Review of the manuscript 'A Global Approach to Defining Flood Seasons' (hess-2015-140)
by D. Lee, P. Ward, and P. Block

General Comments:

The work presented in the manuscript with the title 'A Global Approach to Defining Flood Seasons', aims to develop a methodology that allows defining spatial and temporal characteristics of major flood seasons globally with the help of daily stream flow simulations.

As the work reuses of already existing and published modelled global streamflow data, the central scientific contribution of this manuscript is the development of an approach that allows defining flood seasons globally. With this in mind, it would be valuable for the authors of the manuscript to focus more on how they define floods seasons and to compare their results with other already existing flood season indicators.

Besides focusing on the major flood season globally, the study does also briefly consider minor floods seasons. As the authors point out, the study of minor flood seasons has not obtained much attention at a global scale and therefore merits further investigation.
Therefore, the manuscript would benefit from extending the scope to minor flood seasons not only at the local scale (i.e. with an example from East Africa as presented in the manuscript) but also to the global scale, which would also better match the overall global scope of the manuscript.

In general, the paper need to be clearer about the different meaning and usage of the terms 'peak month (PM)', 'flood season (FS)' (is it always the 3 months (i.e. PM +-1 month)?), and 'P_{AMF}' (Percentage of Annual Maximum Flow (AMF)). Sometimes these variables are used almost interchangeable. On this matter, also see the specific the comments below.

With the general comments above and the specific comments mentioned below, I recommend thoroughly revising the manuscript, as there are several instances that require further clarification, discussion, corrections, and amendments from the authors.

Overall, the paper is well written and has the potential to be of interest to the readership of HESS. Therefore, I suggest resubmitting the manuscript after a major revision.

Specific Comments:

Section Abstract:

P 4596 L4-6: The authors argue in their abstract that 'forecasting systems in the order of months to seasons are a rarity' and that 'dominant flood seasons must be adequately defined' for prediction and disaster preparedness.
I agree that there is a shortage of long-term forecasting systems; however, I would say that in general the flood regime and therefore the flood prone seasons of rivers are locally (the scale at which preparation for disasters take place) well know.

In addition, I presume that the hydrological model, from which the discharge data has been obtained, performs very different at different scale. This is of particular importance, as ungauged basins, for which this type of information would be useful, are often smaller than the grid scale of the model.
Therefore, it is suggested changing the reasoning/focus of the abstract (and also the introduction and conclusions), as the approach to define flood seasons (of 3 months in length) that has been developed here has only a marginal connection with disaster preparedness and flood forecasting.

Section 1 Introduction:

The introduction focuses on long-range seasonal forecasts for guiding decision-making, seasonal predictability of streamflow impacts and the need for linking atmospheric indices with streamflow predictions at global scale. From this introduction, I would expect a paper that aims to PREDICT streamflow patterns, which is very different to the actual scope of the paper.
For that reason, I think the introduction should focus more on the actual topic (i.e. a data based approach concerned with the identification of flood seasons and the second objective of extending the approach to already existing globally modelled streamflow).
To put the work presented in the manuscript into the context of already existing studies of flood seasonality and global streamflow characteristics, the authors may find the following articles useful: For previous work on different method of identifying/classifying flood seasonality see for example Ouarda et.al. 2006, Liu eta 2010 or Chen et.al. 2013. For more information on how the manuscript fits in the context of or differs to other global studies of streamflow characteristics see for example Dettinger and Diaz 2000 or Beck et.al. 2015.

Section 2 Data description:

2.1 Streamflow stations

P 4599 L 5-7: The current selection of the dataset cannot really be considered 'global' and has a particular bias towards to certain regions (particularly northern hemisphere).
Please provide further explanation on the how the stations were selected (see also comments below).

Does 'having at least 20 years of continuous daily streamflow data' mean that all stations that had one measurement missing were excluded, or was a threshold on missing data applied?

