Interactive comment on "Probabilistic modelling of the dependence between rainfed crops and drought hazard" *by* Andreia F. S. Ribeiro et al.

Andreia F. S. Ribeiro et al.

afsribeiro@fc.ul.pt

Received and published: 7 June 2019

We would like to thank the Referee#1 for his/her careful review and constructive feed-back, and also for the opportunity to engage in a stimulating discussion. We truly believe that this process will enhance and clarify the paper's content. Author's responses follow below identified as AR next to the Referee#1 comments.

Referee#1: The study aims at advancing agricultural drought risk management through providing a probabilistic model for assessing the risk of crop loss and drought for two regions of the Iberian Peninsula with rainfed cropping systems. The authors apply the concept of copula models for this purpose and infer probabilities of crop loss under drought and non-drought conditions for the two regions and two crops. Overall, the

authors address an important topic by transferring the use of copula models to agricultural drought risk management. The paper is very well organized, easy to follow, and clearly written. The methods are precisely described, including a flowchart of the concept. Also, adequate reference to existing literature is presented. However, I am concerned about two issues that should be further addressed:

Referee#1- 1) Data and methods: The study focuses on crop loss of rainfed agriculture due to drought. While the data base is 27 years, only 7 years seem to be years with drought conditions according to Figure 3. Since only one value per year is retained for analysis, this seems like a very small sample size for making robust probabilistic inferences, especially regarding the tail of the distribution.

AR: In response to the sample size issue, Fig. 3 (2nd row panels) shows that the number of years under drought conditions varies with the location (cluster 1 or 2) and with the type of crop (wheat or barley). For instance, in the case of wheat in cluster 1, from a set of 27 years more than half of the years correspond to drought conditions (14 years). However, in the remaining cases (wheat in cluster 2 and barley in clusters 1 and 2), the number of drought events in the samples are lower, as pointed out by the Referee. This feature highlights the singularity of the extreme events and subsequently how challenging they are. Nevertheless, we totally agree that the ideal situation would be to have more than 27 years of data, however to increase the time period, we would have to sacrifice the use of the remote sensing of vegetation, which is currently an important source of information, complementary to climate variables. We must also stress that in several regions of the world, namely in developing countries, remote sensing data from the last 30 years is the only data that is available for crop losses studies. In addition, we also agree with the Referee that the importance of sample size increases with the complexity of the models and, for this reason, in page 5 lines 24 to 28, we highlight that we have restricted the analysis to the bivariate case using two-dimensional copulas, because the development of higher dimensional copulas exhibits very complex structures and entails additional requirements. Moreover, in page 6 lines

20 to 23 we point out that "The drawback of the shorter sample size is surpassed by the nonparametric estimation of the margins, which avoids heavy assumptions about their distributions, even when the available sample is rather small (Corder and Foreman, 2011; Fahr, 2017)." Nevertheless, despite the fit being performed with n=27, the great asset of the proposed methodology is to estimate probabilities of crop-loss based on 1000 simulations preserving the dependence structure, overcoming the sample size issue in the estimations of the probabilities of crop-loss.

Referee#1: One could argue that the study is not only focused on drought, since crop loss under non-drought conditions is also investigated; however, if the authors' intention was to assess the risk of crop loss more generally, in my view a stronger focus on other hazards or drivers of crop loss would have been beneficial to include.

AR: Thank you for the valuable comment which wisely promotes a deep and stimulating reflection on the nature of the stated problem. In fact, the understanding of the concurrent or sequential occurrence of different hazards which lead to extreme impacts is currently a challenging topic. The combined condition of more than one hazard would be a more complete approach to address agricultural losses in general and the development of higher dimensional models with even more complex structures is an optimal suggestion to expand the scopes of the current study. However, this is a relevant topic that requires a detailed analysis being now out of the scope of the present work. Here we focus on the influence of drought hazard on agricultural productions because drought is recognized as the main cause of crop loss on a global scale (Lesk et al., 2016), with particular relevance in the Mediterranean region (Lesk et al., 2016; Zampieri et al., 2017). The proposed methodology assesses a hazard that is still currently a hot topic in several sectors, including research, and which is likely to be more frequent. In addition, in page 3 line 16 in the introduction section we state the main objective of the present study: "In this study, a copula-based approach is adopted to model the joint probability density function of crop yield and the drought conditions for probabilistic yield assessment" and in the conclusion section (page 11

