Dear Editorial Board,

Please consider our revised manuscript “Risk of crop failure due to compound dry and hot extremes estimated with nested copulas” by Andreia Ribeiro et al., which we would like to submit for publication in the Biogeosciences journal as an original research article, in the Special Issue “Understanding compound weather and climate events and related impacts”.

We truly appreciate the Reviewer’s and Editor’s feedback about the manuscript, which have been very helpful in improving the manuscript. Below we provide a point-by-point response to the reviews (from the three reviewers and editor) followed by the marked-up manuscript version. Accordingly, the main changes in the manuscript consisted in:

- Replace the word “cluster” to “region”
- Improve labels in Figures 2 and 7
- Clarification regarding the seasons, the vegetative cycle of the crops and the 3-month averaging window of the weather data
- Introducing the copula parameter
- Explain better the estimation of the margins
- Explaining the requirement of positive dependence
- Interpretation of the slightly lower AIC in the case of barley in Region 2
- Explain the reproducibility of the correlation between Tmax and yield
- Improve discussion about uncertainties
- The physiological explanation for the stronger yield reductions under compound dry and hot extremes

Additionally, the authors took the liberty of performing the following changes:

- Clarification about Figure 1
- Shortening of captions in Fig. 6 and Table A1
- Delete two redundant sentences in the abstract

The changes allowed to improve the quality of the paper and in case of publication we will publish the code in http://impecaf.rd.ciencias.ulisboa.pt/.

With my best regards,

Andreia Ribeiro (PhD student)

Anonymous Referee #1

Received and published: 14 April 2020

This study evaluated the risk of crop failure due to compound dry and hot extremes. A copula model is fitted to estimate the response of crop yield with respect to different dry and hot conditions. This manuscript is well crafted with clear structure.

Reply (0): Thank you for this positive assessment.

A few issues need to be addressed before the potential publication of this study.

(1) Selection of the periods Line 92-93: "We used 3-monthly means of Tmax and 3-monthly means of P during spring". Here the selection is based on the correlation analysis, but not the whole growing season, right? Please justify this period. It is easy to understand this from a statistical perspective. Is this selection still valid from a physical perspective?

Reply (1): Thank you for the question. Given the importance of assessing crop's water and temperature requirements at different moments of the vegetative cycle we conducted a correlation analysis between the yield and the 3-monthly means of precipitation and 3-monthly means of maximum temperature during the whole growing season (approximately from September of year n-1 to June of the year n), as shown in Fig. 2.

The identification of the moment of the vegetative cycle of the highest crop's water and temperature requirements was assessed based on the strongest statistically significant correlation value.

Fig. 2 suggests that the greatest influence of P and Tmax in crop yields is observed during spring (in both regions and cereals) corresponding to the moments in which the vegetation is photosynthetically more active. The effects of water content and high temperatures during middle growth stages of the crop life cycle are in accordance with previous studies (Ferrise et al., 2011; García del Moral et al., 2003; Iglesias and Quiroga, 2007; Ribeiro et al., 2019). Hence, this selection is valid both from the statistical and biophysical point of view.

We have addressed this aspect in the revised manuscript in the Weather data section:

"The vegetative cycle of the winter crops in Spain is mainly driven by precipitation and temperature: sowing occurs around autumn, followed by the vegetative phase in winter, reproductive phase (more photosynthetically active phase) in spring ~~(when vegetation is photosynthetically more active)~~ and crop harvest occurs in the early summer. Therefore, monthly precipitation (P) and monthly maximum temperature (Tmax) were extracted from the Climate Research Unit (CRU) TS4.01 dataset (Harris et al., 2014) spanning the same time period. Given the importance of assessing crop's water and temperature requirements at different moments of the vegetative cycle we conducted a correlation analysis between the annual yields and the 3-monthly means of P and 3-monthly means of Tmax during the whole growing season, as shown in Fig. 2. The identification of the moment of the vegetative cycle of the highest crop's water and temperature requirements was assessed based on the strongest statistically significant correlation value (denoted by filled circles in Fig. 2). Figure 2 suggests that the greatest influence of P and Tmax in crop yields is observed during spring (MAM in both regions and cereals) corresponding to the reproductive phase of plant development, when vegetation is photosynthetically more active. Therefore, for the remaining analysis we focus on 3-monthly means of Tmax and 3-monthly means of P during spring (P_{MAM} and $Tmax_{MAM}$, respectively), which has also been identified in previous studies as a growth stage sensitive to the effects of water content and high temperatures (Ferrise et al., 2011; García del Moral et al., 2003; Iglesias and Quiroga, 2007; Ribeiro et al., 2019). This selection of climate variables allows to maximize the dependence between climate conditions and yields as also shown by previous work based on the same data (Ribeiro et al., 2019c)."

References:

Ferrise, R., Moriondo, M. and Bindi, M.: Probabilistic assessments of climate change impacts on durum wheat in the Mediterranean region, Nat. Hazards Earth Syst. Sci., 11(5), 1293–1302, doi:10.5194/nhess-11-1293-2011, 2011.

García del Moral, L. F., Rharrabti, Y., Villegas, D. and Royo, C.: Evaluation of Grain Yield and Its Components in Durum Wheat under Mediterranean Conditions: An Ontogenic Approach, Agron. J., 95, 266–274, 2003.

Iglesias, A. and Quiroga, S.: Measuring the risk of climate variability to cereal production at five sites in Spain, Clim. Res., 34(1), 47–57, doi:10.3354/cr034047, 2007.

(2) Copula implementation Line 161: "Due to the negative dependence between TmaxMAM and both crop yields", The clayton copula does not permit the negative dependence. Is this the reason to "invert the margins of TmaxMAM for copula modelling"? The rationale of this transformation needs to be clarified. Suggest to make it clear to aid the understanding.

Reply (2): Thank you for the comment. The reason for inverting the margins is that 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 potentially with a different value of θ (as we discuss in lines 53-56, 135-40 and 263-264 in other words) and (ii) positively dependent AC models, hence the pairwise rank correlations are required to be non-negative. Therefore, 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 have addressed this in the revised manuscript by moving the referred information in line 161 (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 potentially different value of θ) 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 $Tmax_{MAM}$ and both crop yields and P_{MAM}, we inverted the margins of $Tmax_{MAM}$

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.

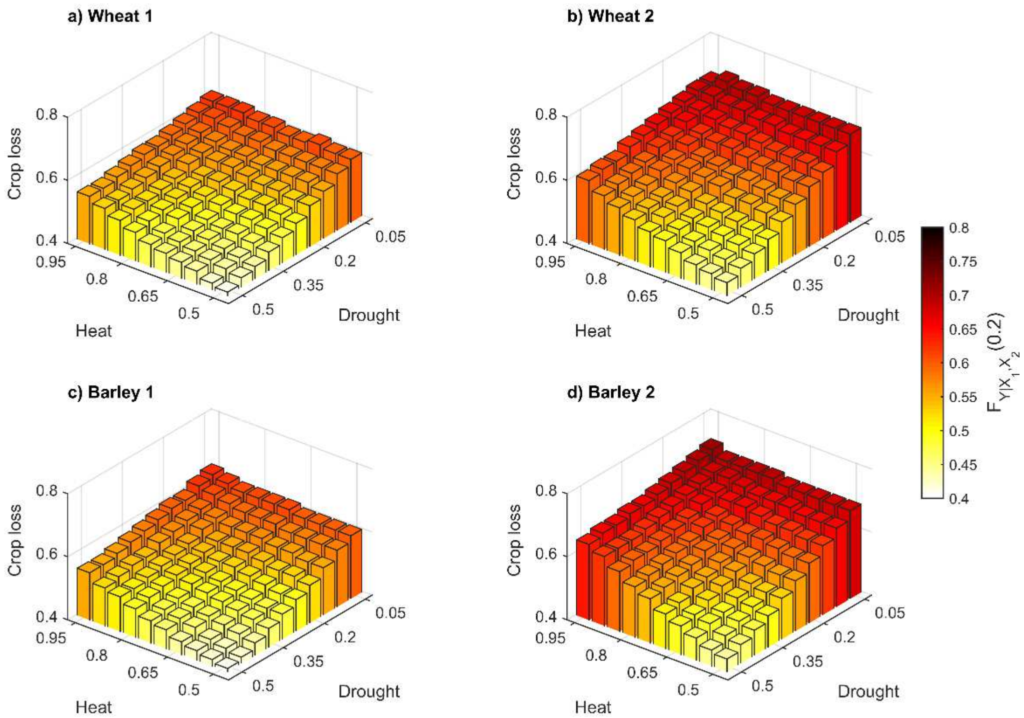
(3) Figure presentation Figure 7: "y-axis indicates the TmaxMAM percentile (Heat)". For heat, should you use the axis with the range like 0.5-0.95? Since for heat, we are interested in high percentile, right? Or if this is related to the aforementioned "inversion of the margins", please clarity this and make it clear.

Reply (3): You are correct. By inverting the Tmax the highest values correspond to the lower quantiles. We have made changes in the Figure y-axis to 0.5 - 0.95 to avoid confusion.

(4) Figure discussion Regarding Figure 7, "When PMAM/TmaxMAM are below/above the median, the probability of crop loss is always higher than 40%." How could you tell this (i.e., above the median?) from the figure? The y-axis for heat stress is below median. Please make it clear.

Reply (4): We agree that this point is not clear. In the text we refer to the Tmax values, rather than the inverted Tmax values as it was supposed to. In other words, when referring to Tmax in the text we keep the concept of exceeding the highest percentiles.

As mentioned above in comment (3), we have made changes in the Figure y-axis to 0.5 - 0.95 to avoid this confusion.



(5) Minor comments: Check the bracket in the caption of Figure 5.

Reply (5): Thank you, we have deleted the extra brackets in Fig. 5 caption.

Anonymous Referee #2

Received and published: 20 April 2020

The manuscript 'Risk of crop failure due to compound dry and hot extremes estimated with nested copulas' uses Archimedian copulas to model trivariate joint distributions between maximum temperature, precipitation and wheat and barley yield deviations. The paper is well structured and well written. It contributes interesting new insights to the literature on compound dry and heat impacts on crop yields.

Reply (0): Thank you for this positive assessment.

A few comments include:

(1) p.4, l.88: I recommend adding the exact months rather than just naming the seasons. As readers of the paper might come from all over the world, it might not be clear which months include the spring time in Spain.

Reply (1): Thank you for the comment. To clarify the season months in the study region we have modified the lines 88 – 90 in the revised manuscript to the following:

"The vegetative cycle of the winter crops in Spain is mainly driven by precipitation and temperature: sowing occurs around autumn (from September through November, SON), followed by the vegetative phase in winter (from December through January, DJF), reproductive (~~photosynthetically more active)~~ phase (more photosynthetically active phase) in spring (from March through May, MAM) ~~(when vegetation is photosynthetically more active)~~ and crop harvest occurs in the early summer (around June)"

(2) p.6, l.132: The copula parameter θ should be introduced

Reply (2): Thank you for the suggestion. We have modified the lines 127-128 in the revised manuscript to:

"AC can be written in terms of the respective generator function φ, which belongs to a parametric family (φ_θ) dependent on the parameter θ, e.g. for the three-dimensional case:"

in lines 130-131 to:

"Due to the symmetry of bivariate AC, the above trivariate form can be expressed in terms of NAC or HAC, where two of the margins are first coupled by their bivariate copula and then coupled with the third margin, via the same generator on each level but different parameters θ_{12} and θ_1, respectively, e.g.:

And in lines 131-132 to:

Equation 8 can also be expressed in terms of the other possible pair copulas $C_{13}(u_1, u_3; \theta_{13})$ and $C_{23}(u_2, u_3; \theta_{23})$ that are coupled with u_2 and u_1 by C_2 and C_3, with expressions $C_2(C_{13}(u_1, u_3; \theta_{13}), u_2; \theta_2)$ and $C_3(C_{23}(u_2, u_3; \theta_{23}), u_1; \theta_3)$, respectively. Like Eq. (8), among each structure of NAC the same generator is required for each level but with potentially different parameter. Hence, both the optimal structure and respective parameters must be determined.

(3) p.8, l.178: Can you interpret what it means that 'C_(u1,u2,u3) is slightly lower than the AIC of C_1 (u3,C_12 (u1,u2))' in terms of compound and single hazards and add a sentence about it?

