Weather Clim. Dynam. Discuss.,
https://doi.org/10.5194/wcd-2020-32-RC1, 2020

Interactive comment on "Impact of North Atlantic SST and Jet Stream anomalies on European Heat Waves" *by* Julian Krüger et al.

Anonymous Referee #1

Received and published: 21 August 2020

The current work considers anomalies of SSTs and the upper tropospheric waviness in the North Atlantic sector in an attempt to estimate their impact on the occurrence of European heat waves. Specific aspects of both factors have been suggested in the past to be conducive to European heat waves, but this paper considers, for the first time, both factors simultaneously. A novelty hereby is the fact that the authors focus on the waviness over the North Atlantic rather than the circumglobal waviness as was common in numerous previous studies.

The manuscript contains a number of interesting aspects, like the lag-correlation analysis in order to tease out a cause-and-effect relationship or the discrepancies with respect to previous work regarding the wave behavior in connection with European heatwaves. There seems to be potential for progress in these directions. At the same

time, I see a number of fundamental weaknesses. First of all, the local double jet index introduced by the authors seems to be incompatible with the idea of quasi-resonance and, therefore, not an adequate metric to quantify the latter. Also, the analysis is mostly based on composites and correlations. A correlation does not imply causation, and therefore the analysis does not live up to the expectation raised by the title ("Impact of .... on ...."), which promises to shed light on causal connections. In addition, many of the shown correlations are pretty weak; in no case do the authors investigate whether the correlations and the patterns on the composite maps are statistically significant or not. Throughout the text the authors relate their results to those of earlier studies, and they sometimes find agreement and sometimes disagreement. In the end this left me lost and I could not find out what I am supposed to learn: are some of the earlier results wrong (because of the detected disagreement), or are the authors' own results an outcome of an inadequate metric (see above) or pure chance (which would be consistent with a potentially missing statistical significance) or both?

Below I provide further details under "major issues" in order to illuminate my concerns. Items 3–5 could be addressed by major revisions. Item 2 seems critical, although one might argue that a subsequent proof of statistical significance (if that turned out to be possible) would save the day. However, item 1 is fundamental in the sense that a substantial part of the analysis is based on the local double jet index, which in my eyes is inappropriate to fulfill its purpose. Addressing all these issues will basically result in a very different manuscript. I thus have to recommend rejecting this manuscript.

I chose not to mention further minor issues.

Major issues

1. The theroy of "quasi-resonance" requires Rossby waves to travel around the entire globe. This is a fundamental aspect of quasi-resonance, because the singularity in the case of no damping (see equation (3) in Petoukhov *et al.* 2013)

results from the constructive interference of an infinite number of waves, which is only possible if the domain is periodic in longitude and if the waves see a zonal waveguide at all longitudes. Correspondingly, previous authors have analyzed the zonally averaged zonal wind and computed the corresponding stationary wavenumber $K_s(\phi)$, which depends on latitude ϕ. According to WKB theory, a relative maximum in $K_s(\phi)$ — which typically occurs in the presence of a double jet structure — is then indicative of a zonal waveguide all the way around the globe.

In the current manuscript, the authors diagnose the double jet structure locally over the North Atlantic ocean. This leaves open two issues: first, does this also correspond to a relative maximum in $K_s(\phi)$? And, second (more fundamentally), even if there were a local (in longitude) relative maximum in $K_s(\phi)$, this alone would not indicate a wave guide all around the globe. Therefore, it seems to me that the authors' local double jet index is inadequate for diagnosing a zonal waveguide as required for the validity of quasi-resonance theory.

2. One key concern is that many of the shown effects (correlations, composites) are rather weak and it is not clear to me whether they are statistically significant of not. It appears necessary to me to determine the statistical significance of all these results before they can be presented and discussed in a scientific paper. For instance: Are the differences shown in Fig. 3 statistically significant? Are the weak correlations shown in Figs. 4 and 5 statistically significant? The same for the correlations shown in Fig. 6, which are even smaller (a correlation coefficient of the order of $r = 0.2$ corresponds to a "described variance" of $r^2 = 0.04 = 4\%$, which is very low and requires a large and high-quality data set to tease it out in a significant manner).

