Reply to reviewers for HESS-2020-306 article

Title: Copulas for hydroclimatic applications – A practical note on common misconceptions and pitfalls

Authors: Faranak Tootoonchi, Jan Olaf Haerter, Olle Räty, Thomas Grabs, Mojtaba Sadegh, and Claudia Teutschbein

Dear Editor and Reviewers

We would like to thank the editor for providing us the opportunity to discuss the content of this paper with the scientific community. We also would like to thank official reviewers and unofficial reviewers (commenters) for taking the time to examine this manuscript. The latter came as a pleasant surprise as it shows the enormously growing interest in this field and – in a humble tone – interest in this paper.

In this study, we provided end-users with a decision support framework to adopt the copula approach to study the dependence structure between hydroclimatic variables (namely, precipitation and temperature). To back up the proposed framework, we performed a reproducible literature review using the Scopus search engine. We also applied the framework to a case study in Sweden and graphically showed the issues that might arise when this framework is not followed carefully.

We received two sets of official reviews: One detailed feedback from anonymous reviewer #1 which acknowledged positive aspects of the paper (e.g., flow and clarity of the text) but also specifies very useful comments that raise fruitful discussions, which will enable further improvement of this paper.

The main comments from the reviewer were focused on:

1-Why we submitted this article as a research paper to HESS?

2-Did we fairly review and give credit to the statistical papers?

3-Did we address most relevant pitfalls?

We also received another extremely positive review from Geoff Pegrem, quote: *'…What a pleasure it was to review this article. This is possibly the best Hydrometeorological paper that I have read in the last few years and is a must-read in this genre. ... It is targeted at authors involved with, or starting off to work with, copulas in time series.'* For which we are extremely thankful. This actually was the motive behind writing this manuscript: To provide the starters in this field with a well-documented, good narrated text to start working with copulas, to be able to later benefit from highly technical papers in this field. We are happy to see that this goal is achieved.

We would like to provide detailed response to reviewers comment here. Afterwards, we also briefly address the points raised by commenters, many of which have also been raised by the reviewers.

Comment/Response to anonymous reviewer #1

General comments

I agree with the authors that copulas are often used in the fields of hydrology and climatology and that there is a lot of room for improvement in when and how they are applied. In my point of few, the most common 'pitfalls' are that (1) the nature of dependence is often not studied before starting to test various copulas; (2) the dependence structure is often reduced to correlation, (3) no proper goodness-of-fit tests are applied to reject inappropriate copulas. While the authors detect several other pitfalls related to P-T analyses, these important pitfalls are not addressed. By looking at the literature I get the feeling that the authors are not very familiar with the basic statistical copula literature [*Nelsen*, 2006; *Joe*, 2015], which may have prevented them from coming up with a comprehensive list. While I am in favor of a piece addressing such pitfalls, I rather see this as a technical note, a review, or a commentary than an independent piece of research because it does in my point of view not present novel concepts, ideas, tools or data. In addition, I think that such a manuscript should properly review and cite the statistical copula literature and acknowledge previous practical guides for copula application in hydrology e.g. by [*Genest and Favre*, 2007]. In addition, it should not create the wrong sense that dependence=correlation because dependence is a wider term including other dependence properties such as symmetry or tail dependence [*Joe*, 2015]. Even though I do not see this manuscript as a paper in the scope of HESS, I provide some suggestions of how to improve it because I think it could be published as a review/commentary in another hydro-meteorological journal after major modifications.

Again we would like to thank anonymous reviewer #1 for their detailed and thorough feedback. We would like to provide an answer to their three major comments. Which are:

1-Why we submitted this article as a research paper to HESS?

2-Did we fairly review and give credit to the statistical papers?

3-Did we address the most relevant pitfalls (pointed by reviewer as numbers 1 to 3)?

For the first point: although we see the potential of this paper as a review paper or a technical note, we argue that we developed a practical decision support framework (= new tool) for the application of copulas to support hydroclimatic researchers, which is illustrated in detail using a particular case study in Sweden (new data). We analysed the data with this approach and discussed the changing nature of dependence within our case study.

