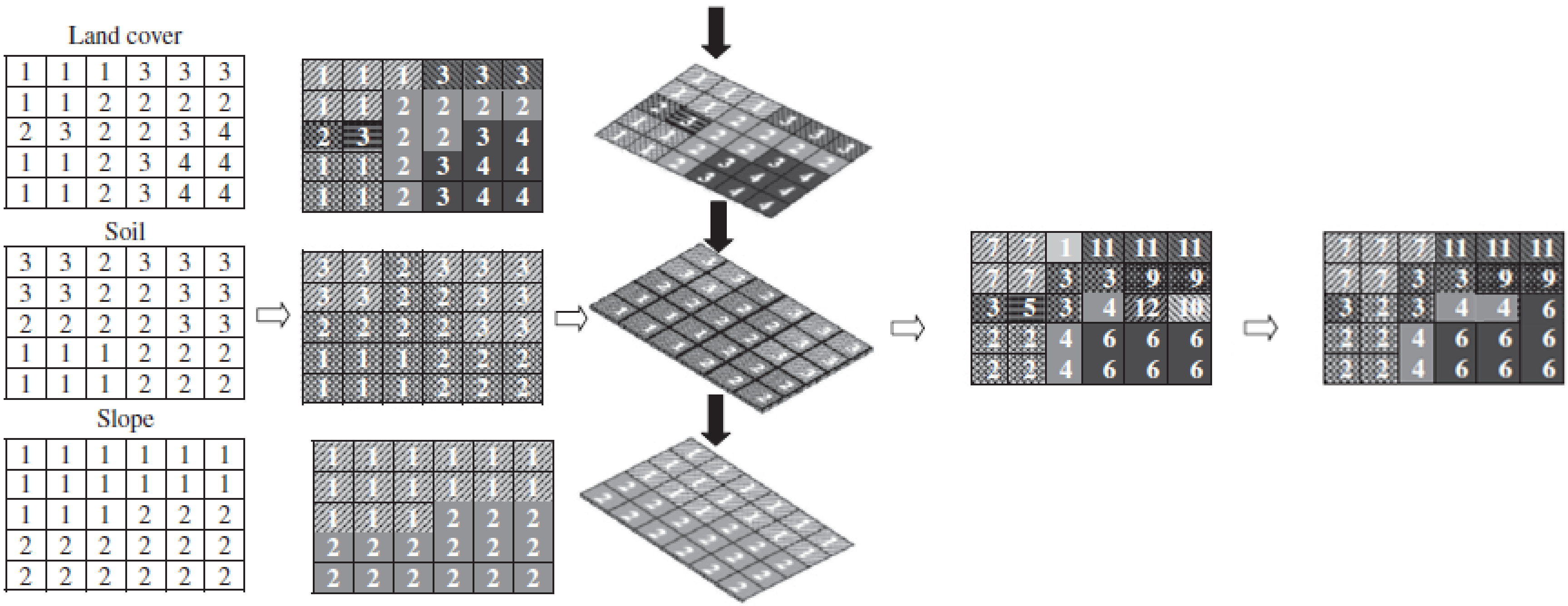


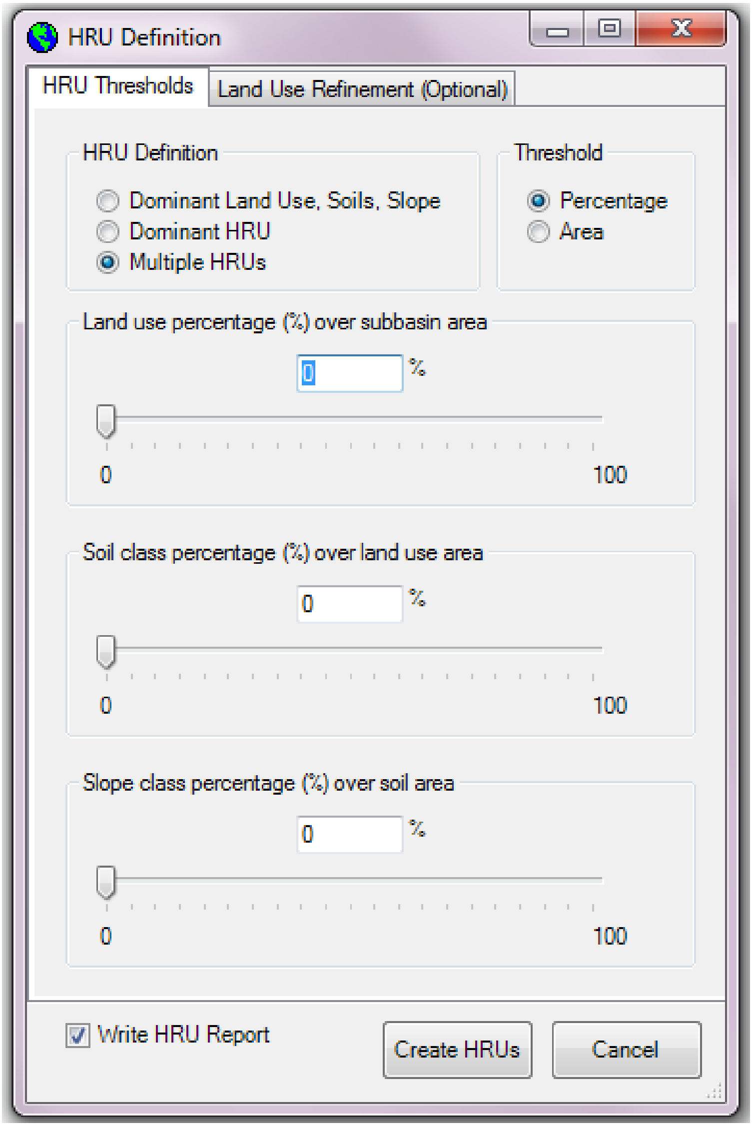
HRU aggregation and its effects on model outputs

Michael Strauch, Robert Otto, Martin Volk

2015 International SWAT Conference
Sardinia, June 24–26, 2015

HELMHOLTZ
ZENTRUM FÜR
UMWELTFORSCHUNG
UFZ

HRU definition in ArcSWAT



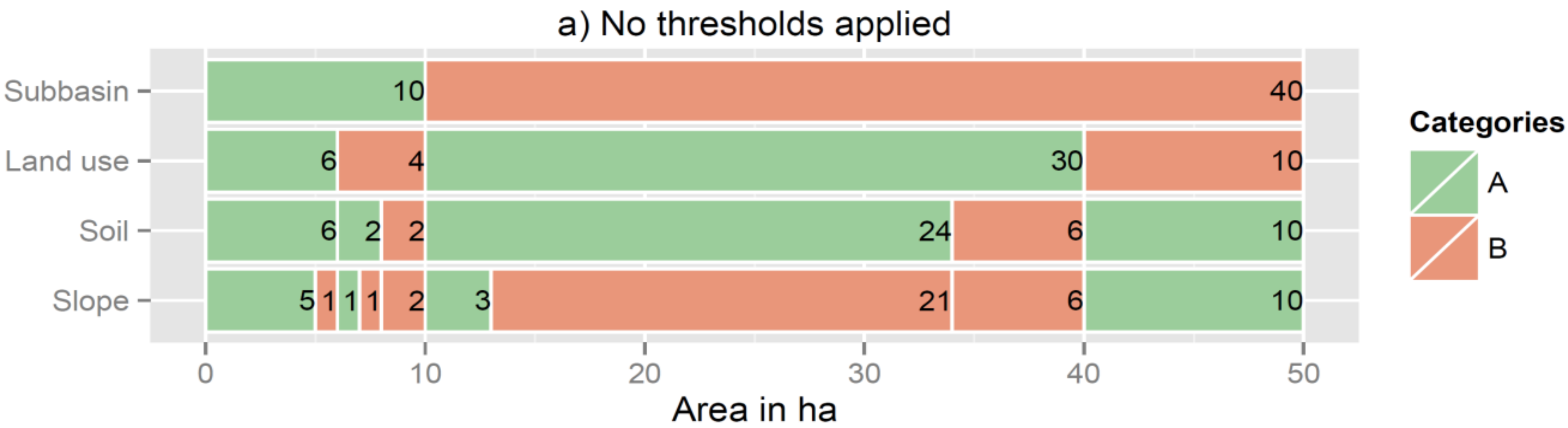
Which threshold values should be used?

”For most applications, the default settings for **land use threshold (20%)** and **soil threshold (10%)** and **slope threshold (20%)** are adequate.” (Winchell et al., 2007, p. 126)

