Received: 22 December 2020 | Revised: 29 June 2021 | Accepted: 9 August 2021

DOI: 10.1002/gdj3.131

DATA PAPER

Geoscience Data Journal RMetS WILEY

# Recovery of early meteorological records from Extremadura region (SW Iberia): The 'CliPastExtrem' (v1.0) database

José M. Vaquero[1,2] | Nieves Bravo-Paredes[2,3] | María Angeles Obregón[2,3] | Víctor M. S. Carrasco[2,3] | Maria Antonia Valente[4] | Ricardo M. Trigo[4] | Fernando Domínguez-Castro[5,6] | Javier Montero-Martín[2,3] | Javier Vaquero-Martínez[2,3] | Manuel Antón[2,3] | José Agustín García[2,3] | María Cruz Gallego[2,3]

[1]Departamento de Física, Centro Universitario de Mérida, Universidad de Extremadura, Mérida, Spain

[2]Instituto Universitario de Investigación del Agua, Cambio Climático y Sostenibilidad (IACYS), Universidad de Extremadura, Badajoz, Spain

[3]Departamento de Física, Facultad de Ciencias, Universidad de Extremadura, Badajoz, Spain

[4]Instituto Dom Luiz (IDL), Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal

[5]Fundación ARAID, Zaragoza, Spain

[6]Departamento de Geografía y Ordenación del Territorio, Universidad de Zaragoza, Zaragoza, Spain

**Correspondence**
José M. Vaquero, Departamento de Física, Centro Universitario de Mérida, Universidad de Extremadura, 06800 Mérida, Spain.
Email: jvaquero@unex.es

**Funding information**
Ministerio de Economía y Competitividad, Grant/Award Number: CGL2017-87917-P and PID2019-108589RA-I00; Ministerio de Ciencia e Innovación, Grant/Award Number: PRE2018-084897; Junta de Extremadura, Grant/Award Number: GR18097, IB16127 and PD18029

## Abstract

In this work, we provide instrumental meteorological data recovered for the Extremadura region (interior SW Iberia), from 1826 to mid-20th century. Meteorological variables such as air temperature, atmospheric pressure, precipitation, wind direction and humidity, among others, were retrieved. In total, more than 750 000 instrumental data in 157 meteorological series belonging to 131 different locations throughout Extremadura were rescued. It must be noted that daily resolution data constitutes 80% of the database. This great effort of digitization and data collection has been carried out with the aim of contributing to a significant expansion of the length of the databases with meteorological information in this region. Therefore, this database will provide a better understanding of climate variability, trends and extreme events of the Extremadura region.

**Dataset**

Identifier: The dataset is published on Pangaea: https://doi.pangaea.de/10.1594/PANGAEA.928037

Creator: J.M. Vaquero, N. Bravo-Paredes, M.A. Obregón, V.M.S. Carrasco, M.A. Valente, R.M. Trigo, F. Domínguez-Castro, J. Montero-Martín, J. Vaquero-Martínez, M. Antón, J.A. García, M.C. Gallego

Title: Early meteorological records from Extremadura region, SW Iberia (CliPastExtrem)

Publisher: PANGAEA. Data Publisher for Earth & Environmental Science

Publication year: 2020

Resource type: Data files

Version: 1.0

**KEYWORDS**
climate, Extremadura, historical data, Iberian Peninsula, meteorological observations

## 1 | INTRODUCTION

Interest in historical climatology has grown in recent decades. Relevant results obtained in this field for the current discussion on anthropogenic climate change (Neukom et al., 2019; Stocker et al., 2014) and studies about specific events such as extreme precipitation and floods (e.g. Domínguez-Castro et al., 2015; Trigo et al., 2014) have contributed to this growing interest. However, the limited availability of long and high-quality time series of relevant meteorological variables hampers a better framed and more accurate understanding, detection and prediction of the global climate variability and change (Brönnimann, 2015; Brönnimann et al., 2019b, 2019a). This limited availability can be abated, or even avoided, by digitizing the wealthy heritage of past climate data that exist, taking advantage of the fact that in Europe, since the mid-19th century and even earlier, the low atmosphere has been regularly monitored (Brönnimann et al., 2019b, 2019a).

There are regions around the world where digitization of long-time series of past climate data has increased significantly. However, Brönnimann et al. (2018) highlighted the needs of a more coordinated and long-term effort in meteorological and climatological data rescue. Some examples of international initiatives to recover climate data are the Atmospheric Circulation Reconstructions over the Earth (ACRE) (Allan et al., 2011, 2016) (http://www.metacre.org/), the International Data Rescue (I-DARE) (https://www.idare-portal.org/) and the Copernicus Climate Change Service (C3S) Data Rescue Service (https://datarescue.climate.copernicus.eu/). Moreover, several particular initiatives have been carried out. For example, Domínguez-Castro et al. (2017) retrieved more than 300 000 meteorological data summarized in 137 series from Latin America and the Caribbean during the 18th and 19th centuries. In Europe, Brugnara et al. (2020) digitized meteorological observations from 40 locations in Switzerland for the period 1708–1873, and Ashcroft et al. (2018) recovered around 8 million observations for Europe and the southern Mediterranean for the period 1877–2002. Other examples of this kind of work are Camuffo et al. (2017) for Italy or Hawkins et al. (2019) for Scotland.

The digitization of early meteorological data in the Iberian Peninsula started after similar endeavours took place in many other European countries, and the vast majority of data were still not digitized at the beginning of this century. In the Iberian Peninsula, there were projects focused on digitizing data (Alcoforado et al., 2012; Barriendos et al., 2000; Barriendos et al., 2002; Brunet & Jones, 2011; Domínguez-Castro et al., 2014a, 2014b). Research carried out in the last two decades in both Iberian countries, namely Spain and Portugal, has increased considerably the number of long-term time series of precipitation and temperature (e.g.: Brunet et al., 2006; Domínguez-Castro et al., 2014a, 2014b; Fragoso et al., 2010; Trigo et al., 2009). In Spain, the three locations with the oldest continuous data are San Fernando (Cádiz), Madrid and Barcelona. Observations in San Fernando began to be continuously recorded in 1786 (Barriendos et al., 2002; Obregón et al., 2020; Rodrigo, 2002). There are data available from Madrid since 1786 (Antón et al., 2014, 2017; Aparicio et al., 2019; Barriendos et al., 1997; Brunet et al., 2006) and although the Barcelona series started in 1780 and is practically continuous since then, the first official observations took place in 1885 (Rodríguez et al., 2001). The observations made by Diogo Nunes Ribeiro in Lisbon between 1 November 1724 and 11 January 1725 were the first known meteorological measurements made in the Iberian Peninsula. Although this is a short series, it was useful to study the intense and unusual tropical storm of November 1724 that caused significant damages in Madeira and continental Portugal (Domínguez-Castro et al., 2013).