Reply (3): This means that in the case of barley in Cluster 2, the trivariate copula fits the data slightly better than the two-parameter NAC C(u3, C12(u1,u2)) in terms of AIC, even though the Cramer-von Mises distance is better for the NAC. This could mean that in this case a NAC structure favouring the dependence between yield and precipitation may be less relevant compared to the other clusters and yields. Drought individually seems to play a less dominant role in the compound event, in comparison to the other cereals and regions.

This interpretation would be consistent with our discussion about Figure 8, where barley in Cluster 2 is also the case with the highest difference between drought and compound dry and hot conditions, hence illustrating that here drought is the least dominant driver of crop loss in comparison to the other cereals and regions.

Following the reviewer's suggestion we have added a sentence in the Results section:

"The only exception is barley in Cluster 2 whose AIC of Cθ(u1,u2,u3) is slightly lower than the AIC of Cθ1 (u3,Cθ12 (u1,u2)) (Table 2). This feature may suggest that a structure favouring the dependence between yield and precipitation (u1,u2) may not be as relevant as in the other clusters and yields due to a less dominant role of drought individually in this case. Nevertheless, in terms of Cramer-von Mises distance (Sn) the nested copula is the closer to the empirical trivariate copula. For this reason, we modelled the trivariate joint distribution based on nested Frank copulas for all cases. (…)"

(4) *Code availability. It is more and more common to publish the code used in scientific publications and I strongly recommend the authors considering to publish their code once the paper is accepted.

Reply (4): In case of publication we agree to publish the code in http://impecaf.rd.ciencias.ulisboa.pt/.

Anonymous Referee #3

Received and published: 5 June 2020

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.

Reply (0): Thank you for this positive assessment.

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?

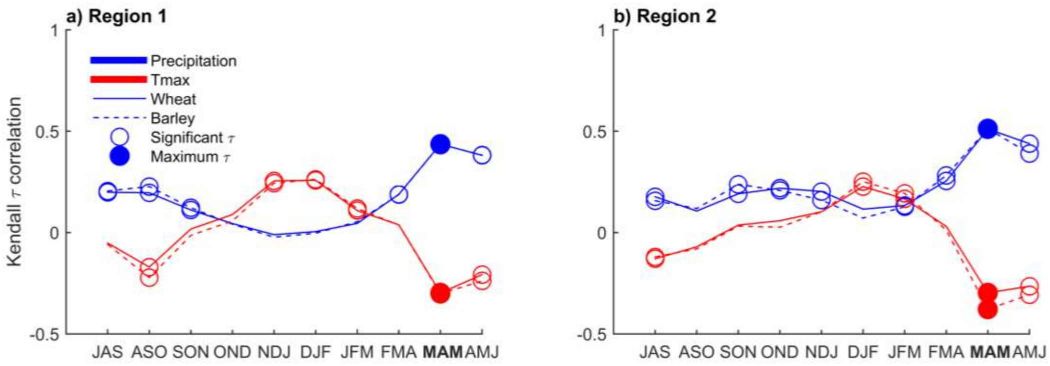
Please specify.

Reply (1): Thanks for recommending the clarification, CRU TS4.01 provides monthly cumulative precipitation, hence we have clarified 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 (…)"

(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.

Reply (2): We appreciate the suggestion and propose to denote MAM in bold text.



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

Reply (3): Thank you for the comment. We used empirical ranks as explained in the methods sections. We have added the following text in the revised manuscript:

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* function in the R package *copula* (Ivan Kojadinovic and Jun Yan, 2010), a common approach for copula modelling."

References:

Ivan Kojadinovic and Jun Yan: Modeling Multivariate Distributions with Continuous Margins Using the copula R Package, Journal of Statistical Software, 34, 1–20, http://www.jstatsoft.org/v34/i09/, 2010.

(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?

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 θ (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).

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.

We have addressed 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 θ) 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 $Tmax_{MAM}$ and both crop yields and P_{MAM}, we inverted the margins of $Tmax_{MAM}$ 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)."

(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?

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 have addressed 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 weakest for both cereals (Table A.1), and likely for this reason it is the pair in Figure 5 and Figure A.1 with observational τ 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 $\tau_{u_1,u_2} > \tau_{u_2,u_3} > \tau_{u_1,u_3}$."

(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.

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 have addressed 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 P_{MAM} is below the 5th percentile in Figure 7 slightly decreases when the threshold of $Tmax_{MAM}$ 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 indeed that compound dry and hot extremes contribute to increase yield loss.

Moreover, following the work by Okhrin and Ristig (2014), (…)"

Bart van den Hurk (Editor)

Received and published: 18 June 2020

The three reviewers agree on the clear purpose and structure of the manuscript and are all supportive to publication of this manuscript subject to minor edits. Some refer to justification of (implicit) choices made, some ask for some extra clarification. I would recommend to reply to all review comments that are raised.

Reply (1): Thank you. We have answered to all the review comments individually and in detail.

For me the most significant comments are:

Additional justification regarding:

• The choice of the 3-month averaging window for the meteorological quantities

• The processing of the negative dependence (why not use a rotated copula, as suggested by a reviewer)

• Availability of the code (and data if that is open)

Reply (2): Thank you. We have addressed these points in detail and in case of publication we agree to publish the code in http://impecaf.rd.ciencias.ulisboa.pt/.

Some further explanation concerning:

• The copula parameter {theta • The implication of C being lower than the AIC

• The choice of the marginals in section 2.3

• The reconstruction of the correlation between yield and temperature for wheat 2

• The physical interpretation, not only of the notion that drought and heat alone can give stronger effects than their combination, but in general sense: why would the combination of environmental drivers lead to stronger yield reductions? Is this physiologically explainable?

Reply (3): Thank you. These points where properly addressed. Moreover, in the general sense, the biophysiological explanation for the combination of environmental drivers leading to stronger yield reductions relates with the crop's requirements of water and thermal conditions during the key phenological stage in analysis. The selection of the climate variables during spring corresponds to the reproductive phase of the plant's and when vegetation is photosynthetically more active, and the combined effect of water and heat stress during this period is critical for crop's health leading to yield decrease. During this stage of formation of the grains the compound dry and hot extremes may accelerate the maturation reducing the size, number and weight of the grains and consequently reducing crop's harvests in quantity and quality (Balla et al., 2011; COPA-COGECA, 2003; Nicolas et al., 1984; Qaseem et al., 2019; Talukder et al., 2014).

In the revised version we have added the following text to the Discussion section:

"Nevertheless, the best estimates (bars in Figures 8 and A.3) show indeed that compound dry and hot extremes contribute to increased yield loss. In the general sense, the biophysiological explanation for the combination of environmental drivers leading to stronger yield reductions relates with the crop's requirements of water and thermal conditions during the key phenological stage in the analysis. The selection of the climate variables during spring corresponds to the reproductive phase of the plant's and when vegetation is photosynthetically more active, and the combined effect of water and heat stress during this period is critical for the crop's health leading to yield decrease. During this stage of formation of the grains the dry and hot extremes may accelerate the maturation affecting the size, number and weight of the grains and consequently affecting the crop's harvests in quantity and quality (Balla et al., 2011; COPA-COGECA, 2003; Nicolas et al., 1984; Qaseem et al., 2019; Talukder et al., 2014).

Balla, K., Rakszegi, M., Li, Z., Békés, F., Bencze, S. and Veisz, O.: Quality of winter wheat in relation to heat and drought shock after anthesis, Czech J. Food Sci., 29(2), 117–128, doi:10.17221/227/2010-cjfs, 2011.

Nicolas, M. E., Gleadow, R. M. and Dalling, M. J.: Effects of drought and high temperature on grain growth in wheat., Aust. J. Plant Physiol., 11(6), 553–566, doi:10.1071/PP9840553, 1984.

Qaseem, M. F., Qureshi, R. and Shaheen, H.: Effects of Pre-Anthesis Drought, Heat and Their Combination on the Growth, Yield and Physiology of diverse Wheat (Triticum aestivum L.) Genotypes Varying in Sensitivity to Heat and drought stress, Sci. Rep., 9(1), 1–12, doi:10.1038/s41598-019-43477-z, 2019.

Talukder, A. S. M. H. M., McDonald, G. K. and Gill, G. S.: Effect of short-term heat stress prior to flowering and early grain set on the grain yield of wheat, F. Crop. Res., 160, 54–63, doi:10.1016/j.fcr.2014.01.013, 2014.

Some editing would be recommended regarding:

• The renaming of "cluster" to "region" (as "cluster" does not have a strong geographical association)

• The labeling of heat percentiles in Figure 7.

Reply (4): Thank you. We have modified "cluster" to "region" in the revised version and relabelled the heat percentiles in Figure 7 properly, as also suggested by the reviews.

Risk of crop failure due to compound dry and hot extremes estimated with nested copulas

Andreia Filipa Silva Ribeiro[1,2], Ana Russo[2], Célia Marina Gouveia[2,3], Patrícia Páscoa[2,3,4], and Jakob Zscheischler[1,5]

[1]Climate and Environmental Physics, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland
[2]Instituto Dom Luiz (IDL), Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749–016, Lisboa, Portugal
[3]Instituto Português do Mar e da Atmosfera, Lisboa, Portugal
[4]Environmental Physics Laboratory (EPhysLab), University of Vigo, Ourense, Spain
[5]Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland

Correspondence: Andreia F. S. Ribeiro (afsribeiro@fc.ul.pt)

Abstract. ~~Drought and heat events stress agricultural systems and may threaten food security.~~ The interaction between co-occurring drought and hot conditions is often particularly damaging to crop's health and may cause crop failure. ~~In this context, traditional univariate analyses may not be adequate for reliable risk assessment of crop failure associated with compound hazards.~~ Climate change exacerbates such risks due to an increase in the intensity and frequency of dry and hot events in many land regions. ~~Here~~ Hence, here we model the trivariate dependence between spring maximum temperature, spring precipitation and wheat and barley yields, respectively, over two province ~~clusters~~ regions in Spain with nested copulas. Based on the full trivariate joint distribution, we (i) estimate the impact of compound hot and dry conditions on wheat and barley loss and (ii) estimate the additional impact due to compound hazards compared to individual hazards. We find that crop loss increases when drought- or heat-stress aggravates to compound dry and hot conditions and that an increase in the severity of compound conditions leads to larger damages. For instance, compared to moderate drought only, moderate compound dry and hot conditions increase the likelihood of crop loss by 8 to 11% while when starting with moderate heat, the increase is between 19 to 29% (depending on the cereal and region). This findings suggest that the likelihood of crop loss is driven primarily by drought stress than by heat stress, suggesting that drought plays the dominant role in the compound event, that is, drought stress does not require to be so extreme as heat stress to cause a similar damage. Furthermore, when compound dry and hot conditions aggravate from moderate to severe or extreme stress, crop loss probabilities increase 5 to 6% and 6 to 8% respectively (depending on the cereal and region). Our results highlight the additional value of a trivariate approach for the estimating the compounding effects of dry and hot extremes on ~~of~~ crop failure risk. Therefore, this approach can effectively contribute to design management options and guide the decision-making process in agricultural practices.

1 Introduction

The assessment of the adverse social, economic and environmental impacts associated with a combination of multiple climate hazards have recently become a focus of high interest (Leonard et al., 2014; Zscheischler et al., 2020). Such compound events often lead to larger impacts compared to when hazards occur separately (Zscheischler et al., 2018). For instance, compound dry and hot conditions reduce carbon uptake more strongly compared to the sum of the individual hazards (Zscheischler et al., 2014). Dry and hot conditions often co-occur. For instance in Europe, the extreme 2003, 2010 and 2018 heatwaves were accompanied by strong soil moisture deficits (Bastos et al., 2014; Schumacher et al., 2019; Buras et al., 2020). In 2010, the compound event was particularly strong in Russia (Schumacher et al., 2019), while in 2003 the extreme drought and heatwave affected mostly central Europe, extending to west Mediterranean countries like Portugal and Spain (Garcia-Herrera et al., 2010), with critical consequences in several sectors. In 2010, widespread crop yield declines and failures occurred over the major grain producing regions of Russia, northeastern Ukraine, and northwestern Kazakhstan (Loboda et al., 2017). Previously, the shortages in crop yields in 2003 have also caused major financial losses in the agricultural sector, and when compared to the previous year, the cereal productions in European Union (EU) have decreased 23 million tonnes (COPA-COGECA, 2003). The decline in the harvests was both in quantity and quality, as was the case of winter cereals whose maturation was accelerated due to compound extreme dry and hot conditions, forming grains with insufficient water content (COPA-COGECA, 2003). The 2018 event strongly impacted pastures and arable land north of the Alps (Buras et al., 2020). As the occurrence of climate extremes such as heatwaves, droughts and compound dry and hot events is expected to increase in intensity and frequency in many land regions due to climate change (IPCC, 2012; Zscheischler and Seneviratne, 2017), associated adverse impacts such as widespread harvest failures threatening global cereals supplies may also increase.

