TOWARD TRANSPARENT, REFUTABLE HYDROLOGIC MODELS IN KANSAS OR OZ

In class:
   Mary C. Hill   U.S. Geological Survey, Boulder
   Olda Rakovec   Wageningen University, Netherlands
Other co-authors:
   Martyn Clark   UCAR, Boulder
   A. H. Weerts   Deltares, Wageningin University
   A. J. Teuling   Wageningin University
   R. Uijlenhoet   Wageningin University

Problem

- Transparency and refutability are championed in articles by Oreskes and others as necessary for meaningful environmental models
- Transparency and refutability are supposed to be served by model analysis methods.
  - Analyze model fit, sensitivity analysis, uncertainty
- Many model analysis methods
  - Nearly every field and even report use a different method.
- Many analysis methods require heroic efforts
  - 10,000s to millions of model runs.
  - Modelers forced to choose between processes and analysis.

- How well are transparency and refutability served?

- Can we do better?

Include computationally frugal methods and organize by purpose

Model Adequacy

- **How to include many data types with variable quality?** Error-based weighting and SOO, MOO*
- **Is model misfit/overfit a problem? Are prior knowledge and data subsets inconsistent?** Variance of weight-standardized residuals, residual graphs and space/time plots, MOO*
- **How nonlinear is the problem?** Modified Beale's measure, Explore objective function*, TSDE*

Sensitivity and Uncertainty

Observations ⟷ Parameters

- **What parameters can be estimated with the observations?** b/SD_b, CSS&PCC, SV, DoE*, IR*, First and total-order effects from OAT*, eFAST, Sobol', RSA* (First-order only)
- **Which observations are important and unimportant to parameters?** Leverage, Cook's D, CV*, MOO*
- **Are any parameters dominated by one observation and, thus, its error?** Leverage, DFBETAS, CV*
- **How certain are the parameter values?** b/SD_b, Parameter uncertainty intervals[#]

Parameters ⟷ Predictions

- **Which parameters are important and unimportant to predictions?** PSS, FAST*
- **How certain are the predictions?** z/SD_z, Prediction uncertainty intervals [#], MMA*
- **Which parameters contribute most and least to prediction uncertainty?** PPR, FAST*, Sobol',* MCMC*

Observations ⟷ Predictions

- **Which existing and potential observations are important to the predictions?** OPR, CV*
- **Which models in MMA are likely to produce the best predictions?** For individual model evaluations: AIC, AICc, BIC, KIC, CV*

Sensitivity analysis

- Exposes what is important to a given situation, test, etc.

- Needed????

"Would you go to a orthopedist who didn't take x-rays and show them to you?"

SimLab web site: http://simlab.jrc.ec.europa.eu

4

Outline

- Some basics about sensitivity analysis
- Sensitivity analysis – Global and local methods
  - Total-order statistics
- Simple test case (analytical example relevant to FUSE) – run local sensitivity analysis in class
- Sensitivity analysis results for the test case
- Some results from FUSE

Some basics about sensitivity analysis

- Decide on model
- Decide on parameters
  - Can include parameters on unusual aspects of the model, like additive or multiplicative parameters for precipitation
  - Decide on ranges for each parameter
  - The defined parameter space has the dimensions of the number of parameters, k. Easy to plot results for 1- and 2-D parameter space. Could easily be 10-D parameter space!
- Decide on performance metric, ψ
- Calculate performance metric variability over parameter space. Measure it with the variance $V(\psi)$

Some basics about sensitivity analysis

- Total-order sensitivity for parameter j S_T^j

  - 1 minus Fraction of $V(\psi)$ explained by variation in all other parameters.

  - Includes of interdependence of parameter j with the other parameters.

- Values of S_T^j range from 0.0 to 1.0.

  - Small value ➜ unimportant parameter

  - Large value ➜ important parameter

Sensitivity analysis – Global method

- Define a ranges for each parameter

- Sample the parameter space R times using Sobol'. For each, calculate performance metric, ψ

- $V(\psi) = \Sigma_{r=1,R} \; \psi_r^2 \; - [\Sigma_{r=1,R} \; \psi_r]^2$

- $S_T^{\,j} \quad = 1 - [U - \Sigma_{r=1,R} \; \psi_r^2 \,] / V(\psi)$

  $U = (1/R)\Sigma_{r=1,R} \, [\psi_r \, \psi_r^{-j}]$

- Number of model runs = R(k+1)

- R typically 10,000

Sensitivity analysis – Local methods

- Sample the parameter space N times. For each, calculate ψ and $\partial\psi/\partial\theta_j$ (θ_j is parameter j)
- $V_L(\psi) = \qquad (\partial\psi/\partial\theta)^T \, V(\theta) \, (\partial\psi/\partial\theta)$
- $S_{T(L)}{}^j = 1 - [(\partial\psi/\partial\theta)^T \, V(\theta)_{(j)} \, (\partial\psi/\partial\theta)] / V_L(\psi)$ \qquad (New)
- $V(\theta)$
  - Here, $V(\theta)$ diagonals equal parameter variances and $V(\theta)_{(j)}$ has a zero for parameter j.
  - Make variances consistent with parameter ranges used for global method. Here, st dev = (max-min)/2.
- Number of model runs = L(2k + 1)
- N typically 1 for local results. We suggest the idea of multi-location local methods, and N>1.

Simple Test Case
(analytical example relevant to FUSE)

- Nonlinear reservoir with 2 parameters
- ODE: $dS/dt = p(t) - q(t)$
- Here, p(t)=0, so: $q(t) = K(S(t)/S_{sc})^c$
  - where K[mm/day] and c[-] are analyzed parameters
  - S_{sc} [mm] is a scaling parameter that is held constant
- Goal: investigate sensitivity of K and c to $\psi = mean(S)$
- Used R script that solves the problem using fixed-time step, implicit Euler.
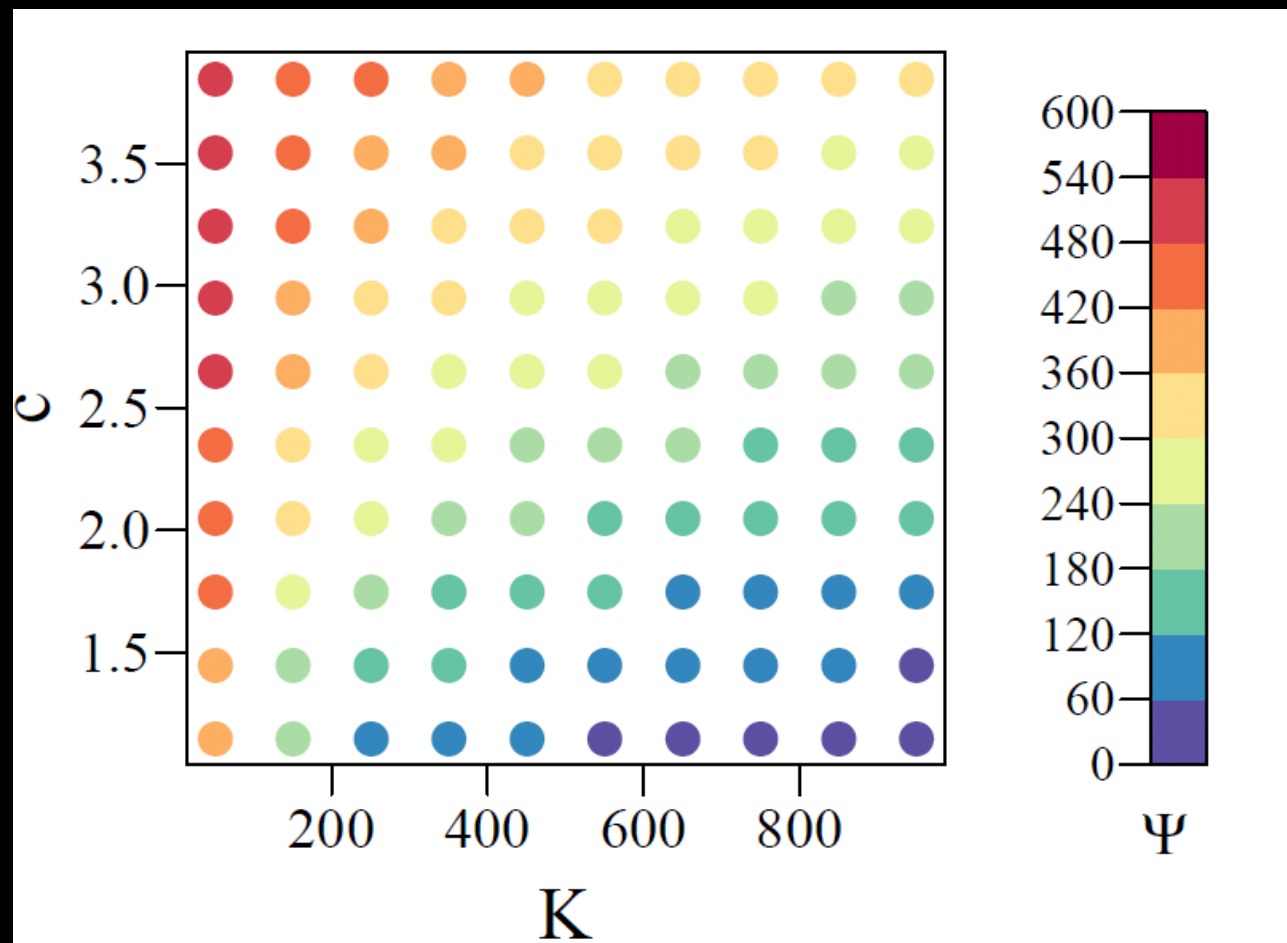- Solution is in R script FUN_Nonlinear_Implicit_v02.r

