Anonymous Referee #3
Received and published: 31 July 2012

The paper presents an extensive analysis of regime curve behavior in the US, performed by means of a downward approach based on the application of models with increasing complexity. The topic is of great interest and also the approach adopted is interesting. Obviously, due to the strong heterogeneity observed in regime curves in all the US, I believe that the authors encountered more than a few problems in trying to reduce such complexity and provide a clear overall picture from results. I also think that some more effort could be devoted in improving the readiness of the paper by clarifying the procedure adopted and the presentation of results.

We appreciate the reviewer's comments and interests. We have made our efforts to address these comments, clarifying the data sources, model development and the presentation of results. We hope this will be adequate.

Comments
1. Page 7039, lines 23-25; the authors should mention what kind of LAI maps they used (daily, monthly, algorithm used, etc)

The LAI used in the model is calculated as the average of the monthly LAI value for each (University of Maryland (UMD) vegetation classification) vegetation type from Mosaic vegetation dataset, weighted by the area fraction of each vegetation type in the catchments. We have included the brief description of how we obtained the LAI in the revised manuscript in the hope that the procedure is now explained more clearly

2. Page 7040, line 5; the authors should introduce here (or before) the definition of the "regime curve" mentioning which kind of year they used (civil, solar, hydrological…?)

The regime curves were presented in civil year (Jan 1st to Dec 31st). We thank the reviewer for helping clarify this.

3. Page 7040, line 10; the algorithm used for separation of base flow should be described since such a separation is important for the use that they do of the different components of flow. In particular they should specify whether it is a physically based filtering or an empirical/numerical procedure.

We agree with the reviewer that we should describe the algorithm more clearly; therefore we have included it as follows, and we hope the reviewer finds it adequate.

"The fast flow and slow flow were separated by a numerical baseflow separation algorithm of Lyne and Hollick (1979),

$$Q_u(i) = aQ_u(i-1) + \frac{1-a}{2}(Q(i) + Q(i-1))$$

(1)

where a is the filter parameter and was set to 0.925 (Brooks et al., 2011). Since the hydrologic partitioning is not strongly sensitive to baseflow separation methods (Troch et al., 2009) we will use this easily-implemented algorithm for baseflow separation.

4. In the caption of figure 1 the AI should be mentioned as part of the figure.

We appreciate the reviewer's suggestion and have added AI in the caption.

5. While it is clear that the curves shown in Figure 1 are chosen as representative of different areas (states) of the US, the authors should mention which are the 9 basins that were chosen and comment how far they can be effectively considered representative of the area with respect to heterogeneities eventually observed within it.

These catchments were selected based on their longitude and latitude to ensure the coverage of the whole continent. We also try to cover as many classes in Koppen climate classification map as possible in order to take into account the most climate characteristics. Therefore,these catchments were considered representative of the climate conditions under the regional similarity assumption (Merz and Blöschl, 2004; Patil and Stieglitz, 2011); we also expect that they can represent the regional dominant processes, especially those related to or evolved with the climate. The results presented in Fig. 12 and 13 also confirmed that catchments that are geographically close share dominant processes. This similarity is constrained by the temperature gradient (i.e. the snowmelt is not important from the south of North Carolina), topography (i.e. the Appalachian mountainous catchments are different from catchments at the bottom of the mountain), and so on.

