Uncertainty in gridded precipitation products: Influence of station density, interpolation method and grid resolution

Authors: Sixto Herrera[1], Sven Kotlarski[2], Pedro M.M. Soares[3], Rita M. Cardoso[3], Adam Jaczewski[4], José M. Gutiérrez[5], Douglas Maraun[6]

[1]Meteorology Group. Dpto. de Matemática Aplicada y Ciencias de la Computación. Univ. of Cantabria, Spain
[2]Federal Office of Meteorology and Climatology MeteoSwiss, Switzerland
[3]Instituto Dom Luiz, Universidade de Lisboa (IDL), Portugal
[4]Institute of Meteorology and Water Management, Poland
[5]Meteorology Group. Instituto de Física de Cantabria, CSIC-Univ. of Cantabria, Spain
[6]Wegener Center for Climate and Global Change, University of Graz, Austria.

Abstract

This work analyses three uncertainty sources affecting the observation-based gridded datasets: station density, interpolation methodology and spatial resolution. For this purpose, we consider precipitation in two countries, Poland and Spain, three resolutions (0.11, 0.22 and 0.44 degrees), three interpolation methods, both areal- and point-representative implementations, and three different densities of the underlying station network (high/medium/low density). As a result, for each resolution and interpolation approach, nine different grids have been obtained for each country and inter-compared using a variance decomposition methodology.

Results indicate larger differences among the datasets for Spain than for Poland, mainly due to the larger spatial variability and complex orography of the former region. The variance decomposition points out to station density as the most influential factor, independent of the season, the areal- or point-representative implementation and the country considered, and slightly increasing with the spatial resolution. In contrast, the decomposition is stable when extreme precipitation indices are considered, in particular for the 50-year return value.

Finally, the uncertainty due to station sub-sampling inside a particular grid box decreases with the number of stations used in the averaging/interpolation. In the case of spatially homogeneous grid boxes, the interpolation approach obtains similar results for all the parameters, excepting the wet day frequency, independently of the number of stations. When there is a more significant internal variability in the grid box, the interpolation is more sensitive to the number of stations, pointing out to a minimum stations' density for the target resolution (6-7 stations).

1. Introduction

There is an increasing need for global and regional climate model data for present and future climates. Global climate models and downscaling methodologies, which include regional climate models and statistical downscaling methods, are the only tools that enable the production of future climate projections and thus, quality assessment of climate models in present climate is a crucial step for augmenting the confidence of the projections and to characterize the inherent uncertainties. Presently, climate models are being run at increasing resolutions, and statistical downscaling methods mostly aim at describing climate properties at local scales. Global climate models are approaching resolutions of 50km in the atmosphere (e.g. Haarsma et al. 2016) and regional climate models are reaching the convective permitting scales of ~2km resolution, where deep convection is expected to be explicitly or at least partially resolved (Prein et al. 2015; Ban et al. 2014).

In evaluation exercises, potential scale mismatches between climate models and observation-based estimates have to be carefully considered. Climate models provide area-averaged quantities of the different surface meteorological variables whilst ground based observations are collected on a local scale. To ensure consistency, grid box model results should be compared to area-averaged observational estimates (Osborn and Hulme 1997) and, as a result, the need to build observational gridded datasets emerged, representing areal-average quantities and allowing fair comparisons and evaluations of model results (Chen and Knutson 2008). In addition, observation-based gridded datasets have been used for the development of bias-corrected climate change scenarios and many other climate impact studies (IPCC-BG3b 2015).

The first global regular gridded datasets proposed were the monthly datasets, at 0.5° × 0.5° latitude/longitude resolution, originally developed by New et al. (1999, 2000), later updated by Mitchell and Jones (2005) and by Harris et al. (2014) – the CRU TS3.10 Dataset. At a European level the most widely used dataset, E-OBS, was developed in the framework of the ENSEMBLES project and includes daily observations for temperature, precipitation (Haylock et al. 2008; Klein Tank et al. 2002; Klok and Klein Tank 2009) and sea level pressure (van der Besselaar et al. 2011). Other continental scale examples include the APHRODITE precipitation dataset for Asia (Yatagai et al. 2012), the AWAP climate datasets for Australia (Jones et al. 2009), the North America regular gridded dataset (Chen et al. 2008 and Daly et al. 2008) and the CLARIS dataset for South America (Menéndez et al. 2010). Recently, some regional and/or national regular gridded datasets were built, often at higher spatial and temporal resolution, incorporating a higher number of local station observations, spanning different periods and covering further variables (e.g. air temperatures, precipitation, cloud cover, relative humidity, etc.). Examples exist for Spain (Herrera et al. 2012, 2016), Portugal (Belo-Pereira et al. 2011), Germany (Rauthe et al. 2013; Frick et al. 2014), Switzerland (Frei et al. 2014, Isotta et al., 2014) or Romania (ROCADA, 2014 and 2015).

Typically, the aim of a gridded dataset is to represent the areal average in a grid box, which requires a sufficiently large number of stations within the grid box to account for the subgrid variability and allow the computation of an accurate mean value. Unfortunately, in many regions the observational network is sparse which poses challenges and difficulties for the building of high quality regular gridded datasets affecting the spatiotemporal structure of not only the mean values, but also the variance and extremes. Moreover, as the target resolution of the gridded dataset increases the effects of the mentioned issues on the quality of the dataset also increase. To the best of our knowledge there are only few studies focusing on these issues and due to their importance more investigation on this topic is needed.

Some previous studies have analyzed the impact of the station density and interpolation method on the quality of the gridded datasets produced regarding mean and extreme values (e.g. Gervais et al. 2014, Avila et al. 2015). In fact, the recognizable good quality of gridded datasets for mean values is not extendable for the variance as shown by Beguería et al. (2016). This study showed that modifications in the sample size result in changes in the variance of the gridded data. Usually gridded products underestimate the variance, even of the area average, leading to flawed inferences about changes in climate variability and extremes. Hofstra et al. (2010) investigated the effect of station network density on distributions and trends in indices of area-averaged daily precipitation and temperature in the E-OBS gridded dataset. By randomly decreasing the number of stations included in some of the grid boxes, the authors found that the fewer stations are used for the interpolation the larger the over-smoothing, for both precipitation and temperature. This smoothing is larger for higher percentiles and thus for extremes and the related extreme indices. In the context of a new method for spatial interpolation of daily surface air temperature from local stations in complex orographic regions, Frei et al. (2014) showed that, in the covered period, as the network became denser the interpolation accuracy improves. Nevertheless with the drawback regarding the long-term homogeneity of the resulting grid dataset, Herrera et al. (2016) used two interpolation methods to produce three regular gridded temperature (maximum and minimum) and precipitation datasets for Spain and tested the sensitivity of the interpolated fields to the use of orography as a covariate in the interpolation. As in Hofstra et al. (2010), the inclusion of orography in the temperature interpolation method is necessary to produce consistent results, while it introduces high precipitation biases with increasing elevation. The high station density also allowed the authors to infer that the precipitation underestimation (mean and extremes) encountered in E-OBS is associated to the density of the underlying station network.

For Australia, Contractor et al. (2015) assessed the spatiotemporal variability of precipitation comparing different gridded datasets, namely AWAP, TRMM, GPCP, in addition to 6 datasets built using diverse interpolation methods (cubic spline, triangulation with linear interpolation, ordinary kriging, natural neighbor interpolation, Barnes objective analysis) and grid station average. Regarding the temporal variability, grids interpolated by ordinary kriging and cubic spline interpolation show regionally larger differences (lower correlations). Additionally, the larger differences are associated with rainfall extremes, which in some locations have differences up to a factor of five. From the gridding methods no one consistently performs better when compared with the station observations, but in particular cubic spline interpolation shows a tendency to “overshoot” in comparison to station data and the other gridded products, particularly in regions with strong spatial gradients. Similar shortcomings appear to occur in particular in regions where the sparseness of the network is larger and important climate variability exists, like e.g. in West Africa (Wagner et al. 2007).

An increasing number of studies assess the observational uncertainty by comparing different observation-based reference datasets for specific variables (e.g., Herold et al., 2016, Berg et al., 2015, Dunn et al., 2014, Hofstra et al. 2009, Isotta et al., 2015, Kyselý and Plavcová, 2010, Palazzi et al., 2013, Schneider et al., 2014) and investigate the influence of uncertainties in gridded observations on regional climate model evaluation (e.g. Prein and Gobiet 2017; Kotlarski et al. 2017).

