



## RESEARCH ARTICLE

## OPEN ACCESS

# Intercomparison of Daily Maximum and Minimum Temperature Gridded Products Over Mainland Spain

Sixto Herrera<sup>1,2</sup>  | Fidel González Rouco<sup>3</sup> | Roberto Serrano-Notivol<sup>4</sup> | Juan Luís Garrido<sup>5</sup> | Santiago Beguería<sup>6</sup>  | José M. Gutiérrez<sup>7</sup> | Pere Quintana-Seguí<sup>8</sup> | Maialen Iturbide<sup>7</sup> | Esteban Rodríguez<sup>2</sup> | Ana Morata<sup>2</sup> | Candelas Peral<sup>2</sup>

<sup>1</sup>Applied Mathematics and Computer Science Department, Universidad de Cantabria, Santander, Spain | <sup>2</sup>Grupo de Meteorología y Computación, Universidad de Cantabria, Unidad Asociada al CSIC, Santander, Spain | <sup>3</sup>Dpto. de Geografía y Ordenación Del Territorio, Instituto Universitario de Ciencias Ambientales (IUCA), Universidad de Zaragoza, Zaragoza, Spain | <sup>4</sup>División de Contaminación Atmosférica, Centro de Investigaciones Energéticas, Medioambientales Y Tecnológicas (CIEMAT), Madrid, Spain | <sup>5</sup>Estación Experimental de Aula Dei (EEAD), CSIC, Zaragoza, Spain | <sup>6</sup>Instituto de Física de Cantabria (IFCA), CSIC-Universidad de Cantabria, Santander, Spain | <sup>7</sup>Observatori de l'Ebre (OE), Universitat Ramon Llull-CSIC, Roquetes, Spain | <sup>8</sup>Spanish Meteorological Agency (AEMET), Madrid, Spain

**Correspondence:** Sixto Herrera ([sixto.herrera@unican.es](mailto:sixto.herrera@unican.es))

**Received:** 20 March 2025 | **Revised:** 22 August 2025 | **Accepted:** 26 August 2025

**Funding:** This study was supported by the MCIN/AEI (RYC2021-034330-I), Agencia Estatal de Investigación (PID2019-111481RB-I00) and European Commission – NextGenerationEU.

**Keywords:** Iberian Peninsula | intercomparison | interpolation | mainland Spain | maximum and minimum temperatures

## ABSTRACT

The sensitivity to the observational reference has been reported in recent studies, highlighting the importance of observational uncertainty in climate research. These studies stress the importance of properly comparing available datasets, recognising their respective strengths and limitations. Here, we conduct a comprehensive analysis of the various datasets of maximum and minimum daily temperatures available for mainland Spain. We examined 10 publicly available daily gridded datasets of maximum and minimum temperatures, analysing multiple evaluation dimensions to identify the key strengths and limitations of each dataset: statistical distribution, extreme events, temporal structure and spells and spatial patterns. We conclude that observational uncertainty is greater for minimum temperatures than for maximum temperatures. This uncertainty is not strictly linked to the type of dataset (interpolation, analysis or reanalysis) or its spatial domain (national, European or global) but rather to specific datasets which vary depending on the analysis dimension. Overall, the most stable dataset across all evaluated indices is STEAD, whereas the PTI-Clima v0 dataset exhibits some underestimation of extremes and spells but performs well in capturing central parameters and temporal correlations.

## 1 | Introduction

Observational data of essential climate variables (ECVs) (Bojinski et al. 2014) are fundamental for evaluating and quantifying climate evolution across global, regional and local scales. Additionally, different vulnerability, impact and adaptation (VIA) communities have specific requirements regarding the type of variables, as well as the spatial and temporal resolution needed to translate climate change signals into locally relevant

effects within their respective sectors (Bedia et al. 2013; Setti et al. 2020; Alexandridis et al. 2021). Consequently, a wide range of datasets has been developed over time using diverse sources and methodologies, including satellite observations, ground-based station networks, reanalysis products and spatially interpolated station data.

In this context, the Spanish climate research community has made significant efforts to develop historical datasets for

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

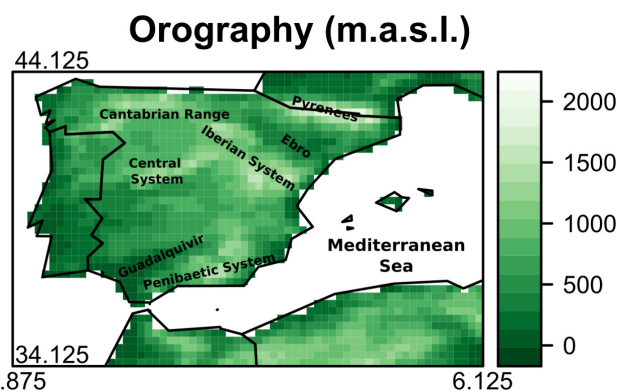
© 2025 The Author(s). *International Journal of Climatology* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society.

key meteorological variables (such as precipitation and temperature) across the national territory, ensuring adequate spatial and temporal resolution. These datasets are derived using a variety of approaches, ranging from purely interpolation methods using observations, such as STEAD (Serrano-Notivoli, Beguería, and De Luis 2019), Iberia01 (Herrera et al. 2019a) or PTI-Clima v0 (Beguería 2025, in preparation), to more sophisticated techniques that combine observations with assimilation techniques such as HUMID01 (Quintana-Seguí et al. 2017) or ROCIO-IBEB (Peral et al. 2017). In addition, global and European datasets cover the target region and complement national products. These include datasets derived from observational interpolation (e.g., E-OBS v27e: Cornes et al. (2018)) as well as datasets based on reanalysis (e.g., ERA5-Land: Muñoz Sabater (2019), CHELSA-W5E5: Karger et al. (2023), CERRA-SFC: El-Said et al. (2021) and EMO-1arcmin: Thiemig et al. (2020)).

Differences between these datasets lead to discrepancies in the analysis of current climate conditions (Burton et al. 2018; Herrera, Kotlarski, et al. 2019; Lledó et al. 2024; Newman et al. 2019; Tanarhte et al. 2012; Thorne et al. 2016), as well as in the calibration and evaluation of climate change projections (Bedia et al. 2013; Kotlarski et al. 2019; Herrera et al. 2020). Therefore, observational uncertainty is a new source of uncertainty that must be considered in climate studies. Uncertainties typically arise from differences in methodology (Quintana-Seguí et al. 2017), the quality and density of observational networks (Herrera et al. 2019b), the temporal resolution and the spatial coverage used in interpolation or numerical modelling, either at the global, European or national level (Kotlarski et al. 2019). Therefore, intercomparison studies that identify the differences, limitations, properties and potential uses/misuses of each available product are essential to interpret the results of climate studies (Newman et al. 2019).

Although most of these datasets are publicly available and intercomparison analyses have gradually emerged in the scientific literature (Burton et al. 2018; Thorne et al. 2016; Tanarhte et al. 2012; Lledó et al. 2024), no studies specifically evaluate their ability to reproduce different temperature regimes in the Iberian Peninsula. As a result, despite the availability of multiple gridded products, there is limited understanding of their differences, strengths and weaknesses. This lack of understanding has significant implications for applying these datasets in climate services and decision-making processes. Furthermore, adherence to the FAIR principles, ensuring that datasets are Findable, Accessible, Interoperable and Reusable (Wilkinson et al. 2016; Iturbide et al. 2022), becomes crucial in this context.

This study seeks to address this gap by comparing available datasets of daily maximum and minimum temperatures across mainland Spain (see Figure 1). To do this, we conduct a comprehensive analysis that takes into account various aspects of the statistical distribution, spatial patterns, temporal structure and extreme events. In Section 2 we describe the datasets used, in Sections 3 and 4 the procedure to homogenise the different datasets and the evaluation measures are described. Finally, Section 5 describes the main results, whereas Section 6 presents the conclusions and discussion with a focus on future work within the framework of the Spanish PTI-Clima.



**FIGURE 1** | Altitude of the Iberian Peninsula as given by the coarser resolution ( $0.25^\circ$ ) of the interpolated dataset E-OBS v27e (Cornes et al. 2018). [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

## 2 | Datasets

In this section, the datasets considered are first described and then the evaluation procedure is defined. Table 1 shows the main properties of the 10 datasets considered. Global and/or European-wide datasets have been highlighted in italics and bold, respectively.

### 2.1 | PTI-Clima v0

PTI-Clima v0 (Beguería 2025, in preparation) is a daily gridded dataset produced by the Climatic Interdisciplinary Thematic Platform (PTI-Clima), an initiative of the Spanish National Research Council (CSIC), as part of its Climate Services Initiative. This dataset includes information on key observational surface variables from the Spanish Meteorological Agency (AEMET) network: maximum and minimum temperature, cumulative precipitation, mean relative humidity, mean wind velocity, total insolation, mean radiation and mean surface air pressure. The dataset covers the whole territory of Spain, including the archipelagos (Balearic Islands and Canary Islands) and the autonomous African cities (Ceuta and Melilla). For daily temperature, the dataset utilises data from 5179 stations covering the period from January 1, 1961, to December 31, 2022. Data underwent a quality control process before daily grids were created using Universal Kriging (Wackernagel 2003), incorporating orography (elevation) and distance to the sea as covariates. The spatial resolution of the grid is  $0.025$  degrees (approximately  $2.5$  km).

The PTI-Clima v0 dataset is currently not disclosed to the public.

### 2.2 | STEAD

The STEAD daily maximum and minimum dataset (Spanish TEMperature At Daily scale, Serrano-Notivoli, Beguería, and De Luis 2019) is based on data from 5056 stations, gathered from various sources—including AEMET and the Spanish Ministry of Environment and Agriculture (MAGRAMA)—to enhance the density of the observational network used in the interpolation process. The interpolation method employs generalised linear mixed models (GLMM) (McCulloch and Neuhaus 2014) and generalised linear models (GLMs) (Hilbe 2011) as a general approach

**TABLE 1** | Gridded maximum and minimum temperature datasets considered in mainland Spain and their main properties (Name (approach), Spatial and Temporal resolution, Number of stations, Period and Reference).

Name (approach)	S-T resolution	N. stations	Period	References
STEAD (I)	5 km; daily	5056	1901–2014	Serrano-Notivol, Beguería, and De Luis (2019)
<b>E-OBS v27e</b> (I)	0.1°; daily	210	1951–present	Cornes et al. (2018)
Iberia01 (I)	0.1°; daily	275	1971–2015	Herrera et al. (2019a)
PTI-Clima v0 (I)	0.025°; daily	5179	1961–2022	Beguería (2025)
HUMID01 (A)	0.05°; hourly	1237	1979–2017	Quintana-Seguí et al. (2017)
ROCIO-IBEB (A)	0.05°; daily	1800	1951–present	Peral et al. (2017)
<i>ERA5-Land</i> (R)	0.1°; hourly	—	1950–present	Muñoz Sabater (2019)
<i>CHELSA-W5E5</i> (R)	0.01°; daily	—	1979–2016	Karger et al. (2023)
<b>EMO-1arcmin</b> (R)	1.5 km; daily	18,964	1990–2022	Thiemig et al. (2020)
<b>CERRA-SFC</b> (R)	5.5 km; 1- and 3-hourly	—	1984–2021	El-Said et al. (2021)

Note: Global and European-wide datasets have been highlighted in italics and bold, respectively. When available, the number of stations used to construct the dataset has been included. The datasets have been grouped according to the methods used to build them: observational interpolation (I), analysis-type (A) and reanalysis-based (R) products.

to estimate daily temperatures, using as support environmental variables and monthly estimates based on daily data of months with complete observations. STEAD also provides, for all daily estimates, their corresponding uncertainties extracted from the individual models. The STEAD dataset is publicly available through the Digital.CSIC repository<sup>1</sup> (Serrano-Notivol, De Luis, and Beguería 2019).

### 2.3 | E-OBS v27e

E-OBS (Haylock et al. 2008; Cornes et al. 2018, 2020) is the main reference gridded dataset for Europe, due to its relatively high spatial resolution (0.1°), the daily resolution and the availability of different meteorological parameters that have been progressively included in the dataset. In addition, the ensemble version lets the user take into account the uncertainty of the estimate. However, the number of stations considered for Spain is very low, with 210 stations approximately for temperature and precipitation.

The dataset is publicly available from the Copernicus Climate Data Store (Cornes et al. 2020) and also from the European Climate Assessment & Dataset (ECA&D). The version considered in this paper is E-OBS v27e, which was downloaded in April 2023, being the latest version available at that time.

### 2.4 | Iberia01

Iberia01 (Herrera et al. 2019a, 2019b) is a daily gridded precipitation and temperature (maximum, minimum and mean) dataset for the Iberian Peninsula and the Balearic Islands. It relies on a dense, quality checked (see details in Herrera 2011; Herrera et al. 2012) observational network—3485 and 275 stations for precipitation and temperatures, respectively—sourced from AEMET, the Portuguese Institute for Sea and Atmosphere (IPMA) and the Portuguese Environmental Agency (APA).

The Iberia01 gridded dataset is made available under the Open Database Licence<sup>2</sup> (Herrera et al. 2019b).

### 2.5 | HUMID01—Spain

The SAFRAN dataset (Système d'Analyse Fournissant des Renseignements Atmosphériques à la Neige) was initially developed to produce the boundary conditions to drive a snow model (Brun et al. 1989) 'considering 1237 precipitation stations'. The HUMID01-Spain dataset is an updated version of SAFRAN-Spain, extending to 2017. HUMID01 combines observations from AEMET with a first guess (i.e., the outputs of a numerical weather prediction model) derived from the ERA-Interim reanalysis (Dee et al. 2011) provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). This integration employs an optimal interpolation algorithm (Gandin 1966; Quintana-Seguí et al. 2017), which was initially applied in France (Quintana-Seguí et al. 2008; Vidal et al. 2010) and subsequently in other regions, including Spain (Quintana-Seguí et al. 2017).

The primary objective of SAFRAN and HUMID01 is to estimate the meteorological screen-level variables required to drive a Land-Surface Model (LSM). They generate gridded datasets that include temperature, precipitation, wind speed and other essential variables. The outputs have an hourly time step; however, HUMID01 only assimilates observed data every 6 h (24 h for precipitation).

Although SAFRAN-Spain is publicly available for the scientific community,<sup>3</sup> HUMID01 has not been published yet.

### 2.6 | ROCIO-IBEB

The Spanish Meteorological Agency (AEMET) has developed several gridded products, some utilising interpolation at varying

resolutions, whereas others employ optimal interpolation that combines observational data with numerical model outputs. In this study, we focus on ROCIO-IBEB, formerly known as AEMET-5km v2 (Peral et al. 2017), the latest release in this series. ROCIO-IBEB is derived from a dense network of 1800 quality-controlled stations and incorporates surface analysis from the operational short-range weather forecast, using optimal interpolation to merge these inputs. This dataset is publicly and freely distributed for use in research work through the AEMET data server.<sup>4</sup>

When needed, this dataset has been used as a reference in this study for two main reasons. First, AEMET is the official national authority for meteorological information and services, and this dataset is currently maintained and used by the institution as a standard reference. Second, since all datasets are primarily based on the observational network managed by AEMET, using that same network for validation would introduce challenges, such as issues of spatial representativeness (e.g., point vs. area-averaged values) (Osborn and Hulme 1997) and varying degrees of dependence among datasets, as many are constructed using the same or highly similar station data. As a result, such evaluations would be difficult to interpret. It is also important to note that the objective of this work is not to independently validate each dataset, but to compare them and highlight their differences. Therefore, the reference dataset is used here for illustrative purposes only.

