MESH Model Calibration Results for IP3 basins: Scotty, Wolf and Reynolds

IP3 Whitehorse Workshop

Nov. 14, 2008

Bryan Tolson

Angela MacLean

Frank Seglenieks

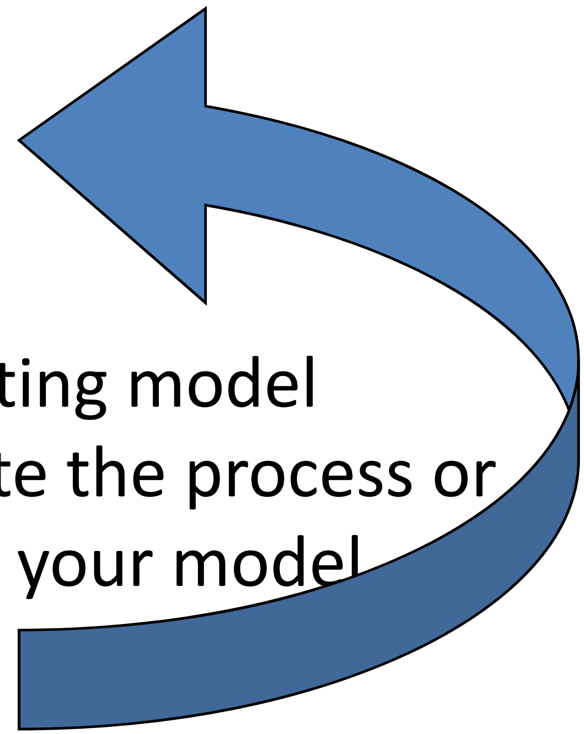
University of Waterloo

Model "Calibration"

→ manual calibration

→ automatic calibration approach using an optimization algorithm

- **DDS algorithm**
- If you are a modeller …
  - When you tire of manually adjusting model parameters and want to automate the process or you just want to know how good your model could do, get in touch with me



Hydrological Prediction at Regional Scale

- Focus on model development for larger basins
  - Scotty: 202 km^2
  - Wolf: 185 km^2
  - Reynolds: 229 km^2
- Ultimate goal is to make Spatially distributed model predictions with MESH 1.2
  - MESH uses CLASS 3.4*+ WATDRAIN + WATROUTE
- Evaluate quality of model predictions by comparing to monitoring data … improve if necessary

Hydrological Prediction at Regional Scale

- STEP 1: Model development for each basin
  - spatial discretization → GRU definition
  - input forcing data (e.g. rainfall)
- STEP 2: Model prediction quality assessment
  - Initial primary focus of performance assessment is on basin outlet streamflow
  - Ultimately we also want to simultaneously focus on:
    - smaller subbasin streamflows
    - Snow Water Equivalent, Snow Covered Area etc.
- STEP 3: Model calibration to improve prediction quality (via optimization)

STEP 1: Model development for each basin

- How do we discretize the spatial domain and define the <u>computational units</u> used in MESH
- Spatial discretization is non-trivial task and critically important … data to do so is not as readily available as say streamflow data
- Initial approach was to develop the simplest (most unrealistic) basin models *without spatial data*:

    1 grid – 1 GRU models which serve as baseline

STEP 1: Model development for each basin … 1 grid – 1 GRU models

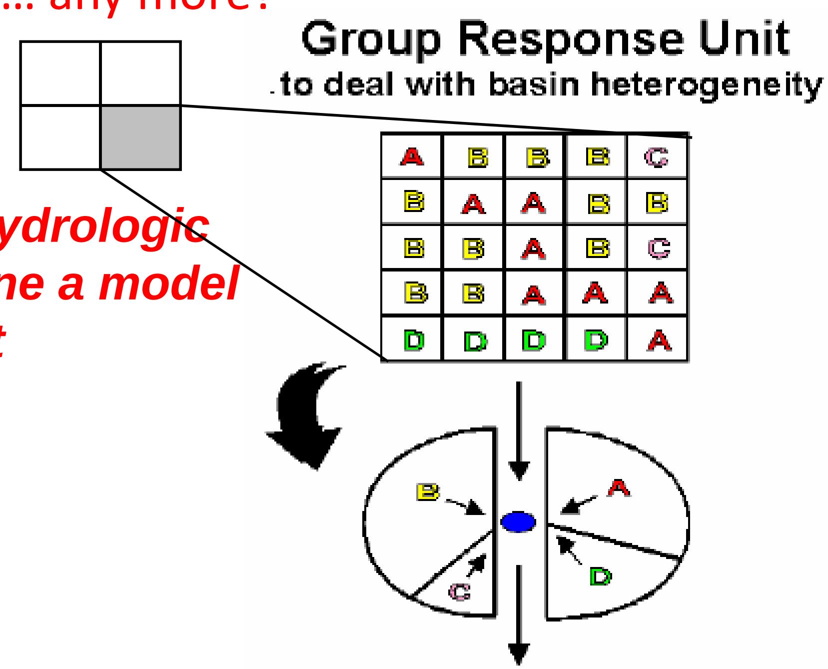
- GRU is not necessarily the computational unit in MESH …
  - each GRU can be divided into 4 possible structural vegetation types
- What we actually have done is define 1 grid – 1 GRU models with 2 vegetation structural classes as follows:
  - 70% of Scotty Creek is defined as Needleleaf vegetation type, 30% of each is Crop vegetation type … Bill?
  - same breakdown of Wolf and Reynolds
  - we are sort of modelling an aggregate vegetation type
- Another discretization strategy (b) would have been to define a 1 grid – **2 GRU model** with each GRU being 100% of a structural vegetation type
- Which strategy should be used?
- Do they generate different predictions?

Spatial Discretization "Terminology" in Modelling

- Grids
- Subgrids
- Grouped Response Unit (GRU)
- Hydrologic Response Unit (HRU)
- Tile
- Inter-landscape unit
- Inter-grid square
- "Mosaic" tiles
- Land class
- Landscape Unit
- Hydrologic Element

- Subarea
- Representative basins
- HRU structure
- Basin element
- Computational unit
- … any more?

areas of uniform hydrologic response that define a model computational unit



Completing STEP 1 for spatially distributed MESH model

- It has become clear that Angela, Frank and I do not have time to do this STEP (process spatial data to come up with land uses, GRU types etc.)
- We will not likely do it properly as we are not in the basin
- **We need lots of help to do this**
- Each basin investigator is best suited to guide the definition of GRUs
  - should closely consider approach by Pablo/Al/John in Wolf
- My interest is in making spatially distributed versions of MESH predict all measured calibration as best as possible in the most time efficient way possible
  - we really can't set up distributed models AND do this