However, as the reviewer pointed out, given the heterogeneities within the USA, we wouldn't expect these nine catchments to be representative of the whole country. For example, the results also indicated that these nine catchments didn't cover the strong anthropogenic influences in Midwestern catchments (Iowa, Illinois, etc.). As our goal is to explore the dominant processes of all the 197 MOPEX catchments across the states, the nine catchments were selected for initial model development. Thus the regional similarity of these nine selected catchments were commented on in the discussion about the process mapping . We have also included a map of the nine catchments as follows. We hope the reviewer finds it sufficient.
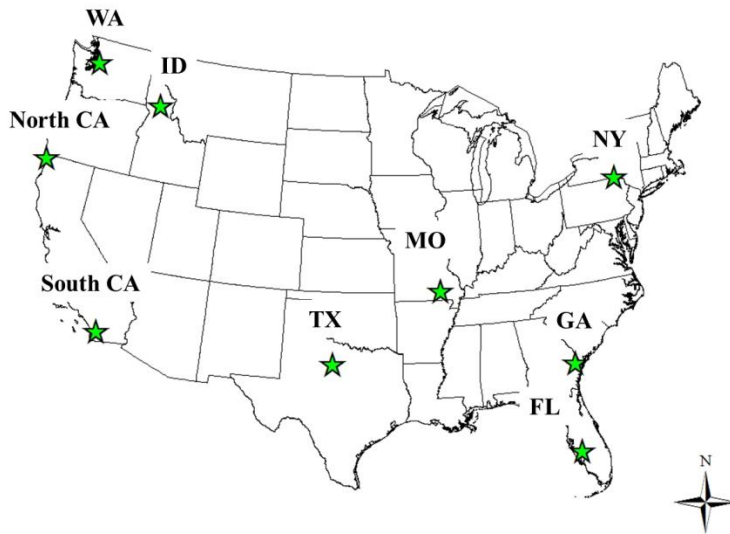
6. Page 7043, line 13; "topographically driven subsurface drainage", the authors with this definition provide an important physical characterization of the subsurface flow process. While in principle I would agree on such a definition, I would also like to know how far the authors detect the influence of topography in their results, considering that topography, as far as I understood, is not explicitly parameterized in their model.

We apologize for the confusion. It is true that as a simple bucket model, the topography (or, slope) was not explicitly parameterized. But it does influence the simulation through the value of t_u, the characteristic response time of the slow flow: the smaller the gradient, the longer it takes to drain the slow flow. As Table 1 indicated, the central catchments, which are located in the flat Plains region, have a much larger t_u than the eastern and western catchments which tend to be more mountainous.

7. Page 7044, bottom line, "e.g. catchments in California", I think that here the authors should specify "North CA".

Thank you for the clarification; we have specified it in the revised manuscript.

8. Page 7047, section 2.2.5, which algorithm or method (data) was used to evaluate PET?

The PET is given from the MOPEX website; it was calculated based on NOAA Pan Evaporation Atlas (NOAA, 1982) using the Penman (1948) method. The solar radiation required in the calculation was estimated from percent sunshine (Hamon et al, 1954).

9. Page 7051 Eq. 16; I do not understand this equation, the authors should please better specify how they obtained their standardized flow values.

We are sorry for the confusion caused during the print; the correct equation is:

$$SQ = \frac{Q - mean(Q_{obs})}{std(Q_{obs})}$$

where Q represents the time series of flows (observed for SQ_{obs} or model-predicted for SQ_{sim}), Q_{obs} is the time series of observed flow, SQ_{obs} is the standardized observed flow, and SQ_{sim} is the standardized simulated flow, both of them are represented by SQ in the equation as they are calculated in the same way..

10. Page 7056, line 11; The authors choose 0.53 as the threshold for detecting a "satisfactory" behavior. This choice is not really motivated, what would happen if they assumed 0.50 as a threshold?

We agree with the reviewer that it is sudden to give 0.53 without an explanation; we have now described how we arrived at it as follows. We hope the reviewer finds it satisfactory.

The initial screening of the model's simulations suggested that even the complete model was insufficient in certain catchments, such as those in the Midwest, where human impacts cannot be ignored. In some catchments, the flow regime curves were bimodal while the model was only able to capture one of the flow peaks. Because it is a simple model, we would not expect it to accommodate anthropogenic activities; therefore, we need to eliminate these catchments where the model performs poorly. To ensure that the model captures the dynamics as well as the volume of the flow, we use MSE as our criterion. The decomposition of the MSE (or Nash-Sutcliffe efficiency) shows that the MSE consists of three components: mean, variance and correlation coefficient (Gupta et al., 2009). However, as the error is scaled by the standard deviation, it could be problematic for comparisons among catchments. To avoid this, we standardized the flow before the MSE calculation. We selected the 90% of the catchments with lowest MSE in fast flow, slow flow and total flow separately and then obtained the intersection of these three sets to determine those catchments that had the lowest MSE in fast flow, slow flow and total flow simulation. These catchments were then considered as satisfactory catchments. The 0.53 value was reported because all these catchments had MSE less than 0.53.