Simple Test Case

- Run R script "Sensitivities_Global_Local_v02.r"
- Time series plots of storage (Files: 1_Plot_S__K900_vs_c.pdf, 1_Plot_S__K100_vs_c.pdf)
- ψ = mean storage over time, so average over each curve
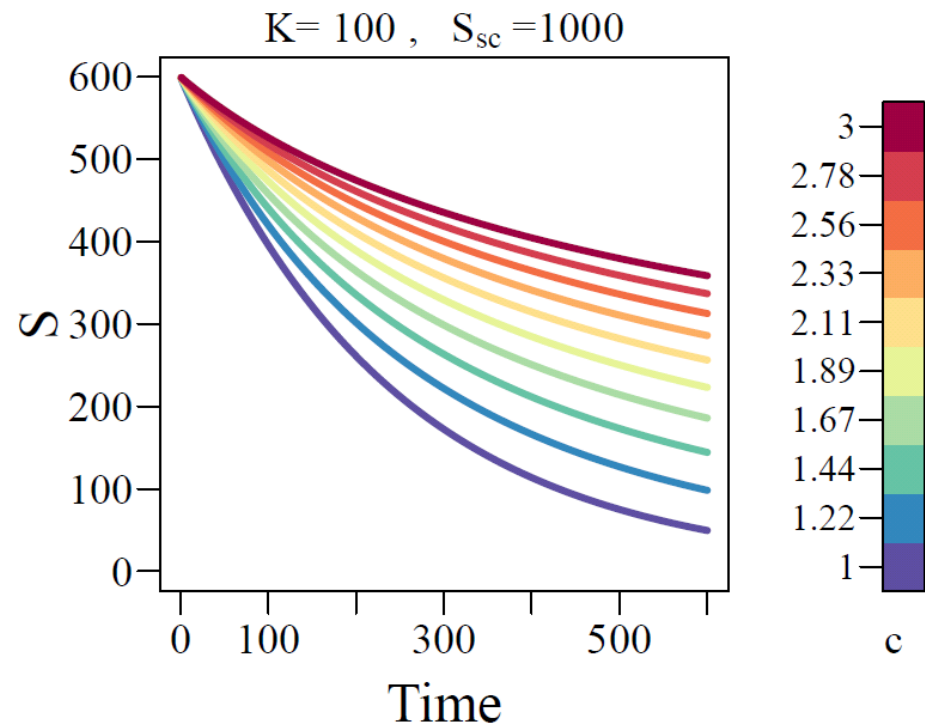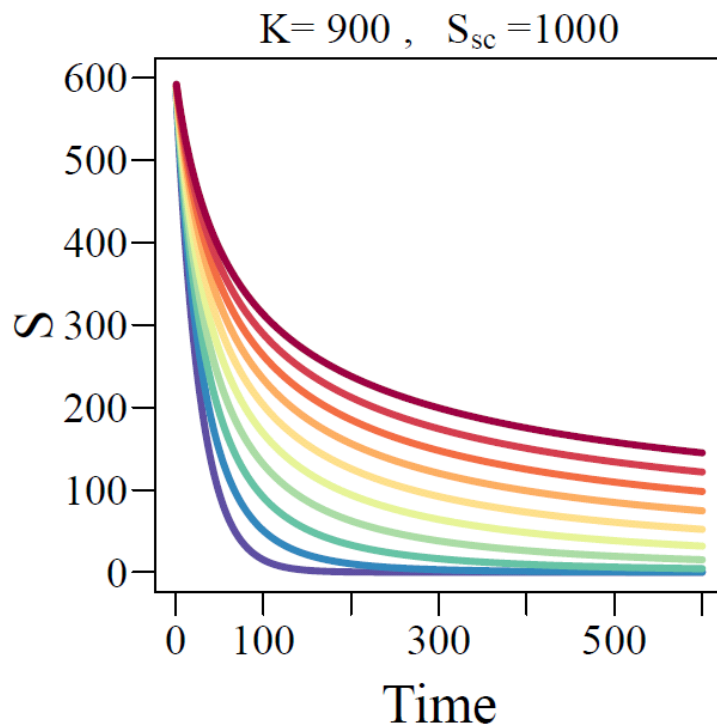
Simple Test Case

Performance metric ψ (mean storage)

File: 2_Image_OF_2D_K_vs_c.pdf

Typical Sobol' Global Sensitivity Analysis

File: 7_Barchart_SobolSens_NONLIN888.pdf

Sobol, whole domain

Sj — First-order sensitivity
St — Total order sensitivity

Global Sobol' suggests the two parameters have the same sensitivity.
10,000 model runs
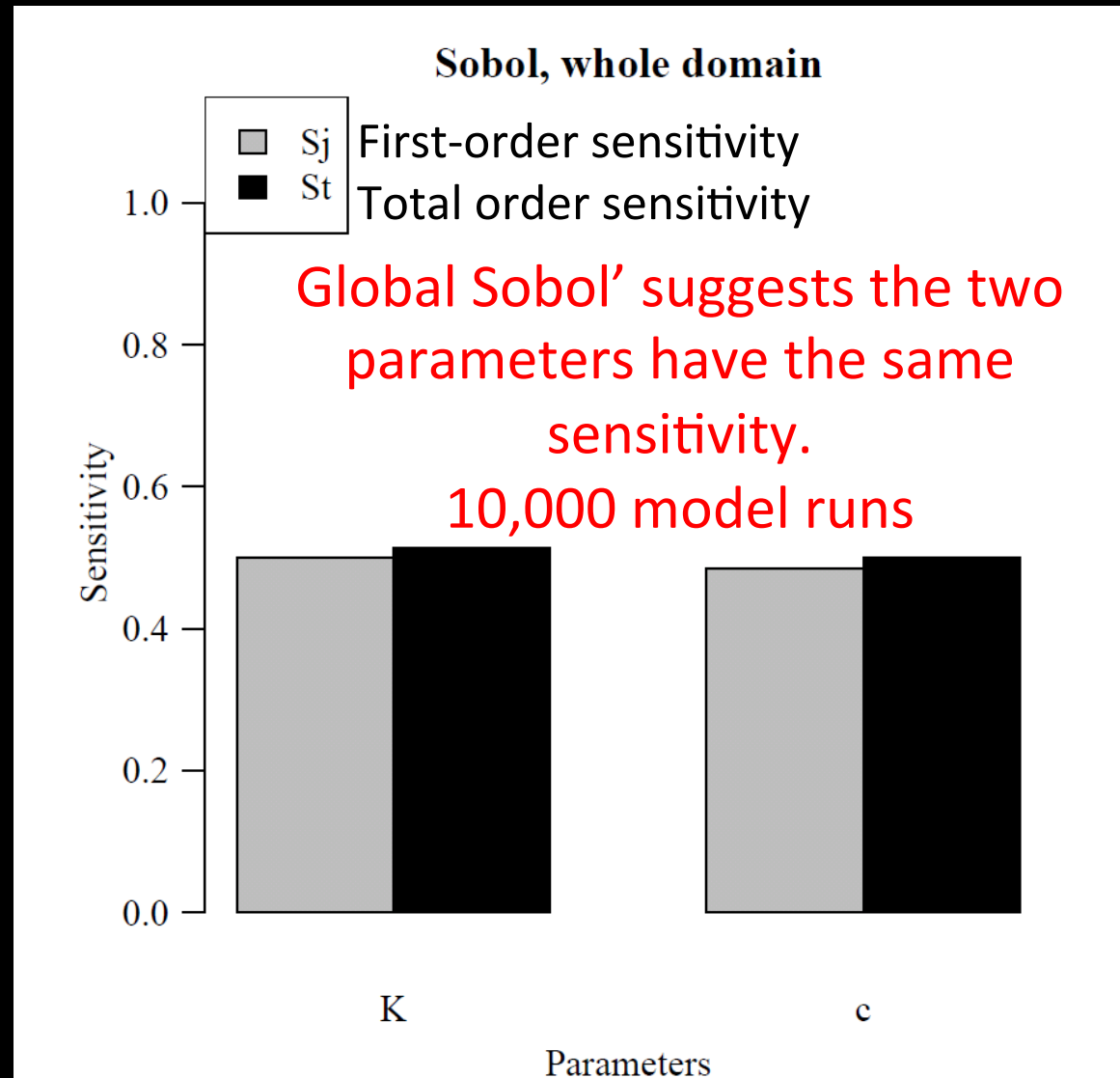
Can we say more?

- Consider evaluations at many locations within the parameter space. Like these 100 locations.
- How does parameter importance change as the solution conditions change?

File: 2_Image_OF_2D_K_vs_c.pdf

Global and local statistics both need V(ψ). What values of V(ψ) do we get at each point using Sobol' and local methods?
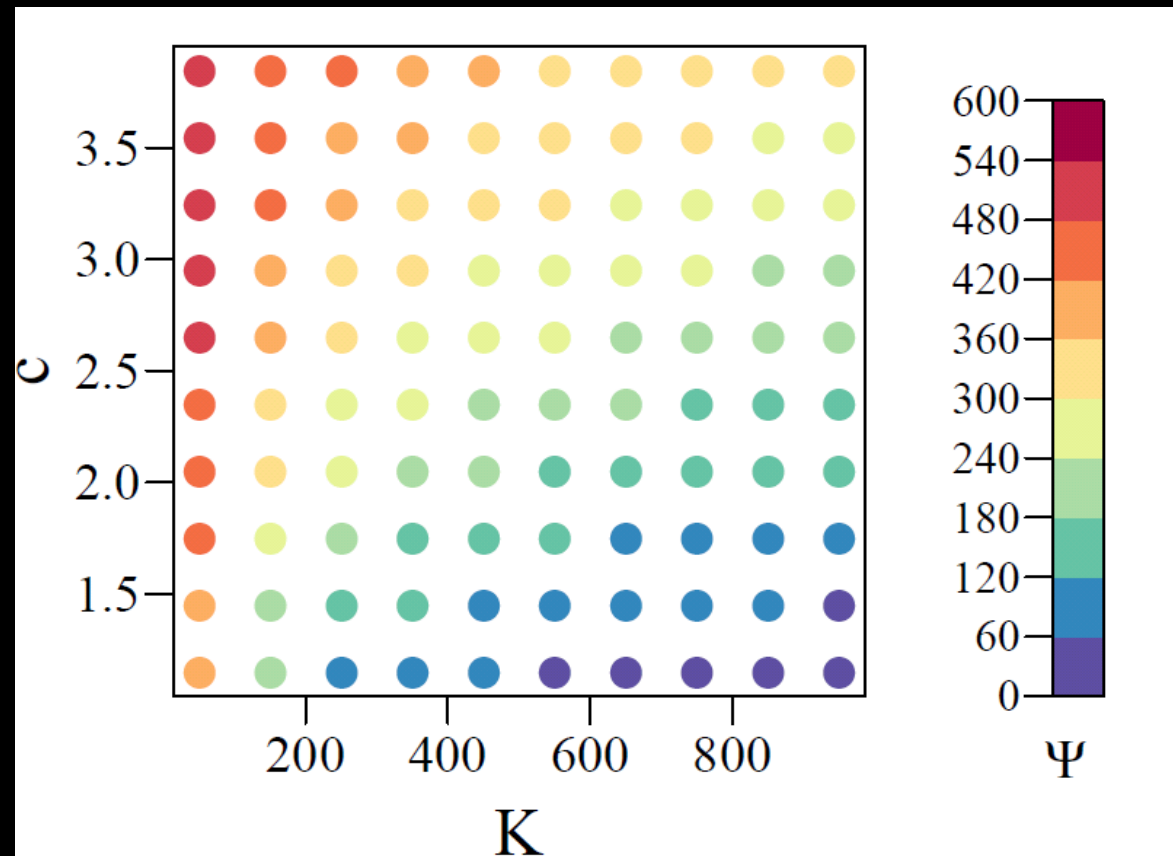
Sobol' Global variance½, V(ψ)$^{1/2}$

