Summer temperature extremes in northeastern Spain: spatial regionalization and links to atmospheric circulation (1960-2006)

Ahmed El Kenawy[1,2], Juan I. López-Moreno[1] and Sergio M. Vicente-Serrano[1]

[1] Instituto Pirenaico de Ecologia- CSIC, Campus de Aula Dei, P.O. Box 13034, Zaragoza, 50059, Spain.

[2] Department of Geography, Mansoura University, Mansoura, Egypt.

(kenawy@ipe.csic.es)

Abstract

A procedure for classifying daily summer temperature extremes in northeastern Spain into homogenous regions has been presented and evaluated. This procedure employed daily temperature series from a dense network of 128 weather stations spanning the period from 1960 to 2006. Characteristics of temperature extremes included temperature frequency (e.g. warm days), intensity (e.g. warmest day), and duration (e.g. maximum length of hot spell). Following the results of the principal components analysis (PCA) and the Ward´s method of clustering, the study area was divided into four homogenous sub-regions in terms of both the geographic and climatic meanings: the Mediterranean region, the mainland and the Cantabrian region, the moderately elevated areas westward and southward, and the mountainous region. Based on an internal cluster validation measure (Silhouette width), the quality of clustering was evaluated and ensured. The temporal evolution of the long-term (1960-2006) temperature extremes clearly showed different behavior amongst these sub-regions. The Mediterranean and the highly elevated regions revealed the strongest signals

in both day-time and night-time extremes. For mainland areas, considerable differences in the behavior of the day-time and night-time temperature extremes were evident. The influence of atmospheric circulation on spatial and temporal variability of temperature extremes was also explored. The variability of summer temperature extremes in NE Spain appears to be mainly driven by the Scandinavian (SCA), the Western Mediterranean Oscillation (WeMO), and the East Atlantic (EA) patterns, with a tendency toward increasing during the positive (negative) phases of the EA (WeMO and SCA) circulation modes. In such a region with complex geography and climate, regionalization of summer temperature extremes can be advantageous for extracting finer scale information, which may prove useful for the vulnerability assessments and the development of local adaptation strategies in areas such as health, ecosystems and agriculture.

Key words: Regionalization, principal component analysis (PCA), cluster analysis (CA), clustering validation, temperature extremes, atmospheric circulation, northeastern Spain.

1. Introduction

Temperature is a climate variable of high importance from the view of various disciplines including hydrological, agricultural and environmental applications. Spatial and temporal variability of temperature have increasingly been the focus of climatic research worldwide (e.g. Jones et al., 1999, Folland et al., 2001). Numerous studies have also observed significant changes in the intensity and frequency of temperature extremes across

44 many regions of the world (e.g. Karl et al., 1995, Alexander et al., 2006). Nevertheless,
45 capturing spatial variability of temperature remains a challengeable task, particularly over
46 areas of complex orography or low density of observatories. In this sense, spatial
47 regionalization can be an effective tool not only to obtain a detailed knowledge on spatial
48 variability of temperature at local/regional scale, but also to identify driving factors that
49 may influence these spatial variations. Multivariate statistical techniques (e.g. principal
50 component analysis [PCA], cluster analysis [CA], and discriminant analysis [DA]) have
51 increasingly gained acceptance to identify homogenous regions in many fields, such as
52 hydrology (e.g. Love et al., 2004), geology (e.g. Reimann et al., 2002), forestry (e.g.
53 Schulte and Mrosek, 2006), soil sciences (e.g. Young and Hammer, 2000), and ecology
54 (e.g. Camiz and Pillar, 2007). In atmospheric research, these techniques have also been
55 carried out at different spatial and temporal scales ranging from regionalization of a
56 specific climate variable such as precipitation (e.g. Wolting et al., 2000), temperature (e.g.
57 Coronato and Bisigato, 1998), and evapotranspiration (e.g. Mohan and Arumugam, 1996)
58 to synoptic classification including air masses (e.g. Bejaran and Camilloni, 2003) and large-
59 scale atmospheric circulation and weather types (e.g. Romero et al., 1999, Esteban et al.,
60 2006). Among studies focusing on severe weather events, Holt (1999), for example,
61 employed PCA to obtain a synoptic classification of extreme surge events along the
62 Western Europe coasts
63

64 The climate of the Iberian Peninsula is largely influenced by the complex interactions of the
65 Mediterranean and the Atlantic configurations. The complex land–sea interactions along
66 with latitude, altitude and orography variations make the local climate very changeable.
67 Topography and continentality may produce more or less complicated patterns of
68 temperature extremes and therefore the impact of extreme events can largely depend on
69 these local and regional conditions. Accordingly, it seems important to capture dominant
70 spatial modes of extreme temperatures at the sub-regional scale. In this context,
71 homogenous climate regions have often been a topic of interest in the Iberian Peninsula.
72 These efforts have mainly been promoted by the increasing concern to understand spatial
73 variability of climate and its physical causes. However, most of these studies gave much
74 more concern to obtain regionalization of precipitation data at both coarse and fine
75 resolution (e.g. Martin-Vide and Gomez, 1999, Muñoz-Diaz and Rodrigo, 2004, Vicente-
76 Serrano, 2006). For instance, Martin-Vide and Gomez (1999) classified Spain into distinct
77 regions based on the length of dry spells over the period 1951-1990. Similarly, Muñoz-Diaz
78 and Rodrigo (2004) divided Spain into relatively homogenous pluviometric regions using
79 seasonal time series covering the period from 1912 to 2000. More recently, Vicente-
80 Serrano (2006) obtained a regionalization of drought in the Iberian Peninsula employing a
81 monthly precipitation dataset. In contrast to precipitation, there has been little work on
82 identification of homogenous regions of either temperature means or extremes, particularly
83 at fine spatial resolution. A recent study by Serra et al. (2010) employed the PCA and CA
84 techniques to daily temperature data in Catalonia (NE Spain) to identify reasonable

85 homogenous regions. According to this study, the spatial regionalization was performed
86 using extreme normalized residuals defined as deviations from the long-term (1950-2004)
87 daily maximum and minimum temperature over the region.
88
89 Recently, numerous years have been identified as anomalously warm with record-breaking
90 temperatures over Europe (e.g. 1995, 1998, 2003, 2005 and 2010) (Ciais et al., 2005,
91 Barriopedro et al., 2011). The environmental, economical and societal impacts of extreme
92 temperatures on physical (e.g. ecology, forest fire, and hydrology) and human
93 environments (e.g. agriculture, mortality, tourism, and energy demand) are more
94 pronounced during these warm events. A representative example over Western Europe is
95 the unrelenting 2003 heat wave causing more than 30,000 human mortalities (at least
96 15,000 in France) (WHO, 2003). Over the study domain, these anomalous summers have
97 also been more frequent during the last decades. One clear example is the Pyrenees where
98 summer maximum temperature reached its maximum value on record in 2003, exceeding
99 the 35°C threshold. A recent study by El Kenawy et al. (2011a) reported that the mean
100 summer surface air temperature in northeastern Spain has increased by about 1.9ºC since
101 1960, with a warming rate of about 0.41°C per decade. Recalling that there are only very
102 few assessments of spatial variability of seasonal temperature extremes over northeastern
103 Iberia, application of spatial regionalization to summertime extremes can be of particular
104 interest for numerous fields, such as agriculture, human health, urban development and
105 planning, and water resources management. With respect to natural environment, these

106 events can also have severe impacts on forestry across the study area as its unique flora,
107 fauna and ecosystems are vulnerable to even slight variations in climate (Pasho et al.,
108 2011a,b). The regionalization of extreme events could thus increase our understanding of
109 the possible roles that mesoscale and regional-scale climate processes can play in this
110 diversity. Also, this regionalization could facilitate the development of appropriate
111 adaptation strategies to cope with changes in these heat events. The outcome of any
112 adaption policy is inherently maximized when considered at local and regional scales.
113
114 Atmospheric circulation is considered one of the main physical processes responsible for
115 long term interannual and interseasonal variability of climate, particularly at regional scale
116 (Vicente-Serrano et al., 2009). Many studies have investigated the linkages between
117 teleconnection patterns and climate variability in the Iberian Peninsula (e.g. Romero et al.,
118 1999, Rodriguez-Puebla et al., 2001; Vicente-Serrano et al., 2009). Nevertheless, most of
119 these works focused on influences of circulation patterns on spatial and temporal variability
120 of precipitation. The possible links between the general atmospheric circulation patterns
121 and daily temperature series are still lacking and worth investigating at both coarse and fine
122 spatial scales. Recently, Rodriguez-Puebla et al. (2009) linked spatial and temporal changes
123 on warm days and cold nights with large-scale atmospheric circulation for the entire Iberian
124 Peninsula on the basis of 26 daily time series.
125

126 The main objectives of the present study are (1) to divide northeastern Spain into regions as
127 homogenous as possible based on the information derived from 14 summertime extreme
128 temperature indices using PCA and CA, and (2) to assess the connections between spatial
129 and temporal variability of temperature extremes in the defined sub-regions on one hand
130 and the main modes of atmospheric circulation over western Europe and the Mediterranean
131 region on the other hand. Classification of temperature extreme events into objectively
132 defined clusters in such a transitional Mediterranean/Atlantic area and the investigation of
133 their connection to large-scale atmospheric circulation could improve our understanding of
134 the regional climate variability and its driving forces.
135

136 ## 2. Study area

137 The study area is located in NE Spain with an area of approximately 160,000 km^2 (Figure
138 1). The territory is comprised between latitudes of 39° 43' N and 43° 29' N and longitudes
139 of 05° 01' W and 03° 17' E. It is enclosed by the Cantabrian Sea (Atlantic Ocean) in the
140 northwest and the Mediterranean Sea in the east. The study domain is characterized by
141 terrain complexity since it is enclosed by the Pyrenees system in the north (>3000 m a.s.l.),
142 the Cantabrian system northwest and the Iberian system southward. The central part is
143 mainly configured by the Ebro valley (200-650 m a.s.l.), with some peaks over 1000 m.
144 The area is mainly affected by the Atlantic and the Mediterranean disturbances.
145 Accordingly, it yields very considerable spatial and temporal variability of temperature.
146

3. Dataset description and methodology

3.1. Dataset:

The current analysis is based on daily maximum and minimum temperature of a dataset of 128 observatories spanning the period from 1960 to 2006. The raw data were provided by *the Spanish Meteorological Agency (AMET)*. The original data were subjected to a vigorous quality control procedure to eliminate any spurious values and then a reconstruction scheme to fill in missing values by linear regression. We examined the station histories and evaluated any potential discontinuities caused by station relocation. A detailed discussion of the methodology behind this dataset is presented in El Kenawy et al. (2011b). In this work, the stations were selected on the basis of data completeness, length and homogeneity. Spatial distribution of the observatories is shown in Figure 1. In general, the observatories are evenly distributed across the study domain. This sounds critical in spatial regionalization studies since low density of observatories could maximize the bias errors affecting spatial patterns (Karl and Koscielny, 1982). The high density of this dataset is advantageous to capture the variations of temperature extremes over short distances, particularly in areas of complex orography. The vertical distribution of the observatories is also quite satisfactory since 24 (18.8 %) of them are located above 900 m a.s.l.