11. Page 7062, line 3; I believe that "self-similarity" is not properly mentioned here. I would rather say "hydrological similarity".

Thank you for the correction; we have made the suggestion change.

12. Page 7064, section 4.6; since the analysis on flow duration curves is reduced to comments in this short section, and also, results are not quite satisfactory, I definitely suggest to move the emphasis of the paper (in the title and in the abstract) from flow duration curves to the regime curves.

We agree with the reviewer that this paper is focused on the regime curves rather than the flow duration curves. We have now modified the title and revised the abstract as well as the beginning of the introduction. Since this manuscript is the second one of the four-part work, also the motivation of this paper was from the study in the flow duration curve (Yokoo and Sivapalan, 2011). We think it would be better to keep the shared part of this four-part work in the title

(Exploring the physical controls of regional patterns of Flow Duration Curves) and change the title of this manuscript. We hope the reviewer find these adequate.

***Title**: Exploring the physical controls of regional patterns of Flow Duration Curves - Part 2: Role of seasonality and associated process controls, from the perspective of Regime Curves*

***Abstract**: The goal of this paper is to explore the process controls underpinning regional patterns of variations of runoff regime behavior, i.e., the mean seasonal variation of runoff within the year, across the continental United States. The ultimate motivation is to use the resulting process understanding to generate insights into the physical controls of Flow Duration Curves (FDC). However, construction of the FDC removes the time dependence of flows. Thus to better understand the physical controls, which may have seasonal dependence, the regime curve (RC), which is closely connected to the FDC, is studied in this paper and later linked back to the FDC. To achieve these aims a top-down modeling approach is adopted; we start with a simple two-stage bucket model which is systematically enhanced through addition of new processes on the basis of model performance assessment in relation to observations, using rainfall-runoff data from 197 United States catchments belonging to the MOPEX dataset. Exploration of dominant processes and the determination of required model complexity are carried out through model-based sensitivity analyses, guided by a performance metric. Results indicated systematic regional trends in dominant processes: snowmelt was a key process control in cold mountainous catchments in the north and north-west, whereas snowmelt and vegetation cover dynamics were key controls in the north-east; seasonal vegetation cover dynamics (phenology and interception) were important along the Appalachian mountain range in the east. A simple two-bucket model (with no other additions) was found to be adequate in warm humid catchments along the west coast and in the south-east, with both regions exhibiting strong seasonality, whereas much more complex models are needed in the dry south and south-west. Agricultural catchments in the mid-west were found to be difficult to predict with the use of simple lumped models, due to the strong influence of human activities. Overall, these process controls arose from general east-west (seasonality) and north-south (aridity, temperature) trends in climate (with some exceptions), compounded by complex dynamics of vegetation cover and to a lesser extent by landscape factors (soils, geology, and topography).*

***Introduction**: This paper is the second paper of a 4-part series (the others being Cheng et al., 2012; Coopersmith et al., 2012; and Yaeger et al., 2012) that attempt to understand the physical controls on regional patterns of variations within hydrological signatures of runoff variability. Instead of exploring the Flow Duration Curve (FDC, a key frequency-based signature of daily runoff variability) like the first paper, we will approach the issue from a different perspective, focusing on another compact signature of runoff variability, namely, the regime curve, which denotes the mean seasonal variation of within-year runoff variability.*

13. Last but not least, as an overall final comment, I believe that the authors should put more emphasis on the procedure that they use for validating their models which is suitable, in my opinion, for the purposes of the paper but need to be better addressed. The purpose of the paper is speculative and not aimed at proposing predictive models. In this perspective I believe that their model validation is a particular case of the "scientific validation" introduced by Biondi et al. (Physics and Chemistry of the Earth, doi:10.1016/j.pce.2011.07.037, 2012) as opposed to

"performance validation" of predictive models. If the paper purpose lies in this second case (performance validation) I would ask for a more rigorous process of calibration\ validation (first: split sample; second: recalibration of parameters when moving towards more (or less) complexity and so on). On the other hand, according to the principles of scientific validation, a good point of the paper, which is not commented at all by the authors, is the multi-objective validation that they perform by separately comparing the fast and slow flows. This procedure actually allows to detect improvements of the model performance that poorly affect the global discharge (see for example Figure 7 and figure 8) but are significant for characterizing the fast flow component and detecting the main control processes.