Among the panoply of multivariate approaches applied to assess the impacts of multiple climate hazards, the use of copulas has become quite popular in studies focused on analysing the social, environmental and economic risks associated with adverse climate conditions (Bokusheva et al., 2016; Gaupp et al., 2019; Madadgar et al., 2017; Ribeiro et al., 2019b, a; Zscheischler et al., 2017). With copulas nonlinear dependency structure can be modelled, which offers more flexibility and possibly a more adequate fit for different dependence types in the extremes. (Durante and Sempi, 2015; Nelsen, 2006; Salvadori and De Michele, 2007; Salvadori et al., 2016). Among all types of copulas described in the literature, the popularity of the class of elliptical copulas comes from the fact that they derive from well known distributions associated to the widely used Pearson's correlation, but the elliptical dependence is only able to capture radial symmetry and the respective mathematical expressions do not have a closed form. One of the copula classes that overcomes this drawback is the Archimedean, which have a simpler mathematical form and can capture different kinds of tail dependence and radial symmetry or asymmetry.

Archimedean copulas (AC) are exchangeable, which means that the copula is the same if we permute the respective margins. For the bivariate case this may not be a limitation, but as the number of dimensions increase, it is unlikely that exchanging across the involved variables allows for the 'true' dependence structure to be well-defined. To avoid exchangeability, nested Archimedean copulas (NAC) have been proposed (Okhrin and Ristig, 2014), also referred to as hierarchical Archimedean copulas (HAC), obtained by nesting lower dimensional Archimedean copulas into each other and/or with marginal distributions. Okhrin and Ristig (2014) introduced NACs where all copulas belong to the same family with a nesting condition that requires decreasing dependence strength from the highest to the lowest hierarchical level. Here we make use of this NAC approach, taking advantage of the balance between flexibility (modelling different types of dependence structures) and usability in higher dimensions (limiting the number of parameters).

The present work aims to identify how ~~risks associated with~~ compound dry and hot conditions affect wheat and barley yields over two ~~clusters~~ regions of provinces in Spain based on the trivariate dependence between precipitation, maximum temperature and yields using a NAC approach. In particular, we are interested in quantifying the additional risk associated with compound dry and hot conditions compared to only dry or only hot conditions. Wheat and barley are chosen as they are two of the major rainfed crops in the Iberian Peninsula (Peña-Gallardo et al., 2019; Vicente-Serrano et al., 2006). Moreover, we here build on prior work which has estimated wheat and barley losses in the same area, but related to a single hazard, namely droughts (Ribeiro et al., 2019a, b).

Using NACs, we estimate the conditional probabilities of crop loss under different severity levels of dry and hot conditions based on the full trivariate joint distribution. We focus on annual wheat and barley yield data at the sub-national scale, thus overcoming drawbacks related to assessing climate related crop risks at the national scale. Based on the proposed approach we (i) characterize the dependence structures between the dry and hot conditions and the crop yields; (ii) estimate the conditional probability of crop loss under different compound dry and hot severity levels; and (iii) evaluate how much the compound dry and hot conditions increase the risk of crop failure in comparison to the individual hazards.

2 Data and methods

2.1 Crop yield data

Wheat and barley yields were obtained for 9 provinces in Spain from the Spanish Agriculture, Fishing and Environment Ministry (available at https://www.mapa.gob.es/es/estadistica/temas/publicaciones/anuario-de-estadistica/, last access on 9 November 2019). ~~The data were assembled in two clusters of provinces~~ Those 9 provinces were aggregated to two distinct regions (Figure 1), which are dominated by rainfed agricultural practices ~~considering~~ following the non-irrigated arable land classification from CORINE Land Cover dataset based on an earlier ~~clustering (Ribeiro et al., 2019c, b).~~ regionalization (Ribeiro et al., 2019c, b) . The provincial regionalization consisted in the application of three main criteria: first the provinces with land use dominated by agricultural practices were identified (Figure 1), and from those provinces, the ones dominated by non-irrigated practices and contiguous in space were selected for analysis (Figure 1 - bold black contours). Figure 1 shows the Iberian provinces with $< 50\%$ agricultural pixels colored in white, the provinces with $> 50\%$ agricultural pixels coloured with the respective agricultural CORINE classes and the selected two ~~clusters~~ regions of contiguous provinces dominated by rainfed agriculture delineated in bold black contours. This aggregation of provinces allowed for the identification of two major breadbaskets where rainfed systems are the predominant crops among the provinces with higher percentage of agricultural land use in the IP.

Crop yields were obtained as the ratio between production and harvested area during the period of 1986–2016. We computed crop yield anomalies by removing longer term trends based on locally estimated scatterplot smoothing (LOESS, a method for local regression) to account for yield increases due to technological development (Ben-Ari et al., 2016). We pooled crop yields from the provinces over each ~~cluster~~region, resulting in samples sizes $N_1 = 155$ for ~~Cluster~~ Region 1 (~~30~~ 31 years of annual data over five provinces) and $N_2 = 124$ for ~~Cluster~~ Region 2 (~~30~~ 31 years of annual data over four provinces). Pooling time series greatly expands the sample size allowing greater robustness in three-dimensional statistical analysis that otherwise would be compromised. This type of assessment is a compromise between the use of a sub-national resolution of crop data and the sample size to evaluate the number of cases of simultaneous occurrence of dry and hot conditions.

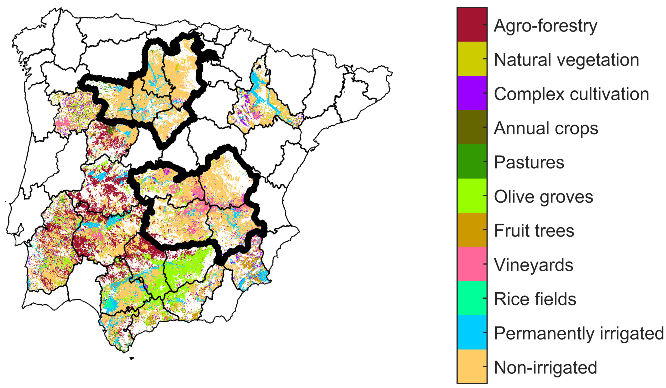


Figure 1. Iberian provinces dominated by agricultural land use ($> 50\%$ agricultural pixels belonging to all agricultural CORINE classes, see legend) according to the CORINE Land Cover dataset and respective categories. The contiguous provinces dominated by rainfed practices ($> 50\%$ non-irrigated pixels in yellow) are delineated in bold black contours and grouped in two ~~clusters~~regions. Northern region (~~Cluster~~ Region 1) provinces: Burgos, Palencia, Segovia, Valladolid, and Zamora. Southern region (~~Cluster~~ Region 2) provinces: Albacete, Ciudad Real, Cuenca, and Toledo.

2.2 Weather data

The vegetative cycle of the winter crops in Spain is mainly driven by precipitation and temperature: sowing occurs around autumn ~~,~~ (from September through November, SON), followed by the vegetative phase in winter ~~, reproductive phase in spring (when vegetation is photo-synthetically more active)~~ (from December through January, DJF), reproductive phase (more photosynthetically active phase) in spring (from March through May, MAM) and crop harvest occurs in the early summer (around June). Therefore, monthly accumulated precipitation (P) and monthly maximum temperature (Tmax) were extracted from the Climate Research Unit (CRU) TS4.01 dataset (Harris et al., 2014) spanning the same time period.~~We used~~ Given the importance of assessing crop's water and temperature requirements at different moments of the vegetative cycle we conducted a correlation analysis between the annual yields and the 3-monthly means of ~~Tmax and~~ P and 3-monthly means of Tmax during the whole growing season, as shown in Fig. 2. The identification of the moment of the vegetative cycle of the highest crop's water and temperature requirements was assessed based on the strongest statistically significant correlation value (denoted by filled circles in Fig. 2). Figure 2 suggests that the greatest influence of P and Tmax in crop yields is observed during spring (MAM in both regions and cereals) corresponding to the reproductive phase of plant development, when vegetation is photosynthetically more active. Therefore, for the remaining analysis we focus on 3-monthly means of Tmax and 3-monthly means of P during spring (P_{MAM} and $Tmax_{MAM}$, respectively), which ~~was identified as the most sensitive time period for crop~~

~~yield based on correlation analysis (Figure 2)~~has also been identified in previous studies as a growth stage sensitive to the effects of water content and high temperatures (Ferrise et al., 2011; Del Moral et al., 2003; Iglesias and Quiroga, 2007; Ribeiro et al., 2019c) . This selection of climate variables allows to maximize the dependence between climate conditions and yields as also shown by previous work based on the same data (Ribeiro et al., 2019c).

We considered three severity levels of dry and/or hot conditions: Moderate (+), Severe (++) and Extreme (+++) based on percentile thresholds as shown in Table 1. Besides these three severity levels, we further considered all combinations of 10 categories of severity levels of dry and hot conditions exceeding the 50th to 5th and 50th to 95th percentiles for P_{MAM} and $Tmax_{MAM}$, respectively. We further considered the 20th percentile of the crop anomaly time-series as lower exceedance threshold for crop failure (Ben-Ari et al., 2016; Ribeiro et al., 2019a, b).

Table 1. Categories of severity levels of dry and hot conditions based on P_{MAM} and $Tmax_{MAM}$ percentiles.

	Moderate (+)	Severe (++)	Extreme (+++)
dry	$P_{MAM} \leq 20^{th}$ percentile	$P_{MAM} \leq 10^{th}$ percentile	$P_{MAM} \leq 5^{th}$ percentile
hot	$Tmax_{MAM} \geq 80^{th}$ percentile	$Tmax_{MAM} \geq 90^{th}$ percentile	$Tmax_{MAM} \geq 95^{th}$ percentile

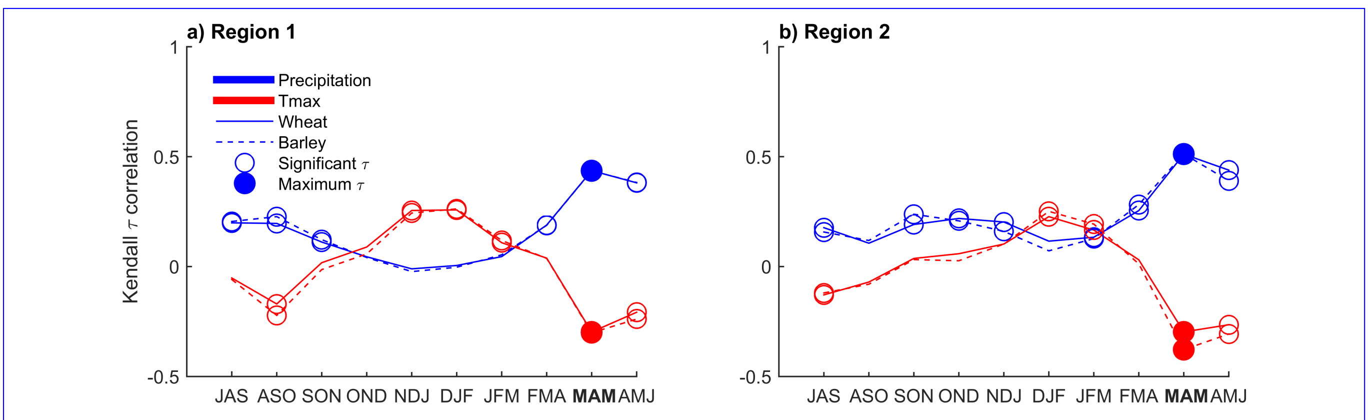


Figure 2. Kendall correlation τ between three-monthly means of maximum temperature (Tmax, red) and precipitation (blue) with wheat (filled lines) and barley (dashed lines) yield anomalies, respectively. Correlations were computed during the crop growing period (September to June) over 1986-2016 for ~~Cluster~~ Region 1 (a) and 2 (b) (Figure 1). The ~~months~~ letters on the x-axis denote the ~~end-month of the~~ three-month averaging ~~period~~periods. Circles indicate statistically significant correlations at $\alpha = 0.05$. The strongest correlation (positive or negative) is denoted by filled circles (P_{MAM} and $Tmax_{MAM}$).

