Changes in the simulation of instability indices over the Iberian Peninsula due to the use of 3DVAR data assimilation

Santos J. González-Rojí[1,2], Sheila Carreno-Madinabeitia[3,4], Jon Sáenz[5,6], and Gabriel Ibarra-Berastegi[7,6]

[1]Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland.
[2]Climate and Environmental Physics, University of Bern, Bern, Switzerland.
[3]Department of Applied Mathematics, Statistics and O.R., University of the Basque Country (UPV/EHU), Vitoria-Gasteiz, Spain.
[4]TECNALIA, Basque Research and Technology Alliance (BRTA), Parque Tecnológico de Álava, Vitoria-Gasteiz, Spain.
[5]Department of Applied Physics II, University of the Basque Country (UPV/EHU), Leioa, Spain.
[6]Plentzia Itsas Estazioa, PIE, University of the Basque Country (UPV/EHU), Plentzia, Spain.
[7]Department of NE and Fluid Mechanics, University of the Basque Country (UPV/EHU), Bilbao Engineering School, Bilbao, Spain.

Correspondence: Santos J. González-Rojí (santos.gonzalez@climate.unibe.ch)

Abstract. The ability of two downscaling experiments to correctly simulate thermodynamic conditions over the Iberian Peninsula (IP) is compared in this paper. To do so, three instability indices are evaluated: TT index, CAPE and CIN. The WRF model is used for the simulations. The N experiment is driven by ERA-Interim's initial and boundary conditions; The D experiment has the same configuration as N, but the 3DVAR data assimilation step is additionally run at 00, 06, 12 and 18 UTC. Eight radiosondes are available over the IP, and the values for these indices directly retrieved from the Integrated Global Radiosonde Archive (IGRA) were chosen as reference in the validation of both simulations. Additionally, vertical temperature and moisture profiles from the radiosondes provided by the University of Wyoming were used to calculate the three instability indices by our own methodology using the R package *aiRthermo*. According to the validation, the correlation, Standard Deviation (SD) and Root Mean Squared Error (RMSE) obtained by the experiment D for all the indices in most of the stations are better than those for N. The different methods produce small discrepancies between the values for TT, but these are larger for CAPE and CIN due to the dependency of these indices on the initial conditions assumed for the calculation of a lifted air parcel. Similar results arise from the seasonal analysis concerning both WRF experiments: N tends to overestimate or underestimate (depending on the index) the variability of the reference values of the indices, but D is able to capture it in most of the seasons. The heterogeneity of the indices is highlighted in the mean maps over the IP. According to those from D, the unstable air masses are found along the entire Atlantic coast during winter, but in summer they are located particularly over the Mediterranean coast. The values of CAPE in those areas at 12 UTC are much higher than at 00 UTC; The convective inhibition is more extended towards inland at 00 UTC in those areas. However, high values are observed near the southeastern corner of the IP (near Murcia) also at 12 UTC.

1 Introduction

Precipitation is one of the most important variables involved in the water balance, and its variability determines the water resources of the planet. Following the definitions of regional models, precipitation can be separated in two categories: large-scale and convective precipitation, associated for example with frontal systems and unstable atmospheric layers respectively. In general, convective precipitation is usually associated with extreme events due to their high intensity and short duration. However, the simulation of these events is a well-known problem in the modelling community (Sillmann et al., 2013) due to restrictions in the resolution, poor representation of complex topography, forecast errors or deficiencies in the microphysics schemes in the numerical models. In order to avoid these problems, as previously done in the literature (Viceto et al., 2017), this paper focuses on the evaluation of the atmospheric conditions favourable for the development of convective extreme precipitation rather than the validation of the simulation of extreme events.

The evaluation of the atmospheric conditions is typically based on the calculation of some instability indices such as Convective Available Potential Energy (CAPE) (Moncrieff, 1981), Convective INhibition (CIN) (Moncrieff, 1981), Lifted Index (LI) (Galway, 1956), K-Index (George, 1960), Total Totals index (TT) (Miller, 1975) or Showalter Index (S) (Showalter, 1953). All of these indices are commonly used in the literature for this kind of studies (Ye et al., 1998; DeRubertis, 2006; Viceto et al., 2017). CAPE and CIN are based on the adiabatic lifting of a parcel, while most of the others are based on differences in the values of several variables at different pressure levels. The deep convection and thunderstorms are triggered by three ingredients: high levels of moisture in the planetary boundary layer (PBL), potential instability and forced lifting (Johns and Doswell, 1992; McNulty, 1995; Holley et al., 2014; Gascón et al., 2015). CAPE and CIN provide information about the first two ingredients (Holley et al., 2014), and both can provide information about the genesis and intensity of the atmospheric convection (Riemann-Campe et al., 2009). However, previous studies (Angus et al., 1988; López et al., 2001) suggest that CAPE should not be used alone, but should be combined with other indices. The final ingredient (lifting) is provided by the orography (Doswell et al., 1998; Siedlecki, 2009), the convergence of horizontal moisture fluxes (McNulty, 1995) or the breezes in coastal regions (van Delden, 2001). Thus, the high spatial and temporal resolution is important for this kind of studies focusing on atmospheric convection, and that is why regional simulations are needed (Siedlecki, 2009).

The probability of occurrence of precipitation extreme events is not the same along the day, and previous studies support that the maximum convection takes place in the afternoon and evening (Siedlecki, 2009; Virts et al., 2013; Piper and Kunz, 2017; Enno et al., 2020). According to van Delden (2001), the preferred time in most of Western Europe is between 18 and 24 UTC, with the exception of the island of Corsica where the sea breeze triggers convection usually between 6-12 UTC. In open sea areas, the lightning activity peaks in the morning (Enno et al., 2020), associated to thunderstorms triggered by land breezes at night (Virts et al., 2013). A regional study focusing over the UK (Holley et al., 2014) suggests that the reduction overnight of CAPE is over 500 J/Kg.

On the global scale, CAPE follows the spatial pattern of specific humidity and surface air temperature, which means that it increases from pole to Equator (Riemann-Campe et al., 2009). The minimums are obtained in arid regions and over areas with cold water up-welling. Focusing on Europe, convective storms develop for lower values than the U.S. (Kaltenböck et al.,

2009), and several studies tried to determine the most active regions. Romero et al. (2007) found that the most unstable region is located along a zonal belt over the south-central Europe, particularly over the west Mediterranean sea and the surrounding areas. This agrees with Brooks et al. (2003), who found that the favourable environment for thunderstorms is developed in southern Europe, and that the highest number of days in such a regime are located over the Iberian Peninsula (hereafter, IP), south of the Alps, and Northern Balkans. However, van Delden (2001) found that the southwestern France and the Basque Country seem to be a preferred region for the formation of severe storms that drift towards the northeast. More recent studies based on lightning data (Enno et al., 2020) and regional climate models using higher resolution (Rädler et al., 2018; Mohr et al., 2015) highlighted the same areas with favourable environments for thunderstorms in Europe. Particularly, over northern Italy (Po Valley), east of the Adriatic Sea (Albania, Bosnia and Serbia) and in the northeastern IP and Southern France (near the Gulf of Lyon).

Over the IP, the seasonality of precipitation is determined by different sources of moisture due to seasonal variations of the global atmospheric circulation and contrasting climatic regions (influenced by the strong topography). Northern and western IP are mainly affected by stratiform precipitation during winter, while eastern and southern IP receive great amounts of precipitation during autumn due to convective activity (Rodríguez-Puebla et al., 1998; Esteban-Parra et al., 1998; Romero et al., 1999; Iturrioz et al., 2007). Maximum precipitation amounts over central IP are measured in early spring (Tullot, 2000).

Previous studies about instability indices over the IP (Viceto et al., 2017) suggest that CAPE shows a high spatiotemporal variability: the values in winter and spring over land are small due to the reduced surface temperature, and the differences between Atlantic and Mediterranean regions are remarkable during summer. According to Siedlecki (2009), the mean values range from below 50 J/kg in the north to between 100 and 200 J/kg in the Mediterranean coast (some events can even reach 1000 J/Kg). As Romero et al. (2007), Viceto et al. (2017) also stated that CAPE is low during autumn in the Atlantic and continental regions, but high in the areas surrounding the Mediterranean sea. This seasonality was also observed for other indices such as K-index or TT, which show maximum values during summer (Siedlecki, 2009). Observations proved that annual precipitation over eastern stations is mostly accumulated during autumn, as a result of the cumulative warming of the Mediterranean sea due to summer insolation (Romero et al., 2007; Iturrioz et al., 2007), and later entry of very hot and humid air into the IP while cold air is present at higher levels (Dai, 1999; Eshel and Farrell, 2001; Correoso et al., 2006). Additionally, September and October are the months with highest frequency of waterspouts and tornadoes near the Balearic Islands (Gayà et al., 2001). Over the northwestern IP, the mean CAPE when hailstorms occur is 360 J/kg, while for thunderstorms it is only 259 J/kg (López et al., 2001). The dispersion of these values is really high (almost 350 J/kg over the whole sample), which is similar to that found in previous studies (Alexander and Young, 1992; Lucas et al., 1994). The values are similar to those observed in other regions of Europe, but lower than those values obtained in studies based on synoptic or lightning data for severe hailstorms (around 500 J/kg) (Kunz, 2007; Púčik et al., 2015; Taszarek et al., 2017). Due to global warming, the conditions necessary for the development of extreme precipitation events will be enhanced (Brooks, 2013; Rädler et al., 2019). The frequency and intensity of climate extremes will be magnified (Diffenbaugh et al., 2013), projecting larger values of CAPE in the Mediterranean during summer and autumn (Marsh et al., 2009; Viceto et al., 2017).

The main objective of this paper is to evaluate the performance of two simulations created by using the Weather and Research Forecasting (WRF) model (Skamarock et al., 2008) (including or not the extra 3DVAR data assimilation step) at reproducing the atmospheric conditions that can trigger convective precipitation over the IP if the third ingredient (e.g. lifting) is fulfilled. We are not restricting our analysis only to extreme events, and the entire period from 2010-2014 will be considered. For the evaluation, the comparison of pseudo-soundings extracted from the model against real observations will be carried out. Additionally, the seasonal patterns of different instability indices will be studied. The novelty of this study lies in the inclusion of data assimilation in the downscaling experiment used for the analysis of some instability indices.

This paper is organised as follows: The details of the configuration of the WRF model used in both experiments are presented in section 2, along with a brief outline of the methodologies used in the study. The main results are presented in section 3, while they are compared against previous studies presented in the introduction. Finally, we conclude with some remarks about our research in section 4.

2 Data and Methodology

2.1 WRF model configuration

Two experiments were carried out using version 3.6.1 of the WRF model for the period 2010-2014. In both simulations, ERA-Interim provides the initial and boundary conditions (Dee et al., 2011). 6-hourly data at 0.75 degrees were downloaded from the Meteorological Archival and Retrieval System (MARS) repository at European Centre for Medium-Range Weather Forecasts (ECMWF). Analyses of temperature, relative humidity, both wind components and geopotential at 20 pressure levels (5, 10, 20, 30, 50, 70, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 850, 900, 925, 950, 1000 hPa) were used to feed WRF. Both simulations were started on the 1st of January, 2009 from a cold start. Following similar methodologies to previous studies (Argüeso et al., 2011; Zheng et al., 2017), the entire year 2009 was selected as the spin-up for the land surface model included in WRF, and consequently, it was omitted in the study presented here.

One of the experiments (hereafter, N) was nested inside ERA-Interim as usual in numerical downscaling experiments, which means that the model is driven by the boundary conditions after its initialization. It is generated running 6-hour long segments that are restarted from the restart file produced at the end of previous segment, which is similar to a continuous WRF run where the boundary conditions are provided to the model every 6-hours after the initialization of the model. The other experiment (D) presents the same parameterizations as N, but with the additional 3DVAR data assimilation step (Barker et al., 2004, 2012) that is run every 6 hours (at 00, 06, 12 and 18 UTC). In this case, 12-hour long segments starting at every analysis time (00, 06, 12 and 18 UTC) are used. The analyses are generated from the outputs of the model at a 6-hour forecast step from the previous segment as first guess in a 3DVAR data assimilation scheme. In both experiments, the outputs are saved every 3 hours, which means that analyses (00, 06, 12 and 18 UTC) and 3-hour forecasts (at 03, 09, 15 and 21 UTC) are included in our results. In the data assimilation step, quality controlled temperature, moisture, pressure and wind observations in PREPBUFR format from the NCEP ADP Global Upper Air and Surface Weather Observations dataset (referenced as $ds337.0$ at NCAR's Research Data

Archive) were included. Only those observations included in a time-window of two-hours centered in the analysis times were assimilated.

As Figure 1 shows, the domain focuses over the IP, but it also includes parts of Europe, Africa and the Atlantic ocean. As stated by previous studies (Jones et al., 1995; Rummukainen, 2010), the set-up of the domain used in this study prevents border-effects affecting our results as mesoscale systems can develop freely. The spatial resolution of both experiments is 15 km, and with 51 vertical levels up to 20 hPa in eta (η) coordinates.

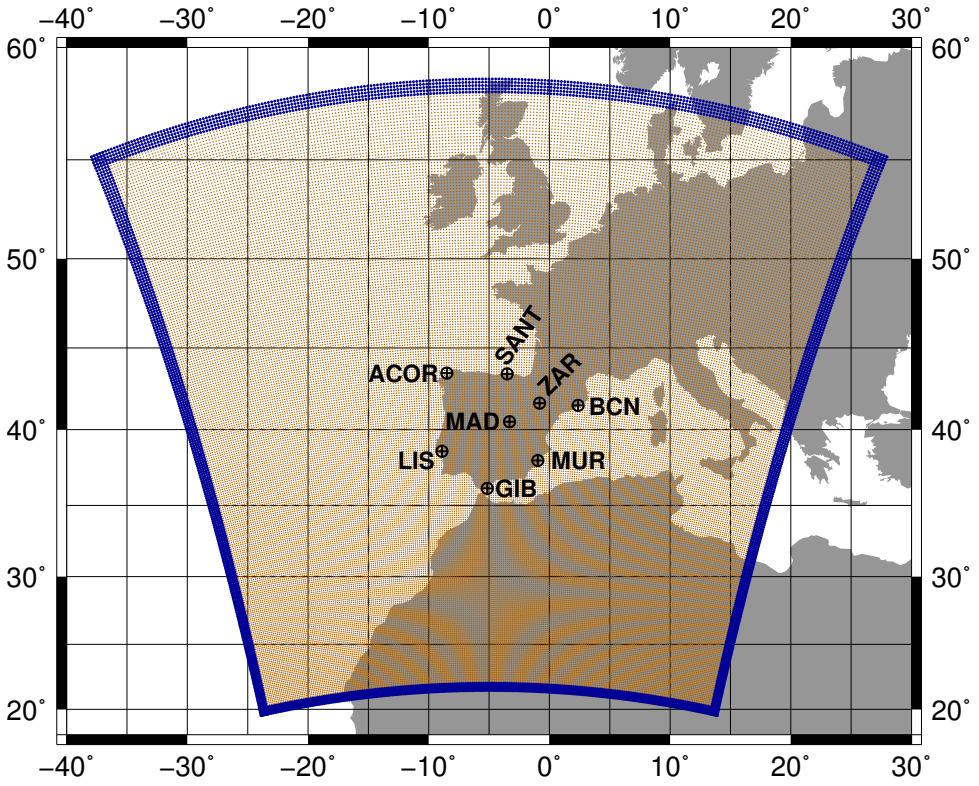

Figure 1. The domain used in both WRF simulations is presented with darkorange dots, while the darkblue region highlights the relaxation zone. The location of all the radiosondes available over the IP is also presented with quartered circles.

