(Semi-) Distributed Modeling

Markus Hrachowitz
m.hrachowitz@tudelft.nl
Room 4.87

What is the issue?

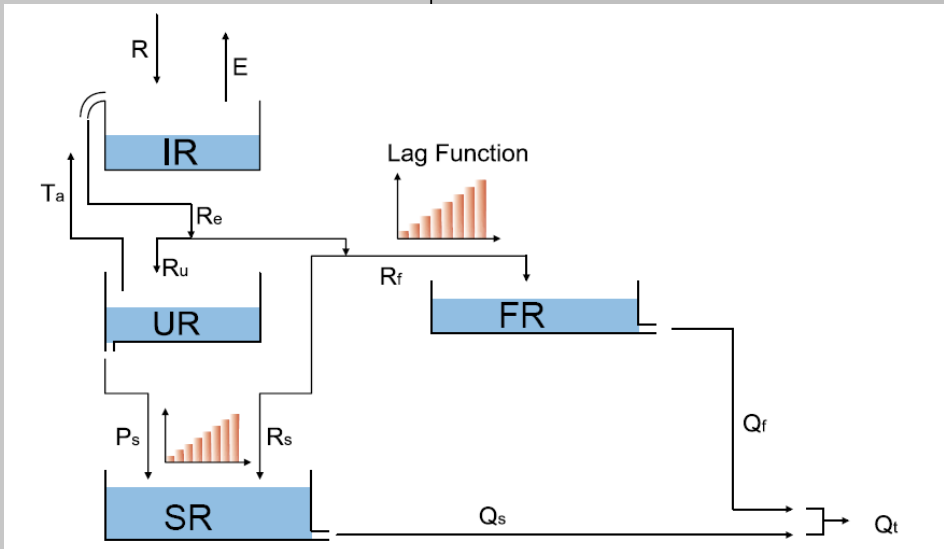
One of the fundamental problems in hydrological modeling are the **conflicting priorities** of process representation and model parsimony:

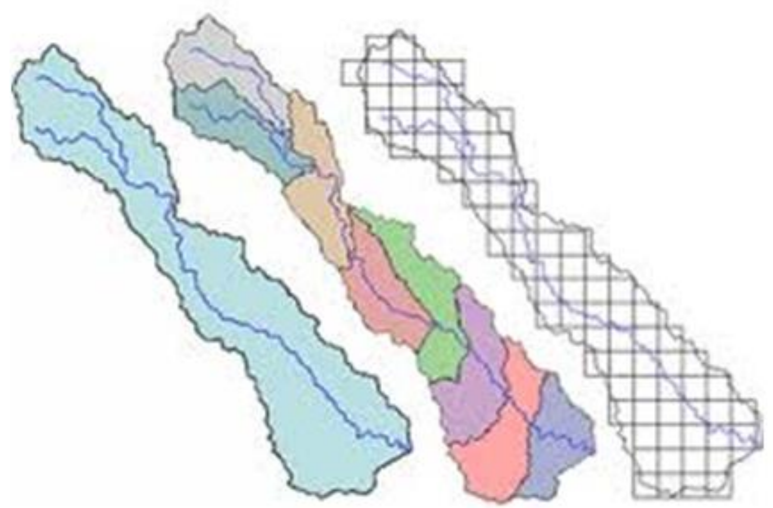
– Lumped conceptual models often too simplistic in process representation

– Distributed ... ensive

Goal: better proce... or more parameters

What does that mean for us?

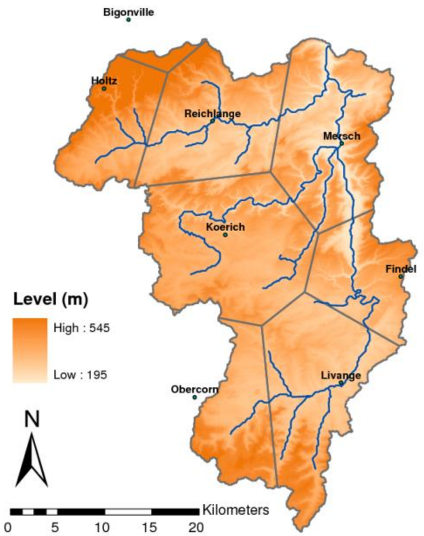


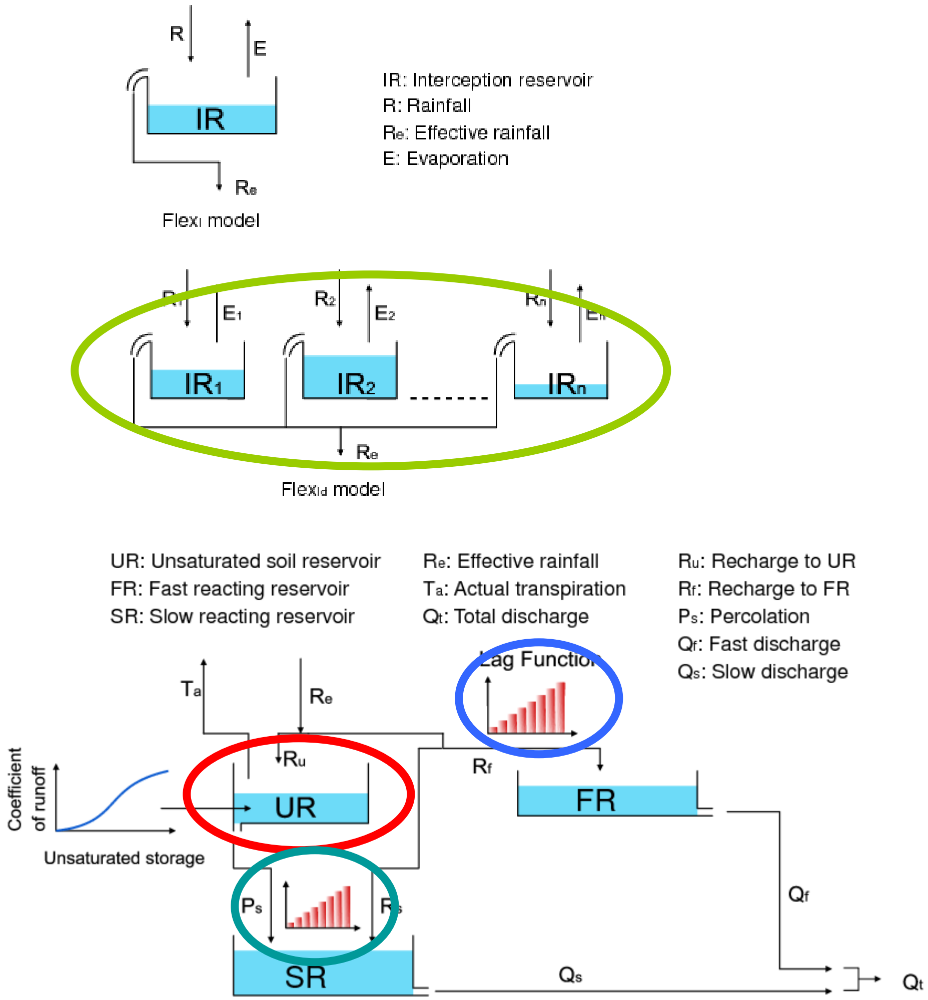


Semi-distributed rainfall accounting

Semi-distributed rainfall accounting

- Distributed **interception**
- Distributed **I**+**UR**
- Distr **I**+**UR**+**LagFR**
- Distr **I**+**UR**+**LagFR**+ **LagSR**





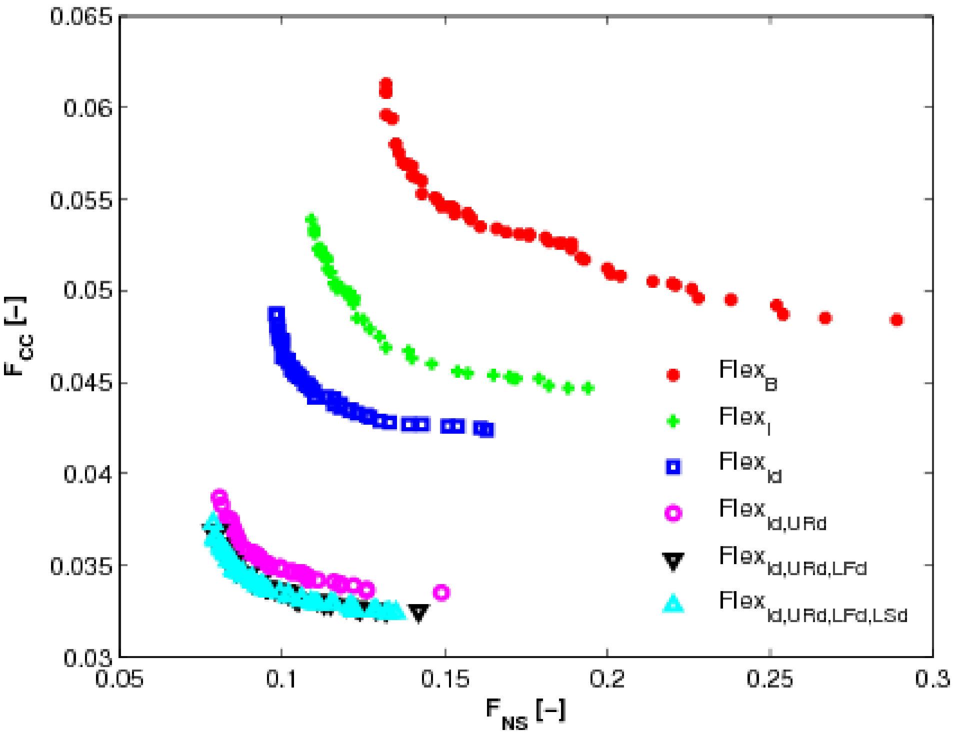
Semi-distributed rainfall accounting

Effects on model performance:

• Distributed interception and unsaturated soil provide a large improvement

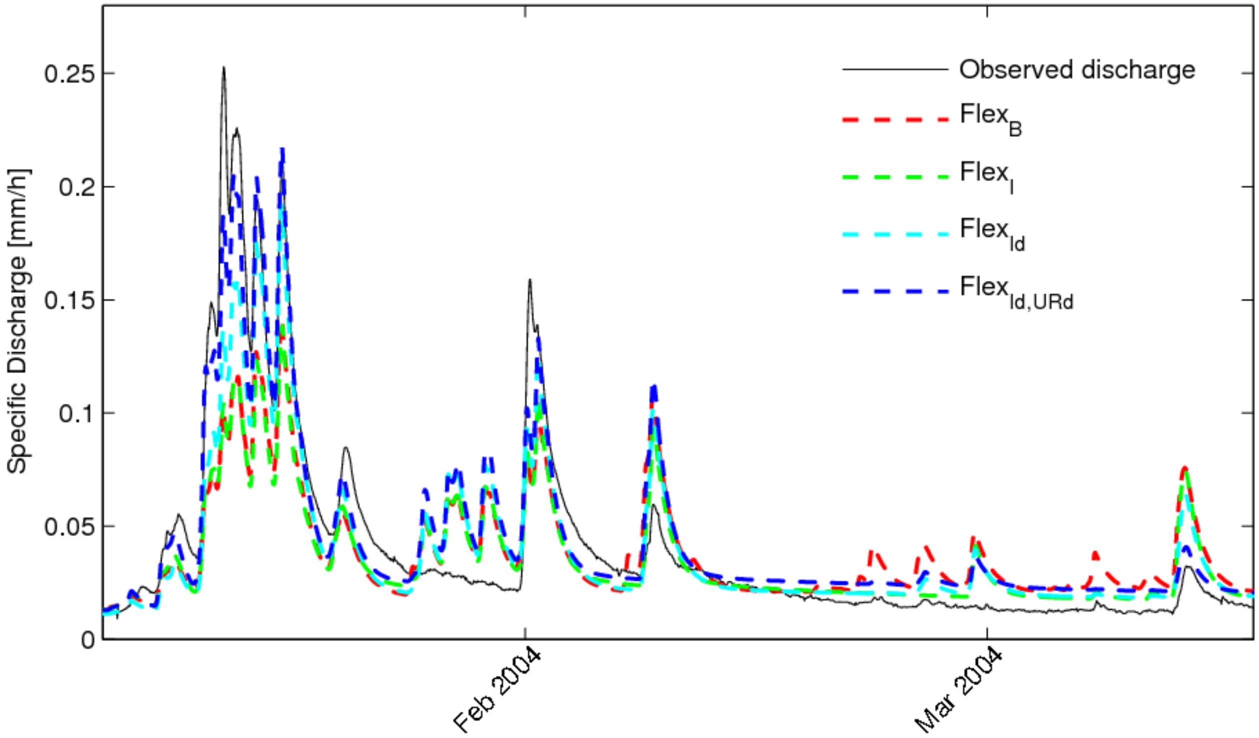
• Distributed lag times associated to fast flow provide a small improvement

• Distributed lag times associated to slow flow provide no improvement



Semi-distributed rainfall accounting

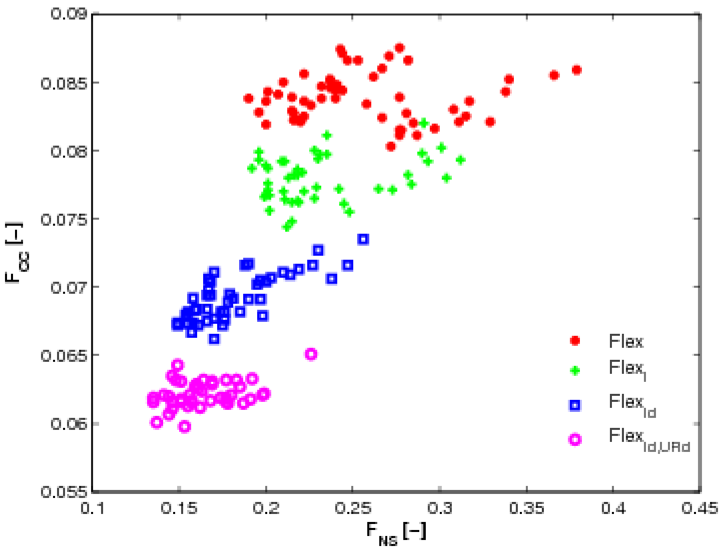
Effect on hydrograph



Semi-distributed rainfall accounting

Model validation:

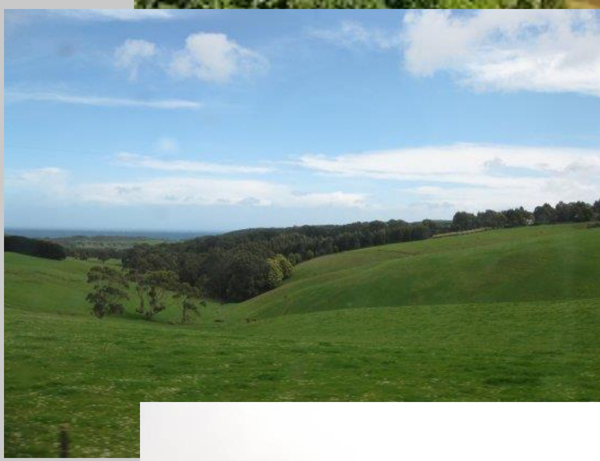
- Effect of errors
- Hierarchy of performance is maintained: no overfitting

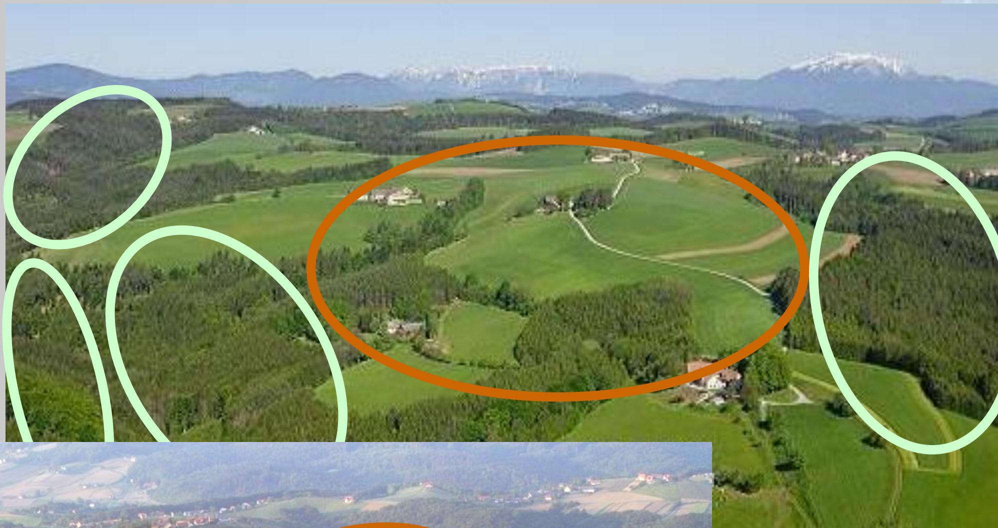


Topography driven conceptual modeling (FLEX-Topo)


Let us have a look at some landscapes…
…can we see some recurring patterns?

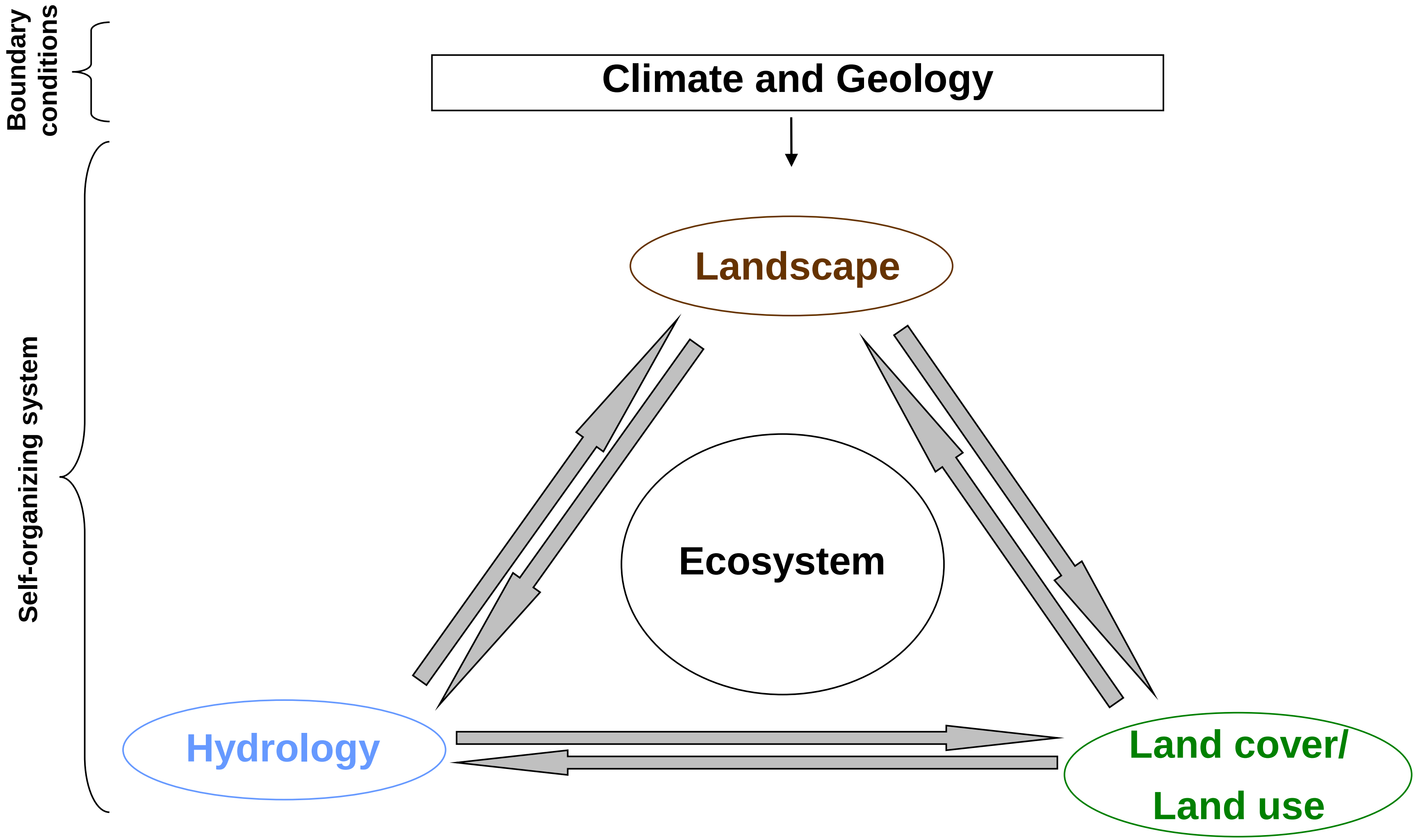




Patterns with similar emergent behaviour:

- agriculturally cultivated **Plateaus**
- forested **Hillslopes**
- **Valley bottoms** used for crop cultivation or grazing


Boundary conditions
Climate and Geology
Self-organizing system
Landscape
Ecosystem
Hydrology
Land cover/
Land use

- Possible way to go is to incorporate available data with knowledge about ecosystem functioning
- A priori definition of landscape units with different function

Forested hillslopes

Two crucial observations:

- **Floods** predominantly generated on hillslopes
- Hillslopes frequently covered by **forests**

Forest can only survive on hillslope because of

(a) moisture **retention** to bridge dry spells

(b) efficient sub-surface **drainage** of excess water to ensure aerated soil and avoid erosion

→ **Preferential flow pathways** most dominant and efficient mechanism – “Storage excess subsurface flow”

→ **Feedback process:** floods are generated on hillslopes because they are forested and forests are located on hillslopes because there are efficient drainage mechanisms

Plateaus

Characterized by low gradients and deep groundwater table

Primary functions:

(a) moisture retention

(b) evaporation

Primary processes predominantly **vertical**

Little contribution to storm runoff generation

Riparian zone / wetlands

Characterized by modest slopes and located close to open water

→ Groundwater table close to ground surface

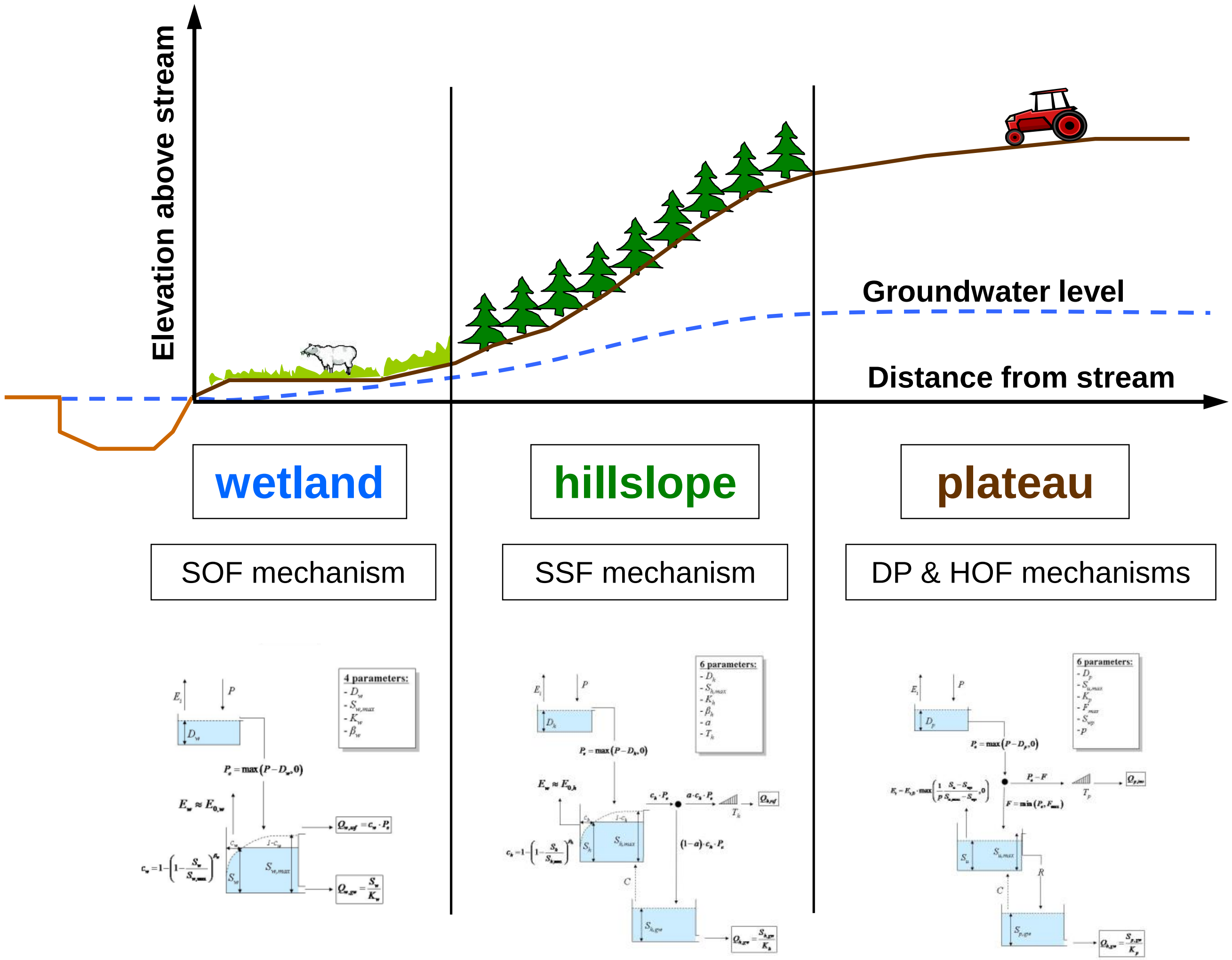
→ Moisture storage potential low

Primary function:

Contribution to flood generation, but due to small areal extent of riparian areas less important than hillslopes

Primary process **saturation overland flow**

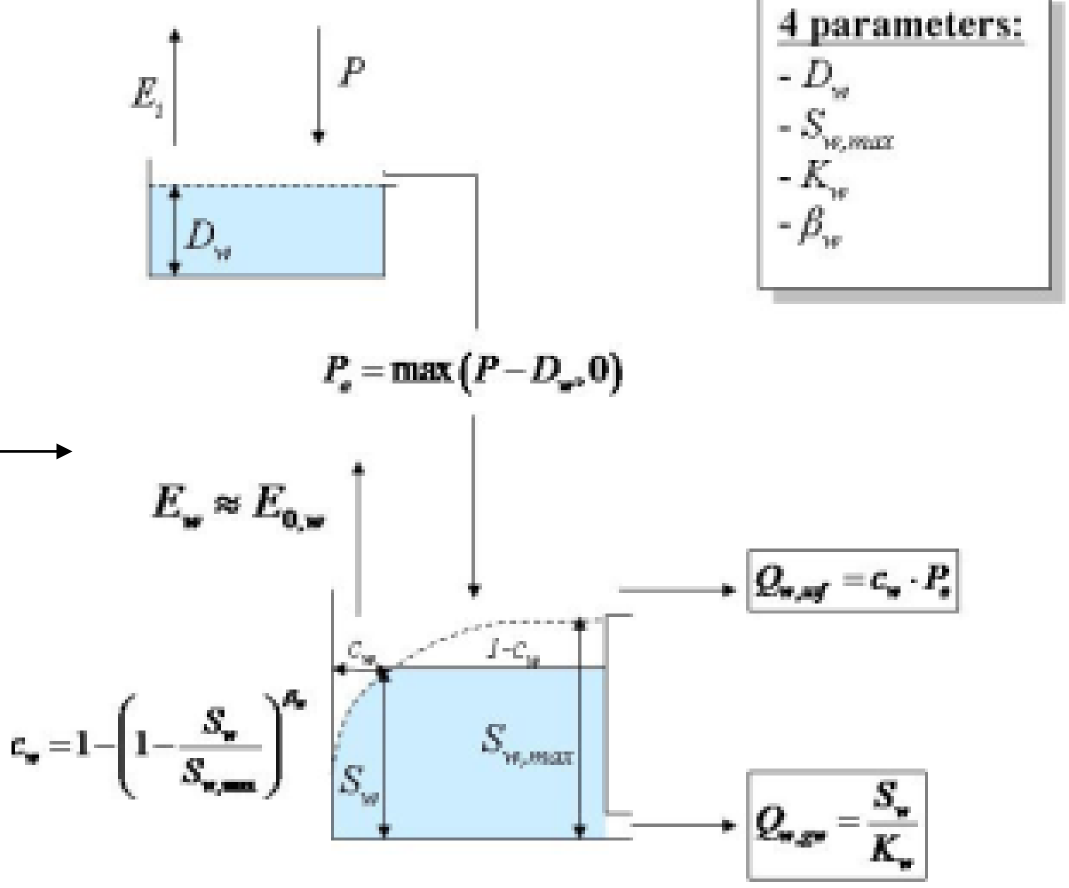
Different landscape units → different hydrological function → different model structure



Riparian zone / wetland

Dominant process:

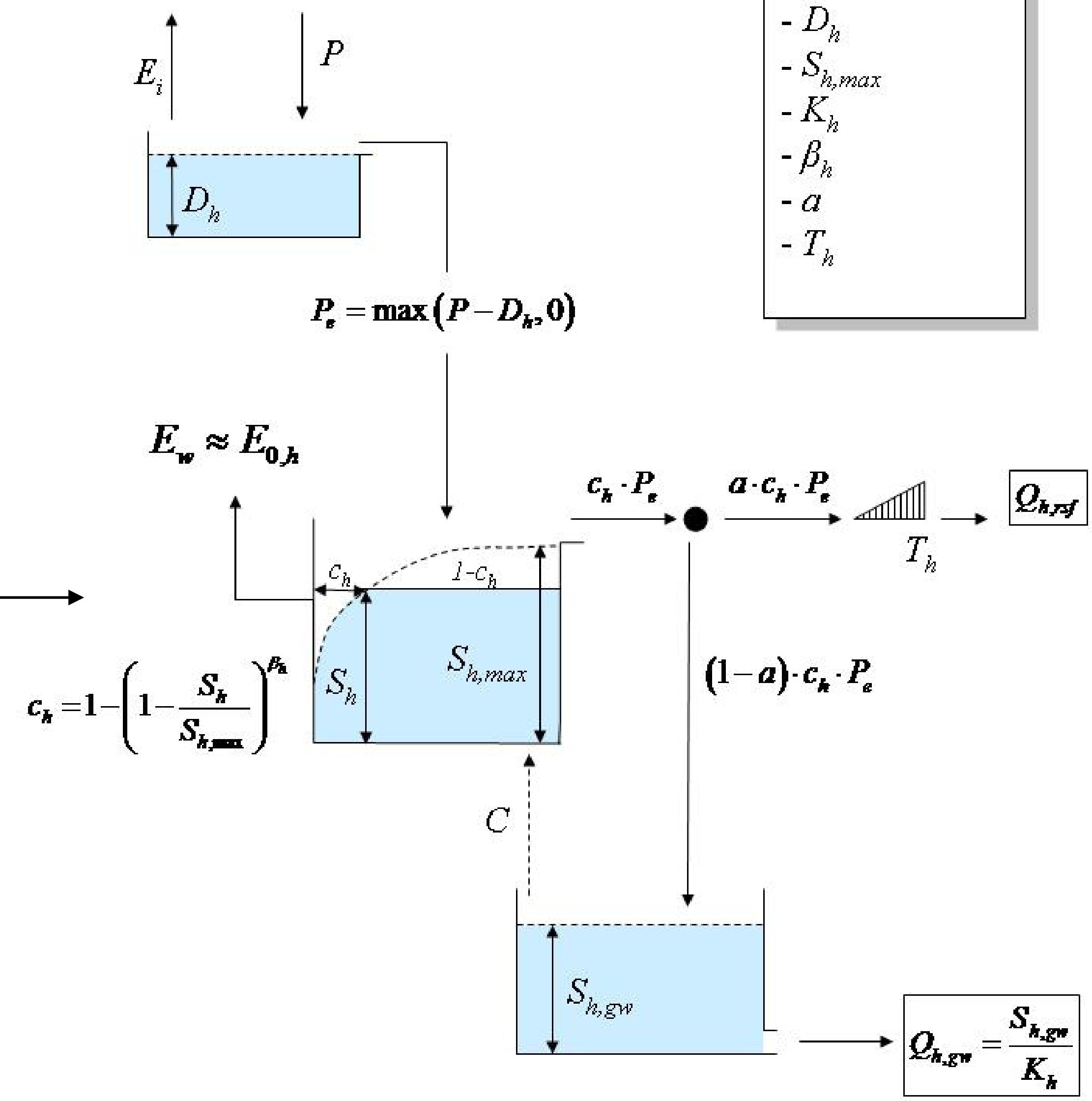
Saturation overland flow



Hillslope

Dominant process:

Shallow subsurface flow



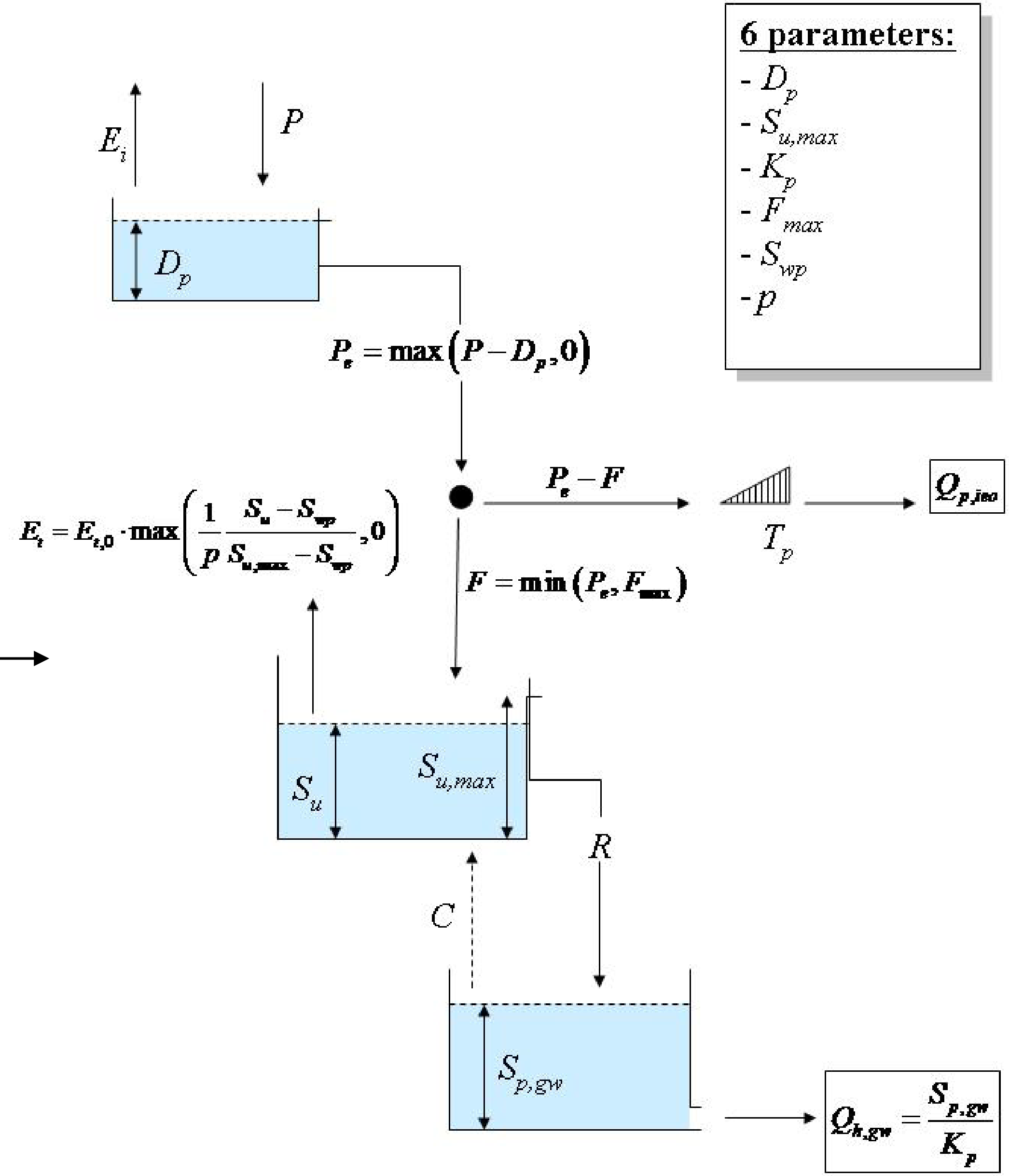
6 parameters:
- D_h
- $S_{h,max}$
- K_h
- β_h
- a
- T_h

Plateau

<u>Dominant processes:</u>

Deep percolation

Evaporation

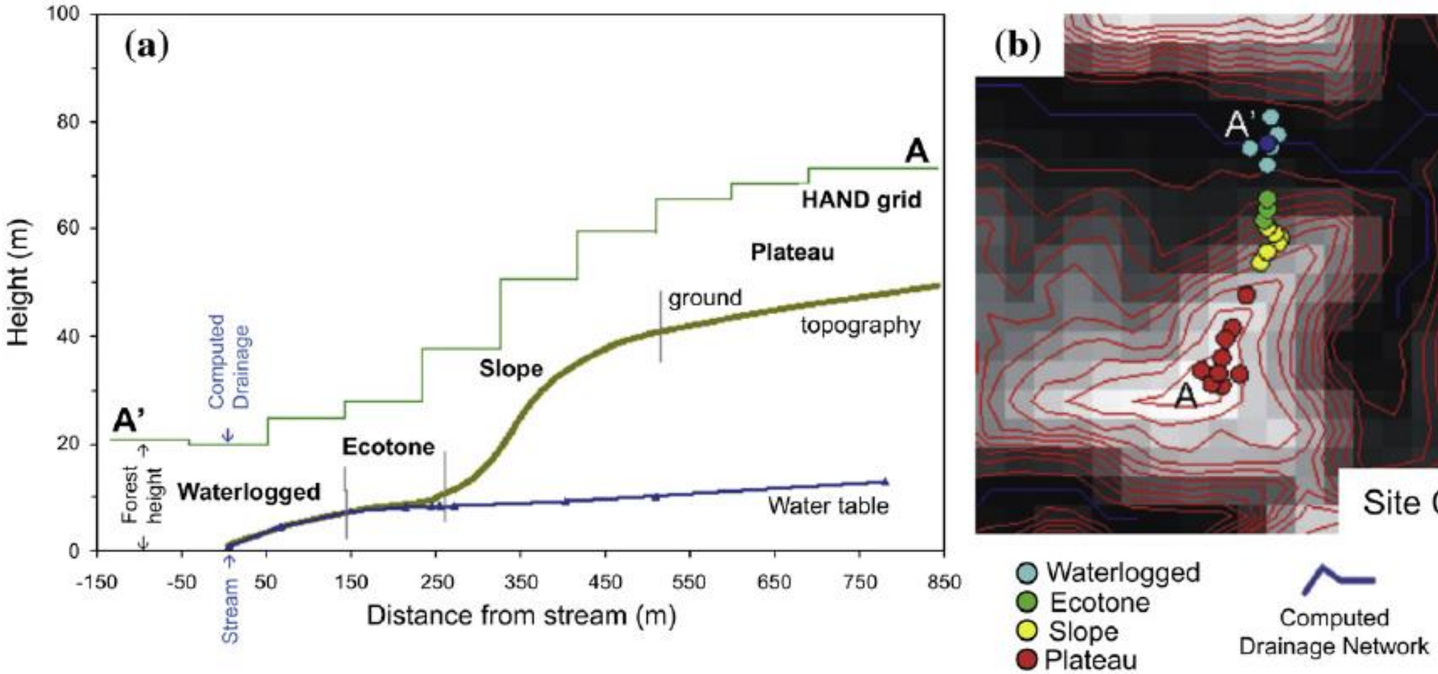


Can we find an **<u>objective measure</u>** to define landscapes which represent different functional units?

