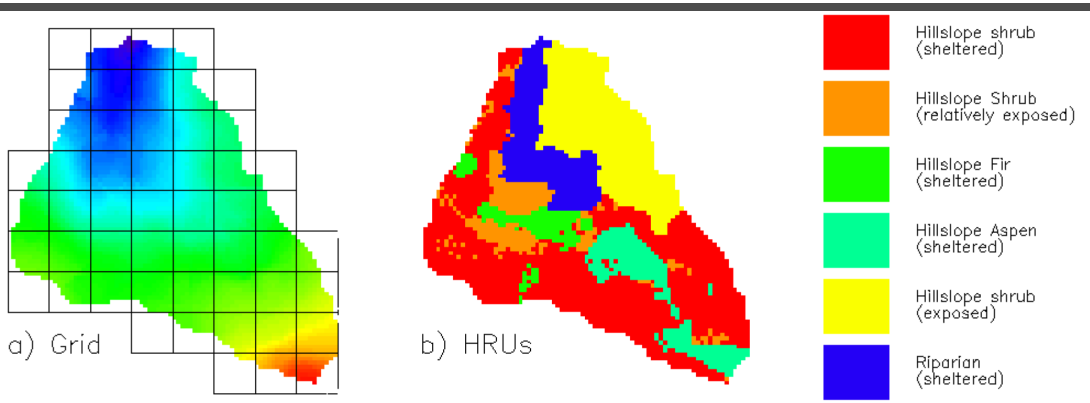
$$q_{sx} = \begin{cases} p\left[1-\left(1-\dfrac{S_1}{S_{1,\max}}\right)^b\right] & \text{. . . . . VIC parameterization} \\ p\int_{\zeta_{crit}}^{\infty} f(\zeta)\,d\zeta, \quad \zeta_{crit} = \lambda_n\left(\dfrac{S_2}{S_{2,\max}}\right)^{-1} & \text{. . . TOPMODEL parameterization} \end{cases}$$

