Inferring model structural deficits by analyzing temporal dynamics of model performance and parameter sensitivity

Dominik Reusser, Erwin Zehe

EGU 2011

reusser@pik-potsdam.de

Outline

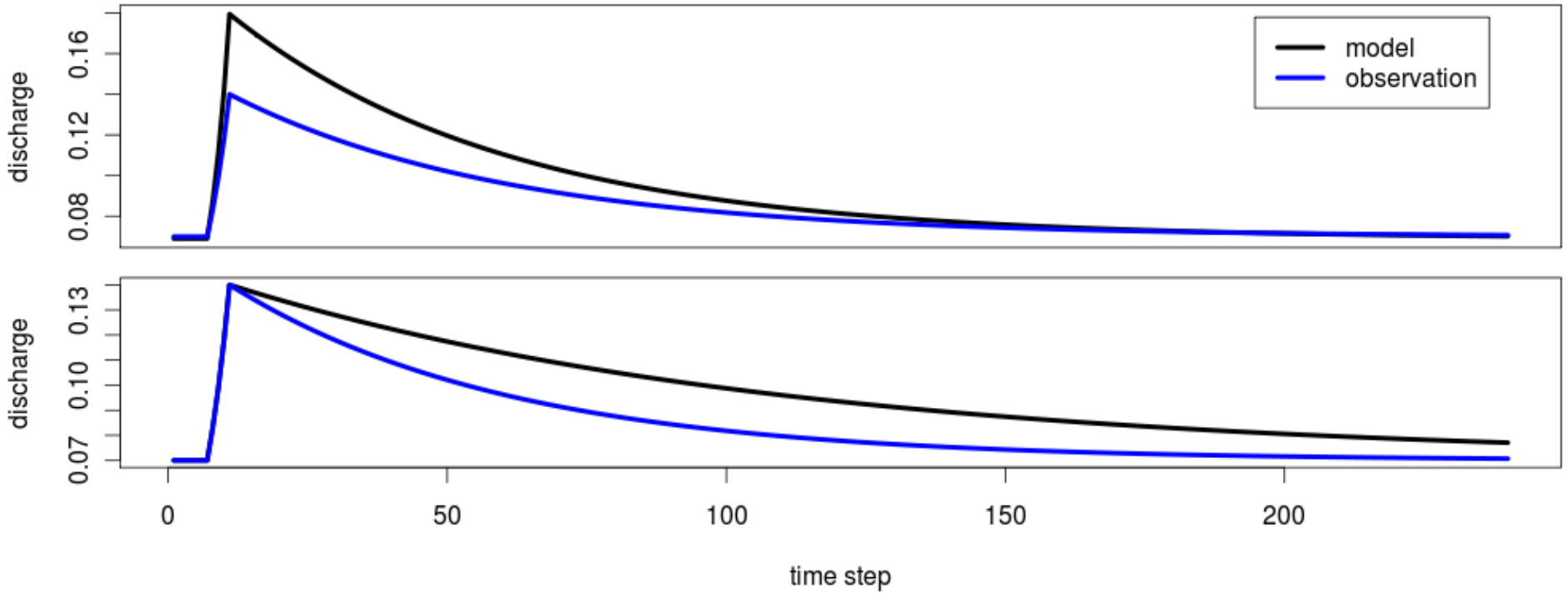
- Finger prints of model performance (TIGER)
  - Results for case study
- Parameter sensitivity (TEDPAS)
  - Results for case study

Reusser et al. HESS 2009,
Reusser et al. WRR (submitted)
Reusser and Zehe WRR (submitted)

Packages available as open source software in R

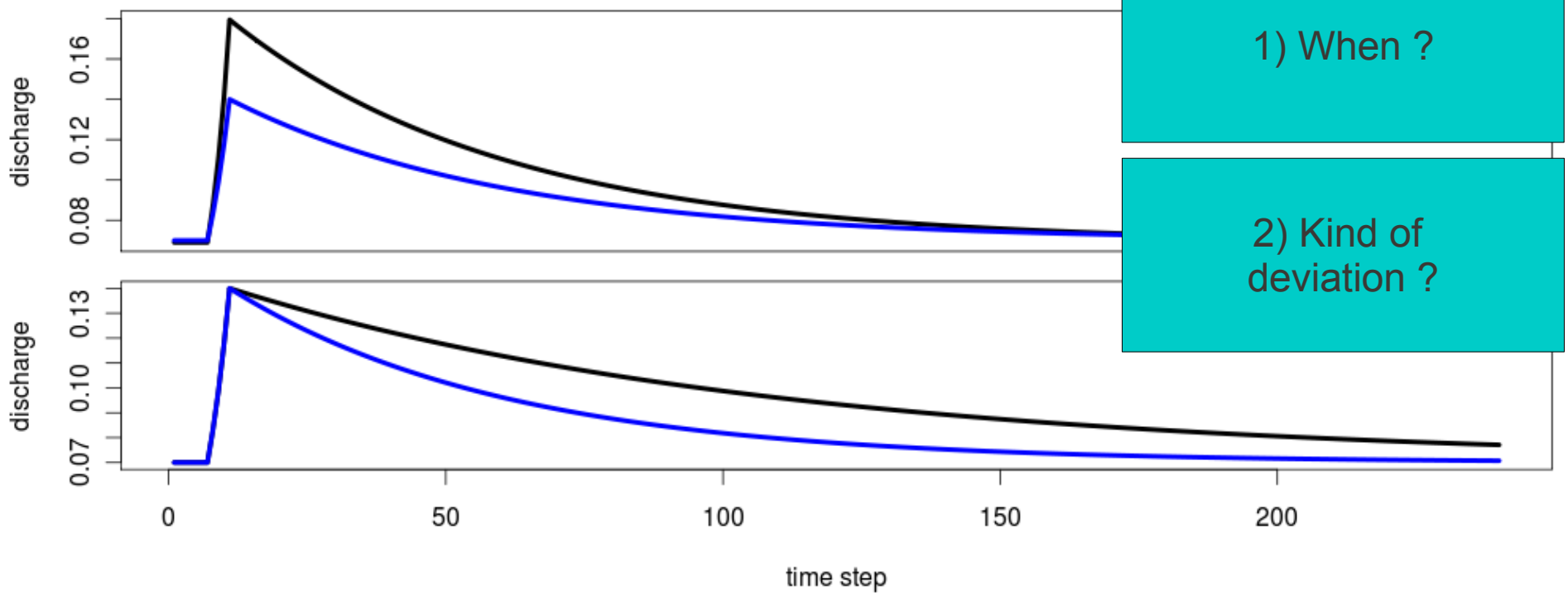
TIGER: Motivation

Time series of grouped errors



TIGER: Motivation

Time series of grouped errors

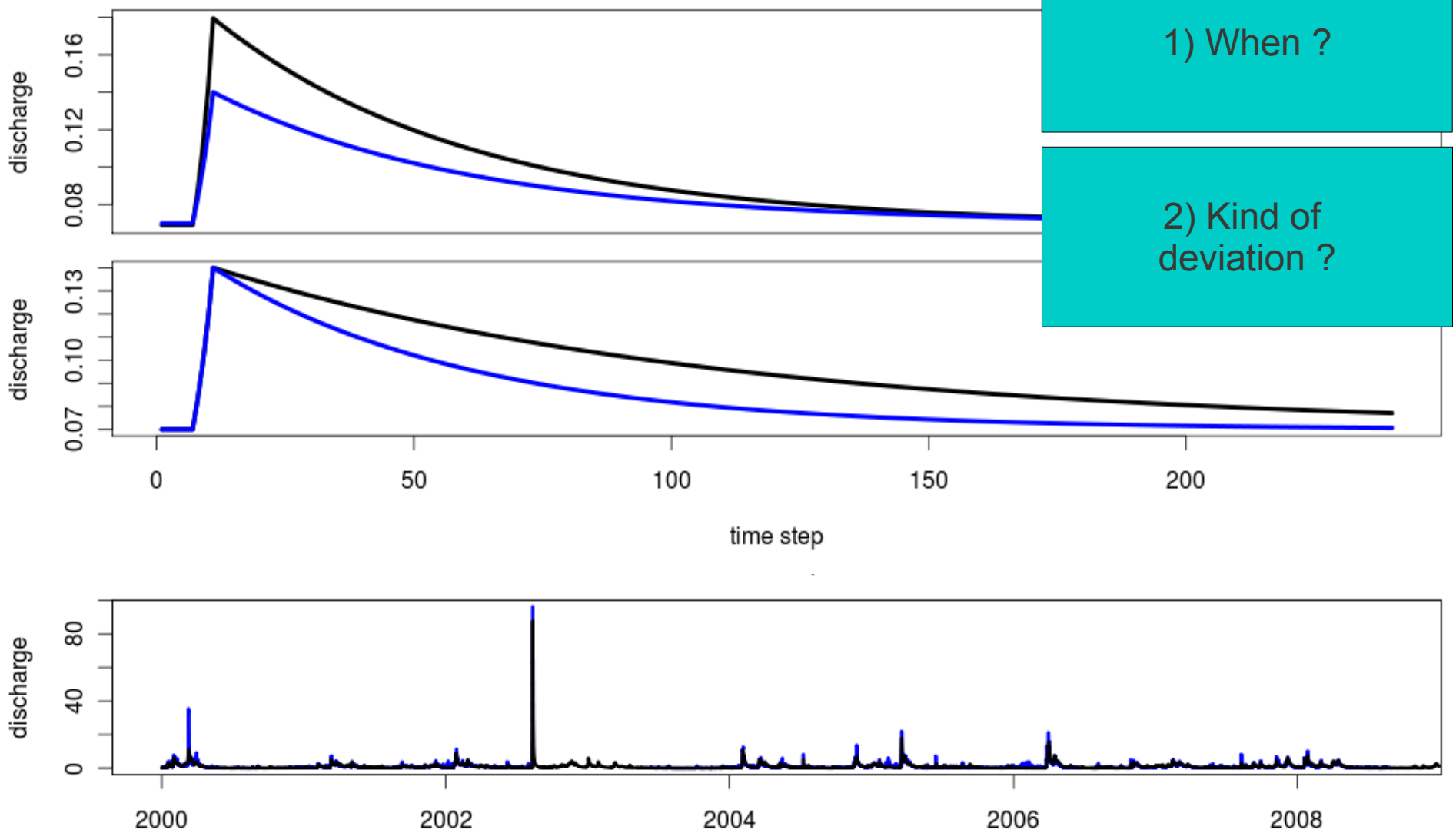


1) When ?

2) Kind of deviation ?

TIGER: Motivation

Time series of grouped errors

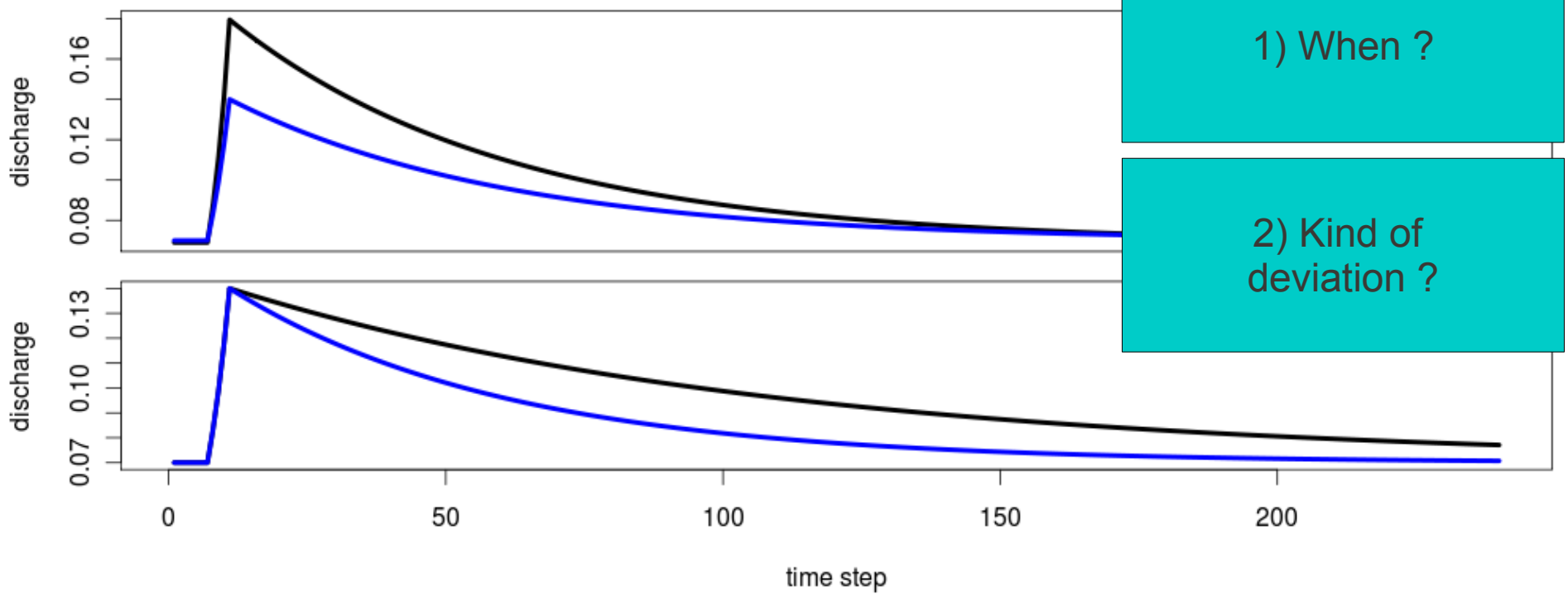


1) When ?

2) Kind of deviation ?

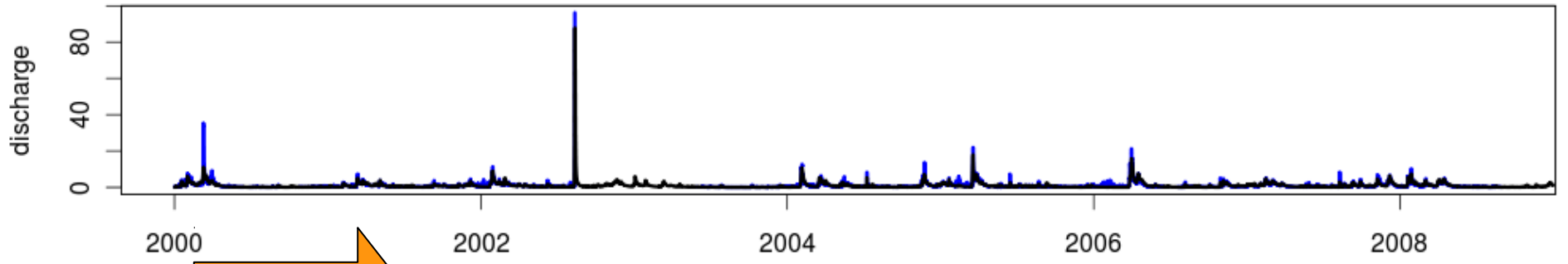
TIGER: Motivation

Time series of grouped errors



1) When ?