One possibility: “Height-above-nearest-drainage” **<u>(HAND)</u>**:

(1) Landscape classified by distance from and elevation above stream

(2) Inflection points indicate changes in landscape features and functional units



Renno et al., 2008, RSE

HAND (cont.):

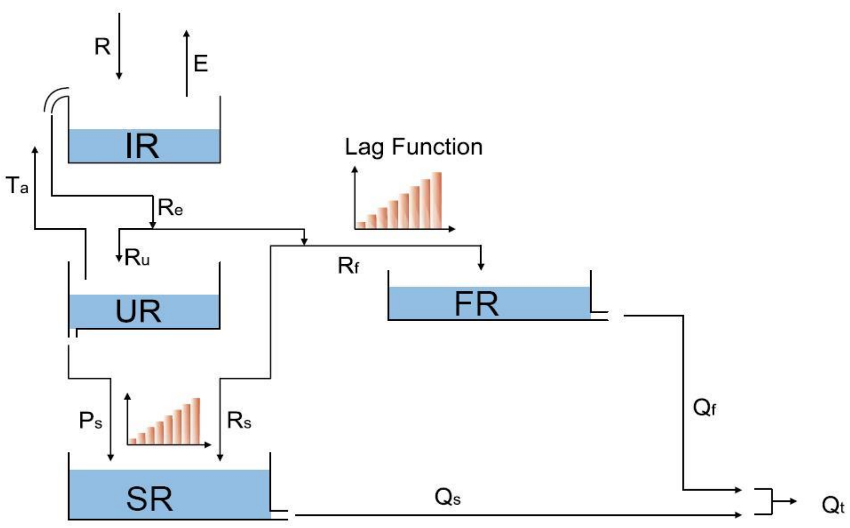
(3) Suitable model structure assigned to each functional unit

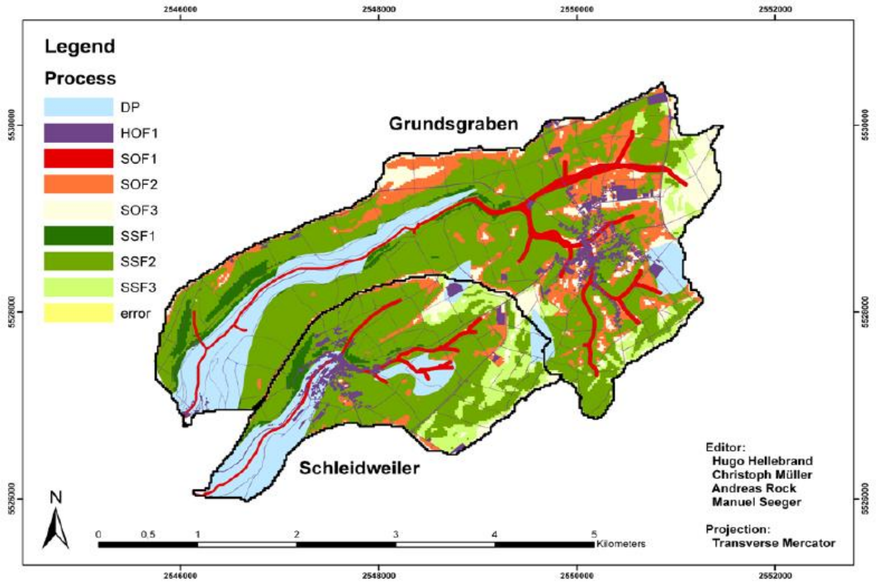
(4) Modeled runoff is a composite of proportional contributions from each functional unit, i.e. pre-assigned model structure

Advantage:

Improved proce model, only on basis of readily availa crease in parameter number

→ Potential for b ges and ensuring better paramete





Has this not been done before?

To a certain extent yes, **BUT:**

(a) Incorporated in process based models, resulting in considerable parameter equifinality (e.g. Uhlenbrook et al., 2004) or

(b) Classification of functional units based on information that is too detailed and therefore mostly not available (e.g. Scherrer et al., 2007; Hellebrand and van den Bos, 2008)

Flex-TOPO

Land classification

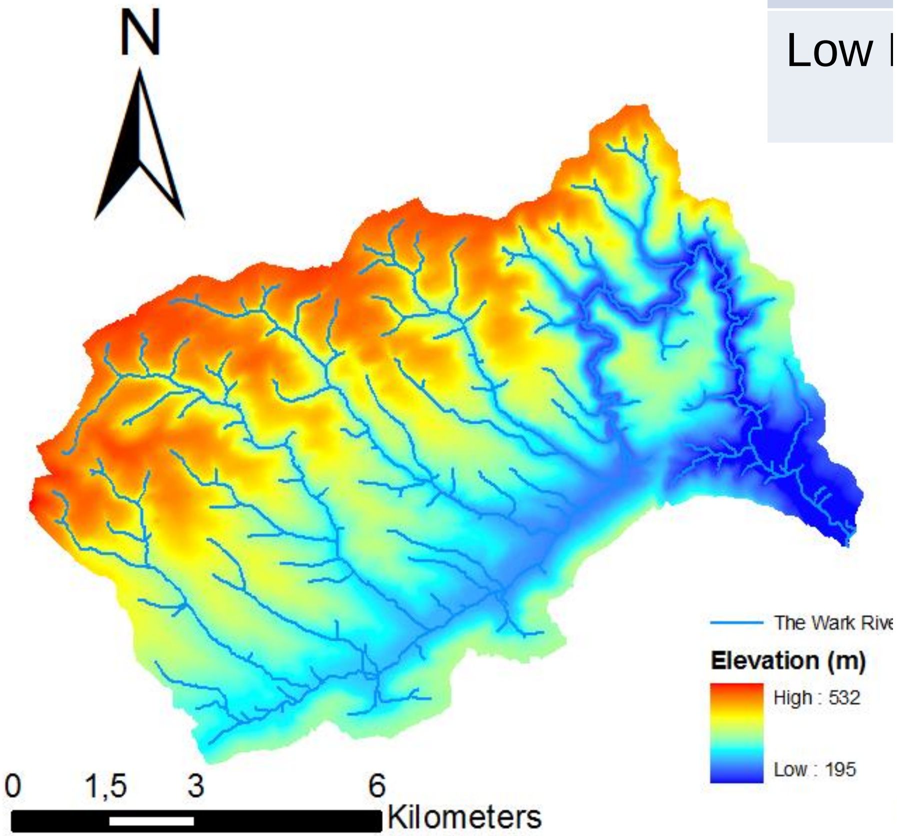
Assigning different hydrological response unit

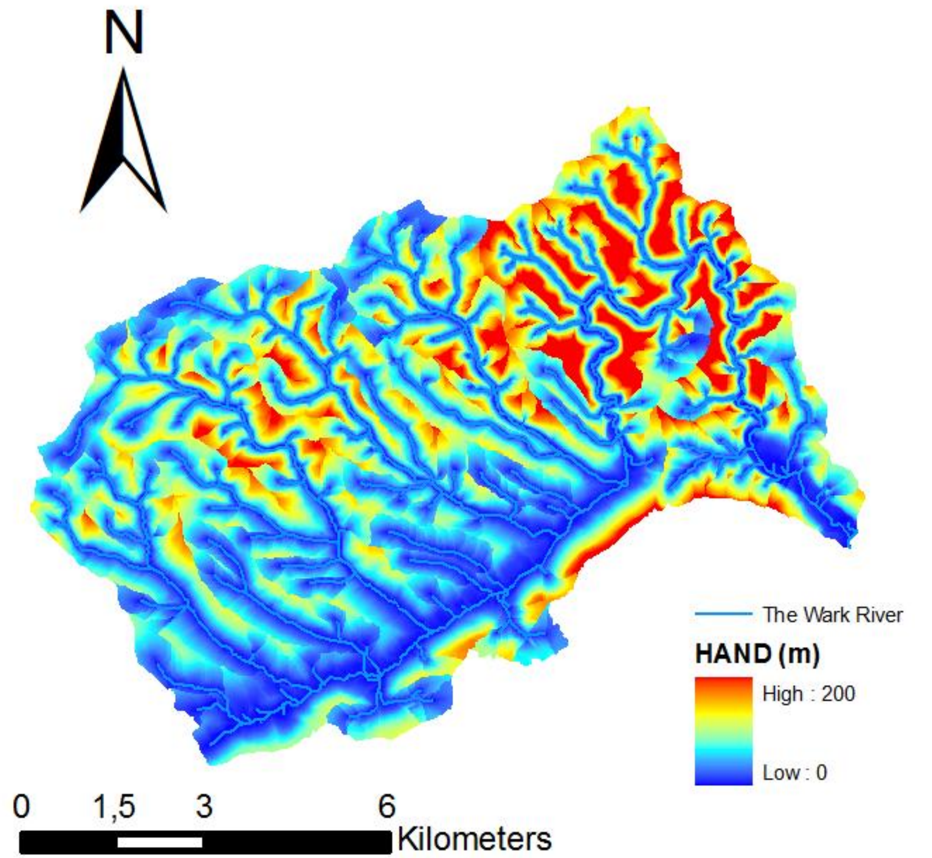
Conceptualization of each response unit

Flex-TOPO

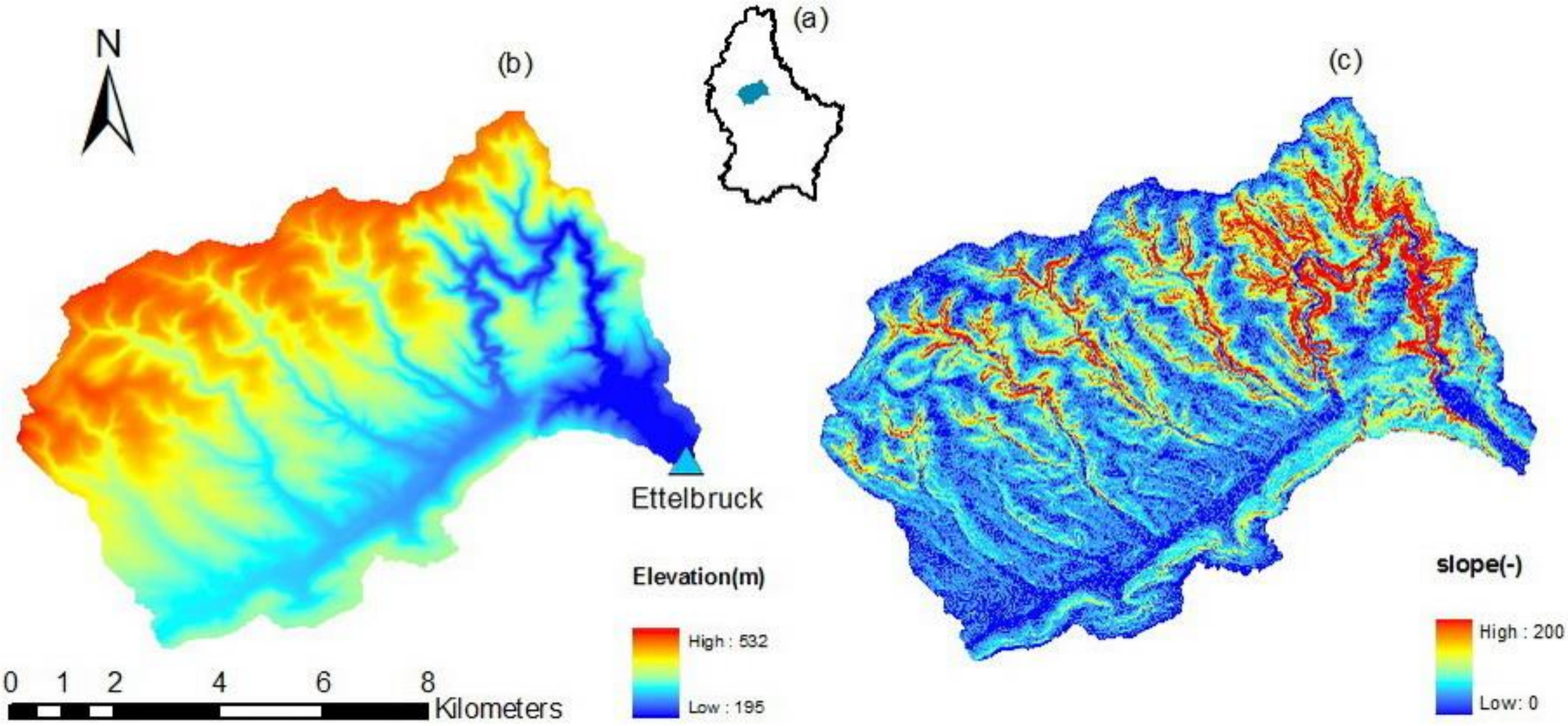
1. Landscape classification v

	High Slope	Low Slope
High HAND	Hillslope (B)	Plateau (A)
Low		





Flex-TOPO: Case study



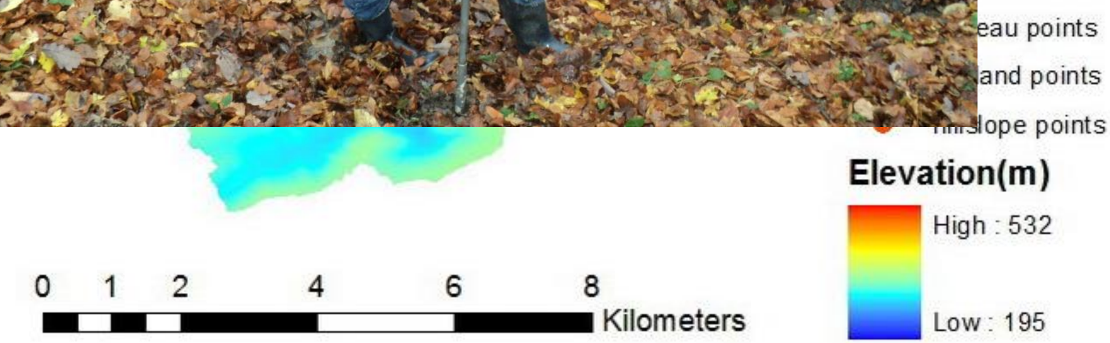
Flex-TOPO: Case study

Landscape classification : calibration

A total number of 5611 points

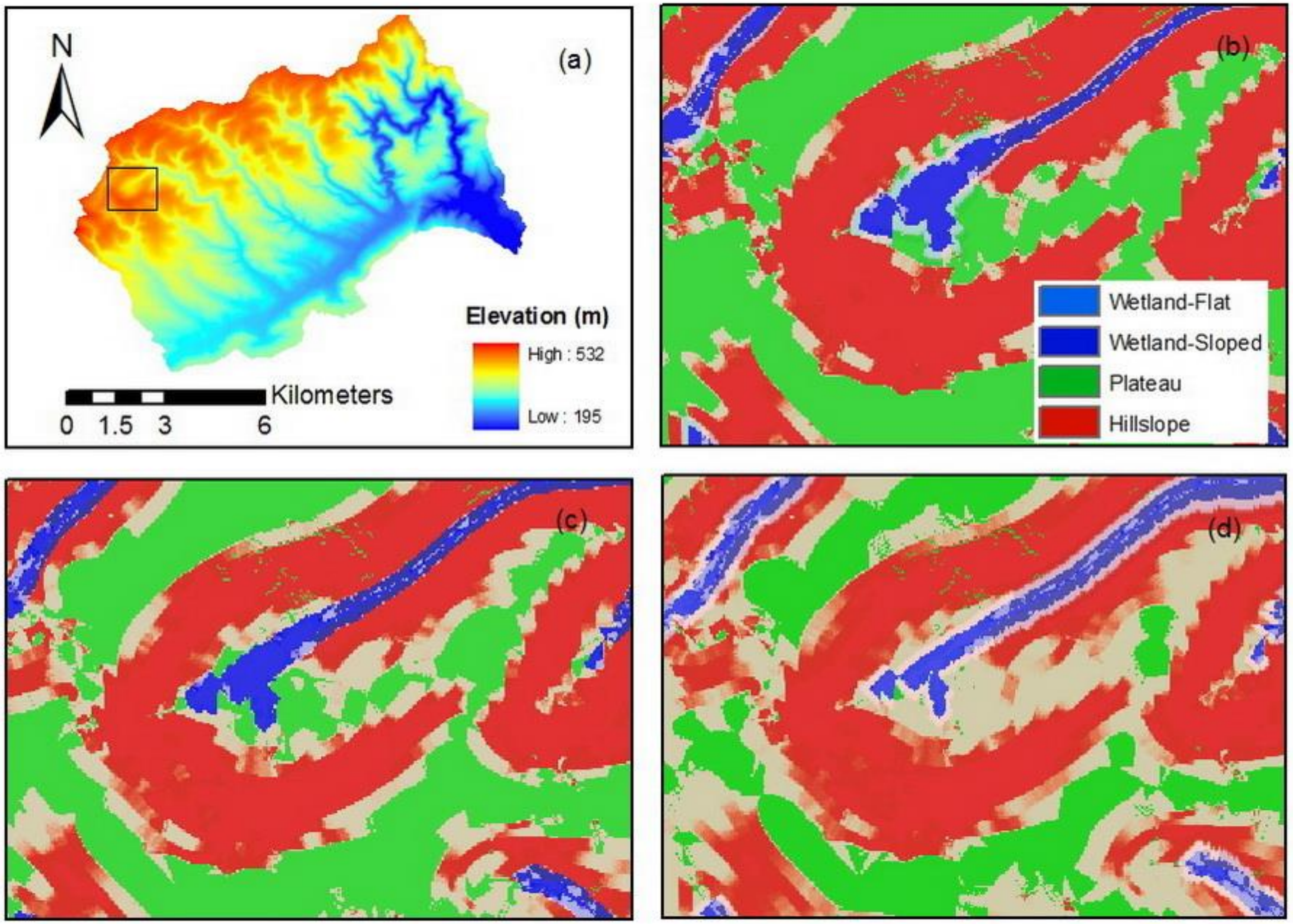
- 1501 (26.8%) Wetland
- 1385 (24.6%) Hillslope
- 2725 (48.6%) plateau

$$\mathrm{ObjFun} = \left[1 - \frac{\sum_{i=1}^{N_H} p_H}{N_H} \right.$$

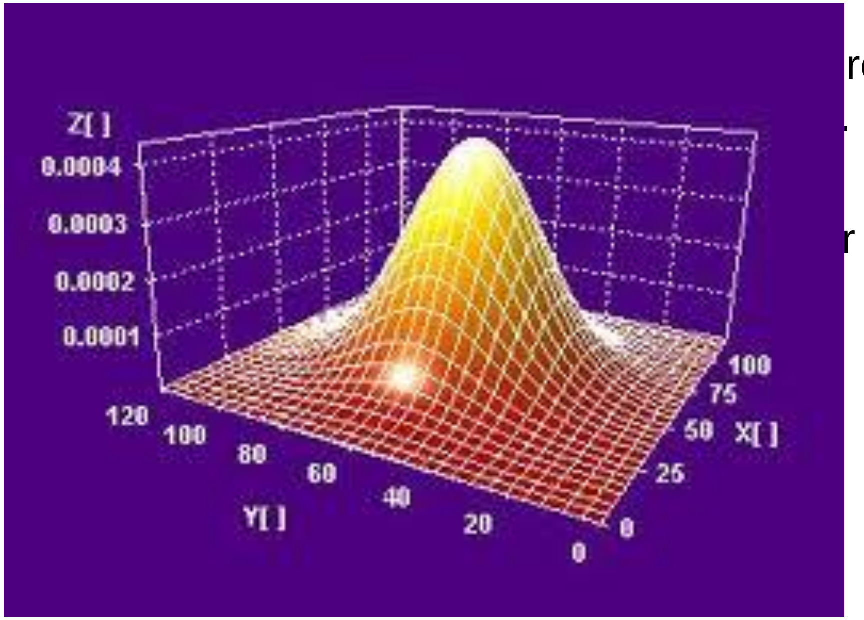


Flex-TOPO: Case study

HAND landscape classification

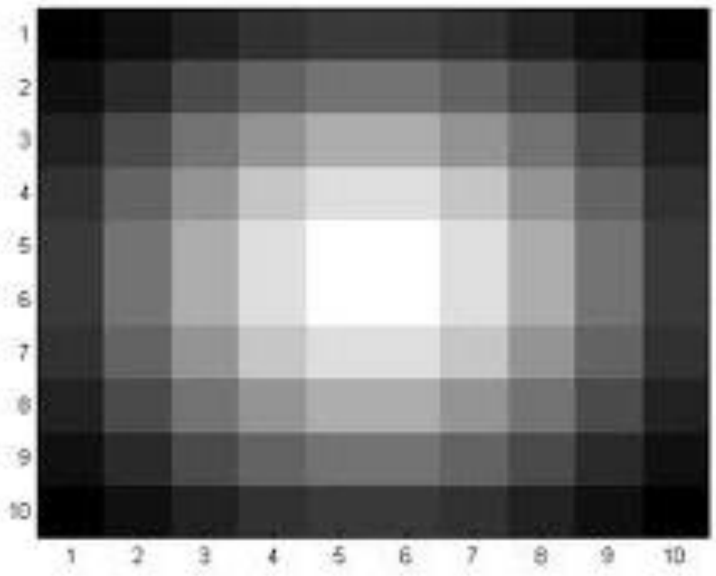


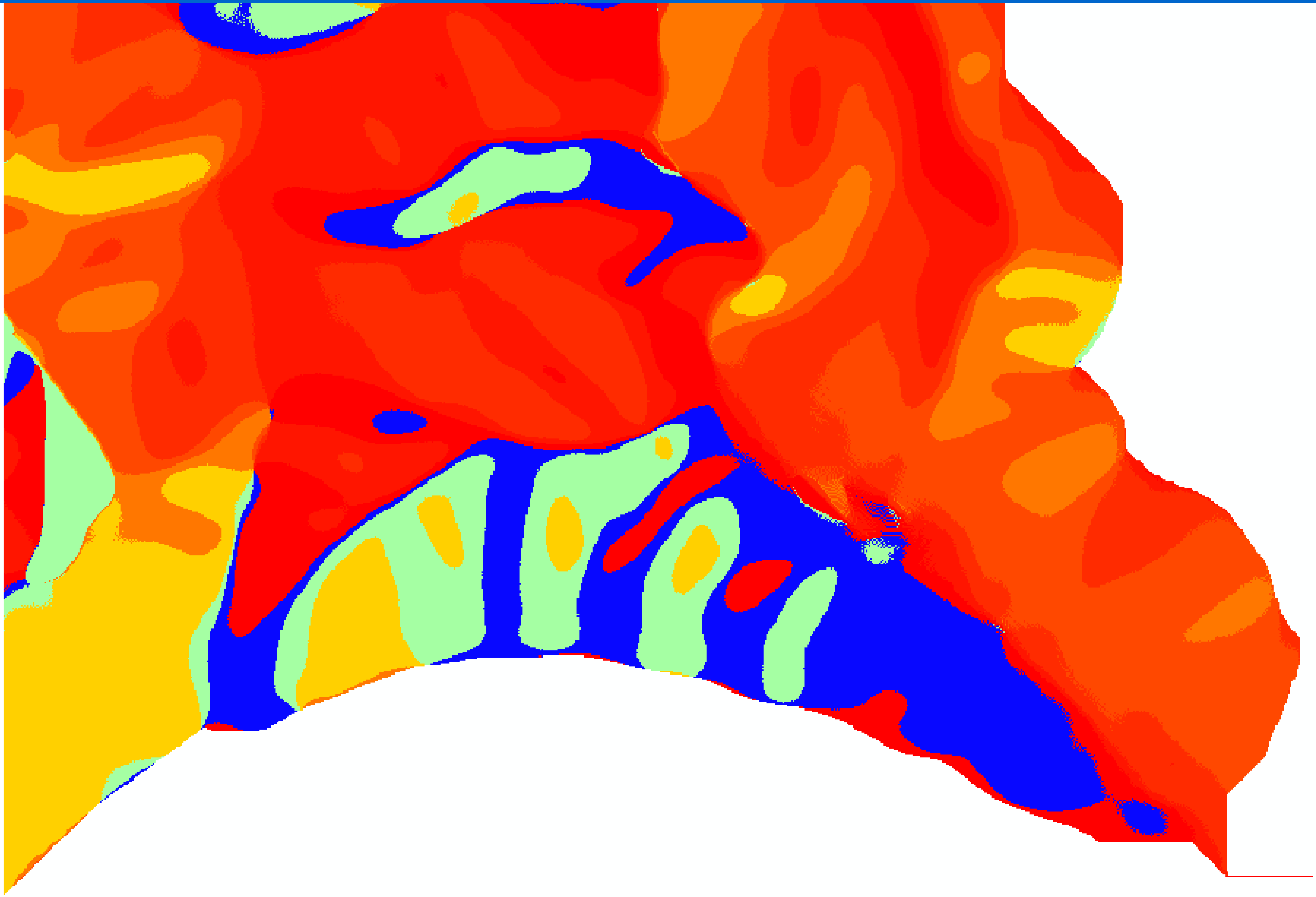
Flex-TOPO: Case study



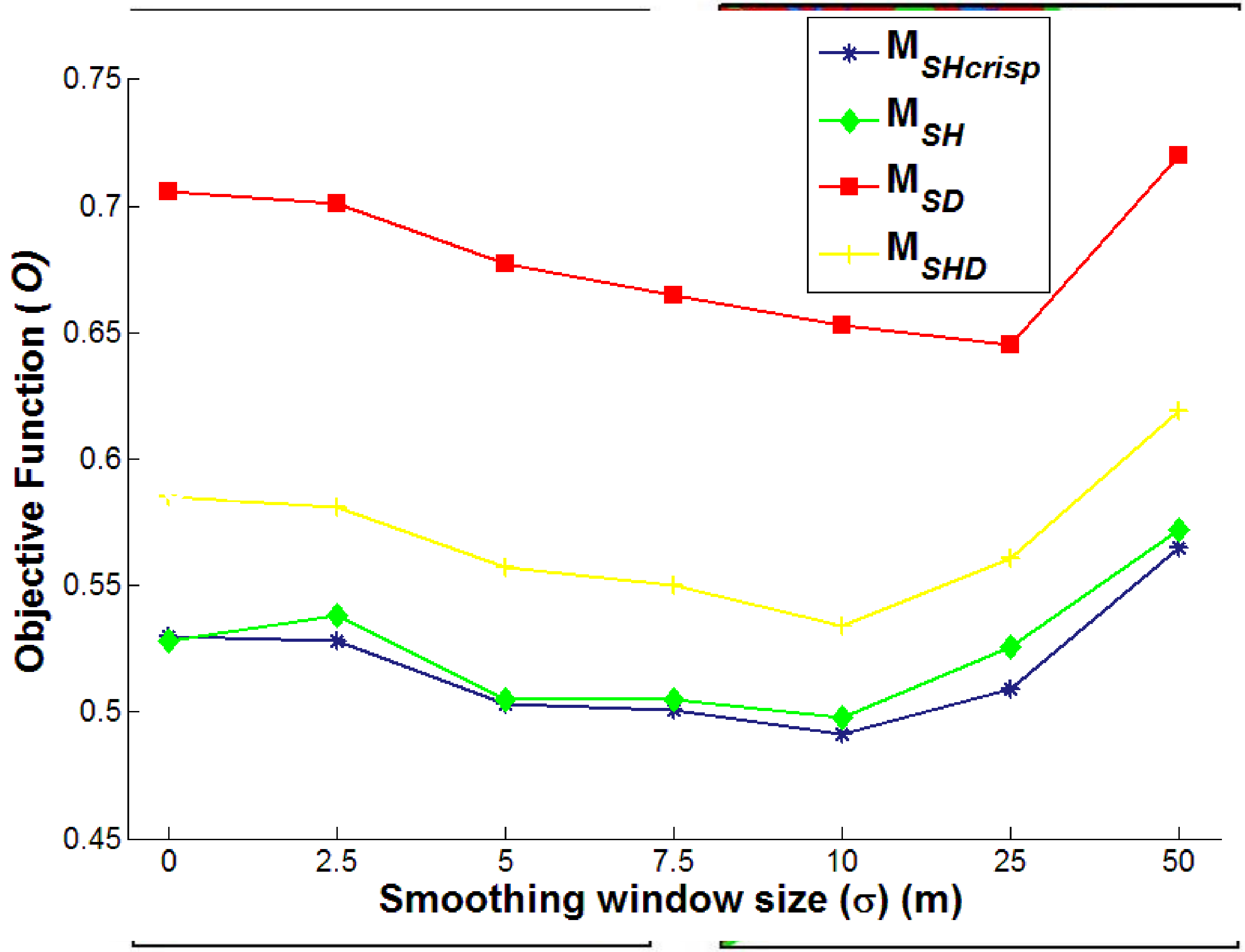
rock outcrop, structures or roads which
.

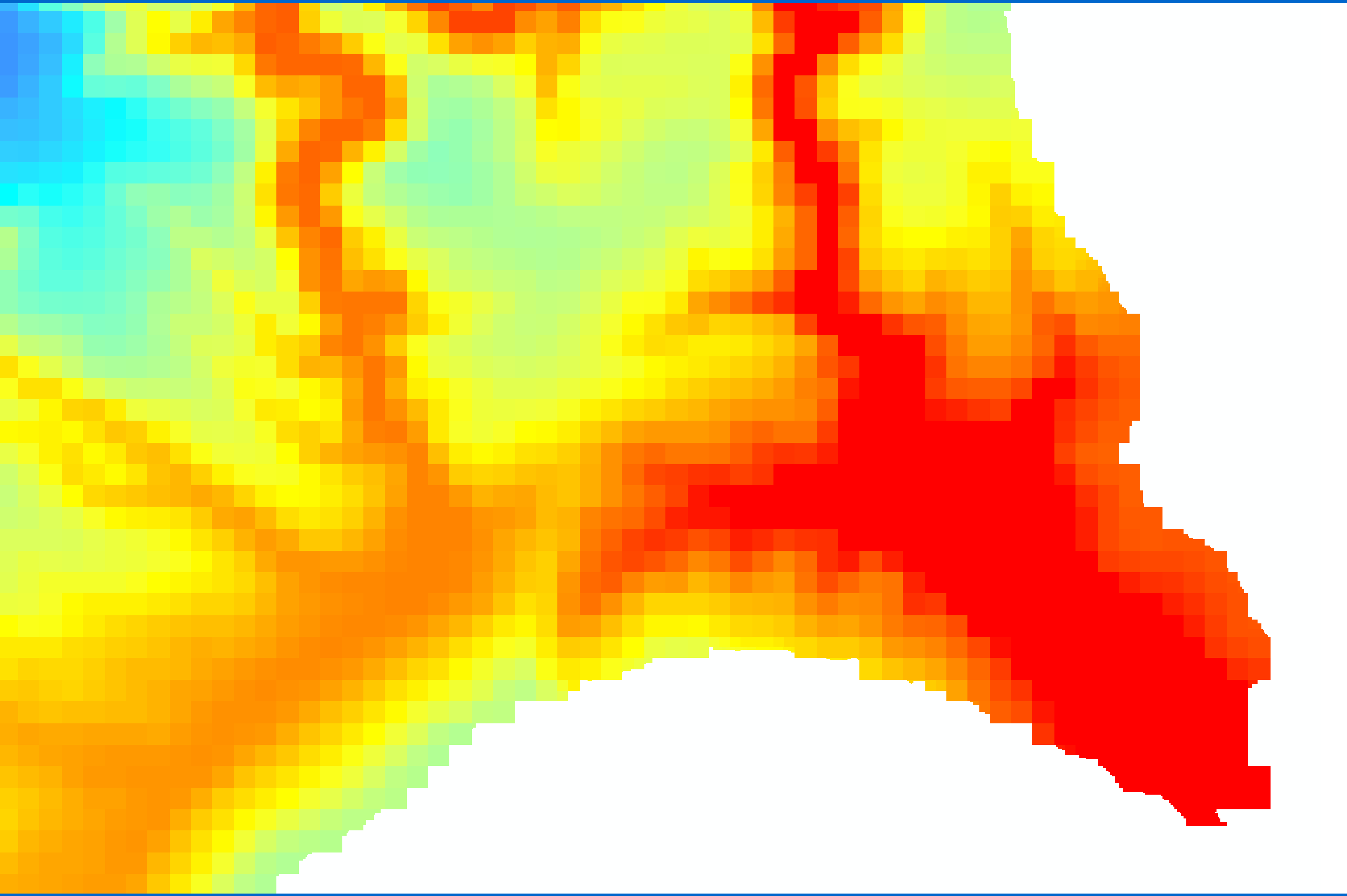
r with different sizes was used.



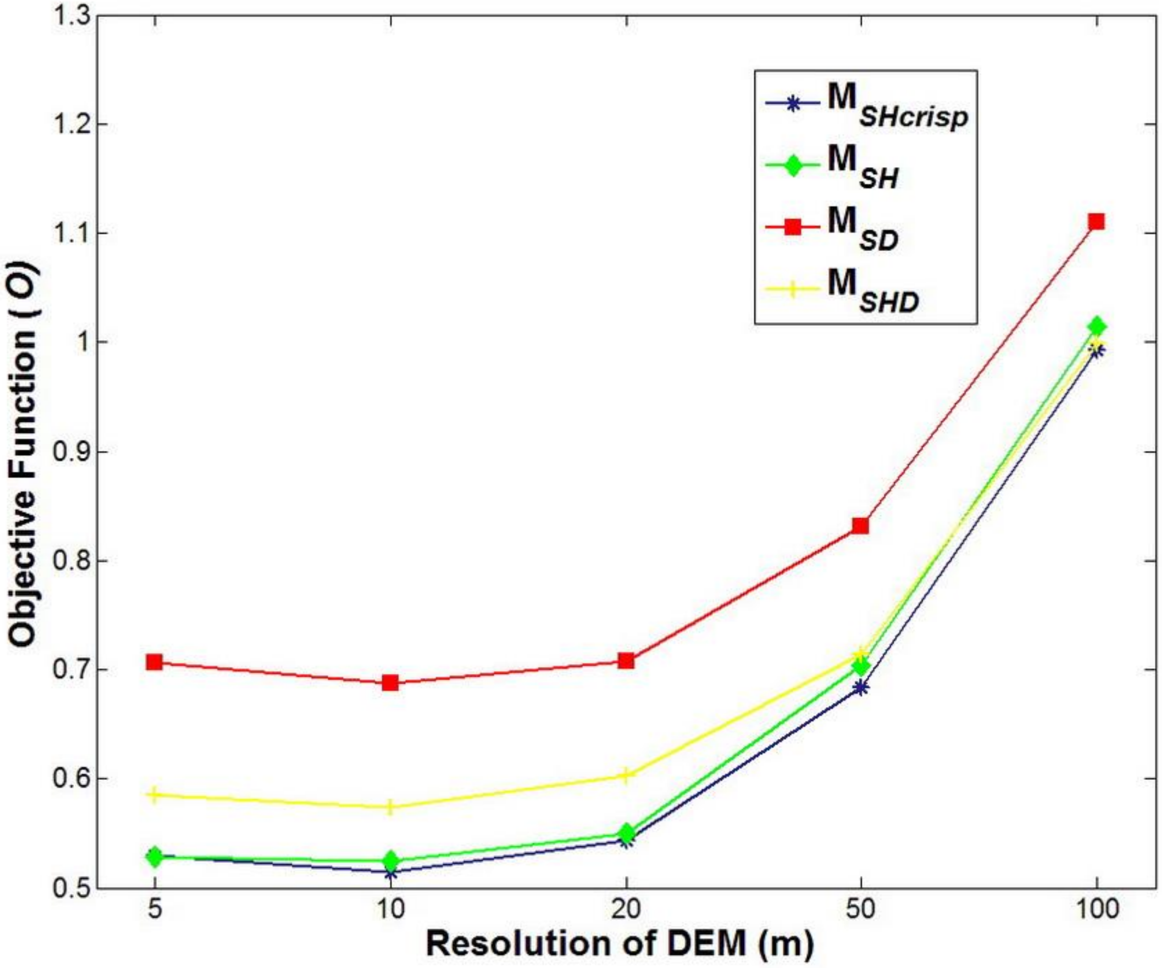


Flex-TOPO: Case study

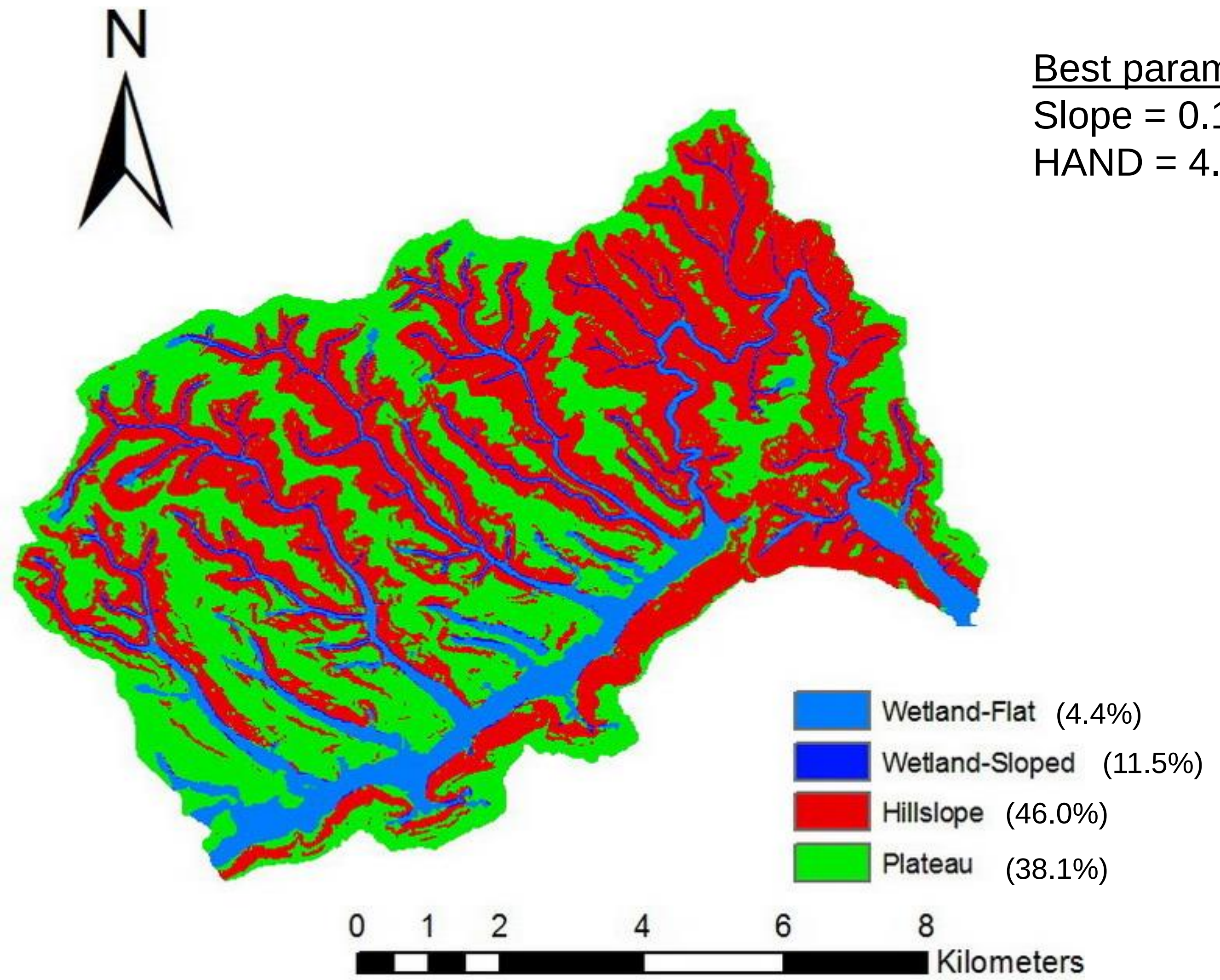




Flex-TOPO: Case study



Flex-TOPO: Case study

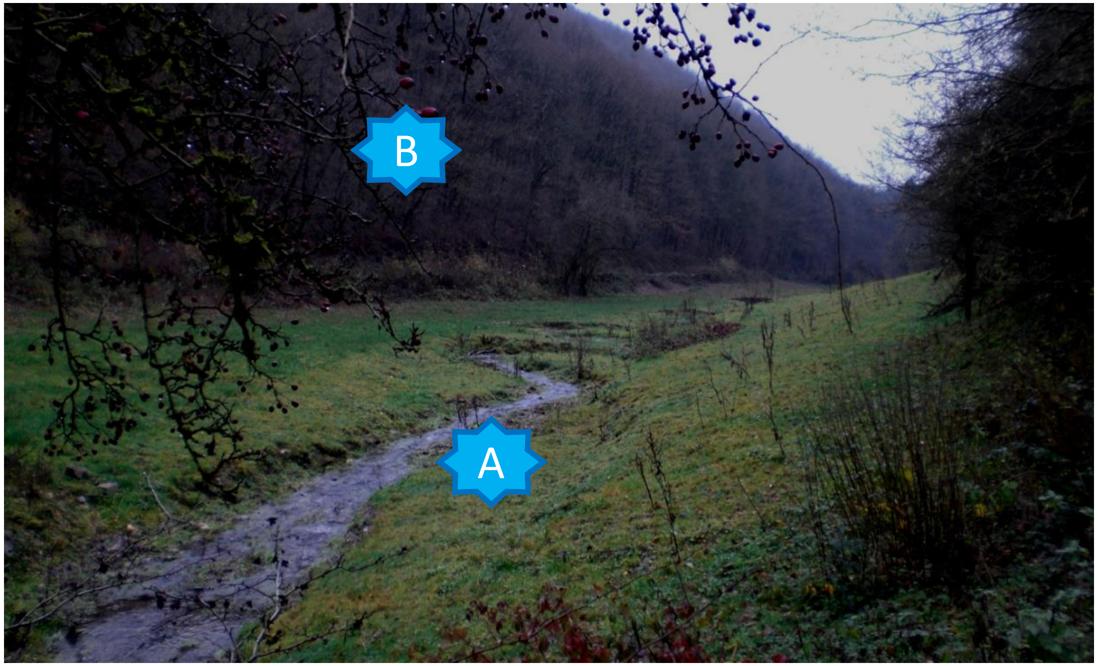


Best parameter set:
Slope = 0.113
HAND = 4.67 m

Flex-TOPO: Case study

Comparison with Topographic Wetness index

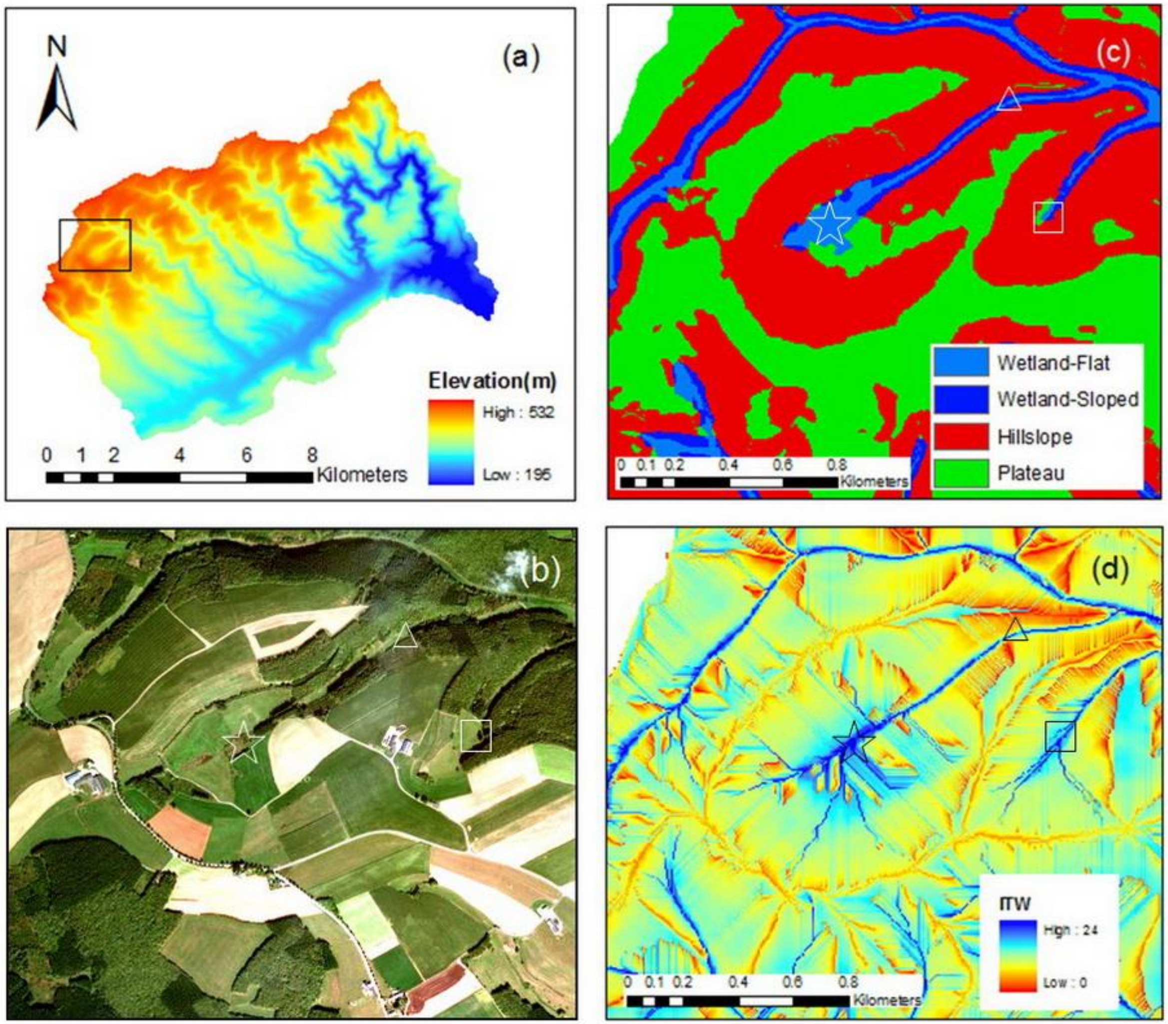
TWI=ln(A/s), TOPMODEL(Beven and Kirkby 1979)