Against this background, a first objective of this study is to assess the sensitivity of gridded products to resolution, observation density and interpolation method, and assess their contribution to the total variance of the resulting gridded datasets. In particular we consider three resolutions (0.11º, 0.22º and 0.44º), three station densities (high, medium and low, the latter equivalent to the density of the E-OBS dataset), and three interpolation methods based on ordinary kriging (alone, or applied to the residuals after considering 2D or 3D splines for the monthly mean values). Moreover, we considered both point- and areal-representative implementations for the interpolation methods. The analysis is carried out for two distinct regions, Poland and Spain, which represent a wide range of European climatic variability. A second objective of this work was to provide suitable information to undertake an analysis of the uncertainty of gridded products for the evaluation of Regional Climate Models (see Kotlarski, 2017, this issue).

This paper is organized as follows. Section 2 describes the observational networks considered to build the gridded datasets for both countries (Poland and Spain), the interpolation methods used in this work and the variance decomposition analysis considered to separate the contribution of each factor to the total variability. Section 3 describes the main results obtained and Section 4 summarizes the results and concludes the paper.

2. Data and Methods

2.1 Observation data for Spain

The Iberian peninsula is located in southwestern Europe (Figure 1a), in the transition zone between extratropical and subtropical influences, spanning a region with complex orography influenced by both Atlantic and Mediterranean climates. The resulting local climate variability (Figure 2c-d) ranges from temperate climates with regular precipitation spread over the whole year in the north; to dry (semiarid) climates with areas with less than 100 mm/year in the southeast; to the Mediterranean coast and part of the Ebro basin where frequent drought periods alternate with heavy rainfall events (see, e.g. Llasat, 2009).

The observational network used in this work is based on the observational datasets described in Herrera et al. (2011). Only stations with at least 40 years in the period 1950-2003 and with at least 90% data availability within each year have been considered, being representative of the climatology observed in the whole period and also for each particular year. None automated missing daily data completion algorithm has been included in the process. Moreover, only the homogeneous stations (at a 95% confidence level) according to the standard normal homogeneity test (SNHT) and Alexandersson test (Alexandersson, 1986) have been considered. The final observational network obtained contains 822 precipitation stations (Figure 1b) in the period 1951-2010 which have been used in the interpolation procedure.

[INSERT FIGURES 1 and 2 AROUND THIS POINT]

2.2 Observation data for Poland

Poland is, for the most of its territory, topographically relatively flat and reaches from the Baltic Sea in the north, to the Carpathian Mountains in the south (Figure 1a). The country is located in a transition zone in the temperate climatic zone (Szwed, 2010). Poland’s climate is influenced by oceanic (Atlantic) and continental air masses approaching from the western and eastern direction, respectively, as well as by polar and tropical air. These factors determine precipitation amounts which, in the annual sum, are largest (greater than 600 mm) in the northwest of the country and in the southern upland and mountainous part (Tatra mountains) where they partly exceed 1000 mm per year (Figure 2a-b). The maximum monthly precipitation sum is observed in July and the minimum in February. There is a tendency for extreme daily totals to increase towards the east (Marosz et al., 2013). Drought is more frequent in the northern and western part of Poland (Kalbarczyk, 2010).

The number of publicly available precipitation station data for Poland is very limited. Even the station network used for public gridded datasets is sparse, resulting in low quality datasets, especially for the extremes (Wibig et al., 2014; Hofstra et al., 2009). The rain gauge network operated by the Institute of

Meteorology and Water Management – National Research Institute in Poland is not very dense in comparison to other European countries of similar scale and comprises about 500 stations. To avoid dealing with missing data and to focus on the sensitivity analysis we have decided to take only stations with more than 98% data available into account, without including automated missing daily data completion in the process. This finally leads to 197 Polish stations in the period 1978-2012 considered in our analysis (Figure 1c). The data has been homogenized and quality controlled by the MASHv3.03 (Szentimrey, 2011) procedure.

Note that the quality control and homogeneity analysis considered for each country are the standard procedures as employed (a) by the Institute of Meteorology and Water Management to be applied to its observational datasets in Poland and (b) in the development of the ensemble of gridded datasets Spain02 (Herrera et al. 2012; 2015). In order to maintain the coherence with the datasets of reference used in both countries we decided to keep both procedures for the present analysis in spite of their differences.

For the sake of the coherence and the comparability of the analysis proposed for both countries, only the common period, 1981-2010, of both observational dataset has been considered in this study. For this period, the 50-year return value for each grid-box was obtained by adjusting a Generalized Extreme Value (GEV) distribution to the annual maximum of daily precipitation (see Herrera et al. 2015, Section 2.3, for a detailed description).

2.3 Target Resolutions and Grids

The objective of a related study was to provide suitable information to undertake an analysis of the uncertainty of gridded products for the evaluation of Regional Climate Models (see Kotlarski, 2017, this issue). Therefore, the grids considered in this study match the rotated grids considered in the EURO-CORDEX initiative at a horizontal resolution of 0.11º, 0.22º and 0.44º (see Figure 3). Note that the latter two grids are degraded versions of the 0.11º one. Note also that the E-OBS grid (v11 was used in this work) matches the 0.22º resolution grid used in this work.

In order to evaluate the uncertainty associated with the density of the station network we defined three different densities for each domain (high, medium and low), considering the full set, half and quarter of all available stations, respectively. The medium and low density network were chosen from the total network with a stratified random sampling based on the location (longitude and latitude) and the mean precipitation. To obtain each sample, a k-means clustering algorithm was applied to these variables (after standardization), considering the same number of clusters as stations used for each target density. Then, one station was randomly chosen within each of the clusters. Note that the lowest density considered (~200/50 stations for Spain/Poland) approximately corresponds to the number of stations used by E-OBS to define the 0.22º rotated grid (~190/30 stations for Spain/Poland). Therefore, the sensitivity analysis undertaken in this study sheds some light on the representativeness of the E-OBS datasets over the two regions under study.

We want to remark that the sensitivity analysis provided in this work focuses on the impact of a relative change in the actual density of stations available (note that the actual density is higher in Spain than in Poland), not at the impact relative to some hypothetically perfectly representative density. This latter aspect is illustratively analyzed with a case study in Sec. 3.3, considering the gridbox with highest density.

[INSERT FIGURE 3 AROUND THIS POINT]

2.4 Interpolation Methods

The present work builds on the methodologies described in a previous study of Herrera et al. (2016). In that study four interpolation methods were applied to obtain a set of daily gridded datasets for precipitation and temperature (mean, minimum and maximum), targeting the three different resolutions presented in Sec. 2.3. In order to obtain gridded products comparable with the RCM direct output we consider the three methods representing the area average (AA) included in the work of Herrera et al. 2016: Ordinary Kriging (AA-OK), and 2- and 3-dimensional thin plate splines (AA-2D and AA-3D). These three methods are based on ordinary kriging. In the first case, the method is directly applied to the observed daily precipitation values, while the other two methods follow a two-step approach: first the 2- or 3-dimensional thin plate splines (considering or not the orography as covariable) is applied to the observed monthly values, and secondly the daily anomalies are interpolated using the ordinary kriging. Finally, both monthly values and daily anomalies are combined to obtain the interpolated daily values. Area-averaged representativeness is achieved by performing the interpolation using an auxiliary 0.01º grid, and averaging the resulting interpolated results in both regions for the three target resolutions (0.11º, 0.22º and 0.44º). In addition, the above interpolation methods were applied using a point-representative implementation (OK, 2D and 3D, according to the previously defined notation) by directly estimating the values for the final grid, not including the interpolation to the auxiliary 0.01º grid. In this case, the resulting grids are not averaged versions of the higher-resolution one.

Note that both point- and area-representative interpolation have been applied in previous studies (Herrera et al. 2012 and 2016) and the optimal approach depends on the application. As an example, for climate models evaluation area-represantive methods should be considered in order to maintain the coherence between both datasets, while for observation data completion a point-representative can be more appropriated.

2.5 Variance Decomposition Analysis

The described experiments lead to a 3x3 (interpolation method and stations density) matrix of gridded datasets for each interpolation approach (point- and areal-representative), resolution (0.11º, 0.22º and 0.44º) and region (Poland and Spain) which allows to assess the contribution of each dimension to the total variance by means of a statistical analysis of variance (von Storch and Zwiers, 1999; Dequé et al 2007, 2011):

$$V = I + D + ID$$

$$I = \frac{1}{3}\sum_{i=1:3}\left(X_{i.} - X_{..}\right)^2; D = \frac{1}{3}\sum_{j=1:3}\left(X_{.j} - X_{..}\right)^2;$$

$$ID = \frac{1}{3}\sum_{i=1:3}\frac{1}{3}\sum_{j=1:3}\left(X_{ij} - X_{i.} - X_{.j} + X_{..}\right)^2,$$

where $\boldsymbol{V}$ is the total variance, $\boldsymbol{I}$ and $\boldsymbol{D}$ are the variance due to the interpolation method and the stations density alone respectively, $\boldsymbol{ID}$ is the variance due to the interaction between both dimensions, $\boldsymbol{X_{ij}}$ is the gridded dataset corresponding to the $\boldsymbol{i^{th}}$ interpolation method and the $\boldsymbol{j^{th}}$ stations density, and the dots (***.***) in the subindex represent the mean in the corresponding dimension. The variance decomposition is carried out for each grid cell independently, and variance contributions are then spatially averaged.