These are the 100 values -- V(ψ)$^{1/2}$ calculated for each of the 100 points plotted two ways.
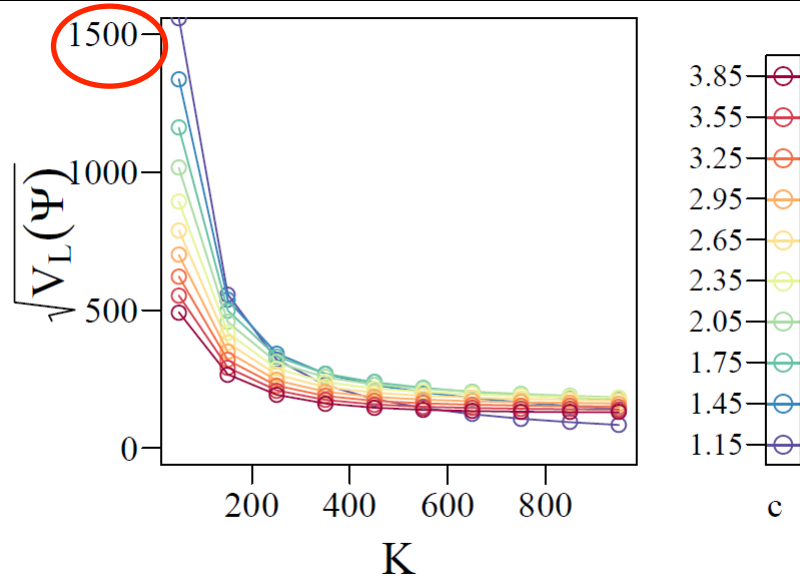Both in file: 3_Variances_Sobol.pdf
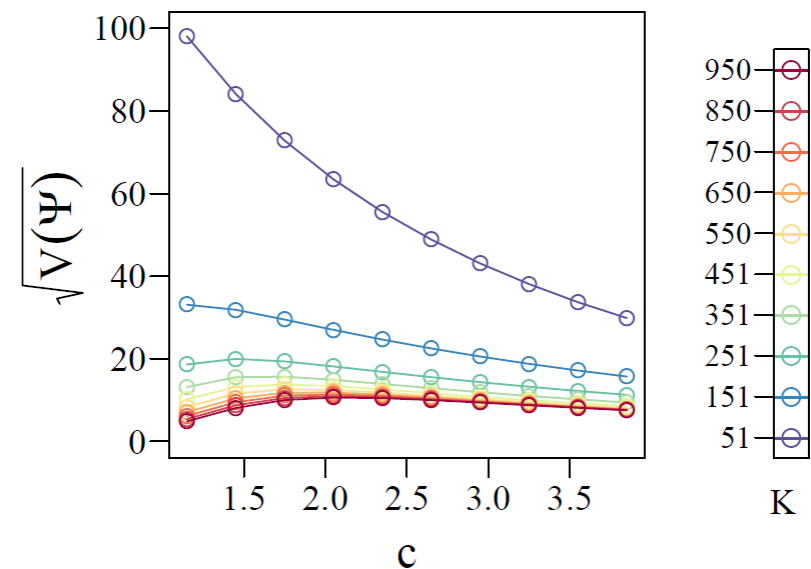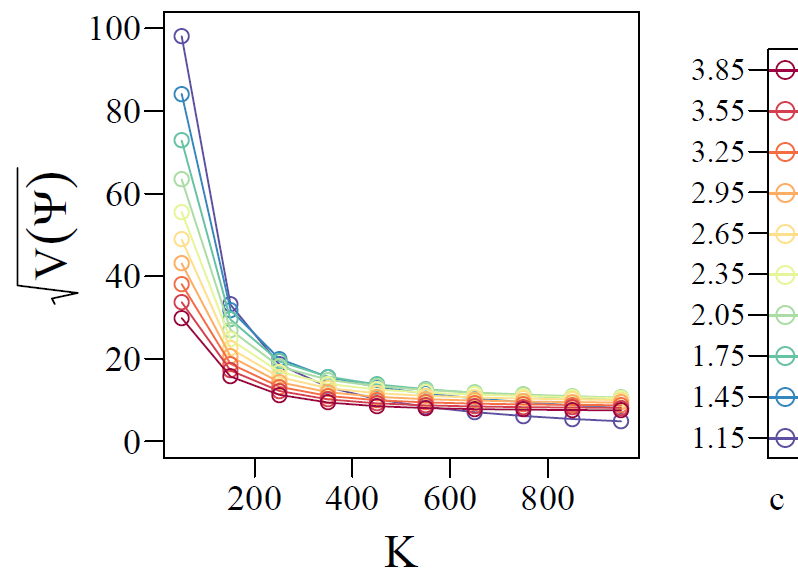
Bootstrapping suggests 3,000,000 model runs are needed to obtain stable results
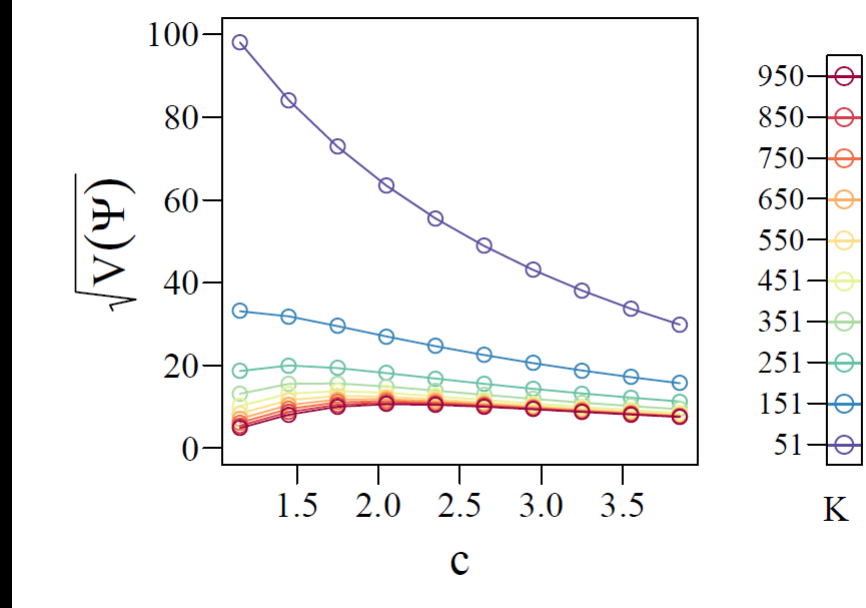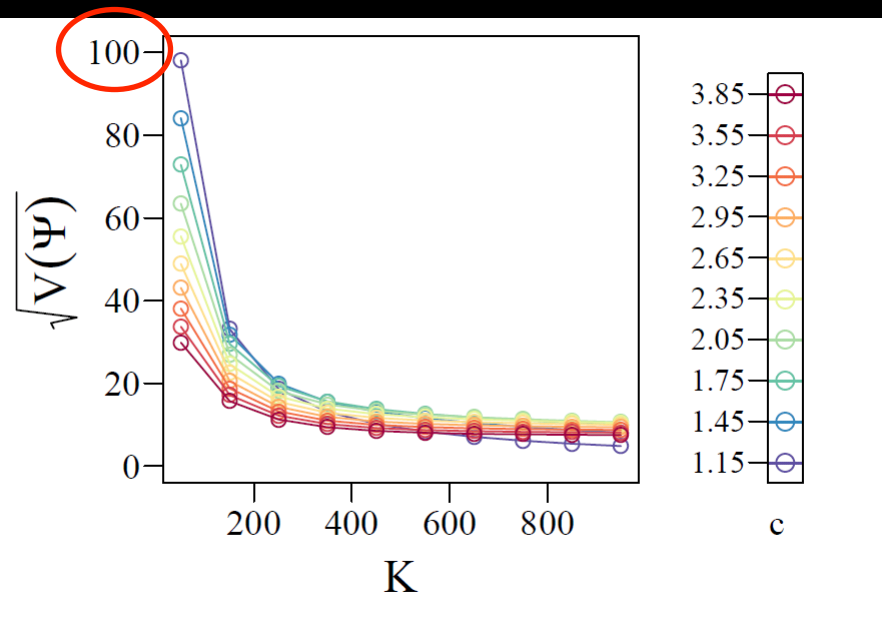
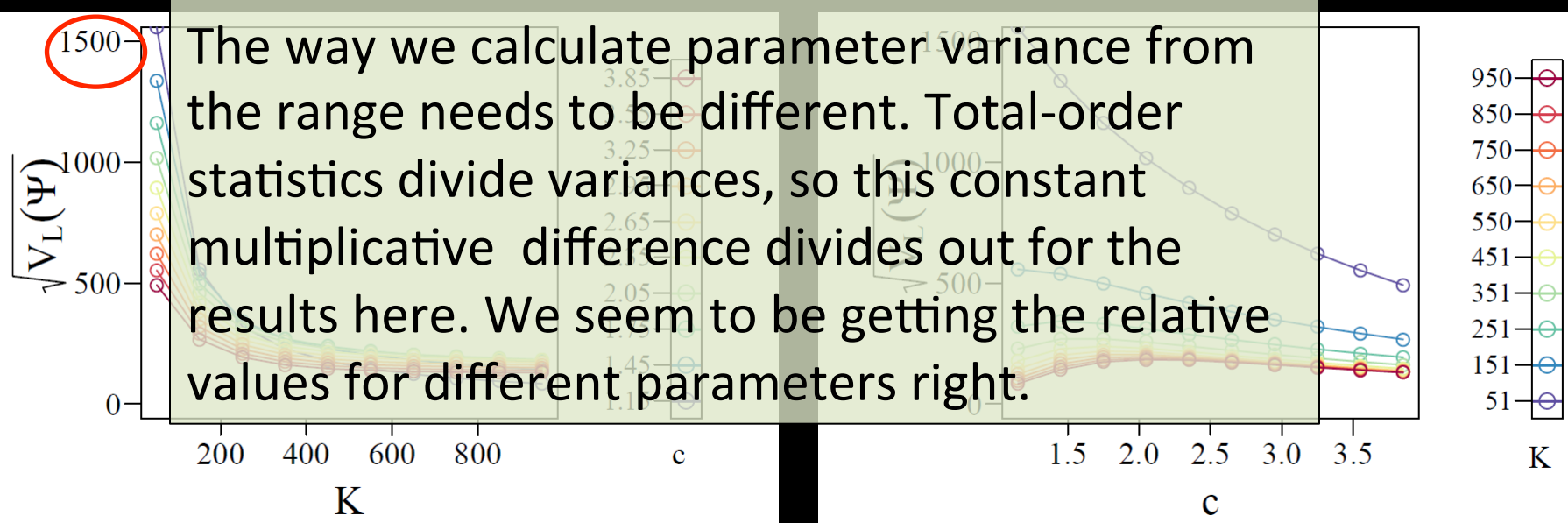Sobol' variance (3,000,000 model runs)   file: 3_Variances_Sobol.pdf
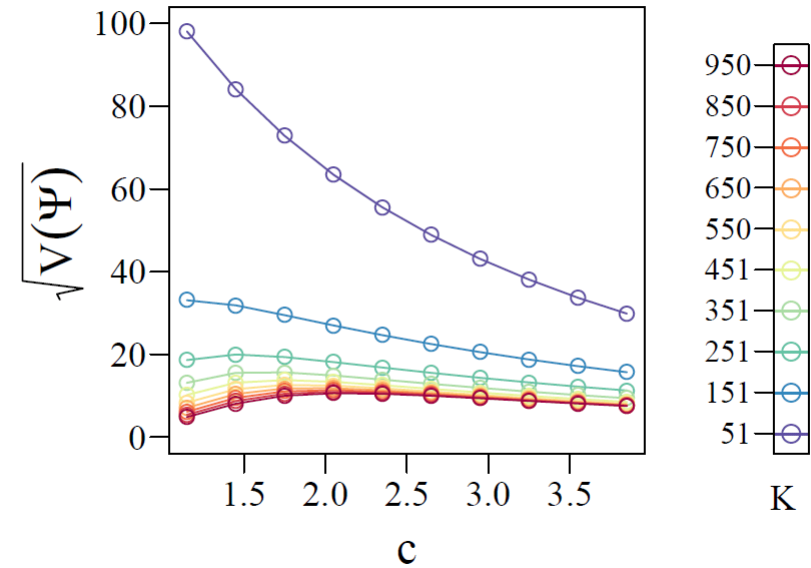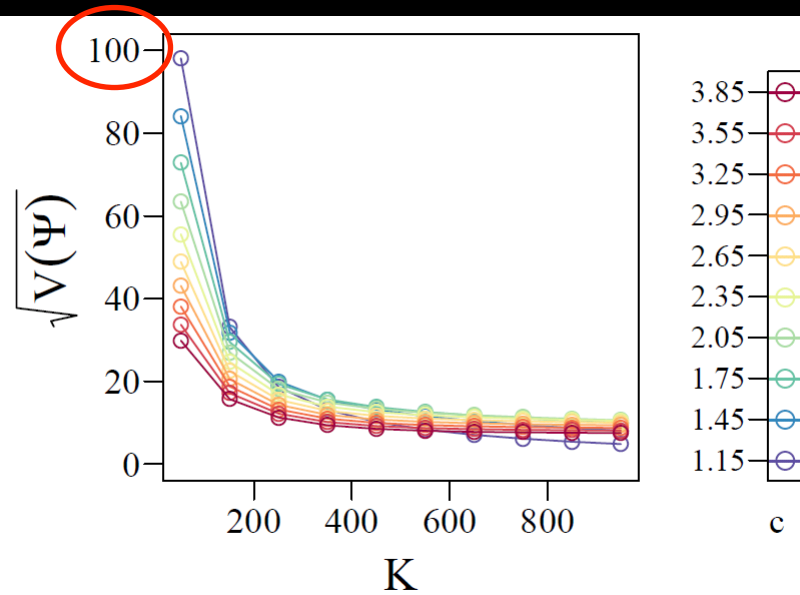
Local variance (300 model runs)   file: 4_Variances_psi_local.pdf