For point A:
WI=ln(A↑/s↓)
WI ↑

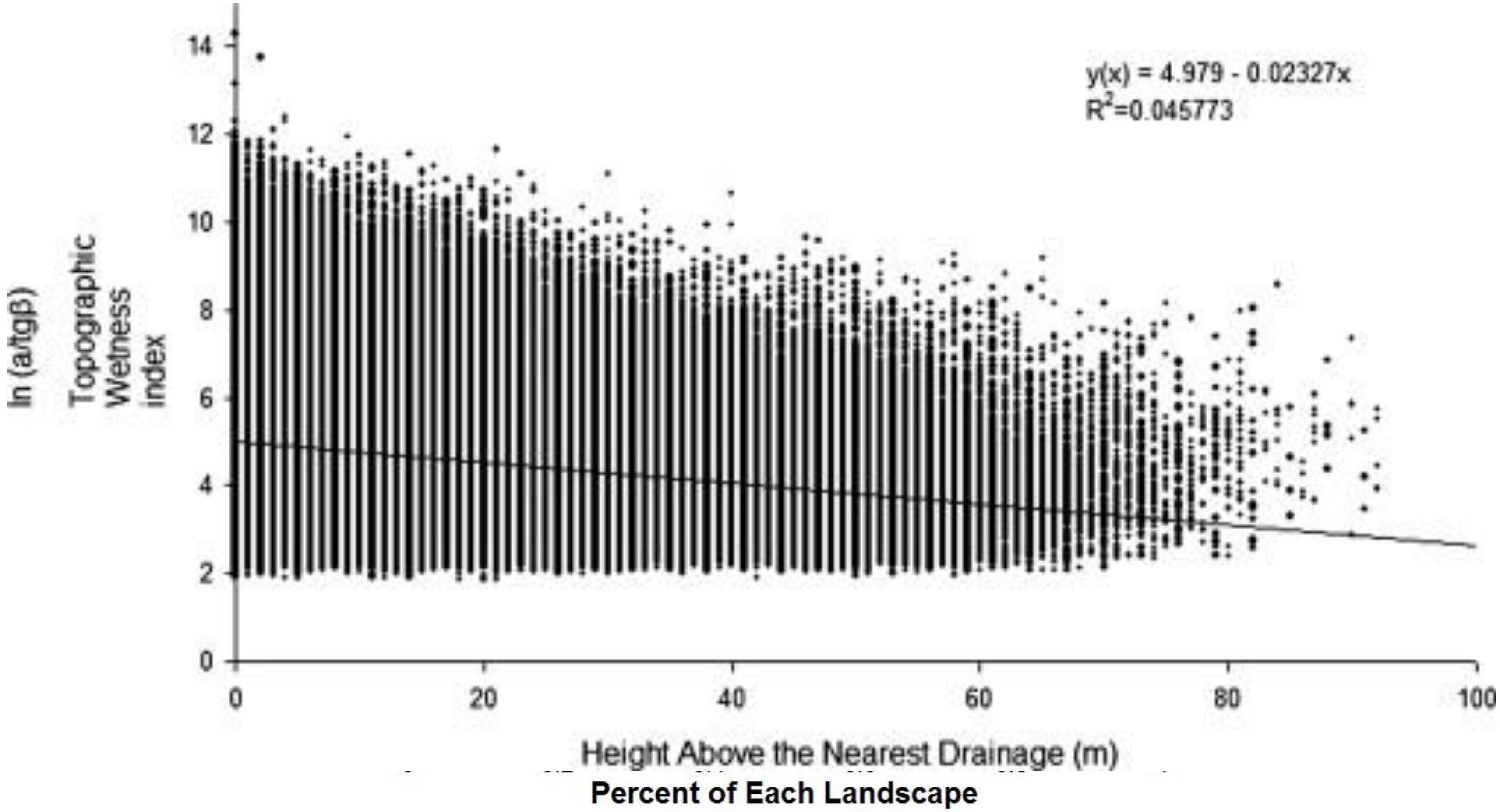
For point B:
WI=ln(A↓/s↑)
WI ↓

Flex-TOPO: Case study



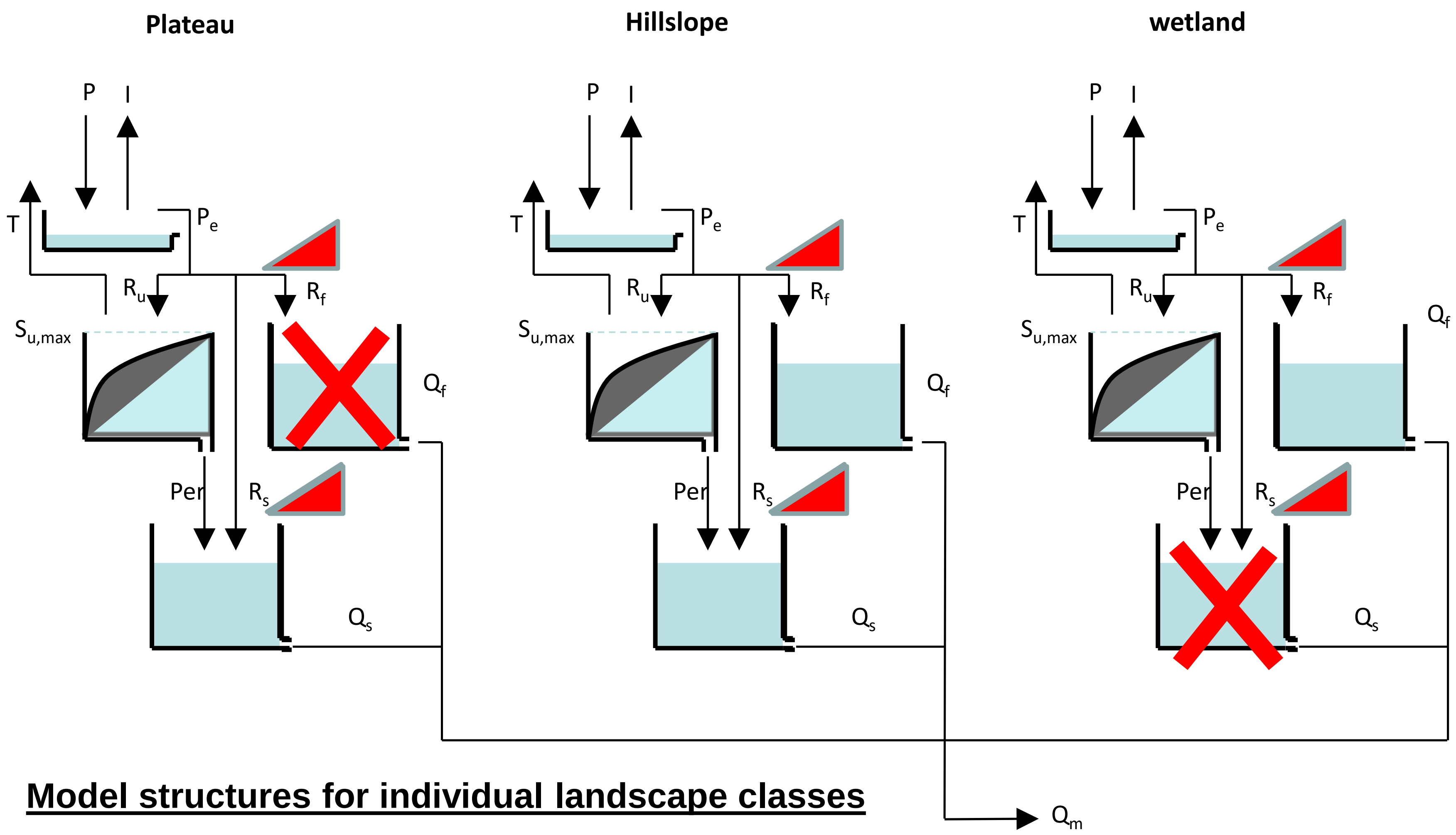
Flex-TOPO: Case study

Comparison with Topographic Wetness index



(Nobre et al., 2011)

Flex-TOPO: Case study



Model structures for individual landscape classes

Flex-TOPO: Case study

<u>What do we know??</u>

Example:

<u>What do we want the model to do?</u>

1 - Wetlands response faster to rainfall than hillslope and plateau:

- The fast reaction of wetland should be triggered earlier than other entities.
- Wetlands are active during dry period.

2 - Hillslope of the Wark catchment are mostly forested:

- Interception threshold of hillslope should be higher than the other landscape entities.

Flex-TOPO: Case study

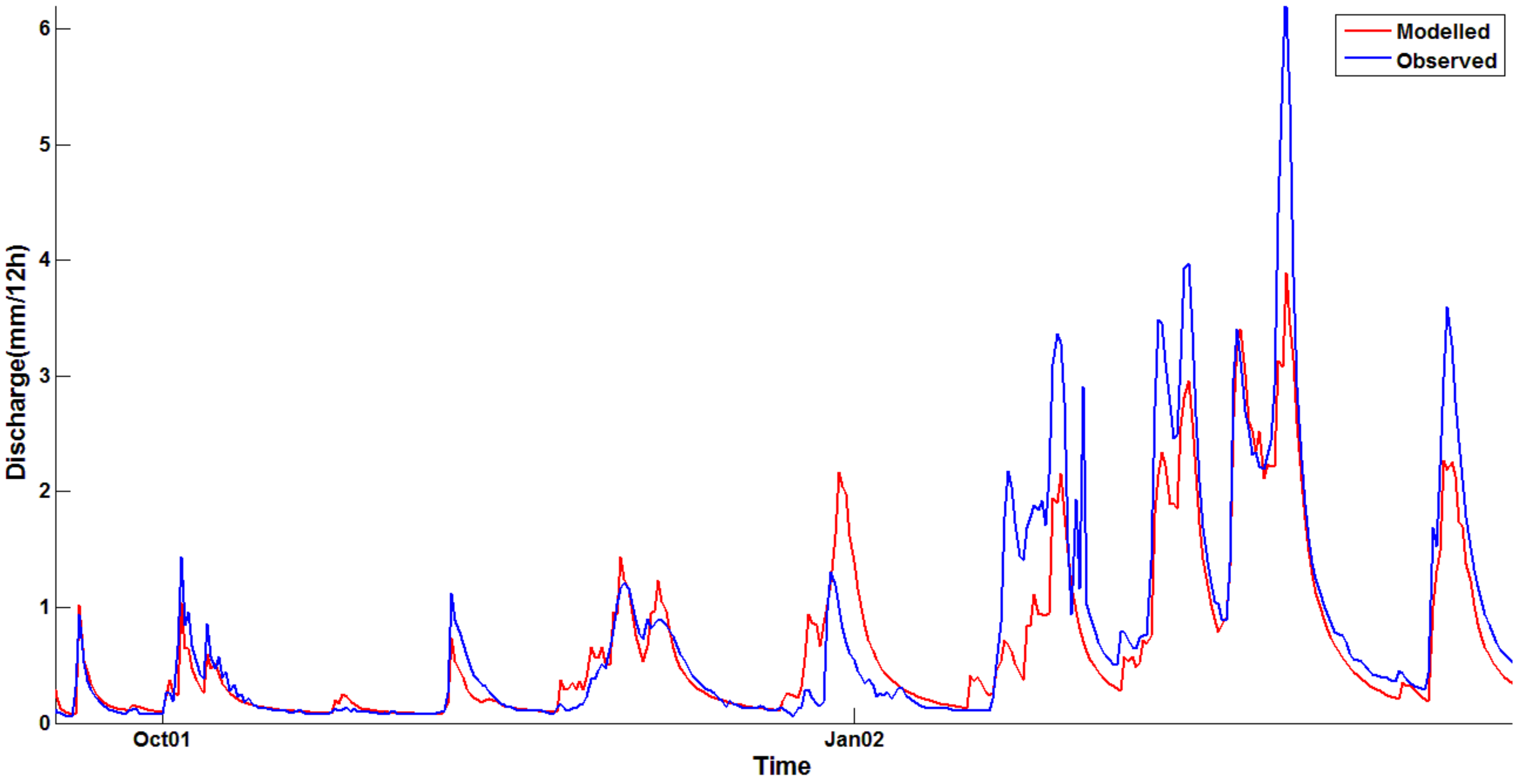
Calibration

- Possible soft criteria were formulated, e.g. Imax hillslope > Imax Plateau

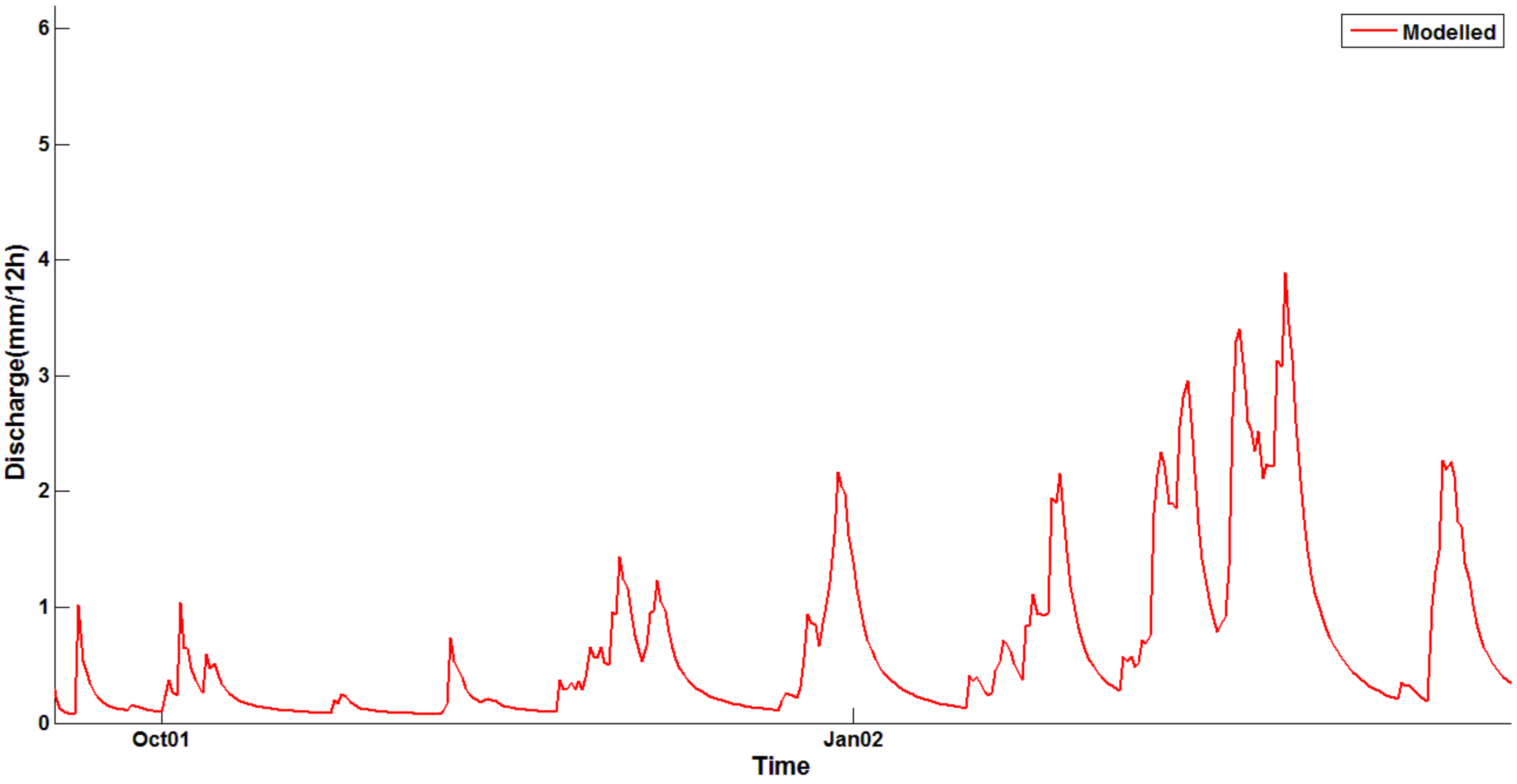
- In a stepwise approach with help of MOSCEM_UA (Vrugt et al.,2003) more and more criteria were satisfied.

- The optimum solution was found in the area of the parameter space which satisfied all soft criteria.

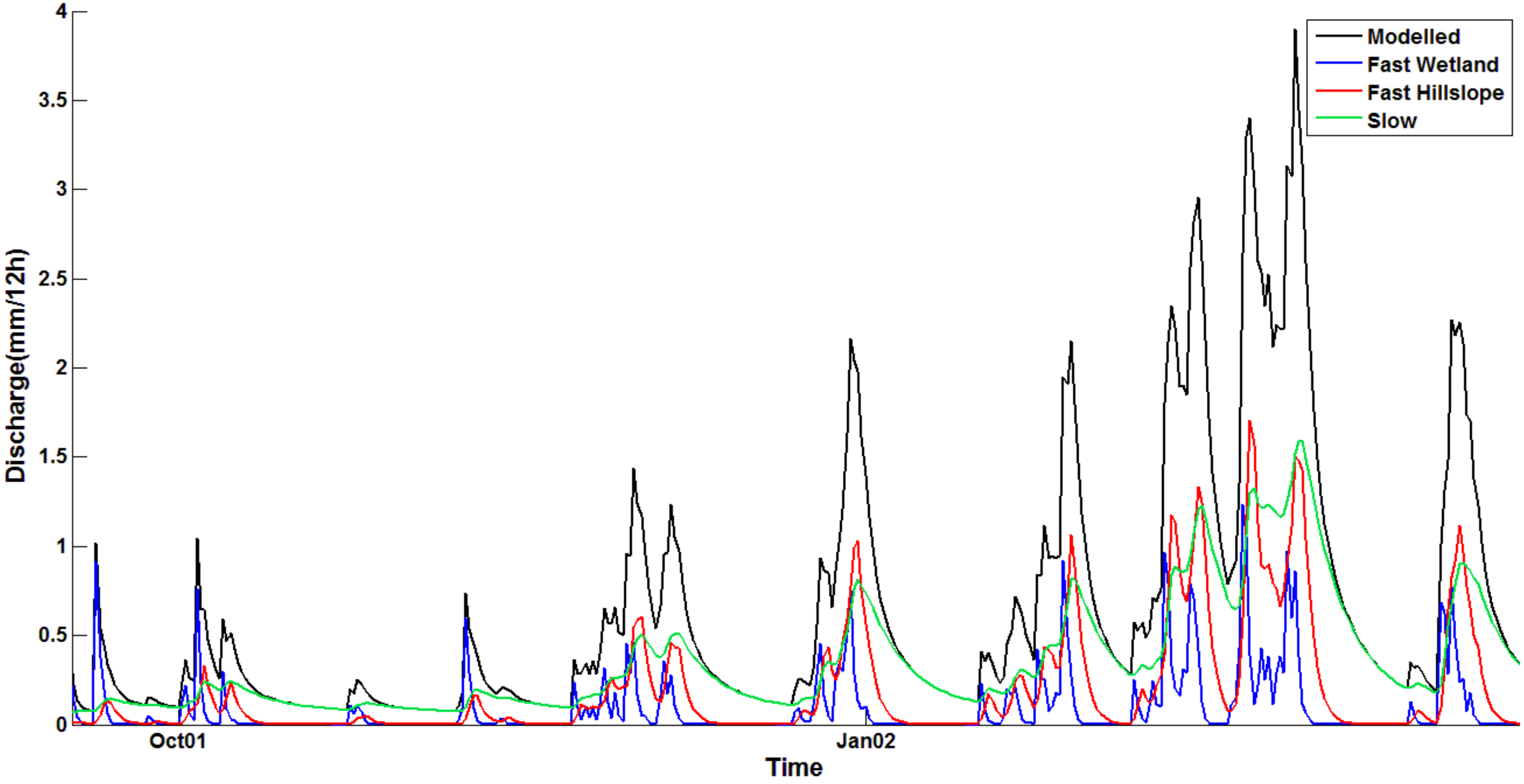
Flex-TOPO: Case study



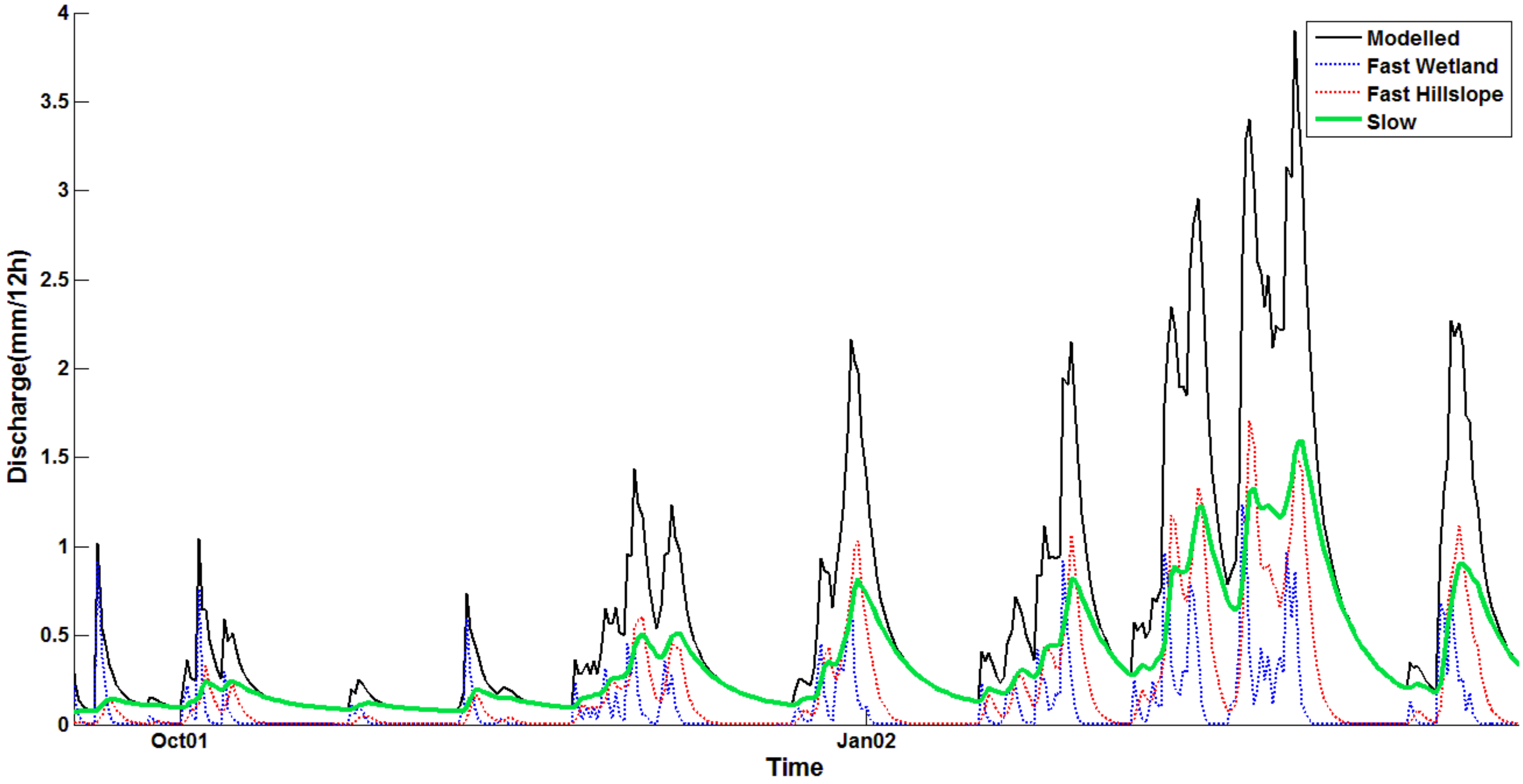
Flex-TOPO: Case study



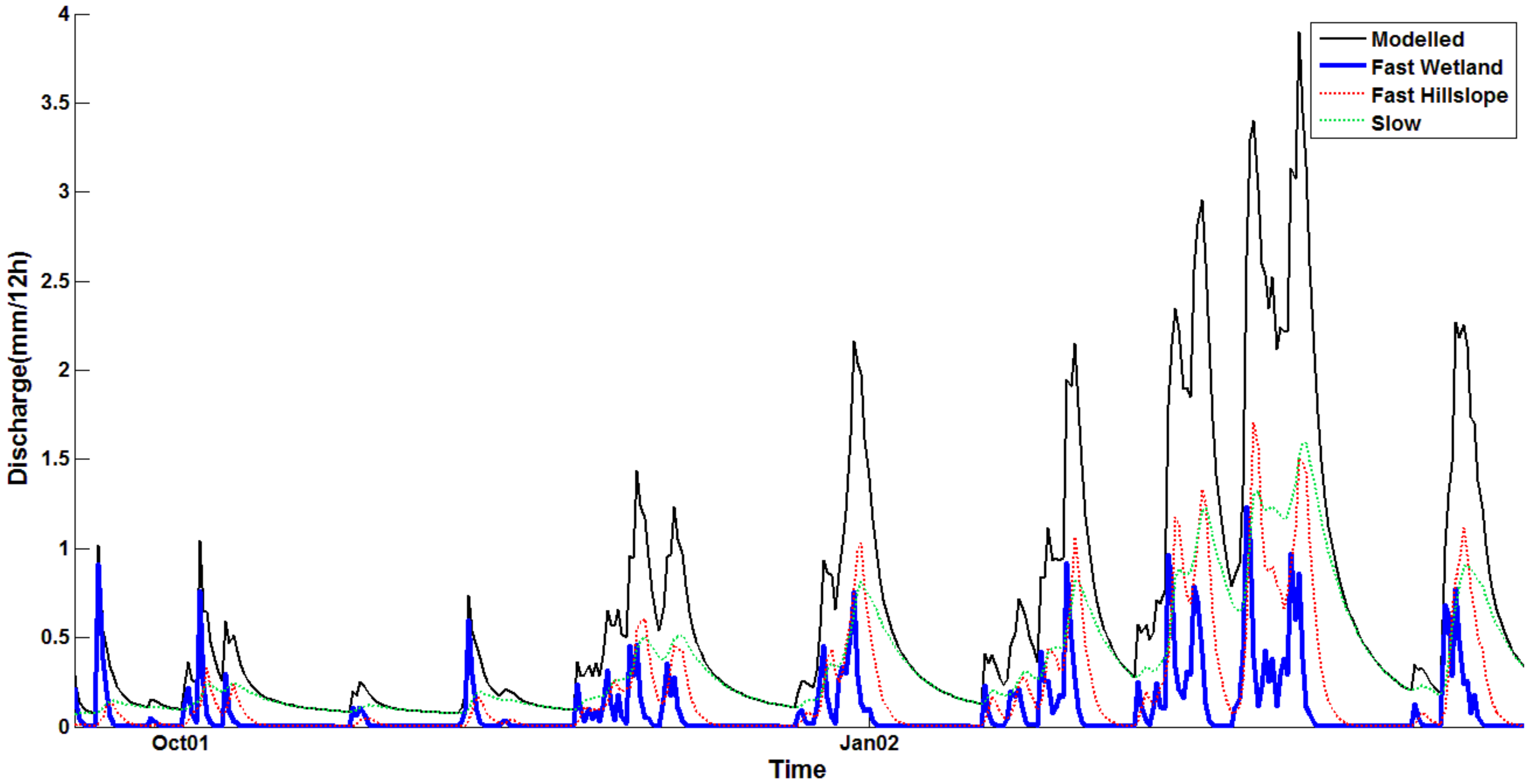
Flex-TOPO: Case study



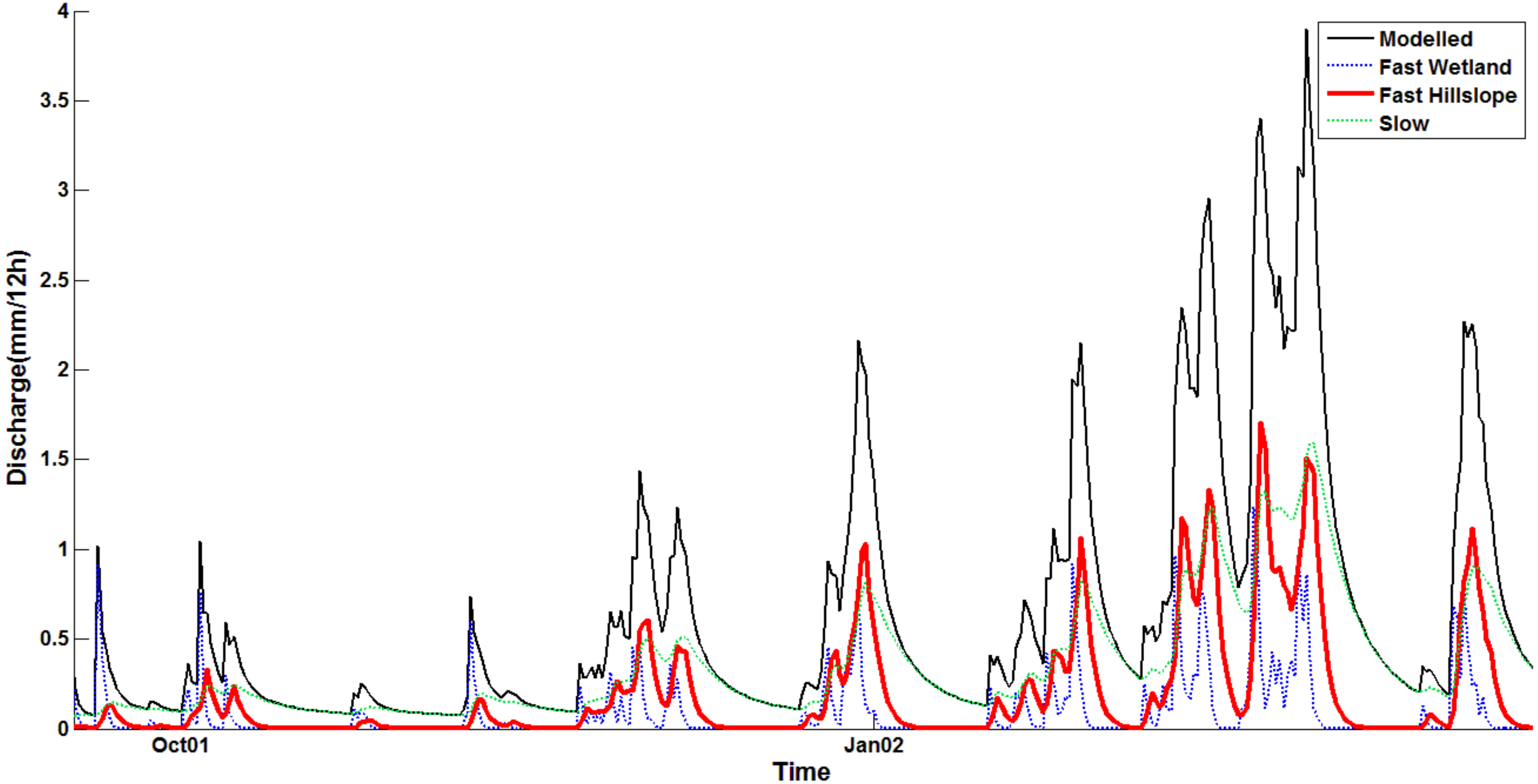
Flex-TOPO: Case study



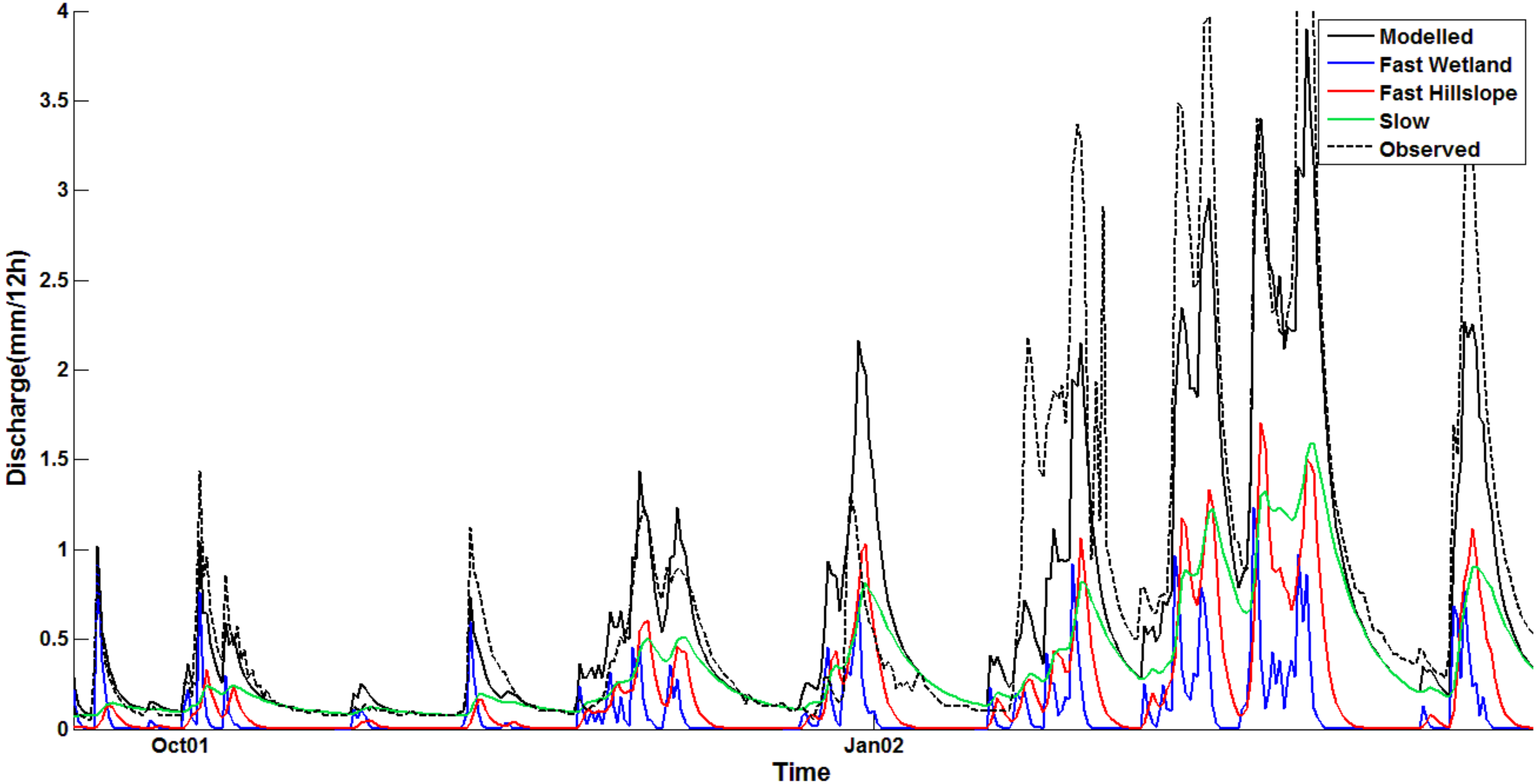
Flex-TOPO: Case study



Flex-TOPO: Case study



Flex-TOPO: Case study



Getting the right answers for the right reasons

Markus Hrachowitz
m.hrachowitz@tudelft.nl
Room 4.87

Getting the right answers for the right reasons

Or: Why does my model produce weird results ???