Please further explain how the selection criterion 'continuous daily streamflow data' influenced the spatial coverage of the data.
To what period does the 'at least 20 years' refer to? Same the stream flow simulations (1958-2000)? Please specify.
With a less stringent selection criterion ('having at least 20 years of continuous daily streamflow data'), could one have obtained a better compromise of spatial coverage and data quality?
Please add a paragraph further elaborates on these choices, as this step is crucial in determining the amount of data and spatial coverage available for method validation.

2.2. PCR-GLOBWB

P 4599 L20: The authors mention that the model was forced with input data from ERA-40, which 'were subjected to a number of corrections'.
Please specify how these corrections might or might not influence the model output.
Please also discuss/analyse the influence of the hydrological model and the grid cell size on the ability of the model to generate the magnitude of hydrological extremes, which will be used as a key variable for the definition of the flood season using the volumetric-based threshold approach.

3. Defining flood seasons

3.1 Methodology for defining grid-cell scale flood seasons

P 4061 L1-7: It is pointed out that it is important to consider not just the magnitude but also volume to define a flood season and that the authors therefore adopt a volume-based threshold technique. However, the authors then select the 'streamflow exceeding the top 5 % of the FDC', which is related to magnitude. If this is the case, it is not clear why the need for/use of a volume-based measurements is highlighted here and several times throughout the document.

Additionally, please elaborate on the decision process of selecting the '5% threshold', as on the previous page the importance on selecting the 'proper threshold for POT' highlighted. Have other thresholds been tested and what was the outcome?

P 4061 L10-12: From the description, it appears that after identifying the peak month, the flood season is defined as the month before and after the peak month. Is this the case or is the flood season related to the three month with the highest number of days above the 5% streamflow threshold?
I could imagine a situation similar to the synthetic streamflow data used in Figure 2, with August (105 days) being the peak months but June with 60 days (instead of the 25 days used in the example)) and July with 75 days. Resulting in the peak month being off centre.
It needs to be clarified, if such a situation had been considered and if not how that will influence the results (including the calculation of the index 'Percentage of Annual Maximum Flow (P_{AMF})').

P 4601 L 20-24: The index P_{AMF} has been created to 'evaluate' the identified flood seasons. Therefore, it is suggested interpreting the 'high' or 'low' values of P_{AMF} in that regard, (e.g. a high P_{AMF} indicates a well represented Flood Season (FS), a low P_{AMF} indicates poorly identified FS?). Additionally, as the index has been created to evaluate the defined FS, please also give an indication of what is considered by the authors of being a good or less acceptable value (i.e. what percentages are considered good ? And for the discussion of the results for what regions is the approach used to define flood seasons not working).

P 4601: Generally, after highlighting the advantages of the POT approach and the disadvantages of the annual maximum flows (AMF) (such as) it is not clear to me why the P_{AMF} method uses AMF to evaluate the defined flood seasons. Has other values instead of the AMF been considered, and if so why has the AMF been chosen?

3.2 Classification techniques

P 4602 L 12-14: Please further elaborate why '1-7 days favour identifying flood magnitude, while 15-30- days favour identifying flood volume'.

P 4602 L 16-21: I am having difficulties in understanding what this section means. Can you please rephrase and explain for what reason 'they may be considered slightly superior'.

How can one calculate the P_{AMF} for the other classification techniques?
Can equation (1) be reformulated to be more generally applicable (see also comment on monthly P_{AMF} below).

General Comment on Section 3.1 and 3.2:

It is not clear to me, why these two sections are separate. I would expect to evaluate the PM that has been identified with the 5% threshold approach together with the other classification techniques and then pick the best indicator (i.e. here apparently the PM) for further analysis. In addition, if applicable compare the performance with other seasonal indicators that have been publish in other studies before and explain why the approach here is superior to the other methods (otherwise the new approach would not be needed).

Having first the 'Methodology for defining grid-cell scale flood seasons' and then having a separate section on '3.2 Classification techniques' is confusing.
Therefore, I suggest combining these two section together with an in depth analyses on the flood season classification approaches (e.g. how do global maps of differ?).

