An Improved Framework for Watershed Discretization and Model Calibration: Application to the Great Lakes Basin

Amin Haghnegahdar (PhD Candidate)

Bryan A. Tolson (Associate Professor)

James R. Craig (Associate Professor)

EC, Dorval, 19 April 2013

civil AND environmental ENGINEERING

University of Waterloo

Acknowledgements

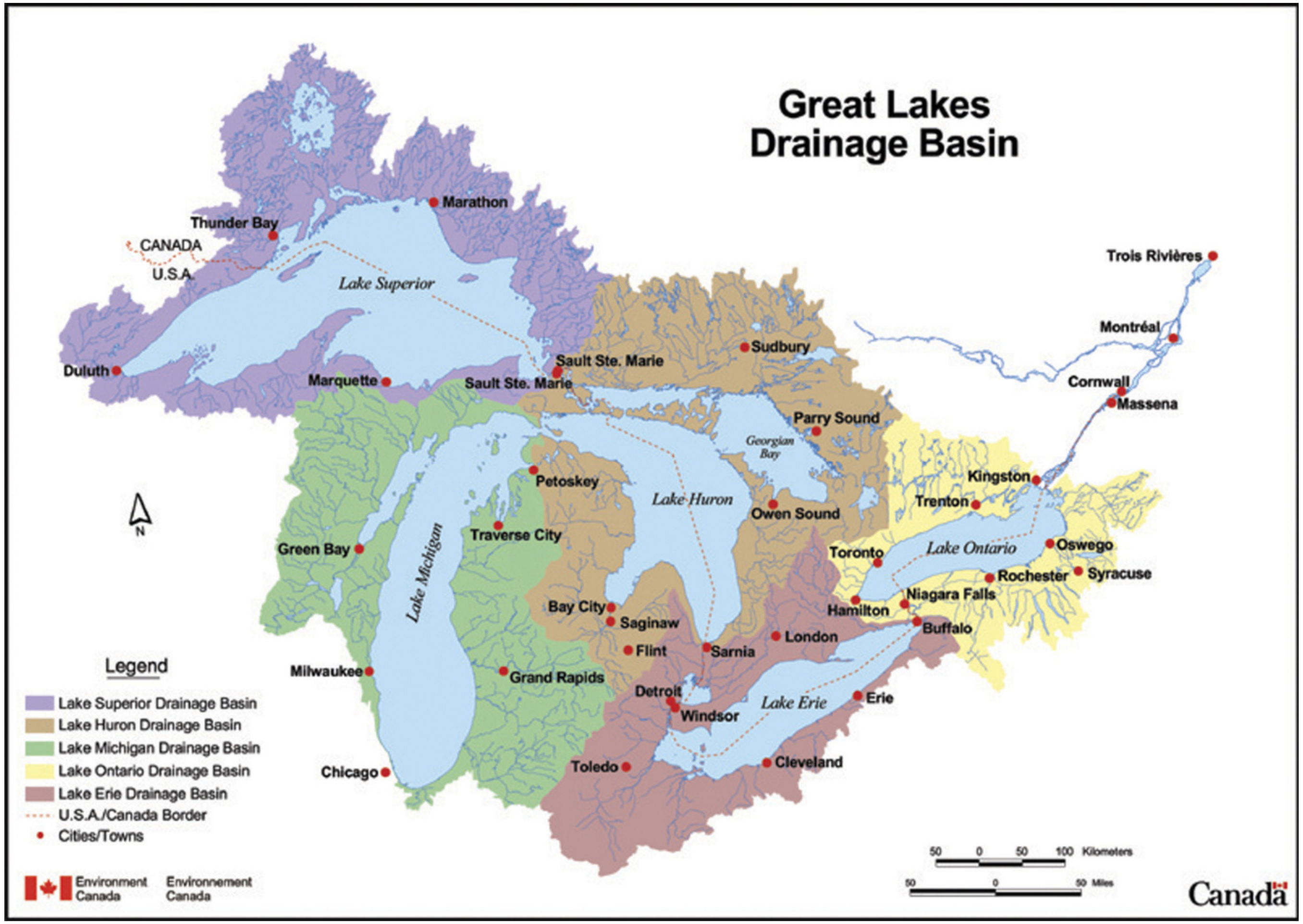
- G & C support from Environment Canada over the past 3 years.
- EC personnel: V. Fortin, B. Davison, F. Seglenieks, E. Klyszejko, P. Pellerin
- Other contributors: Dr. E. D. Soulis, Dr. L. S. Matott, K. Paya, C. Werstruck

Goals and Outline

1. Report on key findings from 3 yrs of G & C Grant support
   – Calibration of stand-alone MESH (also WATFLOOD)
   – Baseline Calibration Results for MESH
   – Comparing 2 discretization schemes:
     • 1 GRU per grid vs. 7 GRUs per grid
   – A new framework to assess discretization decisions in the context of modelling ungauged basins
2. Highlight aspects of the work that *other* EC modellers and scientists may find useful
3. Future plans

The Great Lakes Basin

- Nearly 1 million km^2
- Largest group of fresh surface water on Earth
- 18% of the world supply of fresh surface water
- 84% of North America's supply



http://www.ec.gc.ca/grandslacs-greatlakes/default.asp?lang=En&n=03B3F448

MESH Modeling System for the Great Lakes Basin

- 1st MESH application to Great Lakes Basin was Pietroniro et al. (2007)

- MESH = WATFLOOD + CLASS
  (versions 1.3.003, 1.3.005)

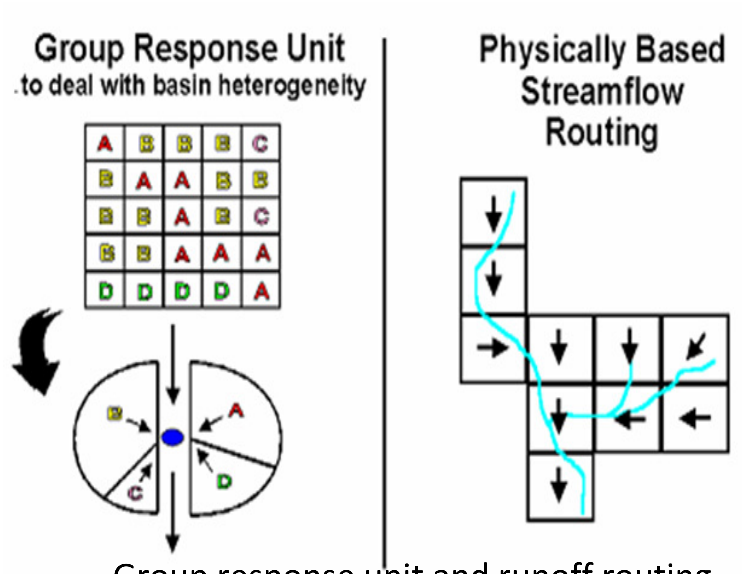
•1/6th degree grid cells (~ 15 x 15 km)

•7 GRUs defined by 7 land classes

Crop, Grass, Deciduous forest

Coniferous forest, Mixed forest

Water and Impervious



Group response unit and runoff routing concept (WATFLOOD Manual)

•Each GRU modelled as a single land cover class

•**Infiltration:** complex multi-layer Green-Ampt "type " formulation

•**Interflow**: estimated from the bulk saturation of each soil layer

•**Baseflow:** Simply is the water percolating out of the bottom of the soil column

•**Routing**: between the grid cells and across the river network for the entire basin

Research Involved Repeated Model Calibration Experiments

- Our focus was only on streamflow simulation only
- Calibration accomplished via automatic calibration (optimization algorithm)
- Core algorithm was the DDS algorithm (Tolson & Shoemaker, 2007)
- DDS and MESH model communication accomplished via Ostrich software developed by Dr. L.S. Matott

- Calibration experiments involved:
  - estimating 11 to 71 model parameters (simultaneously)
  - Evaluating 250 to 5000 model simulations per experiment
  - 1000 model simulations in our experiments requires **18 days of simulation**

DDS- Dynamically Dimensioned Search

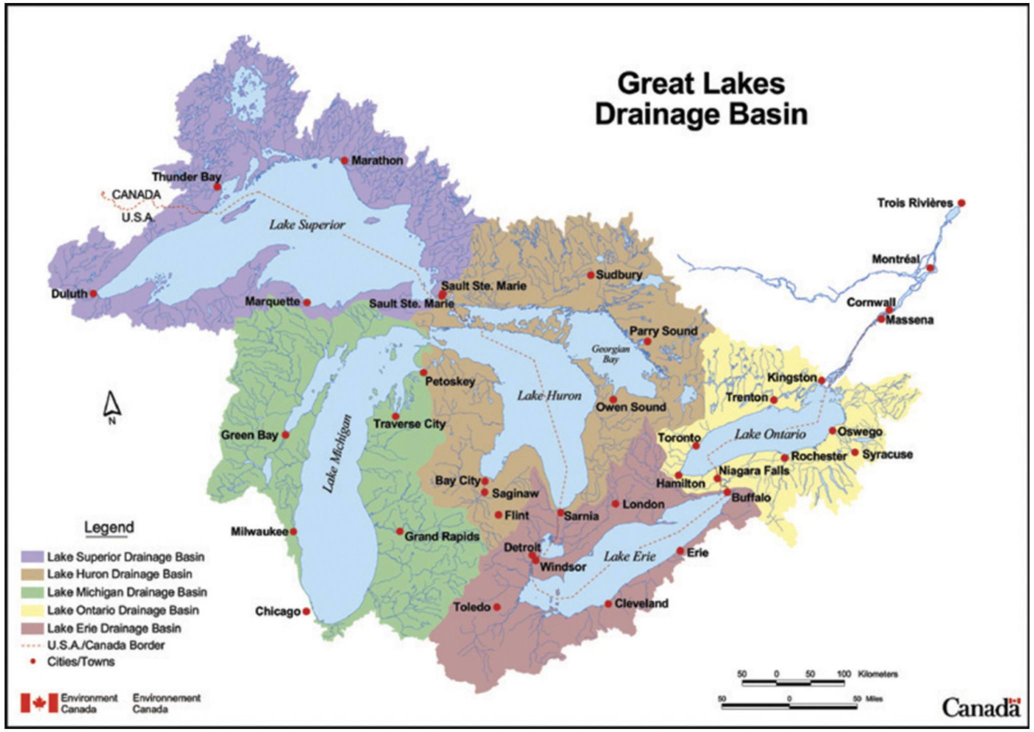
- Tolson & Shoemaker (2007)
- A heuristic stochastic global optimization algorithm
- Originally designed for efficient automatic calibration of environmental models
  - Simple to implement
    - Only one algorithm parameter (neighborhood size perturbation, *r*)
    - Normally no parameter tuning required, default $r=0.2$
- Generates good (not best) results in modeller's time frame