HELMHOLTZ ZENTRUM FÜR UMWELTFORSCHUNG UFZ

HRU definition in ArcSWAT

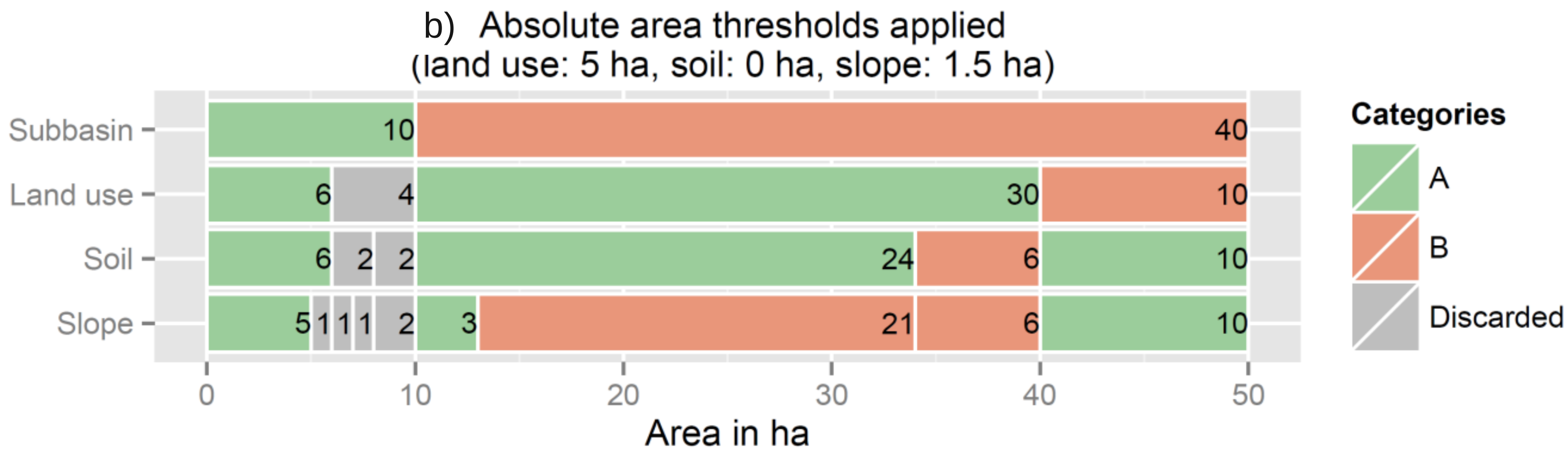
Theoretical example



HELMHOLTZ ZENTRUM FÜR UMWELTFORSCHUNG UFZ

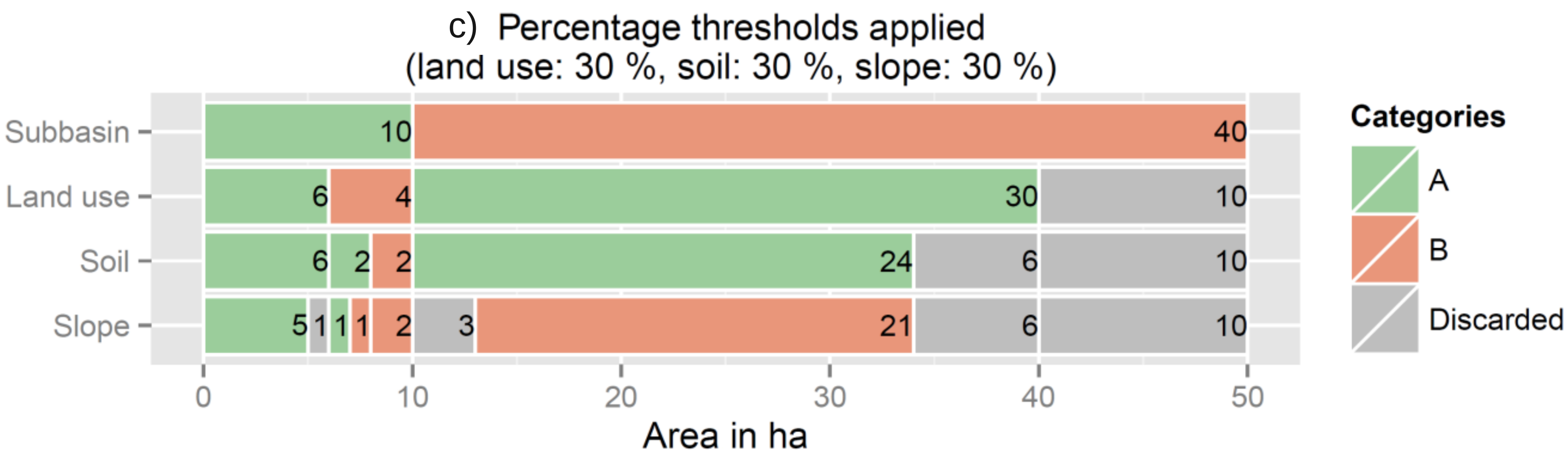
HRU definition in ArcSWAT

Theoretical example



HRU definition in ArcSWAT

Theoretical example



Error measure

Average Absolute Error of Aggregation:

(10 ha (Lu) + 6 ha (Soil) + 14 ha (Slope))/3 = 10 ha

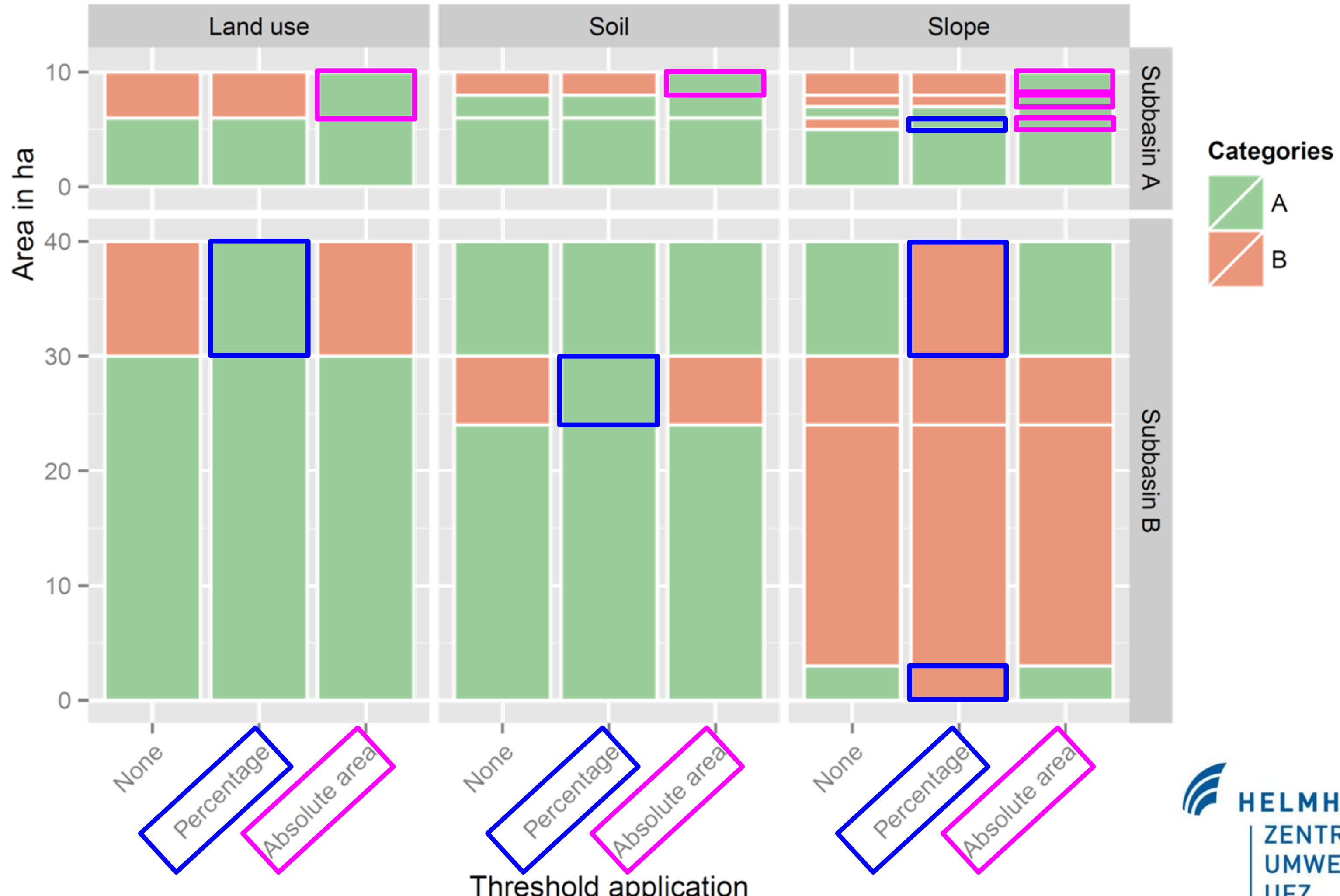
(4 ha (Lu) + 2 ha (Soil) + 4 ha (Slope))/3 = 3.34 ha

Relative Error:

10 ha / 50 ha = 0.2

3.34 ha / 50 ha = 0.067

Theoretical example



HELMHOLTZ ZENTRUM FÜR UMWELTFORSCHUNG UFZ

Working hypotheses

1. ***HRU number reduction can significantly distort input data.***

2. ***The larger the input data error, the larger is the model error*** *(assuming the model equations are right).*

3. ***Different threshold combinations can lead to similar totals of HRUs with different input data errors*** *and thus different model errors.*

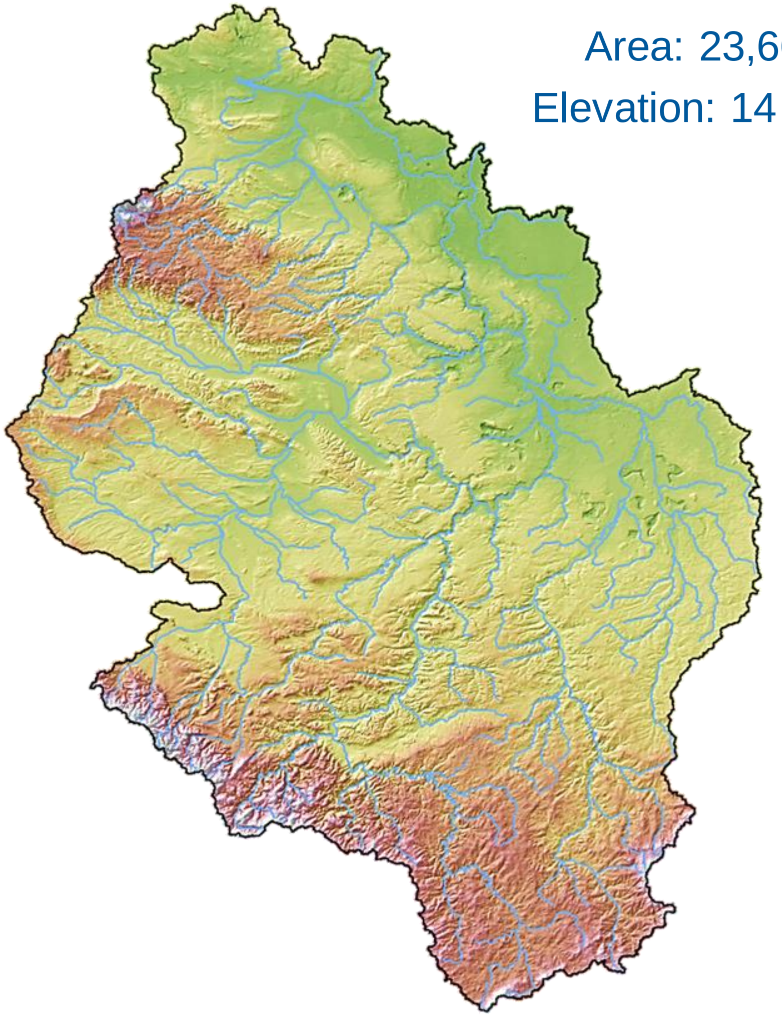
HELMHOLTZ ZENTRUM FÜR UMWELTFORSCHUNG UFZ

