Parallel Calibration of SWAT+ Model at Large Scale Using Dynamically Dimensioned Search Algorithm

Jungang Gao[1], Mike White[2], Jeff Arnold[2], Natalja Cerkasova[1], Peter Allen[3], Kelly Thorp[2], Joon-Hee Lee[2], Marilyn Gambone[2], Sagarika Rath[1], Celray James Chawanda[1]

July 2024

Texas A&M AgriLife

USDA United States Department of Agriculture
Agricultural Research Service
Grassland Soil and Water Research Laboratory (GSWRL)

[1] Blackland Research & Extension Center, Texas A&M AgriLife, 702 E. Blackland Road, Temple, TX 76502, USA;
[2] Agricultural Research Service, US Department of Agriculture, 808 E. Blackland Road, Temple, TX 76502, USA;
[3] Department of Geosciences, Baylor University, 101 Bagby Ave., Waco, TX 76706, USA;

Contents

1. Dynamically Dimensioned Search (DDS) Algorithm

2. Introduction of a Parallel DDS Interface

3. A case study in Upper Mississippi River basin (UMRB)

1. Dynamically Dimensioned Search (DDS) Algorithm

Dynamically Dimensioned Search Algorithm (DDS) Brief

1. The sampling operator, controlled by sampling criteria probability (P), selects some decision variables perturbed from all the decision variables.

$$P(i) = 1 - \ln(i) / \ln(m)$$

2. The perturbing operator perturbs the current best solution, as shown in Figure 3, to generate the candidate solution for a selected decision variable.

3. The decision operator serves to determine whether the candidate solution yields an objective function value inferior to the current best value

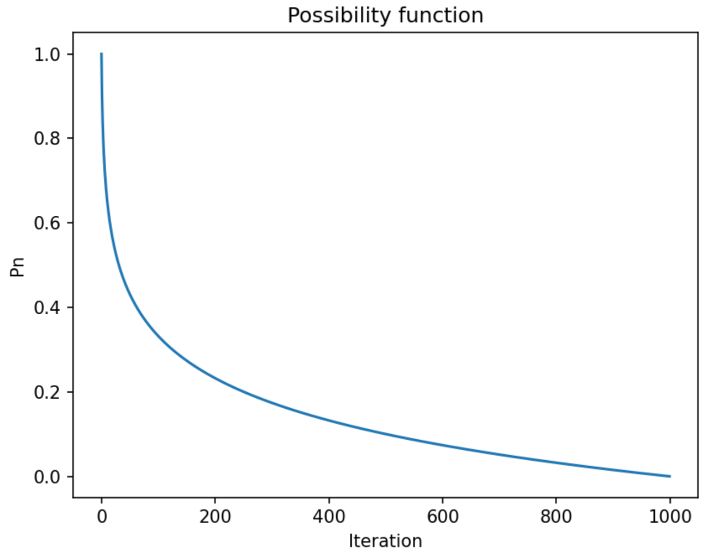
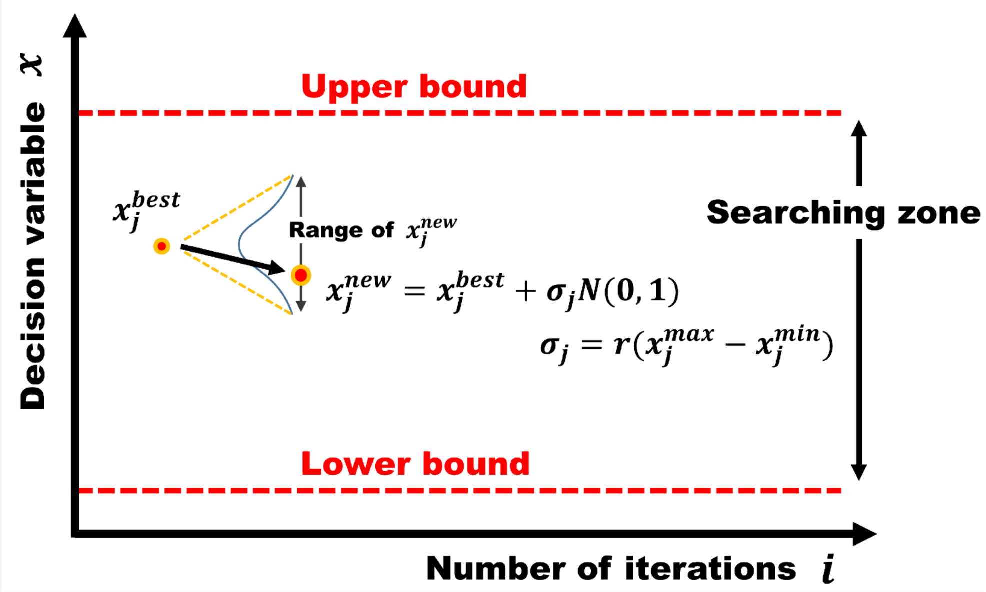


Figure 3. The schematic diagram of the perturbing operator in the DDS.



Jin, Y.; Lee, S.; Kang, T.;Kim, Y. A Dynamically Dimensioned Search Allowing a Flexible Search Range and Its Application to Optimize Discrete edging Rule Curves. Water **2022**, 14, 3633.

2. Introduction of a Parallel DDS Interface

2.1 DDS Parallel Running Concept

Water Resources Research

Regular Article | Free Access

Dynamically dimensioned search algorithm for computationally efficient watershed model calibration

Bryan A. Tolson, Christine A. Shoemaker

First published: 17 January 2007 | https://doi.org/10.1029/2005WR004723 |

Abstract

[1] A new global optimization algorithm, dynamically dimensioned search (DDS), is introduced for automatic calibration of watershed simulation models. DDS is designed for calibration problems with many parameters, requires no algorithm parameter tuning, and automatically scales the search to find good solutions within the maximum number of user-specified function (or model) evaluations. As a result, DDS is ideally suited for computationally expensive optimization problems such as distributed watershed model

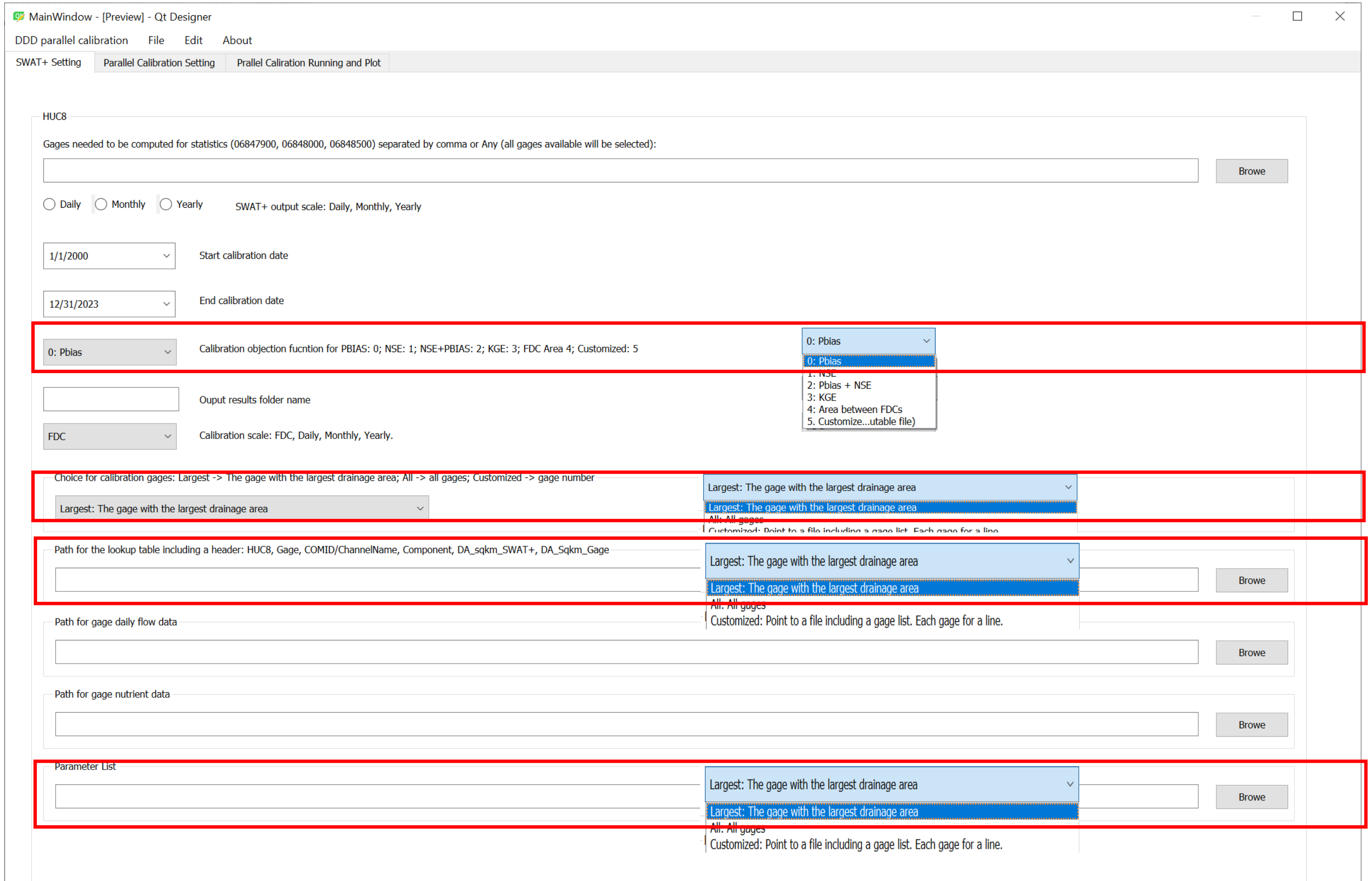
❑ Interface (1/3) – SWAT+ Setting

+ Object Function setting
  - Pbias
  - NSE
  - KGE
  - Pbias + NSE
  - Area between two FDCs (Sim. Vs Obs.)
  - Customized executable file

+ Gage selection setting
  - The gage with the largest drainage area
  - All gages
  - Customized gage selection (excluding outlier gages)

+ Lookup table for gages and channels
  - Sub-model, Gage, ComID/CHA_Name, Components, DA_SWAT, DA_Gage

+ Parameter list
  - Name
  - Change type: absval, abschg, pctchg
  - Lower and upper boundry



