

RAT (Robustness Assessment Test): a straightforward evaluation of hydrological model robustness to a changing climate

V. Andréassian¹, L. Santos¹, T.O. Sonnenborg², A. de Lavenne¹, G. Lindström³, P. Nicolle⁴ & G. Thirel¹

¹ INRAE, UR HYCAR, Antony, France

² GEUS, Copenhagen, Denmark

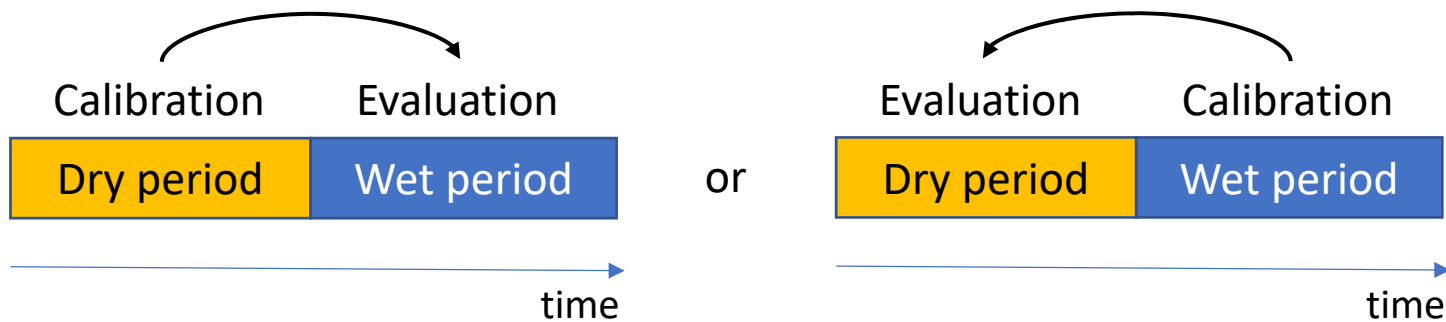
³ SMHI, Norrköping, Sweden

⁴ Université Gustave Eiffel, Bouguenais, France



➤ Context

- Hydrological models: increasingly used under changing climate conditions
- Are hydrological models robust? If not: where? why?
- Traditional robustness assesment: Differential Split Sample Test (DSST)



- But what about models that are calibrated/parameterized manually?
- **Aim of the work:**
 - replace/complement SST by avoiding any calibration step
 - facilitate model comparison