One of the Spanish regions with the lowest number of meteorological data rescued is Extremadura (Southwestern Spain), despite its wealthy heritage of past climate data. Extremadura is a peripheral Spanish region dedicated to the primary sector (agriculture and livestock) with few industrial infrastructures (Tapiador, 2020). In the upper right corner of Figure 1, the situation of Extremadura in the Iberian Peninsula is depicted. The state meteorological network began with the Badajoz station in 1860 but had little development until the end of the Spanish Civil War (1936–1939) and, similarly to the remaining Spanish provinces, was not consolidated until the middle of the 20th century (Anduaga Egaña, 2012) when the density of stations raised in Spain. In this relatively undeveloped framework, it is worth stressing a curious exception related with the establishment of a private network of meteorological stations in the 1880s in the province of Badajoz (approximately the southern half of Extremadura), although it did not last for too long (Vaquero & Gallego, 2000).

The main objective of this work was to describe the 'CliPastExtrem' (v1.0) database that recovers early meteorological readings from numerous locations of the Extremadura region. We have retrieved more than 750 000 instrumental meteorological data corresponding to 130 different locations in the Extremadura region and one in a Portuguese small city (Campo Maior) located very close to the border between Spain and Portugal. From this work, the meteorological

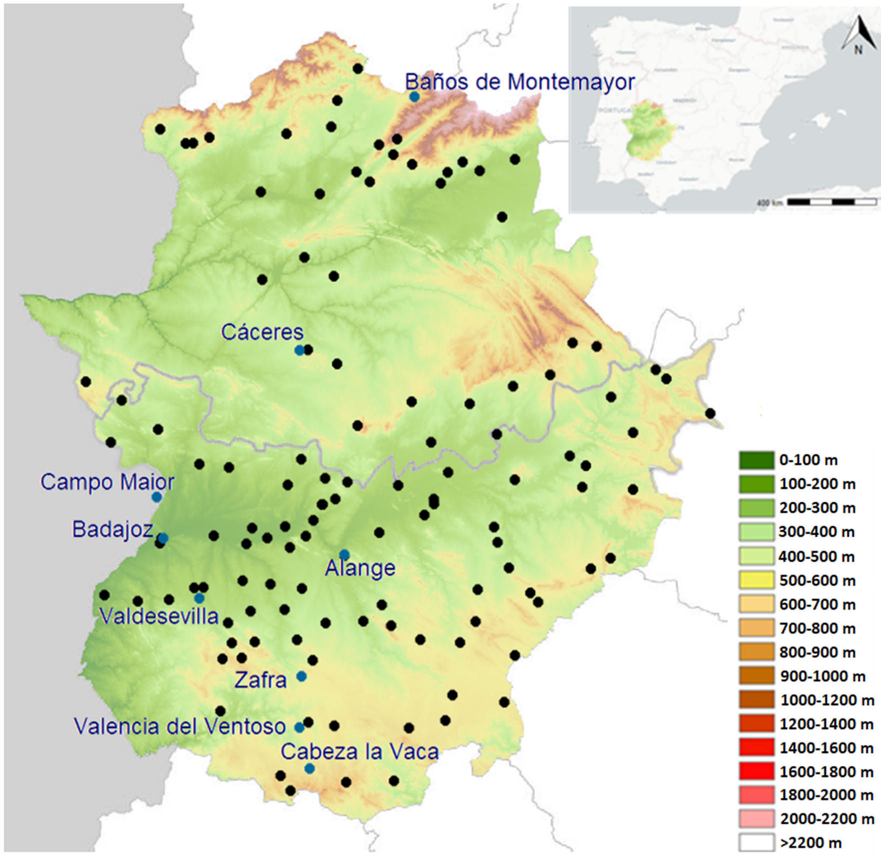



**FIGURE 1** Map of the Extremadura region (courtesy of Infraestructura de Datos Espaciales de Extremadura – IDE Extremadura) containing the locations with data in CliPastExtrem. The blue points indicate locations that appear with the name, and the grey line indicates the division between the Extremadura provinces of Cáceres (north) and Badajoz (south). In the upper right corner, the situation of the region of Extremadura in the Iberian Peninsula is represented

observation datasets for Extremadura can be extended to the early 20th and even mid-19th centuries. A large fraction of the data values digitized was obtained specifically for this work; however, an equally significant component of the data had been obtained previously. Only 1.9% of the data, those corresponding to Badajoz (1830–1833), were already published by Dominguez-Castro et al. (2014a) (Salvá-Sinobas project). Campo Maior data were digitized in 2007 under Portuguese funded project SIGN but has not been published yet. Furthermore, Fernández-Fernández et al. (2015, 2017) retrieved qualitative temperature and precipitation data recorded in Zafra for the period 1750–1840, and Bravo-Paredes et al. (2019) studied actinometric measurements from Cáceres for the period 1913–1920.

## 2 | METHODOLOGY AND DATASET

The CliPastExtrem dataset was obtained by completing the following steps: (1) search for documentary sources with unrecovered meteorological data, (2) digitization of meteorological data, (3) quality control of data and (4) data description.

### 2.1 | Documentary sources

CliPastExtrem was constructed by consulting many documents from different documentary sources. Regarding the same station, in general, there is no overlap in the observations included in the different sources consulted in this work. The types of documentary sources consulted in this research and their descriptions are shown in Table 1. The primary source for data rescue consisted of scientific annals (67%), followed by newspapers (22%) and meteorological records in monographs (11%). Figure 2 shows an example of each kind of documentary source used to retrieve the meteorological observations included in CliPastExtrem: (a) scientific records – Annaes do Observatório do Infante Dom Luiz; (b) newspapers – La Gazeta de Madrid; and (c) monographs – Memoria del año 1877 Sobre el Balneario de Alange. The description of each documentary source is given in a general point of view, not in the particular case of this database. For example, newspapers are primary documentary sources when information about water height of floods in the 19th and 20th centuries is recovered. In this case, the water height was recorded in the newspapers as original information and was not taken from another source. Anyway, to avoid possible misunderstandings, the description of newspapers as primary documentary sources was eliminated in Table 1.