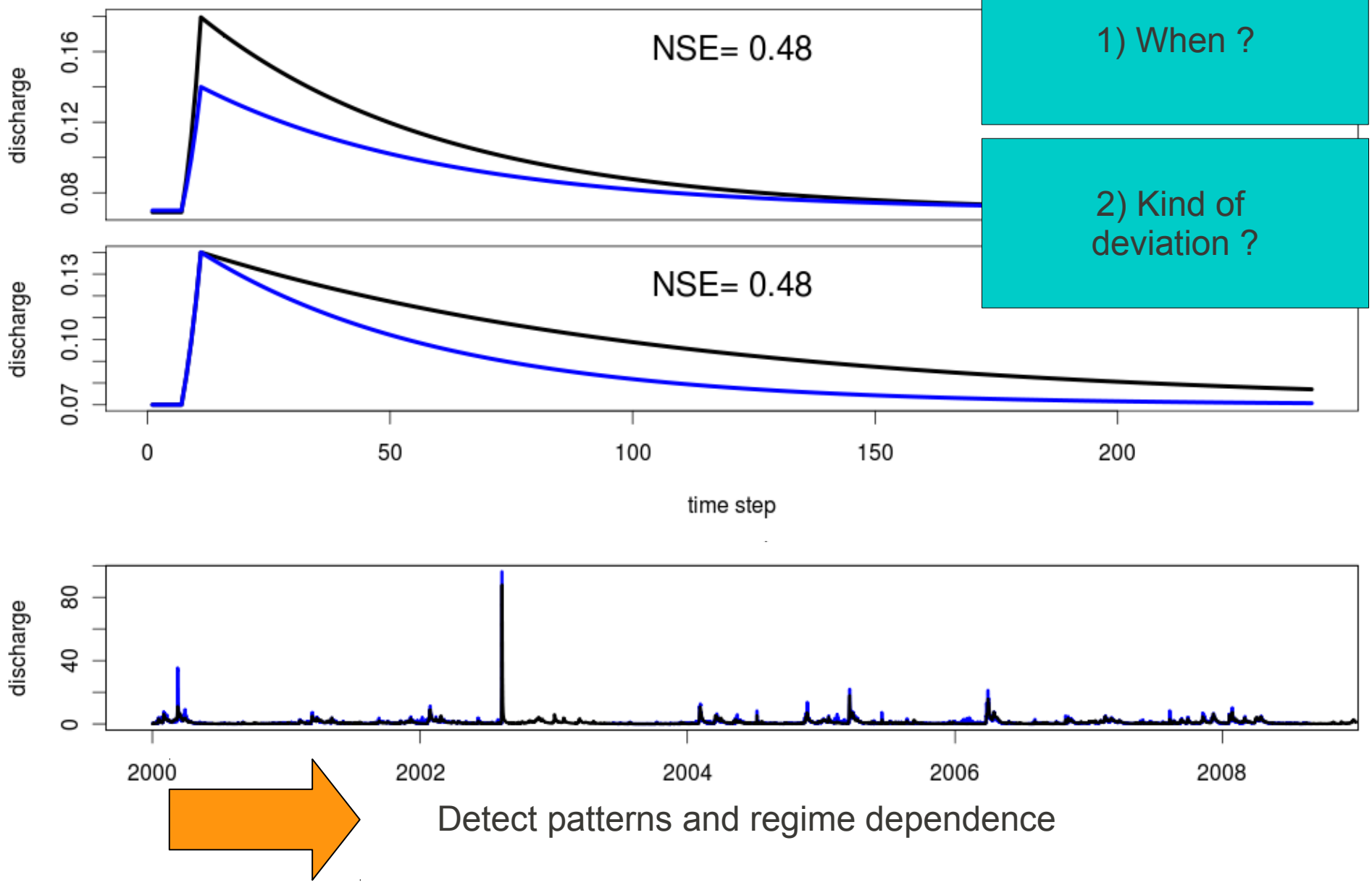
2) Kind of deviation ?



Detect patterns and regime dependence

TIGER: Motivation

Time series of grouped errors

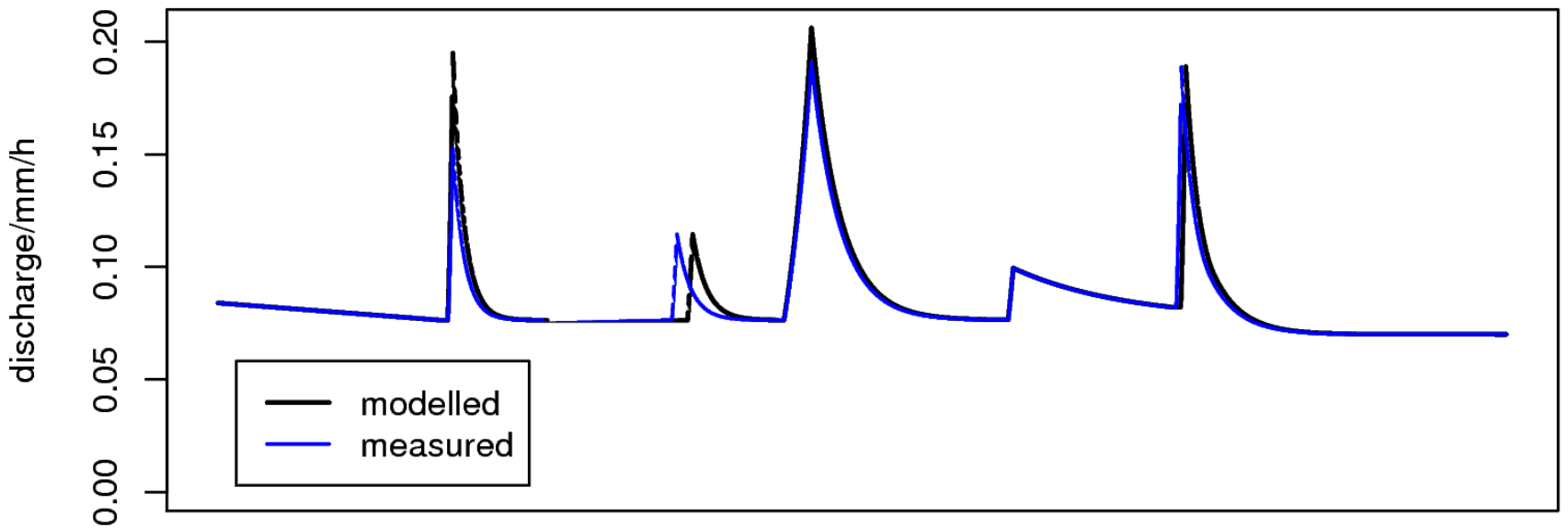


1) When ?

2) Kind of deviation ?

