RheinBlick2050

Hydrological model intercomparison, suitability and robustness studies

K. Görgen
Project Coordinator
CRP-GL / EVA / GEOSAT
Belvaux, Luxembourg
http://www.lippmann.lu
Centre de Recherche Public Gabriel Lippmann

O. de Keizer
Deltares
Enabling Delta Life

J. Beersma
Koninklijk Nederlands Meteorologisch Instituut
Ministerie van Verkeer en Waterstaat

H. Buiteveld
R. Lammersen
Rijkswaterstaat
Ministerie van Verkeer en Waterstaat

P. Krahe
M. Carambia
E. Nilson
bfg Bundesanstalt für Gewässerkunde

C. Perrin
Cemagref
Sciences, eaux & territoires

G. Brahmer
HLUG

D. Volken
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Bundesamt für Umwelt BAFU

Structure of the presentation

CHR
KHR

- How reliable are hydrological models for climate change (CC) impact studies?
- Objectives within RheinBlick2050
- Hydrological models and data

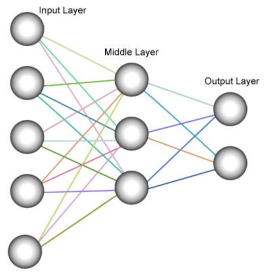
- Testing the reliability of hydrological models
- Testing the stationarity hypothesis
- Comparing sources of uncertainty
- HBV134 reliability on reference period

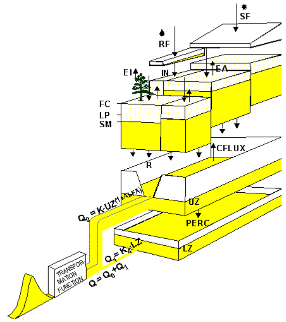
- Main conclusions

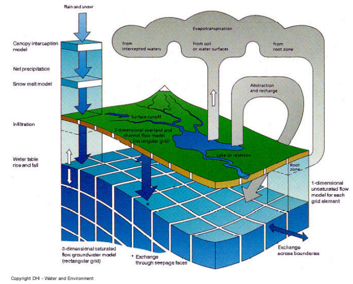
How reliable are hydrological models for CC impact studies?

- A primary goal of hydrological modelling is to **simulate the transformation of precipitations into streamflow** at the catchment scale

- There is a large number of existing hydrological models differing in:
  - structure and parameterization
  - space and time scales
  - data requirements
  - etc.







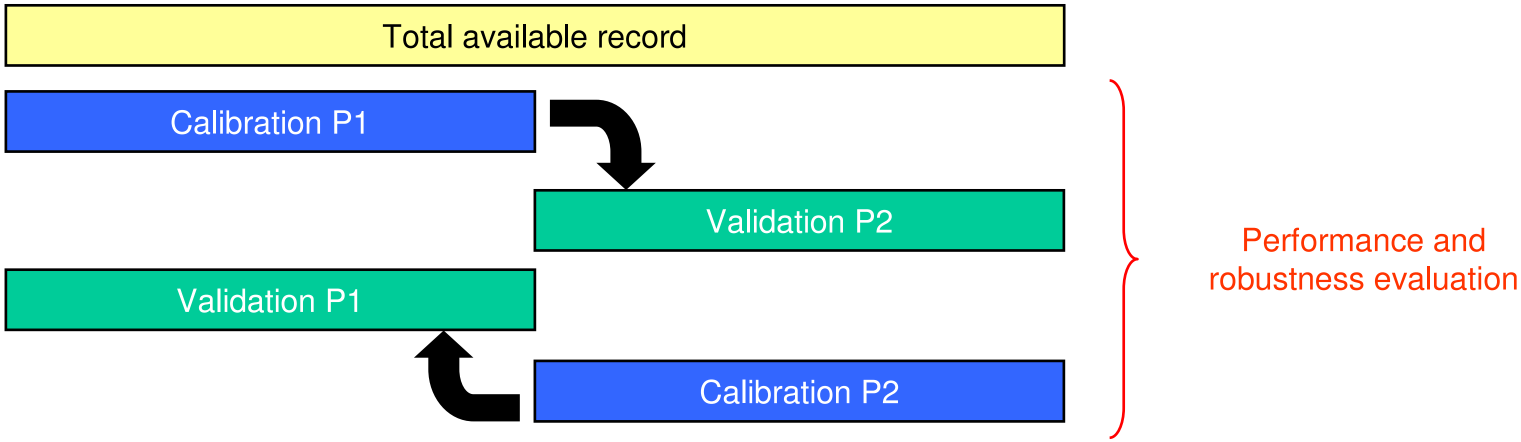
- Hydrological models are essential tools in water related studies… **but**:
  - they are imperfect
  - they introduce uncertainties in results

How reliable are hydrological models for CC impact studies?