3. Results

3.1 Overall interpolation results

In order to analyze the effect of the station density and the interpolation methods on the resulting gridded datasets, we focus both on mean and extreme regimes and consider mean precipitation and 50-year return value of the daily all-year time series for each grid box. Figures 4 and 5 show the results for the nine datasets developed for the high density case considering the three resolutions (columns) and the three interpolation methods (in rows), for Spain (left) and Poland (right). The benchmarking results for the stations and the gridded (0.22º) dataset E-OBS are shown in the upper two panels of each figure. Note that, for the sake of simplicity, the individual map results have been shown only for the areal-representative methods, whereas both approaches will be considered in the analysis of variance in the following section.

Figure 4 shows the results for mean daily precipitation. On the one hand, this figure shows a general underestimation by E-OBS, mainly in regions with complex orography (e.g. Spain or the south of Poland). On the other hand, comparing the interpolation methods, there are in general smaller differences between them for each resolution than with respect to E-OBS, which is related to the stations network considered in both cases and points out to the importance of the stations density in the interpolation process. The main difference is found between the AA-OK and AA-2D metods in Southern Poland.

Regarding the 50-year return value, Figure 5 shows similar results than those described for the mean daily precipitation with a significant underestimation by E-OBS in both countries for all the interpolation methods. In this case, the differences between methods are more significant than for the mean for the three resolutions. In particular, for both countries the main differences appear when the monthly splines-based interpolation is included in the methodology in isolated points/regions (e.g. Northwestern Spain or Center Poland). Note that the spatial pattern of the extreme precipitation is less dependent on the orography which has been partially explained by the relation between the occurrence of extreme events with different circulation patterns (Ramis et al, 2013; García-Ortega, 2007; AEMET, 2011; Herrera, 2012) like the North Atlantic Oscillation (NAO) or the western Mediterranean Oscillation (WeMOi). In particular, despite orography, in the Mediterranean coast the humid air from the Mediterranean Sea favors the development of mesoscale convective systems and convective clouds resulting in hail or heavy precipitations in the eastern coast of the Iberian Peninsula.

[INSERT FIGURES 4 and 5 AROUND THIS POINT]

3.2 Variance decomposition

Figure 6 shows the spatially averaged variance decomposition (considering only those grid boxes with at least one station) for the mean (first three column blocks) and extreme (last block) values, considering both countries and both area- and point-representative implementations of the interpolation methods. For each block, the stacked bar plots show the contribution to the total variance of only station density (blue), only the interpolation methodology (red) and the interaction of both (green), separately for each resolution. Station density appears to be the main variability factor, explaining over 60% of the variance in Spain and over 50% in Poland in all cases. This contribution is particularly relevant for summer, with differences in the range of 5-10% with respect to winter and annual contributions. Moreover, the interaction term is roughly insensitive to the resolution in all cases, so the increment of the relative importance of station density (~10% for all seasons) with resolution leads to a similar decrement of the relative importance of the interpolation method. For the point-representative interpolation methods the results are very similar in Spain, but in Poland the station density component is more important, particularly due to a reduction of the interaction term.

Similar results have been found for the 50-year return value, with the exception of the sensitivity to resolution: in this case the results are roughly insensitive to this factor. This could be partly explained by the fact that extreme precipitation is less dependent on the orography than the mean precipitation (see Figures 4 and 5). Since the main effect of increasing the resolution is to improve the representation of the orography (and the orographyc effects on precipitation), the dependence on the resolution has less impact in the 50-year return value than for the mean precipitation.

Figure 7 shows the spatial variability of the variance decomposition obtained for the area-representative implementation and the highest-resolution grid. In agreement with the obtained results, areas with a large dependence on the interpolation method broadly correspond to station-sparse regions (not included in the bar plot). Note that regions without stations depend on the surrounding network and on the nature of the interpolation method. Subsequently, once the station's network is defined, all the variability comes from the interpolation method as is shown in Figure 7. This information can support the identification of target regions where a clear need for increasing the local representativeness with new station data.

These results are also in agreement with those obtained for the local analysis (see Sec. 3.3 and Figure 9) which shows that, for a particular grid box, the interpolated value strongly depends on the stations surrounding the grid box. Hence, changes in the density modify the nearest stations affecting the grid box and lead to a high variability in the resulting interpolated values.

[INSERT FIGURES 6 and 7 AROUND THIS POINT]

3.3 Analysis of effective resolution

In the previous section we have seen that the station density largely influences the variability of the interpolated grid box values. However, this analysis does not provide information on the effective resolution for a given network of stations or, in other words, the number of stations needed for convergence of interpolated values for a given resolution. In this section we shed some light on this problem by considering two grid boxes from the 0.44º resolution grid in Northern Spain with the largest number of available stations (see the rectangle in Northern Spain in Figure 1a). Figure 8 shows the local grid boxes of the 0.44º resolution grid in this area and the location of the stations within these grid boxes. In this case, for each number of stations and grid box a bootstrapping approach was considered to obtain 100 randomly chosen samples from the 10 stations available within each grid box. For each sample, the point-representative 3D interpolation method was adjusted, considering the stations chosen for both grid boxes, and applied to obtain the interpolated values of both grid boxes for the full period 1981-2010.

Moreover, for comparison we have considered an alternative grid box value estimation method based on simply averaging the station values.

On the one hand, the averaging approach reduces the uncertainty w.r.t. the interpolation due to the station combinations considered for all the parameters and grid boxes considered. Moreover, with the exception of the wet day frequency, in the case of the averaging approach the median value for the realizations is almost independent of the number of stations considered. On the other hand, a greater variability and more dependence on the number of stations have been found for the interpolation. In this sense, the results obtained for the interpolation point out to a minimum stations' density of 6-7 stations per grid box to reach the target resolution (0.44º). Although this result depends on the internal variability of each grid box, as is reflected in figure 9, for both approaches and grid boxes considered all the parameters stabilize around these values.

[INSERT FIGURES 8 and 9]

4. Summary and Conclusions

In this work the sensitivity to different uncertainty sources of a precipitation interpolation methodology has been analyzed considering three interpolation methods (OK, 2D and 3D), three stations density (low, medium and high), two interpolation approaches (areal- and point-representative), three resolutions (0.11º, 0.22º and 0.44º) and two different geographical domains (Spain and Poland).

The main conclusion obtained is the relevance of the stations density, explaining more than 60% of the variance independently of the areal- or point-representativity of the interpolation method, the resolution, season and the country considered. Regarding the spatial distribution of the explained variance, regions with largest percentage due to the interpolation method are located in regions with low stations density.

For the sake of comparison, E-OBS has been considered as well and showed large differences w.r.t. the gridded products built, independently of the method considered, including the AA-3D which corresponds to the method used to develop E-OBS. These differences are more significant for Spain than Poland due to the large climatic variability and the complex orography of the former.

To analyze the local effect of the stations network considered in the interpolation, two Spanish grid boxes containing the largest number of stations (10) for the 0.44 resolution, and the time series corresponding to the interpolation and average of 100 randomly chosen sub-samples have been considered. Although it depends on the internal variability inside the grid box, the uncertainty due to the combinations and number of stations considered is reduced when the spatial average is used instead an interpolation method. Based on the results obtained for both grid boxes and interpolation/averaging approaches, at least 6-7 stations per grid box should be considered to reach the target resolution.

Acknowledgments

VALUE has been funded as EU COST Action ES1102. Participation of S. Herrera and J.M. Gutiérrez was partially supported by the project MULTI-SDM (CGL2015-66583-R, MINECO/FEDER). P.M.M. Soares and R.M. Cardoso wish to acknowledge the projects SOLAR (PTDC/GEOMET/7078/2014) and FCT UID/GEO/50019/ 2013 - Instituto Dom Luiz, both financed by the Fundação para a Ciência e Tecnologia. We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu). We also thank two anonymous reviewers for their useful comments that helped to improve the original manuscript.

Conflicts of Interest

There are no known conflicts of interest associated with this publication, it has not been submitted (or published previously) to other journal and there has been no significant financial support for this work that could have influenced its outcome. All the authors included have contributed to this work, both in the development and in the interpretation of the scientific results. All the authors have consented the submission of this work and are prepared to collect documentation of compliance with ethical standards and send it if it is requested during peer review or after publication.