## 2.7 | EMO—1arcmin

The EMO (European Meteorological Observations, Gomes et al. 2020; Thiemig et al. 2020, 2022) is a high resolution Copernicus Emergency Management Service product that includes, among others, daily maximum and minimum temperature data for Europe at two different spatial resolutions: 5 km (EMO-5) and 1.5 km (EMO-1arcmin). EMO is built by integrating data from 18,964 stations, four high resolution regional observational grids (CombiPrecip, Sideris et al. (2014); ZAMG-INCA, Haiden et al. (2011); EURO4M-APGD, Isotta et al. (2014); and CarpatClim, Antolović et al. (2013); Spinoni et al. (2015)) and one global reanalysis (ERA-Interim/Land, Dee et al. (2011)). The raw observations were interpolated using a modified version of the SPHEREMAP (Willmott et al. 1985) deterministic algorithm to obtain the estimates, whereas the uncertainty was calculated using the formulation defined by Yamamoto (2000).

In this paper, we have used the EMO-1arcmin version, the one currently supported, which was downloaded in May 2023 from the public repository.<sup>5</sup>

## 2.8 | ECMWF—ERA5-Land

ERA5-Land (Muñoz Sabater 2019) is a global reanalysis focused on land areas that improves the spatial resolution of the ERA5 reanalysis (Hersbach et al. 2020) from ECMWF. It combines model data and observations using the laws of physics and can thereby be considered a physical interpolation of observations. To control the simulation, ERA5-Land uses the simulated land

**TABLE 2** | Indices considered for the validation of the different datasets. The reference considered has always been ROCIO-IBEB. The ?? code refers to an index applied to maximum and minimum temperatures.

Code	Description	Units	Dimension
<i>tasm</i> ?? <sub>clim</sub>	Mean of daily values	°C	SD
<i>tasm</i> ?? <sub>std</sub>	Standard deviation of daily values	°C	SD
<i>tasm</i> ?? <sub>50th</sub>	Median of daily values	°C	SD
<i>tasmax</i> <sub>95th</sub>	95th percentile of daily values	°C	SD
<i>tasmax</i> <sub>99th</sub>	99th percentile of daily values	°C	SD
<i>tasmin</i> <sub>5th</sub>	5th percentile of daily values	°C	SD
<i>tasmin</i> <sub>1th</sub>	1th percentile of daily values	°C	SD
<i>rv50y</i>	50-years return value of annual max. and min. temperatures	°C	EE
<i>GEV</i> <sub>shape</sub>	Shape parameter of annual max. and min. temperatures	—	EE
<i>tasmax</i> <sub>13Aug 11Aug</sub>	Mean of daily maximum temperature in event 11–13 Aug. 2003	°C	EE
<i>tasmin</i> <sub>2Feb 26Jan</sub>	Mean of daily min. temperature in event 26 Jan. to 2 Feb. 2005	°C	EE
<i>rho</i>	Correlation	—	TS
<i>rho</i> <sub>annual-cycle</sub>	Correlation of the annual cycle	—	TS
<i>nhw</i> <sub>25</sub> <sup>3</sup>	Number of Summer day Spells (3 days) with threshold 25°C	days	SS
<i>ntn</i> <sub>20</sub> <sup>3</sup>	Number of Tropical night Spells (3 days)	days	SS

Abbreviations: EE, Extreme events; SD, Statistical distribution; SS, Spells; TS, Temporal structure.

fields of ERA5 atmospheric variables as atmospheric forcing. Furthermore, before running ERA5-Land, several input variables (air temperature, air humidity and pressure) are corrected to account for the altitude difference between the resolution of atmospheric forcing and the higher resolution grid of ERA5-Land (lapse rate correction).



**TABLE 3** | Percentage of grid-points with statistically significant (95%) mean differences for maximum temperature (upper triangle) and minimum temperature (lower triangle, in bold).

Approach	Interpolation				Analysis		Reanalysis			
Dataset	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
D1	0.00	86.50	87.11	93.97	96.66	84.77	84.17	83.82	49.73	89.17
D2	<b>88.90</b>	0.00	70.25	67.17	99.33	62.88	96.08	71.03	85.52	99.47
D3	<b>88.86</b>	<b>84.46</b>	0.00	74.01	98.12	73.76	94.86	79.55	84.68	95.75
D4	<b>88.77</b>	<b>74.69</b>	<b>82.76</b>	0.00	100.00	66.80	100.00	72.23	85.64	99.04
D5	<b>99.87</b>	<b>97.73</b>	<b>97.05</b>	<b>98.77</b>	0.00	99.86	74.08	97.06	94.07	71.12
D6	<b>78.19</b>	<b>78.90</b>	<b>88.05</b>	<b>78.70</b>	<b>99.86</b>	0.00	95.84	73.15	77.64	99.32
D7	<b>90.31</b>	<b>81.73</b>	<b>89.17</b>	<b>83.47</b>	<b>94.00</b>	<b>85.71</b>	0.00	93.64	86.80	67.25
D8	<b>91.31</b>	<b>85.45</b>	<b>90.31</b>	<b>86.46</b>	<b>93.58</b>	<b>89.45</b>	<b>83.90</b>	0.00	82.04	94.44
D9	<b>93.67</b>	<b>79.36</b>	<b>84.68</b>	<b>81.12</b>	<b>94.74</b>	<b>88.20</b>	<b>79.18</b>	<b>86.33</b>	0.00	90.48
D10	<b>92.51</b>	<b>86.64</b>	<b>87.38</b>	<b>86.18</b>	<b>89.30</b>	<b>89.73</b>	<b>84.44</b>	<b>86.38</b>	<b>78.82</b>	0.00

Note: Results were obtained with the statistical *t*-test to compare the mean of two distributions. D1: STEAD; D2: E-OBS v27e; D3: Iberia01; D4: PTI-Clima v0; D5: HUMID01; D6: ROCIO-IBEB; D7: ERA5-Land; D8: CHELSA-W5E5; D9: EMO-1arcmin; D10: CERRA-SFC.

**TABLE 4** | Percentage of grid points with statistically significant (95 % confidence) differences in centred daily temperature distributions—that is, after subtracting the long-term mean at each grid point—between each pair of datasets for maximum temperature (upper triangle) and minimum temperature (lower triangle, in bold).

Approach	Interpolation				Analysis		Reanalysis			
Dataset	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
D1	0.00	43.72	61.21	35.89	37.43	39.37	49.25	56.28	43.80	47.73
D2	<b>97.73</b>	0.00	51.20	31.05	47.06	37.26	59.27	39.81	30.03	51.85
D3	<b>96.38</b>	<b>54.85</b>	0.00	34.88	61.48	39.15	75.91	68.66	55.24	56.57
D4	<b>85.07</b>	<b>49.93</b>	<b>47.61</b>	0.00	53.42	16.46	70.52	57.05	42.00	47.47
D5	<b>98.13</b>	<b>49.60</b>	<b>63.36</b>	<b>68.90</b>	0.00	49.93	36.02	46.66	34.23	37.30
D6	<b>82.99</b>	<b>58.77</b>	<b>60.58</b>	<b>11.89</b>	<b>70.37</b>	0.00	67.68	58.08	43.35	49.45
D7	<b>99.18</b>	<b>70.77</b>	<b>77.94</b>	<b>86.64</b>	<b>58.25</b>	<b>84.47</b>	0.00	38.57	57.55	48.99
D8	<b>91.44</b>	<b>52.25</b>	<b>66.93</b>	<b>46.92</b>	<b>63.64</b>	<b>54.79</b>	<b>74.97</b>	0.00	45.44	61.38
D9	<b>98.11</b>	<b>52.14</b>	<b>57.53</b>	<b>44.32</b>	<b>79.11</b>	<b>58.85</b>	<b>94.97</b>	<b>67.56</b>	0.00	45.17
D10	<b>97.19</b>	<b>66.80</b>	<b>72.78</b>	<b>68.13</b>	<b>56.82</b>	<b>68.36</b>	<b>58.86</b>	<b>73.41</b>	<b>83.11</b>	0.00

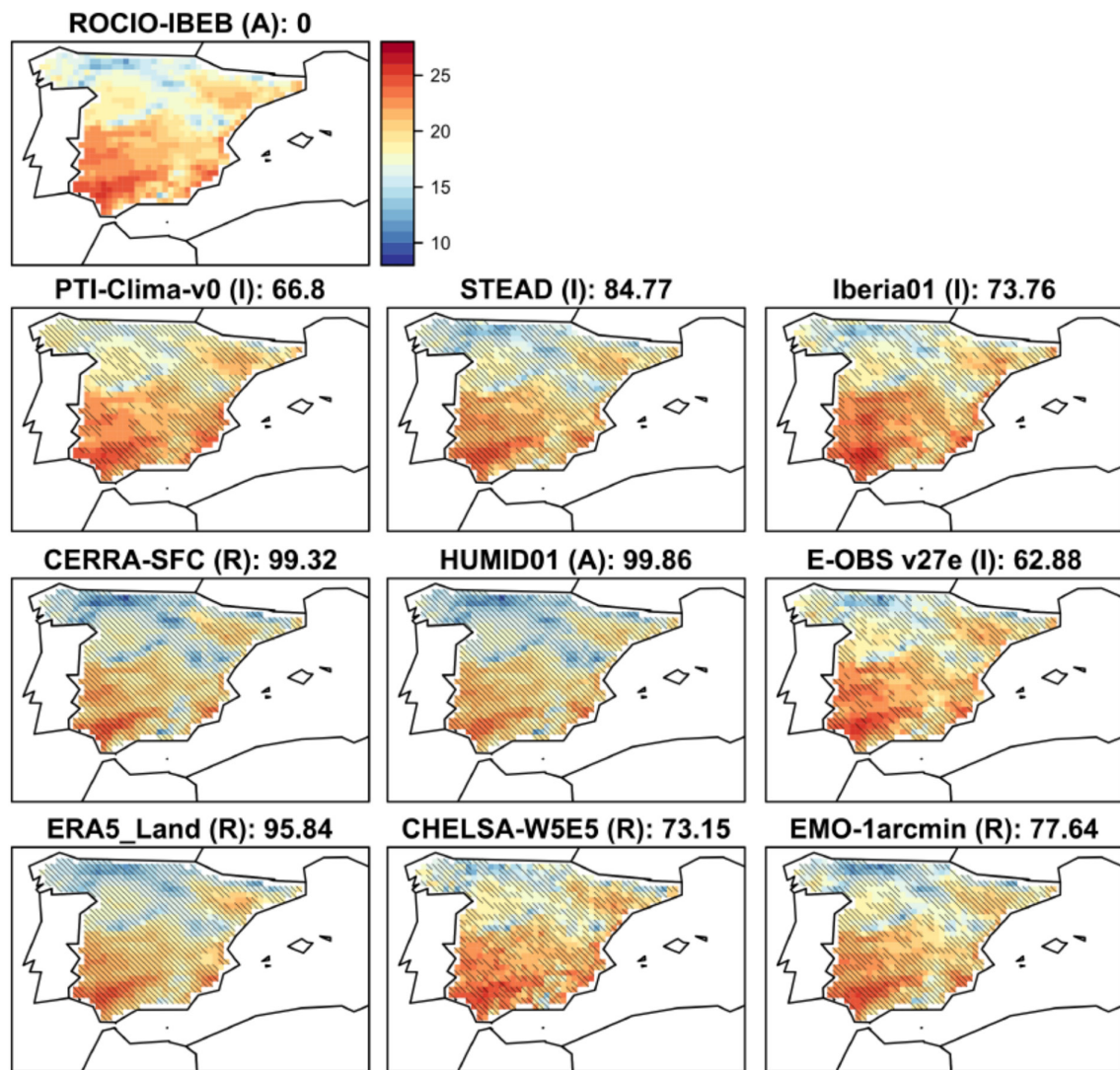
Note: Results were obtained using a two-sided Kolmogorov–Smirnov test on the zero-mean series. D1: STEAD; D2: E-OBS v27e; D3: Iberia01; D4: PTI-Clima v0; D5: HUMID01; D6: ROCIO-IBEB; D7: ERA5-Land; D8: CHELSA-W5E5; D9: EMO-1arcmin; D10: CERRA-SFC.

This dataset is provided through the Copernicus Climate Data Store (Hersbach et al. 2023) under the terms of use defined for the Copernicus Products. The dataset was downloaded in May 2023 from the Copernicus Climate Data Store.<sup>6</sup>

## 2.9 | CHELSA-W5E5

CHELSA (Climatologies at High resolution for the Earth's Land Surface Areas, Karger et al. 2023) is a global climate dataset built to provide free access to high-resolution climate data for a wide range of research and applications related to impact studies. It is

the product of statistically downscaling global reanalysis data—specifically WFDE5 (WATCH Forcing Data methodology applied to ERA5, Cucchi et al. 2020) over land merged with ERA5 (Hersbach et al. 2020) over the ocean (W5E5) v1.0 (Lange 2019)—and takes into account orographic predictors such as wind, topographic exposure and boundary layer height (Karger et al. 2021). In this paper, we used the CHELSA-W5E5 version (Karger et al. 2022), which is a spatially downscaled variant of the dataset W5E5 (Weedon et al. 2014; Lange 2019; Cucchi et al. 2020). This downscaling was performed using the CHELSA V2 topographic downscaling algorithm (Karger et al. 2021). The dataset is publicly available through the Inter-Sectoral Impact Model



**FIGURE 2** | Annual mean of the daily maximum temperature for the period 1990–2014 (°C). Grid-points with statistically significant differences (95%) in the mean w.r.t. ROCIO-IBEB according to the *t*-test results have been hatched. The percentage of the grid-points is also shown for each dataset in the title. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/joc.70111)]

Intercomparison Project (ISIMIP) repository under the CC0 1.0 Universal Public Domain Dedication licence. It was downloaded from the data server in May 2023.<sup>7</sup>

## 2.10 | CERRA-SFC

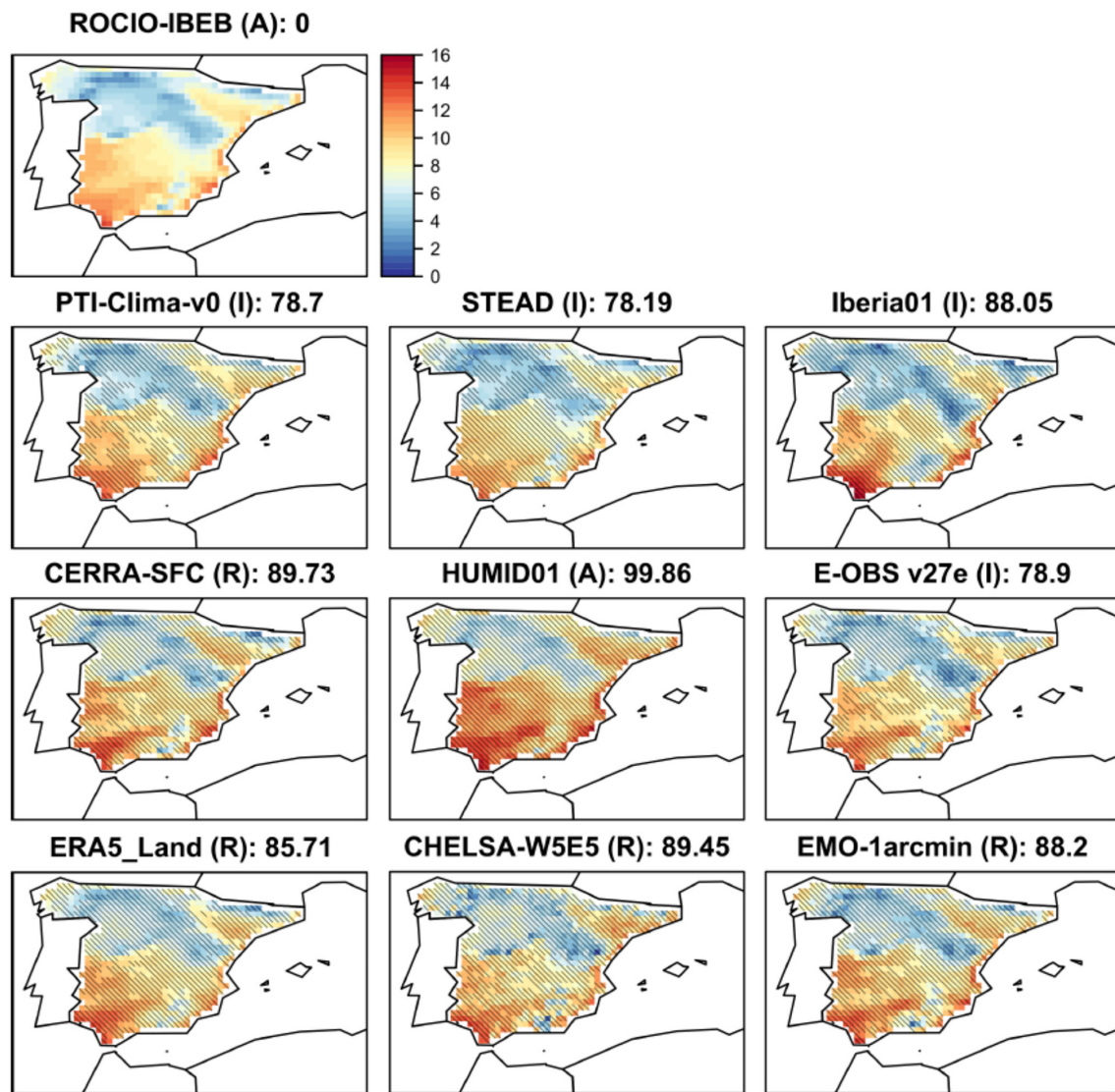
The CERRA (Copernicus European Regional ReAnalysis, El-Said et al. 2021; Schimanke et al. 2021) dataset is a regional reanalysis that uses the HARMONIE-ALADIN (Bengtsson et al. 2017) limited-area numerical weather prediction and data assimilation system, with lateral boundary conditions provided by the ERA5 global reanalysis. In this paper, we focus on the subset corresponding to single levels (atmospheric and surface quantities) and the daily maximum and minimum temperature forecasts obtained with an analysis time of 00:00 and a lead time of 24h. The dataset was downloaded from its data server in February 2024.<sup>8</sup>