- Model selection depends on:
  - modelling objectives
  - data availability
  - user's experience and **confidence in the model**

- Quantification of the confidence in a model by evaluating its ability to **simulate catchment behaviour and answer modelling objectives**

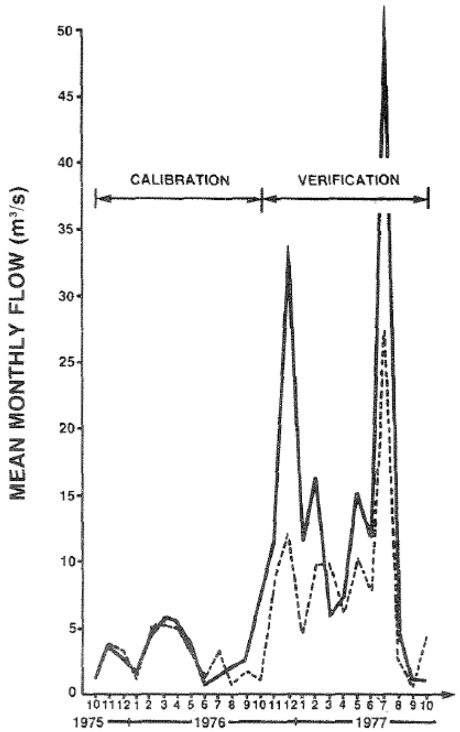
- Klemes (1986) proposed the ***"split sample test"*** scheme to evaluate model transposability in time



How reliable are hydrological models for CC impact studies?

- Two options in the evaluation scheme with tests under:
  1. **stationary conditions** (*split sample test*): periods with similar climatic conditions
  2. **non stationary conditions** (*differential split sample test*): periods with contrasted conditions
- **Option 1 widely used in model testing, not demanding enough** in the case of CC impact studies
- **Necessary use of option 2**: evaluation of the model's ability **to adapt to unknown conditions**

Calibration on dry period and validation on wet period: **change towards wetter conditions**



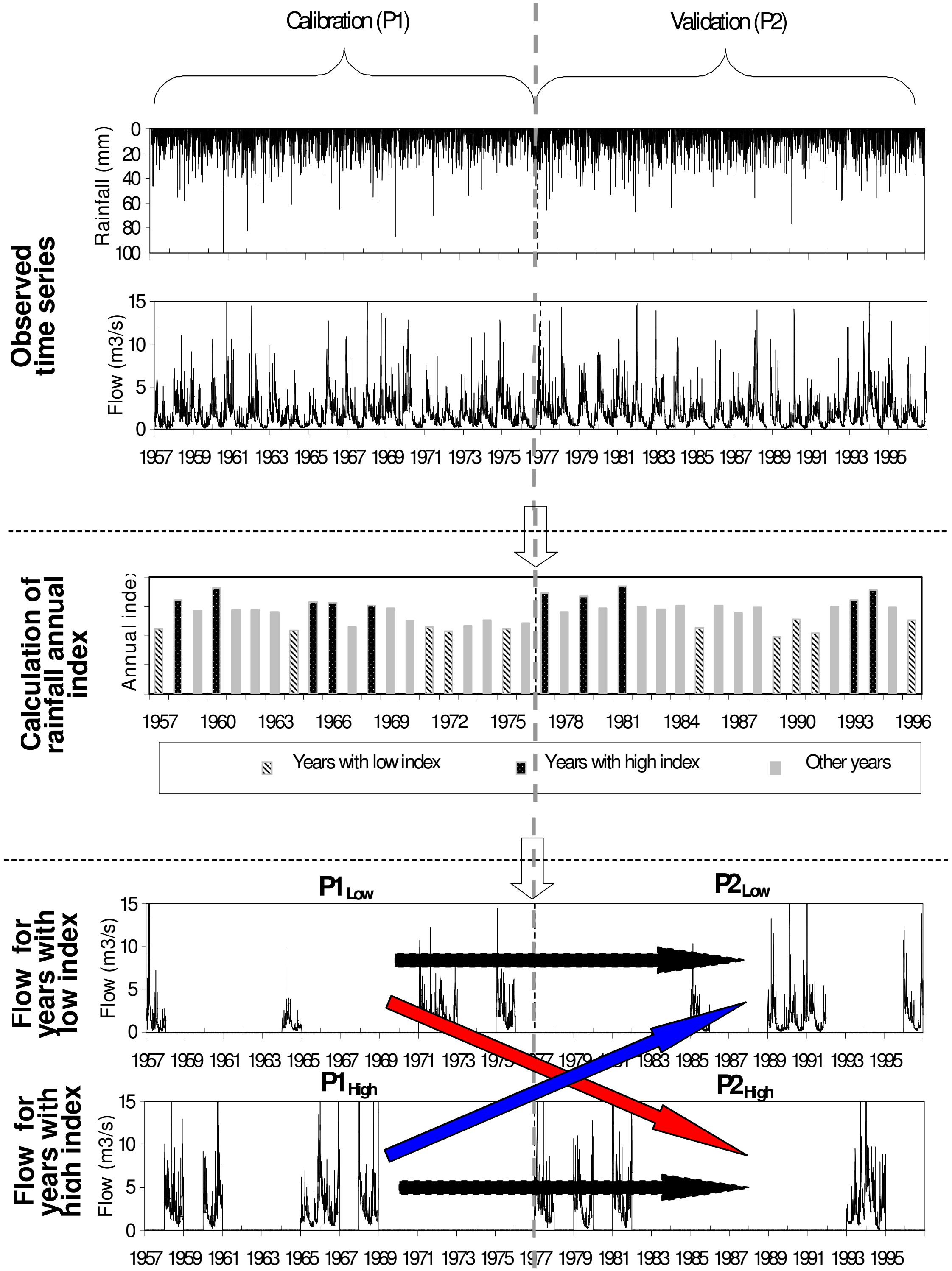
Calibration on wet period and validation on dry period: **change towards drier conditions**