“Horrendogram” of a modelling framework

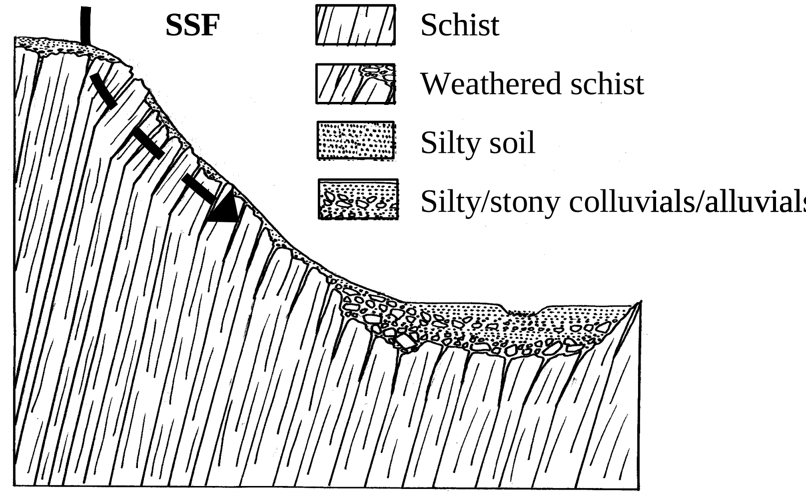


Different representations of spatial variability and hydrologic connectivity

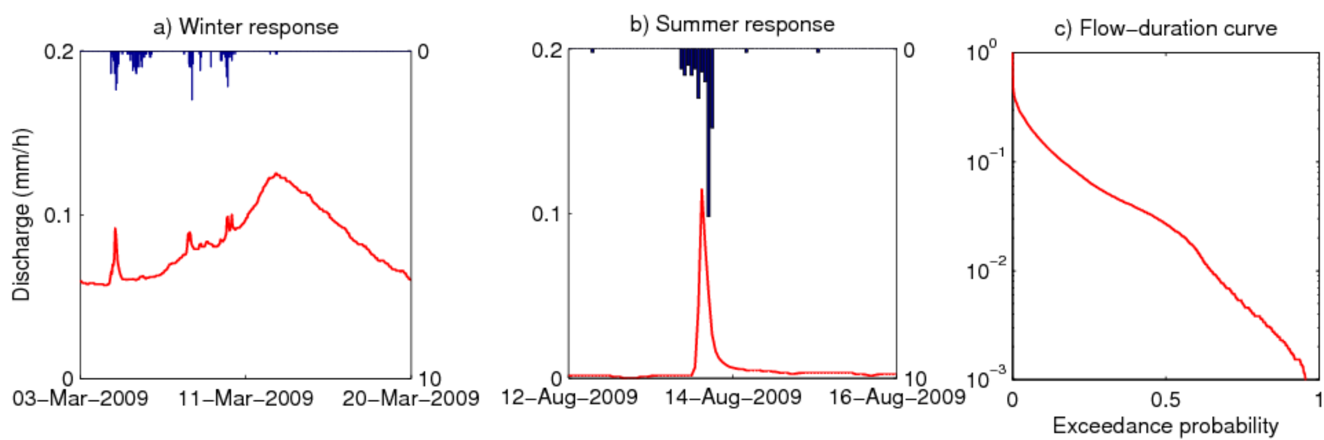




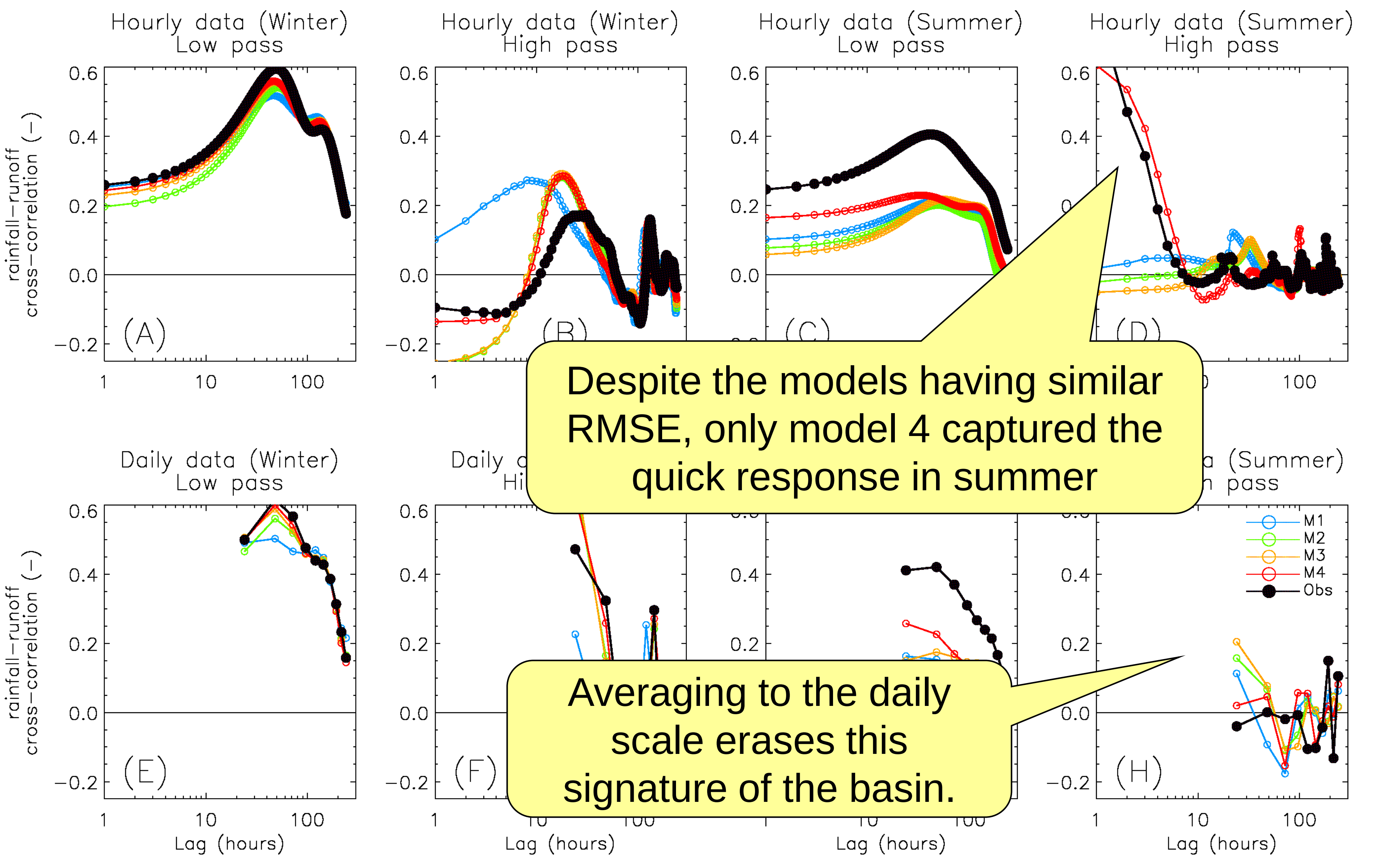
Using signatures instead of “The Nash“



Geological cross section



Model diagnostics: Capturing the rainfall-runoff cross-correlation signature of the Weierbach



Outline of Presentation

1. Some current challenges in hydrologic modeling
   - Data and structural errors
   - Some historical ways we’ve dealt with these challenges
2. Eliminating unnecessary artefacts
   - Robust numerical formulation of hydrological models
   - Impact on hydrological calibration
3. Model development and hypothesis-testing
   - Too many models with too many differences?
   - Towards more systematic model comparison
4. Some conclusions and a view to the future

Some of food (and strawmen?) for thought

- Some reasons why we “can’t do better than TOPMODEL”
  - Numerical artefacts often swamp model simulations, and have affected major historical choices in calibration methods
  - Myriads of uncontrolled differences obscure model comparisons
  - Anything more complex than TOPMODEL is “too complicated”

- Can we move past these difficulties?
  - Robust numerical approximations are just as important in “conceptual” models as they are in “physical” models
  - Model comparisons must proceed in a systematic way. This requires careful model design and case study setups
  - Further dialog between Modeller and Experimentalist

Some of food (and strawmen?) for thought

- Are we suffering “calibration fatigue”?
  - Yes if we try to use calibration as a panacea to all problems
  - Not if we carefully scrutinize the calibration setup
  - Going cold turkey on “calibration” just as extreme as making it the focus of hundreds of hydrological papers

