RESPONSE TO REFEREE#1

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:

The manuscript describes frequency of Extreme Hot Days (EHD) in Spain for the period 1958-2008. The study is about trend detection and attribution of EHD taking into account the trend of circulation types (CTs). The data used are maximum temperature (Tmax) from Spain02, SLP, T850 and Z500 from ERA40 and ECMWF analysis. One of the contributions of this study is that the trend of EHD can be partially attributed to the trend of Circulation Types (CTs). However, other factors such as global warming, soil-atmosphere feedbacks or surface properties changes may be responsible of EHD tendency (mentioned on the manuscript).

Studies on attribution of extreme events attract attention because of the impacts and the interest on the mechanisms of extreme occurrence, which may present regional differences. Among the causes of extremes are the characteristics or air mass. Therefore, the aims of this study are relevant. However, my impression is that the manuscript is not easy to read because the steps to describe the methods are poorly presented. One of the conclusions is that the method can be used for statistical downscaling, however the study is focussed only on the trend component of EHD and CTs, consequently the utility of the downscaling application is limited for not considering other scales of variability.

We would like to thank the interest of the reviewer in our work as well as the positive view. We strongly acknowledge the time devoted in reviewing the manuscript and the constructive comments that would reflect in an improvement of the manuscript.

The authors agree with the reviewer that probably the complexity of some parts of the paper can difficult the reading. This would be specially related to the extension of the methodology. Two techniques of clustering are used, for regionalization and Circulation Types respectively. Furthermore, a procedure for assigning the rest of days into de CTs and a method for trends attribution are presented. We understand that all these methods, together with the results derived could be to much information and therefore readers could have some difficulty to follow the message. A possibility would be to split the paper. We were thinking about this possibility, but there would be more cons than pros. In a revised version we will try to make clearer the procedure followed.

One of the conclusions is that the method can be used for statistical downscaling, however the study is focussed only on the trend component of EHD and CTs, consequently the utility of the downscaling application is limited for not considering other scales of variability.

We agree with the reviewer's comment. Probably the idea of employing this method for downscaling purposes has not been clearly stated. Obtaining the changes of CTs frequency appearance could serve to estimate EHDs variations (due to atmospheric dynamics). This comment will be added to final section of the manuscript.

Specific comments

Section 2

The large-scale dataset used are from ERA40 (1958-2002) and ECMWF (2003-2008) analysis. My recommendation is to consider same data for the entire period, for example ERA-interim reanalysis.

The main reason of not using ERA-interim is the shorter period available (from 1970s onwards). Since temperature data are available from 1950 to 2008 we chose ERA40 starting at 1958, in order to have as much data as possible in the study till 2002. In addition, analysis from 2003 to 2008 have also been considered to take advantage of the data available in the Spain02 dataset.

In a previous paper (García Valero, et al. 2012) we performed an homogenity analysis of the SLP and Z500 series combining both datasets. To do so, the penalized maximal F test (Wang 2008) was used. Results showed that series were homogeneous. For this work homogeneity of T850 has been also verified. Therefore the use of these two data sets should not introduce any relevant source of potential errors.

Section 3

To obtain the regional series k-means clustering procedure is applied to Tmax daily anomalies. Are the anomalies obtained with respect to the seasonal cycle?

Yes, anomalies are obtained with respect to the seasonal cycle, but k-means was not applied directly to the anomalies, but to the loadings of the retained Empirical Orthogonal Functions (page 3329, line 20). Anomalies are only used for the previous Principal Component Analysis. This aspect will be clarified in a new version of the manuscript.

Why the period 1951-2008 is used for the regionalization if the study is applied for the period 1958-2008? How sensitive is the regionalization to the period used?

We understand that using two different periods could lead to some confusion, but this is clear in the manuscript (page:3328, l.15 and page 3333 l.6). The fact of using two periods is simply for having the better possible statistics. We have tested that the regionalization procedure is quite robust and results are the same using 51 years instead of 58.

Would the regionalization be different when the annual time series is considered instead of summer?

Our interest focus on summer and we did not check the regionalization for annual series. Probably, some differences could appear for other seasons. However, we think that perhaps it wouldn't change greatly since the obtained regions follow are related to the main orographic features of Spain.

I think that the study should indicate the degree of homogeneity of each region according to EHD or to inform about the representativeness of EHD at each region by comparing the correspondence between EHD of the Tmax averaged and the EHD at every point in each region.

Good point. We have obtained the degree of homogeneity of each region according to EHD occurrence. It has been obtained by calculating the percentage of grid-points having a value over its 95th percentile when a regional EHD is reported. This value ranges from 52 (E) to 65% (NWs) depending on the region (60% in average). In addition, if the 90th local percentile is considered, more of 80\% of gridpoints overcomes this percentile (88% in average). These results will be included in a new version of the manuscript.

Figures 2 are too small to distinguish the evolution of the time series. How is the agreement between Tmax and EHD for the different regions?