Performing the differential split sample test scheme

Generalization to get more contrasted conditions between calibration/validation by **selecting specific years**:

1. Choose a selection index (e.g. mean precipitation or temperature)
2. Identify years with low/high index values (here selection of 5-year periods)
3. Calibrate/validate models on selected years in contrasted conditions:
   - calib. on low and valid. on high
   - calib. on high and valid. on low
4. Analyze change in model performance

		Validation	
		P2Low	P2High
Calibration	P1Low	Conditions unchanged	**Change towards increase**
	P1High	**Change towards decrease**	Conditions unchanged



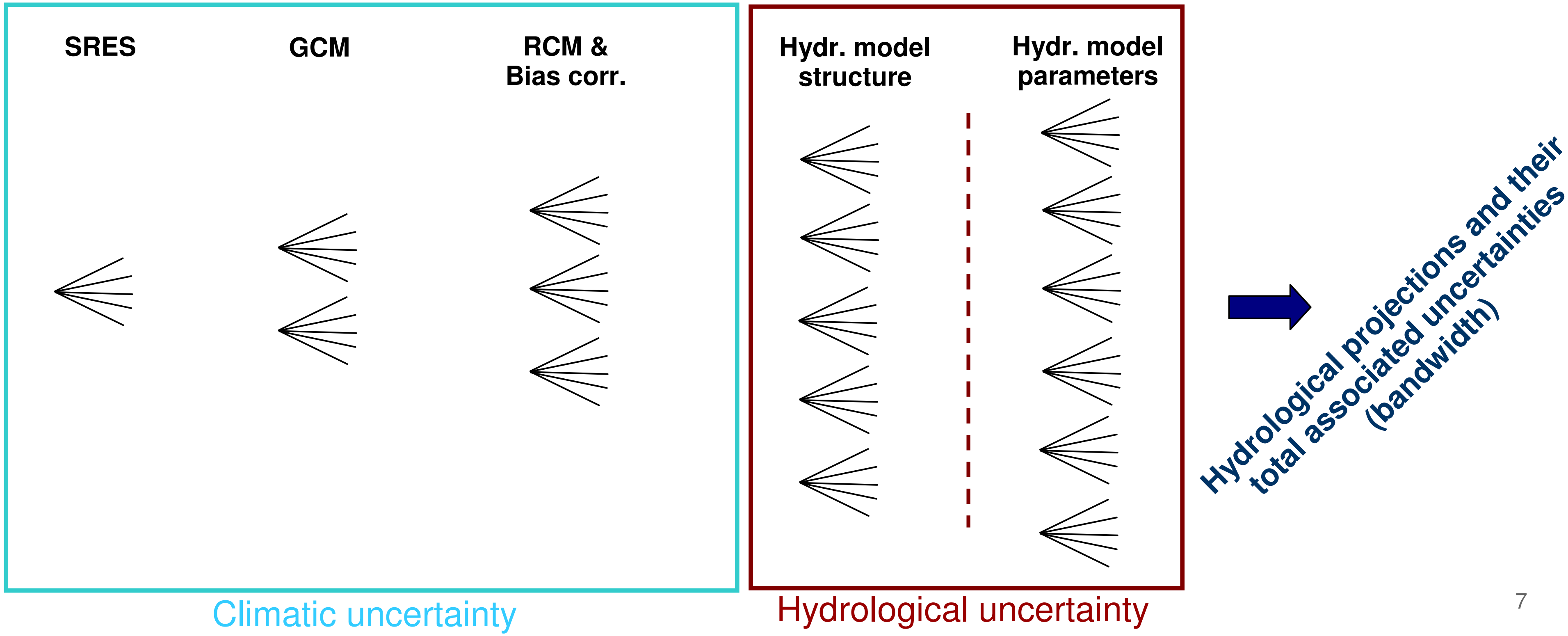
Objectives within RheinBlick2050?