## 3 | Methods

### 3.1 | Data Homogenisation

Each described dataset has its own temporal and spatial resolution (see Table 1), including variations in map projections and temporal references. Therefore, all datasets were aligned to a common grid and a time reference to allow effective comparative analysis. Two exceptions were made to obtain more meaningful results. First, the extreme value analysis was conducted using the entire available period for each dataset. Second, the correlation analysis was performed using the period common to each dataset and the reference dataset, ROCIO-IBEB, rather than the period common to all the considered datasets (1990–2014).

To ensure a fair comparison between the datasets, we established a common grid that none of the datasets had originally used. We opt for a coarser regular grid that covers mainland



**FIGURE 3** | Annual mean of the daily minimum temperature for the period 1990–2014 (°C). Grid-points with statistically significant differences (95%) in the mean w.r.t. ROCIO-IBEB according to the *t*-test results have been hatched. The percentage of the grid-points is also shown for each dataset in the title. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/joc.7011)]

Spain with a resolution of  $0.25^\circ$ . Then, all data sets were re-defined to fit this grid using a conservative interpolation scheme (Zhuang et al. 2023). It is important to note that only the maximum and minimum temperatures were interpolated, excluding the uncertainty/error measures provided in some datasets that were not considered in this paper. For temporal homogenisation, we used the common period across the 10 datasets (1990–2014) at a daily resolution with a consistent time reference.

#### 4 | Intercomparison Parameters

To identify the main characteristics, shortcomings and advantages of the 10 datasets, we considered several parameters (Table 2) that describe different aspects of the statistical distribution, spatial patterns, temporal structure (annual cycles and spells) and extreme events (EE).

Table 2 describes the parameters considered, including the statistics evaluated with them. All these parameters have been estimated for each grid point, obtaining a spatial pattern for each dataset. Two extreme warm/cold events that occurred in the Iberian Peninsula in August 2003 (AEMet 2023a) and January–February 2005 (AEMet 2023b), respectively, have additionally been used to evaluate how each dataset reproduces the intensity and spatial pattern and better understand the uncertainties related to how different datasets can reproduce extreme events.

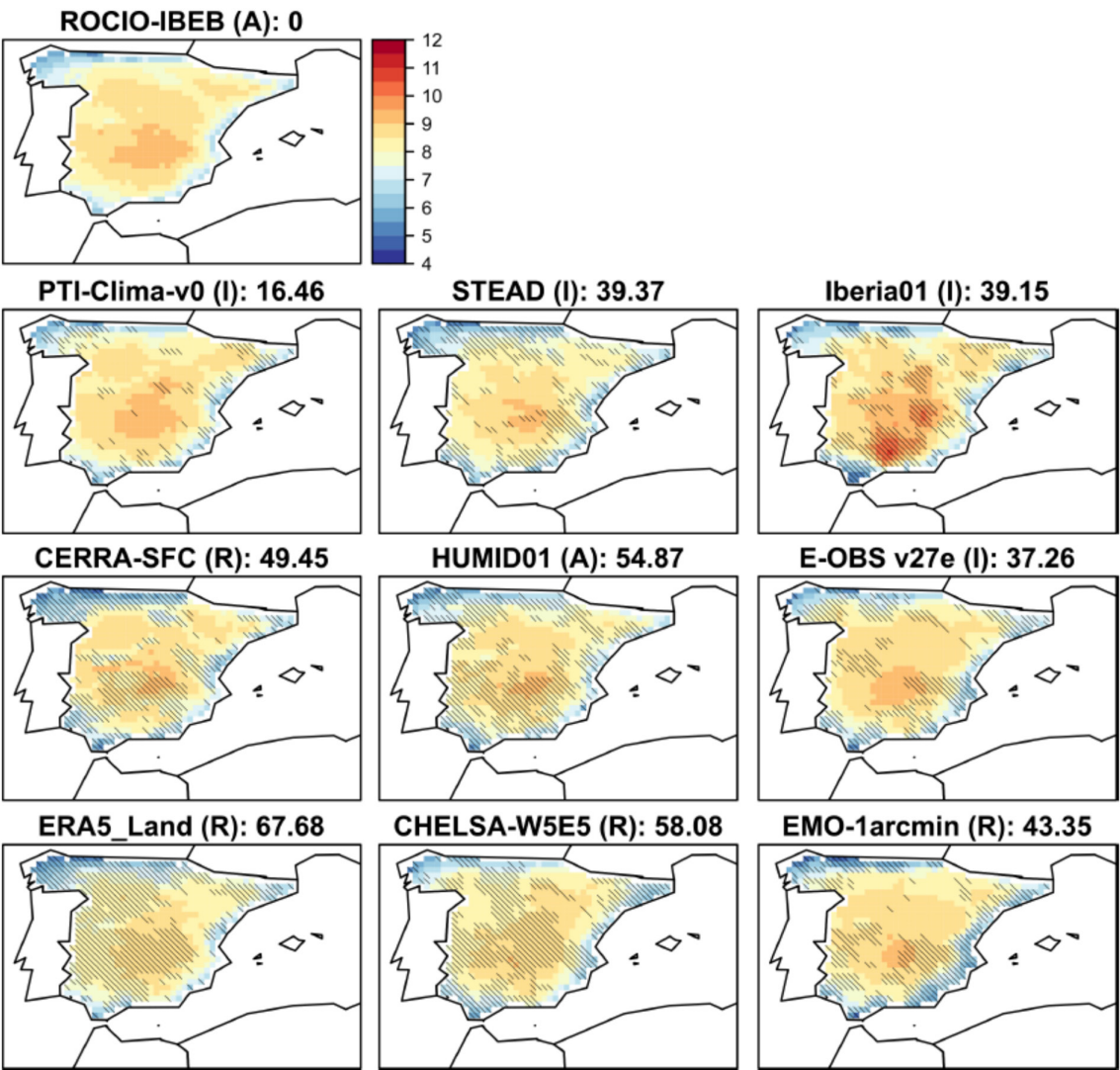
##### 4.1 | Statistical Distribution (SD)

The first seven indices in Table 2 describe the parameters used to compare the statistical distributions provided by the different datasets. To evaluate the central tendency of the distribution, the mean and median daily values were calculated, along with a statistical test to compare the means of two samples (*t*-test).



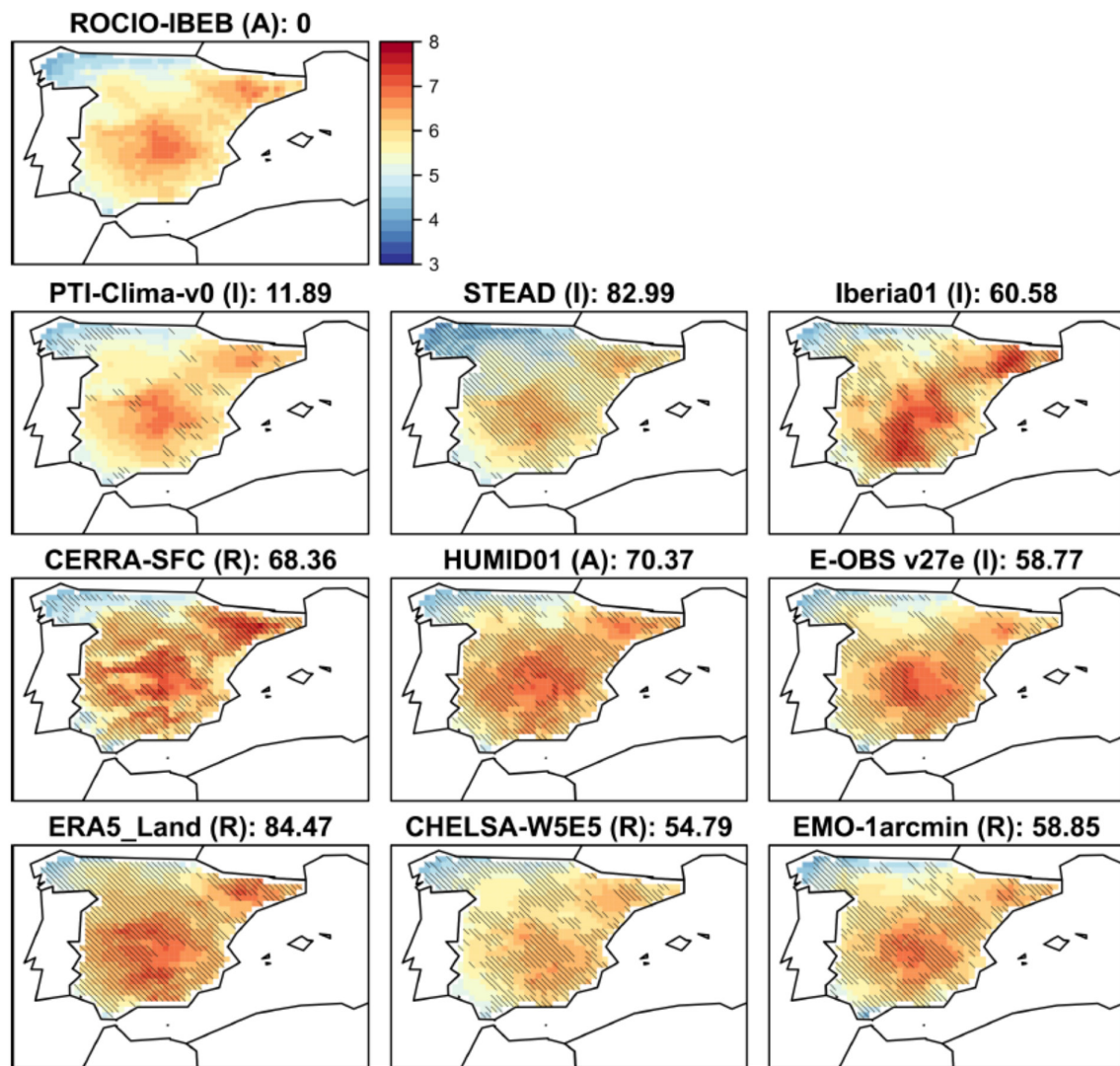
**TABLE 5** | Structural validation of the spatial pattern. Degrees of Freedom (DoF) based on the principal components as defined in Widmann et al. (2019) (columns 1 and 2). It should be noted that the number of cells is 730, which would be the default value in case all the time series were independent. The mean absolute error for the spatial patterns corresponding to the mean (columns Tx and Tn) and the different percentiles considered for the maximum (columns Tx95 and Tx99) and minimum (columns Tn05 and Tn01) temperatures is also provided in columns 3–8. The reference considered has been highlighted in bold letters.

Dataset	DoF Tx	DoF Tn	Tx	Tn	Tx95	Tx99	Tn05	Tn01
STEAD	8.11	11.23	0.61	0.39	1.12	1.35	0.98	0.70
E-OBS v27e	7.56	10.37	0.41	0.54	0.64	0.68	0.85	0.74
Iberia01	10.37	14.11	0.58	0.83	0.74	0.89	1.19	1.03
PTI-Clima v0	7.31	7.81	0.59	0.48	0.67	0.70	0.51	0.47
HUMID01	10.72	12.85	1.58	1.61	1.78	1.51	1.10	1.26
<b>ROCIO-IBEB</b>	4.47	4.45	0.00	0.00	0.00	0.00	0.00	0.00
ERA5-Land	6.52	12.65	1.13	0.66	1.44	1.34	0.98	0.65
CHELSA-W5E5	5.75	8.63	0.69	0.98	0.79	0.82	1.13	1.04
EMO-1arcmin	10.96	14.85	0.61	0.71	1.00	0.87	0.76	0.73
CERRA-SFC	11.94	14.84	1.42	0.86	1.75	1.58	0.73	0.64



**FIGURE 4** | Standard deviation of maximum temperature for the period 1990–2014 (°C). Grid-points with statistically significant differences (95%) in the centred statistical distribution w.r.t. ROCIO-IBEB according to the Kolmogorov–Smirnov test results have been hatched. The percentage of the grid-points is also shown for each dataset in the title. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/joc.70111)]





**FIGURE 5** | Standard deviation of minimum temperature for the period 1990–2014 (°C). Grid-points with statistically significant differences (95%) in the centred statistical distribution w.r.t. ROCIO-IBEB according to the Kolmogorov–Smirnov test results have been hatched. The percentage of the grid-points is also shown for each dataset in the title. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/joc.70111)]

To assess variability, several percentiles of the distribution (1th and 5th were obtained for the minimum temperature, and 99th and 95th for the maximum temperature), together with standard deviations. To assess the statistical significance of the differences between percentiles (1th and 99th for the minimum and maximum temperature, respectively), we applied a block bootstrap implementation of the hypothesis test for percentile difference (Kunsch 1989), using 1000 replicates and a block length of 365 days to account for the autocorrelation of the time series, with the same confidence level used in the other tests (95 %).