To that end, we continue to believe that this manuscript is a perfect fit to be published in HESS, which is aiming "to serve not only the hydrological science community but all earth and life scientists, water engineers, and water managers". Also, because this platform allows for discussions among the scientific community, we believe that our manuscript serves this purpose, because of the clear text that can guide the user in further adoption/adaptation of this framework as it has already been noticed by some researchers. Taking all this into consideration, we think that this manuscript (in an updated version) could potentially attract a wide audience.

For the second point, while we extensively used statistical papers such as Genest and Favre (2007) or Genest, C., B. Rémillard, and D. Beaudoin (2009) and referenced them in the

original manuscript, we can see that further credit is due upon these important papers. We will address this in the revised manuscript.

For the third point, 1- Yes, that is one of our goals to ensure future researchers comprehensively analyze their data and its dependence structure, before using statistical models such as copula.

2- We do agree that the nature of dependence should not be reduced to correlation and we will further discuss this in the revised manuscript. However, correlation – in general sense, including Spearman rank correlation – is an easy and available test to evaluate the dependence – in its informal definition – between variables. Without putting too much burden on the end-users, correlation tests can give a sense of the data. Further, we did not introduce correlation as a requirement for copula analysis, we simply suggest correlation analysis as a first step for the researchers to understand their data before modeling.

3- We argue that we extensively discussed goodness-of-fit tests (we adopted p-value following the procedure discussed in Genest et al. 2009) both in the methodology section (section 2.5) and later when mentioning copula fitting, and rejection of unsuitable copulas. However, we will try to put more emphasis on this in the revised manuscript.

Specific comments

Title: I think that the title is too general. The study only reviews manuscripts related to P-T copula analyses and some of the pitfalls described are very specific to that pair of variables (e.g. ties, zero values). I would rephrase it to something like: 'Copulas for joint precipitation-temperature studies – a practical note on common misconceptions and pitfalls'.

We appreciate this comment. The reviewer is right that our case study is focused on precipitation-temperature relationship; however, the nature of the proposed framework and guide to use copulas is general and can be applied to a myriad of variables.

Abstract: goal, methods, and outcome of study are clearly described. I would probably summarize the different pitfalls identified to summarize the conclusions.

We appreciate the reviewer's positive comment and will conclude the identified pitfalls.

Introduction: The study content is generally well introduced. I would personally jump into P-T analyses a bit more directly by removing the first paragraph because it rises the expectation that different hydro-climatic variables are addressed in the manuscript, which is not the case. Instead, I would extend the section on where previous studies looking at hydro-climatic variables are introduced (l. 89-92) because I think this short section does not do the existing literature justice. I would in particular better introduce the study by [*Genest and Favre*, 2007] that describes 'the various steps involved in investigating the dependence between two random variables and in modeling it using copulas' and illustrate these steps on a hydrological example. I would also highlight what exactly is the benefit of your study compared to this previous one, which had a very similar goal.

We appreciate the reviewer's positive comment. We will address this comment when revising the manuscript.

Step by step copulas: This section in my opinion needs a more solid theoretical/statistical basis. Proper citations to the statistical literature should be provided for all equations and

statements. All variables should be introduced properly and used consistently. I would furthermore expect a discussion of the following points:

When revising the manuscript, we will improve this section as suggested.

(1) Two-dimensional copulas are not popular in hydro-climatology because people are necessarily interested in only two variables but rather because they are easier to apply and visualize than higher dimensional copulas.
We will mention this.

(2) I would add a short section about when to use empirical rather than theoretical copulas and vice versa.
We will add this part.

(3) I would mention that copulas model the form and intensity of dependence between variables. The form can be represented by the choice of the copula function while the copula parameter describes the intensity.

We will mention this.

(4) I would introduce the notion of dependence and clearly state that correlation only describes one particular part of a dependence structure which also comprises tail dependence [*Poulin et al.*, 2007] or symmetry characteristics. I would recommend having a look at Chapter 2 in [*Joe*, 2015]. These additional characteristics are very important for choosing a suitable copula form.

We will add this part.

