RESPONSE TO REFEREE#2

Attributing trends in extremely hot days to changes in atmospheric dynamics, by J.A. García-Valero et al.

Nat. Hazards Earth Syst. Sci. Discuss

The referee's comments are in black, response to the referee's comments in blue.

General comments

This manuscript provides yet another analysis of weather extremes and their relations to atmospheric circulation over the Iberian Peninsula. Even though the topic is important and worth studying, I have doubts about the scientific quality and formal aspects of the paper. I don't think the topic is original - a similar analysis, only using station temperature series instead of gridded data, was performed by Fernandez-Montes etal. (Atmospheric Research, 127, 154–177, 2013).

Firstly we sincerely appreciate and acknowledge the time devoted by the reviewer. The referee's comments will help to improve the quality of the manuscript. However we have some discrepancies respect some of the comments.

We do not agree with the reviewer's comment referring to the non originality of the paper. The differences with the Fernandez-Montes (2013) are notable. The methodology is quite different as well as the variables used (number of Extreme hot days) at regional scale and large scale (Z500, T850 and SLP). The main objectives also strongly differ. An original idea of this paper is the use of regional series to detect links between regional occurrence of EHD and characteristic atmospheric patterns. Regional information is used for defining first the classes of CTs (i.e., only EHDs are classified). This greatly differs form the general CT classifications used for the study of extremes as is the case of Fernandez-Montes et al. (2013), who made use of generalized classification to define the CTs (i.e., considering all days in the period analysed). Such idea is widely argued in the introduction (lines from 63 to 77). Furthermore, in the paper it is presented a sensitivity analysis of the results to the atmospheric variables used in the definition of the CTs, which can be described by the combination of two atmospheric fields that provide a deeper description of situations leading to this kind of events (only SLP is used in Fernandez-Montes et. al 2013).

We agree with the referee that in Fernandez-Montes et al. (2013) an study of trends of the TX90p (90th percentile) index in summer is presented. This is similar to our analysis of the EHD trends (95th percentile), but they are different in the way they were performed. Both works show regional results, thus in Fernandez the trends are based on a PCA analysis applied over the series of the indices obtained from local stations, analysing trends throughout the time series of the PC linked to a given region. In our paper, the regionalization is based on a clustering technique applied first over the tmax daily series that identifies the regions. Next the regional series of tmax are obtained by averaging the tmax grid-points series belonging to a same region and the 95th percentile of the regional series are calculated for the extreme definition. Finally, EHD trends are calculated for the regional series. On the other hand, the fact of using a more dense dataset show

regions tightly connected to the main topographical and geographical characteristics of the Iberian Peninsula (IP). An example of this is the NE region. In Fernandez-Montes this region (EOF1 Fig4) is formed by the regions NE and E of our regionalization, the first within the Ebro valley (with NW-SE direction) largely affected by NW circulation and the second more protected of this flux by the orography. Similar assessments could be said for other regions. The most detailed regionalization (based on a denser dataset) help to deep on the knowledge of the causes related to this spatial variability, as is the case of the influence of the Circulation Types on such variability.

Specific comments

1. The methodology is so complex that it is hard to follow the message. I am not convinced that the methodology is correct. It seems problematic to study extreme events in data that were smoothed twice: first by interpolating station temperature data into a grid, second by averaging data from grid-points in each of the 8 regions. I suppose that the statistical distributions (PDFs) of the daily data are severely distorted, especially in their tails, i.e. extremes. Indeed, the effects of extremely high temperatures are always local and should be studied at a local scale, be it adverse health effects or environmental effects such as drought or wildfires.

We understand that the methodology could be complex in the sense of the using, two clustering procedures (one for the regional series, and other for the CT classification) , analysis of regional series, procedure to find the better CT classifications, allocation procedure and attribution of trends. However we are sure that the methodology is right.

Probably, if the intention is to explain local extreme variability, this methodology will not be the more suitable, and the use of interpolated dataset would be problematic. However, the goal of the paper is to relate regional variability to dynamics. This will be more clearly stated in some parts of the new version of the manuscript (page 3327 lines 5-9; page 3329 lines 9-13). We agree that averaging grid-point series -for the construction of regional series- leads to a smooth of the PDFs, however high percentiles will continue being related to extreme situations. In fact, a new result that will be included in a new revised version, show the percentage of grid-points belonging to a given region that experiment EHD local occurrences when an EHD regional occur (see response to ref# 1. Thus, in general terms, around 60% of the grid-points have EHD local occurrences, increasing this percentage above 80% when the 90th local percentile is considered. This result reinforces the idea that working with regional series avoids local behaviours less controlled by certain atmospheric dynamical conditions.

2. The classification of circulation types in all days when extremely high temperatures occurred anywhere within Spain is again problematic. It would make more sense to perform the classification separately for each of the regions.

This point was addressed at the beginning of this work. It presents some pros and cons. The main reasons that made us to select the classification of

circulation types in all days when extremely high temperatures occurred any where within Spain were:

- To perform the classification separately for each of the regions would greatly difficult the explanation of such situations: 8 regions multiplied by 4 or 5 CTs each one.
- This would require to identify the similarity between the different centroids assigned to each region and therefore would complicate the interpretation of the different response of the regions to similar CTs.

3. The method used to allocate all days without extremely high temperature to circulation types that occur in days with extremely high temperature is either incorrect or improperly described (section 5.1, table 6). I did not understand the selection of thresholds of correlation and distance. How could two fields of an atmospheric variable that are negatively correlated still be considered similar?

Lines 5 to 8 (page 3337) of the manuscript explain that the most objective election of the distance and correlation thresholds (for each cluster) is considering the lowest and highest values, respectively, being these obtained from the days clustered in a same group. Inside some clusters there are some days that are outliers (this occurs mainly because of the election of the number of CTs during the clustering step). These have been included to avoid subjectivity. Inside the clusters, the differences among the correlation coefficients are larger than for distances. So, as it is explained in line 15 of the page 3337, the final criteria to discern if a day is allocated to a given centroid is taken using the euclidean distance.

4. The Conclusions and Discussions section is barely understandable without reading the rest of the article. As I am not convinced that the used methodology is correct, I am simply not able to identify with the presented conclusions.

As mentioned before, we completely trust on the methodology employed and we are totally sure that it is correct. We think that the conclusions of this work are interesting and have an important additional value to study of extreme events and atmospheric dynamics. On the other hand, we propose a method and give some results on the contribution of changes in circulation to the increase of extremely hot days.

We understand that probably the redaction of the paper could be improved in order to make it more readable. We will strongly effort in doing this for its final publication

– **Technical corrections**

1. The level of English should be improved. Common mistakes include incorrect word order, improper use of words, misspellings, mistaken concord. Some sentences do not make sense. This hampers the overall readability of the manuscript.

The English will be carefully revised. The authors acknowledge the advice and apologize for the readability of the manuscript.

2. Please clarify the number of extremely hot days – is it 863 (p. 3331, l. 4) or 784 (p.3333, l. 6)?

In pag 3331 line 4 we explain that the number of EHD refers to the period 1951-2008. It will be clearer stated in the new version.

3. In Table 2 it is not obvious what do the non-diagonal percentages mean, specifically, to which region the percentage pertains. Example: NE and SW regions share 37% of extremely hot days, but does this mean that 37% of EHDs from the NE region are also hot days in the SW region, or vice versa?

Yes. The percentage is for both regions. This is due to the definition of EHD used. All regions have the same number of EHDs. Therefore the percentage of shared EHDs is the same. The matrix is then symmetric.