Anonymous Referee #2

We appreciate the thoughtful and constructive comments from reviewer #2, and our responses are as follows:

Thank you for the opportunity to review the paper titled "Exploring the physical controls of regional patterns of flow duration curves – Part 4: A synthesis of empirical analysis, process modeling and catchment classification." I found the paper to be well-written, the topic of the manuscript relevant to HESS and, although flow-duration curves (FDCs) have received much attention in the literature, the authors present a new application of flow-durations curves that yield interesting insights into the similarity of catchments. I do have several comments about the technical aspects of the paper that could require substantial revision before publication of the manuscript in its final form.

1) Please explain the regime curve in more detail in Section 2.2. Not all readers are familiar with this terminology and it may be difficult to fully understand the results and conclusions without a clear understanding of how the regime curve was determined.

While we do describe briefly in the text what a regime curve (RC) is, we can also add a brief description of how an RC is calculated to Sect. 2.1. There, we briefly describe the FDC, as well as what information is taken from the other papers, and here may be a good place to insert a few lines describing the RC. Section 2.2 describes only those methods unique to the fourth paper in the series. The relevant revised passage from Sect. 2.1 is reproduced below, and will appear in the final paper:

"…From the modeling study (Ye et al., 2012), the dominant model process "class" (i.e., that combination of processes which were found to be necessary for good model prediction for a given catchment) for each of the 197 catchments was extracted. From the catchment classification study (Coopersmith et al., 2012), the class associated with each of these catchments was added to this new database. Lastly, daily regime curves used in the previous two studies will also be used. These are calculated using the entire period of record by finding the average flow for each Julian day of the year, then smoothing with a 30-day circular moving average (see Coopersmith et al, 2012 for further details)."

2) In classifying streamgauges based on climate and catchment process, how do the authors rectify the findings of Wang and Hejazi (2011) that show many of the MOPEX locations are impacted by alteration in the catchment?

The reviewer brings up a valid point that understanding the possible human impacts to these catchments is important when analyzing a synthesis of results obtained from various analyses of these catchments. We do note in this paper that the gamma parameters, the model process classes, and even the catchment classes all highlighted the main agricultural regions of the Midwest as behaving differently, and we have attributed this in part to human impacts. However, we must also note that this paper does not classify streamguages, but rather synthesizes the results from the companion papers in this series. Thus the classes and processes we discuss were *developed in Coopersmith et al (in review) and Ye et al (in review), respectively.*

The algorithm detailed in Coopersmith et al classifies catchments based on four indices that can be determined from the daily regime curve (RC), which has been further smoothed with a 30-day circular moving average. If certain catchments are altered, it will either show in terms of their behavior (alter one of the four indices) or it will not. If the human impacts do not alter one of the four indices, then the algorithm will sort the catchments as if there were no impacts. If one of the four indices is altered by these impacts, however, similarly altered catchments will still be clustered together. Thus for the catchment classification, much of the agricultural Midwest was separated into about 4 unique classes; however, the majority of the NE US, where low climate seasonality dominated and human impacts may not have significantly altered one of the four indices used, fell into only one class.

The modeling study of Ye et al study also focused on the average seasonality (RC) instead of the time series, and thus the human impact signature has been reduced though still remains. In those catchments with a relatively strong climate signal, such as in the Northeast and the Pacific Northwest, the signals from the climate (snow, seasonality of precipitation) and the presence of forest vegetation can override those from human impacts. For those catchments with significant human impacts, such as those in the Midwest, they are singled out both in the classification system and in the modeling work. They belong to individual classes in the former and cannot be modeled accurately by the simple model processes chosen in the latter.

In this paper's conclusions (Sect. 5) we do mention that more detailed modeling studies may be able to fully separate the climate from the catchment processes. We can further add that this would also be better suited to detecting human impacts on these catchments as well. In addition, we can add a citation to Wang and Hejazi (2011) to Sect. 4.1 where the regional patterns found in the first three papers are briefly mentioned and then used to interpret and discuss the findings of this paper. Awareness of the human alterations to the MOPEX catchments will strengthen the analysis of the findings, and we thank the reviewer for bringing this to our attention.