**TABLE 1** Documentary sources consulted and their descriptions

| Type of documentary source | Description | Documents |
|---|---|---|
| Scientific annals | Scientific annals are the documentary sources that record the most complete information. The information presented in these documentary sources is extensive. These works are carried out by academies, universities or scientific institutions. The metadata is very detailed. Usually, annals published monthly information due to the fact that the information is a summary of the original records. Annals contained a large set of meteorological variables, such as pressure, temperature, humidity and wind direction They are primary documentary sources. All the meteorological data of locations in Extremadura were recovered. | *Annaes do Observatório do Infante Dom Luiz* (Lisbon)<br>*Anuario del Observatorio de Madrid*<br>*Estación meteorológica del Colegio de San José* (Villafranca de los Barros - Badajoz)<br>*Anuario del Servicio Meteorológico Español*<br>*Anuario del Observatorio Central Meteorológico* |
| Newspapers | Newspapers published the most important events that happened in a region (in villages and towns) since 18th century. So, early meteorological observations can be found in some newspapers. Usually, newspapers registered daily information. Also, weekly and monthly information can be found in some cases. The meteorological variables registered in newspapers are usually the temperature, pressure, wind direction, precipitation, days of precipitation and the state of the atmosphere. Unfortunately, metadata does not appear in most cases. Many issues of the newspapers are missing. So, it is difficult to reconstruct long meteorological series. Newspapers are secondary documentary sources due to the fact that institutions published their observations on them. | *Revista de Extremadura*<br>*La Gaceta de Madrid*<br>*Diario de Badajoz*<br>*Boletín de la Agencia de Cáceres* |
| Meteorological records in monographs | Some institutions published meteorological observations in monographs. These institutions can be very varied, such as medical institutions. Usually, a brief description of the instruments and the data are given. The information is registered daily or monthly. The meteorological variables recorded are the most common: pressure, temperature, etc. They are primary documentary sources. | *Monografías de las aguas y baños minerales (Catálogo Cisne - UCM)* |
| Manuscripts | Manuscripts include handwritten information by the author. It used to be made on a manageable medium such as paper or parchment. These are primary documentary sources. | *Consultas y Decretos del Ducado de Feria*<br>*Logbooks of Meteorological Observations. Library of Territorial headquarters of the AEMET in Badajoz* |

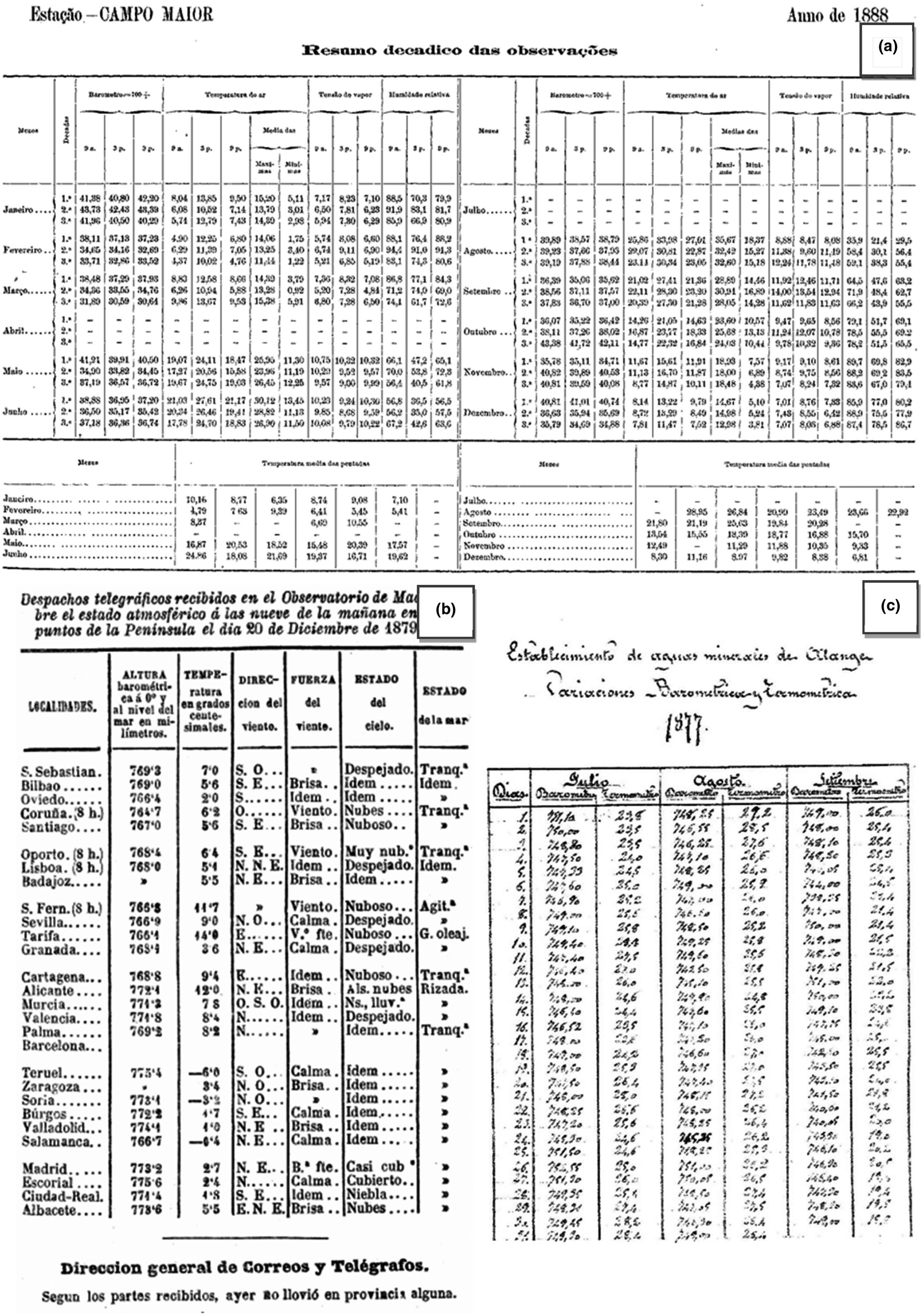
## 2.2 | Digitization

Once the documentary sources were located, they were photographed in situ with the aim of having the possibility to re-verify the data during the digitization process, when the documents were not available. This greatly speeds up the process. The authors, who are familiar with the database and the variables studied because they worked in the libraries and archives, digitized the data. The method used to digitize the data was the 'key entry' by one author for a given number of data (i.e. the data were divided into subsets and each author digitized a subset). Optical character recognition programmes were not used because of the great variety of documentary sources with different formats, layouts and typing, and also, because this type of programmes can lead to errors (Brönnimann et al., 2006; Tan et al., 2004). Furthermore, having digitized the data allowed the authors to correct some errors during digitization, such as inconsistencies in sequential dates or disappearance of decimal points.