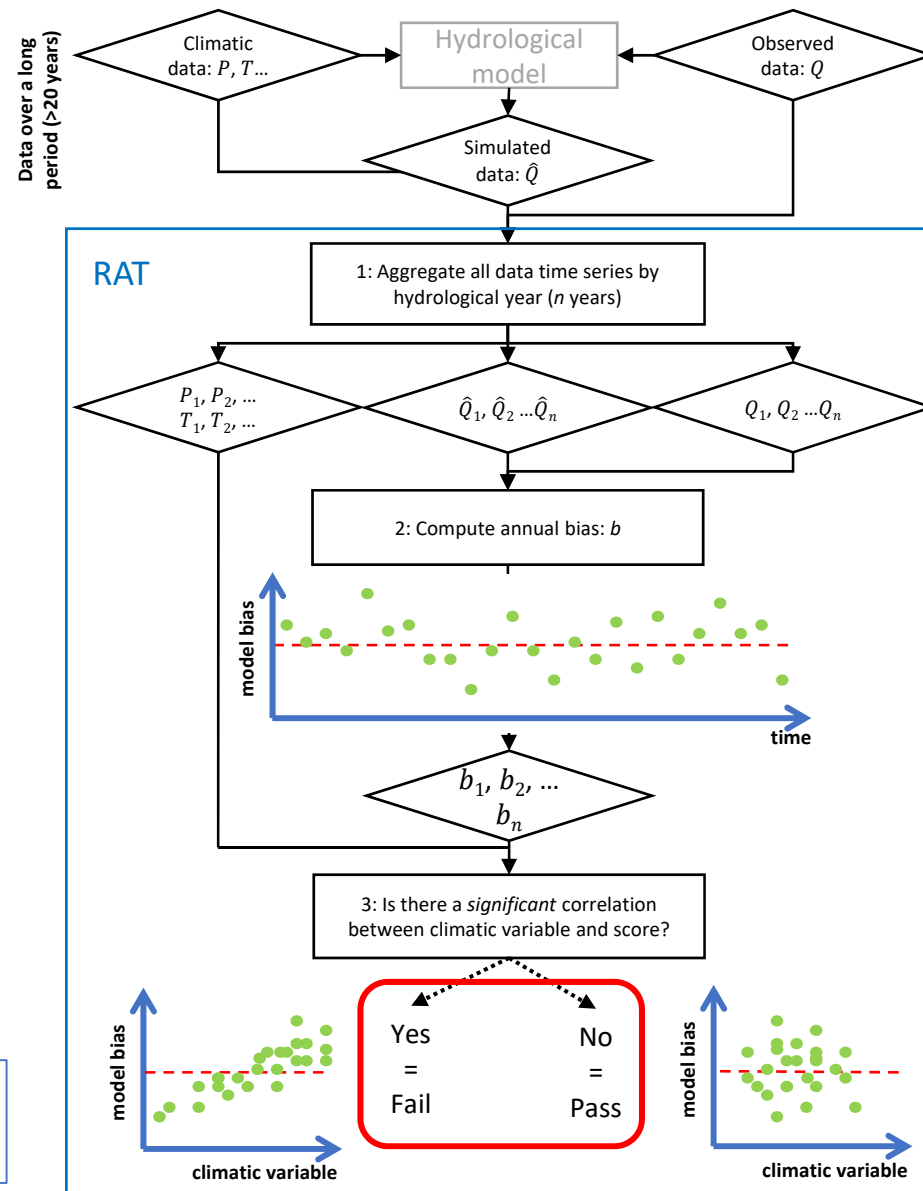
➤ The Robustness Assessment Test (RAT)

Aim: detection of unwanted correlations between bias and a given climatic variable