1. **Testing the reliability of hydrological models**
   by comparative benchmarking
2. **Testing the stationarity hypothesis**
   by performing Klemes's differential evaluation scheme

} using **observed** climatic inputs

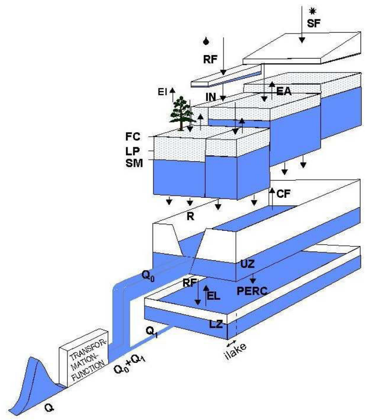
3. **Comparing sources of uncertainty in hydrological projections**
   by comparing bandwidth

} using **climate projections**

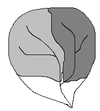


Hydrological models

Semi-distributed HBV134 model



- Divides the catchment into sub-catchments with flow routing

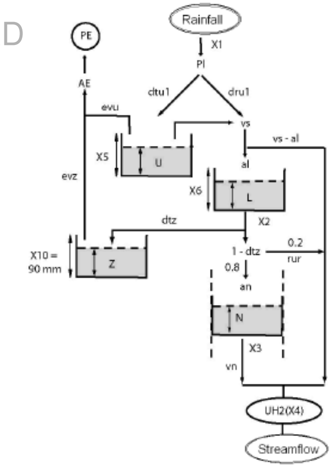


- Widely used on the Rhine basin
- Major prior efforts to optimize the model on the CHR_OBS data set (no re-calibration in RB2050)
- Two model versions: BFG and DELTARES

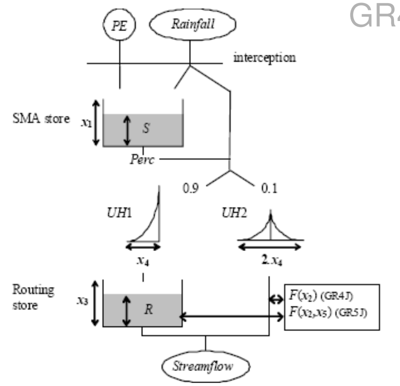
7 lumped hydrological models

(GR4J, GR5J, HBV0, IHAC, MORD, MOHY, TOPM)