OSTRICH - Optimization Software Toolkit for Research Involving Computational Heuristics

- Dr. L. Shawn Matott (2005)
- A model-independent calibration and optimization tool
- Embeds a number of popular optimization algorithms (DDS, GA, gradient-based)
- Available for both Linux and Windows platforms
- Very easy to use for model calibration
- Embedded within the U.S. EPA (Environmental Protection Agency) FRAMES (Framework for Risk Analysis of Multimedia Environmental Systems) modeling system (http://www.epa.gov/athens/research/modeling/3mra.html)

Initial Calibration Experiments with WATFLOOD

- Helped us decide how to pose the calibration problem
  - Which gauging stations to use for calibration and validation?
  - How to transfer calibrated parameters to ungauged basins



Initial Calibration Experiments with WATFLOOD

- **71 parameters** for 7 land classes, 5 river classes:
  - Land class parameters: e.g. upper zone specific retention, the reduction in soil evaporation due to tall vegetation
  - Crops, Grass, deciduous, conifer, Mixed, Water, Impervious
  - River class parameters: e.g. lower zone drainage function, channel roughness
- 1000 model evaluations for calibration
- Calibration period : Jun 04- Sep 06 (27 months)
  - Spin-up period: first 4 months

Calibration objective function:

$$NS = 1 - \frac{\sum(Q_{obs} - Q_{sim})^2}{\sum(Q_{obs} - \overline{Q_{obs}})^2}$$

$$-\infty < NS \leq 1$$

Best performance

WATFLOOD Calibration: Summary of Results

Table 3- **Comparison of the NS values** overall and for the 5 selected stations in different cases of the 2 new experiments (calibration period is June 04-September 06 for all cases)

Station	Original Simulation			Individually Calibrated
BLACK RIVER NEAR WASHAGO (02EC002)	-1.57			0.72
GOULAIS RIVER NEAR SEARCHMONT (02BF002)	-1.76			0.75
PIC RIVER NEAR MARATHON (02BB003)	-0.35			0.63
NITH RIVER NEAR CANNING (02GA010)	-0.11			0.47
TAHQUAMENON RIVER NEAR PARADISE (04045500)	-2.78			0.48
Average (weighted for unregulated stations)	-9.97			-
Max	0.43			-
Min	-194.37			-

WATFLOOD Calibration: Summary of Results

Table 3- **Comparison of the NS values** overall and for the 5 selected stations in different cases of the 2 new experiments (calibration period is June 04-September 06 for all cases)

Station	Original Simulation	All-station Calibration (Not regulated, 133)		Individually Calibrated
BLACK RIVER NEAR WASHAGO (02EC002)	-1.57	0.38		0.72
GOULAIS RIVER NEAR SEARCHMONT (02BF002)	-1.76	0.27		0.75
PIC RIVER NEAR MARATHON (02BB003)	-0.35	0.19		0.63
NITH RIVER NEAR CANNING (02GA010)	-0.11	-0.00		0.47
TAHQUAMENON RIVER NEAR PARADISE (04045500)	-2.78	0.21		0.48
Average (weighted for unregulated stations)	-9.97	-1.42		-
Max	0.43	0.48		-
Min	-194.37	-38.51		-

WATFLOOD Calibration: Summary of Results

Table 3- **Comparison of the NS values** overall and for the 5 selected stations in different cases of the 2 new experiments (calibration period is June 04-September 06 for all cases)

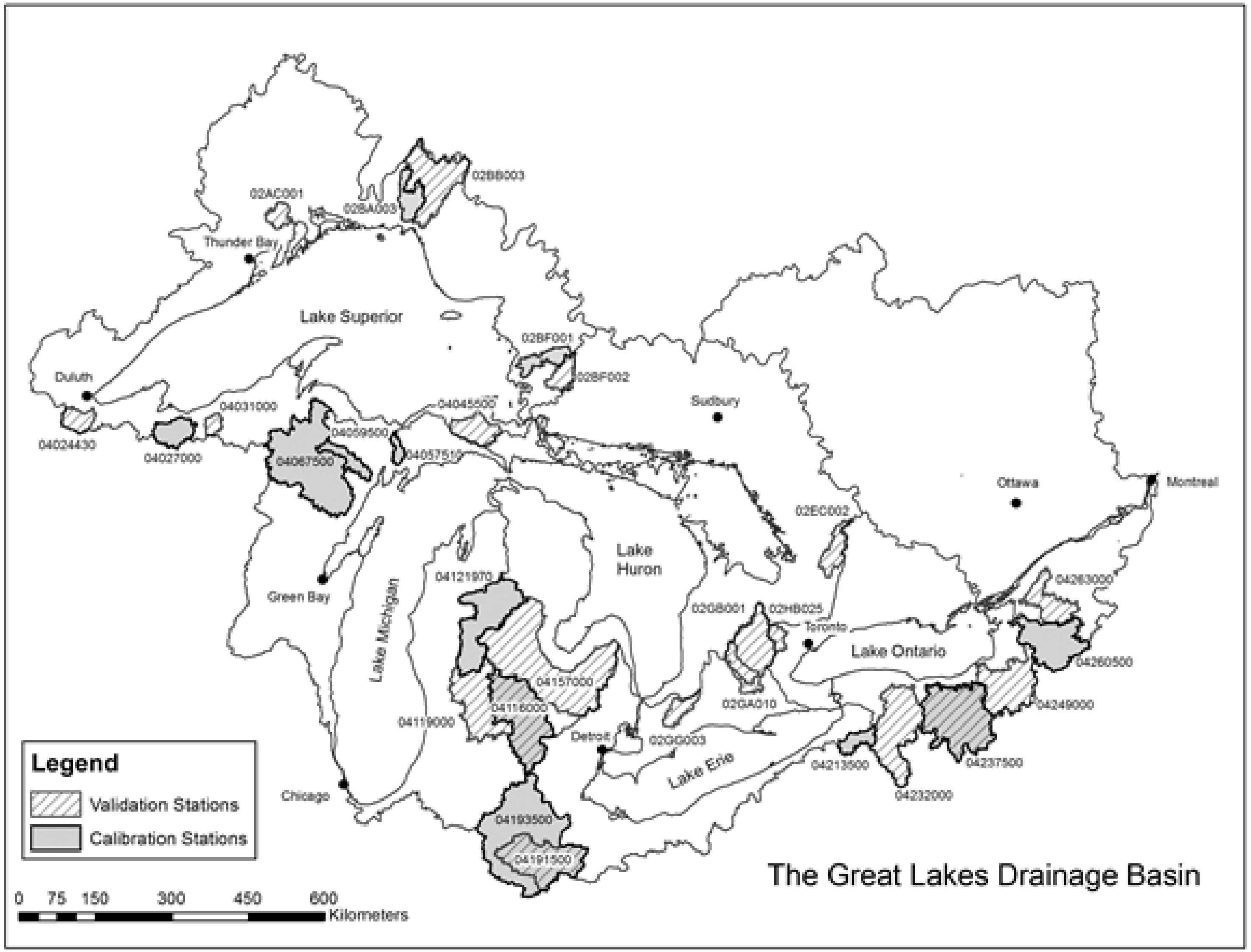
Station	Original Simulation	All-station Calibration (Not regulated, 133)	All-station Calibration (Not regulated/poor, 98)	Individually Calibrated
BLACK RIVER NEAR WASHAGO (02EC002)	-1.57	0.38	0.51	0.72
GOULAIS RIVER NEAR SEARCHMONT (02BF002)	-1.76	0.27	0.38	0.75
PIC RIVER NEAR MARATHON (02BB003)	-0.35	0.19	0.61	0.63
NITH RIVER NEAR CANNING (02GA010)	-0.11	-0.00	0.15	0.47
TAHQUAMENON RIVER NEAR PARADISE (04045500)	-2.78	0.21	0.45	0.48
Average (weighted for unregulated stations)	-9.97	-1.42	-0.01	-
Max	0.43	0.48	0.61	-
Min	-194.37	-38.51	-1.05	-

Experience Gained from WATFLOOD Calibration

- Model calibration is an absolute requirement!
- 71 parameters can be calibrated by DDS given a budget of only 1000 model evaluations (calibration result acceptable)
- Despite roughly 200 streamflow gauges, less than 100 are helpful from model calibration perspective – more care needed
- Global calibration scheme approaches best possible gauge by gauge calibration result