## 2.3 | Quality control

Digitized data often contains errors of different origins and these should be detected and corrected whenever possible. These errors can be due to two main reasons:

- Original error: This results typically from writing an incorrect number in the official document of the institution instead of the real number from the manuscript of the observer (usually not available).




| LOCALIDADES. | ALTURA barométrica á 0° y al nivel del mar en milímetros. | TEMPERATURA en grados centesimales. | DIRECCION del viento. | FUERZA del viento. | ESTADO del cielo. | ESTADO de la mar |
|---|---|---|---|---|---|---|
| S. Sebastian. | 769'3 | 7'0 | S. O... | » | Despejado. | Tranq.ª |
| Bilbao...... | 769'0 | 5'6 | S. E... | Brisa. | Idem..... | Idem. |
| Oviedo...... | 766'4 | 2'0 | S...... | Idem. | Idem..... | » |
| Coruña. (8 h.) | 764'7 | 6'2 | O...... | Viento. | Nubes..... | Tranq.ª |
| Santiago.... | 767'0 | 5'6 | S. E... | Brisa.. | Nuboso.. | » |
| Oporto. (8 h.) | 768'4 | 6'4 | S. E... | Viento. | Muy nub.ª | Tranq.ª |
| Lisboa. (8 h.) | 768'0 | 5'4 | N. N. E. | Idem.. | Despejado. | Idem. |
| Badajoz..... | » | 5'5 | N. E... | Brisa.. | Idem..... | » |
| S. Fern. (8 h.) | 766'8 | 11'7 | » | Viento. | Nuboso... | Agit.ª |
| Sevilla..... | 766'9 | 9'0 | N. O... | Calma. | Despejado. | » |
| Tarifa...... | 766'4 | 14'0 | E...... | V.ª fte. | Nuboso... | G. oleaj. |
| Granada.... | 768'4 | 3'6 | N. E... | Calma. | Despejado. | » |
| Cartagena... | 768'8 | 9'4 | E...... | Idem.. | Nuboso... | Tranq.ª |
| Alicante.... | 772'4 | 12'0 | N. E... | Brisa.. | Als. nubes | Rizada. |
| Murcia...... | 771'2 | 7'8 | O. S. O. | Idem.. | Ns., lluv.ª | » |
| Valencia.... | 774'8 | 8'4 | N...... | Idem.. | Despejado. | » |
| Palma...... | 769'2 | 8'2 | N...... | » | Idem..... | Tranq.ª |
| Barcelona... | | | | | | |
| Teruel...... | 775'4 | —6'0 | S. O... | Calma. | Idem..... | » |
| Zaragoza... | » | 3'4 | N. O... | Brisa.. | Idem..... | » |
| Soria...... | 773'1 | —3'2 | N. O... | » | Idem..... | » |
| Búrgos..... | 772'2 | 4'7 | S. E... | Calma. | Idem..... | » |
| Valladolid.. | 774'1 | 4'0 | N. E.. | Brisa.. | Idem..... | » |
| Salamanca.. | 766'7 | —0'4 | N. E... | Calma. | Idem.... | » |
| Madrid..... | 773'2 | 2'7 | N. E.. | B.ª fte. | Casi cub.ª | » |
| Escorial.... | 775'6 | 2'4 | N...... | Calma. | Cubierto.. | » |
| Ciudad-Real. | 774'4 | 4'8 | S. E... | Idem.. | Niebla.... | » |
| Albacete.... | 773'6 | 5'5 | E. N. E. | Brisa.. | Nubes.... | » |

**Direccion general de Correos y Telégrafos.**

Segun los partes recibidos, ayer no llovió en provincia alguna.

**FIGURE 2** An example of each kind of documentary source used to retrieve meteorological observations: (a) scientific records – Annaes do Observatório do Infante Dom Luiz (Lisbon); (b) newspapers – La Gazeta; (c) monographs – Memoria del año 1877 Sobre el balneario de Alange

- Digitization error: This tends to result from the writing of an incorrect number in the transcription on the digital document instead of the real number from the original document (hand-written or printed).

In some cases, these errors cannot be detected easily (e.g. when the mistake is located in decimal digits). In other cases, it is easier (e.g. when the mistake is made in the first digit of the number, it is easier to find out due to the fact that these numbers seem obvious outliers). A visual check was performed immediately after transcriptions to a machine-readable file in order to detect the errors made in this process.

A basic quality control was carried out in order to detect possible typos, errors or suspicious values in daily data. The reason to check only the daily values lies in the ease of detecting suspicious values. The monthly and annual values are calculated as the average of the daily values (by the observers), and it is more difficult to detect potential errors. Note that, monthly and annual values were visually checked during the digitization process as mentioned previously. This quality control was applied to the following variables: temperature, pressure, humidity and cloud cover. In this work, a similar quality control made by Domínguez-Castro et al. (2017) has been adopted to detect possible errors considering the following steps:

- **Tolerance test**: Outlier values were detected by calculating the mean plus/minus three standard deviations of each variable. These values were flagged.
- **Temporal consistency**: The difference in the values between consecutive days should not exceed a predefined threshold. Values above this threshold could be an error, a suspicious value or an outlier. These values were also flagged. The threshold was calculated from the mean plus/minus three standard deviations for each variable and for each file.
- **Internal coherence test**: Some variables contain maximum, minimum and mean values. In these cases, values that do not keep the condition maximum > mean > minimum were flagged. In addition, some variables are recorded at different hours of the day. The same two steps described previously are also carried out in these cases. For example, generally, the temperature in the first hours of the day should be lower than the temperature at noon. Then, the same process is carried out as previously explained to flag values.
- **Limit test**: Some variables are bounded by upper and/or lower limit values. Tenths of covered sky are one of these variables. The lower limit is 0, and the upper limit is 10. Therefore, values above 10 correspond to errors, and these values were flagged.