After this lecture you will

- know which problems can be introduced by data
- know why data pre-processing is important
- be able to interpret model performance
- know how to diagnose sources of model error
- be able to improve model performance

Model performance vs. realism

Models should reflect real world processes as exactly as possible to ensure high predictive power of a chosen model.

Conceptual models rely on spatially and temporally averaged inputs, effective parameters and calibration.

Frequently the model performance is wrongly inferred from the ability of the model to mimic the observed variable during calibration.

→ Reduces selection of best parameter set to curve fitting.

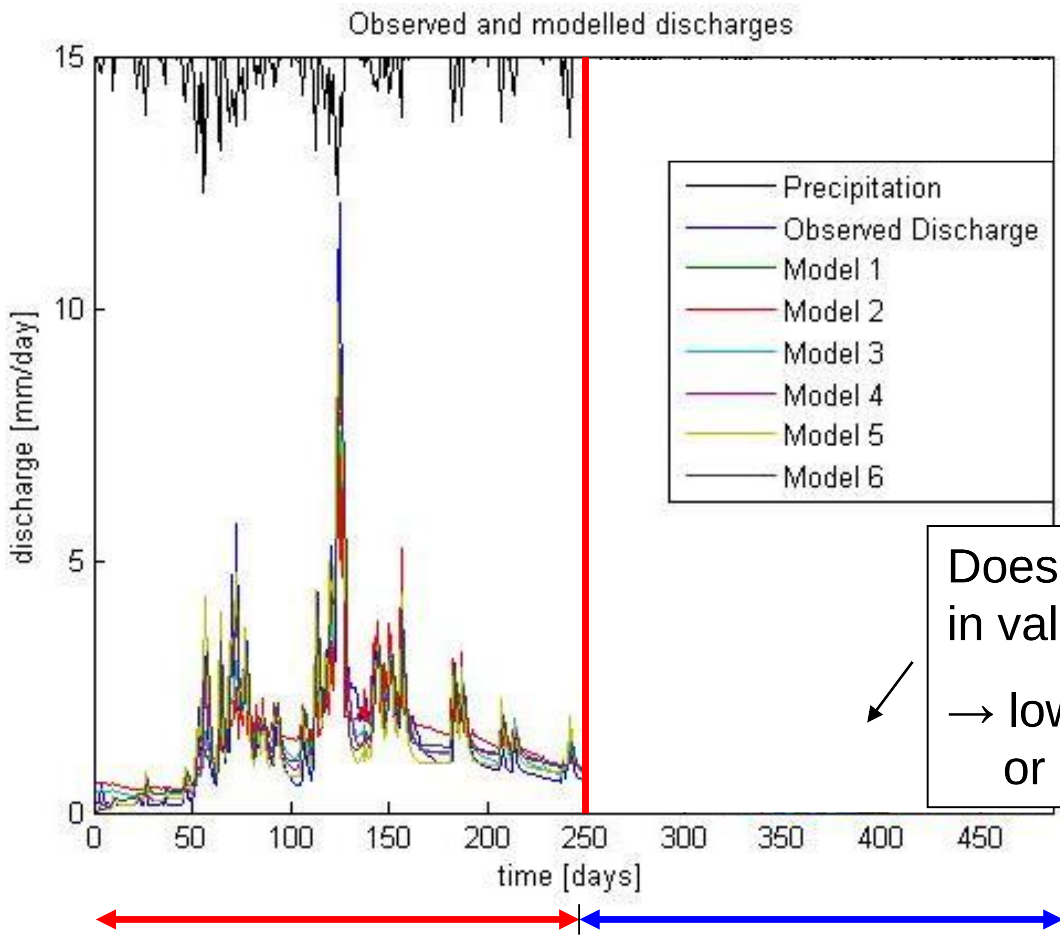
This can result in considerably reduced performance in independent test periods (i.e. validation periods): the high degrees of freedom of a typical conceptual model can produce excellent fits with unsuitable parameter sets: “Crap in, crap out models”

→ Parameters sets with reduced performance can result in higher predictive power!

→ Model performance should be evaluated according to reality checks and assessment of performance during validation periods

→ **“model realism”**

Model performance vs. realism example



Does not pick up peaks in validation period

→ low predictive power or realism

Calibration

Good performance

Validation

Bad performance

Steps towards a useful model

(1) Screening and cleaning of data
(2) Choice of appropriate interpolation methods for input data
(3) Choice of appropriate time-stepping of model (dictated by data and objective)
(4) Plausibility check of observed runoff data (rating curve!)
(5) Choice of an appropriate model spin-up period
(6) Choice of a plausible model structure based on perception of the system
(7) Choice of an appropriate numerical method to solve the water balance
(8) Choice of an appropriate model calibration method
(9) Choice of appropriate objective functions
(10) Testing (validation) of the model to infer most adequate parameterization

Data screening

This is a bit of a no-brainer:

Check the available data for plausibility and obvious errors (e.g. temperature sensor shows -32°C in southern Spain during summer). Important when data come from different sources.

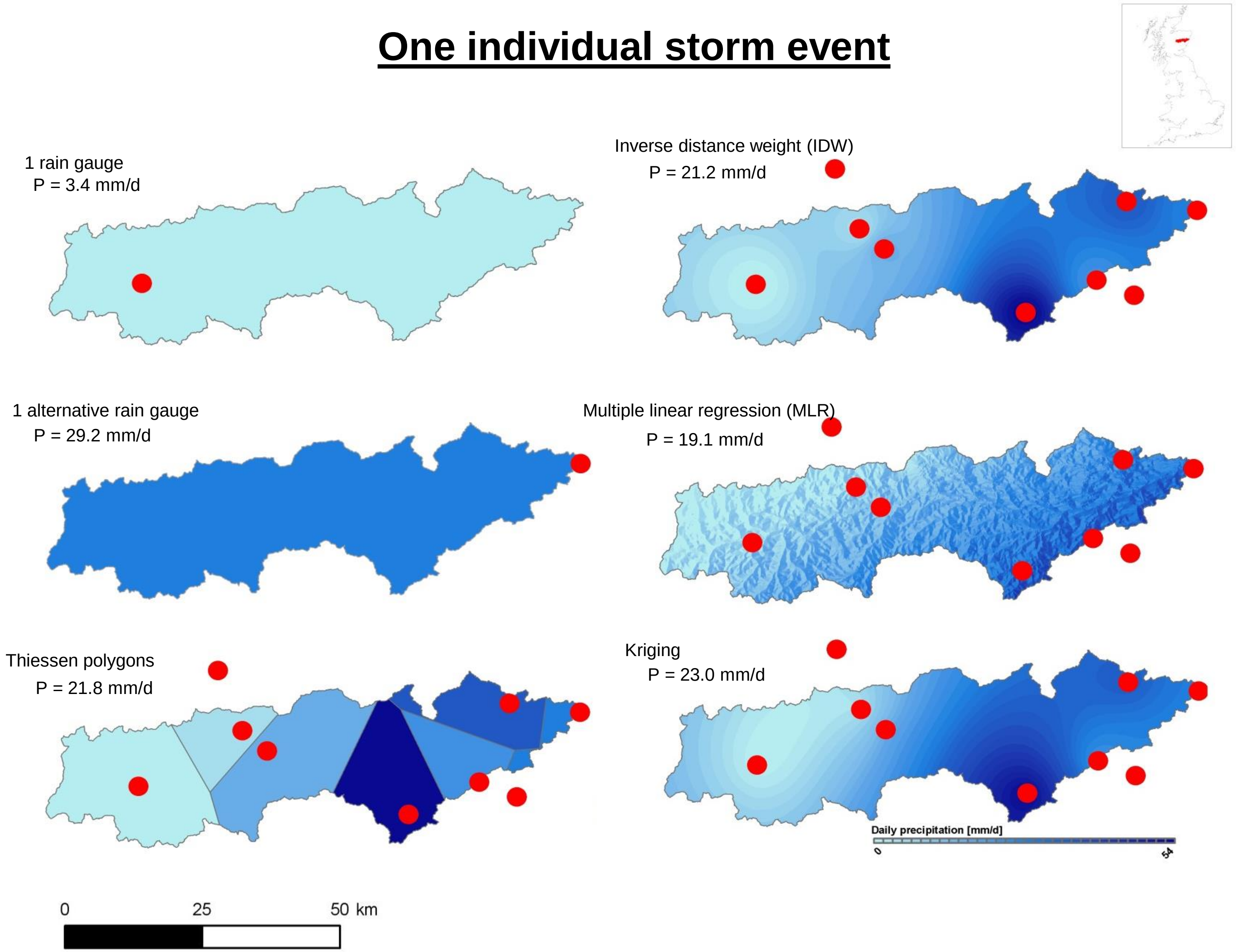
Where possible and appropriate close gaps in the input data series using an adequate method, e.g.:

- Long term averages
- (Non-)Linear interpolation between previous and following available data points
- Spatial interpolation (with or without the use of spatial correlation features) if data from nearby climate stations are available.

Gaps in stream flow observation do not have to be closed (only used for calibration!)

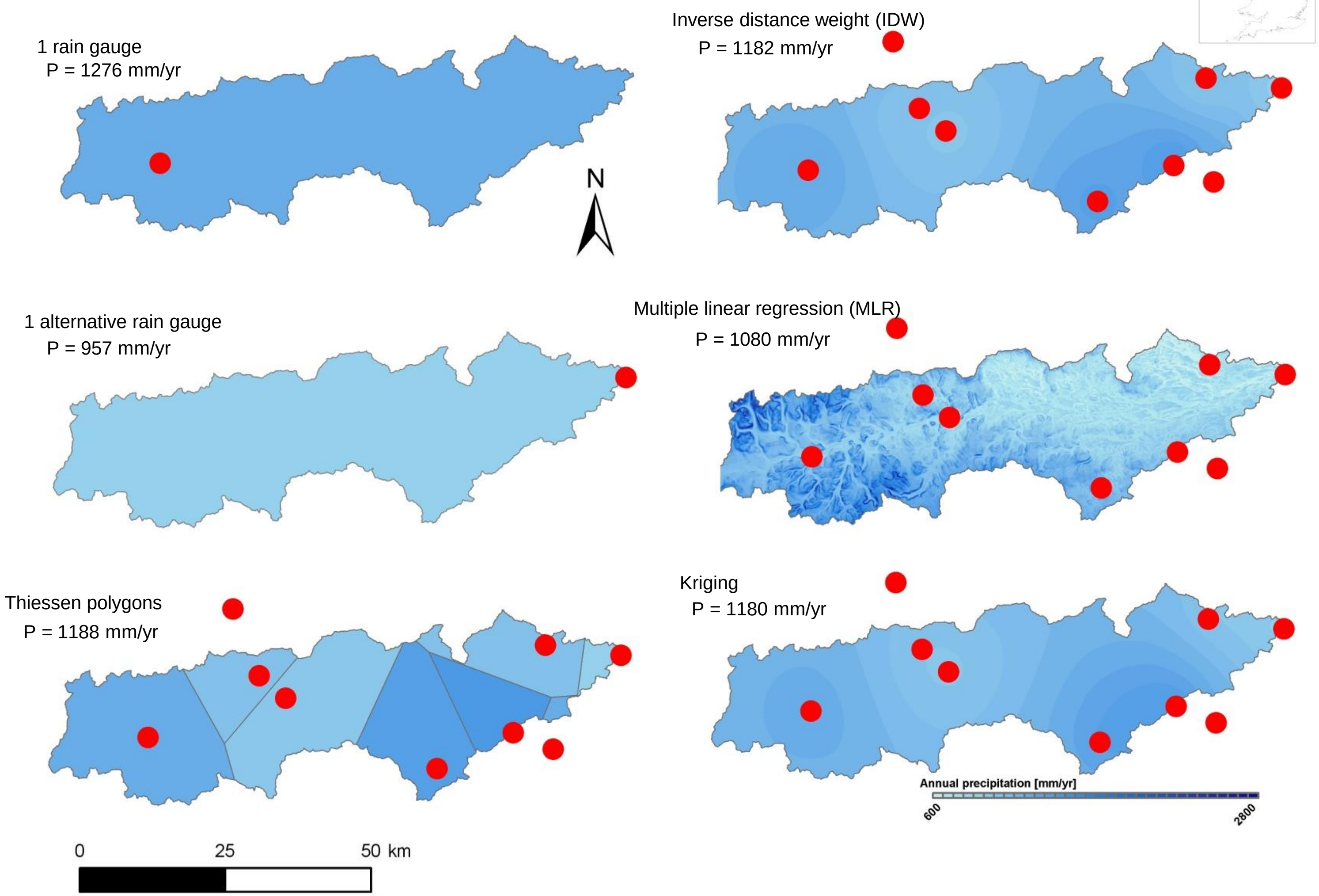
Uncertainty in observed precipitation

One individual storm event

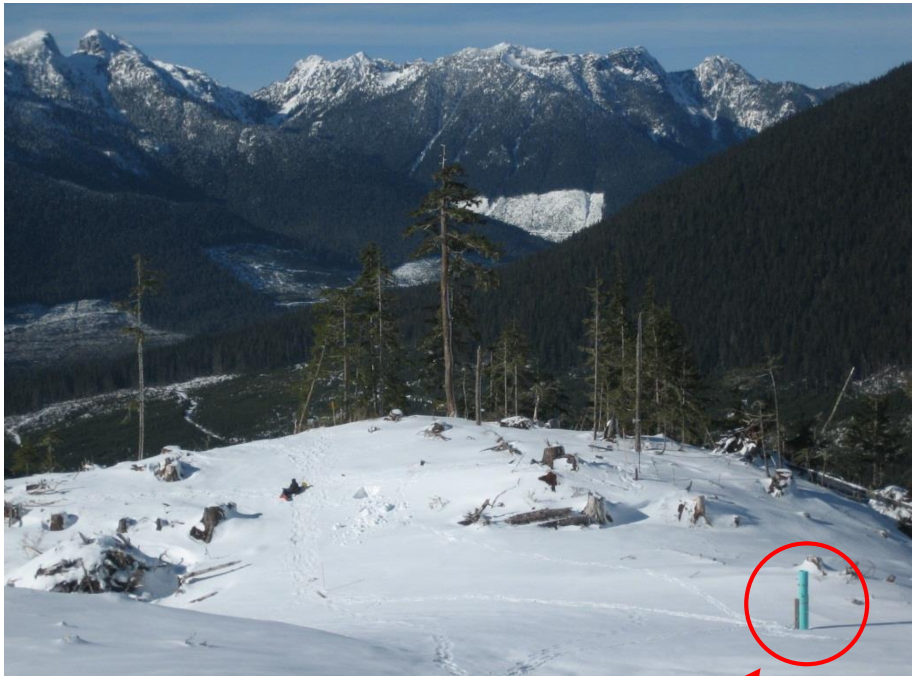


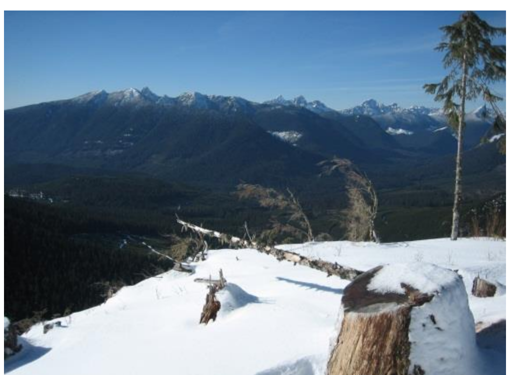
Uncertainty in observed precipitation

Long annual average precipitation



Precipitation uncertainty

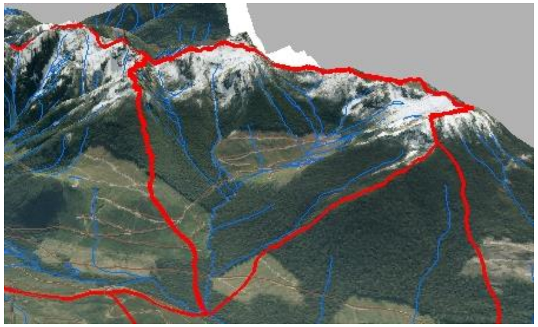


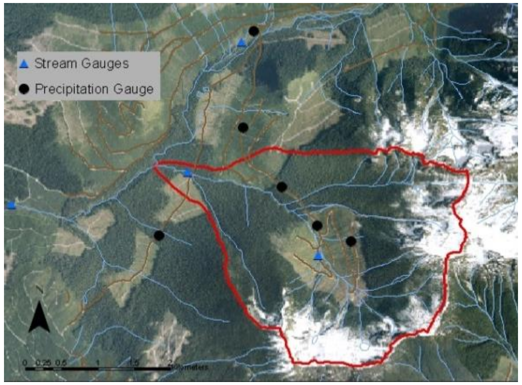


BAD location!

No snow measured, but……

Precipitation uncertainty – case study



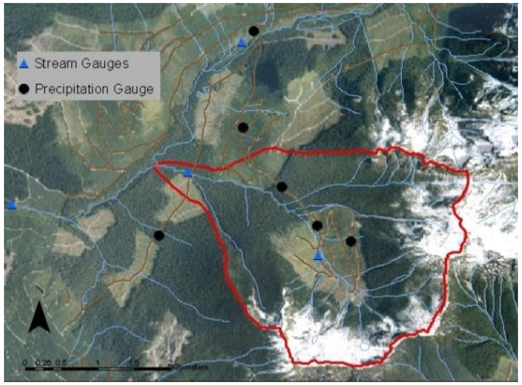


Stephanie Creek experimental catchment

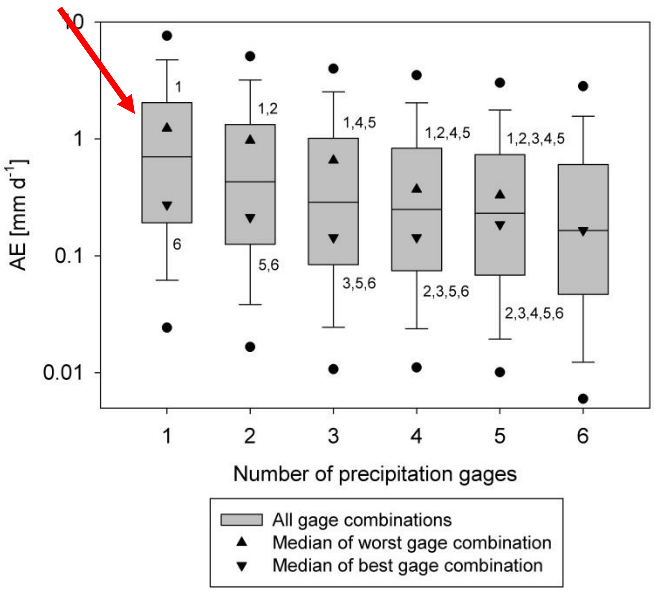
Coastal British Columbia (8 km^2), highly dynamic meteorological conditions, pronounced relief, 2000mm/yr

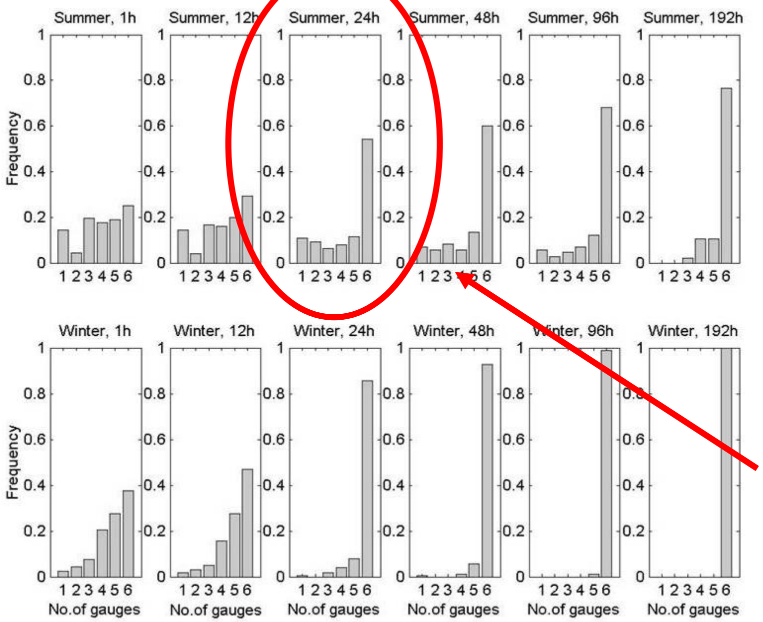
Study on the effect of climate station network design on average catchment precipitation estimates

Precipitation uncertainty – case study



Deviation from best estimates increases with lower number of gauges. Mean deviation for single gauges ~700 mm/yr or 40%!!





In more than 40% of the rain events (summer), not all gauges recorded rain within 24h.

In 10% of the events, only ONE gauge recorded rain, even in this small catchment (8km^2)!!

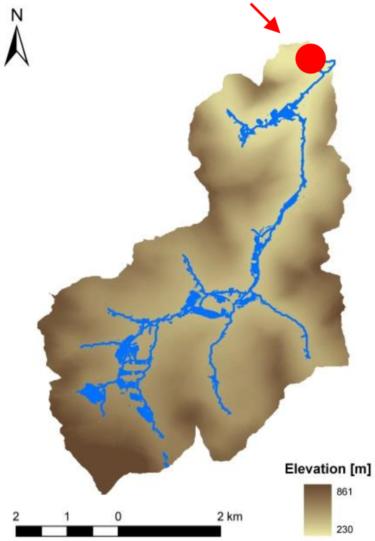
Precipitation uncertainty – effect on model

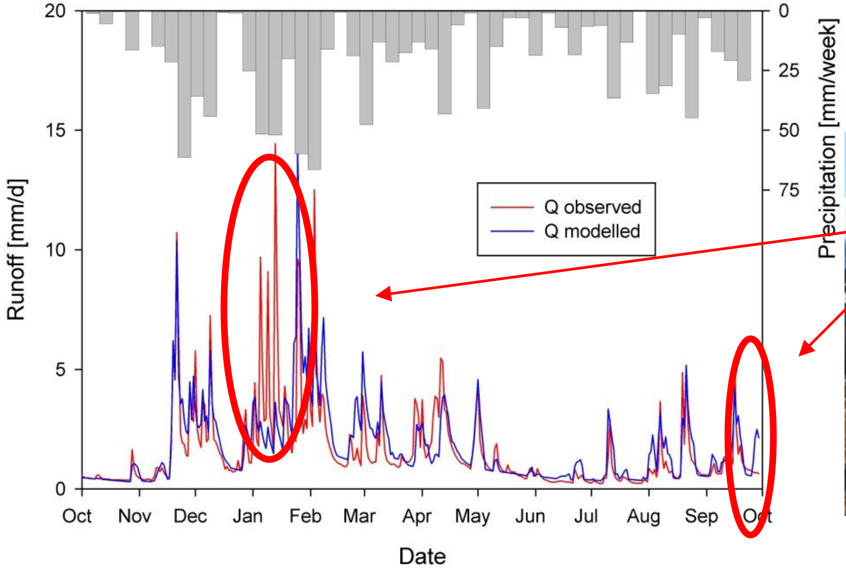
<u>Girnock Burn experimental catchment</u>

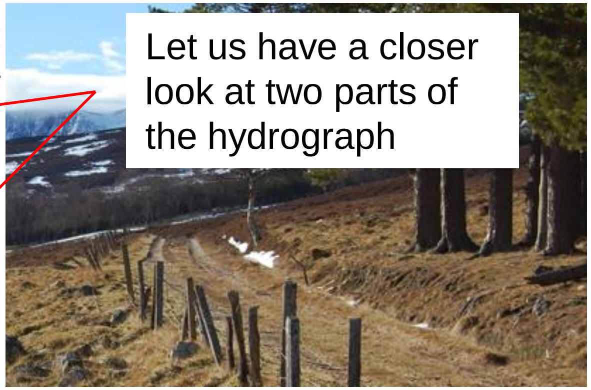
Cairngorm Mountains, Scotland (~30 km²), one(!) climate station, 1100mm/yr

Several models (5 – 12 parameters), NSE ~ 0.7, NSE-log ~ 0.85

Climate station







Let us have a closer look at two parts of the hydrograph

Precipitation uncertainty – effect on model

Girnock Burn experimental catchment

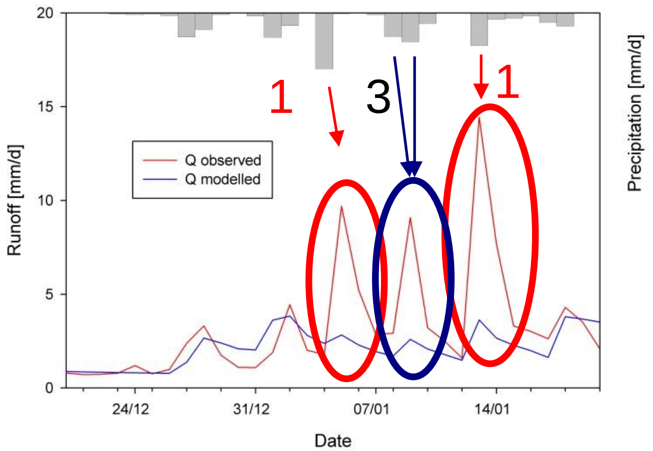
Several different types of potential errors can be identified in these two short periods:

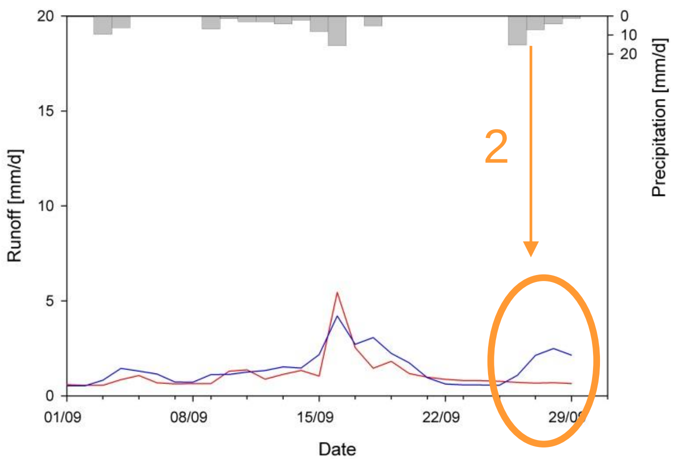
1 – Recorded precipitation amount is too low → Q is underestimated

2 – Recorded precipitation amount is too high → Q is overestimated

Not only related to the number or position of gauges:

3 – Recorded precipitation amount is either too low or rain from the preceding time step t-1 should count towards the peak at t → Q is underestimated





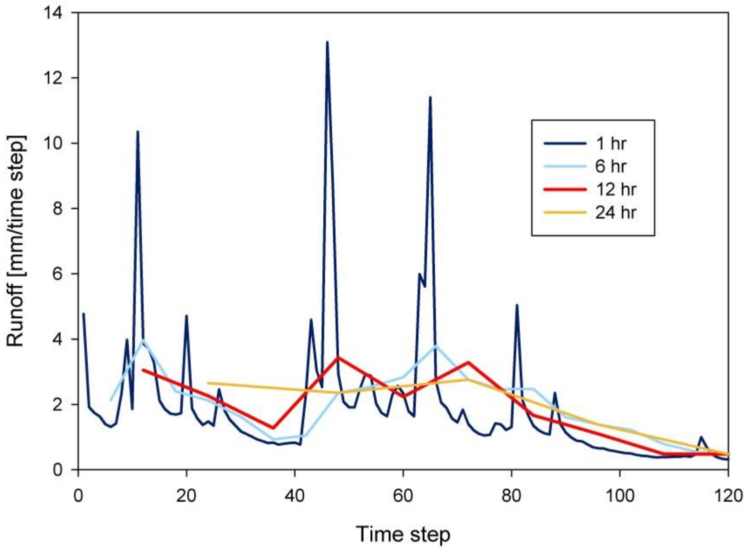
Temporal resolution of data

Observation data resolution and discretization

Problem: continuous processes in nature need to be discretized.

Or in other words: observation technology does not allow us to continuously record variables of interest → everything we are observing (e.g. rain, stream flow,…) is averaged over the observation time interval and likely to be subject to a time lag.

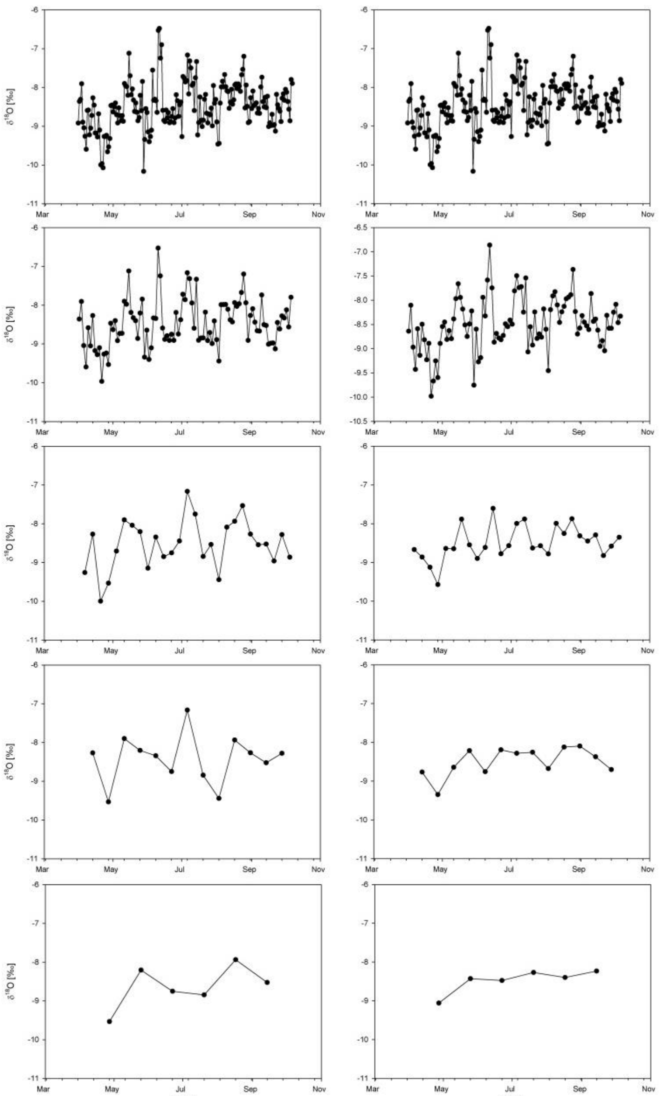
Commercial data loggers can record stream flow with a frequency ≤ 1 second^{-1} (in theory!!) and rainfall ≤ 1minute^{-1} (in theory!!).
In reality observation frequencies between 1 hr^{-1} and 1 day^{-1} are common.



Where only observation frequencies ≤ 1 day^{-1} are available → problematic for peak flow estimates and design values

Temporal resolution of data and sample type

Grab sample	Composite sample	Sampling interval



cf.Kirchner et al. (2004)

1 day

2 days

7 days

14 days

28 days

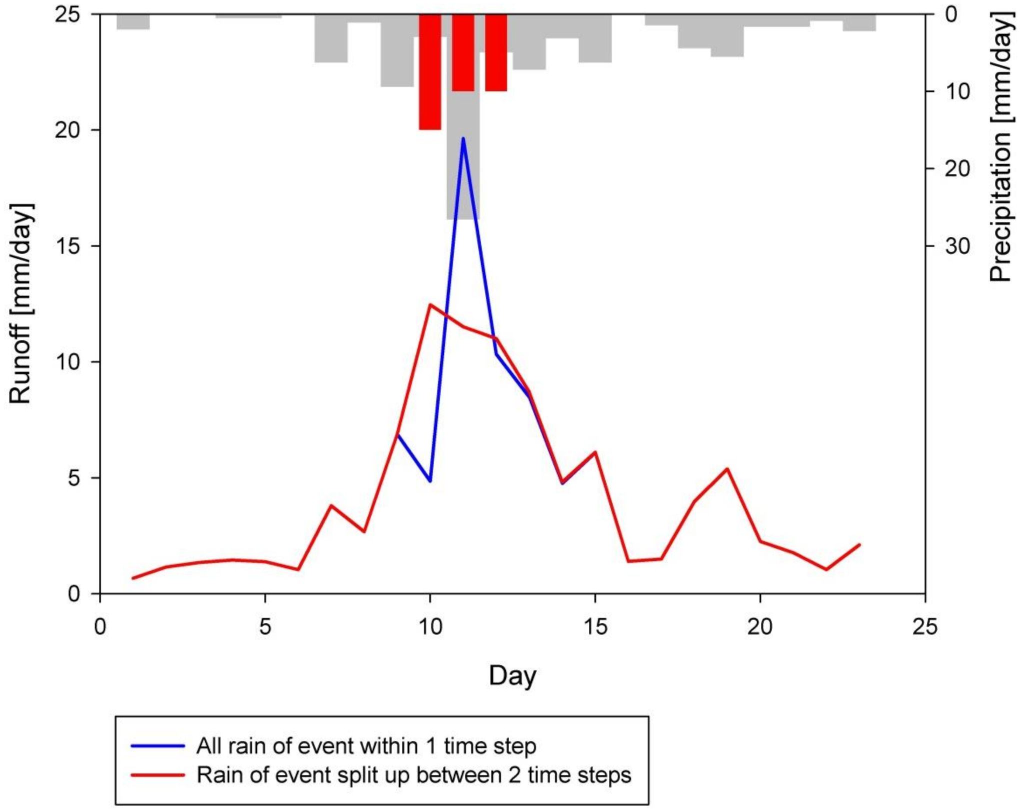
For determining the geochemical or isotopic composition of water, quasi-continuous sampling very difficult and expensive.

Sampling frequencies often < 1 week^{-1}

Note also the difference between different sample types. Precipitation usually composite samples and stream flow grab samples!

Timing of observations

Observ. time	Precipitation [mm/timestep] Hourly	Daily	Daily
00:00	0		
02:00	0		
04:00	3		
06:00	**0**		3
08:00	0		
10:00	0		
12:00	0		
14:00	0		
16:00	0		
18:00	0		
20:00	0		
22:00	**0**	3	
00:00	0		!!!
02:00	0		!
04:00	0		
06:00	**15**		15
08:00	12		
10:00	0		
12:00	0		
14:00	0		
16:00	0		
18:00	0		
20:00	0		
22:00	**0**	27	
00:00	0		!!!
02:00	0		!
04:00	0		
06:00	**0**		12



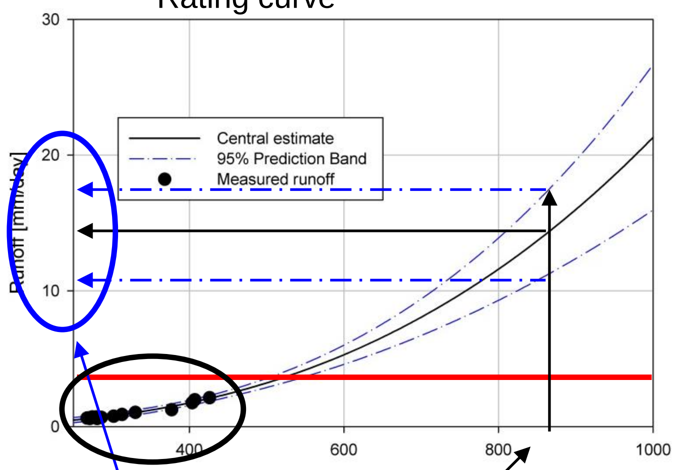
What can be the effect of changing the timing of observations?

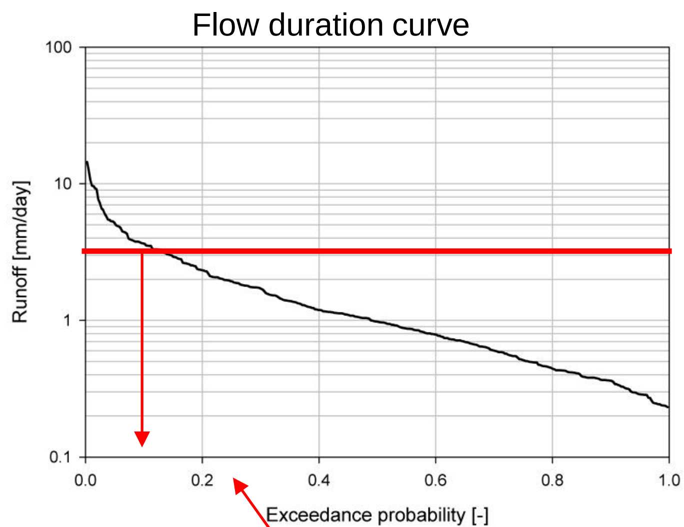
How to reduce errors from data

Attention!

- Be consistent in the use of rain gauges for extrapolation
- Adjust extrapolation method to data available
- Adjust extrapolation method to the objective of the research and to its spatio-temporal scale (e.g. peak flows, groundwater body,…)
- Be careful when using data from different sources
- Be careful if there is a change in the person responsible for operating field equipment
- Be careful if there is a change in the sensors and loggers
- Be careful if there is a change in observation frequency
- Be careful if there is a change in sample type

Uncertainty in runoff observation

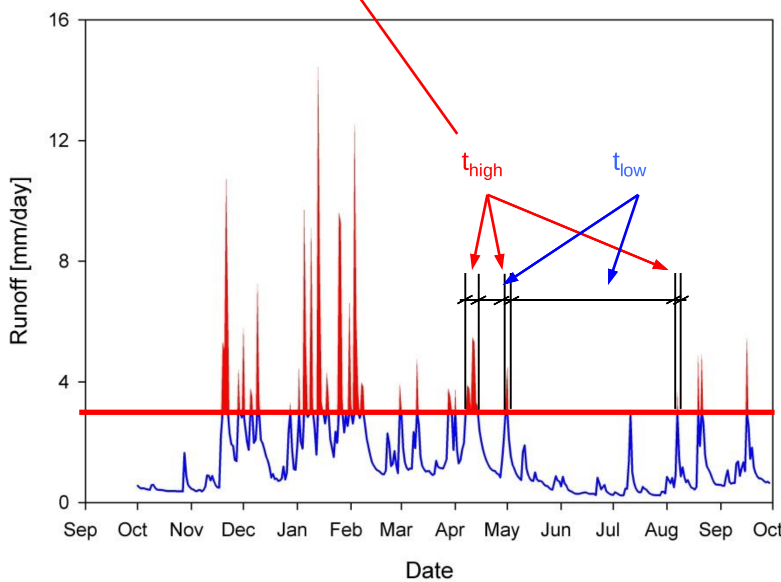




Max. recorded water level

11 < Q < 18 mm d^{-1} !!

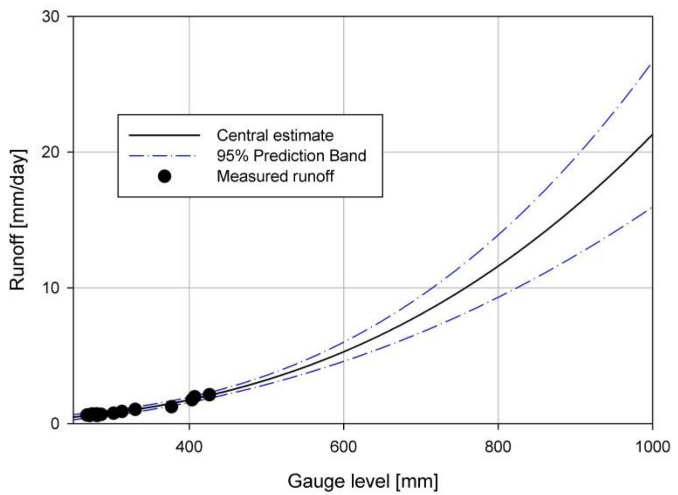
Measured runoff –
biased towards low flow!