# Sobol' variance (3,000,000 model runs)   file: 3_Variances_Sobol.pdf

Local variance (300 model runs)   file: 4_Variances_psi_local.pdf

The way we calculate parameter variance from the range needs to be different. Total-order statistics divide variances, so this constant multiplicative difference divides out for the results here. We seem to be getting the relative values for different parameters right.
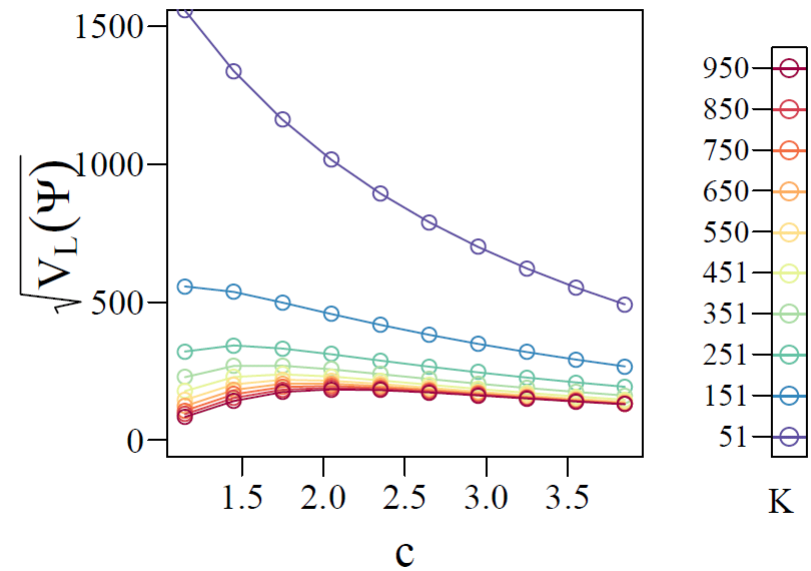
Total sensitivity indices (S_T) at many locations

Results obtained from local ● (300 model runs) and global ■ (3,000,000 model runs)