2.3 Modelling trivariate distributions with nested Archimedean copulas

We model the trivariate relationship between temperature, precipitation and crop yields with nested copulas. Consider a vector of crop yield annual anomalies Y and the climate variables $X_1 = P_{MAM}$ and $X_2 = Tmax_{MAM}$ with marginal cumulative distribution functions (CDF) F_Y, F_{X_1} and F_{X_2}, respectively. We aim to estimate and compare three conditional cumulative distribution functions (CDFs) with the scalars x_1^* and x_2^* corresponding to the dry and hot thresholds, respectively:

$$F_{Y|X_1}(Y|X_1 = x_1^*) = P(Y \leq y|X_1 \leq x_1^*) \tag{1}$$

$$F_{Y|X_2}(Y|X_2 = x_2^*) = P(Y \leq y|X_2 \geq x_2^*) \tag{2}$$

$$F_{Y|X_1,X_2}(Y|X_1 = x_1^*, X_2 = x_2^*) = P(Y \leq y|X_1 \leq x_1^*, X_2 \geq x_2^*) \tag{3}$$

With the above equations we can estimate the ~~agricultural~~ crop yield impacts under dry conditions $F_{Y|X_1}$ (Equation 1), under hot conditions $F_{Y|X_2}$ (Equation 2) and under compound dry and hot condition $F_{Y|X_1,X_2}$ (Equation 3), respectively. In other words, if the compound dry and hot conditions cause more damage than the individual hazards, it is expected that $F_{Y|X_1,X_2}$ suggests higher probabilities of crop loss (i.e., $y = y^*$ for a low y^*) than $F_{Y|X_1}$ or $F_{Y|X_2}$. Furthermore, we can study the relative role of $\mathrm{P_{MAM}}$ and $\mathrm{Tmax_{MAM}}$ ~~in~~ for crop loss with Equations 1 and 2.

To compare the additional impact of compound dry and hot conditions with the impacts caused by the individual hazards, Equations 1~~, 2 and~~ - 3 are used to estimate

$$\text{Relative change from drought-stress} = \frac{F_{Y|X_1=x_1^*,X_2=x_2^*}(0.2) - F_{Y|X_1=x_1^*}(0.2)}{F_{Y|X_1=x_1^*}(0.2)} \tag{4}$$

$$\text{Relative change from heat-stress} = \frac{F_{Y|X_1=x_1^*,X_2=x_2^*}(0.2) - F_{Y|X_2=x_2^*}(0.2)}{F_{Y|X_2=x_2^*}(0.2)}, \tag{5}$$

where 0.2 is the threshold of crop loss (y^*) corresponding to the 20$^{\text{th}}$ percentile of the crop ~~yields~~yield anomalies. These changes can be estimated for different severity levels of dry (x_1^*) and hot (x_2^*) conditions.

Following the theorem of Sklar (1959) we can decompose a multivariate probability distribution into its marginals and a copula C which describes the dependence structure between the margins. To estimate the multivariate distribution $P(Y, X_2, X_3)$, the respective copula C is fitted, which is then a joint CDF whose marginal distributions are uniform in the interval $[0, 1]$ (Durante and Sempi, 2015; Nelsen, 2006; Salvadori and De Michele, 2007). Transforming the margins to uniform variables through their CDFs, that is, $u_1 = F_Y$, $u_2 = F_{X_1}$ and $u_3 = F_{X_2}$, the trivariate CDF can be written as (Sklar, 1959):

$$F(u_1, u_2, u_3) = C(u_1, u_2, u_3). \tag{6}$$

Within the copula families, AC are extensively used due to their flexibility and applicability to a variety of tail dependence structures, as well as their analytical tractability. AC can be written in terms of the respective generator function φ, which belongs to a parametric family φ_θ dependent on the parameter θ, e.g. for the three-dimensional case:

$$C(u_1, u_2, u_3; \theta) = \varphi_\theta(\varphi_\theta^{-1}(u_1) + \varphi_\theta^{-1}(u_2) + \varphi_\theta^{-1}(u_3)) \tag{7}$$

Due to the symmetry of bivariate AC, the above trivariate form can be expressed in terms of NAC~~or HAC~~, where two of the margins are first coupled by their bivariate copula ~~,~~and then coupled with the third margin, via the same generator on each level but different parameters θ_{12} and θ_1, respectively, e.g.:

$$C(u_1, u_2, u_3; \theta_{12}; \theta_1) = C_1(C_{12}(u_1, u_2; \theta_{12}), u_3; \theta_1) \tag{8}$$

Equation 8 can also be expressed in terms of the other possible pair copulas $C_{13}(u_1, u_3; \theta_{13})$ and $C_{23}(u_2, u_3; \theta_{23})$ that are coupled with u_2 and u_1 by C_2 and C_3, with expressions $C_2(C_{13}(u_1, u_3; \theta_{13}), u_2; \theta_2)$ and $C_3(C_{23}(u_2, u_3; \theta_{23}), u_1; \theta_3)$, respectively. Like Equation 8, among each structure of NAC the same generator is required for each level but with potentially different parameters. Hence, both the optimal structure and respective parameters must be determined.

Most structures of ~~NAC~~ NACs require decreasing parameters from the inner to the outer hierarchical level to attain a properly fitted copula. As for most ACs, the larger the parameter the stronger the dependence, this means that most structures of NAC require that the marginal copulas in the inner level should correspond to the pair with the strongest dependence, i.e., satisfying $\theta_{12} \geq \theta_1$ in the case of Equation 8. This requirement applies to NAC with generators from the same family, providing a flexible estimation of the NAC, which allows for specifying the full distribution with at most $d-1$ parameters, where d is the number of copula dimensions or marginal distributions (Okhrin and Ristig, 2014).

In our study we focus on a total of four Archimedean families that capture different kinds of joint dependence structures: Clayton, Gumbel, Frank and Joe. The Clayton, Gumbel and Joe copulas describe an asymmetrical tail behaviour, while the Frank copula, in a similar way to the Gaussian copula, captures joint symmetric dependence. While Gumbel and Joe copulas can represent upper tail dependence, Clayton copulas can represent lower tail dependence. The estimation of the copula parameters is based on maximum likelihood based on the R package ~~HAC~~ *HAC* (Okhrin and Ristig, 2014) .

The main steps of the trivariate approach used in this study can be summarized as follows (Okhrin and Ristig, 2014). First, the marginal distributions u_1, u_2 and u_3 are estimated non-parametrically by simple ranking, using the empirical distribution functions of the data through the *pobs* function in the R package *copula* (Ivan Kojadinovic and Jun Yan, 2010), a common approach for copula modelling. Afterwards, the fit of bivariate copula models is performed to every pair of variables to estimate $C_{\theta_{12}}$, $C_{\theta_{13}}$ and $C_{\theta_{23}}$. For each pair, the copula selection is performed based on the Akaike's information criterion (AIC) and checking the goodness-of-fit by comparing the empirical copula based on the Cramer-von Mises distance (Sn). The bivariate copula with the strongest dependence, with the lowest AIC and the lowest Sn, is selected to define the structure of the NAC. Afterwards, the marginal distribution that is not part of the selected bivariate copula is joined and the parameter of the upper level copula of the same family is estimated (Equation 8). As a final step, the estimated NAC with two parameters is compared with the same Archimedean family with one parameter (Equation 7) in terms of the AIC, which penalizes the number of estimated parameters.

2.4 Diagnostics and uncertainties in the estimation procedure

The visual diagnostics of the quality of the selected models are performed analogously to a QQ-plot by comparing the empirical estimate of the Kendall function (cumulative distribution of the copula) with the theoretical estimate of the Kendall function based on the selected parametric trivariate copulas (Okhrin and Ristig, 2014).

Best estimates of all conditional probabilities (i.e., Equations 1-5) are estimated by drawing $N = 100{,}000$ samples from the fitted trivariate copula. ~~Due to the~~ Using the same single-parameter generator function on each level of NAC (but with a potentially different value of θ) 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 $Tmax_{MAM}$ and both crop yields and P_{MAM}, we inverted the margins of $Tmax_{MAM}$ 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).

Uncertainties of the statistical modelling are estimated by repeated sampling (10,000 times) of the fitted model with sample sizes equal to the number of observations (i.e., N_1 in the case of ~~Cluster~~ Region 1 and N_2 in the case of ~~Cluster~~ Region 2). From these samples, 95% confidence intervals of Kendall's rank correlation are estimated and compared with the observed pairs (u_1, u_2), (u_1, u_3) and (u_2, u_3). This validation step intends to verify if the generated pairs of copula-based samples preserve the level of dependence found in the observations. Furthermore, this approach is used to estimate uncertainties related to the conditional probabilities (Equations 1-5).

3 Results

In both cereals and both ~~clusters~~ regions the most dependent pair of variables corresponds to crop yields and P_{MAM}, hence the pair of variables u_1, u_2 defines the optimal NAC structure (Figure 3). Results for all possible variable pairs and the respective bivariate copulas are shown in Table A.1.

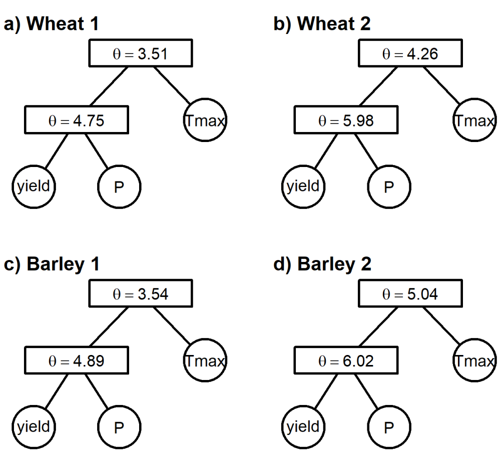


Figure 3. Structure and respective parameters of the selected nested Frank models $C_1(C_{12}(u_1, u_2; \theta_{12}), u_3; \theta_1)$ to model the trivariate joint distributions between crop yields, P_{MAM} and $Tmax_{MAM}$. (a) Wheat in ~~Cluster~~ Region 1. (b) Wheat in ~~Cluster~~ Region 2. (c) Barley in ~~Cluster~~ Region 1. (d) Barley in ~~Cluster~~ Region 2.

Once the bivariate copula $C_{12}(u_1, u_2)$ of yields and P_{MAM} is known, the NAC models are constructed (Table 2). The Frank copula provides the best fit of $C_{12}(u_1, u_2)$ (Table A.1) for both cereals and both ~~clusters~~ regions and thus the parameters of the trivariate nested copulas are all from the Frank family. Nevertheless, despite Frank being the best family to characterize the nested copulas, we also constructed NAC models with Gumbel, Clayton and Joe copulas for comparison, as well as trivariate

Archimedean copulas with one parameter where we selected the best structure between one-parameter and two-parameter AC copulas via the AIC (Table 2). In all but one case the NAC models with Frank copulas is the best model. The only exception is barley in ~~Cluster~~ Region 2 whose AIC of $C_\theta(u_1,u_2,u_3)$ is slightly lower than the AIC of $C_{\theta_1}(u_3,C_{\theta_{12}}(u_1,u_2))$ (Table 2). This feature may suggest that a structure favouring the dependence between yield and precipitation (u_1,u_2) may not be as relevant as in the other regions and yields due to a less dominant role of drought individually in this case. Nevertheless, in terms of Cramer-von Mises distance (Sn) the nested copula is ~~the~~ closer to the empirical trivariate copula. For this reason, we modelled the trivariate joint distribution based on nested Frank copulas for all cases. For all fitted models, the empirical cumulative distribution corresponds well to the theoretical cumulative distributions (Figure 4).

