AGENCE FRANÇAISE POUR LA BIODIVERSITÉ
ÉTABLISSEMENT PUBLIC DE L'ÉTAT

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

MINISTÈRE DE LA TRANSITION ÉCOLOGIQUE ET SOLIDAIRE

The PREMHYCE project: a comparative evaluation of hydrological models for low-flow forecasting on French catchments

P. NICOLLE, V. ANDREASSIAN
pierre.nicolle@irstea.fr; http://webgr.irstea.fr

irstea
www.irstea.fr

Basel, 19 September 2017
CHR Symposium
" Low flows in the Rhine Catchment "

Context

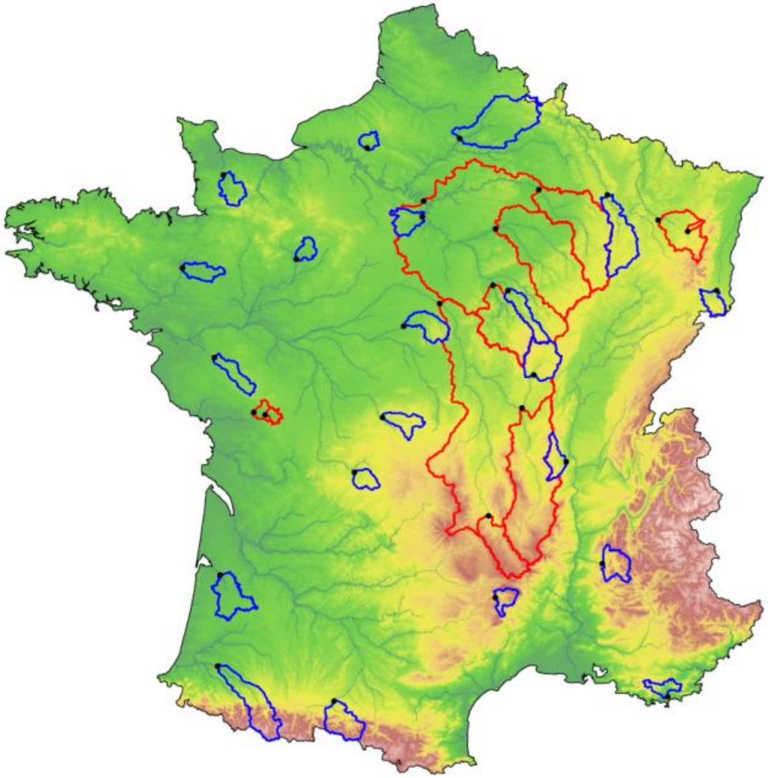
- In France in 2013: 83% of total withdrawals from rivers

- Water uses affected by water shortages in rivers
- Climate change : perspective of more severe summer low-flows
- Lack of forecasting tool at national scale

Early anticipation of low-flow periods needed to improve water management

Study objectives

- Comparing hydrological models for low-flow ensemble forecasting in a common test protocol
- Assessing the ability of forecasting tools to anticipate low-flow situations (magnitude, maximum lead-time)
- Developing operational low-flow forecasting tool



Catchment set

- 21 with no or limited influence (380 to 4 300 km²)
- 11 influenced by dams or water withdrawals (120 to 44 000 km²)

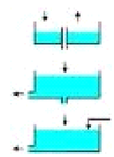
Data set

- Daily streamflow (HYDRO French database): 14 to 36 years (1974-2010)
- Daily P, PE, Temp (SAFRAN climate reanalysis): 51 years (1959-2010)
- Daily influences (dam volume, withdrawals): 11 to 25 years (1999-2010)

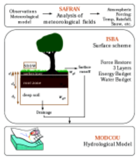
Hydrological models

- Five rainfall-runoff models already used in operational conditions in France
- Daily continuous functioning
- Different modelling approaches (model type, spatial resolution)
- Various number of free parameters
- Influences not systematically taken into account
- Various use of assimilation schemes or statistical correction procedures

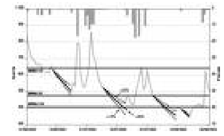
GARDENIA/EROS



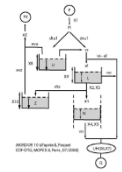
Isba-Modcou



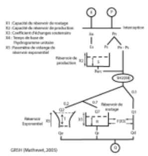
PRESAGES



MORDOR

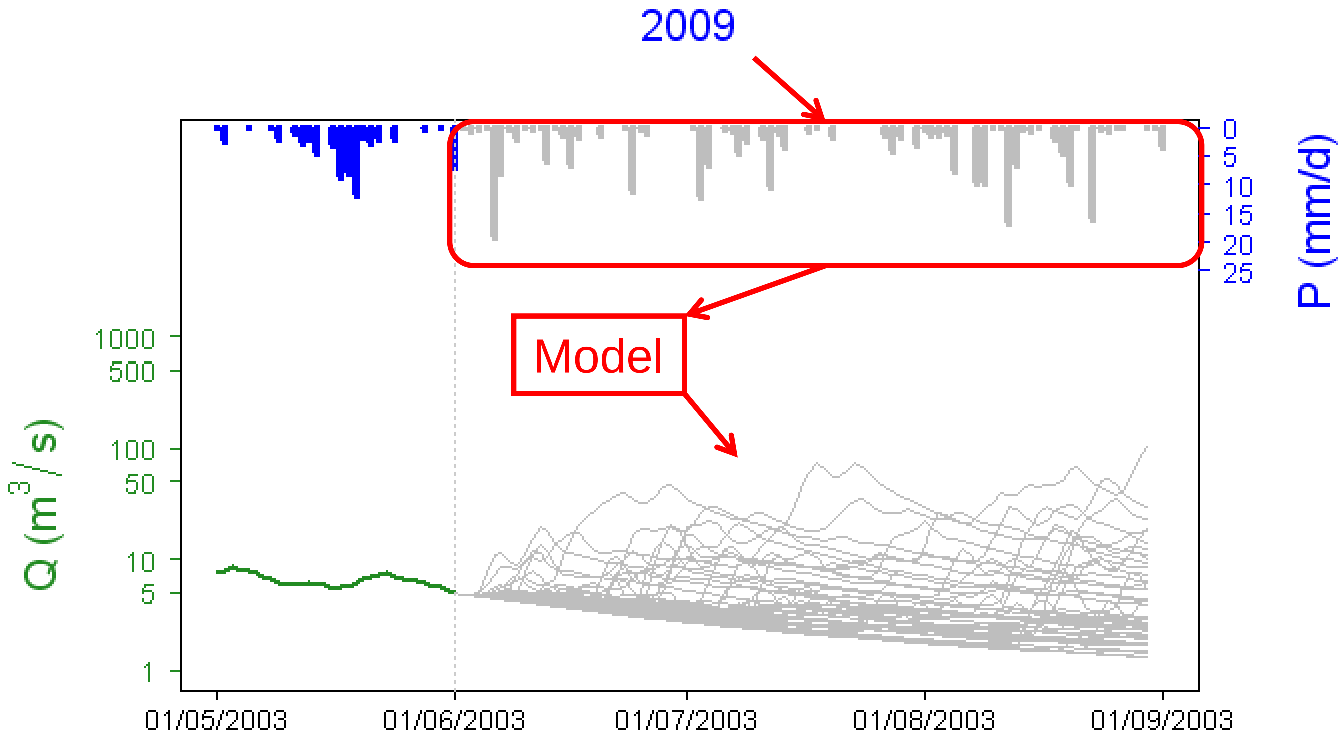


GR6J



Ensemble forecasting

- Future meteorological inputs (P, PE, Temp): climatic archive (50 scenarios) → Get general results, include severe drought conditions



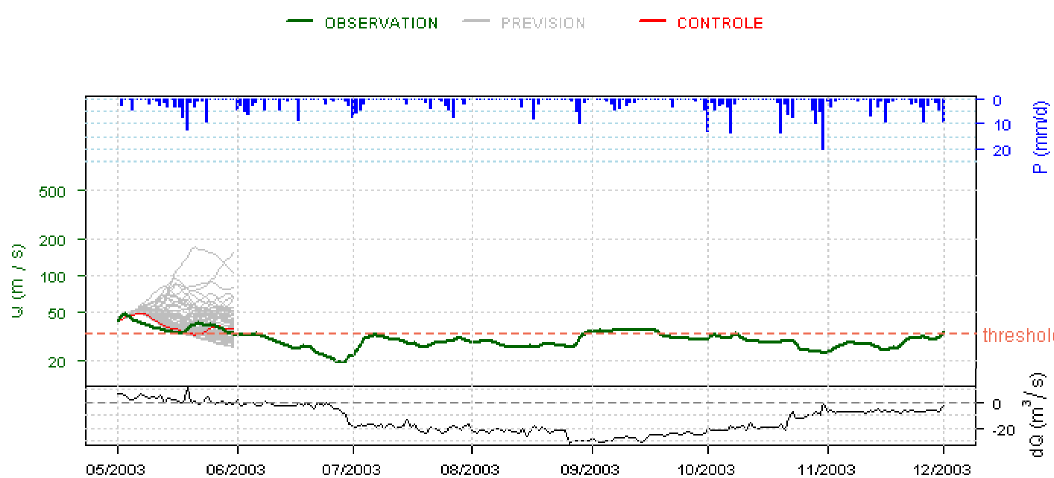
Evaluation method

- Split-sample test approach (incl. 3-year warm-up)