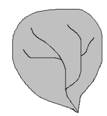
MORD



GR4J



- Consider the catchment as a single unit



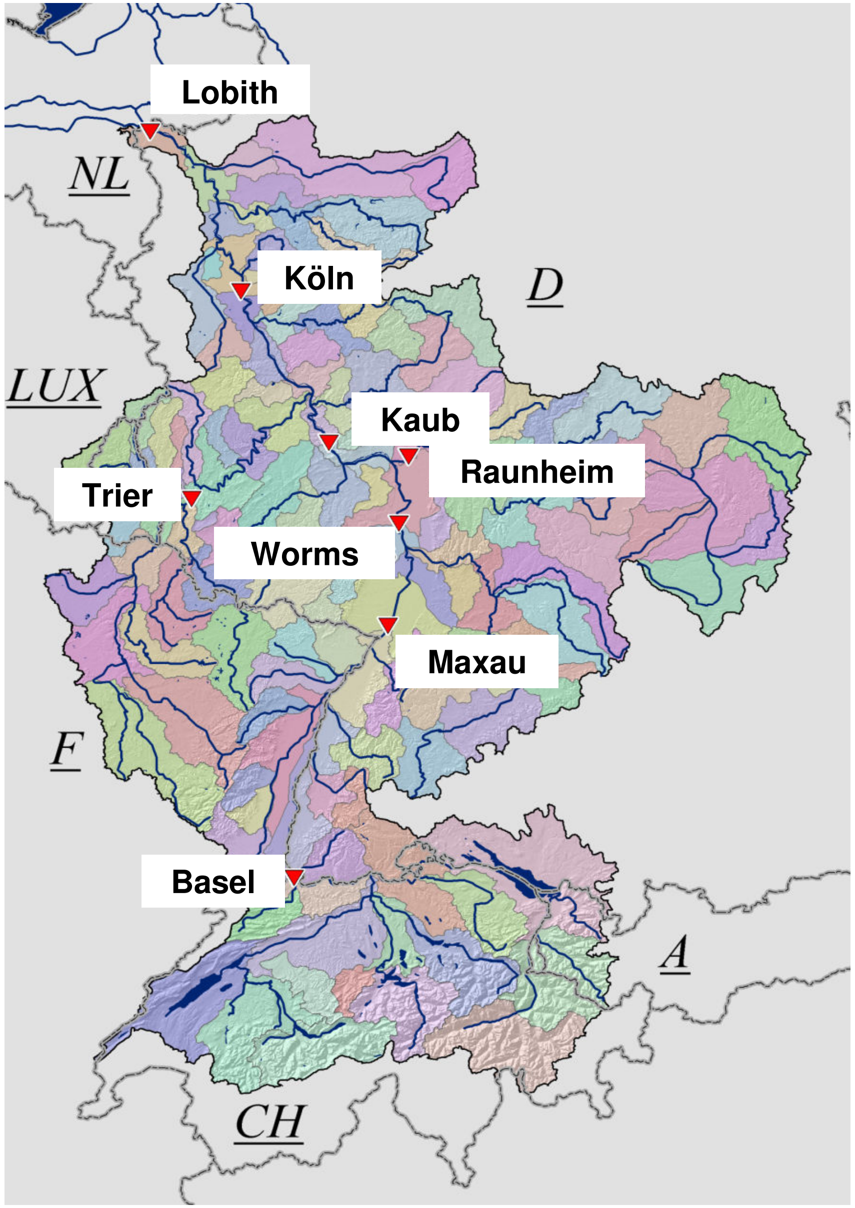
- Models tested in various conditions
- From 4 to 9 free parameters automatically optimized
- Use of a simple degree-day snowmelt module on top of models
- Lumped models used as *benchmarks* for the semi-distributed model

Data and criteria

- CHR_OBS data set for the 1961-1990 reference period
- 8 target stations
- Daily time series of P, Temp., PE and Q
- Two PE formulations tested (Penman-Wendling and Oudin)
- 18 climatic projections

- Model evaluation using 6 performance measures (focusing on regime curve, mean, high and low flows)

- *Difficulty to compare results due to different testing schemes*



Testing the reliability of hydrological models

CHR
KHR

HBV134 model

- Generally high level of performance: more than 90% of observed flow variance explained on most of the 8 target stations (calibration results)
- Limited differences between BFG and DELTARES model versions

Lumped models

- Best models obtained good to very good results on all stations
- Catchment size not a factor limiting model performance (results similar to Merz et al., 2009)
- Significant differences between the worst and best lumped models, i.e. model structure potentially brings significant uncertainty
- Not a single model better everywhere
- MORD best performing on average followed by GR4J/GR5J/MOHY (the three other models are significantly poorer)

Testing the reliability of hydrological models

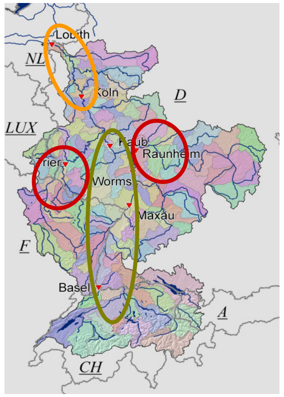
HBV134 model benchmarking *(comparison in calibration)*