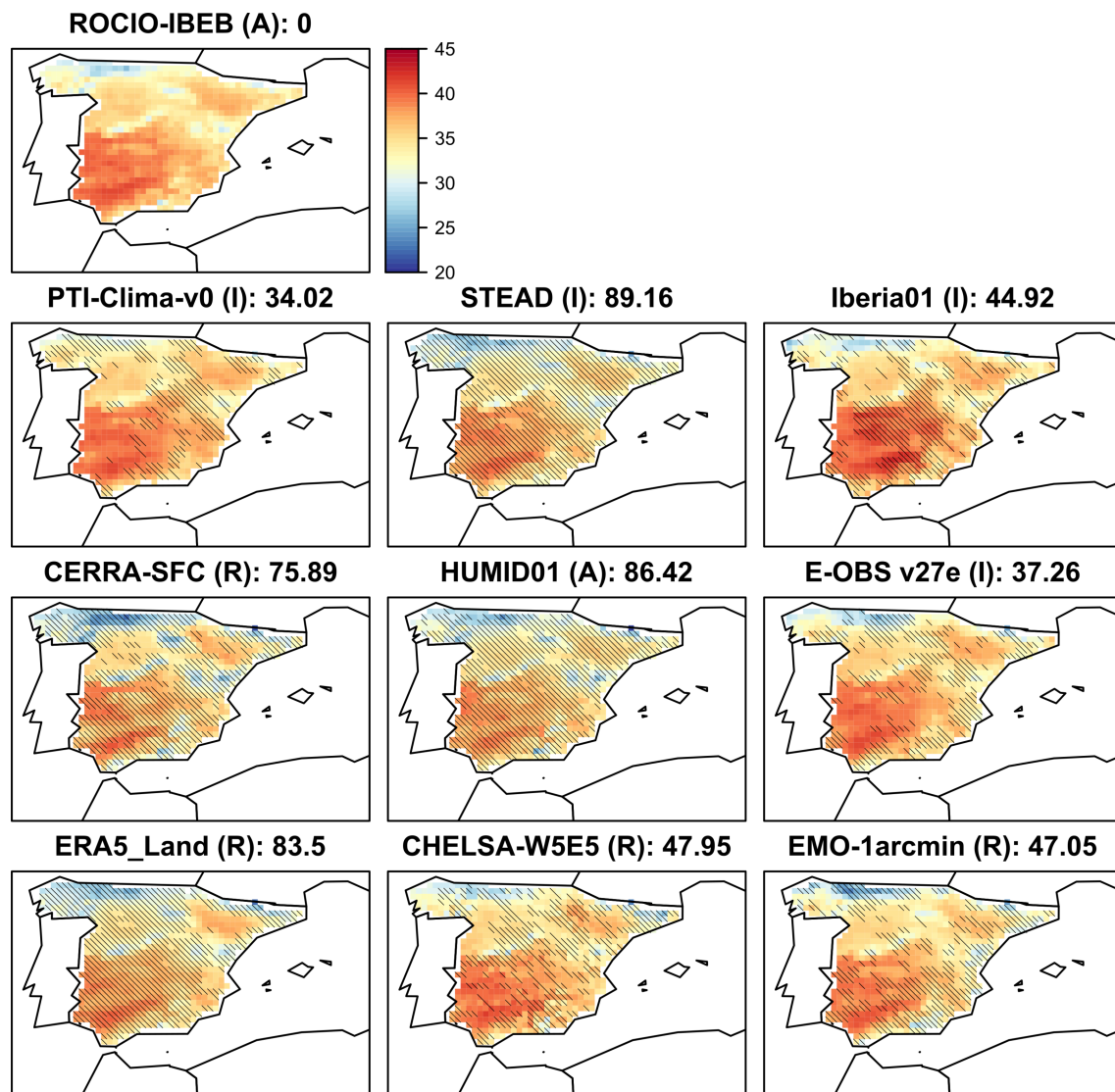
Finally, at each grid point, we subtracted the long-term mean from both the ROCIO-IBEB and comparative dataset time series—yielding zero-mean (centred) distributions—and then applied the two-sample Kolmogorov–Smirnov (KS) test to these centred series. For the *t*-test (on raw series), the KS test (on centred series) and the percentile difference test, we calculated the percentage of grid points showing statistically significant differences at the 95% confidence level. It should be noted that the *t*-test and the KS test were applied to daily data, and

thus their results are influenced by the large sample size (over 9000 days) and by the autocorrelation inherent in the time series. Nonetheless, repeating the analyses with monthly or yearly data (not shown) led to similar conclusions.

#### 4.2 | Extreme Events (EE)

Two approaches have been considered to evaluate EE. On the one hand, we use the classical extreme theory to adjust a Generalised Extreme Value (GEV) distribution to the time series of annual maxima (Rypkema and Tuljapurkar 2021).

Depending on its shape parameter ( $\xi$ ), the GEV distribution includes three distribution families corresponding to different types of tail behaviour (Coles 2001): the Gumbel family (Gumbel 1935), the case  $\xi = 0$ ; the Fréchet distribution (Ramos et al. 2018), with  $\xi > 0$ ; and the Weibull family (Padgett 2011), with  $\xi < 0$  and a bounded tail. In this work, we analyse the shape parameter and the 50-year return value given by the



**FIGURE 6** | 99th percentile of maximum temperature for the period 1990–2014 (°C). Grid-points with statistically significant differences (95%) in the percentile w.r.t. ROCIO-IBEB have been hatched. The percentage of the grid-points is also shown for each dataset in the title. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/joc.7011)]

corresponding quantile of the distribution (Mínguez and Herrera 2023). Note that we work with the transformed values of the minimum temperatures, defined as  $aux_n = -tasmin_n$ , to apply the annual maximum analysis. Consequently, the results are transformed accordingly ( $tasmin_{rv50y} = -tasmin_{rv50y}$ ) to obtain the cold 50-year return value. Furthermore, we consider the warm (11 August 2003–13 August 2003) and cold (26 January 2005–2 February 2005) extreme events identified by AEMet (2023a, 2023b) and compare the spatial patterns and intensities provided by the different datasets.

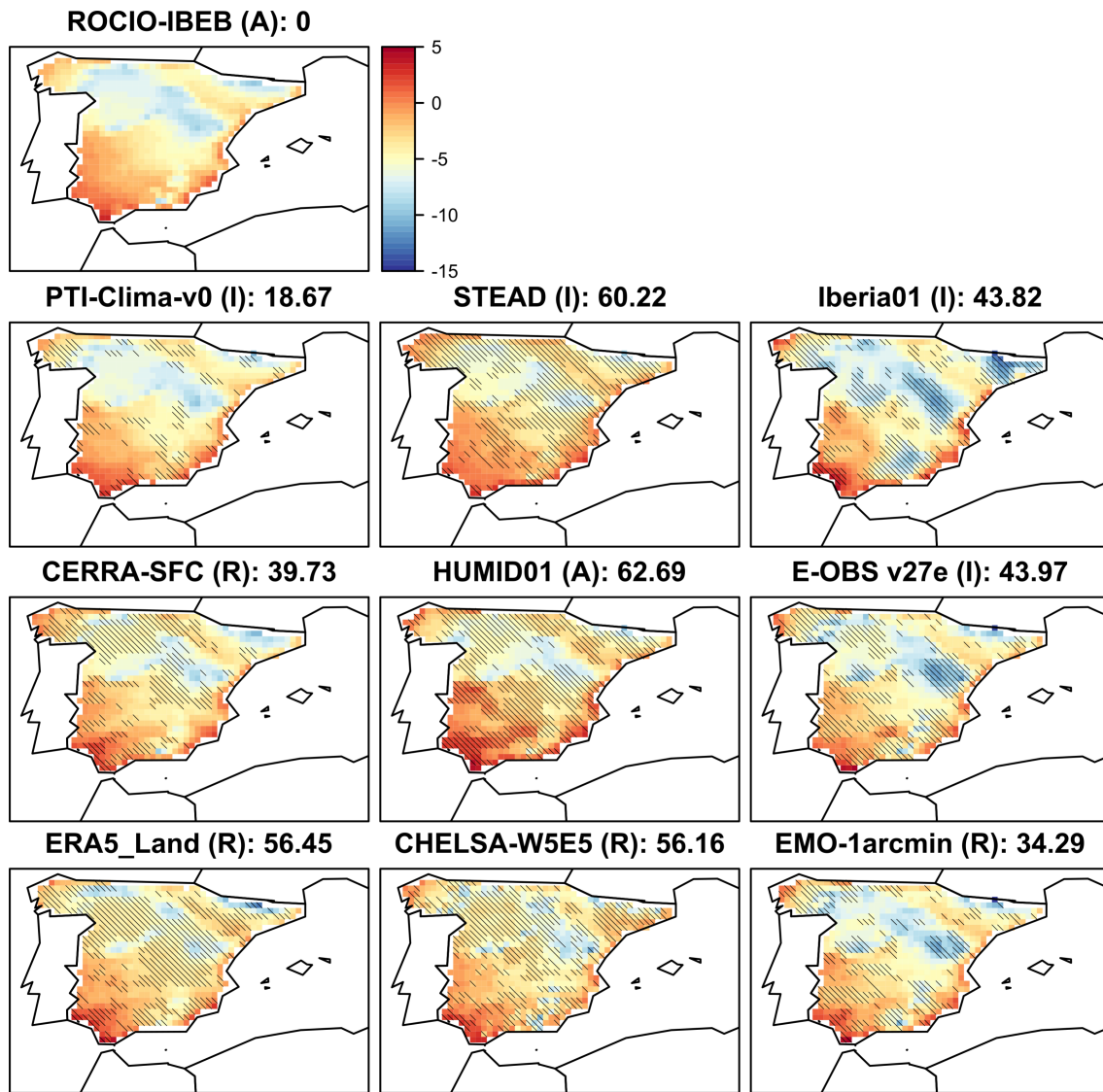
### 4.3 | Temporal Structure (TS) and Spells (SS)

The temporal structure (TS) is evaluated using two correlation indices. The first index is the standard Pearson cross-correlation among the gridded datasets, which considers the annual mean series of maximum and minimum temperatures for each grid point. The second index applies the Pearson correlation to the annual cycle.

We analysed three indices to evaluate the discrepancies and consistencies in 3-day spells among different datasets (Klein Tank et al. 2009): the mean number per year of summer day spells ( $nhw_{25}^3$ ), heatwave spells ( $nhw_{35}^3$ , not shown) and tropical night spells ( $ntn_{20}^3$ ). Each index is determined by specific absolute temperature thresholds (see Table 2), which define the number of days the maximum or minimum temperatures exceed these values ( $tasmax > 25^\circ\text{C}$ ,  $tasmax > 35^\circ\text{C}$  and  $tasmin > 20^\circ\text{C}$ , respectively), allowing us to better understand the variations in persistent extreme temperature events across the datasets.

### 4.4 | Spatial Pattern

Following Widmann et al. (2019), the spatial pattern was evaluated by analysing the number of independent spatial Degrees of Freedom (DoF) based on analysis of the principal components. Note that the common spatial grid in this work is composed of 730 cells, allowing the degrees of freedom (DoF) to vary between



**FIGURE 7** | 11th percentile of minimum temperature for the period 1990–2014 (°C). Grid-points with statistically significant differences (95%) in the percentile w.r.t. ROCIO-IBEB have been hatched. The percentage of the grid-points is also shown for each dataset in the title. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

1—if all the explained variance is attributed to the first principal component—and 730—if all principal components contribute equally to the spatial structure, with the latter representing the highest complexity.

In addition, the spatial patterns of the statistical distributions were compared using several parameters. Spatial bias and mean absolute error were calculated for both mean climatology and percentiles of the distribution. For the mean climatology, the altitude-conditioned bias was also calculated for each dataset. Using the elevation values provided by the E-OBS v27e dataset at its coarser resolution (0.25°),<sup>9</sup> the difference between each dataset and the reference (ROCIO-IBEB) was determined for all grid points within each elevation interval of 50 m (see Figure 14).

Finally, the similarity of the spatial patterns in mean climatology was assessed using the PACO index (Pattern CORrelation, Kotlarski et al. 2019)

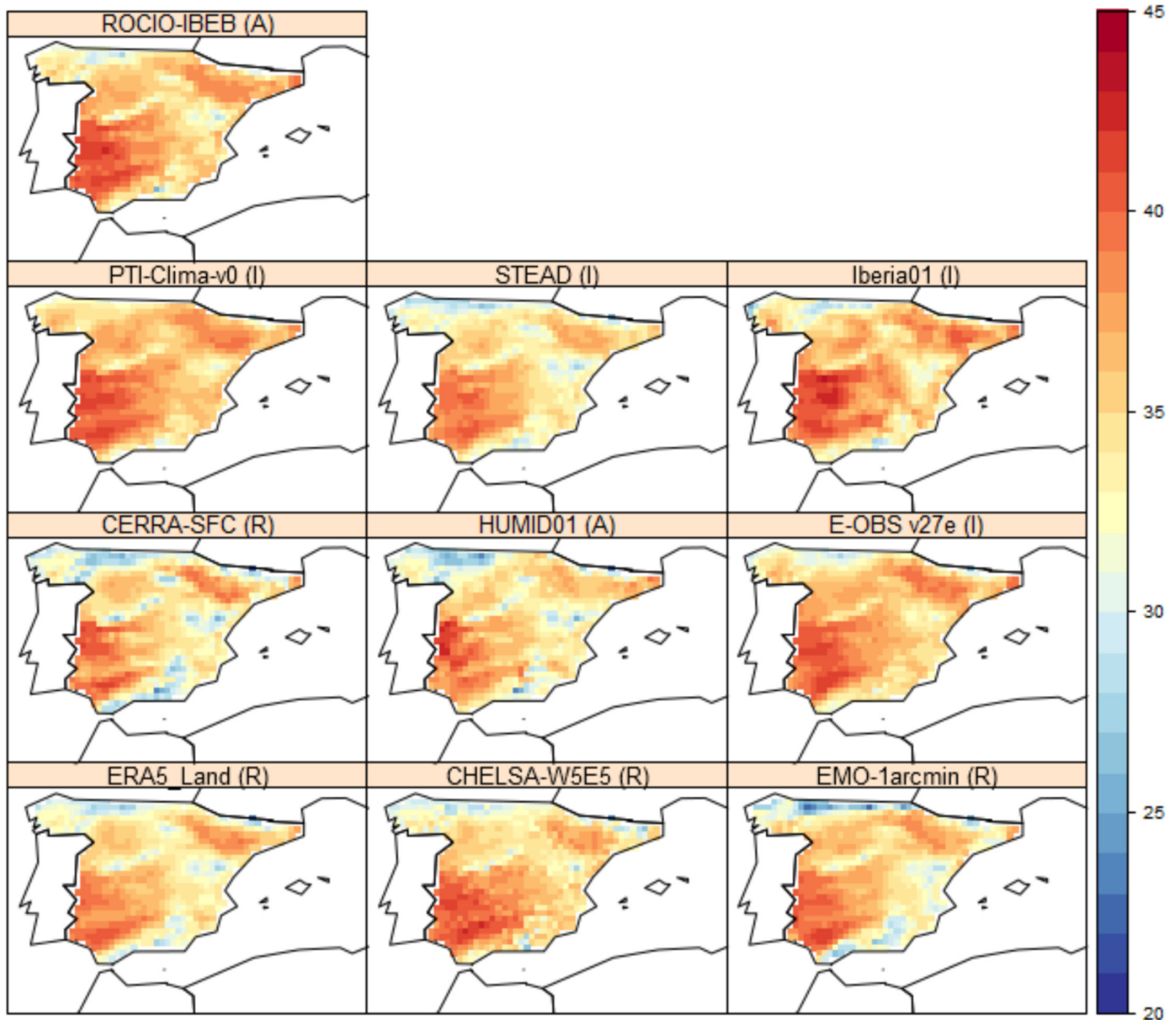
$$\text{PACO} = \frac{\text{cov}(X, O)}{\text{std}(X) * \text{std}(O)}$$

where cov and std represent the spatial covariance and the standard deviation, respectively. Here,  $X$  refers to each of the gridded datasets considered, whereas  $O$  denotes the ROCIO-IBEB dataset developed by AEMET (Peral et al. 2017), which serves as the reference for this study.