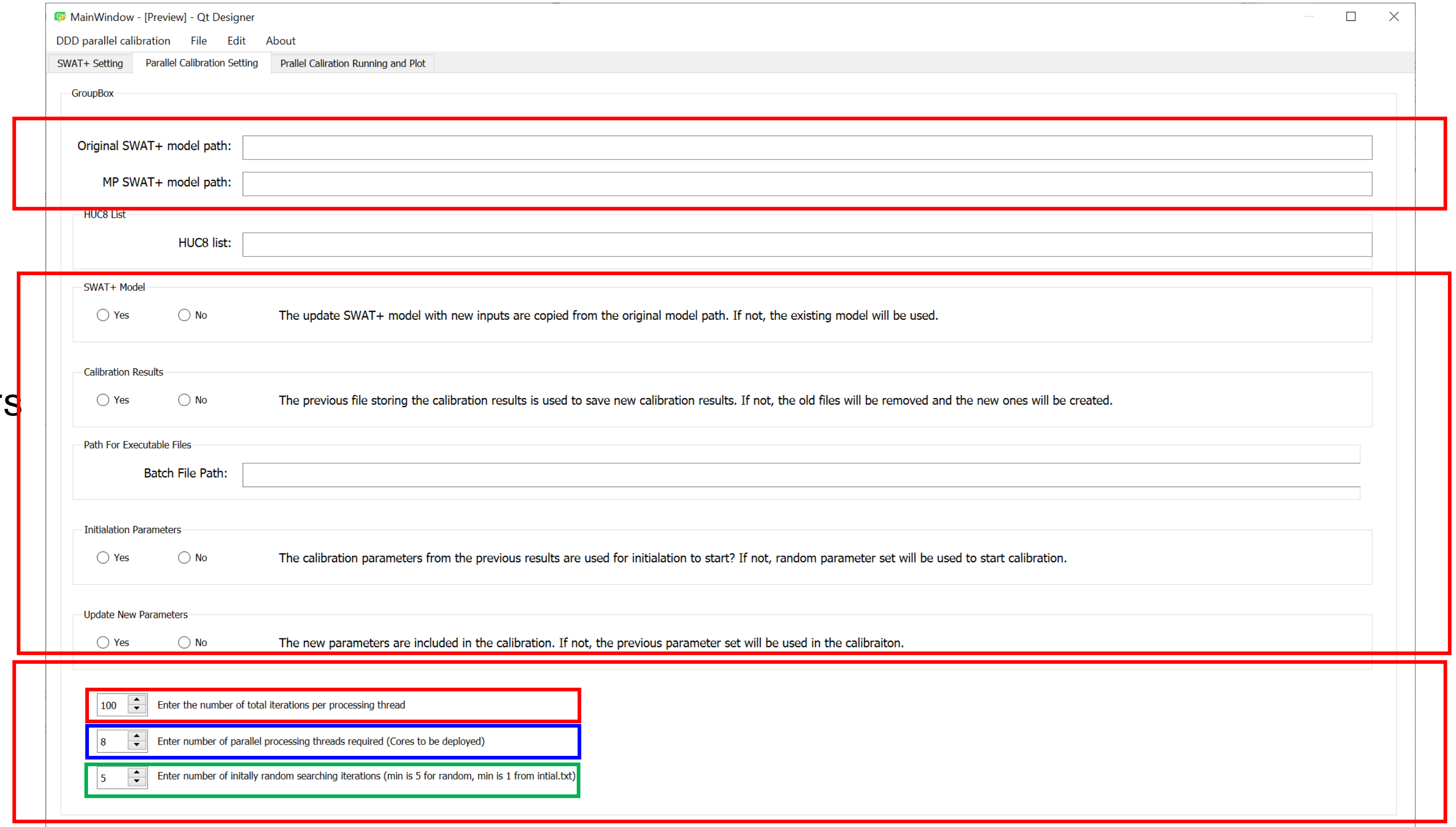
15							
NAME	CHG_TYPE	File	LowerB	UpperB	Interger?	DefalutVAL	UNIT
alpha	absval	aqu.aqu	0.01	1.000	0	0.03	days
flo_min	absval	aquifer.aqu	0	50	0	25	-
revap_min	absval	aquifer.aqu	0	10	0	5	mm

❑ Interface (2/3) – Parallel Setting

+ Original and Parallel running path for SWAT+ models
  - Sub model name list

+ General setting
  - Update SWAT+ files
  - Use or not with the previous parameters
  - Start calibration from random or customized or previous parameter combination
  - Add new parameters

+ Total iterations of all cores = Cores * (Random + Further])



+ 100 Iterations = 5 Cores * (5 Random + 15 Further) → From Random Parameters

+ 100 Iterations = 5 Cores * (1 Random + 19 Further) → From Initial Parameters

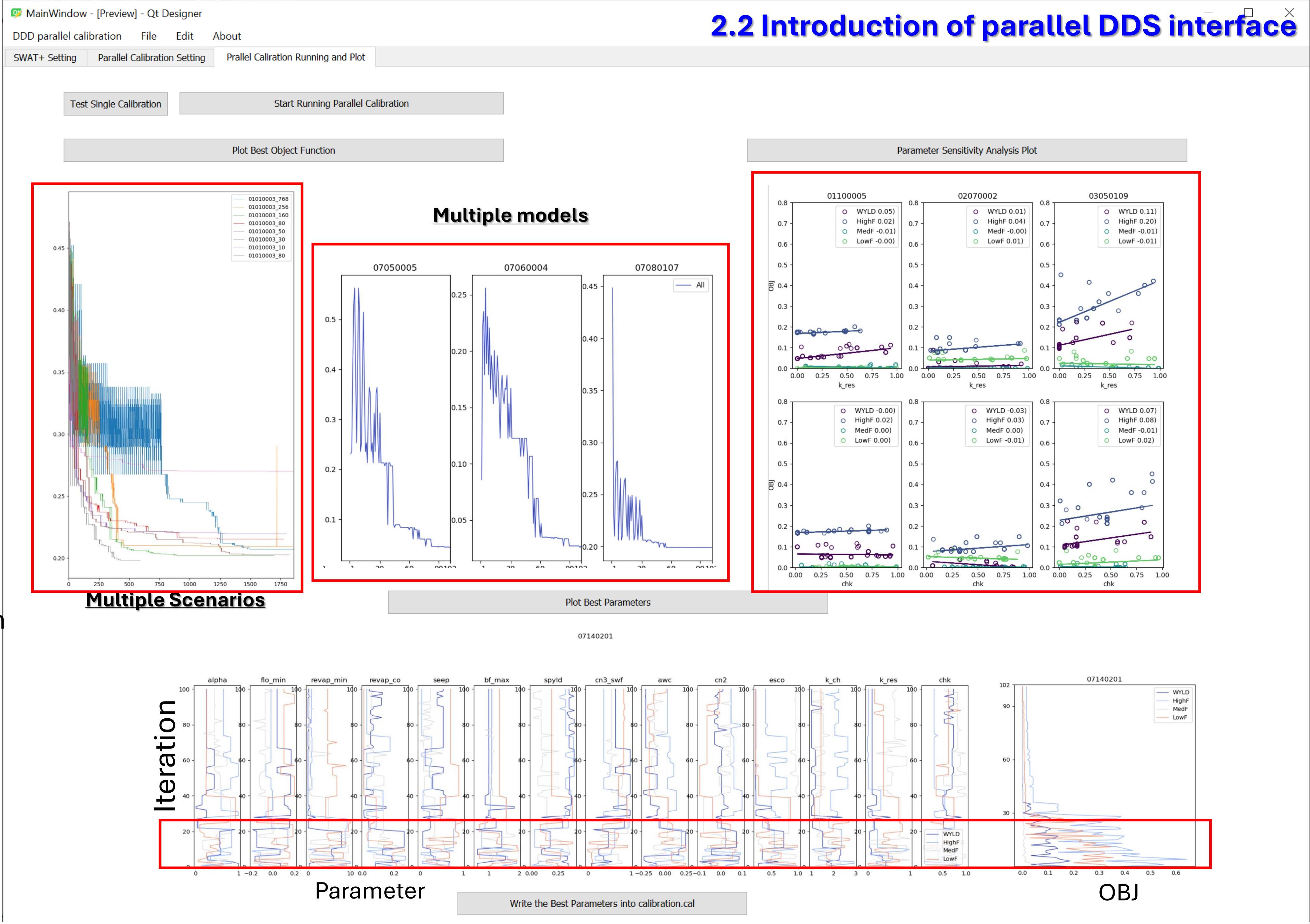
2.2 Introduction of parallel DDS interface

❑ Interface (3/3)

+ Plot Best Object Function
  - Multiple scenarios
  - Multiple models

+ Plot Parameter Sensitivity
  - Multiple scenarios
  - Slope rate

+ Best parameter combination
  - Multiple scenarios
  - Check convergence of each parameter



❑ Statistical results

+ Parameter record
  - Each iteration

Scen	Gage	COMID	CHA_NAM	DA_NSqkm	DA_GSqkm	G0.1	G0.5	G1	…	G98	G99	S0.1	S0.5	S1	…	S98	S99
FDC	7283000	18015234	m_4_1172	633.03	657.86	413.43	265.05	174.29	…	0.20	0.19	63.17	60.21	56.71	…	0.70	0.62

+ Statistics on time series/FDC data
  - NSE, r, R^2, Pbias, KGE, RMSE, RSR, median simulation, median observation, mean simulation, mean observation

Scen	Gage	COMID	CHA_NAM	DA_NSqkm	DA_GSqkm	Count	r	r2	NSE	pBias	RMSE	RSR	Median_sim	Median_obs	Mean_sim	Mean_obs	KGE
Yearly	7285500	18016786	m_6_2745	4039.83	4014.50	16.00	0.81	0.66	-6.94	-99.82	67.38	2.82	0.12	61.73	0.11	63.12	0.81

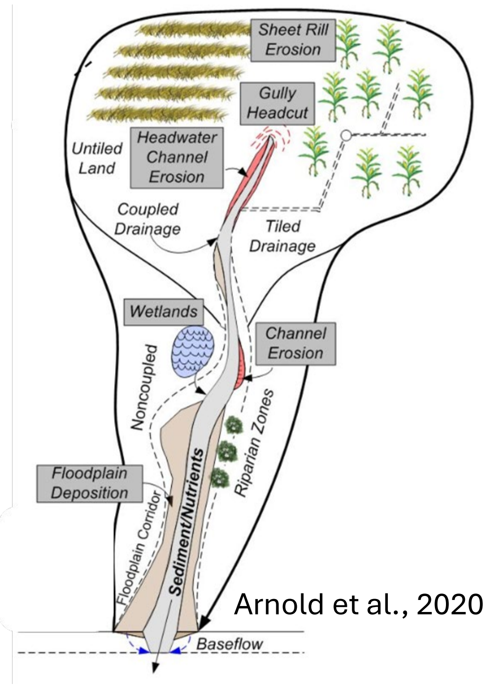
Scen	Gage	COMID	CHA_NAM	DA_NSqkm	DA_GSqkm	Count	r	r2	NSE	pBias	RMSE	RSR	Median_sim	Median_obs	Mean_sim	Mean_obs	KGE
FDC	7285500	18016786	m_6_2745	4039.83	4014.50	27.00	0.97	0.94	-0.33	-80.03	85.27	1.15	6.49	67.68	16.33	81.79	0.66