- <u>On average on the 8 target stations</u>: **HBV134 versions better than all lumped models** (reduction of MORD RMSE values ranging from 2 to 20% for HBV134 model versions)

Average efficiencies (NSE) on the 8 target stations

	Regime	Low flow	High flow	
Model	Efficiency on monthly mean flows	Efficiency on log-transf. flows	Efficiency on flows	
HBV134_BFG_EOU	0.916	**0.919**	**0.911**	Semi-distributed
HBV134_BFG_EWP	0.927	0.910	0.907	Semi-distributed
HBV134_DELTARES	**0.940**	0.915	0.897	Semi-distributed
GR4J	0.868	0.849	0.857	Lumped
GR5J	0.869	0.845	0.862	Lumped
HBV0	0.823	0.814	0.776	Lumped
IHAC	0.795	0.828	0.826	Lumped
MOHY	0.886	0.842	0.857	Lumped
MORD	0.911	0.870	0.870	Lumped
TOPM	0.820	0.841	0.815	Lumped

- <u>On each target stations</u>:
  - HBV134 outperformed lumped models in Basel, Maxau, Worms and Kaub
  - Slight advantage to HBV134 in Köln and Lobith
  - Slight advantage to lumped models at Raunheim (Main) and Trier (Moselle)

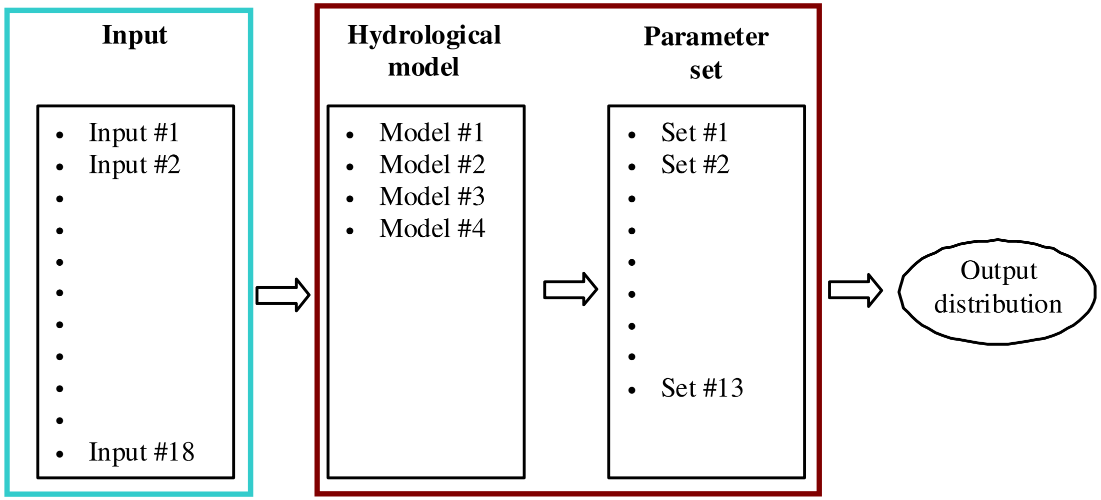


Testing the stationarity hypothesis

- Tests performed only with lumped models
- Main results:
  - Models **less robust** in contrasted conditions
    - ➢ models calibrated under current conditions are **not optimal** to simulate catchment hydrological behaviour under very different future climate conditions
  - Sensitivity of results to calibration conditions **varies between models**
    - ➢ poorest models in split sample testing are also the most sensitive to calibration conditions
  - **Differences much larger between models** with the test under contrasted conditions
    - ➢ differential split sample test scheme more informative on models' limitations
  - Relative errors increase (by about 10% on low flows, a bit less on mean and high flows)
  - **Level of model error found under stationary conditions most likely optimistic**

Comparing sources of uncertainty in hydrological projections

- Cascading (reliable) model options



- Evaluating the bandwith associated with hydrological projections
- Four combinations investigated:
    - Combination #1: 18 clim. proj. × 1 hydr. model × 1 param. set
    - Combination # 2: 18 clim. proj. × 1 hydr. model × 13 param. sets
    - Combination # 3: 18 clim. proj. × 4 hydr. models × 1 param. set
    - Combination # 4: 18 clim. proj. × 4 hydr. models × 13 param. sets