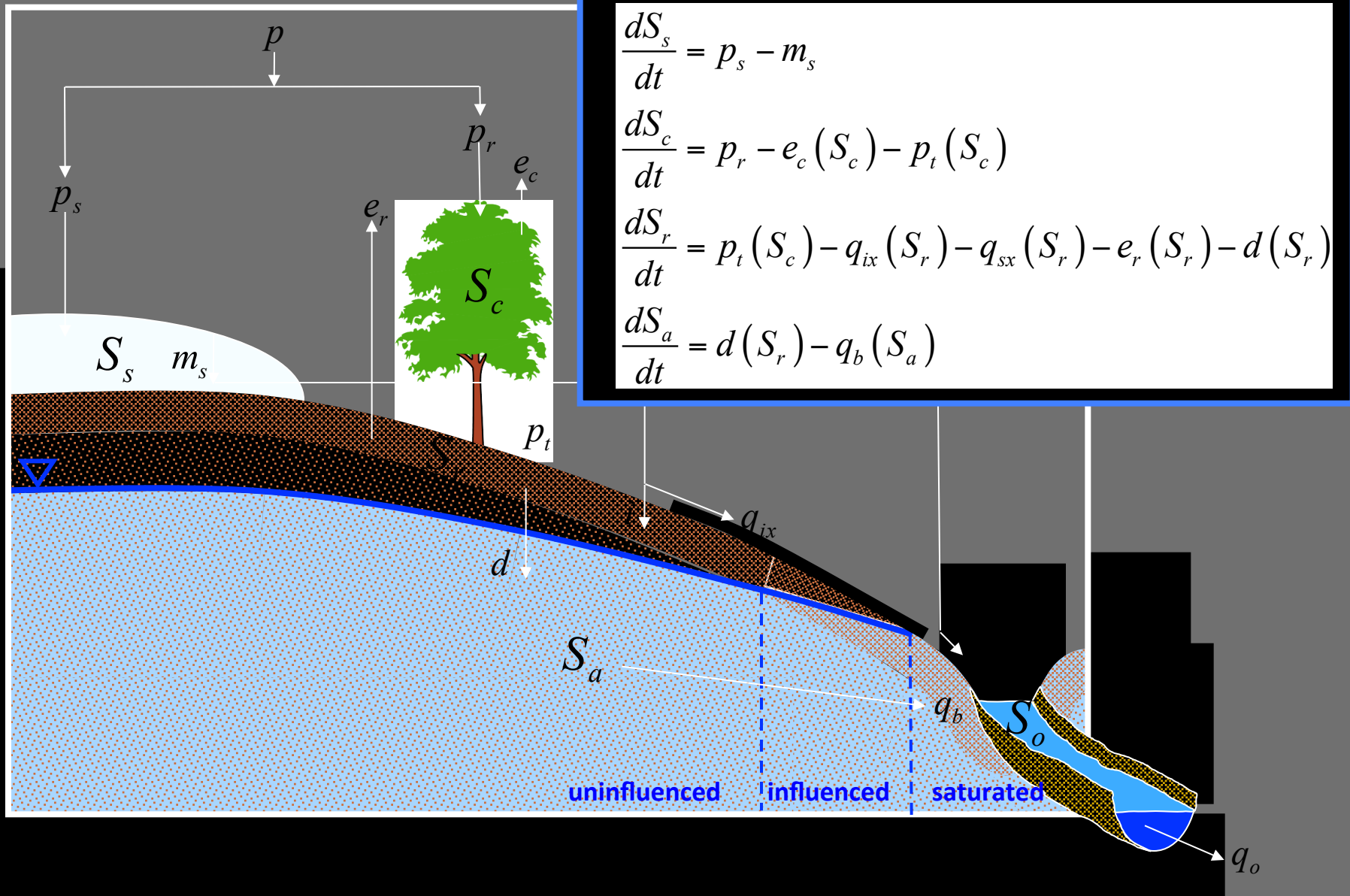File: 5_Image_sens_tot_2D_K_vs_cx.pdf
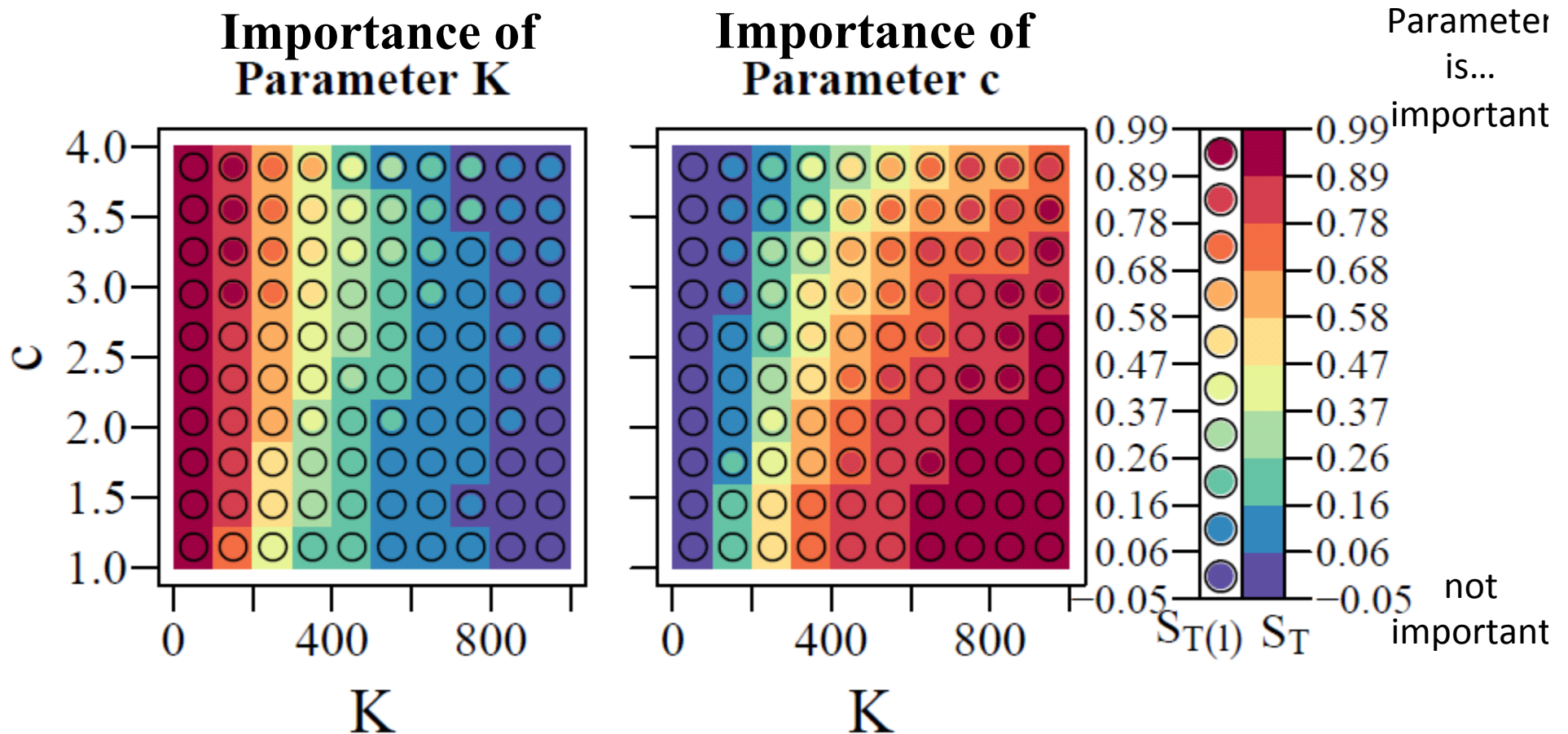
Exercise

- The Sobol' results are provided – the runs take too long for class (about 12 hours).

- You can produce the local sensitivity results in class using an R script written by Olda.
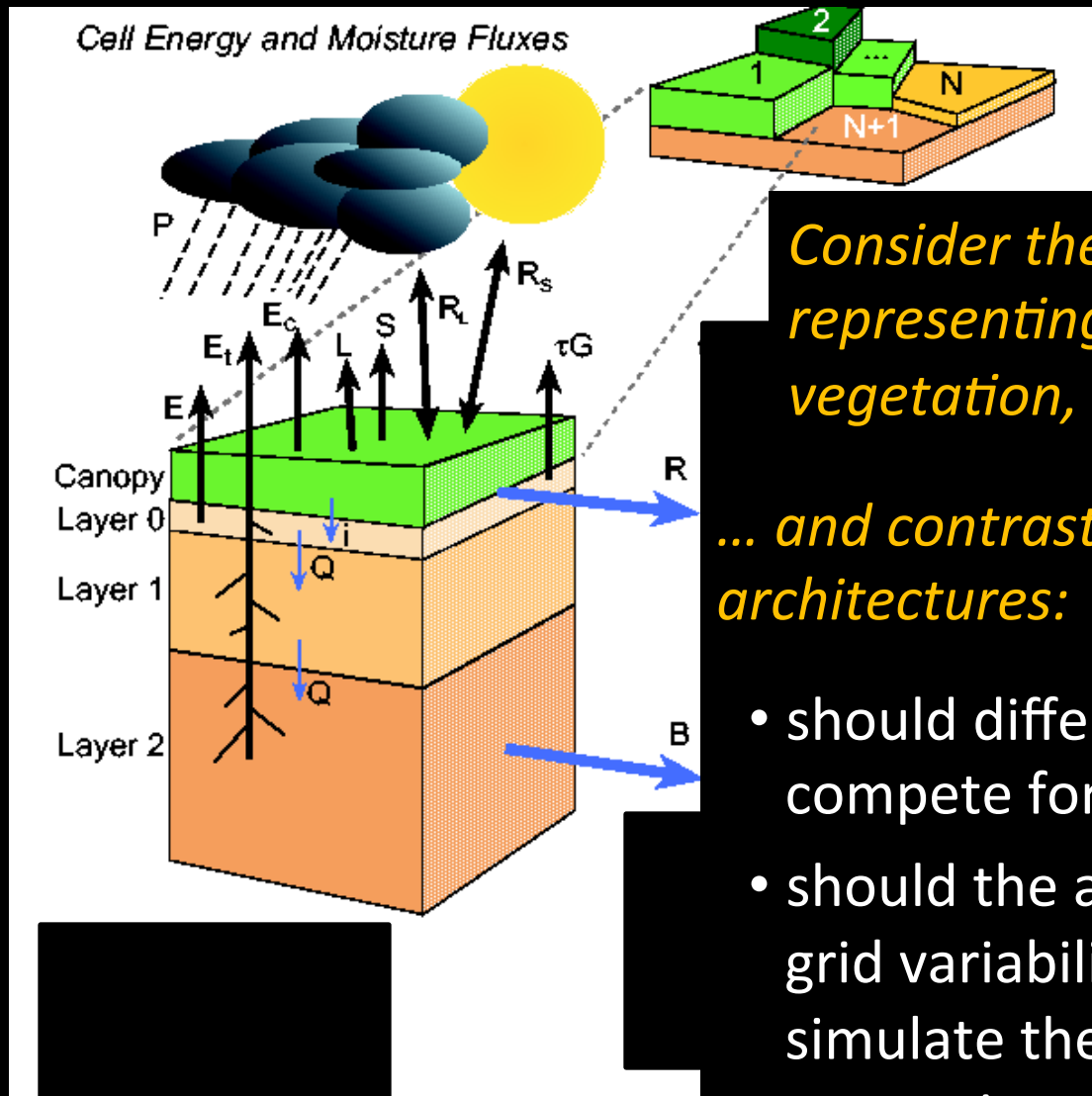
Exercise instructions

- How to run
  - Start R using R-2.15.3/bin/i386/Rgui.exe or R.bat
  - File/Open script/ Sensitivities_Global_Local_v02.r Change path on line 17 to where the R script is. Change any \ to /. Spaces ok.
  - Run program: With R script window active, cntl-a, cntl-r.
- PDF files are produced
  - All the files in the ppt presentation. You did all the calculations for the local method right here.
- Explore
  - For the global method defined parameter ranges (max,min).
  - For local method used st dev=(max-min)/2.
  - R script Sensitivities_Global_Local_v02.r
    - Defines st dev for K as WW1=499.5 and for c WW2=1.5. Use caps to search for variable names.
    - Try changing these values by a constant such as multiplying by 10. Do the figures change?
    - Try changing just one of the values by a factor of 10. Do the figures change?

FUSE (Clark et al, 2008)

FUSE accommodates different decisions regarding process selection and representation

$$\frac{dS_s}{dt} = p_s - m_s$$

$$\frac{dS_c}{dt} = p_r - e_c(S_c) - p_t(S_c)$$

$$\frac{dS_r}{dt} = p_t(S_c) - q_{ix}(S_r) - q_{sx}(S_r) - e_r(S_r) - d(S_r)$$

$$\frac{dS_a}{dt} = d(S_r) - q_b(S_a)$$

p

p_r

e_c

p_s

e_r

S_c

$S_s \quad m_s$

p_t

q_{ix}

d

S_a

$q_b \quad S_o$

uninfluenced   influenced   saturated

q_o

FUSE accommodates different options for model architecture

Cell Energy and Moisture Fluxes

Consider the "mosaic" approach to representing heterogeneity in vegetation, as in the VIC model

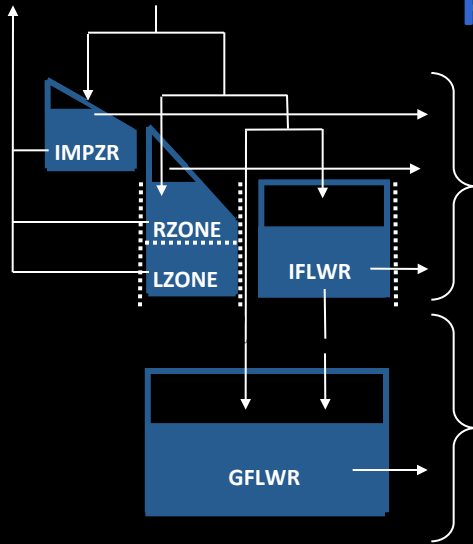*... and contrast against other model architectures:*

- should different vegetation types compete for the same soil moisture?
- should the architecture allow for sub-grid variability in soil moisture, to simulate the co-variability between vegetation type and water availability?