A case study on DDS parallel calibration in UMRB

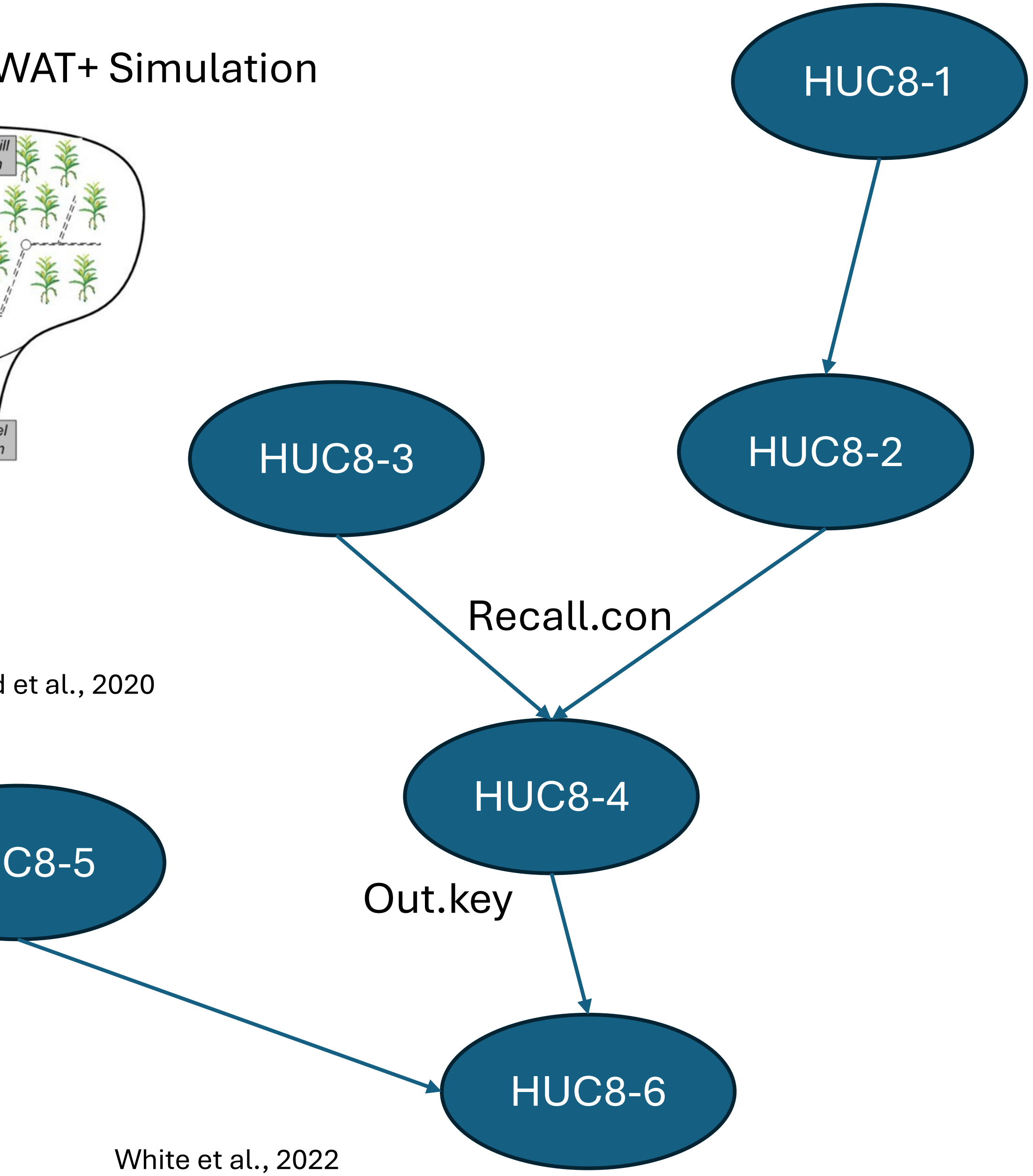
1. NAM model framework
2. NAM inputs
3. DDS parallel calibration in UMRB
4. Calibration results
5. Future work

1. National Agroecosystem Model (NAM) model Framework

Process Based SWAT+ Simulation

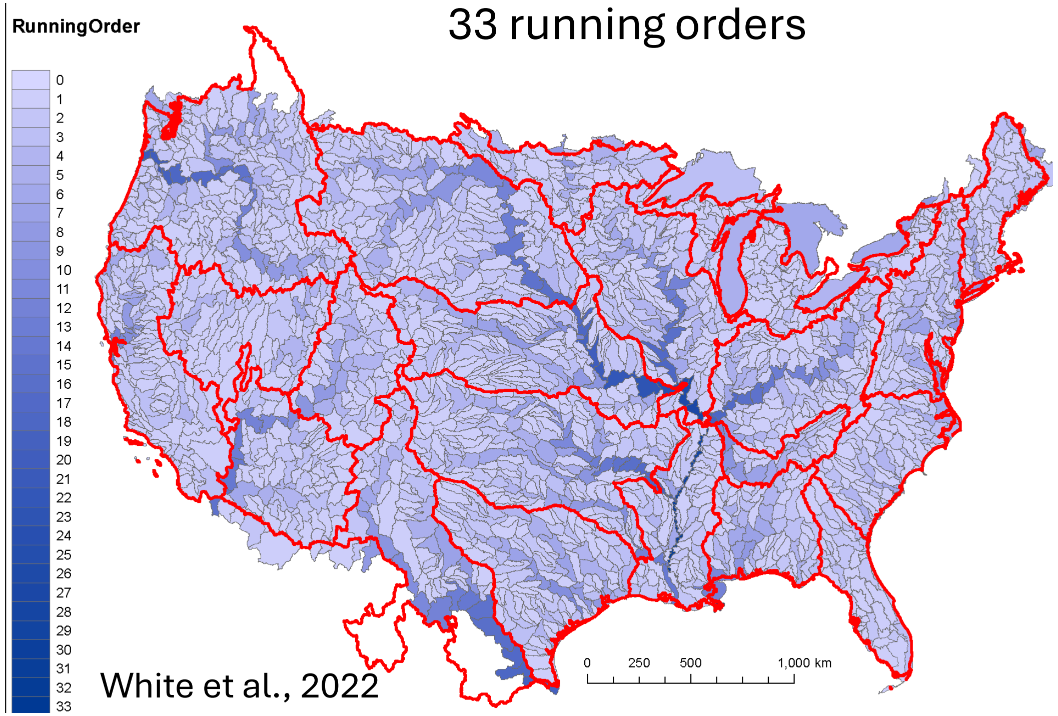


Arnold et al., 2020



White et al., 2022

33 running orders

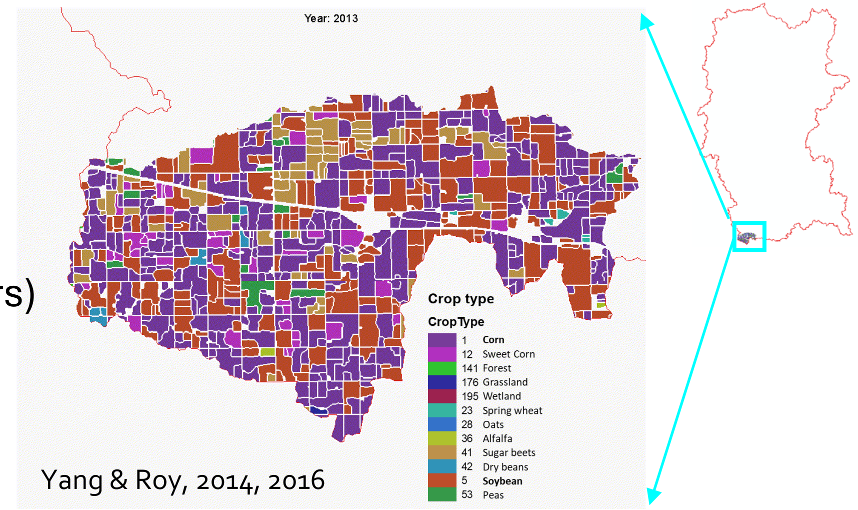


White et al., 2022

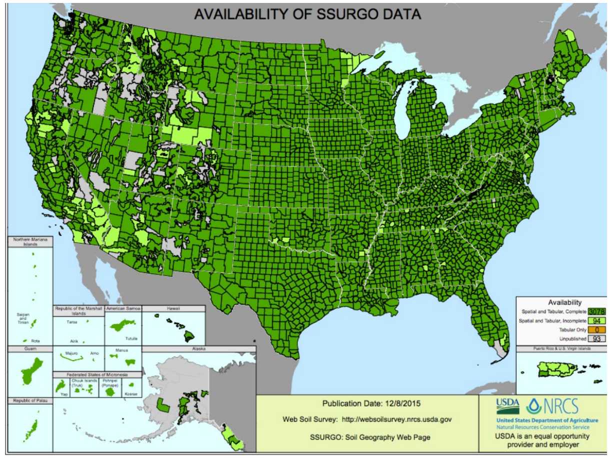
- National Hydrography Dataset V2
  - 2.6 million digitized reaches
  - 2121 HUC8s
- Water bodies
  - Lakes/Reservoirs
  - PL-566
  - Farm Ponds
- Point Sources
  - EPA DMR

2. NAM inputs

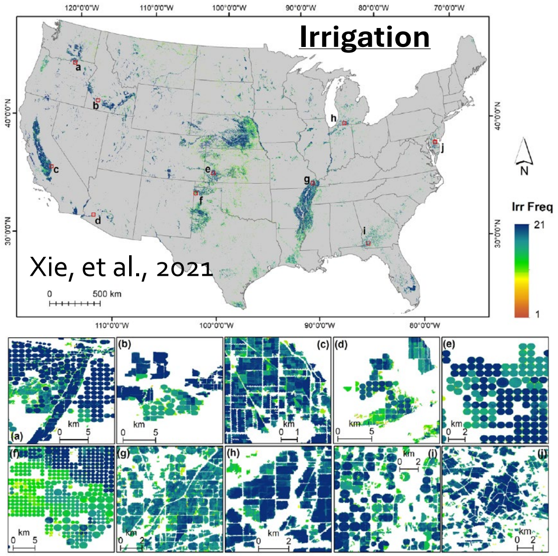
- Land use (CDL) and crop rotations (5 yrs)
- Soil: SSURGO (Vector)
- Slope: DEM (10m)
- Irrigation
- Tile
- Weather data: NEXRAD, PRISM, Station data at HUC12 level
- Tillage: County level
- Fertilization at country level
- Conservation practices
- Decision tables (reservoir, crop mgt, scenario…)
- …