line 20) we reinforce the idea that "This study investigated the usefulness of copula methods in estimating the likelihood of drought risk in wheat and barley cropping systems, when applied to two regions in IP". Nevertheless, the results of barley in cluster 2 suggest that other factor than water stress is the cause of crop failure (page 11 line 6 and 7), and for this reason the assessment of whether the occurrence of other climate extremes is amplifying the impact of droughts on agricultural productions is the right direction to further next research (however out scope in the present study). We hence propose to improve the future work in the last paragraph (in page 12 line 28) based on this suggestion to: "Other potential usefulness of this methodology for future research is the evaluation of its suitability at the province level and the assessment of whether other hazards (such as heat waves) are amplifying the impact of droughts on crop's harvest."

Referee#1: In addition to the sample size issue, the spatial and temporal scale of the selected drought indicators seems rather coarse. Spatial scale: I assume some spatial average or aggregate measure of drought indicators over the two clusters was used (not clear from the manuscript). How representative is such a spatially aggregated drought intensity measure for explaining crop loss, which likely results from small-scale interactions of hazard, vulnerability of the plants, soil type, additional irrigation systems in parts of the clusters, etc.?

Regarding to the spatial scale issue, we agree that the use of spatial averages over the clusters is not clearly written in the manuscript, and in the below Detailed comments section 2.1 of the Referee#1 we give a more detailed answer with rephrasing suggestions in order to improve this issue in the manuscript. Regarding the representativeness of the used spatial scale, in fact, most of the studies are at the national scale or local case studies. However, despite the small-scale interactions between crops and other factors, as pointed out by the Referee, the availability of fine-scale harvest data with common time periods in order to have a crop-specific database for the Iberian Peninsula is very limited. The use of longer time series at the province level is a good

compromise and explains rather well the response of the cereals to the drought conditions as shown in a previous study by (Páscoa et al., 2017b). Moreover, drought usually affects large areas, and in the Iberian Peninsula, drought events affecting more than 40% of the territory accounted for 33.8% of all the events in the period 1975-2012 (Páscoa et al, 2017a). The spatial averages over the clusters are the result of the exposure analysis to ensure that we are addressing the annual yield response in the regions dominated by rainfed conditions, which are more vulnerable to droughts (Ribeiro et al., 2018).

Referee#1- Temporal scale: monthly SPEI (as used for cluster two) is a commonly used indicator; however, for this specific purpose, wouldn't sub-monthly drought conditions be more suitable given the effect during different growth stages, as the authors point out themselves? Given these data issues, I am a bit concerned about the meaningfulness of the inferred results on probability of crop loss under drought and non-drought, as presented in section 3.2.

Regarding the temporal scale, despite the short-term response of the plant's activity to the dry conditions (here considered by the weekly TCI and VHI) the integrated effect of droughts at the monthly scale or longer (here considered by the monthly SPEI) is also relevant to quantify the mechanisms that can lead to annual crop loss. Moreover, the methodology considered for selection of one drought indicator for each case reflects the importance of the different temporal responses of the cereals to climate conditions along the vegetative cycle and their impact on the cereal's annual yield. In page 4 line 30 to 33, and page 5 line 1 is written in the manuscript: "First, a pool of the drought indicators better related with wheat and barley yield were chosen based on stepwise regression (95% confidence level), selecting the most statistically significant time scales and months of SPEI, together with the weeks of VHI, VCI and TCI. Afterwards, from the pool of selected drought indicators, the one with largest contribution to the yield's variance was selected (Table 1)." Since this aspect is not clear in the article, we propose to rewrite to the following (in page 5 lines 4 to 10): "This selection of