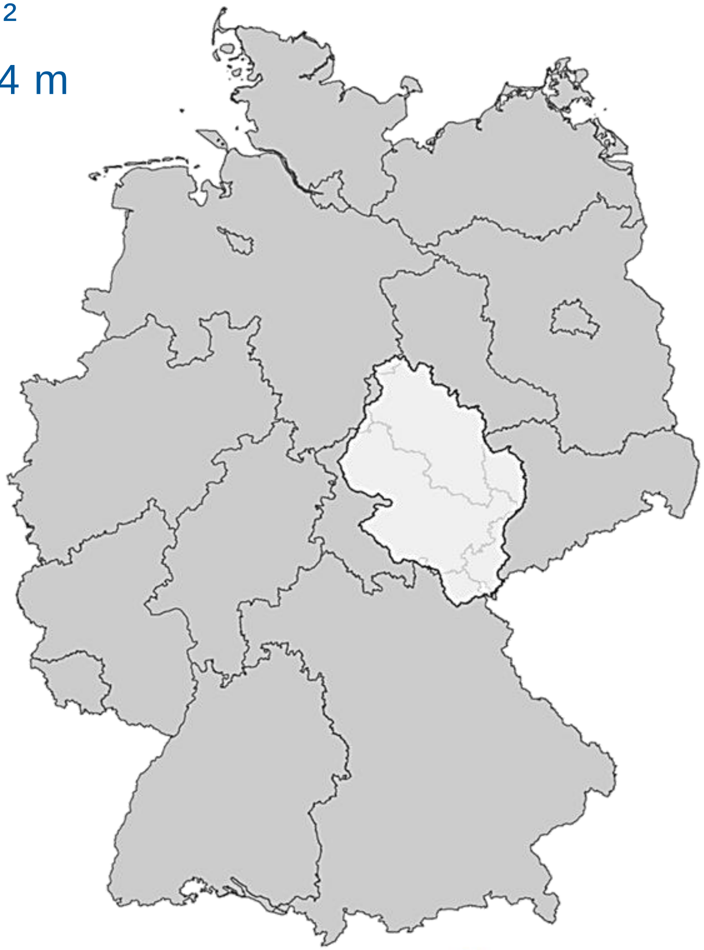
Study catchment 1

Saale River Basin

Area: 23,666 km²

Elevation: 14 – 1144 m





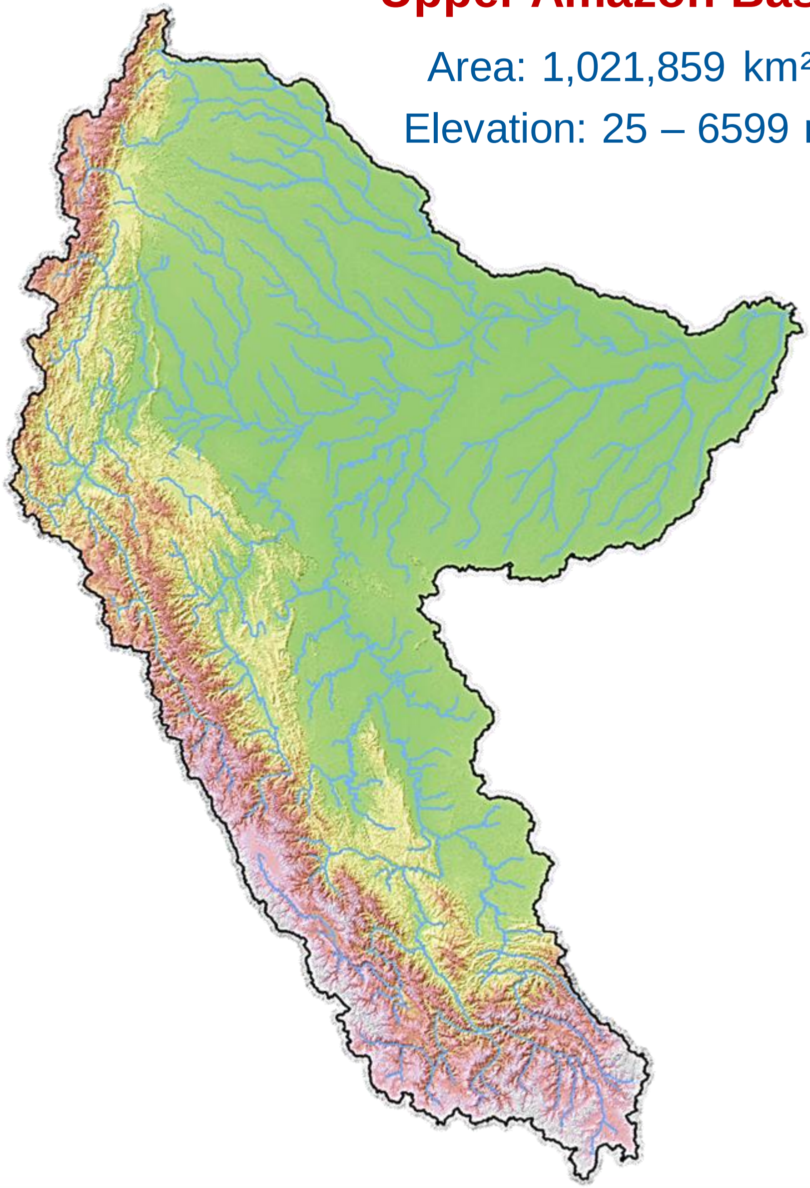
HELMHOLTZ
ZENTRUM FÜR
UMWELTFORSCHUNG
UFZ

Study catchment 2

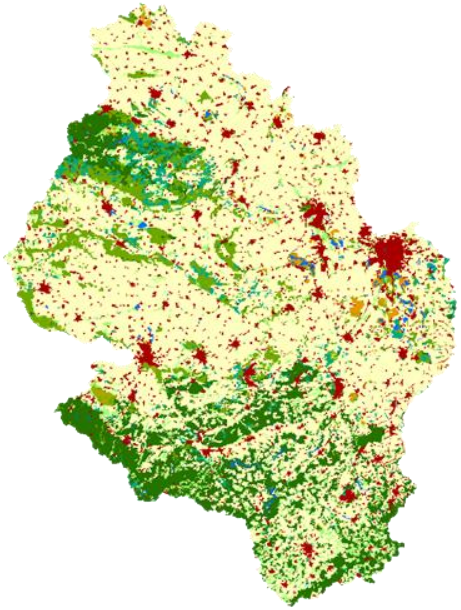
Upper Amazon Basin

Area: 1,021,859 km²

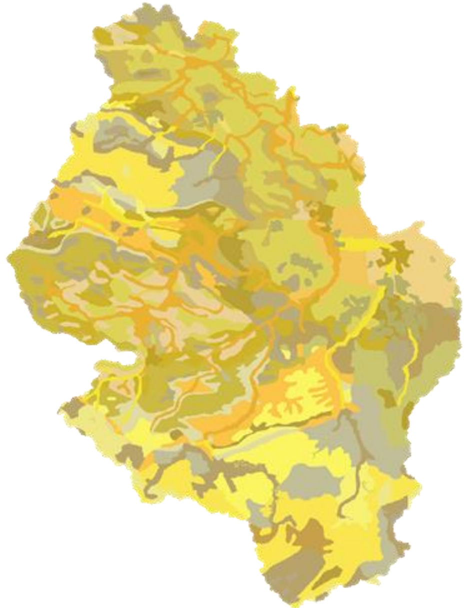
Elevation: 25 – 6599 m



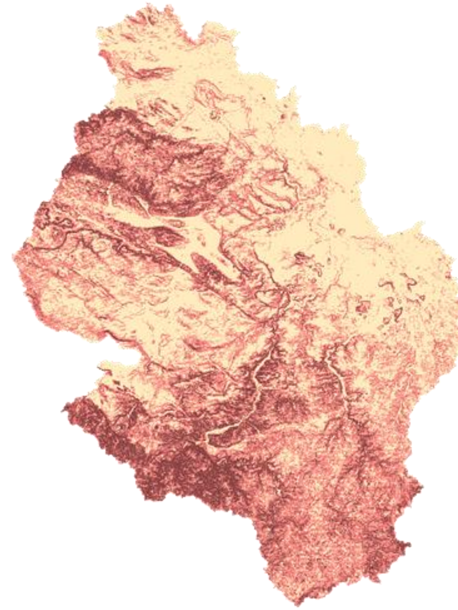
Land use (Corine)



Soils (BÜK 1000)



Slope (n=3)



Subbasins (n=40)

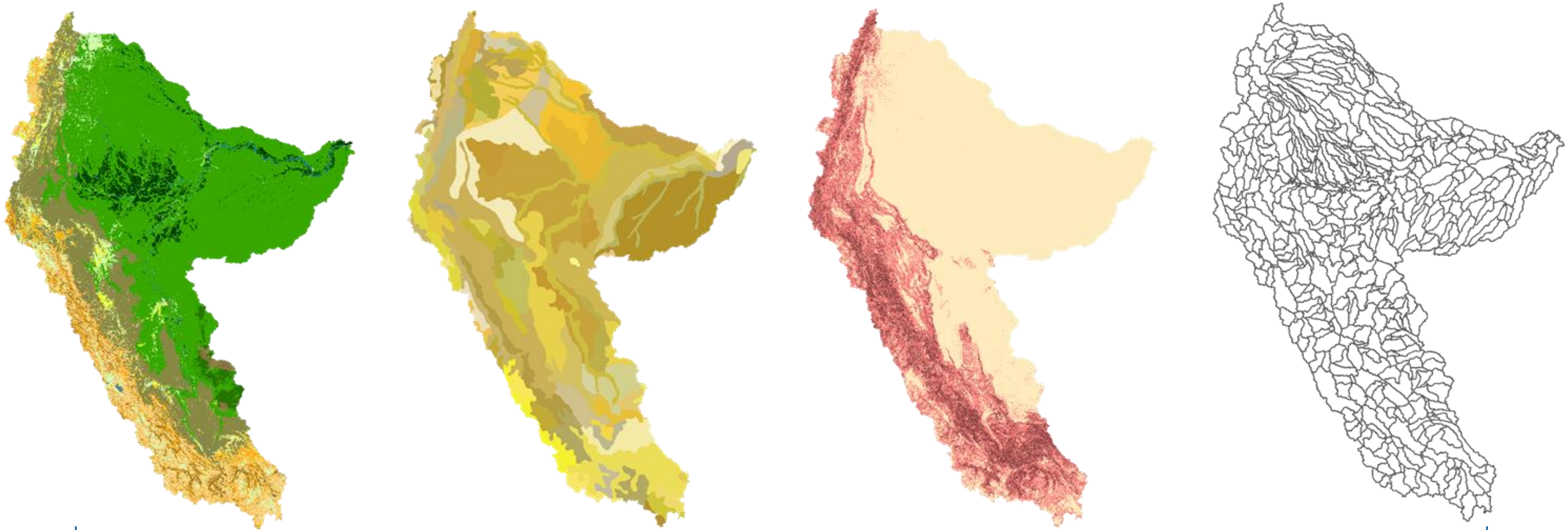
Saale River Basin

full HRU Model

HRUs (n=5,256)

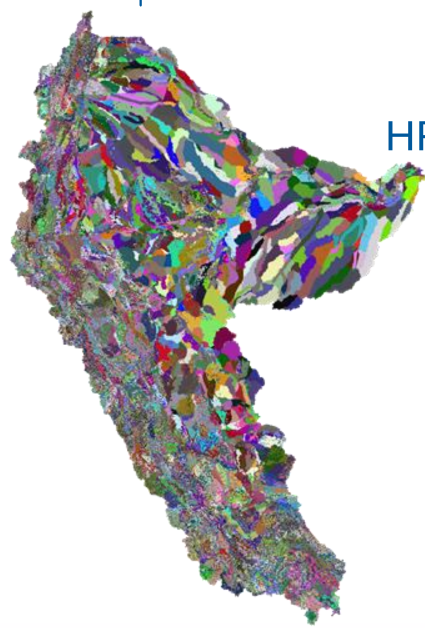
HELMHOLTZ
ZENTRUM FÜR
UMWELTFORSCHUNG
UFZ

Land use (GLC)   Soils (HWSD)   Slope (n=3)   Subbasins (n=539)



Upper Amazon Basin

full HRU Model



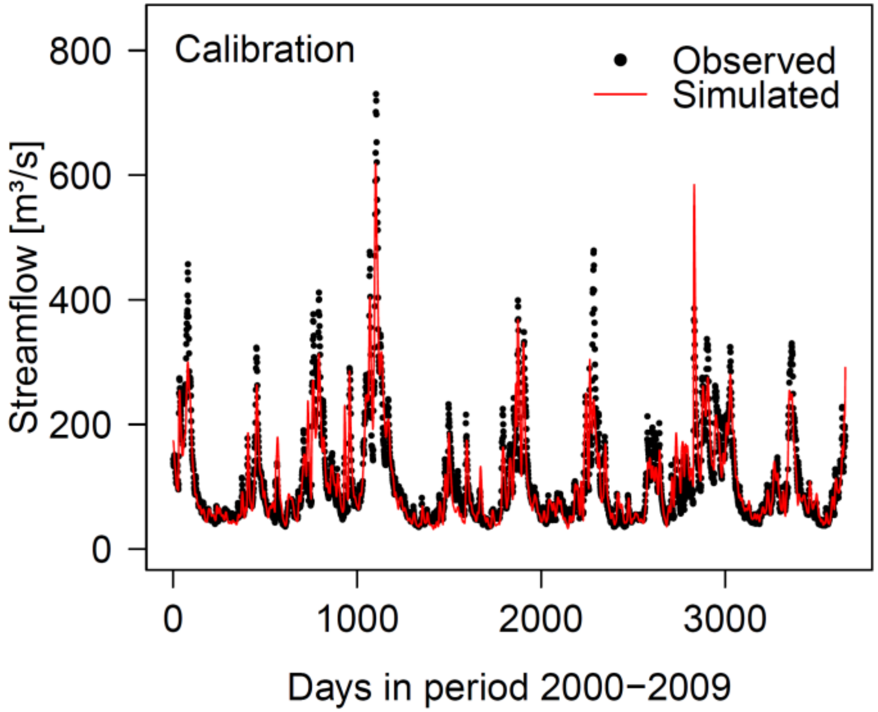
HRUs (n=15,989)