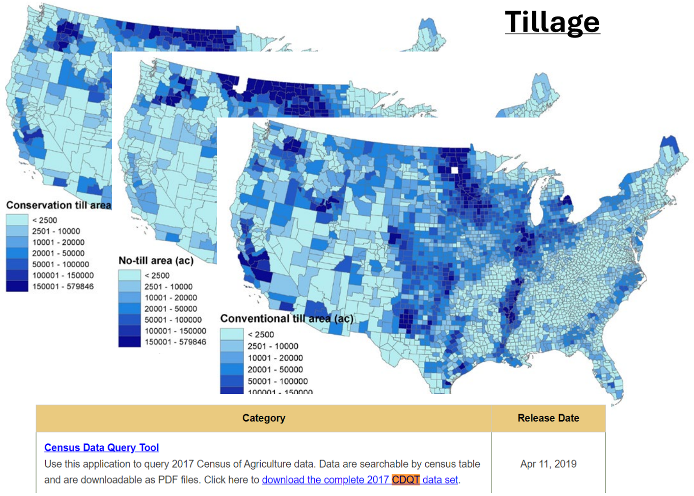


Five years (2013-2017) of crop rotation in HUC10: 0701020505

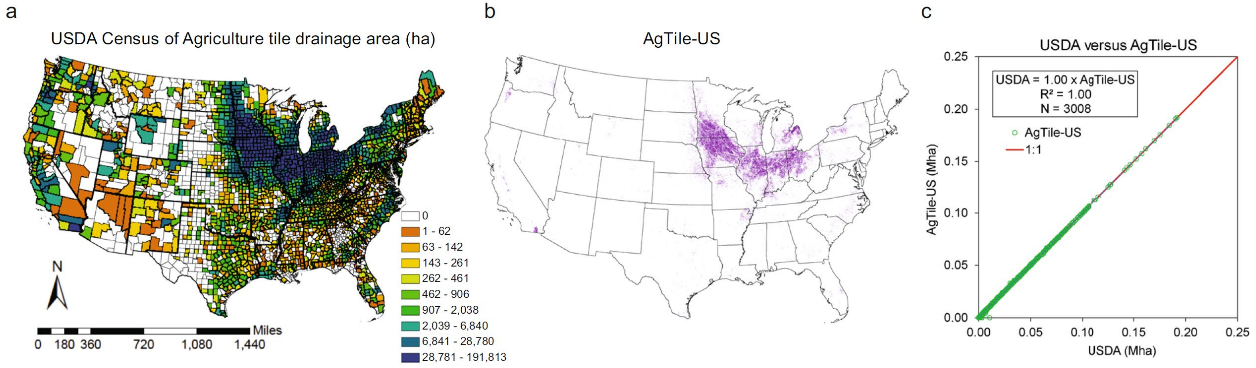


SSURGO





USDA National Agricultural Statistics Service



Alayamkunnath, et al., 2020

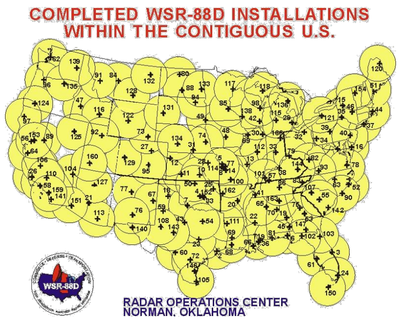
Tile

PRISM CLIMATE GROUP

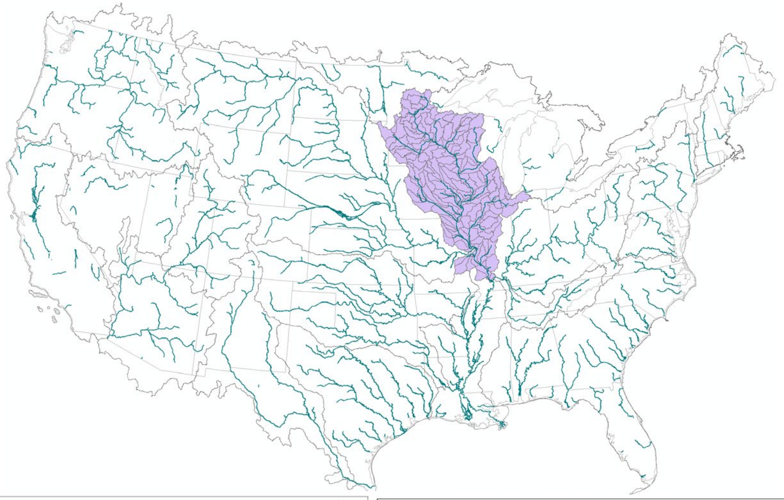
Temperature

Weather Data

Precipitation NEXRAD radar data

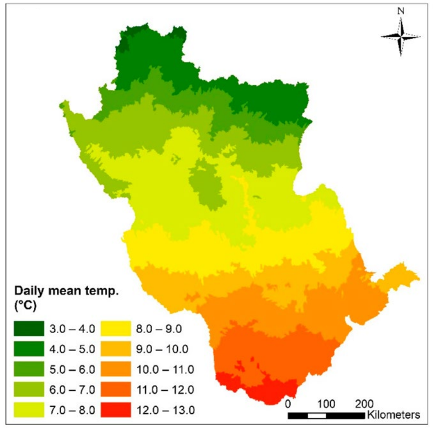


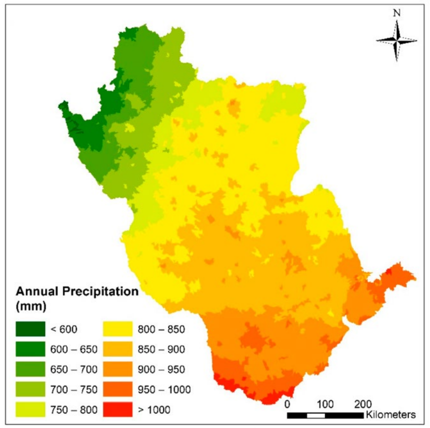
3. DDS parallel calibration in UMRB

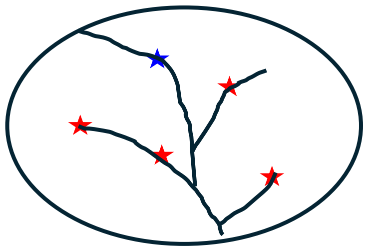


3.1 Background

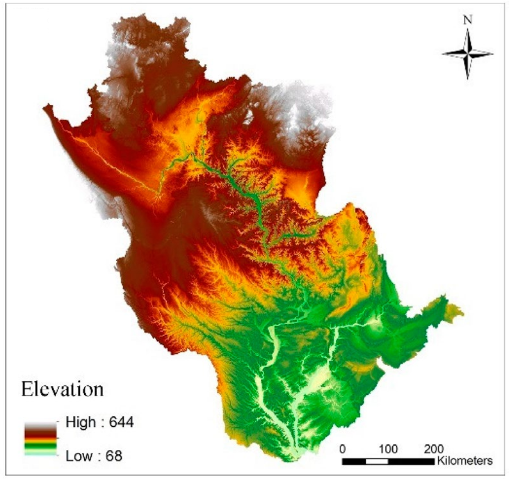
- A sub-model of (National Agroecosystem Model) NAM (1/19)
- Drainage area: 431,000 km^2
- HUC8s: 131
- HUC12s: 6,731
- HRUs: 1,048,800
- Channels: 187,161
- Calibration period: 1997-1999 (warm-up) & 2000-2018 (calibration)







- All local gages (red) in each HUC8
- Relative Error (RE) between median FDCs at gages and corresponding channels simulated by NAM
- Mean RE in High flow (RE_{HF}, Q1-10), Mean RE in Median flow (RE_{MF}, Q10-60) and Mean RE in Low flow (RE_{LF}, Q60-99).
- Relative Error (RE_{WYLD} & RE_{Tile}) in Water yield (WYLD) and tile flow (Tile)
- ...



- OBJ = $w_1 * RE_{WYLD} + w_2 * RE_{HF} + w_3 * RE_{MF} + w_4 * RE_{LF} + w_5 * RE_{Tile} + ...$

Developed by Dr. Mike White (USDA ARS GSWRL)

3.2 Determination of random initial trial and total iteration

1) The trade-off initial random start vs. further iteration number

16 cores * (48 random + 68 further) = **1856** Random: 768

16 cores * (16 random + 100 further) = **1856** Random: 256

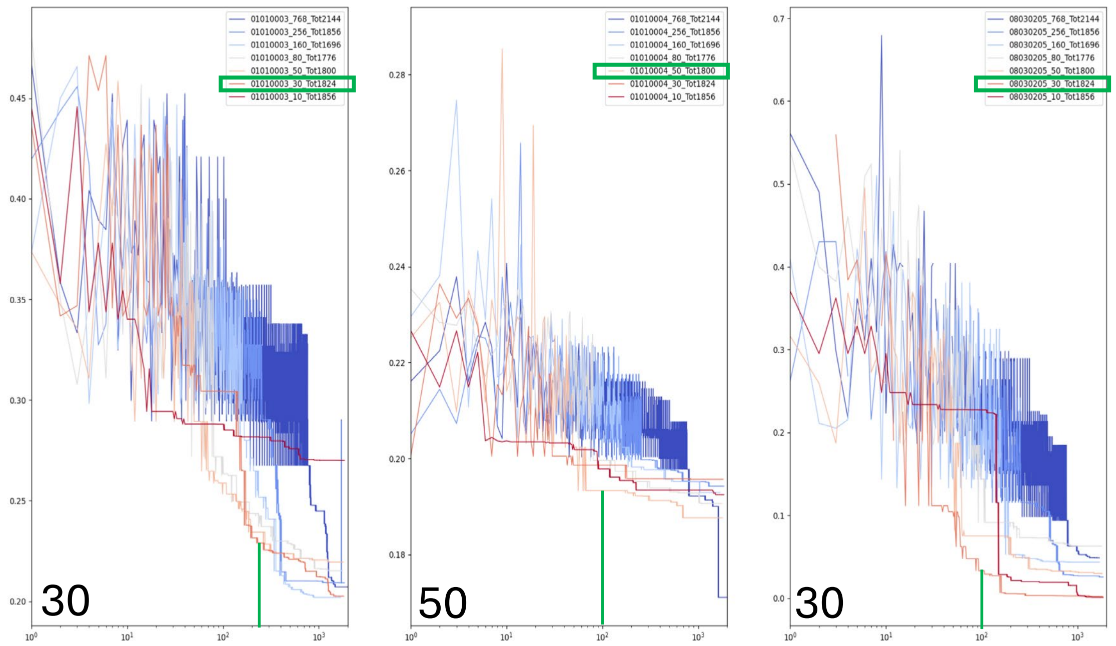
16 cores * (10 random + 106 further) = **1856** Random: 160

16 cores * (5 random + 111 further) = **1856** Random: 80

10 cores * (5 random + 180 further) = **1850** Random: 50

6 cores * (5 random + 304 further) = **1854** Random: 30

2 cores * (5 random + 923 further) = **1856** Random: 10



2) The trade-off total iteration number (Max vs. Efficient)

