MULTIMODEL EVALUATION OF TWENTY LUMPED HYDROLOGICAL MODELS UNDER CONTRASTED CLIMATE CONDITIONS

Seiller, G.[1], Anctil, F[1]. and Perrin, C.[2]

[1] Chaire de recherche EDS en prévisions et actions hydrologiques, Université Laval, Département de génie civil et de génie des eaux, 1065, avenue de la Médecine, Québec, Qc, G1V 0A6, Canada
[2] Irstea, Hydrosystems and Bioprocesses Research Unit (HBAN), 1, rue Pierre-Gilles de Gennes, 92761 Antony Cedex, France

Introduction

There is a large consensus that the bulk of the adaptation strategies to climate change will be driven by water issues. Already, some components of the water cycle are of concern, such as precipitation frequency and intensity, snow cover, soil moisture, surface runoff, atmospheric water pressure, evapotranspiration, and others (Bates et al., 2008). These findings stress the importance of quantifying the **impacts of climate change on the hydrologic cycle** and evaluating related **uncertainties**.

The most common way assessing the impact of climate change on water resources combines the use of climate projections and hydrological modelling. Four main steps must be considered in such impact studies (Boé et al., 2009):
(1) Constructing gas emission/concentration scenarios,
(2) Modelling global climate,
(3) Downscaling and bias correcting the meteorological projections, and
(4) Estimating impact with hydrological models.

All these chained steps have associated uncertainties whose relative importance may differ between climate conditions and catchment characteristics.

Hydrological modelling in a climate change context

Difficulty in using hydrological models in these impact studies stems from the **non-stationary** nature of climate. Common practice usually assumes that parameters associated to the hydro-climatic conditions of the calibration data set remain valid in other test periods, making implicit the assumption of the stationarity of the rainfall-runoff transformation.

However, in a climate change context, the contrasts of climate conditions between the calibration and projection periods are important, thus questioning the stationarity hypothesis.

Hence **model transposability in time under contrasted conditions** must be analysed in details and could even become a criterion for the selection of modelling tools to be used in impact studies. To this end, demanding validation methods must be designed.

Objectives

The main objective is to explore the structural uncertainties of a selection of twenty lumped conceptual models and to quantify their robustness when climate conditions strongly differ between calibration and validation, following two application modes: individual and collective (ensemble).

Our analysis mainly addresses the following two questions:

- What is the level of appropriateness of each selected model, in terms of transposability in time (i.e. performance and robustness) under contrasted conditions?

- Is there any added-value using all these models together or a subset of them based on their performance and transposability in time?

Material & methods

To answer these questions, the twenty hydrological models are evaluated individually and collectively under the Differential Split Sample Test (DSST) framework on two catchments, Au Saumon in Canada and Schlehdorf in Germany.