All of the above considered for MESH calibration experiments

Stand-Alone MESH Baseline Calibration: Calibration & Validation Stations



MESH Baseline Calibration: Calibration Strategy

- Calibration & Validation period:
  - Defined based on available forcing data (CaPA)
  - Calibration: October 2004 to September 2005
    - Model spin-up: 4 months (June 04 – Sep. 04)
  - Validation: October 2005 to June 2009
    - Also spatially for the full period at various stations

$$NS = 1 - \frac{\sum (Q_{obs} - Q_{sim})^2}{\sum (Q_{obs} - \overline{Q_{obs}})^2}$$

$$-\infty < NS \leq 1$$

Best performance

- Objective function:
  - Weighted Nash-Sutcliffe (NS) of daily flows
    - Weights based on average flow magnitudes
- Calibration parameters:
  - Selected & modified based on developers' opinion and past studies
  - 51 total parameters (independent):
    - Water and impervious parameters were not calibrated
    - 9 per 5 GRUs + 1 per 5 river roughness + 1 basin-wide parameter (soil depth)
- 1000 model runs for calibration
  - More than 18 days for 11 stations (~26 min/model simulation) per calibration
  - A series of calibration runs

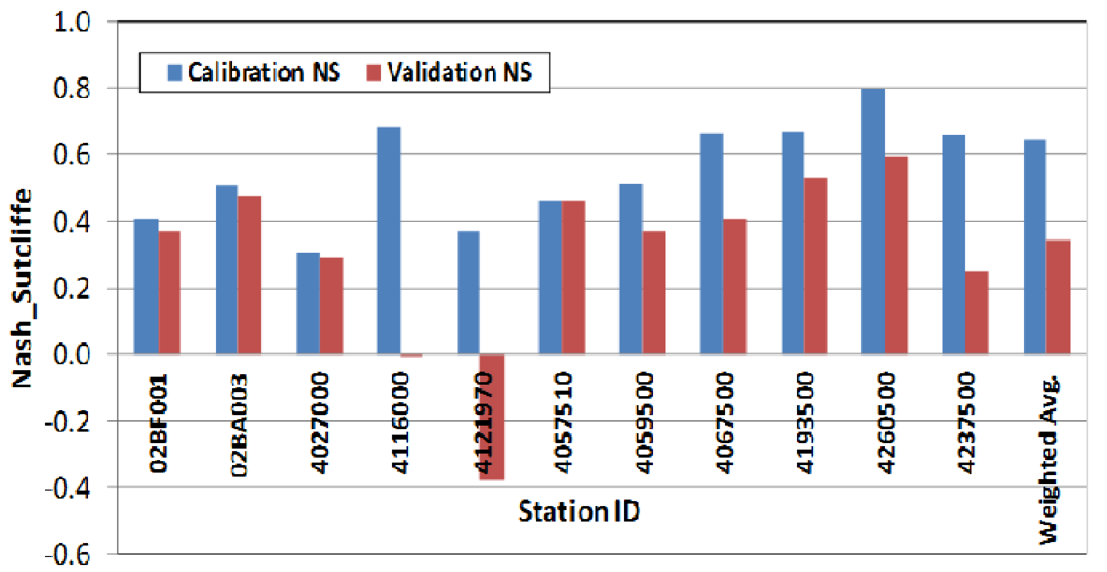
MESH Baseline Calibration: Parameters in Ungauged Basins

- GRU (HRU) approach natural way to assign model parameters in ungauged basins:
  - E.g., Calibrated LAI parameters for deciduous forest assigned to all deciduous forest GRUs

- Ungauged basin model preformance critical because approximately half of Great Lakes basin is ungauged

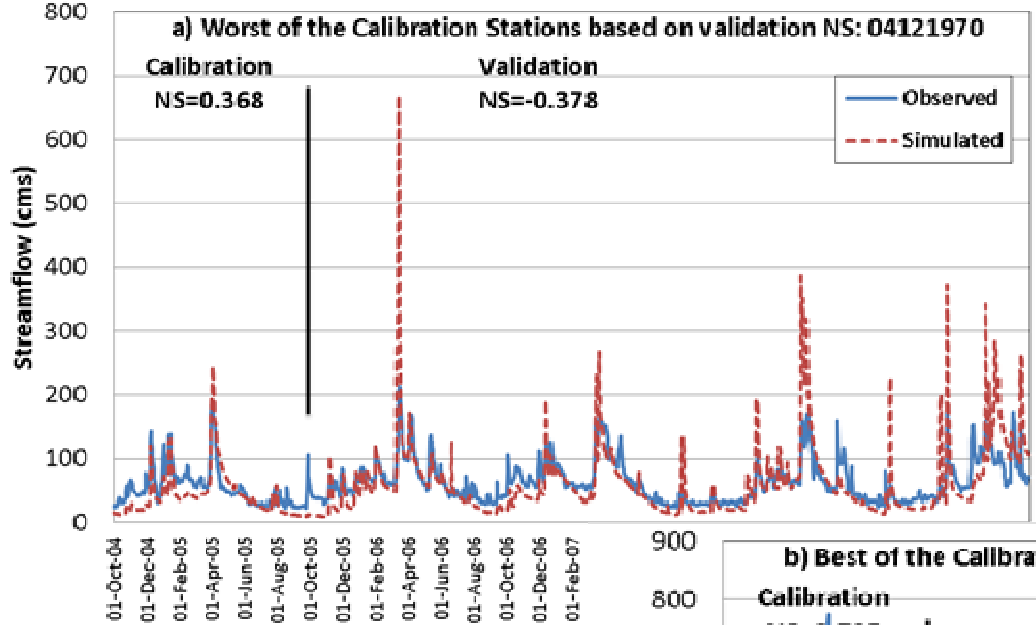
MESH Baseline Calibration Results: Calibration and *Temporal* Validation

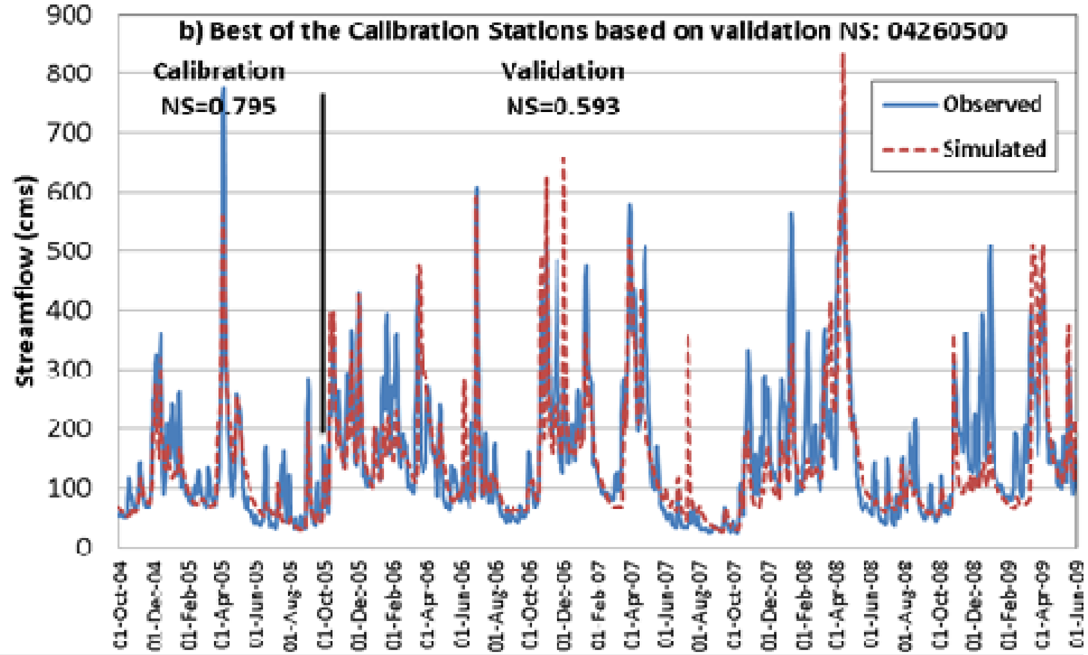
- Global Calibration:
  - Calibrating to a number of stations simultaneously and for all land-classes (GRUs)
    - 51 total parameters calibrated
    - 11 Calibration & 16 Validation stations



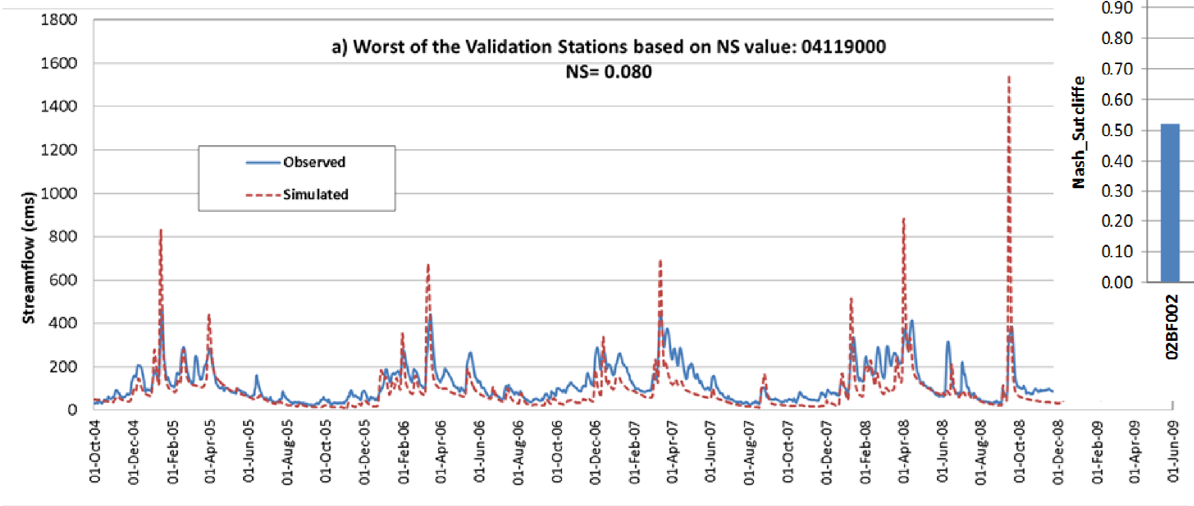
NS	Calib. Period	Valid. Period
Min	0.31	-0.38
Max	0.80	0.59
Median	0.51	0.37