Bivariate dependencies as measured by Kendall's τ 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 weakest for both cereals (Table A.1), and likely for this reason it is the pair in Figure 5 and Figure A.1 with observational τ closest to the lower bound of the 95% confidence intervals (Figure 5f,h 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 $\tau_{u_1,u_2} > \tau_{u_2,u_3} > \tau_{u_1,u_3}$.

Table 2. Trivariate Archimedean copulas (AC) parameters (θ) with nested structure with two-parameters $C_1(C_{12}(u_1,u_2;\theta_{12}),u_3;\theta_1)$ and with one-parameter $C(u_1,u_2,u_3;\theta)$ and respective Akaike's Information Criteria (AIC) and Cramer-von Mises distance (Sn). Fit based on maximum pseudo-likelihood (Gumbel (G), Clayton (C), Frank (F) and Joe (J) copulas). Smaller values of AIC and Sn indicate the selected copula for each cereal and ~~cluster~~ region (bold).

		Region 1					Region 2				
			G	C	F	J		G	C	F	J
Wheat	$C(u_1,u_2,u_3;\theta)$	θ	1.41	0.66	3.22	1.53	θ	1.53	0.75	3.88	1.72
		AIC	-74.16	-79.89	-99.02	-49.16	AIC	-89.12	-74.67	-106.14	-69
		Sn	0.15	0.21	0.07	0.31	Sn	0.14	0.31	0.07	0.27
	$C_1(C_{12}(u_1,u_2;\theta_{12}),u_3;\theta_1)$	θ_{12}	1.37	0.9	3.51	1.41	θ_{12}	1.57	0.91	4.26	1.76
		θ_1	1.59	0.93	4.75	1.73	θ_1	1.88	1.37	5.98	2.11
		AIC	-79.69	-71.27	**-102.84**	-54.29	AIC	-99.7	-79.76	**-112.93**	-78.49
		Sn	0.12	0.11	**0.03**	0.3	Sn	0.08	0.18	**0.03**	0.19
Barley	$C(u_1,u_2,u_3;\theta)$	θ	1.43	0.66	3.25	1.57	θ	1.58	0.81	4.12	1.8
		AIC	-80.8	-78.91	-101.84	-57.51	AIC	-105.59	-85.87	-118.55	-83.54
		Sn	0.12	0.21	0.07	0.26	Sn	0.16	0.36	0.08	0.3
	$C_1(C_{12}(u_1,u_2;\theta_{12}),u_3;\theta_1)$	θ_{12}	1.38	0.87	3.54	1.43	θ_{12}	1.72	1.05	5.04	1.99
		θ_1	1.7	0.92	4.89	1.92	θ_1	1.94	1.41	6.02	2.21
		AIC	-95.8	-72.07	**-107.17**	-73.98	AIC	-112.52	-86.85	-116.31	-90.86
		Sn	0.09	0.12	**0.04**	0.22	Sn	0.08	0.21	**0.03**	0.19

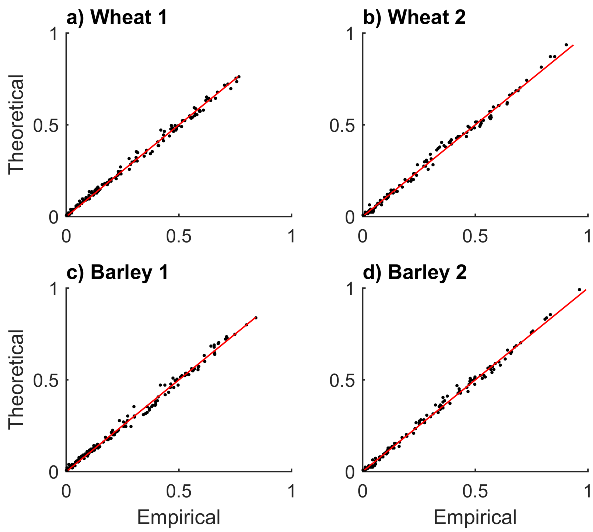


Figure 4. Empirical versus theoretical probability distributions based on the nested Frank copula models. (a) Wheat in ~~Cluster~~ Region 1. (b) Wheat in ~~Cluster~~ Region 2. (c) Barley in ~~Cluster~~ Region 1. (d) Barley in ~~Cluster~~ Region 2.

The cumulative conditional probabilities of yield under moderate (+), severe (++) and extreme (+++) compound dry and hot conditions demonstrate that the probability of crop loss increases with the severity of compound dry and hot conditions for both ~~clusters~~ regions and both cereals (Figure 6a-d). Moreover, the likelihood of crop loss is higher in ~~Cluster~~ Region 2 for both cereals, particularly in the case of barley. Under extreme dry and hot conditions (+++dry+++hot, purple), the likelihood of crop loss is 68% and 71% for wheat and barley, respectively, in ~~Cluster~~ Region 2, in contrast to 62% and 63% in ~~Cluster~~ Region 1 (Figure 6e, purple bars). In addition, the differences in crop loss are higher between moderate (+dry+hot) and severe (++dry++hot) conditions compared to the differences between severe and extreme (+++dry+++hot) conditions. More precisely, when the compound dry and hot conditions aggravate from moderate to severe stress, crop loss increases 5 to 6% and when the compound dry and hot conditions aggravate from moderate to extreme stress, crop loss increases 6 to 8% (depending on the cereal and region). For comparison, conditional cumulative probability distributions for single stressors compared with the compound stressors are shown in Figure A.2 for all three severity levels.

While Figure 6 illustrates the same severity levels for the different hazards, Figure 7 illustrates crop loss for a range of different combinations of severity levels of dry and hot conditions (e.g., extreme dry conditions combined with moderate, severe and extreme hot conditions, and vice-versa) starting from the 50^{th} percentile of P_{MAM} and $Tmax_{MAM}$. When P_{MAM}/$Tmax_{MAM}$ are below/above the median, the probability of crop loss is always higher than 40%. Similarly to Figure 6, the increase of crop loss with the severity of drought- and heat-stress is evident (Figure 7). The higher likelihood of crop loss in ~~Cluster~~ Region 2, particularly for barley, is also consistent with Figure 6. Moreover, the results indicate that droughts are typically associated with higher probabilities of crop loss than heatwaves at the same severity level. This finding suggests that drought stress causes more damage to crop yields than heat stress, even for lower values of stress.

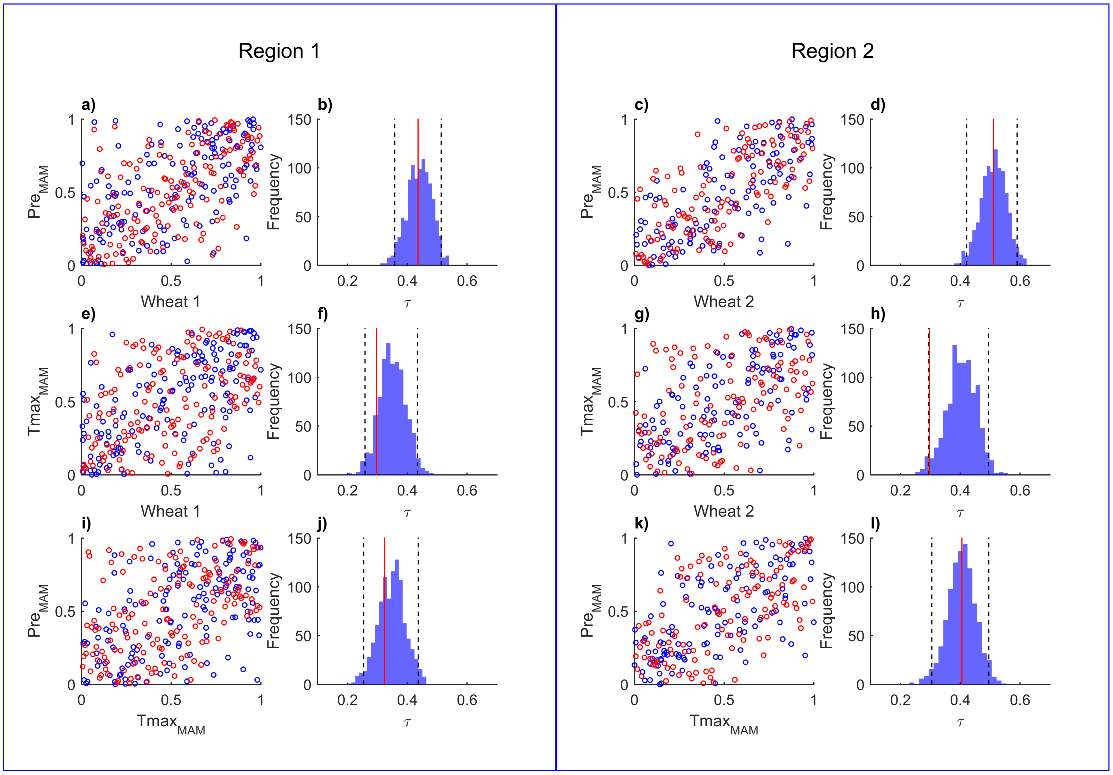


Figure 5. Scatterplots of copula-based samples (blue) compared with ranked observations (red) of crop anomalies with climate variables (P_{MAM} and $Tmax_{MAM}$ ~~)~~ (a~~)~~, c~~)~~, e ~~)~~ and g) ~~)~~ and P_{MAM} against $Tmax_{MAM}$ (i ~~)~~ and k)~~)~~, for both ~~clusters~~regions. The histograms (b~~)~~, d~~)~~, f~~)~~, h~~)~~, j~~)~~, l) ~~)~~ correspond to the Kendall rank correlation of each pair based on 10,000 simulations with the same sample size of the observational sample. The 95% confidence intervals are shown with dashed lines. The red lines indicate the Kendall rank correlation of the observations.

In all cases, the additional effect of compound dry and hot conditions is larger when starting from only hot conditions, compared to when starting from only dry conditions (Figure 8 for moderate stress, Figure A.3a and b for severe and extreme stress). The estimates are based on Equations 4 and 5. Depending on the cereal and region, the difference from drought stress to compound conditions may vary from 8% (barley in ~~Cluster~~ Region 1) to 11% (barley in ~~Cluster~~ Region 2). In contrast, the difference from heat stress to compound conditions may vary between 19% (barley in ~~Cluster~~ Region 2) to 29% (wheat in ~~Cluster~~ Region 2). Uncertainties are large for these estimates and increase with the severity of the events (Figure A.3). Consistent with Figure 7 these findings suggest that drought stress is the major driver of crop loss associated with compound drought and heat.

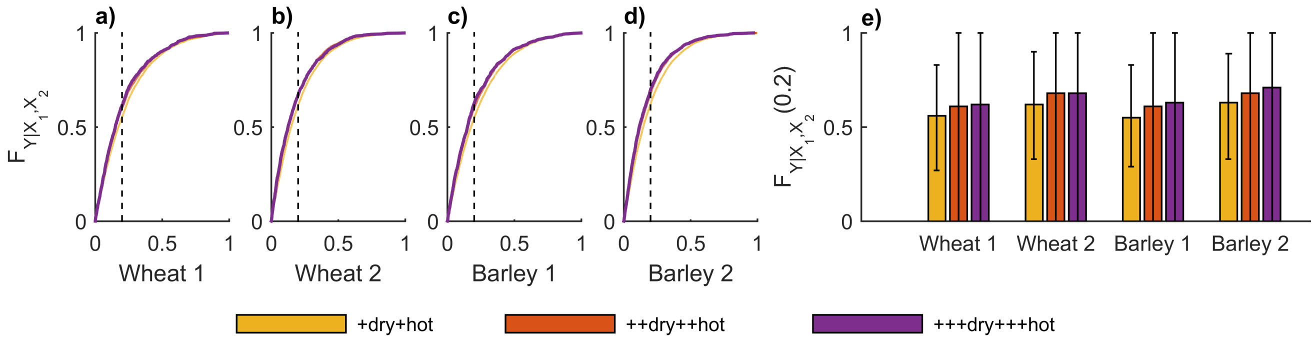


Figure 6. Conditional probability distributions of crop yield anomalies $F_{Y|X_2,X_2}$ over each ~~cluster~~ region of provinces (wheat in ~~Cluster~~ Region 1 (a), wheat in ~~Cluster~~ Region 2 (b), barley in ~~Cluster~~ Region (1) and barley in ~~Cluster~~ Region 2 (d)) under moderate (+dry+hot, yellow~~, P_{MAM} below the 20th and Tmax$_{MAM}$ above the 80th percentile~~), severe (++dry++hot, orange~~, P_{MAM} below the 10th and Tmax$_{MAM}$ above the 90th percentile~~) and extreme (+++dry+++hot, purple~~, P_{MAM} below the 5th and Tmax$_{MAM}$ above the 95th percentile~~) compound dry and hot conditions (see Table 1). (e) Conditional probabilities of non-exceeding the crop loss threshold (20th percentile – vertical dashed line in a-d)) for each severity level of compound dry and hot conditions given by $F_{Y|X_1,X_2}(0.2)$. Uncertainty ranges illustrate the 95% confidence intervals.