Period 1	Period 2
Calibration →	Validation

- Test in hindcasting mode : retrospective run at each time step of period, forecast as in real time
- Calibration method and objective function:
  choice of the modeller based on his experience with his model

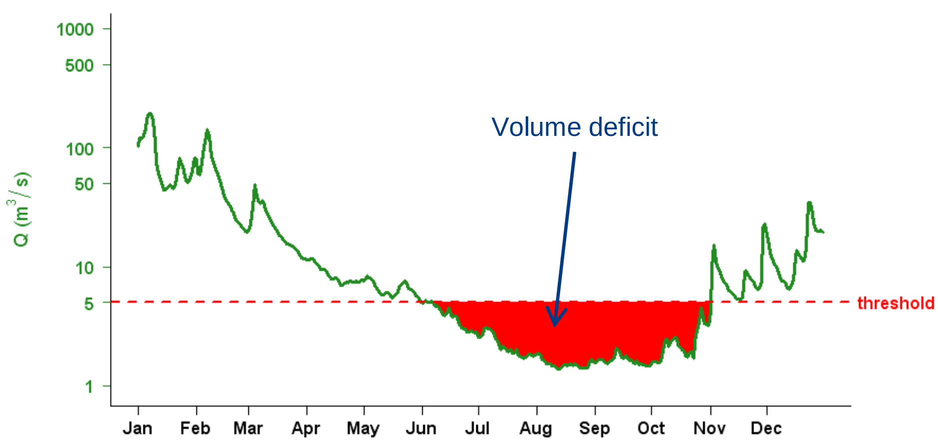
The Seine river at Pont-sur-Seine

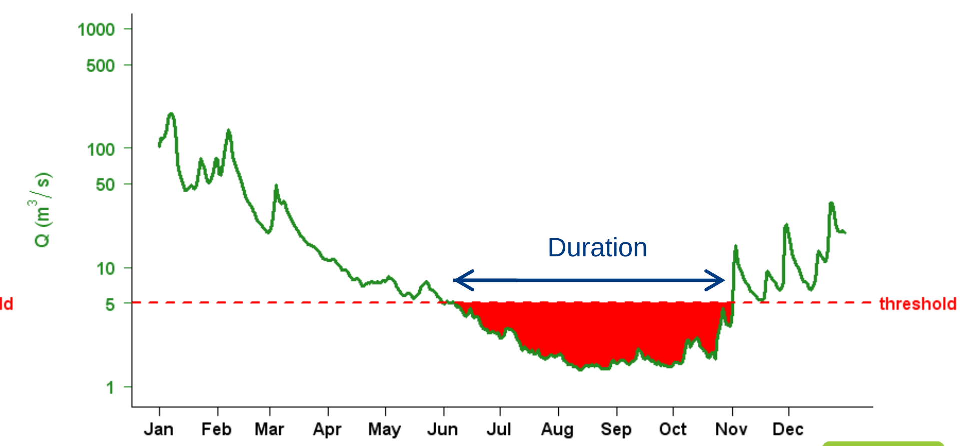


nent "

Target variables for low-flows

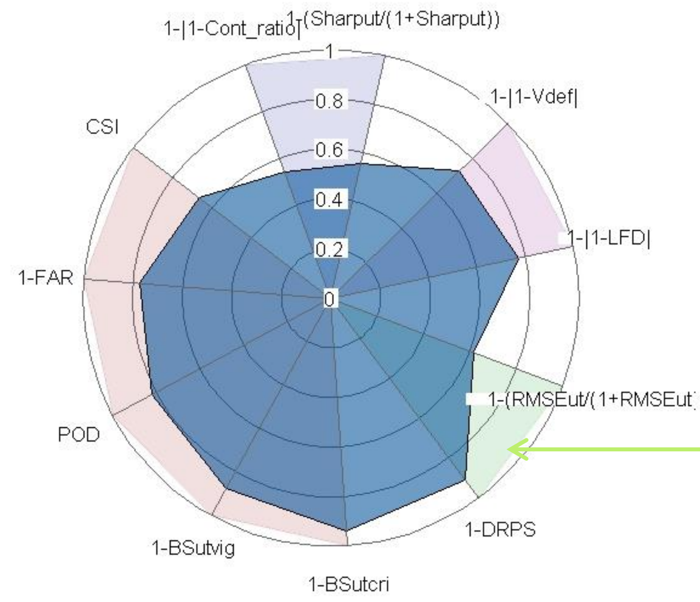
- Moving average streamflow over 3 days
- Streamflow threshold: Q80 (80% of streamflows above the threshold)
- Low-flows characteristics:
  - ➢ **Volume deficit**
  - ➢ **Low-flow duration**





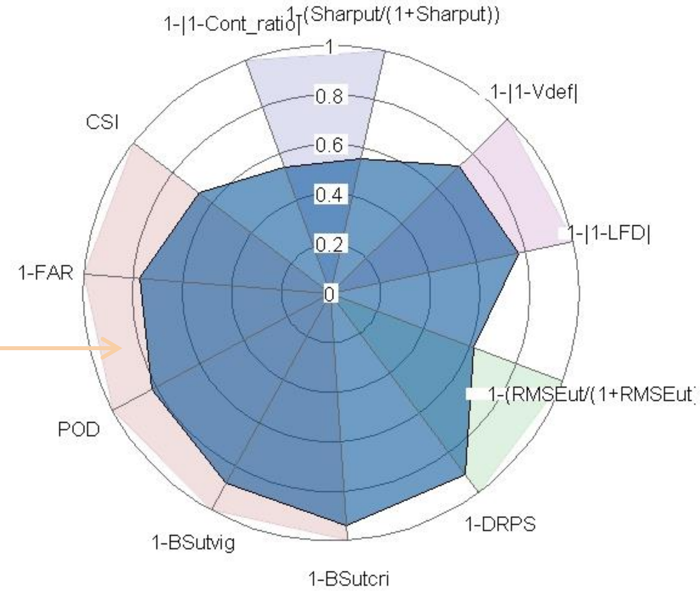
Benchmark and evaluation criteria

- Large selection of efficiency criteria: evaluation of different qualities of hydrological models for forecasting
    - Range of flows (all and low-flows)
    - Relative to the cross of threshold
    - Low-flow characteristics
    - Sharpness
    - Reliability



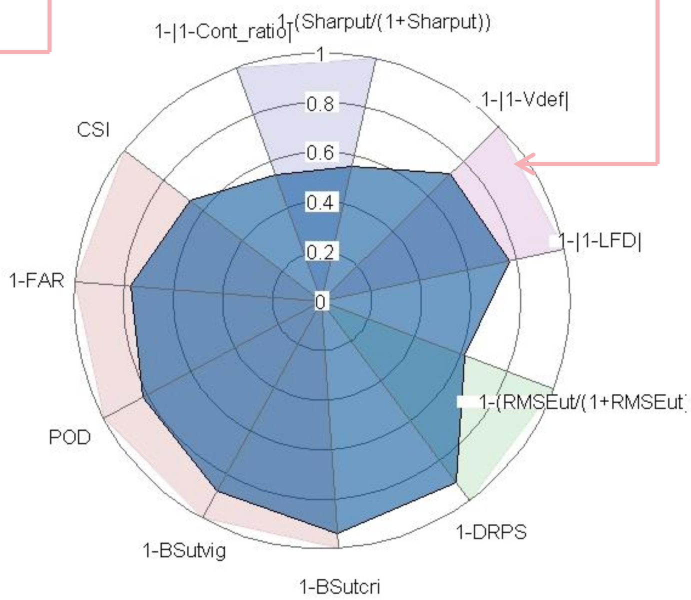
Benchmark and evaluation criteria

- Large selection of efficiency criteria: evaluation of different qualities of hydrological models for forecasting
  - Range of flows (all and low-flows)
  - Relative to the cross of threshold
  - Low-flow characteristics
  - Sharpness
  - Reliability



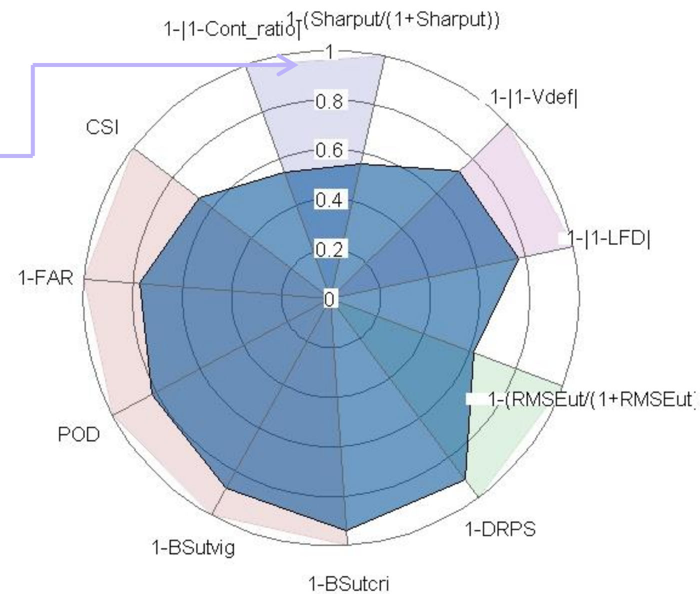
Benchmark and evaluation criteria

- Large selection of efficiency criteria: evaluation of different qualities of hydrological models for forecasting
  - Range of flows (all and low-flows)
  - Relative to the cross of threshold
  - Low-flow characteristics
  - Sharpness
  - Reliability



