# Review of: "Saharan warm air intrusions in the Western Mediterranean: identification, impacts on temperature extremes and large-scale mechanisms" by Cos et al.

## Summary and General Assessment

Warm air from the Sahara can be advected to the Mediterranean and possibly cause extreme temperatures in the Mediterranean and beyond. Cos et al. use an identification method based on the geopotential thickness between 1000 and 500 hPa, and the mean potential temperature between 925 and 750 hPa to investigate the so-called Saharan warm air intrusions into the Western Mediterranean. Using this method, they calculate a catalogue of past intrusions. The authors perform a clustering of intrusion days based on the distribution of geopotential thickness and mean potential temperature, and find three to five distinct clusters per season, that mainly differ in their geographical location. The transition probabilities between the clusters, the correlations with the teleconnection indices, and the composites of meteorological fields are computed per cluster per season. Extreme temperatures are often observed during intrusion days. Based on the conditional probabilities the authors argue that intrusions have a great impact on heat events, but not such a big contribution, i.e. intrusions often cause heat, but heat is not always caused by intrusions. The composites of the meteorological fields show that the different locations of the clusters are due to different locations of the driving pressure patterns.

The study is motivated by increasing temperatures in the western Mediterranean and the mention of such intrusions in the literature. While the connection between the intrusions and heat is not too surprising, the concept of Saharan warm-air intrusions into the western Mediterranean gives a new perspective on extreme temperature events in the region.

The overall structure of the study is well-organised, however some parts need improvement as they are not so easy to follow. The figures support the presented results mostly well, but their readability could be improved. Below I have major, minor, and figure-specific concerns, comments, and questions that should be addressed.

## Major comments

*Section 3*

In this section, you describe the algorithm used to identify the intrusions, based on the two indicator variables $\Delta GH_{500-1000}$ and $\overline{\theta}_{750-925}$. You refer to the work of Sousa et al. (2019), however, I still lack a motivation/explanation for why these levels are useful to determine the air mass. Especially considering that daily and seasonal variations may change the height of the boundary layer height drastically, for example. Have you looked at how sensitive the results are to your choice of indicators?

*l 91-105*

This paragraph is very difficult to follow. It can be improved by rewriting/rephrasing, but since it is a crucial paragraph to the paper, I consider this a major comment. Let me summarise what I understood:

Monthly climatologies of $\Delta GH$ and $\overline{\theta}$ show that in summer values are maximised over the Sahara, but during the rest of the year they decrease slowly and constantly between 15 and 35°N.

Then, you choose a region as representative for the Sahara. You average over this region, which was chosen so that you meet the values described in literature. I am missing the exact extent of this region in the text.

Then, you calculate 31-day rolling means, averaged over the region, to define thresholds. Why do you chose 31d in particular and how sensitive is your result to this choice?

The air mass is then defined as all grid cells that exceed both threshold values. Does this always yield a coherent air mass as depicted in Fig 1? I would imagine there might be single cells within the air mass that are not identified, or single cells outside of it that are?

Please improve this paragraph by explaining step by step what you did and why.

In lines 109-115 you mention again that in the colder seasons, the values in the Sahara are not as distinct, since there is a constant gradient present. First, this is a little bit repetitive to what was mentioned before (l 92 following). Second, you do not really explain what the consequences of this are. I would imagine that the sensitivity to the threshold values varies a lot across seasons because of this? Does this mean your method works better in summer than in the rest of the year?

*l 107*
"air masses that are displaced from the Sahara". Does that mean the air mass has to cover parts of the Sahara at all times to ensure the origin?
Is the visual inspection necessary for your detection algorithm in general, or was this just a proof of concept in this case? Because if there is doubt about the origin of the air masses, and this step is always necessary, this is a big issue in the applicability of the method to other cases.

*l 116-118*
As far as I understand, the development of the catalogue and analysis of seasonal and inter-annual variations is a crucial part of your study. I was therefore disappointed to be referred to the Supplement for such a central information. Why don't you show an average number of events per month and their spread across years (e.g. as a boxplot). Then your description of the seasonal variability can refer to this plot.

*Fig 3*
From the cross-sections it becomes obvious that the air mass tilts up towards the edges, especially in the North and West. This will not be captured by your methodology, which focuses on the lower levels only. I think, however, this warm, dry air mass might still have considerable effects in those regions. This goes in the same direction as the question about how sensitive your results are to the choice of the indicator variables and their respective levels. Could you discuss this a little bit further?

*l 153 and following and Fig 4*
I was wondering how similar the clusters between the seasons are. The different colour scales make it difficult to compare by eye, but there seem to be great similarities (referring to the figures in the appendix). I would be interested whether one kind of IT appears in all season, or another exists only in summer, or only in MAM and SON, or something like that. This then raises the question, if similar/the same clusters appear across seasons, is it necessary to split the seasons for the clustering?