We have also carried out a quality control by comparing data from neighbouring stations such as BABADA2 versus BABADA4, CCCACE2 versus CCCACE5 and BAVALD1 versus BAVALD2. Comparing data from these stations, no point is out of the range defined by the best linear fit plus/minus three times the standard deviation. Instead, some relevant discrepancies between the values of the measurements taken at neighbouring stations can be found. For example, that occurred in January 1882, when the monthly average of the temperature recorded at BABADA2 was 5.0°C and it was 8.7°C according to BABADA4, and that in January 1881, when the monthly mean temperature calculated from BAVALD1 was 4.6°C and that provided by BAVALD2 was 9.1°C. Note that, these values are not outliers and, for that reason, no flag was indicated in the database. These differences can be due to different kind of instruments used to measure, particular conditions in the locations of the stations, different number of observation days in a month, etc. In total, 188,140 values were analysed by this quality control procedure. This represents around 24% of the total dataset. The incorrect or suspicious values are flagged by an asterisk in the digitized files. In total, there are 807 values flagged (0.4% of the total data analysed). The percentage of these suspicious values in each variable is as follows: 62% for pressure values, 35% for temperature, 2% for humidity and 1% for cloud cover. These values were not corrected due to the fact that the real value is not known. Moreover, no unit change was carried out (also in the entire dataset). Taking into account, the set of 807 values flagged, and the percentage of suspicious values in each variable is as follows: 62% for pressure values, 35% for temperature, 2% for humidity and 1% for cloud cover. We note that the percentage of the number of data analysed in this work by the quality control with respect to the total data included in the dataset, and the percentage of the incorrect or suspicious values detected is similar to those obtained in other works. For example, Domínguez-Castro et al. (2017) analysed a quality control around 30% of data and detected 0.3% of problematic values.

## 2.4 | Data description

A large number of meteorological observations were retrieved from 131 locations of the Extremadura region and from the location of Campo Maior (located in the border of Portugal – around 20 km northwest of Badajoz, see Figure 1). In total, 772,253 meteorological observations were recovered. These records include daily, monthly, annual and decadal observations. In some stations, such as Cáceres and Badajoz, five sub-daily meteorological observations were recorded at 7 a.m., 8 a.m., 1 p.m., 4 p.m. and 6 p.m. The earliest observations included in this database correspond to the observations recorded at Zafra in 1826, while the most recent observations were those of cloud cover and sunshine duration made in Badajoz in 1955. Campo Maior data series cover the longest period of time and have the highest number of data points. In total, the Campo Maior series is composed

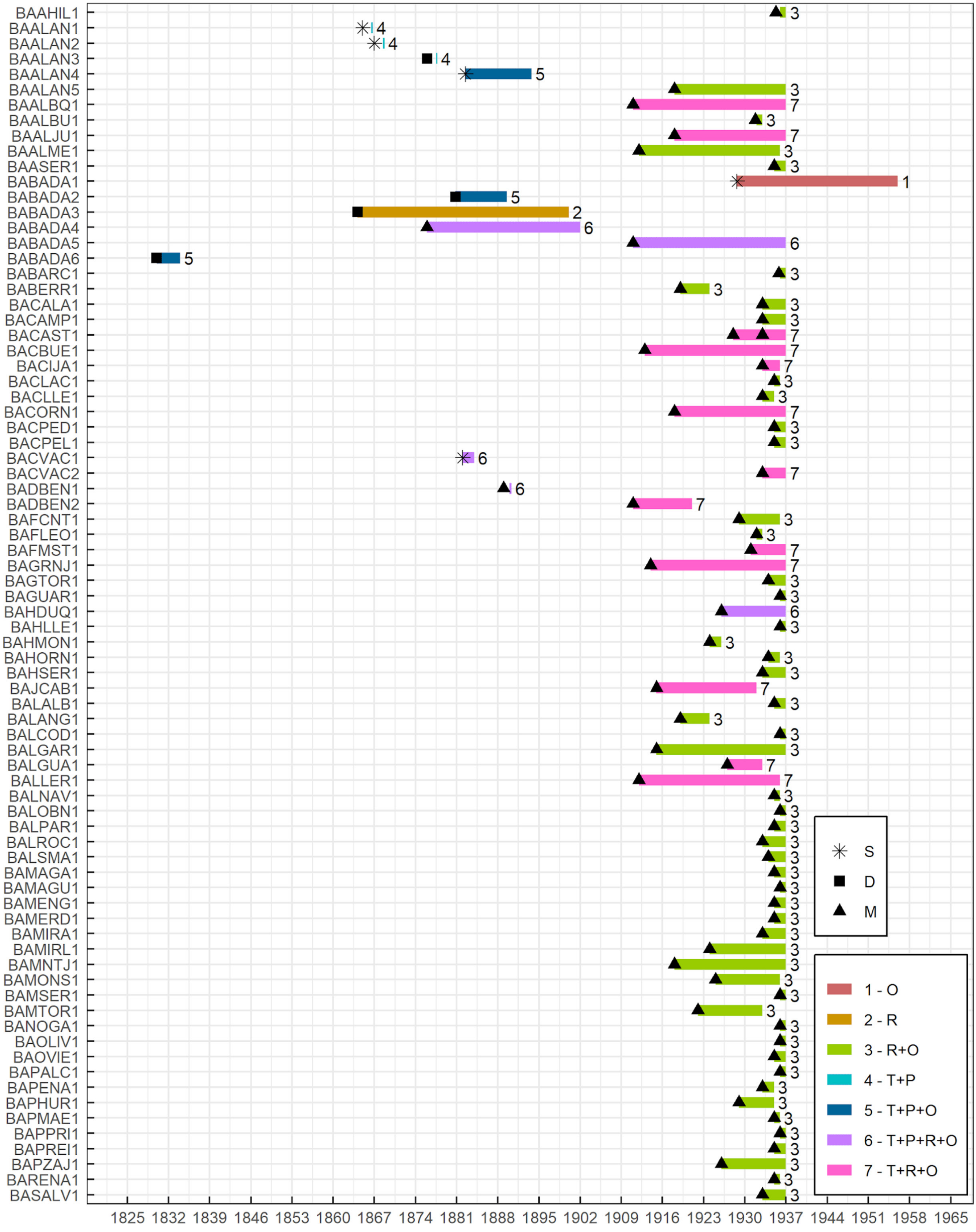



**FIGURE 3** Observational period covered for each station. The meteorological variables recorded at each station and the temporal resolution is also shown. The legend indicates the temporal resolution: D – daily; M – monthly; Dc – decadal; Y – annual. The letters of the colour legend indicate the meteorological variables recorded: T – temperature; P – pressure; R – rain; O – others