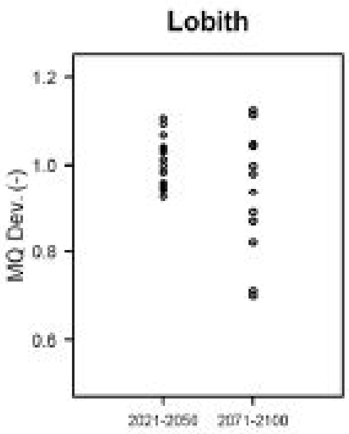
Increase in uncertainty due
to model structure
to parameterization

Comparing sources of uncertainty in hydrological projections

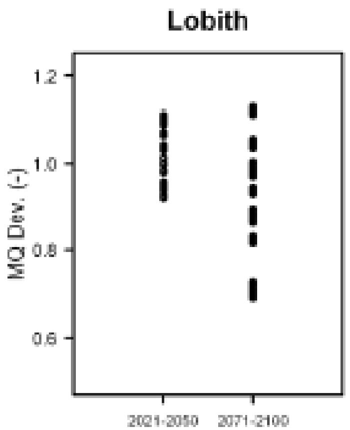
CHR
KHR

- Example results for Lobith for mean flow deviation for the 2021-2050 and 2071-2100 periods

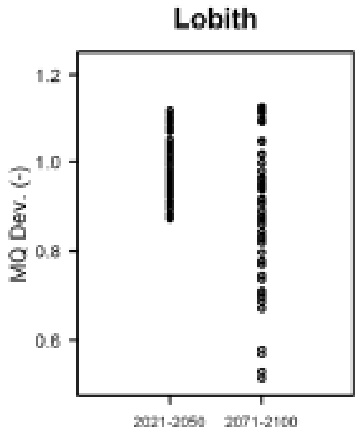
Combination #1
18 cp × 1 ms × 1 ps



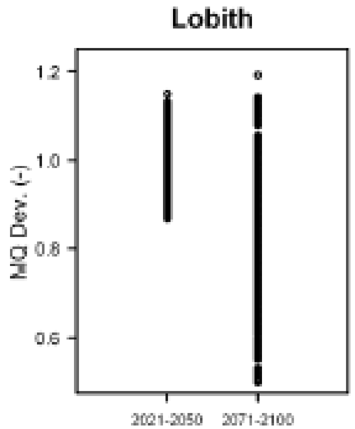
Combination #2
18 cp × 1 ms × 13 ps



Combination #3
18 cp × 4 ms × 1 ps



Combination #4
18 cp × 4 ms × 13 ps



(cp: climate projection; ms: model structure; ps: parameter set)

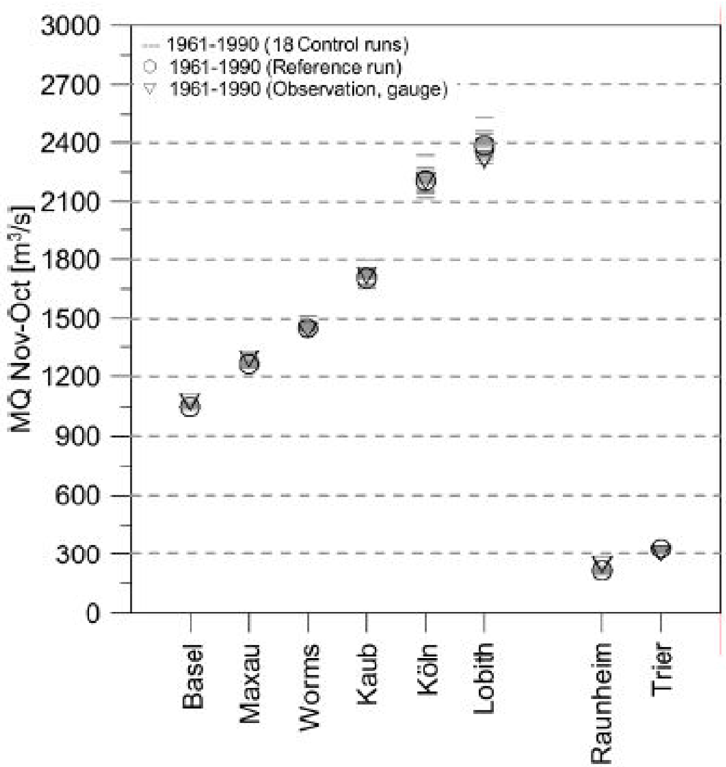
- Structural uncertainty larger than parameter uncertainty
- Most of the spread in the bandwidth brought by climatic inputs
- Results similar for other stations and criteria