Apart from the ERA-Interim data, sea surface temperature (SST) of the model was updated on a daily basis using the high-resolution dataset *NOAA OI SST v2* (Reynolds et al., 2007). Additionally, the following parameterizations for the physics of the model were included in both WRF simulations: five-class microphysics scheme (WSM5) (Hong et al., 2004), MYNN2 planetary boundary layer scheme (Nakanishi and Niino, 2006), Tiedtke cumulus convection scheme (Tiedtke, 1989; Zhang et al., 2011), RRTMG scheme for both long and shortwave radiation (Iacono et al., 2008), and NOAH land surface model (Tewari et al., 2004).

The background error covariance matrices were created before running the simulation with 3DVAR data assimilation. To do so, the CV5 method included in WRFDA (Parrish and Derber, 1992) was used. A separate simulation initialized at 00 and

12 UTC and spanning 13 months (from January 2007 to February 2008) was necessary for the calculation of these matrices. Independent matrices were created for each month, and each of them was calculated taking into account a 90 days period centered on each month.

Both simulations were already presented and validated in previous studies by the authors. Precipitable water, precipitation and evaporation over the IP were validated against station measurements and gridded datasets in González-Rojí et al. (2018), and the outputs produced by D were always superior to N and the driving reanalysis ERA-Interim (for the latter, at least comparable for some variables). The closure of the water balance was also better for D. Additionally, the precipitation from D exhibited similar capabilities to the one downscaled with statistical methods (González-Rojí et al., 2019). Furthermore, the wind field from D showed also improvements compared to ERA-Interim, and consequently, that data were used for the calculation of the offshore wind energy potential in the west Mediterranean (Ulazia et al., 2017). Afterwards, that study was extended to every coast of the IP (Ulazia et al., 2019). The moisture recycling over the IP was also evaluated in González-Rojí et al. (2020), highlighting its importance in the Mediterranean coast during spring and summer.

2.2 Radiosonde data

Atmospheric radiosonde data were downloaded from the server of the University of Wyoming (freely accessible at http://weather.uwyo.edu/upperair/sounding.html). Even if the University of Wyoming does not apply any quality control to the data, this dataset was already used in previous studies by the authors and none of the values were taken as erroneous. Additionally, data from the Integrated Global Radiosonde Archive (IGRA) created by NOAA was also included in this study. This dataset is available online after several quality control procedures.

Only eight radiosondes are available over the IP: A Coruna (ACOR), Santander (SANT), Zaragoza (ZAR), Barcelona (BCN), Madrid (MAD), Lisbon (LIS), Gibraltar (GIB) and Murcia (MUR). The location of each station is presented in Figure 1. Measurements are carried out every day at midday and midnight (00 and 12 UTC, 02 and 14 LT), with the exception of Lisbon where they are only available at 12 UTC (13 LT). Additionally, the amount of data available for Gibraltar is extremely scarce since August 2012.

Temperature and mixing ratio were retrieved at all the available pressure levels at each location. Moreover, the values of the TT, CAPE and CIN as calculated by the University of Wyoming were also retrieved. Additionally, vertical profiles of pressure, temperature and mixing ratio download from the University of Wyoming were also used to calculate TT, CAPE and CIN following our own methodology using the *aiRthermo* R package (further details can be found in the next subsection). The comparison between the original values of the indices retrieved and our results can give us information about whether their discrepancies are only due to differences in the calculation procedure. The values of pressure, temperature and moisture at each level from IGRA were used to calculate TT, CAPE and CIN following our own methodology employing *aiRthermo*. The values computed from the IGRA dataset were assumed as the reference in our analysis.

It must be said that all the radiosondes presented here were assimilated during the 3DVAR data assimilation step in WRF. The effect of the data assimilation step was measured by the analysis increments (analysis minus background) in González-Rojí et al. (2018). The effect of the data assimilation is more intense at 12 UTC compared to the other times, and particularly

for summer (see their Figure 13). The spatial analysis of these values highlighted that the effect of data assimilation is not homogeneous over the IP, and it concentrates mainly in the southeastern IP and both Guadalquivir and Ebro basins. This is related to the well-known cold bias observed in the IP in summer in WRF simulations (Fernández et al., 2007; Argüeso et al., 2011; Jerez et al., 2012), as the data assimilation is able to make it much smaller.

2.3 Methodology

2.3.1 Calculation of instability indices

For both simulations, the nearest grid point to the real latitude and longitude of each radiosonde was determined, and the corresponding pseudo-sounding (pressure, temperature and mixing ratio) at 00 and 12 UTC was obtained at WRF's original η levels. We did not consider the average value of several grid points (e.g., an array of 3×3 grid points) as we would be taking into account an area of 2025 km^2 (45 km $\times$ 45 km), and that would be too much for a comparison against radiosonde data. Additionally, most of the vertical levels up to 6 km are already measured for a drifting distance of 7.5 km, independently of a clear or cloudy day (Xu et al., 2015). That means that our spatial resolution is suitable for the direct comparison of the nearest points against radiosonde data, as we do not neglect the horizontal drift of the radiosoundings.

This procedure was tested for reanalysis (Lee, 2002) and model data (Molina et al., 2020), and it showed that these pseudo-soundings are able to reproduce reasonably well the atmospheric conditions measured by the soundings. However, as highlighted by Holley et al. (2014), this procedure takes into account a stationary column at a fixed time, which can influence the comparison to real radiosonde data as these measurements are not instantaneous and not in a straight vertical line (the balloons take many minutes to measure the profile of the atmosphere and they are deviated because of wind).

In order to calculate the instability indices TT, CAPE and CIN using the pseudo-soundings from the model, the R package *aiRthermo* was used (Sáenz et al., 2019). The most recent version was used (version 1.2.1), which is publicly available in the CRAN repository (https://cran.r-project.org/package=aiRthermo). Both CAPE and CIN are calculated by means of the vertical integrals using discrete slabs defined by the resolution of pressure in the soundings (using all the available levels). The integrals for each of the slabs enclosed by linear profiles are computed analytically, and the energy corresponding to each slab is accumulated, producing the final value of CAPE or CIN. The virtual temperature was used in every integral (Doswell and Rasmussen, 1994). Further details about the functions used for the calculation of the vertical evolution of the air parcels can be found in Sáenz et al. (2019) and also in the manual of the R package *aiRthermo* associated to that publication.

Additionally, in order to calculate CAPE and CIN in the most similar way to the University of Wyoming with the aim of reducing the differences between the values due to different calculation procedures (Siedlecki, 2009), the average of the lower vertical levels was set as the initial representative parcel (Craven et al., 2002; Letkewicz and Parker, 2010). As in Siedlecki (2009), the averaged values from the lowest 500 m were used in this study. Furthermore, in order to avoid that the averaged initial parcel state is still too hot compared to the ambient conditions (in that case, CIN will never be computed as the parcel is already artificially buoyant), an isobaric precooling was applied if needed. To do that, the parcel was cooled along an isobar until it crossed the sounding so that it was not buoyant.

The TT index was calculated following the definition from Miller (1975). It is defined as:

$$TT = (T_{850} - T_{500}) + (D_{850} - T_{500}) \quad (1)$$

where T_{850} and T_{500} are the temperatures at 850 and 500 hPa, and D_{850} is the dewpoint temperature at 850 hPa. According to the ECMWF (Owens and Hewson, 2018), thunderstorms are likely when the values for this index are above 44 °C. However, other values can be found in the literature: 48.1°C for southern Germany (Kunz, 2007), 46.7°C for the Netherlands (Haklander and Van Delden, 2003) or 46°C for Switzerland (Huntrieser et al., 1997).

It can be seen that this index is not highly dependant on the initial conditions for its calculation as it only depends on temperature at two discrete pressure levels, while CAPE and CIN are very sensitive to the initial conditions used for the simulated ascent. TT avoids this problem, but the results can suffer from errors due to inversion layers (Siedlecki, 2009). It must be pointed out that the dew point temperature is needed for TT, and that it is highly important for the calculation of the Lifted Condensation Level (LCL) while calculating CAPE and CIN. In the case of the radiosonde data, the indices are calculated using the measured dew point temperature at 850 hPa when is needed, while in our method, this variable is calculated from the temperature and mixing ratio at that pressure level. This can trigger small differences in the results even if the same original radiosonde data are used.

Further indices could be calculated from the pseudo-soundings obtained from the outputs of the model or real observations, but they were omitted because they can provide similar information to CAPE, CIN and TT. K-index is also based on temperature at different pressure levels, so it suffers from the same problems as TT. Additionally, previous studies reported a strong correlation between CAPE and LI (Blanchard, 1998; López et al., 2001). In order to avoid these connections between indices, we restricted this study to TT, CAPE and CIN indices.

2.3.2 Analysis

Once TT, CAPE and CIN are calculated at the nearest grid points to radiosonde locations of both simulations (N and D), and also those using the original sounding data (labelled as 'aiRthermo' in the results), we obtain a time series with a 12-hourly temporal resolution for each index. These values can be compared against the reference values of the indices retrieved directly from IGRA and the University of Wyoming (labelled as 'Reference' and 'Wyoming' in the next figures). The comparison of 'Wyoming' vs 'aiRthermo' aims to achieve an estimation of the error/differences due to the different methods applied by both sources of results. This comparison was based on independent locations over the IP (separated by several kilometers), so a Taylor diagram was chosen as the best option to show Pearson's correlation (r), Root Mean Squared Error (RMSE) and Standard Deviation (SD) of each experiment in the same plot. In order to determine which experiment is doing the best job at simulating the reference values of the indices, the procedure explained by Taylor (2001) was followed: the dots that lie nearest to the reference on the *X* axis represent variables that agree well with observations (high correlations and low RMSEs), and those lying near the highlighted arc will present comparable standard deviations to the observations.

Additionally, the bootstrap technique with resampling was applied to the results in order to represent an estimation of the sampling errors from each experiment (Efron and Gong, 1983; Wilks, 2011). In our case, the original time series used in the

Taylor diagrams consist of 60 values, each of them for the corresponding month along the period 2010-2014 (12 months × 5 years). For the bootstrap, we created 1000 perturbed time series taking into account different samples of the data. 67 % of the new time series (2/3 of the length of the original time series – 40 values in our case) is made from the original data, and the remaining 33 % (1/3 – 20 values) is chosen from those values already taken from the original data. For each correlation calculated, the same samples are taken from all datasets and experiments. The variability of the Pearson's correlations obtained with these synthetic time series were shown by means of Box-Whiskers plots. To complete the statistical analysis of the bootstrap, the Spearman correlations were also added to these Box-Whiskers plots.

Then, the seasonal analysis of each index at each location was carried out. In this case, the variability of the results is showed by different Box-Whiskers plots. Each season was defined as follows: winter is defined from December to February (DJF), spring from March to May (MAM), summer from June to August (JJA) and autumn from September to November (SON).

Finally, the calculation of each index was extended to every grid point included in a mask defined for the land points of the IP over model's domain. The spatial distribution of the mean values of each index at 00 and 12 UTC during winter and summer was calculated. These maps show the spatial distribution of TT, CAPE and CIN over the land grid points in the IP which are more prevalent in each season.

3 Results

Taylor diagrams for the TT index calculated for each radiosonde of the IP are shown in Figure 2. The Box-Whiskers associated to the correlations (both Pearson's and Spearman's) obtained for each of the 1000 time-series created with the bootstrap technique are also included. According to the Taylor diagrams, as expected, the best experiment reproducing the reference values is Wyoming, followed by aiRthermo (the real measurements of temperature, mixing ratio and pressure from the sounding were used to calculate TT with our methodology), D and later by N. Wyoming obtains the closest values to the observations in all the stations. The results from aiRthermo are quite similar to Wyoming, except for Murcia where D is better reproducing the reference data. The correlations are always above 0.99 for Wyoming, 0.98 for aiRthermo, 0.97 for D and 0.75 for N. The observed SD is really well simulated by Wyoming, aiRthermo and D, but N underestimates it in most of the stations as it is only able to reproduce the one in Santander and A Coruna. The RMSE is below 0.6 °C for aiRthermo, below 1 °C for D and below 2.5 °C for N.

The bootstrap analysis is consistent with the results obtained in the Taylor diagrams, and it shows that the Pearson's correlations are always above 0.99 for Wyoming and 0.98 for aiRthermo (again, with the exception of Murcia where they are above 0.9). The correlations are always above 0.95 for D. In the case of N, the spread of the values is much larger than for aiRthermo and D, and their median values are obtained between 0.8 and 0.9. If we change to Spearman's correlations, we can see that values are similar but with a small worsening of the values (particularly in Gibraltar, Murcia and Madrid).