3.2. Definition of extreme events:

Extreme weather refers to infrequent, but significant, departures from the normal weather conditions. Indeed, climate change is expected to impact the frequency, intensity and

duration of extreme events. In this work we selected 14 indices/variables which likely represent most of the variability of summer temperature extremes. These indices potentially include many aspects of changing climate conditions including frequency, intensity and persistence. It is believed that inclusion of different definitions of temperature extremes can make the climate classification more representative and objective, particularly for capturing micro-climatic characteristics. In addition, the information derived from different indices could be valuable for climate change impact studies. Indeed, the influence of anthropogenic climate variability on changes in a single extreme index (e.g. warm nights) cannot be clearly distinguished in space. This is simply because any given characteristic of an extreme event still has a probability to be influenced by natural variability. Moreover, inclusion of only one feature could be spatially biased or might have a limited effect on both natural and human environments. For example, summer days (SU25) may not be viewed as extremes over lowlands with gentle topography. For these reasons, using various characteristics of extreme events could improve our understanding of their changing likelihood under global warming. The different aspects and properties of extreme temperature on the one hand, and the various applications of these characteristics for climate impact and assessment studies on the other hand, can be brought together best if a statistical scheme including the various aspects of extremes is used. A list of the indices used in this study is given in Table 1.

189 Overall, the indices were retrieved from the daily dataset corresponding to summers over
190 the period from 1960 to 2006. In this study, summer season is defined as June to August.
191 Much of the warm extremes typically occur during these June–August peak season months,
192 for which the amount and distribution of temperature generally determine the overall
193 severity during this season. An exploratory analysis of the local distribution of daily
194 maximum and minimum temperatures suggested that little number of warm extremes (e.g.
195 hot days) may remain for the late spring (May) or early autumn (September). Thus,
196 evaluating the behavior of warm extremes for June–August is employed to assess extreme
197 temperature variations.
198
199 In particular, the indices were defined in various ways varying from arbitrary-based
200 approaches such as summer days (SU25) to statistically-based threshold approaches such as
201 warm days (TX90p) and warm nights (TN90p). For instance, the number of summer days
202 (SU25) is defined as the total number of days per summer with maximum temperature (T
203 max) over 25 °C. Overall, using a fixed threshold is not superior in the study domain where
204 the variability of daily temperature is high. On the other hand, the definition of extreme
205 events based on percentiles is advantageous because it allows direct comparison among
206 regions with different climates. In this work, the percentiles were calculated based on the
207 local daily distribution of maximum and minimum temperatures during summers of the
208 period 1960-2006. For instance, in order to reveal the behavior of frequency change, warm
209 days (TX90p) index was defined based on days in which maximum temperature exceeding

the climatological 90th percentile of the daily temperature distribution at each location from 1960 to 2006. In order to measures the severity of heat stress during summer season, the maximum duration of consecutive hot days was considered using a centered moving window of consecutive (*n*) days exceeding the 90th percentile of daily maximum distribution. For each observatory, the temporal evolution of a 47-yrs (1960-2006) time series for each particular index was assessed using the ordinary least squares method. The magnitude of the trend was obtained from the slope of the linear regression.

3.3. Statistical analysis:

In this section, a detailed description of a two-step statistical procedure to obtain homogenous regions of extreme temperature is provided. First, the climatic information as summarized by the trends of different temperature indices is reduced using factor analysis. The second step examined the spatial regionalization of the scores of the retained factors gained in the first step. This two-step procedure has previously been recommended by several climate regionalization studies including, among others, Baeryswil and Rebetez (1997), Romero et al. (1999) and Papadimas et al. (2011). As has been proposed by these studies, it was necessary to apply factor analysis prior to cluster analysis to minimize autocovariance in the dataset.

In order to reduce the multidimensionality associated with the large number of input variables (i.e. 14 variables by 128 cases), PCA (S-mode) is applied for the magnitude of

231 trends in the defined extreme indices. This analysis is a statistical method commonly used
232 in climate research to analyze large multivariate datasets and derive the main spatial
233 patterns of climate variables. Herein, the raw data were standardized by their mean and
234 standard deviation in order to facilitate comparison between input variables of different
235 scale units. Also, recalling that temperature parameters are standardized and normally
236 distributed, as being tested in our dataset; a data matrix based on the inter-station
237 correlation was obtained to characterize the levels of the relationships among the input
238 variables. The correlation matrix is favored compared to the covariance matrix because it
239 gives equal weights to all years involved in the analysis. This seems to be important
240 because the covariance matrix is expected to give more weighting to the warmer years
241 (events) during recent decades. Selection of appropriate PCs that can adequately represent
242 most information of the original dataset is another important decision in PCA. According to
243 the Kaiser criterion, only PCs with eigenvalues greater or equal to 1.0 were extracted. The
244 retained PCs were then rotated by means of the varimax orthogonal technique in order to
245 reduce data dimensionality. This procedure facilitates spatial reasoning of the PCs that later
246 became important to objectively cluster the variables. Then, the observatories were
247 assigned to factors based on their maximum factor loadings. In this regard, it is important to
248 indicate that the PCA results were validated to ensure the stability of the obtained factors
249 and their scores. Given that this work is based on employing a set of indices (variables) for
250 regionalization purpose, it was necessary to ensure that the obtained results are not variable
251 dependent. For this reason, we performed a sensitivity analysis of the input dataset to see

252 whether the obtained factors and their explained variance will change when one or more
253 variable are removed.
254
255 Cluster analysis is a multivariate technique commonly used to classify observations into
256 groups according to similarity in their quantitative characteristics (DeGaetano, 2001). The
257 outcome of clustering analysis is heavily dependent on the pre-processing procedures, such
258 as selection of a "best" clustering algorithm, similarity function, number of clusters, and
259 weights of input variables. A comprehensive review of the cluster analysis algorithms is
260 given by Gong and Richman (1995). In this work, the standardized PC scores were used as
261 a basis for CA procedure in order to detect the best spatial classification of summer
262 temperature extremes. There have been numerous studies comparing the performance of
263 hierarchical and non-hierarchical algorithms of clustering (e.g. Kalkstein et al., 1987).
264 Since the accuracy of non-hierarchal methods (e.g. k-means) is very sensitive to the
265 selection of the centroids points and also the order in which data are processed; our
266 preference was given to the hierarchical techniques. These methods are particularly
267 preferred when a priori knowledge of data structure is inadequate. In this sense, the
268 hierarchical Ward´s method was chosen. In terms of their statistical accuracy, the Ward
269 algorithm has been found superior to other methods in various climatic applications (e.g.
270 Romero et al., 1999, DeGaetano, 2001). This algorithm is an ANOVA-type approach which
271 explicitly minimizes the within-group similarity and maximizes the between-group
272 similarity (Bonell and Summer, 1992). Another important decision in the clustering

procedure was to define the “accurate” number of clusters to be retained. A greater number of clusters are not desirable for practical uses and may introduce noisy patterns that could not be justified in terms of the climatological reasoning. For this reason, a consideration of only spatially prolonged patterns that could have the most significant environmental, economic and social impacts is more preferred. Other factors than these spatially large-scale patterns are likely to reveal very local modes of extreme events. In the same sense, an inadequate number of clusters may cause missing of valuable information. Indeed, there seems no uniform criterion to decide on the number of clusters. Milligan and Cooper (1985) introduced 30 different statistics to define a relatively appropriate number of clusters. To ensure the reliability of the defined number of clusters, multiple statistics are desirable to check for agreement between results. In our case, two statistics were used: the agglomeration coefficient of squared Euclidean distance and the Wilk´ Lambda test. The agglomeration coefficient reveals change in squared Euclidean distance between the two most dissimilar observatories in combined clusters at each stage. A large increase in agglomeration coefficient shows the optimal number of clusters as it indicates that inhomogeneous clusters are being merged. On the other hand, the Wilk´ Lambda is a multivariate statistic of variance, defined as the ratio of the within group variance to the total variance (Everitt and Dunn, 1991). Given that lower values of this statistic suggest that the source of total variation in the dataset is due to the between-groups variance, it can therefore be employed to test the best cluster solution.

With hierarchical methods, clustering algorithms may incorrectly aggregate some observations into misclassified clusters. In fact, most clustering algorithms do not provide estimates of clusters significance. For example, the Ward algorithm allows crisp clustering in which each observation is assigned to a unique partition (cluster) and cannot be reassigned to alternative cluster whenever more appropriate (Gong and Richman, 1995). Given that our clustering procedure is unsupervised as the number of clusters is defined objectively, it is important to verify its goodness of fit. In this work, the Silhouette width index, mainly based on proximity matrix, was chosen to evaluate the homogeneity of the final clusters. This index (Rousseeuw, 1987) is defined as:

$$S_i = \frac{b_i - a_i}{\max\{a_i, b_i\}}$$

where a_i is the intracluster distance (average distance of the observatory (i) to all observatories in the same cluster), while b_i refers to the intercluster distance (the average distance of the same observatory to observatories in another cluster). The Silhouette average width was simply calculated by averaging the coefficients of observatories belonging to each independent clustering. This index is favorable compared to other clustering validity measures (e.g. Dunn index and Davies-Bouldin index) for two reasons. First, it does not only indicate validity of the entire clustering, but it also provides a measure of the extent to which each individual observatory closely matches its cluster. Second, this statistic is robust to outliers, noisy observatories and number of clusters (Rousseeuw, 1987). According to this measure, it is possible to re-assign an observatory to

an alternative cluster to satisfy the homogeneity conditions. Overall, values of the Silhouette width are limited to the interval [-1, 1]. Values close to 1 correspond to clusters that are compact and well separated from other clusters. When the intercluster distance of the observatory is equal or less than the intracluster distance, a decision was objectively made to reassign these observatories to another cluster with the lowest intercluster distance. Afterwards, the silhouette coefficient was re-calculated. This process was applied iteratively until all observatories have a positive value of the silhouette coefficient.

3.4. Temporal evolution and links to atmospheric circulation

To account for the temporal evolution of temperature extremes time series in the established clusters, we employed a regional series for each cluster. This procedure is twofold. First, it helps to compare trends in temperature extremes from both spatial and temporal perspectives. Second, it allows spatial patterns to be linked with underlying mechanisms, such as atmospheric circulation. To provide a more proper definition of the regional series, we used a weighted average of all observatories belonging to a particular cluster based on their Silhouette coefficients, giving larger weights to observatories that are close in their similarity. Simply, high value of the Silhouette coefficient indicates that the observatory can better reflect the overall characteristics of its cluster (sub-region). This procedure also overcomes the problem of discontinuities at cluster boundaries, particularly when the coefficient is close to 0. Linear trends in the defined indices time series were analyzed for each particular sub-region by means of the ordinary least squares method

(OLS) and the statistical significance was assessed using the Mann-Kendall statistic at the 95% level of significance. This statistic is robust to outliers and does not assume an underlying probability distribution of the data time series.