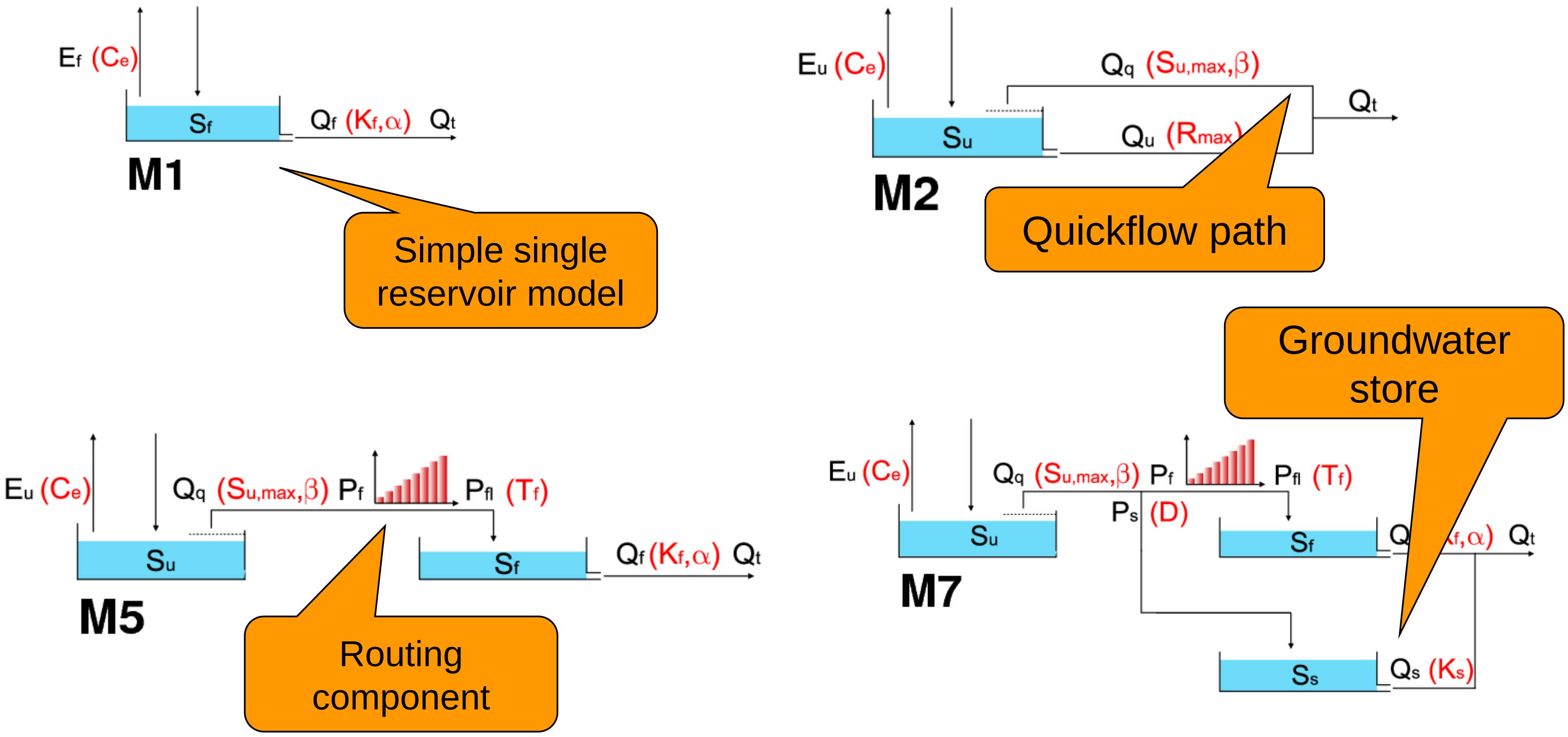
- Are we building a Modelling Tower of Babel?
  - “Models and modelling method” proliferation
  - Eventually confuses users, decision-makers … and modellers too!
  - Innovation is great … as long as we properly understand <u>already</u> existing methods
  - “Systematic” vs “ad hoc” development – there is a difference!


Questions?

Outline of Presentation

1. Some current challenges in hydrologic modeling
    - Data and structural errors
    - A more systematic treatment of distinct errors
2. Eliminating unnecessary artefacts
    - Robust numerical formulation of hydrological models
    - Impact on hydrological calibration
3. Bayesian Total Error Analysis (BATEA)
    - Concepts and motivation
    - Case studies and insights
4. Hypothesis-testing in environmental modelling
    - Concepts and motivation
    - Case studies and insights
5. Conclusions and a view to the future

Using SUPERFLEX for hypothesis-building



Used to explore/compare representations of specific catchments

Eg, double-peak schistose dynamics, snow modules, etc

=> BUT: needs to be evaluated to maintain parsimony

General mathematical formulation

- Consider the mathematics of hydrological models
- Sets of (coupled) differential equations

$$\mathrm{d}\boldsymbol{S}(t)/\mathrm{d}t = \boldsymbol{g}_S\left(\boldsymbol{S}(t), \boldsymbol{P}(t) \mid \boldsymbol{\theta}\right) \quad \ldots\ldots\ldots \text{(a)}$$

$$Q(t) = \boldsymbol{g}_Q\left(\boldsymbol{S}(t), \boldsymbol{P}(t) \mid \boldsymbol{\theta}\right) \quad \ldots\ldots\ldots \text{(b)}$$