3) The paper fits a mixed gamma distribution to flow-duration curves; however, no probability plots or goodness of fit metrics demonstrating the appropriateness of this distribution is presented in the manuscript. At a minimum, the method of parameter estimation should be included in Section 2.2. What are the bounds of the mixed gamma? Are they such that the lower bound cannot generate streamflows below zero? I can understand that the mixed gamma distribution provides a reasonable fit to the FDC for much of the curve; however, I wonder how closely the three parameters are able to capture the tail behavior, where catchment response may differ more across the study region.

As previously noted, this paper synthesizes results obtained from the work done in the companion papers. Thus this paper does not fit any distributions, but rather uses the fitted parameters obtained in Cheng et al (in review). The revised version of that paper will have clear descriptions of fitting methodology, distribution bounds, metrics, and plots of goodness-of-fit, and we will amend this manuscript to point the reader to the Cheng et al paper for the specific details of the curve fitting. The synthesis paper is focused on what can be learned from the results of the previous papers and assumes that those results have withstood a rigorous peer *review.*

To answer the reviewer's questions regarding the bounds of the distribution, we have provided the following details in this response:
(1) The mixed gamma distribution cannot generate negative values.
(2) The FDC is the inverse complementary cumulative distribution function (CCDF) of the daily streamflow. In Cheng et al, the inverse form of the CCDF of the mixed gamma distribution is used to fit the shape of different duration curves. The non-zero segment of the fitted duration curves was bounded within the exceedance probability from 0 to 1 minus alpha, where alpha is the probability of a zero rainfall or streamflow record in the observed daily data.
(3) We were not interested in determining whether the FDC is drawn from the mixed gamma probability distribution; rather we simply used the mixed gamma distribution to fit the FDC. The differences between observed duration curves and fitted duration curves were quantitatively measured by goodness of fit (R^2) and the Nash-Sutcliffe coefficients (Ens). Based on R^2 and Ens, most of the duration curves are estimated well.

The reviewer does bring up an excellent point in that the tails of the FDC, which in this paper is where we found the most difference between catchments, may have also been where the fitting was poorest. Cheng et al (in review) noted that the visual fit of the lower tail appeared worse because the logarithmic scale of the y-axis exaggerated the differences between the fitted and observed FDCs, while the actual differences were very small. They also note that the mixed gamma distribution slightly underestimated the highest flows in the upper tail. However, by their metrics of goodness of fit, the mixed gamma distribution provided a satisfactory estimate of the FDC. In the revised version of this paper, we can briefly discuss in Sect. 4.1 how this may affect the groupings we have done of the gamma parameters.

4) Following on comment 3, I wonder if the FDC slope is providing different information than the parameters of the mixed gamma? The conclusions appear to be similar for both analyses. I appreciate the authors' synthesis approach to this problem; however I wonder – given the length of the manuscript – whether both analyses are needed.

The κ parameter of the mixed gamma distribution addresses the "shape" of the distribution, as does the slope of the middle of the FDC, albeit in a more limited fashion. The alpha parameter also contains the zero-flow information, which in arid catchments can comprise a significant portion of the lower tail, which in turn affects the slope of the FDC. Thus it is not surprising that similar conclusions arise from analysis of each. Theoretically, however, the slope of the middle of the FDC would also provide different information than the gamma parameters, because the gamma parameters are obtained from the entire FDC, including both tails. In particular, the slope of the FDC, as calculated in this paper, considered only the middle third of the FDC. From the work of Yokoo and Sivapalan, the middle third of the FDC is thought to represent the regime curve (RC). Thus we use the slope of the FDC as a surrogate for the RC, which the middle two papers address specifically, and for the synthesis, this provides a further link between the FDC and the results of these two papers.

Given the length of the paper, as noted by this reviewer, and suggestions from another reviewer of the manuscript, we will be condensing the initial introduction of the previous three papers as *well as eliminating the long recap of the same that is found in Sect 4.1, and this will shorten the paper a great deal. We feel that the inclusion of both analyses in this paper lays valuable groundwork for future research. If a better fitting of another (or the same) distribution is found, the differences in the information given by the fitted parameters and the FDC slope may become more clear, and this could lead to better insights into the processes controlling the FDC.*

References: Wang, D. and M. Hejazi (2011), Quantifying the relative contribution of the climate and direct human impacts on mean annual streamflow in the contiguous United States, Water Resour. Res., 47, W00J12, doi:10.1029/2010WR010283.