6 cores * (5 random + 79 further) = **504** Random: 30

6 cores * (5 random + 62 further) = **402** Random: 30

6 cores * (5 random + 45 further) = **300** Random: 30

6 cores * (5 random + 32 further) = **222** Random: 30

6 cores * (5 random + 29 further) = **204** Random: 30

6 cores * (5 random + 25 further) = **180** Random: 30

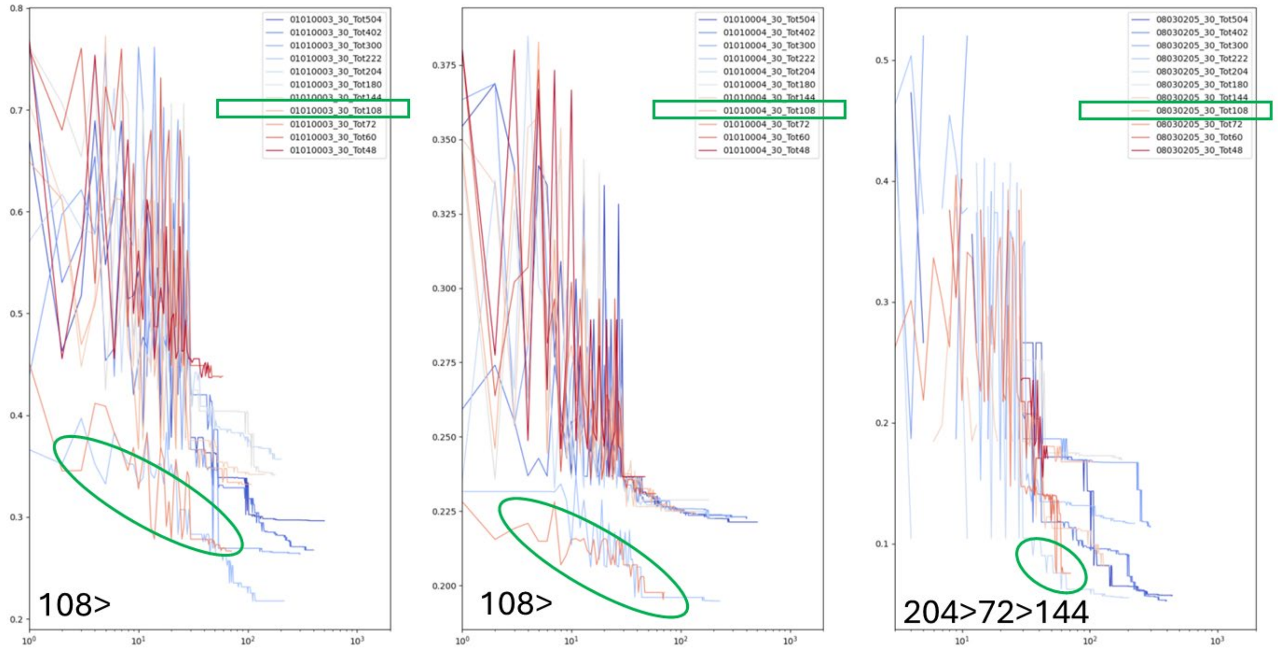
6 cores * (5 random + 19 further) = **144** Random: 30

6 cores * (5 random + 13 further) = **108** Random: 30

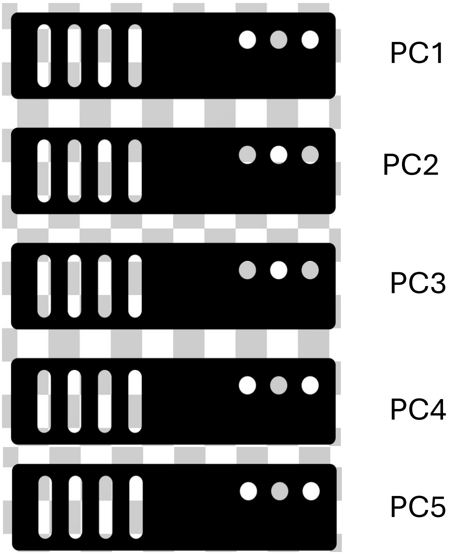
6 cores * (5 random + 7 further) = **72** Random: 30

6 cores * (5 random + 5 further) = **60** Random: 30

6 cores * (5 random + 3 further) = **48** Random: 30



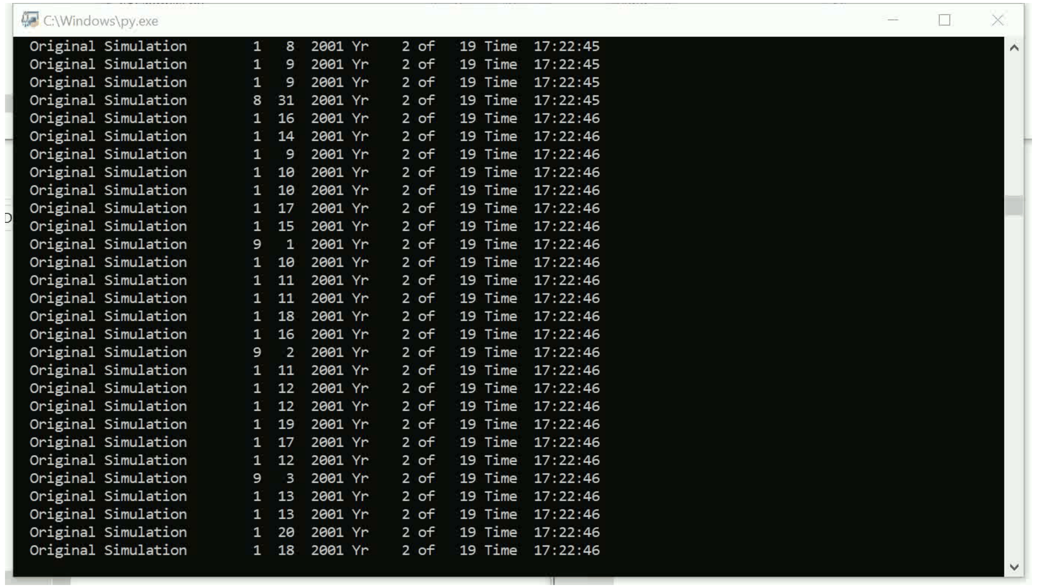
3.3 Calibration process



Around 700 physical cores available

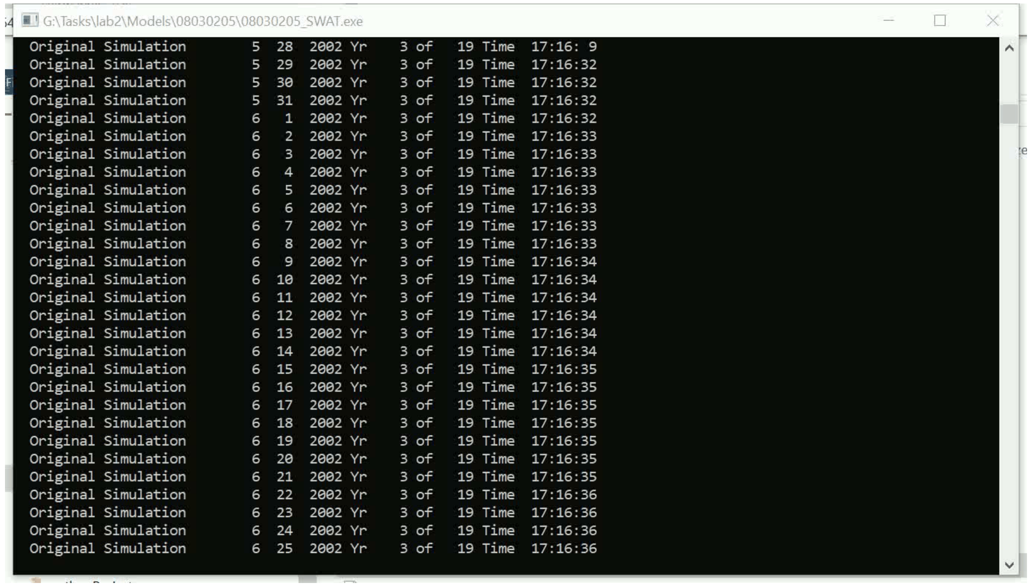
- 5 super computers (average cores 140 for each pc)
- Runs at the same time: 6 parallel threads for each HUC8 model * 131 HUC8s = 786
- Total runs for all HUC8s: 131*102=13,362
- Consuming Time: 10 hours

6 cores parallel running



```
C:\Windows\py.exe
Original Simulation       1   8  2001 Yr    2 of   19 Time  17:22:45
Original Simulation       1   9  2001 Yr    2 of   19 Time  17:22:45
Original Simulation       1   9  2001 Yr    2 of   19 Time  17:22:45
Original Simulation       8  31  2001 Yr    2 of   19 Time  17:22:45
Original Simulation       1  16  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  14  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1   9  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  10  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  10  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  17  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  15  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       9   1  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  10  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  11  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  11  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  18  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  16  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       9   2  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  11  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  12  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  12  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  19  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  17  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  12  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       9   3  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  13  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  13  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  20  2001 Yr    2 of   19 Time  17:22:46
Original Simulation       1  18  2001 Yr    2 of   19 Time  17:22:46
```