FUSE Separates the hypothesized model equations from their solutions

- Many hydrologic models use ad-hoc approaches to add/subtract model fluxes from model stores in a pre-determined sequence

- For example, the Sacramento model implementation

- **Critically**: Lack of attention applied to numerical error control

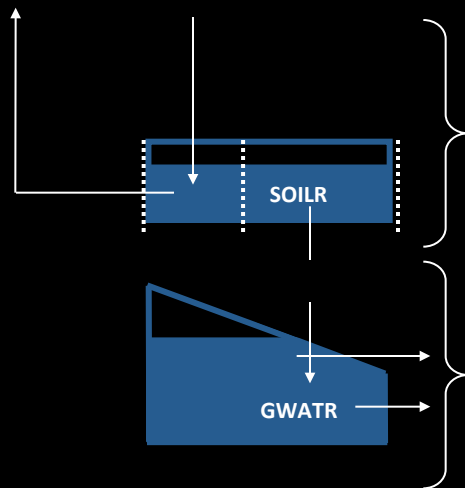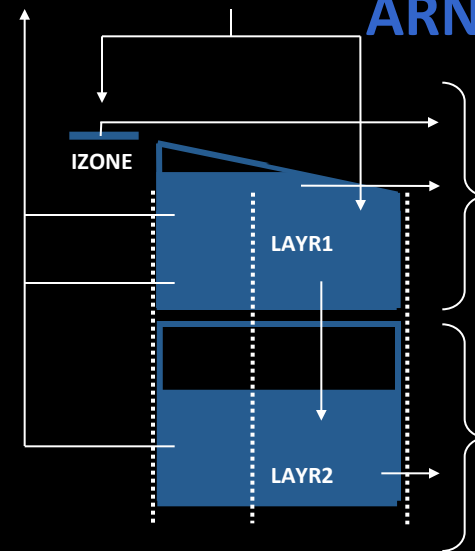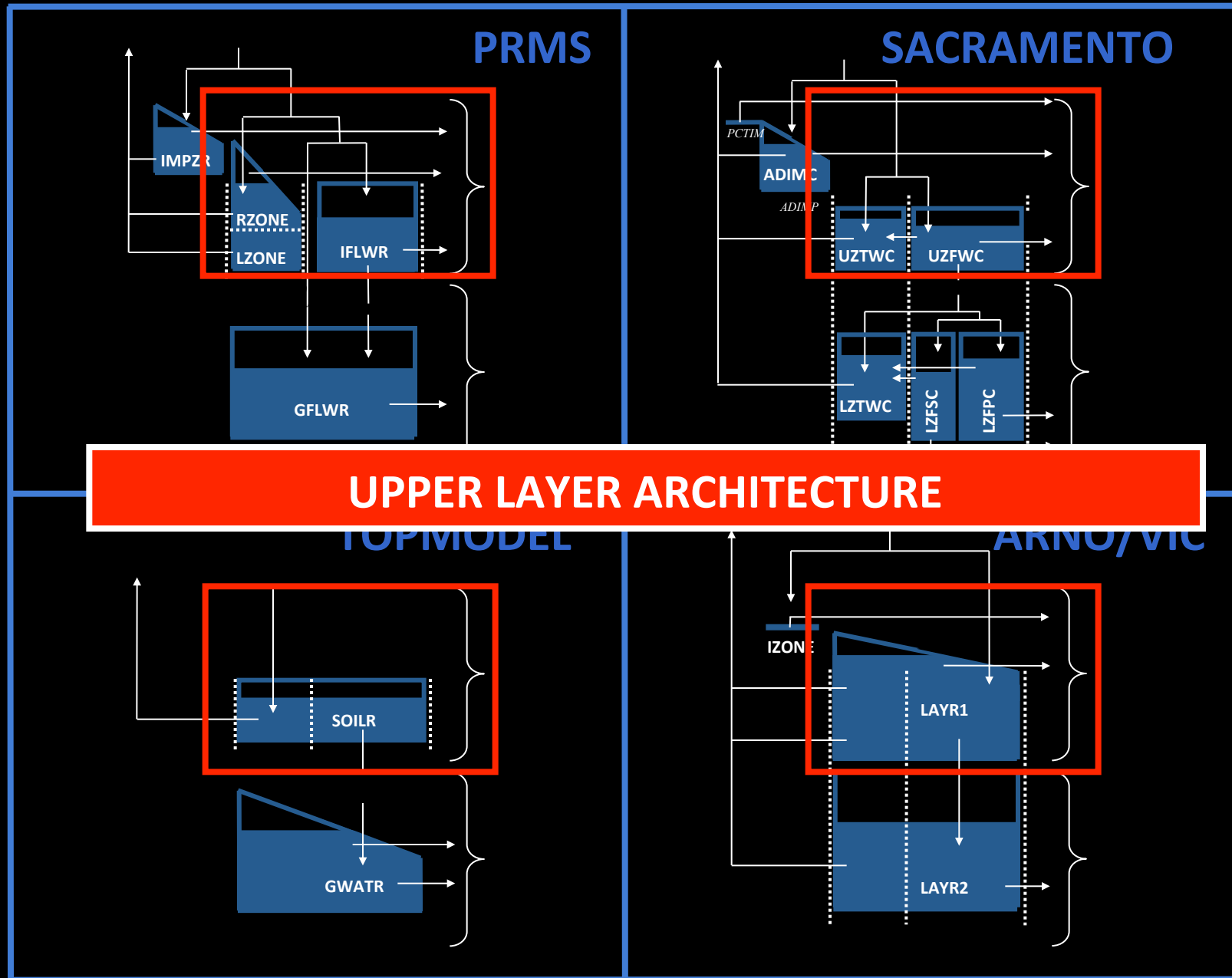# FUSE supports quantifying uncertainty using multiple models...
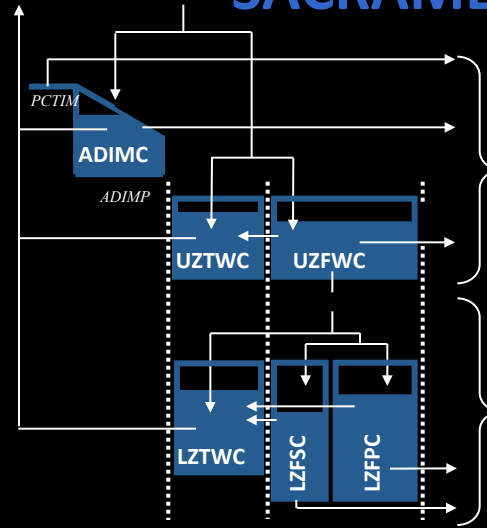
Define model building decisions—1