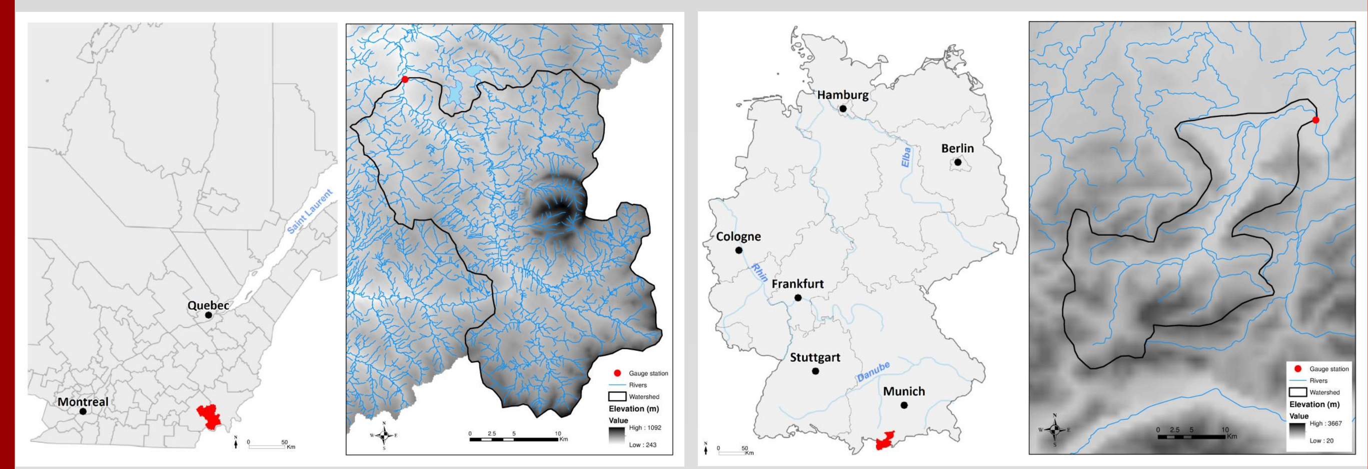

Figure 1 : Location of the studied catchments

Model intercomparison has been identified as a convenient mean approaching this project. This may provide information on model complementarity and thus open ways to create **multimodel combinations** with improved efficiency.

Twenty lumped conceptual hydrological models were selected, to get a wide variety of conceptualizations of the rainfall-runoff relationship. The choice of these models is mainly based on known performance and structural diversity, i.e. 4 to 10 free parameters, and 2 to 7 storage units.

Most of these model versions originate from the works by Perrin et al. (2001).

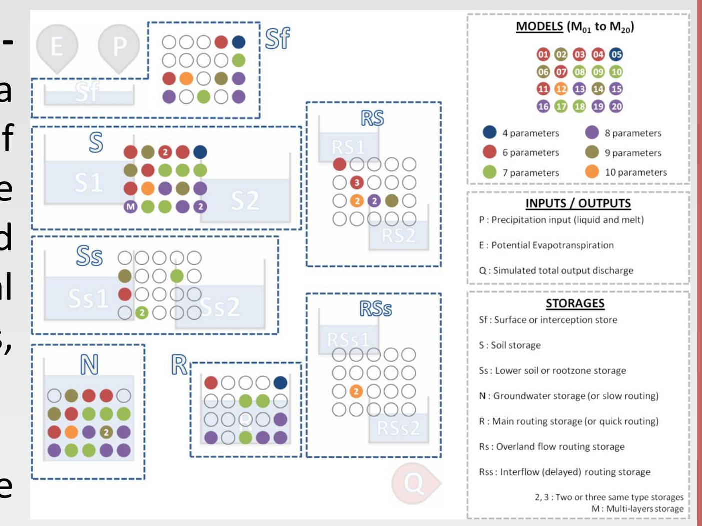

Figure 2 : Illustration of model structural diversity

The transposability in time of hydrological models is assessed and used as a criterion for the selection of appropriate projection tools based on **Differential Split Sample Tests** (Klemeš, 1986). The idea is to calibrate the model on a time series with selected characteristics and to validate it on contrasted time series, placing the model in a demanding situation.

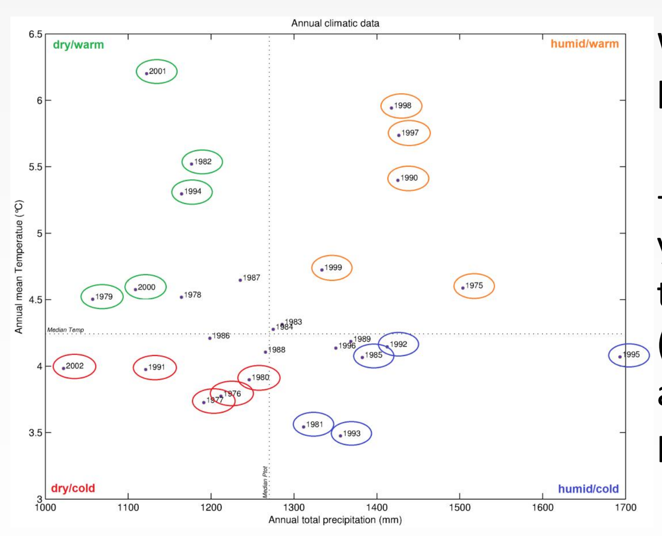

Figure 3 : Time series clustering results for Au Saumon

We applied the three-step procedure below to our set of twenty models:

- Select five non-continuous hydrologic years (1st Oct. to 30th Sept.) for four contrasted climate conditions: dry/warm (DW), dry/cold (DC), humid/warm (HW), and humid/cold (HC), based on annual precipitation and temperature,

- Calibrate and validate on contrasted time series: DW→HC (calibration on DW and validation on HC), HC→DW, DC→HW, HW→DC.,

- Evaluate model performance using preselected criteria (i.e. Nash-Sutcliffe efficiency) and comparatively assess the relative transposability of the tested models in the various configurations: DW→HC, HC→DW, DC→HW, HW→DC.