HELMHOLTZ ZENTRUM FÜR UMWELTFORSCHUNG UFZ

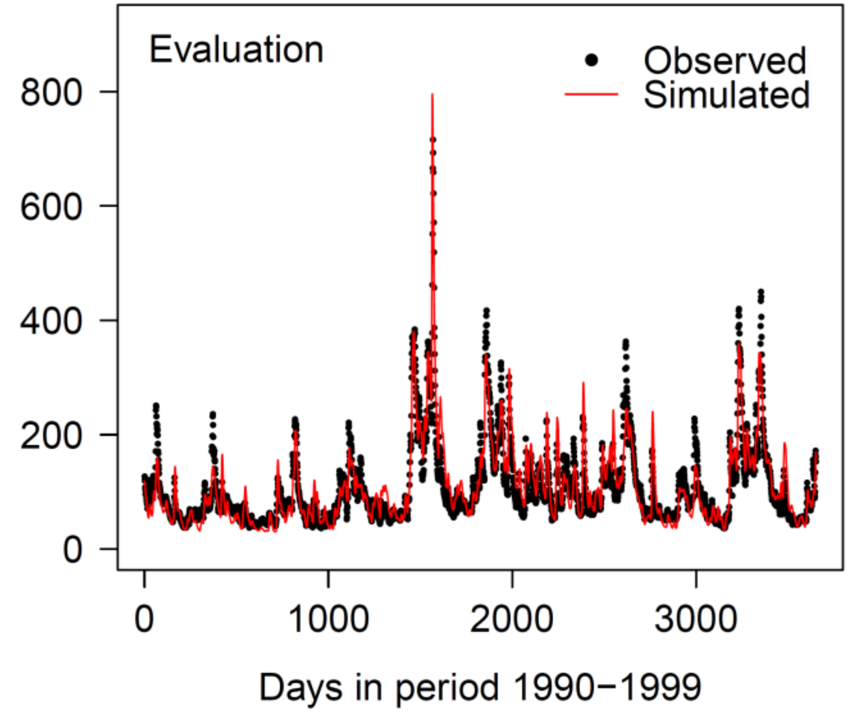
Model calibration (full HRU model)

Saale

NSE: 0.84
PBIAS: 0.2



NSE: 0.82
PBIAS: 3.4



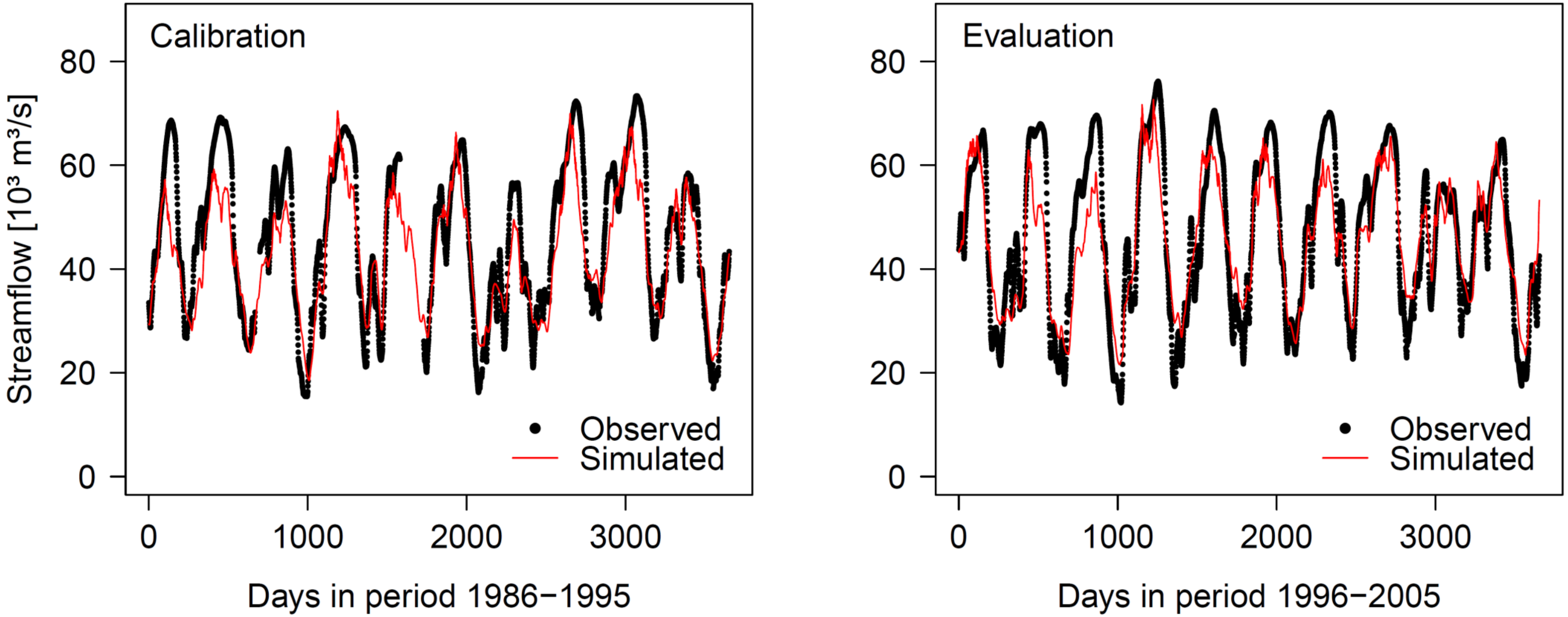
HELMHOLTZ
| ZENTRUM FÜR

Model calibration (full HRU model)

Upper Amazon

NSE: 0.75
PBIAS: -6.5

NSE: 0.77
PBIAS: -2.6



Method:

Define HRUs with thousands of different threshold combinations using R

⇒ **Input for R-Script:** *hrus* table (project.mdb) without thresholds

Table

hrus

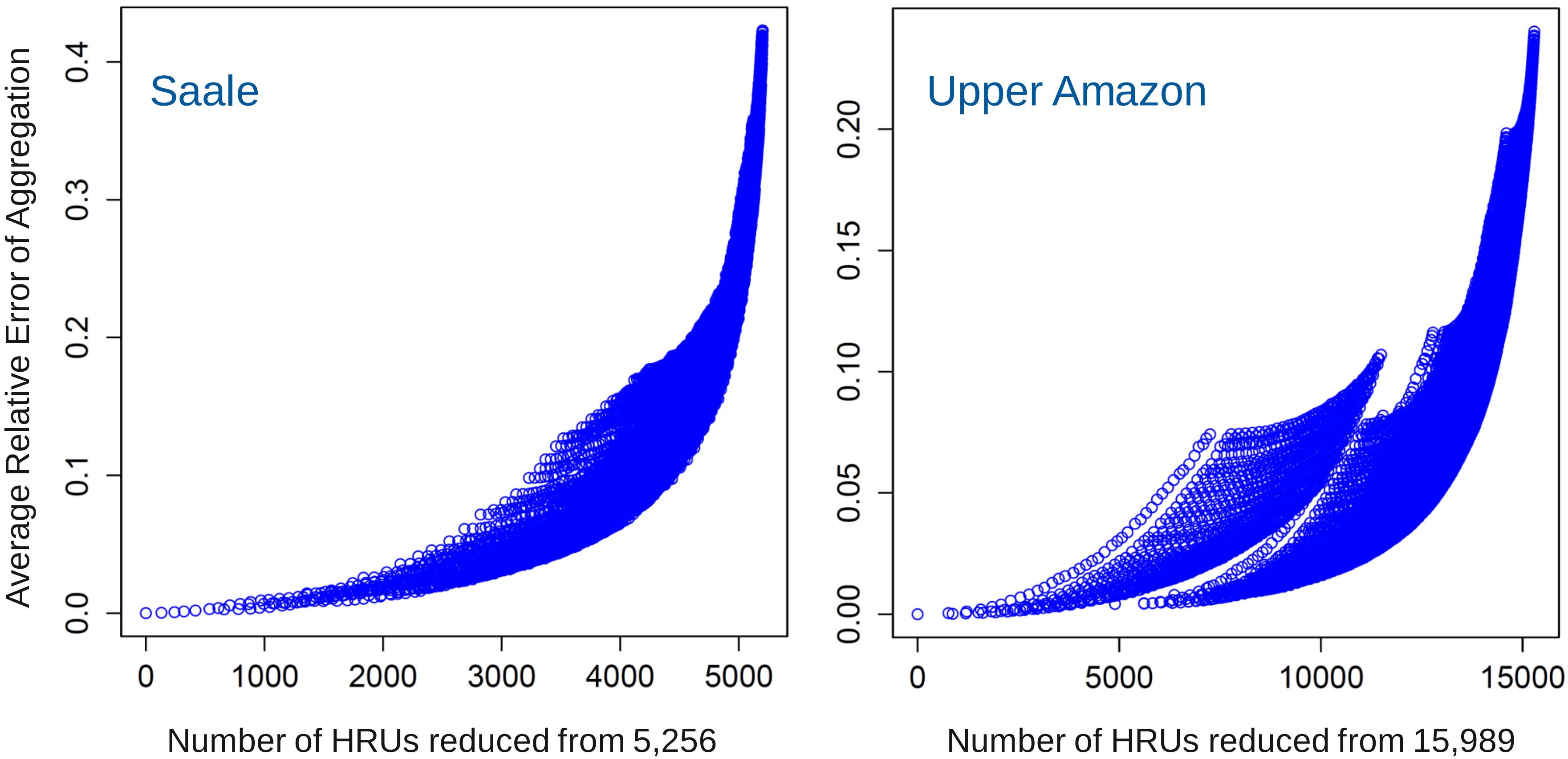
OID *	SUBBASIN *	ARSUB	LANDUSE	ARLU	SOIL	ARSO	SLP	ARSLP	SLOPE	UNIQUECOMB	HRU_ID	HRU_GIS
420	76	3493797.677604	RNGE	1522435.270746	PE52	273096.551882	0-15.45	119163.134603	9.122732	76_RNGE_PE52_0-15.45	420	000760015
421	76	3493797.677604	RNGE	1522435.270746	PE52	273096.551882	15.45-43.54	153933.417279	25.671921	76_RNGE_PE52_15.45-43.54	421	000760016
422	76	3493797.677604	RNGE	1522435.270746	PE53	172976.854437	0-15.45	172976.854437	7.956806	76_RNGE_PE53_0-15.45	422	000760017
423	76	3493797.677604	RNGE	1522435.270746	PE54	528817.545667	15.45-43.54	321808.514489	26.727308	76_RNGE_PE54_15.45-43.54	423	000760018
424	76	3493797.677604	RNGE	1522435.270746	PE54	528817.545667	0-15.45	207009.031178	9.140409	76_RNGE_PE54_0-15.45	424	000760019
425	76	3493797.677604	RNGE	1522435.270746	PE55	375618.211774	15.45-43.54	375618.211774	27.959536	76_RNGE_PE55_15.45-43.54	425	000760020
426	76	3493797.677604	RNGE	1522435.270746	PE65	53555.228567	15.45-43.54	53555.228567	28.941126	76_RNGE_PE65_15.45-43.54	426	000760021
427	76	3493797.677604	RNGE	1522435.270746	PE78	118370.878419	15.45-43.54	118370.878419	28.321695	76_RNGE_PE78_15.45-43.54	427	000760022