… let me describe the calibration experiments I have and want to run

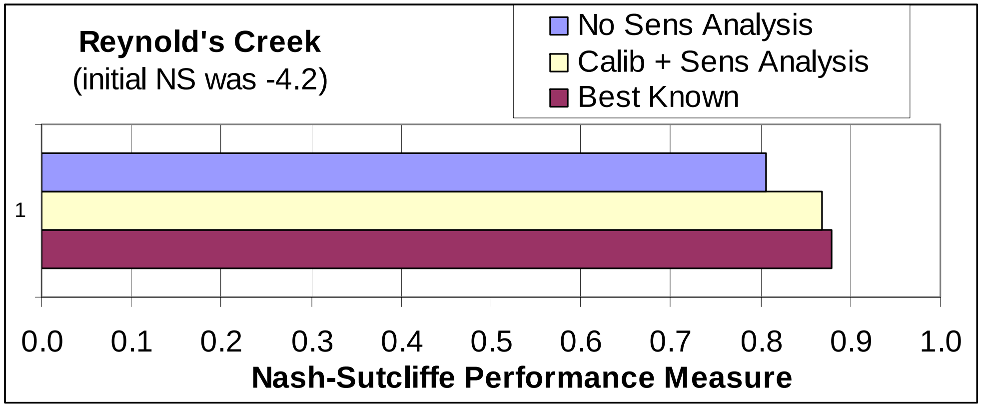
Calibration Experiments

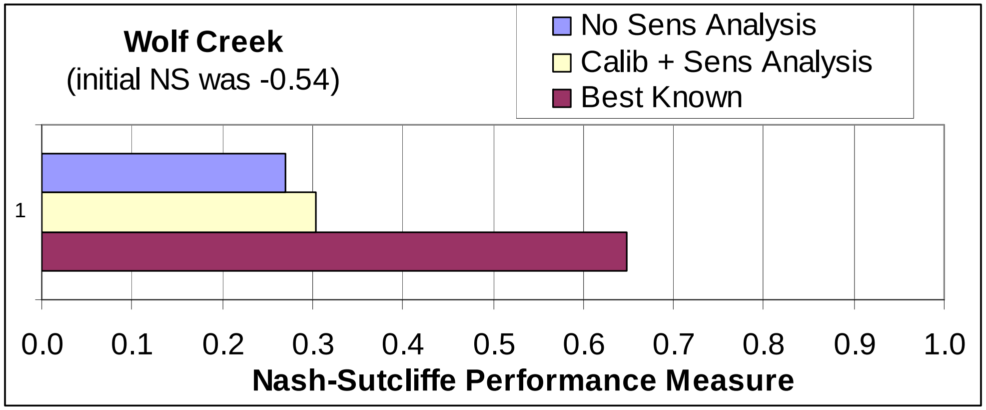
- **When faced with a new model calibration problem and a defined computational budget, how should one go about performing calibration?**
  - **e.g. MESH has 50+ model parameters**
- Strategy 1 (No Sens. Analysis):
  - calibrate all MESH parameters at once with entire computational budget.
- Strategy 2 (Calib. + Sens. Analysis):
  - briefly calibrate all MESH parameters to get to a decent solution
  - repeat 2-3 times to arrive at multiple solutions
  - roughly determine insensitive parameters based on sensitivity measured around the 2-3 decent solutions
  - continue calibration of only the sensitive parameters from best current solution
- Strategy 3: suggestions from model developers welcome … e.g. only calibrate these parameters …

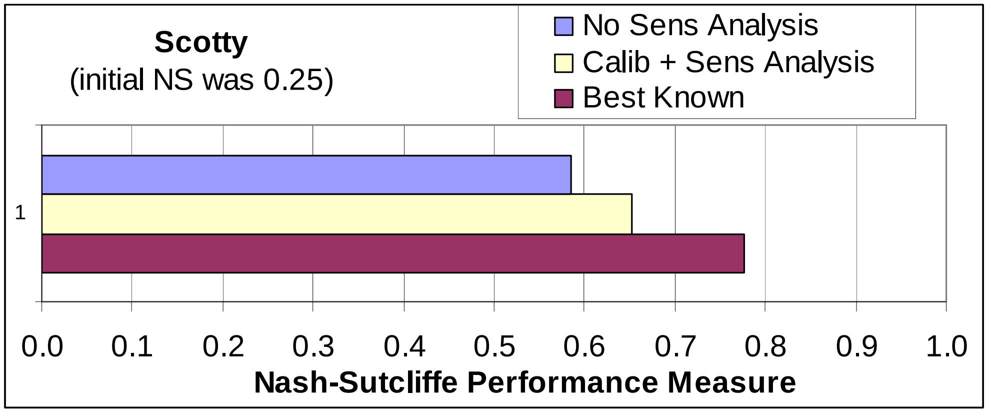
Calibration Experiments

- Compare strategies under the following conditions:
  - Use 1-grid, 1-GRU models of Scotty, Wolf and Reynolds
  - use computational budget of 4000 model simulations
  - calibrate to basin outlet streamflows only

