Detailed response to the comments of M. Demirel

We want to thank M. Demirel for his valuable comments and suggestions about our paper. In this author comment, we give our answers to its remarks.

[1] We agree that spatial metrics, such SPAEF proposed by M. Demirel, may be very usefull to calibrate distributed models on spatial observations. SPAEF was developed to compare two maps (for example, actual evapotranspiration over the basin) but here, we are evaluating the streamflow at some pseudo-ungauged locations but not maps. Therefore, we apply the KGE over the whole time series of these pseudo-ungauged stations to assess the performance of the regionalization.

[4] The model is conceptual since its parameters do not stand for physical properties of the catchment. They are parameters controlling the several processes of the hydrological cycle described in a conceptual way. On the other hand, the catchment is spatially distributed as it is divided into hydrological meshes (= small subcatchments). Every hydrological mesh is associated with its own set of conceptual parameters and its own climatic inputs and its own state variables.

[5] We are aware that the terminology of "semi-distributed" and "distributed" varies from one hydrologist to another. In this paper, the term "semi-distributed" means that climatic inputs are distributed but not the hydrology (the parameters are identical over the whole catchment, and the spatialization is not done according to hydrological sub-catchments : we cannot simulate the streamflow at interior points). On the contrary, the term "distributed" means that both the climate and the hydrology are distributed. The catchment is divided into elementary sub-catchments according to a target size (of 100 km²) and a Digital Elevation Model. In the distributed model, a streamflow simulation can be obtained at any mesh oulet of the model. Then, even if we do not use regular meshes (square or triangular grids) but irregular meshes derived from the DEM, the model may be qualified as spatially distributed.

[6] The target size of the meshes is 100 km². To divide the catchment into hydrological meshed, an algorithm then tries to minimize the variance of the area of the meshes around the target size. Therefore, the hydrological meshes are sub-catchments of approximately 100 km².

[7] To assess the spatial performance, we apply the KGE over validation stations whose streamflow is never used to calibrate the parameters. We assess the performance at pseudo-ungauged stations over their whole streamflow time series. Therefore, since theirs streamflow time series are never used to calibrate the parameters, theirs KGE quantify the performance of the parameter regionalization.

[8] We agree with M. Demirel that use of other observations, as AET estimations, may help to better identify some of the parameters. However, our experience is that on the considered regions, such observations suffer from significant biases. Therefore, including them in the calibration process may degrade parameter estimation and model performances. It is why we concentrate on streamflow data. Moreover, the four signatures are about runoff but they are related to different hydrological processes. They allow to evaluate the model both in seasonality, current flows, low flows and high flows.

[9] The IDPR appears in figure 5 to facilitate the comparison with the other spatial patterns. This index is used to prescribe one of the parameters (as explained in section 4.2.2). So, the IDPR is the basis of the regionalization of one parameters.

[10-11-12-13-14] Details about the sensitivity analysis are not are not given in this paper to avoid overloading the article since the sensitivity analysis would deserve a whole paper. However, we can

say that the sensitivity analysis is conducted with a quasi-Monte Carlo method over the four KGE criteria (daily runoff, seasonality, flood and low flow) plus the KGE over the fractional snow cover. A total of 45.056 combinations were tested over each basin and required days of computation.

[15] Before (Exp1), the crop coefficient was uniform. In the paper, we propose to make it spatially variable (Exp2), based on the NDVI map obtained by satellite. Kc is therefore spatially heterogeneous. "*Previously*, it was obtained through a one-parameter formulation where the parameter was set uniformly to its default value. *From now on*, the inter-annual time series of NDVI (16-day and 1-km² resolution) is aggregated at the mesh scale and then used to prescribe Kc at the same scale."

[2-3-16-17-18] Thank you for the papers suggested, which are really relevant. We added some of them in our paper.