3.3 Methodology for defining sub-basin scale flood seasons

P 4603 L 21-22: I understand that under certain circumstances, the P_{AMF} can be useful to indentify managed dams or reservoirs, but if the dams are managed in accordance with what is considered the 'natural flow regime', this will not help. Additionally, why not use the dataset mentioned a few lines above to find the location of the dams? Without the factual knowledge of the presence/absence of a dam, one will have difficulties in determining if the low values of the P_{AMF} indicator obtained from the modelled data are due to management or due to difficulties of the model to represent the hydrological characteristics of that region.

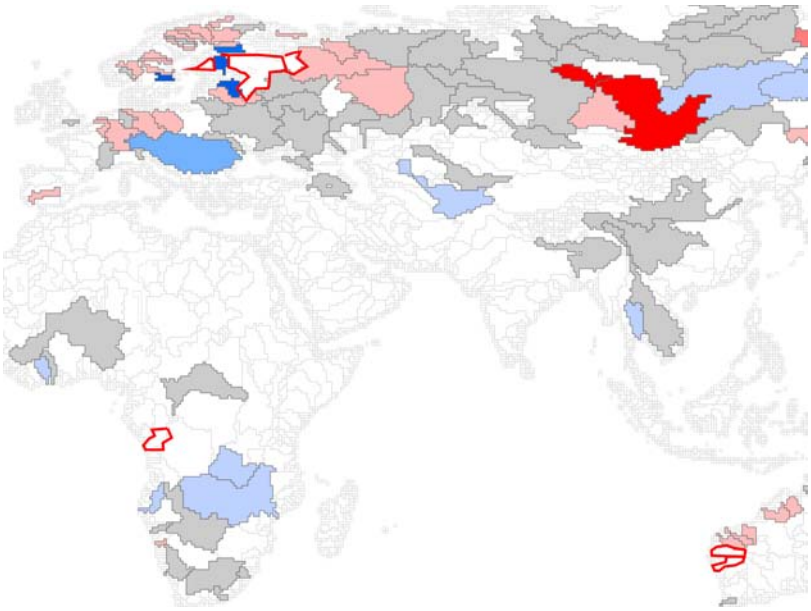
4. Verification of selected flood seasons

P 4604: I suggest adding the characteristics of the data obtained from the DFO (such as available period...) here, instead of having it in section 4.2 (P 4606), where I would focus on discussing the results.

4.1 Observed vs. Modelled flood seasons

P 4605 L 6: How are the temporal differences calculate? Is it based on the (central?) peak months or on the entire 3 month long flood season? I.e. if I have an observed Flood season June to August, and a modelled season September to November, is the difference three or just one months?
I would suggest to calculate the differences not for the FS but for the PM (if the PM is centred in 3 month flood season (see also discussion on the definition of FS above)) and the P_{AMF} respectively and then add a panel showing the differences in Figure 4-6 respectively directly, allowing a direct comparison (instead of having them separately in Figure 7).

P 4605 L 6: In Figure 7 (P 4624), the temporal differences are shown.
However, the colour scale of the Figure seems to omit basins with differences larger than +- 4 months (see catchments highlighted with red boarders in the Figure excerpt below)!
Please check again, why these catchments are not shown. If these catchments actually have such extreme differences in the FS, please do not omit them from the discussion in section 4.1.
This is an important part of the analysis, which is currently not apparent to the reader and should be highlighted and discussed!



P 4605 L 7: Is there is a mix-up with the % of stations used in the text to describe the bar plot. 62% and 44% seem to refer to the % of stations of the entire dataset but I rather think that the percent should read according to the height of the bars ~ 35% and ~50%. Please check.

P 4606 L 6-7: I'm not sure how the authors come to the conclusion of 'Europe exemplifying a constant-flow region'. From my knowledge of the flood hydrology in Europe, I would say that most of the regions in Europe have a well-defined seasonal flood regime. Could the authors please better explain how this had been concluded.