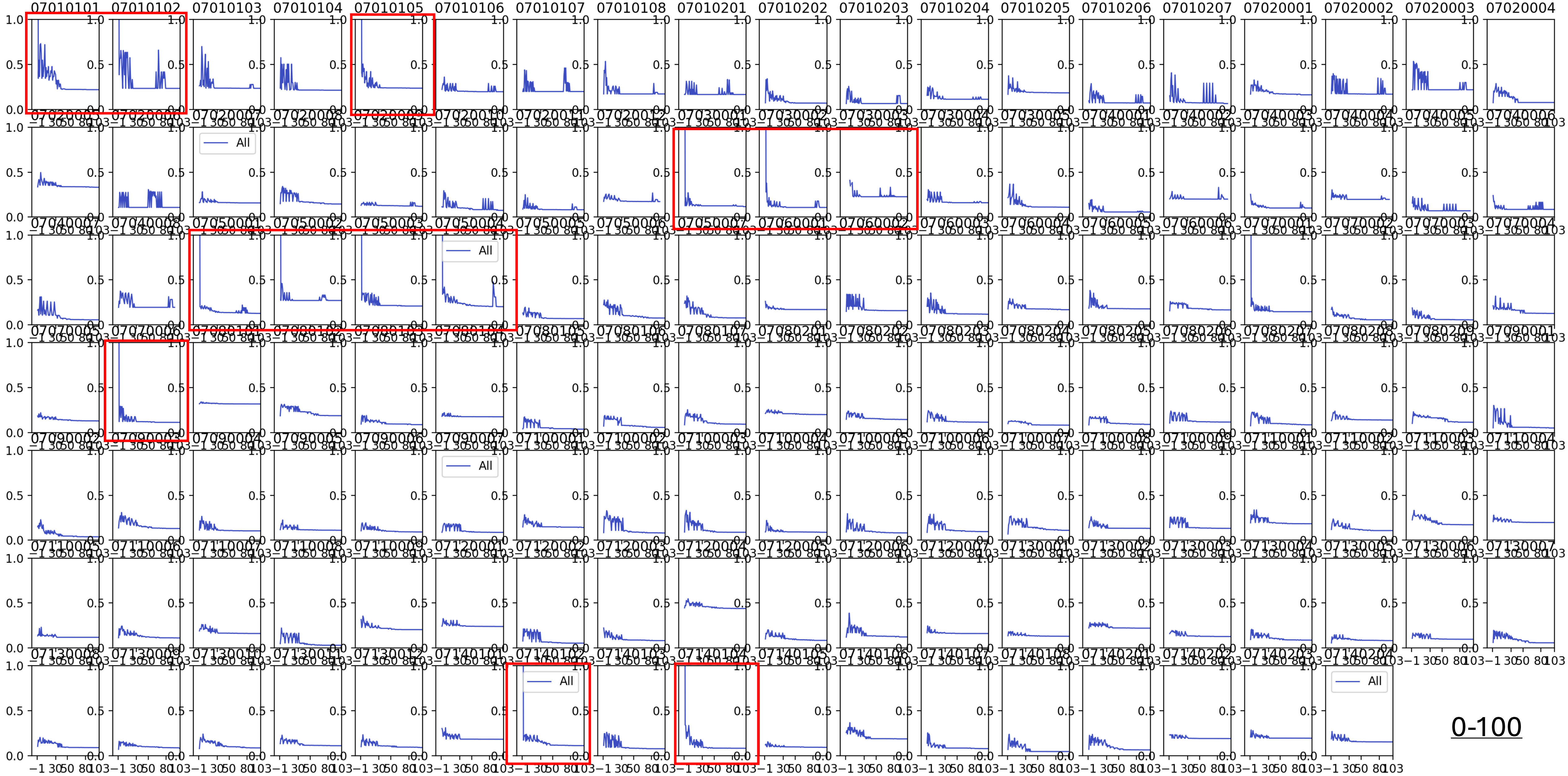
Single core running



```
G:\Tasks\lab2\Models\08030205\08030205_SWAT.exe
Original Simulation       5  28  2002 Yr    3 of   19 Time  17:16: 9
Original Simulation       5  29  2002 Yr    3 of   19 Time  17:16:32
Original Simulation       5  30  2002 Yr    3 of   19 Time  17:16:32
Original Simulation       5  31  2002 Yr    3 of   19 Time  17:16:32
Original Simulation       6   1  2002 Yr    3 of   19 Time  17:16:32
Original Simulation       6   2  2002 Yr    3 of   19 Time  17:16:33
Original Simulation       6   3  2002 Yr    3 of   19 Time  17:16:33
Original Simulation       6   4  2002 Yr    3 of   19 Time  17:16:33
Original Simulation       6   5  2002 Yr    3 of   19 Time  17:16:33
Original Simulation       6   6  2002 Yr    3 of   19 Time  17:16:33
Original Simulation       6   7  2002 Yr    3 of   19 Time  17:16:33
Original Simulation       6   8  2002 Yr    3 of   19 Time  17:16:33
Original Simulation       6   9  2002 Yr    3 of   19 Time  17:16:34
Original Simulation       6  10  2002 Yr    3 of   19 Time  17:16:34
Original Simulation       6  11  2002 Yr    3 of   19 Time  17:16:34
Original Simulation       6  12  2002 Yr    3 of   19 Time  17:16:34
Original Simulation       6  13  2002 Yr    3 of   19 Time  17:16:34
Original Simulation       6  14  2002 Yr    3 of   19 Time  17:16:34
Original Simulation       6  15  2002 Yr    3 of   19 Time  17:16:35
Original Simulation       6  16  2002 Yr    3 of   19 Time  17:16:35
Original Simulation       6  17  2002 Yr    3 of   19 Time  17:16:35
Original Simulation       6  18  2002 Yr    3 of   19 Time  17:16:35
Original Simulation       6  19  2002 Yr    3 of   19 Time  17:16:35
Original Simulation       6  20  2002 Yr    3 of   19 Time  17:16:35
Original Simulation       6  21  2002 Yr    3 of   19 Time  17:16:35
Original Simulation       6  22  2002 Yr    3 of   19 Time  17:16:36
Original Simulation       6  23  2002 Yr    3 of   19 Time  17:16:36
Original Simulation       6  24  2002 Yr    3 of   19 Time  17:16:36
Original Simulation       6  25  2002 Yr    3 of   19 Time  17:16:36
```

4. Calibration Results

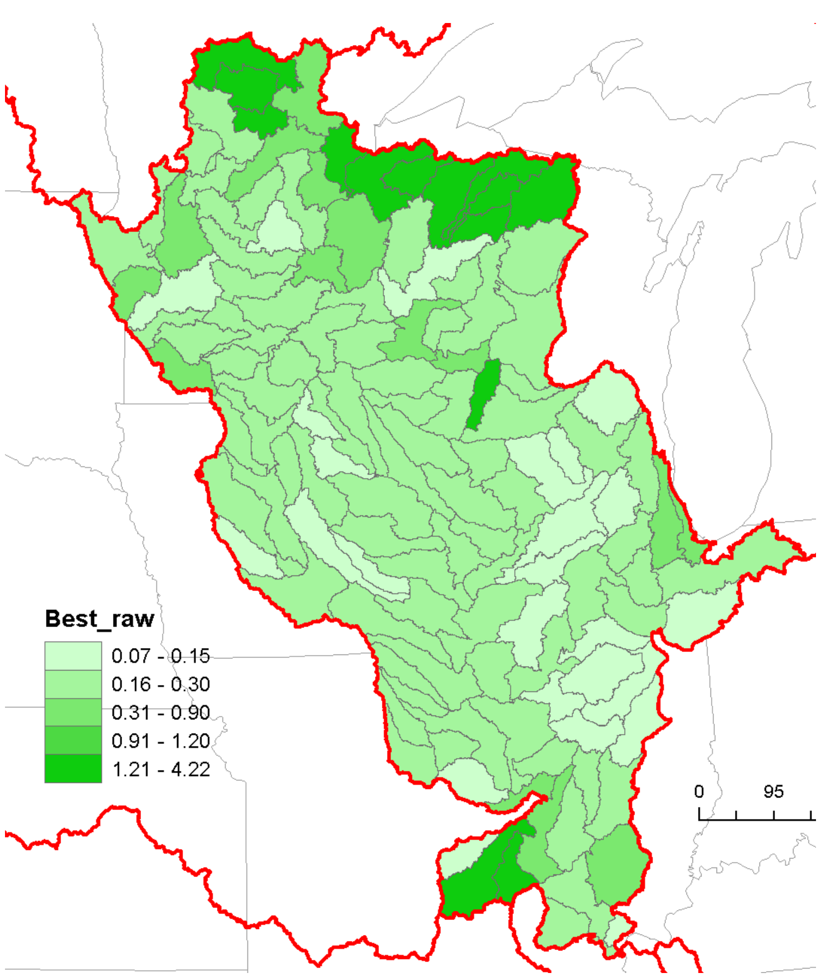
4.1 Objection function convergence in around 100 iterations for 131 HUC8s



