We would like to thank reviewer 2 for his/her positive feedback and the detailed suggestions of how to make the manuscript easier to read and to understand.

A reduction of the number of examples is certainly a good starting point. The toy examples will be removed. The synthetic ARMAX and HYMOD examples have been introduced to illustrate the method for simple models that are widely known and for time series not showing a strong periodic component. We will remove the HYMOD example and keep the ARMAX example for this purpose.

Specific comments:
i) We will simplify figures 1, 7 and 8 as suggested and remove figure 4 and the corresponding text.

ii) For experiment 3 and 4, only the parameters for the non glacier part of the catchments were used, which explains why Table 3 does not show all the parameters of the model. This will be specified.

iii) Meaning of Fig. 7b. Model calibration starts with some prior plausible parameter ranges, which are then reduced based on the available information for a given catchment (the parameter ranges are updated). This reduction depends on the chosen calibration method. In the extreme case, we will reduce the prior ranges to a single value for each parameter (in the case of single-objective optimization). Here, for illustration purposes, we have chosen a Monte Carlo simulation approach; we have randomly drawn a large number of parameter sets in the prior ranges and then retained only the ones that give an acceptable bias. This shows where the physically meaningful parameter sets lie within the prior ranges (light grey triangles in Fig. 7b). Within this "updated" range, we further highlighted the best 100 parameter sets under the time-domain performance criterion (black triangles) and under the wavelet domain criterion (red circles). The best 100 time-domain parameters cover a sub-space of the physically meaningful parameters, i.e. the use of a time-domain performance criterion further constrains the parameter ranges. This is a known result as mentioned in the manuscript (p. 2476, line 21). The 100 wavelet-domain parameters do not further reduce the physically meaningful parameter range, as illustrated in Fig. 7b, where the red circles cover about the same range as the grey triangles. This and the corresponding conclusions will be better explained in the text. We will keep this part since it illustrates, to our view, very well how the proposed method can be used to obtain additional insight into how the model, the data and the performance criterion interact during model calibration.

iv) Fig. 5 (why does this figure show different parameters for the real-world and the synthetic case?): we wanted to give a "sample" of the different possible relationships
between the objective function and the parameter values. Since the real-world case study uses a very different model set-up (catchment with glacier, very strong annual cycle) than the synthetic case (no glacier, much less seasonality), the relationship between a given parameter and the objective function for one or the other case study would not be comparable. We will specify this.

v) Other technical comments: we will correct the manuscript as suggested.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 6, 2451, 2009.