Results of Calibration Experiments







Initial NS based on initial*/default parameter set …

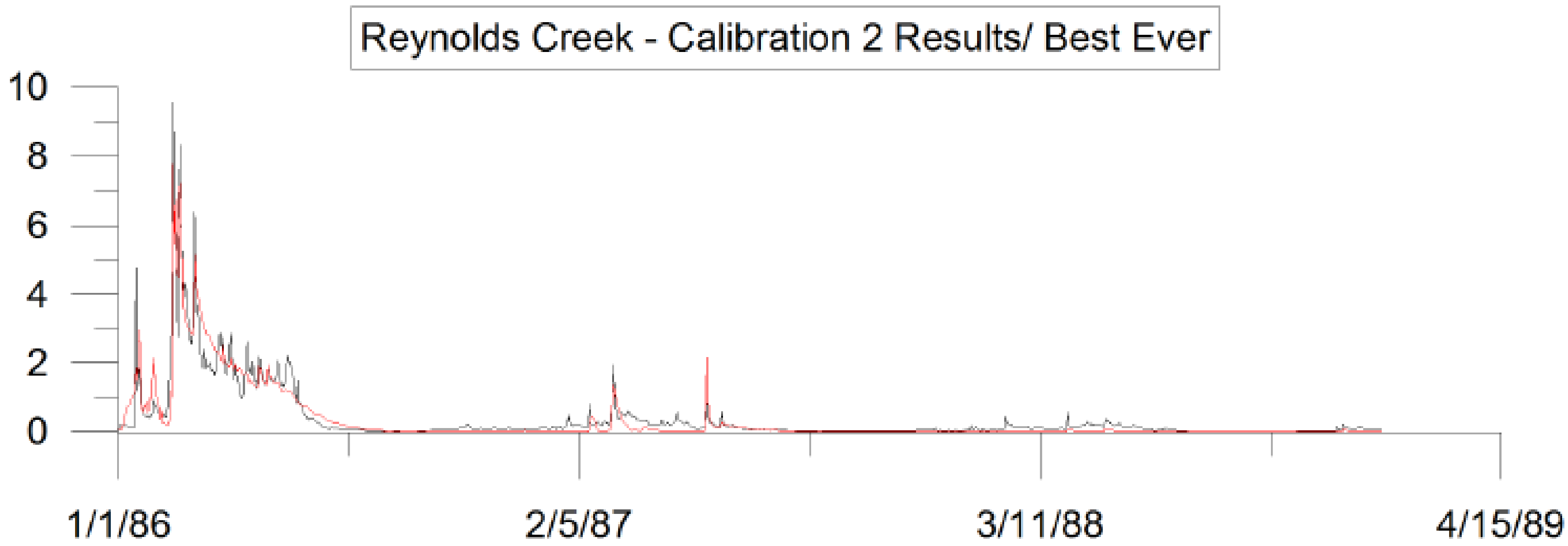
1. MESH requires calibration to achieve good performance.

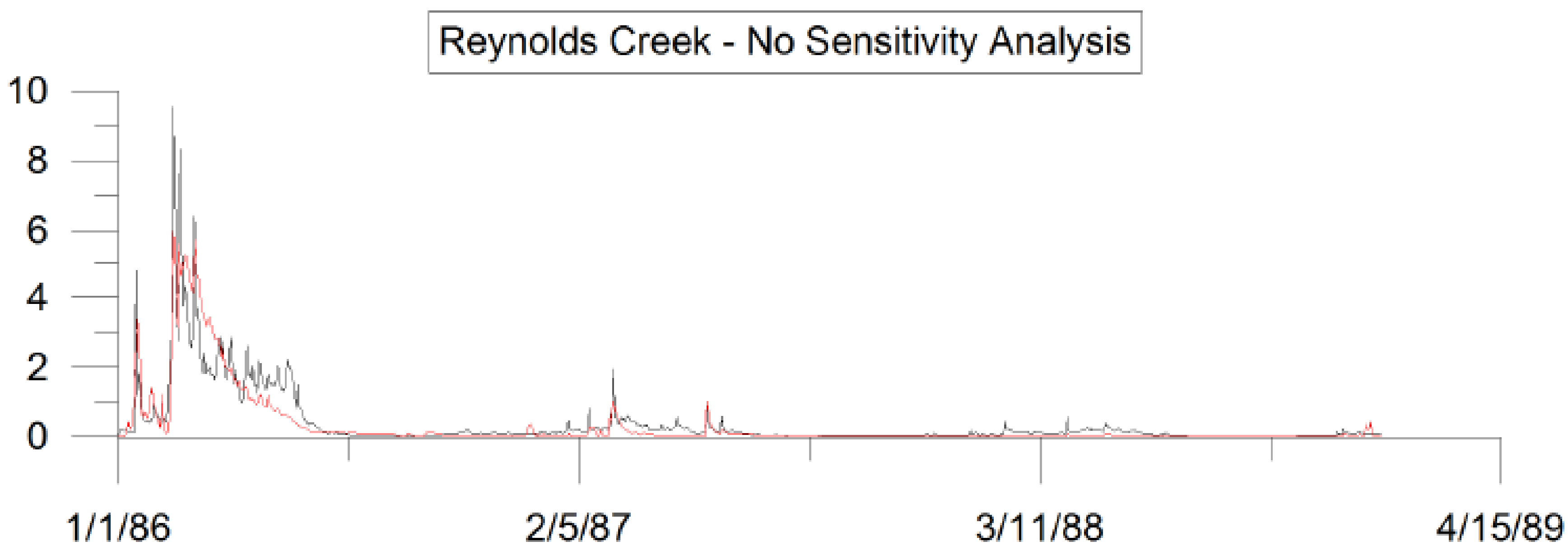
Best known based on all optimization runs ever conducted (100,000+).