(5) The particular copulas introduced seem a bit random. Why are extreme value copulas not introduced? They are very important when looking at joint P-T extremes. At least, it should be specified how you determined 'the five most widely used copulas'.

We did not include extreme value copulas as we looked at the dependence between complete range of precipitation and temperature. We do agree with the reviewer that justification is needed for the selection of five most widely used copulas. For this, we analyzed copulas that have been used in the studied literature returned from Scopus. This will be reflected in the revised manuscript.

(6) Some Archimedean copulas have more than one parameter and Archimedean copulas have the disadvantage that the same degree of dependence is assumed for all pairs of variables.
We will further assess this.
(7) Elliptical copulas have the advantage that they can handle the same degree of dependence for different variables pairs but they have symmetric dependence structures which may be a disadvantage in some cases [*Favre et al.*, 2018].
We will add this part.
(8) I would treat parameter estimation methods separately and mention why maximum likelihood in some cases can be computationally very expensive and may be replaced by pseudo-maximum likelihood estimation [*Han and De Oliveira*, 2019].
We will include this part.
(9) I think equation 16 is wrong. Where does it come from?

We double-checked and can confirm that this equation is correct. Please refer to eq. 1-6 in Shojaeezadeh et al. 2018

(10) A goodness-of-fit test never ‘accepts’ a hypothesis but rather ‘rejects’ it. ‘non-rejection’ does not imply ‘acceptance’.

Agreed.

(11) It is new to me that NSE and RMSE can be used as copula evaluation metrics (l. 242-245). NSE is used to evaluate time series rather than distributions. I do not see the link to the dependence structure (except that correlation is evaluated as part of NSE) and neither is the statement underlined by a reference.

NSE is used herein in accordance with its counterpart Coefficient of Determination that is used for statistical models. This can provide a sense of how close the copula-predicted probabilities are to their empirical counterparts, as opposed to p-values that do not show this. In the studied literature, we found that NSE and RMSE have been used to evaluate copulas. NSE and RMSE are two widely used metrics in hydrological studies. We, however, showed in table 2 that relying solely on these two metrics, might be misleading. We will include more discussion about proper copula evaluation metrics.

Common issues, misconceptions and pitfalls: I would move the methods description (l. 248-264) to some methods section. I would also describe how the ‘six aspects’ investigated (l.255-264) were determined. As mentioned in my general remarks, I would also look at whether authors characterized the nature of dependence (e.g. by looking at rank scatterplots or by computing different dependence metrics including tail dependence) and I would look at whether they performed a proper goodness-of-fit test [*Genest et al.*, 2009]. Furthermore, I would suggest to illustrate the different concepts on your case study example in a separate section called ‘Application’.

Thank you for the detailed suggestions. We would like to re-emphasize that we had used p-value as a formal goodness-of-fit measure for copulas in the original manuscript. In revising the manuscript, we will emphasize the importance of rank correlation and rank plots as an appealing approach in determining presence or lack of dependence between variables. When revising the manuscript, we will make some adjustment to the structure.

Spatio-temporal scale: It remains unclear to me why exactly this matters unless you wanted to model spatial dependencies. I would introduce the case study in the newly created methods section as suggested above (l. 278-280).

Spatio- temporal scale can be of importance, as the significance of correlation might change depending on the scale (fig 8). This generally speaks to the importance of understanding the physical processes that govern the data generation, before using statistical models for any analysis.

Correlation: I would call this section ‘Dependence’ and discuss dependence aspects going beyond correlation as assuming dependence=correlation is a pitfall in itself (see also my earlier comments).

We agree and will further discuss this.

How should correlation be independent of the selected sample?

We illustrated this in figure 5. We meant here that the correlation must not be merely random (e.g. number of divorces increases as the number of marriages increases, but is marriage a cause for divorce?). We will change this sentence to clarify.

(l. 288-289) By 'generate copula co-dependence structure' do you mean 'fitting a copula structure'?