PRMS

SACRAMENTO

IMPZR

RZONE

LZONE

IFLWR

GFLWR

PCTIM

ADIMC

ADIMP

UZTWC

UZFWC

LZTWC

LZFSC

LZFPC

UPPER LAYER ARCHITECTURE

TOPMODEL

ARNO/VIC

SOILR

GWATR

IZONE
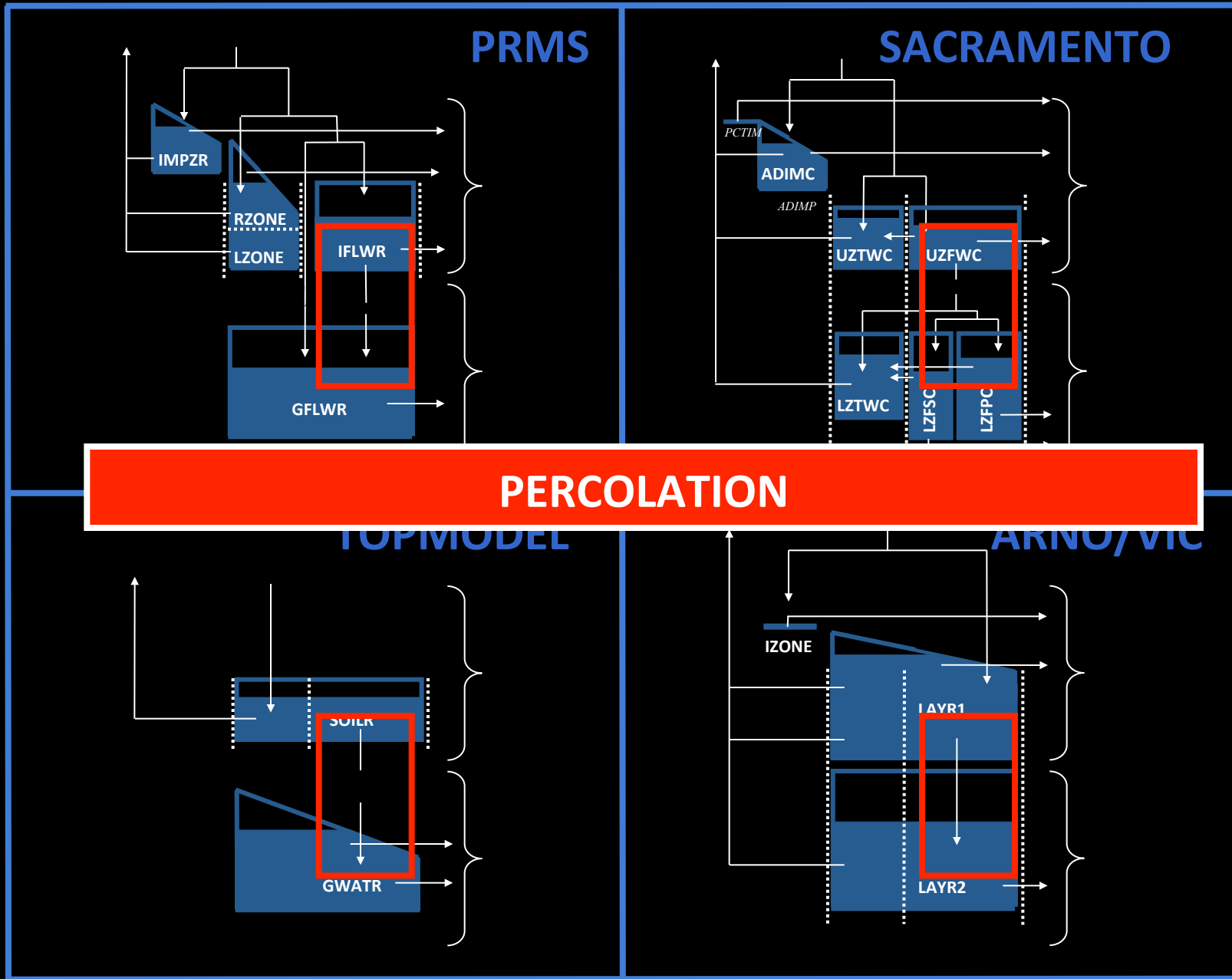
LAYR1

LAYR2

Define model building decisions—2

PRMS

SACRAMENTO

TOPMODEL

ARNO/VIC

LOWER LAYER ARCHITECTURE / BASEFLOW
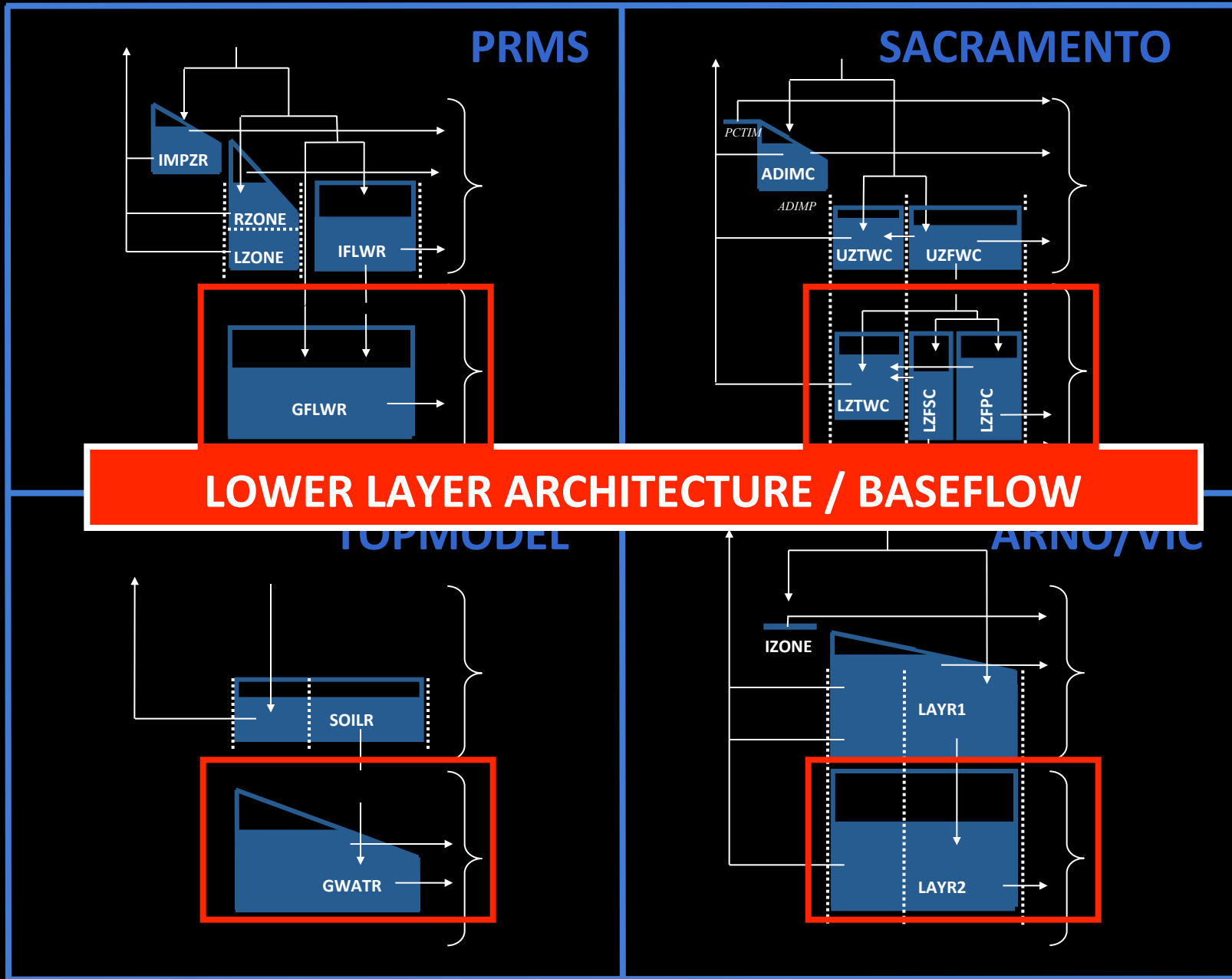
Define model building decisions—3

Define model building decisions—4

PRMS

IMPZR
RZONE
LZONE
IFLWR
GFLWR

SACRAMENTO

PCTIM
ADIMC
ADIMP
UZTWC
UZFWC
LZTWC
LZFSC
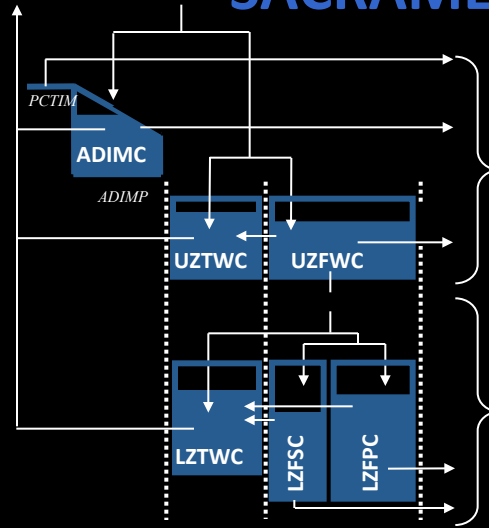LZFPC

SURFACE RUNOFF

TOPMODEL

SOILR
GWATR

ARNO/VIC

IZONE
LAYR1
LAYR2
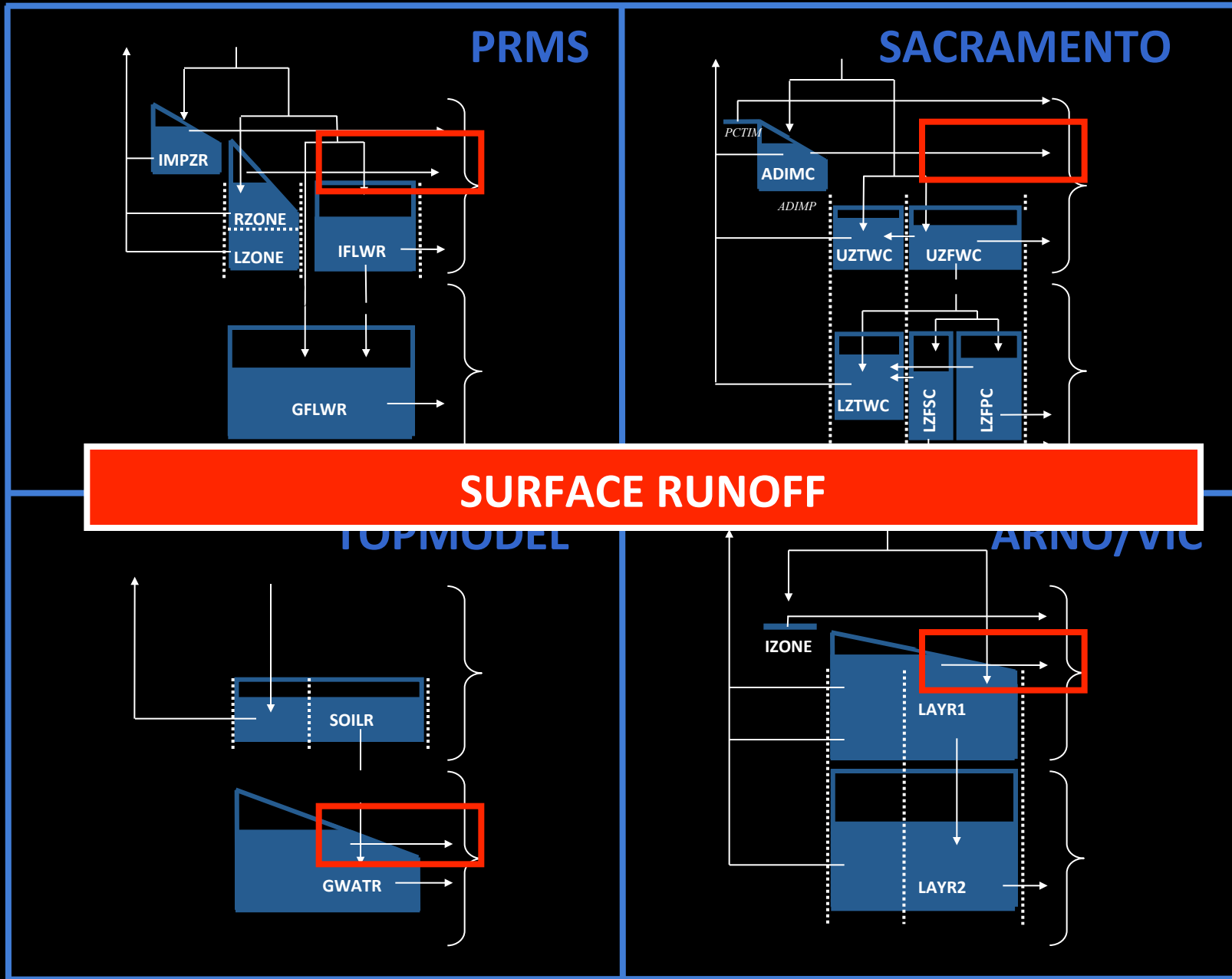
and mix architecture + parameterizations…

