Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2020-116-AC3, 2020


# *Interactive comment on* “Risk of crop failure due to compound dry and hot extremes estimated with nested copulas” *by* Andreia Filipa Silva Ribeiro et al.

**Andreia Filipa Silva Ribeiro et al.**

afsribeiro@fc.ul.pt

Received and published: 24 June 2020

Reviewer: This is a well-written manuscript that investigates the compound effects of precipitation and temperature on crop failure in two provinces in Spain using nested copulas. It contributes to a better understanding of these type of compound events and it therefore deserves publication.

Author’s Reply: Thank you for this positive assessment.

Reviewer: I have just some minor comments that I recommend the authors to address before publication: (1) p.4 line 90. Is it monthly daily mean precipitation or monthly



accumulated precipitation? Please specify.

Author’s Reply (1): Thanks for recommending the clarification, CRU TS4.01 provides monthly cumulative precipitation, hence we will clarify in the revised manuscript:

“Therefore, monthly accumulated precipitation (P) and monthly maximum temperature (Tmax) were extracted from the Climate Research Unit (CRU) TS4.01 dataset (. . .)”

Reviewer: (2) p.5 Figure 2. I would highlight the month MAM as it is the choice to of this study to perform the compound event analysis.

Author’s Reply (2): We appreciate the suggestion and propose to denote MAM in bold text (please see updated Figure 2 further below).

Reviewer: (3) p 7. line 152. I think this Section should explain what the choice for the marginals is and why.

Author’s Reply (3): Thank you for the comment. We used empirical ranks as explained in the methods sections. We suggest to add the following text:

The main steps of the trivariate approach used in this study can be summarized as follows (Okhrin and Ristig, 2014). First, the marginal distributions u1, u2 and u3 are estimated non-parametrically by simple ranking, using the empirical distribution functions of the data through the pobs R function, a common approach for copula modelling.”

Reviewer: (4) p 7. line 162 The authors say they invert the margins due to negative dependence between temperature and precipitation. Why not use a rotated copula that can represent negative correlation instead?

Author’s Reply (4): Thank you for the question. A similar comment was raised by the reviewer 1. As we also answered to reviewer 1, the required complete monotonicity of the ACs generators to construct NAC following Okhrin and Ristig (2014) implies (i) that the same single-parameter generator function is used on each level of NAC (i.e. same family), but with a different value of $\theta$ (as we discuss in lines 53-56, 135-40 and 263-

264 by other words) and (ii) positively dependent AC models, hence the pairwise rank correlations are required to be non-negative. Therefore, nested rotated copulas are not covered by the NAC approach following Okhrin and Ristig (2014). For this reason, in order to model positive dependencies among all possible pairs, we considered the inverted values of Tmax (i.e. multiplication by $-1$). For more details on complete monotonicity of the ACs generators and NAC constructions see e.g. Górecki et al. (2017).

We will clarify this in the revised manuscript by moving the referred information (as the required complete monotonicity of the AC generators implies both conditions) and improving to:

"Using the same single-parameter generator function on each level of NAC (but with a different value of $\theta$) satisfies the required complete monotonicity of the ACs generators to construct NAC following Okhrin and Ristig (2014), which also implies that the possible pairs are positively dependent. Therefore, due to the negative dependence between TmaxMAM and both crop yields and PMAM, we inverted the margins of TmaxMAM for copula modelling (i.e. multiplication by $-1$). For more details on complete monotonicity of the ACs generators and NAC constructions see e.g. Górecki et al. (2017)."

References: Górecki, J., Hofert, M. and Holeňa, M.: On structure, family and parameter estimation of hierarchical Archimedean copulas, J. Stat. Comput. Simul., 87(17), 3261–3324, doi:10.1080/00949655.2017.1365148, 2017.

Reviewer: (5) p.8 line 183 For wheat 2, it seems that the statistical model tend to produce a larger kendal correlation between temperature and wheat (FIgure 5g) than what is seen in the observations (the observations are almost outside the confident interval obtained from simulations). Could the authors explain why the performance for this specific case seems to be worse?