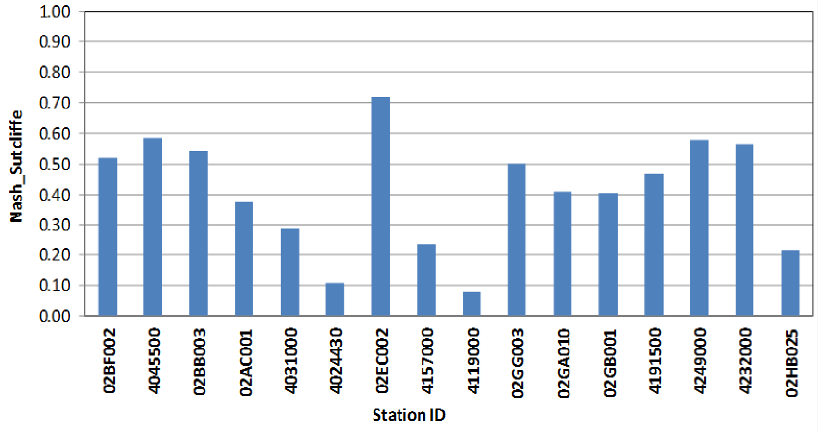
MESH Baseline Calibration Results: Calibration and *Temporal* Validation Hydrographs

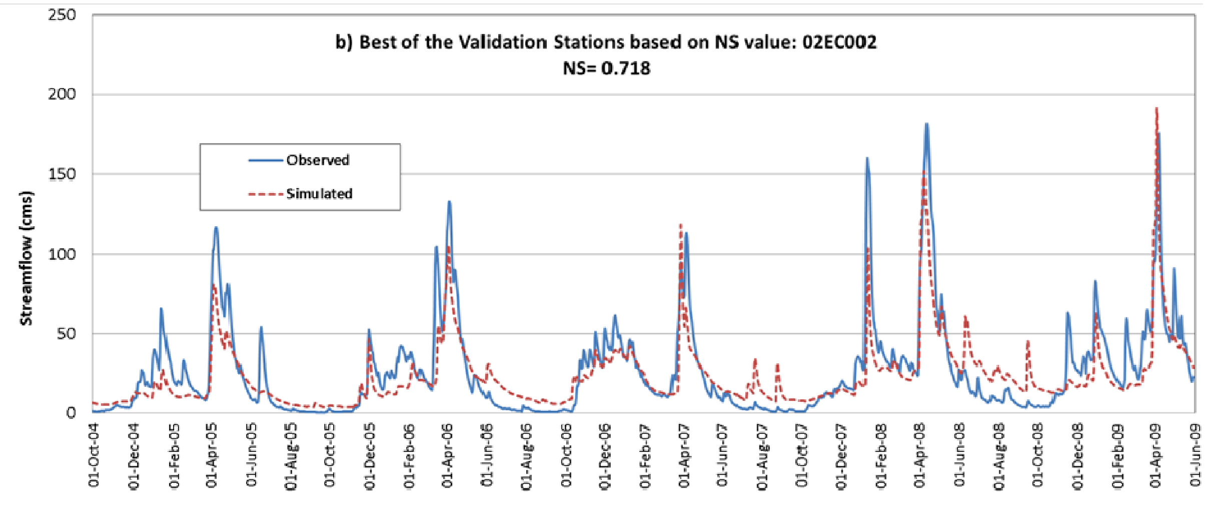




MESH Baseline Calibration Results:
Spatial Validation Performance





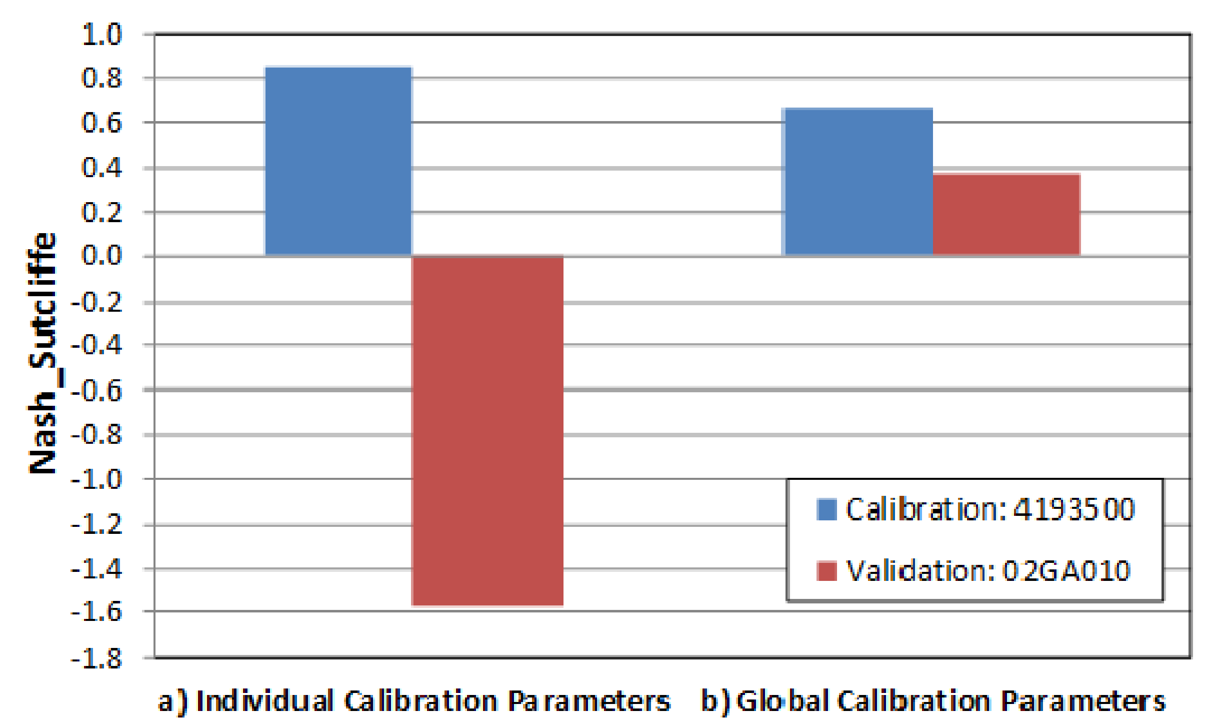


NS	Full Period
Min	0.08
Max	0.72
Median	0.44

MESH Baseline Calibration Results: Compare Against Alternative Strategy

- ***Individual*** Calibration Strategy:
  - Calibrating to a single sub-basin with one dominant land-class (GRU) and then transfer parameters to similar land-classes elsewhere
  - Reasonable approach, other MESH modellers have done this in the past
  - Tested against the global strategy for 2 GRUs

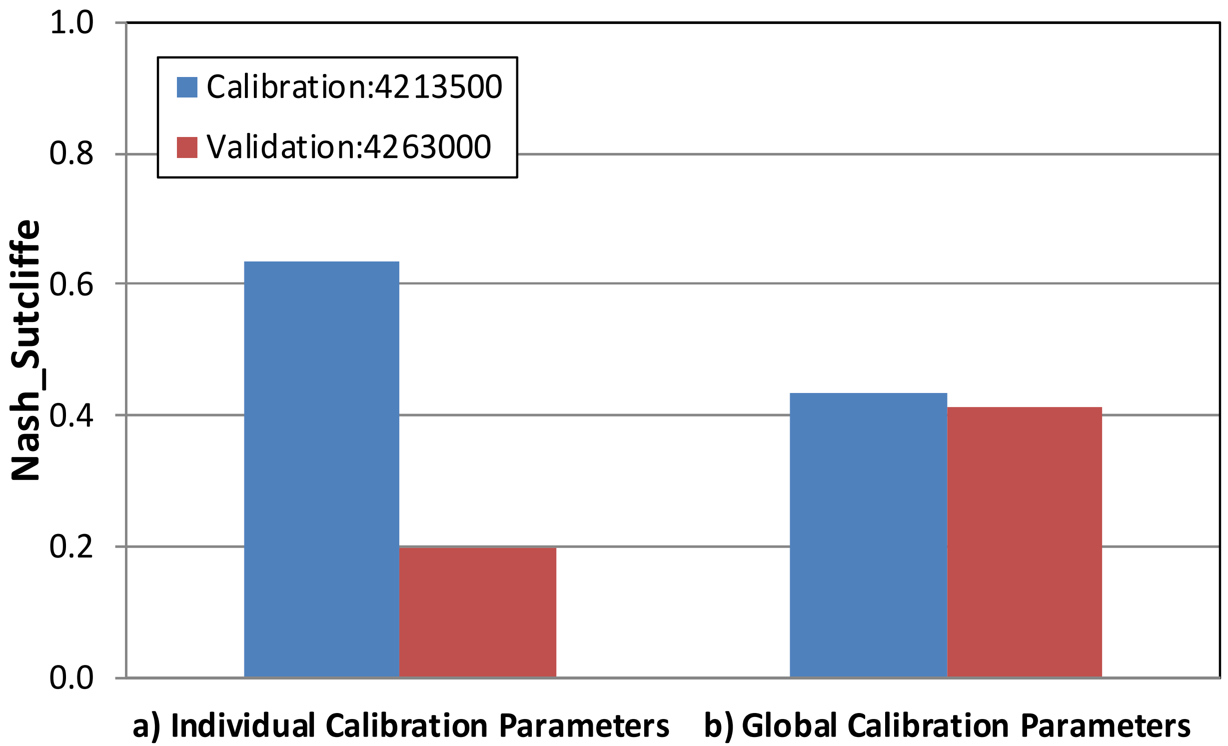
MESH Baseline Calibration Results: Global vs. Individual (Crop)