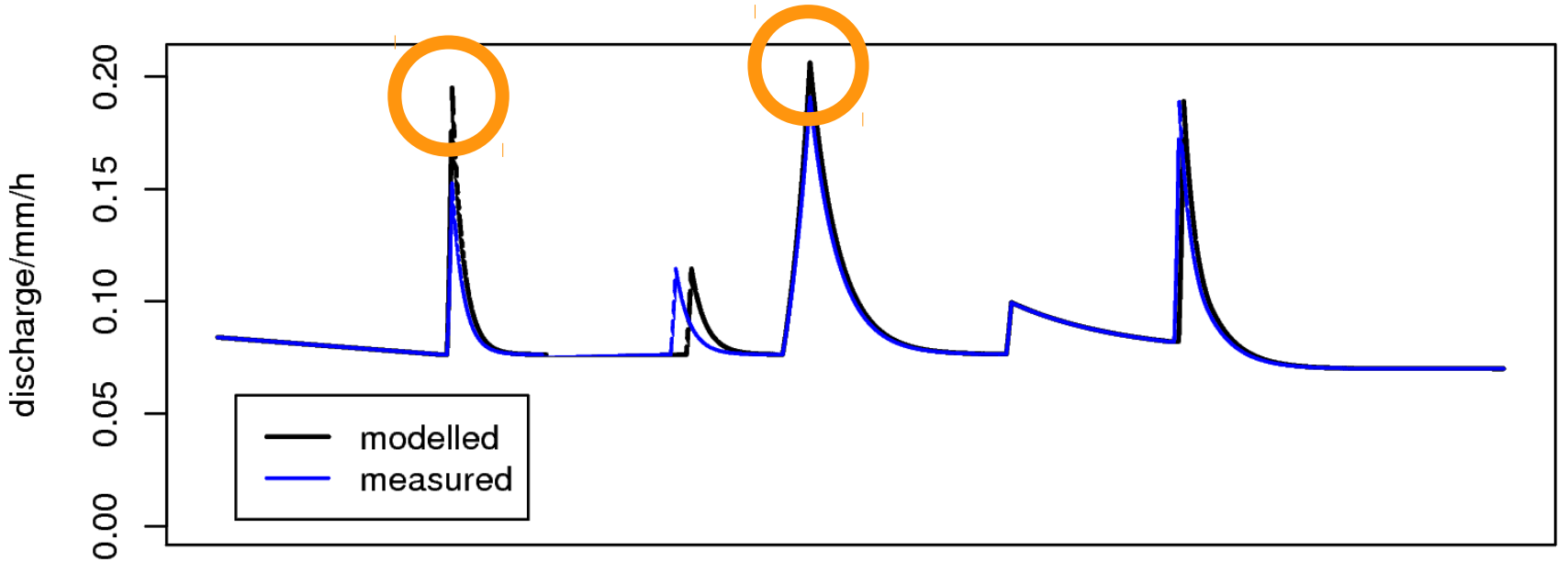
Detect patterns and regime dependence

TIGER method



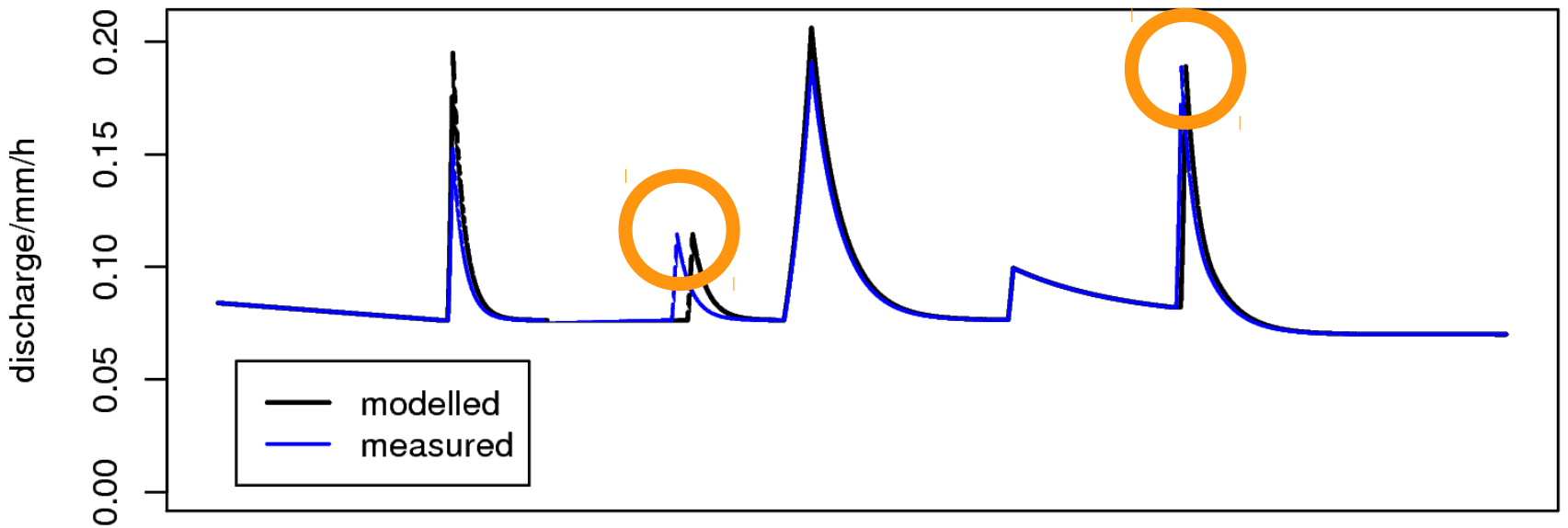
Reusser et al., HESS 2009

TIGER method



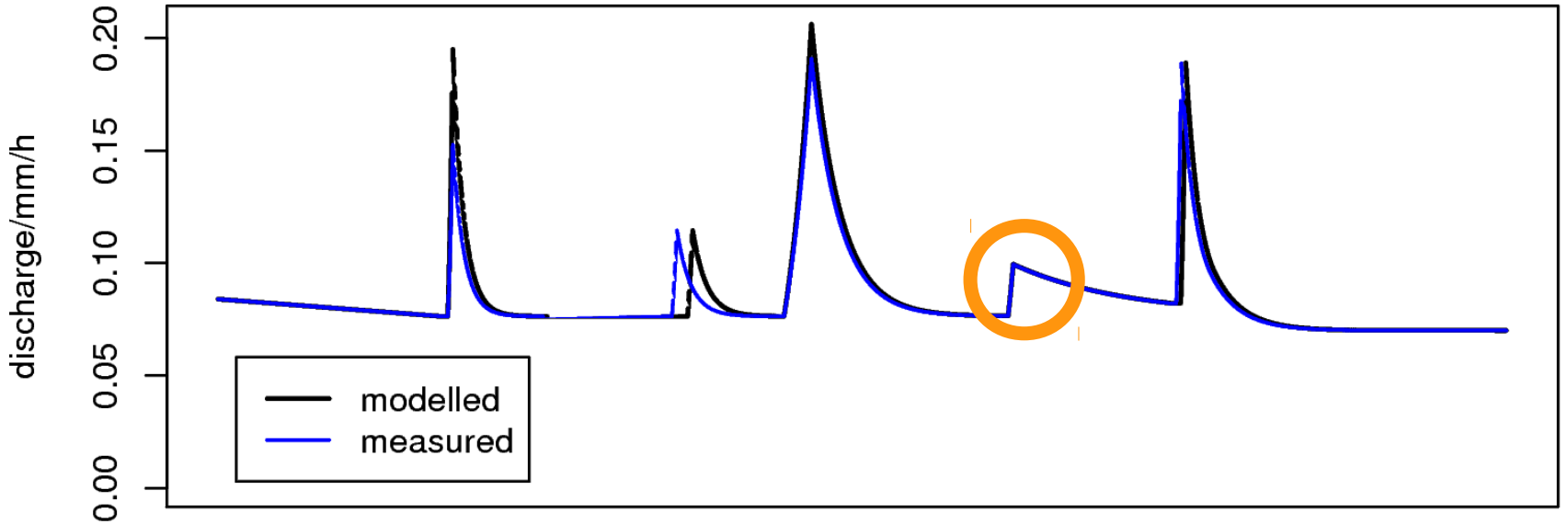
Reusser et al., HESS 2009

TIGER method



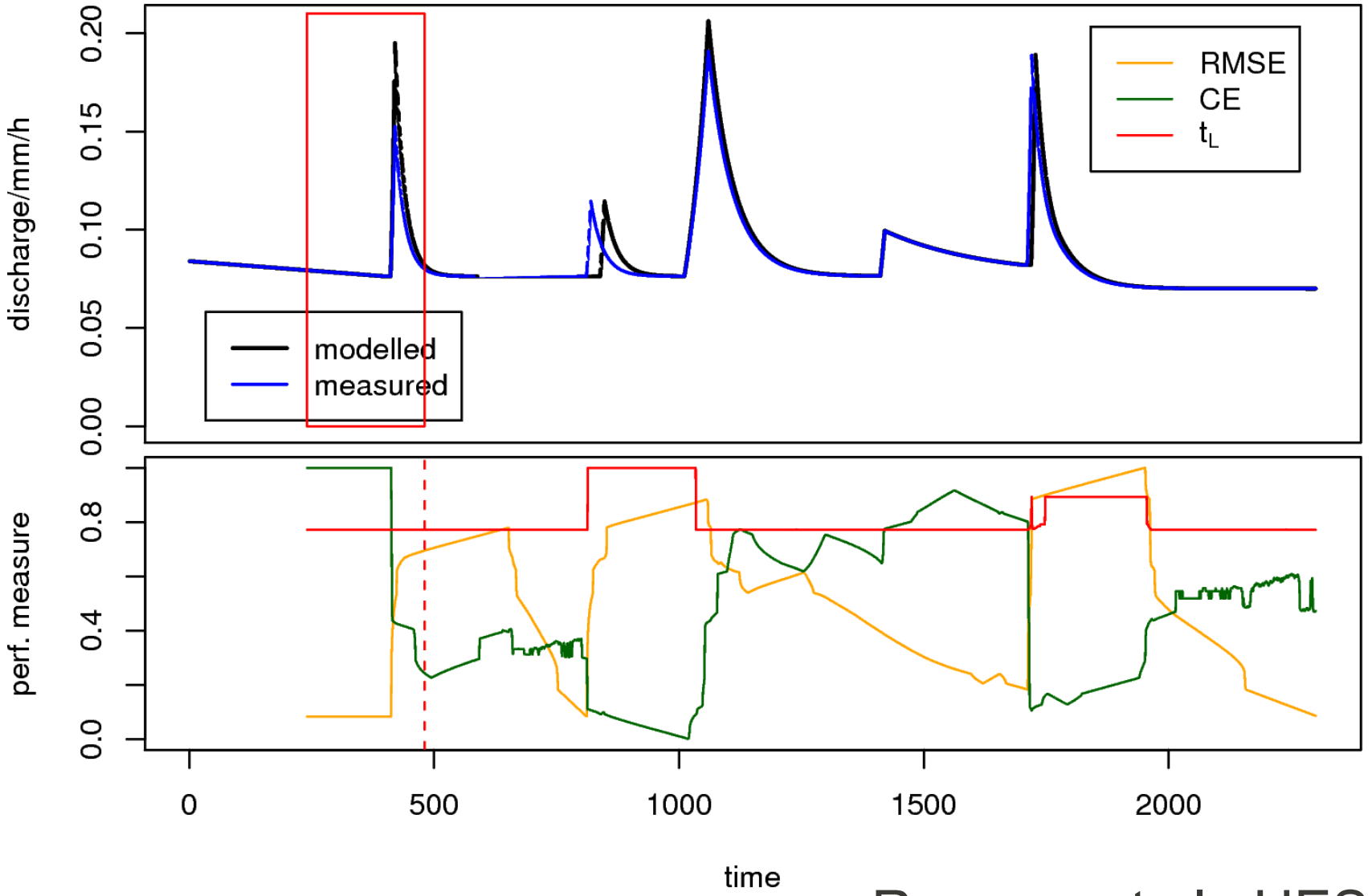
Reusser et al., HESS 2009

TIGER method



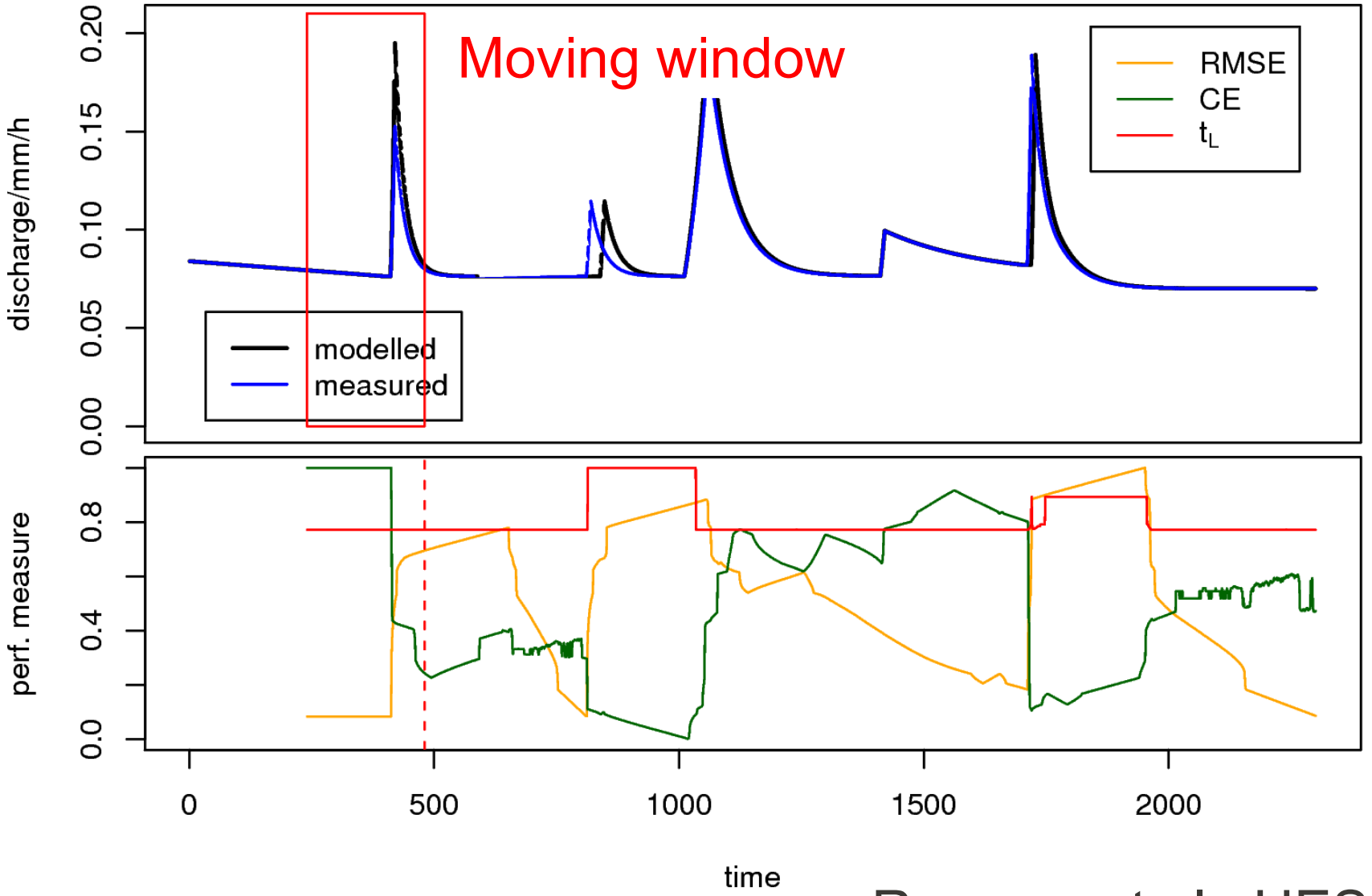
Reusser et al., HESS 2009

TIGER method



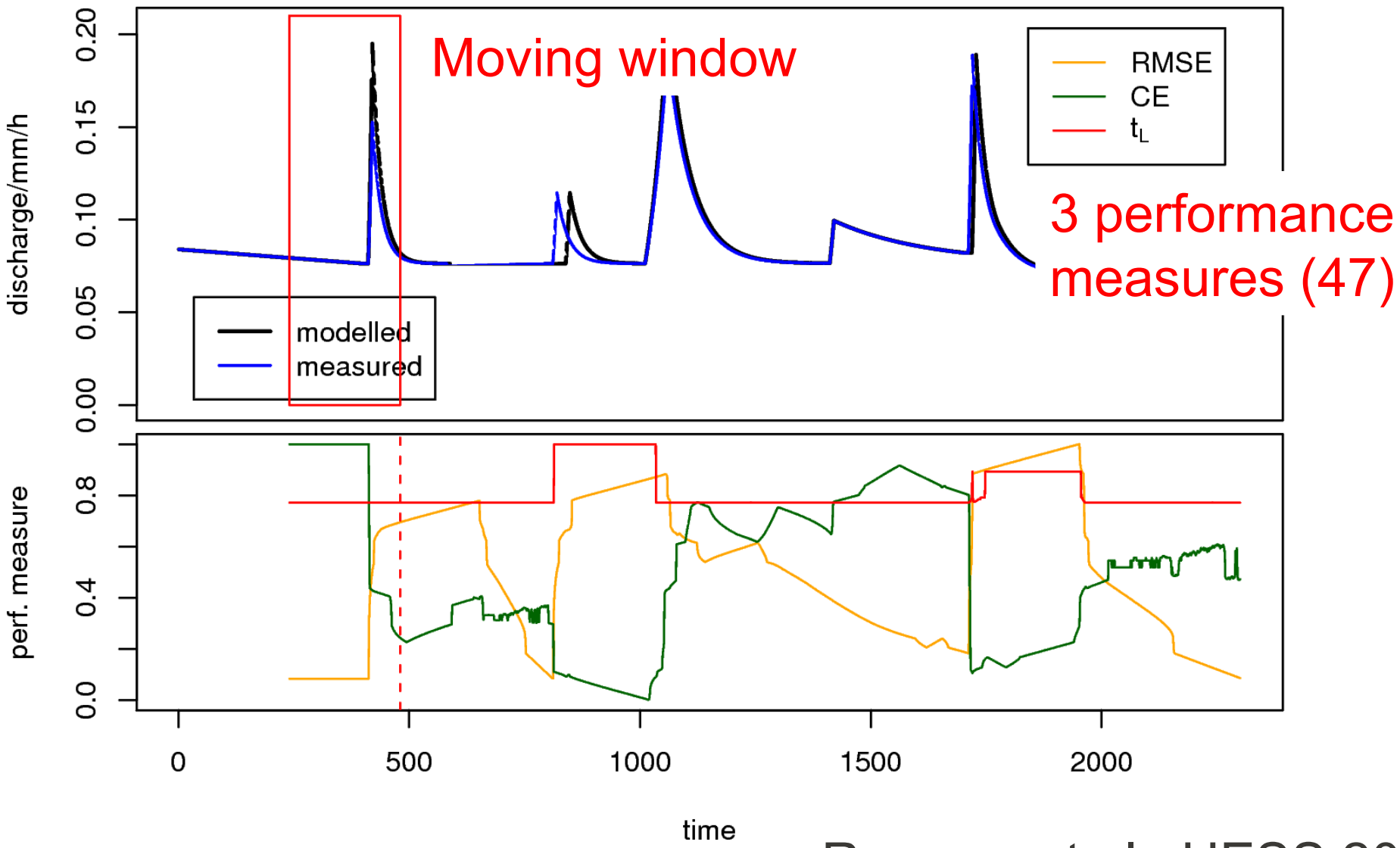
Reusser et al., HESS 2009

TIGER method



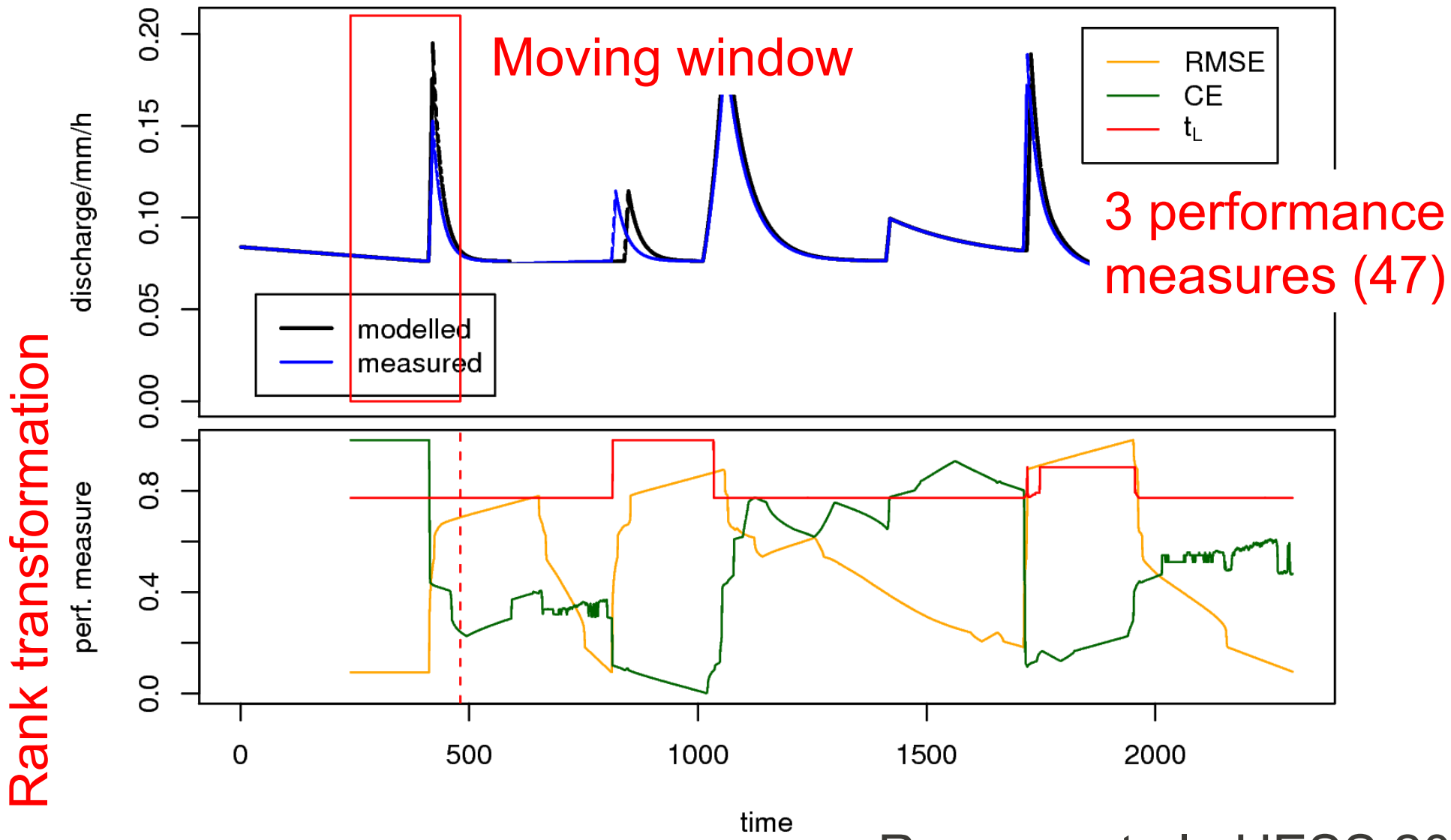
Reusser et al., HESS 2009