⇒ **the R-script then…**

- defines HRUs using *ha* or % threshold values for all threshold combinations (land use / soil / slope) within a pre-defined range
- calculates the average Relative Error of Aggregation (aREA) compared to the full HRU distribution

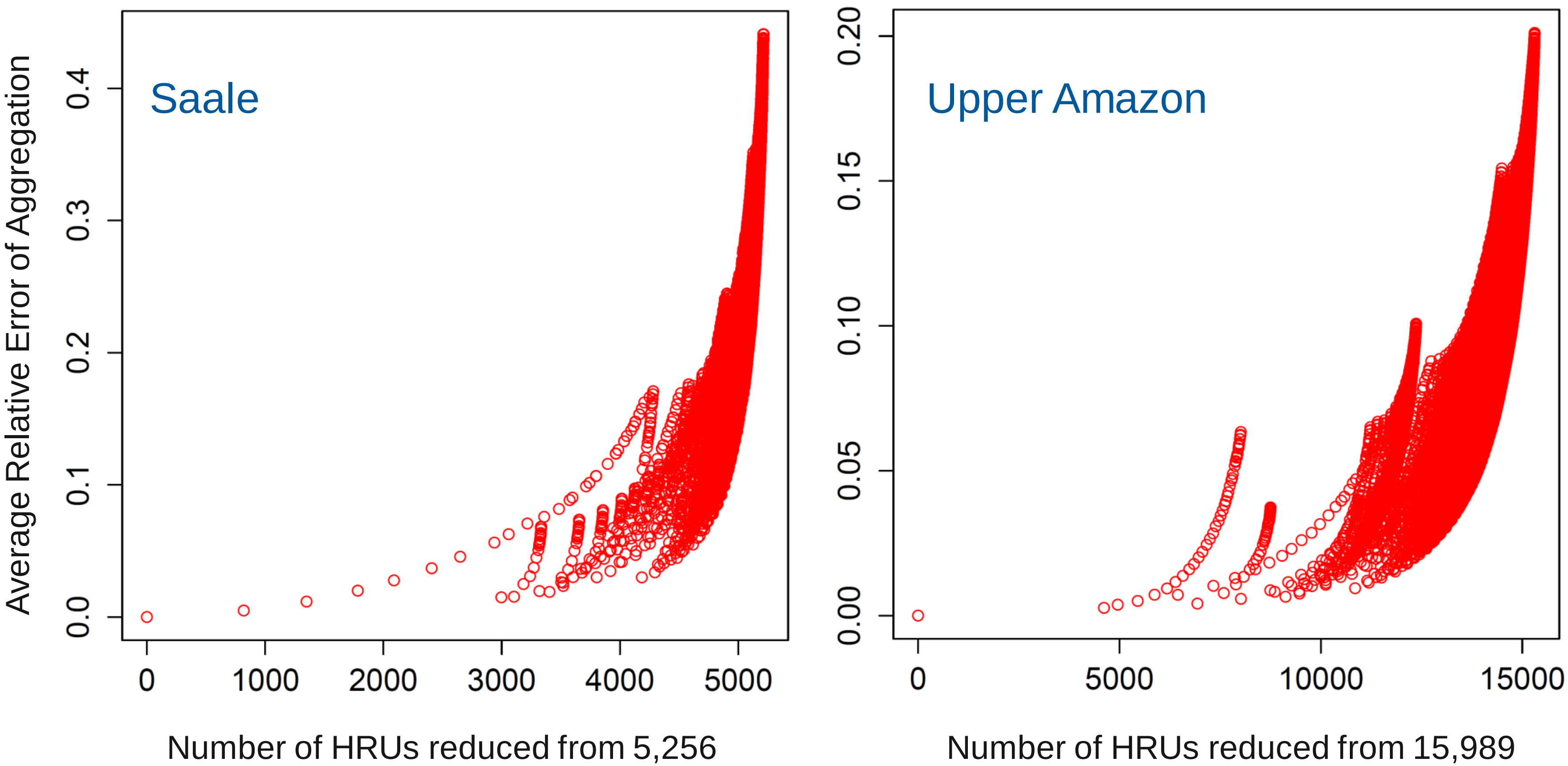
HRU-Definition using R

- percentage method
- land use 0-39%, soil 0-39%, slope 0-39%
- increment 1% (64,000 solutions)



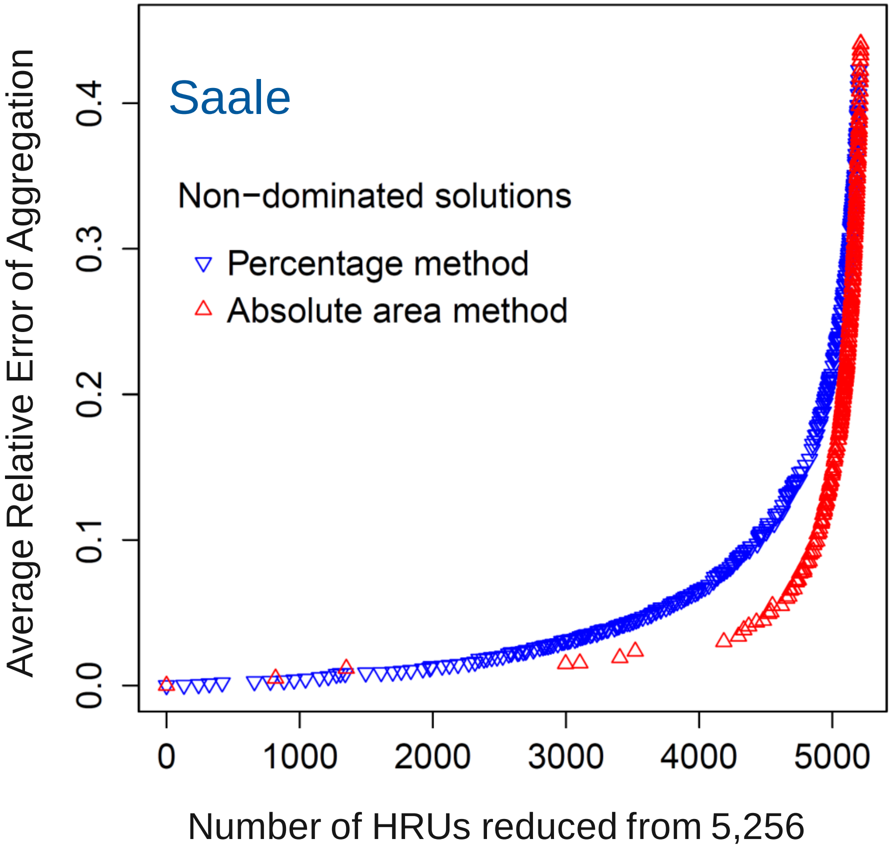
HRU-Definition using R

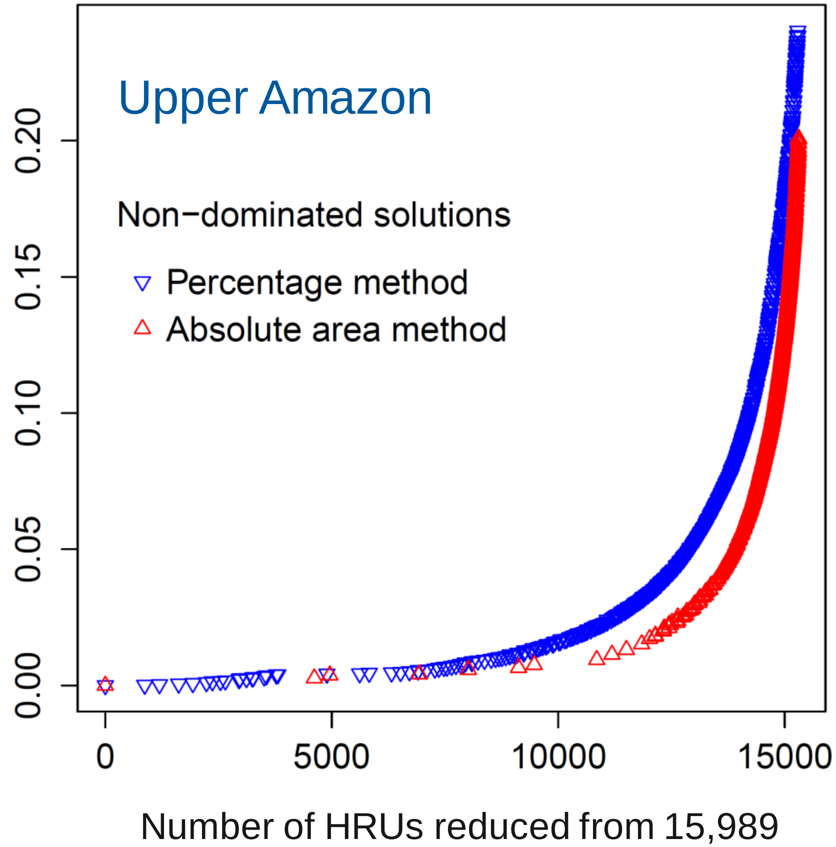
- absolute area method
- range 0-23,010 ha (Saale) / 0-74,100 ha (Amazon)
- increment 590 ha (Saale) / 1,900 ha (Amazon) = 64,000 solutions



HRU-Definition using R

- non-dominated solutions as best solutions
- absolute area method is preferable over percentage method