- Ind.: Very high calibration performance, very poor validation performance

- Global: Relatively high calibration and reasonable validation performance

MESH Baseline Calibration Results: Global vs. Individual (Deciduous Forest)



- Ind.: Very high calibration performance, low validation performance

- Global: Relatively high calibration and validation performance

CaPA vs GEM based simulations (June 04-09)

NS-Values	GEM	CaPA
02EC002	0.77	0.72
02BF002	0.49	0.50
02BB003	0.54	0.51
02GA010	0.36	0.38
04045500	0.72	0.60
04119000	0.58	0.16
AVG	**0.58**	**0.48**
Min	0.36	0.16
Max	0.77	0.72

South shore of Superior & Central Michigan

Summary of MESH Baseline Calibration Findings

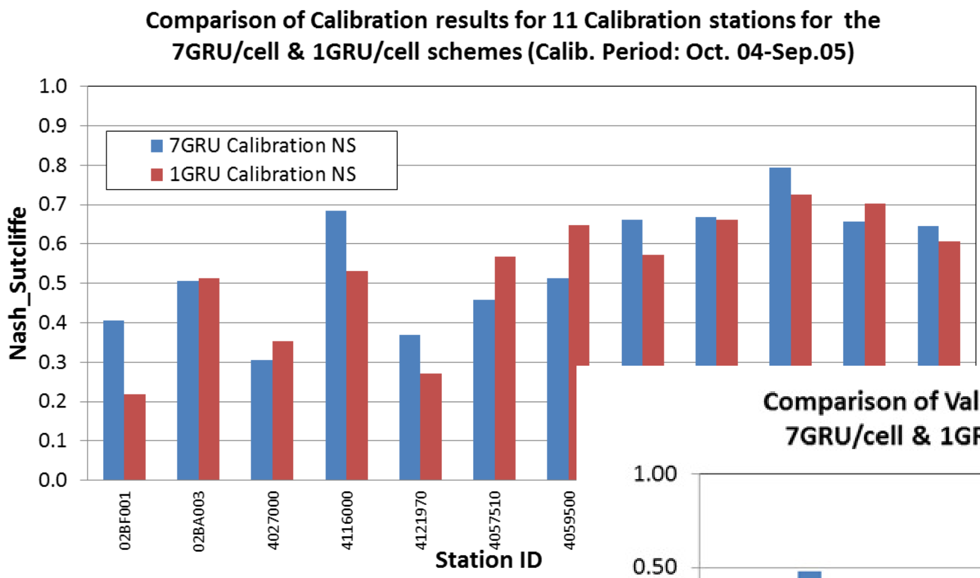
- Global Calibration generates reasonable quality hydrographs, especially in ‘ungauged’ basins
- Evaluating validation performance critical (as opposed to calibration performance)
- These results and the corresponding model setup are referred to from this point as “7-GRU results”

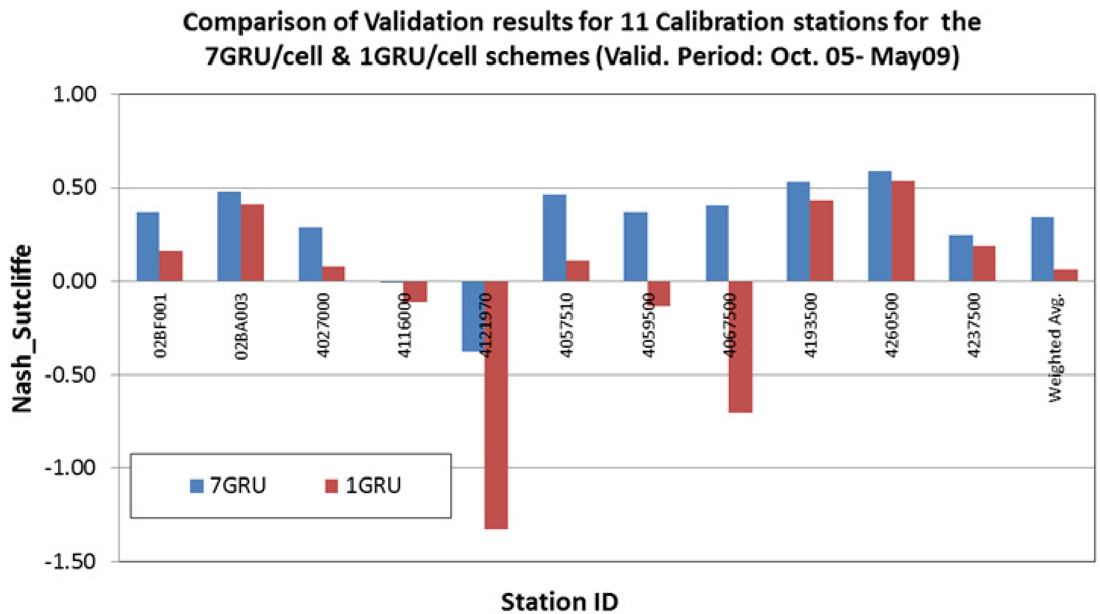
…

- *Next question is what happens if we change discretization decisions?*

Comparing two discretization schemes: 1-GRU vs. 7-GRU at Calibration Stations

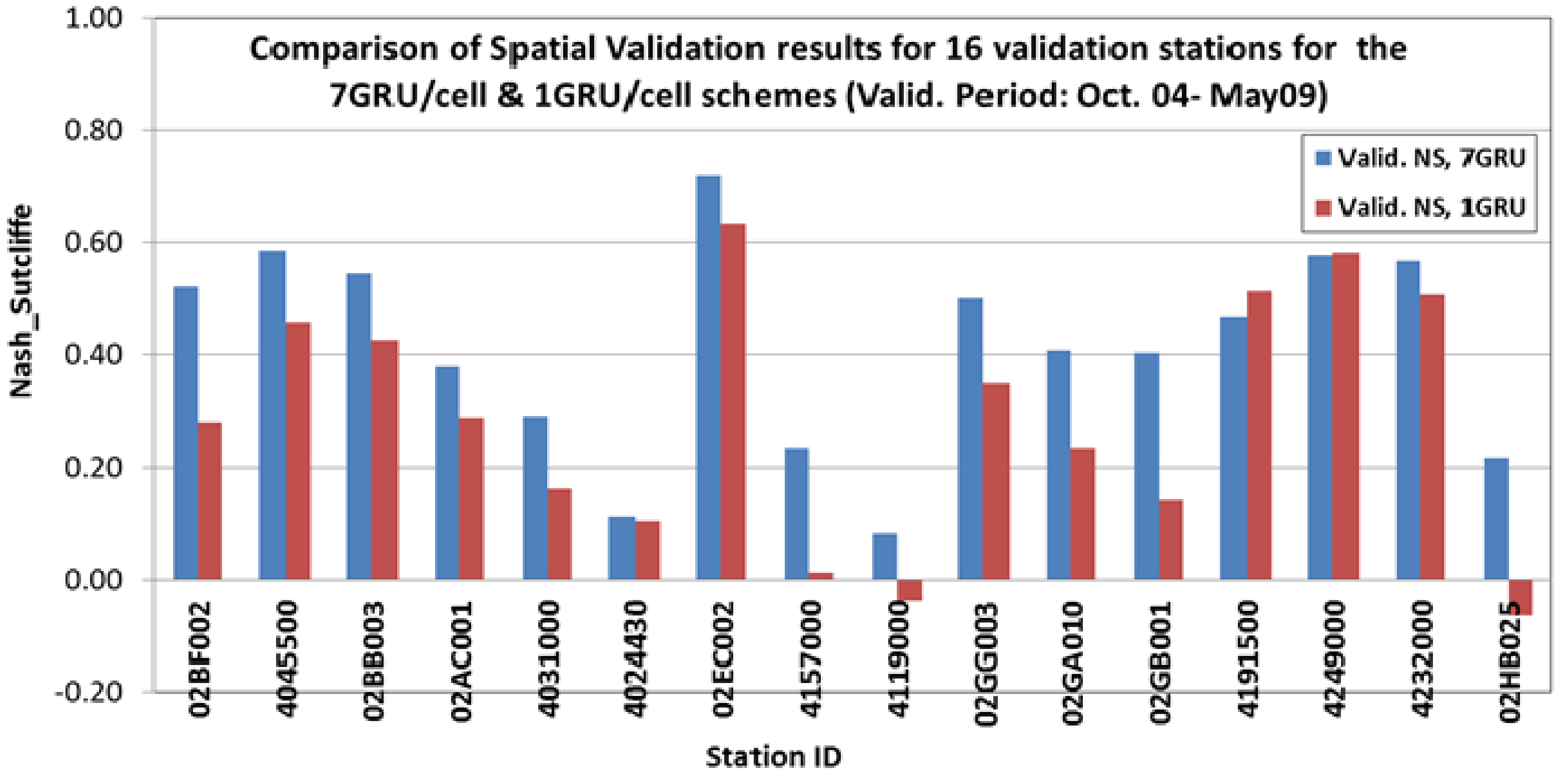
- 1-GRU: each grid cell modelled as single dominant GRU (recalibrated model!)
- 7-GRU: each cell splits into up to 7 different GRUs (higher runtime)





In temporal validation, 7-GRU better everywhere!