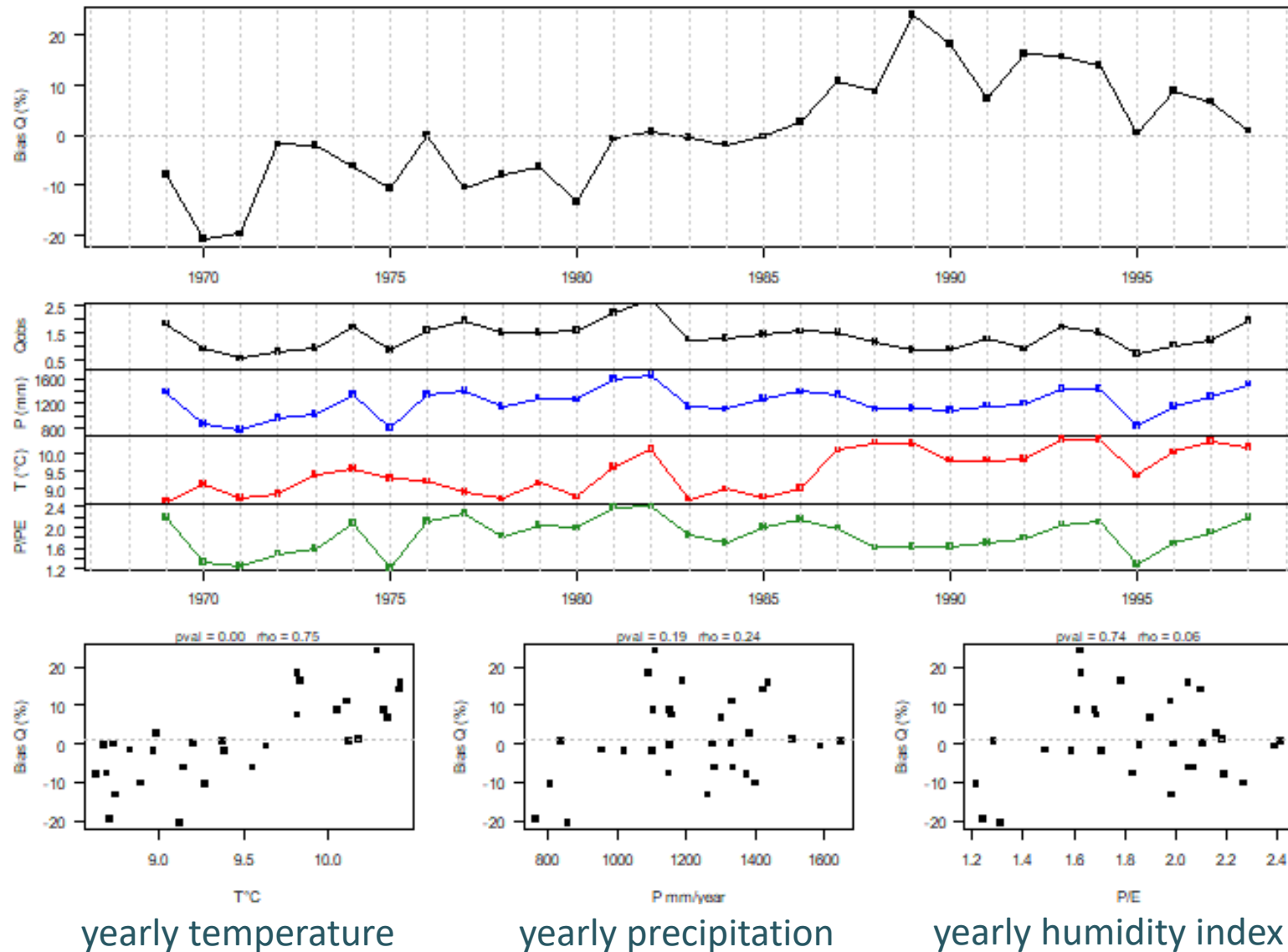
- **Step 1:** split variable by hydrological year
- **Step 2:** compute bias by hydrological year
- **Step 3:** test correlation between bias and variable