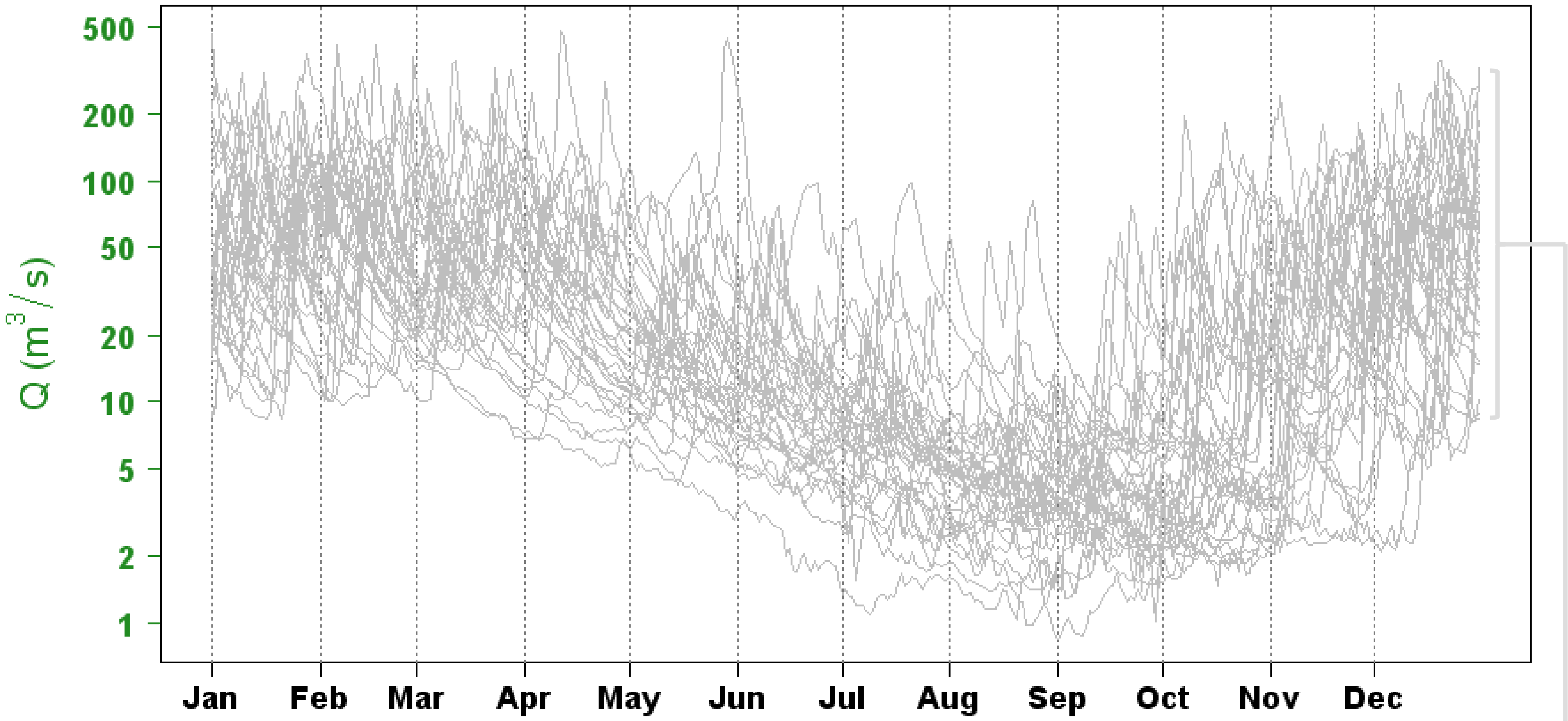
Benchmark and evaluation criteria

- Large selection of efficiency criteria: evaluation of different qualities of hydrological models for forecasting
  - Range of flows (all and low-flows)
  - Relative to the cross of threshold
  - Low-flow characteristics
  - Sharpness
  - Reliability



- Mean performances on all catchments
- Target lead times for model evaluation: 7 days and 30 days

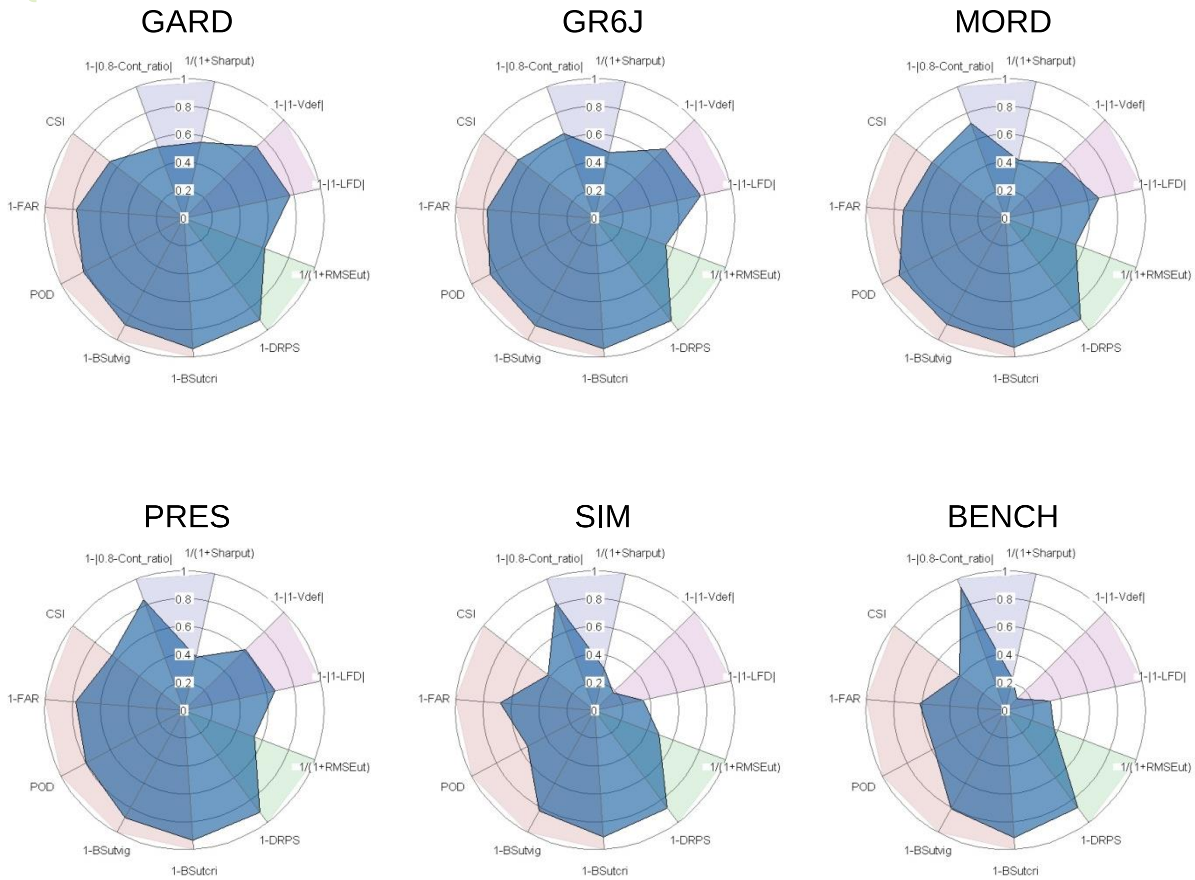
Benchmark and evaluation criteria



- Models to be compared to benchmark:
  - Natural variability of observed streamflow (Bench)