Comparing two discretization schemes: 1GRU vs. 7GRU *Spatial* Validation Performance



The 7GRU scheme
outperforms
the 1GRU scheme in all
stations except 2

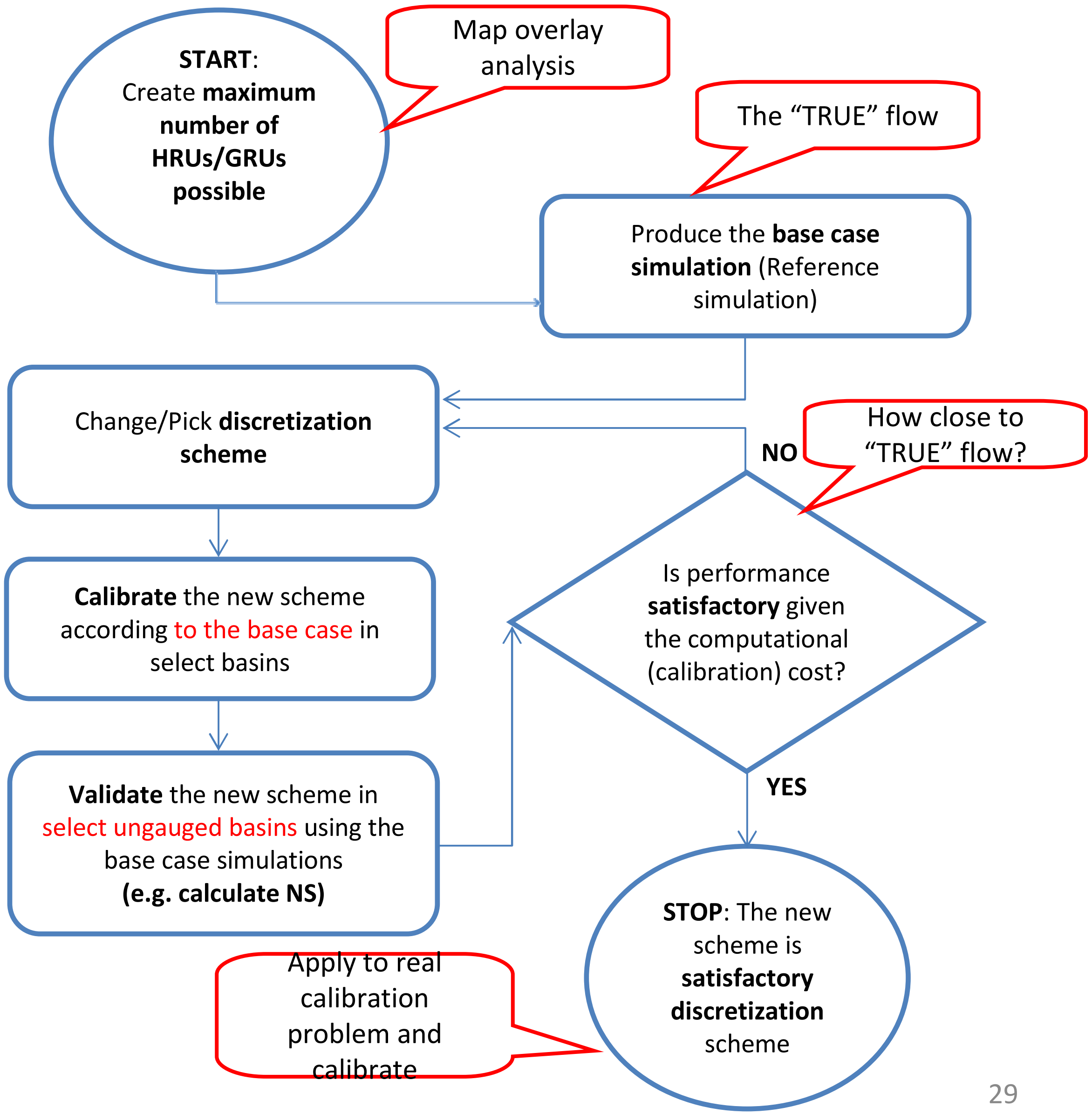
Summary of 1-GRU vs 7-GRU Discretization Strategy

- 1-GRU a 7-GRU strategies per grid appear quite similar in quality based only on calibration results
- We can only see 7-GRU is superior when validation performance is considered
- There are obviously more discretization options possible …

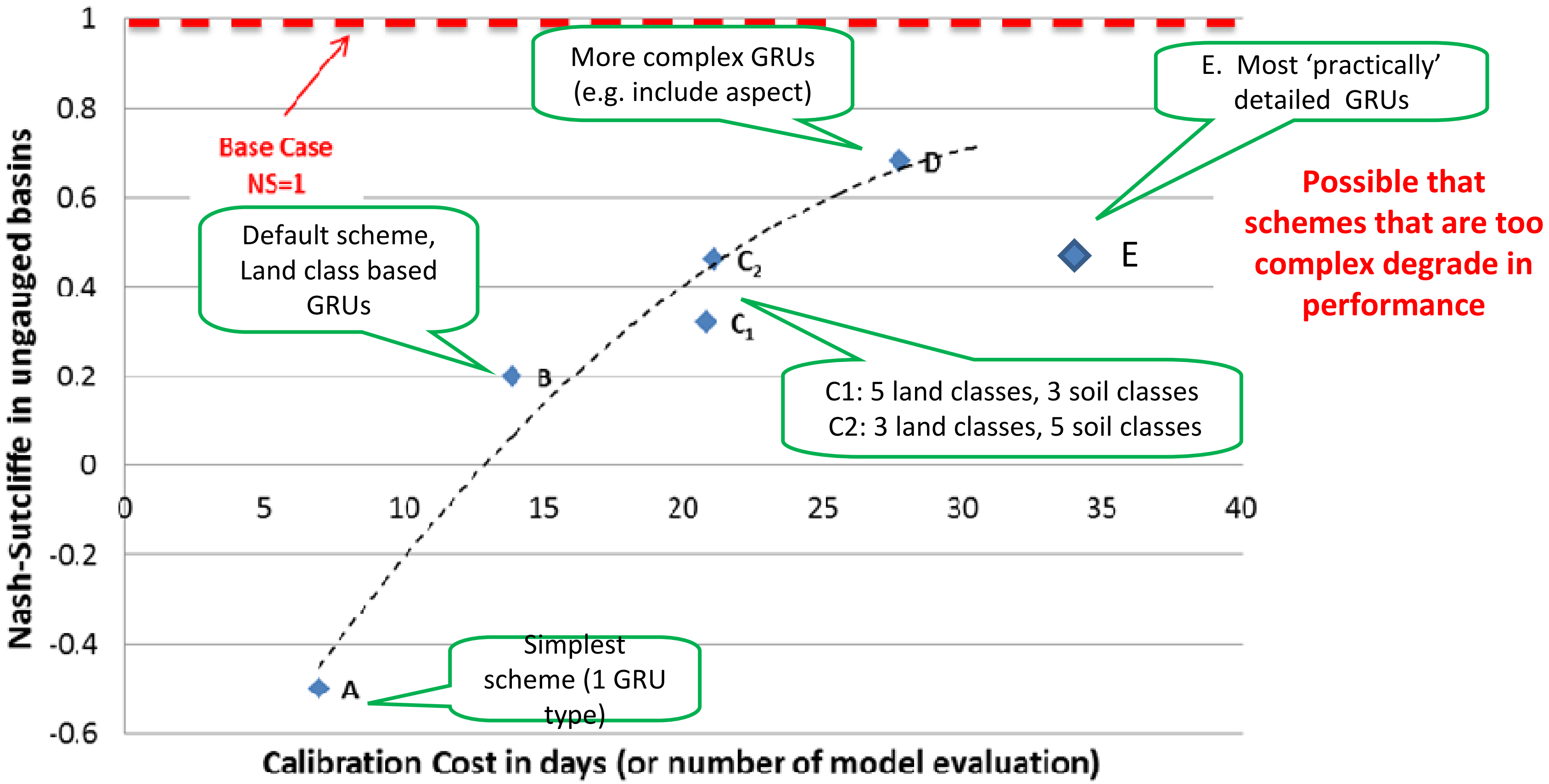
- *Next question is how to formalize discretization decisions as part of the calibration process?*

A new framework to assess alternative discretization schemes

- Synthetic framework to compare various discretization schemes
- All conditions like real case:
  - Calibration & Validation strategy
  - Forcing data
  - Basin/Sub-basins