We thank the reviewer for bringing to our attention the paper on scientific validation and the distinction between performance and scientific validation. We agree that our goal is not to deliver precise predictions of the streamflow time series, but rather to gain a general understanding of first order impacts of different processes on flow generation mechanisms along the climatic or other gradient. For this reason, the "scientific validation" (Biondi et al, 2012) suits our work better. We have added the following explanation of the validation, including the multi-objective validation, to the revised version of the paper. Again, we appreciate the reviewer's suggestions, and hope the reviewer finds this satisfactory.

"The scientific validation (Biondi et al, 2012) could be used in the assessment of model hypotheses: the identification of integral processes for which the model should account. This was shown in the model development section where we initially applied the base model to the nine selected catchments, assessed the model performance, and then added four processes one by one based on catchment characteristics in order to improve the model's predictions. This systematic model development procedure itself helps to validate the importance of each remaining process.

The other goal of scientific validation is to "provide the proof of model adequacy to the representation of the real world" (Biondi et al, 2012). Since a model could produce good results with a wide range of specific parameter values, it is important to consider the parameter set as a combined set (Freer et al., 1996). The Bayesian framework we used is able to find optimum parameter sets by giving greater weight to the better simulations. These parameter sets and predictions can then be chosen as more likely than others.

In addition to the assessment of model hypotheses and parameters, a multi-criteria approach can also be used to verify model performance. In this work, we calibrate the parameters to optimize both the fast flow and the slow flow simultaneously. This multi-objective check helps provide information regarding whether individual subsystems or processes are performed in the catchments. For example, some processes may not affect the total discharge, but could influence the quantities of observed fast flow (Fig. 7 and 8). This multi-objective calibration enables us to detect those improvements in model performance that negatively affect the global discharge but are beneficial for characterizing the fast flow component and detecting the main control processes."

References

Brooks, P. D., Troch, P. A., Durcik, M., Gallo, E., and Schlegel, M.: Quantifying regional scale ecosystem response to changes in precipitation: Not all rain is created equal, Water Resour. Res., 47, W00J08, doi:10.1029/2010WR009762, 2011.

Hamon, R. W., Weiss, L. L., and Wilson, W. T.: Insolation as an Empirical Function of Daily Sunshine Duration, Monthly Weather Rev., 82 (6), 141-146, 1954.

Merz, R. and Blöschl, G.: Regionalisation of catchment model parameters. J. Hydrol., 287(1–4), 95–123, 2004.

NOAA Technical report NWS 33: Evaporation atlas for the contiguous 48 United States, Washington, D.C, 1982.

Patil, S. and Stieglitz, M.: Hydrologic similarity among catchments under variable flow conditions. Hydrol. Earth Syst. Sci., 15, 989–997, 2011.

Penman, H. L.: Natural Evaporation from Open Water, Bare Soil and Grass, Proceedings of the Royal Society of London, Ser. A, Vol. 193, No. 1032, PP120-145, April 1948.

Troch, P. A., Martinez, G. F., Pauwels, V. R. N., Durcik, M., Sivapalan, M., Harman, C., Brooks, P. D., Gupta, H., and Huxman, T.: Climate and vegetation water use efficiency at catchment scales, Hydrol. Processes, 23(16), 2409–2414, doi:10.1002/hyp.7358, 2009.

Yokoo, Y. and Sivapalan, M.: Toward construction of flow duration curves in ungauged basins: Flow duration curve separation and sensitivity to watershed characteristics, Hydrol. Earth Syst. Sci., 15, 2805–2819, 2011.