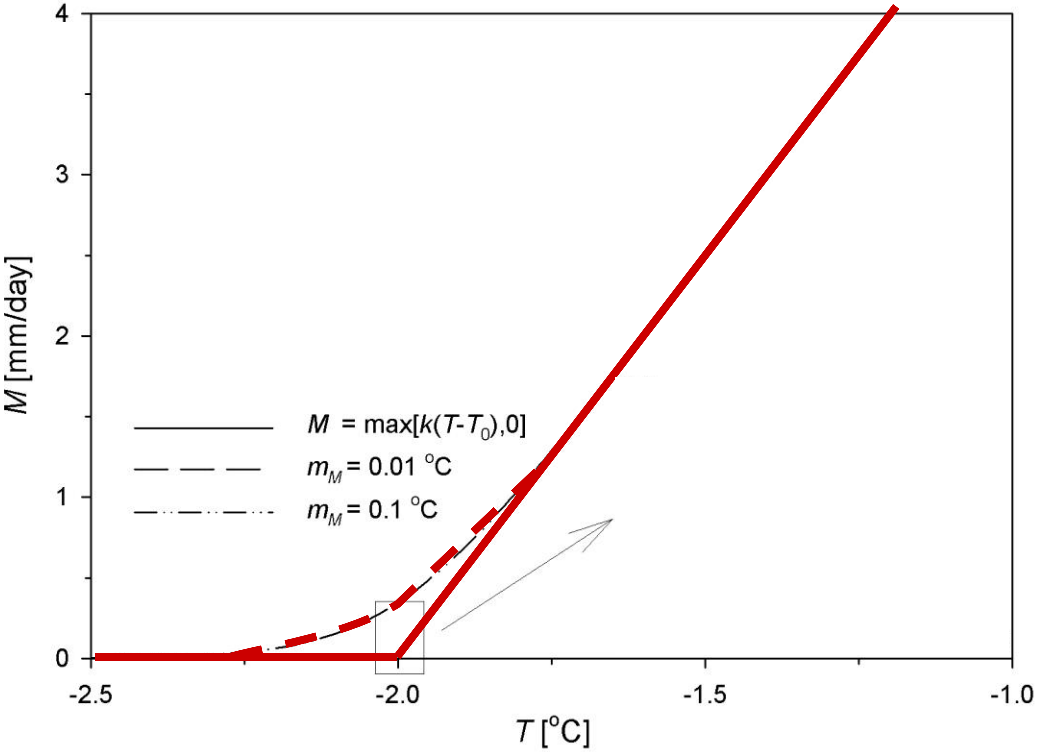
S = states, θ = parameters, *P* = forcings, *Q* = responses

- For example, VIC model (Wood, 1992)

$$\mathrm{d}S(t)/\mathrm{d}t = P\left[1 - S / S_{\max}\right]^{\alpha} - kS^{\beta} - E(S)$$

Another numerical aspect: Model smoothing

Replace discontinuities in the model fluxes (with respect to parameters and states) with smooth transitions



Wide selection of smoothers:

1) Splines
2) Special functions

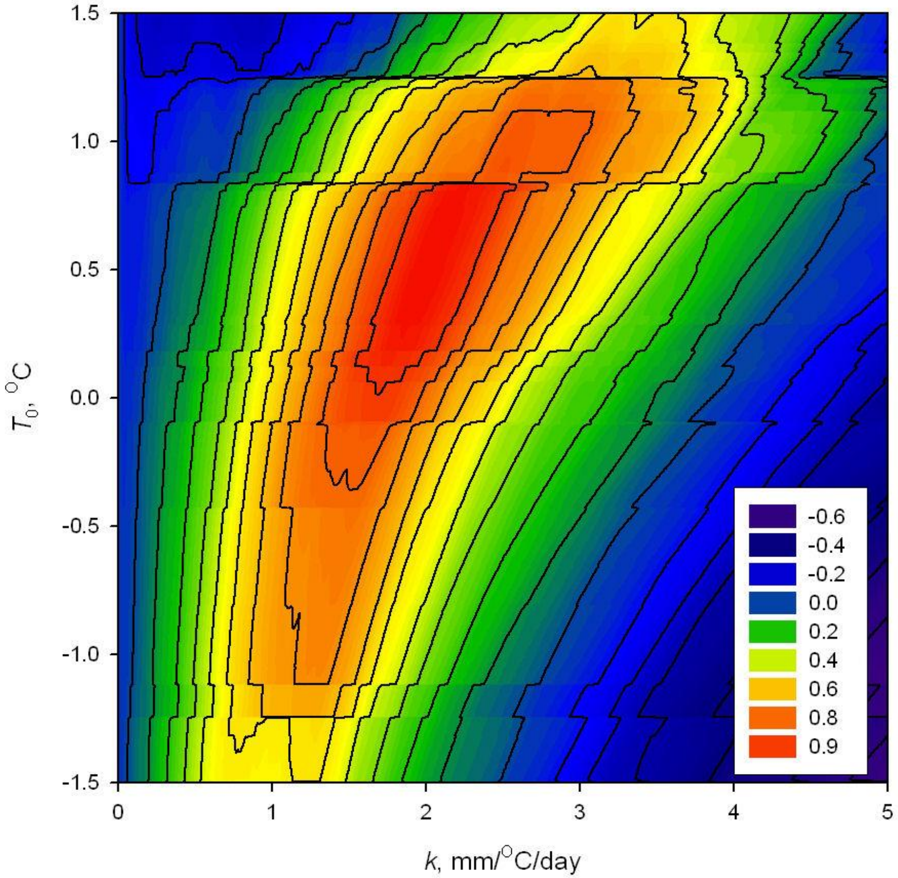
$$M = km\left(\bar{T} + \ln\left[1 + \exp\left(-\bar{T}\right)\right]\right)$$