3. Forecasting results

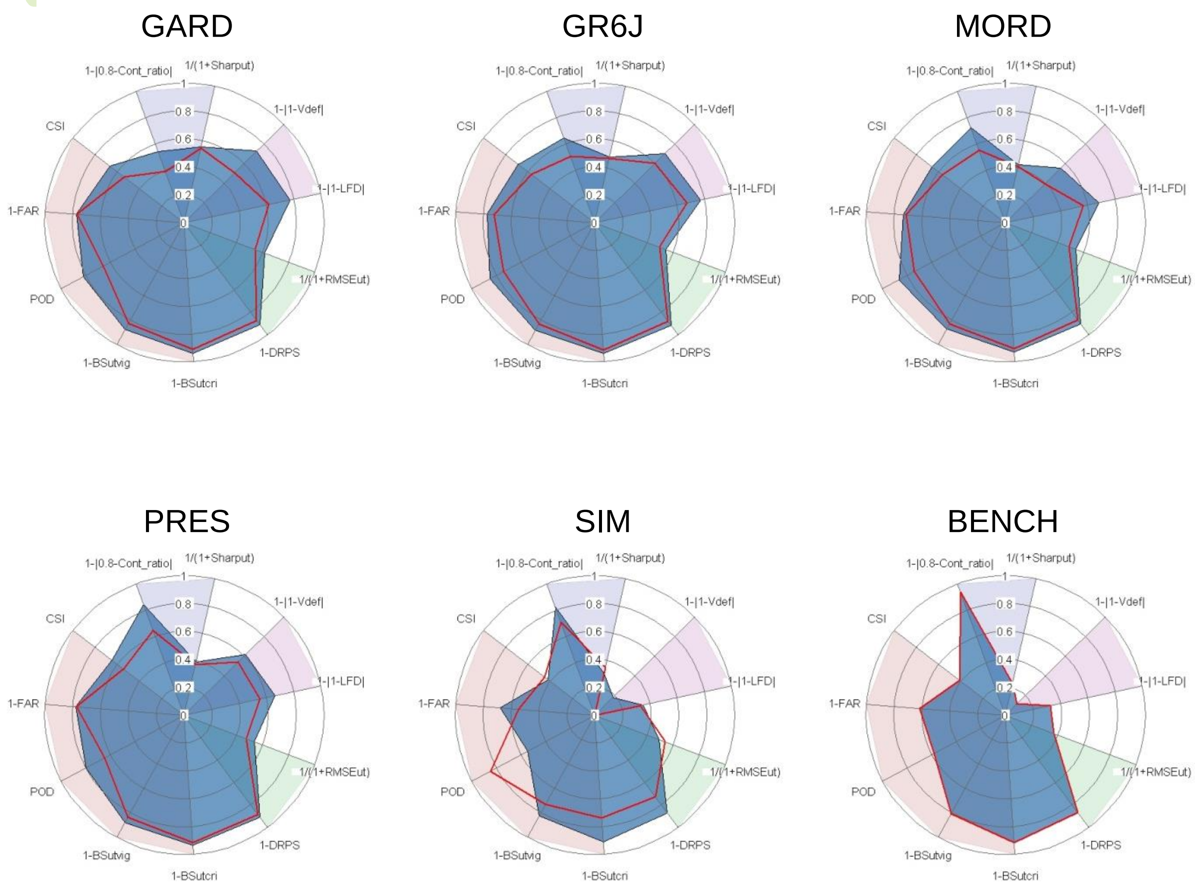
Lead time: 7 days



- Differences for a few criteria, but difficult to identify a better model
- Significant gain compared to the benchmark

3. Forecasting results

Lead time: 7 days



No assimilation or post-correction method

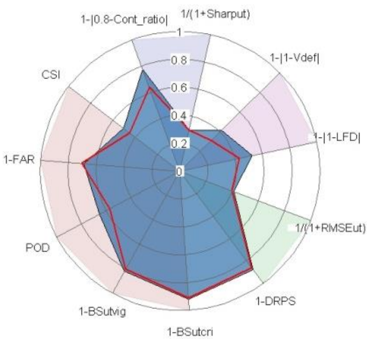
Use of assimilation or post-correction method

- Differences for a few criteria, but difficult to identify a better model
- Significant gain compared to the benchmark
- Significant gain when using streamflow assimilation or post-correction methods

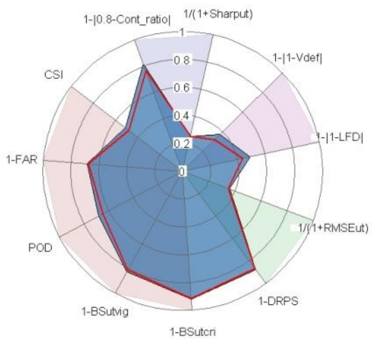
3. Forecasting results

Lead time: 30 days

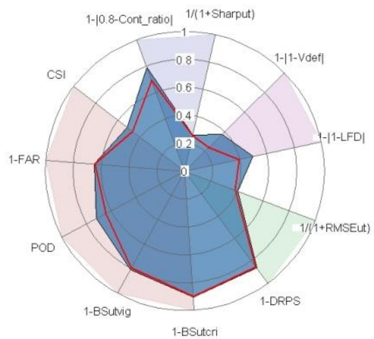
GARD



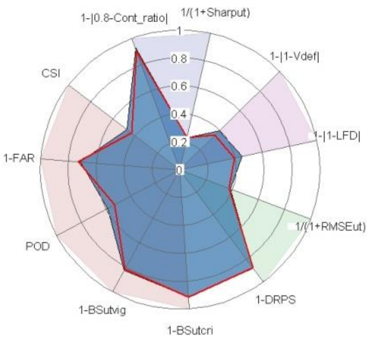
GR6J



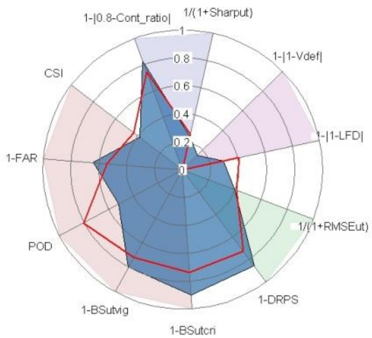
MORD



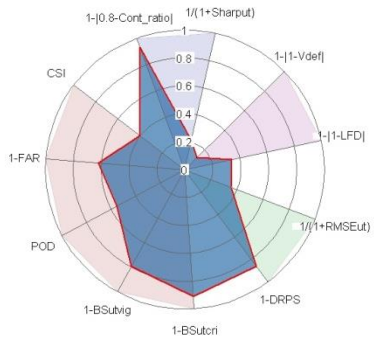
PRES



SIM



BENCH



No assimilation or post-correction method

Use of assimilation or post-correction method

- Assimilation or post-correction methods less useful with increasing lead-time
- Performance loss with increasing lead-time
- Closer than benchmark but still better

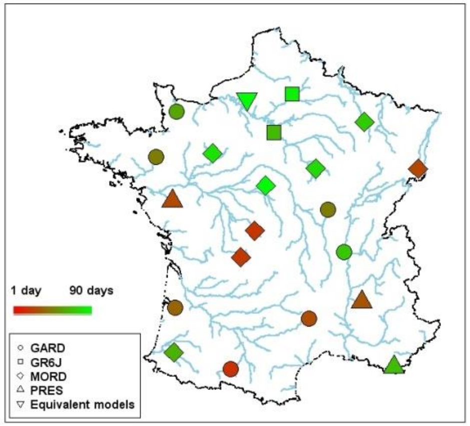
irstea

Evaluation of the Useful Forecasting Lead time (UFL)

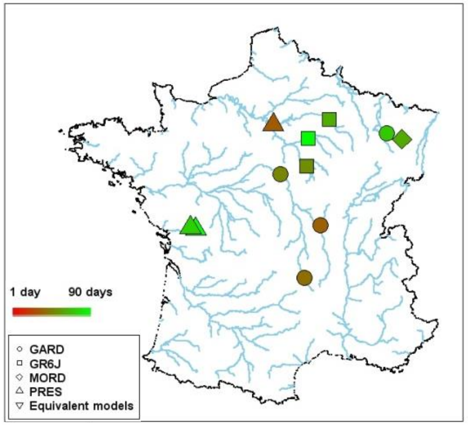
- Definition:
  Lead time beyond which the model does not bring valuable information compared to the benchmark (natural variability of streamflow)
- Here valuable information if model efficiency at least 20% better than the benchmark efficiency

- UFL depends on efficiency criteria
- UFL varies between:
  - Catchments, but no relation between UFL and low-flow or catchment characteristics
  - Models