drought indicators highlights that the response of crop yields to climate conditions vary according to the location, type of crop, moment of the vegetative cycle and the temporal scale. While annual yield anomalies in cluster 1 are better characterized by short-term responses to the drought conditions based on the weekly values of TCI and VHI, the annual yield anomalies in cluster 2 are better characterized by the monthly response to the dry conditions based on the SPEI. Moreover, in terms of predictability, the effects of temperature (TCI) and vegetation health (VHI) during late growth stages (weeks 23 and 22 correspond approximately to end of May and beginning of June, respectively for wheat and barley) are the most influential conditions in the northern cluster. On the other hand, the yields in cluster 2 are influenced by drought conditions described by SPEI much earlier, in the beginning of the intermediate growth stages (February and April with 5 and 1 month of time-scale, respectively for wheat and barley). In this way, the importance of including multiple drought response time scales is evidenced for predictability purposes and assessment of drought-related crop-losses." Moreover, to meet the suggestions of Referee#2 regarding this issue, we also propose to move Table 1 and the previous paragraph to the results section.

Referee#1- 2) Aim of the study versus methodological setup: The aim of the study is to quantify the risk of crop-loss under drought and non-drought conditions for two regions and two crop types; for this purpose the authors model the dependence between drought conditions and crop anomalies. What actually is the practical value of assessing the probability of rainfed crop loss under non-drought conditions with this setup, where the focus is solely on drought indicators?

AR: The indicators used in the present study enable the measurement of both dryness and wetness in the case of SPEI, and both stress and good-health of vegetation in the case of VCI and TCI. Hence, the present study is focused on the losses associated to the measures of water balance indicated by SPEI (in the case of cluster 2) and on the losses associated to the measures of vegetation health indicated by VCI and TCI (in the case of cluster 1). As shown by Fig. 6 boxplots, the yield anomalies reduce sig-

nificantly during drought conditions, while higher yield anomalies values are expected during non-drought conditions (which are characterized by the drought indicators as well). In addition, Fig. 7 shows the distributions of the yield anomalies given different climatic conditions and that droughts lead to major crop losses. A similar setup was considered by Madadgar et al. (2017) in Australia, which assessed droughts based on Standardized Precipitation Index (SPI) and the Standardized Soil-moisture Index (SSI), and have estimated the probability of crop yields exceeding its annual average (while we estimate the probability of non-exceedance -1 standard deviation, i.e., the probability of more extreme crop losses) in dry conditions (SPI or SSI< - 0.5) as compared to normal/wet conditions (SPI or SSI> - 0.5). Similarly, Madadgar et al. (2017) found that a shift from wet (SPI> -0.5) to dry (SPI<-0.5) causes yields of Australian rainfed crops to decrease.

Referee#1: The aim of the study one the one hand and the deployed data and statistical methods one the other hand seem a bit separated or not well aligned at times. While for drought conditions the data does not seem fully suitable for analysis (see issue 1 on sample size), for non-drought conditions the value of the analysis is not entirely clear to me, at least how it is presented currently. It would be beneficial if the authors could elaborate further on the applicability of their method for agricultural drought risk management.

AR: As well stated by the Referee, the fitting is performed with drought years and non-drought years and the probabilities of crop-loss, estimated based on 1000 simulations, distinguish between drought and non-drought conditions. The importance of estimating CPNE|non-drought is to quantify the losses associated to mechanisms not related to dryness directly (for instances flash extreme events during non-drought conditions that lead to crop-loss). In order to discuss whether drought is the main determinant of the crop loss, the following decomposition is required: PNE = (CPNE | drought) * Prob(drought) + (CPNE | no- drought) * Prob (no drought). If drought is the main conditioner then the 1st term is determinant, otherwise it will be the 2nd as it happens

in the case of barley in cluster 2 and almost in the case of wheat in cluster 2. This quantification is important and therefore it is of all relevance to include the probability CPNE | non-drought. In addition, it allows to quantify that the risk of crop-loss increases 32%- 55% (depending on the cereal and cluster) when drought conditions are below the applied severity thresholds. The consideration of other drought severity thresholds for comparison is also a good suggestion to pursue in future agricultural drought risk studies.

Detailed comments: Referee#1: 2.1 A bit more background how the drought indicators per cluster and crop were calculated and selected in the authors' previous study would be useful (especially how the spatially variable indicators were regionally averaged). In contrast, details about SPEI and vegetation condition indicator calculation could be shortened by pointing to existing references.