   The issue of statistical significance and the interpretation of a correlation in terms of causality can be illustrated in Figure 2. Apparently, some heatwaves over Europe are associated with cold SSTs, some are not, and sometimes there is

a cold SST anomaly but no heat wave. Overall this is a very mixed bag, and I believe that the correlation between two respective time series is very low and probably not significant. On the other hand, if cold North Atlantic SSTs were really "drivers" ("impact", "influcence", etc.) of European heatwaves, wouldn't this correlation have to look more convincing?

3. The authors talk about "impact of ... on ...." (title), about "drivers of ...." (second paragraph in the introduction), about the "influence on ...." (section 3.1 heading), "A promotes B...." (conclusion section) etc. All these terms suggest that this paper is about identifying causes for European heat waves. Correspondingly I started to read the paper with high expectations and was disappointed later on, because in my eyes the kind of analysis presented in the manuscript does not allow one to determine the causes for heat wave occurrence. As I said above, correlation does not imply causation. More specifically, I do not believe that the "influence of ... on ..." can be "evaluated by using a composite study" (as promised, e.g., on line 96). An analysis based on composites or on correlations is at best suggestive, but certainly not conclusive. To give an example: the plot shown in Fig. 3 indicates an association between North Atlantic SST and the upper tropospheric flow over that region in two particular years. The questions that remains unanswered: are these two years just a fortuitous coincidence (see my earlier issue about statistical significance), does the ocean "drive" the atmosphere, or does the atmosphere "drive" the ocean? In summary: given what the authors show in the manuscript, I think that the language is inflated and the analysis does not live up to the expectations raised by some of the formulations.

4. Some of the reasoning seems rather casual. For instance, the authors extend the analysis of a previous study to the year 2018 and show that the European heat wave was associated with a simultaneous cold SST anomaly over the North Atlantic similar as in that previous study for a previous year. So far so good. On the other hand, Fig. 2 shows that some other extreme heat waves such as the

2003 or the 2010 heat wave were *not* associated with strong cold SST anomalies over the North Atlantic. So what am I supposed to learn from that? Does it mean that the association between SST anomalies and European temperatures is a random event? The authors do not provide any guidance to the reader as to how to interpret these results.

5. I have an issue with the lag correlation analysis (Figs. 4, 5, and 6) in combination with preprocessing the involved variables through a 15-day running mean. A running mean is a *non-causal filter* which projects information backwards in time. I do not say that the results shown in Fig. 6 are invalid, I am just skeptical and it would be more convincing if the same results could be obtained without a non-causal filter. Also, the four plots in Figs. 4 and 5, respectively, are very similar to each other, and it appears to me that this is in essence a result of the 15-day running mean filter preprocessing: the latter smears out the data in the time domain such that a time lag of 4 or 8 (or even 12) days becomes almost invisible.

   In the end the issue goes back to the question of statistical significance (see one of my previous items). For instance, the effective number of degrees of freedom in Fig. 6 is likely to be very low as a result of the 15-day running mean. This makes me believe that a difference in the correlation coefficient of as little as 0.2 between a 15 day lead and a 15 day lag could be statistically insignificant.

References

Petoukhov, V., S. Rahmstorf, S. Petri, and H.-J. Schellnhuber 2013. Quasiresonant amplification of planetary waves and recent Northern Hemisphere weather extremes. *Proceedings of the National Academy of Sciences 110*(14), 5336–5341, doi:10.1073/pnas.1222000110.

Interactive comment on Weather Clim. Dynam. Discuss., https://doi.org/10.5194/wcd-2020-32, 2020.