1 General comments

The submitted manuscript covers important dynamical aspects on the evolution of heat waves and tries to link those to understand the underlying mechanisms. The authors focus on previously identified potentially important drivers of heat waves, namely the wave amplification according to the quasi-resonance mechanism and Atlantic precursor SST anomalies. Those feature have been previously shown to be linked to some heat waves. The authors further investigate the role of jet stream anomalies as link between those features and try to answer the question about their importance on heat waves. The study nicely combines the investigation of two case studies (heat waves in 2015 and 2018) and a more general climatological perspective of a 40 years period using ERA5.

The study is nicely motivated and it picks up two (three) mechanisms which are further investigated in detail and the authors tried to investigate a possible link between them. This setup of the paper appears quite exciting, however, the manuscript still needs some substantial improvement to more convincingly support the claims they present with their results. One major issue I have with the study is that their quick conclusions about causal relationships is not made very convincing. I don't think that the presented results are enough to support their causal relationships (specific comments about this further below). This seems to be even more problematic that the results are not supported by any means of statistical significance test. As mentioned further below for the specific comments I did struggle to understand if the shown patterns are really meaningful enough to come to the conclusions by the authors. I think applying some measure of statistical significance could be helpful for the authors to make their point more convincingly and convince the reader that their results are indeed meaningful. A related issue I have with this is the missing understanding of sensitivity of the results to some of their choices. I made this point also in the specific comments, like for example for the choice of double jet index. I would suggest that the authors should either refer to another study where this index has been used, or at least motivate their choices or tell the reader about the sensitivity of the results on those subjective choices. This issue appears several time during the manuscript and is also enhanced by the authors themselves, by showing that changing the averages boxes for SST or temperature can lead to different results. This raises the question, how much the following results can be trusted or if a meaningful interpretation is difficult due to this sensitivity. I would suggest that the authors give the reader further insight into the impact of these changes.

Further, I did struggle a bit to understand what the actual real goal of this paper is, related to wave amplification mechanisms. The quasi-resonance mechanism is a possible mechanism relevant for heat waves, but it seems that the authors are interested in the dynamics between different contributors in the Atlantic to support the evolution of heat waves, SST anomalies, jet structure and evolving wave patterns. This impression is further strengthened by the authors choices to deviate from the procedure of the quasi-resonance mechanism on a more locally restricted scale (in terms of extension of double jet structure). But why then not using a more general wave measure? Previous literature has shown that not all heat waves are associated with global scale waves, but more locally restricted ones. I assume the reason is that the authors were only interested in this sub-class of waves associated with heat waves, which is fine, but I think that the authors then should make this point more clearly and highlight that they specifically are only focusing on such waves. This includes also to clarify why the authors nevertheless deviate in their search for dynamical links to locally restricted jet anomalies. So the authors should clarify their choices, that they investigate potentially important aspects that can lead to heat waves, but also why they link those without clarifying why the authors think it is reasonable to link and consider explicitly those features together?

In general, I think the manuscript represents a nice study which tries to understand specific aspect and their dynamical links that can lead to European heat waves. However, at the moment the conclusions of the paper are not presented convincingly. This could be improved by better explain to the reader the motivation behind some of their diagnostics, including some support from including some kind of statistical significance test, comments about the sensitivity of the results and being a bit clearer on the aim of this paper (related to the applied diagnostics).

2 Specific comments

2.1 Section 1 (Introduction)

lines 60-62:
The aim of this paper is not fully clear to me and I am slightly confused about the focus between SST anomalies, heat waves and QRA. These two sentences are a good summary for this slight confusion, because it seems like the authors have some underlying assumptions that are not explicitly made clear to the reader. According to the second sentence (lines 61-62), the research objective is to investigate the connection between European heat waves and North Atlantic SST anomalies with the double jet structure as possible dynamical link between those two (with the associated wave properties that can be linked to the jet structure). Therefore the authors explicitly want to investigate the link between Duchez et al. (2016) and Kornhuber et al. (2017). I find it difficult to understand this reasoning. If I understand the mentioned studies correctly, Duchez et al. (2016) try to link the SST anomaly to heat waves, resulting from some form of atmospheric wave signal that can be linked to the SST anomaly. The QRA proposes a specific subset of waves which are assumed to be important for the occurrence of heat waves. However, the QRA is a subset of waves and there are also more general studies linking atmospheric waves to European heat waves. As the authors are interested in linking SST, jet structure, waves and European heat waves it is not clear why the study attempts to use the QRA as the underlying wave mechanism of relevance. E.g. in Wolf et al. (2018) the authors showed on the one hand a possible relevant contribution from the QRA to extreme positive temperature anomalies over Europe during NH summer, but also mentioned on the other hand that a specific contribution of the waves connected with extreme events were a rather local feature which does not so well fall into the QRA

perspective. As the study of Duchez et al. (2016) is not depending on the QRA nor did they make this connection, it is not clear why the authors focus explicitly on this link. So the main objective is to understand the QRA and trying to identify possible underlying processes for this QRA, for which the authors focus on the SST anomalies highlighted by Duchez et al. (2016)?

