Authors Reply to the Anonymous Referee Comment #3 on "Are streamflow recession characteristics really characteristic?" by M. Stoelzle et al.

This paper compares different methods used for estimating recession parameters based on the work of Brutsaert and Nieber [1977]. The paper is timely and interesting. It may provide a good operational guideline on data selection and model fitting for other researchers who conduct recession analysis. It will be better if the authors can revise the paper with an emphasis along this line. I agree with other reviewers that the title needs to be changed. I also have three major comments:

We thank the Anonymous Referee #3 for the thoughtful comments and the helpful suggestions on our manuscript. Please see below for our detailed answers and suggested revisions (in blue).

1. The three methods to fit the data points of –dQ/dt ~ Q are different. Particularly when the lower envelope method is compared with the "reg" and "bin" methods. It may be better to discuss the advantage or disadvantage of the three methods or the conditions applicable. I encourage the authors to discuss the data selection criteria and the methods to fit the data instead of concluding that all the methods should be used to represent the uncertainty of parameter estimation. For example, it would be great that the authors can recommend a set of recession data selection criteria that will reduce the bias on parameter estimation. The set of recession data selection criteria can be from the three methods (VOG, BRU, KIR) or others.
Authors Reply: We agree that the methods based on different recession extraction procedures and different fitting models are based on different perceptual models. Interestingly we found that the distribution of slopes of the REG-method is comparable to the distribution of slopes of the LE-method, when $b < 1.5$ and that the distribution of slopes of the REG-method is comparable to the distribution of slopes of the BIN-method, when $b > 1.5$ (figure not shown). Thus, none of the methods is entirely distinguishable from the others. We have discussed advantages and disadvantages of different extraction procedures and parameterization methods (cf. 5.2 discussion paper) and found that each method has its strength and weakness. But as the "true" values for the recession model parameterization are unknown, we cannot choose one best method or recession extraction procedure. However, we strongly believe that the paper provides enough evidence to the reader that for a specific question a selected suite of extraction procedures and fitting methods should be used to get a good idea about the differences among the methods.

2. For the lower envelope, it may be better to use two lines for early and late recessions [Brutsaert and Nieber, 1977]. As shown in the figure below, the red line is for the early recession. The slope in early recession is larger than that in late recession [Brutsaert and Nieber, 1977]. This is also reflected in the "bin" method, particularly in the figure of VOG.bin. From this perspective, it is a disadvantage of "reg" method to assume constant recession slope. Meanwhile, it should be noted that b values for the three fitting methods in Figure 1 are based on different recession stages: b for "le" is based on late recession, b for "reg" is based on entire recession, and estimation of b for "bin" excludes the very high values of Q. This may contribute to the variations of estimated b values.
Authors Reply: We are in accordance with the suggestions of the referee, but as we want to compare the parameterization of different RAMs we cannot implement two envelopes for early and late recessions since this would not allow us to compare this method with the other methods. Further on, we are aware of biases in the estimated b values and will discuss this issue more detailed in the revised paper.

3. Lines 16-20 on page 10577, the author mentioned the difference between individual recession event and collected recession events. The authors may discuss this with more

details since this paper is focused on the methods of estimating the recession parameters. For example, the slope of individual recession event is affected by initial groundwater condition and evapotranspiration and may vary from event to event. Analysis based on a collective recession events may represent the average recession behavior, particularly, the lower envelope has strong physical basis, i.e., the minimum recession rate is corresponding to the groundwater discharge.

Authors Reply: This is a very interesting point by the referee. We extracted single recession events for some catchments (results not shown) and plotted the single –dQ/dt-Q-events over the recession plot and derive slope b with linear regression. According to our expectations we found both recession with larger and smaller values of b compared to the "main" b derived by the recession plot with all collected recession in it. Interestingly in some cases the "lower envelope" of the recession plots were formed by the last days of different recession events rather than by one lower envelope with a distinct slope b (Shaw and Riha, 2012). However, if we assume that individual recession events are driven by different hydrological conditions and processes (i.e. groundwater condition, evapotranspiration, seasonal drainage behavior and contribution of storages), then one must consider that this heterogeneity may be concealed by collected recession events in one recession plot (Shaw and Riha, 2012). Accordingly, one may use different RAMs to estimate storage-outflow parameterization for different stages of recession (e.g. binning for recession's early stages and lower envelopes for long time aquifer response), thus to characterize catchment-specific storage-outflow behavior with more detail (i.e. seasonality).

Nevertheless, the scope of this paper is not a comparison between individual and collective analysis of recession events; instead we compare established recession analysis methods to illustrate the inconsistence (uncertainty/bias) of a single method approach and we want to verify whether the provided methods found recession characteristics as a function of catchment characteristics or not.