Yes

Stationarity of correlation: I would not say that the detection of non-stationarity per se precludes a copula analysis. However, it requires the use of a proper non-stationary model [e.g. *Ahn and Palmer*, 2016]. I do not see the value of the resampling experiment (l. 319-327). Why should this be useful to detect non-stationarity? Why not just test how mean and variance change over time?

We would like to thank the reviewer by pointing out this very interesting point. We do agree that 'stationarity of the hydroclimatic' can be addressed differently, thus should be discussed more cautiously. In fact, this is a hot topic in hydroclimatic research and we will add more discussion, referencing the papers mentioned by the reviewer.

However, we would like to argue that the motive behind writing this part was not disapproving any use of statistical methods in presence of climate change. Rather by this part and further by illustration in figure 5, we wanted to point out the importance of careful analysis of dependence within a timeframe. Accordingly, in figure 5b we showed that correlation between precipitation-temperature – and thus copula parameter – may exhibit significant variation within a time frame, based on the selected sample. In other words, we wanted to point out that it is necessary to check if the observed correlation strength was purely by chance. Therefore, the conclusion that can been drawn from copula parameter – as a measure of strength of dependence – may be deceptive in certain cases.

Autocorrelation: 'time series are dependent on a delayed copy of themselves'?

We will change this sentence to make it clearer.

Correlation of data with the same rank: use the term 'ties'. I would remove subsection 3.5.1 because there is just one subsection at that hierarchy level.

Agreed.

Selecting suitable copula families: I would remove the NSE and RMSE part (l. 399-400).

P-value was used as the statistically sound goodness-of-fit metric in the original manuscript. As mentioned earlier, we do agree that NSE and RMSE are not traditional copula performance measures but they have been widely used in hydroclimatic applications. On the contrary to p-value that only speaks to whether the data are from a certain copula family, NSE and RMSE provide a sense of how close are empirical and copula-predicted probability levels. These metrics are not to replace p-value, but to provide further information to the user.

Decision support framework for applying copulas: would add 'to jointly model P and T' because some of the points are very specific (particularly the one on ties). I would re-order the different steps and put pre-treatment steps such as removal of autocorrelation (2), testing for non-stationarity (3) and ties (4) before dependence assessment (2) and copula fitting (6). I

would also include two additional steps: (x) visual inspection of dependence structure and (x) goodness-of-fit testing. My new suggested order is the following: (1) scale (if this is even important), (2) removal of autocorrelation, (3) removal of ties, (4) testing for non-stationarity, (5) visual dependence assessment, (6) computation of dependence metrics, (7) copula fitting, (8) goodness-of-fit tests. Maybe you could even have two main parts called (A) pre-treatment and (B) copula analysis.

We highly appreciate reviewer's suggestion on re-ordering the steps. We found dividing the paper in two main parts very useful. We will consider changing the structure when revising the manuscript.

Concluding remarks: I would discuss which parts of the decision framework are transferable to other variable pairs and which ones are specific to P-T analyses.
We will include such a discussion.

Structure and language: The manuscript generally has a nice flow and would profit from some editing.
Thanks, will be done.

References: Some additions from the statistical literature required as specified above.

We will add further statistical references to this manuscript.

Figures:
In general, I would recommend the use of subplot labels (a, b, c) to facilitate referencing.
Will be added where missing.

Figure 2: What do these turquoise bars on the left and lower part of the figure to the right mean?
Uniformly distributed histograms of the margin. We will mention this.

Figure 3: would remove the grey borders in the figures to the right (point clouds).
Not sure what the reviewer means here: they are x and y borders.

Figure 4: would use distinct colors in the different subplots (different shades of turquoise are used for lakes).
We will adjust this.

Figure 5: As mentioned above, I do not see the value of the analysis presented in 5b.
Through this figure, the changing dependence strength and significance can be shown, when a combination of different years are selected to form a decade and compute the correlation.

Figure 6: remove random black borders and increase legend (one should be enough).
We will adjust this.

Figure 7: Would recommend to add isolines to the scatterplots in 6b.
Fig. 6b shows autocorrelation. We are not sure how to add scatterplot to this figure.

Figure 8: Would recommend to restructure figure according to the steps order suggested above.
We will consider this option when revising and restructuring the manuscript.