*l 168 and following* The paragraph starting with "Kendall-tau correlations..." seems to be quite interesting but lacks depth. First, I think there should be a new paragraph, i.e. linebreak. Second, the topic is not introduced at all. It would be helpful to mention the teleconnection indices and then explain that you calculated correlations. Also, you mention you use "some" of the indices but do not motivate which ones. Third, the acronyms used in the equations are not introduced, and the dashes as bullet points can be mistaken for minus signs. Most importantly: I am lacking a discussion of what your resulting numbers mean. I.e. does the first equation imply that IT2 in summer occur more often in a positive AMO phase? Is this reasonable/expected or not? Please elaborate on these results more.

*Sect 5*
The intrusions are identified based on temperature thresholds, the conditional probability $P(heat|intrusion)$ can therefore be expected to be extremely high, at least in the area that is covered by the affected air mass.
Your approach adds to this as it shows any area that is impacted. However, since the information shown in Fig 7 is a combination of many intrusion days, it cannot be distinguished between areas that are hot "by definition" just because the air mass is present, and areas that are hot on an intrusion day, but not covered by the intrusion itself. Could you please elaborate on this. One idea could be to show Fig 7 for cell-wise calculation of the impact: $P(heat\ in\ this\ cell|intrusion\ in\ this\ cell)$ in order to see how big the difference is between this direct vs your "also -remote" impact.

Another thing about this section is the way you look at the contribution. I spent a long time thinking about this. My immediate question was why don't you show $P(heat|no\ intrusion)$, as this would be the intuitive measure to compare $P(heat|intrusion)$ to.
I see how your version of the contribution answers the question "In how many percent of heat days was an intrusion present?" and the intuitive conclusion would be "in this many percent of the heat days, an intrusion was present and therefore played a role, i.e. contributed". However, since the heat cannot be the cause of the intrusion, it is difficult to interpret this in terms of causality.
I would definitely ask you to also show $P(heat|no\ intrusion)$. This shows you how often it was hot, despite no intrusion being present. Contrasting $P(heat|intrusion)$ and $P(heat|no\ intrusion)$ by looking at a difference,

relative risk, odds ratio or log-odds ratio will give insight into how the presence of an intrusion increases or decreases the chance of heat. I.e. when the two probabilities are similar, intrusions do not influence heat very much, while when they are different, you can see the effect of the presence of an intrusion.

*l 259* Here you conclude, that Saharan intrusions in the cold months could be similar to subtropical intrusions. To a reader who is not totally familiar with the literature, it is not obvious why and how Saharan intrusions differ from subtropical intrusions. Please make this more clear.

*l 273* Here you argue that the increasing number of intrusions can be explained by thermodynamic changes as the warming of the troposphere. I cannot follow that argument. An increase in intrusions according to your detection algorithm could either be due to an increase in the synoptic situations causing the advection, or due to a general warming of the atmosphere in the WMed region. The former is not necessarily connected to thermodynamic changes, but rather to dynamic changes in a warming climate. The latter is closely tied to thermodynamics and a warming troposphere, but it might be a false trend due to the detection method. If your WMed is generally warmer and the latitudinal gradient weaker, your indicator variable thresholds might be exceeded without having advection from the Sahara present. Please elaborate on this in more detail.

# Minor Comments

*Throughout the manuscript:*
Please be consistent in your spelling and the usage of the terms for the air mass (Saharan warm air intrusion, warm-air intrusion, warm intrusion, intrusion, air mass, etc.). It will help the readability if you use one clear term throughout the manuscript. Examples, e.g. 50,53,57 or the title of Section 5.
Similarly: Use cross-section (or a precise term of your choosing), as “cut” can be misunderstood (you use cut, vertical section,...).

*l 14* You are referring to changes over what time period? Climatological short(er) term changes?

*l 38* this should be “forest fires in Portugal”

*l 45* gain insight *into*

*l 46* applicable

*l 50*
I don’t understand this sentence. These are the aims of the study, no? What do you mean by “objectives to characterise intrusions into the WMed”?

*l 65-66* This was already mentioned very similarly in l. 43 and following.

*l51/69* Please define the “historical period” in terms of years somewhere.

*l 69* Do you mean daily resolution?

*l 73* Daily maximum temperature at what level?

*l 74* Please define the WMed and Euro-Mediterranean region in terms of lon/lat

*l 77* For better readability I suggest:
“We adapt the definition of Sousa et al. (2019), which takes into account...”

*l 81-82* Please reference equations 1 and 2

*l 90* the gas constant of *dry* air

*l 106* "the detected events"

*l 120* Why are you showing August 2006 here and May 2015 in Fig 1? Are they specifically representative? But why chose 2 different ones?