2.2 Section 2 (Data and Methods)

lines 66-69:
The objective of the study is to understand the dynamical link between SST, jet structure and temperature anomalies. Wouldn't it therefore make sense to subtract the climatological trend from the underlying variables to not run into danger to get a bias into the results for more extreme warm events occurring later in the investigated periodic?

lines 74-76:
As previously mentioned, I find it difficult to understand on the one hand the focus on the QRA, but on the other hand deviating from it to investigate rather more local processes. This somehow does a bit contradict with explicitly focusing on the global QRA feature. The authors should probably be clearer on what their study is aiming for - what is it exactly they try to understand.

lines 82-83:
In the previous paragraph the authors explained how they identify the three climatological branches of the jet structure for May/June and July/August from the climatological state. So the branches define these three step functions (black lines with dots) in Fig. 1, correct? In lines 82-83 the authors explain how they calculate the branches for each month, but I thought the branches are derived from the climatological fields. Do the authors intended to explain how they calculate the associated U-values for each month based on the climatological branches? Maybe the authors could make this description a bit clearer.

line 94:
Is there a reasoning behind the specific choice of DJI or do the specific not really matter? It is some kind of curvature (Unorth - 2*Umid + Usouth), but with more weight to the absoulte maxima of the jet maxima. Would the use of the curvature instead have an impact on the results?

line 97:
How do the authors remove the seasonal variability? Do they use a daily based climatology? I think the authors should specifically mention this.

lines 104-105:
What does this exactly mean (daily-based detection of the jet stream)? Is the maximum calculated from the composite fields and daily based refers to applied timelags for this

composite, or is the location of the maximum represented by the average of the locations for each maximum at every day used for the composite?

line 110 (but also valid for the other figures):
I would suggest to order the figures in the way they are refered to in the text.

line 110:
Is there a reason to apply a 15 day running mean to one variable, but not the other? Why is this running mean necessary. Using the 0.25/0.75 quantiles (or even 0.1/0.9) should include enough cases to get a smooth line for the time-lagged correlations.

line 125:
Maybe the authors could spend a bit more time on explaining this index. What does this measure of a trough or ridge, relative to the zonal mean tells the reader. An overall southward or northwards shifted jet in the Pacific would by this measure have an impact on zanom in the Atlantic. This could result in an increase or decrease of the relative strength of the Atlantic ridge or trough, even for identical conditions in the Atlantic. How can this interpreted or what is it aimed for to explain? Also in the following the definition of a trough/ridge seems a bit surprising. Isn't this region in a climatological sense represented by higher z-values, which would mean that by this definition one could identify ridges (zanom¿0.5) for cases when a gridpoint based climatological anomaly would indicate a negative anomaly (trough).

2.3 Section 3 (Impact of North Atlantic SST anomalies)

lines 155-159:
What is the conclusion here? Must the hypothesis be rejected that cold Atlantic are leading/associated with European heat waves? Or is it necessary to analyze the spatial pattern shift in T2m and SST anomalies? The authors should be clear on their interpretation of their results. This missing interpretation also raises further questions in the following paragraph. The authors indicate that there is no clear connection between Atlantic SST and European T2m, but now they investigate if the SST is responsible for the phase-locking of the jet leading to summer extremes in Europe, which they said are not linked. I therefore suggest that the authors include some further description and interpretation of their results to avoid this confusion.

lines 191-192:
How do the authors come to this conclusion? Jet stream position/wind and SST anomalies are maybe not surprisingly associated, therefore a change in the jet stream position will be also associated with SST anomalies, or those SST anomalies could also be driven by a shift in the jet stream position. The authors have not investigated the causal relationship and therefore should rephrase their statement that it does not lead to a misleading interpretation of a causal relationship. The author should be clear on this and rather explicitly refer to their following lead-lag-correlations to gain further insight

into this connection.