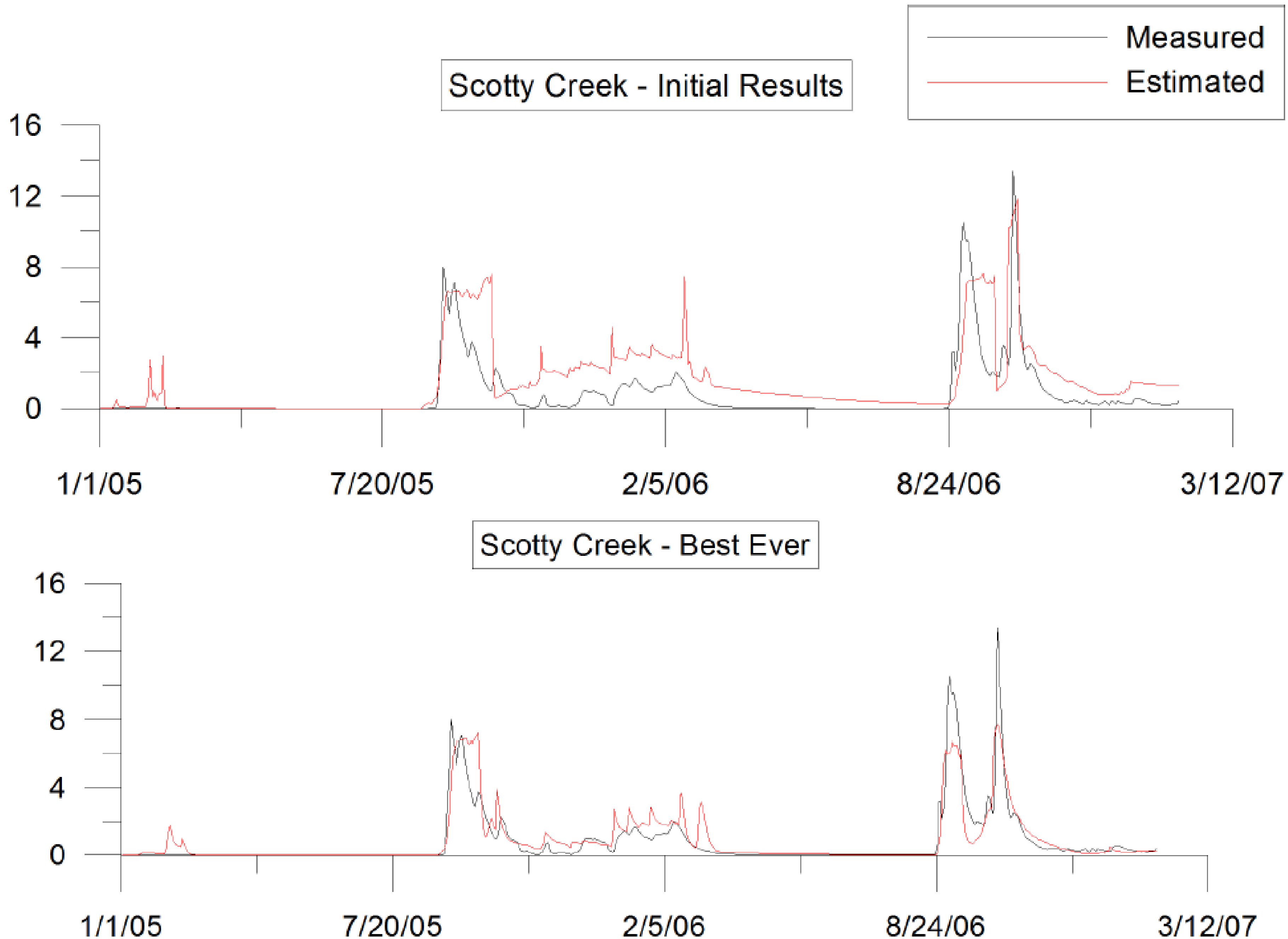
2. Calibration may require much more than 4000 model evaluations.

