Nat. Hazards Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/nhess-2017-187-AC2, 2017

Natural Hazards and Earth System Sciences
Discussions

Interactive comment on "Developing drought impact functions for drought risk management" *by* Sophie Bachmair et al.

Sophie Bachmair et al.

sophie@bachmair.info

Received and published: 13 August 2017

Thank you for the detailed and constructive feedback to our study. We appreciate the minor editorial and clarification comments and will present a reply to all minor comments with the revision. For the online reply, we focus on the major/general comments only:

"1. The study currently only looks at counts of impacts and not the actual height of impact. I wonder whether the authors checked the relation between the count and actual height of impacts, whether this relation is positive of negative. Could the authors elaborate further on this and on the question what would happen with the impact functions if height of impact is taken into account?"

We presume that by "height of impact" the reviewer means some numerical measure of impact severity. This type of information is not available in the text-based drought impact reports per se. We use different methods of quantifying the coded drought impacts: one is the use of binary impact information (presence versus absence of an impact per month) and another is the count how many impacts were reported for a given month. The latter method considered different ways of counting (see methods section 2.3). By counting the number of impacts reported we obtain a measure of impact severity, although this reflects the comprehensiveness of the drought impacts rather than relating a particular value of the SPI/SPEI to, say, a particular value of agricultural yield loss, or similar. Yield data would provide an objective measure of impact severity, but it would only reflect a very limited range of drought impact types.

Other options to assign a severity measure to impact data were not found reasonable. The EDII database provides impact report data that are based on text reports and coded into a very detailed system of impact categories and subtypes. The subtypes in some, but not in all categories may represent different levels of severity of an impact (for water supply they range from awareness, to bans, to actual supply restrictions, for example), as discussed in Stahl et al. (2016). However, there is no additional severity coding in the EDII database, i.e. no information about impact severity in a standardized/objective way is currently available from the database (see Stahl et al. (2016)).

During early stages of this project we in fact conducted a small test, asking a group of people to rate the severity of selected drought impact reports according to severity classes (low-medium-high). This small test revealed the complexity of how impact severity is perceived depending on the impact category, knowledge about drought impacts, affected area, and level of detail in the report. Given the subjectivity of text-based data that are available, we used different impact quantification methods, which have not been addressed so far. Exploring different methods for counting the number of impact reports as we did in our study is in our opinion the best way currently to somehow

address impact severity. We will clarify this in the paper.

"2. The authors fit "damage functions" based on a 'leave-on-out' principle. To me it isn't a surprise that with such an approach high correlations/good results are being found. I'd suggest the authors to do some extra sensitivity testing on this issue: e.g. leaving out more variables in the fitting. Could the authors elaborate more on how stable this relation is then? Up to what level (# of points left out) are still reasonable results achieved?"

The aim of the study was to test and compare three data-driven models for linking drought intensity with drought impacts, applying each model in the best possible way with regard to 1) number of data points for fitting, and 2) choice of predictor variables. The reviewer seems to be concerned about both issues, i.e. number of data points for fitting (addressed in this comment) and selection of predictor variables (see minor comment 12 about "Could you elaborate a bit further on whether taking these variables make sense from a physical point of view? And how about double-counting of drought mechanisms?").

Number of data points for fitting: Leave-one-out cross-validation is a very common approach. Given the issue of impact data scarcity, we think that we should use the largest dataset possible for model fitting. Clearly, less data points for fitting will deteriorate results. If our focus was on designing impact functions for operational use, we agree that further testing of the effect of sample size would be needed. However, for the purpose of comparing different approaches for drought impact functions (i.e. relative to each other) in our opinion a leave-one-out cross-validation is appropriate. We will add some extra text to the discussion explaining the reasoning of this issue.

Choice of predictor variables: The selected predictor variables for each model make sense from a physical point of view. In most models a combination of shorter-term and longer-term time scales of SPI or SPEI was selected, and the month or year of impact occurrence (the variables M and Y are introduced on P4 L 20-22). While for the RF

model all predictors are used, the above-named ones were also identified as the most important ones. Finding and interpreting 'best-predictors', i.e. physical indicator relating most strongly to impact occurrence, has been the focus of preceding studies, e.g. Stagge et al. (2015b) or Blauhut et al. (2016). The aim here was not to repeat again the previous results like 'longer duration water deficits relate more to water resources impacts because of longer response times of such systems', which have been previously shown. The aim was to compare the different methodological approaches and to initiate the idea to work towards potential impact function derivation. Nevertheless, we agree that a short paragraph on the differences of the selected predictors will likely be useful to the reader and we suggest to add this into the discussion section. We are not entirely sure what is meant by 'double counting'. Possibly the interrelation of predictors in a multiple predictor model? In our view this is not an issue given the way we selected the predictors excluding highly correlated ones (P 7 L 5-7).

"3. From reading the methods it does not become clear to me how exactly you coupled a gridded product (SPI/SPEI) to counted impacts over the basin in SSE. And this you use an equal time-period to establish the fits for the different accumulation times? Please elaborate further on this."

Thank you for pointing out that this is currently not clearly described. For each month we calculated the regional average of all E-OBS grid cells falling within the polygon covering South East England. The regional average was chosen since Bachmair et al. (2015) found little difference between the performances of different regional indicator metrics (e.g. mean vs. minimum vs. maximum etc.). The SPI/SPEI accumulation durations reflect the water deficit accumulated in the SEE area over that duration, and we relate this to the number of impacts occurring in the single month following the SPI/SPEI accumulation duration. We will add this information to the Methods section.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-187, 2017.