Response to reviewer 1

We thank the reviewer for their positive feedback and constructive suggestions to improve the manuscript. The reviewer's main points relate to (i) the need to improve the overall readability of the manuscript and (ii) the need to clarify the contribution of the specific parameter estimation method used on the prediction error of the process-based method. We will address these two issues in the revised manuscript as follows

Manuscript presentation

We agree with the reviewer that the readability of manuscript would benefit from a simpler and more explicit structure. We will modify the reviewed manuscript as follows:

- **Methods**: we will add a flow chart to graphically describe the cross-validation and Monte Carlo analyses that we ran. This will allow Section 2.3 (p9779, l5 –p 9781 l10) to be substantially shortened, and will help clarify the methods.

- **Discussion**: The procedures and results related to the complementary analyses currently presented in the discussion (Figures 5 and 6 and Appendix C) will be moved to supplementary material. These analyses will only be referred to in the Section 4 (which will be substantially shortened) to support the discussed hypotheses.

Parameter estimation of the process-based model

We agree with the reviewer that the analyses presented in the manuscript do not disentangle errors arising from the physical model from parameter estimation uncertainties. The former type of error (physical model) was assessed extensively in Muller 2014 for several seasonally dry climates (Nepal, Western Australia and California) and showed that the process-based model performed well, when its parameters were estimated using observed streamflow. In contrast, the purpose of this study is to assess the model's ability to be used as a streamflow prediction tool in *ungauged* basins. From this practical standpoint, parameter estimation uncertainties are an integral part of the prediction error of the assessed models. Arguably, process-based models allow parameter estimation uncertainties to be substantially decreased because they relate streamflow to observable parameters (e.g., rainfall statistics). By comparing the process model to a baseline statistical model, we wanted to investigate, whether this advantage outweighs the more restrictive process assumptions in operational situations.

To emulate operational streamflow predictions for the specific case of Nepal, we selected the parameter estimation method that we found to be most adapted to the local context (data availability and spatial scale of climate and runoff processes). In the spirit of assessing the operational performance of the models in ungauged basins, parameter estimation uncertainties and modeling errors are lumped in the results presented in Section 3. However, we do attempt to disentangle the two sources of errors in section 4.1, where we discuss likely sources of prediction uncertainties in both models. For the process-based model, errors on high flows are driven by the estimation of rainfall parameters, while errors on low flows arise from simplifying assumptions on the seasonal recession.

To make these points more explicit, we will add the following points in the revised manuscript:

- Introduction:
    - We will clarify the fact that the predictive ability of the process-based model *in gauged basins* has already been assessed. We will provide a more extensive overview of the related results discussed in Muller 2014.
    - The purpose of this is to evaluate its performance in *ungauged basins*. In particular, we will assess whether the advantages of the process based method (observable parameters) outweighs its drawback (process assumptions), and allows it to outperform a baseline method that purely relies on parameter estimation.
    - It follows that (i) we consider the parameter estimation techniques that are most adapted and likely used in the context of Nepal and (ii) we don't disentangle model errors from parameter estimation errors in the results presented in section 3, although we discuss error sources in Section 4.1.

- Conclusion:
    - We will note that prediction performances of both approaches (statistical and process based) are strongly affected by the method selected to estimate the model parameters.
    - Although we believe that the selected methods are appropriate to compare process-based and statistical approaches for practical PUB application in Nepal, this study cannot be interpreted as a general benchmark to compare these approaches at a global level.
    - Substantial research remains to be done to compare these approaches in other parts of the world, where locally appropriate methods should be carefully considered.

Response to specific comments

1. We changed the title to:

   Comparing statistical and process-based flow duration curve models in ungauged basins and changing rain regimes"

2. Thanks for your comment. It is true that prediction uncertainty errors on high flows are driven by rainfall estimation error. Yet in can be argued, that errors arising from the spatial scaling of rainfall are, in a way, a violation of the model assumption of (spatially) homogenous precipitations.

3. That is a very valid point: thanks for bringing it to our attention. We will remove p9769, l. 12 from the revised manuscript.