Non-influenced

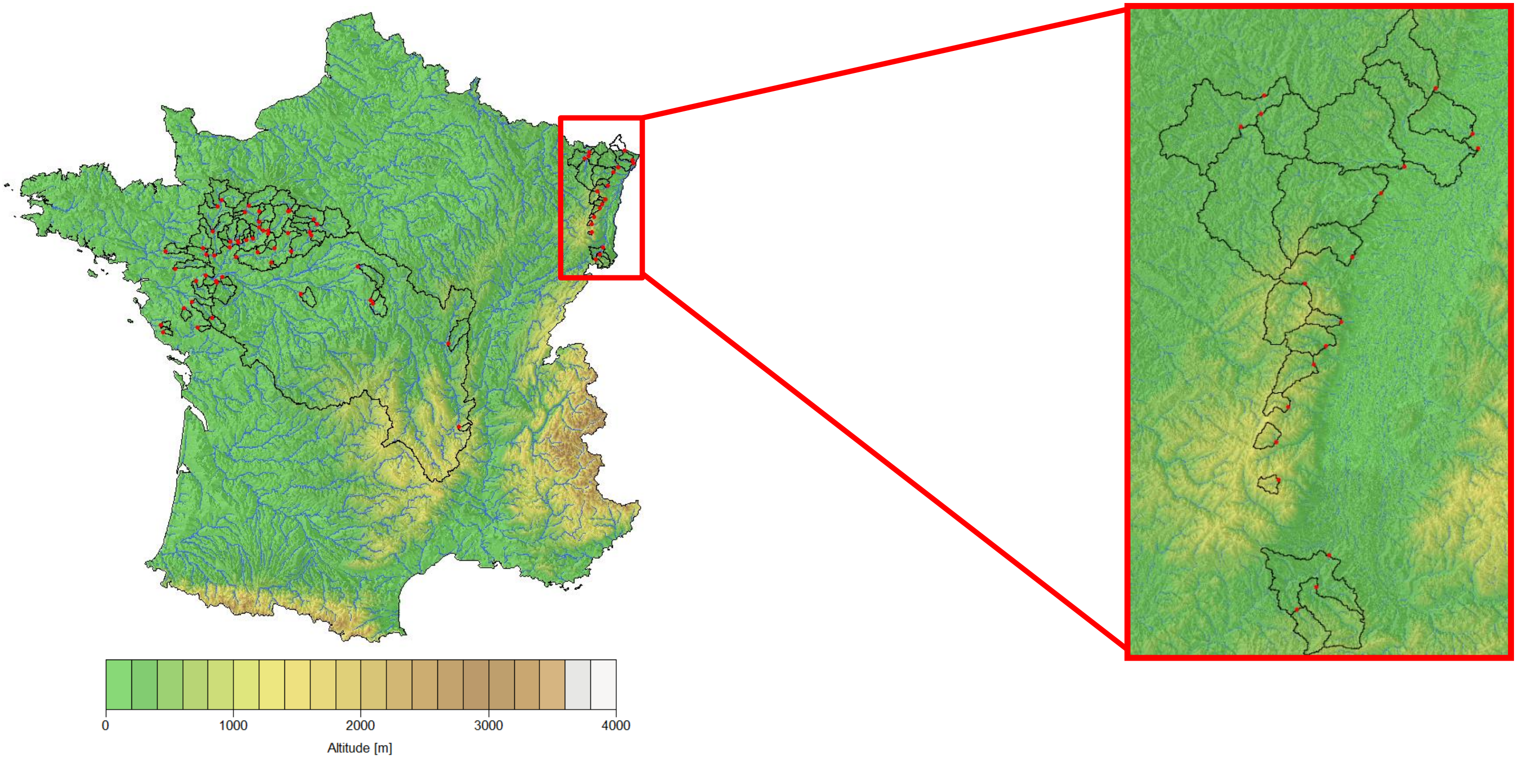


Influenced

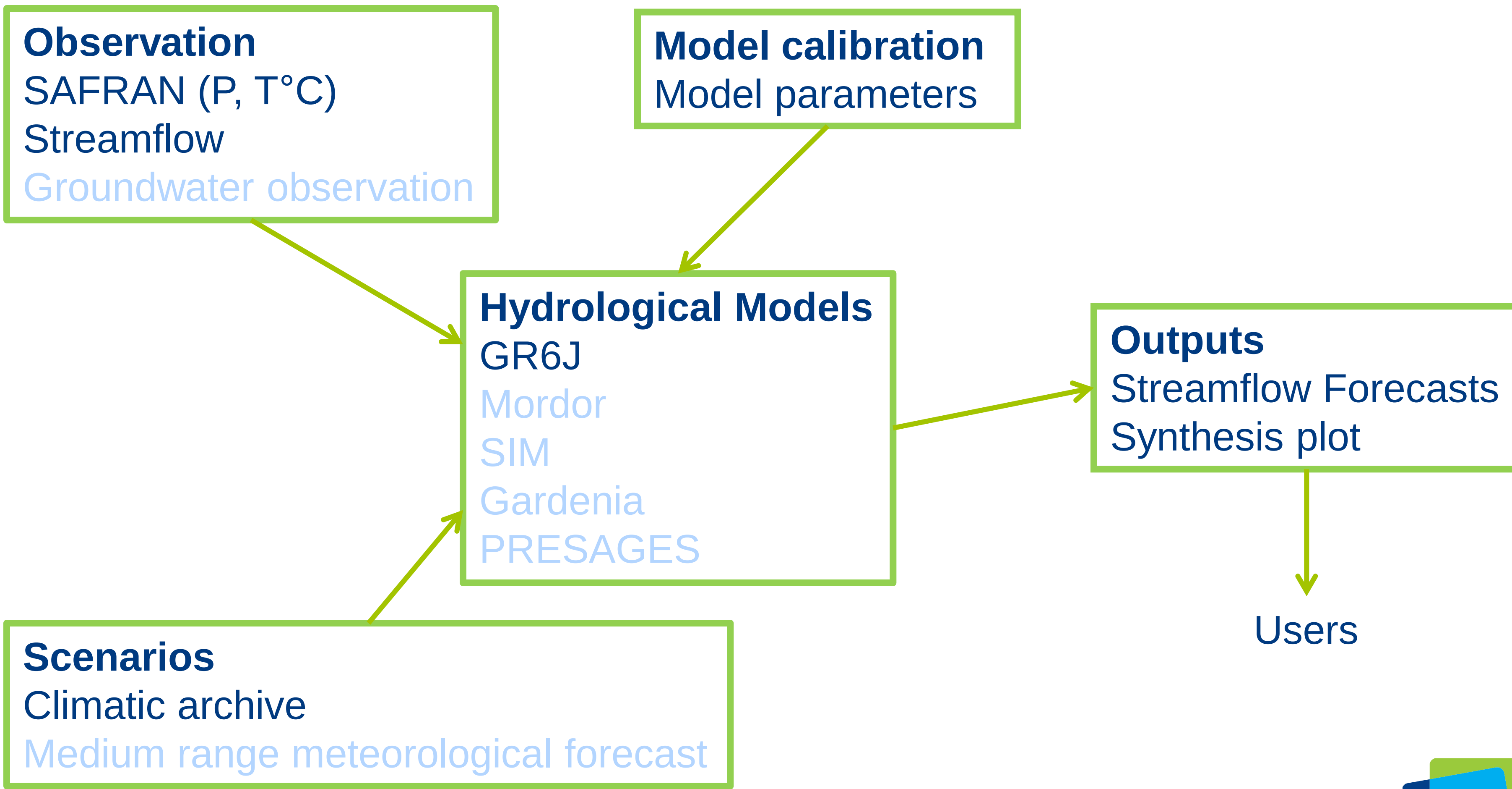


Operational implementation of forecasting tool

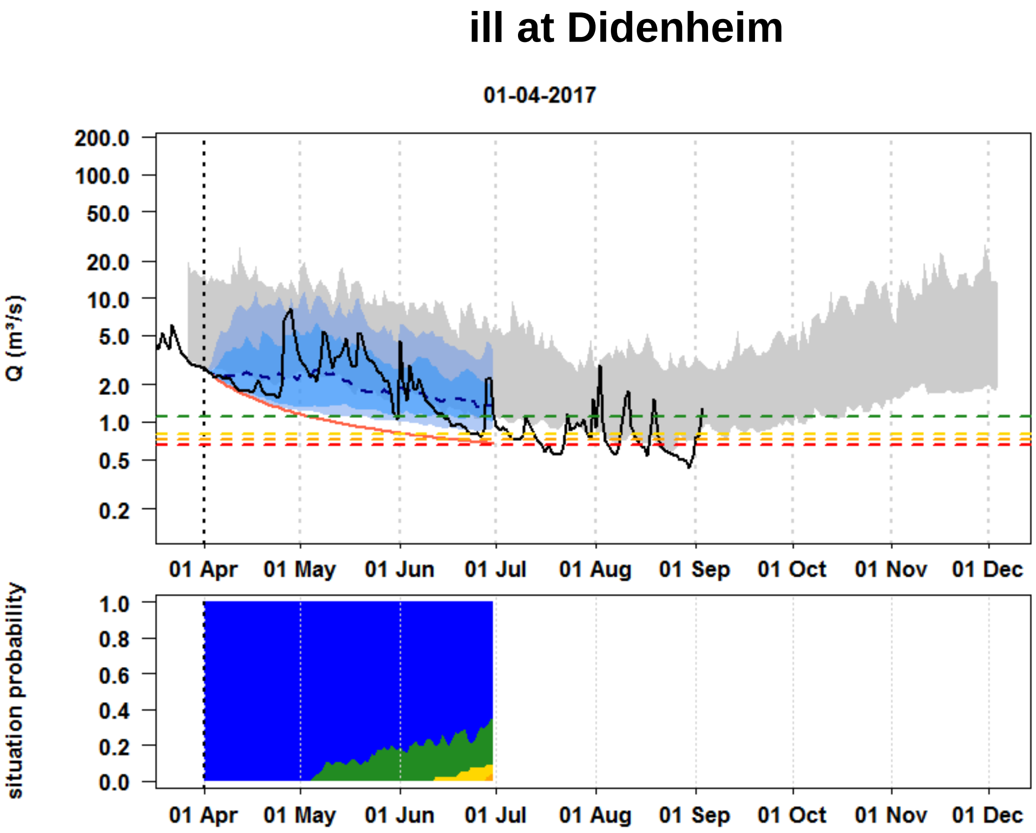
- Beta version with GR6J since July 2017
- Ensemble low-flow forecasting at 90 days lead-time
- Test on 70 catchments, 19 on the French part of the Rhine catchment

