## Model testing for improving flow simulation and uncertainty quantification: what can we learn from large catchment sets?

The downward approach that learns from observations the main features of the catchment hydrological response has long been recognized as a way to develop hydrological models at the catchment scale (Klemes, 1983). In this approach, a link is made between rainfall inputs and flow outputs using the mathematical tools found the most efficient to reproduce catchment behaviour. This approach received more attention in the last decade, as some limitations of the upward approach were identified (Sivapalan et al., 2003).

However, model structures developed with this downward approach at the catchment scale are often difficult to generalize, i.e. difficult to transpose to other catchments. Indeed they are often over-adapted to the specific features of the catchment on which they were developed. Generalization is a major problem in current hydrological modelling (Sivakumar, 2008). This is potentially a major drawback for the application of such models to the case of ungauged catchments.

We argue that a better way to develop the structure of hydrological models following a downward approach is to place model development at the level of large set of catchments and not only at the level of a single catchment (Andréassian et al., 2006). This way of developing models will force them to be general, i.e. more transposable in space. They will capture the essential features of the rainfall-runoff transformation common between catchments. This way of developing models also gives a number of opportunities in the analysis of the spatial patterns of model failures, or the empirical quantification of uncertainty by intensive model testing. However, this requires well-defined testing schemes and criteria for model evaluation. We will illustrate the advantages (and possible limitations) of this approach using examples drawn from our past and current research activities based on large data sets. A focus will be given on the diagnosis of model structures and the quantification of corresponding uncertainty, discussing issues of model complexity and the current level of reliability of hydrological models (Michel et al., 2006).