Thus, as expected, we obtain the most similar results to those calculated from IGRA (Reference) with the values from Wyoming and those calculated with the real measurements from the soundings (that is, aiRthermo). However, we can still

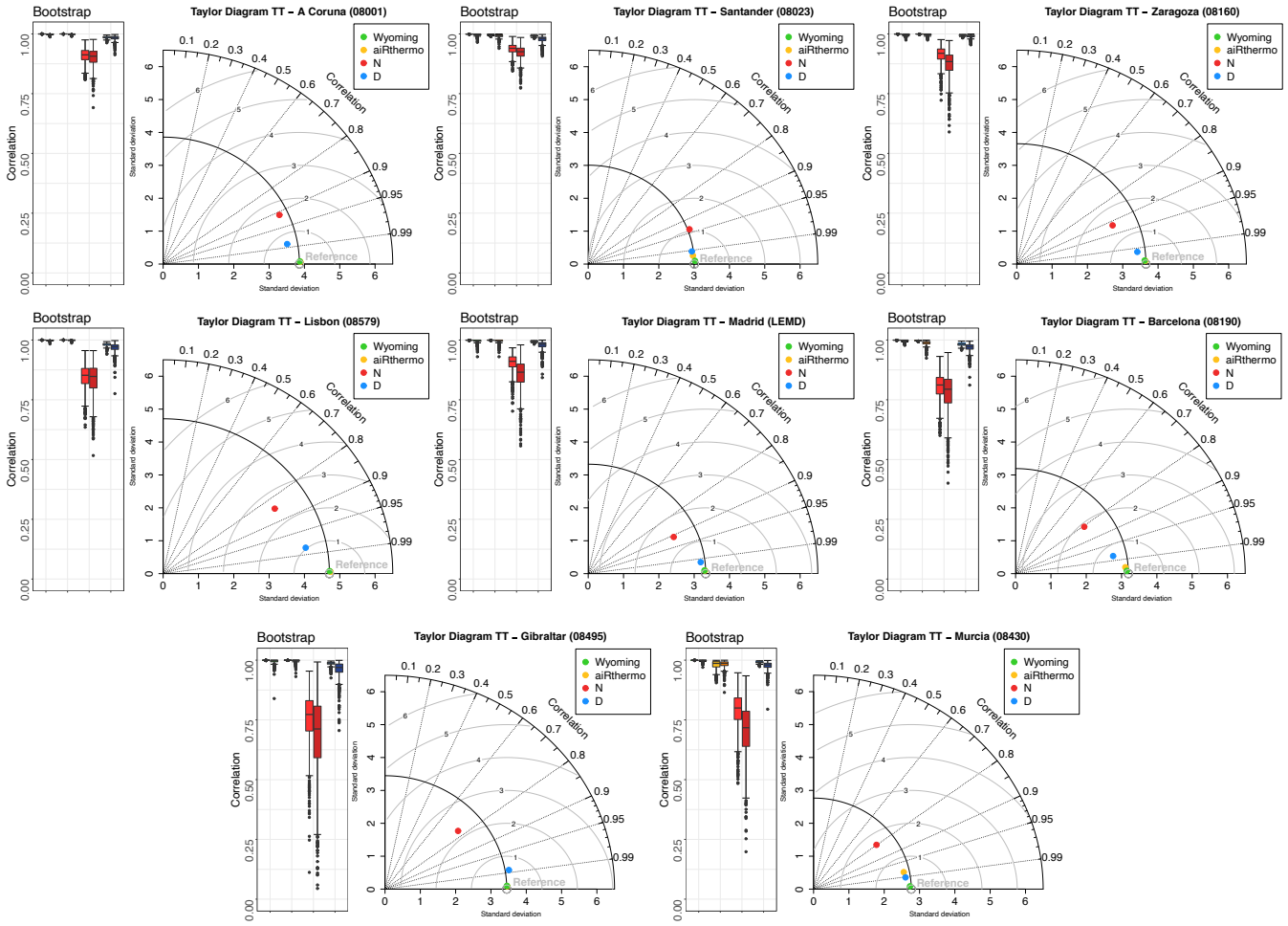

Figure 2. Taylor diagrams showing the r, RMSE and SD values for Wyoming, aiRthermo, N and D compared to TT values retrieved directly from IGRA (Reference). On the left side of each Taylor diagram, a Box-Whiskers plot is added in order to show Pearson's and Spearman's correlations between each experiment and the reference data (lighter and darker colours, or first and second columns of the Box-Whiskers respectively). The bootstrap technique with resampling was used to create 1000 synthetic time-series. Wyoming, aiRthermo, N and D are plotted in green, orange, red and blue respectively.

measure small differences between the values of the datasets due to the use of measured dew point temperature in Wyoming, but the calculation of it from temperature and mixing ratio in IGRA and aiRthermo. These differences are remarkable in Murcia. Between both WRF experiments, it is clear that the experiment including the 3DVAR data assimilation is able to outperform the standard simulation only driven by the reanalysis data at the boundaries of the domain. The differences between both WRF simulations are highlighted, particularly in those stations located in the Mediterranean coast (Barcelona, Murcia and Gibraltar) and in Lisbon.

In the case of CAPE, the validation results are presented in Figure 3. The best experiment reproducing the results is aiRthermo, followed by Wyoming, D and finally by N. The correlations are in all the stations above 0.99 in aiRthermo and 0.95 for Wyoming, while for D they are above 0.9 and above 0.7 for N. A similar behaviour is observed for SD and RMSE. The largest RMSEs are obtained in Barcelona and Murcia (both in the Mediterranean region).

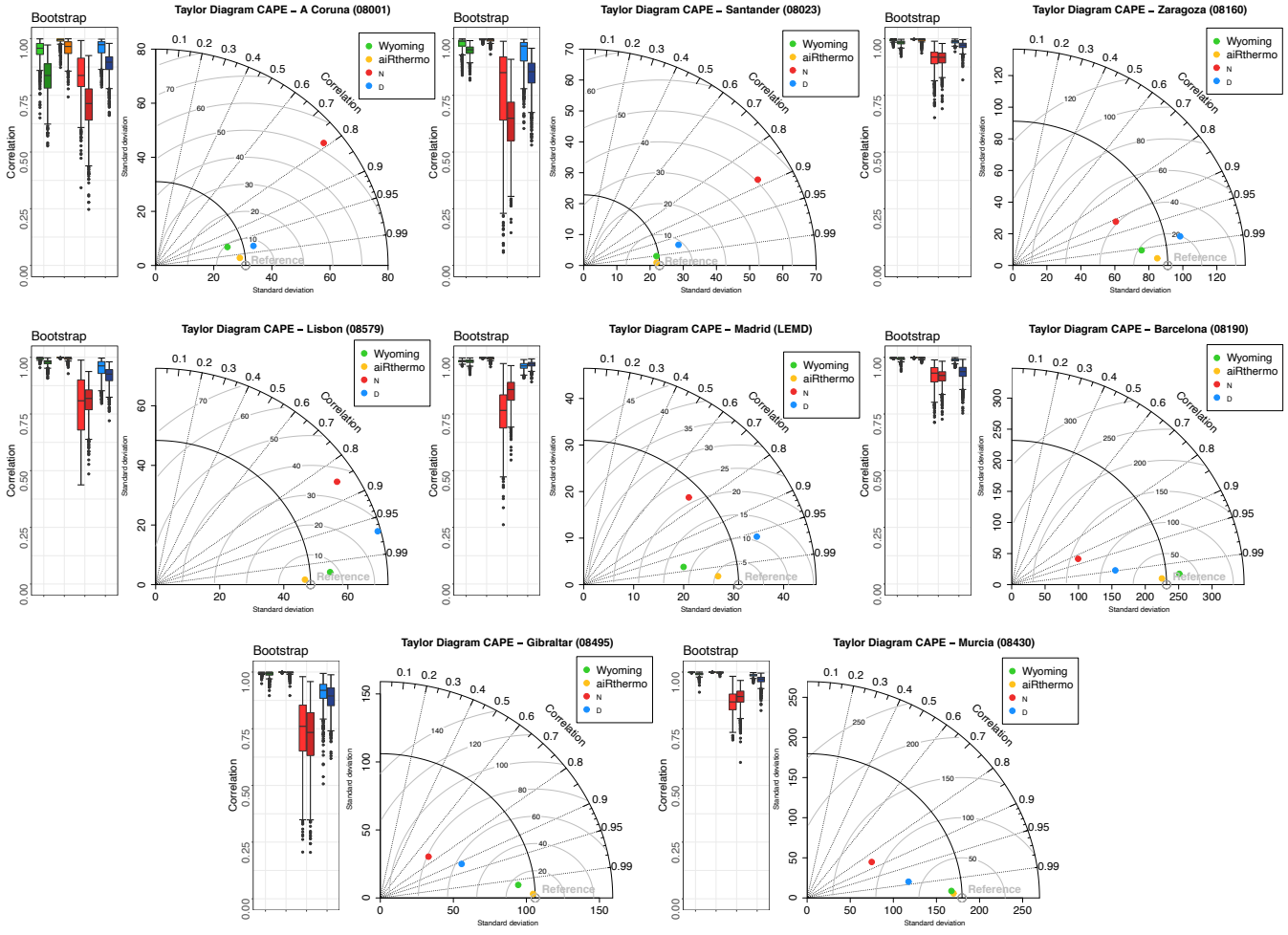

Figure 3. Same as Figure 2 but for CAPE.

The bootstrap analysis shows that the highest Pearson's correlations are obtained by aiRthermo and Wyoming, but followed really closely by D. As for TT index, N presents the worst performance and the largest spread. If we consider, instead, the use of Spearman's correlations, we can see that the values are similar in most of the stations and only in A Coruna and Santander there is a strong worsening of the values.

As stated before in section 2.3.1, the calculation of CAPE is more difficult than TT index, and this is highlighted in the validation of these results. Even if the same data are used for the calculation of CAPE (Wyoming and aiRthermo used the

same measurements), it is clear that small differences in the initial conditions can result in serious discrepancies between both methods as stated by Siedlecki (2009). The largest RMSEs for aiRthermo can be found in Barcelona and Murcia. As for the TT index, two stations in the Mediterranean coast present the largest differences between the experiments. However, while the computation of TT from both WRF simulations produces standard deviations similar to the observed ones, the results for CAPE substantially overestimate the variance of Atlantic sites (A Coruna, Santander, Lisbon and Madrid) or overestimate it in the Mediterranean (Barcelona, Murcia and Gibraltar). Anyway, it can be seen that data assimilation improves the simulation of CAPE over the IP.

Finally, the validation of CIN is presented in Figure 4. As for CAPE, the best results are obtained again by aiRthermo, followed by Wyoming, D and N (with the exception in Gibraltar where D and N are really similar). aiRthermo obtains in every station correlations above 0.97, followed by Wyoming, D and N with correlations above 0.93, 0.85 and 0.65 respectively. Both WRF experiments overestimate or underestimate it depending on the station (particularly N in Lisbon, Madrid, Murcia and Zaragoza). The RMSE is always larger for N, and the values are remarkable in Murcia and Gibraltar for all the experiments.

The bootstrap analysis presents the same results as for CAPE (Figure 3). However, for Gibraltar, as shown in the Taylor diagram, both WRF experiments produce similar Pearson's correlation values during the bootstrap. If we consider, instead, Spearman's correlations, the worsening of the values is perceptible in A Coruna, Santander, Murcia and Gibraltar. However, in Gibraltar, differences between both WRF experiments arise: WRF D obtained better correlations than WRF N as in the other stations. In contrast to previous results, the poorest correlations for CIN are obtained in stations located in the Atlantic coast as Lisbon and A Coruna.

As for CAPE, the differences between aiRthermo and Wyoming are highlighted here. This result supports the idea that small differences in the initial conditions of the lifted air parcel (and the determination of the LCL due to differences in the dew point temperature) can trigger large differences in the values of CIN even if the same values of temperature, mixing ratio and pressure are used for its calculation. Again, the differences between both WRF experiments are important and the experiment including data assimilation (D) presents generally closer results to the observed ones.

The seasonal analysis of the five datasets (Reference, Wyoming, aiRthermo, N and D) for TT index is presented in Figure 5. In this case, Wyoming, aiRthermo and D are able to correctly simulate the reference seasonal variability of TT index in all the stations and all the seasons. However, N tends to overestimate the values of TT in every season and for most of the stations over the IP.

A Coruna and Santander present the largest values during winter, which is in concordance with the regions where more precipitation is accumulated over the IP in that season (Rodríguez-Puebla et al., 1998; Esteban-Parra et al., 1998; Romero et al., 1999; Iturrioz et al., 2007). Higher values than in winter are observed during spring, but the maximum is recognizable in Madrid. This station is the only one located over central IP and it is in concordance with the maximum in precipitation in that region during that season (Tullot, 2000). However, the other stations also present values above 38 °C. During summer, central, eastern and southern stations (Madrid, Barcelona, Zaragoza and Murcia) are the ones presenting higher values. In that season, the Atlantic stations (A Coruna, Santander and Gibraltar) and Gibraltar present values below 40 °C. The values of TT in summer on those stations are shorter than the ones in winter, which occurs mainly due to the combined effect of the high

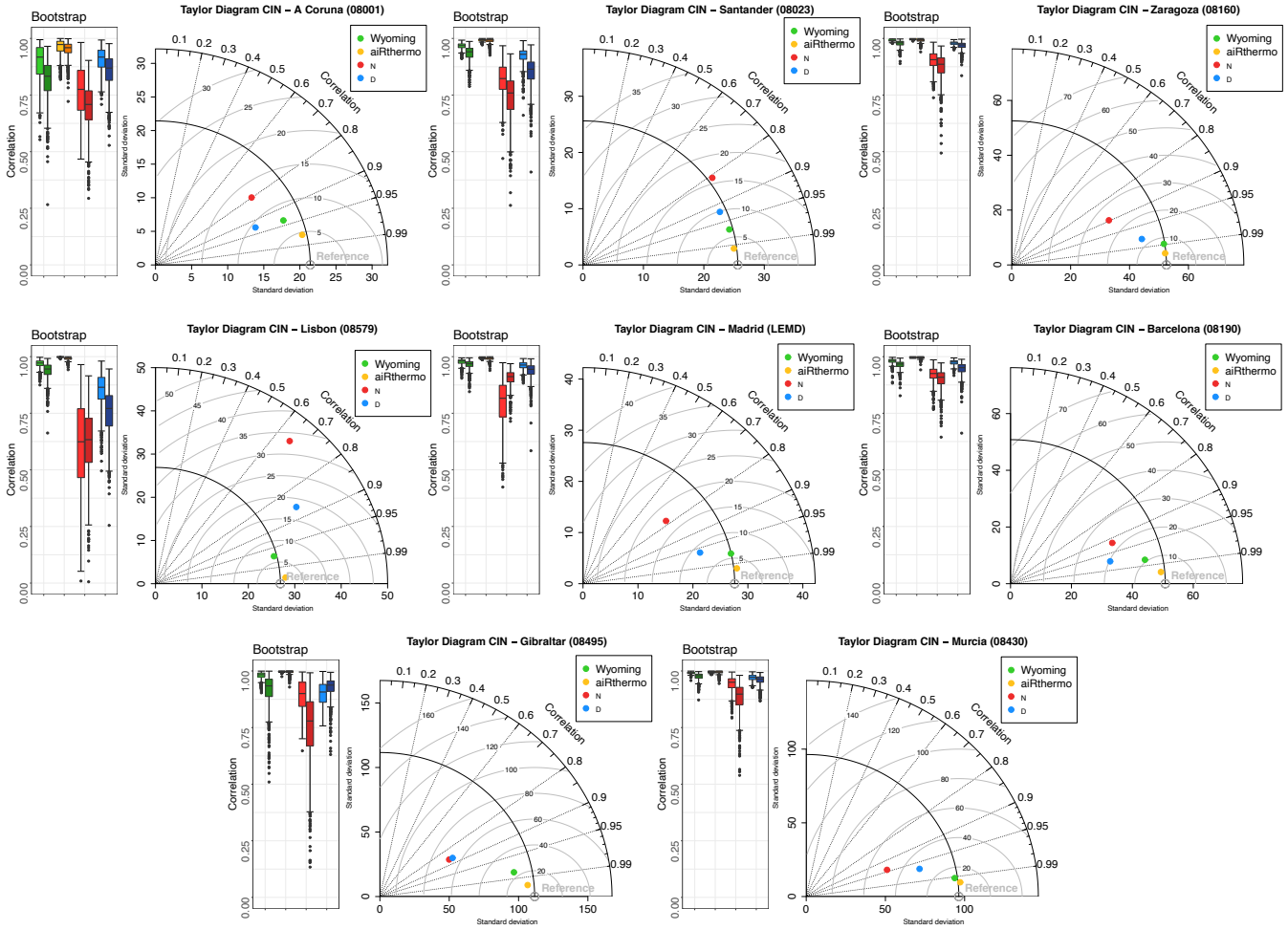

Figure 4. Same as Figures 2 and 3 but for CIN.

increasing values of temperature at 850 and 500 hPa (about 15 and 10 degrees respectively) and the shorter increase of dew point temperature (only a few degrees) in those regions from winter to summer (see Figures A1, A2 and A3 from the Appendix). This pattern is also in agreement with previous studies highlighting that precipitation is important near the Mediterranean coast during that season (Rodríguez-Puebla et al., 1998; Esteban-Parra et al., 1998; Romero et al., 1999; Iturrioz et al., 2007). Finally, all the stations show similar values in autumn, with the exception of Gibraltar where the values are smaller.