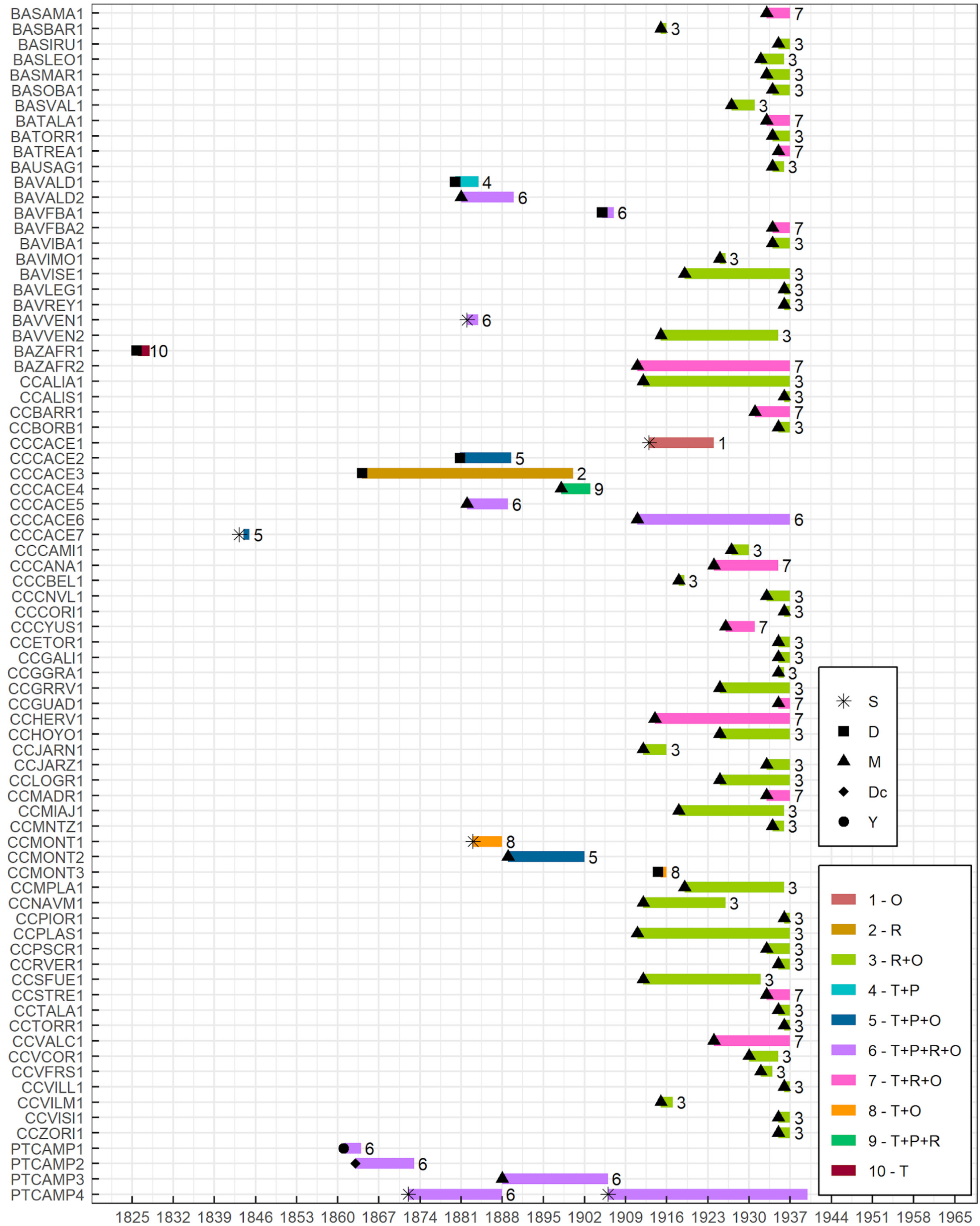



**FIGURE 3** (Continued)

of 481,692 meteorological observations. Figure 1 shows the geographical distribution of the locations with data included in the database CliPastExtrem.

The meteorological variables recorded in each station were not necessarily the same. Figure 3 shows information of the period covered for each station. Data from 12 different categories of variables, shown in Table 2, were recorded in the stations included in the dataset. These categories are pressure, temperature, wind, humidity, precipitation, evaporation, visibility, clouds, state of sky, sunshine duration, ozone and others. These 12 categories of variables contain 256 sub-variables, such as number of rainy days, maximum daily precipitation, temperature oscillation and number of frost days. Temperature, wind and pressure are the categories that encompass the largest number of sub-variables (179 almost 70%). Table 2 also shows the percentage of each variable makes up of the full dataset. As can be seen, the variable with the highest number of data is wind (21.40%), followed by temperature (19.45%) and pressure (15.81%). A general description of units and instruments used in the stations is given in the following paragraphs. Details for each station are provided in the station data files.

**TABLE 2** Variables recorded and their units. Note that the variables recorded in each station were not necessarily the same

| Variable | Units | Number of sub-variables | Percentage |
|---|---|---|---|
| Pressure | mmHg, inches | 38 | 15.81 |
| Temperature | °C, Reamur | 90 | 19.45 |
| Wind | km/h | 51 | 21.40 |
| Humidity | % | 10 | 7.29 |
| Precipitation | mm | 10 | 10.51 |
| Evaporation | mm | 2 | 2.50 |
| Visibility | - | 7 | 0.35 |
| Clouds | Several | 18 | 12.60 |
| State of Sky | - | 13 | 3.17 |
| Sunshine duration | h | 3 | 3.13 |
| Ozone | ozonometric degrees | 1 | 0.73 |
| Others | - | 13 | 3.01 |

The temperature was measured with thermometers, such as standard thermometers (e.g. Casella and Fuess) or extreme thermometers (Tonnelot or Six's thermometer), and it is expressed in Celsius or Reaumur degrees. According to the metadata, the monthly and annual averaged temperatures were calculated as the average between the maximum and minimum daily temperatures.

The atmospheric pressure is generally expressed in mm of mercury at 0°C and is corrected by capillary effects of the mercury tubes. When the station level pressure is available (there are stations without this information), information is given in the text files. The atmospheric pressure was measured with barometers, generally with the Tonnelot barometer, and although in Campo Maior, it was measured with the Kew barometer. A description of these instruments is given by Brombacher et al. (1960). Barographs were also available at Badajoz and Cáceres stations, such as Tonnelot or Weight Richard Barograph.