Reasons:

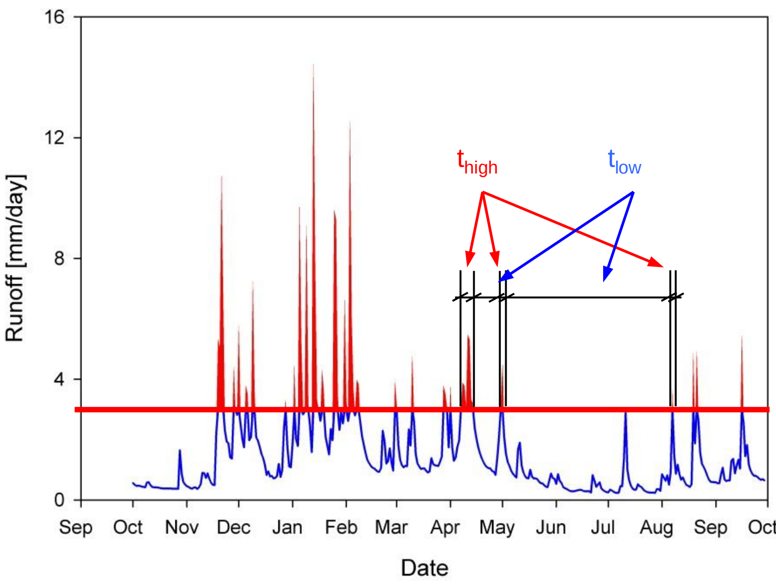
- Too few high flow measurements
- Uncertainty in runoff measurement (i.e. wide confidence interval)
  → problematic for extrapolation of high flows

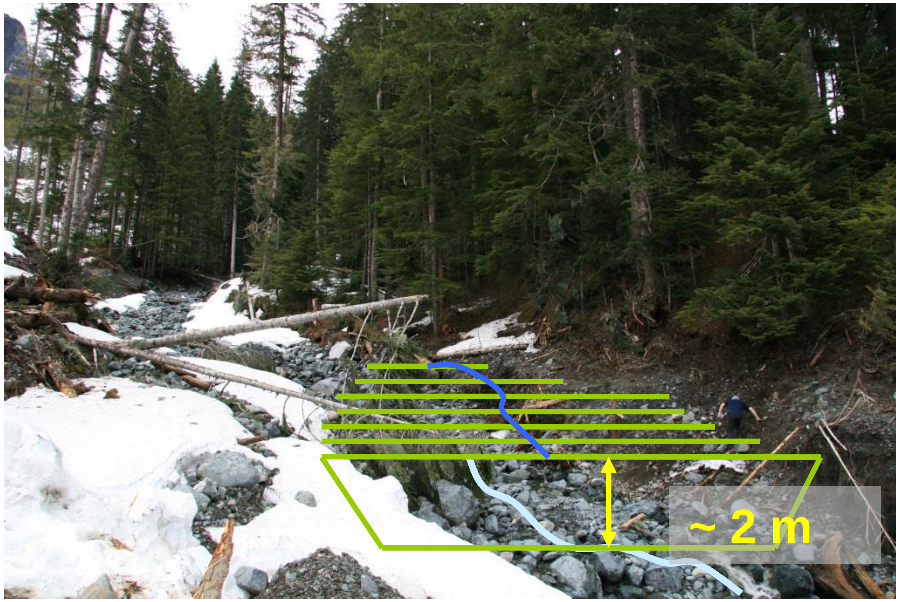
Uncertainty in runoff observation

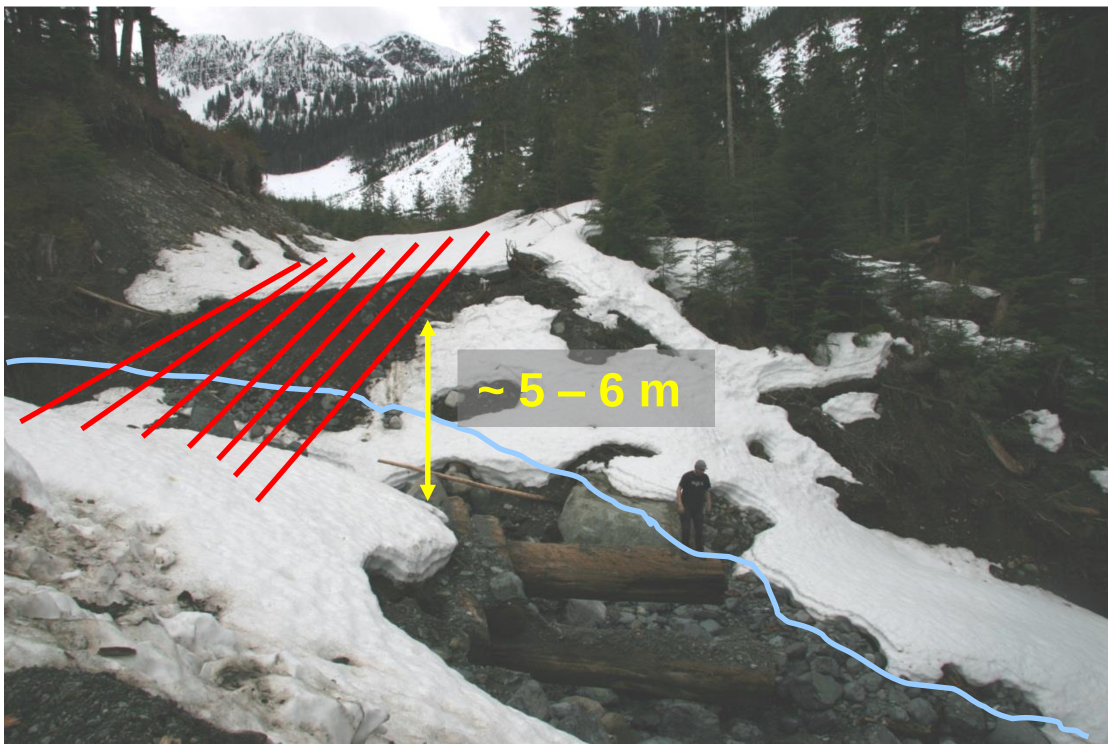


<u>Other reasons:</u>

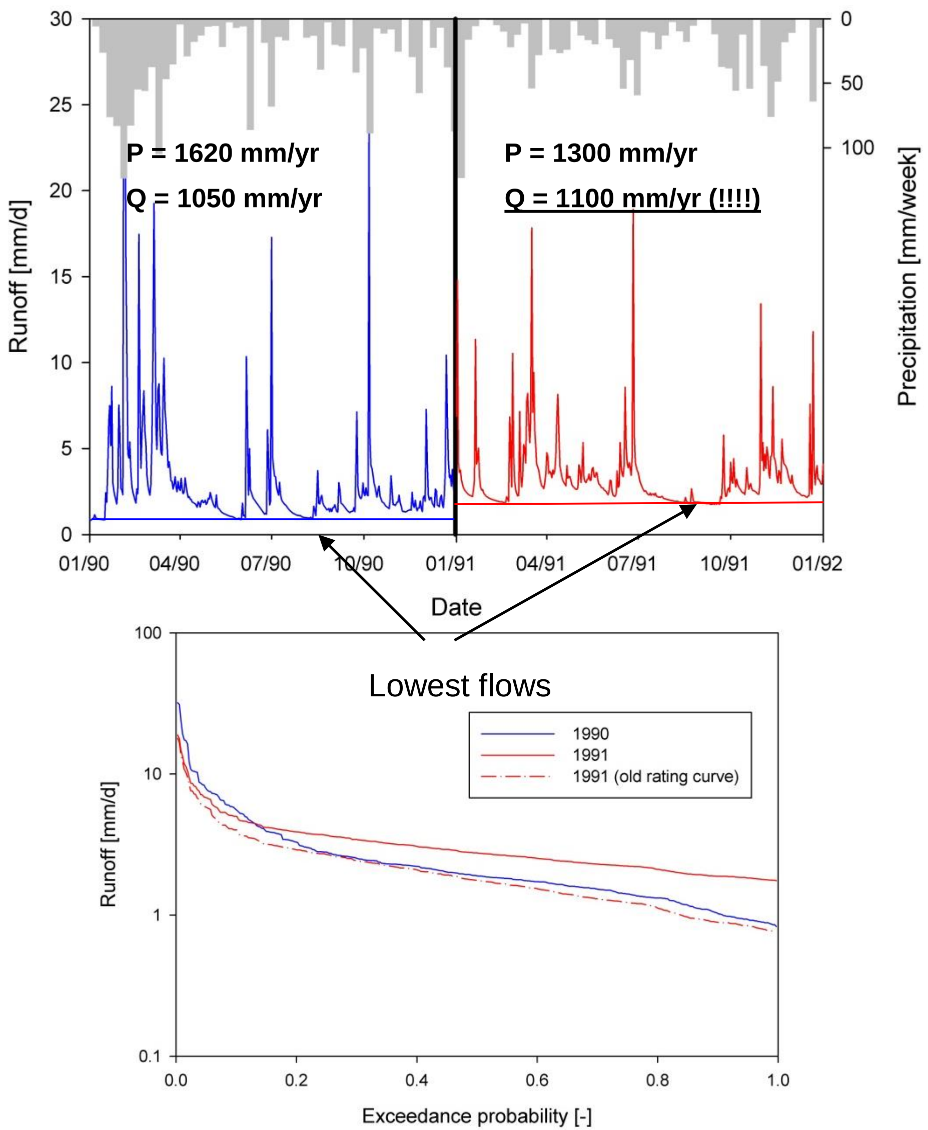
- Gradual or instantaneous change of stream cross section by bedload redistribution, river bank erosion or debris caused obstruction



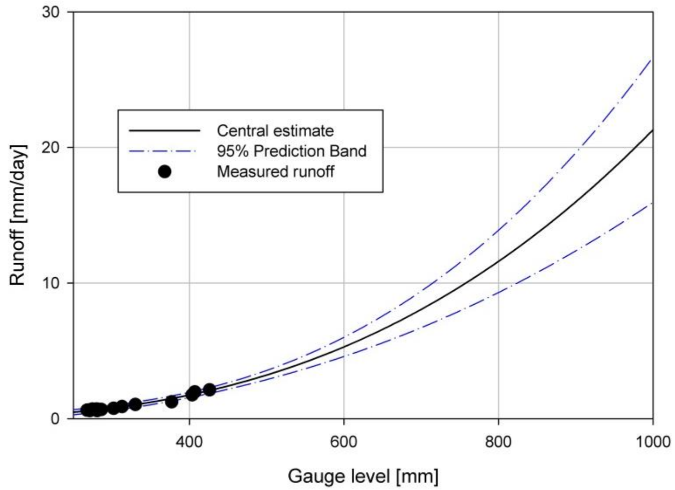

~ 2 m


~ 5 – 6 m

Uncertainty in runoff observation



Uncertainty in runoff observation



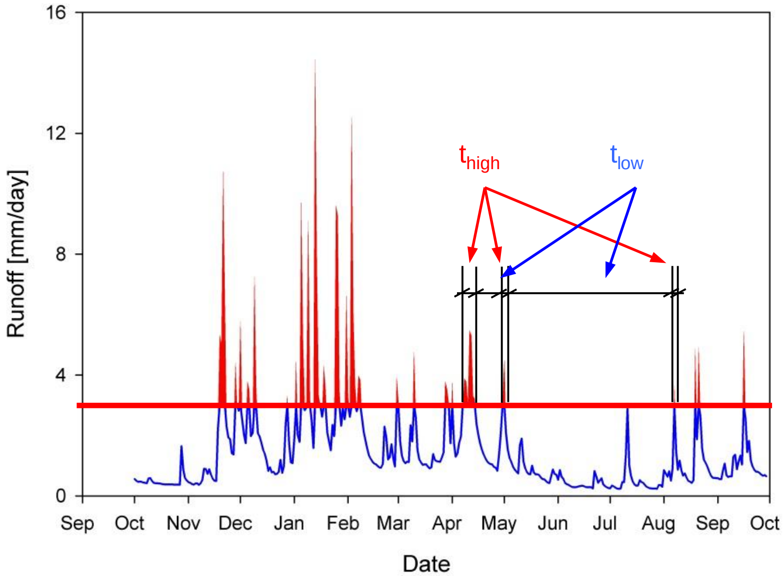
Other reasons:

- Gradual or instantaneous change of stream cross section by bedload redistribution, river bank erosion or debris caused obstruction

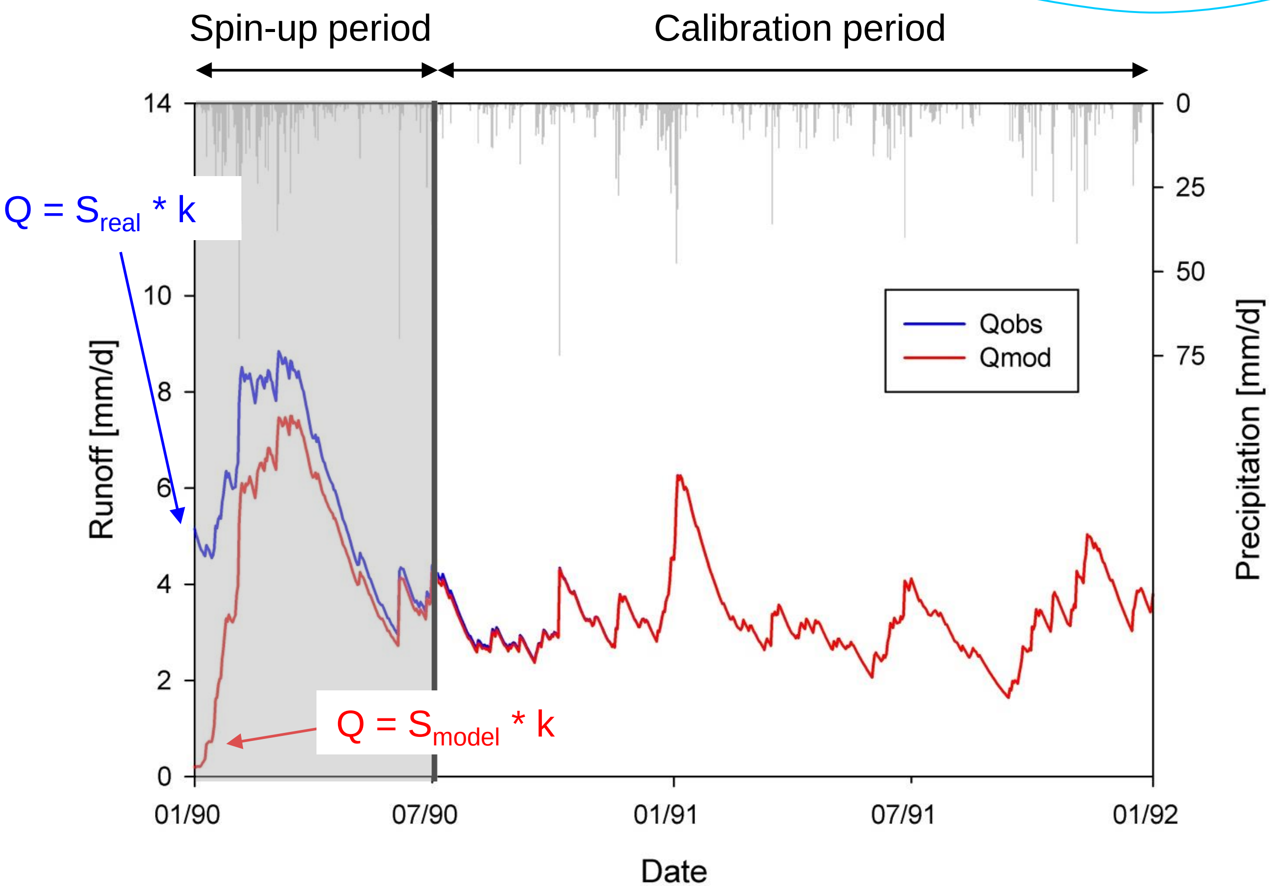
- Agencies concerned with data collection (e.g. National Meteorological Services) adapt rating curves in certain intervals. This can result in abrupt shifts in recorded flow.

- Unless fixed to a solid structure, water level sensors can be subject to movement (e.g. after maintenance, by bedload movement,…)

- Time averaging



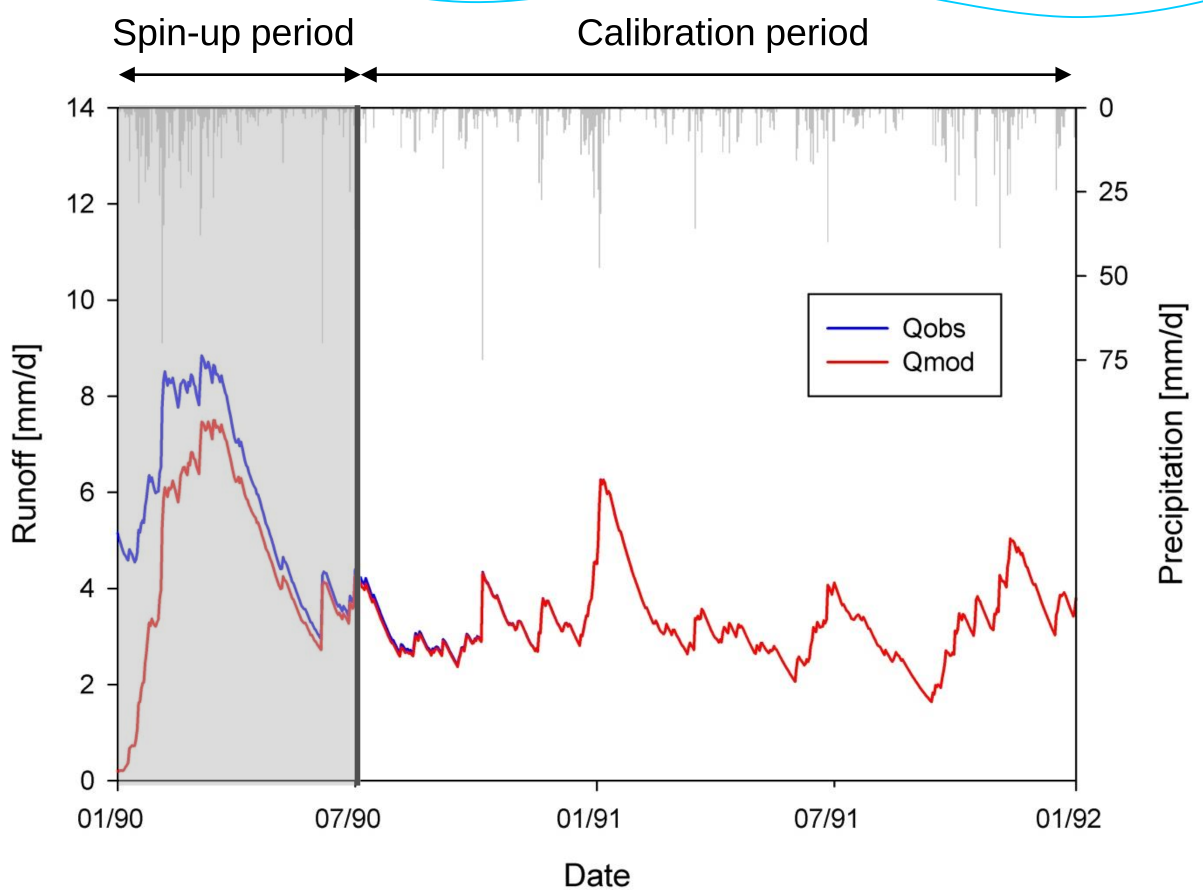
Model spin-up period



Simple linear reservoir: Q = S*k

<u>Initial condition</u> S(0) is <u>unknown</u> and arbitrarily chosen, here: S(0) = 0.
Model needs certain period ("warm-up", "spin-up", "burn-in" period) to bring the system states to values reflecting real world processes, here: storage S needs to fill up to "real-world" level.

Model spin-up period

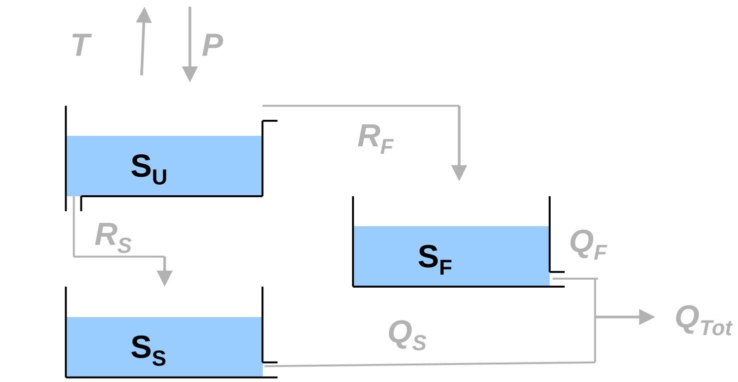


Spin-up period is <u>not considered for calibration.</u> Otherwise, model is forced to compensate for errors due to incorrect S(0)!!

Length of spin-up period is dependent on the catchment but is typically 1 year.

To reduce uncertainty, the calibration period should start in dry season, when S ~ 0

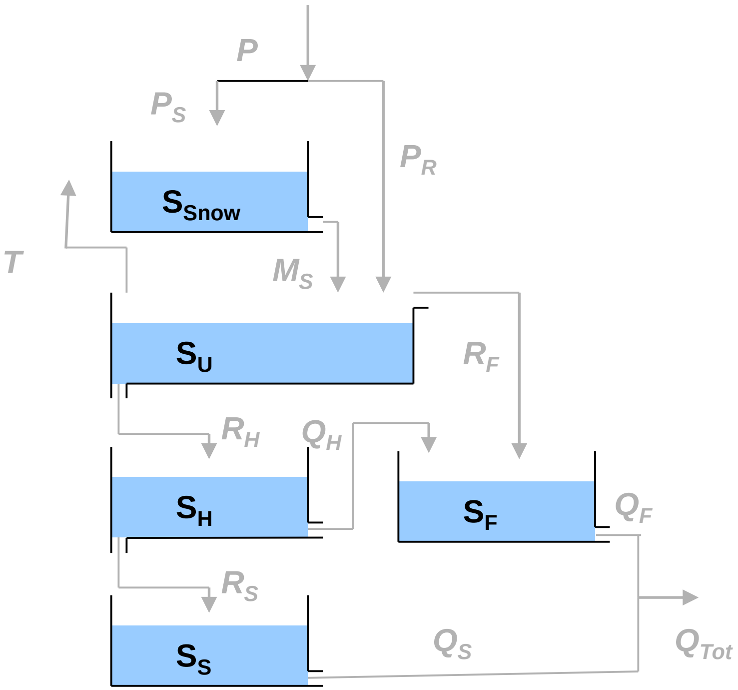
Model structure



Model 1

5 Parameters,

3 states
(Unsaturated, fast, slow responding storage)



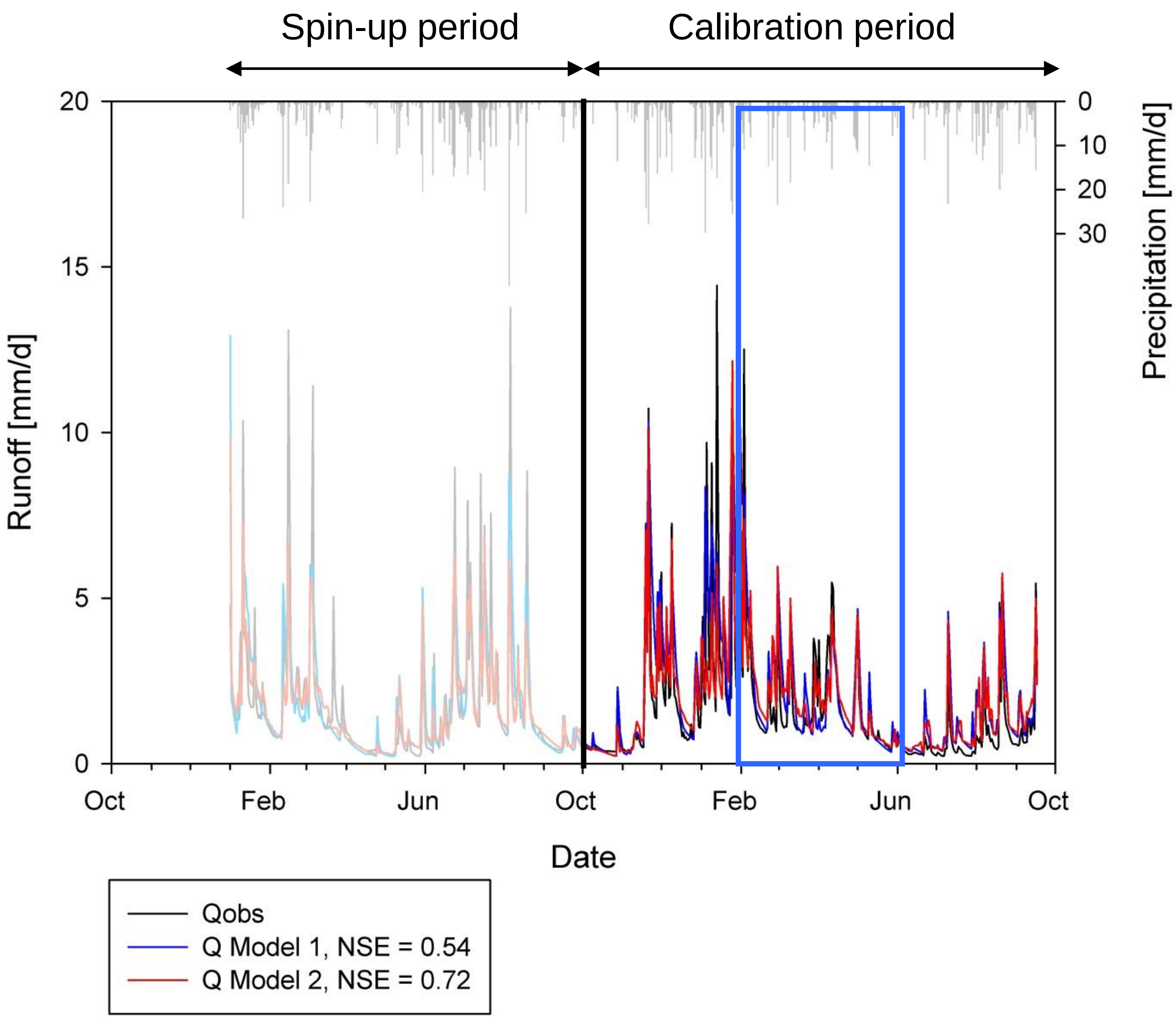
Model 2

12 Parameters,

5 states
(Snow, unsaturated, hillslope, wetland and slow responding storage)

P = precipitation, P_S = snow, P_R = rain, M_S = snow melt, T = transpiration, R_F = recharge of fast/wetland storage, R_S = recharge of slow storage, R_H = recharge of hillslope storage, Q_F = flow from fast/wetland storage, Q_S = Flow from slow storage, Q_H = flow from hillslope storage, Q_{Tot} = total runoff; S_{Snow} = snow storage, S_U = unsaturated storage, S_F = fast/wetland storage, S_S = Slow storage, S_H = hillslope storage

Model structure



Model selection

<u>Which model is now the "better" model??</u>

<u>Bayesian Information Criterion BIC</u>

- Used for selecting the most adequate model among a finite set of models.

- It is a relative metric → Does not give information on the absolute performance of the model

- Most adequate model is the one with the lowest BIC

ˈmance (NSE = 0.54) but might
; snow storage and melt
(!) parameters more.

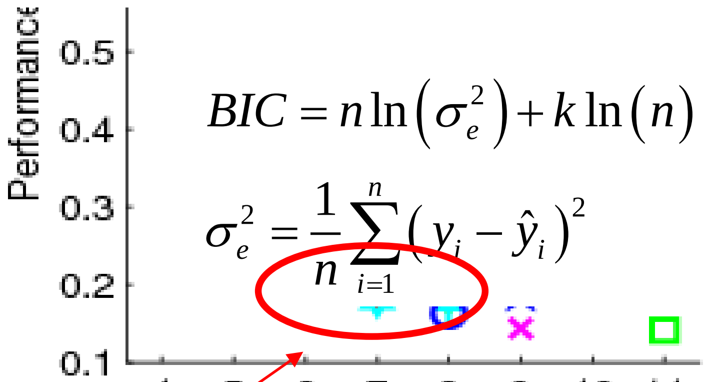
because more
parameters = I

There is no rea ... ıelp in the decision:

(1) Check th ... ʼhen certain components are activ ... for the hydrograph in this periods c ... e significant

(2) Do additi ... arameter identifiability? If yes, they ... m climates)

(3) If testing ... ˈameters against performance to decide for best structure



$$BIC = n\ln\left(\sigma_e^2\right) + k\ln\left(n\right)$$

$$\sigma_e^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2$$

Where n = sample size, k = number of parameters, σ_e^2 = the error variance, y_i = observed data, $\hat{y}_i$ = modeled output

Models in this region are likely to be the most

As a rule of thumb: if $BIC_i - BIC_{min} > 5$ there is a difference in the performance of the models

(5) Independently test the model (e.g. split sample test → validation)

Numerical implementation

Hydrological models are sets of coupled differential equations of the form

$$\frac{dS(t)}{dt} = f_S\left(S(t), F(t) \middle| \theta\right)$$
$$Q(t) = f_Q\left(S(t), F(t) \middle| \theta\right)$$

where S = states, F = forcings, θ = parameter set, Q = output

Consider a simple **linear reservoir**:

In the special case of F = 0 (i.e. no precipitation, evaporation or other gains or losses) the system can be solved analytically

$$\frac{dS}{dt} = -Q$$
$$Q = \frac{S}{k}$$

where k = timescale

Rearranging:

$$\frac{dS}{dt} = -\frac{S}{k}$$
$$\frac{dS}{S} = -\frac{1}{k}dt$$
$$\ln(S) = -\frac{1}{k}t + C$$
$$S = e^{-\frac{t}{k}+C}$$
$$S = e^{-\frac{t}{k}}e^{C} \quad \text{with } e^C = S_0$$
$$S(t) = S_0 e^{-\frac{t}{k}}$$

Rearranging and substituting:

$$S = Qk$$
$$Qk = S_0 e^{-\frac{t}{k}}$$
$$Q = \frac{S_0}{k} e^{-\frac{t}{k}}$$
$$Q(t) = Q_0 e^{-\frac{t}{k}}$$

Numerical implementation

For the special case F = 0, there exists also an analytical solution for **<u>nonlinear reservoirs</u>** of the form

$$\frac{dS}{dt} = -Q \qquad Q = \frac{S^{\alpha}}{k} \longrightarrow S(t) = \left(-kt + kt\alpha + S_0^{\,1-\alpha}\right)^{-(\alpha-1)^{-1}}$$

where S = states, F = forcings, Q = output, k = timescale and α = non-linearity parameter

However, generally F ≠ 0, for which no analytical solution exists so far

$$\frac{dS}{dt} = P - E - Q \qquad Q = \frac{S^{\alpha}}{k} \longrightarrow$$

Numerical solutions must be applied:

- explicit methods
- implicit methods

Explicit methods

Explicit methods compute the state of the system at a later time based on the state of the system at the present:

$$S(t+\Delta t) = F(S(t))$$

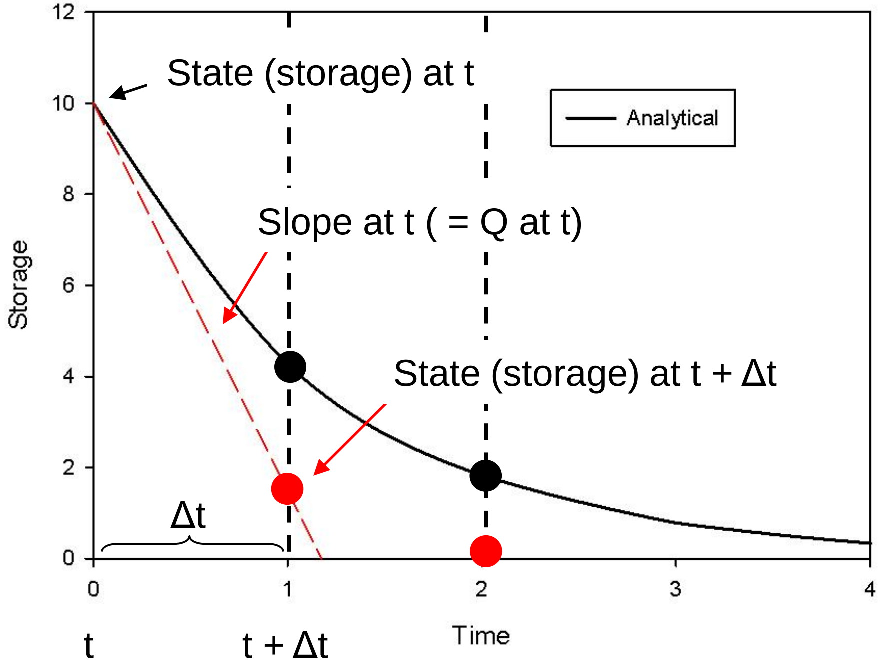
The simplest explicit method is the **forward Euler method**

$$\frac{dS}{dt} = -\frac{S(t)^{\alpha}}{k}$$

$$\frac{S(t+\Delta t)-S(t)}{\Delta t} = -\frac{S(t)^{\alpha}}{k}$$

$$S(t+\Delta t) = S(t) - \frac{S(t)^{\alpha}}{k}\Delta t$$

where S = states, t = time, k = timescale and α = non-linearity parameter



+ can be computed with one iteration

- problematic for large time steps and/or steep gradients

Explicit methods

Better results can generally be obtained by using sub-stepping methods, such as the **classic Runge-Kutta** scheme or **RK4**.

RK4 uses a weighed average of gradients to estimate the state at t + Δt instead of only the gradient at t as the foreward Euler method does.

$$\frac{dS}{dt} = -\frac{S(t)^{\alpha}}{k}$$

$$S(t+\Delta t) = S(t) - \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

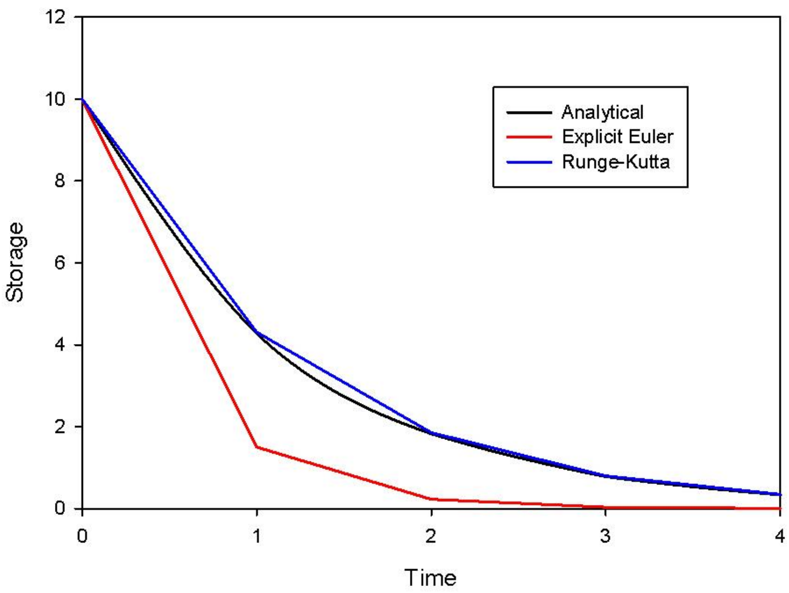
where S = states, t = time, k = timescale and α = non-linearity parameter and

$$k_1 = -\frac{S(t)^{\alpha}}{k}$$

$$k_2 = S(t) - k_1\frac{\Delta t}{2}$$

$$k_3 = S(t) - k_2\frac{\Delta t}{2}$$

$$k_4 = S(t) - k_3\Delta t$$



Implicit methods

Implicit methods compute the state of the system at a later time based on the state of the system at the present AND the state of the system at a later time:

$$G\left(S(t), S(t+\Delta t)\right) = 0$$

The simplest implicit method is the **backward Euler method**

$$\frac{dS}{dt} = -\frac{S(t)^{\alpha}}{k}$$

$$\frac{S(t+\Delta t) - S(t)}{\Delta t} = -\frac{S(t+\Delta t)^{\alpha}}{k}$$

$$S(t+\Delta t)^{\alpha}\frac{\Delta t}{k} + S(t+\Delta t) - S(t) = 0$$

where S = states, t = time, k = timescale
and α = non-linearity parameter

- cannot be computed with one iteration and needs to be solved using root finding algorithms (e.g. Newton method)

+ better approximations

+ more stable

Reduce parameter uncertainty

The choice of unsuitable parameters leads to prediction errors → low model realism

<u>Parameter uncertainty is largely caused by:</u>

(1) Imperfect model structures → parameters compensate for errors

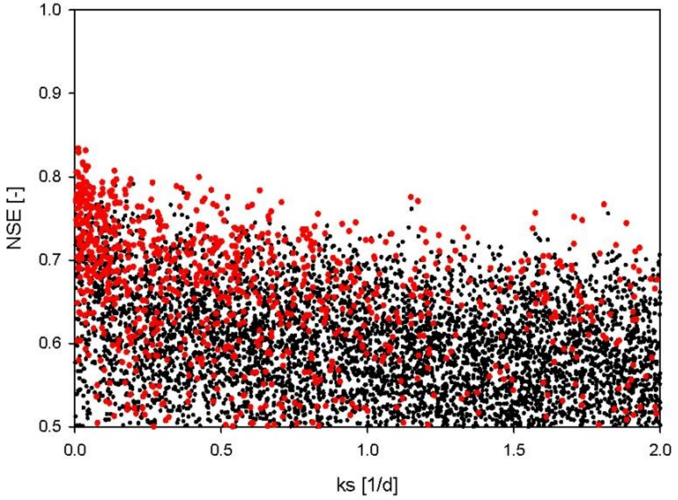
(2) Overfitting of the model → to few constraints, i.e. to many degrees of freedom

(3) Temporally changing parameter values, e.g. extent of saturated area

<u>Way forward by increasing number of constraints:</u>

(1) Multi-objective calibration → more than one objective function

(2) Multi-criteria calibration → more than one variable to be calibrated (e.g. stream flow and tracer concentration)



Reduce parameter uncertainty

Example of the effects of combined multi-objective and multi-criteria calibration:

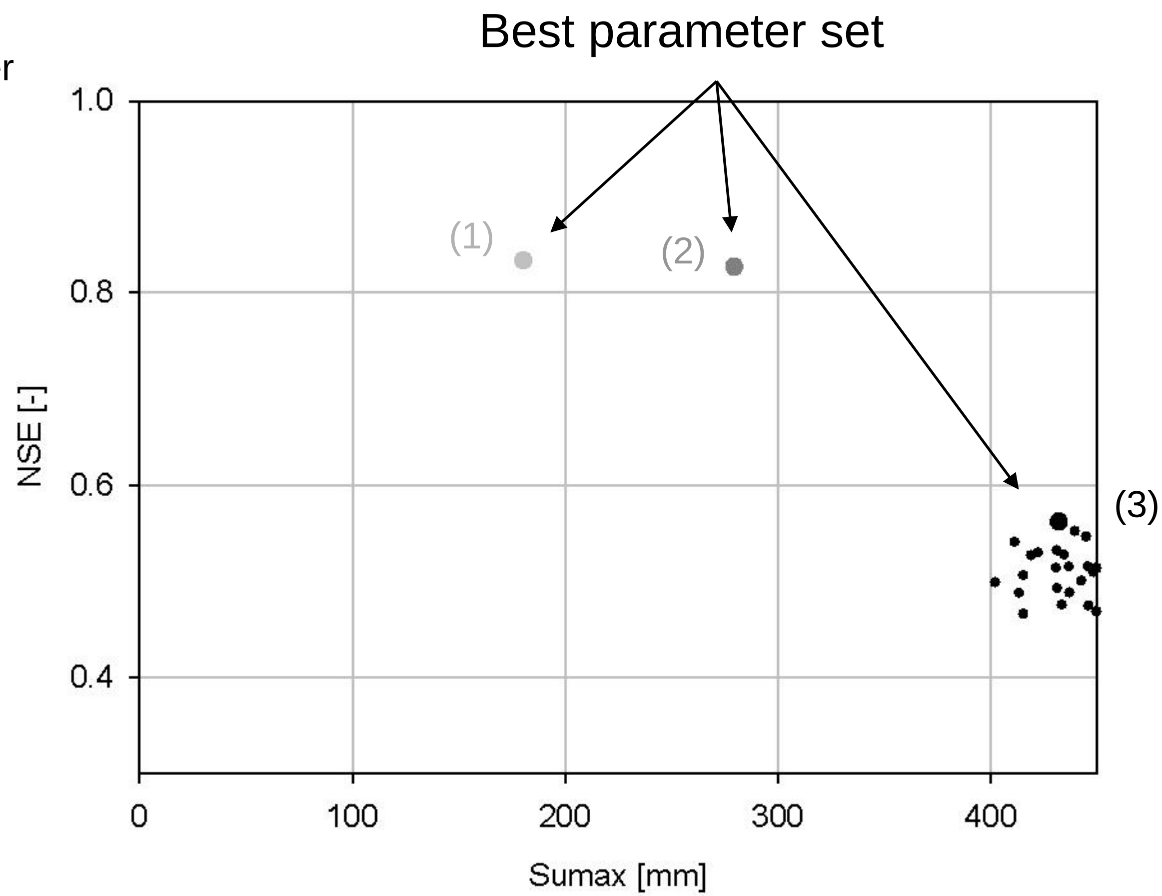
Model calibrated against (1) stream flow using NSE, (2) NSE of log flows and (3) NSE of tracer concentration in stream

Feasible parameter sets reduced from

(1) 5077

To (2) 3857

and (3) 24 (!!!)