References

Alexandersson H 1986. A homogeneity test applied to precipitation data. Journal of Climatology 6: 661–675.

AEMET, 2011 Atlas climático ibérico/Iberian climate atlas, Agencia Estatal de Meteorología (AEMET), Madrid, Spain, 2011.

Avila F.B., Dong S., Menang K.P., Rajczak J., Renom M., Donat M.G. and Alexander L.V., 2015, Systematic investigation of gridding-related scaling effects on annual statistics of daily temperature and precipitation maxima: A case study for south-east Australia, Weather and Climate Extremes, 9 6-16, DOI: https://doi.org/10.1016/j.wace.2015.06.003

Ban N., Schmidli J., and Schär C. 2014. Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations, J. Geophys. Res. Atmos., 119, 7889–7907, doi:10.1002/2014JD021478.

Beguería S. Vicente-Serrano SM. Tomás-Burguera M. Maneta M. 2016. Bias in the variance of gridded data sets leads to misleading conclusions about changes in climate variability. International Journal of Climatology 36:3413-3422. doi:10.1002/joc.4561

Belo-Pereira M., Dutra E., Viterbo P. 2011. Evaluation of global precipitation data sets over the Iberian Peninsula. Journal of Geophysical Research 116:D20101. doi:10.1029/2010JD015481.

Berg N., Hall A., Sun F., Capps S., Walton D., Langenbrunner B., Neelin D. 2015. Twenty-First-Century Precipitation Changes over the Los Angeles Region. Journal of Climate 28(2):401-421. doi:10.1175/JCLI-D-14-00316.1.

Birsan MV. and Dumitrescu A. 2014. ROCADA: Romanian daily gridded climatic dataset (1961-2013) V1.0. doi:10.1594/PANGAEA.833627, https://doi.pangaea.de/10.1594/PANGAEA.833627, PANGAEA, Administratia Nationala de Meteorologie, Bucuresti, Romania

Chen CT., Knutson T. 2008. On the verification and comparison of extreme rainfall indices from climate models. Journal of Climate 21:1605-1621. doi:10.1175/2007JCLI1494.1

Chen M., Xie P., and Co-authors 2008. CPC Unified Gauge-based Analysis of Global Daily Precipiation. Western Pacific Geophysics Meeting, Cairns, Australia, 29 July - 1 August, 2008.

Contractor S., Alexander LV., Donat MG., Herold N. 2015. How well do gridded datasets of observed daily precipitation compare over Australia? Advances in Meteorology 325718, 15, http://dx.doi.org/10.1155/2015/325718

Dahlgren P., Gustafsson N. 2012. Assimilating Host Model Information into a Limited Area Model. Tellus A 64: 15836, doi:10.3402/tellusa.v64i0.15836

Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., and Pasteris, P.A. 2008. Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. International Journal of Climatology, 28:2031-2064. doi: 10.1002/joc.1688

Déqué M., Rowell DP., Lüthi D., Giorgi F., Christensen JH., Rockel B., Jacob D., Kjellström E., de Castro M., van den Hurk B. 2007. An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. Climatic Change 81(1):53-70. doi:10.1007/s10584-006-9228-x

Déqué M., Somot S., Sanchez-Gomez E., Goodess CM., Jacob D., Lenderink G., Christensen OB. 2012. The spread amongst ENSEMBLES regional scenarios: regional climate models, driving general circulation models and interannual variability. Climate Dynamics 38(5):951-964. doi:10.1007/s00382-011-1053-x

Dumitrescu A. and Birsan MV. 2015. Supplement to ROCADA: a gridded daily climatic dataset over Romania (1961-2013) for nine meteorological variables. Natural Hazards 78(2):1045-1063. doi:10.1007/s11069-015-1757-z

Dunn RJH., Donat MG., Alexander LV. 2014. Investigating uncertainties in global gridded datasets of climate extremes. Climate of the Past 10(6):2171-2199. doi:10.5194/cp-10-2171-2014.

Durand Y., Brun E., Merindol L., Guyomarc'h G., Lesaffre B., Martin E., 1993, A meteorological estimation of relevant parameters for snow models, Ann. Glaciol. 18:65–71.

Frei C. 2014. Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances. International Journal of Climatology 34:1585-1605. doi:10.1002/joc.3786.

Frick C., Steiner H., Mazurkiewicz A., Riediger U., Rauthe M., Reich T., Gratzki A. 2014. Central European high-resolution gridded daily data sets (HYRAS): Mean temperature and relative humidity. Meteorologische Zeitschrift 23(1): 15-32. doi:10.1127/0941-2948/2014/0560

García-Ortega E, Fita L, Romero R, López L, Ramis C, Sánchez J., 2007, Numerical simulation and sensitivity study of a severe hail-storm in northeast Spain. Atmospheric Research 83: 225–241, DOI: 10.1016/j.atmosres.2005.08.004.

Gervais M., Tremblay LB., Gyakum JR., Atallah E. 2014. Representing Extremes in a Daily Gridded Precipitation Analysis over the United States: Impacts of Station Density, Resolution, and Gridding Methods. Journal of Climate 27(14):5201-5218. doi:10.1175/JCLI-D-13-00319.1

Haarsma RJ., Roberts MJ., Vidale PL., Senior CA., Bellucci A., Bao Q., Chang P., Corti S., Fučkar NS., Guemas V., von Hardenberg J., Hazeleger W., Kodama C., Koenigk T., Leung LR., Lu J., Luo JJ., Mao J., Mizielinski MS., Mizuta R., Nobre P., Satoh M., Scoccimarro E., Semmler T., Small J., and von Storch JS. 2016. High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. Geosci. Model Dev. 9:4185-4208. doi:10.5194/gmd-9-4185-2016.

Harris I., Jones PD., Osborn TJ., Lister DH. 2014. Updated high-resolution grids of monthly climatic observations-the CRU TS3.10 dataset. Int. J. Climatol. 34:623-642. doi:10.1002/joc.3711

Haylock MR., Hofstra N., Klein Tank AMG., Klok EJ., Jones PD., New M. 2008. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. Journal of Geophysical Research 113:D20119. doi:10.1029/2008JD010201.

Herold N., Behrangi A., Alexander LV. 2017. Large uncertainties in observed daily precipitation extremes over land. J. Geophys. Res. Atmos. 122:668-681. doi:10.1002/2016JD025842.

Herrera S. 2011. Desarrollo, Validación y Aplicaciones de Spain02: Una Rejilla de Alta Resolución de Observaciones Interpoladas Para Precipitación Y Temperatura en España. PhD thesis (in Spanish). Universidad de Cantabria. Cantabria, Spain. Available online in: http://www.meteo.unican.es/tesis/herrera

Herrera S., Gutiérrez JM., Ancell R., Pons MR., Frías MD., Fernández J. 2012. Development and analysis of a 50-year high-resolution daily gridded precipitation dataset over Spain (Spain02). Int. J. Climatol. 32(1):74-85, doi:10.1002/joc.2256.

Herrera S., Fernández J., Gutiérrez JM. 2016. Update of the Spain02 gridded observational dataset for EURO-CORDEX evaluation: assessing the effect of the interpolation methodology. Int. J. Climatol. 36:900-908. doi:10.1002/joc.4391

Hofstra N., Haylock M., New M., Jones PD., Frei C. 2008. Comparison of six methods for the interpolation of daily European climate data. Journal of Geophysical Research 113:D21110. doi:10.1029/2008JD010100.

Hofstra N., New M. 2009. Spatial variability in correlation decay distance and influence on angular-distance weighting interpolation of daily precipitation over Europe. Int. J. Climatol. 29:1872-1880. doi:10.1002/joc.1819

Hofstra N., Haylock M., New M., Jones PD. 2009. Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature. Journal of Geophysical Research 114:D21101. doi:10.1029/2009JD011799.

Hofstra N., New M., McSweeney C. 2010. The influence of interpolation and station network density on the distribution and extreme trends of climate variables in gridded data. Clim Dyn 35:841–858. doi:10.1007/s00382-009-0698-1

IPCC-BG3b, 2015: Workshop Report of the Intergovernmental Panel on Climate Change Workshop on Regional Climate Projections and their Use in Impacts and Risk Analysis Studies [Stocker TF., Qin D., Plattner GK., Tignor M. (eds.)]. IPCC Working Group I Technical Support Unit, University of Bern, Bern, Switzerland, (pp. 21-23).

Isotta FA., Vogel R., Frei C., 2014. Evaluation of European regional reanalyses and downscalings for precipitation in the Alpine region. Meteorologische Zeitschrift 24(1):15-37. doi:10.1127/metz/2014/0584.

Jones DA., Wang W., Fawcett R. 2009. High-quality spatial climate data-sets for Australia. Australian Meteorological and Oceanographic Journal 58 (4):233.248.