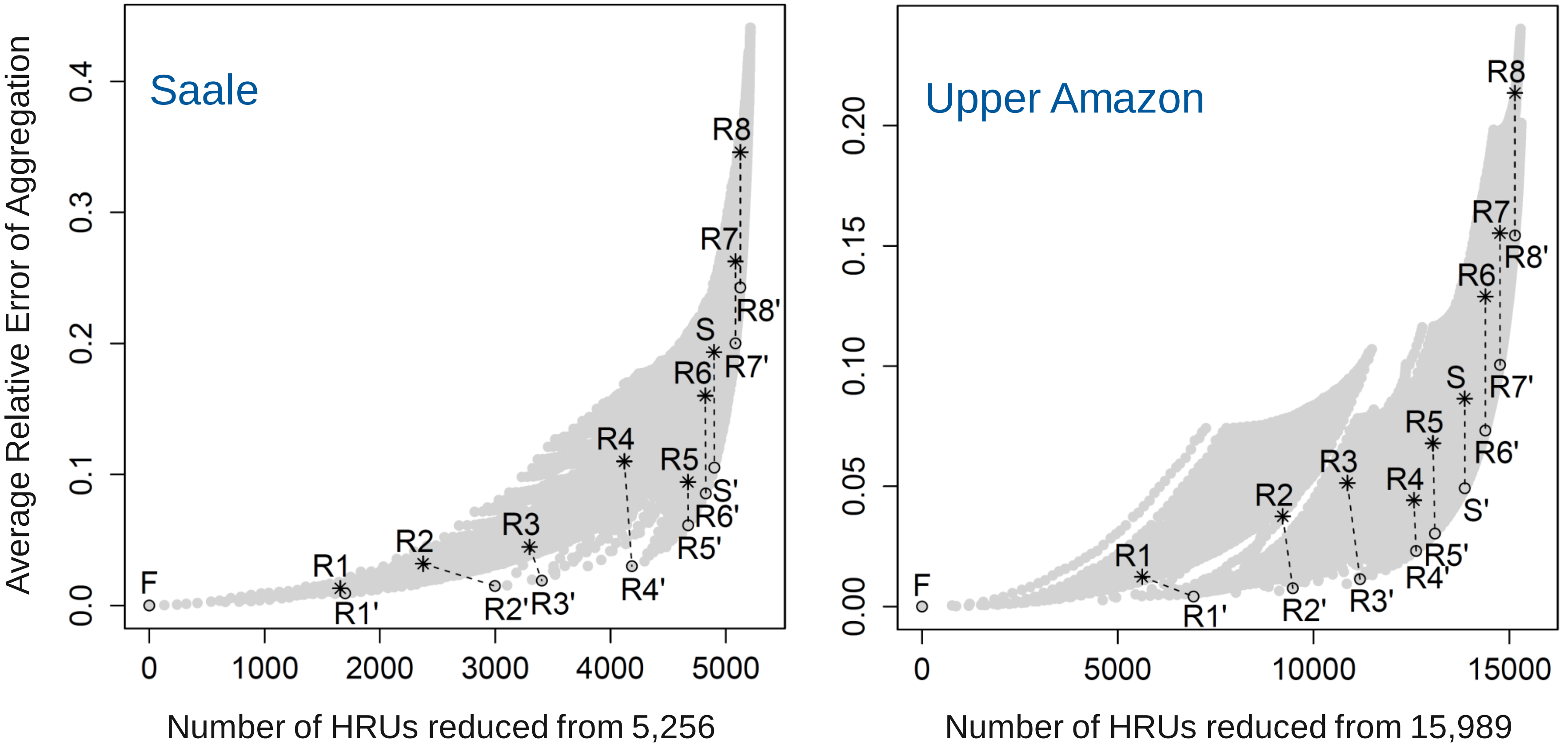


Study design to test effects on model outputs

- test 18 solutions out of 128,000 (all percentage & absolute area solutions)

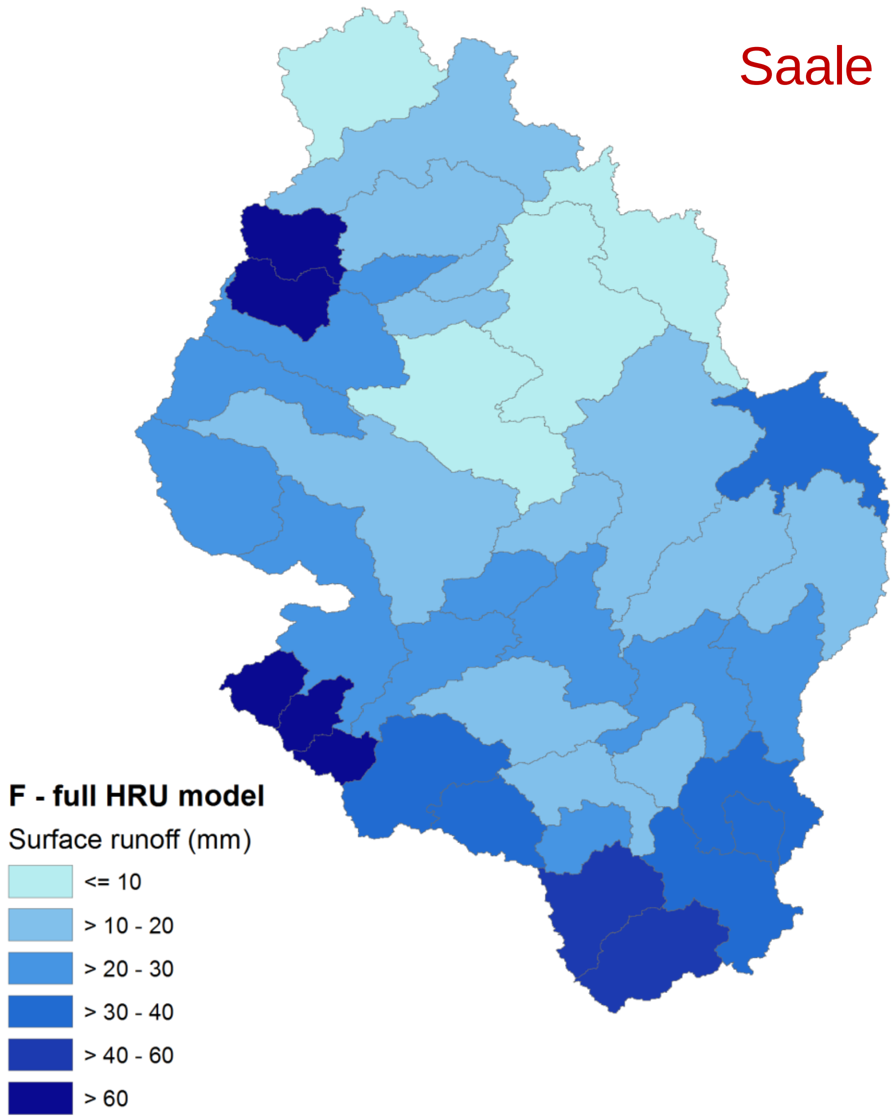
⇒ S = standard values solution (20%/10%/20%), S’ = respective best solution

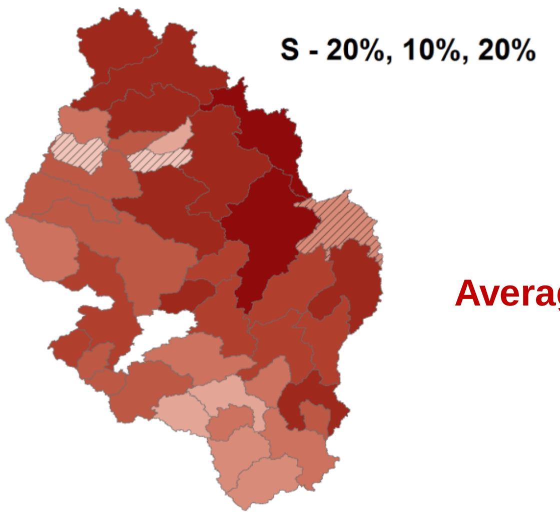
⇒ R1-R8 = stratified random samples, R1’-R8’ = respective best solutions



	Saale ($2.37\ 10^6$ ha)					Upper Amazon ($1.02\ 10^8$ ha)				
	Threshold combination					Threshold combination				
Solution	Land use	Soil	Slope	n HRUs	aREA	Land use	Soil	Slope	n HRUs	aREA
F	0%	0%	0%	5256	0.000	0%	0%	0%	15989	0.000
S	20%	10%	20%	359	0.194	20%	10%	20%	2126	0.087
R1	0%	3%	4%	3600	0.013	0%	1%	19%	10363	0.013
R2	0%	5%	8%	2881	0.032	1%	23%	0%	6775	0.038
R3	1%	5%	15%	1958	0.045	16%	0%	0%	5125	0.052
R4	17%	1%	2%	1136	0.110	26600 ha	0 ha	3800 ha	3422	0.044
R5	1180 ha	590 ha	13570 ha	585	0.094	15%	16%	4%	2938	0.068
R6	9%	7%	37%	435	0.160	36%	7%	21%	1600	0.129
R7	20060 ha	2360 ha	20060 ha	173	0.263	17%	36%	35%	1226	0.156
R8	31%	38%	4%	131	0.346	39%	39%	23%	843	0.214
S'	2360 ha	2360 ha	2360 ha	356	0.105	7600 ha	15200 ha	15200 ha	2124	0.049
R1'	1%	1%	2%	3557	0.010	0 ha	0 ha	1900 ha	9054	0.004
R2'	0 ha	0 ha	590 ha	2259	0.015	1900 ha	0 ha	1900 ha	6518	0.008
R3'	590 ha	590 ha	0 ha	1852	0.019	1900 ha	3800 ha	1900 ha	4801	0.012
R4'	0 ha	590 ha	590 ha	1071	0.030	3800 ha	5700 ha	5700 ha	3368	0.023
R5'	590 ha	1180 ha	1770 ha	583	0.061	5700 ha	7600 ha	7600 ha	2887	0.031
R6'	1180 ha	2360 ha	1770 ha	430	0.086	15200 ha	24700 ha	19000 ha	1599	0.073
R7'	3540 ha	7080 ha	4130 ha	173	0.200	20900 ha	32300 ha	43700 ha	1224	0.101
R8'	4720 ha	8850 ha	4720 ha	131	0.243	39900 ha	66500 ha	62700 ha	843	0.155