Results

Results are analysed following two points of view: individual (models comparison) and collective (ensemble analysis).

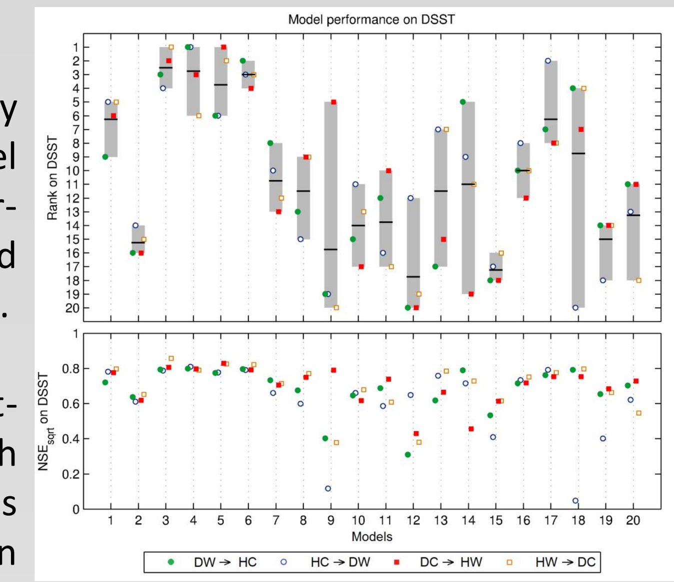

Figure 4 : Performance and rank in validation for the DSST—Schlehdorf catchment

Individual results illustrate the difficulty in identifying a single lumped model that could behave well in terms of performance and robustness, when tested under all possible contrasted conditions.

Our tests allow identifying best-compromise individual models for each catchment. This better robustness is quite difficult to explain solely based on the analysis of model structure components and clearly dependent on the test catchment.

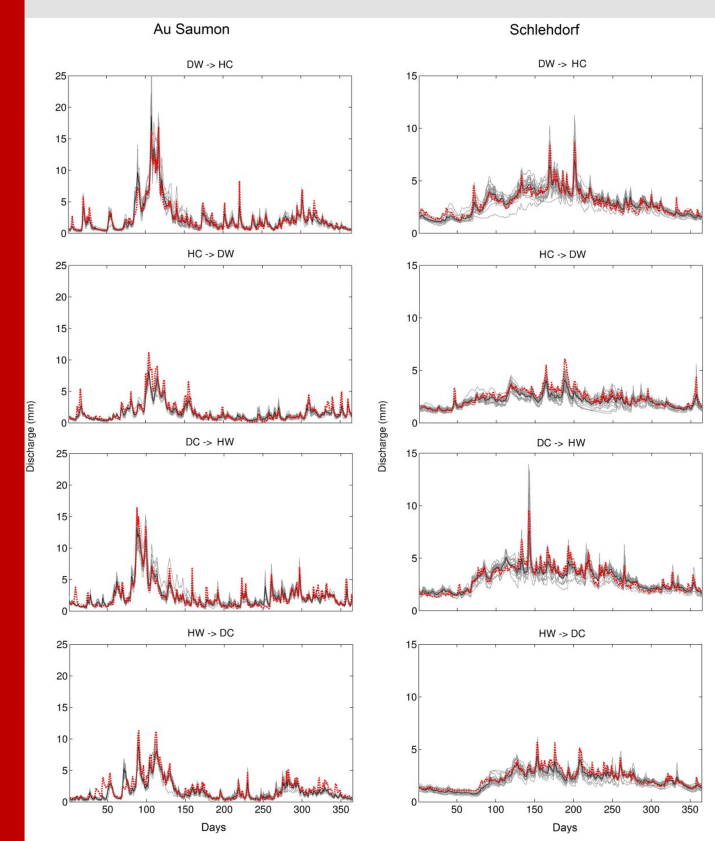

Figure 5 : Mean daily interannual discharges for all the DSST. Individual models (grey), twenty-member ensemble (black) and observed discharge (red)

A deterministic **multimodel** ensemble analysis, taking the average of simulated streamflow series as output, was next performed. We explored the simple twenty-member ensemble and all other models combinations.

The **twenty-member ensemble** gives better results than the best individual model for all DSSTs on the Au Saumon catchment. Although the improvement is not large, it is substantial in all cases.