To explore the extent to which atmospheric circulation determines spatial patterns of summer temperature extremes, three different northern hemisphere atmospheric circulation indices were obtained from the Climate Prediction Center, NOAA/NCEP, USA (http://www.cpc.noaa.gov/data/teledoc/telecontents.shtml) between 1960 and 2006. These indices included: the East Atlantic (EA), the Scandinavian (SCA), and the Eastern Atlantic-Western Russian (EAWR) patterns. In addition, the North Atlantic Oscillation (NAO) index provided by the Climate Research Unit of the University of East Anglia, the UK (http://www.cru.uea.ac.uk/cru/data/nao/) was considered. The NAO has a north-south dipole, with one centre over Iceland and the other with opposite sign over the mid-latitudinal Atlantic (Hurrell, 1995). On the other hand, the EAWR pattern has two main centers: the first is located in the Caspian Sea and the second is found in Western Europe. The EA resembles the NAO in terms of the geographical domain. However, it has a more southward shift toward low latitudes (Canary Islands: 25ºN, 25ºW). The SCA is a dipole with a main center over Scandinavia and minor centers across Western Europe. Recalling that variability of temperature across the Mediterranean can be linked to regional atmospheric regimes, such as the Western Mediterranean Oscillation (WeMO) and the Mediterranean Oscillation (MO) indices as noted by many authors (e.g. Martin-Vide and Lopez-Bustins, 2006, Vicente-Serrano et al., 2009), these two later indices were also

356 considered in our analysis. The WeMO index was defined by Martin-Vide and Lopez-
357 Bustins (2006) as the difference in surface atmospheric pressure between Padua, Italy (45º
358 24'N, 11º 47'E) and San Fernando, Spain (36º 17'N, 06º 07'W). Palutikof (2003)
359 calculated the MO index as the difference in the SLP anomalies between Gibraltar (Spain)
360 and Lod (Israel). Overall, the selected indices are among the most dominant atmospheric
361 configurations in the western Mediterranean and the North Atlantic basin. They
362 summarized a wide variety of flows that affect climate variability in the Iberian Peninsula
363 during summertime. The Pearson correlation coefficient (*r*) was simply computed between
364 the regional series obtained for each sub-region and the time series of the atmospheric
365 circulation at *p* value < 0.05. However, it is noteworthy that the time series were detrended
366 prior to compute correlation calculation to remove the possible influence of the time series
367 monotonic trend and interannual variability on the strength and significance of correlation.
368 Then, the series were standardized for the base period 1960-12006 to zero mean and unit
369 variance. This was principally performed before calculating the correlation to confirm all
370 the time series are equally weighted.
371
372 For the leading circulation modes which showed significant influence on summertime
373 temperature variations, we performed the canonical correlation to assess the relationship
374 between SLP anomalies (independent variables) and summertime temperature (dependent
375 variables). The daily SLP data are provided by the NCEP/NCAR reanalysis data compiled
376 by the NOAA/OAR/ESRL PSD (http://www.nws,noaa.gov). Summertime (JJA) datasets of

SLP and temperature corresponding to the period from 1960 to 2006 were used for this analysis. Factor analysis was first applied on both fields to reduce the dimensionality of the original data sets. Then, canonical correlation analysis was performed on factors that explained more than 5% of total variance for each dataset. More details on the canonical correlation theory are given by Dillon and Goldstein (1984). Although the study domain has a relatively small area (approximately 160,000 km^2), a large window (30ºW-30ºE and 20ºN-80ºN) has been considered in this analysis. This spatial domain is large enough to encompass all regions that include forcings and circulations which directly affect temperature climate over the study domain. For each particular circulation mode, the statistically significant canonical functions that explain a large proportion of summertime temperature variance were retrieved.

4. Results and discussion

4.1. PCA results:

Kaiser supports the use of the PCA to reduce data dimensionality when the KMO value is greater than 0.5 (Norusis, 1988). The overall KMO statistic among the input variables reached 0.81 (at the 99 % significance level), suggesting a high level of dependence among those variables. The Scree plot showed that the magnitude of the eigenvalues dropped sharply after a stage of 3 PCs. Those three PCs explained together 84.32 % of the total variance in the original data. Due to the small-explained variance (15.68 %) of the high-

398 order modes and their small differences, they are not considered in this work. This amount
399 of variance can be originated to very local factors that are not physically meaningful and
400 thereby difficult to interpret. The loadings of the PC1 (29.46 % of the variance) reveal
401 aspects on the spatial variability of the night-time summer temperature (e.g. Min_summer,
402 TNn, TN90p, and TNx). On the other hand, the PC2 (27.43 % of the variance) has the
403 variables that are mostly correlated with the day-time summer temperature (e.g.
404 Max_summer, TXx, TX90p, SU25 and WD). The first component can therefore be seen as
405 a measure of “minimum” summer temperature extremes; meanwhile the second component
406 can be interpreted as a measure of “maximum” summer temperature extremes. The PC3
407 (27.34 % of the variance) does not distinctly explain specific pattern since it combined
408 effects of both maximum and minimum temperature (e.g. DTR and Tsums).
409
410 **4.2. Clustering results:**

411 In order to group objectively those observatories presenting a similar temporal evolution of
412 summertime extremes, the dominant distribution patterns are identified by applying CA to
413 the retained factor scores. The results of both the agglomeration coefficient and the Wilks´
414 Lambda statistic suggest 4 solutions as the optimum number of clusters. Thus, we finally
415 decided to partition the clusters at 4 groups. Using more than 4 clusters may not be useful
416 because no meaningful spatial patterns can be explored. These small-scale patterns are
417 often attributed to local factors that are difficult to interpret. In topographically complicated
418 areas, differences in land surface characteristics such as vegetation canopy and surface

albedo can induce local changes in surface heat exchanges with the atmosphere and in turn cause local disturbances in the dominant patterns of extreme events. Overall, the resulting number of clusters represents the range of climatological conditions, which can conventionally be accepted over the region. In other words, the spatial modes suggested by the CA are thought of as representing the main modes of the regional climate regimes according to our previous knowledge. Therefore, those four sub-regions that exhibited similar characteristics of summer temperature extremes were identified. Hereafter we referred to these sub-regions as: CL1, CL2, CL3 and CL4.

Application of the silhouette coefficient to test homogeneity of the obtained clusters shows that the clustering partitions are not completely homogenous. It was found that 16 (12.5%) observatories were assigned to “inappropriate” clusters. Figure 2 compares the Silhouette width for the defined clusters before and after validation. In general, the results indicate how the Silhouette width coefficient has improved after validating the clustering, specifically for CL1 and CL4. The coefficient has risen from 0.54 (0.23) to 0.62 (0.39) for CL1 (CL4). This can be seen as an indicator of higher between-group variation and lower within-group variation. In other words, the observatories belonging to each cluster are well-separated from other clusters and also compacted within their clusters.

Figure 3 shows the spatial distribution of the observatories corresponding to the delineated clusters. As noted, the number of stations varies considerably among clusters. CL2 represents the densest cluster (45, 3% of observatories) with broad spatial distribution,

440 followed by CL1 (25 %), CL3 (21.1%), and CL4 (8.6%). In general, there are remarkable
441 geographic and climatic contrasts between the defined sub-regions. They have a
442 geographical feature, with relatively clear physiographic boundaries. Presumably, the
443 defined clusters show marked distinction between inland and coastland regions as well as
444 lowlands and highly elevated areas. In few cases, there is a low consistency among close
445 observatories. This can probably be due to the joint effect of local topography and synoptic
446 conditions. As illustrated in Figure 4, it can be clearly seen that elevation and distance to
447 the sea significantly contribute to the spatial variability of summer temperature extremes in
448 the study domain. As depicted, spatial variations of summertime extremes match well with
449 changes in elevation. This is particularly the case for observatories belonging to CL3 and
450 CL4 whose mean elevation is 773.1 and 1101.4 m respectively. On the other hand, CL1
451 suggests a joint effect of topography and land-water interaction. In addition to their
452 proximity to the sea with average distance of approximately 14.8 km, those observatories
453 are generally located at lowlands with mean elevation of 174.9 m. With respect to
454 continentality effects, both CL3 and CL4 exhibited clear continental influences (average
455 distance to the sea =143.3 km and 122.6 km respectively), compared with the more
456 continental CL1 and CL2. This finding can be clearly seen in Figure 3, where the
457 orientation of the clusters is likely controlled by continentality and orography. For instance,
458 CL1 is generally meridian along the Mediterranean coast, whereas CL4 has an east-west
459 orientation along the Pyrenees Mountains. Spatially, CL1 is mainly situated along the
460 coastline of the Mediterranean Sea capturing possible maritime influences of the

461 Mediterranean, whereas CL2 is located close to the Atlantic Ocean and penetrates eastward
462 through the Ebro valley in central proportions. This clustering mostly includes leeward sites
463 of the main mountain regimes (i.e. the Pyrenees system in the north and the Cantabrian
464 system in the west). On the other hand, the CL3 corresponds to stations located in the
465 moderately elevated areas southward and westward. This partition of stations is
466 characterized by a relatively low maximum and minimum temperature as a consequence of
467 the local orographic effects. A cluster encompassing mainly the Pyrenees Mountains,
468 together with some scattered sites mostly located in the highly elevated areas in the south,
469 was also distinctly identified (CL4). Although the network density in these areas is
470 generally irregular given that only 14.4 % of the observatories are located above 1000m,
471 compared with other data-rich regions (e.g. the coastal and mainland areas), the clustering
472 procedure skillfully captured the variability of temperature extremes at these high
473 elevations. This can probably be explained by the free-air advection at the mountains
474 summits and along the free-drainage slopes. Pepin and Lundquist (2008) confirmed this
475 finding for elevated sites with annual 0º isotherm across the globe, suggesting that inter-site
476 variance of temperature at those sites is expected to be lower than moderate and low
477 elevation sites, as a consequence of the weak influence of local factors near surface such as
478 land use changes.
479
480 **4.3. Temporal evolution of summertime extreme indices:**
481

Table 2 summarizes the linear trends in summertime temperature extremes time series for the established 4 sub-regions over the period from 1960 to 2006. In general, the overall tendency in temperature extremes is toward warming for all sub-regions. However, this warming has a spatial component. Figure 5 illustrates the temporal evolution of a set of temperature extremes time series, as indicative examples. As can be seen, the strongest signals were found in the most elevated areas (CL4) and in the Mediterranean (CL1) for both high and low temperature indices. For instance, warmest day (WD) has significantly increased at a rate of 0.71º and 0.44ºC per decade for CL1 and CL4 respectively. Over the period 1960-2006, the percentage of warm nights (TN90p) per summer has also shown uptrend, with an increase of 3.88 and 2.86 % per decade for CL1 and CL4 respectively. This suggests that orography and distance to the Mediterranean Sea play a key role in the temporal evolution of summer extremes in NE Spain. On the other hand, CL2 (mainly located in the Ebro valley and along the Cantabrian sea) only showed statistically significant warming trend in night-time extremes (e.g. TNn, TN10p and TNx), whilst trends in day-time extremes were generally insignificant at the 95% level. Contrarily, day-time extremes exhibited a remarkable significant trend in CL3 (denoted to moderately elevated areas), while night-time extremes have no trend.

The temporal evolution of summer extremes along CL1 coincides with the observed changes found across many Mediterranean areas (e.g. Klein Tank and Können, 2003, Kostopoulou and Jones, 2005, Hertig et al., 2010). For example, Klein Tank and Können

503 (2003) found upward trend in warm temperature extremes over the Mediterranean region
504 from 1976 to 1999. For the western Mediterranean, Hertig et al. (2010) recently observed a
505 strong warming trend in summer maximum temperature, more intense over the Iberian
506 Peninsula. The same finding has also been previously confirmed by Brunet et al. (2007)
507 who assessed variability of extreme temperatures in Spain, providing evidence of larger
508 changes in warm temperature extremes during the 20th century, as compared to cool
509 extremes. Given that the Mediterranean is a close basin, temperature variability at coastal
510 sites seems to be closely linked to Sea Surface Temperature (SST) variations. Santoleri et
511 al. (1994) found an increase of 1.5ºC in mean SST across the western Mediterranean,
512 mostly faster during summer and winter compared to spring and autumn. More recently,
513 Salat and Pascual (2007) showed a similar upwarding trend in SST along the Catalan coast
514 (NW Mediterranean).
515
516 Our results confirm that high mountain areas (CL4) respond more rapidly to the global
517 warming compared to both mainlands (CL2) and moderately elevated areas (CL3). This can
518 be clearly seen in the indices of Max_summer (0.71ºC/decade), TXn (0.85ºC/decade), TXx
519 (0.59ºC/decade), SU25 (3.63 days/decade), Min_summer (0.60ºC/decade), TNn
520 (0.64ºC/decade), TNx (0.60ºC/decade) and Tsums (125.02ºC/decade). Our results can be of
521 particular importance in the context of the possible impacts of the global climatic change on
522 behavior of temperature extremes in areas of complex topography. Indeed, it is still unclear
523 whether these regions are undergoing a global warming, particularly with low density of

524 observatories in much of these areas. Only few studies have been undertaken to assess
525 elevation dependency on climate change signals in areas of large temperature gradient.
526 Among these few studies are Giorgi et al. (1997) and Beniston et al. (1997) for the Alps,
527 Fyfe and Flato (1999) for the Rocky, Chen et al. (2003) for the Tibetan Plateau, and
528 Coronato and Bisigato (1998) for the Andes. In the study domain, the rapid warming at
529 high elevations appears to be of great importance for ecological and hydrological systems
530 and climate impact assessment. Rapid changes could lead to significant changes in their
531 natural vegetation and thus impacts on ecosystems and biodiversity. Moreover, these
532 mountain environments are more likely to be affected by climate change and therefore they
533 can be an early indicator of climate variability and change for the nearby low elevated
534 areas.
535
536 **4.4. Influences of large-scale circulation on summertime extreme temperature:**
537 Figure 6 depicts Pearson correlation values between the general atmospheric circulation
538 and the regional time series of extreme temperature in the period 1960-2006. The
539 connection between the circulation patterns and temperature extremes is found statistically
540 significant ($p<0.05$) only for the EA, SCA and WeMO patterns. Among them, the WeMO
541 represents an east-west dipole; meanwhile the EA and SCA are north-south dipoles.
542 Though the relationship between the NAO and winter temperature over large proportions of
543 the Mediterranean and Europe is confirmed (Hurrell, et al., 2003), the NAO seems to be a
544 weak predictor for temperature extremes during summer season. This finding agrees well

with Trigo and Palutikof (2001) who found that the NAO poorly explained variability in atmospheric circulation during summer months, as compared with other seasons. Similarly, the role of the MO pattern sounds irrelevant. Variability of temperature extremes is mainly controlled by atmospheric circulation during the positive mode of the EA and the negative modes of the SCA and WeMO.