Model outputs – example surface runoff

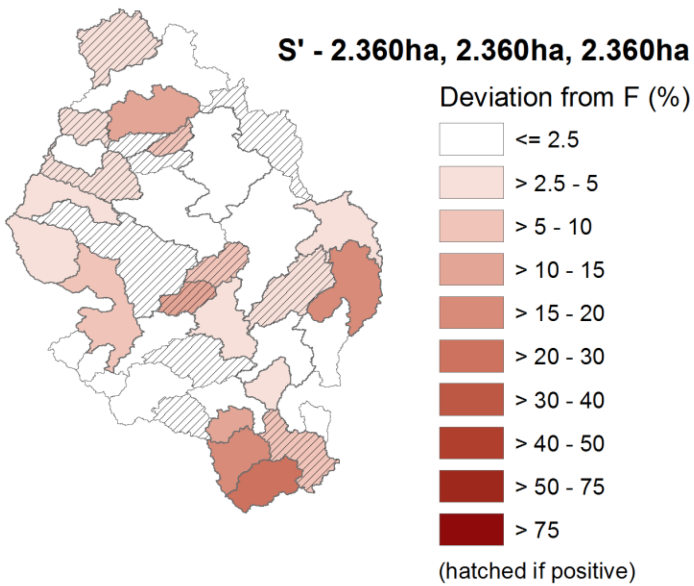




Average absolute deviation

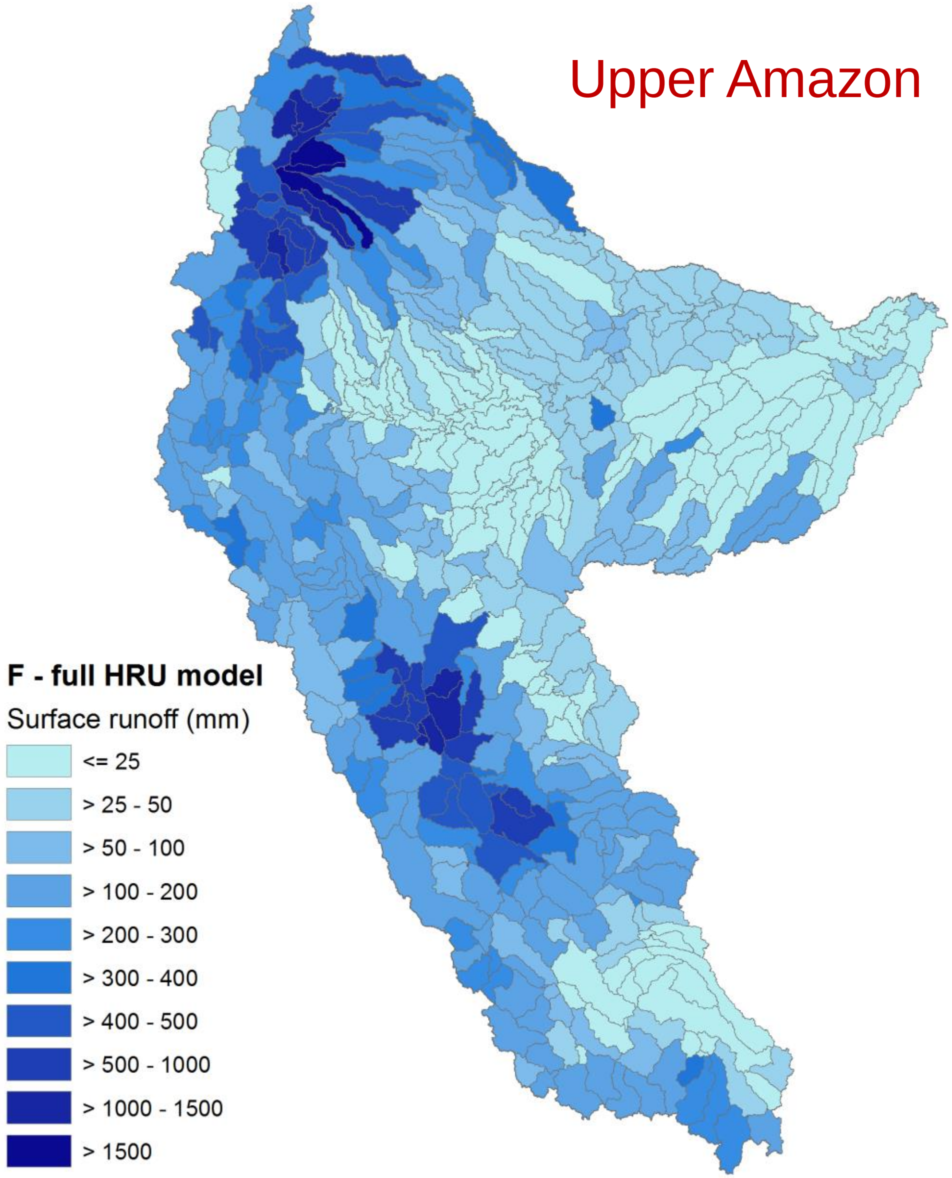
S: 43.2 %

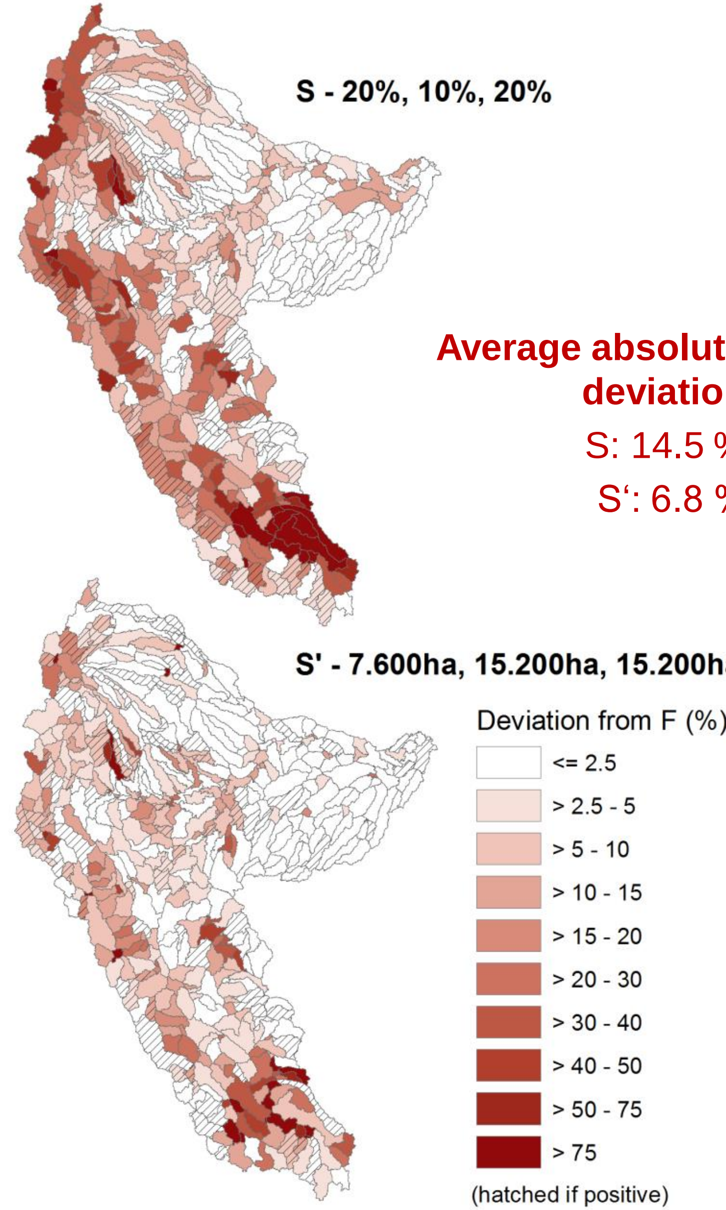
S‘: 5.4 %



Model outputs – example surface runoff

Upper Amazon





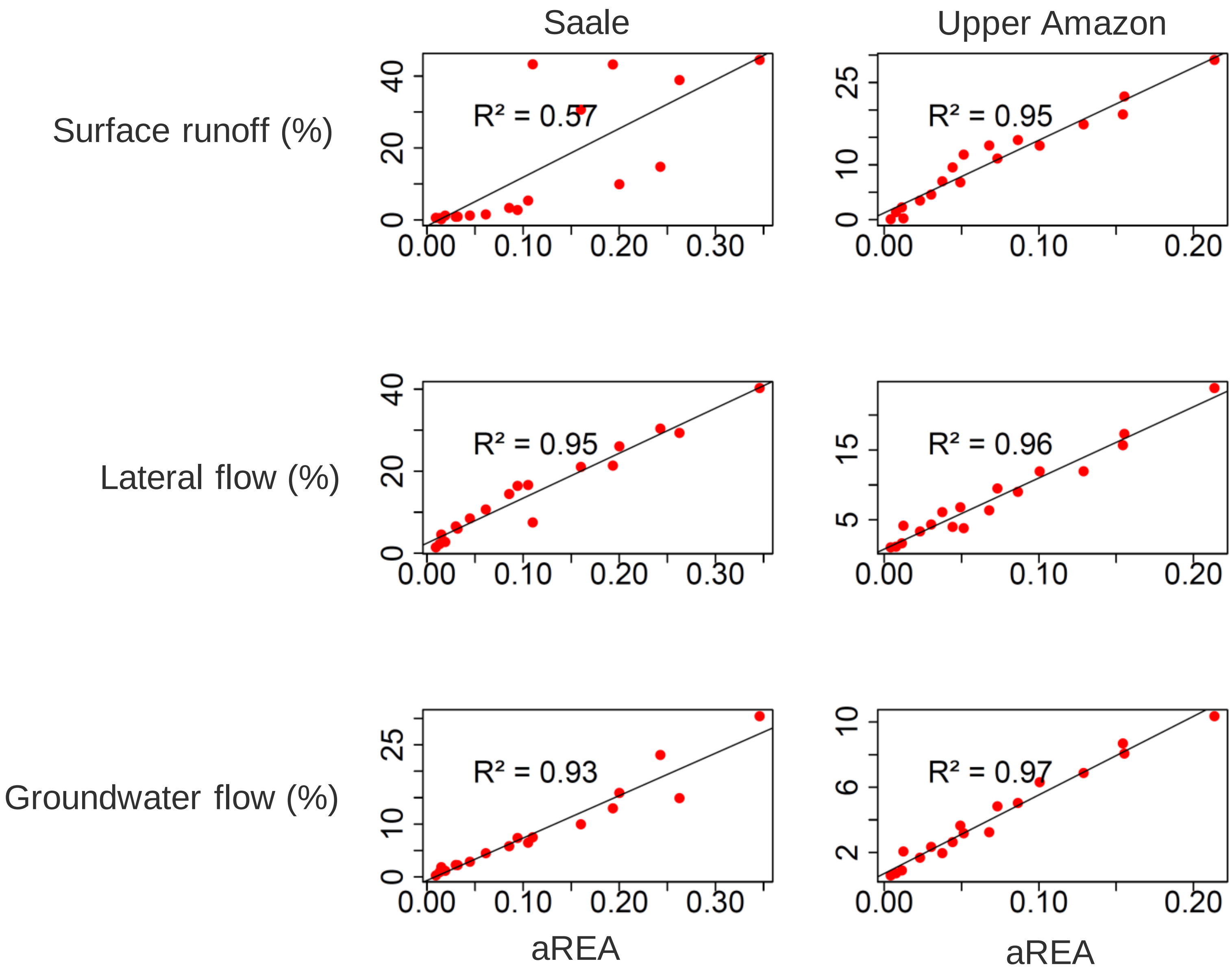
Average absolute deviation

S: 14.5 %

S‘: 6.8 %

Output ~ Input error:

Deviation from F averaged for basin ~ average Relative Error of Aggregation (aREA)



Back to the hypotheses: conclusions

1. ***HRU number reduction can significantly distort input data.***

- aREA as a direct measure for input data distortion, e.g. 19 % average error for the Saale Basin when using standard thresholds (5256 => 359 HRUs)
- Input data error increases exponentially with HRU number reduction

2. ***The larger the input data error, the larger is the model error*** *(assuming the model equations are right).*