## 5 | Results

The results of the intercomparison between the datasets considered are presented below, focusing on several key aspects: statistical distribution, extreme events, temporal structure and spells and spatial pattern. Tables 3 and 4 group the results of different datasets according to their methodological origin (i.e., interpolation, analysis, reanalysis), facilitating a structured





**FIGURE 8** | Extreme event of max. temperature occurred between August 11th and August 13th 2003. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]

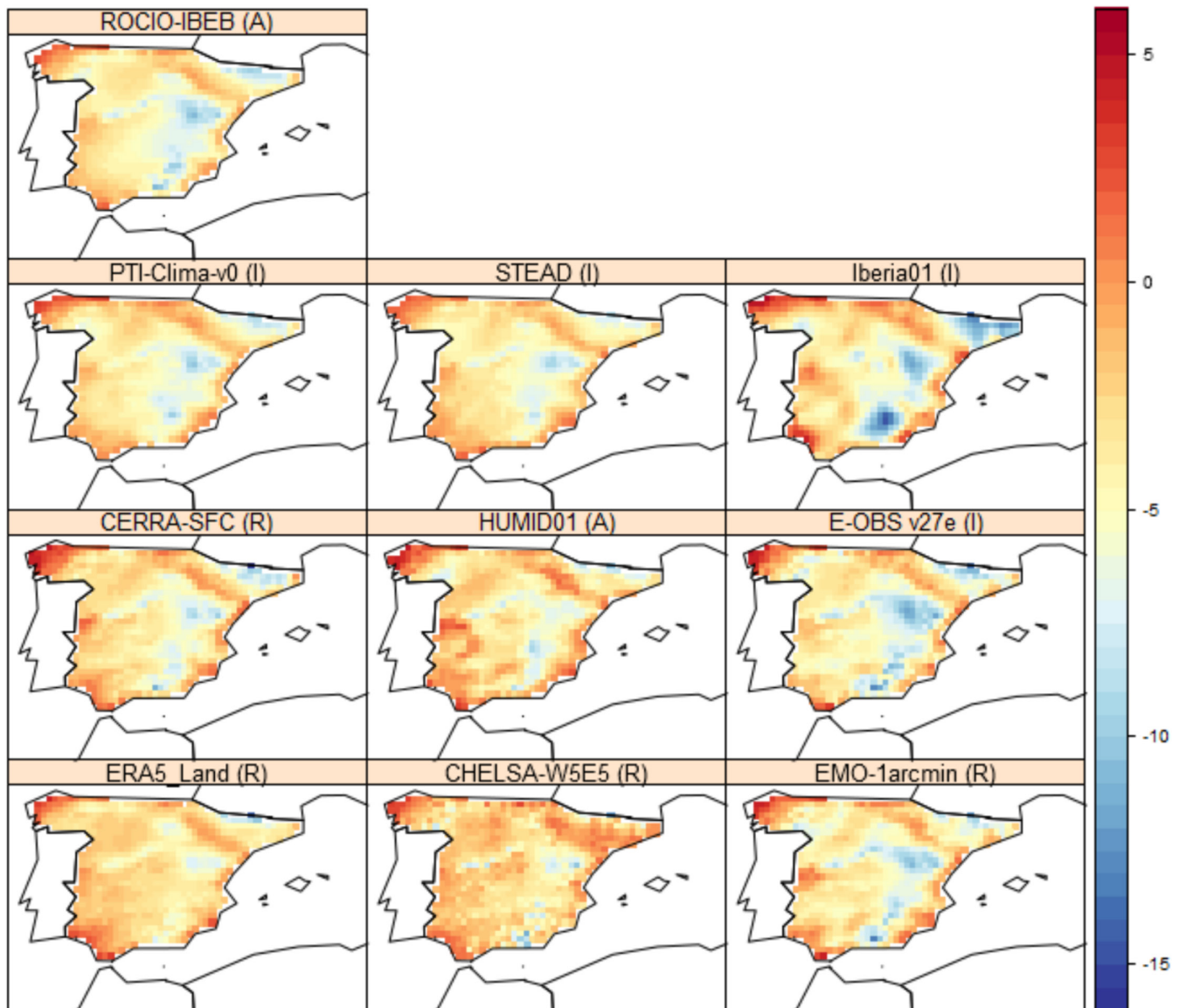
analysis of the results in light of this aspect. Note that the values corresponding to the minimum temperature (lower triangle of both tables) have been highlighted in bold.

### 5.1 | Statistical Distribution (SD)

Table 3 presents the percentages of grid cells with statistically significant (95%) differences between the means of two distributions for all combinations of the 10 datasets, for maximum and minimum temperatures. Among the datasets developed using the same methodology, the greatest differences are found between those created using the analysis approach (HUMID01 and ROCIO-IBEB), with a percentage of 99.86 for both variables. Within the interpolation methods, except for the comparison between STEAD and PTI-Clima v0, the variations in minimum temperature are more pronounced than those in maximum

temperature. In contrast, the reanalysis products exhibit the opposite behaviour, with more variation in maximum temperature. This may be due to the minimum temperature being influenced by other factors that are not fully captured by the interpolation methods, such as continentality (Peña-Angulo et al. 2016) or local factors such as land use (García-Martín et al. 2021). However, in the case of reanalysis datasets, differences in resolution result in varying orographic influences that primarily affect the maximum temperature (Dunn et al. 2022). As an example, ERA5-Land and CERRA-SFC, which is driven by ERA5, present the lowest value among the reanalysis products for maximum temperature (67.25%), whereas EMO-1arcmin, based on ERA-Interim and CHELSA-W5E5, which uses a higher resolution orography, present higher differences for maximum temperature (82.04%). The differences in minimum temperature among the reanalysis products range from 78.82% to 86.38%, showing greater homogeneity than those for maximum temperature.



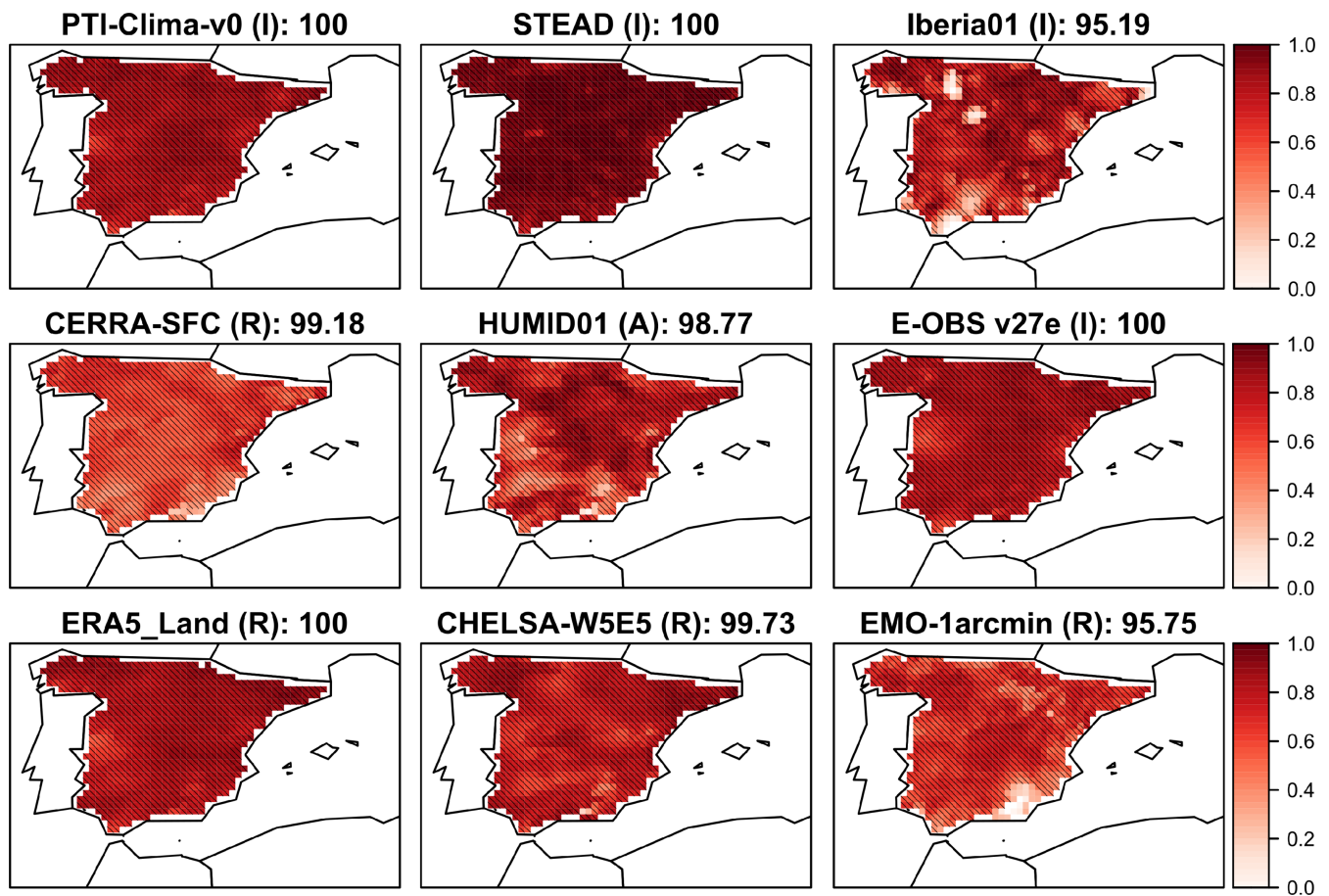


**FIGURE 9** | Extreme event of min. temperature occurred between January 26th and February 2th 2005. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]

As shown in the maps of Figure 2, two groups are distinguished. On the one hand, interpolated datasets and CHELSA-W5E5 overestimate the maximum temperature of ROCIO-IBEB in most of the Iberian Peninsula. On the other hand, HUMID01, ERA5-Land and CERRA-SFC underestimate the maximum temperature in the region, with the EMO-1arcmin dataset being more similar to the interpolation datasets than to the analysis and reanalysis ones, due to the methodology used to build this dataset. Note that HUMID01 (an analysis dataset) considers the ERA-Interim reanalysis (Dee et al. 2011) as the first guess during the optimal interpolation process, leading to the shown underestimation. The most noticeable differences occur along the northern coast, where PTI-Clima v0 displays a uniform bias (i.e., similar differences from ROCIO-IBEB at all grid points), whereas the other datasets show more spatially variable cold biases in this region, as well as in the south and southeast of the Iberian Peninsula (Figure 2).

For minimum temperature (Figure 3), the differences between the datasets are most pronounced, particularly in the northern half of Spain, where the plateau and continentality have a strong influence, and in the southwestern Iberian Peninsula, where certain datasets, such as Iberia01, report significantly higher values than the others. HUMID01 appears to be the most divergent dataset with almost all the grid-points presenting significant differences (99.85% in Table 3), whereas CHELSA-W5E5 exhibits the noisiest pattern, with neighbouring regions showing markedly different behaviours—for example, over Sierra Nevada in the southeast of the Iberian Peninsula.

The discrepancies observed at both maximum and minimum temperatures for the mean extend to the median of the empirical distribution (Figures S1 and S2). The complexity of the minimum temperature compared to the maximum temperature becomes



**FIGURE 10** | Pearson correlation between ROCIO-IBEB and each dataset for the annual mean of maximum daily temperature for the period 1990–2014. Grid-points with significant correlation (95%) have been hatched. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/joc.70111)]

evident when examining the structural validation of the spatial pattern (Table 5). Here, minimum temperature consistently exhibits more Degrees of Freedom (DoF) than maximum temperature across most datasets, except for ROCIO-IBEB, where the DoFs are comparable.

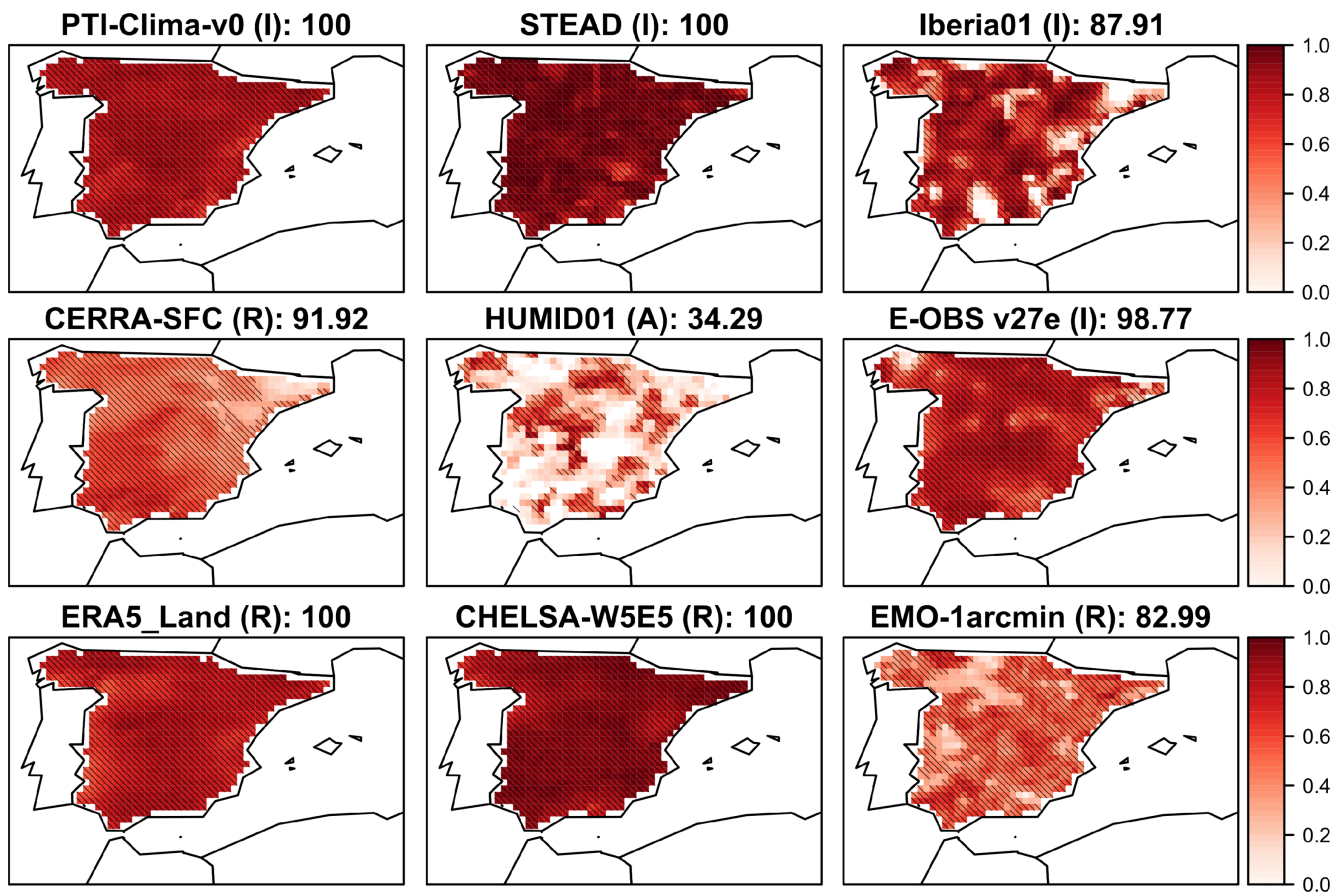
Although the differences in maximum temperature diminish considerably when analysing the centred distribution (Table 4), the differences in minimum temperature persist, particularly for the interpolated data and the ERA5-Land reanalysis. The standard deviation maps (Figures 4 and 5) highlight these differences, especially in the Iberia01 dataset for maximum temperature over the Strait of Gibraltar, in the southern Iberian Peninsula. In contrast, PTI-Clima v0 shows the greatest similarity to the reference dataset, ROCIO-IBEB, for both variables, reflecting that most of the differences between them are related to the mean of the distribution. The remaining datasets also show significant differences in the shape of the distribution (standard deviation, kurtosis, skewness, etc.). For minimum temperatures, the variability differences are higher than for maximum temperatures, but more spatially homogeneous, covering most of the domain, as shown in the corresponding standard deviation maps (Figure 5).

The differences observed in the climatologies (Figures 2 and 3) are also reflected in the percentiles (Figures 6 and 7). However, in the case of maximum temperature, the dominance of

orographic patterns tends to exacerbate the local discrepancies related to the limits of mountain ranges and main river basins compared to minimum temperature, which mostly intensifies the bias observed for the mean value. For minimum temperature, significant differences are observed in the Pyrenees in the northeast, the Sierra Nevada mountain range in the south and the northern half of the plateau.