Kalbarczyk R. 2010. Temporal and spatial diversity of the occurrence of atmospheric drought in Poland (1966-2005) and its effect of yield of pickling cucumber (Cucumissativus L.). Spanish Journal of Agricultural Research 8:1147-1162. doi:10.5424/sjar/2010084-1405

Klein Tank AMG., Wijngaard JB., Können GP., Böhm R., Demarée G., Gocheva A., Mileta M., Pashiardis S., Hejkrlik L., Kern-Hansen C., Heino R., Bessemoulin P., Müller-Westermeier G., Tzanakou M., Szalai S., Pálsdóttir T., Fitzgerald D., Rubin S., Capaldo M., Maugeri M., Leitass A., Bukantis A., Aberfeld R., van Engelen AFV., Forland E., Mietus M., Coelho F., Mares C., Razuvaev V., Nieplova E., Cegnar T., Antonio López J., Dahlström B., Moberg A., Kirchhofer W., Ceylan A., Pachaliuk O., Alexander LV. and Petrovic P. 2002. Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. Int. J. Climatol. 22:1441–1453. doi:10.1002/joc.773

Klok EJ, Klein Tank AMG. 2009. Updated and extended European dataset of daily climate observations. Int J Climatol 29:1182–1191. doi:10.1002/joc.1779.

Kotlarski S., Szabó P., Herrera S., Räty O., Keuler K., Soares PM., Cardoso RM., Bosshard T., Pagé C., Boberg F., Gutiérrez JM., Isotta FA., Jaczewski A., Kreienkamp F., Liniger MA., Lussana C., Pianko-Kluczyńska K., Szepszo G. 2017. Observational uncertainty and regional climate model evaluation: A pan-European perspective. International Journal of Climatology: VALUE special issue.

Lakatos M., Szentimrey T., Bihari Z., Szalai S. 2013. Creation of a homogenized climate database for the Carpathian region by applying the MASH procedure and the preliminary analysis of the data. Idojaras 117:143–158

Marosz M., Wójcik R., Pilarski M., Miętus M. 2013. Extreme daily precipitation totals in Poland during summer: the role of regional atmospheric circulation. Clim Res 56:245-259. doi:https://doi.org/10.3354/cr01155

Menéndez CG., De Castro M., Sörensson A., Boulanger JP. 2010. CLARIS project: towards climate downscaling in South America. Meteorologische Zeitschrift 19 (4):357-362.

Mitchell TD., Jones PD. 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. International Journal of Climatology 25: 693–712. doi:10.1002/joc.1181

New M., Hulme M., Jones PD. 1999. Representing twentieth-century space-time climate variability. Part I: development of a 1961-90 mean monthly terrestrial climatology. Journal of Climate 12: 829–856. doi: 10.1175/1520-0442(1999)012<0829:RTCSTC>2.0.CO;2.

New M., Hulme M., Jones PD. 2000. Representing twentieth century space-time climate variability. II: development of 1901–1996 monthly grids of terrestrial surface climate. Journal of Climate 13: 2217–2238. doi: 10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2.

Nikulin G., Bosshard T., Yang W., Barring L., Wilcke, R., Vrac M., Vautard, R. and Noel T., Gutiérrez JM., Herrera S., Fernández J., Haugen JE., Benestad R., Landgren OA., Grillakis M., Ioannis T., Koutroulis A., Dosio A., Ferrone A., Switanek M. 2015. Bias Correction Intercomparison Project (BCIP): an introduction and the first results, EGU General Assembly Conference Abstracts 17:2250-2250. http://adsabs.harvard.edu/abs/2015EGUGA..17.2250N

Osborn TJ., Hulme M. 1997. Development of a relationship between station and grid-box rainday frequencies for climate model evaluation. J Clim 10:1885–1908. doi:10.1175/1520-442(1997)010<1885:DOARBS>2.0.CO;2

Palazzi E., von Hardenberg J., Provenzale A. 2013. Precipitation in the Hindu-Kush Karakoram Himalaya: Observations and future scenarios. J. Geophys. Res. Atmos. 118:85-100. doi:10.1029/2012JD018697.

Plavcová E., Kyselý J., 2010. Relationships between sudden weather changes in summer and mortality in the Czech Republic, 1986-2005. International Journal of Biometeorology 54:539-551. doi:10.1007/s00484-010-0303-7

Prein AF., Langhans W., Fosser G., Ferrone A., Ban N., Goergen K., Keller M., Tölle M., Gutjahr O., Feser F., Brisson E., Kollet S., Schmidli J., van Lipzig NPM., Leung R. 2015. A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. Reviews of Geophysics. 53:323–361. doi:10.1002/2014RG000475.

Prein AF. and Gobiet A. 2016. Impacts of uncertainties in European gridded precipitation observations on regional climate analysis. Int. J. Climatol. 37:305-327. doi:10.1002/joc.4706

Quintana-Seguí P., Turco M., Herrera S., Miguez-Macho G. 2017. Validation of a new SAFRAN-based gridded precipitation product for Spain and comparisons to Spain02 and ERA-Interim. Hydrology and Earth System Sciences 21:2187-2201. doi:10.5194/hess-21-2187-2017

Ramis C, Homar V, Amengual A, Romero R, Alonso S., 2013, Daily precipitation records over mainland Spain and the Balearic Islands, Nat. Hazards Earth Syst. Sci., 13, 2483-2491, https://doi.org/10.5194/nhess-13-2483-2013

Rauthe M., Steiner H., Riediger U., Mazurkiewicz A., Gratzki A. 2013. A central european precipitation climatology part i: Generation and validation of a high-resolution gridded daily data set (hyras). Meteorologische Zeitschrift 22(3):235–256, doi:10.1127/0941-2948/2013/0436.

Schneider U., Becker A., Finger P., Meyer-Christoffer A., Ziese M., Rudolf B. 2014. GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. Theor. Appl. Climatol. 115(1):15-40. doi:10.1007/s00704-013-0860-x.Szentimrey T. 2011. Manual of homogenization software MASHv3.03. Hungarian Meteorological Service.

Szwed M. 2010. Uncertainty of the climate model projections based on the Poland’s territory example. In: 6th Alexander von Humboldt International Conference on Climate Change, Natural Hazards, and Societies, held March 15-19 in Merida, Mexico. http://meetings. copernicus. org/avh6, id. AvH6-51.p 51

van den Besselaar EJM., Haylock MR., van der Schrier G., Klein Tank. AMG. 2011. A European Daily High-resolution Observational Gridded Data set of Sea Level Pressure. J. Geophys. Res. 116:D11110. doi:10.1029/2010JD015468

von Storch H and Zwiers FW. 1999. Statistical Analysis in Climate Research, Cambridge University Press, Cambridge, doi:10.1017/CBO9780511612336

Yatagai A., Kamiguchi K., Arakawa O., Hamada A., Yasutomi N., Kitoh A. 2012. APHRODITE constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. Bulletin of the American Meteorological Society 93 (9) pp. 1401–1415.

Wagner S, Kunstmann H, Bardossy A. 2007. Uncertainties in water balance estimations due to scarce meteorological information: a case study for the White Volta catchment in West Africa. Proceedings of Symposium HS2004 at IUGG2007, Perugia: Quantification and Reduction of Predictive Uncertainty for Sustainable Water Resources Management. IAHS Publ. 313.

Wibig J., Jaczewski A., Brzóska B., Konca-Kędzierska K., Pianko-Kluczyńska K. 2014. How does the areal averaging influence the extremes? The context of gridded observation data sets. MeteorologischeZeitschrift 23:181–187.

