THE UNIVERSITY OF ARIZONA.
DEPARTMENT OF GEOSCIENCES

Geosciences Dept.
University of Arizona
1040 E. Fourth St.
Tucson, AZ 85721
June 14, 2016

Dear Editor:

Thank you for the chance to revise our manuscript and for providing clarifying comments and guidance on how best to revise the manuscript. Below is a point-by-point discussion of the changes we have made to address the AE's comments (and by association those provided by both reviewers). We have also provided a "tracked-changes" draft of the main body of the paper and figures.

<u>AE's Comments</u>

1. Though I understand the concern of reviewer #1 regarding the estimation of recurrence intervals that can be compromised by the regionalization technique used in the study, I tend to agree with the authors in that the use of a regionalization technique (i.e., space for time) that pools data from neighbouring basins can be useful to construct an extended record. This type of approach can potentially be the only way to extend the rainfall or discharge time series for frequency analysis in areas with short precipitation records, and has therefore been used in a number of published studies (in many cases for practical applications, see for example Soong et al., 2004; among other studies). It will be useful to appropriately acknowledge and briefly mention some of the previous research (using space for time approaches) in the paper, establishing similarities and differences to these previous procedures. The limitations of applying this type of technique need to be also appropriately acknowledged (see also Soong et al., 2004; Stedinger et al., 1992). For example, one of the assumptions the index-flood procedure (that uses regional data) is that observations at different sites must be independent (i.e., not subjected to the same rainfall event). However, as discussed in previous work, this assumption is unlikely to be satisfied for frontal rainfall events as adjacent sites can be correlated (Hosking and Wallis, 2005). But even if this is a limitation, Hosking and Wallis (1988) showed regional frequency analysis can (in some cases) be more accurate than at-site analysis even when inter-site (spatial) dependence is present. You can probably add these or alternative references.

As you have stated, we believe the regionalization technique or space-for-time substitution is a viable way to extend the record for frequency analysis. We agree with your comment concerning providing context and background on this method and have added text to the introduction of the manuscript to: (1) provide examples of previous studies using the same technique, and (2) acknowledge and discuss the main limitations of this method. The following text was added to the introduction of the paper:

"To mitigate the uncertainty caused by short and incomplete flood discharge records, this study uses a space-for-time substitution (e.g. regionalization) to lengthen the record for a given contributing area. Previous studies have employed similar methods, including the index-flood procedure first described by Dalyrymple (1960) and expanded upon by many subsequent authors. The index-flood method uses data from multiple sites within a region to construct more accurate flood-quantile estimates than would be possible using a single site (Stedinger et al., 1993 and Hosking and Wallis, 2005). This method can also be used on precipitation data where it is referred to as the station-year method (Buishand, 1991). The index-flood method is based on two major assumptions: (1) that observations from two or more basins are independent; and (2) that observations follow the same distribution (Wallis et al., 2007).

Here we use a regionalization method similar to the index-flood method in order to calculate rainfall intensity values associated with specific recurrence intervals. The assumption of statistical independence of rainfall (and associated flood) observations is one that we assume in this study but understand may not be true for all samples in our natural data set. This assumption is difficult to definitively prove with natural data (Hosking and Wallis, 2005). For example, a large rainfall event may affect two basins in a similar way and therefore create correlated maximum rainfall intensity values. This spatial correlation is difficult to avoid and may cause biased results. However, it has been shown that the index-flood method can be used in the absence of fully statistically independent observations and still give robust results (Hosking and Wallis, 1988; Hosking and Wallis, 2005). The assumption that observations are sampled from the same distribution is also somewhat difficult to prove with natural data, but by knowing the study areas well a researcher can identify regions with similar rainfall and flood mechanisms. Many examples of this type of area analysis can be found in the literature, including Soong et al. (2004) who separated rural streams in Illinois into hydrological regions based on basin morphology and soil characteristics. Soong et al. (2004) used regionalization in their study to increase the amount of flood data available for frequency analysis. Wallis et al. (2007) employed a similar regionalization method to identify hydroloclimatic regions in their study of precipitation frequency in Washington. It should be noted that FEC's in general use this type of regionalization approach to analyze maximum flood data for hydroclimatic regions with similar flood mechanisms. In this study we similarly attempt to analyze regions based on their basic rainfall mechanisms, in this case by separating the Upper and Lower Colorado River basins."