Reduce parameter uncertainty

Models should in general aim to predict stream flow with adequate accuracy.

High predictive power is strongly related to how well the model represents reality → model realism

Best parameter set in calibration period not necessarily the most realistic parameter set → performance of parameters in independent test periods ("validation periods") holds crucial information on model realism

Several strategies to include information from validation periods, e.g.:

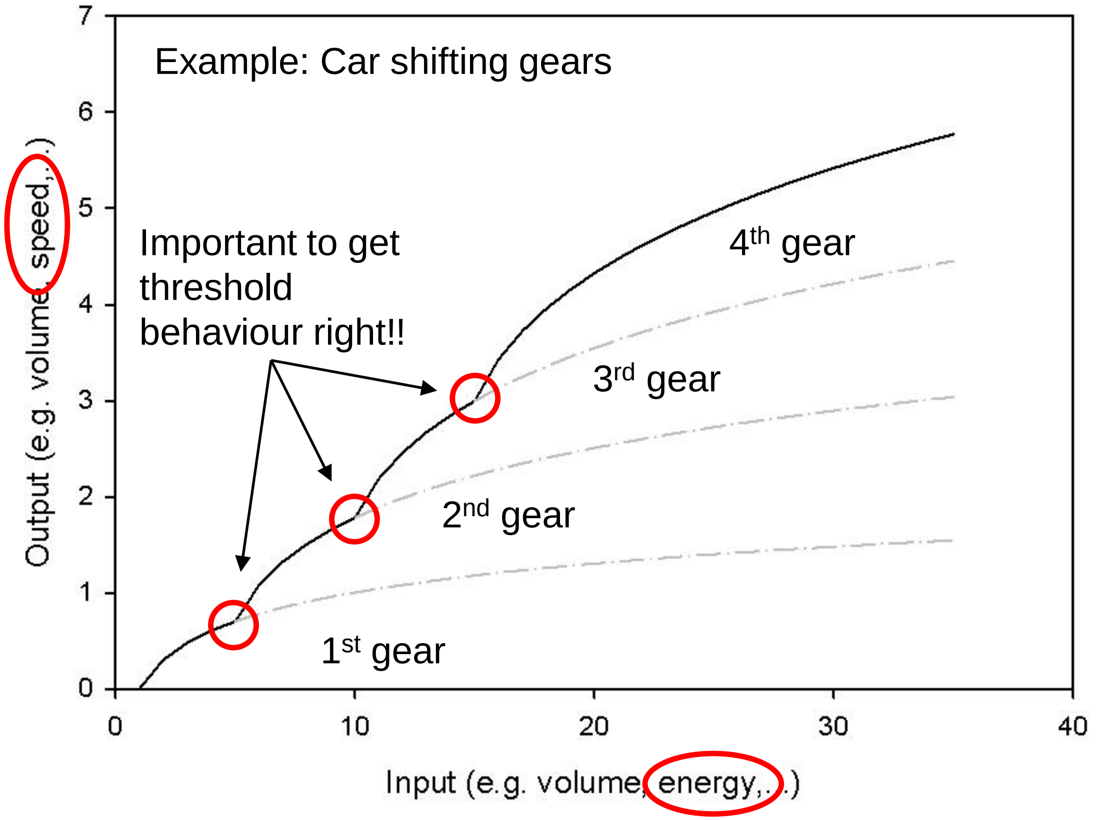
(1) Split sample test: train model using the first part of the available data (e.g. 50%) and then test the model with the remaining part. Frequently applied, very simple, however only limited additional information

(2) Ordinary differential split sample test: calibrate the model on a dry period and validate it for a wet period, if the model is used to predict wet periods and the other way round if it is to be used for dry period.

(3) Extended split and differential split sample tests: sequentially train the model on more than one calibration period and test it for the remaining period (e.g. calibrate year 1 - validate year two, calibrate year 2, validate year 1). Use the parameter sets that perform best during ALL calibration and validation periods.

Feedback and thresholds

Modelling complex feedback systems can frequently result in surprising results as they are characterized by thresholds. The system response can thus tip from one regime to another with only VERY SMALL changes in the forcings and states (e.g. like horses switch gaits, cars switch gears or distinct storages in a hydrological system are activated and deactivated)



Feedback and thresholds

Very simple example to illustrate how highly sensitive to changes in the initial system state a feedback system can be:

Verhulst population dynamics model

$$p_n = p_{n-1} + r p_{n-1}\left(1 - p_{n-1}\right)$$

Where p_n = size of population as proportion of the maximal population supported by the system and r = growth rate

<u>Case 1:</u>

$p_0 = 0.05$, $r = 3$

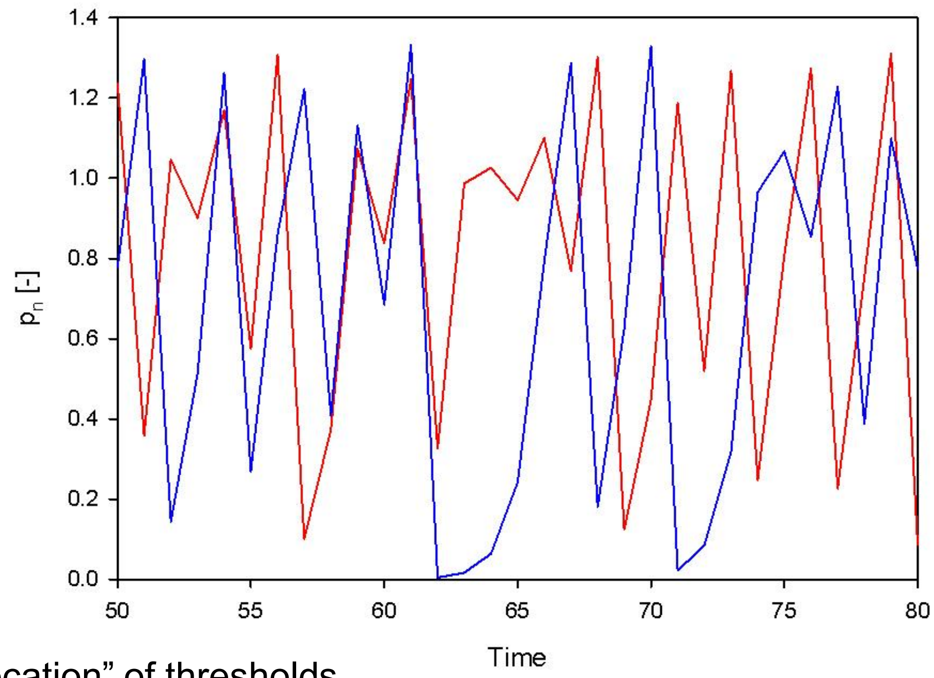
<u>Case 2:</u>

$p_0 = 0.05 * 0.999999$, $r = 3$

<u>Case 3:</u>

$p_0 = 0.05$, $r = 3$, but:

$$p_n = \left(1 + r\right) p_{n-1} - r {p_{n-1}}^2$$



→ Input, state of the system, “location” of thresholds, computational precision, etc. can strongly affect results!!!!

Programming essentials

(1) Prepare a concept and follow it

(2) Insert as many comment lines as possible, describing in detail what you are doing and why

(3) Put date and author in comment lines

(4) Do not include system dependent items in the main code, but isolate these to make changes easier

(5) Keep old lines as comment lines if you make changes to the code describe why you changed it

(6) Test the code extensively, if possible let also others test it – “There is no code without a bug!”

(7) Get your units right

(8) Be accurate

(9) Avoid rounding and truncation of numbers

(10) Be careful if the results look “too” good! The better the results seem, the harder bugs are detected

The value of more efficiently extracting information from data

OR

Why bother to calibrate??

Simplicity – Complexity - Information

Where do we stand?

- There is need to balance process heterogeneity and model parsimony
- **Objective:** more realistic representations of the system
  -> Reduced uncertainty -> Increased predictive power
- Model *performance* vs. model *consistency*
- BUT: in most cases we do not have enough information to improve models!

STOP!!

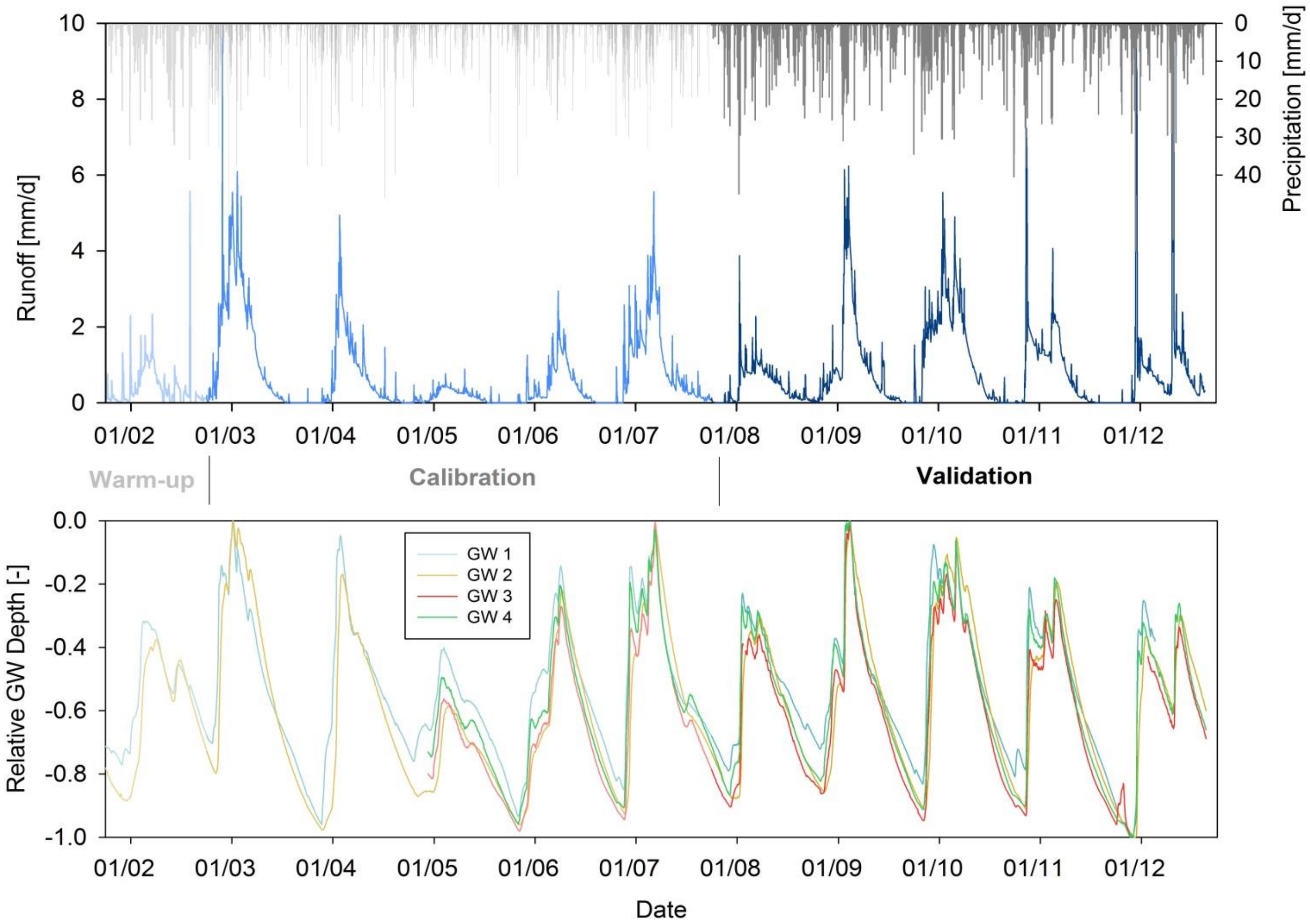
Kerbernez Research Catchment

Small catchment in French Brittany

- Catchment area ~11ha
- Average precipitation 1100 mm/yr
- Average Penman Evaporation 700 mm/yr
- Deep, relatively free draining soils
- Underlain by weathered Granite
- Mainly grass- and cropland
- Agriculturally used
- First order, intermittent stream
- Long-term data record
- 5 Piezometers along 2 transects

Kerbernez Research Catchment

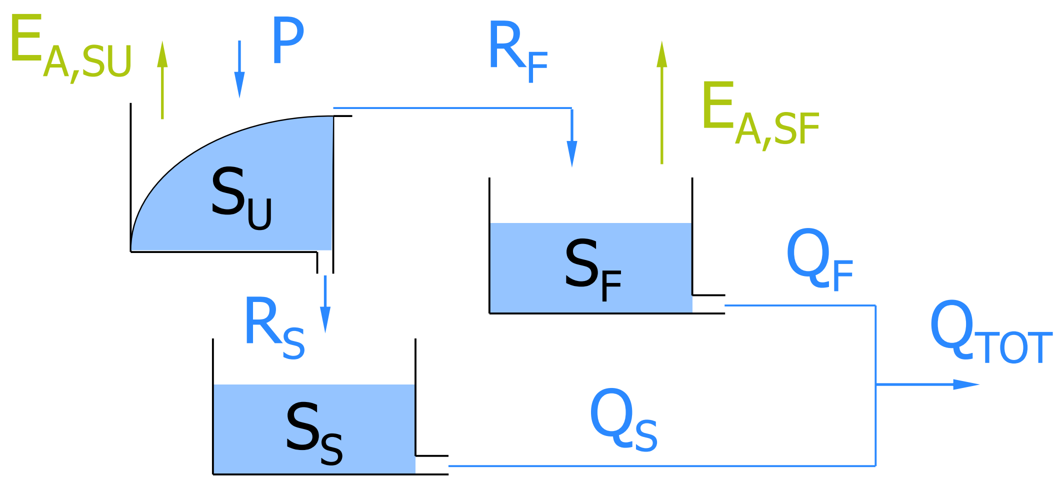
Small catchment in French Brittany



M1- Benchmark Model

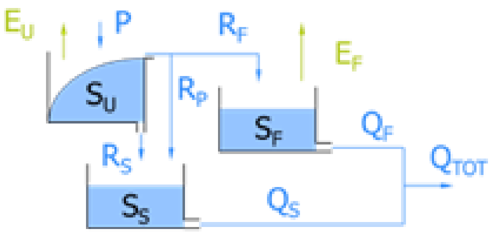
Model 1

- Ordinary lumped, 3 box HBV-type model
- 6 free calibration parameters
- No constraints
- Multi-objective calibrations:
  NSE Q, NSE log(Q), VE Q and NSE FDC



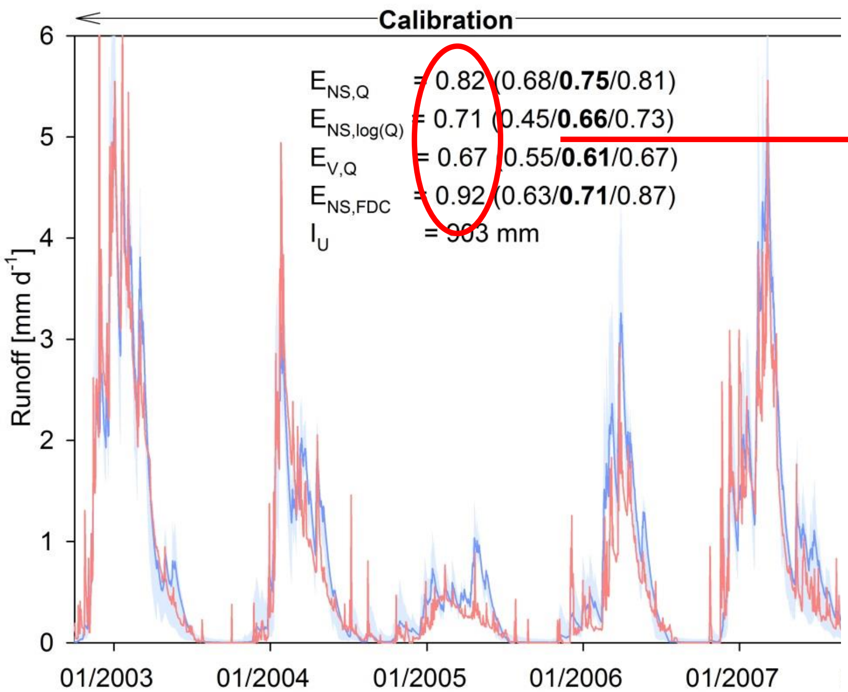
m.hrachowitz@tudelft.nl

M1- Benchmark Model



Model 1: 6 parameters, no constraints

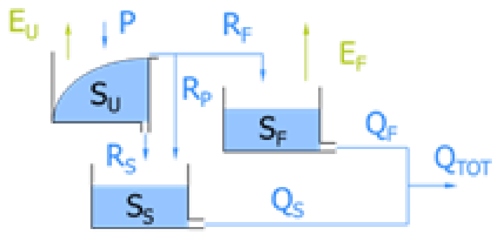
- Good calibration performance
- Well constrained by 4 objective functions
- ~~Yippie – We have a working model!!!~~
- **<u>But:</u> performance in validation decreased**

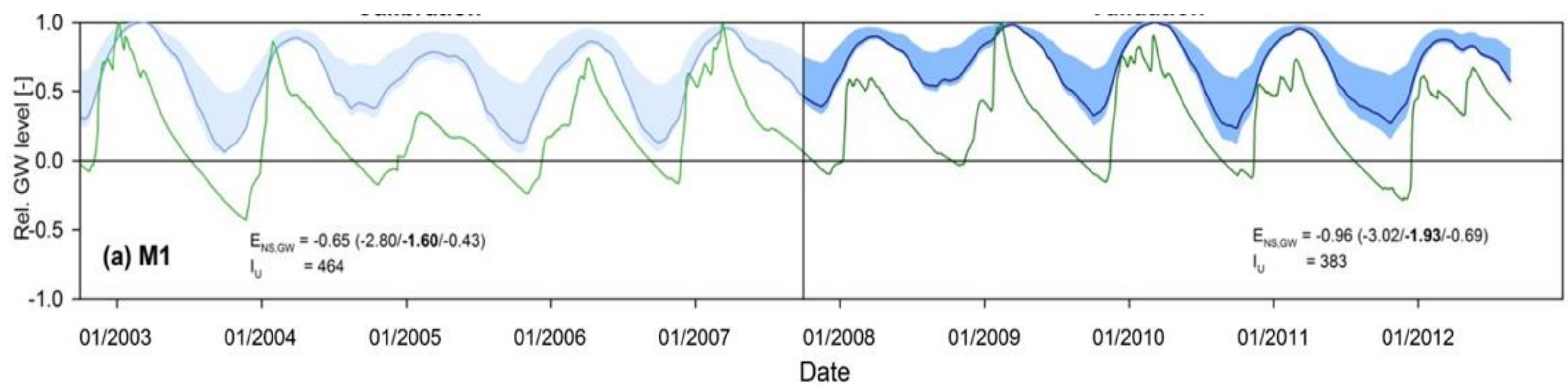


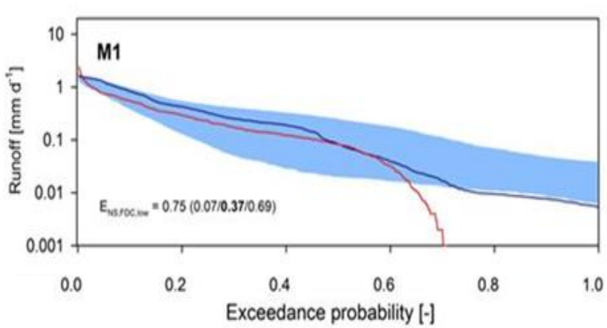
Hrachowitz et al., 2014, WRR

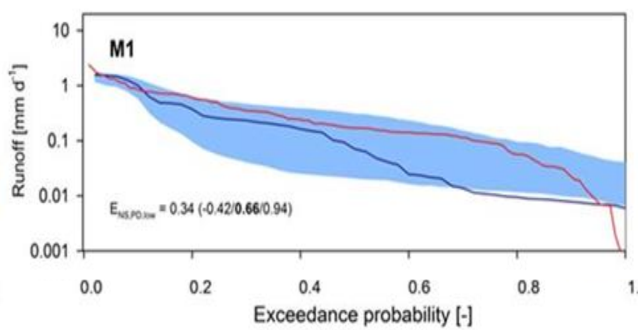
M1- Benchmark Model

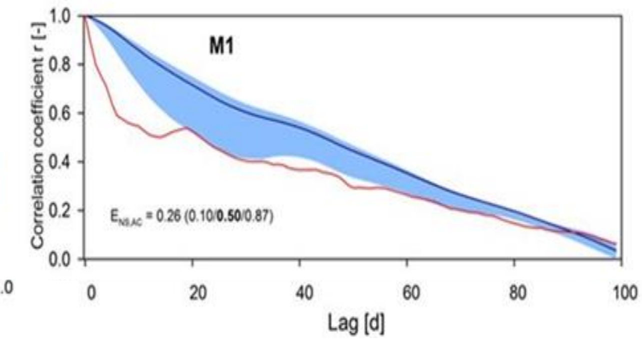
Model 1: 6 parameters, no constraints







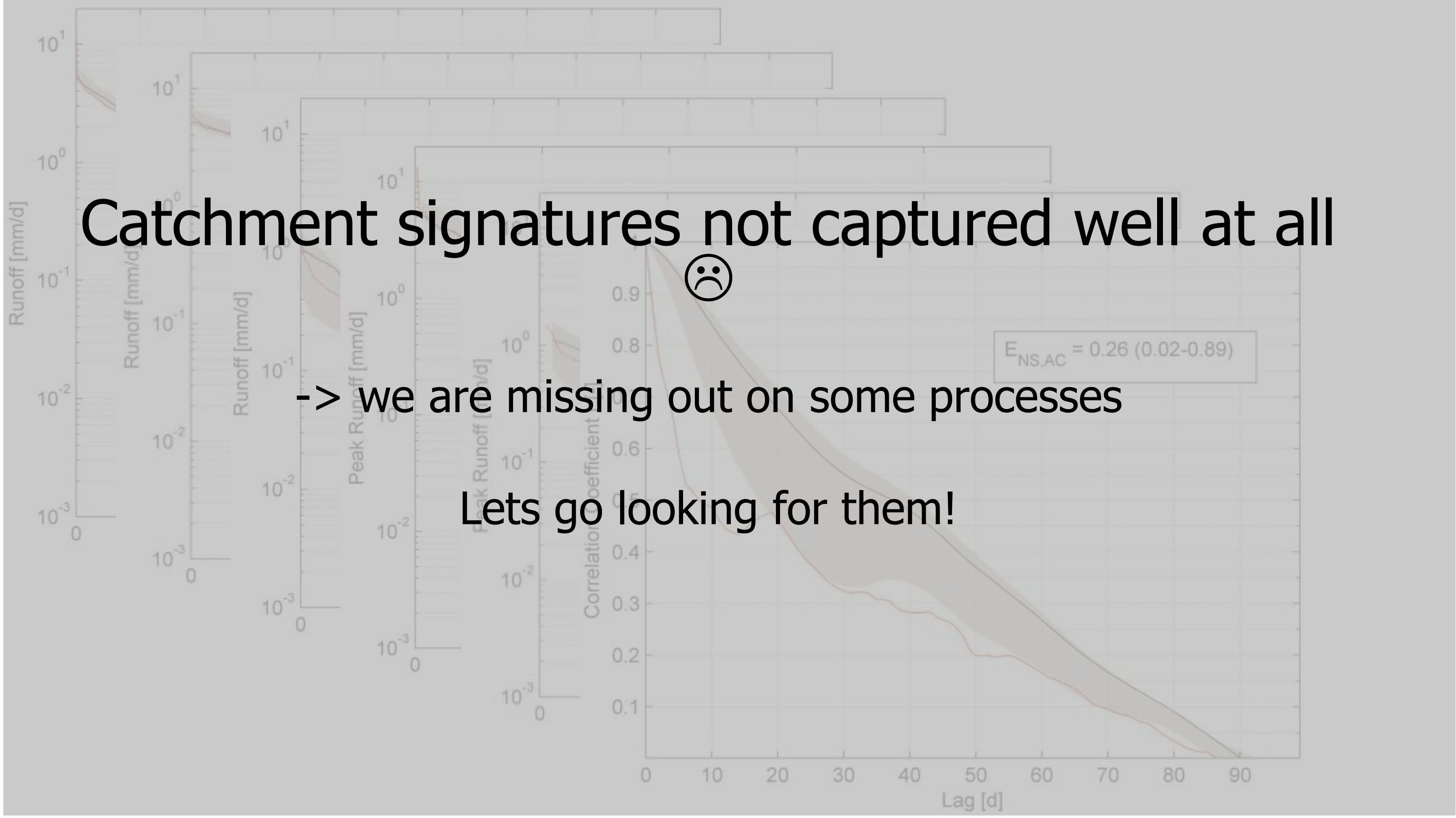




Hrachowitz et al., 2014, WRR

M1- Benchmark Model

Model 1: 6 Parameters, no constraints



Catchment signatures not captured well at all ☹

-> we are missing out on some processes

Lets go looking for them!

m.hrachowitz@tudelft.nl

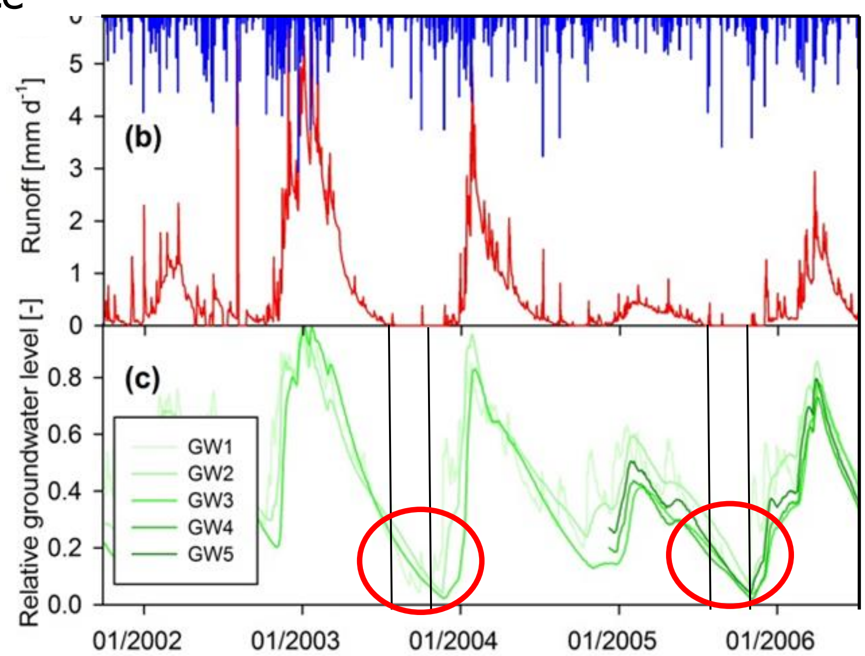
Additional processes and landscape heterogeneity

What do we know about the catchment?

(a) Long-term water balance deficit

- Bit of a no-brainer: check water balance
  -> find substantial losses
- Add loss term to model

P	1035 mm
- Q	- 274 mm
- E_P	- 692 mm
Loss	68 mm



Hrachowitz et al., 2014, WRR

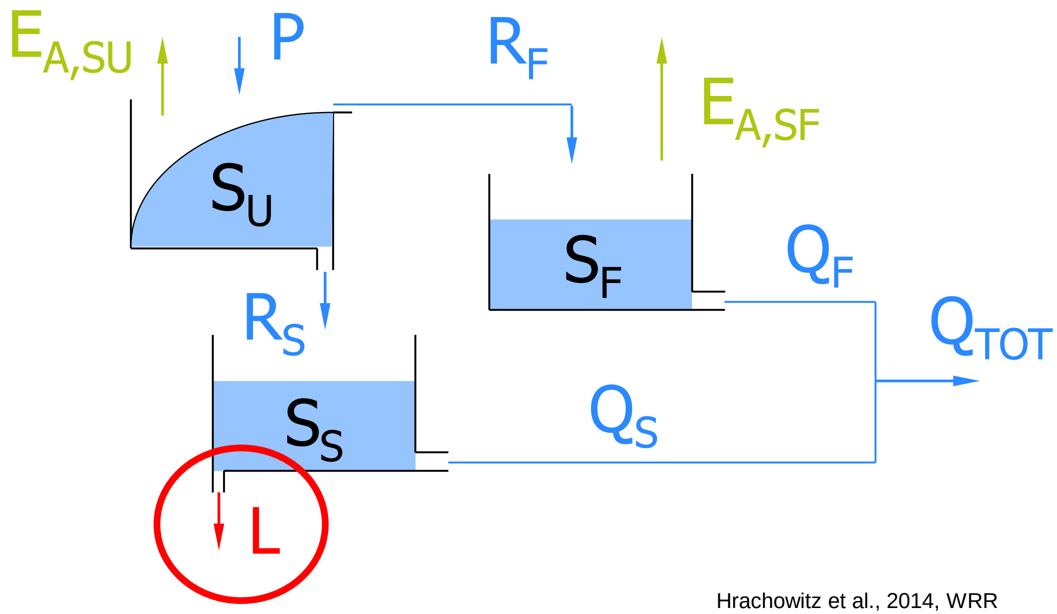
Additional processes and landscape heterogeneity

What do we know about the catchment?

(a) Long-term water balance deficit

- Bit of a no-brainer: check water balance
  -> find substantial losses
- Add loss term to model

P	1035 mm
- Q	- 274 mm
- E_P	- 692 mm
Loss	68 mm



Hrachowitz et al., 2014, WRR

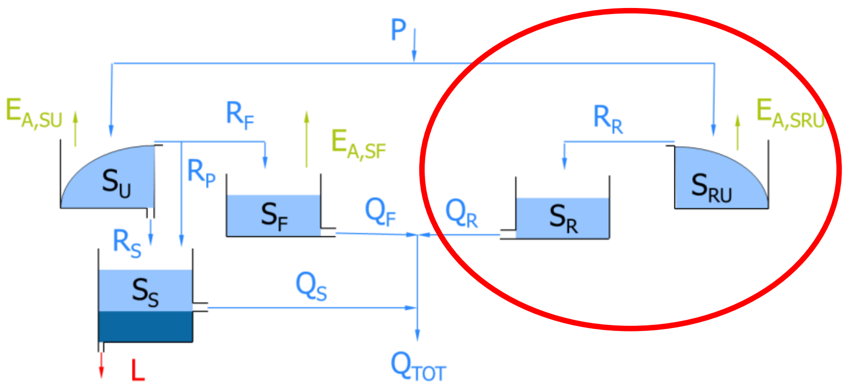
Additional processes and landscape heterogeneity

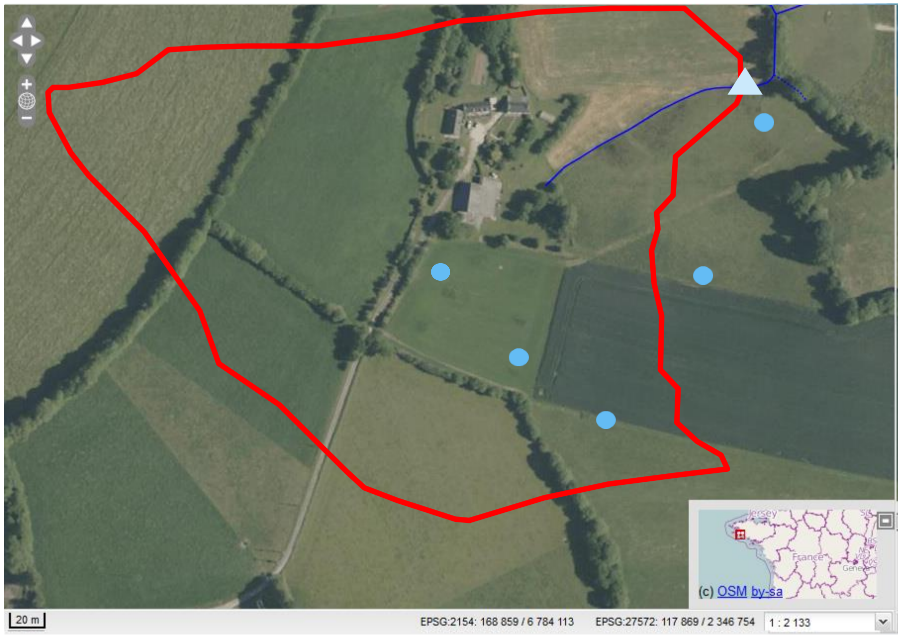
What do we know about the catchment?

(a) Long-term water balance deficit

(b) Presence of extended riparian zone / wetland

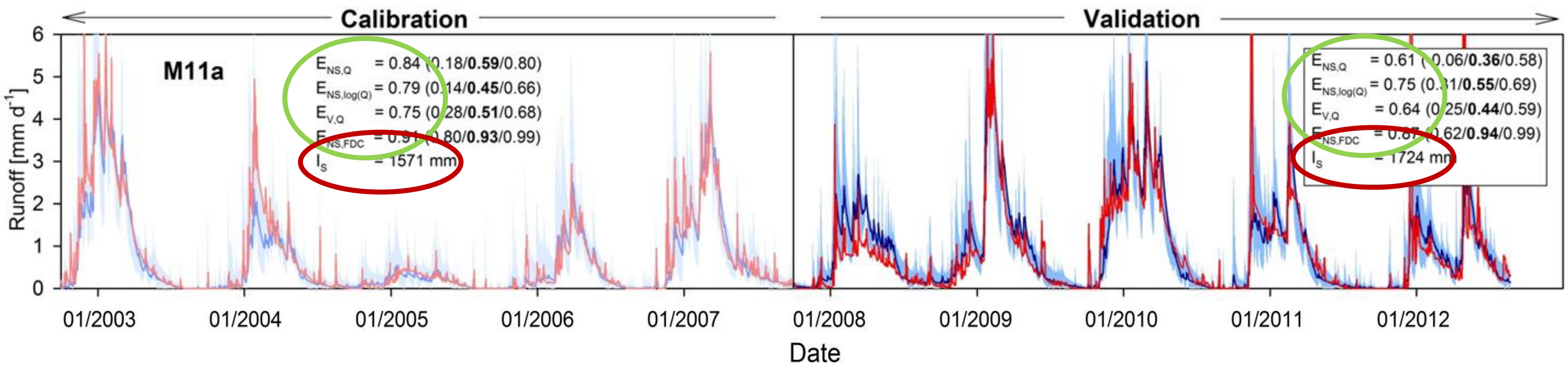
- Wetlands often show distinct process dynamics than hillslopes.
- Add wetland component to model

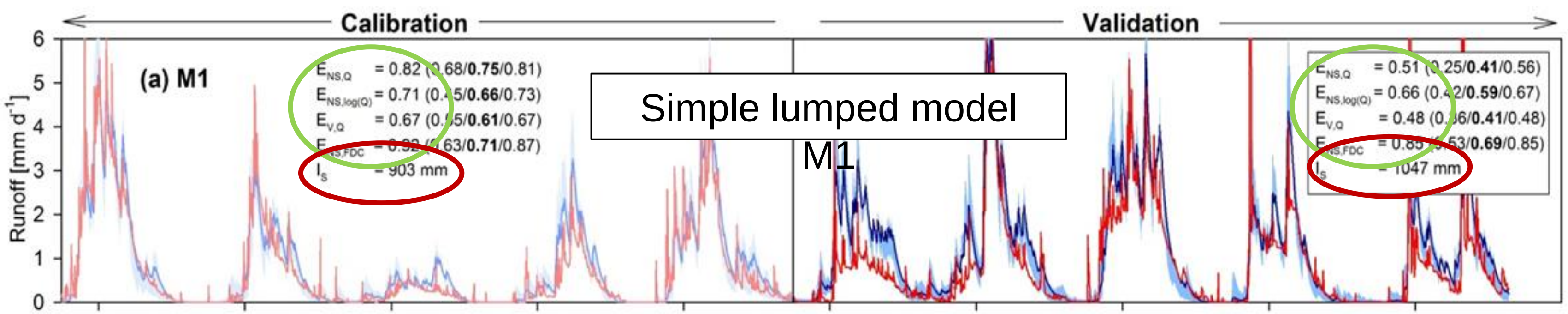




Model evaluation

M11a – Improved performance but higher uncertainty



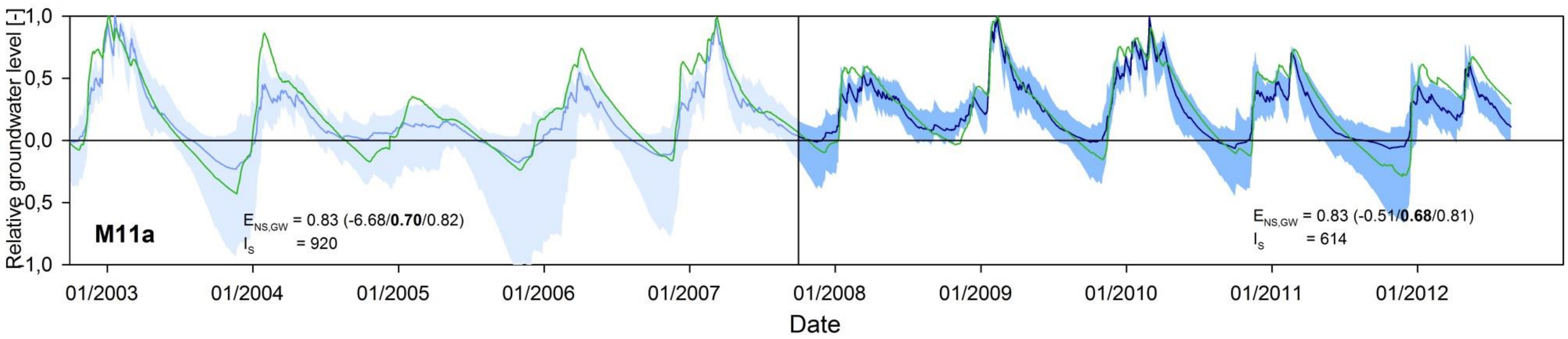


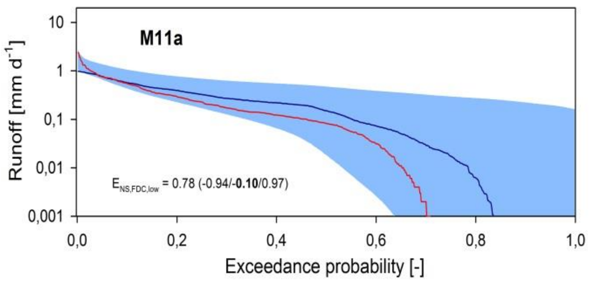
Simple lumped model M1

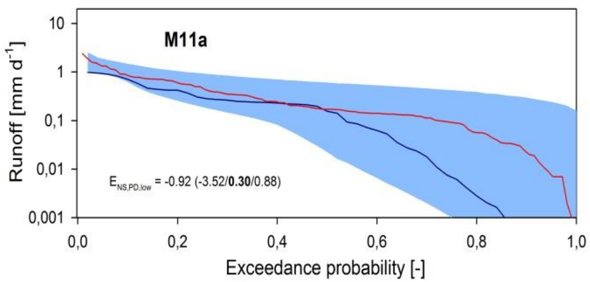
Hrachowitz et al., 2014, WRR

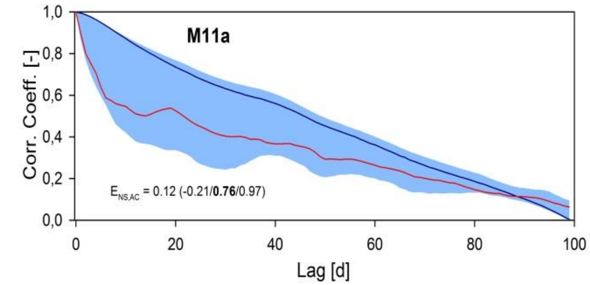
Model evaluation

M11a – Improved performance but higher uncertainty









Hrachowitz et al., 2014, WRR

Simplicity – Complexity - Information

Let us have a closer look

- Traditional steps: data -> model -> calibration
- Nothing wrong with that
- **Except:** even simple models are usually over-parameterized

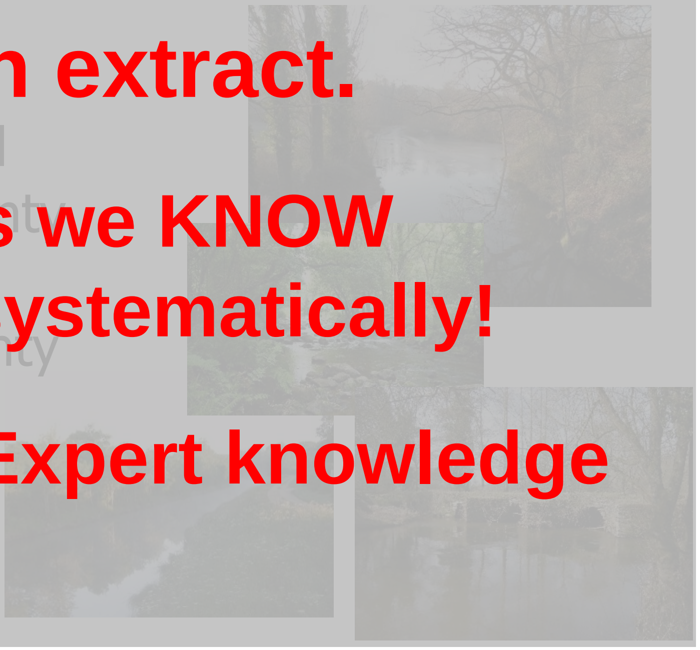
-> Too many degrees of freedom
or
not enough criteria to constrain model
-> Good performances but high uncertainty
or
bad performances and lower uncertainty

What can we do????