## 5.2 | Extreme Events (EE)

Figure 8 shows the results of various datasets in reproducing the extreme heat event of August 2003. All datasets accurately place the highest maximum temperatures ( $t_{max} \geq 45^\circ\text{C}$ ) in southwestern Spain, with a secondary hotspot over the Ebro river basin (AEMet 2023a). The main differences lie in how the event is represented in mountainous regions and in the overall intensity of the heat event. Due to the limitations of observational networks and the methodologies used, interpolation datasets tend to extend the maximum temperatures to the Pyrenees (e.g., E-OBS v27e, Iberia01) and the northern coast (PTI-Clima v0), or lose some local effects, such as those in the Sierra Nevada (E-OBS v27). In contrast, analysis methods and reanalysis exhibit a strong orographic pattern, spreading local effects across wider areas (e.g., ERA5-Land and EMO-1arcmin over the Sierra Nevada), which also results in the loss of local anomalies.



**FIGURE 11** | Pearson correlation between ROCIO-IBEB and each dataset for the annual mean of minimum daily temperature for the period 1990–2014. Grid-points with significant correlation (95%) have been hatched. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/joc.70111)]

The January–February 2005 cold event affected 31 of the 50 provinces in Spain (AEMet 2023b), with temperatures reaching a minimum of  $-7^{\circ}\text{C}$ . However, the spatial patterns of the minimum temperatures shown in Figure 9 do not clearly reflect an extreme event of this magnitude. Only Iberia01, ROCIO-IBEB, E-OBS v27e, and to some extent EMO-1arcmin identify three main centres of cold: one in the northeastern Iberian Peninsula, one below the Ebro River basin and another in the Sierra Nevada, in the southeastern Iberian Peninsula. The other datasets replicate a similar spatial pattern but do not capture the intensity of the event. In addition, CHELSA-W5E5 and ERA5-Land show unusually high temperatures in midwestern Spain, the Ebro river basin and the Pyrenees.

In the GEV analysis, Figures S4 and S6 show the shape parameter for maximum and minimum temperatures. Note that for maximum temperature, most grid boxes exhibit a shape parameter below zero, indicating a Weibull distribution with a bounded tail. In the case of minimum temperature, there are some isolated unstable grid boxes with positive values, and only ERA5-Land shows structures with a positive shape parameter, corresponding to a Fréchet distribution. This results in a lower 50-year return value for the minimum temperature in these regions compared to the other datasets (Figure S5). For maximum temperature (Figure S3), PTI-Clima-v0 overestimates the 50-year return value in the northern Iberian Peninsula, where the other datasets show the lowest values in

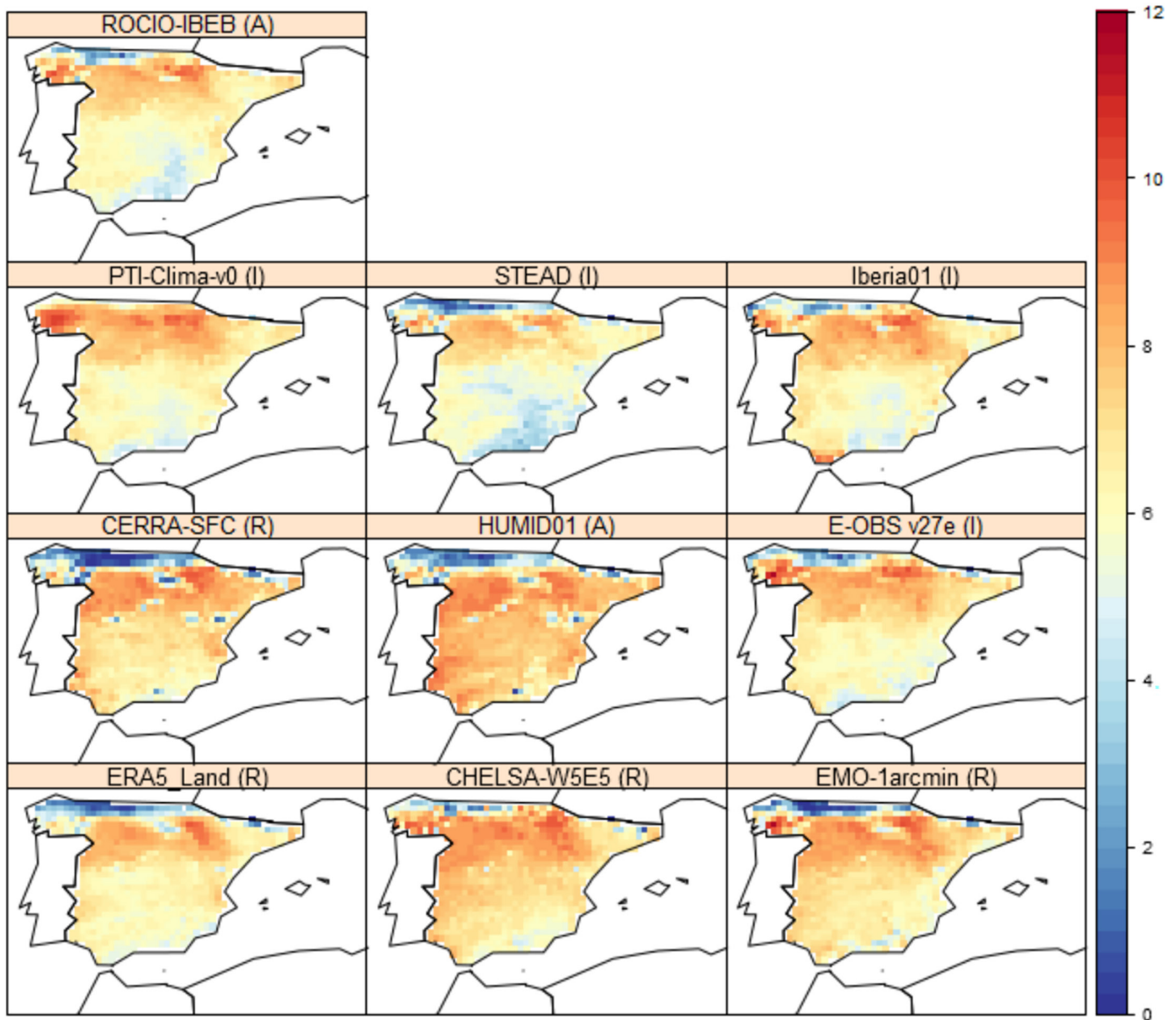
the region, with a secondary minimum over the Pyrenees in some cases.

### 5.3 | Temporal Structure (TS) and Spells (SS)

Figures 10 and 11 display the Pearson correlation between the reference, ROCIO-IBEB, and each dataset for the annual mean of maximum and minimum daily temperatures, respectively, including the statistical significance of the coefficient (hatched grid-boxes). The correlation for maximum temperature is higher, more statistically significant and more spatially homogeneous than for minimum temperature. Some datasets exhibit localised regions where the correlation is either not significant or close to zero. Specifically, for maximum temperature, this occurs in Iberia01, HUMID01 and EMO-1arcmin; for minimum temperature, it is observed in Iberia01, CERRA-SFC, HUMID01, E-OBS v27e and EMO-1arcmin, although the problem in the case of HUMID01 affects most of the Iberian Peninsula.

Although E-OBS v27e, CERRA-SFC and CHELSA-W5E5 also show low correlations in some locations, these are limited to relatively small areas. These problems are even more pronounced for the minimum temperature, especially in the HUMID01 dataset, with variations observed in the affected regions. Additionally, the CERRA-SFC dataset shows lower correlations in the Pyrenees. However, the pattern correlation (Table S1) and





**FIGURE 12** | Mean number of summer days ( $tasmax \geq 25^\circ\text{C}$ ) spells (3 days) for the period 1990–2014. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/joc.70111)]

the seasonal cycle correlation (not shown) are nearly 1 for all variables and datasets, indicating that the inconsistencies between datasets stem from the interannual variability.

The number of summer day spells (Figure 12) reaches a minimum along the northern coast—except in the case of PTI-Clima v0—followed by a latitudinal band exhibiting the maximum values across the Iberian Peninsula and then a downward gradient of decreasing values. The main differences between datasets concern the intensity of the index and the spatial extent and distribution of a secondary minimum, which appears in the southern and/or southeastern regions in some datasets. HUMID01 displays the minimum only in the north and presents a relatively homogeneous pattern across the rest of the territory.

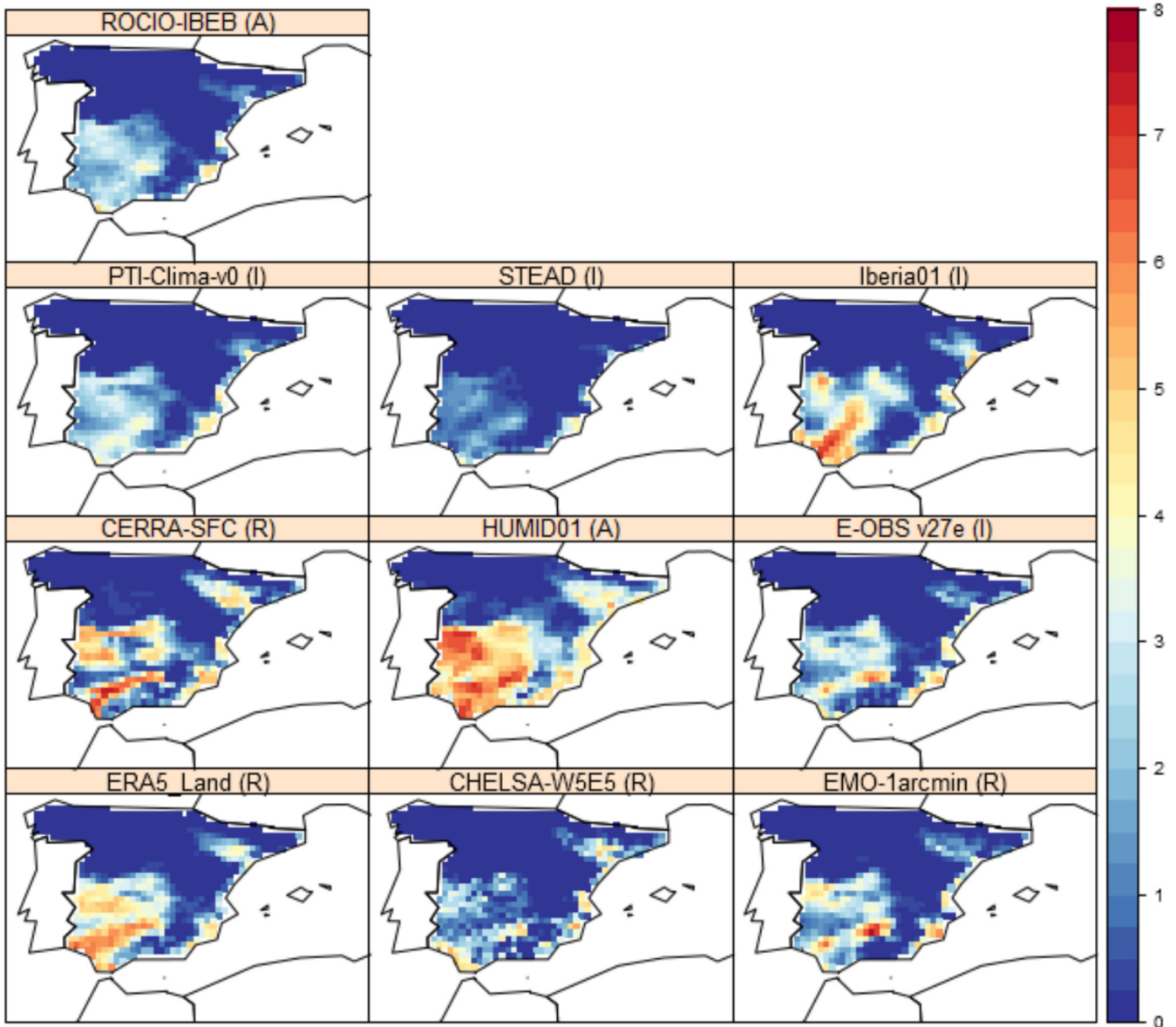
Unlike summer day spells, the highest number of tropical night spells (Figure 13), with values ranging approximately between 5

and 10 spells per year, is concentrated in the southwest, with a secondary hotspot in the Ebro River basin, consistent with other indices discussed earlier. Although the overall pattern is similar across datasets, there are differences in the intensity and extent of the maximum in the southwest. In addition, areas with higher values for tropical night spells are also observed along the eastern coast.

#### 5.4 | Spatial Pattern

From the analysis done in the previous sections and the results in the first two columns of Table 5, we can conclude that the minimum temperature is more complex than the maximum temperature, regardless of the dataset or the approach used. This difference may be attributed to the interaction of various factors, such as continentality (Peña-Angulo et al. 2016) or local factors (García-Martín et al. 2021), that affect the minimum





**FIGURE 13** | Mean number of tropical nights ( $tasmin \geq 20^{\circ}\text{C}$ ) spells (3 days) for the period 1990–2014. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]

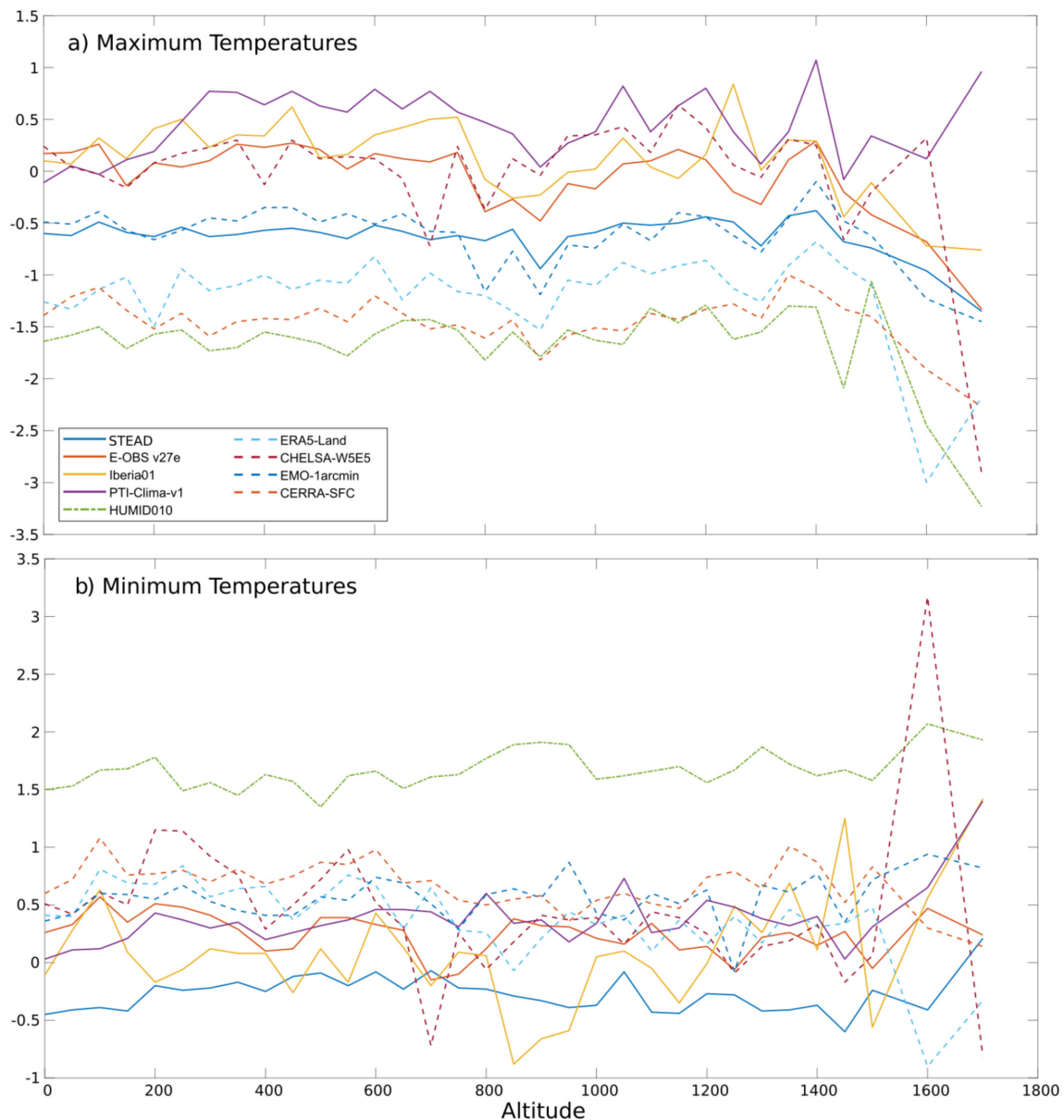
temperature, in contrast to the strong dependence of the maximum temperature on orography, which is often the main factor in interpolation methods. Only ROCIO-IBEB and PTI-Clima v0 produce similar results for both variables.