The seasonal analysis for CAPE is presented in Figure 6, and it highlights the spatial and temporal heterogeneity of the areas where unstable air masses can be observed over the IP, as also shown by Holley et al. (2014). Wyoming and aiRthermo are able to reproduce (as expected) the variability of the reference values, and D is able to capture the spread of the values in most of the stations during winter, summer and autumn. However, both WRF experiments (particularly D) overestimate CAPE in most of the stations in spring due to the differences in the virtual temperature in lower levels compared to reference data (colder

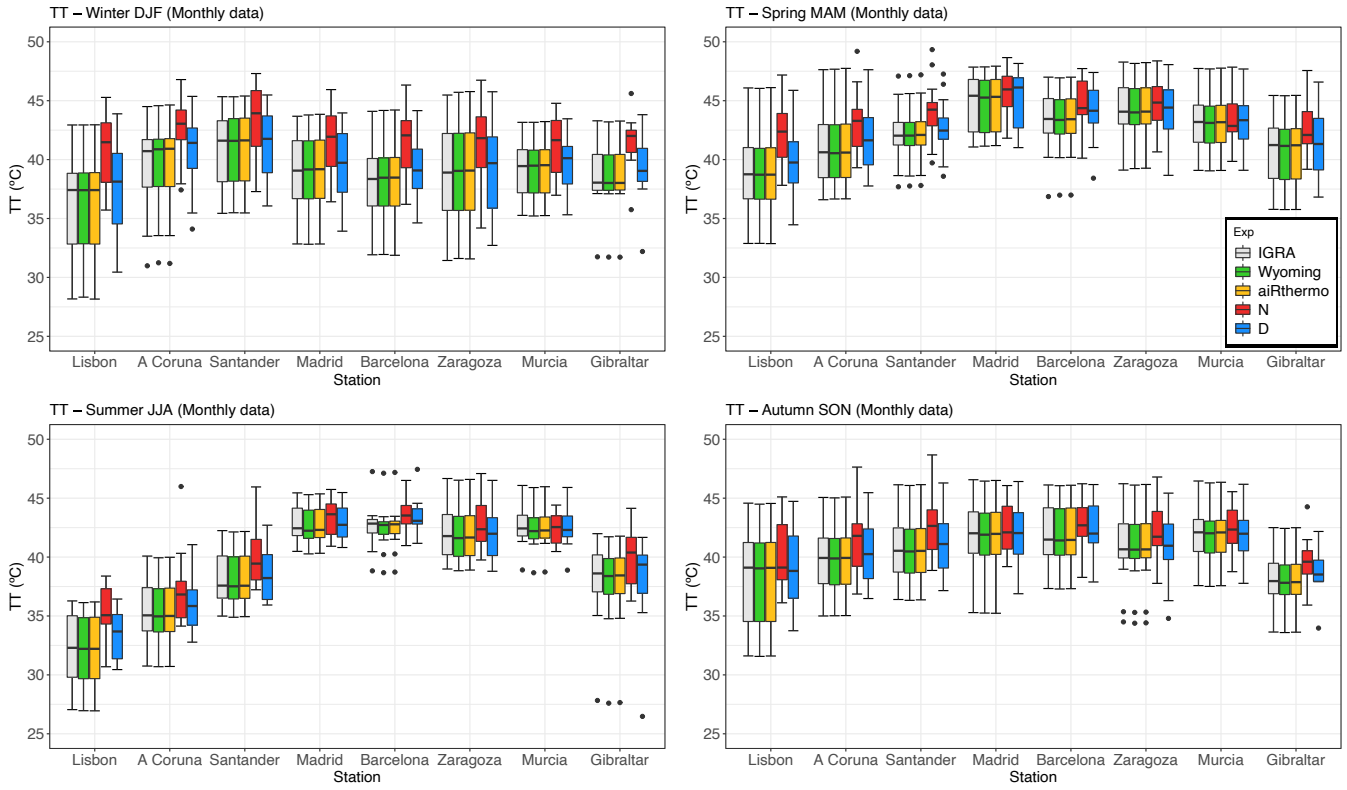

Figure 5. TT index for the reference data (grey), Wyoming (green), aiRthermo (orange), N (red) and D (blue) computed at each station for every season: winter (top panel, left), spring (top panel, right), summer (bottom panel, left) and autumn (bottom panel, right).

near surface and warmer near 800 hPa) and with lifted trajectories for D slightly warmer than the reference ones and N (see Figure A4).

The experiment without data assimilation (N) tends to overestimate CAPE in winter and underestimate it in summer. In winter, this overestimation is caused mainly by colder soundings and warmer lifted trajectories (particularly in Lisbon, A Coruna and Santander - See Figure A5). On the contrary, in summer, the underestimations are caused by warmer sounding levels compared to Reference, which produces that the lifted trajectory crosses earlier than D the sounding, and consequently, underestimating CAPE (particularly in Barcelona, Murcia and Gibraltar - See Figure A6). During spring and autumn, the underestimations or overestimations of N depend on the station and a clear pattern is not observed.

The lowest values of CAPE are obtained during winter (below 50 J/kg in all the stations), and the largest ones are observed in summer (reaching 500 J/kg in some stations). However, as stated before, the distribution of CAPE is not homogeneous and different regions are prone to higher values during each season. During winter, the three Atlantic stations (A Coruna, Santander and Lisbon) and Gibraltar present the highest values of CAPE over the IP. In general, the values are below 50 J/Kg, but some events can get CAPEs over 100 J/kg. During spring, the distribution of CAPE is quite homogeneous over the IP and

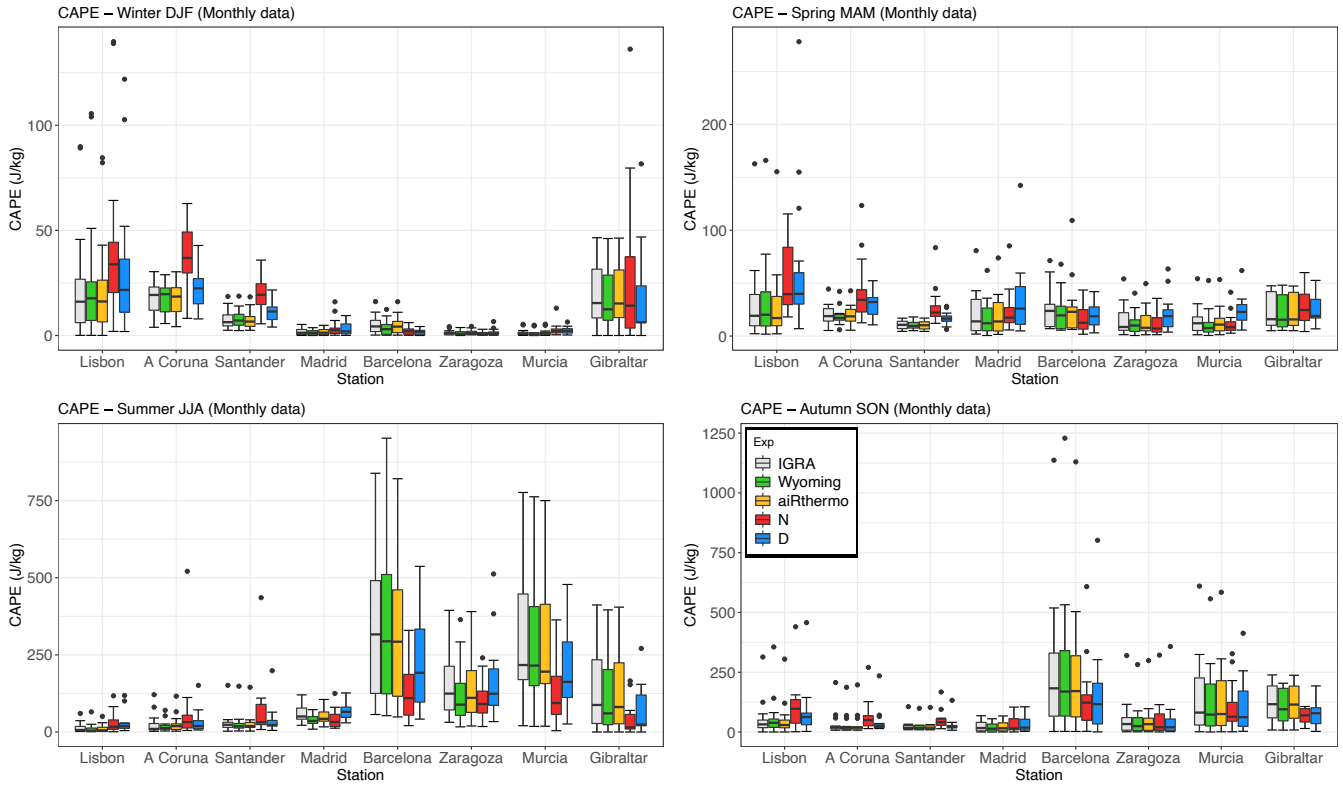

Figure 6. Same as Figure 5 but for CAPE.

only stations such as Lisbon, Madrid or Gibraltar present slightly higher values of CAPE than the other stations. In summer, only the stations located in the eastern and southern parts of the IP present remarkable values of CAPE. Particularly, the most active ones are those located in the Mediterranean coast (Barcelona and Murcia). Finally, during autumn, the regions with high CAPEs are extended towards the inland of the IP, such as Madrid and Zaragoza. During this season, some extreme events can reach values over 1000 J/kg over the Mediterranean coast. This feature was already observed by Siedlecki (2009).

Finally, the seasonal analysis for CIN is presented in Figure 7, and it highlights the stations where the inhibition is important. In general, Wyoming tends to underestimate the values of CIN in most of the stations and in every season, while aiRthermo is able to capture it. Both WRF simulations (but particularly the experiment without data assimilation) tend to underestimate the observed variability.

The values of CIN are smaller in winter and spring, and the maximum is observed in summer. During winter, CIN is higher in Gibraltar and in the Atlantic stations (Lisbon, A Coruna and Santander) than in the other stations from the IP. However, these values are small compared to those from other seasons. During spring, the values are higher than in winter, and similar values are observed in most of the stations (around 10 J/kg), with the exception of Barcelona where the CIN reaches values of 20 J/kg. In summer, the values are higher in all the stations, but particularly in those from the eastern and southern IP (Barcelona,

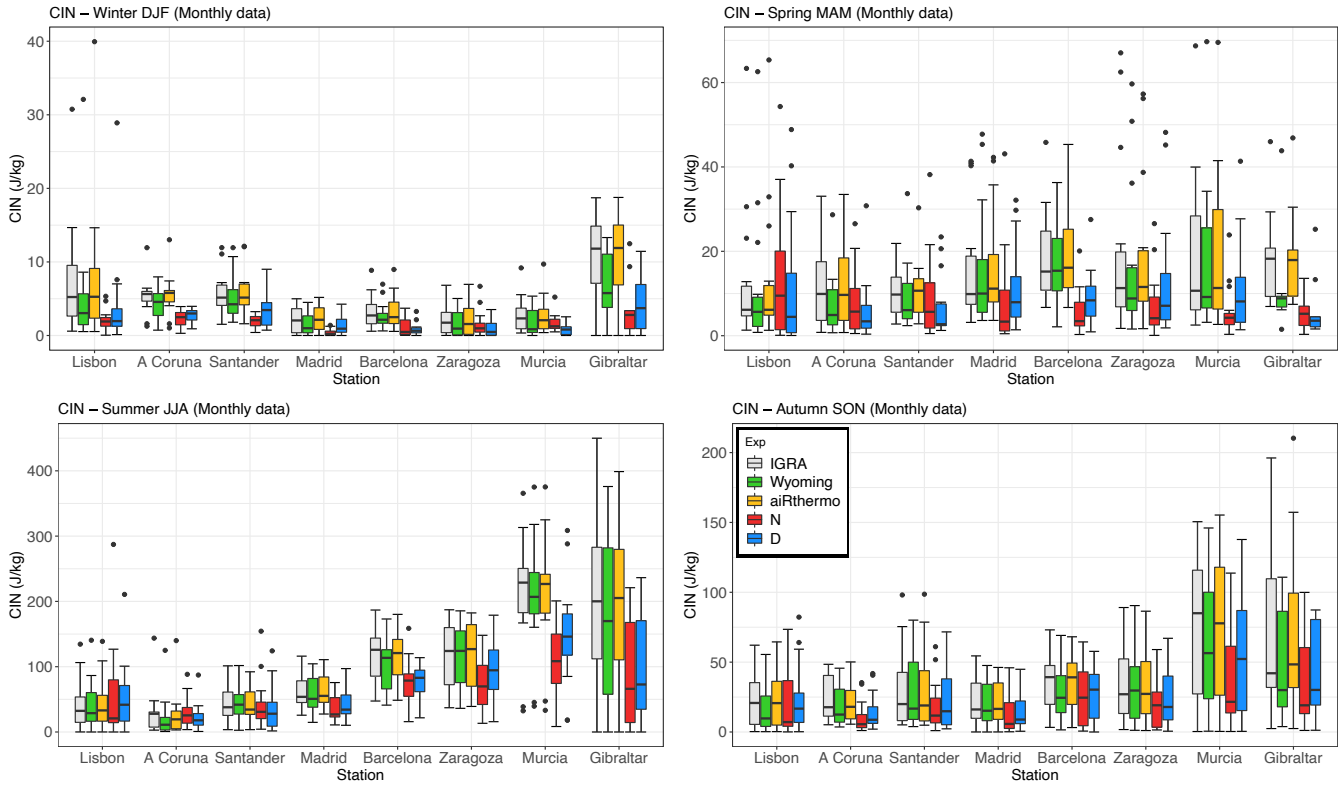

Figure 7. Same as Figures 5 and 6 but for CIN.