*l 121*
Sardinia is not at 17°E? Do you mean the air mass penetrates *east* of Sardinia?
also: "intense trough in the North Atlantic"

*l 132* "different expressions" is not a precise term, maybe rather use "properties"?

*l 137* I am not sure whether I understand this clustering correctly. You use $\Delta GH$ and $\overline{\theta}$ in the WMed region on days that were identified as intrusion days? And then you want to cluster those days into different types. Your results seem as if the position of the anomalously high indicator values defines the clusters. But would the clustering also pick up on two clusters that affect the same region but with different magnitude?

*l 151* Looking at the plot for JJA, I cannot really see how you conclude on 3 clusters. The elbow seems to be higher and the high value in the Silhouette score looks a bit like an irregularity?

*l 151* The last sentence in this paragraph "Finally,..." is really inherent to the clustering method and not necessary here.

*l 155* You are referring to Fig S6, not 5 here.

*l 162* I think this should be DJF and SON, you are referring to panels a,c.

*Fig 6* How do you handle the duration of events, where a long period of intrusion is interrupted by a very short period, where the 5% criterion is just not met for a short time? If you handle these as separate events, this would reduce the mean duration very much compared to when you use a filling or filter.

*l 167* from Fig S8, to me it looks more like IT1 and IT2 often transition to IT3, not IT1 to IT2 and IT3? Or are you interpreting 3 in more than 20 as often?

*l 178* This is not the first time you use the term Euro-Mediterranean. Please introduce the acronym and the exact definition when you mention it first (in the data section?)

*l 180* the 90th percentile of what variable (there is simply a word missing I think).

*l 197* What do you mean with "largest contributions *come from* JJA"?

*l 212* subtracting them *from* the intrusion days' composite

*l 217* "for" instead of "from" in the parentheses

*l 223* What do you mean with "the [...] circulation and SLP anomalies are coherent"?

*l 224* What is the tilde over the 3?

l 225 alway*s*

*l 228* negative sea level pressure *anomaly*

*l 243* Your description earlier implied you are detecting intrusions that originate in the entire Sahara, why are you mentioning only the western Sahara here now?

*l 251* You do not really show that the threshold you choose is robust, or how sensitive result are to it's choice.

*l 256* What is a "climatological air mass"? I understand what you are referring to, but this can be misleading. Also "the latitudinal band" is not precise.

*l 262* Please leave out the "backward", as the studies you cite use forward and backward trajectories.

*l 282* What do you mean by "historical contribution"?

*l 292* remove parenthesis, "which" instead of "and" afterwards

## Comments about Figures

While many of your figures are nice in general, all have very small font sizes on the axis labels, the axis ticks, etc. Please change this for better readability. Often the use of shared axes or removal of figure titles that are redundant with the caption can create space.

*MAJOR: Fig S1* Caption is wrong, the red box is not the Saharan region. Also here the figure title is redundant. Axis ticks are not necessary in all rows and columns but only at the bottom and left, which makes space for bigger labels ad better readability. Please mention what contours and shading represent in the caption.

*Fig 1:* The green and cyan lines are not so easily distinguishable, please use additional different linestyle. The black dashed box is supposed to be an example for how big 5% are, this needs to be stated very clearly, otherwise it is not obvious that this is just exemplary.

*Fig 2* I do not see the additional value of this figure. When it is mentioned in line 96 it is rather confusing. As far as I understand the helpful information here is that the averages over the Saharan region are in the high tails of the distribution of the WMed region, I think this can be mentioned without showing the figure. One other question this figure raised for me is the spread in the blue dots in February.

*Fig 3* Add dates to the caption when the exact time of the event is, and what you refer to as before, during and after.

*Fig 4* Caption could be clearer. These are composites of the intrusion days of the respective type in the JJA season? The labels in the contour lines are difficult to read and interpret.

*Fig 5* As IT 1-3 are categorical values, a non-sequential colourmap would make the perception easier. Especially in connection with the figures in the Supplement, where more ITs exist. Then the question is also, relating to an earlier comment of mine, whether all IT1 across seasons are similar to each other or not - using the same colour for all IT1 implies that they are more similar to each other than to others.

*Fig 6* Do I understand it correctly, that the orange bar can never be larger than the blue? Consequently, when I only see orange, this means it is a single event?

*Fig 7/8/9* Can you add the area of influence as a contour. You discuss the relation between the impact and the area of influence (e.g. l 192), but it is difficult to see that by comparing differnt plots on different pages.

*Fig 9* be constistent in use of SLP and PSL, I think it should be wind speed *anomalies* in the caption here. And what arrows are you referring to in panels (a) that are not the quivers? This caption is not easy to follow. It might also help to add numbers to the climatological contours for reference. And the quivers are way too small to see their direction properly.
Is Fig 9 panel (a) the same as Fig (4) panel (b) (except for the significance masking)???