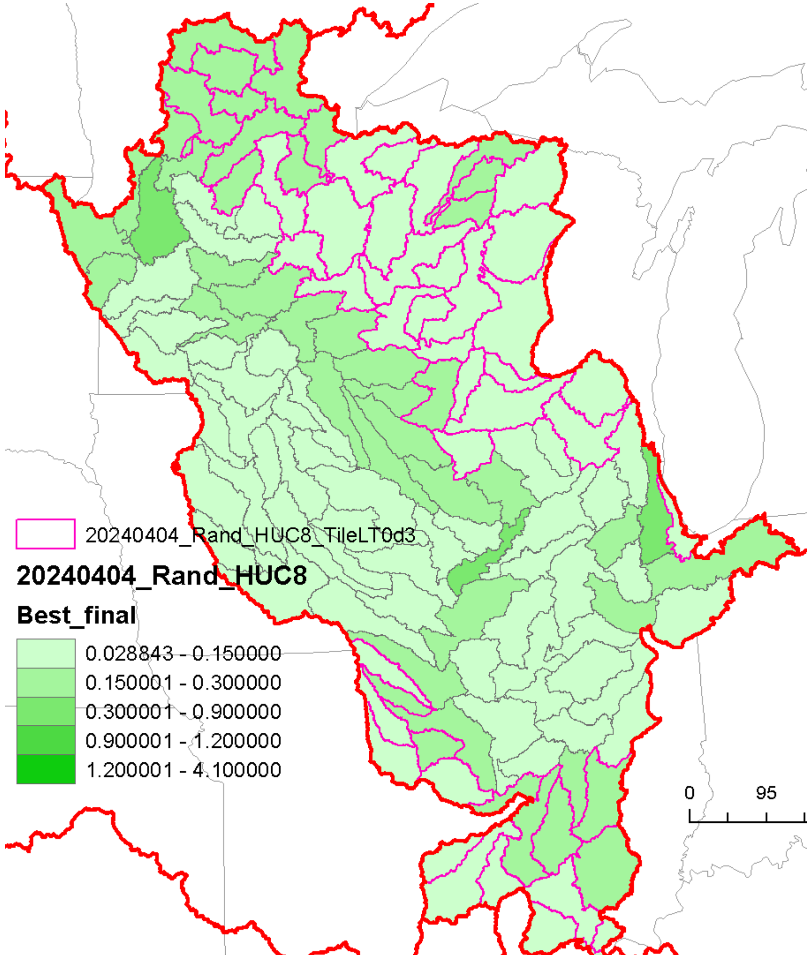
0-100

4.2 Final improvement by DDS

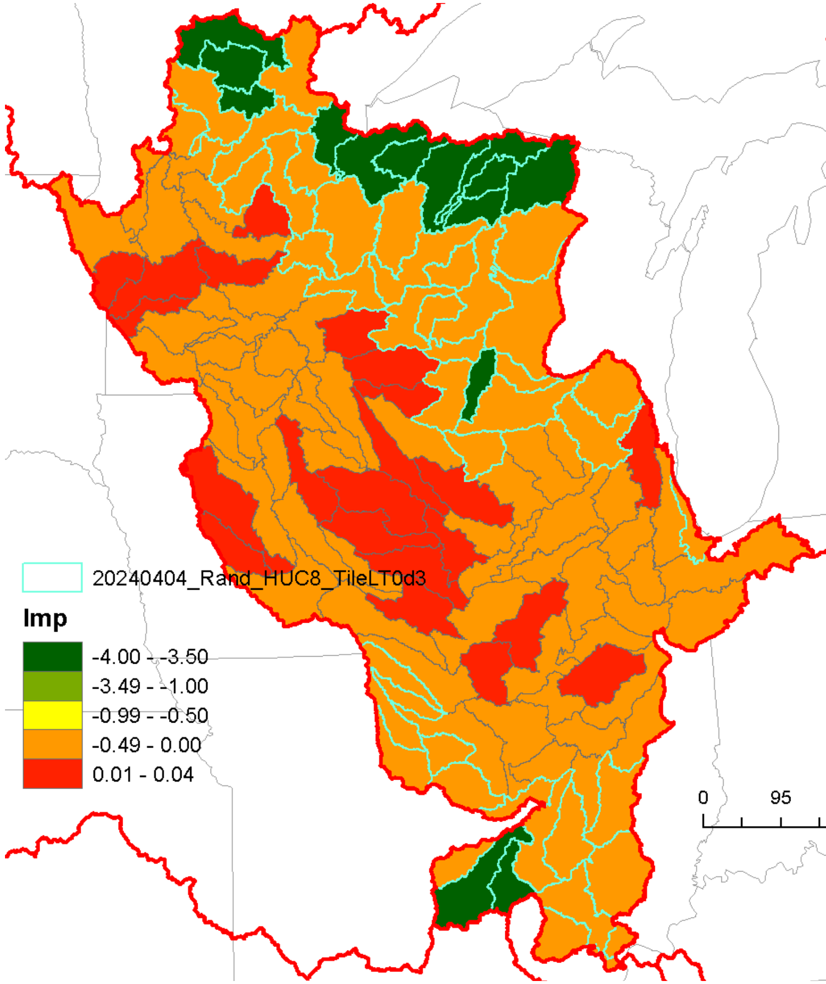
Initial



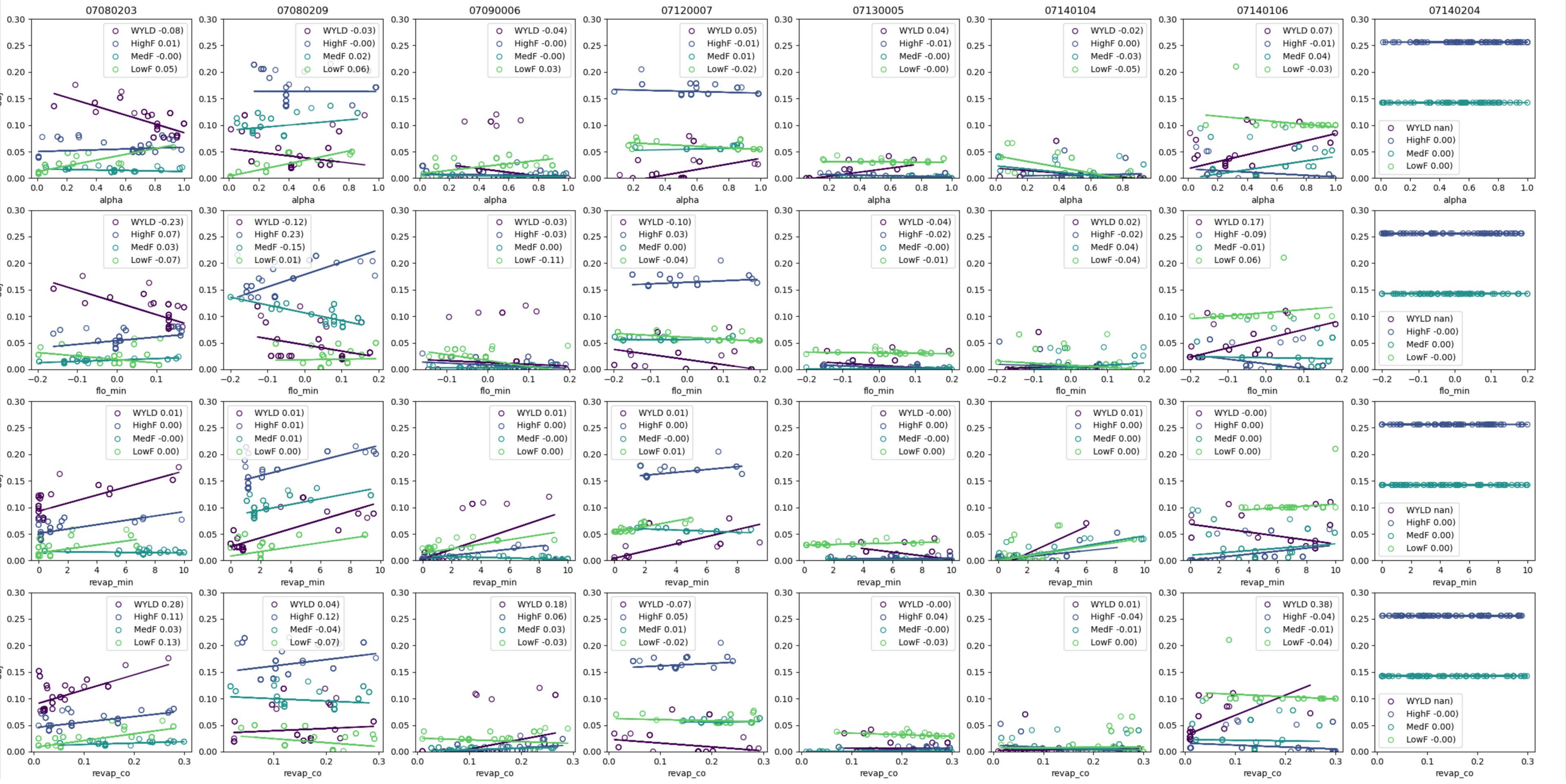
Best



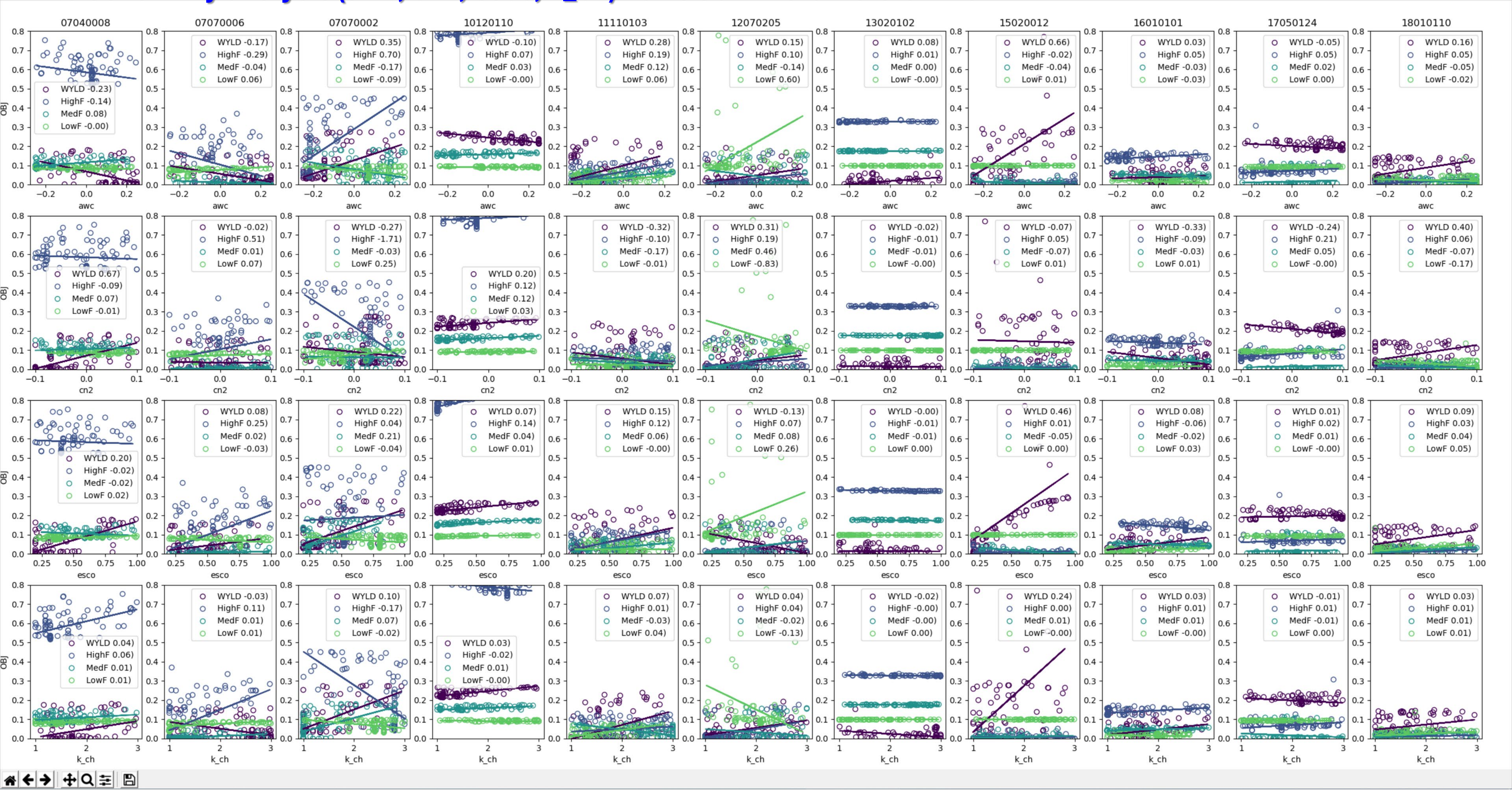
Improvement



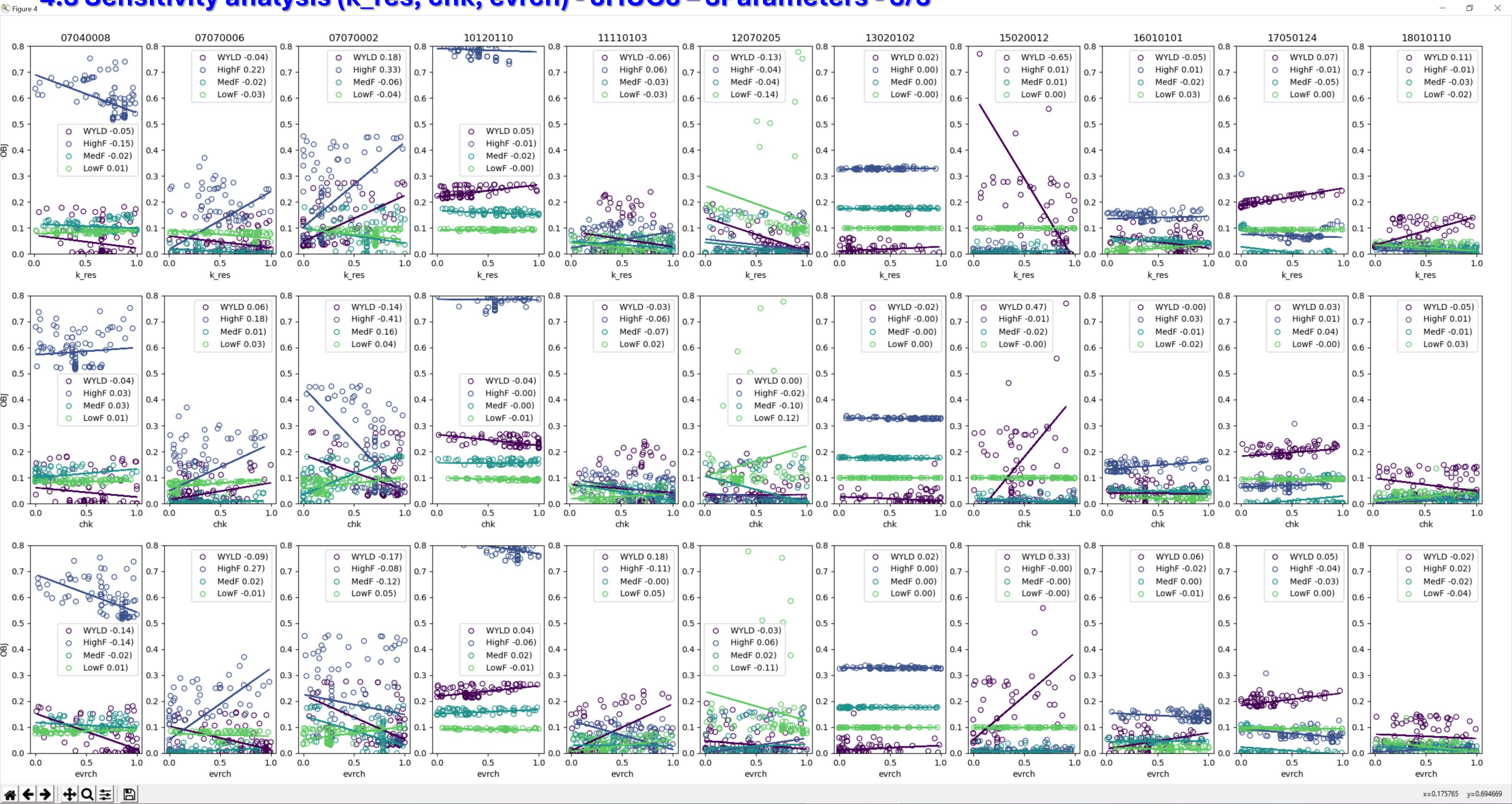
4.3 Sensitivity analysis (alpha, flow_min, revap_min & revap_co) - 8HUC8 – 4Parameters - 1/3



4.3 Sensitivity analysis (awc, cn2, esco, k_ch) - 8HUC8 – 4Parameters - 2/3



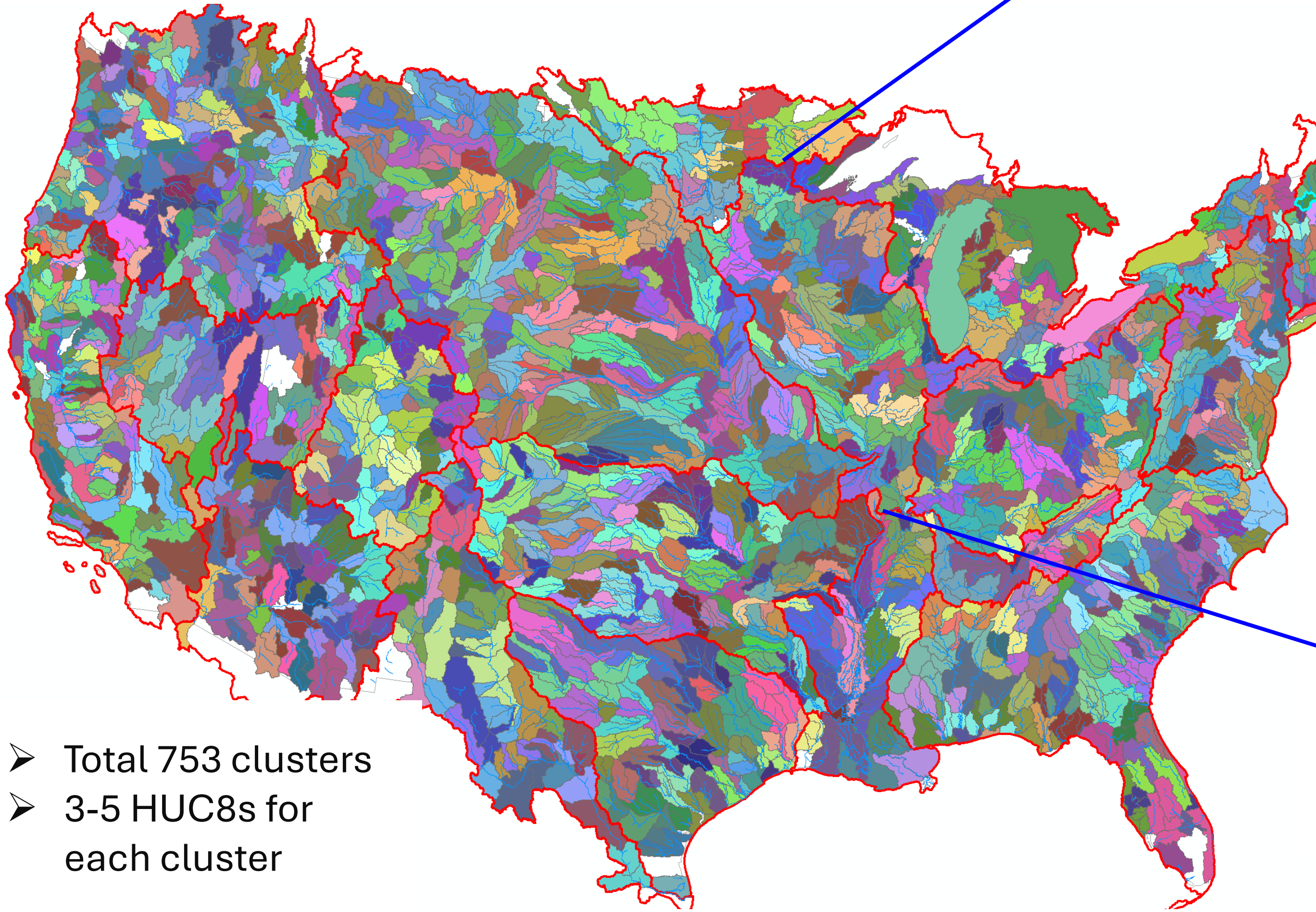