4 Discussion

We have modelled the trivariate relationship between Tmax$_{MAM}$ and P_{MAM} and wheat and barley yields in two regions in Spain using nested copulas. We found that the likelihood of crop loss increases with the severity of the compound dry and hot conditions and that compound drought and heat always increases the likelihood of crop loss. Moreover, our findings suggest that drought stress does not require to be so extreme as heat stress to cause the same adverse impact on crop yields. Hence drought is the more stressful driver of crop loss, when considering compound drought and heat.

Although the use of different methodologies, spatio-temporal scales and the focus on different cereals and regions makes a comparison between studies difficult, our findings are consistent with previous work. Using bivariate return periods of combined climate conditions, Zscheischler et al. (2017) have shown how linear models based directly on precipitation and temperature (and not the respective bivariate return period) may underestimate the explained variability of crop yields and that in several countries maize yields decrease with dry and hot conditions. Based on a meta-Gaussian model at the national level, Feng et al. (2019) have also shown that compound dry and hot extremes lead to larger impacts on maize yields than the individual hazards over five major maize-producing countries.

In terms of the relative contributions of drought and heat conditions, a variety of studies at the national scale have found that the response varies from country to country. Feng et al. (2019) have found that China, France and Romania expect higher chances of maize loss under dry conditions with normal temperatures (rather than under hot conditions with normal precipitation), while USA and Argentina expect higher chances of maize loss under hot conditions with normal precipitation (rather than under dry conditions with normal temperatures). In contrast, Zscheischler et al. (2017) have found that countries such as

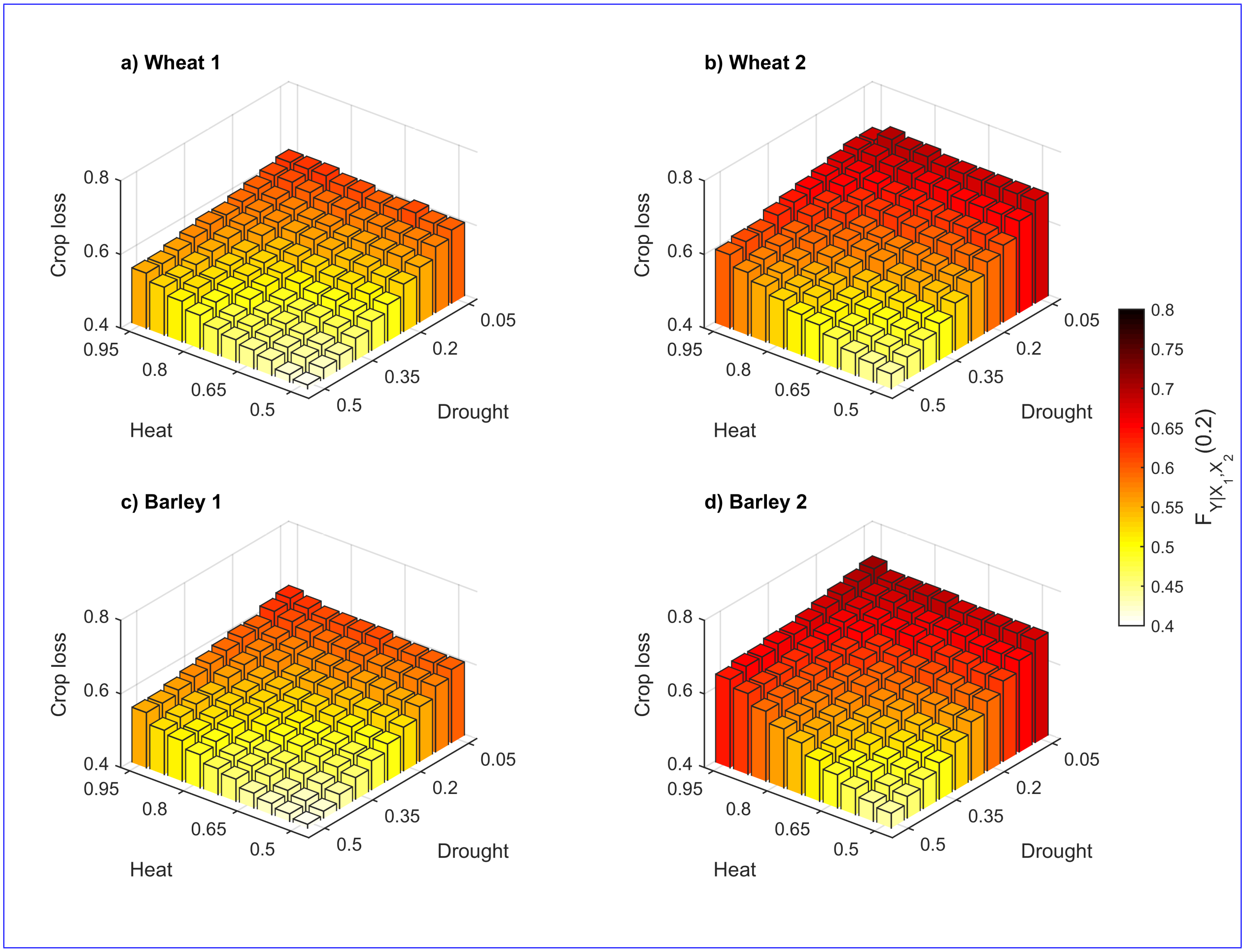


Figure 7. Conditional probability of crop loss given by $F_{Y|X_1,X_2}(0.2)$ (bar height) for both ~~clusters~~ regions and cereals (wheat in ~~Cluster~~ Region 1 (a), wheat in ~~Cluster~~ Region 2 (b), barley in ~~Cluster~~ region (1) and barley in ~~Cluster~~ Region 2 (d)) for different combinations of severity levels of dry and hot conditions. The x-axis indicates the P_{MAM} percentiles (Drought) and y-axis indicates the $Tmax_{MAM}$ percentile (Heat).

Lithuania, Luxembourg, and the UK, maize yields increase under hot and wet conditions, likely because of the importance of summer precipitation for the crop vegetative cycle and the relatively cooler climate in those countries.

Although previous studies have discussed that maximum temperature might be the best predictor variable for yield variability in most countries (Zscheischler et al., 2017), our study highlights that in Spain crop loss of wheat and barley is more sensitive to dryness than to hot conditions. This finding agrees with the rainfed practices adopted in the wheat and barley cultivation in Spain. In fact, the nesting structure of the trivariate models adopted in the present study privileges the stronger dependency between yields and precipitation, rather than between yields and temperature or between precipitation and temperature (Figure 3). Though irrigated crops typically produce higher yields, the pressure in water resources is already increasing the deficit between

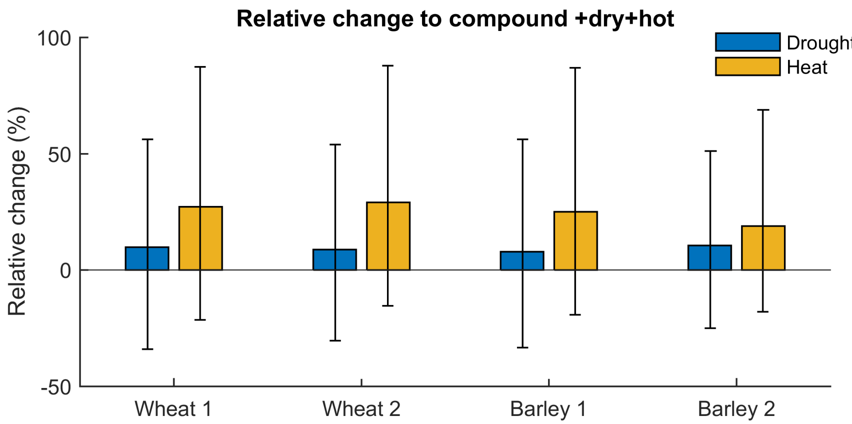


Figure 8. Difference in probability of crop loss from dry (blue) and hot (orange) to compound dry and hot conditions in wheat (left) and barley (right) for ~~Cluster~~ Region 1 and 2. Shown are the best estimates for moderate dry and hot (+dry+hot) conditions (bar height) and associated 95% confidence intervals.

water supplies and water demand in Spain (Rodríguez Díaz et al., 2007). Hence, understanding climate risks for rainfed crops is crucial to address the current water management challenges for agricultural practices in Mediterranean regions.

Higher probabilities of crop loss under drought and/or heat stress are generally expected in the southern region of Spain, in comparison to the northern region (Figures 6 and 7), in agreement with the higher temperatures and lower rainfall amounts observed in the south (Ribeiro et al., 2019a; IM and AEMET, 2011). In the case of wheat losses, this finding is in agreement with previous work which focused on drought risks for the same crops and the same region (assessed based on remote sensing and hydro-meteorological drought indicators, Ribeiro et al. (2019b). However, Ribeiro et al. (2019b) identified a higher likelihood of barley loss with drought in the northern ~~cluster~~region. This discrepancy underlines importance of addressing the interaction between compound dry and hot conditions and the associated impacts on vegetation. For instance, compound dry and hot conditions have a larger impact on the carbon uptake potential than the sum of the individual impacts (Zscheischler et al., 2014), highlighting the relevance of interactions between multiple stressors.

We found that for barley in ~~Cluster~~ Region 2, drought is the least dominant driver in comparison to the other cereals and regions. Barley in ~~Cluster~~ Region 2 shows the highest difference between drought and compound dry and hot conditions, and the lowest difference between heat stress and compound conditions (Figure 8). This suggests that among both cereals and both regions, barley in ~~Cluster~~ Region 2 is the case where the compound and possibly interacting effects of drought and heat are most relevant. Note that in this case also the ~~ECDF~~CDF's between the dry and hot and dry or hot conditions are more differentiated from each other for the severe and extreme stress (Figure A.2). This is consistent with a recent study at the province level, which recommended that crop production in Spain should focus more on wheat production given that most provinces displayed lower levels of wheat loss with drought in comparison to barley loss (Ribeiro et al., 2019a). This finding is also consistent with Figures 6 and 7.

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, 2020). For the same reason, the wheat loss in ~~Cluster~~ Region 2 when P_{MAM} is below the 5th percentile in Figure 7 slightly decreases when the threshold of $Tmax_{MAM}$ 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 ~~6 and 8~~A.3) due the rapid decrease of available samples in the corners of the three-dimensional probability distribution. ~~Moreover, following the~~ Nevertheless, the best estimates (bars in Figures 8 and A.3) show indeed that compound dry and hot extremes contribute to increase yield loss. In the general sense, the biophysiological explanation for the combination of environmental drivers leading to stronger yield reductions relates with the crop's requirements of water and thermal conditions during the key phenological stage in the analysis. The selection of the climate variables during spring corresponds to the reproductive phase of the plants and when vegetation is photosynthetically more active, and the combined effect of water and heat stress during this period is critical for the crop's health leading to yield decrease. During this stage of formation of the grains the dry and hot extremes may accelerate the maturation affecting the size, number and weight of the grains and consequently affecting the crop's harvests in quantity and quality (Balla et al., 2011; COPA-COGECA, 2003; Nicolas et al., 1984; Qaseem et al., 2019; Talukder et al., 2014).