Overall, it seems that the SCA is a key controller of temperature extremes in all sub-regions, with correlation coefficients ranging between -0.20 and -0.64. However, it can be noted that the influence of the SCA on day-time extremes (e.g. max_summer, WD and TX90p) is much stronger than its influence on night-time extremes (e.g. min_summer, TXn, TNn and TN90p). In particular, correlation coefficients with day-time extremes lie in the range (-0.41 and -0.54), whereas they varied from -0.29 to -0.49 for night-time extremes. A quick view of the association between the SCA circulation pattern and temperature extremes at sub-regional scale reveals some significant spatial differences. The influence of the negative SCA on most of summer temperature extremes is generally less marked in the Mediterranean region (CL1) and, in contrast, much stronger in the highly elevated regions (CL3 and CL4). A clear example corresponding to DTR showed a significant correlation with the negative SCA in all the defined sub-regions, but with higher values in CL3 (r=-0.60) and CL4 (r=-0.56), compared with CL1 (r=-0.43) and CL2 (r=-0.38). This spatial pattern likely resembles that of the Spell index, which exhibited the strongest relationship with the SCA negative phase in CL4 (r=-0.64) and CL3 (r=-0.56), compared with CL1(r=-0.41) and CL2(r=-0.52).

566 The correlation of summer temperature extremes with the positive EA was generally
567 positive although it does not necessarily reach the statistical significance threshold, as being
568 the case with the WD, INTR, SU25 and Spell indices. In contrast to the SCA pattern, the
569 strongest association between the EA positive phase and temperature extremes was
570 markedly apparent along the Mediterranean coast (CL1) for the majority of the indices.
571 Contrarily, this influence declined in mainland and over complex terrain sub-regions.
572 Similar to the SCA, the WeMO correlated better with maximum temperature indices than
573 with minimum temperature indices. Spatially, it can be noted that the impacts of the WeMO
574 on temperature extremes are more pronounced in continental and low elevated areas than in
575 coastal and highly elevated regions.

576

577 **4.5. Co-variability between summertime temperature and SLP**

578 The obtained results suggest that the behavior of temperature extremes during summer is
579 mainly driven by atmospheric circulation during the positive EA, and the negative modes of
580 the SCA and WeMO. However, the impacts of these configurations seem to have a spatial
581 structure, with clear regional contrasts among the defined sub-regions. Accordingly, it
582 seems important to trace the influence of these leading circulation modes on regional
583 temperature variability during summer season. In this section, we explain the co-variability
584 between SLP as independent variable and summertime (JJA) temperature as dependent
585 variable during the period from 1960 to 2006. The canonical functions summarizing this

586 relationship during the most significant circulation patterns are depicted in figures from 7 to
587 9.

588 Figure 7-a shows the averaged anomalies of SLP over western and southwestern Europe
589 during summers with positive EA values. As illustrated, the positive mode of the EA
590 pattern is mainly associated with a strong dipole over the North Atlantic (approximately
591 50ºN, 30ºW) and dominance of anticyclones over the Mediterranean and central and
592 Eastern Europe. This situation also corresponds to an increase in the anticyclonic activity
593 over the Iberian Peninsula, while the Azores High extends northward. Under this SLP
594 configuration, the advection of the westerly and southwesterly winds over the Iberian
595 Peninsula is enhanced, while the easterly flows transport moisture from the Mediterranean
596 along a SE-NW gradient. Correspondingly, there is a weak moisture transport from the
597 north Atlantic during summer months, which induces a decrease in precipitation over the
598 study domain. The anomalous low precipitation showed positive feedback with soil
599 moisture anomalies. The impacts of the westerly, southwesterly and easterly flows on the
600 study domain are largely constrained by local terrain. For example, the central proportions
601 (e.g. the Ebro valley) are weakly affected by these flows as a consequence of the
602 surrounded orographic lifting (i.e. the Iberian and Cantabrian systems). Contrarily, the EA
603 exerts more important influence along the Mediterranean coast and highly elevated areas,
604 particularly with the increase in the warm and moist flows from the Mediterranean and the
605 mid of the Atlantic. Conversely, owing to the weak contrast between SLP over the
606 Mediterranean Sea and closing land areas during the positive EA as revealed by SLP

607 isopleths, the influences of these inflows can not extend to mainland, particularly with
608 prevalence of high pressure anomalies in most of the peninsula as a consequence of local
609 heating effects (Figure 7-a). In short, the EA positive phase corresponds to a remarkable
610 increase in frequency and intensity of summer temperature extremes along the
611 Mediterranean coast and high elevation sites. In their study on the entire Europe, Beranova
612 and Huth (2008) found the strongest connections between the EA mode and temperature in
613 southern France and NE Spain.
614
615 Figure 7b and c illustrate the two main canonical functions that explain the large proportion
616 of variability in both SLP and temperature during the positive EA mode. As shown, the
617 canonical correlation coefficients of the first and second functions are 0.78 and 0.60
618 respectively. It is also indicated that SLP variate on the first function is responsible for 20.4
619 % of summertime temperature variance. Spatially, it was found that temperature over
620 central and western areas of the study domain are mainly controlled by SLP over the
621 Mediterranean region, particularly the western basin (r = -0.7), whereas it is anticorrelated
622 with SLP anomalies over the northern Atlantic. SLP variate on the second function
623 accounted for 11.9 % of temperature variance. This function also suggested that
624 temperature variations over the eastern proportions of the study area are significantly
625 influenced by SLP over central Europe and northern Africa, while they correlate negatively
626 with SLP over the eastern Atlantic (Figure 7c). Taken together, it can be noted that higher
627 temperatures during summers of positive EA values are mainly associated with positive

628 SLP anomaly over the Mediterranean, central Europe and North Africa on one hand and the
629 negative SLP anomaly centered on the east Atlantic on the other hand.
630
631 Figure 8-a illustrates the composite of SLP anomalies during summers of negative SCA
632 values over the period from 1960 to 2006. As depicted, the negative phase of the SCA
633 pattern produces similar configuration to that of the positive EA, with clear high pressure
634 anomalies over central Europe. This high pressure anomaly extends westward to cover the
635 Iberian Peninsula and southward to include the Mediterranean region and North Africa. On
636 the other hand, a low pressure system predominates over the north Atlantic and vast areas
637 of northern Europe and Scandinavia. This atmospheric circulation indicates that the
638 advection of northern flows over the peninsula is largely restricted during the negative
639 mode of SCA, as a consequence of dominance of anticyclonic conditions over the European
640 mainland, giving rise to increased above-temperature. This situation comes clearly in
641 contrast with SLP features during summers of positive SCA values, where the anticyclonic
642 anomalies predominate over the Scandinavian region favoring advection of northern cold-
643 air flows toward the Iberian Peninsula (Bueh and Nakamura, 2007). This result agrees well
644 with Blackburn and Hoskins (2001) who found that the positive SCA in Western Europe is
645 mainly associated with predominance of the cyclonic conditions over the region. This
646 behavior enhances the advection of the surface westerlies from the Atlantic toward southern
647 Europe and the Mediterranean. The westerlies cause the transport of mild and humid
648 airflow that could weaken heat severity during summer months. The canonical correlation

649 results suggest two main functions to explain the co-variability between SLP anomalies and
650 temperature during summers of the negative SCA (Figure 8b and c). The first canonical
651 function accounted for 18.3% of temperature variance, with a canonical correlation
652 coefficient being 0.74. On the other hand, the second function captured 17.3% of
653 temperature variance, with a canonical correlation coefficient of 0.72. A quick comparison
654 between Figures 8b and c suggests that higher temperature along the northeastern parts of
655 the study domain are mainly controlled by the high pressure anomaly over northern Africa
656 (r between -0.2 and -0.8) and central Europe (r=-0.3). Similarly, the anomalous high
657 temperatures over western areas are significantly connected to the SLP positive anomaly
658 located over the Mediterranean region and central and southern Europe. These spatial
659 patterns clearly match those patterns observed during the EA positive mode.
660
661 Spatial variations of SLP anomalies corresponding to summers of negative WeMO values
662 are illustrated in Figure 9-a. As shown, the atmospheric circulation corresponding to the
663 negative phase of the WeMO is mainly associated with a remarkable increase in the zonal
664 circulation activity. Obviously, summers of negative WeMO values are characterized by a
665 strong positive SLP anomaly over the whole Europe with a maximum over the
666 Northwestern Europe and in particular the British Isles (up to +1.0 hPa). In contrast, SLP
667 shows a slight negative pressure anomaly over western parts of Iberia. This is mainly
668 associated with the northward shift of the Azores high. In contrast to the meridional flows
669 which are clearly constrained during the negative WeMO as a consequence of the European

670 blockings, zonal airflows from the Mediterranean Sea are promoted causing above-normal
671 temperature over eastern proportions of the study area. Figure 9b and c depict the main
672 spatial patterns of SLP-temperature co-variability during the negative WeMO mode. The
673 results suggest two main functions that explain together 40.1 % of temperature variance
674 (the first= 20.6% and the second= 19.5%). The spatial association between temperature
675 variations and centers of action of SLP anomaly resembles well those detected for both the
676 positive EA and the negative SCA patterns.