A new framework to assess alternative discretization schemes: conceptual example

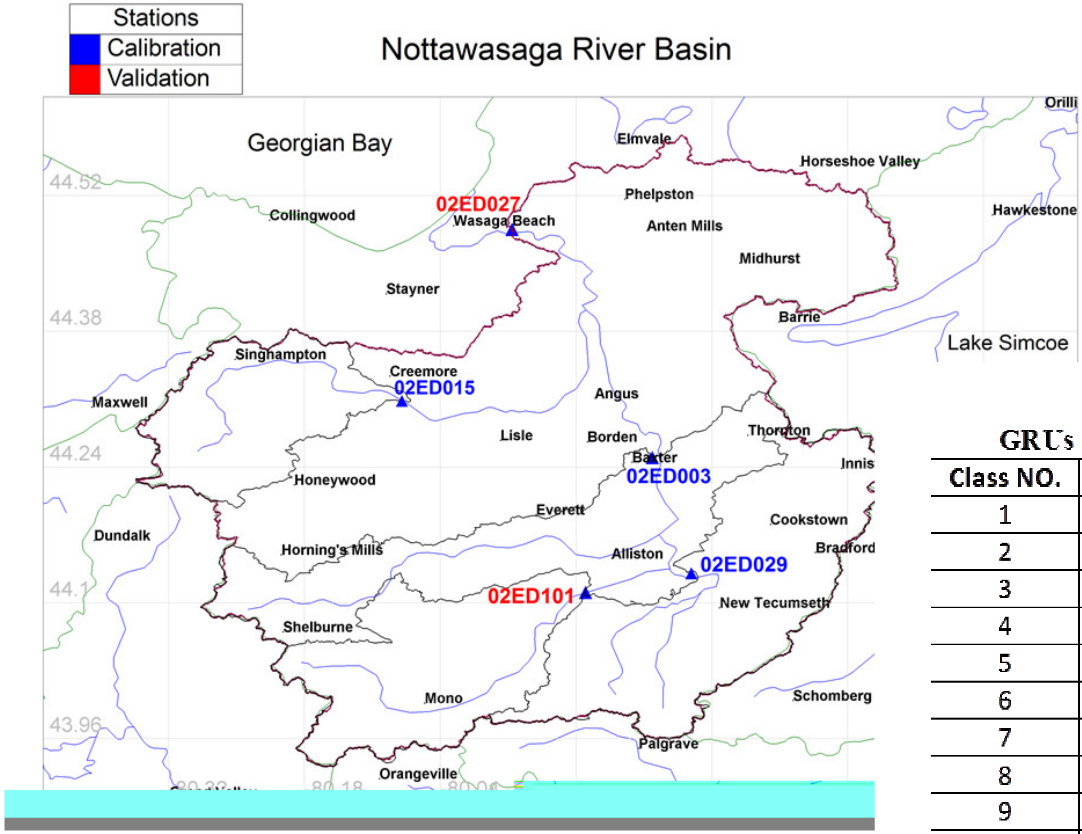


- Comparison of different discretization schemes according to the proposed **synthetic** framework
- *Could do this for real calibration problem as well but this is likely more time consuming*

Application of Framework to Example Basin: Nottawasaga Subbasin

- **3000 sq. km and drains into Georgian Bay**
- **Multiple nested gauges**
- **Many of the same model decisions as previous MESH runs:**
  - **1 yr calibration period**
  - **15 x 15 km grids**
  - **Models calibrated simulate 7 land class based GRU types**
  - **Same parameter types calibrated (except 1)**
- **Some differences include**
  - **Calibrate only two dominant GRU types (80% of the subbasin area)**
  - **Validate spatially only (same time period as calibration)**
  - **Much quicker run time!**

Case Study For New Framework: Nottawasaga River Subbasin of Great Lakes Basin



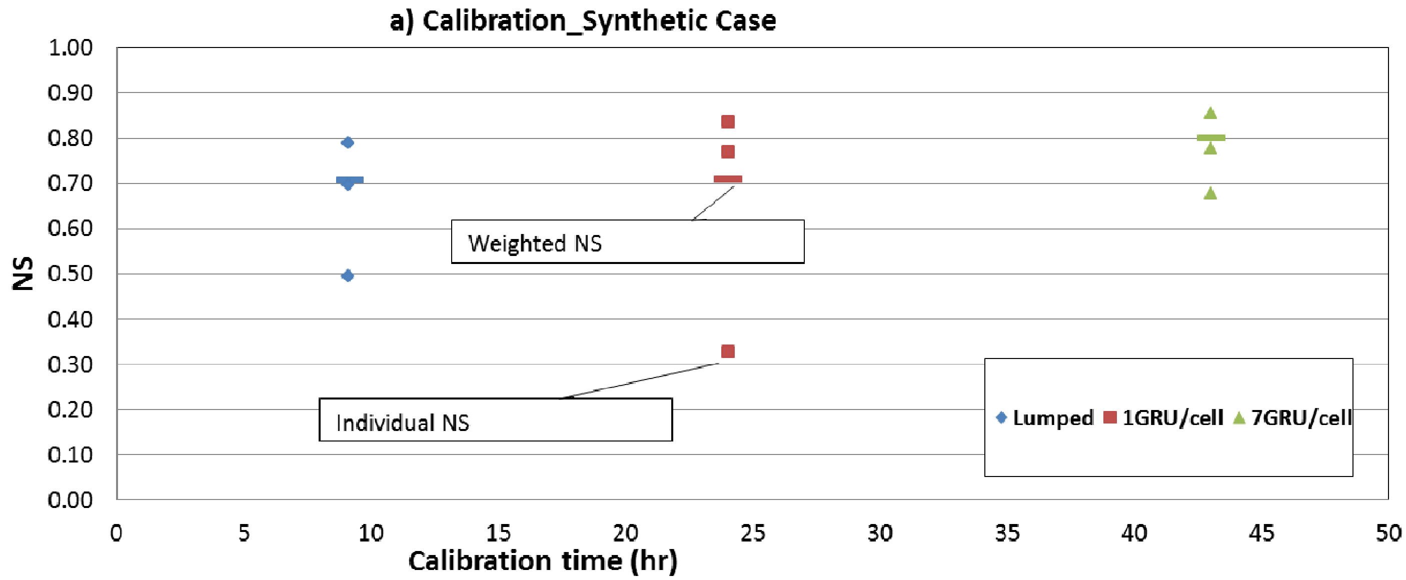
Define complex synthetic reality (pick a corresponding 'true' parameter set):

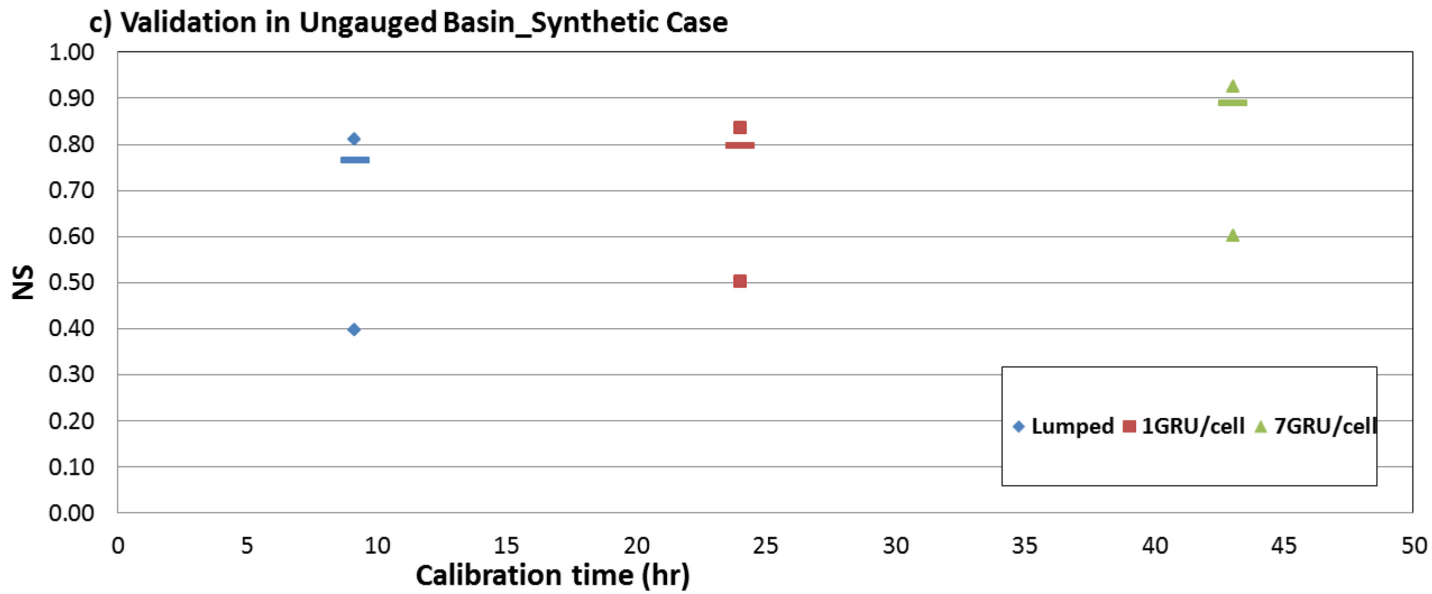
GRUs considered in creating the "TRUE" flow by MESH

Class NO.	Name	Area (km^2)	Percentage
1	Cropland, Clay Loam	196.5	7.3
2	Cropland, Organic	109.3	4.1
3	Cropland, Sandy	974.6	36.3
4	Cropland, Silty	490.2	18.3
5	Deciduous Forest, Clay Loam	12.1	0.5
6	Deciduous Forest, Organic	54.4	2.0
7	Deciduous Forest, Sandy	232.6	8.7
8	Deciduous Forest, Silty	103.9	3.9
9	Grassland, Clay Loam	18.6	0.7
10	Grassland, Organic	9.3	0.3
11	Grassland, Sandy	199.8	7.4
12	Grassland, Silty	99.9	3.7
13	Mixed Forest, Clay Loam	3.4	0.1
14	Mixed Forest, Organic	36.2	1.3
15	Mixed Forest, Sandy	111.9	4.2
16	Mixed Forest, Silty	32.8	1.2

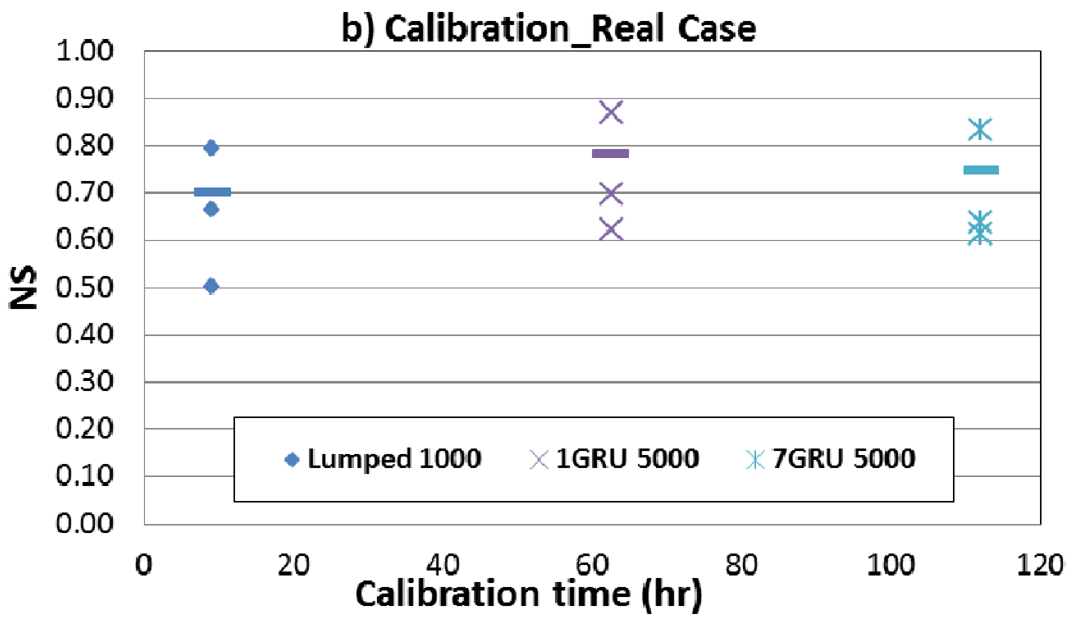
… Using alternative discretization schemes approximating synthetic reality, how close to sythetic reality can we get?