TIGER method



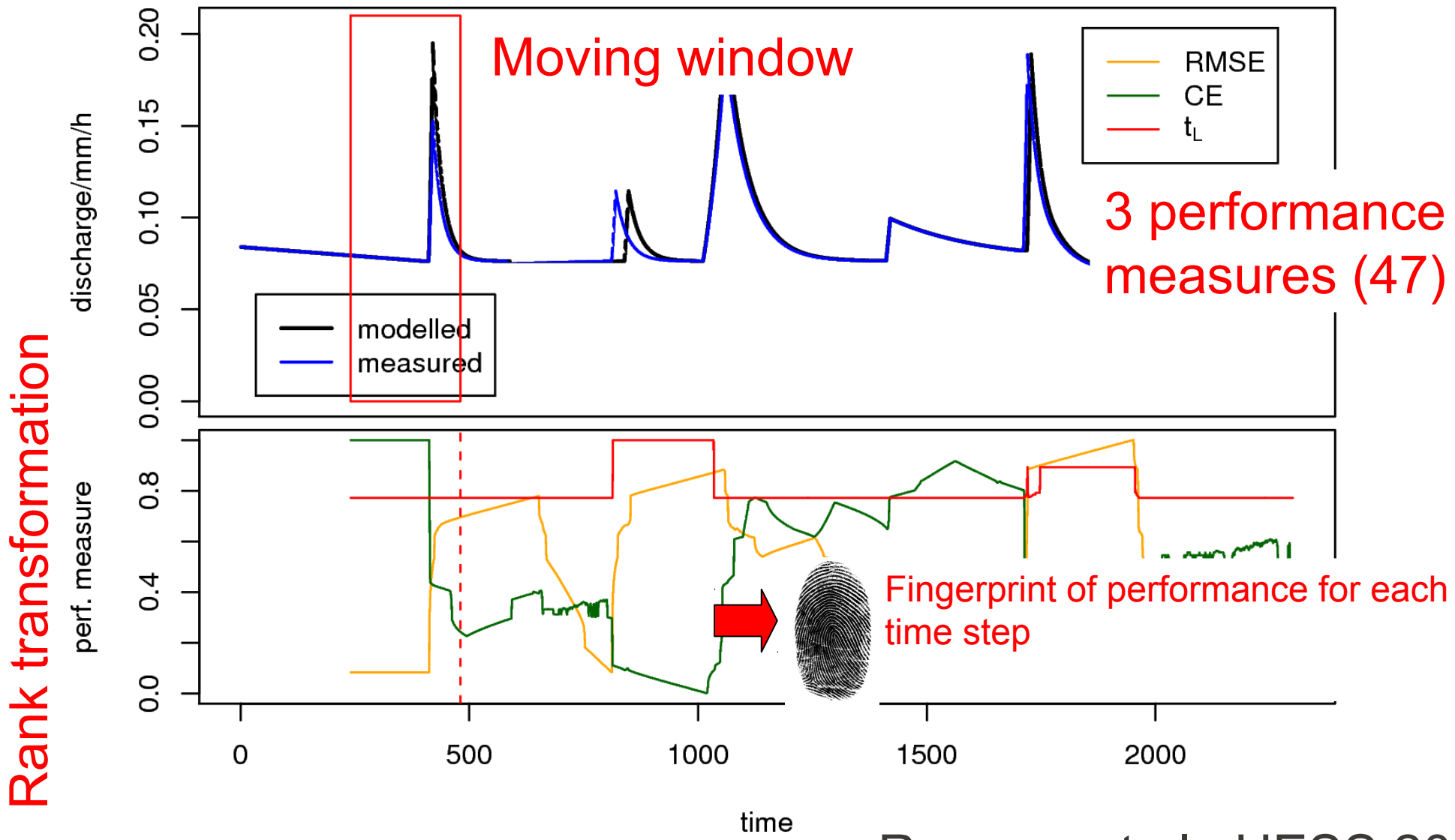
Reusser et al., HESS 2009

TIGER method



Reusser et al., HESS 2009

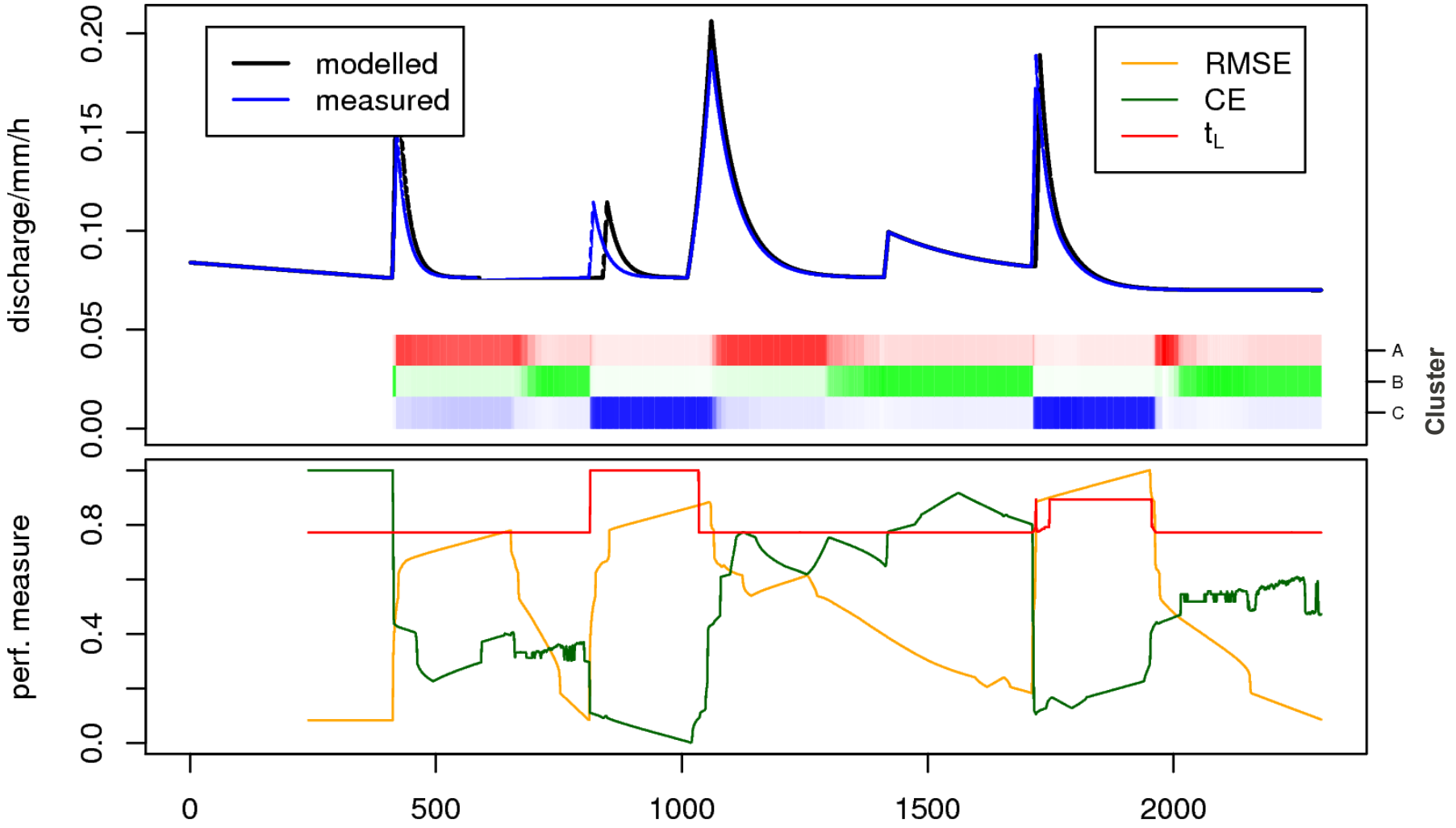
TIGER method



Reusser et al., HESS 2009

TIGER method

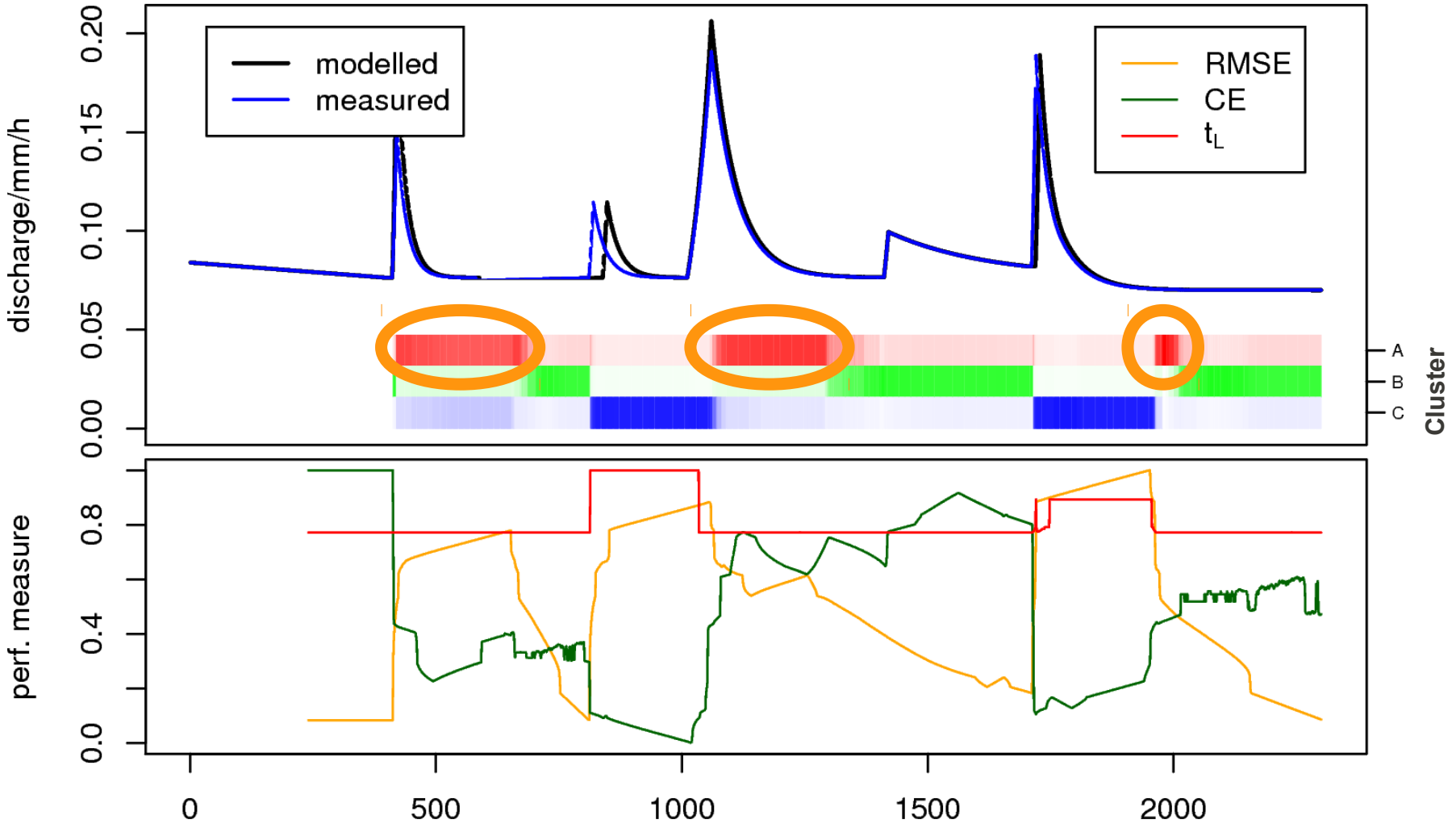
Pattern detection with SOM and fuzzy clustering



Reusser et al., HESS 2009

TIGER method

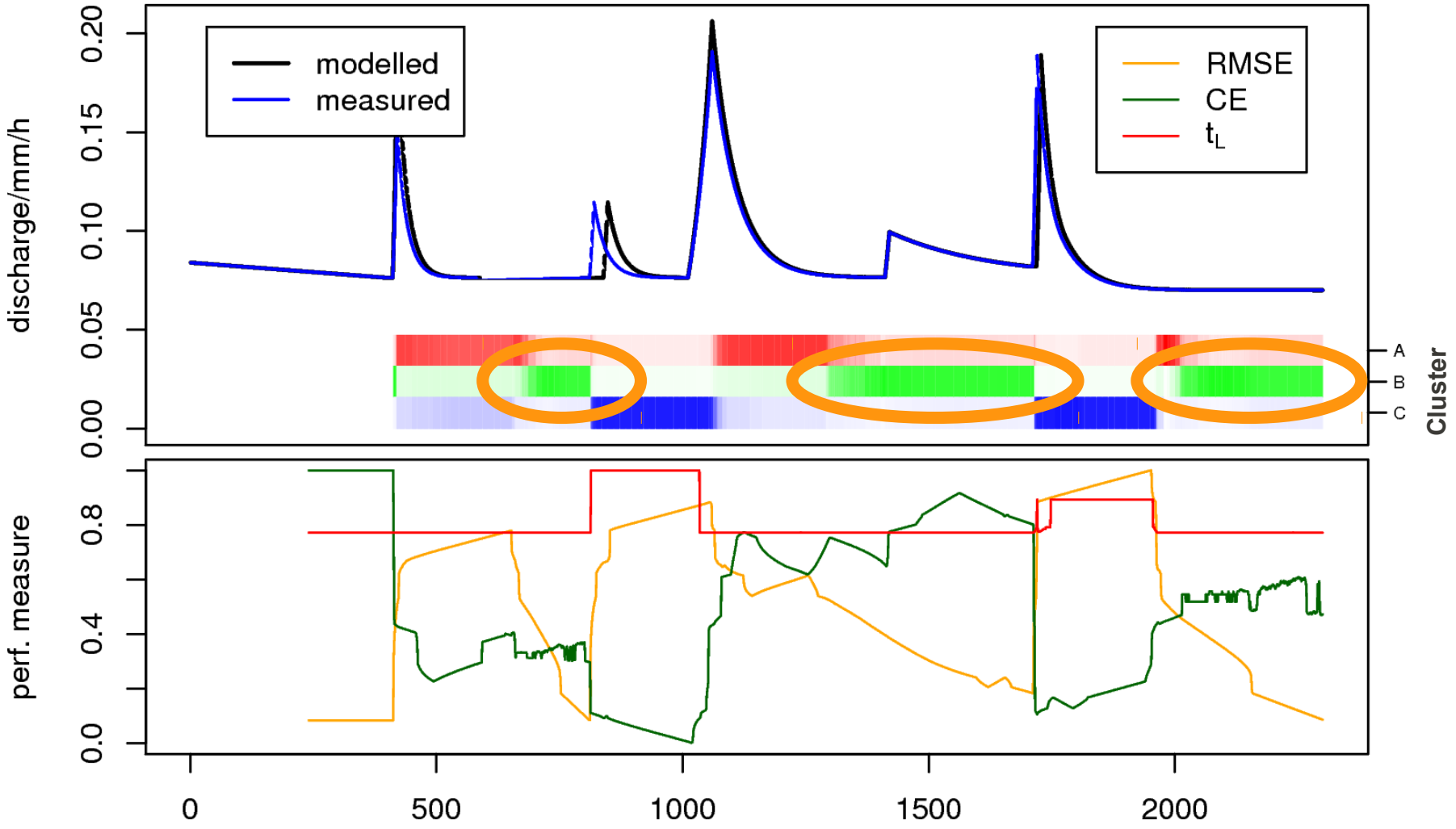
Pattern detection with SOM and fuzzy clustering



Reusser et al., HESS 2009

TIGER method

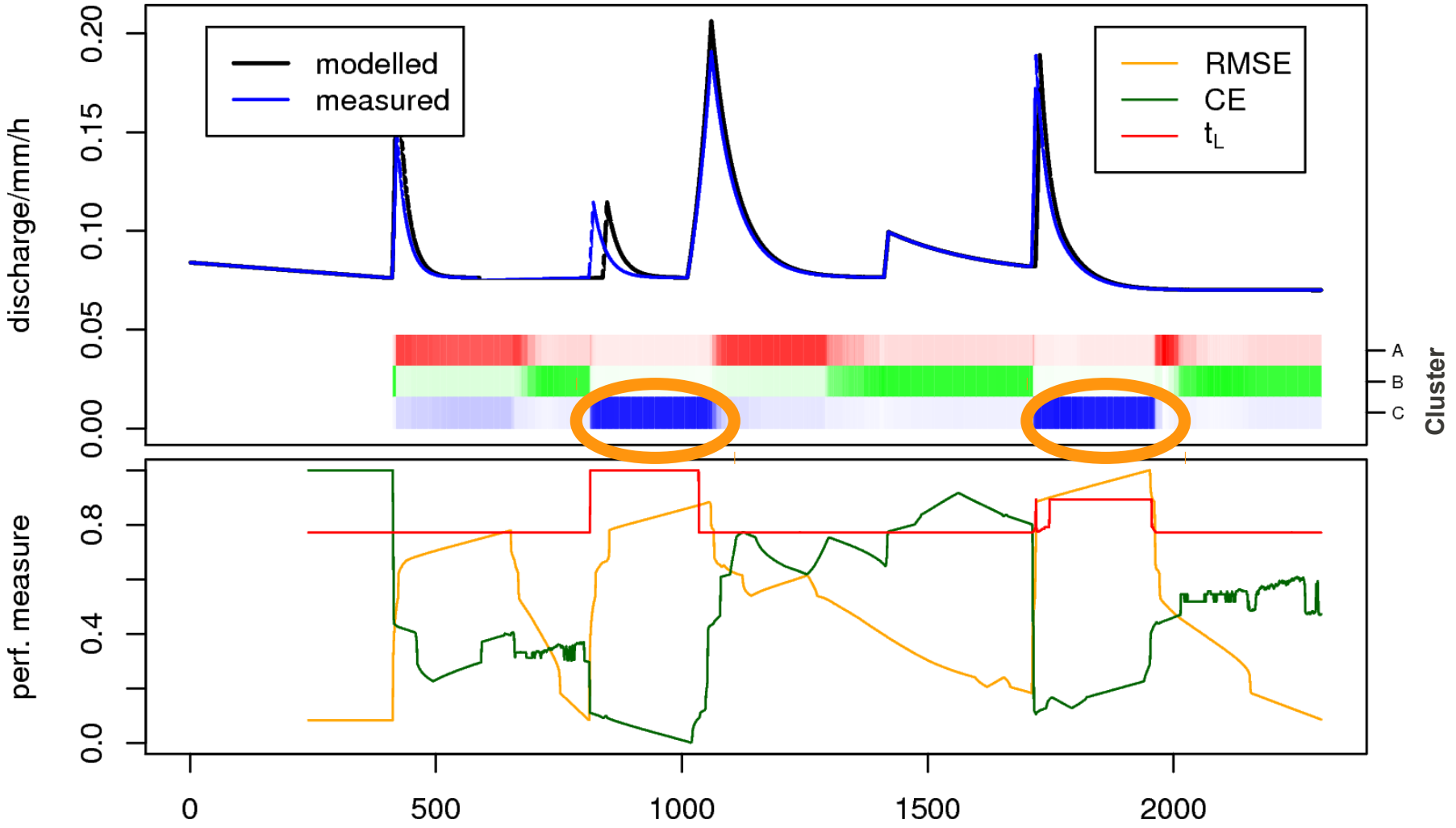
Pattern detection with SOM and fuzzy clustering