FIGURES

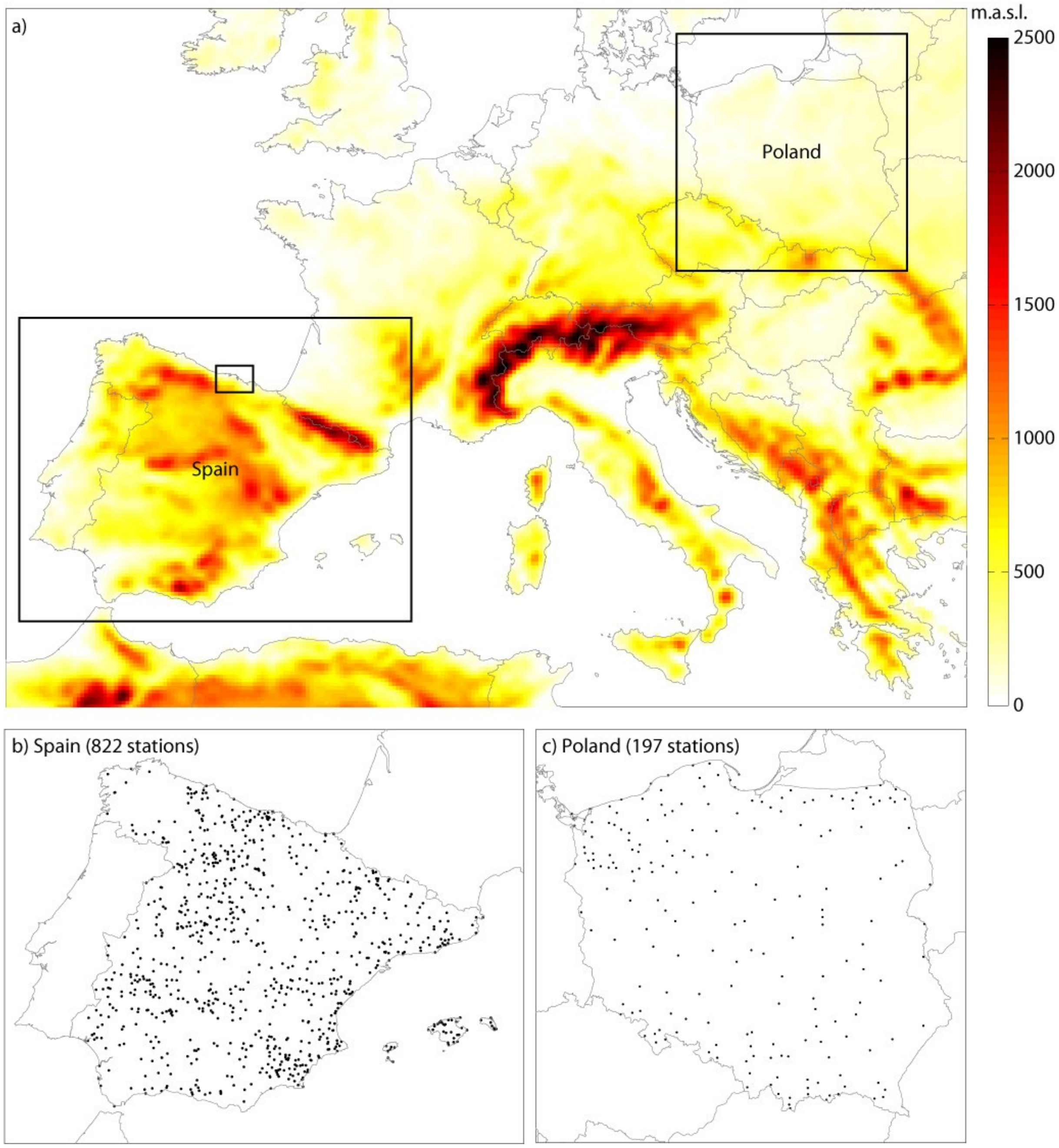


Figure 1: (a) Orography and location of the European regions analysed in this work (Spain and Poland), together with the considered station networks for (b) Spain and (c) Poland. The rectangle in northern Spain shows the region containing the two grid-boxes with the largest station density, which were considered for the local analysis in Sec. 3.3.

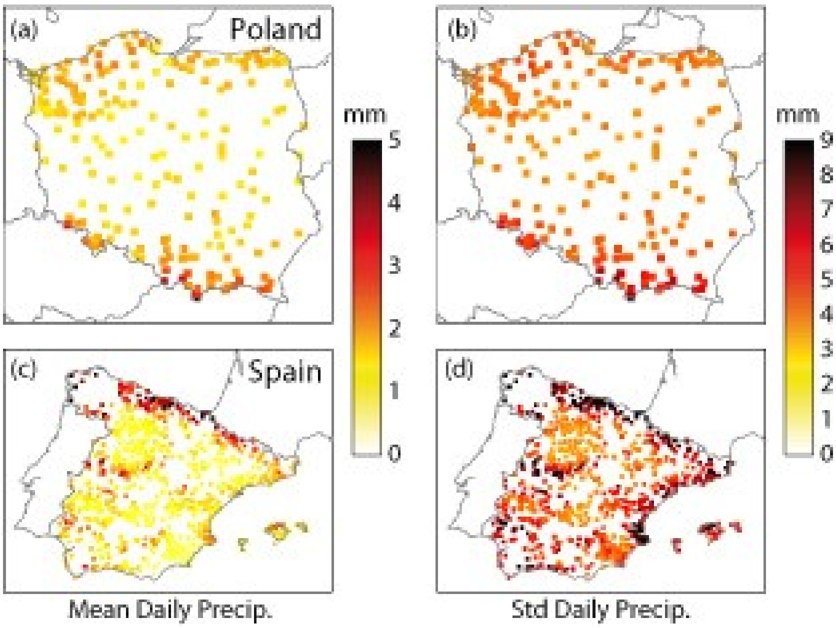


Figure 2: Mean (a-c) and standard deviation (b-d) of daily precipitation of the network considered for Poland (a-b) and Spain (c-d) for the period 1981-2010.

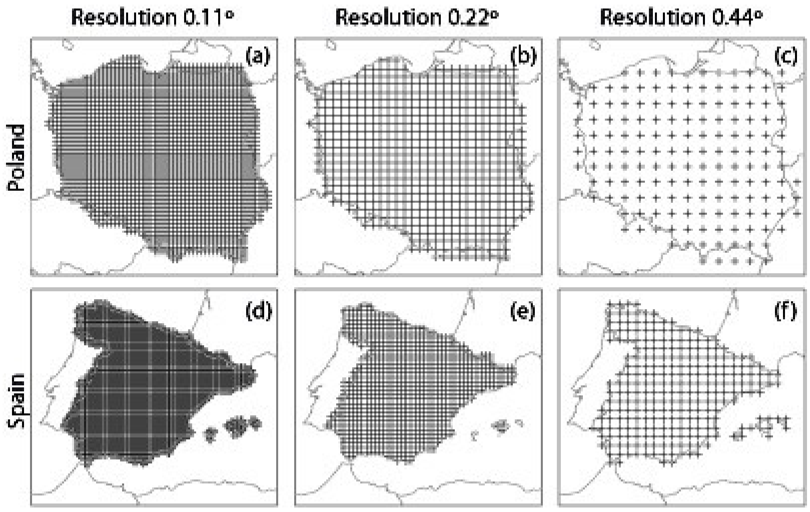


Figure 3: Grids defined for Poland (a-c) and Spain(d-f) for the three resolutions. The two extreme resolutions (0.11 and 0.44) were used in EURO-CORDEX, whereas the intermediate (0.22) was used in E-OBS.