PRMS

IMPZR
RZONE
LZONE
IFLWR
GFLWR

SACRAMENTO

PCTIM
ADIMC
ADIMP
UZTWC
UZFWC
LZTWC
LZFSC
LZFPC

TO CREATE HUNDREDS OF MODELS,
ALL WITH DIFFERENT STRUCTURE

O/VIC

SOILR
GWATR

IZONE
LAYR1
LAYR2

FUSE-016

PRMS

IMPZR
RZONE
LZONE
IFLWR
GFLWR

SACRAMENTO

PCTIM
ADIMC
ADIMP
UZTWC
UZFWC
LZTWC
LZFSC
LZPC

TOPMODEL

SOILR
GWATR

ARNO/VIC

IZONE
LAYR1
LAYR2
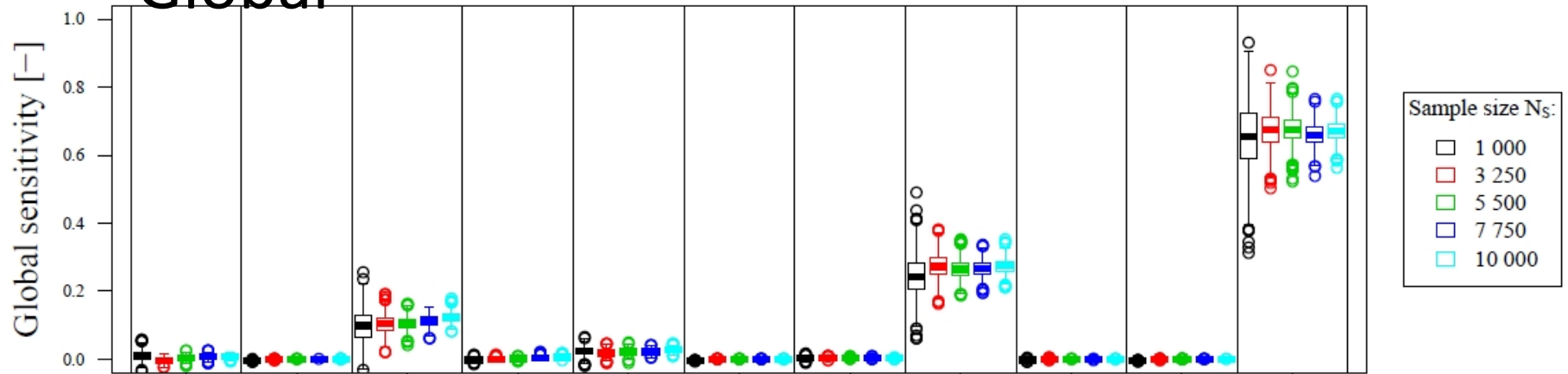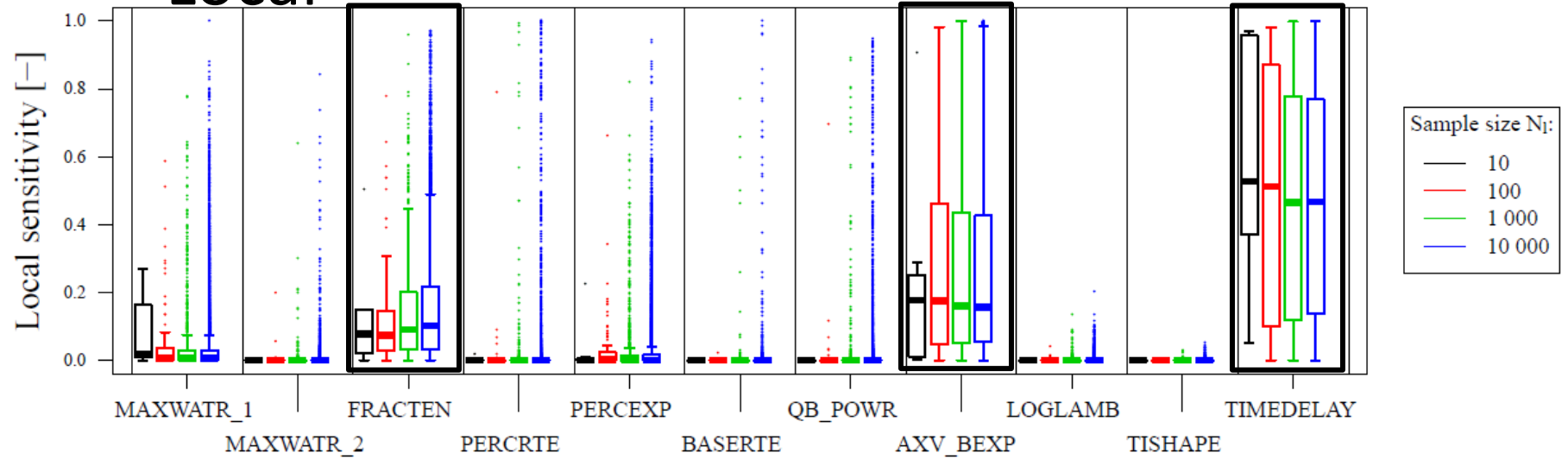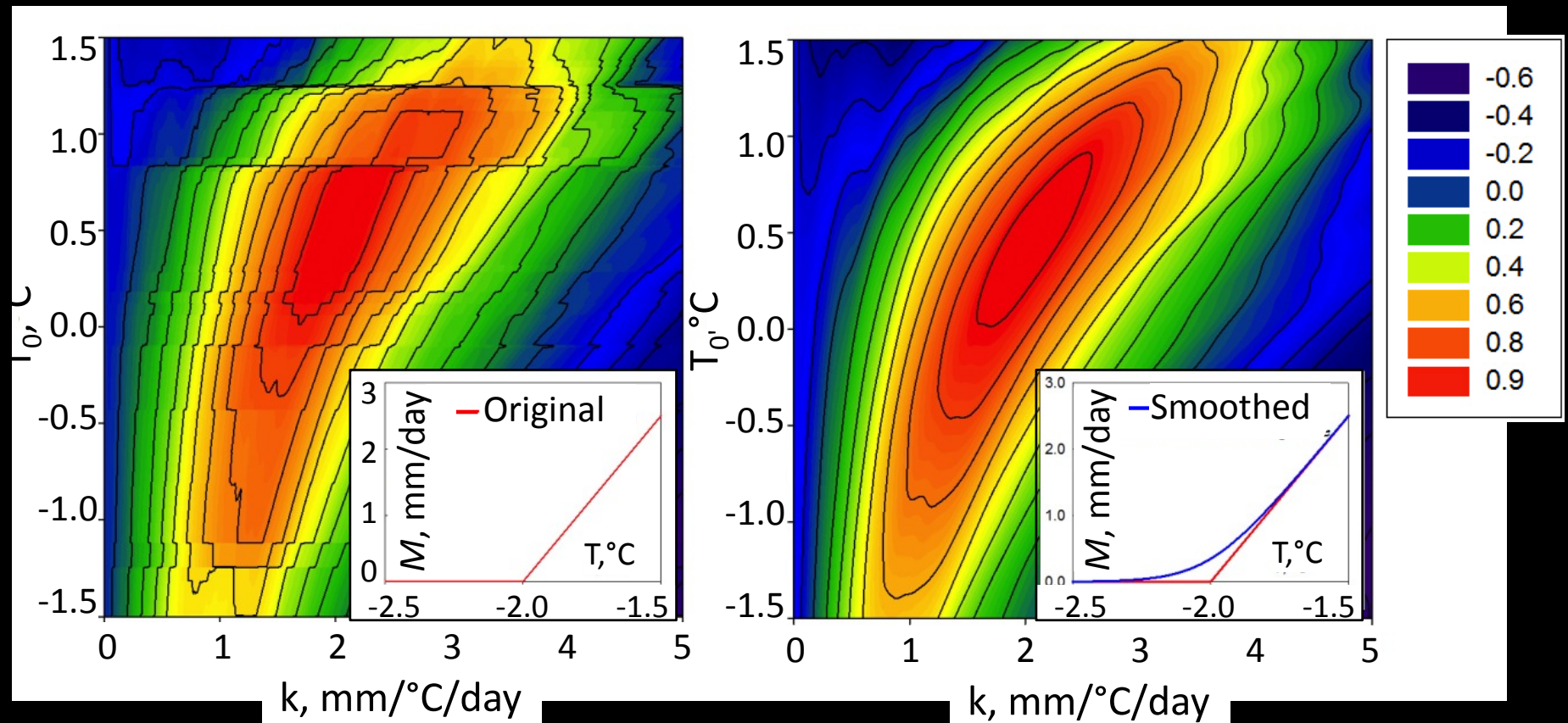
Total sensitivity results from FUSE-016

"Numerical Daemons"
(Clark, Kavetski, 2010; Kavetski and Clark, 2010, WRR)
produced by thresholds can produce large local sensitivities

Kavetski, Kuczera 2007 WRR

Organizing computationally frugal and demanding methods by purpose

Model Adequacy

- **How to include many data types with variable quality?** Error-based weighting and SOO, MOO*
- **Is model misfit/overfit a problem? Are prior knowledge and data subsets inconsistent?** Variance of weight-standardized residuals, residual graphs and space/time plots, MOO*
- **How nonlinear is the problem?** Modified Beale's measure, Explore objective function*, TSDE*

Sensitivity and Uncertainty

Observations ⟷ Parameters

- **What parameters can be estimated with the observations?** b/SD_b, CSS&PCC, SV, DoE*, IR*, First and total-order effects from OAT*, eFAST, Sobol', RSA* (First-order only)
- **Which observations are important and unimportant to parameters?** Leverage, Cook's D, CV*, MOO*
- **Are any parameters dominated by one observation and, thus, its error?** Leverage, DFBETAS, CV*
- **How certain are the parameter values?** b/SD_b, Parameter uncertainty intervals[#]

Parameters ⟷ Predictions

- **Which parameters are important and unimportant to predictions?** PSS, FAST*
- **How certain are the predictions?** z/SD_z, Prediction uncertainty intervals[#], MMA*
- **Which parameters contribute most and least to prediction uncertainty?** PPR, FAST*, Sobol',* MCMC*

Observations ⟷ Predictions

- **Which existing and potential observations are important to the predictions?** OPR, CV*
- **Which models in MMA are likely to produce the best predictions?** For individual model evaluations: AIC, AICc, BIC, KIC, CV*

TOWARD TRANSPARENT, REFUTABLE HYDROLOGIC MODELS IN KANSAS OR OZ

In class:
  Mary C. Hill   U.S. Geological Survey, Boulder
  Olda Rakovec   Wageningen University, Netherlands
Other co-authors:
  Martyn Clark   UCAR, Boulder
  A. H. Weerts   Deltares, Wageningin University
  A. J. Teuling   Wageningin University
  R. Uijlenhoet   Wageningin University