Reusser et al., HESS 2009

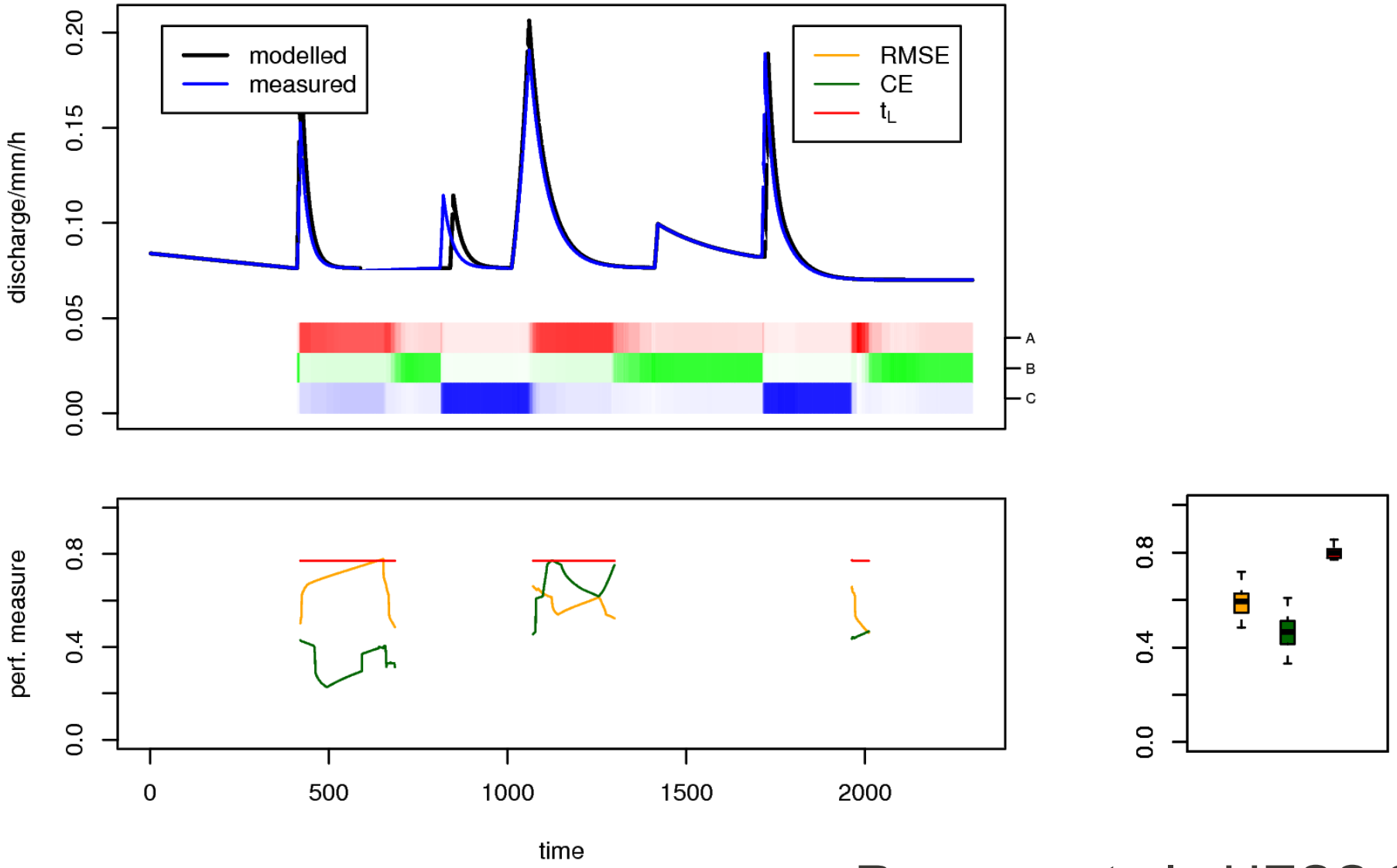
TIGER method

Pattern detection with SOM and fuzzy clustering



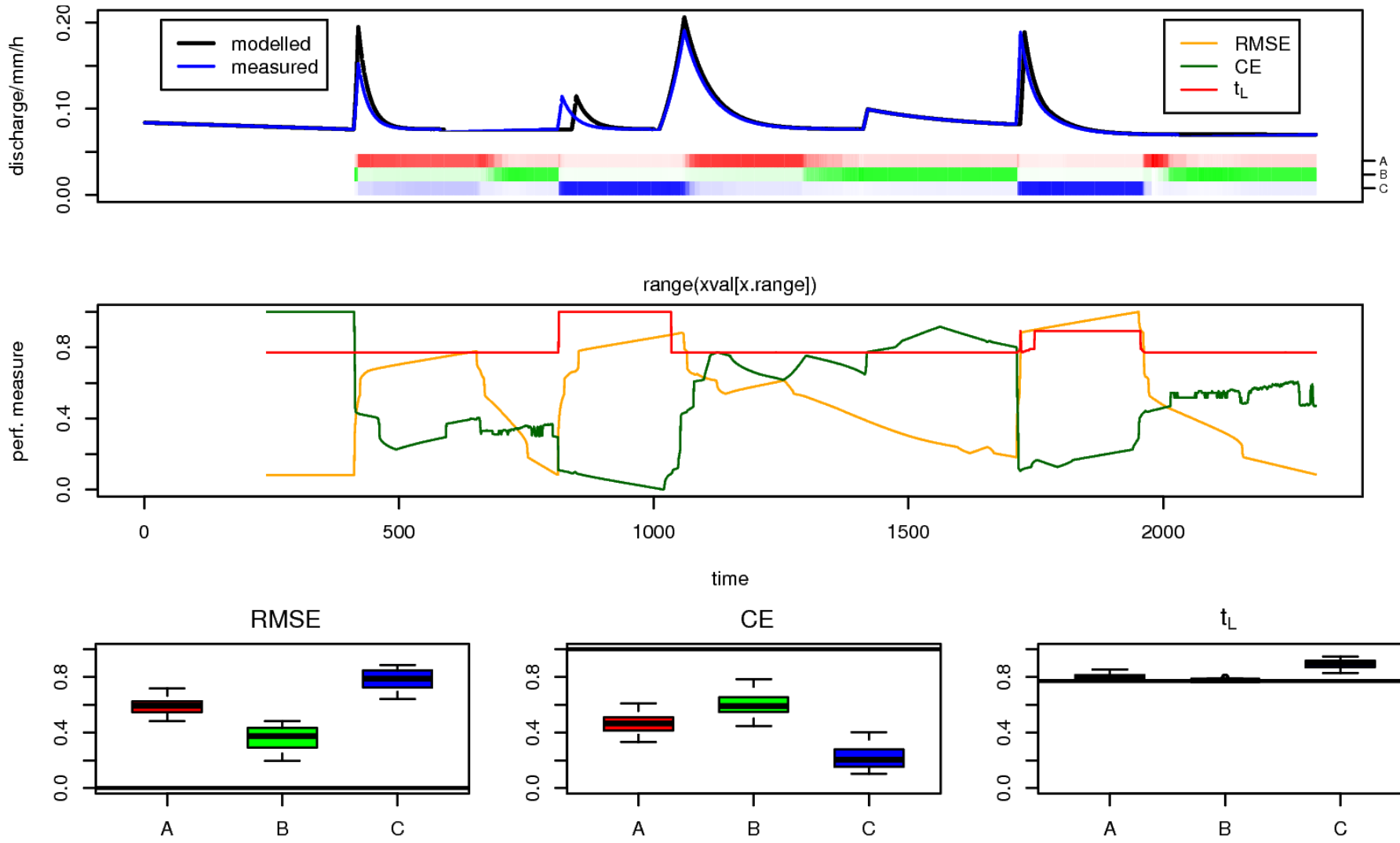
Reusser et al., HESS 2009

Cluster characterization 1



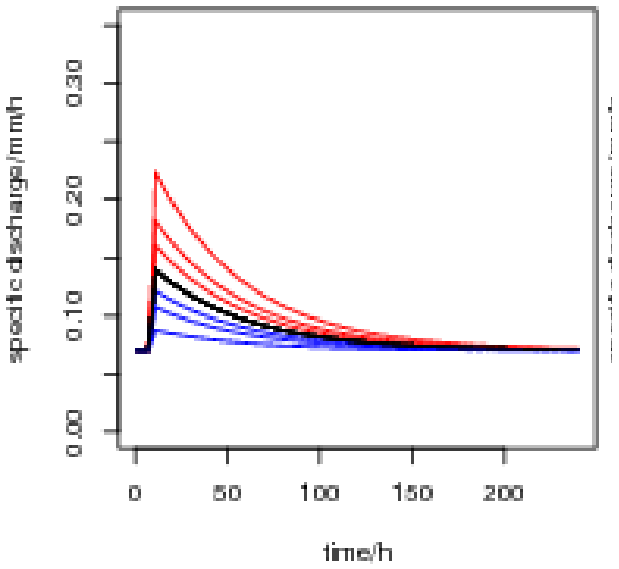
Reusser et al., HESS 2009

Cluster characterization 1



Reusser et al., HESS 2009

Cluster characterization 2



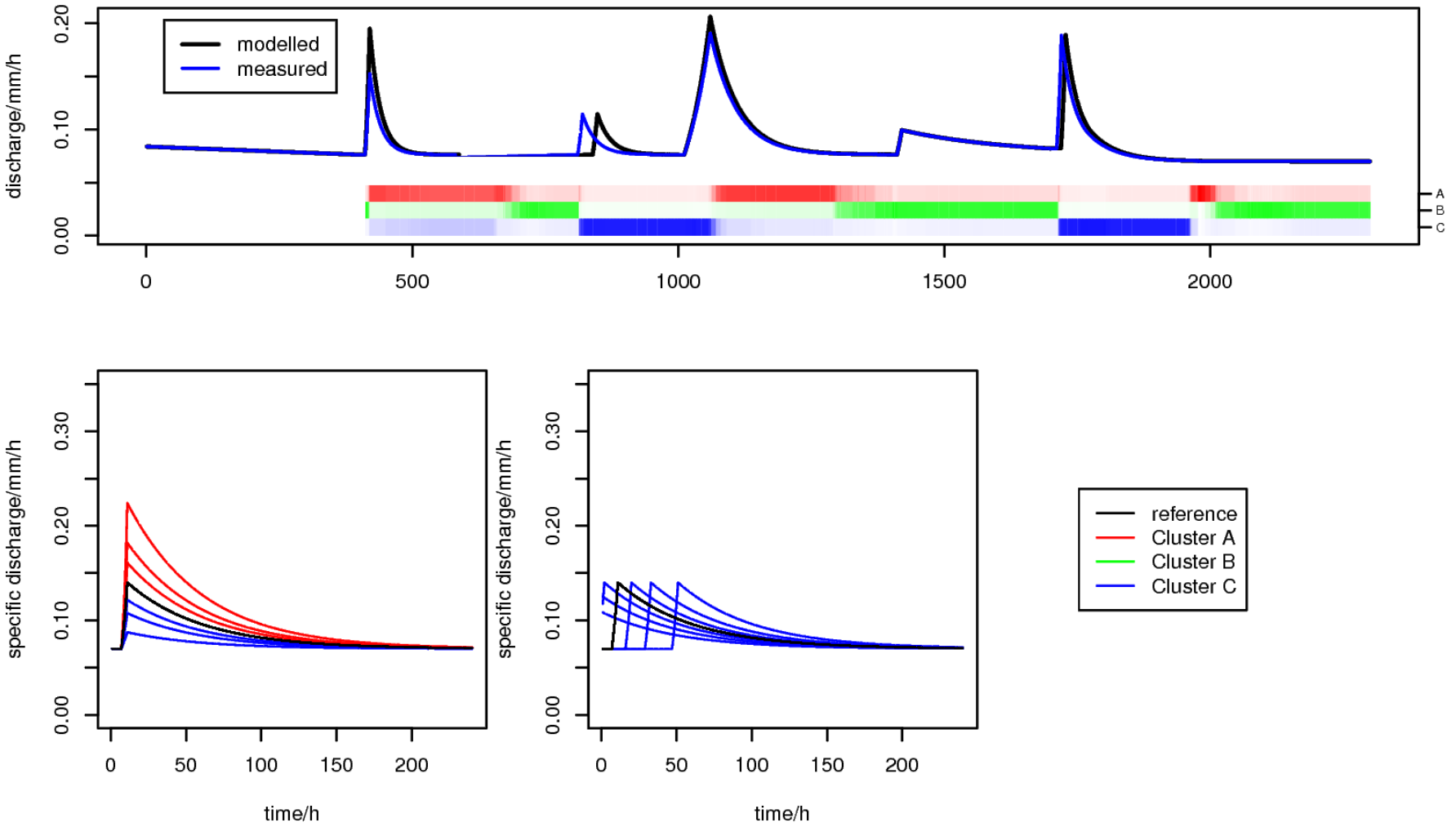
Synthetic peaks for window size

Performance finger print for synthetic peaks

Best matching cluster

Reusser et al., HESS 2009

Cluster characterization 2



Reusser et al., HESS 2009

Poster: Synthetic peak errors

Poster A188

Typology of hydrological metrics based on synthetic errors for single peak events

Dominik E. Reusser
(University of Potsdam, Institute of Earth and Environmental Sciences, Potsdam Institute for Climate Impact Research)
Erwin Zehe
(KIT Karlsruhe, Chari for Hydrology)

Evaluation of metrics

A careful evaluation of metrics allows the selection of the appropriate metric for a given application
Two approaches:
1) benchmark values for simple reference models (Schaefli and Gupta 2007)
2) synthetic errors:
- synthetic errors based on modified observed events (Krause et al. 2005 and Dawson et al. 2007)
- errors for spatial forecasts (Cloke and Pappenberger 2008)

Drawback of modified observed events: only a limited number of modifiactions possible with acceptable effort

Synthetic errors for single peak events

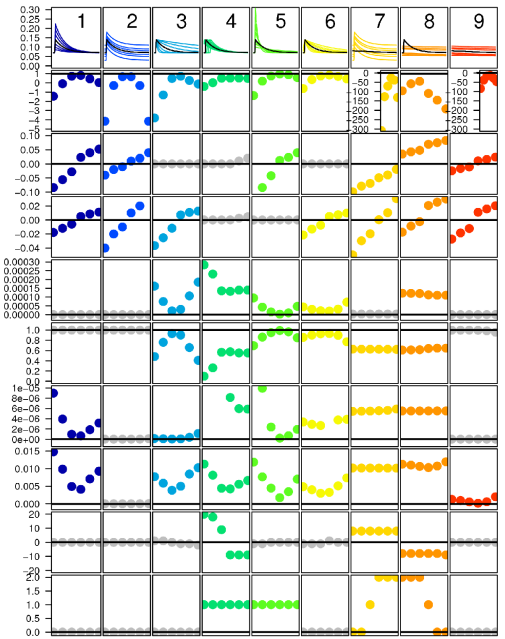
We test a large number of metrics with synthetic errors to learn about their specificity. The advantage compared to modified observed events: more control!

$$Q(t)=\begin{cases} Q_b & t<t_0 \\ Q_b + e^{(t-t_0)\cdot k_r} & t_0<=t<t_{max} \\ Q_b\,(1+e^{(-k_{max}\cdot(t_0-t_{max}))}-e^{((t-t_{max})\cdot k_r)}) & t_{max}<=t \end{cases}$$

k_r recession constant (negative)
k_r constant for rise phase
Q_b base flow
t time
t_0 event starting time
t_{max} peak time

We group the metrics according to their specificity into Error Response Groups (ERG).

ERG	Description of the group	List of metrcs
1 NSCE	Measures react to all peak errors.	many*
2 PDIFF	Insensitive to error types do not change the maximum of the peak.	PDIFF, PEP
3 ME	Show no or only little sensitivity peak errors with correct total volume.	ME, MRE, MSRE, RVE, t_test
4 SMSE	Insensitive to errors that can be corrected with rescaling.	SMSE
5 Rsqr	Insensitive to errors that can be corrected with linear transformation.	Rsqr, RSMSE, RMSOE, MAOE, MSOE, SMAE, LCS
6 MSDE	Insensitive to errors related to shifts and shifts during the late recession phase.	MSDE, CMSE
7 SMSLE	Not sensitive for the shift.	CMAE, RCMSE, SMSLE, slope error
8 lag time	Only sensitive to the lag time and the missing/false peak.	lag time direction error
9 NSC	Only sensitive if the sign of the error changes along the time series.	NSC



Top row: Discharge of the synthetic peaks against time. Others: Metric value against degree of error starting from left with strong understimation via no error in the middle to strong overestimation on the right.