lines 202-207:
These findings (Fig. 4) are a bit surprising, maybe the authors can give some further insight and interpretation of these findings. With synoptic systems usually dominating the midlatitudes (and representing the GPH fields), I would not necessarily expect such a persistent correlation pattern. As the correlation maps look nearly constant, I would also expect the same correlations for 20 to 30 days. This however would make me doubt this is a good measure to investigate the prediction of some dynamical features as the build up of a wave pattern. And does the correlation goes towards zero for negative lags? I have the same issue also with Fig. 5. The patterns suggest some longterm connections, like large scale NAO patterns which can be associated with large scale flow and temperature anomalies. But to investigate such features, a 4 day lag seems a bit strange, which would rather suggest the authors try to investigate the quick atmospheric reply to the cold SST anomalies in amplifying the wave pattern downstream. I think the authors should better describe what it is exactly they are trying to investigate and with it explain their choice of time lags. Do those kind of plots show that the SST anomalies are extremely persistent and define the overall large scale flow pattern for Europe for the whole summer. If so, shouldn't the authors also show negative lags to support their statement of a causal relationship? I could also imagine a wave pattern building up the SST anomalies, with the wave pattern amplifying itself further, which could also lead to the shown correlations, but which would also have the same correlation pattern for negative lags.

lines 241-243:
I think more discussion is necessary and discrepancies cannot be disregarded due to a small sample size. First, if the mechanism is relavant, it should also be visible in the more extreme cases - I thought this is the whole point, presenting a mechanism for extreme summer events. Second, the sample size is still huge as it represents 20 percent of the whole 40 year period. Not even does the higher percentile not support the previous highlighted connection, it further suggests that T2m is leading the SST, raising the question how reliable the claim is that the cold Atlantic SSTs can be seen for a precursor of warm European temperature anomalies. Additionally, this includes the description of lines 226-240 (Fig. 6), the authors should give further interpretation to these correlation plots. It seem a bit surprising that for the cold and warm cases correlation with T2m goes towards 0 towards higher positive lags (stronger for higher percentile in Fig S2) or positive for negative lags, whereas the climatological signal is strongly negative (at high positive lags) or has the opposite sign for negative lags. As for Fig. 6, warm and cold cases already represent 50 percent of the data, this means that the more climatological cases (in SST anomalies) should represent strong negative correlations (?). Wouldn't this somehow contradict the importance of this dynamical connection associated with strong Atlantic SST anomalies? What about cold/warm cases in T2m (instead of SST), if those would be included in the correlation, would they agree with the given interpretation for the importance/impact of SST anomalies?

line 249:
I think the SST anomaly was previously not refered to as subpolar. Maybe it makes sense to be consistent in using the same terms, as subpolar probably is misrepresenting the very specific cold/warm blob anomaly in the Atlantic.

lines 251-253:
I think I mentioned this earlier already, but isn't for the QRA a more global double jet structure necessary whereas here the author are limiting this to a very regional feature? So if the authors want to limit their analysis to a regional double jet structure, why do they pick specifically the QRA as underlying wave response, or if they want to investigate the impact of the QRA, why limiting the double jet structure locally?

line 254:
Maybe refering to the "extended" summer season, as MJJA is not the usually used range for summer season.

lines 258-259:
Could the authors include in their correlation diagnostics maybe also some estimate for statistical significance, something like combining correlation with associated p-values? Also how much depends this picture (Fig. 7) on the choice of the DJI, which seems (as previously mentioned) to be chosen subjectively (or is this a usual way to measure the DJ structure?). But anyway, in general it would be interesting to know how sensitive the results are on the choices for the DJI (to define the index as well as the chosen longitude range) or maybe as well the SST average area (what about the second, newly introduced slightly shifted area?).

caption Fig. 8:
figure caption only for upper panels, not lower panels. Further, how do the quantiles match? (title 0.25-0.75 quantiles, caption 10% most extreme cases). And what data is actually used, I am not sure I fully understand the description, mentioning the coldest 10% with the 10% highest DJI. Are the anomalies calculated from the dates either being in the 10% coldest cases or 10% highest DJI, or in both highest 10%? I think this need to be made clearer.

lines 272-274:
Coming back to one of my previous point about the correlation patterns of Fig. 4 and 5 (as well as Fig. 6). The way I understand those correlation patterns, they show very low frequency variability in all fields, low SST anomalies leading to persisten high pressure or high temperature anomalies, indicating a stationary or quasi-stationary wave response. Even Fig. 6 seems to indicate very persistent T2m/GPH responses of at least 2 weeks. How does this agree with the findings shown in Fig. 8, which indicate that the relevant connection comes from waves with phase speeds 5-10 m/s. In terms of more general definition of waves, this could be explained by propagating and breaking waves,

transferring wave activity to a lower wavenumber wave or a localized blocking high, but how would this agree with the persistent pattern of Fig. 4?