Zaragoza, Murcia and Gibraltar). The same regime is observed for autumn, but the values are smaller than in summer. These values during summer and autumn are in agreement with Siedlecki (2009), who found CIN means above 100 J/Kg in the west Mediterranean sea and surrounding countries.

As stated before, in the final phase of this study, the same procedure for the calculation of the instability indices at each station was extended to each grid point included in the IP. The mean winter and summer spatial patterns at 00 and 12 UTC were calculated for both WRF experiments. The spatial distribution for TT is shown in Figure 8, which highlights the heterogeneity of the results. The differences between both simulations are observable, but also those between day and night. Additionally, it can be seen that TT cannot be calculated in most of the mountain regions of the IP because the 850 hPa layer is near the surface or below ground.

During winter, the TT maps show that N yields higher values than D, which is in agreement with the overestimation observed in Figure 5. At 00 UTC, according to D, the regions where unstable air masses are observed are those in the Cantabrian coast and in the southeastern IP. Both regions are surrounded by remarkable mountainous systems such as the Cantabrian Range and the Baetic system, so the lifting that can trigger convection can appear. For N, these areas are also extended to the rest of the IP, with the exception of the southwestern corner where the values are small. At 12 UTC, after solar irradiance has

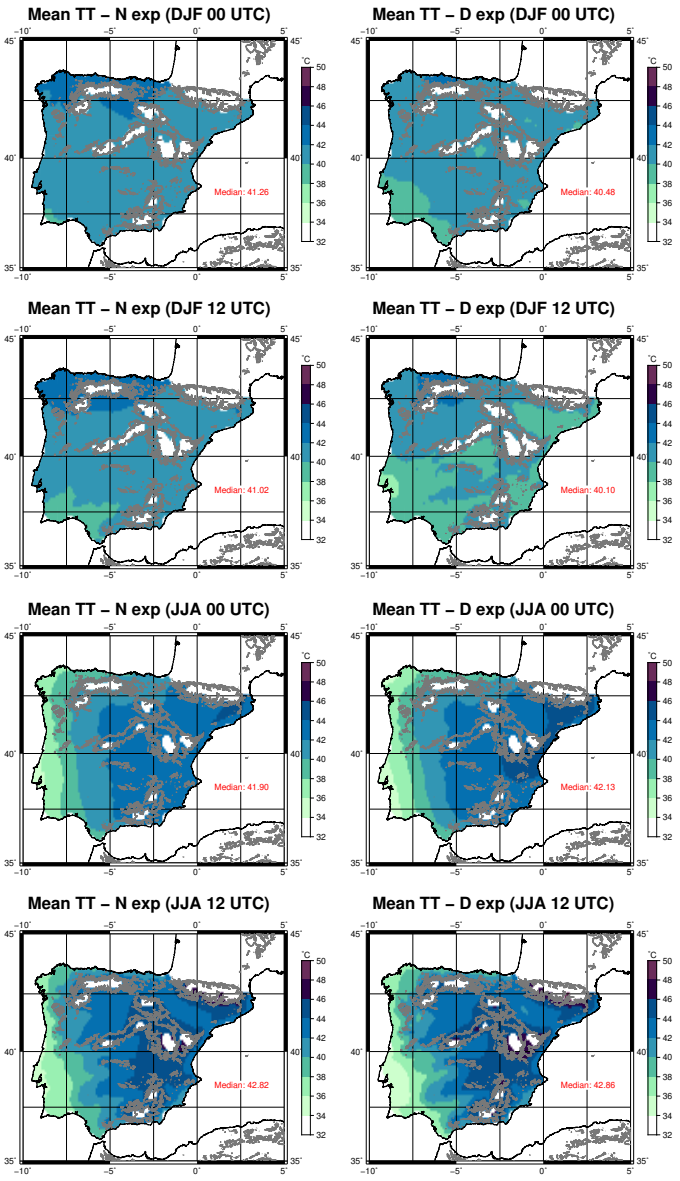

Figure 8. Spatial distribution of mean TT for period 2010-2014 over the IP as computed from N (first column) and D (second column) for winter (rows 1 and 2) and summer (rows 3 and 4) at 00 (rows 1 and 3) and 12 UTC (rows 2 and 4). The median value (°C) of each map is presented in the bottom right corner of the plots.

started heating up the land, the regions extend towards inland areas. In the experiment with data assimilation (D), most of the northern plateau presents high values of TT, and the lowest values are observed near the coastal valleys of the southwestern corner and the Mediterranean coast (like the Ebro basin or Murcia). In the case of N, the lowest values are observed mainly in the southwestern IP near the Guadalquivir valley. According to Figures A1, A2 and A3, the lowest values of TT observed near

the coastal valleys of both WRF experiments are originated due to the low values for dew point temperature observed in those regions, which at the same time are originated by low mixing ratio values.

As in Figure 5, much higher TT values are obtained during summer over the IP, particularly at 12 UTC. At 00 UTC, a west-east gradient is observed in both WRF simulations. However, the values depicted in the Mediterranean coast are higher for the experiment including data assimilation (D). At 12 UTC, the regions with unstable air masses extend towards the central area. In this case, they are located near the mountains of the IP. Particularly, in the southern slope of the Pyrenees and in the proximity of the Iberian, Central and Baetic systems. The minimum TTs are observed in the western part of the IP, but particularly near Lisbon. The intensity of the most extreme values of TT is higher in D (with data assimilation).

The maps for CAPE at 00 and 12 UTC during winter and summer are presented in Figure 9. During winter, as shown in Figure 6, the N experiment presents higher values than D. At 00 UTC, the patterns are really similar for both WRF experiments. The main difference between them is observed in the western Atlantic coast of the IP, where higher CAPE values are obtained for D. At 12 UTC, the unstable air masses are found in the western coast of the IP in both simulations, and they extend further inland that at 00 UTC, particularly near the Tagus and Guadalquivir rivers. Again, the values are higher for N, but the pattern is similar in both experiments.

Compared to what is observed during winter, CAPE is higher during summer for the experiment including data assimilation. This is in agreement with the station analysis from previous Figure 6. At 00 UTC, the area with higher CAPE is observed in the northern and eastern IP, but particularly near the Mediterranean coast. However, at 12 UTC, this area with high values (over 250 J/kg) extends towards the interior and in the experiment including data assimilation it also covers the southern part of the Pyrenees. Additionally, high values are observed in most of the IP (except the southwestern corner for N), but particularly in the simulation including data assimilation.

Finally, regarding CIN, the maps for the mean values at 00 and 12 UTC during winter and summer are presented in Figure 10. In reverse to what we found for CAPE, CIN is usually higher at 00 UTC than at 12 UTC (with the exception of Murcia in summer at 12 UTC). During winter, at 00 UTC, both simulations show small values over the IP, and only some high values are observed in the western and southwestern corner of the IP (and particularly for N). At 12 UTC, the areas are confined to those coastal regions, but they also extend to the Mediterranean coast in the D experiment.

During summer, at 00 UTC, the most remarkable values are obtained in both simulations along the Ebro basin and near the Mediterranean coast. However, the CIN inland is higher for D. At 12 UTC, less inhibition is observed in the eastern valleys of the IP (with the exception of Murcia, which presents extremely high values of CIN). At the same time, an increase in the convective inhibition over the Guadalquivir basin is shown. The extension towards the interior is again higher for D (including data assimilation).

Comparing the results shown in Figures 5 and 6 (or in Figures 8 and 9), it seems that there is a discrepancy between the results from TT and CAPE since maximal values of these indices are not observed in the same regions. However, it must be taken into account that these results for CAPE (and CIN) are obtained from the entire series of 12 hourly values obtained during 2010-2014. Thus, these values are not restricted to highly convective events, and consequently, Figures 6 and 7 (also Figures 9 and 10) must be compared in combination to the values of TT. Additionally, TT and CAPE are indices for atmospheric instability, but

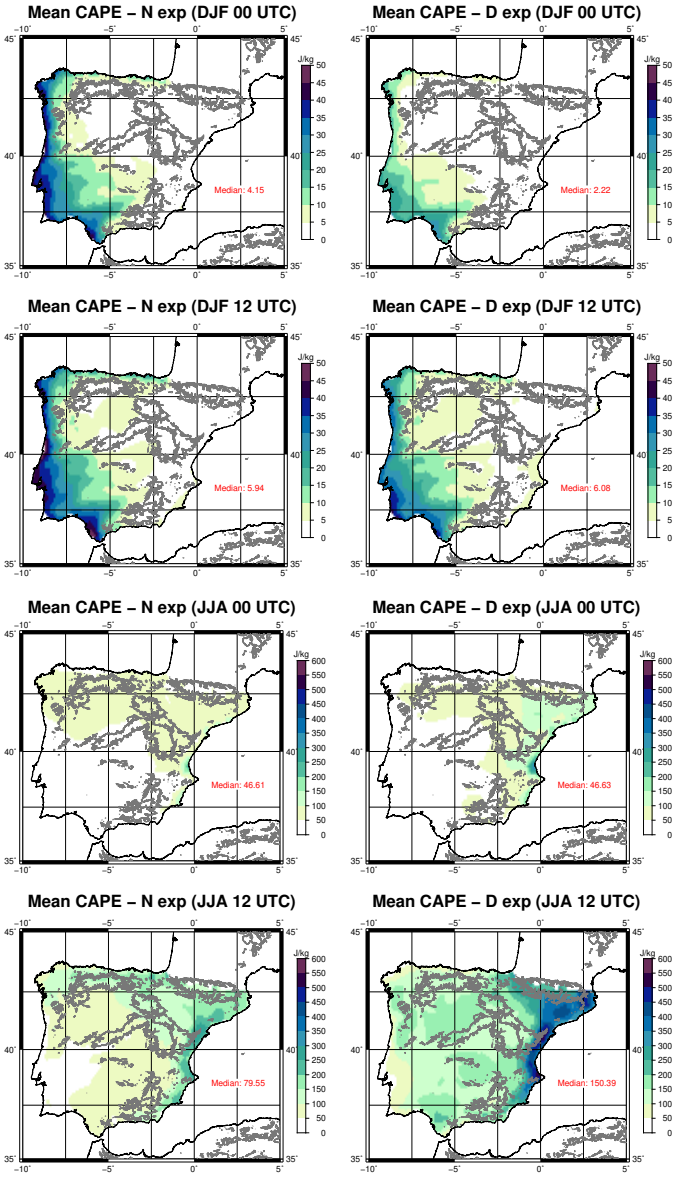

Figure 9. Same as Figure 8 but for CAPE.

they are not related through a simple linear relationship (R^2 below 0.2 for all the stations and seasons), particularly for stable or neutral atmospheres. Thus, since CAPE and CIN are dependent on the entire profile of the atmosphere, they should be considered more reliable than TT, which only takes into account two pressure levels.

Taking in consideration the information presented above, some clear patterns arise from these results. During winter, the areas with unstable air masses are located in the Atlantic coast of the IP, and the instability is more intense during the afternoon as CIN is really high in those region until 12 UTC. However, during summer, the unstable areas are located to the north of the

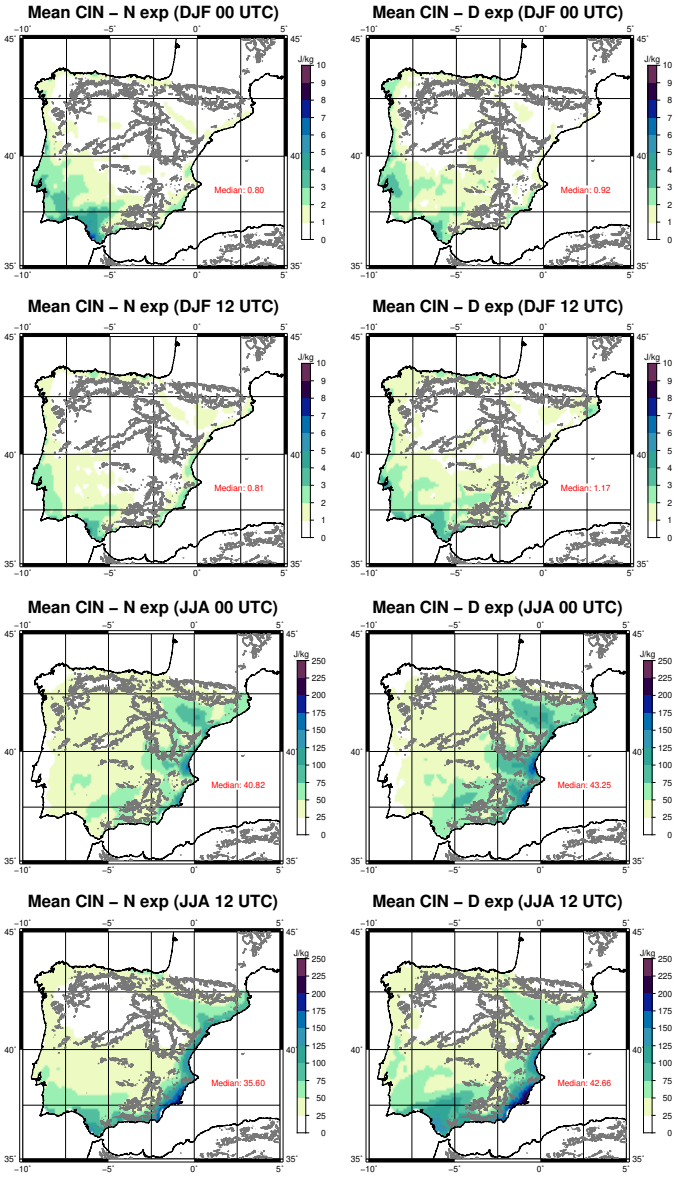

Figure 10. Same as Figures 8 and 10 but for CIN.