This holds for only one of the four Schlehdorf DSSTs. Nonetheless, the multimodel approach remains a valuable alternative since the best model is different for each DSST, a sign of a lack of climate transposability.

Results also reveal that many other **model combinations** (sub-selections) provide better performance than the twenty-member ensemble.

We identified model combinations that also provide enhanced robustness relative to the DSST.
With these efficient and robust ensembles, we evaluated the collective interest of each model, (i.e. the added-value of the structure for an ensemble approach in a climate change context for each catchment).

A link exists between individual and collective interest for Schlehdorf catchment but not for Au Saumon.

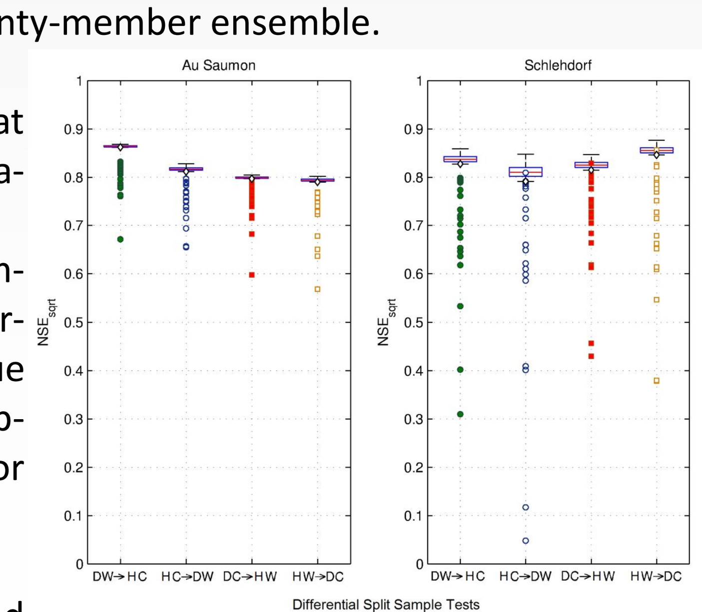

Figure 6 : Individual and multimodel DSST validation performance. Boxplots (best multimodel combinations), diamonds (twenty-model ensemble), and the circles and squares, the individual models

Conclusion

Evaluating hydrological model behaviour under contrasted conditions is, in our opinion, a **pre-requisite** to climate change applications. This approach based on Differential Split Sample Tests allowed climate transposability evaluation of all twenty individual models, along with their collective qualities.

The analysis of the individual value of each lumped model showed that it is unsafe to rely on a single lumped model, unless it is handpicked for each specific catchment as highlighted by **best-compromise models**.

Taken together, the twenty models offered better climate transposability, as if the many model structures compensate for one another's weaknesses, as illustrated by several results. More, this is the only approach that was successful for both catchments, indicating a strong potential for catchment transposability.
Pushing further the **ensemble philosophy**, almost all possible model combinations have been explored. Many combinations were found to provide increased performance over the twenty-member ensemble.

It is also noteworthy that even if best performing models may more likely contribute to the ensemble, worse-performing individual models can also successfully contribute to the ensemble.

Acknowledgements

The authors acknowledge NSERC, Ouranos, and Hydro-Québec for support, as well as the other partners in the QBIC[3] project.

References

Bates, B., Kundzewicz, Z. W., Wu, S., and Palutikof, J., 2008. Le changement climatique et l'eau - Rapport du Groupe d'Experts Intergouvernemental sur l'Évolution du Climat (p. 237).

Boé, J., Terray, L., Martin, E., and Habets, F., 2009. Projected changes in components of the hydrological cycle in French river basins during the 21st century. Water Resources Research, 45(8), 1-15.

Klemeš, V., 1986. Operational testing of hydrological simulation models. Hydrological Sciences– Journal– des Sciences Hydrologiques HSJODN, 31(1), 13-24.

Perrin, C., Michel, C., and Andreassian, V., 2001. Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. Journal of Hydrology, 242(3-4), 275-301.

Seiller, G., Anctil, F. et Perrin, C., 2012. Multimodel evaluation of twenty lumped hydrological models under contrasted climate conditions. Hydrology and Earth System Sciences, Accepted.

Contact

Grégory SEILLER, PhD Student
Chaire de Recherche EDS en prévisions et actions hydrologiques
Université Laval, QUÉBEC, CANADA
gregory.seiller.1@ulaval.ca