The most frequently observed wind direction at the time of observation was considered the dominant wind direction for that day. The daily wind path is the distance tracked by the wind from a certain time of the previous day to the same time of the date of the daily record. In case, data have other time scales (monthly or annual), and the wind path is accumulated during that period. This information was obtained by anemometers and wind vanes.

The rain was generally measured, according to the information displayed on the documentary sources, with Hellmann rain gauges. However, it was also measured with Hervé Magnon rain gauges (e.g. at Granja Badajoz station) and with Babinet udometer (e.g. at Campo Maior station). This variable is expressed as the height that the precipitated water would reach on the horizontal ground if it did not run, evaporate or filter. One day is considered as rainy when the rain measured was equal to or greater than 0.1 mm. Snowy days are those days in which this phenomenon occurred, even if it rained, not counting them as rainy days. The daily amount of rain is the fall measured from a certain time of the previous day to the same time of the date of the daily record.

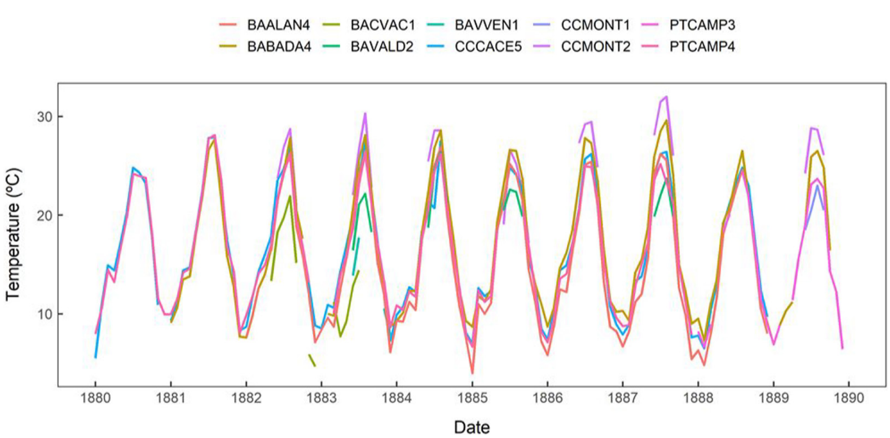



**FIGURE 4** Mean temperature observations recorded in all the stations with data for the period 1880–1889 retrieved in this work. Each colour represents one station

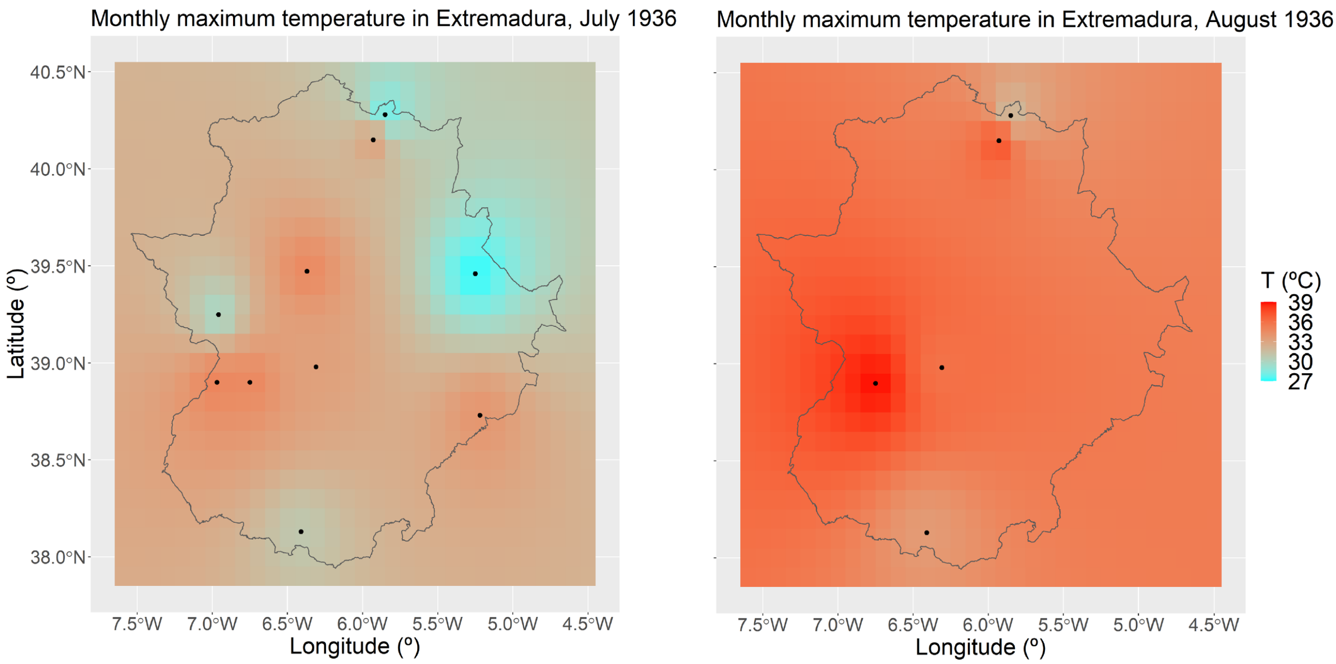



**FIGURE 5** Monthly maximum temperature recorded in Extremadura in July (left panel) and August (right panel) 1936

```
/* DATA DESCRIPTION:
Citation:       Vaquero, JM et al. (year): "Título", https://doi.org/10.1594/PANGAEA.xxxxx,
        Supplement to: Vaquero, JM et al. (In prep): Meteorological records from Extremadura (Spain) during the 19th and 20th cen
Abstract:       xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
Coverage:       LATITUDE: 38.784444 * LONGITUDE: -6.243611
        DATE/TIME START: 1868-07-01T09:00:00 * DATE/TIME END: 1868-09-30T21:00:00
Event(s):       BAALAN2 * LATITUDE: 38.784444 * LONGITUDE: -6.243611 * LOCATION: Spain * METHOD/DEVICE: Observation (OBSE)
Parameter(s):       DATE/TIME (Year) * GEOCODE * PI: Vaquero, José Manuel (https://orcid.org/0000-0002-8754-1509)
        DATE/TIME (Month) * PI: Vaquero, José Manuel (https://orcid.org/0000-0002-8754-1509)
        DATE/TIME (Day) * PI: Vaquero, José Manuel (https://orcid.org/0000-0002-8754-1509)
        DATE/TIME (Hour) * PI: Vaquero, José Manuel (https://orcid.org/0000-0002-8754-1509)
        Pressure [mmHg] (P) * PI: Vaquero, José Manuel (https://orcid.org/0000-0002-8754-1509)
        Temperature [deg] (T) * PI: Vaquero, José Manuel (https://orcid.org/0000-0002-8754-1509) * COMMENT: Reaumur thermometer
        Humidity [%] (H) * PI: Vaquero, José Manuel (https://orcid.org/0000-0002-8754-1509) * COMMENT: Saussure higrometer
License:        Creative Commons Attribution 4.0 International (CC-BY-4.0)
Size:       1104 data points
*/
Year    Month   Day     Hour    P [mmHg]        T [deg] H [%]
1868    7       1       9       639.2870        39.00201        78.23446
1868    7       1       12      639.7099        39.43131        78.75854
1868    7       1       15      640.2081        39.85553        79.33671
1868    7       1       21      640.6226        40.10165        79.82380
1868    7       2       9       640.9161        40.16490        80.14725
1868    7       2       12      641.1417        40.07497        80.39779
```