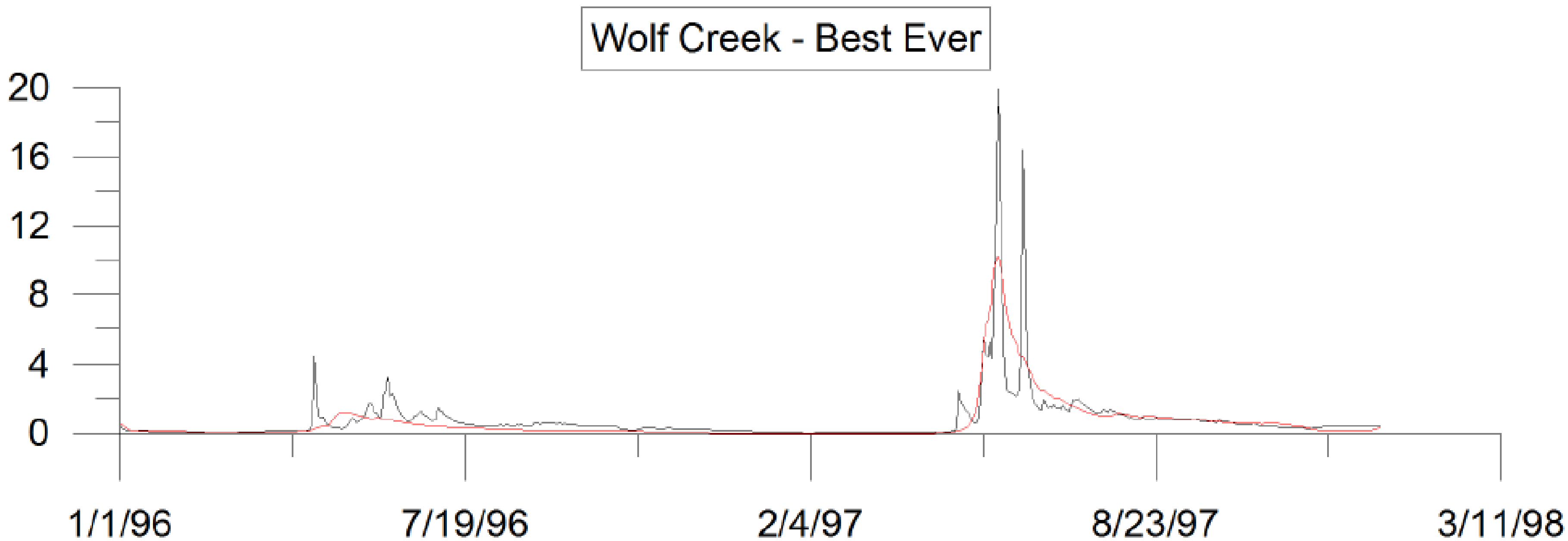
3. Brute force approach to calibration (optimize all, no sens analysis) can be improved upon … somewhat

- unclear whether sens. analysis responsible for this

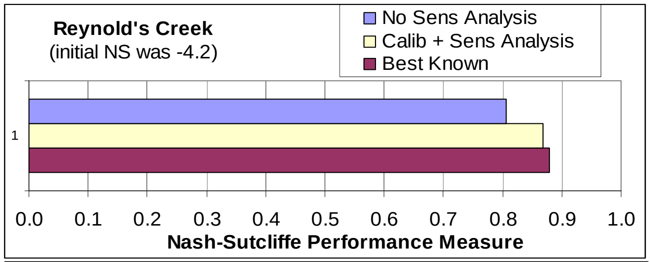

Reynolds Creek - Calibration 2 Results/ Best Ever
10
8
6
4
2
0
1/1/86
2/5/87
3/11/88
4/15/89

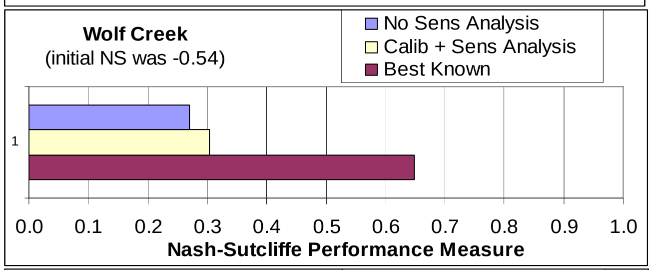

Reynolds Creek - No Sensitivity Analysis
10
8
6
4
2
0
1/1/86
2/5/87
3/11/88
4/15/89

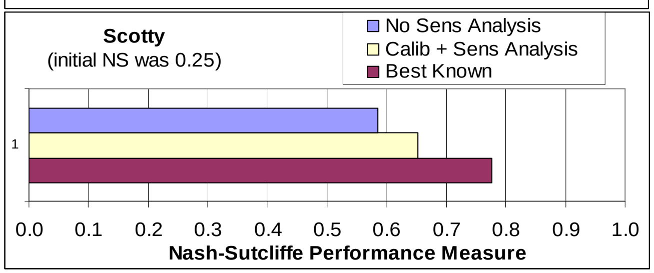

Measured
Estimated
Scotty Creek - Initial Results
16
12
8
4
0
1/1/05
7/20/05
2/5/06
8/24/06
3/12/07
Scotty Creek - Best Ever
16
12
8
4
0
1/1/05
7/20/05
2/5/06
8/24/06
3/12/07