AR: Thank you for the suggestion, we agree that the use of spatial averages over the clusters is not clearly written, and we propose to rectify that in a revised version. In order to meet this comment from Referee#1 and other suggestion from Referee#2 we propose to remove from Page 4 line 15 the sentence "Further spatial averages were computed for each cluster of provinces." and rephrase in Page 4 line 28 "The spatial averages of VCI, TCI, VHI and SPEI were computed for each provincial cluster and used for further modelling of the joint probability between the drought hazard and cereal yield anomalies. " Moreover, in order to shorten the data section we propose to rewrite from page 4 line 3 to 27 to the following: "Drought conditions were investigated using two types of indices: the hydro-meteorological drought indicator SPEI and the satellite-based Vegetation Condition Index (VCI) (Kogan 1990), the Temperature Condition Index (TCI) (Kogan 1995) and the Vegetation Health Index (VHI) (Kogan 1995). The monthly drought index SPEI gridded values, with spatial resolution of $0.5°$, were computed based on precipitation and temperature values from the Climate Research Unit (CRU TS3.21), using a variety of time scales (1 to 12 months). The weekly global maps of VCI, TCI, and VHI were retrieved at 4km spatial resolution from NOAA's ftp server

(ftp://ftp.star.nesdis.noaa.gov/pub/corp/scsb/wguo/data/VHP_4km/geo_TIFF/). While SPEI computation uses climatic water balance anomalies incorporating the role played by the evaporative demand on the occurrence of dry events (Vicente-Serrano et al., 2010), the remote sensing indices characterize the moisture, through the VCI, the temperature induced stress through the TCI and health of vegetation, through the VCI."

Referee#1- Fig. 3: What is the dotted horizontal line?

AR: Thank you for pointing this out. We propose to rewrite the caption of Figure 3 to the following: Figure 3 – Empirical cumulative distribution functions (ECDF, blue points), kernel density estimation of the CDF (red line), crop-loss and drought thresholds (dotted black vertical line), respective marginal probabilities of crop-loss and drought occurrence (dotted black horizontal line), and pseudo-observations (scatter) of the margins on the interval [0,1].

Referee#1- Fig. 4 bottom row: Due to the filled marker style it is hard to discern details; I'd suggest making a density plot of points.

AR: Thank you for the suggestion, we propose the changes presented in the attached Figure and the following changes in the caption: Figure 4 – Selected joint Probability Distribution Functions (PDF) where u and v are scalar values on the interval [0,1] (top), contours showing the two-dimensional view of PDFs (middle) and observed (red triangles) and copula-based simulations (density squares) scatter plots of crop yields and drought indicators (bottom).

Referee#1- References: duplicate reference Ribeiro 2018a and b

AR: Thank you for pointing that out.

References Zampieri et al., Wheat yield loss attributable to heat waves, drought and water excess at global, national and sub-national levels. Environmental Research Letters, 12(6), 064008, https://doi.org/10.1088/1748-9326/aa723b, 2017. Lesk et al., influence of extreme weather disasters on global crop production. Nature, 529,

p.84-87, https://doi.org/10.1038/nature16467, 2016.

Please also note the supplement to this comment:
https://www.nat-hazards-earth-syst-sci-discuss.net/nhess-2019-37/nhess-2019-37-AC1-supplement.pdf

———————————————

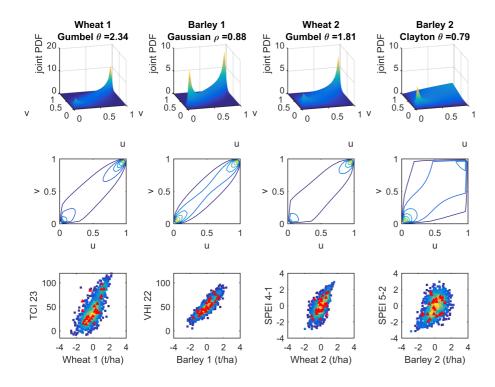Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2019-37, 2019.

Fig. 1. Please see respective rephrased caption of the manuscript Fig. 4 in Author's responses