Mediterranean coast. The intensity increases during the afternoon in those regions, but some unstable areas can also appear before 12 UTC even if the inhibition is high during that period (CAPE is also high). These features are in agreement with the precipitation dynamics observed in previous studies over the IP (Rodríguez-Puebla et al., 1998; Esteban-Parra et al., 1998; Romero et al., 1999; Iturrioz et al., 2007). The patterns for CAPE observed during winter and summer are similar to those obtained by the regional analysis performed by Viceto et al. (2017). However, their values are comparable to those obtained

by our experiment without data assimilation (N). In this case, the data assimilation (D) produces higher values, but much more realistic than the ones from those simulations (N) without it (according to Figure 6).

4 Conclusions

The main purpose of this paper is to evaluate the ability to simulate unstable conditions over the IP with two high-resolution simulations created with the state-of-the-art model WRF. One of these simulations is driven by the boundary conditions provided by ERA-Interim reanalysis (N experiment), while the second one presents the same configuration but including the additional 3DVAR data assimilation step every 6 hours (D experiment). Three instability indices were evaluated: TT index, CAPE and CIN. All of them were calculated from the outputs of the model using the publicly available R package *aiRthermo*, also developed by the authors.

In order to validate these indices, their values were downloaded from the University of Wyoming server for the eight radiosondes available over the IP. Additionally, temperature, mixing ratio and height at all the available pressure levels from the radiosondes were also retrieved. In that case, also from IGRA from NOAA. These variables were used to calculate with *aiRthermo* again these three indices. Comparing these new values with the ones retrieved directly from the University of Wyoming, the small differences which can only be attributed to different methodologies can be obtained.

First, the correlation, SD and RMSE were computed for each of the indices for both simulations and for the ones calculated with the data at pressure levels from the University of Wyoming following our own methodology. The four of them were compared against the reference values (the ones from IGRA) by means of Taylor diagrams. According to these results, small differences can be observed due to the different methods used for the calculation of the indices (particularly associated to the use of measured or calculated dew point temperature). However, these differences are more important for CAPE and CIN than for TT because they are highly dependant on the initial conditions for the calculation of the vertical integrations, while TT index is only dependant on two discrete pressure levels. Between both WRF simulations, the most accurate results are produced by the experiment with data assimilation (D). The bootstrap analysis with resampling also supports this result.

Then, the seasonal analysis was carried out for each index. For TT, the differences between methods are really small in every season. Between both WRF experiments, N tends to overestimate the reference variability, while D is able to capture it. In the case of CAPE and CIN, the differences between methods are larger, but not as those within both WRF experiments. However, D is able to produce closer values to the reference than N. During winter, the unstable air masses are located mainly over the stations from the Atlantic coast of the IP. All the stations are quite active during spring. However, the regime changes to the Mediterranean coast during summer, and also autumn but with less intensity.

Finally, the calculation of the indices was also carried out for every grid point over the IP in both WRF simulations, particularly for winter and summer and at 00 and 12 UTC. All the three indices agree highlighting the heterogeneity of the patterns observed. The D experiment, which is the most accurate one according to the previous analysis, shows that during winter the unstable areas are found along the entire Atlantic coast, but particularly in the southwestern corner of the IP when the instability is extended towards inland regions. During summer, this feature is reversed, and the region most prone to develop unstable air

masses is located in the Ebro basin and the Mediterranean coast. The inhibition (high values of CIN) is strong at 00 UTC in those regions, but that is highly reduced at 12 UTC.

Data availability. These results can be reproduced using the postprocessed outputs from the model available in https://doi.org/10.5281/zenodo.3611343.

Appendix A: Supplementary figures

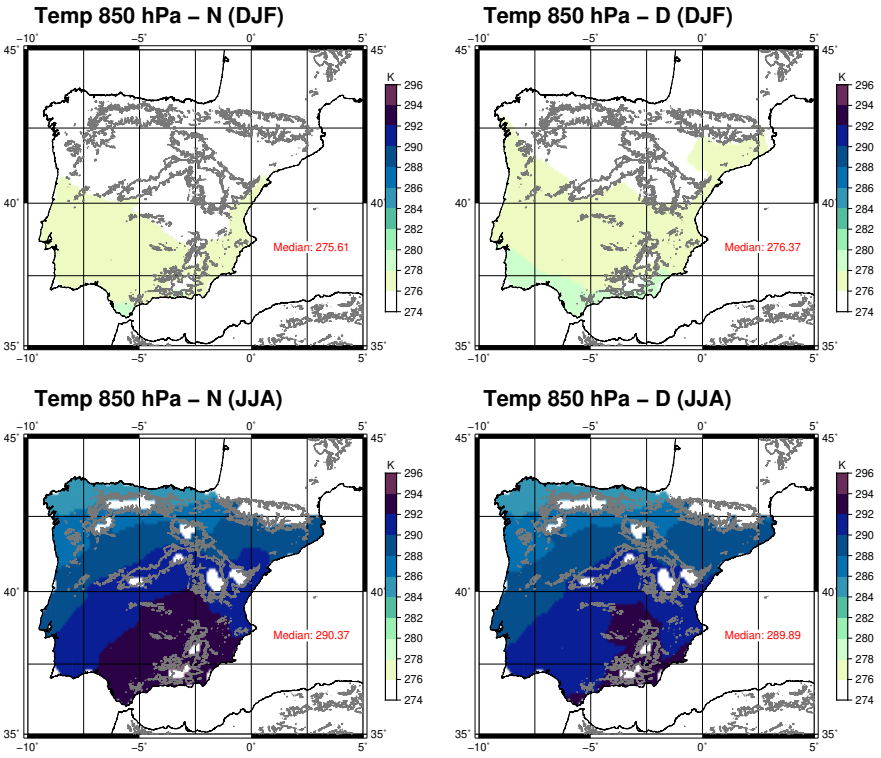

Figure A1. Spatial distribution of mean temperature at 850 hPa for period 2010-2014 over the IP as computed from N (first column) and D (second column) for winter and summer. The median value (K) is in the bottom right corner of the plots.

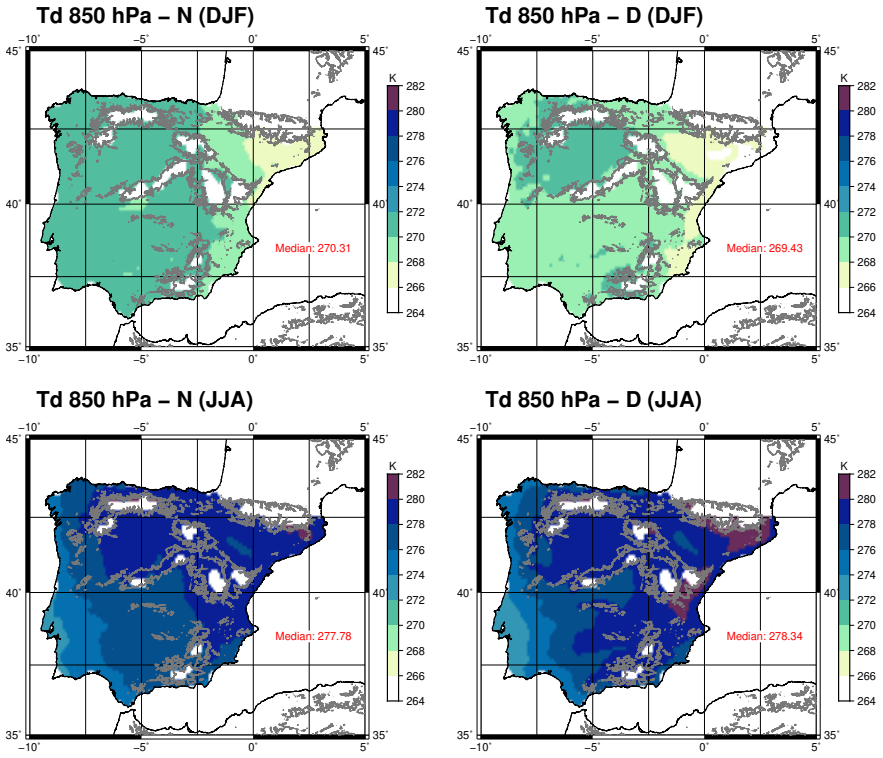

Figure A2. Same as Figure A1 but for dew point temperature at 850 hPa.

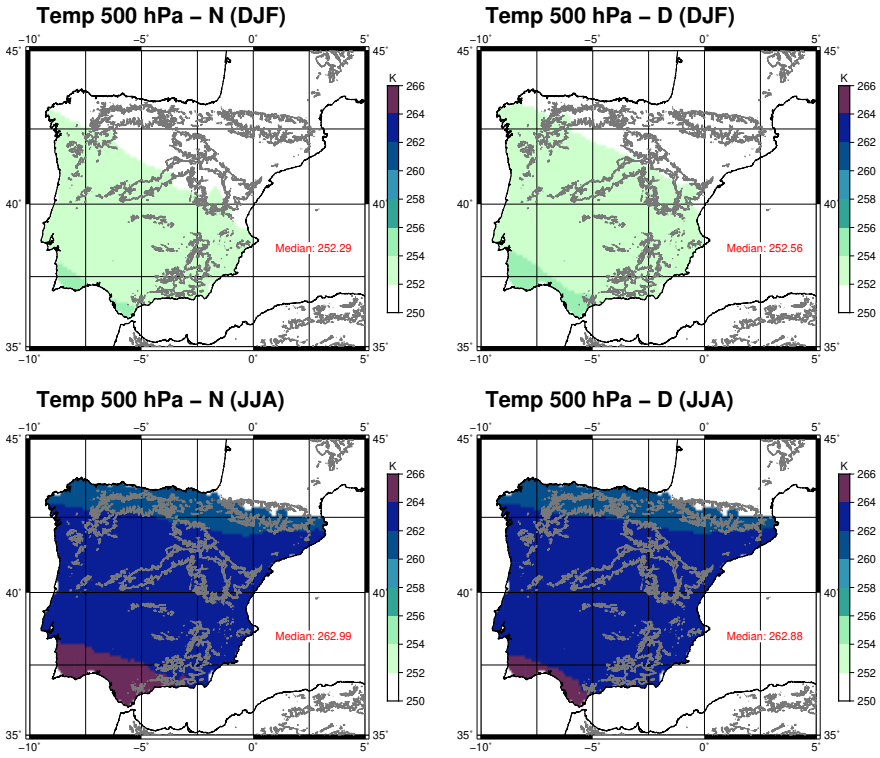

Figure A3. Same as Figures A1 and A2 but for temperature at 500 hPa.

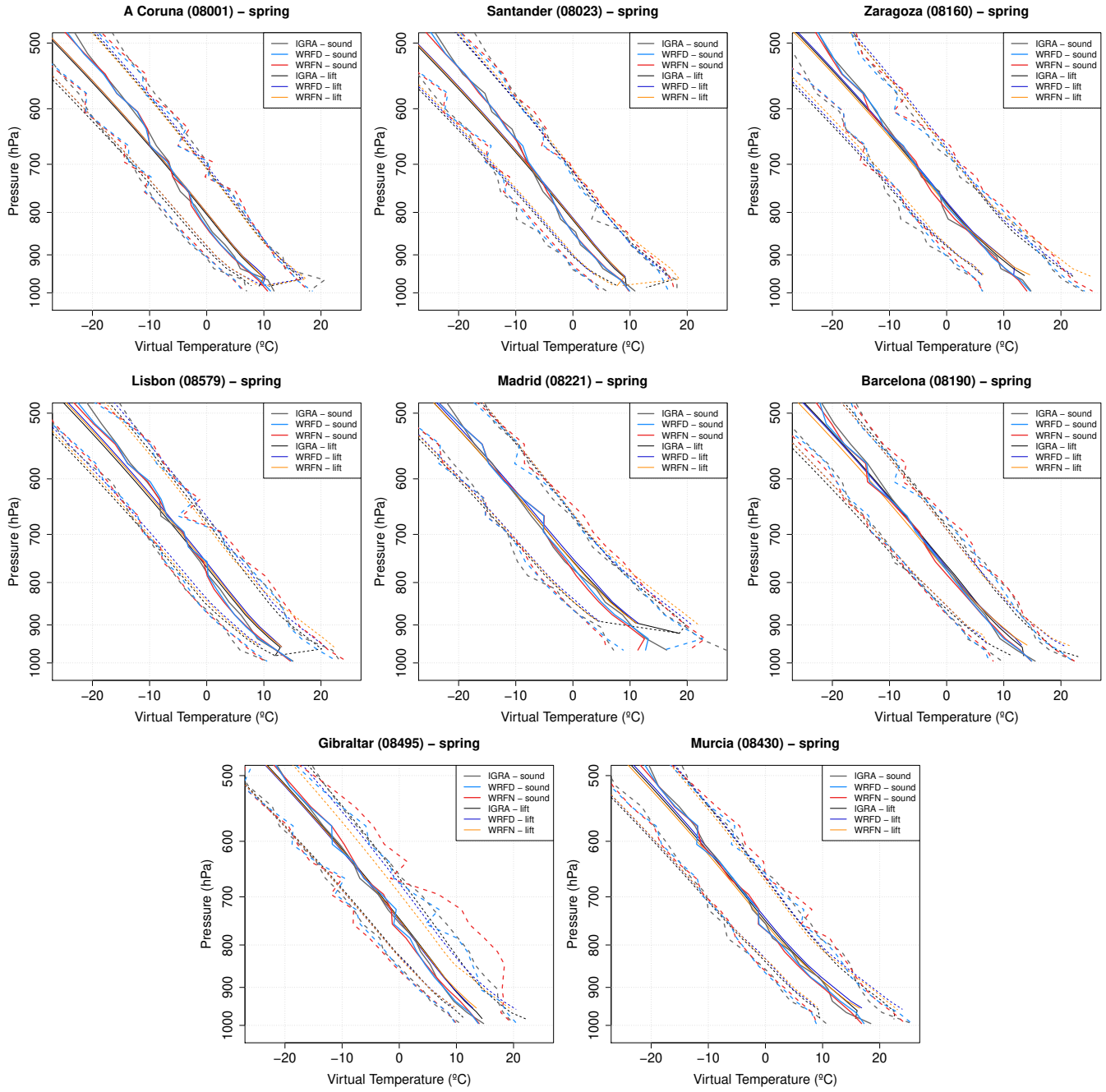

Figure A4. Vertical profiles of virtual temperature for the sounding levels and for the lifted parcel during spring. The dashed lines represent the 5 and 95 percentiles, and the solid lines the median.