Yes, we must improve this figure. Regarding the trends (Table 1) Tmax and EHD are in concordance. Most regions show positive and significant trends for both variables. However, there is not a direct correspondence between the hottest year and the one with the highest occurrence of EHDs. This fact will be mentioned in the new version of the manuscript.

Sub-section 3.3

863 EHD are obtained for the period 1951-2008, however the trend and attribution correspond to 1958-2008. The reasons to consider two periods should be clarified because it causes some misunderstanding. Is the percentile computed for the period 1958-2008 or 1951-2008?

We understand that the use of these two periods could cause some misunderstanding. The period 1951-2008 was used for the regionalization of tmax series while the other was used for the characterization of EHD Circulation Types (1958-2008). Percentiles were computed for the period 1951-2008. It is worth to mention that they do not change too much using the shortest period. This would be clarified in the new version of the manuscript.

According to results depicted on Table 1, the significance of the Tmax and EHD trend is similar but I recommend including the confidence interval of the Sen's test.

Confidence intervals of the Sen's test trends have been calculated for both kinds of trends. They will be included in Table 1 of the new version of the paper.

Section 4

The classification of Circulation is obtained for 784 days over a small window that covers the Iberian Peninsula. In my opinion a wider area would represent better the advection effects associated to EHD.

The size and resolution of the clustering window are two parameters that should be adequate to the objectives of the classification. In Garcia Valero et al 2012 a classification of CTs over the Iberian Peninsula was obtained, employing the same window than here. For that work, a sensitivity analysis

of CTs obtained from several windows, all of them centered over the Iberian Peninsula, was carried out. The best results, in relation to the quality of CT classifications as well as to the explained variance of precipitation and temperature by these classifications, were obtained using the window employed here (the smaller among the three windows considered). In our experience, the use of smaller windows avoids the inclusion on the classifications modes of variability that affect remote areas but not the one we are interested in.

The centroids (CTs) represented over a larger window shows coherent synoptic structures. This indicates that the classifications are also representative of a larger area. Logically, they will be better for explaining the variability over the target area (Spain) than over other regions located also inside this larger window. In summary, we think that the centroids shown here represent well the effects of advection on EHD occurrence. Some examples of this are the CT5 of SLP-T850 and SLP-Z500 over the NWw region and CT8 (of the three classifications) over the E region.

Finally, the fact of obtaining CTs of best quality helps to obtain more reliable results in the assignation process and therefore more reliable trends in the frequency of CTs.

My understanding is that the procedures applied to obtain CTs, the evaluation with effectiveness index are not clear. Probably other easier methods would inform about features of air mass characteristics to explain the extremes at each region. This opinion is based on as the authors comment on page 3327 "the atmospheric patterns associated with extreme events represent a small number of days".

In my opinion composite maps of the atmospheric circulation corresponding to the extreme Tmax days at each region would be more informative on the characteristics of air mass that cause the extreme. Then, the correlation maps between the composite circulation and the every day circulation would allow to give the frequency of the circulation type that cause the extreme in each region.

Right. The methodology proposed by the referee would be useful and easier. However, such exercise would not permit to clearly distinguish among the different synoptic situations related to the occurrence of an EHD. This is that we are interested in. Our method is a kind of composite but classifies those patterns related to the EHD at any (of all) region of Spain. This allows getting different CTs that can be of relevance for a given region (ie, more than only one) and also the same CT can be linked to EHD to several regions at the same time (more intense CTs). Furthermore, CTs obtained by classifying the atmospheric circulations are more homogenous than a composite pattern (which may cluster quite different configurations), and the attributting exercise should be more robust

The main objective of the manuscript is to obtain attribution of the EHD trend. For doing this on page 3339 line 14 is written: " a simple attribution model of the EHDs trends to the trends in the CTs frequency appearance is presented". Here the trend of CTs is computed by applying a linear function to the frequency of CTs, while for EHD the Sen's test has been used. Is the approach attribution a comparison of trend obtained using Sen (EHD) and trend using linear method (CTs)?

Trends in the frequency of CTs (line 4 on page 3339) and trends in the regional EHD (lines 21 and 22, page 3331) are obtained (in both cases) using the Sen's test. The method for attributing EHD trends to the dynamics consist of reproducing the observed EHD trends (obtained with the Sen's method) using a linear combination of trends in CTs (combinations of Sen's trends), so we think they are no inconsistency on this.

I suggest to filter out high frequency components of EHD and CTs, and to compare the low frequency component, what would allow investigating the influence of the changes of trend component of CTs on EHD.

We have filter out high frequencies by using running means of 31 years in the seasonal EHD and CT series. Then, Sen's trends have been obtained for both kinds of series and the attribution equation have been used. The results show attributions slightly lower than those obtained without filtering. This is mainly due to the larger EHD trends observed at low frequencies than those detected at high frequencies. The percentage of attributions varies from 20% (SW, Cs and NWs regions) to about 40% (E and N regions). This would be considered to be included in the new version of the manuscript.

Technical corrections

On page 3327, line 11 is written "the study is applied over the IP". The authors should have considered that the maximum temperature used corresponds only to Spain. The title should include the area of the study.

We will take into account this recommendation in the new version of the manuscript.