Responses to comments: Anonymous Referee 2

The paper by Hoey et al., entitled 'Flood estimation for ungauged catchments in the Philippines' aims at delivering design equations to estimate flood magnitudes and frequency in ungauged catchments across the Philippines.

The paper addresses an interesting topic, the design of equations to predict floods in data scarce region. The authors do enormous work to analyse the data. However, the methodology is not well defined/structured. Overall, the paper lacks clarity and needs to be restructured for better clarity.

General comments:

The title of the paper does not reflect the content of the paper appropriately. The authors propose and evaluate the index flood approach and multi-variate regression to estimate flood. The approach is evaluated only at catchments with data used to fit the regression equations. In a context of flood estimation in ungauged catchments, the paper is lacking the following steps: 1) a clear definition of the methodology used for regionalization to ungauged catchment, 2) a cross-validation to evaluate the performance of the designed equations in ungauged catchments. It is necessary to evaluate how well the proposed approach would perform in ungauged catchments. Therefore, I suggest the authors consider using some cross-validation approach (i.e., leave-one-out or k-cross validation) where a set of catchments are used to fit the design equations; then the remaining catchment(s) are used as pseudo-ungauged catchment(s) to evaluate the accuracy of the designed equations in ungauged catchments. In addition, the methodology used is not clearly described (see specific comments below).

Response: These are all very good suggestions, based on the title of the paper. As noted in our response to Referee 1, this title is not accurate and we propose to amend the title to fit the contents of the paper. The reviewer's suggestions are exactly what we would do if the paper was to retain its current title, and there is merit in presenting such an approach and flow chart to guide readers in the use of the equations. However, that would lengthen the current paper excessively and we consider that this is a separate piece of work.

The paper could benefit from a flow diagram that describes the steps followed to estimate floods at ungauged catchments.

Response: This could be added as part of a short (approx 400 words) new section 6.5 'Applying the equations to ungauged catchments', although this would extend the paper. We would welcome the editor's view on this addition – it would be easy to produce, but would extend an already lengthy manuscript.

Specific comments:

- The authors fit and compare the accuracy of the three distributions showing that the choice of the distribution influences flow estimates and no distribution performs well at all sites. Then, the GLO distribution is selected to predict high flow magnitude in section 4.2 with no justification of this choice.

  Response: There is an error in line 215, in that the Q_T values used were from the **best-fit curve** at each site, and not only from the GLO results. This will be corrected.

  In addition, Equation 5, the factorial standard error for the GLO distribution, is only applicable when number of records is at least 20 years. The study region has only 71

sites (line 245), with more than 20 years data, among the 466 sites retained for the analysis (Line 234) . Why choose the GLO distribution if it cannot be applicable to all sites? Why choose only GLO instead of using the best fitted distribution at each site as shown in table 2?

Response: See previous comment – clarifying the curves used will address all of these points.

- The last paragraph in section 4.1, the authors lists three ways the construct growth curves by combining the curves fitted from each site by region, climate type or catchment areas. Therefore, I am assuming that the resulting curves are regionalized curves where catchments from within each group(region or climate type, catchment area bin) would be represented by a single curve. Correct In other words, catchments of each group would have a single Q_T value estimate from the regionalized growth curve. How the author use these results to perform a correlation analysis with catchment properties in Section 4.2 (Line 212: values of Q_T provided by the GLO analysis previously, were correlated with catchment properties")?

  Response: The analysis reported in 4.2 is for all catchments individually, and does not used the regionalised growth curves. Section 4.2 does not follow from the final paragraph in section 4.1 as the reviewer has assumed. We will revise the wording in lines 205-213 to make it clear why we are producing combined growth curves and how these can be used, and can similarly edit line 215 to clarify that section 4.2 is concerned with individual catchments.

- The authors use precipitation dataset covering only 17 years (from 1998-2015) which is not sufficient to derive climatology of the catchments. In addition, this dataset are not available for a period that overlaps with flows records period. As suggestion, the author could use precipitation datasets from reanalysis products (e.g. ERA5) which are available from 1940 to present around the globe. It would cover the same period as flow records for most catchments and provide sufficient long timeseries to derive climatology for precipitation variables used in Table 3.

  Response: This is an interesting suggestion. We are aware that none of the available precipitation datasets is perfect for our analysis, as we are combining river flow data from different time periods. We used APHRODITE V2 which, as noted, covers the period 1988-2015. Compared with the other precipitation reanalysis products that are available, APHRODITE has the advantage for our study of being focused on Asia and so is likely to be more reliable in that region. The V2 product is regarded as an improvement over APHRODITE V1 in its representation of extreme events which is another benefit for using it in our study. For both of these reasons, we are satisfied that this data set is appropriate, but note that a comparison between different precipitation estimates would be an interesting area for future work.

- Why the fitted curves from individual sites are not used to fit a relationship between flood magnitude and catchment area?

  Response: Table 5 presents the results of these fits and compares them with equations that also include a precipitation term.

- It is unclear how the individual short historical records are combined to generate large dataset that produce consistent results.

Response: This is covered in section 6.3, and we can make it more explicit. In summary, the absence of systematic differences between results from the different data sets (eg Figure 5a-c) provides confidence that we can treat the data from all data sources as a single data set for the analyses presented in the paper.

- Results presented in Section 5.3.3 (Eq6a-6c) would be better presented in a table with equation and associated R^2 and RMSE in the same way as in Table 5.

  Response: That appears to be a good suggestion for consistency within the paper.

Overall, the way the methodology and the results are presented is incoherent and lacks clarity. Other than the specific comments above, it is difficult to know what the reviewer is referring to here. Both reviewers have made specific comments that will improve the clarity of the results. The mis-leading title of the paper may have caused some confusion, and clarifying this will hopefully also aid the clarity of the paper.