lines 276:
Is this true, does a DJI triggers QRA events and are not only a necessary condition for QRA. But apart from this, why is this really surprising, since a very local DJI cannot really be linked to the global QRA mechanism.

lines 280-281:
I cannot follow the authors motivation here why they separate the distribution into troughs and ridges occuring over the SST box. What are the reasons for doing this? Following their previous results, cold SST are rather associated with a trough signalin this region. So it would be interesting to know for their separation of how much percentage of the time there is a ridge or trough. From the kind of distribution shown in Fig. 9, this cannot be derived. And wouldn't it also be helpful to show the distribution for the troughs or ridges during this cold SST period instead of the climatology already given in Fig. 8? I assume, the overall distribution looks a bit different, as there are anomalies that are even larger than the underlying climatological distribution (most obvious for w6). Additionally, what do the authors mean by the trough is slowing down the eastward propagating waves. A mechanism of slowing down the wave cannot be derived from Fig. 9, correct? And the trough is part of the wave, so if there is a propagating wave moving across the SST box I would assume that they had more or less the same phase speed if they both are part of the same wave.

line 287:
Maybe I misinterpret the plots, but how can the difference plots between troughs and ridges can be used to identify an imprint of high-amplitude waves? The figure doesn't show the underlying distribution of either case, ride or trough. How can it therefore be used to identify the occurrences of high-amplitude waves?

lines 289-290:
How does this statement makes sense, that the authors are including also medium-scale waves (4-8), whereas Kornhuber et al. (2017) only investigated wave numbers 6-8. So they also included medium-scale waves. I think the authors should be clearer here on what they mean by that.

lines 294-295:
How can Fig. 10 be seen as a measure of wave amplitude? If I understand it correctly, Fig. 10 shows some sort of ridge/trough GPH anomaly for which the amplitude strongly depends on the position of the trough/ridge relative to the SST box. This is not a measure wave amplitude.

lines 303-305:
Isn't this a statement about the general deviation of the climatological pattern: increase

in trough associated with lower SST, decrease of the trough higher SST. I cannot really see the link to the general associated wave patterns, above all not for a link to propagating waves.

2.4 Section 4 (Conclusions)

lines 329-331:
This statement does suggest a strong causal relationship between precursor jet stream and resulting wave pattern. But the authors did not show this causal relationship. The DJ structure could be just an imprint of the underlying wave pattern with a blocking flow feature over Europe, therefore maybe a stronger Jet tilt over the Atlantic with possible increase U values south of it where no blocking flow is present (northward shift of high geopotential height values). I would say that this point was not shown convincingly.

lines 332-334:
As mentioned in a previous comment, I think the authors should clarify some points associated with this conclusion and make their point clearer how they get to this conclusion.

lines 335-337:
The authors deviate in the method/description for QRA to Kornhuber et al. (2017) and their DJ structure is very regional restricted, it is therefore not so simple to make this link or relate their results to this mechanism and derive conclusions from it.

lines 338-341:
As mentioned in my previous notes, I think this needs further clarification and better description.

References

A. Duchez, E. Frajka-Williams, S. A. Josey, D. G. Evans, J. P. Grist, R. Marsh, G. D. McCarthy, B. Sinha, D. I. Berry, and J. J.-M. H. Hirschi. Drivers of exceptionally cold north atlantic ocean temperatures and their link to the 2015 european heat wave. *Environmental Research Letters*, 11(7):074004, jul 2016. doi: 10.1088/1748-9326/11/7/074004. URL `https://doi.org/10.1088%2F1748-9326%2F11%2F7%2F074004`.

K. Kornhuber, V. Petoukhov, S. Petri, S. Rahmstorf, and D. Coumou. Evidence for wave resonance as a key mechanism for generating high-amplitude quasi-stationary waves in boreal summer. *Climate Dynamics*, 49(5):1961–1979, 2017. doi: 10.1007/s00382-016-3399-6. URL `https://doi.org/10.1007/s00382-016-3399-6`.

G. Wolf, D. J. Brayshaw, N. P. Klingaman, and A. Czaja. Quasi-stationary waves and their impact on european weather and extreme events. *Quart. J. Roy. Meteor. Soc.*, 144(717):2431–2448, 2018. doi: 10.1002/qj.3310.