677 In summary, it can be concluded that the variability of summer temperature extremes in NE
678 Spain is particularly related to the circulation modes that produce high pressure anomalies
679 over much of Europe and the Mediterranean Sea. Numerous studies found a statistically
680 positive trend in SLP over the whole Mediterranean and most of the continental Europe
681 during warm summers of recent decades (e.g. Reddaway and Bigg, 1996, Xoplaki, 2002).
682 For example, Xoplaki (2002) noted an upward trend in both surface pressure and different
683 geopotential heights over the eastern Atlantic and most of continental Europe west of 30ºE.
684 However, the canonical functions suggest that spatial variability of temperature anomalies
685 over northeastern Spain varies considerably according to SLP anomaly variations, which
686 markedly differs its position, strength and influence domain from one prominent mode to
687 another

688 **Conclusion**

689 In this study, the full procedure used to classify daily temperature extremes in NE
690 Spain during summer season is described. The main objective was to delineate spatially

691 coherent regions employing 14 temperature-based extreme indicators derived from a 47-
692 year of daily information (1960-2006). Multivariate statistics (i.e. PCA and CA) did an
693 adequate job in providing a useful classification that gave insights into spatial variability of
694 summer temperature extremes. The goodness of the clustering was evaluated by means of
695 the Silhouette index. Four sub-regions with climatic and geographic meanings were
696 indentified: the Mediterranean region, the Cantabrian region and the inland region, the
697 moderately western and southern areas, and the highly elevated areas were identified.
698

699 The temporal evolution of summer temperature extremes for the established sub-regions
700 was examined. In general, a warming trend was exhibited for both maximum and minimum
701 temperatures, being more pronounced at high elevation sites and along the Mediterranean
702 coastland. The observed changes could be of particular importance for implications in areas
703 of environmental and hydrological assessment and monitoring, and climate impact studies.
704 These findings can also provide a basis to understand the hydrological processes possibly
705 altered by climate change, such as evaporation, surface runoff, drought conditions and
706 water availability. Based on the defined homogenous regions, the regional time series of
707 temperature extremes were also linked with those of teleconnection patterns from 1960 to
708 2006. The correlation of the large-scale atmospheric circulation with temperature extremes
709 showed that temperature extremes in NE Spain are mainly explained by three
710 configurations (SCA, WeMO and EA). The negative mode of the SCA pattern proved to be
711 capable of explaining most of variability in summer temperature extremes at sub-regional

712 scale, with more influence in the highly elevated areas. In contrast, the impact of the EA
713 positive phase is more highlighted in the Mediterranean region, as compared to mainland.
714
715 While the spatial domain of this study is quite limited (≈160,000 km^2), the applied
716 procedure did indicate a high degree of inter-regional variability in characteristics of
717 temperature extremes (i.e. frequency, intensity and persistence). However, recalling that the
718 influence of SLP on temperature may be interrupted by certain local conditions (e.g.
719 vegetation canopy, land use changes and topography), a detailed study on variability of
720 weather types persistence based on classification of atmospheric circulation at various
721 pressure heights (i.e. surface sea, 200hPa, 500hPa and 700hPa) could be of interest in
722 future. This work represents one of the first attempts to explore temporal characteristics of
723 temperature extreme events and their driving mechanisms at sub-regional scale in the
724 Iberian Peninsula. Given that the study area is characterized by complicated mountainous
725 terrain, more information on the behavior of temperature extremes at the regional or local
726 scale is essential for different impact assessment applications. In this regard, this work can
727 provide an insight into the possible mechanisms that may relate spatial and temporal
728 variability in summer temperature extremes with atmospheric circulation. Our results
729 therefore could be meaningful for various applications related to hydrological modeling,
730 agroclimatology and drought monitoring.
731
732

Acknowledgements

We would like to thank the Spanish Meteorological State Agency (AEMET) for providing the temperature database used in this study. This work has been supported by the research projects CGL2008-01189/BTE, CGL2011-27574-CO2-02 and CGL2011-27536 financed by the Spanish Commission of Science and Technology and FEDER, EUROGEOSS (FP7-ENV-2008-1-226487) and ACQWA (FP7-ENV-2007-1- 212250) financed by the VII Framework Programme of the European Commission, "Efecto de los escenarios de cambio climático sobre la hidrología superficial y la gestión de embalses del Pirineo Aragonés" financed by "Obra Social La Caixa" and the Aragón Government and Influencia del cambio climático en el turismo de nieve, CTTP01/10, Financed by the *Comisión de Trabajo de los Pirineos.*

References

Alexander, L. V. and Coauthors (2006): Global observed changes in daily climate extremes of temperature and precipitation, J. Geophys. Res. Atmos., 111, D05109, doi: 10.1029/2005JD006290.

Baeryswil, P.A. and Rebetez, M. (1997). Regionalization of precipitation in Switzerland by means of principal component analysis, Theoretical and Applied Climatology, 58, 31-41.

Barriopedro, D., Fischer, E.M., Luterbacher, J., Trigo, R.M., García-Herrera, R. (2011). The hot summer of 2010: Redrawing the temperature record map of Europe. Science 332: 220-224.

Bejaran, R.A. and Camilloni, I.A. (2003). Objective method for classifying air masses: an application to the analysis of Buenos Aires (Argentina) urban heat island intensity. Theor. and Appl. Climatol., 74: 93–103.

Beniston, M., Diaz, H. F., Bradley, R. S. (1997). Climatic change at high elevation sites: an Overview, Clim. Change, 36, 233–251.

Beranova, R. and Huth, R. (2008). Time variations of the effects of circulation variability modes on European temperature and precipitation in winter, Int. J. Climatol., 28, 139-158.

Blackburn, M. and Hoskins, B.J. (2001): The UK-record breaking wet autumn 2000., pp:38-40, in UK Universities Global Atmospheric Modeling Programme, UGAMP Newsletter 24, Meteorological Department, University of Reading, UK.

Bonell, M. and Summer, G. (1992). Autumn and winter daily precipitation areas in Wales,

1982–1983 to 1986–1987. Int. J. Climatol., 12, 77–102.
Brunet, M., Jones, P.D., Sigró, J., Saladié,O., Aguilar, E., Moberg, A., Della-Marta, P.M., Lister, D., Walther, A., López, D. (2007). Temporal and spatial temperature variability and change over Spain during 1850–2005, J. Geophys. Res., 112, D12117, doi:10.1029/2006JD008249.
Bueh, C. and Nakamura, H. (2007). Scandinavian pattern and its climatic impact. Quarterly Journal of the Royal Meteorological Society, 133: 2117–2131. doi: 10.1002/qj.173.
Camiz, S., Pillar, V.D. (2007). Comparison of single and complete linkage clustering with the hierarchical factor classification of variables. Community Ecology 8(1), 25-30.
Chen, B.D., Chao, W.C. and Liu, X.D. (2003). Enhanced climatic warming in the Tibetan Plateau due to doubling CO2: a model study, Clim. Dyn. 20,401–413.
Ciais, Ph., Reichstein, M., Viovy, N., et al., (2005): Europe-wide reduction in primary productivity caused by the heat and drought in 2003. Nature, 437: 529-533.
Coronato, F. and Bisigato, A. (1998). A temperature pattern classification in Patagonia, Int J Climatol, 18 (7), 765-773.
DeGaetano, A. T. (2001). Spatial grouping of United States climates stations using a hybrid clustering approach, Int. J. Climatol., 21, 791–807, 2001.
Dillon, W. R., and M. Goldstein (1984), Multivariate Analysis: Methods and Applications. New York: Wiley.
El Kenawy, A., López-Moreno, J. I., Vicente-Serrano, S.M. (2011a). Trend and variability of surface air temperature in northeastern Spain (1920–2006): Linkage to atmospheric circulation, Atmospheric Research (2011), doi: 10.1016/j.atmosres.2011.12.006.
El Kenawy, A., López-Moreno, I., Vicente-Serrano, S.M. and Stepanek, P. (2011b). An assessment of the role of homogenization protocol in the performance of daily temperature series and trends: application to northeastern Spain, Int. J. Climatol., doi: 10.1002/joc.3410.
Esteban, P., Martin-Vide, J., Mases, M. (2006). Daily atmospheric circulation catalogue for Western Europe using multivariate techniques, Int. J. Climatol. 26:1501–1515.
Everitt, B.S. and Dunn, G. (1991). Applied Multivariate Data Analysis, Edward Arnold, London, 342p.
Folland, C.K., Karl, T.R., Christy, J.R., Clarke, R.A., Gruza, G.V., Jouzel, J., Mann, M.E., Oerlemans, J., Salinger M.J. and Wang, S.W. (2001): Observed Climate Variability and Change. In: Climate Change (2001): The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, Houghton J.T. et al. (ed.), Cambridge Univ. Press.
Fyfe, J.C, and Flato, G.M. (1999). Enhanced climate change and its detection over the rocky mountains, J. Climate 12, 230–243.
Giorgi, F., Hurrel, J.W., Marinucci, M.R. and Beniston, M. (1997). Elevation dependency of the surface climate change signal: a model study, J. Climate 10, 288–296.
Gong X. and Richman, M.B. (1995). On the application of cluster analysis to growing season precipitation data in North America east of the Rockies, J. Climate, 8, 897–

931.
Hertig, E., Seubert, S. and Jacobiet, J. (2010). Temperature extremes in the Mediterranean area: trends in the past and assessments for the future, Nat. Hazards Earth Syst. Sci., 10, 2039-2050.
Hurrell, J.W. (1995). Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation. Science 269, 676-679.
Hurrell, J.W., Kushnir, Y., Ottersen, G., Visbeck, M. (2003). An overview of the North Atlantic Oscillation. In: Hurrell, J.W., Kushnir, Y., Ottersen, G., Visbeck, M. (eds) The North Atlantic Oscillation. American Geophysical Union, Washington, DC, p 1–35.
Jones, P.D., New, M., Parker, D.E., Martin, S., Rigor, I.G. (1999). Surface air temperature and its changes over the past 150 years. Rev. Geophys. 37, 173-199.
Holt, T. (1999), A classification of ambient climatic conditions during extreme surge events off Western Europe. International Journal of Climatology, 19: 725–744.
Kalkstein, L. S., Tan, G., Skindlov, J.A. (1987). An evaluation of three clustering procedures for use in synoptic climatological classification. J. Climate Appl. Meteor., 26, 717–730.
Karl, T. R. and Koscielny, A. J. (1982). Drought in the United States: 1895-1981, J. Climatol. 2: 313-329.
Karl, T.R., Knight, R.W., Plummer, N. (1995): Trends in high-frequency climate variability in the twentieth century, Nature, 377, 217 – 220.
Klein Tank, A.M.G. and Können, G.P. (2003). Trends in indices of daily temperature and precipitation extremes in Europe, 1946–99. J. Climate, 16, 3665-3680.
Kostopoulou, E. and Jones, P. (2005). Assessment of climate extremes in Eastern Mediterranean, Meteorol. Atmos. Phys., 89:69–85.
Love, D., Hallbauer, D., Amos,A., Hranova, R. (2004): Factor analysis as a tool in groundwater quality management: two southern African case studies, Physics and Chemistry of the Earth, 29(15-18):1135-1143.
Martin-Vide, J., and Gomez, L. (1999). Regionalization of peninsular Spain based on the length of dry spells, Int. J. Climatol., 19, 537–555.
Martin-Vide, J., and Lopez-Bustins, J.A. (2006). The Western Mediterranean Oscillation and rainfall in the Iberian Peninsula, Int. J. Climatol., 26 (11), 1455-1475.
Milligan, G. W. and Cooper, M. C. (1985). An Examination of procedures for determining the number of Clusters in a data Set, Psychometrika, 50, 159–179.
Mohan, S. and Arumugam, N. (1996). Relative importance of meteorological variables in evapotranspiration: Factor analysis approach, Water Resources Management, 10(1):1-20.
Muñoz-Díaz, D. and Rodrigo, F. S. (2004). Spatio-temporal patterns of seasonal rainfall in Spain (1912-2000) using cluster and principal component analysis: comparison, Ann. Geophys., 22, 1435-1448.
Norusis, M.J. (1988). The SPSS guide to data analysis, SPSS/PC+ Advanced statistics, V20, SPSS Inc., Chicago, 420p.

Palutikof, J.P. (2003). Analysis of Mediterranean climate data: measured and modelled. **In** Mediterranean Climate Variability and Trends, Bolle HJ (ed.). Springer-Verlag: Berlin; 133–153.

Papadimas, C. D., Bartzokas, A., Lolis, C. J., Hatzianastassiou, N. (2011). Sea-level pressure-air temperature teleconnections during northern hemisphere winter, Theoretical and Applied Climatology, DOI: 10.1007/s00704-011-0523-8.