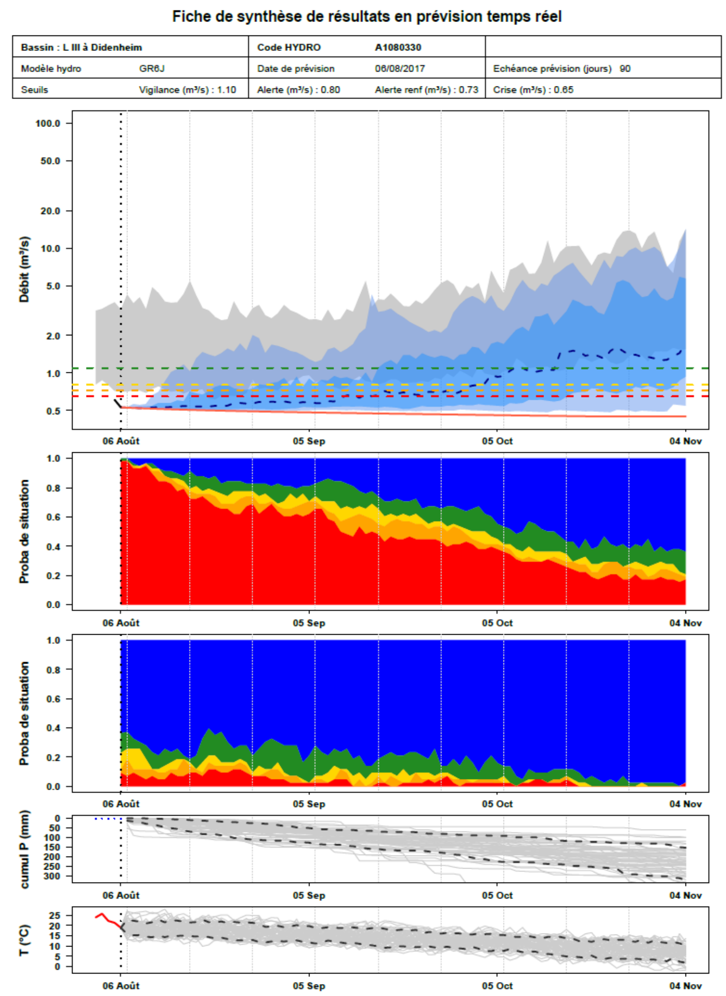


Schematic representation of real-time functioning



4. Operational Forecasting tool





Conclusion

- Common protocol to compare and evaluate hydrological models for low-flow forecasting
- No superior model on all catchments or criteria, comparison with benchmark: quantification of the actual value of low-flow forecasting by hydrological models
- Using assimilation or post-correction method less interesting with increasing lead-time
- Simple method to determine Useful Forecasting Lead-time
- Performances quite good on influenced catchments, with various simple methods to account for influences

Perspectives

- Deployment in operational services in 2018
- Integrating other models in operational tool
- Multi-model approach

Thank you !

Further details in:

Nicolle, P. *et al.*, *Benchmarking hydrological models for low-flow simulation and forecasting on French catchments*, Hydrol. Earth Syst. Sci., 18, 2829-2857, doi:10.5194/hess-18-2829-2014, 2014.

Contact:

pierre.nicolle@irstea.fr; http://webgr.irstea.fr

Acknowledgements:

Meteo-France, EDF, BRGM, University of Lorraine, AFB

Models

Short name used here	GARD	GR6J	MORD	PRES	SIM
Full name	GARDENIA	GR6J	MORDOR	PRESAGES	SIM
Reference on model structure	Thiéry (2013)	Pushpalatha (2011, 2013)	Garçon et al. (1999); Andréassian et al. (2006)	Lang et al. (2006a , 2006b)	Habets et al. (2008)
Type	Conceptual	Conceptual	Conceptual	Conceptual	Physically-based
Spatial distribution	Semi-distributed	Lumped	Lumped	Lumped	Distributed
Number of free-parameters	4 to 9 (+2 to 4 for snowmelt)	6 (+2 : snow routine)	11 (+4: snow routine)	7 (+3 : snow routine)	0
Calibration method	Automatic calibration: Rosenbrock method	Automatic calibration: local research method (step by step)	Automatic calibration: Shuffled Complex Evolution Method and Pareto Front Exploitation	Automatic calibration: simplex method with multistart	No calibration
Calibration criteria	RMSE with ln(Q)	$(KGE + KGE_i)/2$	$(KGE + KGE_i)/2$	Nash–Sutcliffe with $Q^{0.2}$	
Post-correction method (simulation)	Not used	Not used	Not used	Empirical method (Berthier, 2005)	Quantile/quantile post-treatment
Assimilation method (forecast)	When a flow discrepancy appears, the model tanks are updated proportionally to their variance	Correction based on error at first time step before forecast, with decreasing effect when lead time increases	Correction based on errors at previous time steps before forecast, with decreasing effect when lead time increases. No update of model stores.	Update of gravitary routing store	No assimilation method but a quantile/quantile post-treatment
Structure overview: production	Actual evapotranspiration is computed using a non-linear soil capacity. GW exchange is a proportion of the GW flow	A rainfall interception by PE, a non-linear SMA store, an intercatchment GW exchange function	A rainfall excess/soil moisture accounting store ; an evaporating reservoir ; an intermediate store and a deep store	A soil store, rainfall interception by PE	
Structure overview: transfer	A non linear tank distributes the effective rainfall into runoff and GW recharge. The aquifer is represented by a linear tank.	Two unit hydrograph, two parallel nonlinear routing stores	Direct, indirect and baseflow components are routed using a unit hydrograph (Weibull law)	Two unit hydrographs, two linear routing stores : one for streamflow recession, one for interflow	
References on simulation applications in France	800 to 1000 rivers simulated in France		Garavaglia (2011); Paquet et al. (2013)	Lang et al. (2006a, 2006b)	Vidal et al. (2010b) Habets et al. (2008)
References on low-flow forecasting applications in France		Pushpalatha (2011, 2013)	Mathevet et al. (2010)	Lang et al. (2006a, 2006b)	Céron et al. (2010) Soubeyroux et al. (2010) Singla et al. (2012)

Criteria

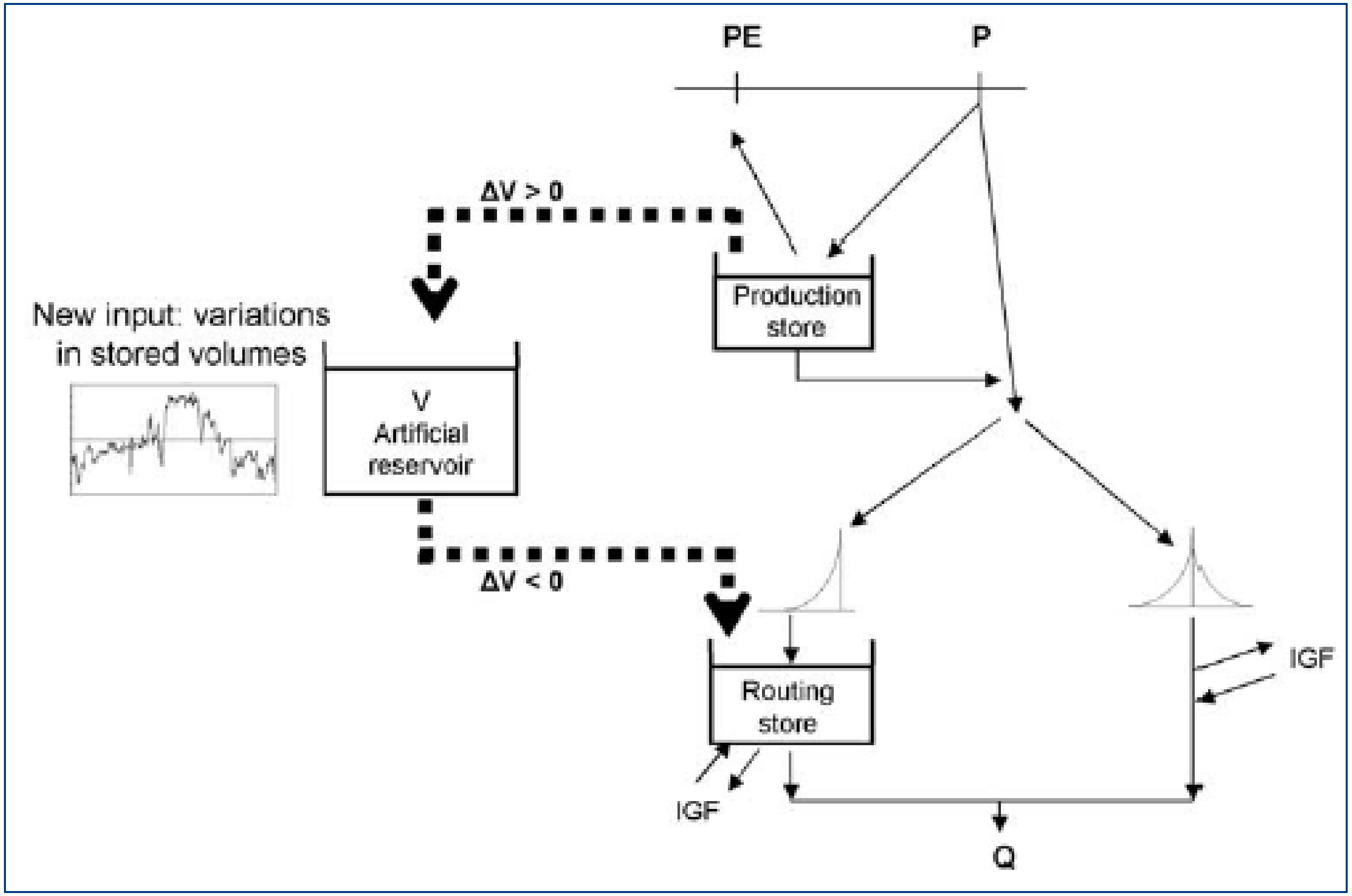
Name	Description
Quadratic criteria	
KGE	Kling-Gupta Efficiency
C2M	Nash-Sutcliffe Efficiency bounded in]-1 ; 1]
Low-flow quadratic criteria	
$C2M_i$	Nash-Sutcliffe Efficiency calculated with 1/Q and bounded in]-1 ; 1]
$RMSE_{ut}$	Root mean square error calculated when observed streamflow is less than Q_{80} threshold
Volume-based and temporal criteria	
Vdef	Ratio of observed and simulated cumulative annual volume deficits
LFD	Ratio of observed and simulated cumulative low-flow duration
DatSt	Relative difference between observed and simulated start of annual low-flow period
DatEn	Relative difference between observed and simulated end of annual low-flow period
Threshold criteria	
POD	Probability of detection, based on contingency table
FAR	False alarm rate, based on contingency table
CSI	Critical success index, based on contingency table