2. According to the rebuttal letter (in the response to reviewer #1), you have now modified (in the revision) the methodology that you use for the frequency analysis to include the "Peak Over Threshold" (POT) analysis. I find this explanation in the rebuttal letter unclear (and even less clear in the revised paper). The POT method is based on utilizing peak events of a given time series exceeding a specified threshold. The selection of the threshold value in the POT method is not straight-forward, and you are not providing enough information of this analysis. You seem to be using a

threshold of zero (?); this means that you are using all the data? By considering several peak events in each year instead of annual maxima, the number of data points for statistical processing increases (which is the whole idea of the POT, Wallis et al., 2007). However, the basic assumption of the POT method is that the threshold is high relative to typical values, and a lowering of the threshold to increase the data may lead to bias (see for example results for varying thresholds in Davison and Smith, 1990). This needs to be clarified; the explanation in the rebuttal letter (and paper) does not seem to agree with the use of POT. If POT is not being used, then please make clear how are selecting values to calculate ARIs for rainfall (the maximum annual?). How is the threshold helping avoid double-counting?

You are correct that we have modified the methodology to include a POT-like approach and we understand that the methodology was not clearly explained in the previous revision. In the discussion paper we used all values in the time series, which had the effect of "double counting" some events (i.e., data series with separate peaks that were caused by one event. We realized that this was incorrect after discussing the reviewer comments. We adopted a POT-like approach where our threshold is zero. This approach makes it so a string of non-zero values in the time series (separated from other strings of non-zero values by zeros on each end) are interpreted as being one event. The maximum/peak value of each event is then taken for our analysis. However, we understand that this does not help exclude values in different, but adjacent basins caused by the same rainfall event (in the case of a frontal-type storm that covers more than one basin). Text to explain the POT method has been added to Section 3.3 of the methods section:

"Rainfall data was sampled temporally by taking the maximum value of each storm event. Storm events were identified as consecutive non-zero rainfall intensity values separated by instances of zero values in time for each temporal scale. This allows for multiple maximum rainfall values in time to be sampled within a year and throughout the entire 9-year study period. This sampling method is similar to that used in the Peak Over Threshold (POT) method typically used on discharge data where a minimum threshold value is set and maximum peaks above the threshold value are recorded as maximum events. Here we set the minimum threshold value to zero and hence the maximum values of all individual storm events are considered in the analysis."

3. In addition, I find that the methods in the paper are still unclear: what is the probability function used in the analysis, are you using the empirical values directly? Please explain and justify if needed.

We are using the rainfall intensity values directly in the frequency analysis. We rank the rainfall observations and use a frequency-rank relationship (Equation 2) to find the specific rainfall value that is associated with a specific frequency or recurrence interval. We chose this method, and not to fit a more complex distribution to the data, in order to use the simplest approach to the calculation of recurrence interval. Text has been added to Section 3.4 of the methods to make this clear:

"Here the recurrence interval is prescribed (10, 50, 100, and 500 yr), then the rank associated with this recurrence interval is computed using the frequency-rank relationship (Equation 2)."

4. Please note that the response to Q10 of reviewer 1 is still unclear. Please clarify the methodology used (and its rationale in terms of exceedance probability calculation of rainfall/discharge) both in the response to this reviewer and in the revised manuscript.

Q10 from reviewer 1 commented, "Additionally, in the case of precipitation, a fixed duration is utilized in extreme value theory to compute the recurrence interval (e.g. the 100 year rainfall intensity for 1-h duration). In this paper, the authors find the maximum intensity recorded for different aggregations times, chosen arbitrarily. This choice has to be supported as well."