Pasho, E., Camarero, J.J., de Luis, M. and Vicente-Serrano, S.M., (2011a): Spatial variability in large-scale and regional atmospheric drivers of Pinus halepensis growth in eastern Spain. Agricultural and Forest Meteorology. 151, 1106-1119.

Pasho, E., J. Julio Camarero, Martín de Luis and Vicente-Serrano, S.M. (2011b) Impacts of drought at different time scales on forest growth across a wide climatic gradient in north-eastern Spain. Agricultural and Forest Meteorology. 151: 1800-1811.

Pepin, N.C. and Lundquist, J.D. (2008). Temperature trends at high elevations: patterns across the globe. Geophysical Research Letters, 35, L14701, doi:10.1029/2008GL034026.

Reddaway, J., and Bigg, G. (1996). Climatic change over the Mediterranean and links to the more general atmospheric circulation. Int. J. Climatol. 16: 651-661.

Reimann, C, Filzmoser, P., Garrett, R.G. (2002). Factor analysis applied to regional geochemical data: problems and possibilities. Applied Geochemistry 17, 185-206.

Rodríguez-Puebla, C., Encinas, A.H., Saenz, J. (2001). Winter precipitation over the Iberian peninsula and its relationship to circulation indices, Hydrol. Earth Syst. Sci., 5, 233–244.

Rodríguez-Puebla C., Encinas, A.H., García-Casado, L.A. and Nieto, S. (2009). Trends in warm days and cold nights over the Iberian Peninsula: relationships to large-scale variables, Clim. Change, DOI 10.1007/s10584-009-9721-0.

Romero, R., Sumner, G., Ramis, C., Genoves, A. (1999). A classification of the atmospheric circulation patterns producing significant daily rainfall in the Spanish Mediterranean area, Int. J. Climatol., 19, 765–785.

Rousseeuw, P.J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, J. Comput. and Appl. Mathematics, 20, 53-65.

Salat, J. and Pascual, J. (2007). Climatological trend from 32 years of observations at L'Estartit station, near the Catalan coast (NW Mediterranean), Rapp. Comm. Int. Mer Medit., 38: 196.

Santoleri, R., Böhm, E. and Schiano, M.E. (1994). The sea surface temperature of the western Mediterranean Sea: Historical satellite thermal data , in *Seasonal and Interannual Variability of the Western Mediterranean Sea, Coastal Estuarine Stud.*, vol. 46, edited by P. E. La Viollette *et al*., pp. 155-176, AGU, Washington, D. C.

Schulte, A. and Mrosek, T. (2006). Analysis and assessment of the forestry and wood-processing industry cluster in the State of North Rhine-Westphalia, Germany. *Forstarchiv* 4, 136–141.

Serra, C., Martínez, M.D., Lana, X., Burgueño A. (2010). Extreme normalized residuals of daily temperatures in Catalonia (NE Spain): sampling strategies return periods and

clustering process, Theor. and appl. Climatol., 101: 1-17.
Trigo, R.M. and Palutikof, J.P. (2001). Precipitation scenarios over Iberia: a comparison between direct GCM output and different downscaling techniques, J. of Climate, 14: 4422-4446.
Vicente-Serrano, S.M. (2006). Spatial and temporal analysis of droughts in the Iberian Peninsula (1910-2000). Hydrological Sciences Journal. 51: 83-97.
Vicente-Serrano, S. M., Beguería, S., López Moreno, J.I., El Kenawy, A., Angulo, M.M. (2009). Daily atmospheric circulation events and extreme precipitation risk in northeast Spain: Role of the North Atlantic Oscillation, the Western Mediterranean Oscillation, and the Mediterranean Oscillation, J. Geophys. Res., 114, D08106, doi: 10.1029/2008JD011492.
Wolting, G., Bouvier, C.H., Danloux, J., Fritsch, J.M. (2000). Regionalization of extreme precipitation distribution using the principal components of the topographical environment, J. Hydrol. 233, 86–101.
WHO (2003). The health impacts of 2003 summer heat-waves. Briefing note for the Delegations of the fifty-third session of the WHO (World Health Organization) Regional Committee for Europe. 12 pp.
Xoplaki E. (2002). Climate variability over the Mediterranean. PhD thesis, University of Bern.
Young, F.J. and Hammer, R.D. (2000). Defining geographic soil bodies by landscape position, soil taxonomy, and cluster analysis. Soil Sci. Soc. Am. J. 64:989–998.

Table 1: List of the indices of summer temperature extremes and their definitions.

Index	Description	Symbol	Unit
Max_monthly_min	Maximum value of monthly minimum temperature in summer.	TNx	º C
Max_monthly_max	Maximum value of monthly maximum temperature in summer.	TXx	º C
Min_monthly_min	Minimum value of monthly minimum temperature in summer.	TNn	º C
Min_monthly_min	Minimum value of monthly minimum temperature in summer.	TXn	º C
Diurnal temperature range	Monthly mean difference between TX and TN.	DTR	º C
Intra-annual extreme range	Difference between maximum TX and minimum TN in summer.	INTR	º C
Temperature sums	Sum of Tmax days >17º c – days Tmax < 17 ^{o}C	Tsums	º C
Warmest day	Highest daily maximum temperature.	WD	º C
Spell	Maximum length of consecutive days with daily maximum temperature higher than the 90th percentile	Spell	days
Summer days	Number of days with maximum temperature >25 ^{o}C during summer.	SU25	days
Warm days	Percentages of days with maximum temperatures higher than the 90th percentile.	TX90p	days
Warm nights	Percentages of days with minimum temperatures higher than the 90th percentile.	TN90p	days
Max_ summer	Highest daily maximum temperature in summer months	Max_Summer	º C
Min_summer	Lowest daily maximum temperature in summer months	Min_Summer	º C

Table 2. Trends of extreme temperature indices for the defined four sub-regions, as defined by the cluster analysis results. Trend magnitude is expressed in unit per decade following the ordinary least squares method (numbers in bold are significant at the 95% level assessed using the Mann-Kendall statistic).

Index	CL1	CL2	CL3	CL4	Unit
Max_summer	**0.56**	**0.24**	**0.47**	**0.71**	ºC
TXn	**0.55**	0.34	**0.66**	**0.85**	ºC
TXx	**0.62**	0.18	**0.34**	**0.59**	ºC
SU25	**2.77**	1.06	**2.42**	**3.63**	days
TX90P	**4.11**	0.89	**1.75**	**2.61**	days
WD	**0.71**	0.09	**0.34**	**0.44**	ºC
Min_summer	**0.44**	**0.46**	0.11	**0.60**	ºC
TN90P	**3.88**	**2.83**	0.40	**2.86**	days
TNn	**0.44**	**0.48**	0.16	**0.64**	ºC
TNx	**0.52**	**0.48**	0.09	**0.60**	ºC
Spell	**1.29**	0.24	**0.41**	**0.54**	days
Tsums	**51.65**	24.76	**53.11**	**125.02**	ºC
DTR	**0.12**	**-0.22**	**0.36**	0.11	ºC
INTR	**0.45**	-0.29	0.35	-0.11	ºC

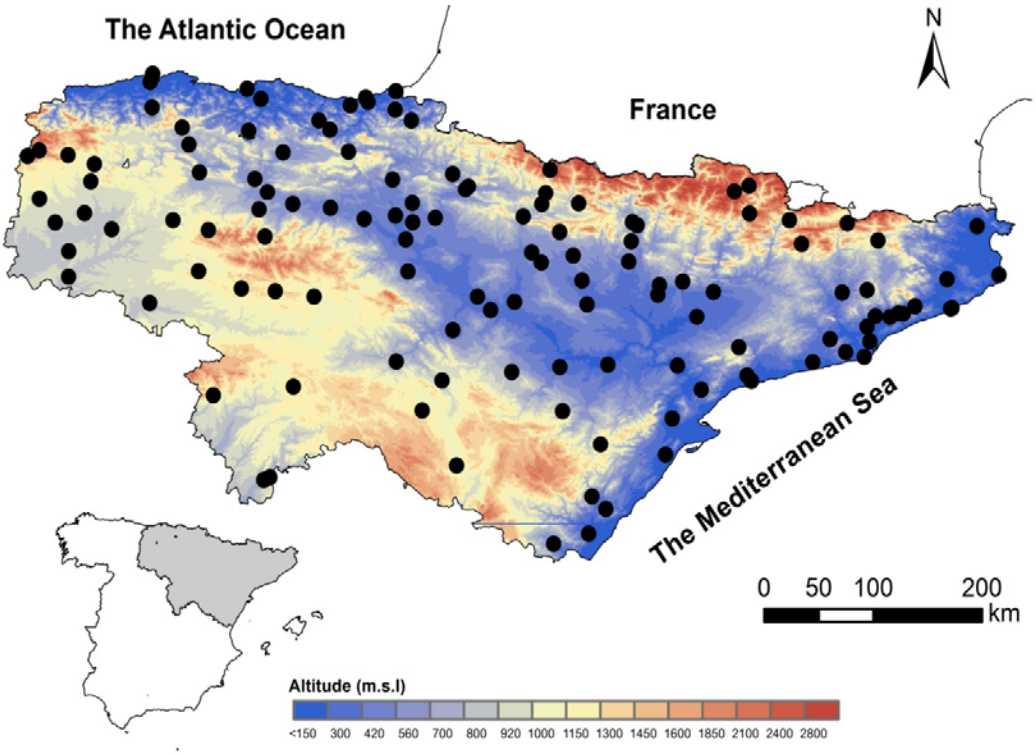

Figure 1: Terrain elevation and spatial distribution of the meteorological observatories employed in this study.

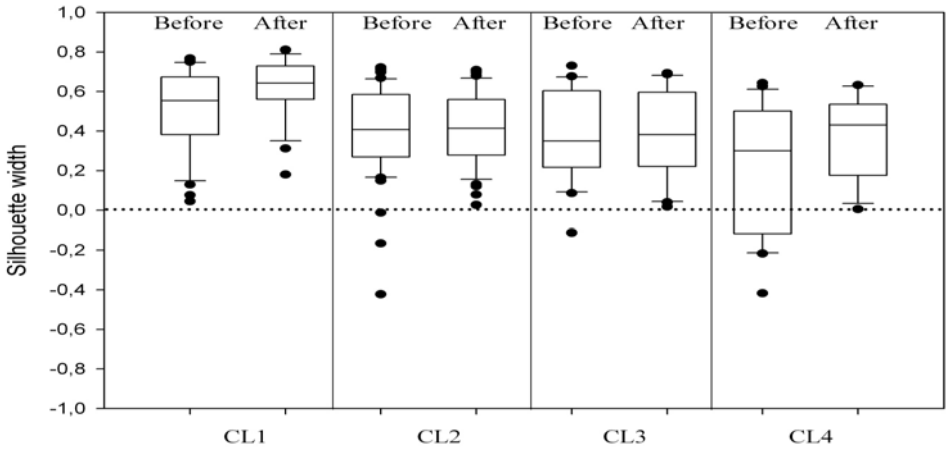

Figure 2: Silhouette width for the defined clusters before and after clustering validation.