* MSE, RMSE, IRMSE, R4MS4E, CE, PI, AME, MAE, RAE, MARE, MdAPE, loAd, MSDE, NSC, MAPE, MALE, MSLE, RMSLE, MAGE, RMSGE, SMALE, SMAGE, RSMSGE, RSMSLE, recession error

Discussion

- Synthetic errors are useful for the characterization of metrics
- MSE - based metrics behave similarly
- The error response groups may guide the selection of metrics depending on the required sensitivity

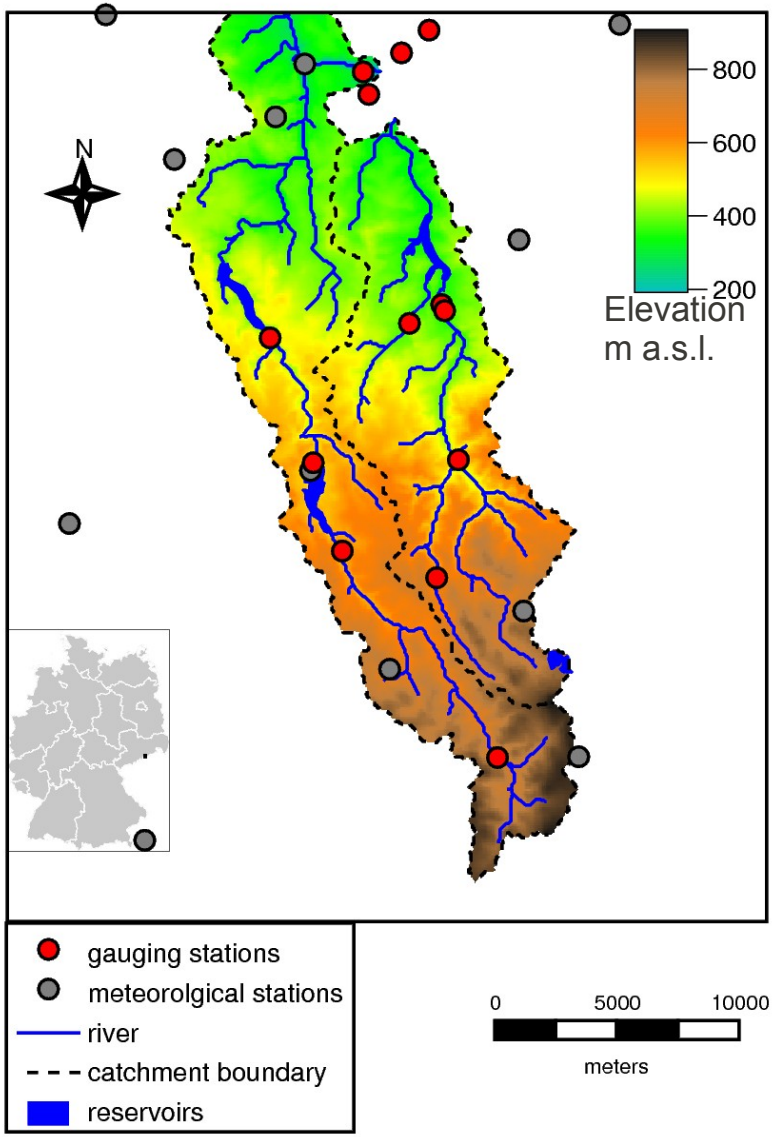
Available as R packages (free software)

- tiger package provides functions to create synthetic peak errors and some metrics
- qualV package implements the remaining metrics

References

Cloke, Hannah, and Florian Pappenberger. "Evaluating forecasts of extreme events for hydrological applications: an approach for screening unfamiliar performance measures." Meteorological Application 15 (2008): 181-197.
Dawson, C W, R J Abrahart, and L M See. "HydroTest: A web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts." Environmental Modelling & Software 22, no. 7 (2007): 1034-1052.
Jachner, Stefanie, K G van Den Boogaart, and T Petzoldt. "Statistical Methods for the Qualitative Assessment of Dynamic Models with Time Delay (R Package qualV)." Journal of Statistical Software 22 (2007): 1-30.
Krause, P, D P Boyle, and F Bäse. "Comparison of different efficiency criteria for hydrological model assessment." Advances in Geosciences 5 (2005): 89-97. http://www.adv-geosci.net/5/89/2005/.
Reusser, Dominik E., Theresa Blume, Bettina Schaefli, and Erwin Zehe. "Analysing the temporal dynamics of model performance for hydrological models." Hydrology And Earth System Sciences 13 (2009): 999-1018.
Schaefli, Bettina, and Hoshin V Gupta. "Do Nash values have value?" Hydrological Processes 21 (2007): 2075-2080.

Weißeritz catchment



Rote and Wilde
Weißeritz: 400km²

Forests
and
agriculture

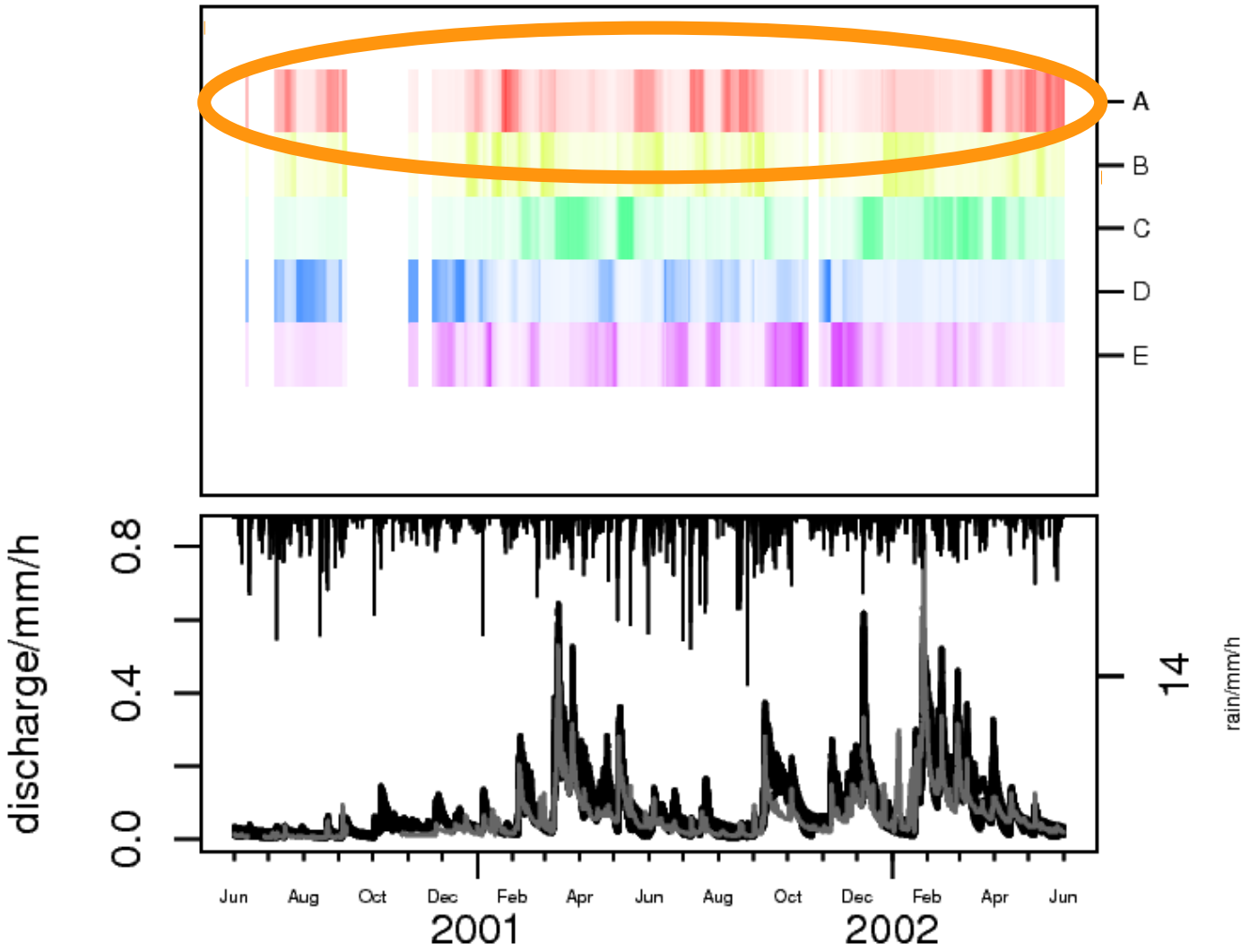
3 Reservoirs:
Drinking water

Major flooding
in 2002

Model

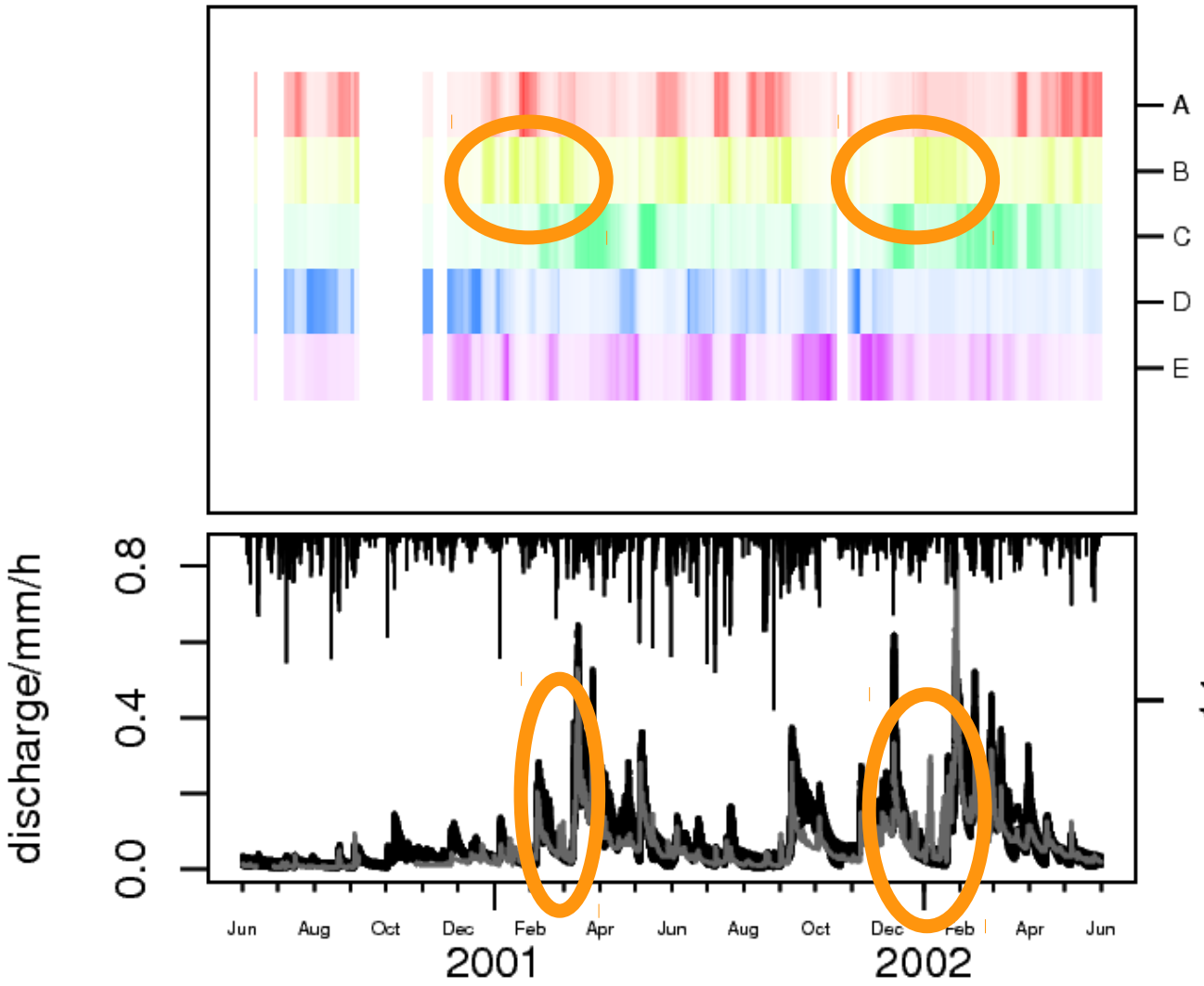
- WaSiM-ETH (TOPMODEL approach)
- 11 parameters:
  - Direct flow recession
  - Interflow recession and storage size
  - Baseflow recharge and discharge
  - TOPMODEL regionalization factor m
  - Snow: Degree day factor, temperatures, fraction for direct flow
  - Infiltration rate threshold
- 2 years of simulation (2000-2002)