4.3 Sensitivity analysis (k_res, chk, evrch) - 8HUC8 – 3Parameters - 3/3

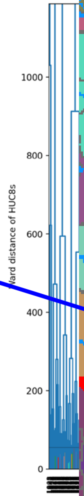


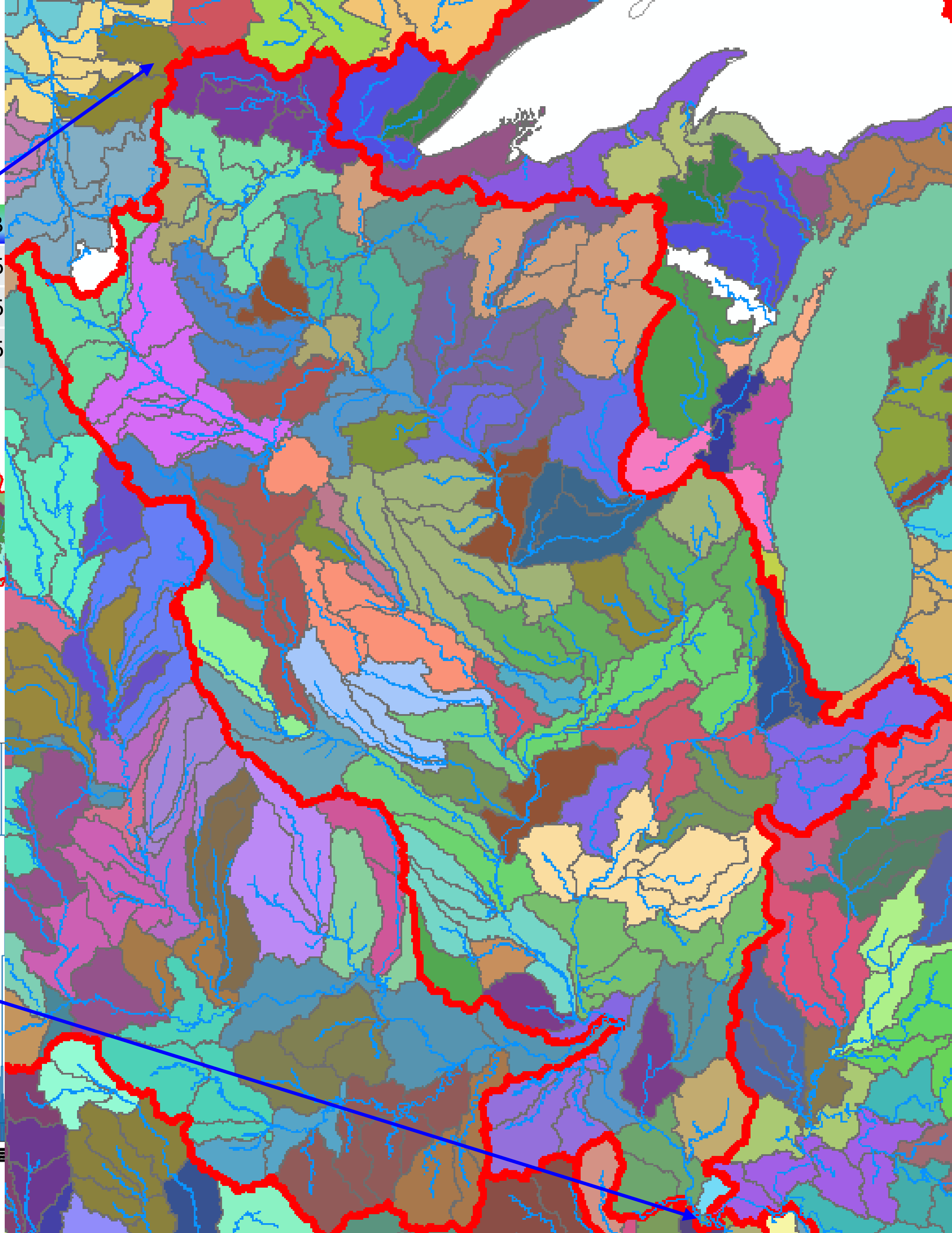
5. Future Work

Clustering HUC8s for the whole

HUC8	BSN_Slope	Cult_Fraction	Urbn_Fraction	Frst_Fraction	BFI	WB
1010002	7.2	0.0	1.6	82.2	49.7	5
1010003	8.4	2.3	2.3	65.6	53.7	5
1010004	5.7	4.1	2.8	66.0	52.8	5







- Total 753 clusters
- 3-5 HUC8s for each cluster

Thank you!

CONTACT
Jungang Gao
Blackland Research and Extension Center
Texas A&M AgriLife
Email: Jungang.gao@agnet.tamu.edu

Questions?