Wolf Creek - Best Ever
20
16
12
8
4
0
1/1/96
7/19/96
2/4/97
8/23/97
3/11/98

Future Calibration Experiments







1. **Repeat this with ideal spatial discretization**
2. **Consider more calibration strategies.**
3. **Compare model performance under alternative spatial discretization strategies**

   Automatic calibration or objective calibration framework is critical to any such comparisons. How can we compare (a) and (b) below?

   a) Calibrated1-grid, 1 GRU model

   b) sort of calibrated 3x3 km grid, 5 GRU model

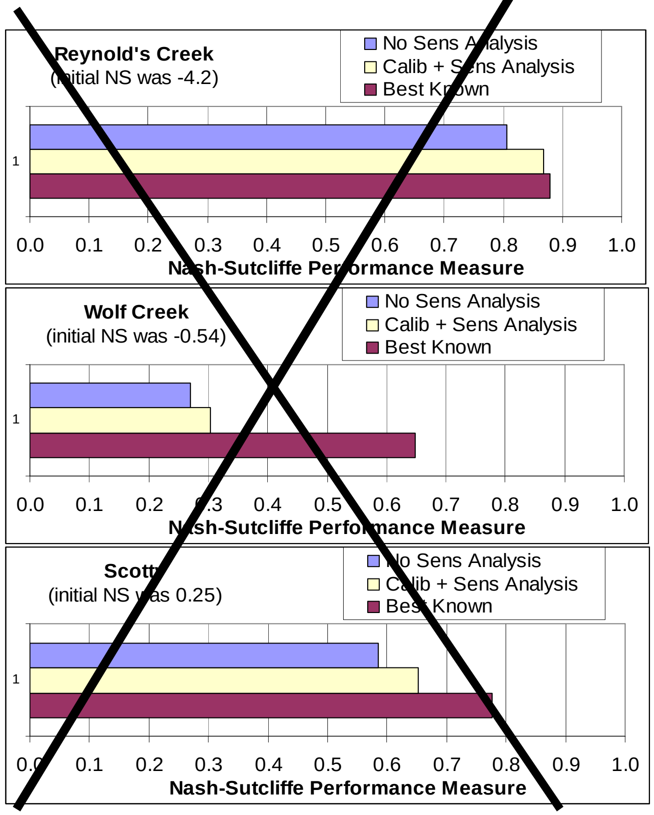
Benefits to the rest of IP3 researchers:

- you can use the best known calibrated models

Benefits to scientific community at end of IP3:

- Guidance on ideal MESH development and calibration strategy available for new basins where the model is to be calibrated

Future Calibration Experiments



4. **Repeat or simply conduct experiments 1-3 within a multi-objective calibration framework:**

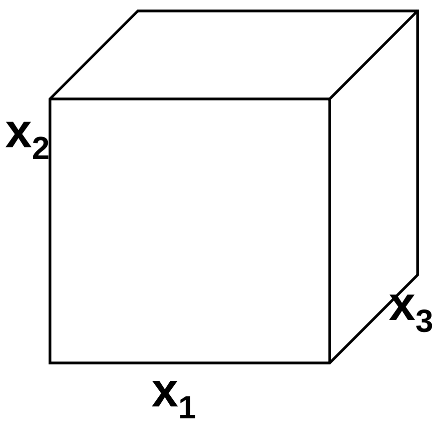
   - we want MESH to predict streamflow, SWE, SCA, sensible heat flux, and any other monitoring data all of you are collecting in the field

   - optimal streamflow prediction parameter set will not likely be optimal for sensible heat flux so we really want a parameter set that predicts both in an adequate (not optimal) way

   - multi-objective calibration is **more difficult** than single objective calibration

