Review of: Constraining frequency-magnitude-area relationships for precipitation and flood discharges using radar-derived precipitation estimates: example applications in the Upper and Lower Colorado River Basins, USA

Authors: C. A. Orem and J. D. Pelletier

General comments:

The manuscript describes and applies a methodology to generate estimates of frequency-magnitude-area (FMA) relationships, using precipitation products from weather radars (the NEXRAD network) in the Upper and Lower Colorado River basins (CRBs). The methodology involves a lumped, event-based rainfall-runoff transformation of constant rainfall intensities derived from the radar in idealized basins. Specifically, if I correctly understood, the methodology of the rainfall-runoff transformation is based on the following steps:

a. The gridded rainfall radar products are aggregated in space based on powers of 2 from 16 to 11664 km^2. This operation creates several squares that are then clipped based on the basin boundaries (as shown in Fig. 2).
b. Each square is an idealized basin where the diagonal is the river network.
c. The mean precipitation in the basin, Q_p, is calculated for different time aggregations.
d. Runoff coefficients variable with the area is used to account for the losses.
e. A triangular normalized area function A(x) is applied to transfer the net precipitation.
f. The diffusion-wave routing scheme is applied to route the water in the idealized channel (the diagonal of the square). In the application of this routing scheme, empirical equations linking slopes and channel width with basin area developed for the Colorado River are used.
g. The method is tested against the FEC curves published for LCRB and U.S.

Even if the paper's motivations are well explained, I have several important concerns on the methodology, both regarding hydrology and frequency estimation, and its validation. As a result, at this stage, my recommendation for this paper is rejection.

Specific comments:

My concerns on the hydrology are described in the following points.

1) The estimation of the losses via a runoff coefficient computed elsewhere is a significant assumption that requires validation in real watersheds of the study area (see also point 3) by comparison with observed discharge. Since the authors have used real precipitation events and not synthetic ones, this could be done. As it is, I have very little confidence in the results of the methodology (even if they may be correct).
2) It is not clear how the runoff coefficient is included in the calculations and the symbols and equations introduced never mentioned it.
3) How are the wet, medium and dry conditions taken into account? This was not explicitly described.
4) The same problem applies for the assumption of a triangular shape of the transfer function: it requires validation.
5) The only validation performed is against the FEC curves published for LCRB and U.S., which are based on observed discharges, after post-processing the results via the frequency analysis. Figure 7 shows significant differences (the axis is logarithmic) between the FEC curves and those generated

via the FMA method, which are based on observed precipitation. The authors have not explained the reasons of these discrepancies and it is hard to have confidence on these results, especially considering the potential use of these curves for flood-related management and design purposes.

6) Why have not the authors considered real basins with real stream networks? The basin shape (that affects the rainfall effectively fallen in the basin) and the stream network organization are known to have critical importance on the flood timing and magnitude.
7) Given the simplified nature of the method, no contribution of snow and snowmelt was considered. This has to be stated. Regarding the snow contribution, I have also doubts about what has been stated on p. 11759, line 28 and p. 11760, line 1: are the authors assuming that NEXRAD products provide snowfall (which they don't)?
8) The description of the methodology is not complete and some details not well explained. I think that more symbols and equations should be introduced to explain better each step, along with a figure that shows a schematic of the approach and an example of a basin (I found Fig. 2 not informative at all).

My concerns on the frequency estimation are described in the following points.

9) In extreme value theory, recurrence intervals are calculated for independent events, either deriving annual maxima or through the peak over threshold approach. In both cases, a time series of a variable observed at a location or a basin is used. In the paper under review, the computation of the recurrence interval accounts for all events observed in all basins of the same drainage area. Assuming that we have N basins with the same area (e.g., 64 km^2) included in the Upper or Lower CRBs, this implies that the recurrence interval is calculated by pooling together N time series of a variable. Through this method, the authors could present discharge values for the 500-year return period, using 10 years of rainfall records. However, since storms may have happened at the same time in contiguous basins, the events may not be statistically independent, as they are originated from the same weather pattern. In other words, increasing the sample size with records of contiguous basins is not a trivial operation, which requires careful evaluation. This may contradict the principle of extreme value theory. Addressing this issue is crucial to build the FMA curves and the authors have not provided any justification.
10) 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 aggregation times, chosen arbitrarily. This choice has to be supported as well.