Simplicity – Complexity - Information

Let us have a closer look

- Traditional steps: data -> model -> calibration
  Nothing wrong with that
- Except: even simple models are usually over-parameterized
  - -> Too many degrees of freedom
    or
    not enough criteria to constrain model
  - -> Good performances but high uncertainty
    or
    bad performances and lower uncertainty



We need to be more efficient in exploiting our knowledge to maximize the information we can extract.

There are many things we KNOW
Need to use them more systematically!

Signatures – Constraints – Expert knowledge

m.hrachowitz@tudelft.nl

Simplicity – Complexity - Information

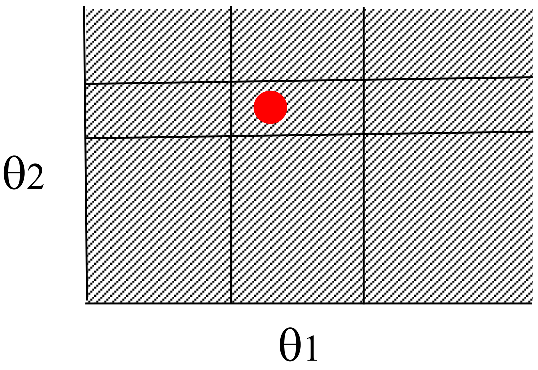
The value of prior information as model constraints

Models often ill-posed, inverse problems, i.e. no unique solution can be found from a pre-defined model space.

<u>Strategies to reduce the problem:</u>

- Reduce number of parameters
- Evaluate against multiple objectives/criteria
- Prior constraints
  a) Prior information on parameter range (e.g. K_s, S_{umax})
  b) Regularization/Regionalization

But there are more possibilities….

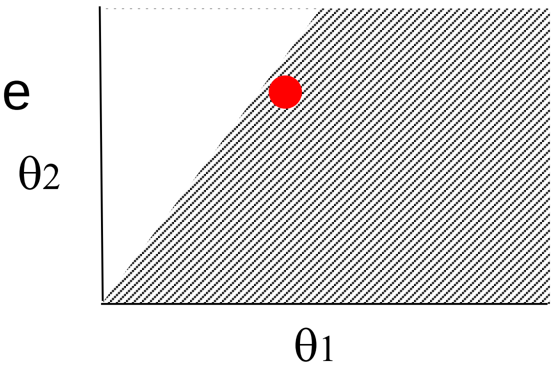


Prior Constraints

Relational **parameter constraints**

Similar to regularization, but:

- Based on semi-quantitative expert knowledge
- No additional data required
- Easy implementation
- Allows overlapping prior distributions



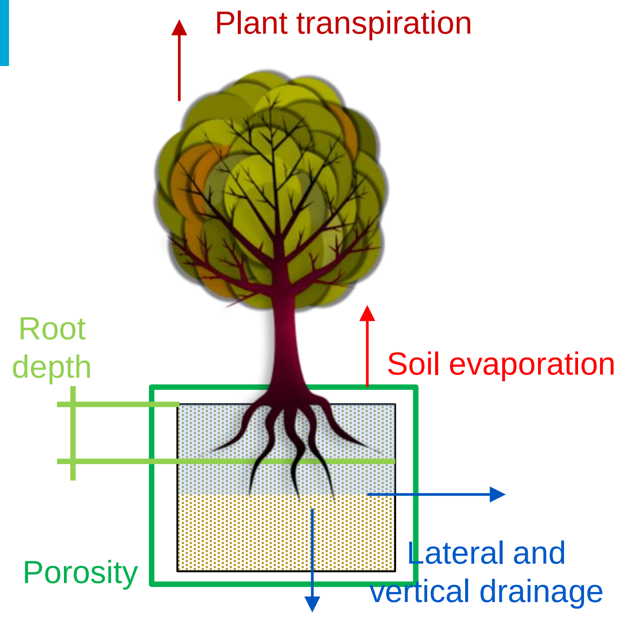
Examples:

- Storage coefficient of slow reservoir smaller than of fast reservoir: $K_s < K_f$
- Root zone storage capacity larger in hillslopes than in wetlands: $S_{umax,H} > S_{umax,W}$
- Interception capacity higher in forests than in grasslands (e.g. based on LAI differences): $I_{max,F} > I_{max,G}$

Gharari et al., 2014, HESS
Gao et al., 2014, HESS

Prior Constraints

Root zone storage capacity - How to quantify at catchment-scale?



Water holding capacity in the root zone S_R is a core parameter in any hydrological system:

(1) It provides storage below field capacity
→ principal source of non-linearity in the system

(2) It controls how water fluxes are partitioned into evaporation/transpiration and drainage

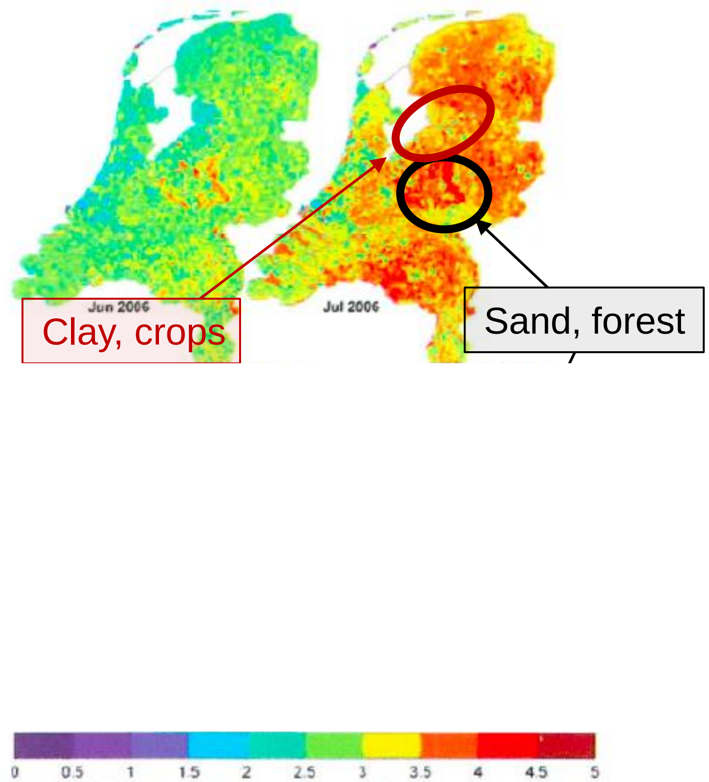
How is S_R determined?

(1) Water filled porosity at field capacity

(2) Average root depth

Prior Constraints

Root zone storage capacity - Effect of the traditional method



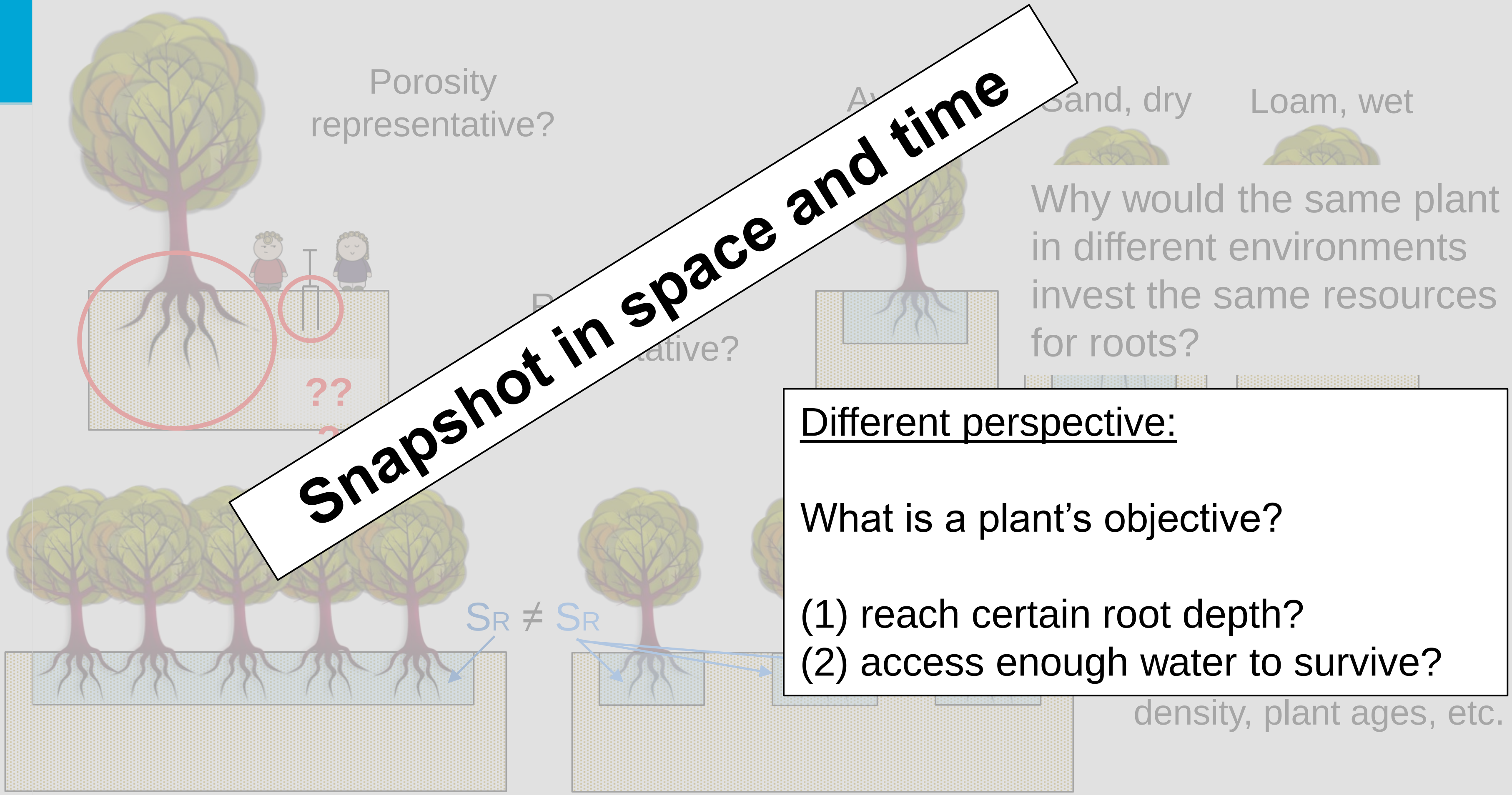
Observed (top) and modelled (below) monthly evaporation (mm/d)

Beekman et al., 2014, Stromingen

- Observed evaporation: upscaled eddy-covariance and lysimeter observations
- Modelled evaporation: physically-based, distributed NHI model
- Difference in spatial pattern
- Similar vegetation (mostly forest)
- Model only reflects differences in soils
- In sand, deeper roots are necessary to satisfy water demand during water limited conditions!

Prior Constraints

Root zone storage capacity - Effect of the traditional method



Snapshot in space and time

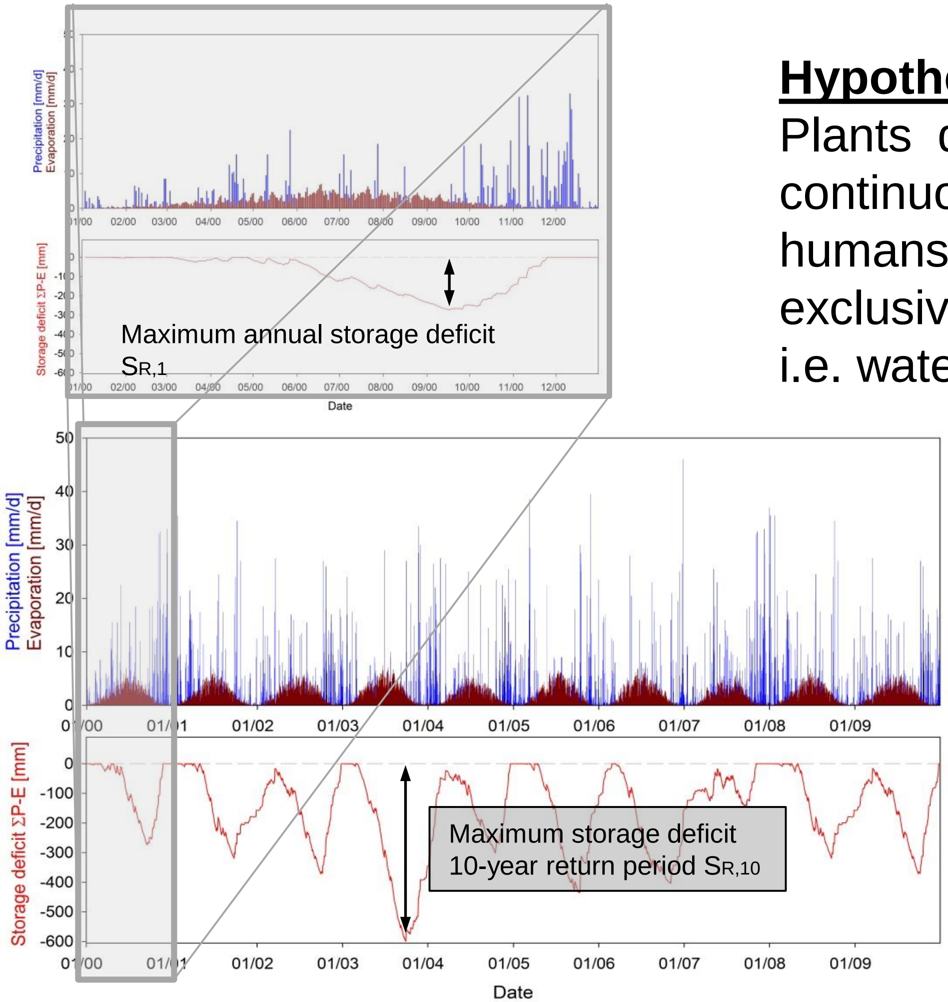
<u>Different perspective:</u>

What is a plant's objective?

(1) reach certain root depth?
(2) access enough water to survive?

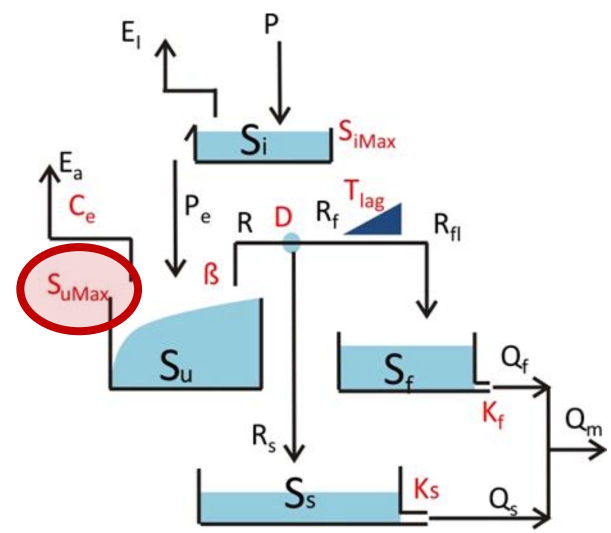
Prior Constraints

Root zone storage capacity - Estimation from water balance



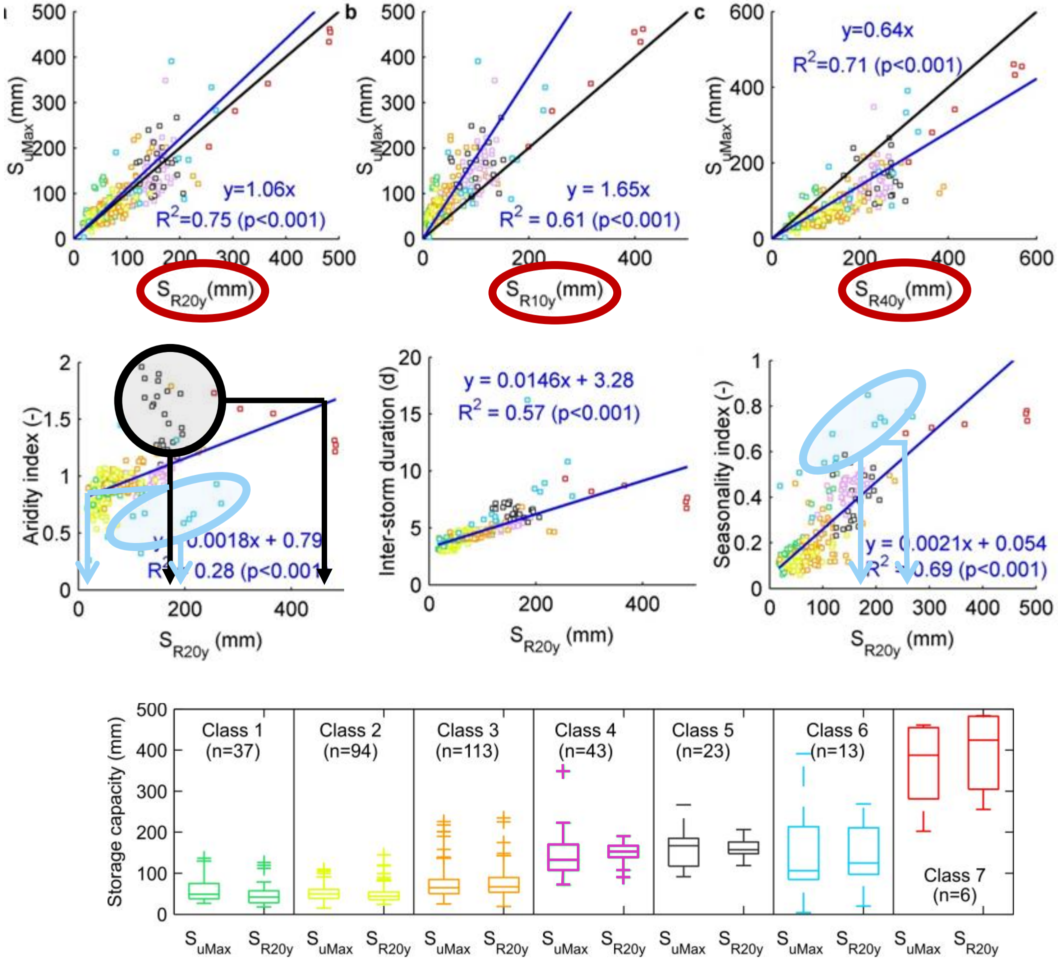
Hypothesis:

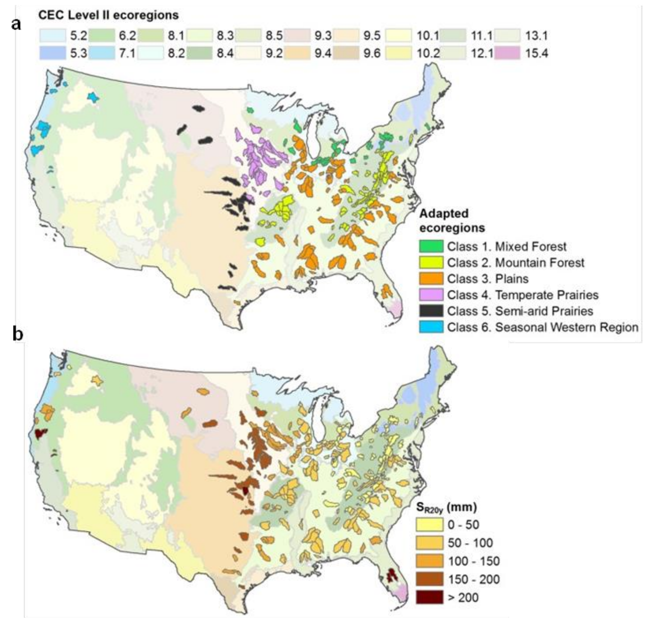
Plants design root systems to guarantee continuous access to water similar to how humans design water reservoirs, exclusively based on supply and demand, i.e. water balance.



Prior Constraints

Root zone storage capacity - What is controlling?

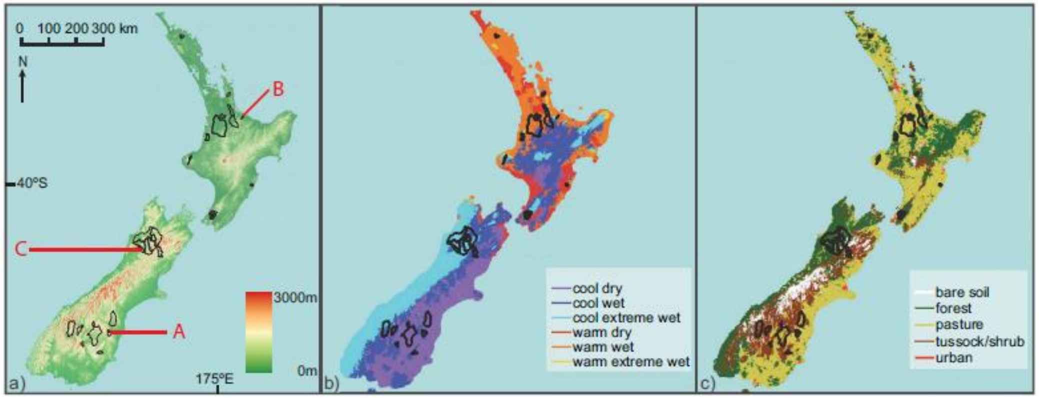


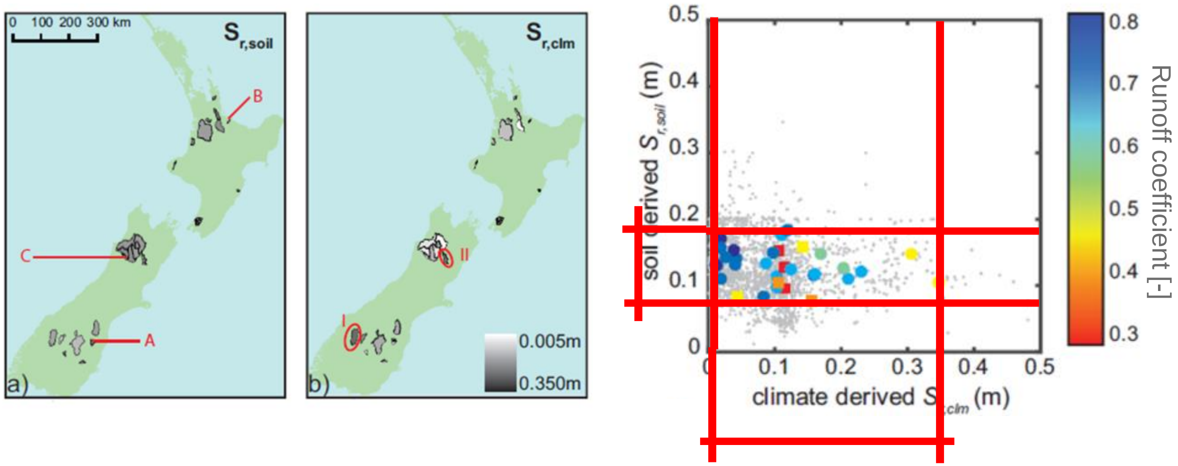


Gao et al., 2014, GRL

Prior Constraints

Root zone storage capacity - And what about soil/root data?



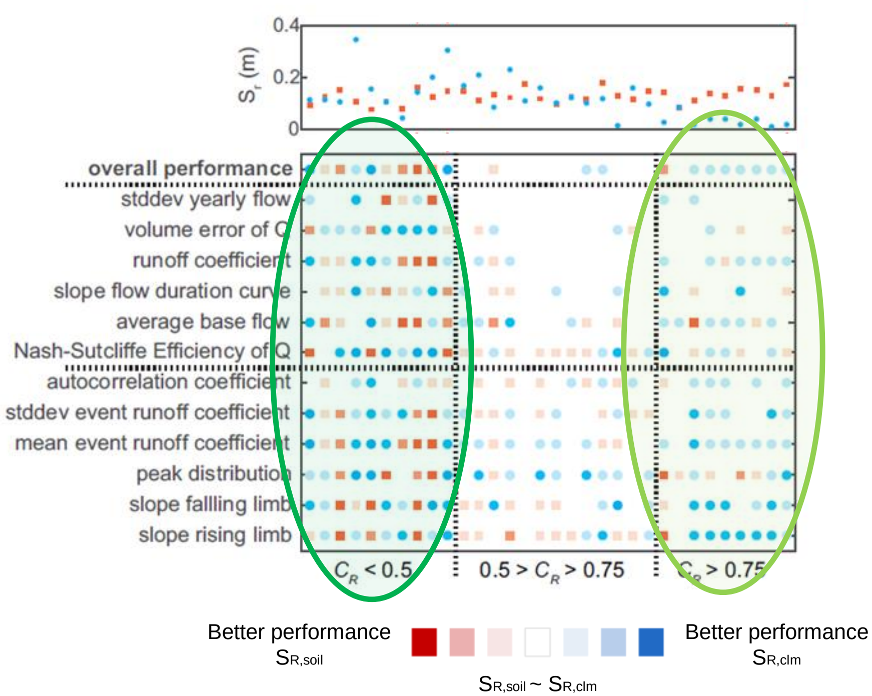


- 30 catchments in contrasting regions in New Zealand
- Higher variability in water balance derived $S_{R,clm}$ estimates than in soil derived ones $S_{R,soil}$
- The lower runoff coefficient the higher S_R

deBoer-Euser et al., 2016, WRR

Prior Constraints

Root zone storage capacity - Water balance vs. soil/root derived estimates



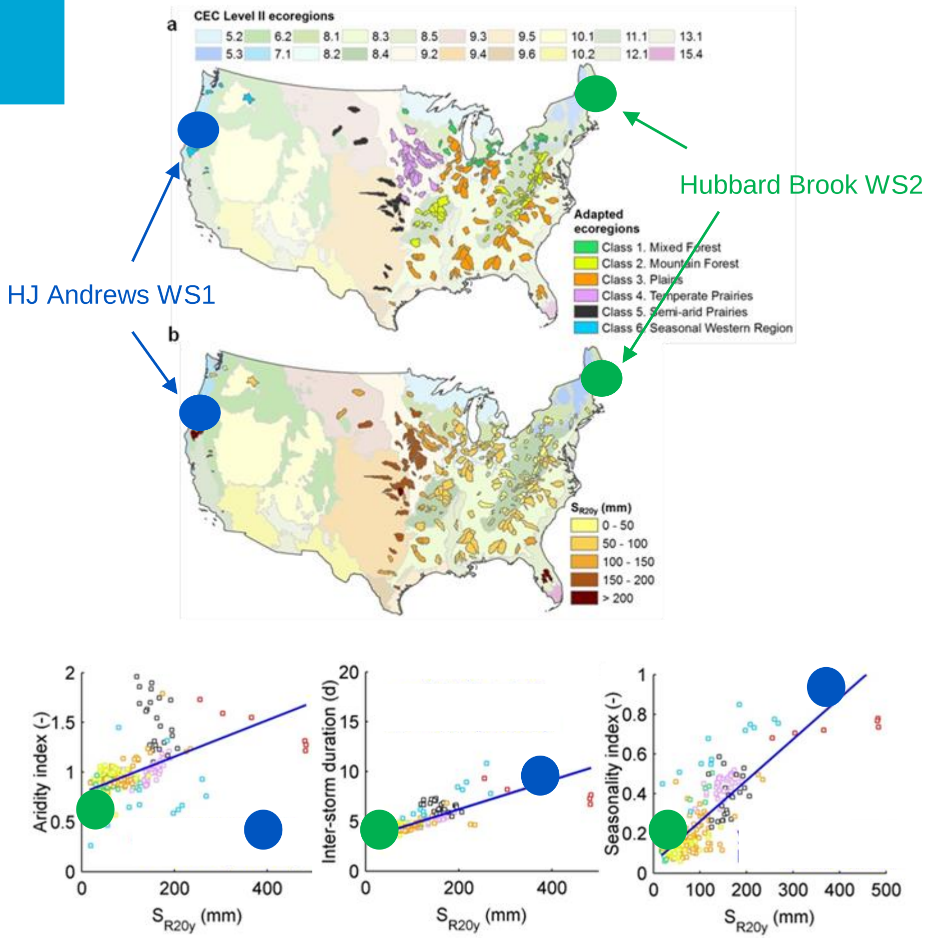
- TopNet model
- UNCALIBRATED!
- Compare model's ability to reproduce wide range of hydrological signatures with $S_{R,soil}$ and $S_{R,clm}$

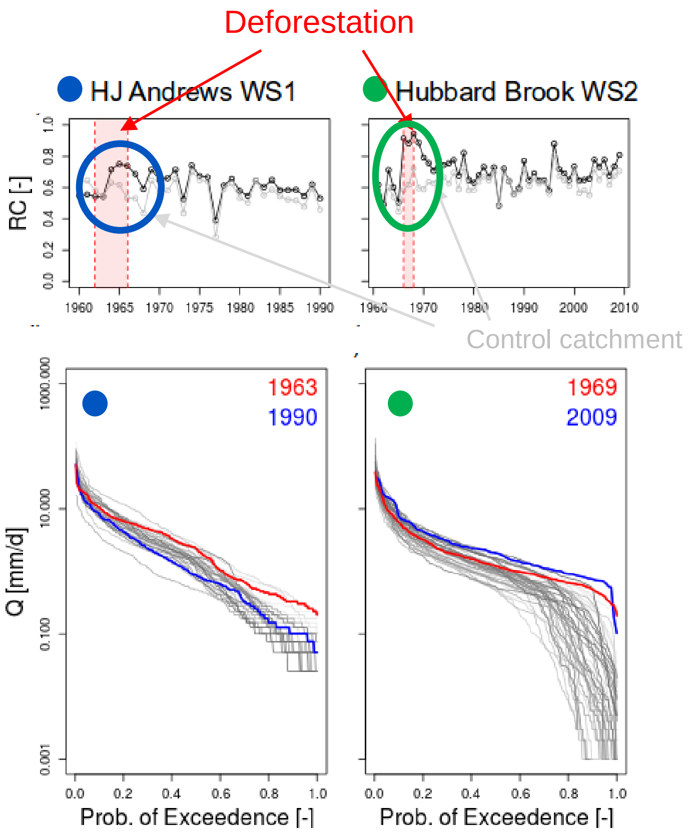
On balance, model performances with $S_{R,clm}$ at least as good as with $S_{R,soil}$, in regions with high runoff coefficients even considerably better.

deBoer-Euser et al., 2016, WRR

Prior Constraints

Root zone storage capacity – Influence of deforestation?

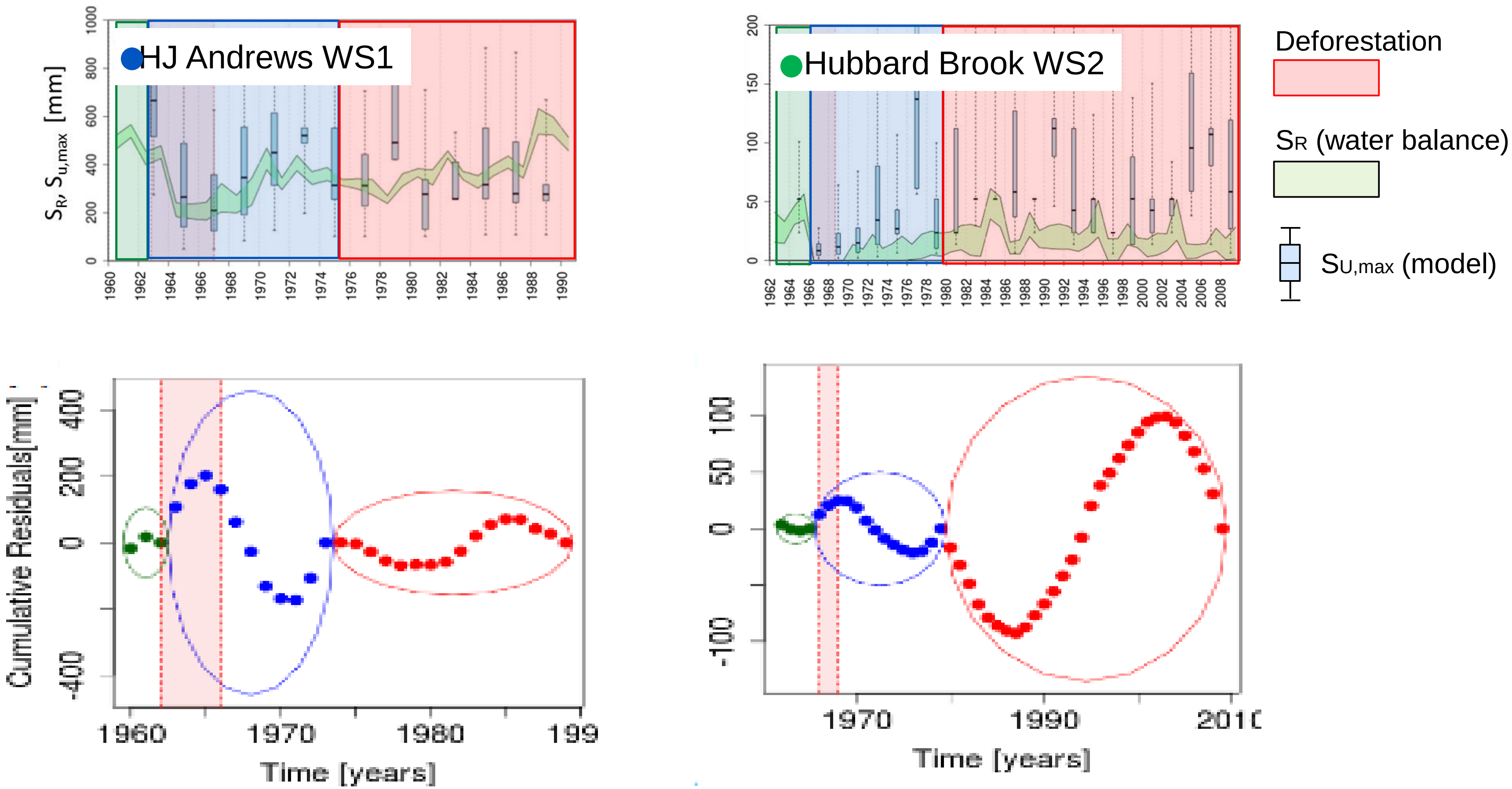




Nijzink et al., 2016, HESS

Prior Constraints

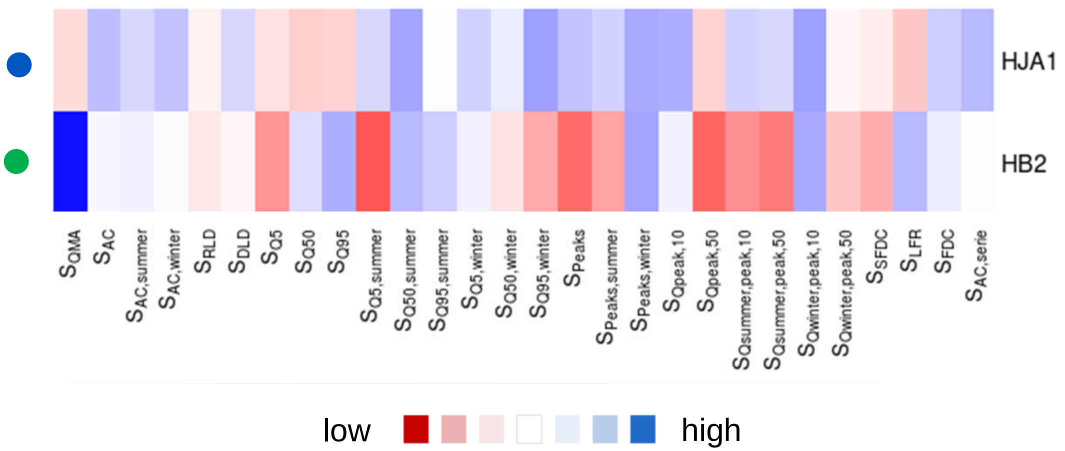
Root zone storage capacity – Influence of deforestation?