Author's Reply (5): As a matter of fact, a similar feature occurs in the case of barley



(Fig. A.1 - h). This applies for the correlation between Tmax and the crop yields. The explanation for this feature may be related to the construction of the NAC models, which is defined by the pair (P,yield) in the inner level due to their stronger correlation (Table A.1). In addition, the correlation between Tmax and yield, is also lower than the correlation between P and Tmax (Table A.1). For this reason, Tmax and yield is the pair with lowest correlation and hence the model is likely to struggle in its representation. Nevertheless, in both cases (wheat and barley), the simulated level of dependence is inside the 95% confidence level and the magnitude of correlations among the pairs is also preserved i.e., such that $\tau$_(u_1,u_2 )>$\tau$_(u_2,u_3 )>$\tau$_(u_1,u_3 ).

We will clarify this on Results section of the revised manuscript:

"Bivariate dependencies as measured by Kendall's are captured well by the fitted models (Figure 5 for wheat, Figure A.1 for barley). Among all possible pairs, the correlation between Tmax and yield is the lowest for the case of both cereals (Table A.1), and for this reason it is the pair in Figure 5 and Figure A.1 with observational $\tau$ closest to the lower bound of the 95% confidence intervals (Figure 5f,h and Figure A.1f,h). Nevertheless, in both Figure 5 and Figure A.1, the simulated level of dependence is inside the 95% confidence level and the magnitude of correlations among the pairs is also reasonably preserved by the models i.e., such that $\tau$_(u_1,u_2 )>$\tau$_(u_2,u_3 )>$\tau$_(u_1,u_3 )."

Reviewer: (6) p. 13 line 256 The authors say that in some cases, draught or heat alone may cause more damage than concurrent drought and heat. I see this is the case for wheat 2 (Figure 7b). Any physical explanation to this? I would have assumed that regardless of draught playing a greater role, extreme values of these variables would both contribute to increase yield loss.

Author's Reply (6): Thank you for the question. The best estimates (bars in Figures 8 and A.3) show indeed that compound dry and hot extremes contribute to increase yield loss. Nevertheless, the lower bound of the 95% confidence intervals in Figures 8 and

A.3 show that drought or heat alone may cause more damage than concurrent drought and heat due to uncertainties associated to the parametric statistical model. This is associated with the uncertainties in the estimation procedure, which may be particularly large for extreme values and it would be difficult to find a physical explanation for such a feature.

We will clarify this in the Discussion section in respect to uncertainties:

“The uncertainties associated to the parametric statistical model were assessed with a large number of sampled distributions with the same sample size as the observations. In some of these distributions, drought or heat alone may cause more damage than concurrent drought and heat (lower uncertainty bound is below 0 in Figures 8 and A.3). This highlights the challenges of estimating the likelihood of rare events in two- or three-dimensional probability distribution with limited sample size (Serinaldi, 2013, 2016; Zscheischler and Fischer, in review). For the same reason, the wheat loss in Cluster 2 when PMAM is below the 5th percentile in Figure 7 slightly decreases when the threshold of TmaxMAM change from the 10th percentile to the 5th percentile (while an increase would be expected like in the other cases). These features are associated with the uncertainties in the estimation procedure, which may be particularly large for extreme values and it would be difficult to find a physical explanation for such a feature. Note that the uncertainties increase with the increasing severity of the compound dry and hot conditions (Figure A.3) due the rapid decrease of available samples in the corners of the three-dimensional probability distribution. Nevertheless, the best estimates (bars in Figures 8 and A.3) show that compound dry and hot extremes contribute to increase yield loss.”

---

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2020-116, 2020.



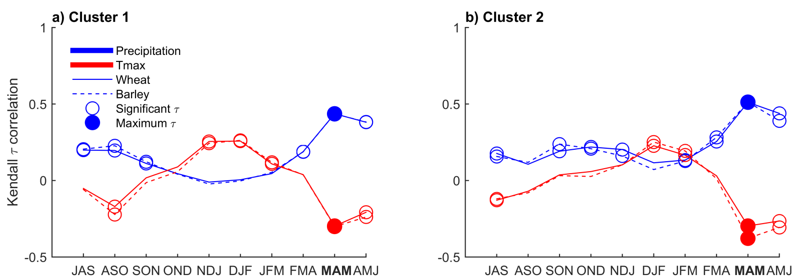


**Fig. 1.** Updated Figure 2