Following the work by Okhrin and Ristig (2014), here we considered nesting copulas of the same family only, as more complex structures would be difficult to implement in general. Vine copulas might offer an alternative that is also appropriate for higher dimensions (Bevacqua et al., 2017), when considering for instance more driver variables. Nevertheless, in comparison with previous studies based on bivariate models only, we argue that the statistical modelling based on NAC is a good compromise between complexity and the trivariate dimension.

5 Conclusions

The present study assessed how compound drought and heat enhance losses of wheat and barley in two major dryland areas in Spain. We showed that nested Archimedian copulas can successfully model the trivariate joint distribution between spring maximum temperature, spring precipitation and yields to estimate conditional probabilities of crop loss under different severity levels of hot and dry conditions. The strongest dependence exists between spring precipitation and yields and is best captured by a Frank copula. Our results demonstrate that the probability of crop loss increases with the severity of compound dry and hot conditions. Furthermore, the likelihood of wheat and barley loss increases when drought or heat, respectively, aggravate to compound dry and hot conditions in both regions. Overall, the likelihood of crop loss in the southern region is larger, in particularly for barley. For both cereals and regions, the likelihood of crop loss increase more with increasing drought stress than with heat stress, suggesting that drought plays a dominant role in the compound event. Our results illustrate the additional value of using trivariate copula modelling to estimate the compounding effects of dry and hot extremes on the risk of crop failure. In operational practice, this research will allow contributing to design supporting tools and provide guidance in the decision-making process in agricultural practices to minimize crop losses related to climate hazards.

Code and data availability. The statistical analysis was performed under R software using the packages *copula* (Ivan Kojadinovic and Jun Yan, 2010) and *HAC* (Okhrin and Ristig, 2014). The R scripts are available at http://impecaf.rd.ciencias.ulisboa.pt/. The precipitation- and maximum temperature gridded values are publicly available from the Climate Research Unit (CRU) TS4.01 dataset by Harris et al. (2014). The Spanish crop yield is published by the Spanish Agriculture, Fishing and Environment Ministry in their Statistical Yearbooks, which can be consulted at https://www.mapa.gob.es/es/estadistica/temas/publicaciones/anuario-de-estadistica/ (last access on 9 November 2019). The CORINE Land Cover datasets are publicly available at https://land.copernicus.eu/pan-european/corine-land-cover.

Author contributions. A. F. S. Ribeiro analyzed the data, produced the figures and drafted the manuscript. J.Zscheischler supervised the overall work with an emphasis on the design of the statistical framework. A. Russo and C. M. Gouveia helped to supervise the work and conceived the original idea together with A.F.S.Ribeiro. P. Páscoa instructed the acquisition and analysis of the crop yield data. All the authors discussed the results, provided critical feedback and helped shape the research, analysis and contributed to the final manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. A.F.S. Ribeiro is thankful to the COST Action CA17109 for a Short Term Scientific Mission (STSM) grant ~~to develop the present work and to the Climate and Environmental Physics Department of~~ and the Young Academics Support from the Faculty of Science of the University of Bern ~~which was the host institution~~to develop the present work. This work was also partially supported by Portuguese funds through FCT (Fundação para a Ciência e a Tecnologia, Portugal) under the projects CLMALERT (ERA4CS/0005/2016) and IMPECAF (PTDC/CTA-CLI/28902/2017). A.F.S. Ribeiro also thanks FCT for grant PD/BD/114481/2016. J. Zscheischler acknowledges funding from the Swiss National Science Foundation (Ambizione grant 179876).

Appendix A: Supplementary material

Table A.1. ~~Kendalls' correlation (τ) between~~ As in Table 2 respecting the possible bivariate pairs of crop yield (u_1), ~~P_{MAM}~~ precipitation (u_2) and ~~$Tmax_{MAM}$~~ maximum temperature (u_3), and corresponding Kendall's' correlation (τ). Maximum values of τ are denoted in bold for each cereal and ~~cluster~~ region indicating the pair of variables with the strongest relationship.~~Bivariate copulas parameter (θ), Akaike Information Criteria (AIC) and Cramer-von Mises distance (Sn) to respective empirical copula, considering the possible pairs of variables. Fit based on maximum pseudo-likelihood (Gumbel (G), Clayton (C), Frank (F) and Joe (J) copulas). Smallest values of AIC and Sn indicate the selected copula for each pair (bold).~~

		Region 1		Region 2									
		τ		G	C	F	J	τ		G	C	F	J
Wheat	$C(u_1,u_2)$	**0.44**	θ	1.59	0.93	4.75	1.73	**0.51**	θ	1.88	1.37	5.98	2.11
			AIC	-51.43	-47.28	**-69.71**	-35.58		AIC	-71.04	-64.6	**-81.22**	-53.26
			Sn	0.06	0.14	**0.01**	0.17		Sn	0.04	0.11	**0.02**	0.13
	$C(u_1,u_3)$	0.30	θ	1.28	0.71	2.73	1.27	0.30	θ	1.31	0.53	2.88	1.38
			AIC	-14.3	-31.71	-28.51	-4.07		AIC	-13.83	-13.07	-23.77	-8.08
			Sn	0.09	0.04	0.03	0.18		Sn	0.08	0.1	0.03	0.13
	$C(u_2,u_3)$	0.32	θ	1.4	0.58	3.27	1.51	0.41	θ	1.66	0.77	4.28	1.98
			AIC	-28.45	-21.74	-38.13	-20.41		AIC	-52.05	-27.27	-48.85	-47.27
			Sn	0.07	0.11	0.03	0.13		Sn	0.04	0.14	0.03	0.08
Barley	$C(u_1,u_2)$	**0.44**	θ	1.7	0.92	4.89	1.92	**0.51**	θ	1.94	1.41	6.02	2.21
			AIC	-66.25	-47.07	**-72.18**	-53.18		AIC	-78.79	-68.34	**-81.99**	-61.18
			Sn	0.02	0.13	**0.02**	0.08		Sn	0.03	0.1	**0.02**	0.1
	$C(u_1,u_3)$	0.30	θ	1.3	0.69	2.77	1.31	0.38	θ	1.46	0.69	3.73	1.61
			AIC	-16.34	-30.27	-29.9	-6.11		AIC	-29.33	-22.43	-38.56	-21.43
			Sn	0.08	0.06	0.04	0.16		Sn	0.09	0.15	0.04	0.16
	$C(u_2,u_3)$	0.32	θ	1.4	0.58	3.27	1.51	0.41	θ	1.66	0.77	4.28	1.98
			AIC	-28.45	-21.74	-38.13	-20.41		AIC	-52.05	-27.27	-48.85	-47.27
			Sn	0.07	0.11	0.03	0.13		Sn	0.04	0.14	0.03	0.08

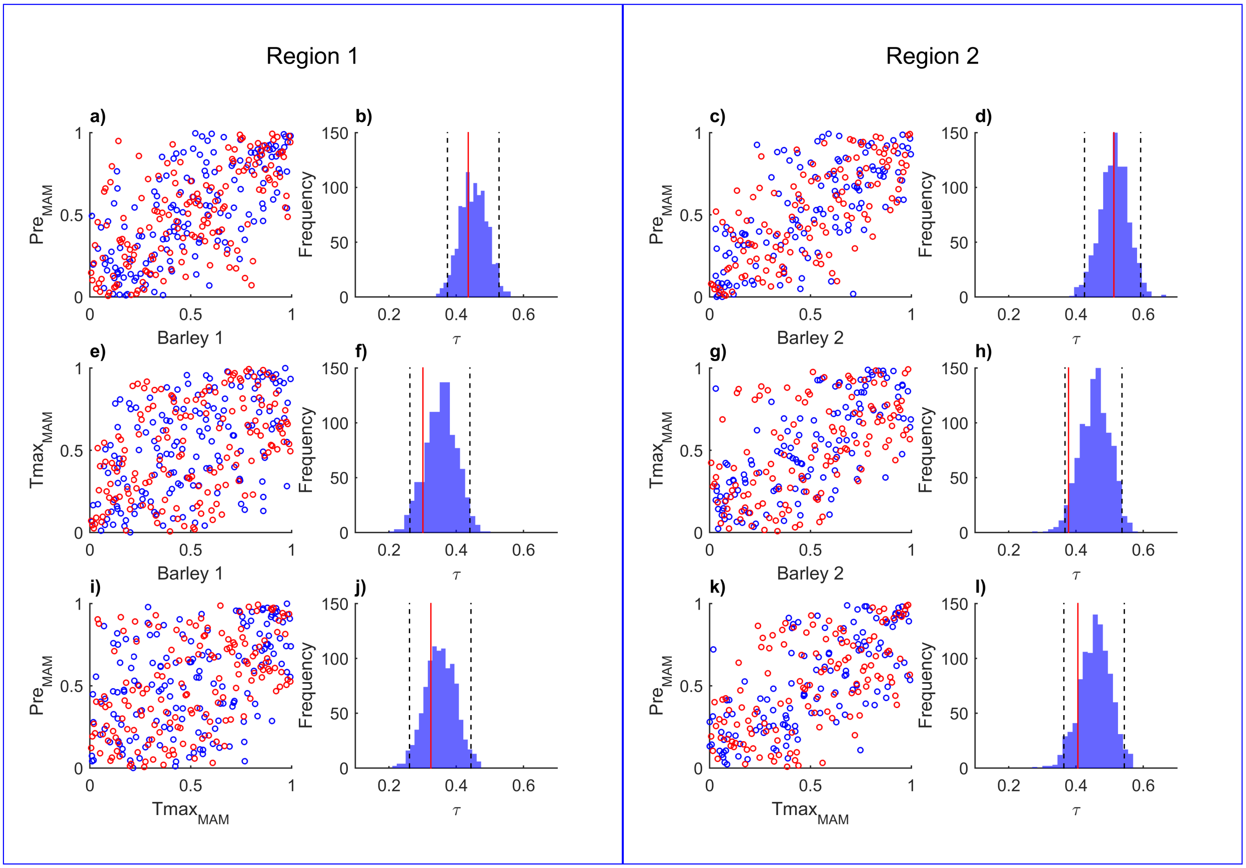


Figure A.1. Same as Figure 5 but for barley.

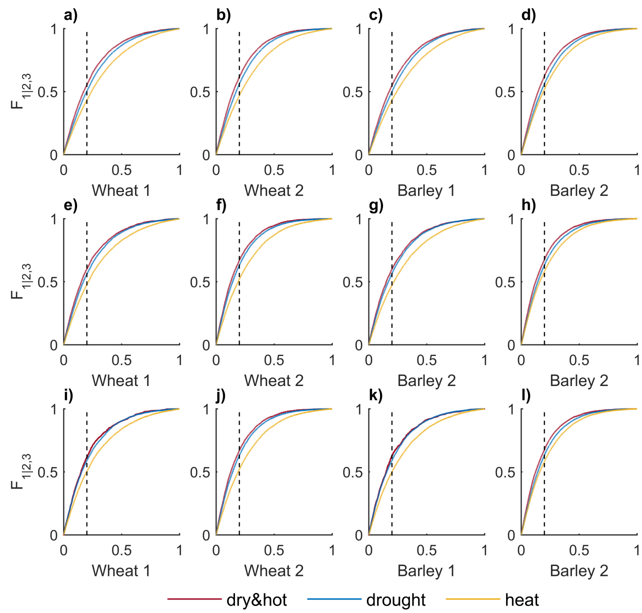


Figure A.2. Conditional probability distributions of crop yield anomalies over each ~~cluster~~ region under hot (yellow), dry (blue) or compound dry and hot (purple) under moderate (a) - d)), severe (e) - h)) and extreme conditions (i) - l)).

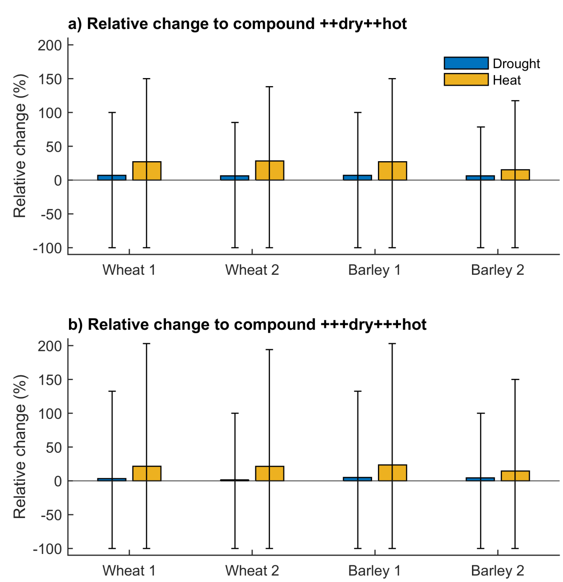


Figure A.3. Same as Figure 8 but for severe (a) and extreme (b) conditions.