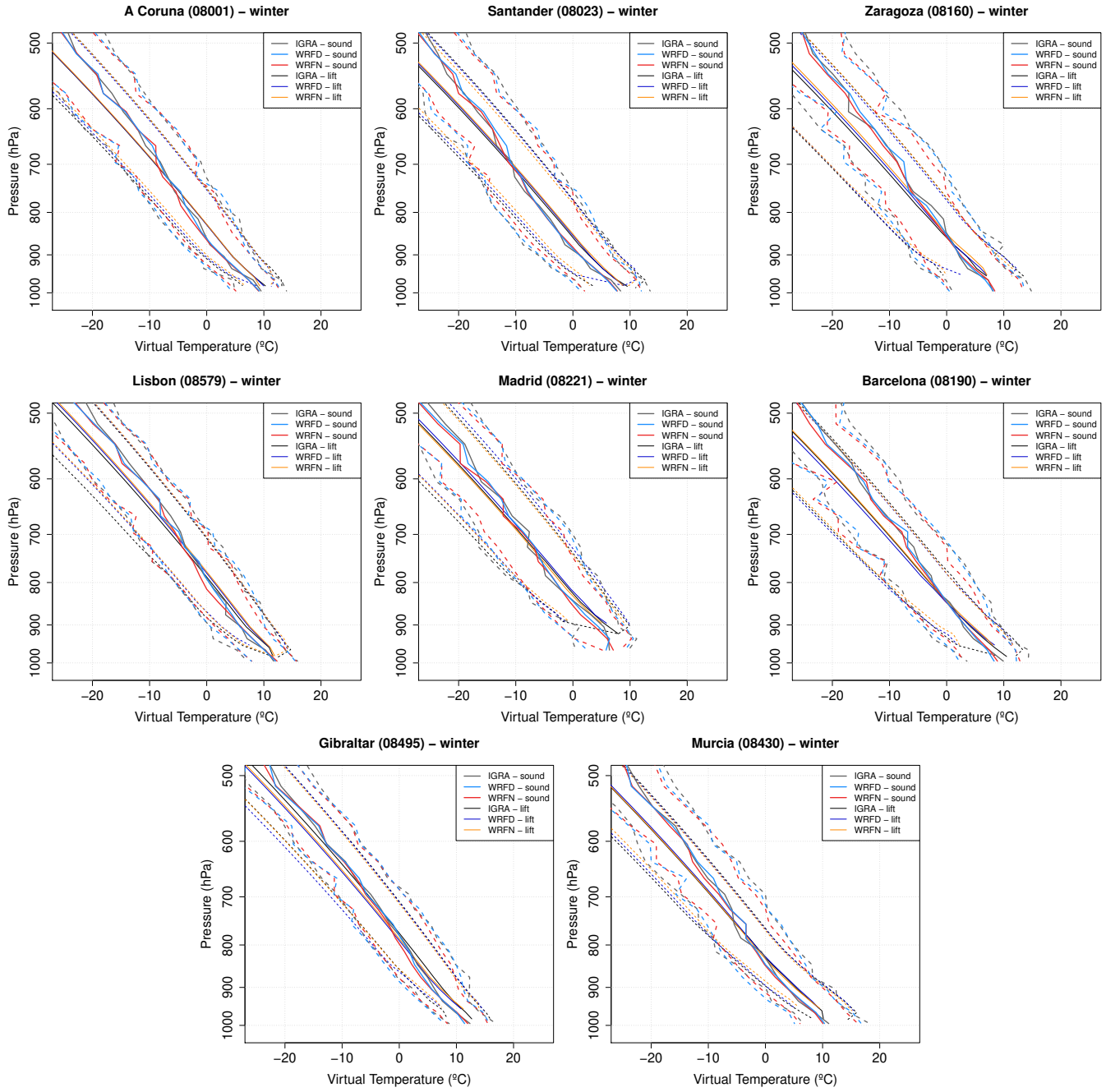

Figure A5. Same as Figure A4 but for winter.

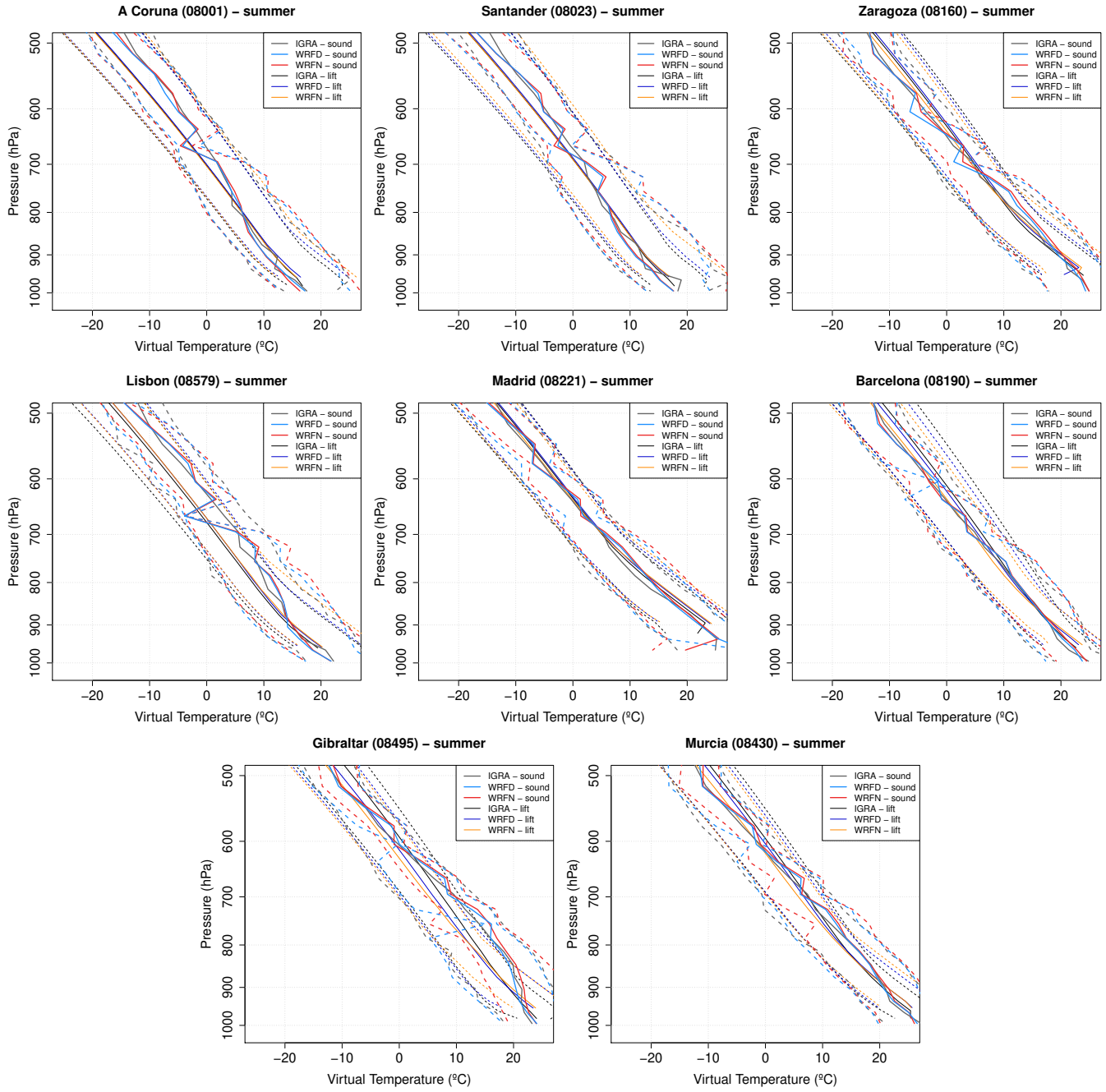

Figure A6. Same as Figures A4 and A5 but for summer.

Author contributions. The methodology and the software was developed by S.J.G.-R., S.C.-M. and J.S.; The conceptualization, preparation of datasets and analysis was carried out by all the authors; The original draft of the paper was written by S.J.G.-R., but all the authors took part in the edition and revision of it.

Competing interests. The authors declare no conflict of interest.

Acknowledgements. S.J.G.R. is now funded by the Oeschger Centre for Climate Change Research (OCCR), but during his PhD he was supported by a FPI Predoctoral Research grant (MINECO, BES-2014-069977). This study was also supported through MINECO project CGL2016-76561-R from the Spanish Government (MINECO/ERDF, UE) and the grant GIU 17/002 from the University of the Basque Country. The computational resources were provided by I2BASQUE, and the authors thank the creators of WRF/ARW and WRFDA systems. Authors thank the anonymous reviewers for their comments, which have helped to improve the paper. Finally, most of the calculations were carried out with R (R Core Team, 2018), and the authors want to thank all the authors of the packages used for it.

References

Alexander, G. D. and Young, G. S.: The Relationship between EMEX Mesoscale Precipitation Feature Properties and Their Environmental Characteristics, Monthly Weather Review, 120, 554–564, https://doi.org/10.1175/1520-0493(1992)120<0554:TRBEMP>2.0.CO;2, 1992.

Angus, P., Rasmussen, S., and Seiter, K.: Short-term prediction of thunderstorm probability and intensity by screening observational and derived predictors, in: Reprints AMS 15th Conference on Severe Local Storms, Baltimore, MA, pp. 368–371, 1988.

Argüeso, D., Hidalgo-Muñoz, J. M., Gámiz-Fortis, S. R., Esteban-Parra, M. J., Dudhia, J., and Castro-Díez, Y.: Evaluation of WRF Parameterizations for Climate Studies over Southern Spain Using a Multistep Regionalization, Journal of Climate, 24, 5633–5651, https://doi.org/10.1175/JCLI-D-11-00073.1, 2011.

Barker, D., Huang, X.-Y., Liu, Z., Auligné, T., Zhang, X., Rugg, S., Ajjaji, R., Bourgeois, A., Bray, J., Chen, Y., Demirtas, M., Guo, Y.-R., Henderson, T., Huang, W., Lin, H.-C., Michalakes, J., Rizvi, S., and Zhang, X.: The Weather Research and Forecasting Model's Community Variational/Ensemble Data Assimilation System: WRFDA, Bulletin of the American Meteorological Society, 93, 831–843, https://doi.org/10.1175/BAMS-D-11-00167.1, 2012.

Barker, D. M., Huang, W., Guo, Y.-R., Bourgeois, A. J., and Xiao, Q. N.: A Three-Dimensional Variational Data Assimilation System for MM5: Implementation and Initial Results, Monthly Weather Review, 132, 897–914, https://doi.org/10.1175/1520-0493(2004)132<0897:ATVDAS>2.0.CO;2, 2004.

Blanchard, D. O.: Assessing the Vertical Distribution of Convective Available Potential Energy, Weather and Forecasting, 13, 870–877, https://doi.org/10.1175/1520-0434(1998)013<0870:ATVDOC>2.0.CO;2, 1998.

Brooks, H.: Severe thunderstorms and climate change, Atmospheric Research, 123, 129 – 138, https://doi.org/https://doi.org/10.1016/j.atmosres.2012.04.002, 2013.

Brooks, H. E., Lee, J. W., and Craven, J. P.: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data, Atmospheric Research, 67-68, 73 – 94, https://doi.org/https://doi.org/10.1016/S0169-8095(03)00045-0, 2003.

Correoso, J. F., Hernández, E., García-Herrera, R., Barriopedro, D., and Paredes, D.: A 3-year study of cloud-to-ground lightning flash characteristics of Mesoscale convective systems over the Western Mediterranean Sea, Atmospheric Research, 79, 89 – 107, https://doi.org/https://doi.org/10.1016/j.atmosres.2005.05.002, 2006.

Craven, J. P., Jewell, R. E., and Brooks, H. E.: Comparison between Observed Convective Cloud-Base Heights and Lifting Condensation Level for Two Different Lifted Parcels, Weather and Forecasting, 17, 885–890, https://doi.org/10.1175/1520-0434(2002)017<0885:CBOCCB>2.0.CO;2, 2002.

Dai, A.: Recent changes in the diurnal cycle of precipitation over the United States, Geophysical Research Letters, 26, 341–344, https://doi.org/10.1029/1998GL900318, 1999.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

DeRubertis, D.: Recent Trends in Four Common Stability Indices Derived from U.S. Radiosonde Observations, Journal of Climate, 19, 309–323, https://doi.org/10.1175/JCLI3626.1, 2006.

Diffenbaugh, N. S., Scherer, M., and Trapp, R. J.: Robust increases in severe thunderstorm environments in response to greenhouse forcing, Proceedings of the National Academy of Sciences, 110, 16 361–16 366, https://doi.org/10.1073/pnas.1307758110, 2013.

Doswell, C. A. and Rasmussen, E. N.: The Effect of Neglecting the Virtual Temperature Correction on CAPE Calculations, Weather and Forecasting, 9, 625–629, https://doi.org/10.1175/1520-0434(1994)009<0625:TEONTV>2.0.CO;2, 1994.

Doswell, C. A., Ramis, C., Romero, R., and Alonso, S.: A Diagnostic Study of Three Heavy Precipitation Episodes in the Western Mediterranean Region, Weather and Forecasting, 13, 102–124, https://doi.org/doi:10.1175/1520-0434(1998)013<0102:ADSOTH>2.0.CO;2, 1998.

Efron, B. and Gong, G.: A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation, The American Statistician, 37, 36–48, https://doi.org/10.1080/00031305.1983.10483087, 1983.

Enno, S.-E., Sugier, J., Alber, R., and Seltzer, M.: Lightning flash density in Europe based on 10 years of ATDnet data, Atmospheric Research, 235, 104 769, https://doi.org/https://doi.org/10.1016/j.atmosres.2019.104769, 2020.

Eshel, G. and Farrell, B. F.: Thermodynamics of Eastern Mediterranean Rainfall Variability, Journal of the Atmospheric Sciences, 58, 87–92, https://doi.org/10.1175/1520-0469(2001)058<0087:TOEMRV>2.0.CO;2, 2001.

Esteban-Parra, M. J., Rodrigo, F. S., and Castro-Diez, Y.: Spatial and temporal patterns of precipitation in Spain for the period 1880–1992, International Journal of Climatology, 18, 1557–1574, https://doi.org/10.1002/(SICI)1097-0088(19981130)18:14<1557::AID-JOC328>3.0.CO;2-J, 1998.

Fernández, J., Montávez, J. P., Sáenz, J., González-Rouco, J. F., and Zorita, E.: Sensitivity of the MM5 mesoscale model to physical parameterizations for regional climate studies: Annual cycle, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/10.1029/2005JD006649, 2007.

Galway, J. G.: The lifted index as a predictor of latent instability, Bulletin of the American Meteorological Society, 37, 528–529, 1956.

Gascón, E., Merino, A., Sánchez, J., Fernández-González, S., García-Ortega, E., López, L., and Hermida, L.: Spatial distribution of thermodynamic conditions of severe storms in southwestern Europe, Atmospheric Research, 164-165, 194 – 209, https://doi.org/https://doi.org/10.1016/j.atmosres.2015.05.012, 2015.