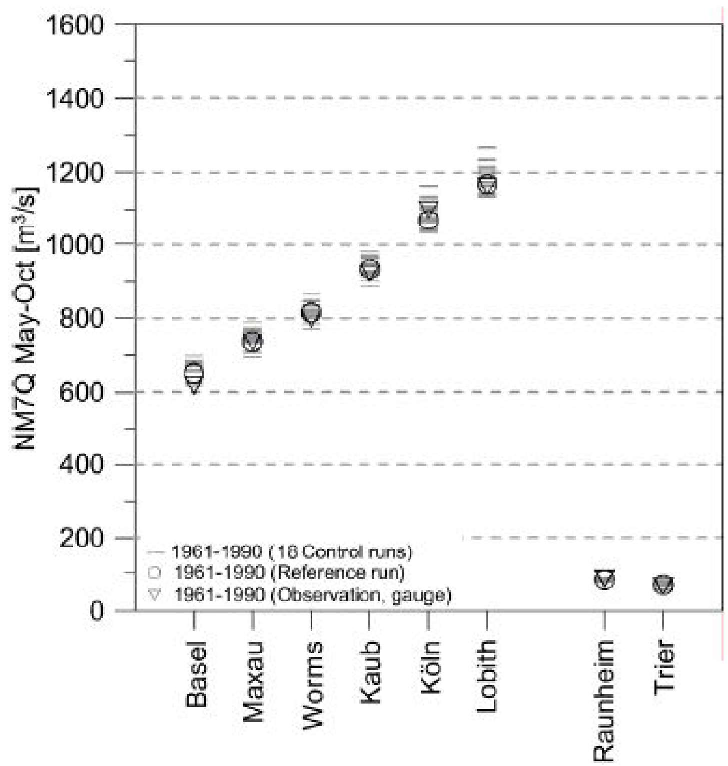
On the 8 target stations:

Climatic uncertainty > Hydr. model structure uncertainty > Hydr. model parameter uncertainty

HBV134 reliability on reference period

- Complete model chain run with the HBV134 to reproduce flows on 1961-1990
- Good confidence on mean flows, larger uncertainties on extreme flows
- Model chain is considerably better for gauges with winter high flows (Lobith, Köln, Kaub, Raunheim and Trier) than for gauges where high flows occur mostly in summer (Basel, Maxau and Worms) (the latter not further investigated)





Main conclusions

CHR
KHR

- To which extent can we trust hydrological models in CC impact studies?
  - This can be objectively tested (although there are limitations to the evaluation framework used here)
  - **Classical split sample test insufficient** to evaluate model robustness
  - New solutions should be found to **cope with the problem of parameter nonstationarity** (ongoing works by Chiew et al. in Australia; Merz et al. in Austria; Brigode et al., Coron et al. in France; etc.)
  - Performance in current conditions are probably over-optimistic (i.e. model potential errors in future conditions underestimated)
  - Hydrological models still to be improved!

- On the Rhine basin:
  - HBV134 model more reliable on average than the lumped "benchmark" models (*but without full testing scheme*); **HBV134 selected for further investigations in RB2050**
  - Differences more significant on stations strongly influenced by the Alps

Main conclusions

In terms of uncertainty:

- On our target stations, significant uncertainty brought by model structure and (to a lesser extent) model parameters
- But uncertainty associated with climate forcings is larger
- Uncertainty underestimated due to the stationarity hypothesis
- Similar results had been obtained by other authors (e.g. Wilby, 2005)
- Considerable bandwidth associated with hydrological projections
- Might be difficult to handle in decision making

- Uncertainty may be reduced if **severe enough evaluation scheme** are applied to accept/reject models of the modelling chain: need to develop appropriate **"crash tests"** (see Andréassian et al., 2009)

RheinBlick2050

http://www.chr-khr.org > Projects > RheinBlick2050

Dr. Charles Perrin

Cemagref
Hydrosystems and Bioprocesses Research Unit (HBAN)

e-mail: **charles.perrin@cemagref.fr**
phone: +33 1 40 96 60 86
web-address: **http://www.cemagref.fr/webgr**

Parc de Tourvoie, BP 44, 92163 Antony Cedex, France