General Comment on Section 4.1:

As the main aim of the paper is to define flood seasons globally, I recommend a more in depth analysis of the obtained differences in the PM or the FS.
For example, it would be valuable to analyse if the differences between the observed and modelled PM and FS are systematically linked to station/sub-basin characteristics such as catchment size, latitude/ longitude or altitude. The results will then give a better fell on the reliability of the modelled PM and FS not only in light of possible human influences (e.g. dams or reservoirs) as discussed in the manuscript

4.2. Modelled flood seasons vs. actual flood records

As mentioned before, I would move the characterisation of the DFO data into section 4 and focus here on a more quantitative assessment of the differences.

P 4606 L 23: To me the there is no 'striking similarity' between the DFO and the modelled data. Maybe if the authors summarise the gridded model data to the same sub-basin scale as the DFO, similarity may becomes more apparent. I therefore suggest to also providing some sort quantification (not only qualitatively discussing the maps), before calling it 'striking'.

5. Defining minor flood seasons

Defining minor flood seasons is a very relevant research topic that has obtained little attention in global studies, as the authors point out (P 4607 L 14-21).
This is where I would see a great contribution of this manuscript in advancing the scientific understanding of flood seasons.
Unfortunately, this aspect is only covered briefly and appears to be appended to the main analysis, currently with limited added value.

P 4607 L 23: Please explain how monthly P_{AMF} values were calculated. (I suggest using a more general formula for equation 1).

P 4608 L 14-16: after describing in length the methods used to define the minor PM, the authors only show an example of East Africa. Here, I would have expected a global map showing regions where such minor flood seasons are existing and if possible indicating the PM as well on global maps.

6. Conclusions and discussions

P 4608 L 20-23: The authors highlight that the streamflow model was evaluated 'to define dominant and minor flood seasons globally'. This has only been partly archived for the case of the dominate season, not for the minor seasons (see comments for section 5 above).

P 4609: As already mentioned in the comments to the introduction, the conclusion of the manuscript focuses on many other aspects surrounding the topic of prediction and links to global and regional climate links, which has little to do with the main focus of the manuscript in the current form.
Therefore, I would suggest, revising this section.

#####

Comments to Tables and Figures

P 4616: Table 1 Please add the '5% ' to the threshold column heading.

P 4618: Figure 1. When printing the manuscript on my printer (printer-friendly version form the HESSD website), the background polygons are not visible. Please check.

P 4621: Figure 4: When printing the manuscript, the colour code for the points does not allow me to identify the different months properly. For example, I cannot distinguish points indicating April from March or May. Please use a different colour scheme.

P 4624 Figure 7: Adjustment of the plotting procedure is necessary to accommodate basins that have differences larger than 4 months and therefore currently are hidden and not visible at all.

P 4624 and P 4624: Please add to the Figure captions, what the meaning or + and - are. (i.e. Do positive red values mean that the observed PM/FS is x months earlier and negative blue values indicate that the observed PM occur x months later?)

References.

Ouarda, T., Cunderlik, J., St-Hilaire, A., Barbet, M., Bruneau, P. and Bobée, B.: Data-based comparison of seasonality-based regional flood frequency methods, Journal of Hydrology, 330(1), 329–339, 2006.

Liu, P., Guo, S., Xiong, L. and Chen, L.: Flood season segmentation based on the probability change-point analysis technique, Hydrological Sciences Journal, 55(4), 540–554, 2010.

Chen, L., Singh, V. P., Guo, S., Fang, B. and Liu, P.: A new method for identification of flood seasons using directional statistics, Hydrological Sciences Journal, 58(1), 28–40, 2013.

Dettinger, M. D. and Diaz, H. F.: Global characteristics of stream flow seasonality and variability, Journal of Hydrometeorology, 1(4), 289–310, 2000.

Beck, H. E., de Roo, A. and von Dijk, A.: Global maps of streamflow characteristics based on observations from several thousand catchments, Journal of Hydrometeorology, published online, doi: 10.1175/JHM-D-14-0155.1; 2015.