The mean absolute error of the spatial pattern for both the mean and percentiles of maximum and minimum temperatures (see columns 3–8 in Table 5) shows that the median values among datasets are similar (0.61 for maximum temperature and 0.71 for minimum temperature). However, the data varies more for maximum temperature ( $Tx_{sd} = 0.42$ ) than for minimum temperature ( $Tn_{sd} = 0.36$ ). This difference in variability is especially noticeable at the extremes ( $Tx_{95_{sd}} = 0.45$ ,  $Tx_{99_{sd}} = 0.36$ ,  $Tn_{05_{sd}} = 0.22$ ,  $Tn_{01_{sd}} = 0.55$ ).

Analysing bias as a function of altitude (Figure 14), we find that, despite some fluctuations, the differences remain relatively stable

up to the highest elevations, where they increase in one direction or another. Observational datasets, except for STEAD, tend to overestimate both temperatures, more noticeably for the maximum temperature. However, STEAD consistently underestimates temperatures regardless of height. Reanalysis datasets show greater variability in maximum temperature than in minimum temperature, generally underestimating maximum temperatures and overestimating minimum temperatures. The differences observed between the analysis datasets extend across all elevation levels, but are more stable for minimum temperature than maximum temperature, where these differences become more pronounced at higher grid-box elevations.

Consistent with previous findings, the PACO index (Table S1) reveals more pronounced differences among datasets and approaches for minimum temperatures compared to maximum temperatures, with the role of observations in the methodology



**FIGURE 14** | Annual mean bias of (a) maximum and (b) minimum temperatures depending on the height of the grid cell. All biases have been obtained using ROCIO-IBEB as a reference. D1: STEAD, D2: E-OBS v27e, D3: Iberia01, D4: PTI-Clima v0, D5: HUMID01, D7: ERA5-Land, D8: CHELSA-W5E5, D9: EMO-1arcmin, D10: CERRA-SFC. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]

increasing these differences. The analysis and the interpolated products are more similar to each other than to the reanalysis products for both variables. However, for maximum temperature, all datasets are more similar to each other than they are for minimum temperature.

## 6 | Conclusions and Discussion

This study responds to the need for an intercomparison analysis of existing datasets to evaluate their advantages, limitations, potential applications and risks of misuse. It is the first of its kind to comprehensively analyse the main datasets available for mainland Spain, considering critical factors such as statistical distribution, as well as temporal and spatial

structures, whereas highlighting key differences among them. In doing so, it fills a gap in the current literature and offers valuable insights for the scientific community, supporting researchers in identifying the most suitable datasets for the specific objectives of their studies.

However, in some cases, greater effort is required to clearly explain how these datasets were developed and to effectively communicate the uncertainties associated with them. A major limitation of this study is the absence of error and/or uncertainty measures in certain datasets, as well as the varying approaches used to define and manage these uncertainties. These factors, along with the homogenisation processes applied to the datasets, constrain the study and limit the interpretability of the results.

The findings reveal that differences in minimum temperature are more pronounced than those in maximum temperature, particularly in datasets based on interpolation. Although discrepancies in maximum temperature primarily affect the mean, variations in minimum temperature extend to higher moments, influencing the overall shape of the statistical distribution. These differences are not confined to a single dataset or methodological approach but are observed across all datasets. Furthermore, most variations occur at spatial scales larger than the resolution of the datasets themselves, raising questions about the actual benefits of higher resolution. These discrepancies become even more significant when analysing extreme events or the upper and lower tails of the distribution.

Regarding extreme events, specific warm and cold episodes were analysed, along with a distribution-based approach, which highlighted the challenges that datasets face in accurately reproducing extremes. As previously noted, differences between datasets are more pronounced for minimum temperature than for maximum temperature. However, some datasets—such as PTI-Clima v0, ROCIO-IBEB and EMO-1arcmin—show similar spatial patterns and intensities for extremes, whereas others tend to underestimate or overestimate the 50-year return values and/or display significant discrepancies in the shape parameter of the GEV distribution. Particularly notable are the issues observed in ERA5-Land and CHLSA-W5E5 with minimum temperature extremes, where specific events are underestimated and the GEV distribution is poorly fitted. As a result, although national datasets might be expected to outperform European or global ones, particularly for local events and extremes, this study demonstrates that performance varies significantly depending on the type of analysis, making it difficult to draw a definitive conclusion.

In terms of the temporal correlation of annual mean time series, datasets such as Iberia01, HUMID01, EMO-1arcmin and CERRA-SFC exhibit problematic regions for at least one variable. Notable differences were also observed in indices based on the duration of heat waves and cold spells for both maximum and minimum temperatures. Although differences are relatively minor for the  $n_{hw}^3$  index, mainly affecting intensity, they become more pronounced for the  $n_{hw}^3$  and  $ntn^3$  indices, impacting both intensity and spatial patterns.

Based on indices analysing spatial patterns, minimum temperature appears to be more complex than maximum temperature, although maximum temperature shows greater variability in the mean absolute error of the spatial pattern. This complexity results in larger differences between datasets and methods for minimum temperatures compared to maximum temperatures. Consequently, this study highlights the need to consider additional factors—such as continentality and land use—during the interpolation process to better capture the complexity of minimum temperature, whereas these factors do not appear to add significant value for maximum temperature.

Overall, the most stable dataset across all evaluated indices is STEAD, whereas PTI-Clima v0 exhibits some issues with underestimating extremes and spells but performs well in capturing central parameters and temporal correlation.

## Author Contributions

**Sixto Herrera:** conceptualization, software, data curation, writing – original draft, formal analysis, methodology, investigation, writing – review and editing. **Fidel González Rouco:** writing – review and editing, conceptualization, writing – original draft, methodology, formal analysis. **Roberto Serrano-Notivol:** conceptualization, methodology, writing – review and editing, writing – original draft, data curation, formal analysis. **Juan Luis Garrido:** conceptualization, methodology, writing – review and editing, writing – original draft, formal analysis. **Santiago Begueria:** conceptualization, methodology, data curation, writing – review and editing, formal analysis. **José M. Gutiérrez:** conceptualization, methodology, writing – review and editing, formal analysis. **Pere Quintana-Seguí:** conceptualization, methodology, data curation, writing – review and editing. **Maialen Iturbide:** conceptualization, methodology, writing – review and editing. **Esteban Rodríguez:** conceptualization, methodology, writing – review and editing, data curation. **Ana Morata:** conceptualization, methodology, data curation, writing – review and editing. **Candelas Peral:** conceptualization, methodology, data curation, writing – review and editing.

## Acknowledgements

This research work was funded by the Spanish Ministry for Ecological Transition and Demographic Challenge (MITECO) and the European Commission NextGenerationEU (Regulation EU 2020/2094), through CSIC's Interdisciplinary Thematic Platform Clima (PTI-Clima).

## Conflicts of Interest

The authors declare no conflicts of interest.

## Data Availability Statement

In this study, no new data have been generated. PTI-Clima v0 and HUMID01 are currently not disclosed to the public; STEAD and Iberia01 are publicly available through the Digital.CSIC repository. Iberia01 is made available under the Open Database Licence; the ROCIO-IBEB dataset is freely distributed for use in research work, and commercial use is not permitted (for information on the commercial use of this grid you must contact AEMET). The work and products carried out, in whole or in part, with this data must also be public and must be made available to the scientific community free of charge. Likewise, to guarantee access to the updated original source, redistribution of data to third parties is not permitted; the EMO-1arcmin dataset is provided under the licence Creative Commons Attribution 4.0 International; CHLSA-W5E5 is publicly available through the ISIMIP Repository under the CC0 1.0 Universal Public Domain Dedication licence; E-OBS v27e, ECMWF-ERA5-Land and CERRA-SFC datasets are publicly available from the Copernicus Climate Data Store under the terms of use defined for the Copernicus Products.

## Endnotes

<sup>1</sup> <https://doi.org/10.20350/digitalCSIC/8622>.

<sup>2</sup> <http://hdl.handle.net/10261/183071>.

<sup>3</sup> <https://www.obsebre.es/en/en-safran>.

<sup>4</sup> [https://www.aemet.es/es/serviciosclimaticos/cambio\\_climat/datos\\_diarios?w=2](https://www.aemet.es/es/serviciosclimaticos/cambio_climat/datos_diarios?w=2).

<sup>5</sup> [https://jeodpp.jrc.ec.europa.eu/ftp/jrc-opendata/CEMS-EFAS/meteorological\\_forcings/EMO-1arcmin/](https://jeodpp.jrc.ec.europa.eu/ftp/jrc-opendata/CEMS-EFAS/meteorological_forcings/EMO-1arcmin/).

<sup>6</sup> <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land>.

<sup>7</sup> <https://chelsa-climate.org/>.



<sup>8</sup><https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-cerra-single-levels>.

<sup>9</sup>[https://knmi-ecad-assets-prd.s3.amazonaws.com/ensembles/data/Grid\\_0.25deg\\_reg\\_ensemble/elev\\_ens\\_0.25deg\\_reg\\_v27.0e.nc](https://knmi-ecad-assets-prd.s3.amazonaws.com/ensembles/data/Grid_0.25deg_reg_ensemble/elev_ens_0.25deg_reg_v27.0e.nc).

## References

- AEMet. 2023a. "Olas de Frío en España Desde 1975." Technical Report-, AEMet- Servicio de Banco Nacional de Datos Climatológicos. [https://www.aemet.es/documentos/es/conocermas/recursos\\_en\\_linea/publicaciones\\_y\\_estudios/estudios/Olas\\_frio/olas\\_frio\\_actualizacionoctubre2023.pdf](https://www.aemet.es/documentos/es/conocermas/recursos_en_linea/publicaciones_y_estudios/estudios/Olas_frio/olas_frio_actualizacionoctubre2023.pdf).
- AEMet. 2023b. "Olas de Calor en España Desde 1975." Technical Report-, AEMet- Área de Climatología y Aplicaciones Operativas. [https://www.aemet.es/documentos/es/conocermas/recursos\\_en\\_linea/publicaciones\\_y\\_estudios/estudios/Olas\\_calor/olas\\_calor\\_actualizacionoctubre2023.pdf](https://www.aemet.es/documentos/es/conocermas/recursos_en_linea/publicaciones_y_estudios/estudios/Olas_calor/olas_calor_actualizacionoctubre2023.pdf).
- Alexandridis, T. K., G. Ovakoglou, I. Cherif, et al. 2021. "Designing Agricultures Services to Support Food Security in Africa." *Transactions in GIS* 25, no. 2: 692–720. <https://doi.org/10.1111/tgis.12684>.
- Antolović, I., V. Mihajlović, D. Rančić, D. Mihić, and V. Djurdjević. 2013. "Digital Climate Atlas of the Carpathian Region." *Advances in Science and Research* 10: 107–111. <https://doi.org/10.5194/asr-10-107-2013>.
- Bedia, J., S. Herrera, and J. M. Gutiérrez. 2013. "Dangers of Using Global Bioclimatic Datasets for Ecological Niche Modeling. Limitations for Future Climate Projections." *Global and Planetary Change* 107: 1–12. <https://doi.org/10.1016/j.gloplacha.2013.04.005>. <https://www.sciencedirect.com/science/article/pii/S0921818113000957>.
- Beguiria, S. 2025. "Undefined. Undefined, in Preparation."
- Bengtsson, L., U. Andrae, T. Aspelien, et al. 2017. "The Harmonie-Arome Model Configuration in the Aladin-Hirlam Nwp System." *Monthly Weather Review* 145, no. 5: 1919–1935. <https://doi.org/10.1175/MWR-D-16-0417.1>.
- Bojinski, S., M. Verstraete, T. C. Peterson, C. Richter, A. Simmons, and M. Zemp. 2014. "The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy." *Bulletin of the American Meteorological Society* 95, no. 9: 1431–1443. <https://doi.org/10.1175/BAMS-D-13-00047.1>.
- Brun, E., E. Martin, V. Simon, C. Gendreau, and C. Coleou. 1989. "An Energy and Mass Model of Snow Cover Suitable for Operational Avalanche Forecasting." *Journal of Glaciology* 35: 333–342.
- Burton, C., S. Rifai, and Y. Malhi. 2018. "Inter-Comparison and Assessment of Gridded Climate Products Over Tropical Forests During the 2015/2016 el Niño." *Philosophical Transactions of the Royal Society, B: Biological Sciences* 373, no. 1760: 20170406. <https://doi.org/10.1098/rstb.2017.0406>.
- Coles, S. 2001. "An Introduction to Statistical Modeling of Extreme Values." *Journal of the American Statistical Association* 97: 1204. <https://api.semanticscholar.org/CorpusID:19678794>.
- Cornes, R., G. van der Schrier, E. J. M. van den Besselaar, and P. Jones. 2018. "An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets." *Journal of Geophysical Research: Atmospheres* 123: 9391–9409. <https://doi.org/10.1029/2017JD028200>.
- Cornes, R., G. van der Schrier, E. J. M. van den Besselaar, and P. Jones. 2020. "E-OBS Daily Gridded Meteorological Data for Europe From 1950 to Present Derived From In-Situ Observations." Technical Report, Copernicus Climate Change Service, Climate Data Store.
- Cucchi, M., G. P. Weedon, A. Amici, et al. 2020. "Wfde5: Bias-Adjusted ERA5 Reanalysis Data for Impact Studies." *Earth System Science Data* 12, no. 3: 2097–2120. <https://doi.org/10.5194/essd-12-2097-2020>.
- Dee, D. P., S. M. Uppala, A. J. Simmons, et al. 2011. "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>.
- Dunn, R. J. H., M. G. Donat, and L. V. Alexander. 2022. "Comparing Extremes Indices in Recent Observational and Reanalysis Products." *Frontiers in Climate* 4. <https://doi.org/10.3389/fclim.2022.989505>.
- El-Said, A., P. Brousseau, M. Ridal, and R. Randriamampianina. 2021. "A New Temporally Flow-Dependent Eda Estimating Background Errors in the New Copernicus European Regional Re-Analysis (Cerra)." *Earth and Space Science Open Archive*: 1–28. <https://doi.org/10.1002/essoar.10507207.1>.
- Gandin, L. S. 1966. "Objective Analysis of Meteorological Fields." *Quarterly Journal of the Royal Meteorological Society* 92: 447. <https://doi.org/10.1002/qj.49709239320>.
- García-Martín, A., L. L. Paniagua, F. J. Moral, F. J. Rebollo, and M. A. Rozas. 2021. "Spatiotemporal Analysis of the Frost Regime in the Iberian Peninsula in the Context of Climate Change (1975–2018)." *Sustainability* 13: 8491. <https://doi.org/10.3390/su13158491>. <https://www.mdpi.com/2071-1050/13/15/8491>.
- Gomes, G., V. Thiemeig, J. O. Skoien, et al. 2020. "Emo: A High-Resolution Multi-Variable Gridded Meteorological Data Set for Europe." European Commission, Joint Research Centre (JRC). <https://doi.org/10.2905/0BD84BE4-CEC8-4180-97A6-8B3ADAAC4D26>.
- Gumbel, E. J. 1935. "Les Valeurs Extrêmes Des Distributions Statistiques." *Annales de l'Institut Henri Poincaré* 5, no. 2: 115–158. [http://www.numdam.org/item/AIHP\\_1935\\_\\_5\\_2\\_115\\_0/](http://www.numdam.org/item/AIHP_1935__5_2_115_0/).
- Haiden, T., A. Kann, C. Wittmann, G. Pistotnik, B. Bica, and C. Gruber. 2011. "The Integrated Nowcasting Through Comprehensive Analysis (Inca) System and Its Validation Over the Eastern Alpine Region." *Weather and Forecasting* 26: 166–183. <https://doi.org/10.1175/2010WAF2222451.1>.
- Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New. 2008. "A European Daily High-Resolution Gridded Data Set of Surface Temperature and Precipitation for 1950–2006." *Journal of Geophysical Research* 113: D20119. <https://doi.org/10.1029/2008JD010201>.
- 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, Universidad de Cantabria.
- Herrera, S., R. M. Cardoso, P. M. M. Soares, F. Espírio-Santo, P. Viterbo, and J. M. Gutiérrez. 2019a. "Iberia01: A New Gridded Dataset of Daily Precipitation and Temperatures Over Iberia." *Earth System Science Data* 11: 1947–1956. <https://doi.org/10.5194/essd-11-1947-2019>.
- Herrera, S., R. M. Cardoso, P. M. M. Soares, F. Espírio-Santo, P. Viterbo, and J. M. Gutiérrez. 2019b. "Iberia01: Daily Gridded (0.1° Resolution) Dataset of Precipitation and Temperatures Over the Iberian Peninsula." Digital.CSIC. <https://doi.org/10.20350/digitalCSIC/8641>.
- Herrera, S., J. M. Gutiérrez, R. Ancell, M. R. Pons, M. Frías, and J. Fernandez. 2012. "Development and Analysis of a 50-Year High-Resolution Daily Gridded Precipitation Dataset Over Spain (Spain02)." *International Journal of Climatology* 36: 74–85. <https://doi.org/10.1002/joc.2256>.
- Herrera, S., S. Kotlarski, P. M. M. Soares, et al. 2019. "Uncertainty in Gridded Precipitation Products: Influence of Station Density, Interpolation Method and Grid Resolution." *International Journal of Climatology* 39, no. 9: 3717–3729. <https://doi.org/10.1002/joc.5878>.
- Herrera, S., P. M. M. Soares, R. M. Cardoso, and J. M. Gutiérrez. 2020. "Evaluation of the Euro-Cordex Regional Climate Models Over the Iberian Peninsula: Observational Uncertainty Analysis." *Journal of Geophysical Research: Atmospheres* 125, no. 12: e2020JD032880. <https://doi.org/10.1029/2020JD032880>.