4. Thanks for your comment: this point will be clarified in the reviewed manuscript.

5. We will add on p9772 l.9:

   "The model admits an additional input parameters the scale a of the of the power-law, dry season recession, which we showed in Muller 2014 can be expressed as a function of k, b, λ_P and α_P"

6. That is a very valid point. We chose not to regionalize mean flow because that would make the process-based method very similar in principle to the considered statistical method that relied on the interpolation of index flows. We wanted to specifically assess the performance of the process based method as an approach that relies on observed rainfall statistics instead. Also, regionalizing rainfall instead of streamflow allows us to be consistant with the analysis conducted in Section 2.3.2 relating the prediction over change, which is based on the generation of synthetic rainfall. We based our decision to use a reasonable constant value for ET and SSC based on previous studies (cited on p9772, l. 21), including our own previous experience in Nepal (Muller 2013).

7. We will add the range of catchment size in the revised text (p9774, l.15). Interestingly, we found no significant relation between catchment size and prediction performance.

8. We agree with the reviewer that the method used to aggregate rainfall does not properly account for spatial correlation. Indeed, the challenges involved in scaling spatially heterogenous and correlated rainfall are a main weakness of the process-based approach. While spatial correlation would be better accounted for by aggregating rainfall time series *before* computing their statistics, using aggregating rainfall statistics instead allows using gauges

non-overlapping rainfall observation periods, which is important in the context of Nepal. Unfortunately, the low density of rain gauges within the considered basins does not allow us to formally account for spatial correlation when aggregating frequencies. However, in a previous study (Muller 2013) we observed large spatial correlation ranges on rainfall occurrence in Nepal (125km during the monsoon). Under these conditions the selected method stands out as the most parsimonious approach to utilize multiple, yet sparse, rainfall observations. We will clarify these points by modifying p9776 l.3-13 in the reviewed manuscript.

9. Thank you – eq 3 p9778 will be corrected in the revised manuscript.

10. To clarify the procedure we will replace l.24-3 p 9780-9781 as follows:

"The statistical model uses a linear regression over a cross section sample of observations to predict mean flow based on mean rainfall. The regression may fail to capture a variety of unobserved characteristics affecting both rainfall and streamflow (e.g., local topographic features), and hence may not capture the causal relation between the two variables. The extent of this bias cannot be quantified a priori, so we considered the two extreme cases of zero and infinite bias."

11. We believe that changing rainfall statistics while keeping recession params constant will not create a paradox because the latter are exogenous parameters that describe watershed response independent from climate. A key assumption of the considered process-based model is that recession constants are exogenous and unaffected by climate – this assumption may be debatable, but that goes beyond the scope of this study. Under these conditions, while it is true that changing the rainfall statistics will affect the initial conditions (Qo) of the dry season, it will *also* affect all flows (Q(t)) during the dry season. In fact, it is precisely (and only) by affecting Qo that wet season rainfall characteristics affect the dry season recession. It follows that fitting Eq. 1 using streamflow values (both Qo *and* Q(t)) generated in current and future rainfall conditions will provide identical estimates of a and b.

12. We agree with the reviewer that the performance of the process-based method for short observation records is ultimately driven by the method chosen to estimate k. Yet again, the purpose was here to assess the operational performance of the method in the specific context of Nepal, so we lumped errors arising from parameter estimation uncertainties and the modeling errors themselves. We will make this point more explicit in the revised manuscript by modifying Sections 2.1.1 and 4.1.3 as described above.

13, 14 Thanks for the suggestion; it will be incorporated in the revised manuscript.

15. Here we would argue that the statement applies to the process-based model itself (as opposed to the parameter estimation method), because catchment homogeneity and the assumed recession behavior (linear in wet season and non-linear in dry season) are fundamental assumptions of the model.

16. We will change the sentence to:

A similar model was used in ***non-seasonal climates*** *by Botter et al (2013) to relate the resilience* ***of the probability density function of streamflow*** *to observable catchment characteristics.*

17. We will change p9789 l.25-28 as follows:

We use the relations derived ***in the*** *stochastic dynamic framework (Appendix B) to infer the effect of rainfall and recession characteristics on the resilience of flow regimes.* ***This will allow*** *the reliability of the statistical models to be assessed for predictions under change.*

18. Thanks for the suggestion, this will be incorporated in the revised manuscript.

19. Thank you very much for the insight. We will make that point clear by adding the following sentence on p9789 l15:

We measure streamflow resilience as the change in the flow duration curve (in terms of differences in NSC) resulting from a change in rainfall \footnote{This contrasts with Botter 2013, who considers the effect of rainfall regime changes on the probability distribution function of streamflow. While the general idea is the same, the numerical results can be different.}