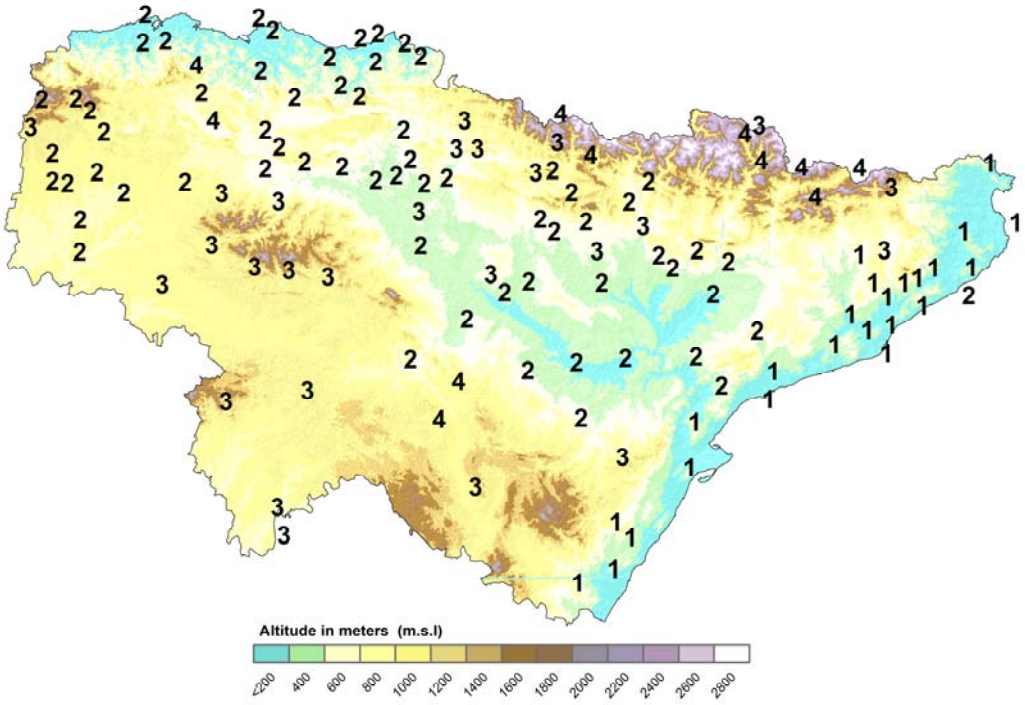

Figure 3: Spatial distribution of the observatories delineated to the four homogenous sub-regions following the cluster validation results.

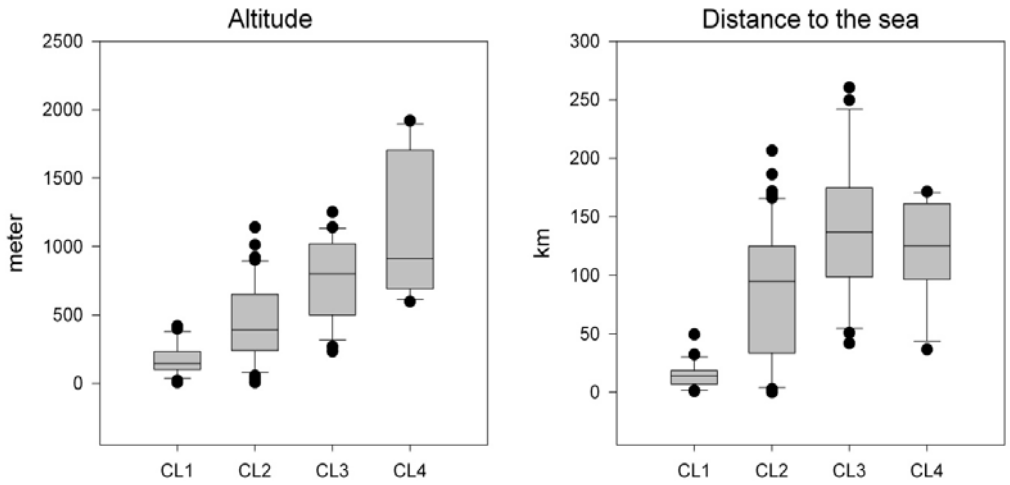

Figure 4: Boxplots showing intracluster differences as a function of altitude (left panel) and distance to closing water bodies (right panel). The median, 10th, 25th, 75th and the 90th percentiles as vertical boxes are plotted with errors bar.

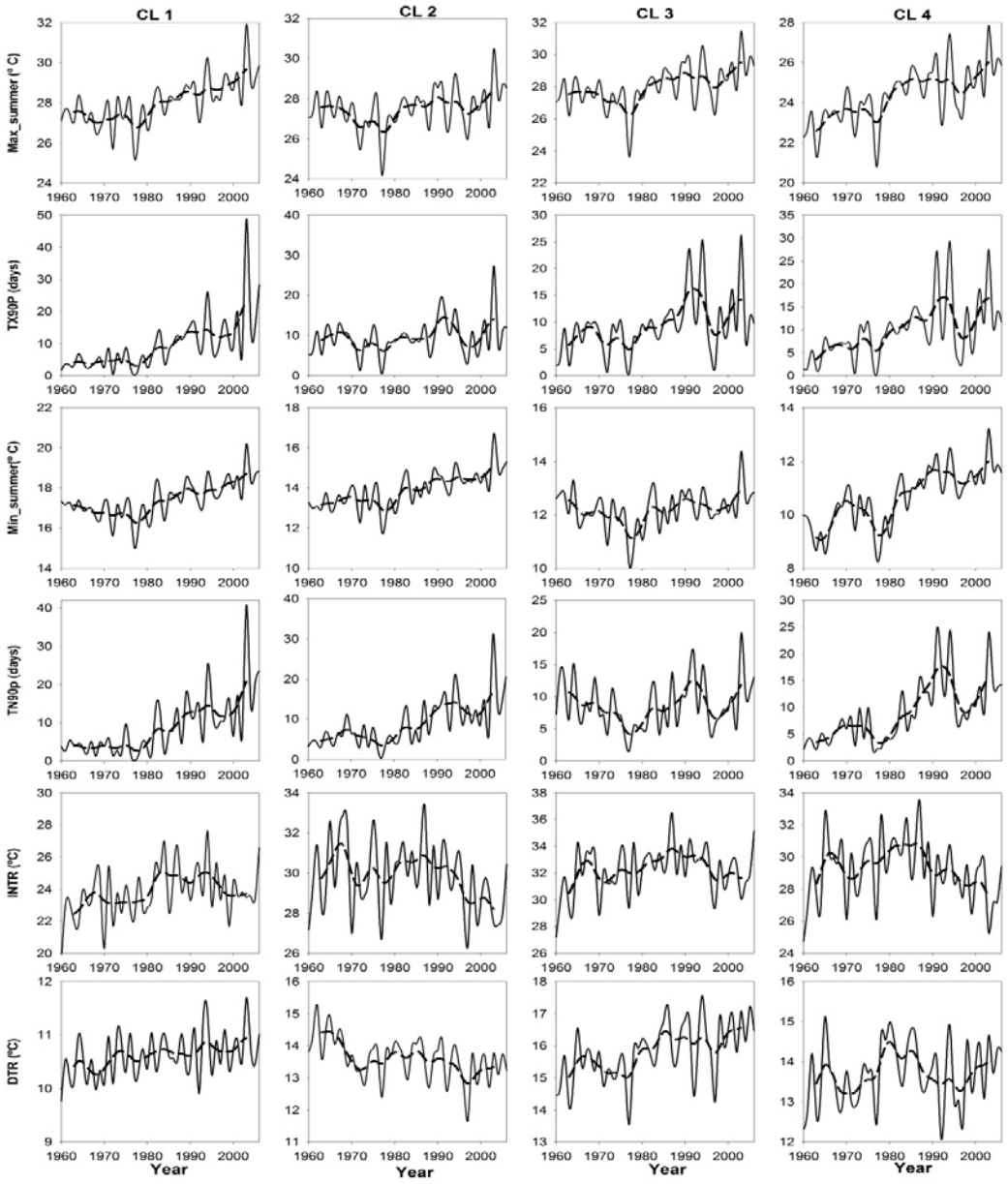

Figure 5: Temporal evolution of a set of temperature indices time series from 1960 to 2006. Max_summer (TX90p) represents an example of intensity (frequency) index of day-time temperature, while Min_summer (TN90p) represents an example of intensity (frequency) index of night-time temperature. INTR and DTR represent intensity indices of temperature variability. Dashed line refers to a 9-year low Gaussian filter.

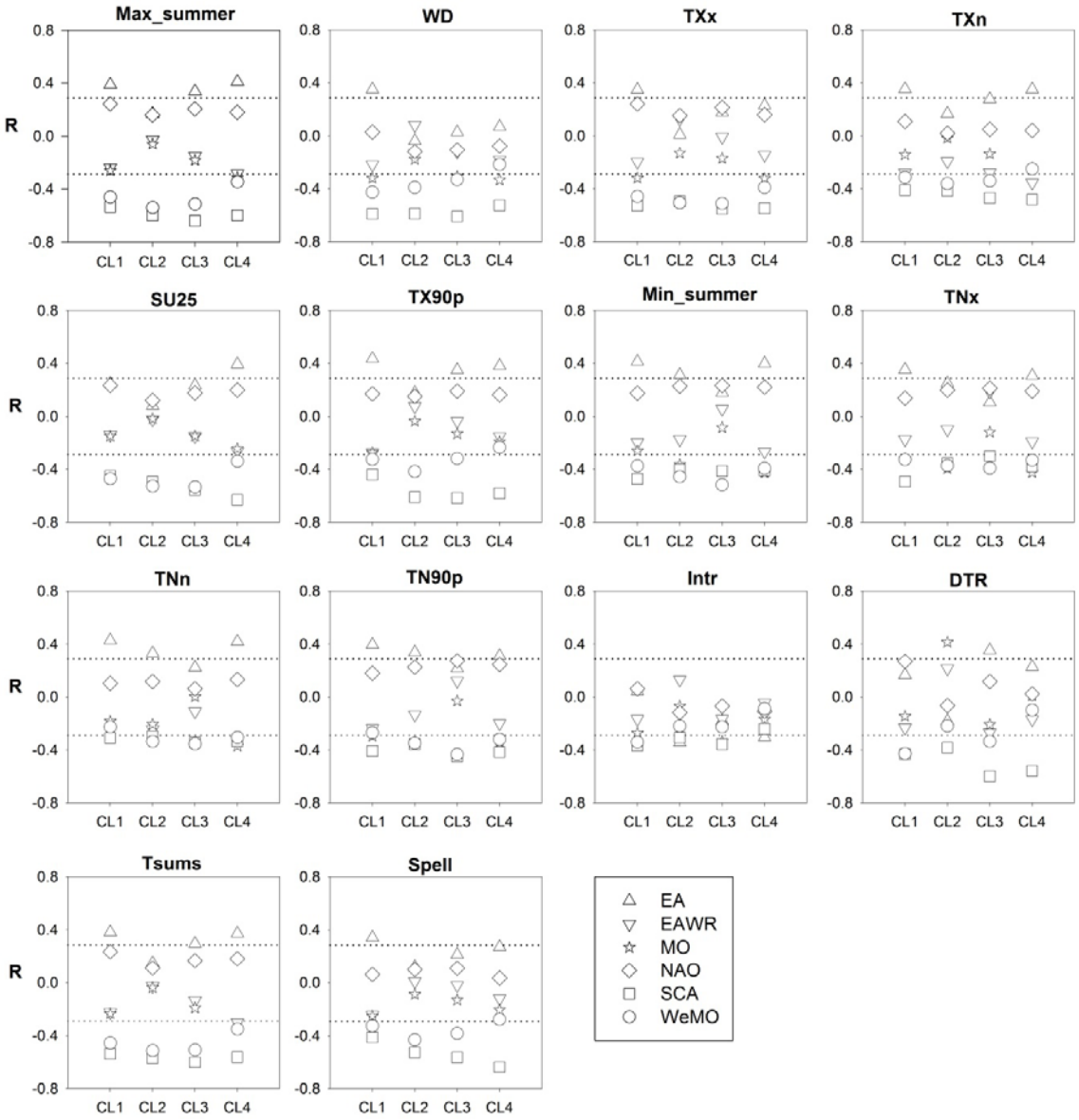

Figure 6: Pearson correlation coefficients between detrended time series of summer extremes and the main modes of atmospheric circulation over the period 1960-2006. Dotted lines show the upper and lower limits of the 95% level of significance.