Simulation

Name	Description
Continuous low-flow quadratic and probabilistic criteria	
$RMSE_{ut}$	Root mean square error calculated when observed streamflow is less than Q_{80} threshold
DRPS	Discrete Ranked Probability Score
Volume-based and temporal criteria	
Vdef	Ratio of observed and simulated cumulative annual volume deficits
LFD	Ratio of observed and simulated cumulative low-flow duration
Sharpness/reliability	
Sharp	Mean width of interval defined by 10% and 90% percentiles of forecast distribution when observed streamflow is less than Q_{80} threshold
Cont_ratio	Percentage of observation in the 80% forecasted confidence interval when observed streamflow is less than Q_{80} threshold (80% of observed streamflow should be included in the interval)
Threshold criteria	
POD	Probability of detection, based on contingency table
FAR	False alarm rate, based on contingency table
CSI	Critical success index, based on contingency table
BS_{vig}, BS_{cri}	Brier Score with vigilance threshold (Q_{80}) or crisis threshold (Q_{95})

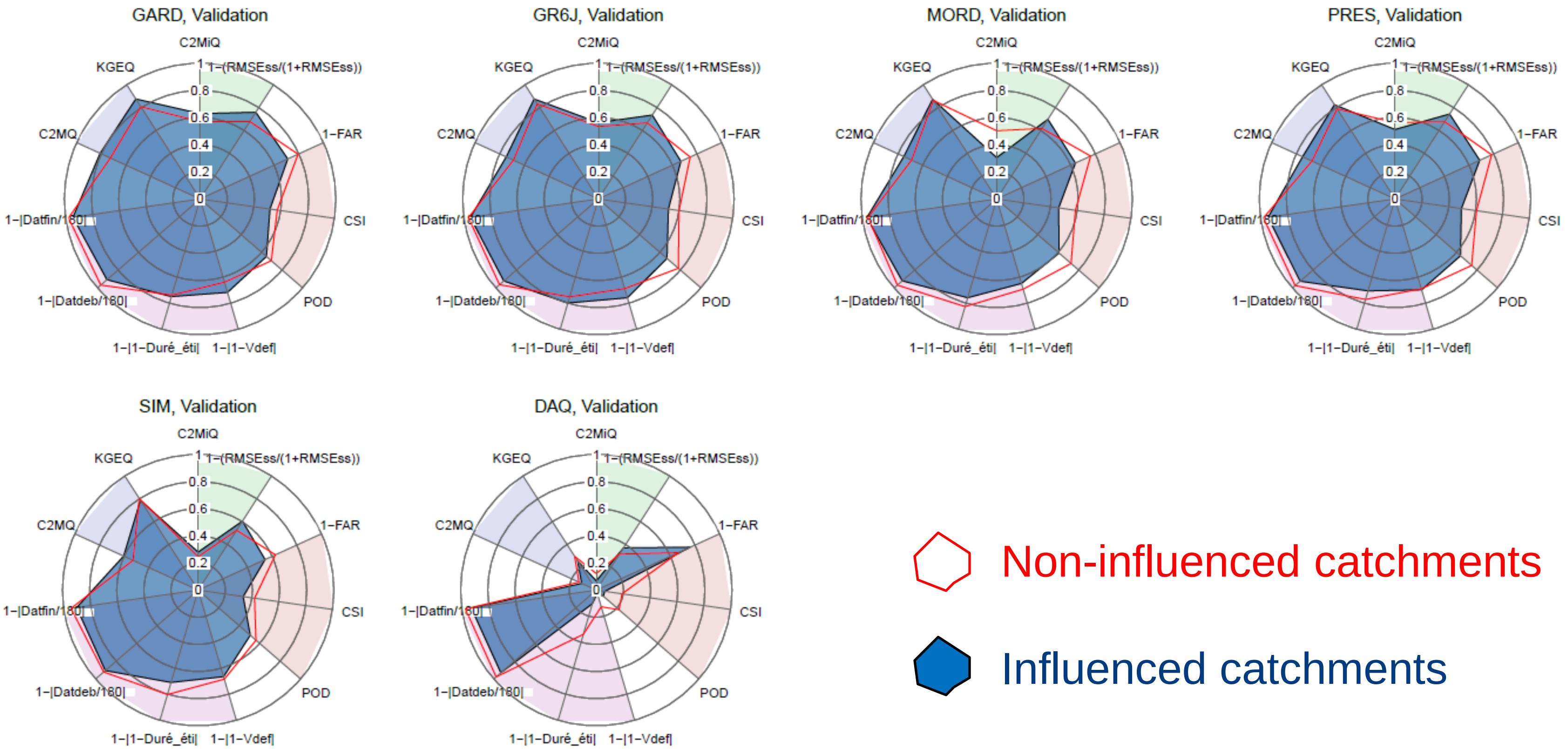
Forecasting

Modelisation of influences



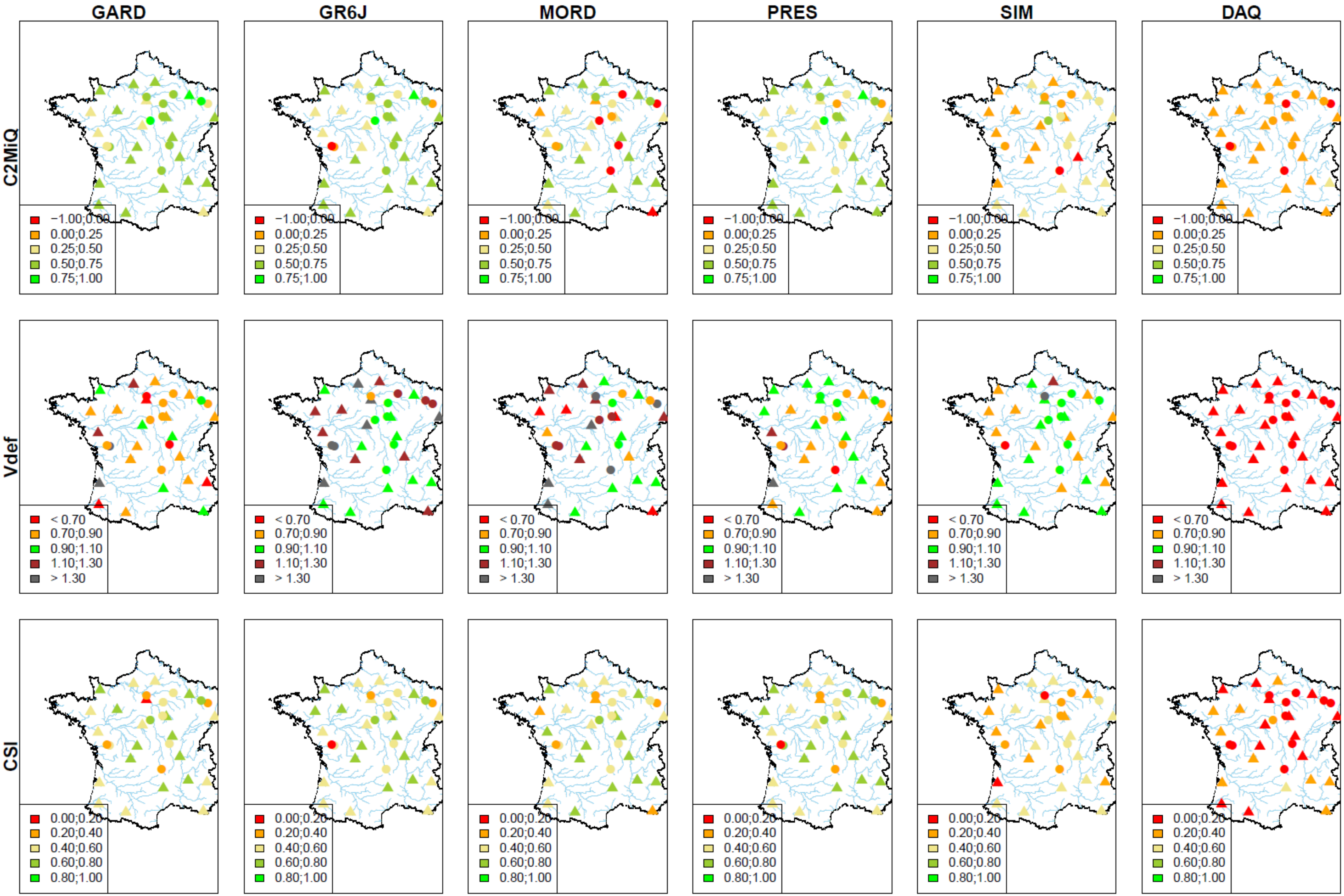
Source: Payan et al., 2008

Simulation results



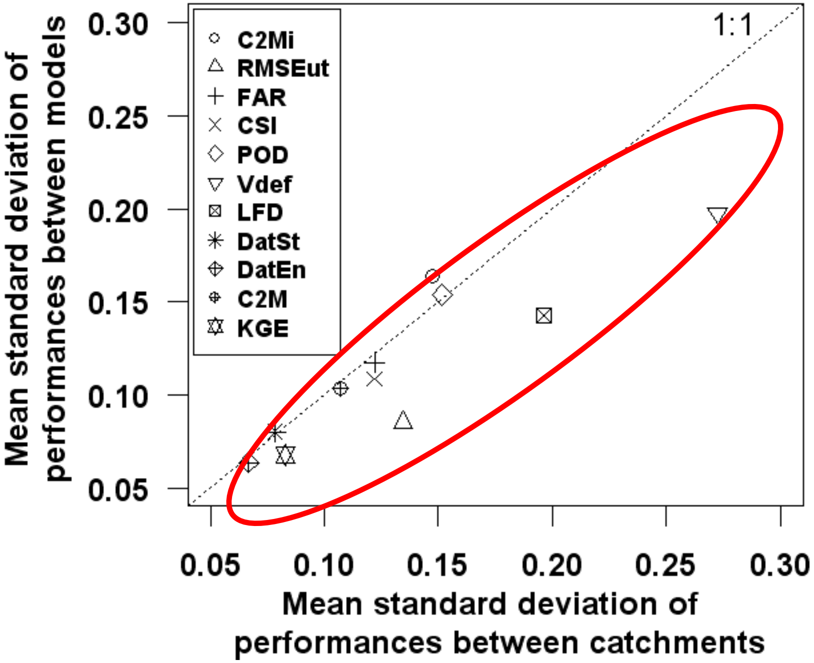
- Models similar on average, more difficulties for SIM
- Performances slightly better on non-influenced than influenced catchments
- Significatif gain compared to the benchmark