- Hersbach, H., B. Bell, P. Berrisford, et al. 2020. "The ERA5 Global Reanalysis." *Quarterly Journal of the Royal Meteorological Society* 146: 1999–2049. <https://doi.org/10.1002/qj.3803>.
- Hersbach, H., B. Bell, P. Berrisford, et al. 2023. "ERA5 Hourly Data on Single Levels From 1940 to Present. Technical Report." Copernicus Climate Change Service (C3S) Climate Data Store (CDS).
- Hilbe, J. M. 2011. *Generalized Linear Models*, edited by M. Lovric, 591–596. Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-04898-2\\_273](https://doi.org/10.1007/978-3-642-04898-2_273).
- Isotta, F. A., C. Frei, V. Weigluni, et al. 2014. "The Climate of Daily Precipitation in the Alps: Development and Analysis of a High-Resolution Grid Data Set From Pan-Alpine Rain-Gauge Data." *International Journal of Climatology* 34: 1657–1675. <https://doi.org/10.1002/joc.3794>.
- Iturbide, M., J. Fernández, J. M. Gutiérrez, et al. 2022. "Implementation of Fair Principles in the Ippc: The Wgi ar6 Atlas Repository." *Scientific Data* 9, no. 1: 629. <https://doi.org/10.1038/s41597-022-01739-y>.
- Karger, D. N., S. Lange, C. Hari, et al. 2023. "Chelsa-w5e5: Daily 1 Km Meteorological Forcing Data for Climate Impact Studies." *Earth System Science Data* 15: 2445–2464. <https://doi.org/10.5194/essd-15-2445-2023>.
- Karger, D. N., S. Lange, C. Hari, C. P. O. Reyer, and N. E. Zimmermann. 2022. "Chelsa-w5e5 v1.0: W5e5 v1.0 Downscaled With Chelsa v2.0." Technical Report, ISIMIP Repository [Data Set].
- Karger, D. N., A. M. Wilson, C. Mahony, N. E. Zimmermann, and W. Jetz. 2021. "Global Daily 1 Km Land Surface Precipitation Based on Cloud Cover-Informed Downscaling." *Scientific Data* 8: 307. <https://doi.org/10.1038/s41597-021-01084-6>.
- Klein Tank, A. M. G., F. W. Zwiers, and X. Zhang. 2009. "Guidelines on Analysis of Extremes in a Changing Climate in Support of Informed Decisions for Adaptation, Climate Data and Monitoring." Technical Report WCDMP-No72. (World Meteorological Organization). <https://library.wmo.int/viewer/48826/#page=1&viewer=picture&o=bookmarks&n=0&q=>.
- Kotlarski, S., P. Szabó, S. Herrera, et al. 2019. "Observational Uncertainty and Regional Climate Model Evaluation: A Pan-European Perspective." *International Journal of Climatology* 39, no. 9: 3730–3749. <https://doi.org/10.1002/joc.5249>.
- Kunsch, H. R. 1989. "The Jackknife and the Bootstrap for General Stationary Observations." *Annals of Statistics* 17, no. 3: 1217–1241. <https://doi.org/10.1214/aos/1176347265>.
- Lange, S. 2019. "Wfde5 Over Land Merged With ERA5 Over the Ocean (w5e5). v. 1.0." Technical Report, GFZ Data Services.
- Lledó, L., T. Haiden, and M. Chevallier. 2024. "An Intercomparison of Four Gridded Precipitation Products Over Europe Using the Three-Cornered-Hat Method." *EGUspere*. <https://doi.org/10.5194/egusphere-2024-807>.
- McCulloch, C. E., and J. M. Neuhaus. 2014. *Generalized Linear Mixed Models*. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118445112.stat07540>.
- Mínguez, R., and S. Herrera. 2023. "Spatial Extreme Model for Rainfall Depth: Application to the Estimation of Idf Curves in the Basque Country." *Stochastic Environmental Research and Risk Assessment* 37, no. 8: 3117–3148. <https://doi.org/10.1007/s00477-023-02440-1>.
- Muñoz Sabater, J. 2019. "ERA5-Land Hourly Data From 1950 to Present." Technical Report, Copernicus Climate Change Service, Climate Data Store.
- Newman, A. J., M. P. Clark, R. J. Longman, and T. W. Giambelluca. 2019. "Methodological Intercomparisons of Station-Based Gridded Meteorological Products: Utility, Limitations, and Paths Forward." *Journal of Hydrometeorology* 20: 531–547. <https://doi.org/10.1175/JHM-D-18-0114.1>.
- Osborn, T. J., and M. Hulme. 1997. "Development of a Relationship Between Station and Grid-Box Rainday Frequencies for Climate Model Evaluation." *Journal of Climate* 10, no. 8: 1885–1908. [https://doi.org/10.1175/1520-0442\(1997\)010<1885:DOARBS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<1885:DOARBS>2.0.CO;2).
- Padgett, W. J. 2011. *Weibull DSSSSistribution*, edited by M. Lovric, 1651–1653. Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-04898-2\\_611](https://doi.org/10.1007/978-3-642-04898-2_611).
- Peña-Angulo, D., R. M. Trigo, N. Cortesi, and J. C. González-Hidalgo. 2016. "The Influence of Weather Types on the Monthly Average Maximum and Minimum Temperatures in the Iberian Peninsula." *Atmospheric Research* 178: 217–230. <https://doi.org/10.1016/j.atmosres.2016.03.022>. <https://www.sciencedirect.com/science/article/pii/S0169809516300722>.
- Peral, C., B. Navascués, and P. Ramos. 2017. "Serie de Precipitación Diaria en Rejilla Con Fines Climáticos." Technical report, AEMET. [https://www.aemet.es/documentos/es/conocermas/recursos\\_en\\_linea/publicaciones\\_y\\_estudios/publicaciones/NT\\_24\\_AEMET/NT\\_24\\_AEMET.pdf](https://www.aemet.es/documentos/es/conocermas/recursos_en_linea/publicaciones_y_estudios/publicaciones/NT_24_AEMET/NT_24_AEMET.pdf).
- Quintana-Seguí, P., P. Le Moigne, Y. Durand, et al. 2008. "Analysis of Near-Surface Atmospheric Variables: Validation of the Safran Analysis Over France." *Journal of Applied Meteorology and Climatology* 47: 92–107. <https://doi.org/10.1175/2007JAMC1636.1>.
- Quintana-Seguí, P., M. Turco, S. Herrera, and G. Miguez-Macho. 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, no. 4: 2187–2201. <https://doi.org/10.5194/hess-21-2187-2017>.
- Ramos, P. L., F. Louzada, E. Ramos, and S. Dey. 2018. "The Frechet Distribution: Estimation and Application an Overview." <https://arxiv.org/abs/1801.05327>.
- Rypkema, D., and S. Tuljapurkar. 2021. "Chapter 2—Modeling Extreme Climatic Events Using the Generalized Extreme Value (Gev) Distribution." In *Data Science: Theory and Applications, 44 of Handbook of Statistics*, edited by A. S. R. Srinivasa Rao and C. R. Rao, 39–71. Elsevier. <https://doi.org/10.1016/bs.host.2020.12.002>. <https://www.sciencedirect.com/science/article/pii/S0169716120300511>.
- Schimanke, S., M. Ridal, P. Le Moigne, et al. 2021. "Cerra Sub-Daily Regional Reanalysis Data for Europe on Single Levels From 1984 to Present." Technical Report, Copernicus Climate Change Service, Climate Data Store.
- Serrano-Notivoli, R., S. Beguería, and M. De Luis. 2019. "Stead: A High-Resolution Daily Gridded Temperature Dataset for Spain." *Earth System Science Data* 11: 1097–1111. <https://doi.org/10.5194/essd-11-1171-2019>.
- Serrano-Notivoli, R., M. De Luis, and S. Beguería. 2019. "Stead (Spanish Temperature at Daily Scale) [Dataset]."
- Setti, S., R. Maheswaran, V. Sridhar, K. K. Barik, B. Merz, and A. Agarwal. 2020. "Inter-Comparison of Gauge-Based Gridded Data, Reanalysis and Satellite Precipitation Product With an Emphasis on Hydrological Modeling." *Atmosphere* 11: 1252. <https://doi.org/10.3390/atmos11111252>.
- Sideris, I. V., M. Gabella, R. Erdin, and U. Germann. 2014. "Real-Time Radar-Rain-Gauge Merging Using Spatio-Temporal Co-Kriging With External Drift in the Alpine Terrain of Switzerland." *Quarterly Journal of the Royal Meteorological Society* 140: 1097–1111. <https://doi.org/10.1002/qj.2188>.
- Spinoni, J., S. Szalai, T. Szentimrey, et al. 2015. "Climate of the Carpathian Region in the Period 1961–2010: Climatologies and Trends of 10 Variables." *International Journal of Climatology* 35: 1322–1341. <https://doi.org/10.1002/joc.4059>.
- Tanarhte, M., P. Hadjinicolaou, and J. Lelieveld. 2012. "Intercomparison of Temperature and Precipitation Data Sets Based on Observations

in the Mediterranean and the Middle East.” *Journal of Geophysical Research: Atmospheres* 117, no. D12: 2011JD017293. <https://doi.org/10.1029/2011JD017293>.

Thiemig, V., G. N. Gomes, J. O. Skøien, et al. 2022. “Emo-5: A High-Resolution Multi-Variable Gridded Meteorological Dataset for Europe.” *Earth System Science Data* 14: 3249–3272. <https://doi.org/10.5194/essd-14-3249-2022>.

Thiemig, V., G. N. Ramos Gomes, J. O. Skøien, et al. 2020. “Emo-1 Arcmin: A High-Resolution Multi-Variable Gridded Meteorological Data Set for Europe (1990–2021).”

Thorne, P. W., M. G. Donat, R. J. H. Dunn, et al. 2016. “Reassessing Changes in Diurnal Temperature Range: Intercomparison and Evaluation of Existing Global Data Set Estimates.” *Journal of Geophysical Research: Atmospheres* 121, no. 10: 5138–5158. <https://doi.org/10.1002/2015JD024584>.

Vidal, J. P., E. Martin, L. Franchistéguy, M. Baillon, and J. M. Soubeyroux. 2010. “A 50-Year High-Resolution Atmospheric Reanalysis Over France With the Safran System.” *International Journal of Climatology* 11, no. 11: 1627–1644. <https://doi.org/10.1002/joc.2003>.

Wackernagel, H. 2003. “Geostatistical Models and Kriging.” *IFAC Proceedings Volumes* 36, no. 16: 543–548. [https://doi.org/10.1016/S1474-6670\(17\)34818-8](https://doi.org/10.1016/S1474-6670(17)34818-8). 13th IFAC Symposium on System Identification (SYSID 2003), Rotterdam, the Netherlands. <https://www.sciencedirect.com/science/article/pii/S1474667017348188>.

Weedon, G. P., G. Balsamo, N. Bellouin, S. Gomes, M. J. Best, and P. Viterbo. 2014. “The Wfdei Meteorological Forcing Data Set: Watch Forcing Data Methodology Applied to ERA-Interim Reanalysis Data.” *Water Resources Research* 50, no. 9: 7505–7514. <https://doi.org/10.1002/2014wr015638>.

Widmann, M., J. Bedia, J. M. Gutiérrez, et al. 2019. “Validation of Spatial Variability in Downscaling Results From the Value Perfect Predictor Experiment.” *International Journal of Climatology* 39, no. 9: 3819–3845. <https://doi.org/10.1002/joc.6024>.

Wilkinson, M. D., M. Dumontier, I. J. Aalbersberg, et al. 2016. “The Fair Guiding Principles for Scientific Data Management and Stewardship.” *Scientific Data* 3, no. 1: 160018. <https://doi.org/10.1038/sdata.2016.18>.

Willmott, C., C. Rowe, and W. Philpot. 1985. “Small-Scale Climate Maps: A Sensitivity Analysis of Some Common Assumptions Associated With Grid-Point Interpolation and Contouring.” *American Cartographer* 12: 5–16. <https://doi.org/10.1559/152304085783914686>.

Yamamoto, J. K. 2000. “An Alternative Measure of the Reliability of Ordinary Kriging Estimates.” *Mathematical Geology* 32: 489–509. <https://doi.org/10.1023/A:1007577916868>.

Zhuang, J., R. Dussin, D. Huard, et al. 2023. “Xesmf: Universal Regridder for Geospatial Data.”

(lower triangle, in bold). D1: STEAD, D2: E-OBS v27e, D3: Iberia01, D4: PTI-Clima v0, D5: HUMID01, D6: ROCIO-IBEB, D7: ERA5-Land, D8: CHELSA-W5E5, D9: EMO-1arcmin, D10: CERRA-SFC.

## Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Figure S1:** Median of maximum temperature for the period 1990–2014 (°C). **Figure S2:** Median of minimum temperature for the period 1990–2014 (°C). **Figure S3:** 50-years return value of maximum temperature for the period 1990–2014 (°C). The first row corresponds to the reference dataset, ROCIO-IBEB, the mean and the standard deviation of the other nine maps. **Figure S4:** Shape parameter of the GEV-distribution of annual maxima of maximum temperature for the period 1990–2014. The first row corresponds to the reference dataset, ROCIO-IBEB, the mean and the standard deviation of the other nine maps. **Figure S5:** 50-years return value of minimum temperature considering the whole period available for each dataset (°C). **Figure S6:** Shape parameter of the GEV-distribution of annual minima of minimum temperature considering the whole period available for each dataset (°C). **Table S1:** PACO—Pattern Correlation of maximum temperature (upper triangle) and minimum temperature