Firstly, the time intervals we chose to use in this study were not arbitrary. We chose to use time intervals of powers of 2 in hours (1, 2, 4, 8, 16, 32, and 64 hr) to simplify the approach and, at the same time, use a range of time intervals that were representative of the storm length of the different types of storms in the area. This range was chosen to include short-duration rainfall events such as convective-type and/or monsoon storms (typically high intensity, short duration summer storms in the UCRB and LCRB) and long-duration rainfall events that last on the order of days such as frontal-type storms (typically lower intensity, long duration winter storms in the UCRB and LCRB).

Secondly, the methods used in this paper calculated the different rainfall intensity values for each specified frequency, area, and time interval separately. That is to say that the frequency analysis completed on the rainfall intensities were completed using a fixed time interval and repeated for each time interval separately. This is a necessary step because small basins are sensitive to short-duration events while larger basins are sensitive to longer-duration events. We aimed to capture this relationship between basin area and storm durations in order to properly capture the largest events for a given basin area. Because we are interested in the maximum values for a given basin area (due to the focus on flood envelope curves), we then take the maximum rainfall intensity for each area and frequency, basically taking the largest rainfall intensity of all the different time intervals. These rainfall intensities are then used in the discharge calculations. We understand that the following point was not clearly stated in the methods and the following text has been added to Section 3.4:

"To determine the rainfall-intensity values with a user-specified recurrence interval, maximum rainfall intensities of storm events sampled from the NEXRAD data for each contributing-area and time-interval-of-measurement class were first ranked from highest to lowest (Fig. 2A, Step 2)."

This point is also clearly shown in the pseudocode of Figure 2A within the Step 2: Frequency Analysis section.

5. I wonder if there are any available rainfall records for at least one of the catchments that you can use to compare the results of the ARI (annual recurrence interval) obtained from the radar (even if just for that catchment).

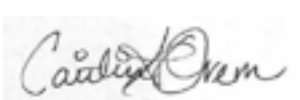
We did not compare rainfall records for specific catchments in this paper because: (1) the rainfall data is typically rain gage data and is therefore different from the radar data; and (2) because we are not focusing on specific catchments in the study. However, there is rainfall recurrence interval data for regions within the UCRB and LCRB that can be compared to the rainfall intensity values of given recurrence intervals and time intervals calculated here. We have added a paragraph to Section 4.2 of the text:

"Maximum precipitation intensities associated with the four defined recurrence intervals are similar to previously published values. In general the values we calculate for the LCRB and the UCRB for the 10-, 50-, and 100-year recurrence intervals are on the order of 10s of mm h^{-1}. This is similar to the spread in values reported on precipitation intensity maps for the same duration and recurrence interval in Hershfield (1961). However, the values reported by Hershfield (1961) are slightly higher (by less than 20 mm h^{-1}) in the LCRB for the three recurrence intervals and in the UCRB for the 10-year recurrence interval than values calculated in this study. The values calculated here are also broadly consistent with presented precipitation frequency estimates for points within the LCRB and UCRB provided by the NOAA Atlas 14 Point Precipitation Frequency Estimates website (http://hdsc.nws.noaa.gov/hdsc/pfds/pfds_map_cont.html). Due to the difference in how precipitation intensities are measured and how the frequencies are calculated the values are expected to be slightly different but within the same order of magnitude."

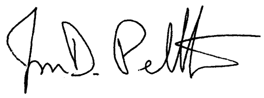
6. As mentioned above, the methods in the original paper submitted for discussion were not well explained. You have partially addressed this issue but there is still a need to further clarify the points stated above. Perhaps a general schematic or a summary of methodological steps could improve the clarity of the methodological approach.

The methods section of the paper has been rewritten as per the comments made above and reorganized to follow the exact calculations made in the codes. The schematic figure added in the last round of revisions has been reworked to follow the methods section more precisely through the use of pseudocode and add further clarification to the text. Please see updated Figures 2A and 2B.

Thank you,



Caitlin A. Orem, oremc@email.arizona.edu



Jon D. Pelletier, Professor, jdpellet@email.arizona.edu