$$T = \left(T - T_0\right)/m$$

Kavetski and Kuczera, 2007

$$M = \frac{k}{2}\left(T - T_0 + \sqrt{\left(T - T_0\right)^2 + m}\right)$$

Objective function: Before and After model smoothing

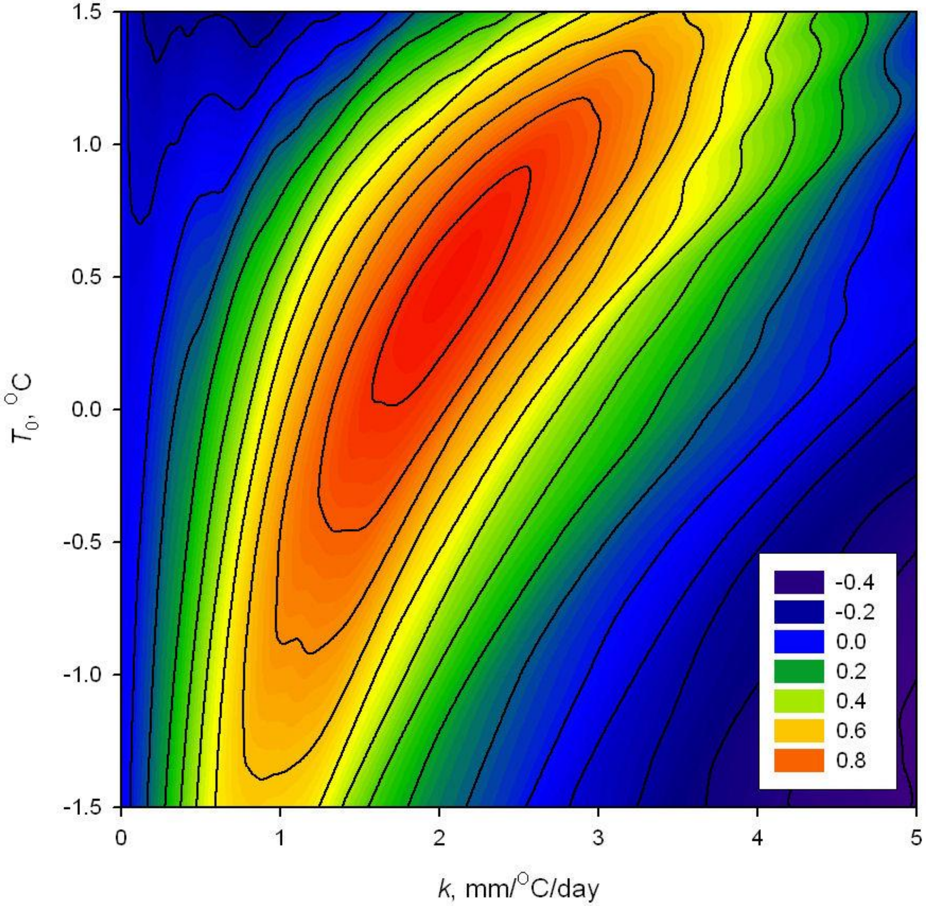


Original model

- Nonsmooth objective function
- Multiple optima

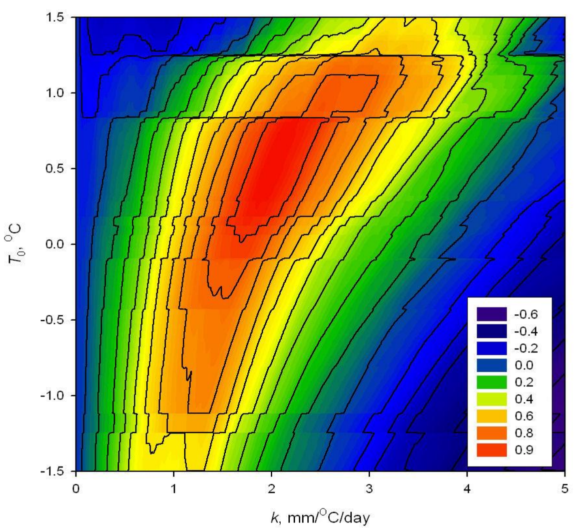
Smoothed model

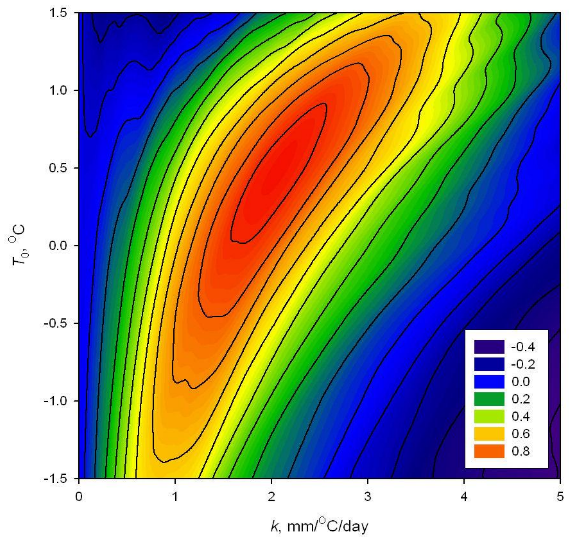
- Smooth objective function
- Single near-elliptic optimum



Kavetski and Kuczera, WRR2007

Objective function: Before and After model smoothing





in hydrology, numerical artefacts appear at least partially responsible for the shift from fast Newton-type methods to robust but much slower global optimizers

Newton ("ideal")	Newton-type (quasi-Newton, Gauss-Newton, etc)	Global optimizers (eg, SCE, genetic, annealing)
Classical algorithm	Common in applied mathematics	Designed for "tough" problems
Seldom applicable in hydrology	With some exceptions, abandoned in hydrology	Believed to be necessary in hydrology
1 min runtime	2-10 min runtime	hrs – weeks

Just a scratch at the surface: More examples and discussion:

- Numerical artefacts:
  - Clark and Kavetski (WRR2010) Ancient numerical daemons of hydrological modeling. Part 1 – Fidelity and efficiency.
  - Kavetski and Clark (WRR2010) Ancient numerical daemons of hydrological modeling. Part 2 – Impact on model application
  - Kavetski and Clark (HP2011) Numerical troubles in conceptual hydrology: Approximations, absurdities and hypothesis-testing

- Time resolution effects (with discussion of causes):
  - Kavetski, Fenicia and Clark (WRR2011) Impact of data resolution on conceptual hydrological modeling: Experimental insights