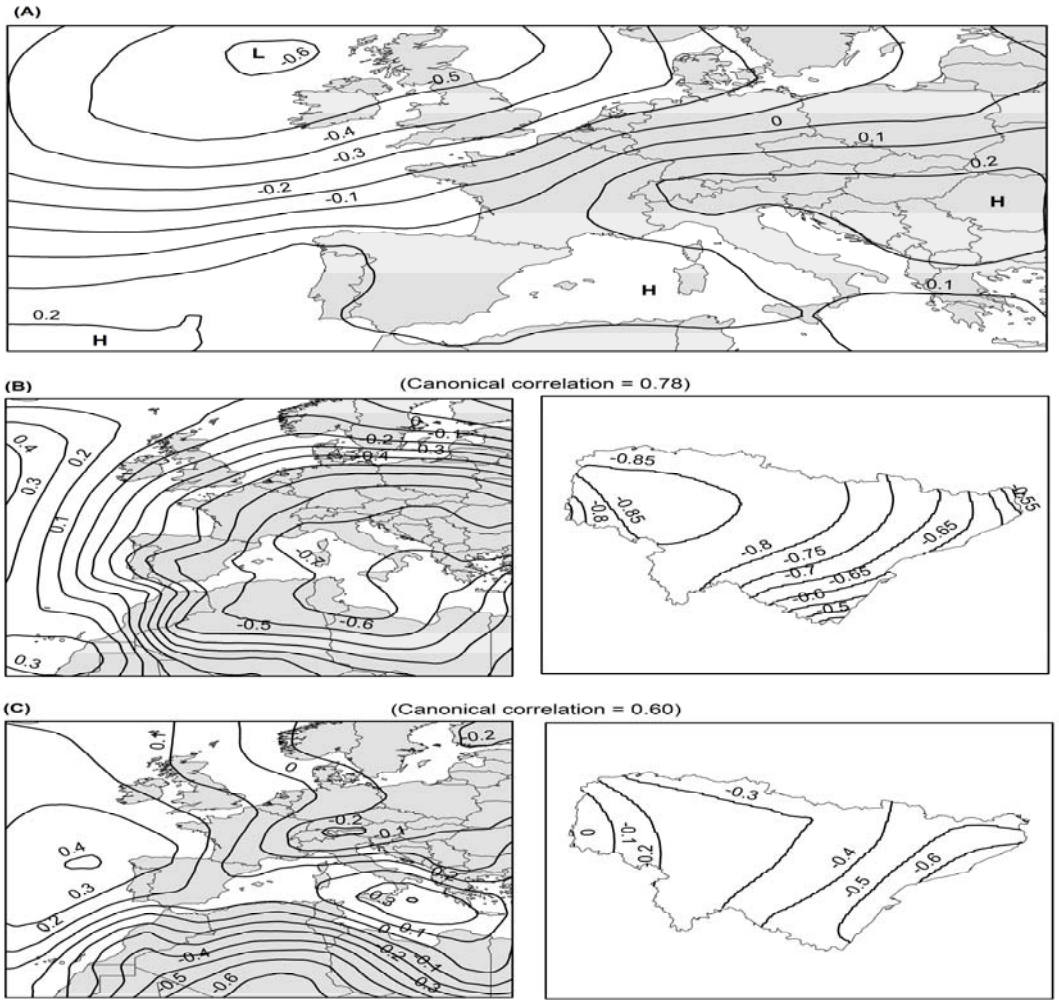

Figure 7: (A) Composite of sea level pressure (SLP) anomalies (hPa) corresponding to summers with positive EA values from 1960 to 2006. SLP anomalies at each grid are computed using the monthly mean and standard deviation for the long-term period (1960–2006). (B) The first leading canonical function of SLP and temperature anomalies co-variability during summers of the positive EA mode, (C) the same as (B) but for the second leading function. Isopleths show the Pearson correlation coefficient. Coefficients above 0.23 are statistically significant at the 95% level.

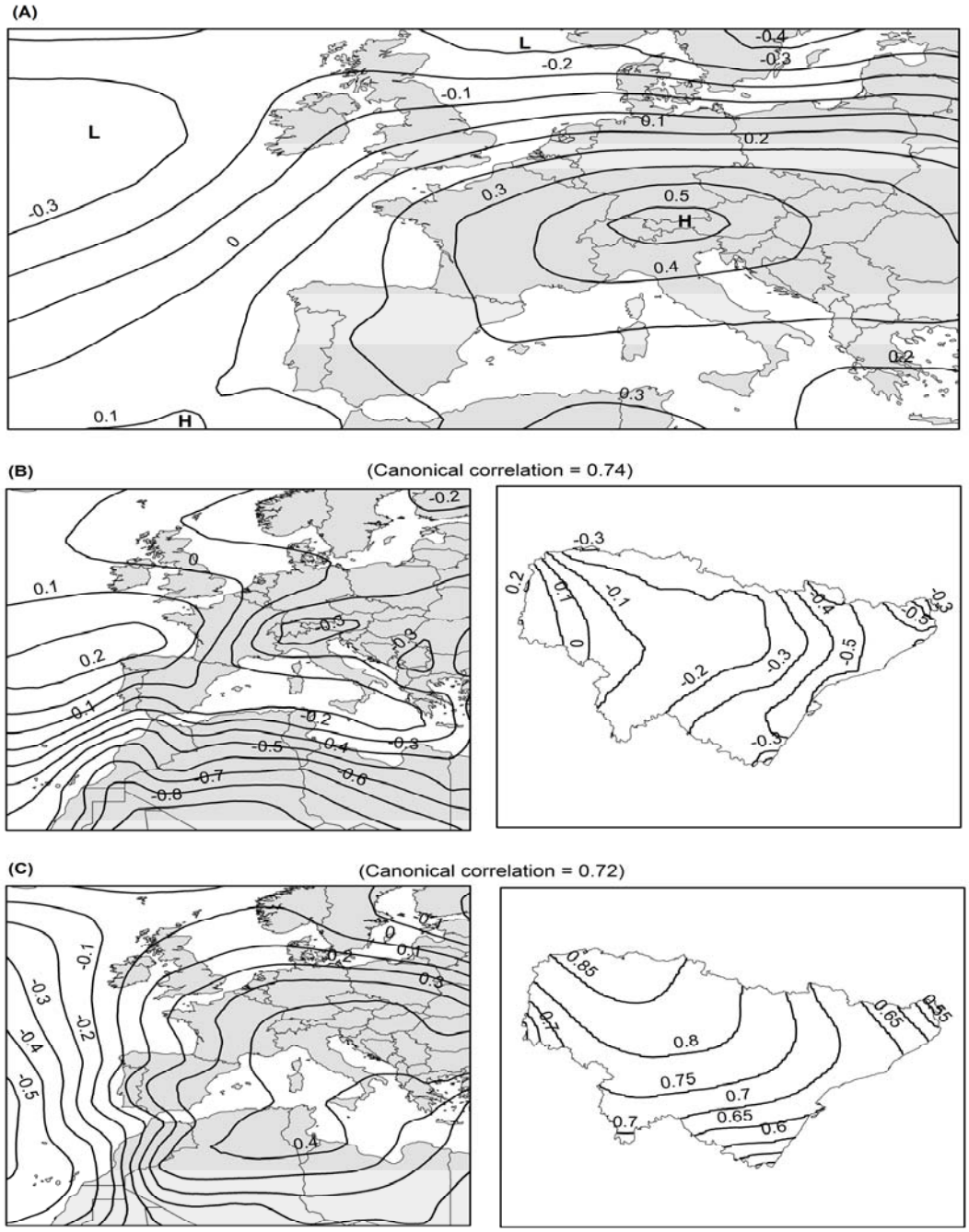

Figure 8 : The same as figure 7, but for summers with negative SCA values.

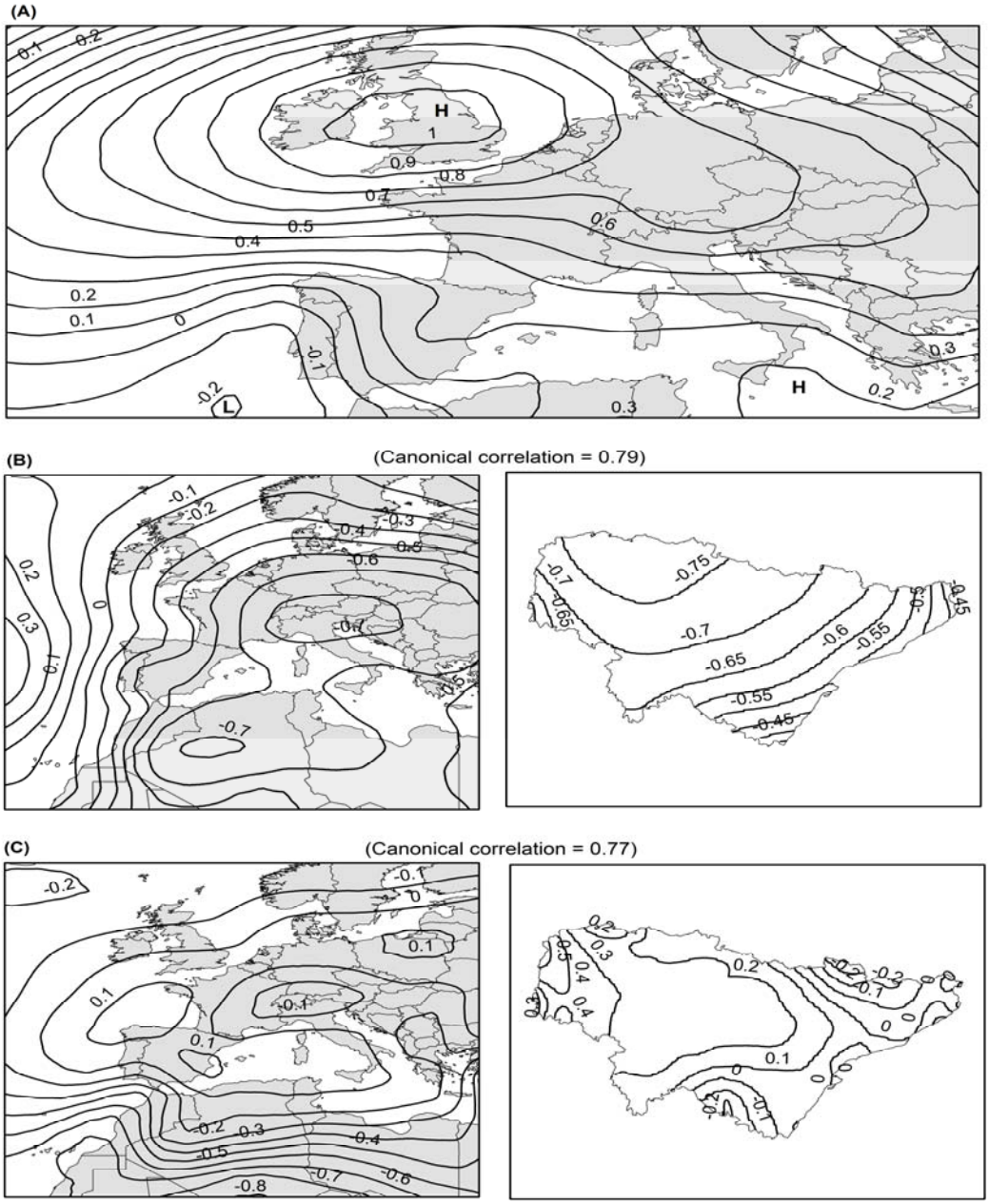

Figure 9: The same as figure 7, but for summers with negative WeMO values.