Dear Prof. Efstratiadis,

thank you very much for your helpful comments and corrections. We tried to edit the manuscript accordingly in the best possible way.

Kind regards
Thomas Krauße

Answers to the (minor) comments and technical corrections

1. Some citations, indicated with ““?”” are missing.

Done

2. The right semicolon of p. 3 and the left semicolon of p. 5 are almost identical.

The first duplicate paragraph was a leftover of the old version. The text was edited accordingly.

3. Page 6, line 8: The transferability of a specific set of optimized model parameters to other catchments is, to my opinion, questionable. Under which premise is this transferability possible?

Bardossy & Singh (2008) and Singh (2010) showed that robust parameter vectors (estimated by drawing deep points from a set of good parameter vectors estimated calibrating several catchments) showed a better prerformance being used for parameter regionalisation. However, this aspect is no focus topic of this paper. Thus, we added an explanatory footnote.

4. Page 6, last paragraph: Change to read ““….… is also known as Turkey depth (Turkey, 1975).””

Done.

5. Page 6, last paragraph: Change x_1 by ␣1.

Done.

6. Page 7, section 2.3: Put Algorithm 3 on the top of the page.

Done.

7. Page 14, paragraph 3: Change to read ““The advantages of the depth based sampling””.

Done.

8. Section 4 is too long. I propose to split it into three subsections, the first one dealing with model and study area description, the second one presenting the 2-criteria calibration and the third one presenting the 3-criteria case study.

The manuscript was adapted accordingly.

9. Page 15, Figure 8: Part of the caption is unclear (w uncertainty. cater plot….…).

The caption was edited accordingly.

10. Page 17, Table 9, caption: Instead of ““number of observation points”” (which is ambiguous,

since it may refer to a number of gauging stations), use "number of observations" or "data length".

Done.

11. Page 17, paragraph 2: "The model parameters considered for calibration are the conceptual model parameters k_d, k_i and d_r. ……… Furthermore, we considered the conceptual model parameter k_{rec}………". In Fig. 13, two additional parameters are shown, i.e. $␣_{SL}$ and $␣_{SiL}$. Yet, according to Table 8, these are not (scalar) parameters but vectors of parameters.

The vector of soil hydraulic parameters are aggregated by a scaling parameter – one per soil. This is done using a procedure described by Grundmann (2010) as given in the paper. The scaling parameters can be easily „re-mapped" to a vector of soil hydraulic parameters. Therefore, they can be easily used as calibration parameters.

12. Page 18, last paragraph: Change to read "……… is almost convex………"

Done.

13. Page 19, first paragraph: ""Lower values of k_d increase the dynamics of the generated direct runoff, a higher value of k_{rec} decreases the effective saturated conductivity of deeper soil layers. This leads to a faster generation of direct runoff. However, within the Rietholzbach catchment direct runoff on the surface has hardly ever been observed, even not during large and intensive convective storm events. Thus these results are already an evidence for shortcomings in the model structure."" Although it is known that the direct runoff component does not play an important role in the runoff mechanism, I cannot understand why the authors used a so extended range for the corresponding parameter k_d, instead of restricting the feasible search space to prohibit the generation of low values of k_d. This is rather a bad calibration practice, leading to parameter values that are inconsistent with their physical interpretation, and not a structural deficiency of the model (cf. Nalbantis et al., 2011).

The calibration of WaSiM for flood events in the Rietholzbach catchments adapts the paratemers in such a way that the drainage runoff is represented by the direct runoff component. We had intensive discussions with other research groups with a lot of modelling experience with WaSiM. According to the consolidated results of all the different opinions we decided to set the parameter boundaries as given in the paper. It is very difficult to assess the „real feasible parameter range" especially for runoff events with high dynamics. Tests with different parameter boundaries showed that our first guess is at least a resaonable way to calibrate WaSiM. The estimated parameter vectors and model performances (in both calibration and validation) did not differ significantly.

14. Page 22, paragraph 2: Change to read "The model performance………"

Done.

15. Page 25, line 3: Change to read "in not non-dominated".

Done.

16. Page 26, first paragraph: ""We suggest further studies in operational flood forecasting studies for fast responding medium scale catchments where such information is available."" Is this a realistic option? In operational flood forecasting, the basin area may be of hundreds or even thousands of km^2. Even at the 3 km^2 scale of the Rietholzbach catchment, a single soil moisture measurement point proved to be insufficient. How much dense should be this information for larger basins?

We agree that typical basins with an operational flood forecasting cover hundreds or even thousands of square kilometers. They are usually represented by semi-distributed models or grid models with a significantly larger cell size. In relation to the used grid resolution already a few soil moisture measurements would provide much more information than on single measuremt in the Rietholzbach catchment with 50mx50m cell size. Furthermore spatially distributed satelite measurements could be used to interpolate the loacal measurements.

Nonetheless, we agree that a good simultaneous representaion of both runoff and the soil moisture dynamics is a very challenging task. Our comment should just point into this direction without neglecting the possible difficulties of such studies.