Minor points

- I am less familiar with the term ‘co-dependence’ than ‘interdependence’. Evtl. reword? (e.g. l. 12).

  Agreed.

- 11-13. I would restructure the sentence and start with the subject ‘Several multivariate analysis approaches have….. to account for precipitation….

  Agreed.

- L. 55: I would talk about ‘joint’ instead of ‘compound’.

  Agreed.

- The use of commas could be improved, e.g. l. 70 ‘At the annual resolution,…’ or l. 73: ‘However,…’

We will check this.

- L. 136: ‘can be’ instead of ‘needs to be’

Agreed.

- .139: ‘provide’ instead of ‘provides’

Agreed.

Comment/Response to reviewer #2: Geoff Pegrem

What a pleasure it was to review this article. This is possibly the best Hydrometeorological paper that I have read in the last few years and is a must-read in this genre. It is targeted at authors involved with, or starting off to work with, copulas in time series. The difficulty that presents itself when analyzing time series characterized by serial correlation, is that that for analysis, modelling or forecasting, the leading question is: 'how do I get a handle on this problem?' The beauty of the paper is that it a distillation of ideas into a rubric for preparing an analysis of one or more time series, to be finished off with a flow chart for guidance.

We appreciate the reviewer for their very positive feedback and recognizing the value of this work. The motive behind writing this manuscript was to provide starters in this field with a guideline to find their way through many-existing non-statistical literature. We are very happy to see, this was perceived by the reviewer as well.

It is an important reminder and guide for time series analysis, and is not only tutorial, but is wisely, simply, and authoritatively compiled. In my judgement this should be published once some small issues have been dealt with. For example, the authors should attend to some cosmetic suggestions to fix the few spelling and grammatical errors, as well as embellishments in the figure and table captions to make them more readable. Again, stylistically, it would improve the readability if you either add a space between all paragraphs or indent the leading line. Also, I could not find 'saturation water vapor mixing ratio' in this paper. My more pertinent remarks follow.

We again appreciate the reviewer's positive comment. We will carefully revise the manuscript for potential grammatical errors and otherwise. We will change the referenced paper where we explained 'saturation water vapor mixing ratio' from (Vaghefi et al., 2014) to (Vaghefi et al. 2019). Where it is stated: 'A sequence of processes due to increasing greenhouse gasses, could be summarized as (i) increases in air temperature and its capacity to hold more water…'

In section 3, line 326, you state: 'We adopted a copula framework on July, because it has significant correlation at both daily and monthly resolutions (Fig. 5(a)-(b)).' Did you try lagging the daily precipitation and streamflow data? Surely the delay depends on the size of catchment.

We only focused on dependence structure between precipitation and temperature variables for this manuscript. Size of catchment can potentially have an impact on the dependence structure between precipitation and temperature too, and we will clarify the text.

Line 338: 'It is important to consider the degree of autocorrelation in the studied data.' Without pushing my co-authorship of a relevant paper, to check the effect of autocorrelation, you might like to look at: Sugimoto, Takayuki, András Bárdossy, Geoffrey G.S. Pegram, and Johannes Cullmann (2016), Investigation of hydrological time series using copulas for detecting catchment characteristics and anthropogenic impacts, Hydrol. Earth Syst. Sci., 20, 2705 -2720, doi:10.5194/hess-20-2705-2016.

We thank the reviewer for his suggested reference and will consider this paper.

In Figure 6: What is the spread of the confidence intervals - 95%? In (d) it looks like

100%

It is on 95% interval. We will state that in the figure.

In Figure 8: that's a very helpful flow-chart - especially the '!!'

Thanks again for recognizing the value of this chart.

In Table 2: please define the symbols in the caption to help the reader: ERMS, ENS,Sn and Dn.

We will modify. Thank you.

There are also a few minor suggestions that I have made for alteration, so I am uploading my marked-up copy of the paper with this review for the authors' information.

We appreciate the reviewer for his effort in evaluating the paper thoroughly and providing valuable comments.

Well done!

Thanks again.