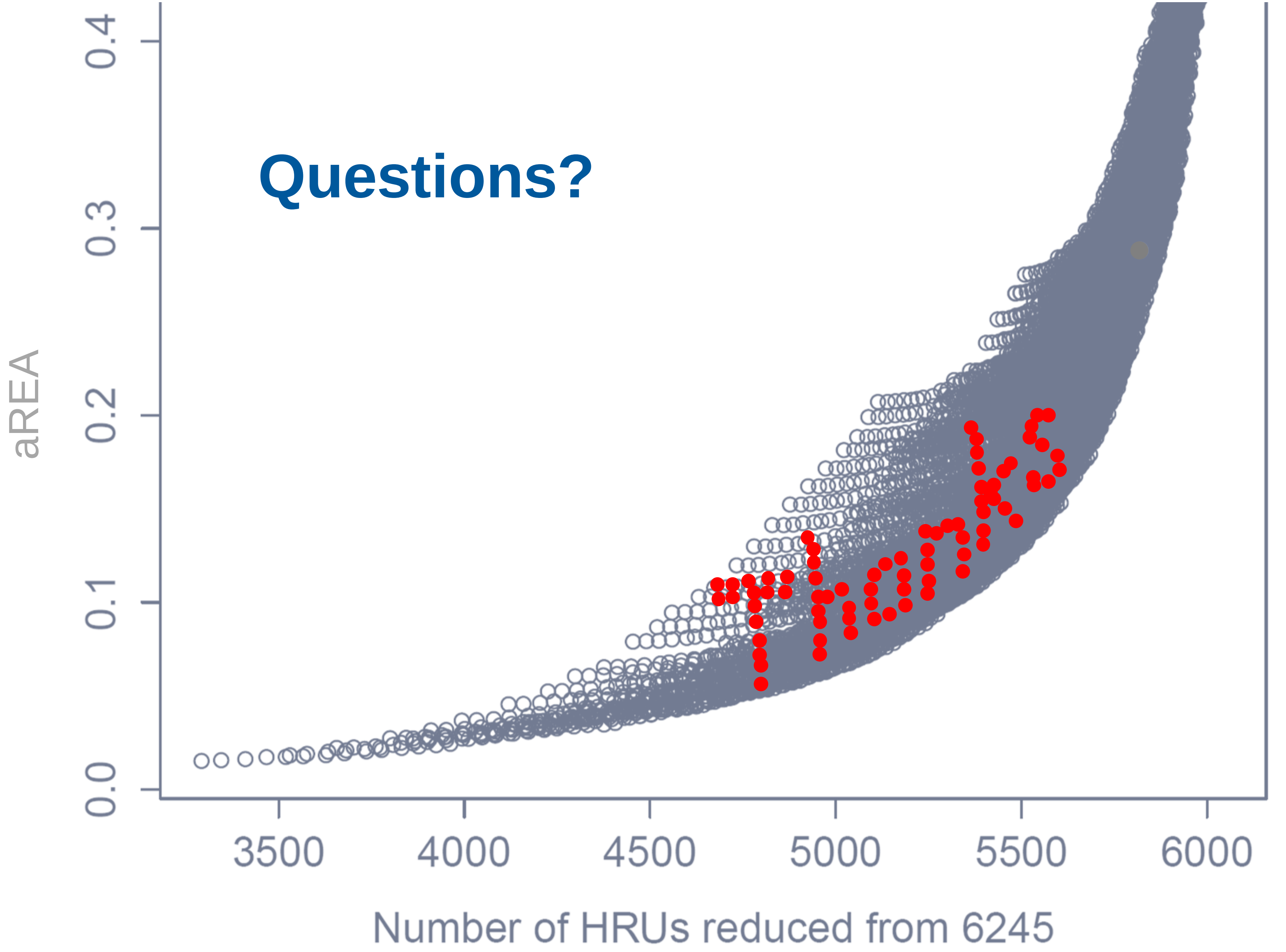
- Linear correlation between input data distortion and output error
- Remarkably high errors for single flow components

3. ***Different threshold combinations can lead to similar totals of HRUs with different input data errors*** *and thus different model errors.*

- Average input error can vary strongly (e.g. 5-20%) for similar totals of HRUs
- Choose thresholds wisely! (=> use R-script!)

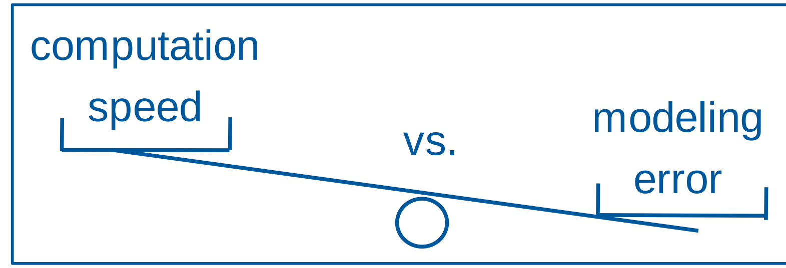
HELMHOLTZ
| ZENTRUM FÜR

Questions?


Thanks
aREA
0.0
0.1
0.2
0.3
0.4
3500
4000
4500
5000
5500
6000
Number of HRUs reduced from 6245

HRU aggregation

- a necessary evil for efficient simulations
- two ways for reducing complexity

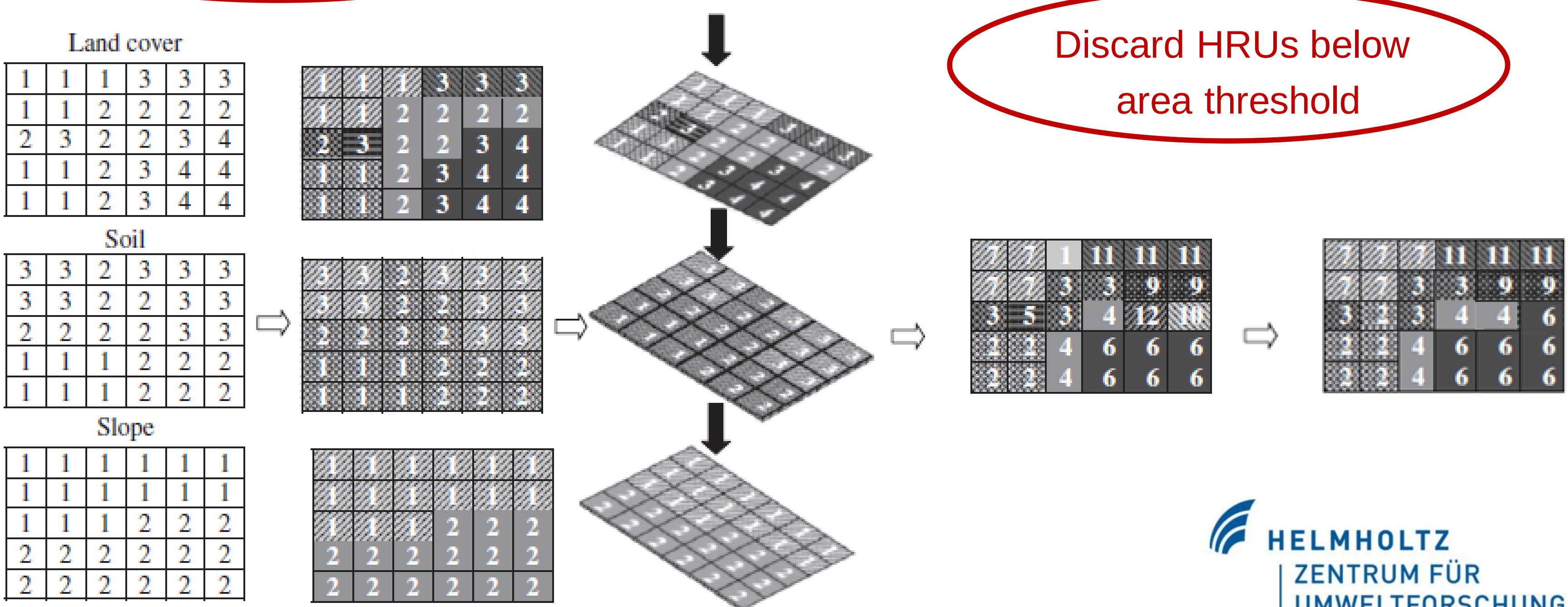


1

2

Generalize (reclassify) single inputs

Discard HRUs below area threshold



HELMHOLTZ
ZENTRUM FÜR UMWELTFORSCHUNG UFZ

How to measure input data distortion?

average Relative Error of Aggregation (*aREA*)

= average of the Relative Error of Aggregation for land use (REA_{lu}), soil (REA_{soil}), and slope (REA_{soil}).

$$REA_{lu} = \frac{0.5 \times \sum_{i,k=1}^{I,K} \left| y_{agg_{i,k}} - y_{ref_{i,k}} \right|}{A}$$

y = HRU area (agg: aggregated, ref: reference)
i = land use category
k = subbasin number

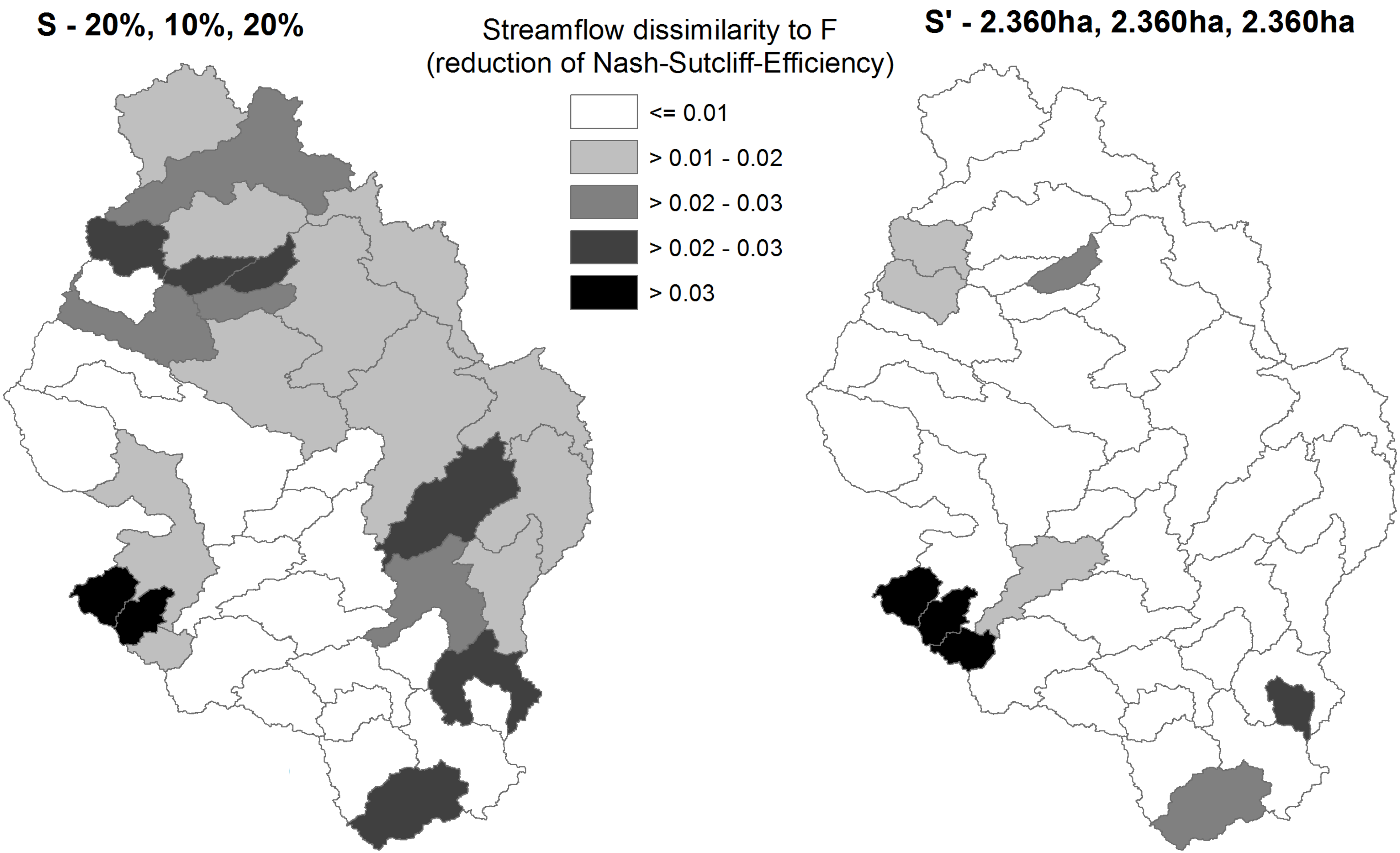
REA_{soil} and REA_{soil} are calculated analogously with i as soil and slope category, respectively.

aREA is an overall indicator for the level of aggregation and directly interpretable in any watershed.

aREA can range between 0 and 1. Multiplied with 100 it is the average percentage of the modified area for land use, soil, and slope on the total watershed area.

Model outputs – example streamflow

Saale



Model outputs – example streamflow