TIGER Results



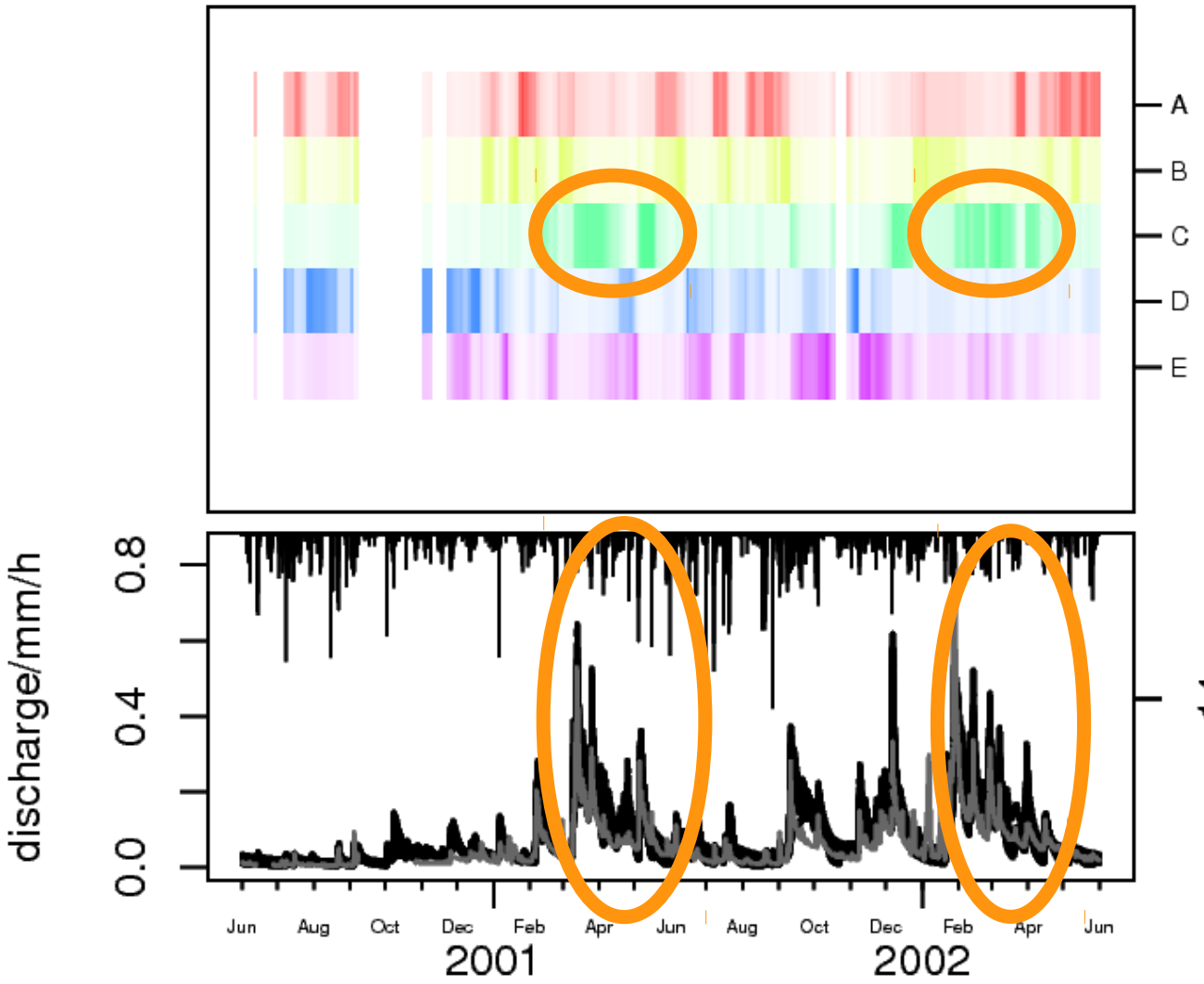
Reusser et al., HESS 2009

TIGER Results



Reusser et al., HESS 2009

TIGER Results

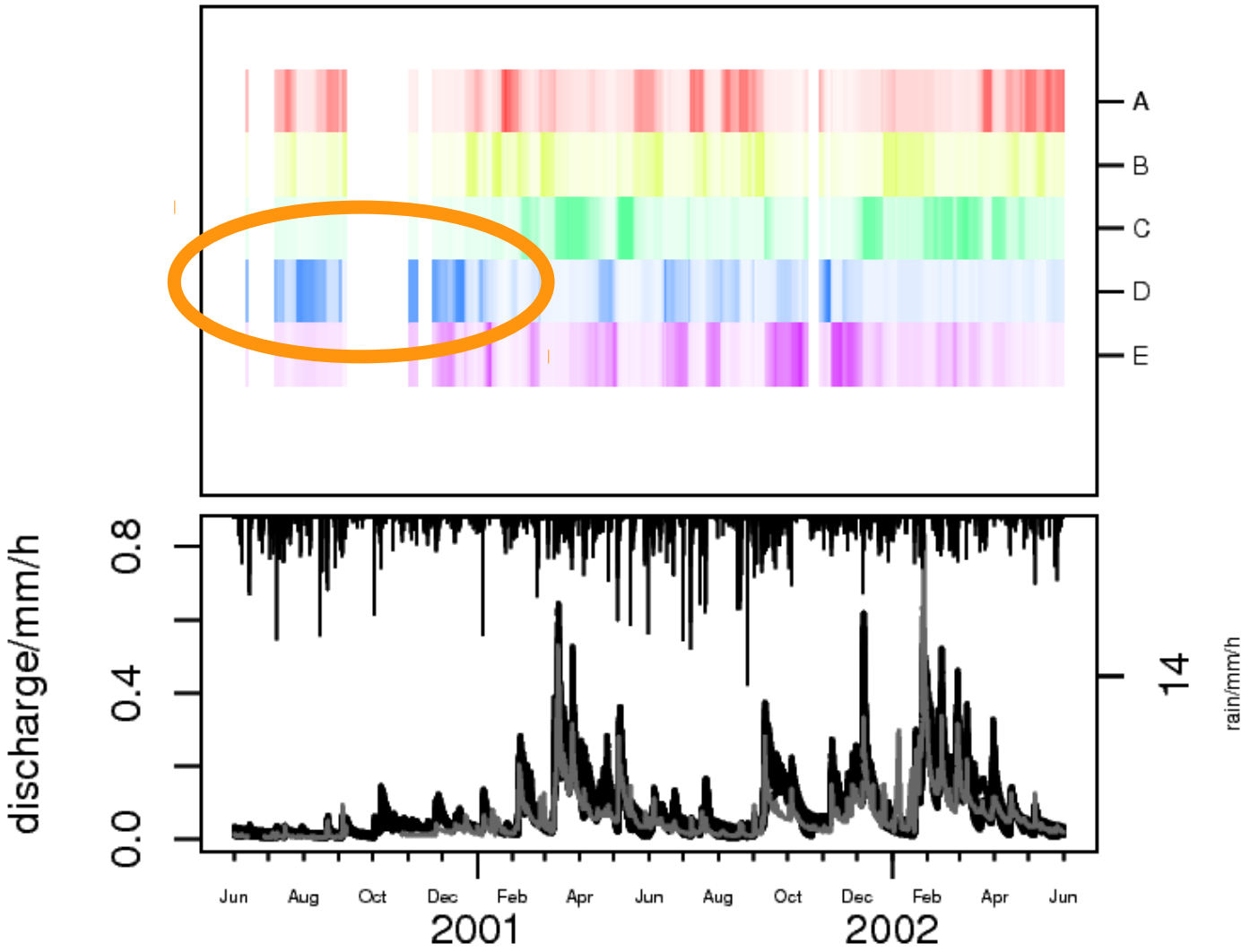


14

rain/

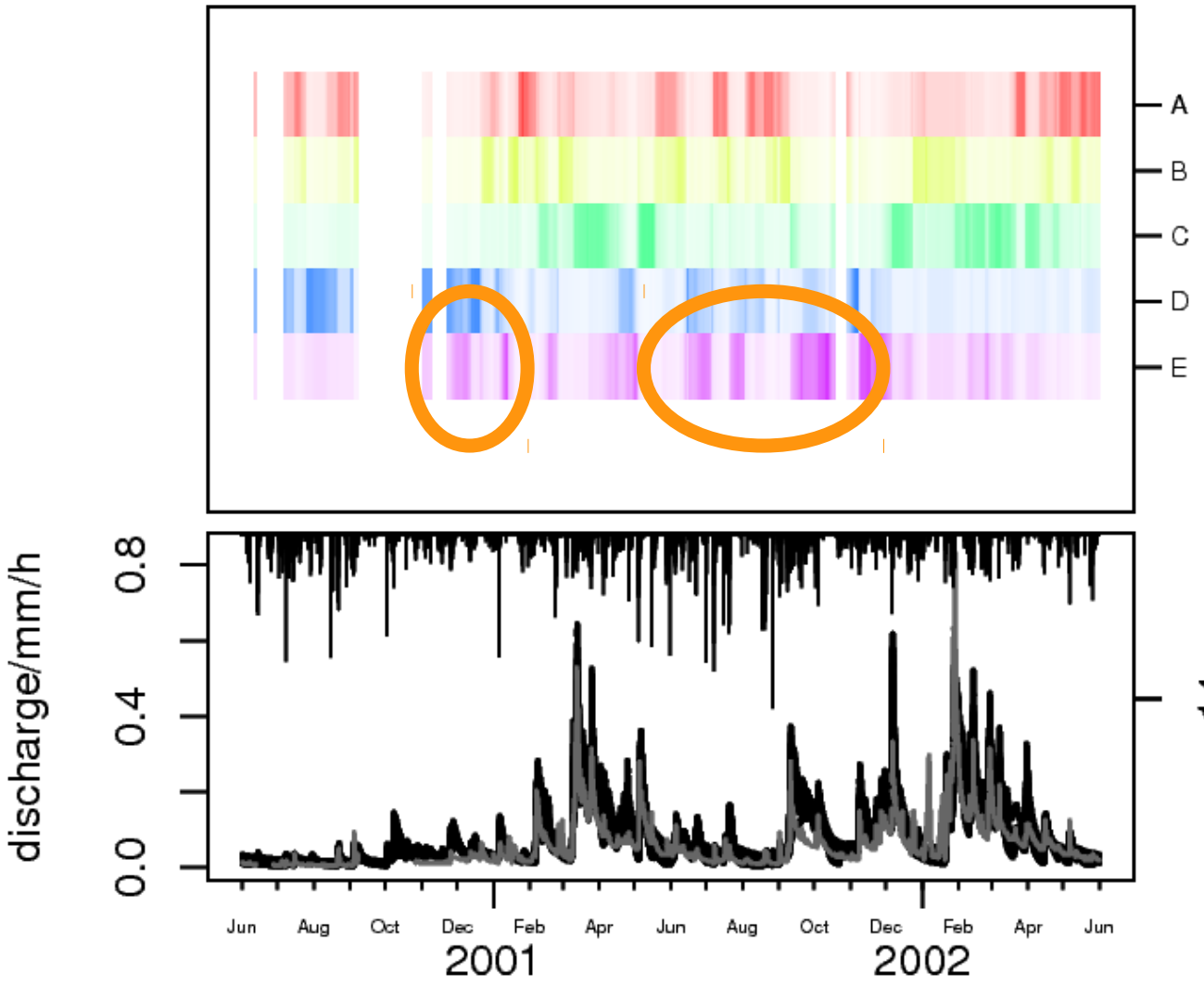
Reusser et al., HESS 2009

TIGER Results



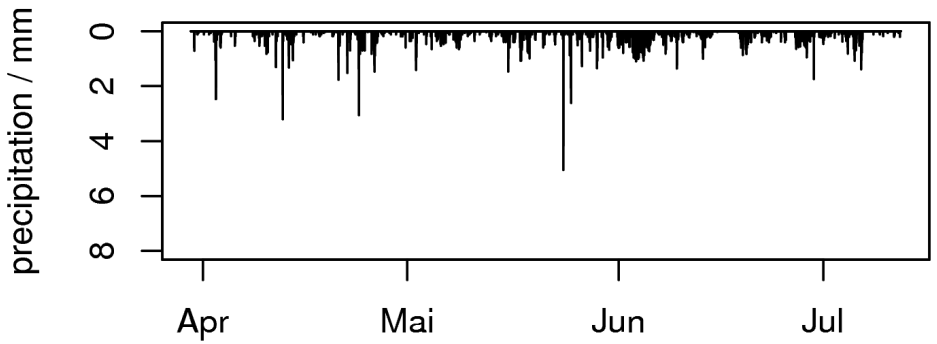
Reusser et al., HESS 2009

TIGER Results



Reusser et al., HESS 2009

Questions of Interest

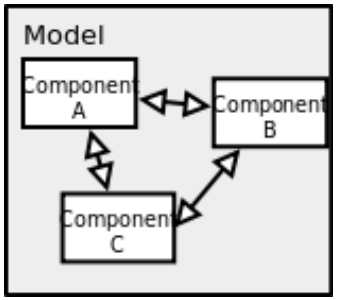


1) When ?

2) Kind of deviation ?

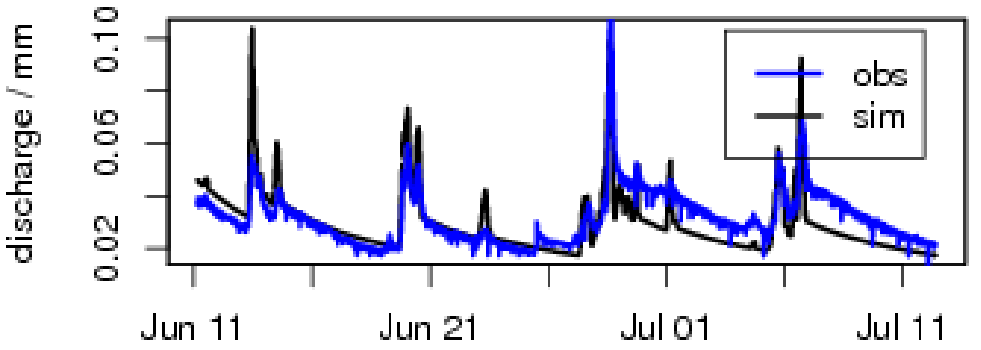
Parameters

Input

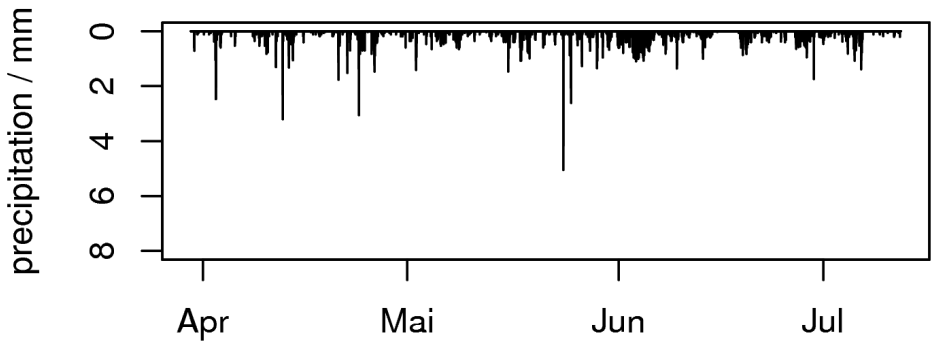


Output

3) Which component ?



Questions of Interest

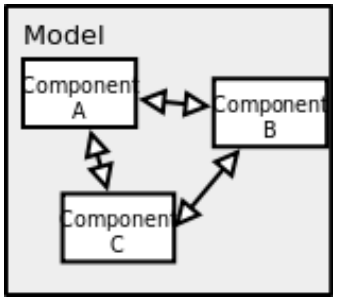


1) When ?

2) Kind of deviation ?

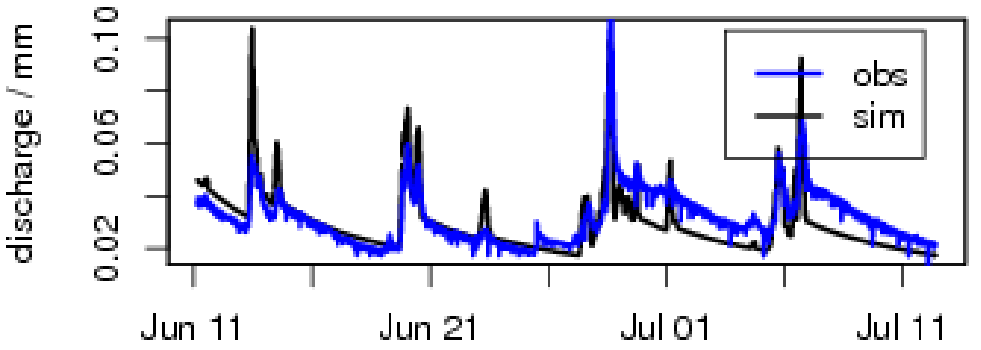
Parameters

Input



Output

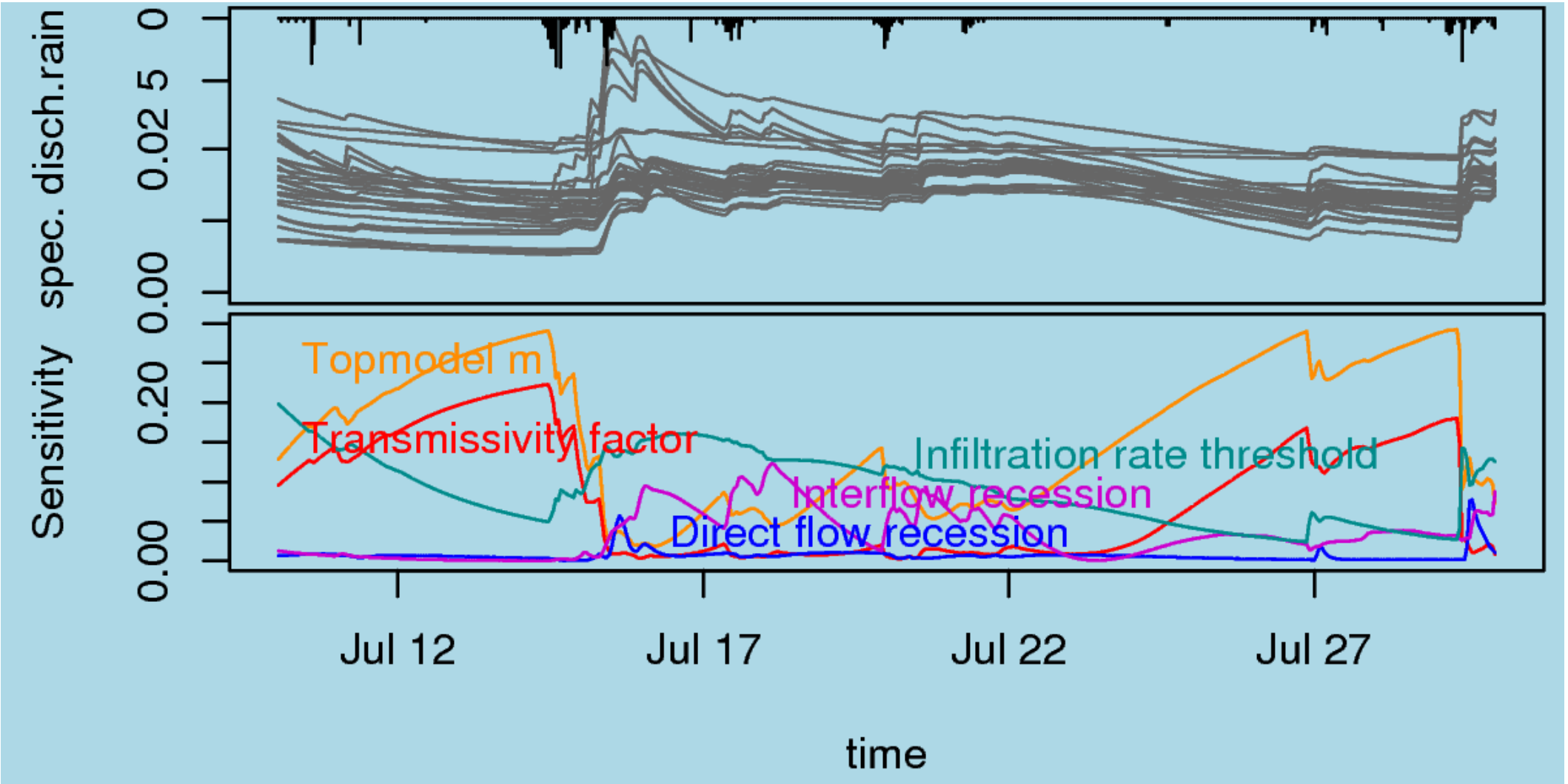
3) Which component ?