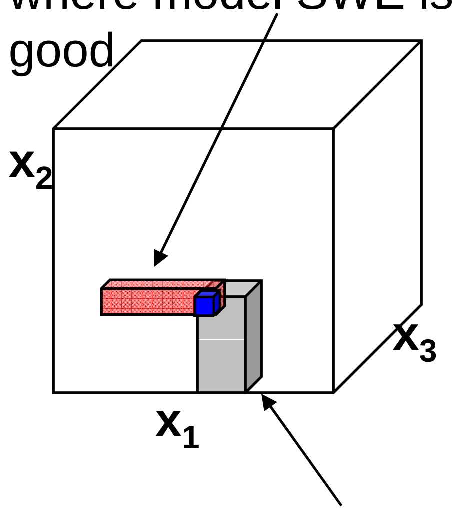
Difficulty of Multi-Objective Calibration:

Consider Q & SWE as objectives, 3 calibration parameters



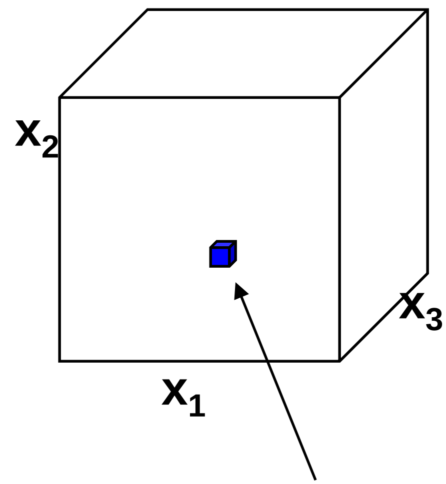
Parameter space hypercube in 3-D … parameter values constrained within this cube

Calibrate to SWE: Parameter region where model SWE is good



Calibrate to Q:

Parameter region where model streamflow is good

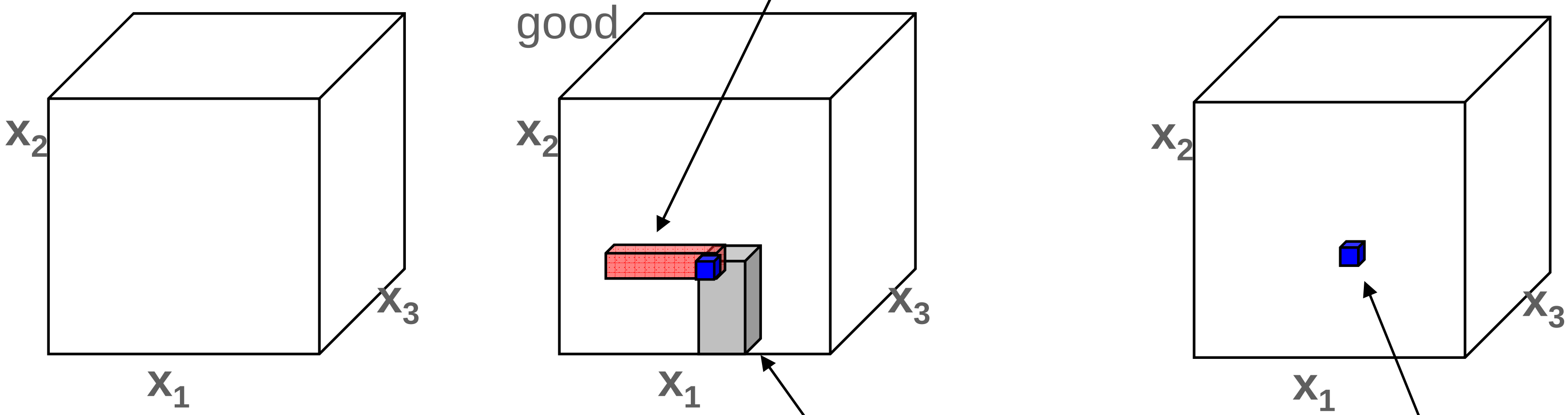


Calibrate to both: Parameter region where both streamflow and SWE are good

What happens when we have another objective like SCA?

Importance of Multi-Objective Calibration

Calibrate to SWE: Parameter region where model SWE is good



Parameter space hypercube in 3-D … parameter values constrained within this cube

Calibrate to Q:

Parameter region where model streamflow is good

Calibrate to both: Parameter region where both streamflow and SWE are good

Parameter sets in the blue cube should be more transferable to ungauged basins

Next Steps

- Who in IP3 are Angela, Frank and I working with to develop spatially distributed MESH models?
  - we have done our own spatial data analysis and GRU definition of Reynolds
  - others might be:
    - Phil Marsh for TVC (sensible heat flux at the grid square measured)?
    - Chris Spence for Baker?
    - ?