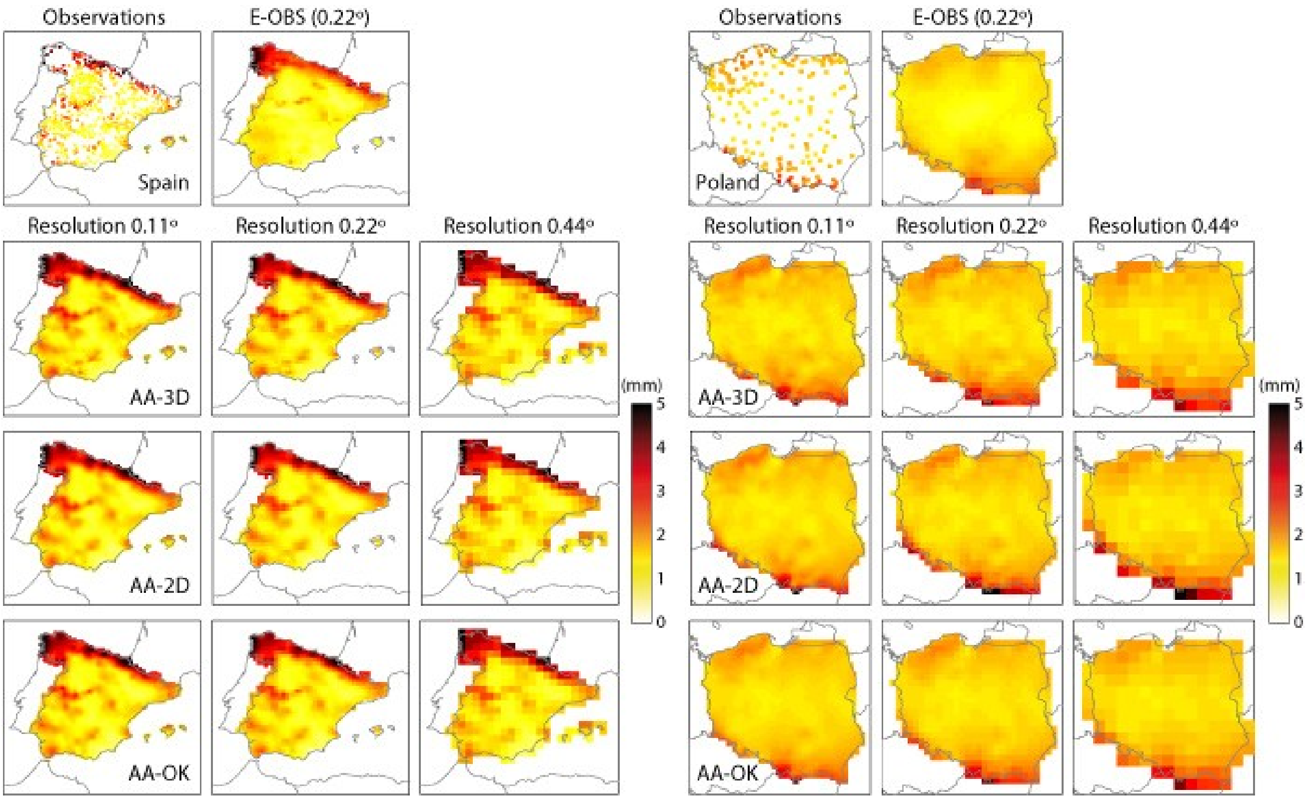


Figure 4: Annual mean daily precipitation for the period 1981-2010 for the observations and the nine (3 interpolation methods x 3 resolutions) area-representative grids built. The 0.22 resolution of E-OBS has been included as reference.

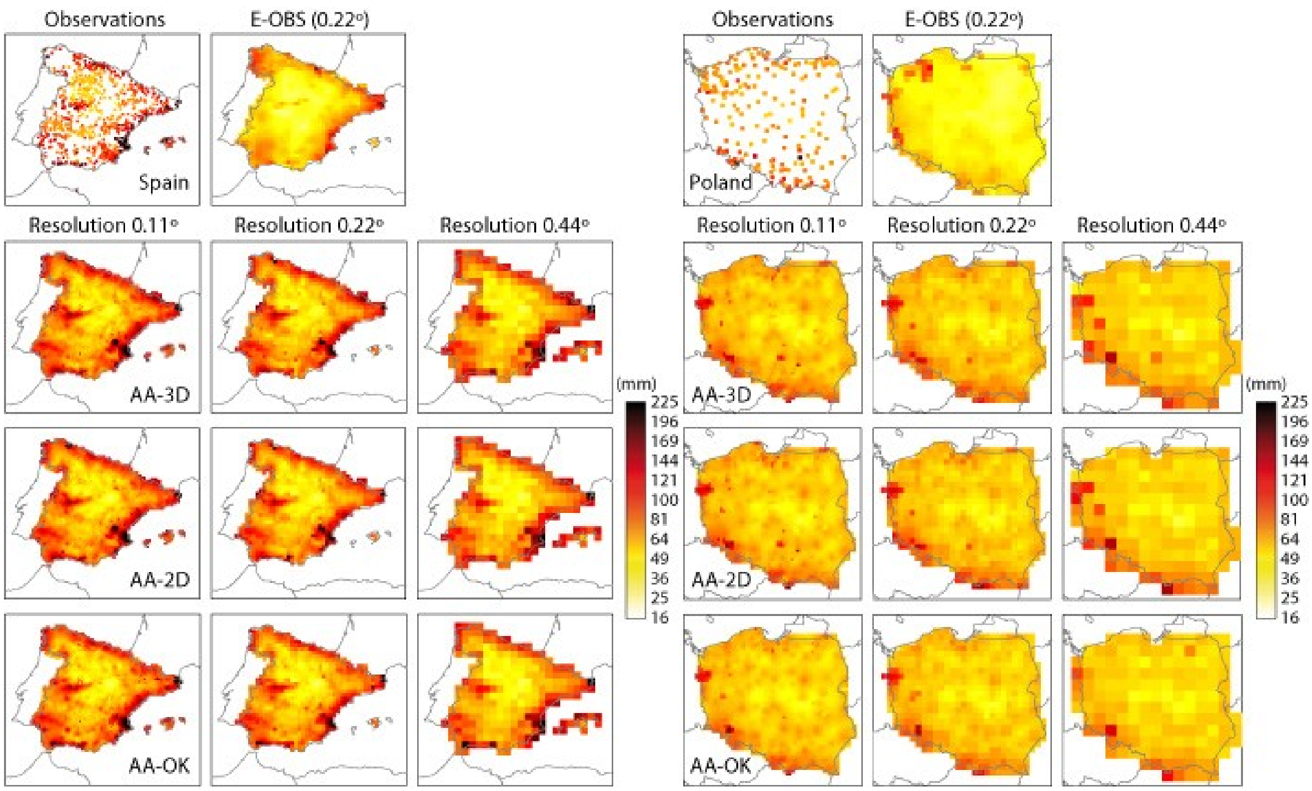


Figure 5: 50-years return value of daily precipitation based on the annual series for the period 1981-2010 for the observations and the nine (3 interpolation

methods x 3 resolutions) grids built. The 0.22 resolution of E-OBS has been included as reference.

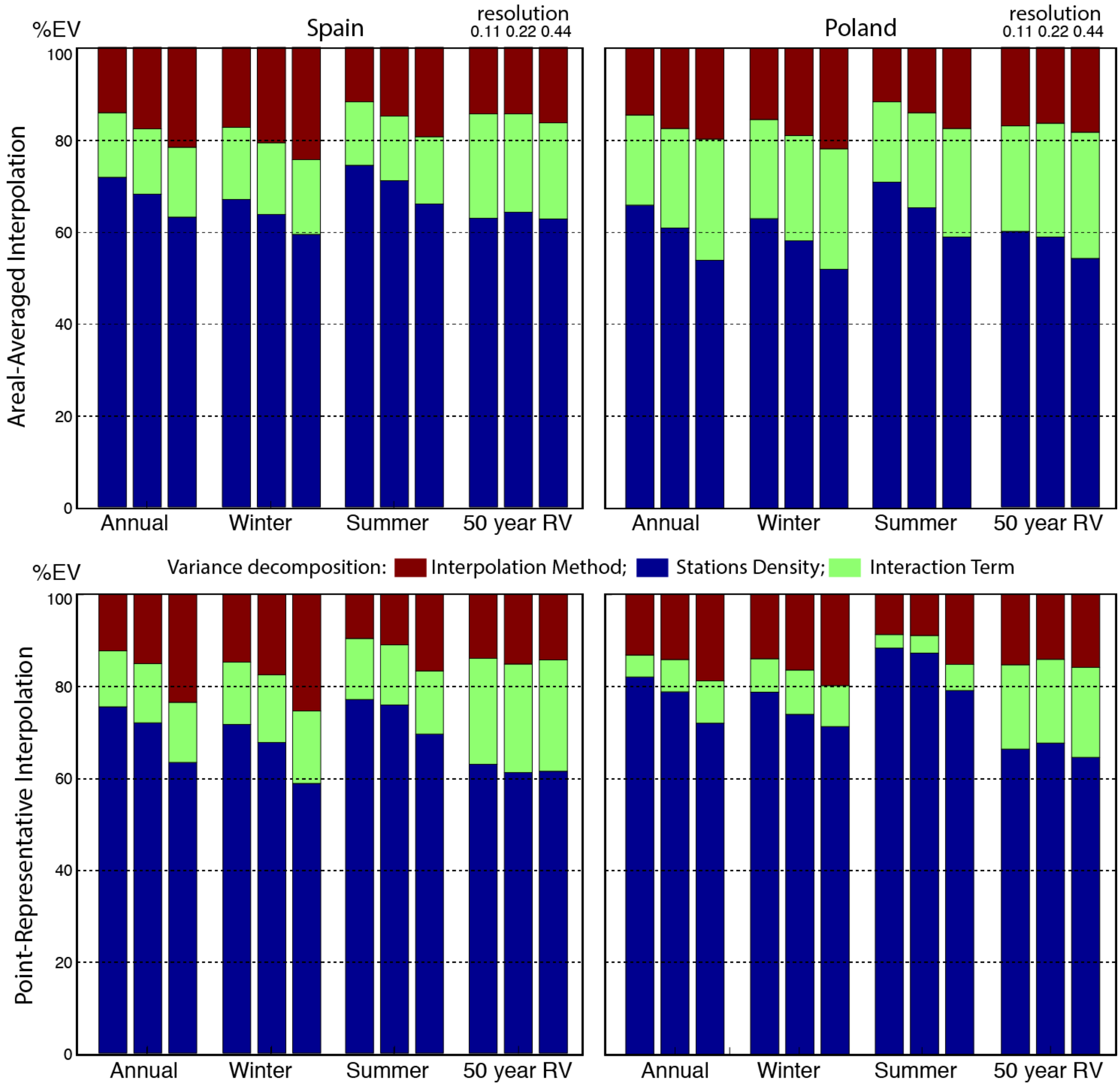


Figure 6: Percentage of explained variance for the two components: station density (blue) and interpolation method (red) for the area-averaged (top) and point-representative (bottom) interpolation methods over Spain (left) and Poland (right). Each panel shows the results for mean annual, winter and summer, and 50-year annual return values for three different resolutions (0.11, 0.22 and 0.44 deg).