Gayà, M., Homar, V., Romero, R., and Ramis, C.: Tornadoes and waterspouts in the Balearic Islands: phenomena and environment characterization, Atmospheric Research, 56, 253 – 267, https://doi.org/https://doi.org/10.1016/S0169-8095(00)00076-4, 2001.

George, J. J.: Weather forecasting for aeronautics, Academic Press, San Diego, page 411, 1960.

González-Rojí, S. J., Sáenz, J., Ibarra-Berastegi, G., and Díaz de Argandoña, J.: Moisture balance over the Iberian Peninsula according to a regional climate model: The impact of 3DVAR data assimilation., Journal of Geophysical Research: Atmospheres, 123, 708–729, https://doi.org/10.1002/2017JD027511, 2018.

González-Rojí, S. J., Wilby, R. L., Sáenz, J., and Ibarra-Berastegi, G.: Harmonized evaluation of daily precipitation downscaled using SDSM and WRF+WRFDA models over the Iberian Peninsula, Climate Dynamics, 53, 1413–1433, https://doi.org/10.1007/s00382-019-04673-9, 2019.

González-Rojí, S. J., Sáenz, J., Díaz de Argandoña, J., and Ibarra-Berastegi, G.: Moisture Recycling over the Iberian Peninsula: The Impact of 3DVAR Data Assimilation, Atmosphere, 11, 19, https://doi.org/10.3390/atmos11010019, 2020.

Haklander, A. J. and Van Delden, A.: Thunderstorm predictors and their forecast skill for the Netherlands, Atmospheric Research, 67-68, 273 – 299, https://doi.org/https://doi.org/10.1016/S0169-8095(03)00056-5, 2003.

Holley, D. M., Dorling, S. R., Steele, C. J., and Earl, N.: A climatology of convective available potential energy in Great Britain, International Journal of Climatology, 34, 3811–3824, https://doi.org/10.1002/joc.3976, 2014.

Hong, S.-Y., Dudhia, J., and Chen, S.-H.: A Revised Approach to Ice Microphysical Processes for the Bulk Parameterization of Clouds and Precipitation, Monthly Weather Review, 132, 103–120, https://doi.org/10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2, 2004.

Huntrieser, H., Schiesser, H. H., Schmid, W., and Waldvogel, A.: Comparison of Traditional and Newly Developed Thunderstorm Indices for Switzerland, Weather and Forecasting, 12, 108–125, https://doi.org/10.1175/1520-0434(1997)012<0108:COTAND>2.0.CO;2, 1997.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2008JD009944, 2008.

Iturrioz, I., Hernández, E., Ribera, P., and Queralt, S.: Instability and its relation to precipitation over the Eastern Iberian Peninsula, Advances in Geosciences, 10, 45–50, https://doi.org/10.5194/adgeo-10-45-2007, 2007.

Jerez, S., Montavez, J. P., Gomez-Navarro, J. J., Jimenez, P. A., Jimenez-Guerrero, P., Lorente, R., and Gonzalez-Rouco, J. F.: The role of the land-surface model for climate change projections over the Iberian Peninsula, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/10.1029/2011JD016576, 2012.

Johns, R. H. and Doswell, C. A.: Severe Local Storms Forecasting, Weather and Forecasting, 7, 588–612, https://doi.org/10.1175/1520-0434(1992)007<0588:SLSF>2.0.CO;2, 1992.

Jones, R. G., Murphy, J. M., and Noguer, M.: Simulation of climate change over europe using a nested regional-climate model. I: Assessment of control climate, including sensitivity to location of lateral boundaries, Quarterly Journal of the Royal Meteorological Society, 121, 1413–1449, https://doi.org/10.1002/qj.49712152610, 1995.

Kaltenböck, R., Diendorfer, G., and Dotzek, N.: Evaluation of thunderstorm indices from ECMWF analyses, lightning data and severe storm reports, Atmospheric Research, 93, 381 – 396, https://doi.org/https://doi.org/10.1016/j.atmosres.2008.11.005, 2009.

Kunz, M.: The skill of convective parameters and indices to predict isolated and severe thunderstorms, Natural Hazards and Earth System Sciences, 7, 327–342, https://doi.org/10.5194/nhess-7-327-2007, 2007.

Lee, J.: Tornado Proximity Soundings from the NCEP/NCAR Reanalysis Data, Ph.D. thesis, University of Oklahoma, 2002.

Letkewicz, C. E. and Parker, M. D.: Forecasting the Maintenance of Mesoscale Convective Systems Crossing the Appalachian Mountains, Weather and Forecasting, 25, 1179–1195, https://doi.org/10.1175/2010WAF2222379.1, 2010.

López, L., Marcos, J. L., Sánchez, J. L., Castro, A., and Fraile, R.: CAPE values and hailstorms on northwestern Spain, Atmospheric Research, 56, 147 – 160, https://doi.org/https://doi.org/10.1016/S0169-8095(00)00095-8, 2001.

Lucas, C., Zipser, E. J., and LeMone, M. A.: Convective Available Potential Energy in the Environment of Oceanic and Continental Clouds: Correction and Comments, Journal of the Atmospheric Sciences, 51, 3829–3830, https://doi.org/10.1175/1520-0469(1994)051<3829:CAPEIT>2.0.CO;2, 1994.

Marsh, P. T., Brooks, H. E., and Karoly, D. J.: Preliminary investigation into the severe thunderstorm environment of Europe simulated by the Community Climate System Model 3, Atmospheric Research, 93, 607 – 618, https://doi.org/https://doi.org/10.1016/j.atmosres.2008.09.014, 2009.

McNulty, R. P.: Severe and Convective Weather: A Central Region Forecasting Challenge, Weather and Forecasting, 10, 187–202, https://doi.org/10.1175/1520-0434(1995)010<0187:SACWAC>2.0.CO;2, 1995.

Miller, R. C.: Notes on analysis and severe-storm forecasting procedures of the Air Force Global Weather Central, vol. 200, AWS Technical Report, 1975.

Mohr, S., Kunz, M., and Geyer, B.: Hail potential in Europe based on a regional climate model hindcast, Geophysical Research Letters, 42, 10,904–10,912, https://doi.org/10.1002/2015GL067118, 2015.

Molina, D. S., Fernández-González, S., González, J. C. S., and Oliver, A.: Analysis of sounding derived parameters and application to severe weather events in the Canary Islands, Atmospheric Research, 237, 104 865, https://doi.org/https://doi.org/10.1016/j.atmosres.2020.104865, 2020.

Moncrieff, M. W.: A theory of organized steady convection and its transport properties, Quarterly Journal of the Royal Meteorological Society, 107, 29–50, https://doi.org/10.1002/qj.49710745103, 1981.

Nakanishi, M. and Niino, H.: An Improved Mellor–Yamada Level-3 Model: Its Numerical Stability and Application to a Regional Prediction of Advection Fog, Boundary-Layer Meteorology, 119, 397–407, https://doi.org/10.1007/s10546-005-9030-8, 2006.

Owens, R. G. and Hewson, T.: ECMWF Forecast User Guide, ECMWF, https://doi.org/10.21957/m1cs7h, 2018.

Parrish, D. F. and Derber, J. C.: The National Meteorological Center's Spectral Statistical-Interpolation Analysis System, Monthly Weather Review, 120, 1747–1763, https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2, 1992.

Piper, D. and Kunz, M.: Spatiotemporal variability of lightning activity in Europe and the relation to the North Atlantic Oscillation teleconnection pattern, Natural Hazards and Earth System Sciences, 17, 1319–1336, https://doi.org/10.5194/nhess-17-1319-2017, 2017.

Púčik, T., Groenemeijer, P., Rýva, D., and Kolář, M.: Proximity Soundings of Severe and Nonsevere Thunderstorms in Central Europe, Monthly Weather Review, 143, 4805–4821, https://doi.org/10.1175/MWR-D-15-0104.1, 2015.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/, 2018.

Rädler, A. T., Groenemeijer, P., Faust, E., and Sausen, R.: Detecting Severe Weather Trends Using an Additive Regressive Convective Hazard Model (AR-CHaMo), Journal of Applied Meteorology and Climatology, 57, 569–587, https://doi.org/10.1175/JAMC-D-17-0132.1, 2018.

Rädler, A. T., Groenemeijer, P. H., Faust, E., Sausen, R., and Púčik, T.: Frequency of severe thunderstorms across Europe expected to increase in the 21st century due to rising instability, Npj Climate and Atmospheric Science, 2, 30, https://doi.org/10.1038/s41612-019-0083-7, 2019.

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., and Schlax, M. G.: Daily High-Resolution-Blended Analyses for Sea Surface Temperature, Journal of Climate, 20, 5473–5496, https://doi.org/10.1175/2007JCLI1824.1, 2007.

Riemann-Campe, K., Fraedrich, K., and Lunkeit, F.: Global climatology of Convective Available Potential Energy (CAPE) and Convective Inhibition (CIN) in ERA-40 reanalysis, Atmospheric Research, 93, 534 – 545, https://doi.org/https://doi.org/10.1016/j.atmosres.2008.09.037, 2009.

Rodríguez-Puebla, C., Encinas, A. H., Nieto, S., and Garmendia, J.: Spatial and temporal patterns of annual precipitation variability over the Iberian Peninsula, International Journal of Climatology, 18, 299–316, https://doi.org/10.1002/(SICI)1097-0088(19980315)18:3<299::AID-JOC247>3.0.CO;2-L, 1998.

Romero, R., Ramis, C., and Guijarro, J.: Daily rainfall patterns in the Spanish Mediterranean area: an objective classification, International Journal of Climatology, 19, 95–112, https://doi.org/10.1002/(SICI)1097-0088(199901)19:1<95::AID-JOC344>3.0.CO;2-S, 1999.

Romero, R., Gayà, M., and Doswell, C. A.: European climatology of severe convective storm environmental parameters: A test for significant tornado events, Atmospheric Research, 83, 389 – 404, https://doi.org/https://doi.org/10.1016/j.atmosres.2005.06.011, 2007.

Rummukainen, M.: State-of-the-art with regional climate models, Wiley Interdisciplinary Reviews: Climate Change, 1, 82–96, https://doi.org/10.1002/wcc.8, https://onlinelibrary.wiley.com/doi/abs/10.1002/wcc.8, 2010.

Sáenz, J., González-Rojí, S. J., Carreno-Madinabeitia, S., and Ibarra-Berastegi, G.: Analysis of atmospheric thermodynamics using the R package aiRthermo, Computers & Geosciences, 122, 113 – 119, https://doi.org/https://doi.org/10.1016/j.cageo.2018.10.007, 2019.

Showalter, A. K.: A Stability Index for Thunderstorm Forecasting, Bulletin of the American Meteorological Society, 34, 250–252, 1953.

Siedlecki, M.: Selected instability indices in Europe, Theoretical and Applied Climatology, 96, 85–94, https://doi.org/10.1007/s00704-008-0034-4, 2009.

Sillmann, J., Kharin, V. V., Zhang, X., Zwiers, F. W., and Bronaugh, D.: Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate, Journal of Geophysical Research: Atmospheres, 118, 1716–1733, https://doi.org/10.1002/jgrd.50203, 2013.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G.: A Description of the Advanced Research WRF Version 3, NCAR Technical Note NCAR/TN-475+STR, https://doi.org/10.5065/D68S4MVH, 2008.

Taszarek, M., Brooks, H. E., and Czernecki, B.: Sounding-Derived Parameters Associated with Convective Hazards in Europe, Monthly Weather Review, 145, 1511–1528, https://doi.org/10.1175/MWR-D-16-0384.1, 2017.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, Journal of Geophysical Research: Atmospheres, 106, 7183–7192, 2001.

Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M., Mitchell, K., Ek, M., Gayno, G., Wegiel, J., and Cuenca, R.: Implementation and verification of the unified NOAH land surface model in the WRF model, in: 20th conference on weather analysis and forecasting/16th conference on numerical weather prediction, vol. 1115, 2004.

Tiedtke, M.: Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models, Monthly Weather Review, 117, 1779–1800, https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2, 1989.

Tullot, I. F.: Climatología de España y Portugal, vol. 76, Universidad de Salamanca, 2000.

Ulazia, A., Sáenz, J., Ibarra-Berastegui, G., González-Rojí, S. J., and Carreno-Madinabeitia, S.: Using 3DVAR data assimilation to measure offshore wind energy potential at different turbine heights in the West Mediterranean, Applied Energy, 208, 1232 – 1245, https://doi.org/https://doi.org/10.1016/j.apenergy.2017.09.030, 2017.

Ulazia, A., Ibarra-Berastegi, G., Sáenz, J., Carreno-Madinabeitia, S., and González-Rojí, S. J.: Seasonal Correction of Offshore Wind Energy Potential due to Air Density: Case of the Iberian Peninsula, Sustainability, 11, https://doi.org/10.3390/su11133648, 2019.

van Delden, A.: The synoptic setting of thunderstorms in western Europe, Atmospheric Research, 56, 89 – 110, https://doi.org/https://doi.org/10.1016/S0169-8095(00)00092-2, 2001.

Viceto, C., Marta-Almeida, M., and Rocha, A.: Future climate change of stability indices for the Iberian Peninsula, International Journal of Climatology, 37, 4390–4408, https://doi.org/10.1002/joc.5094, 2017.

Virts, K. S., Wallace, J. M., Hutchins, M. L., and Holzworth, R. H.: Highlights of a New Ground-Based, Hourly Global Lightning Climatology, Bulletin of the American Meteorological Society, 94, 1381–1391, https://doi.org/10.1175/BAMS-D-12-00082.1, 2013.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, vol. 100, Academic Press, 2011.

Xu, G., Xi, B., Zhang, W., Cui, C., Dong, X., Liu, Y., and Yan, G.: Comparison of atmospheric profiles between microwave radiometer retrievals and radiosonde soundings, Journal of Geophysical Research: Atmospheres, 120, 10,313–10,323, https://doi.org/10.1002/2015JD023438, 2015.

Ye, B., Del Genio, A. D., and Lo, K. K.-W.: CAPE Variations in the Current Climate and in a Climate Change, Journal of Climate, 11, 1997–2015, https://doi.org/10.1175/1520-0442-11.8.1997, 1998.

Zhang, C., Wang, Y., and Hamilton, K.: Improved Representation of Boundary Layer Clouds over the Southeast Pacific in ARW-WRF Using a Modified Tiedtke Cumulus Parameterization Scheme, Monthly Weather Review, 139, 3489–3513, https://doi.org/10.1175/MWR-D-10-05091.1, 2011.

Zheng, D., Van Der Velde, R., Su, Z., Wen, J., and Wang, X.: Assessment of Noah land surface model with various runoff parameterizations over a Tibetan river, Journal of Geophysical Research: Atmospheres, 122, 1488–1504, https://doi.org/10.1002/2016JD025572, 2017.