Simulation results

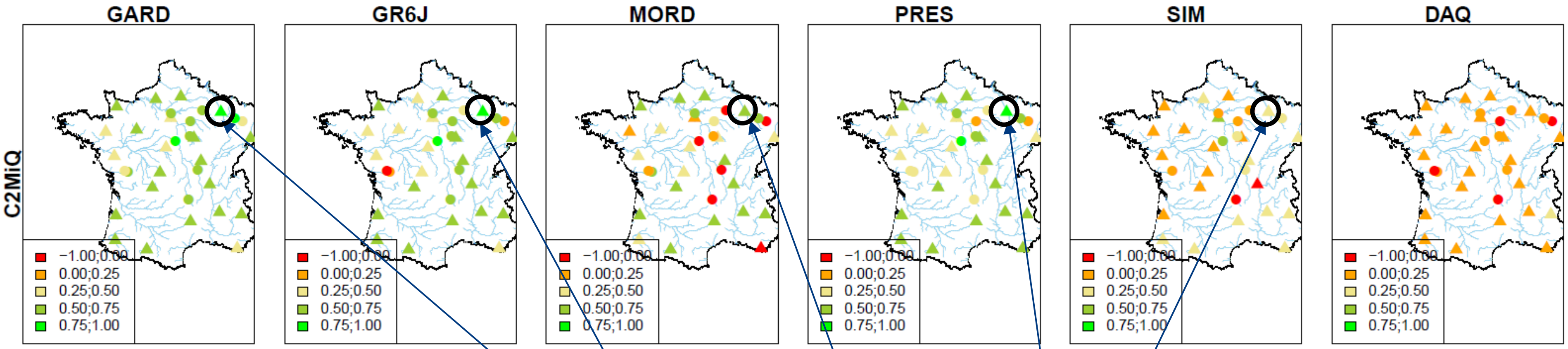


Simulation results

- Mean variability of performances between models:
  - ➢ For each catchment, standard variation of performances of models (sdm)
  - ➢ Mean of sdm
- Mean variability of performances between catchments:
  - ➢ For each model, standard variation of performances on catchments (sdc)
  - ➢ Mean of sdc



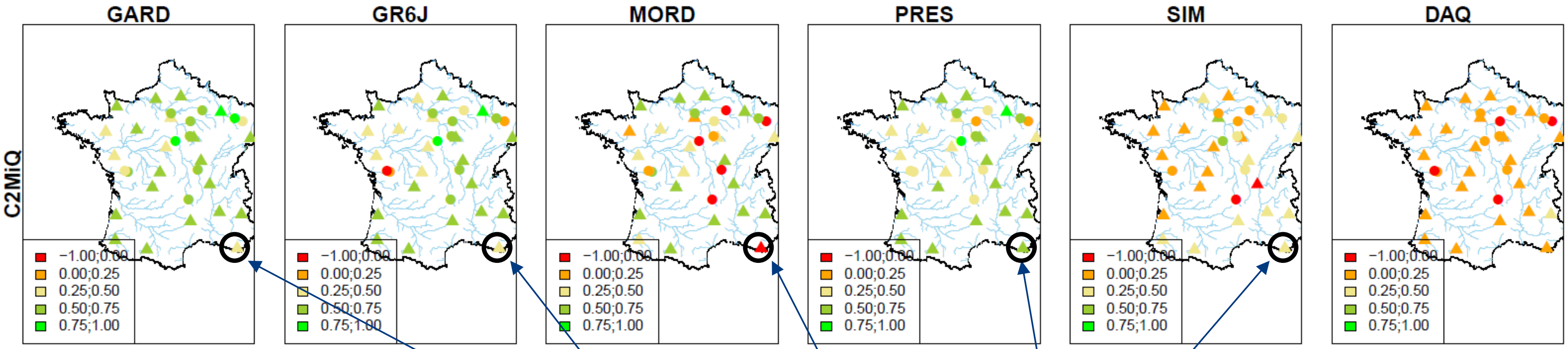
3. Simulation results



The Meuse river at St-Mihiel

- Catchments where all models simulate overall well streamflows

Simulation results

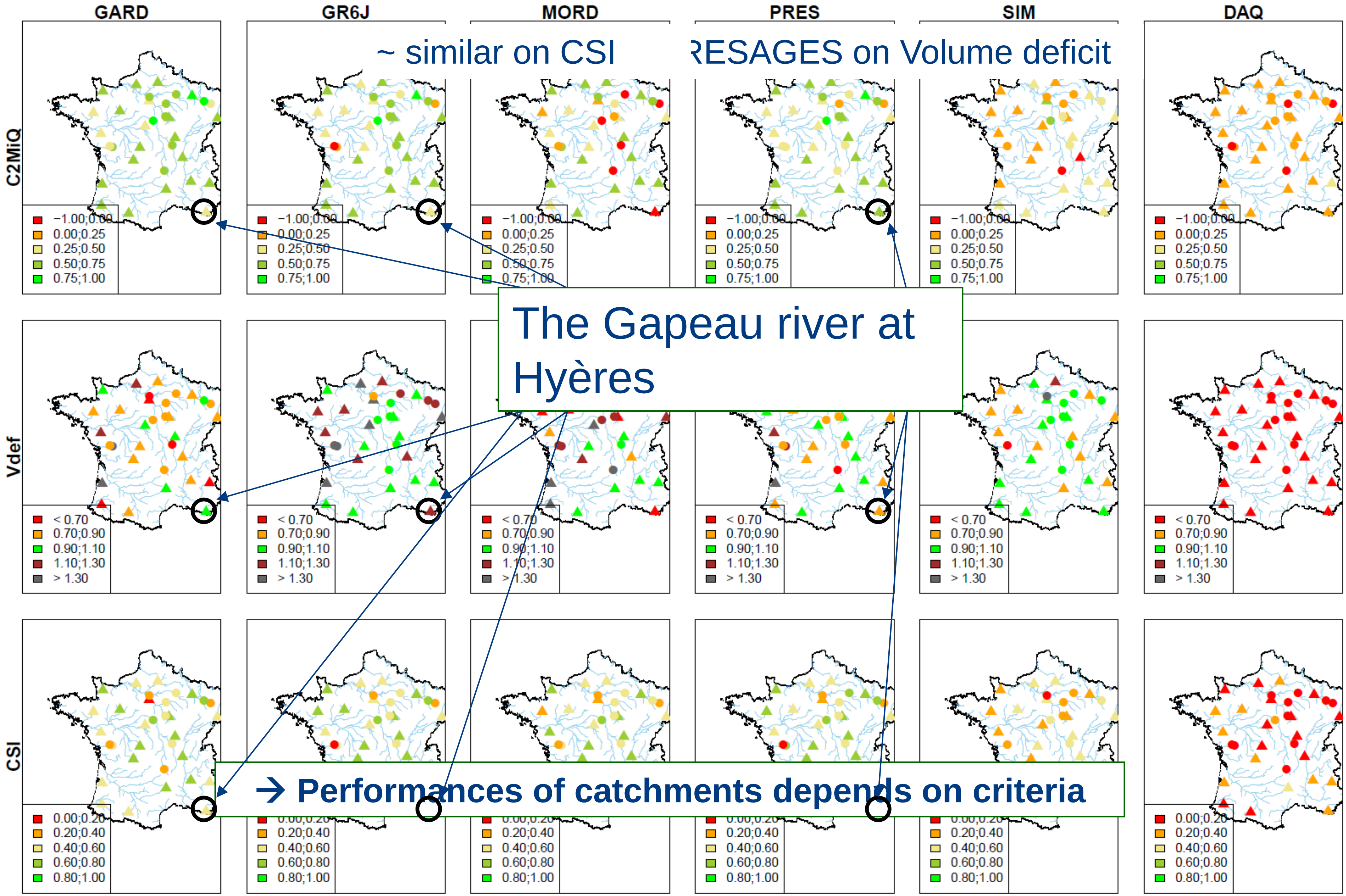


The Gapeau river at Hyères

- Other catchments where performances depends more on the models

→ Performances depends on catchments and models

Simulation results



~ similar on CSI

RESAGES on Volume deficit

The Gapeau river at Hyères

→ **Performances of catchments depends on criteria**