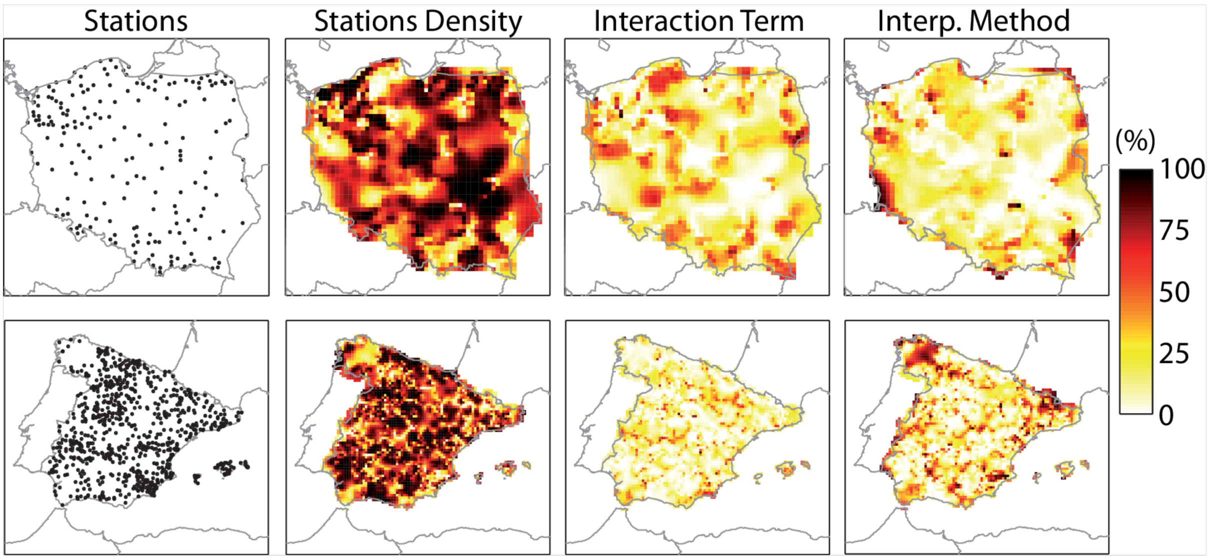


Figure 7: Spatial distribution of the variance components between the two components, interpolation method and stations density, for Poland (top) and Spain (bottom). As a reference the stations network is included (first column).

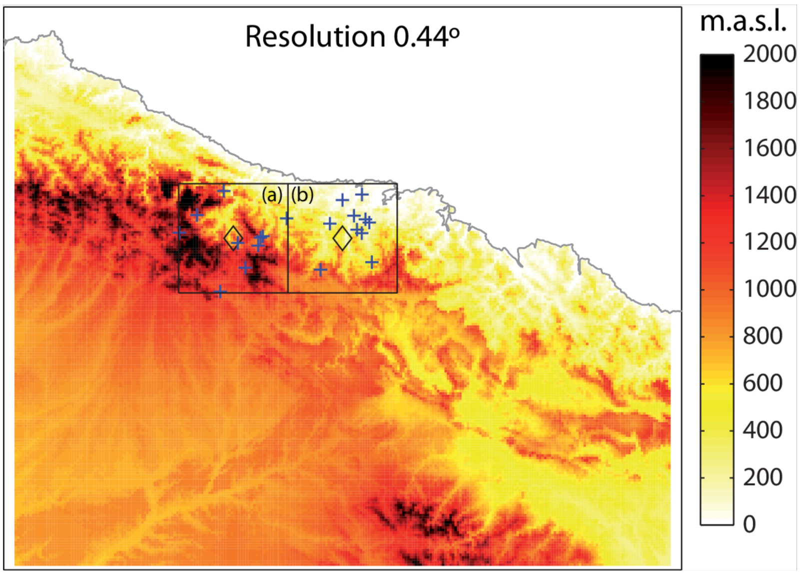
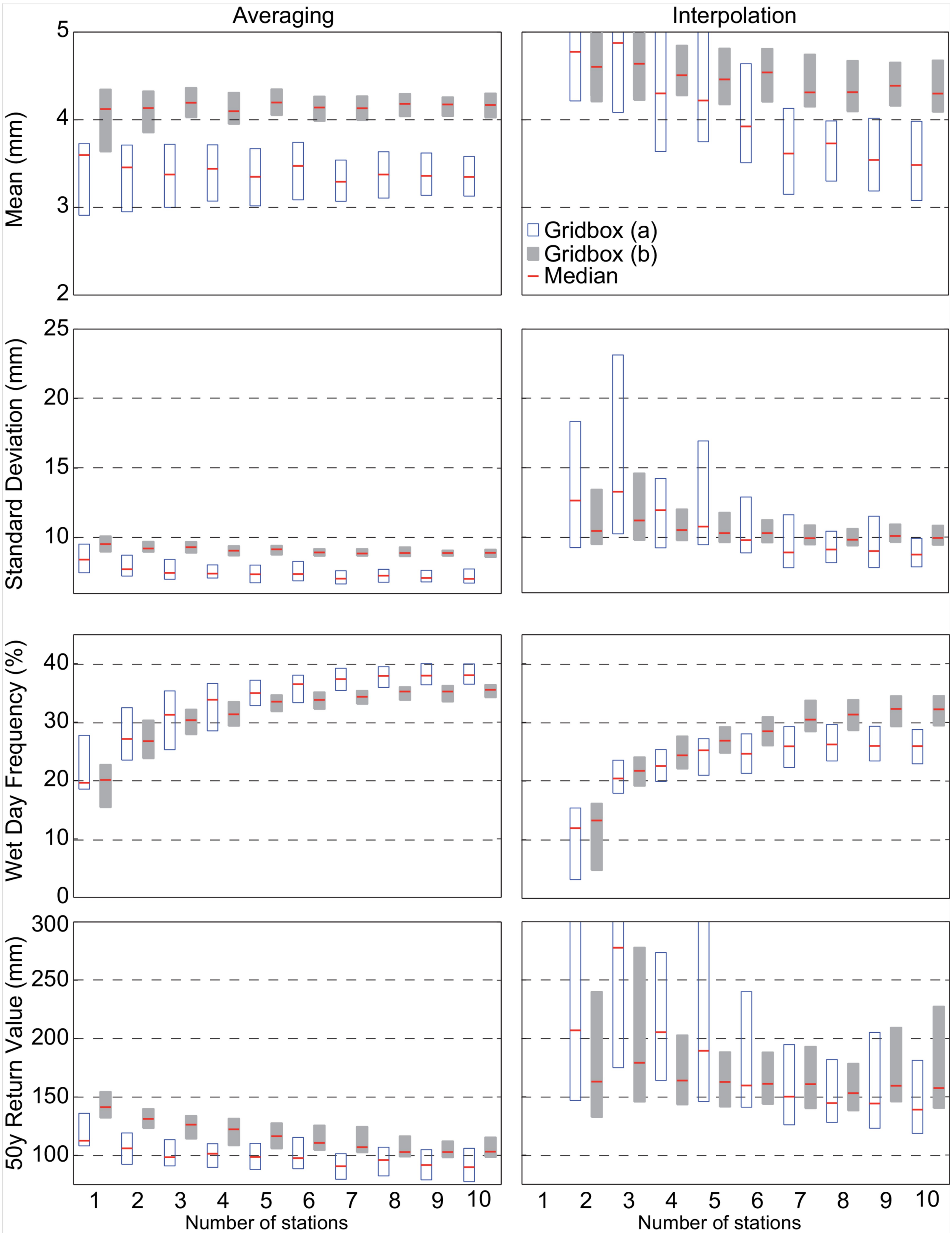


Figure 8: Highest-density grid boxes at 0.44 in Northern Spain and the stations available inside them. Colours represent the orography of the region.



743 **Figure 9:** Effective resolution analysis for the two grid boxes shown in Figure 8.
744 Box-plots show, for both the averaging (left) and interpolation (right) methods,
745 the variability of different statistics (mean daily precipitation, standard deviation
746 of daily precipitation, wet day frequency and 50-year return values, in rows) for
747 the period 1981-2010 due to the different combinations of stations considered,
748 for an increasing number of stations (m), ranging from one to ten (X-axis).

749