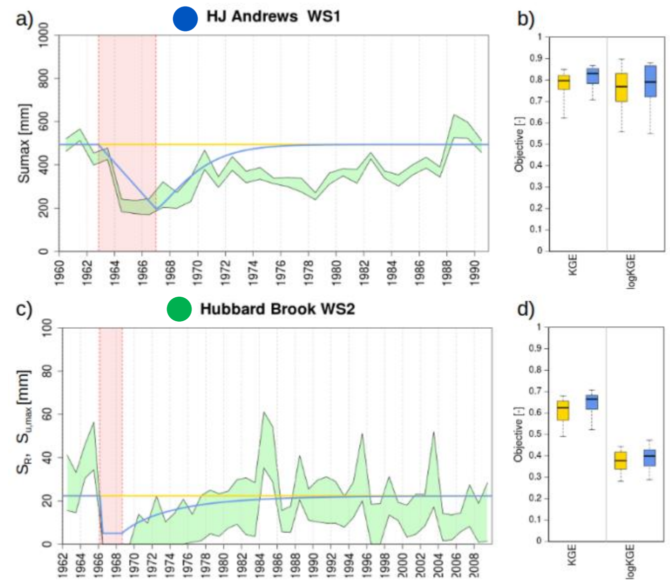
Nijzink et al., 2016, HESS

Prior Constraints

Root zone storage capacity – A step towards dynamic models?



Probability of better model performance with time-dynamic S_R



- Effect stronger in HJ Andrews WS1 → higher absolute S_R
- <u>What is needed?</u>

  (1) Climate projections

  How will precipitation and potential evaporation change?

  (2) Understanding of how vegetation adapts to climatic conditions

  (trade time vs. space!)

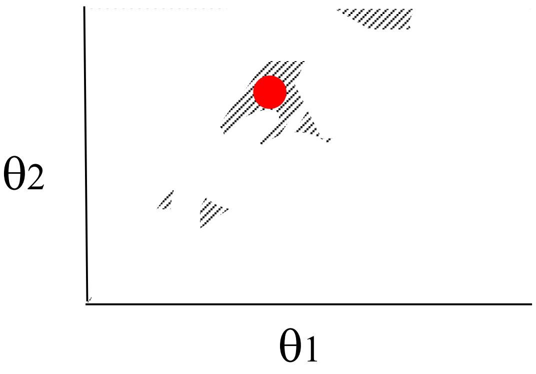
Nijzink et al., 2016, HESS

Prior Constraints

Relational **process constraints**

Limit of Acceptability approaches to extend multiple objective/criteria strategies:

- Expert knowledge and/or semi-quantitative, anecdotal information sufficient
- Ensures plausible (internal) model dynamics
- Easy implementation



Examples:

- Modelled long-term actual evaporation should reflect estimate from Budyko curve: $E_{A,Bud,low} < E_{A,m} < E_{A,\ Bud,high}$
- Modelled long-term base flow contribution should reflect digital filter baseflow estimates: $C_{B,low} < C_{B,m} < C_{B,high}$
- During dry periods higher peak flow contribution from riparian zones than from hillslopes: $Q_{R,dry,avg} > Q_{H,dry,avg}$

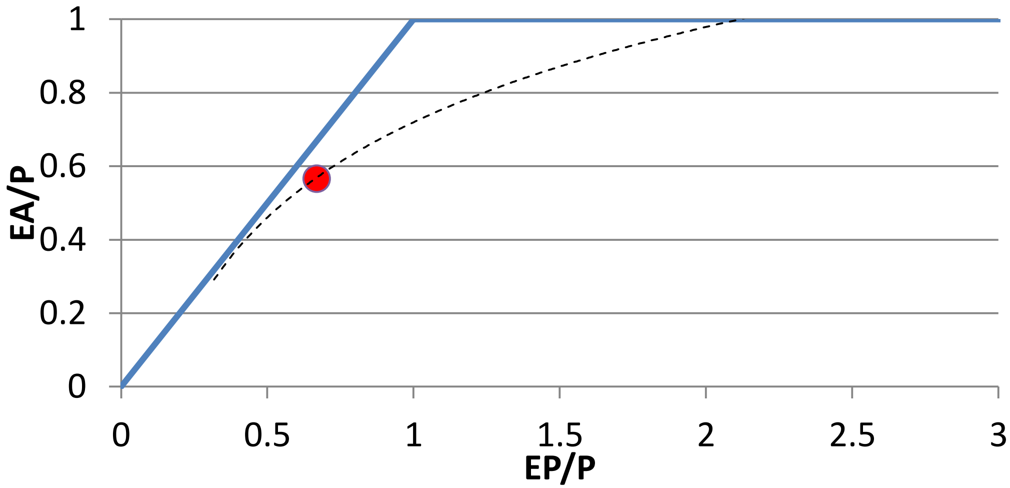
Gharari et al., 2014, HESS
Gao et al., 2014, HESS

m.hrachowitz@tudelft.nl

What to do next?

Model constraints

- Use long-term water balance to constrain parameterization and reduce model uncertainty
- We know: Mean annual runoff coefficient must be satisfied.
- Where there not enough data: estimate actual evaporation from Budyko framework
- From that we can get even more: upper and lower bounds on losses!

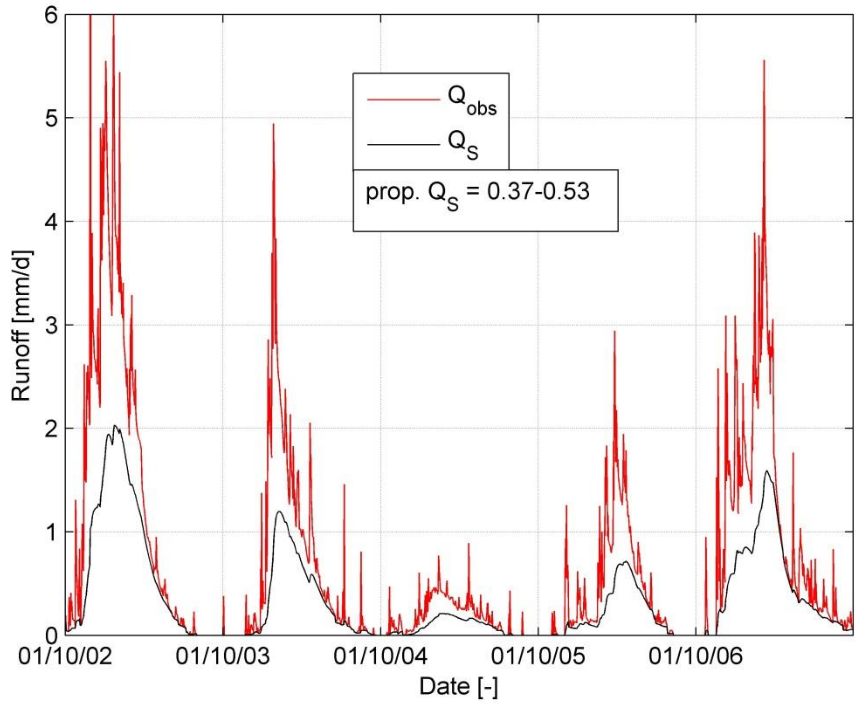


P	1035 mm	1035 mm
- Q	- 274 mm	- 274 mm
- E_P	- 692 mm	
- E_A		- 584 mm
Loss	68 mm	176 mm

Constrain contribution of GW

Model constraints

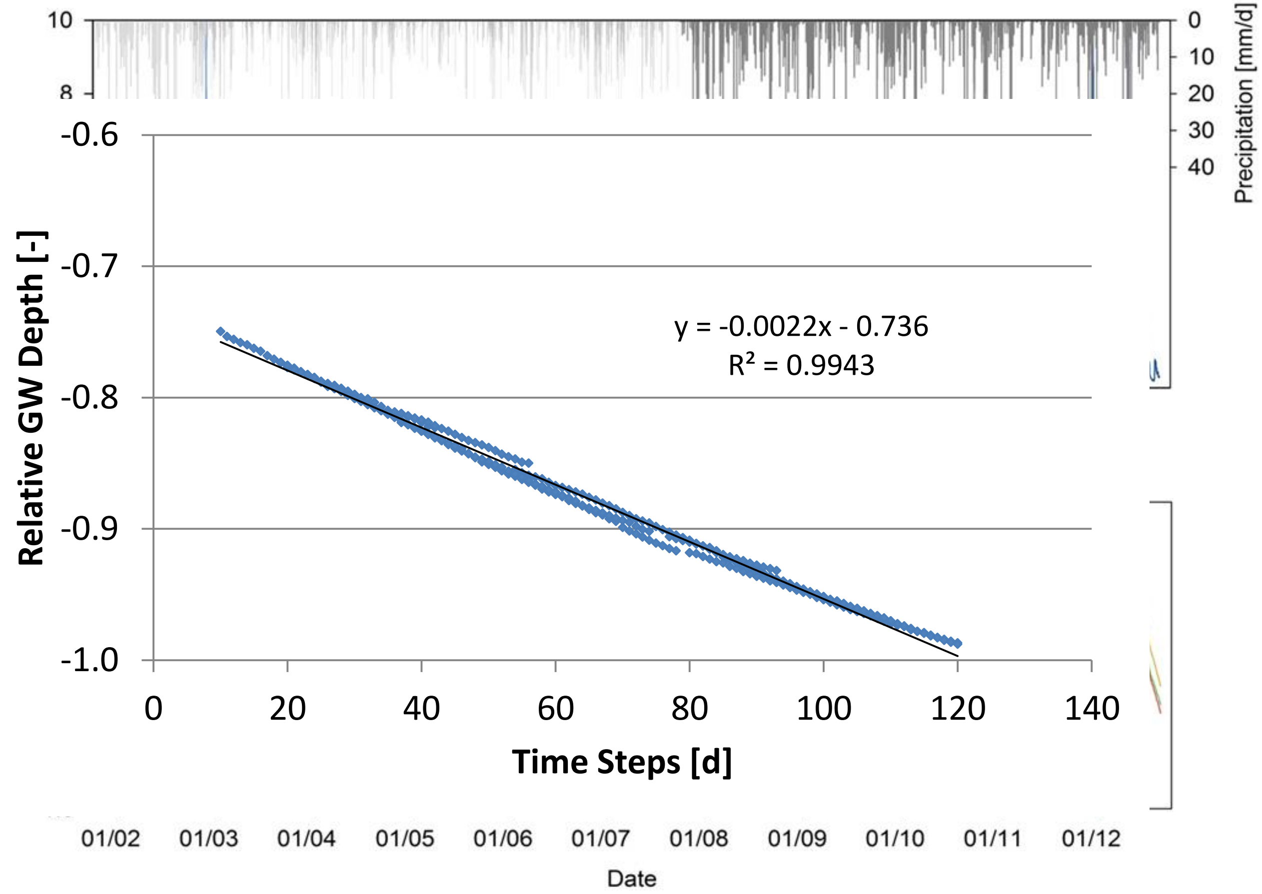
- Use of digital filter or other techniques to get rough estimate of Qs
  -> upper/lower bounds



But there is more we know….

Model constraints

- Lucky: at scale of interest GW falling at an approx. constant rate



m.hrachowitz@tudelft.nl

…and yet more!

Model constraints

- From fact that we have 0-flow periods, constant loss rate can be directly estimated
- When loss rate is known, K_S can also be estimated using a Master Recession Curve
- -> reduce free calibration parameters by 3 (K_S, K_L, S_{uP})

Do not forget landscape heterogeneity!

Model constraints

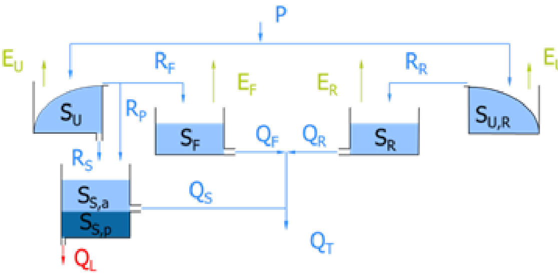
- Wetland in catchment: different process
  -> increase process heterogeneity
- Expert knowledge

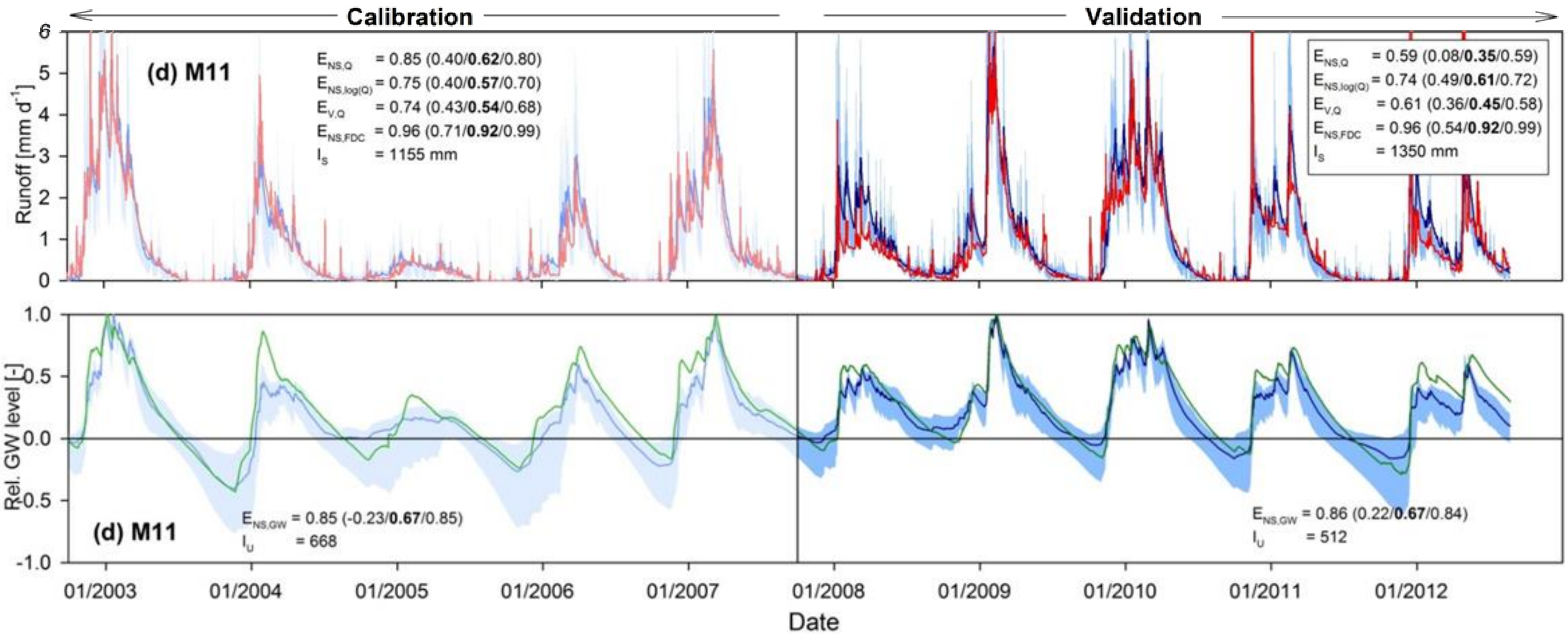
- **Parameter Constraints:**
  1) proportion f of wetland known
  2) $S_{Umax} > S_{Umax,WL}$

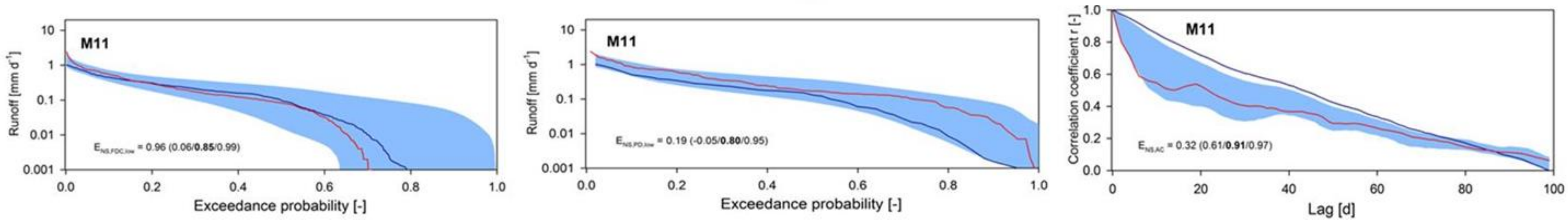
- **Process constraints:**
  3) $\Sigma Q_{WL} > \Sigma Q_F$ in dry periods
  4) $\Sigma Q_F > \Sigma Q_{WL}$ in wet periods

M11 – Nonlinear wetland

Model 11: 12 parameters, 8 constraints



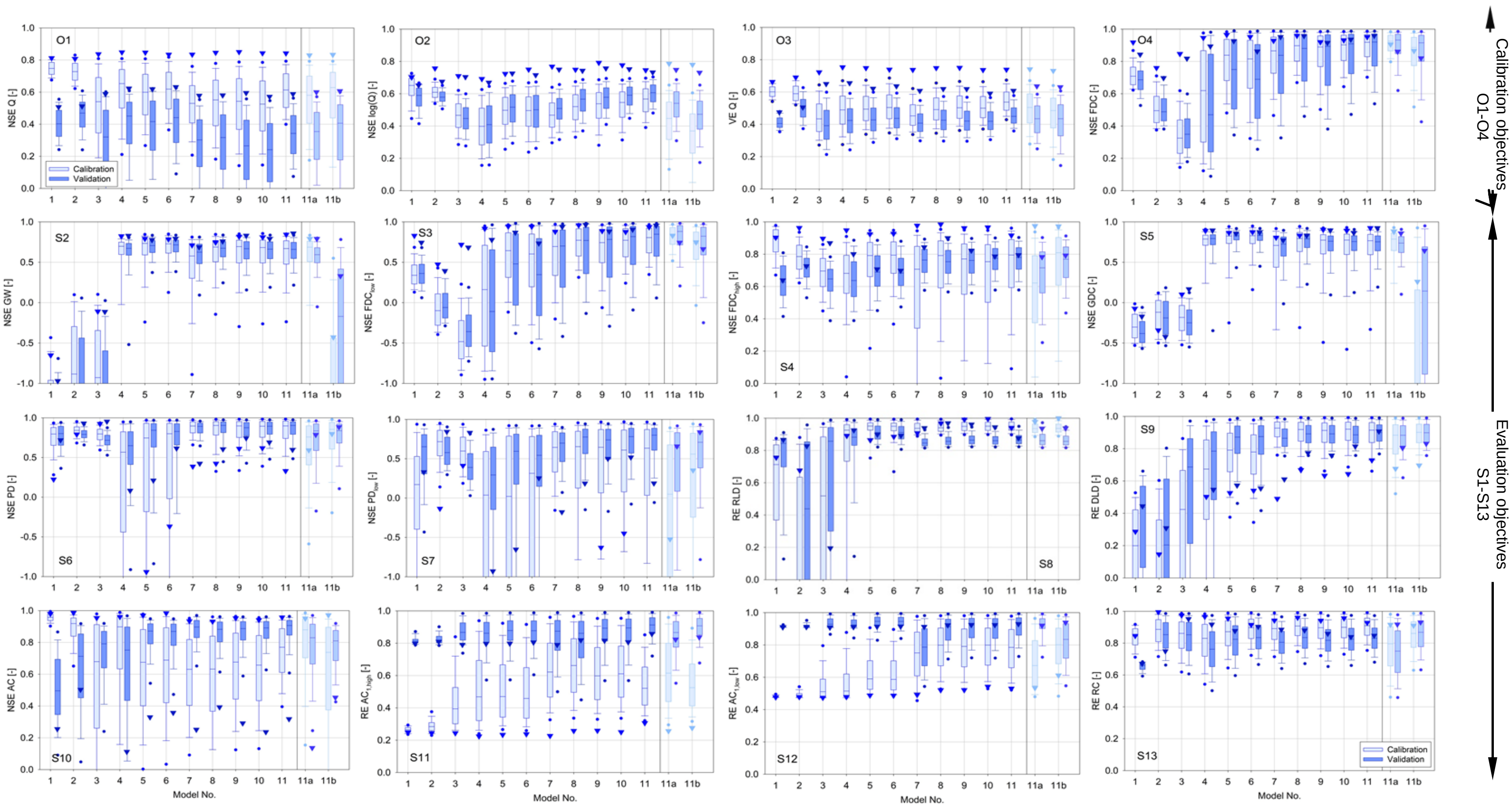




Hrachowitz et al., 2014, WRR

Model evaluation

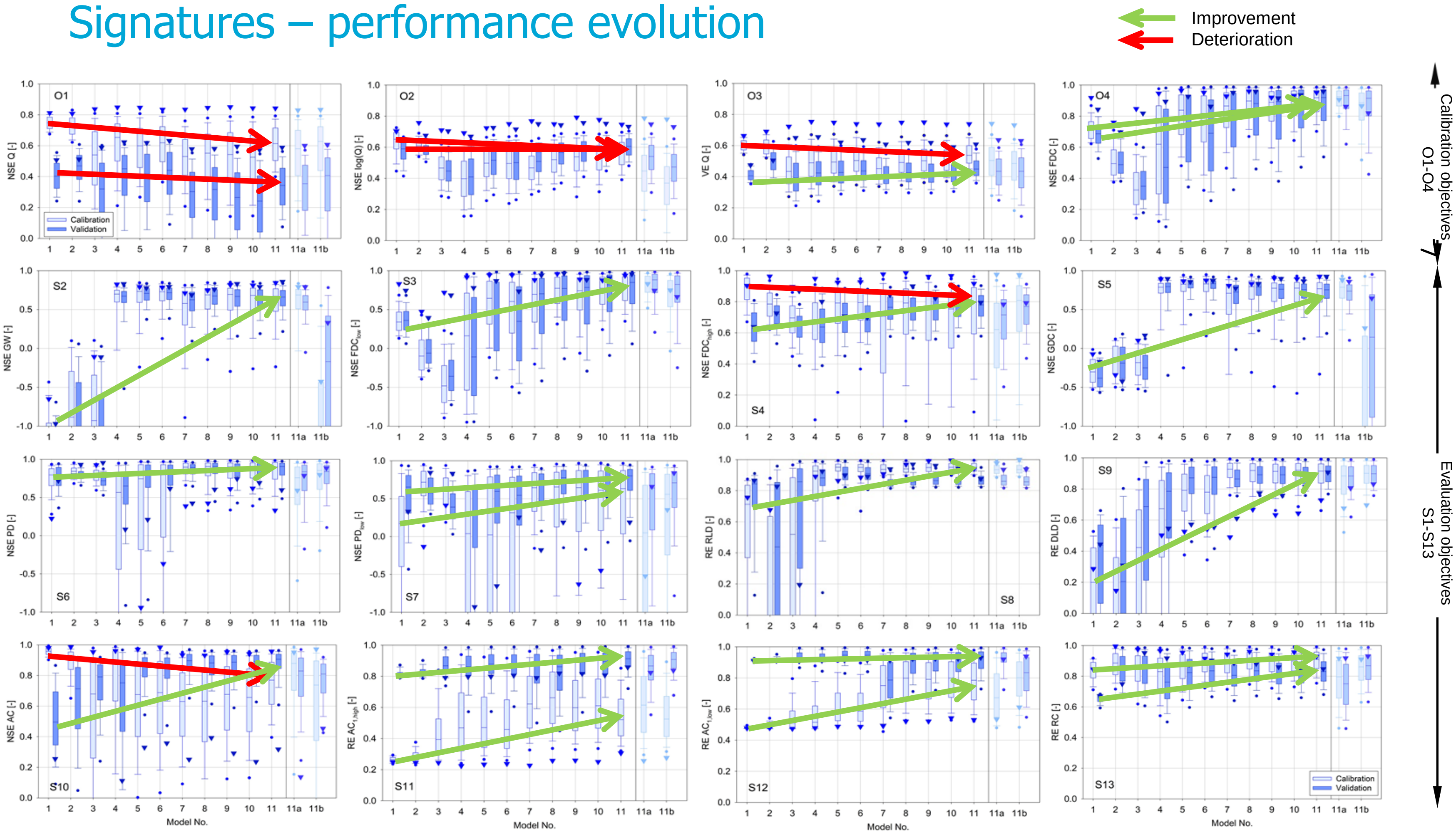
Signatures



Hrachowitz et al., 2014, WRR

Model evaluation

Signatures – performance evolution

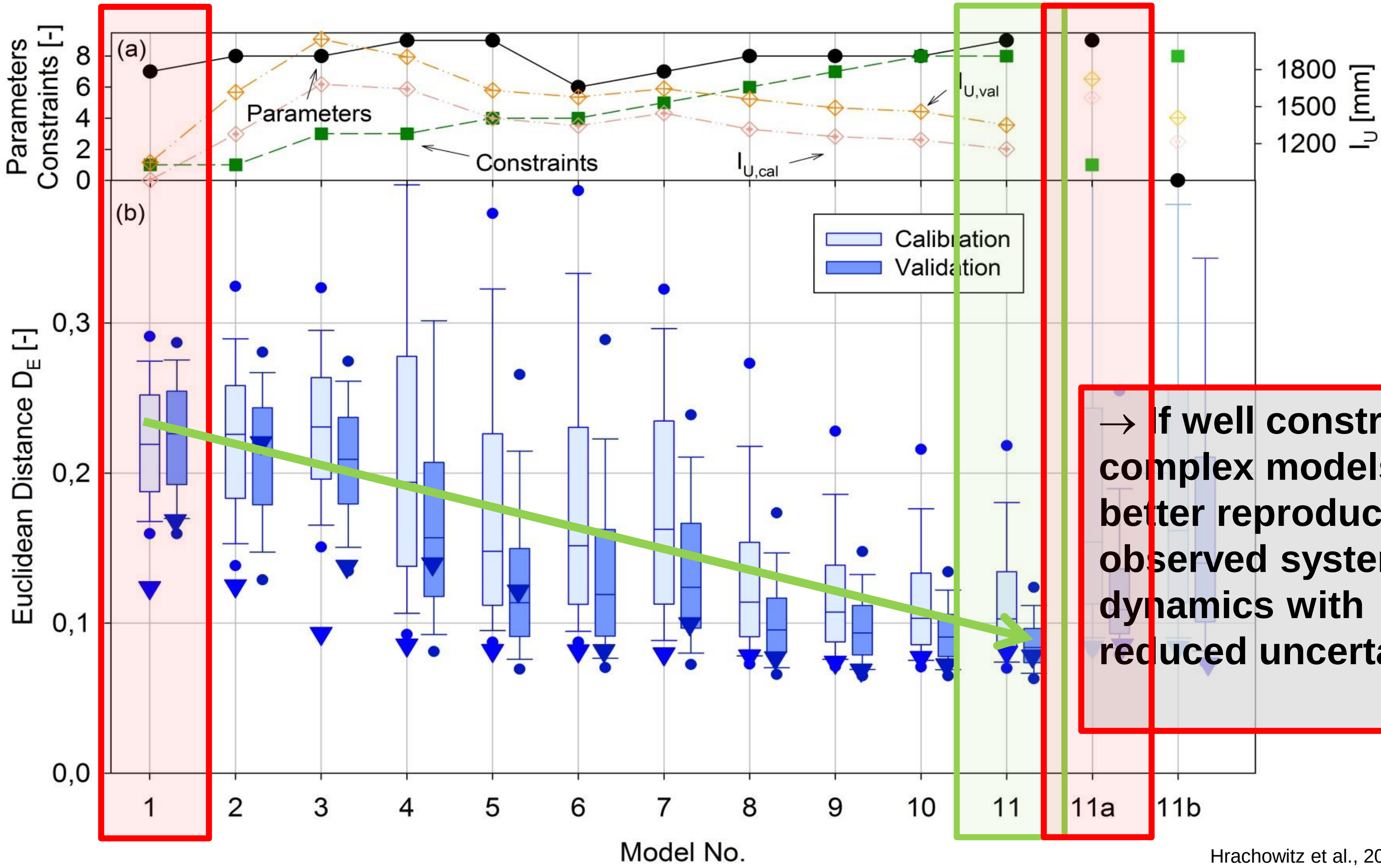


Hrachowitz et al., 2014, WRR

m.hrachowitz@tudelft.nl

Model evaluation

Overall performance evolution using 20 system signatures

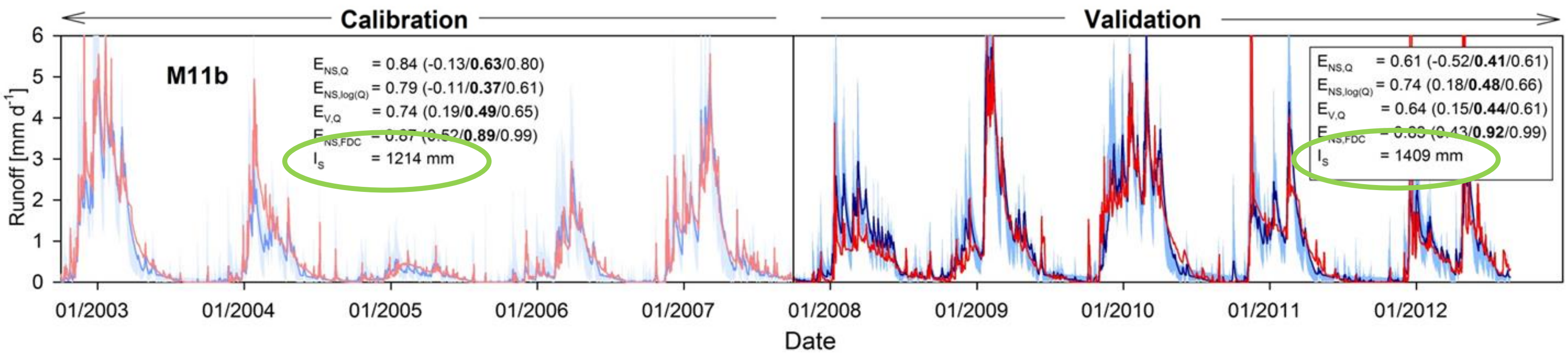


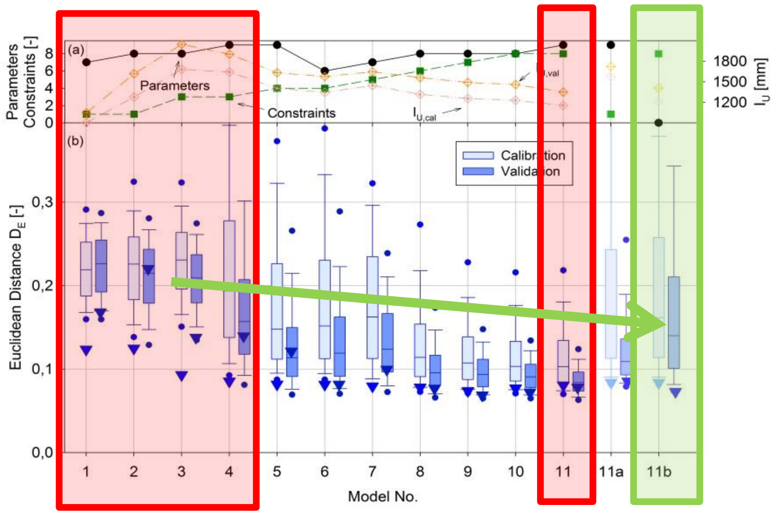
→ If well constrained, complex models can better reproduce observed system dynamics with reduced uncertainty

Hrachowitz et al., 2014, WRR

Model evaluation

M11b – Uncalibrated but all constraints applied





- Uncalibrated but constrained, higher complexity model M11b clearly outperforms standard 3-box model M1

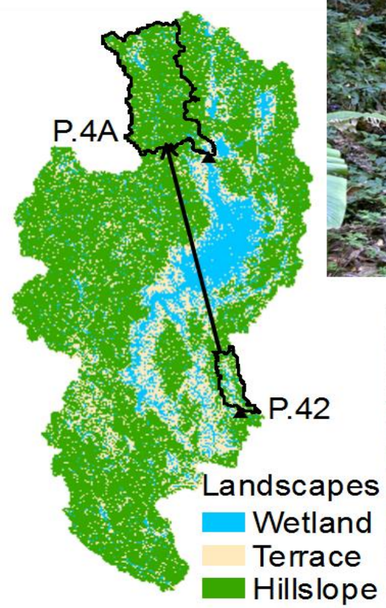
→ **Applying a sufficient number of relatively loose constraints on complex models can minimize the need for calibration – PUB!**

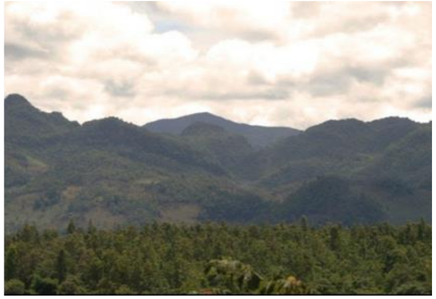
Hrachowitz et al., 2014, WRR

Model transferability

How well can a calibrated model be transferred to another catchment?

- Relevant where no data for calibration are available
- Strong test for a model's ability to capture dominant processes



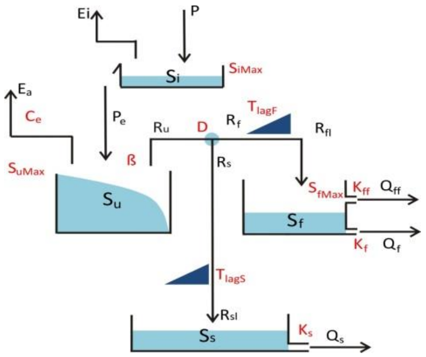


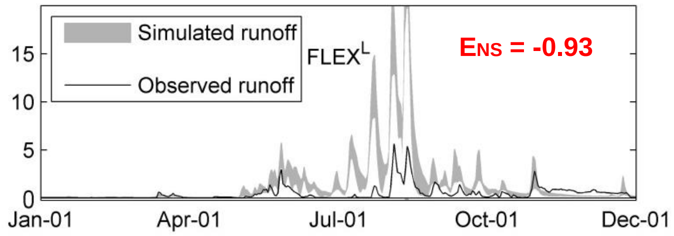
Gao et al., 2016, WRR

Model transferability

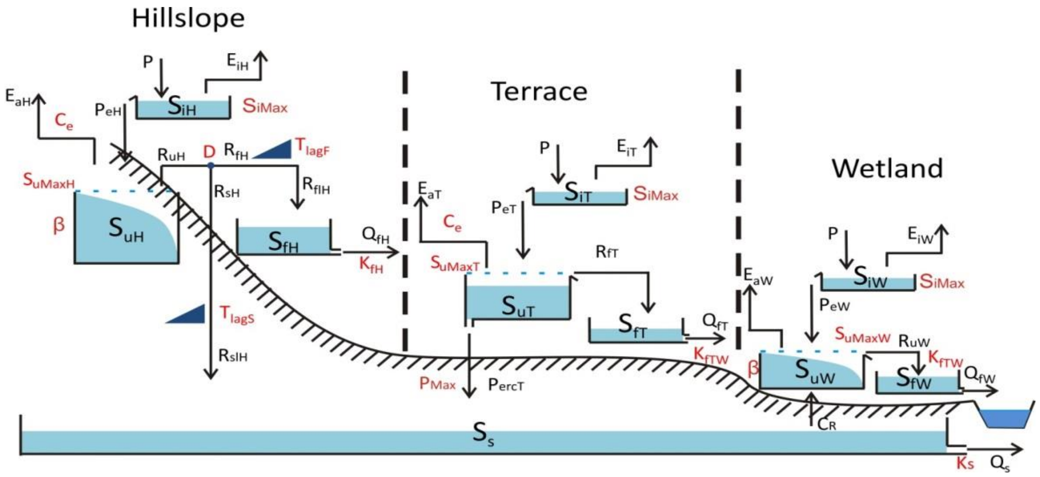
The value of a more complete system representation

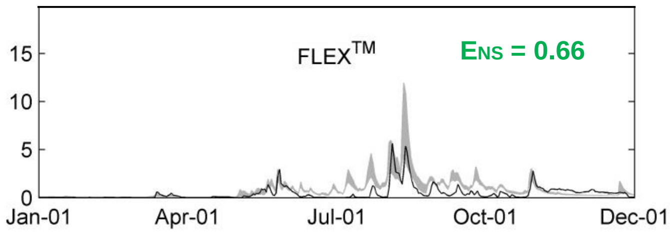
Lumped model





Semi-distributed, constrained model

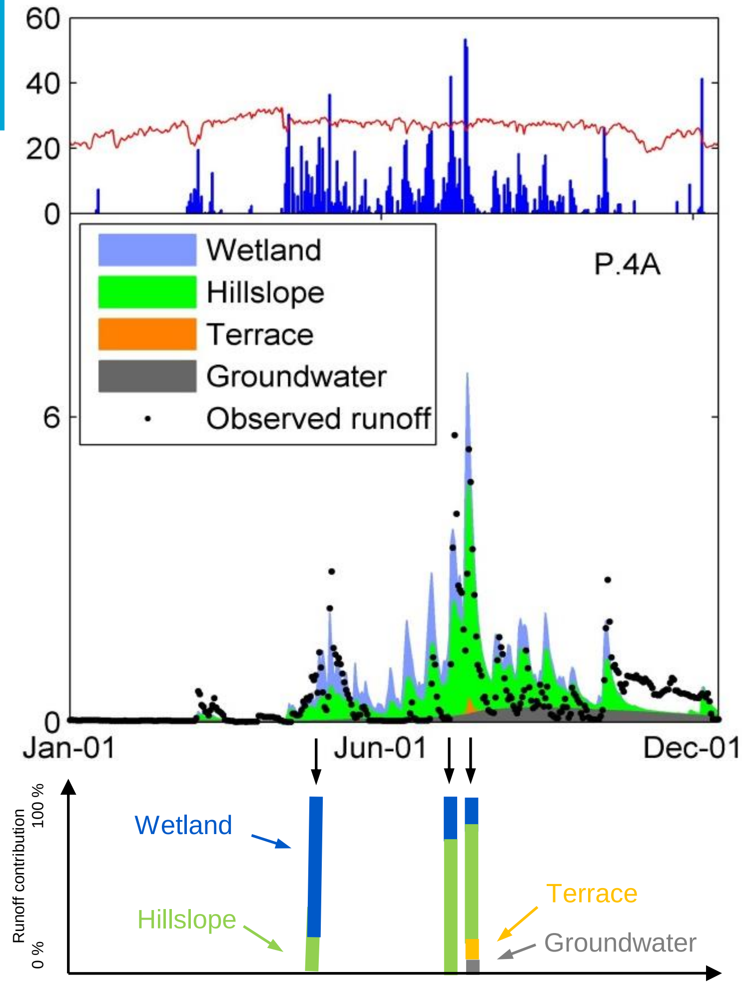


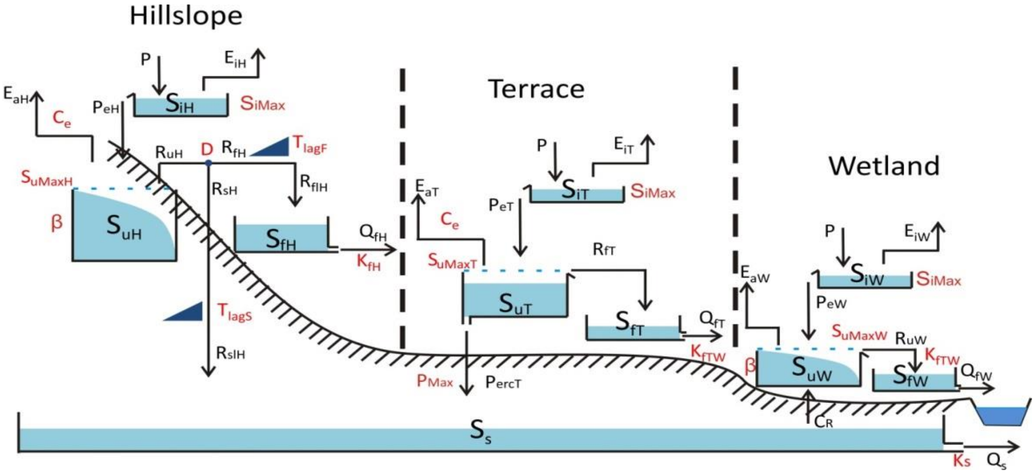


Gao et al., 2016, WRR

System internal dynamics

The value of a more complete system representation





Gao et al., 2016, WRR

In a nutshell

- Need for more realistic representations of the system
- Value of carefully and efficiently exploiting available data to more efficiently extract relevant information
- Combination of
  (1) process heterogeneity
  (2) hydrologic signatures
  (3) expert knowledge and eventually
  (4) additional data critical for improving models
- Need to have a closer look at available information!
- We KNOW more than we typically think we do!

THANK YOU!!