A correlation is a signal for a robustness problem

See our paper in discussion in HESSD:
<https://doi.org/10.5194/hess-2021-147>



➤ An example with 3 climatic variables



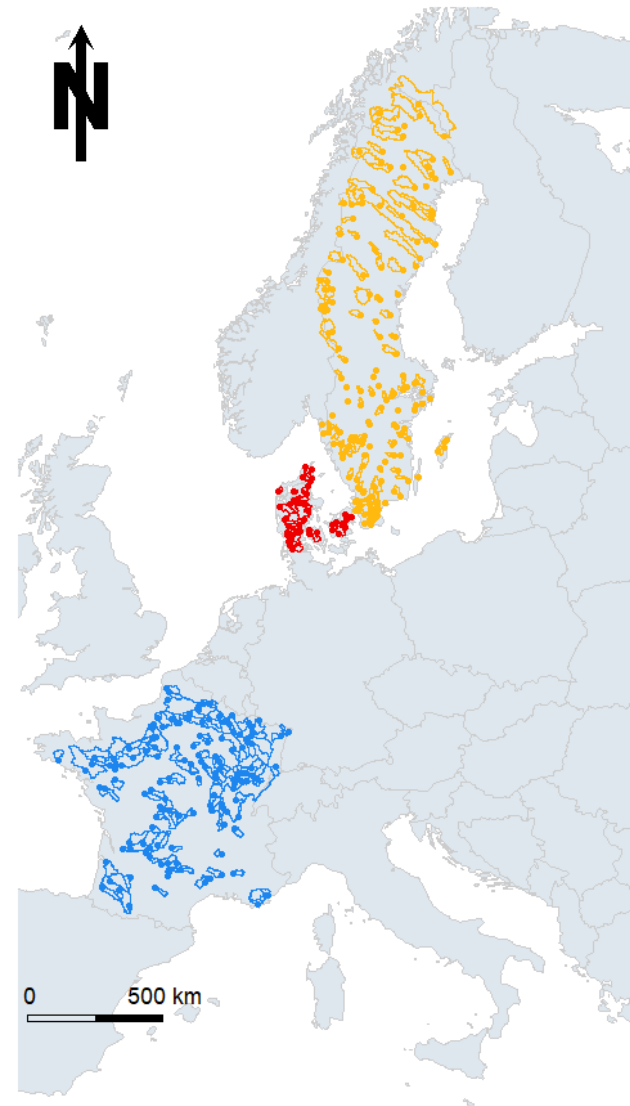
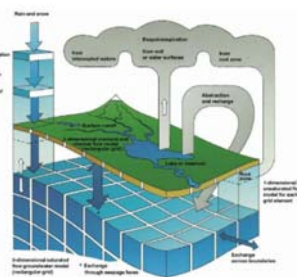
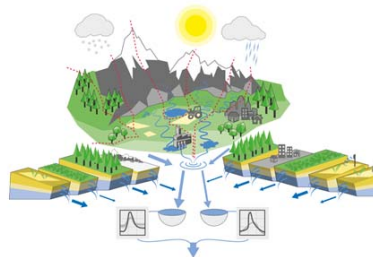
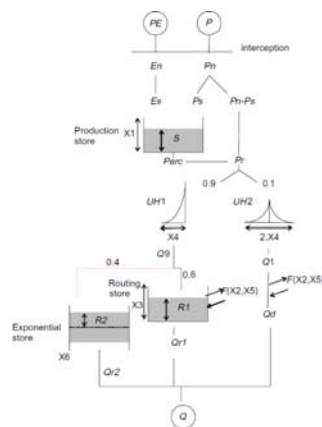
➤ Test on a large catchment set

352 catchments:

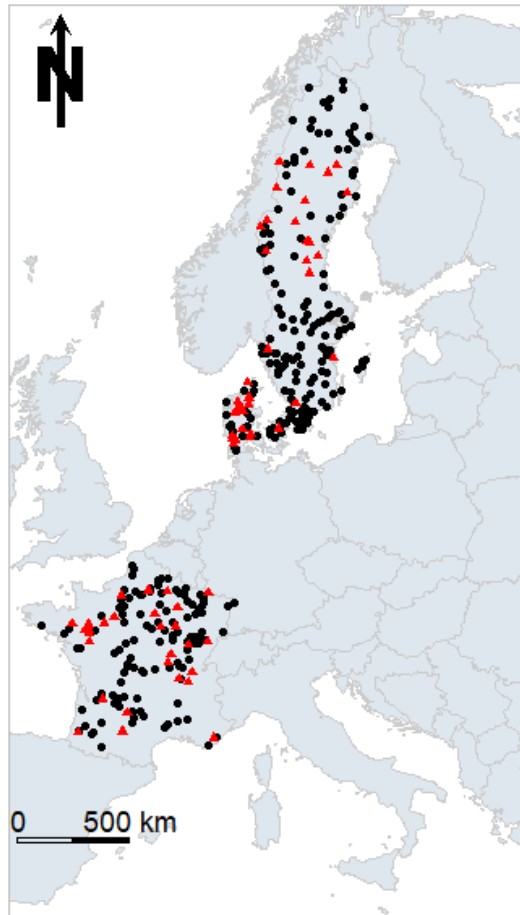
- 163 in Sweden
- 43 in Denmark
- 146 in France

3 models compared:

- GR6J: lumped
- S-HYPE: semi-distributed
- MIKE-SHE: distributed



➤ GR6J over the entire catchment set



64 reactive catchments

Factors correlated to GR6J's lack of robustness

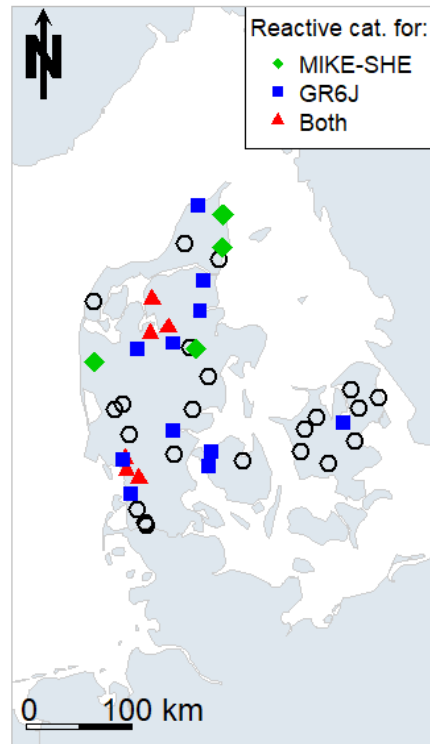
- Low-flows ↑
- High-flows =
- Flow variability =
- Flashiness ↓



Specific robustness issues identified for: slow catchments with high values of low-flows

- GR6J issue on groundwater simulation?
- Due to water balance adjustment?

➤ MIKE-SHE vs GR6J in Denmark



10 reactive catchments for MIKE-SHE

17 reactive catchments for GR6J

Factors for GR6J lack of robustness

- Low-flows 
- High-flows 
- Flow variability 
- Flashiness 

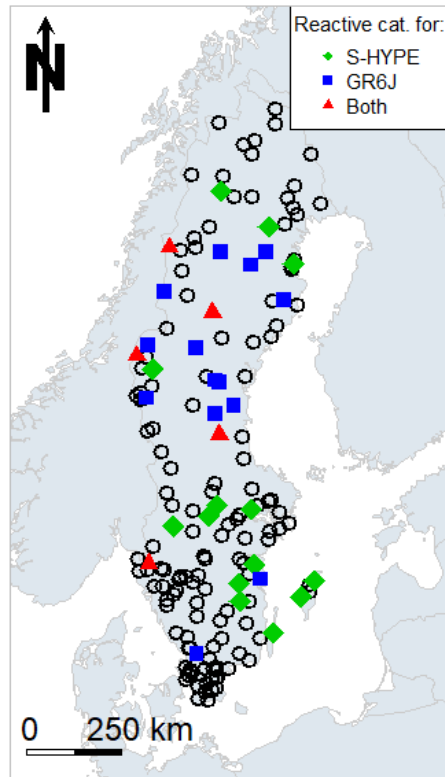
Factors for MIKE-SHE lack of robustness

- Low-flows 
- High-flows 
- Flow variability 
- Flashiness 

MIKE-SHE has less reactive catchments

Both models have similar factors correlated with robustness issues

➤ S-HYPE vs GR6J in Sweden



19 reactive catchments for S-HYPE

19 reactive catchments for GR6J

Same number of reactive catchments but no explicative factors identified for S-HYPE

Factors for GR6J lack of robustness

- Low-flows 
- High-flows 
- Flow variability 
- Flashiness 

Factors for S-HYPE lack of robustness

- Low-flows 
- High-flows 
- Flow variability 
- Flashiness 

➤ Conclusion



- RAT: a simple robustness test, with similar results to DSST/GSST, but easier to apply (does not require multiple calibrations)
- All 3 models showed robustness issues on a significant number of catchments
- The factors associated with these robustness issues are model-specific:
 - MIKE-SHE and GR6J seems to have similar issues in Denmark
 - S-HYPE and GR6J react on very different catchments in Sweden

More details at HS2.2.1 session on Wednesday 28 April 2020 at 14:01

See also Nicolle et al. paper, under review in HESS (<https://doi.org/10.5194/hess-2021-147>)