References

Balla, K., Rakszegi, M., Li, Z., Bekes, F., Bencze, S., and Veisz, O.: Quality of winter wheat in relation to heat and drought shock after anthesis, Czech Journal of Food Sciences, 29, 117–128, 2011.

Bastos, A., Trigo, C., Trigo, R., and Running, S. W.: Analysing the spatio-temporal impacts of the 2003 and 2010 extreme heatwaves on plant productivity in Europe, Biogeosciences, 11, 3421, 2014.

Ben-Ari, T., Adrian, J., Klein, T., Calanca, P., Van der Velde, M., and Makowski, D.: Identifying indicators for extreme wheat and maize yield losses, Agricultural and Forest Meteorology, 220, 130–140, https://doi.org/10.1016/j.agrformet.2016.01.009, 2016.

Bevacqua, E., Maraun, D., Hobæk Haff, I., Widmann, M., and Vrac, M.: Multivariate statistical modelling of compound events via pair-copula constructions: analysis of floods in Ravenna (Italy), Hydrology and Earth System Sciences, 21, 2701–2723, 2017.

Bokusheva, R., Kogan, F., Vitkovskaya, I., Conradt, S., and Batyrbayeva, M.: Satellite-based vegetation health indices as a criteria for insuring against drought-related yield losses, Agricultural and Forest Meteorology, 220, 200–206, https://doi.org/10.1016/j.agrformet.2015.12.066, 2016.

Buras, A., Rammig, A., and Zang, C. S.: Quantifying impacts of the 2018 drought on European ecosystems in comparison to 2003, Biogeosciences, 17, 1655–1672, https://doi.org/10.5194/bg-17-1655-2020, https://www.biogeosciences.net/17/1655/2020/, 2020.

COPA-COGECA: Assessment of the impact of the heat wave and drought of the summer 2003 on agriculture and forestry, Committee of Agricultural Organisations in the European Union General Committee for Agricultural Cooperation in the European Union, Brussels, p. 15, 2003.

Del Moral, L. G., Rharrabti, Y., Villegas, D., and Royo, C.: Evaluation of grain yield and its components in durum wheat under Mediterranean conditions: an ontogenic approach, Agronomy Journal, 95, 266–274, 2003.

Durante, F. and Sempi, C.: Principles of Copula Theory, taylor & f edn., https://doi.org/10.1201/b18674, 2015.

Feng, S., Hao, Z., Zhang, X., and Hao, F.: Probabilistic evaluation of the impact of compound dry-hot events on global maize yields, Science of the Total Environment, 689, 1228–1234, https://doi.org/10.1016/j.scitotenv.2019.06.373, 2019.

Ferrise, R., Moriondo, M., and Bindi, M.: Probabilistic assessments of climate change impacts on durum wheat in the Mediterranean region, Natural Hazards and Earth System Sciences, 11, 1293, 2011.

Garcia-Herrera, R., Díaz, J., Trigo, R. M., Luterbacher, J., and Fischer, E. M.: A review of the european summer heat wave of 2003, Critical Reviews in Environmental Science and Technology, 40, 267–306, https://doi.org/10.1080/10643380802238137, 2010.

Gaupp, F., Hall, J., Hochrainer-stigler, S., and Dadson, S.: Changing risks of simultaneous global breadbasket failure, Nature Climate Change, https://doi.org/10.1038/s41558-019-0600-z, 2019.

Górecki, J., Hofert, M., and Holeňa, M.: On structure, family and parameter estimation of hierarchical Archimedean copulas, Journal of Statistical Computation and Simulation, 87, 3261–3324, 2017.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset, International Journal of Climatology, 34, 623–642, https://doi.org/10.1002/joc.3711, 2014.

Iglesias, A. and Quiroga, S.: Measuring the risk of climate variability to cereal production at five sites in Spain, Climate Research, 34, 47–57, 2007.

IM and AEMET: Iberian climate atlas. Air Temperature and Precipitation (1971–2000), Instituto de Meteorologia Portugês (IM) and Agencia Estatal de Meteorologia Española(AEMET), 2011.

IPCC: Managing the risks of extreme events and disasters to advance climate change adaptation, https://doi.org/10.1596/978-0-8213-8845-7, 2012.

Ivan Kojadinovic and Jun Yan: Modeling Multivariate Distributions with Continuous Margins Using the copula R Package, Journal of Statistical Software, 34, 1–20, http://www.jstatsoft.org/v34/i09/, 2010.

Leonard, M., Westra, S., Phatak, A., Lambert, M., van den Hurk, B., Mcinnes, K., Risbey, J., Schuster, S., Jakob, D., and Stafford-Smith, M.: A compound event framework for understanding extreme impacts, Wiley Interdisciplinary Reviews: Climate Change, 5, 113–128, https://doi.org/10.1002/wcc.252, 2014.

Loboda, T., Krankina, O., Savin, I., Kurbanov, E., and Hall, J.: Land management and the impact of the 2010 extreme drought event on the agricultural and ecological systems of European Russia, in: Land-Cover and Land-Use Changes in Eastern Europe after the Collapse of the Soviet Union in 1991, pp. 173–192, Springer, 2017.

Madadgar, S., AghaKouchak, A., Farahmand, A., and Davis, S. J.: Probabilistic estimates of drought impacts on agricultural production, Geophysical Research Letters, 44, 7799–7807, https://doi.org/10.1002/2017GL073606, 2017.

Nelsen, R. B.: An Introduction to Copulas, Springer Series in Statistics, 53, 276, https://doi.org/10.1017/CBO9781107415324.004, 2006.

Nicolas, M. E., Gleadow, R. M., and Dalling, M. J.: Effects of drought and high temperature on grain growth in wheat, Functional Plant Biology, 11, 553–566, 1984.

Okhrin, O. and Ristig, A.: Hierarchical archimedean copulae: The HAC package, Journal of Statistical Software, 58, 1–20, https://doi.org/10.18637/jss.v058.i04, 2014.

Peña-Gallardo, M., Vicente-Serrano, S. M., Domínguez-Castro, F., and Beguería, S.: The impact of drought on the productivity of two rainfed crops in Spain., Natural Hazards & Earth System Sciences, 19, 2019.

Qaseem, M. F., Qureshi, R., and Shaheen, H.: Effects of pre-anthesis drought, heat and their combination on the growth, yield and physiology of diverse wheat (Triticum aestivum L.) genotypes varying in sensitivity to heat and drought stress, Scientific reports, 9, 1–12, 2019.

Ribeiro, A. F., Russo, A., Gouveia, C. M., and Páscoa, P.: Copula-based agricultural drought risk of rainfed cropping systems, Agricultural Water Management, 223, https://doi.org/10.1016/j.agwat.2019.105689, 2019a.

Ribeiro, A. F. S., Russo, A., Gouveia, C. M., Páscoa, P., and Pires, C. A. L.: Probabilistic modelling of the dependence between rainfed crops and drought hazard, Natural Hazards and Earth System Sciences, 19, 2795–2809, https://doi.org/10.5194/nhess-19-2795-2019, https://www.nat-hazards-earth-syst-sci.net/19/2795/2019/, 2019b.

Ribeiro, A. F. S., Russo, A., Gouveia, C. M., and Páscoa, P.: Modelling drought-related yield losses in Iberia using remote sensing and multiscalar indices, Theoretical and Applied Climatology, 136, 203–220, https://doi.org/10.1007/s00704-018-2478-5, 2019c.

Rodríguez Díaz, J. A., Weatherhead, E. K., Knox, J. W., and Camacho, E.: Climate change impacts on irrigation water requirements in the Guadalquivir river basin in Spain, Regional Environmental Change, 7, 149–159, https://doi.org/10.1007/s10113-007-0035-3, 2007.

Salvadori, G. and De Michele, C.: On the Use of Copulas in Hydrology: Theory and Practice, Journal of Hydrologic Engineering, 12, 369–380, https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(369), 2007.

Salvadori, G., Durante, F., De Michele, C., Bernardi, M., and Petrella, L.: A multivariate copula-based framework for dealing with hazard scenarios and failure probabilities, Water Resources Research, 52, 3701–3721, https://doi.org/10.1002/2015WR017225, 2016.

Schumacher, D. L., Keune, J., Van Heerwaarden, C. C., de Arellano, J. V.-G., Teuling, A. J., and Miralles, D. G.: Amplification of mega-heatwaves through heat torrents fuelled by upwind drought, Nature Geoscience, 12, 712–717, 2019.

Serinaldi, F.: An uncertain journey around the tails of multivariate hydrological distributions, Water Resources Research, 49, 6527–6547, https://doi.org/10.1002/wrcr.20531, 2013.

Serinaldi, F.: Can we tell more than we can know? The limits of bivariate drought analyses in the United States, Stochastic Environmental Research and Risk Assessment, 30, 1691–1704, https://doi.org/10.1007/s00477-015-1124-3, 2016.

Sklar, A.: Fonctions de Répartition à n Dimensions et Leurs Marges, Institut Statistique de l'Université de Paris, 8, 229–231, 1959.

Talukder, A., McDonald, G. K., and Gill, G. S.: Effect of short-term heat stress prior to flowering and early grain set on the grain yield of wheat, Field Crops Research, 160, 54–63, 2014.

Vicente-Serrano, S. M., Cuadrat-Prats, J. M., and Romo, A.: Early prediction of crop production using drought indices at different timescales and remote sensing data: application in the Ebro Valley (north-east Spain), International Journal of Remote Sensing, 27, 511–518, https://doi.org/10.1080/01431160500296032, 2006.

Zscheischler, J. and Fischer, E.: The record-breaking compound hot and dry 2018 growing season in Germany, Weather and Climate Extremes, https://doi.org/doi:10.1016/j.wace.2020.100270, 2020.

Zscheischler, J. and Seneviratne, S. I.: Dependence of drivers affects risks associated with compound events, Science Advances, 3, e1700 263, https://doi.org/10.1126/sciadv.1700263, 2017.

Zscheischler, J., Michalak, A. M., Schwalm, C., Mahecha, M. D., Huntzinger, D. N., Reichstein, M., Berthier, G., Ciais, P., Cook, R. B., El-Masri, B., Huang, M., Ito, A., Jain, A., King, A., Lei, H., Lu, C., Mao, J., Peng, S., Poulter, B., Ricciuto, D., Shi, X., Tao, B., Tian, H., Viovy, N., Wang, W., Wei, Y., Yang, J., and Zeng, N.: Impact of large-scale climate extremes on biospheric carbon fluxes: An intercomparison based on MsTMIP data, Global Biogeochemical Cycles, 28, 585–600, https://doi.org/10.1002/2014GB004826, 2014.

Zscheischler, J., Orth, R., and Seneviratne, S. I.: Bivariate return periods of temperature and precipitation explain a large fraction of European crop yields, Biogeosciences, 14, 3309–3320, https://doi.org/10.5194/bg-14-3309-2017, 2017.

Zscheischler, J., Westra, S., Van Den Hurk, B. J., Seneviratne, S. I., Ward, P. J., Pitman, A., Aghakouchak, A., Bresch, D. N., Leonard, M., Wahl, T., and Zhang, X.: Future climate risk from compound events, Nature Climate Change, 8, 469–477, https://doi.org/10.1038/s41558-018-0156-3, 2018.

Zscheischler, J., Martius, O., Westra, S., Bevacqua, E., R., C., Horton, R. M., van den Hurk, B., AghaKouchak, A., Jézéquel, A., Mahecha, M. D., Maraun, D., Ramos, A. M., Ridder, N., Thiery, W., and Vignotto, E.: A typology of compound weather and climate events, Nature Reviews Earth and Environment, 1, 333–347, https://doi.org/10.1038/s43017-020-0060-z, 2020.