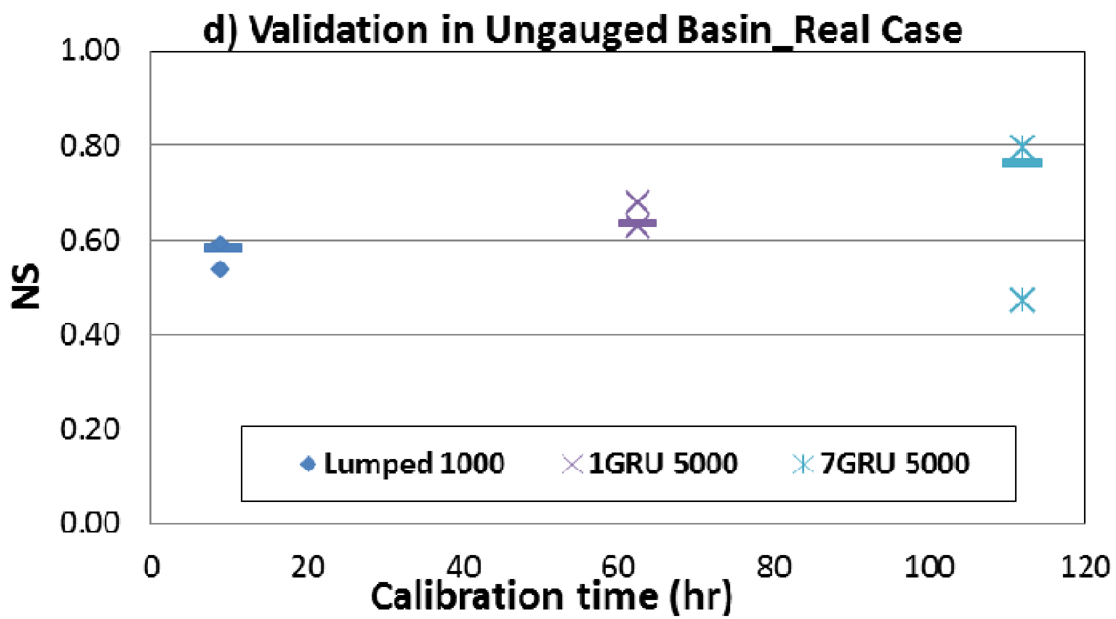
Synthetic Nottawasaga Results





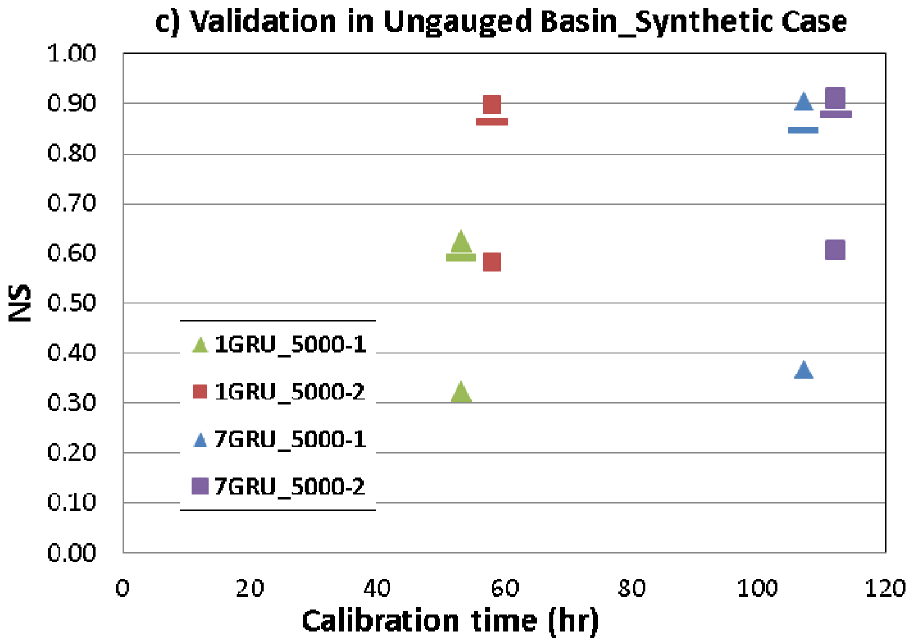
Real Calibration Nottawasaga Results

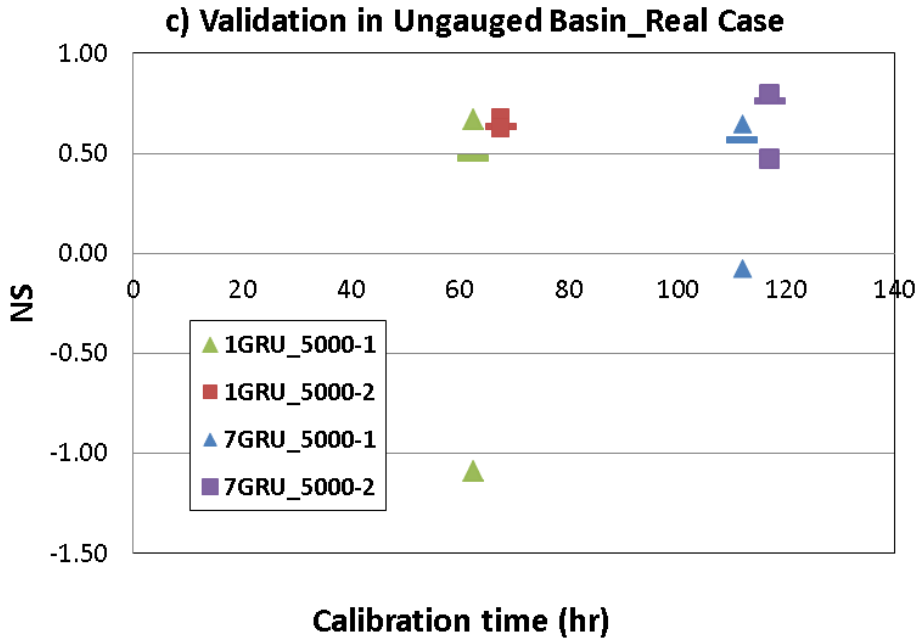




Sampling Variability Complicates the Analysis!!

- Repeat calibration experiment, just change random seed of DDS (initial solution)





Summary of Framework Application Findings

- Synthetic experiments are helpful in terms of bounding performance expectations in real calibration
- Calibration station performance not relevant – validation, validation!
- More experiments needed:
  - More detailed discretization in terms of smaller grid sizes and increased number of GRU types
- Overall findings show promise of framework but efficiency issues must be worked out

Other G & C Outputs

- Spatial datasets for Great Lakes and ultimately a suite of increasing complexity GRU maps
- Dr. Craig's RAVEN model improved to help run complex discretization framework examples

→ A conversion tool to take MESH inputs and produce corresponding RAVEN model inputs

Current Users of Research Products

- EC modellers (ASTD and MSC) use 7-GRU calibrated version of MESH
- PhD student Amin building his PhD around this framework for making discretization decisions

- My research group is also using model to:
  - test soil moisture data assimilation strategies for stand-alone MESH in Great Lakes Basin (X. Xu & Dr. J. Li) *
  - Test new Parallel Pre-emptable DDS algorithm * (SHARCNET programming support grant to end of June)
  - Compare uncertainty-based multi-criteria calibration methods such as Bayesian and informal methods (M. Shafii)

THANKS

- Questions?
- btolson@uwaterloo.ca

Appendix

Table 1- Parameters and their corresponding ranges used in calibration of MESH for the Great Lakes Basin

Parameter	Description	Upper Bound	Lower Bound
ROOT	Annual maximum rooting depth of vegetation category [m]	1.0	3.5
QA50	Reference value of incoming shortwave radiation [W m^{-2}]	30.0	50.0
RSMN	Minimum stomatal resistance of vegetation category [s m^{-1}]	50	300
DRN	Soil drainage index	0.00	1.00
DDEN	Drainage density	2.0	100
XSLP	The average overland slope of a given GRU [%]	0.0001	0.10 (0.04 for crop)
WFCI	Saturated surface soil conductivity [m/s]	0.00	1.02
SAND	Percentage sand content	0.00	50.0
CLAY	Percentage clay content	0.00	50.0
SDEP	Soil permeable depth [m]	0.35	4.1
WFR2	WATFLOOD channel roughness coefficient	0.02	2.0