Sensitivity analysis

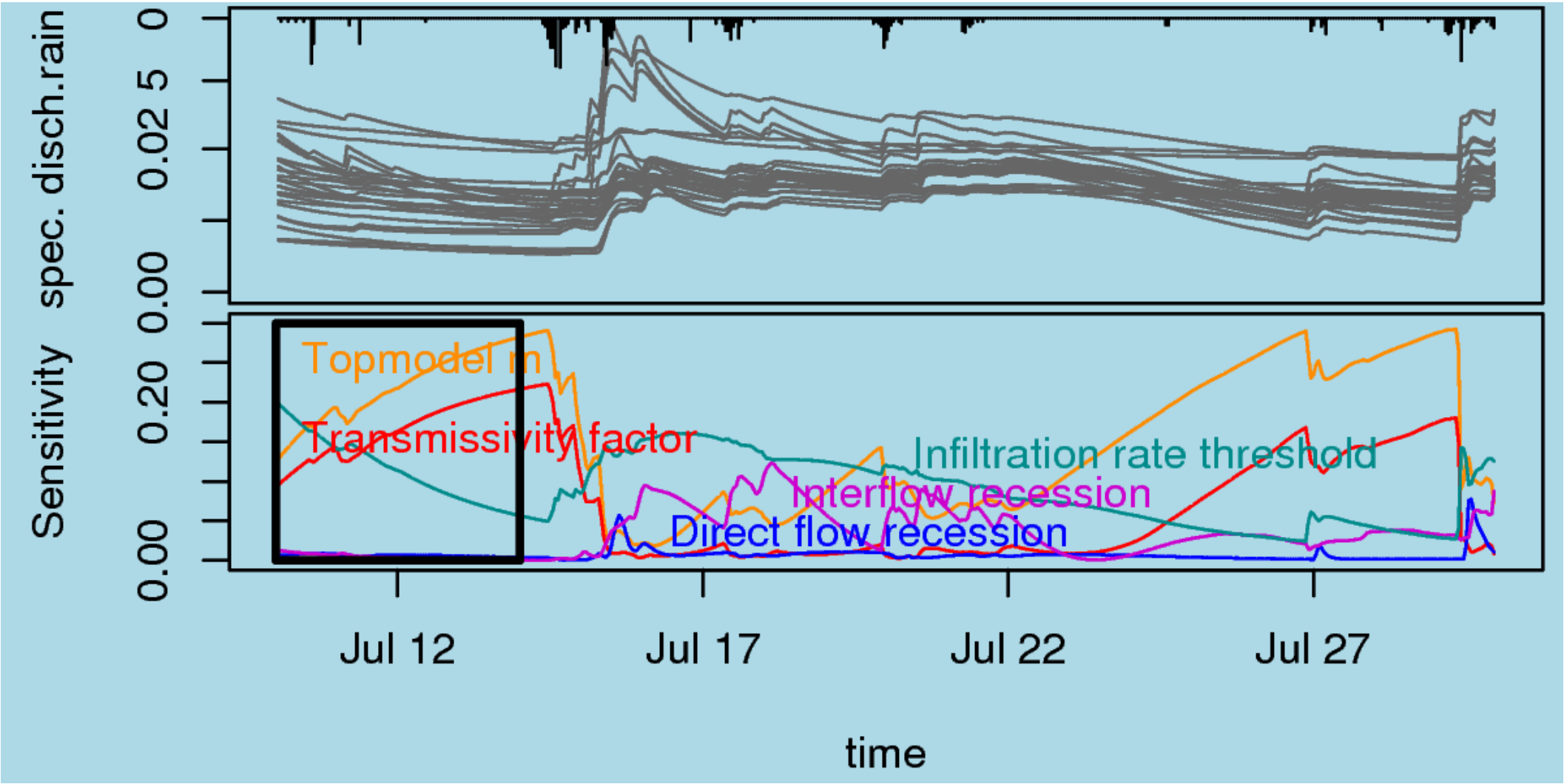
- Fourier Amplitude Sensitivity Test (FAST) (Cukier et al., 1977)
- results equivalent to Sobol's method (Sobol, 1994)
- equivalence shown (Saltelli and Bolado 1998)
  - FAST computationally much more efficient.
  - FAST can not compute interaction terms.

Temporal dynamics of parameter sensitivity: TEDPAS



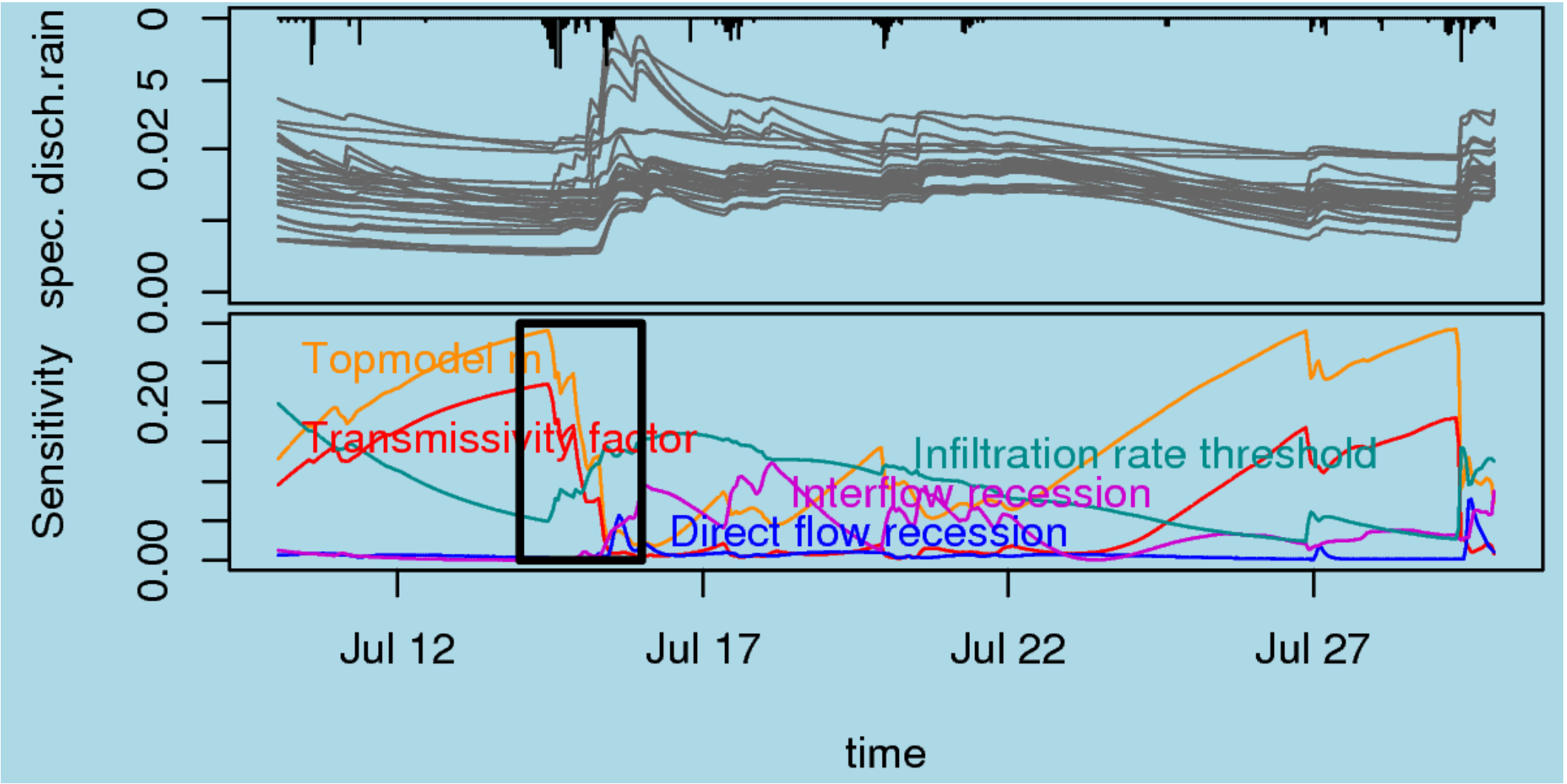
Reusser et al. WRR (submitted)

Temporal dynamics of parameter sensitivity: TEDPAS



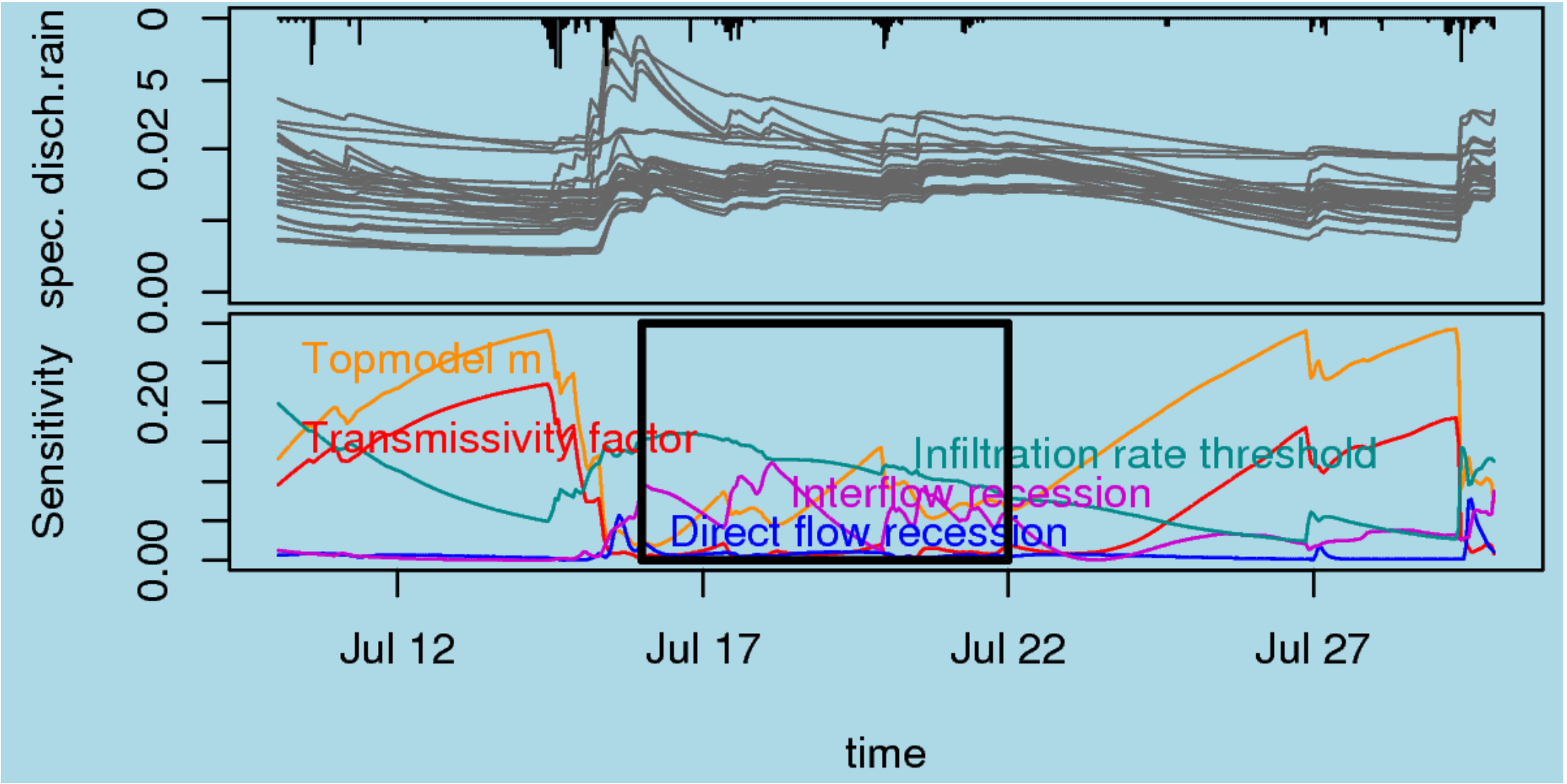
Reusser et al. WRR (submitted)

Temporal dynamics of parameter sensitivity: TEDPAS



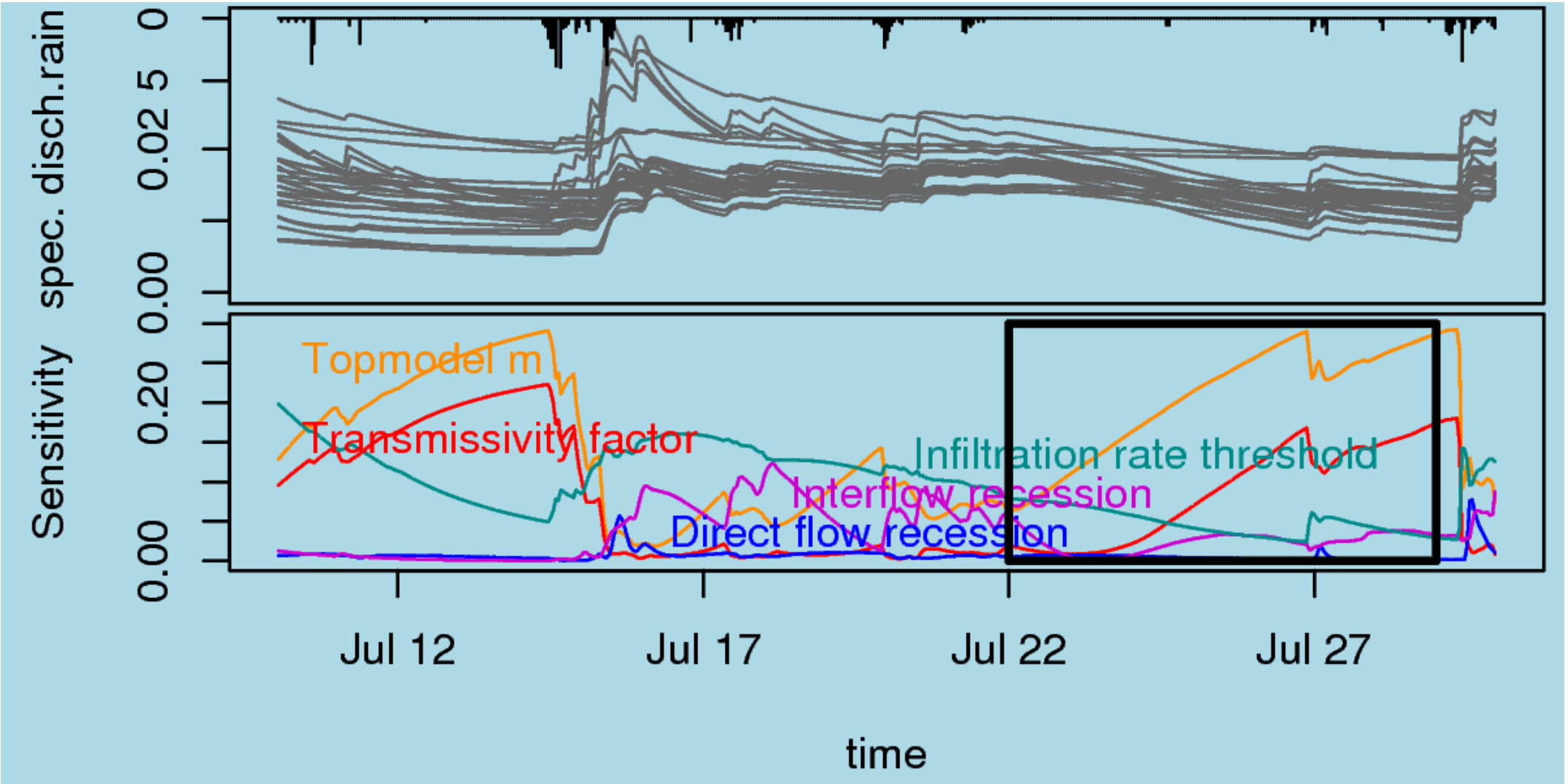
Reusser et al. WRR (submitted)

Temporal dynamics of parameter sensitivity: TEDPAS



Reusser et al. WRR (submitted)

Temporal dynamics of parameter sensitivity: TEDPAS



Reusser et al. WRR (submitted)

Summary

- Finger prints of model performance (TIGER)
- Parameter sensitivity (TEDPAS)
- Combination of temporal diagnostic methods

  ⇒ model and catchment comparison

- Results presented for the Weißeritz case study

Curious? Test it yourself!

Packages available as open source software in R

Reusser et al. HESS 2009, Reusser et al. WRR (submitted), Reusser and Zehe WRR (submitted)

Dominik Reusser, PIK Potsdam
Erwin Zehe, KIT Karlsruhe

reusser@pik-potsdam.de

Poster A188

Discussions with Bettina Schaefli and Theresa Blume were helpful during the initial phase of this research

Funding:
BMBF Germany
OPAQUE-Project within the RIMAX-Förderschwerpunkt