**FIGURE 6** The BAALAN2 station file showing an example of the file structure

In case, data have other time scales (monthly or annual), and the amount of rain is accumulated during that period.

The water vapour pressure is expressed in mm of mercury, and the relative humidity as a percentage of the amount of vapour that would be required to saturate the environment for a given temperature. Both were generally calculated from temperatures measured with psychrometres. However, in the particular case of Alange station during the years 1866 and 1868, the Saussure hair hygrometer was used. Evaporation was measured by Piche evaporimeters and is expressed in mm such as the rainfall data. The actinometric measurements were recorded with an Arago actinometer at Cáceres station during the period 1913–1923. The unit is expressed in Celsius degrees. The ozone was measured at Campo Maior station with a James ozonometer. The unit is expressed in ozonometric degrees.

On the one hand, some stations expressed the cloud cover in tenths, octas or quarters of clouds. On the other hand, other stations used the terms 'clear', 'cloudy' and 'overcast' in order to estimate the amount of 2/10clouds that covered the sky. As an example, regarding the measurements in tenths of clouds, the days of average cloud cover lower than 2/10 of the total sky were considered by observers as clear days, the days above 8/10 as overcast and the intermediate ones as cloudy. The days of persistent fog were considered as overcast, and if the fog disappeared during the morning, leaving the sky without clouds, as clear.

All the meteorological variables recorded at each station are shown in Figure 3. The letters of the legend indicate the meteorological variables recorded: T – temperature; P – pressure; R – rain; and O – others. The temporal scales of the different documentary sources are as follows: 79.80% for daily data, 17.32% for monthly data, 0.02% for annual data and 2.85% for decadal data. There are series covering periods of one year or less, but these series only represent 0.20% of the data. Regarding the different variables, 33.55% of the series contain temperature, 13.81% of the series contain pressure and most of the data series contain rainfall readings.

## 3 | DATA APPLICATIONS

As an example of the potential utilization of this database, Figure 4 shows the temporal evolution of the monthly mean temperature values recorded in all the stations with data during the period 1880–1889. It must be highlighted that the stations CACACE5, PTCAMP4 and, in particular BAALAN4, recorded a mean temperature in the winter of 1885 lower than in most winters of that decade and the summer of that same year was generally colder according to all stations. These facts could be still a consequence of the great eruption of Krakatoa volcano in August 1883 (Dörries, 2003; Obregón et al., 2020). However, this conclusion should be taken with caution taking into account the 2-year lag. Other examples of this are the cases of the data from Zafra between 1826 and 1827, which served to complement a study of precipitation and temperature indices from 1750 to 1840 using visual meteorological descriptions (Fernández-Fernández et al., 2014), and the data from Badajoz between 1830 and 1833, which were used to evaluate a great landslide occurred in 1831 trigger by meteorological conditions (García-Garrido et al., 2020).

Another example of possible use of the database can be found in Figure 5 where we depict the monthly average of the maximum temperature recorded at Extremadura in July and August 1936. We selected this year because its summer was very warm around the world (Conwan et al., 2017; Donat et al., 2016). There are ten Extremadura stations with data available in July 1936. Unfortunately, only five stations recorded data in August 1936. In order to obtain a spatial representation of these monthly temperature fields, we have interpolated using the inverse distance weighted interpolation scheme. A significant warm August also occurred in Extremadura (Figure 5), where it can be seen that the monthly maximum temperature in Talavera la Real (38.9°N, 6.75°W) was 38.9°C, and it was 37.4°C in Barrado (40.2°N, 5.9°W). In any case, note that the data are in their original units in this database.

## 4 | DATASET ACCESS

All the recovered records are publicly available at the World Data Center PANGAEA at https://doi.pangaea.de/10.1594/PANGAEA.925807. The dataset is provided in a zip file containing 157 text files. Each file contains data and metaheader. The metaheader is located in the first few lines of the file. The structure of the metaheader is composed of different elements: ID of the station, location, latitude and longitude, observational period covered by the series, altitude of the stations when this information is available, temporal scale and meteorological variables. The meteorological data are presented in the following lines of each file. The first columns provide the temporal information (year, month, day and hour), and the meteorological variables are represented in the following columns. The file names are composed by the ID of the station following a simple structure: The first two letters indicate the province of Extremadura where the station is located (CC for Cáceres and BA for Badajoz), and the next four letters indicate the name of the place where the station is located and then a number is shown. The number is due to the fact that there are stations with more than one file. A similar nomenclature has been applied to the file names of Campo Maior station, but in this case, the first two letters indicate the country of Portugal (PT). Figure 6 depicts the BAALAN2 station file, that is the data file number two for Alange (Badajoz), showing an example of the file structure.

## 5 | SUMMARY AND CONCLUSIONS

A dataset including more than 700,000 meteorological observations made in 131 locations of Extremadura (southwest of Spain) from 1826 to mid-20th century is presented in this work. Meteorological variables such as temperature, pressure and precipitation, inter alia, were retrieved from these observations. These records include sub-daily, daily, monthly, annual and decadal series. Note that, around 80% of the recovered data correspond to daily records. Different kinds of documentary sources such as scientific annals, newspapers and monographs were consulted to carry out this work.

After digitizing all data, we have applied a basic quality control in order to find outliers or suspicious values. Thus, we have analysed around a quarter of all data, detecting 0.43% of problematic values with respect to the analysed data. Both percentages are similar to those obtained in other works

(Domínguez-Castro et al., 2017). The dataset is publicly available at the World Data Center PANGAEA (https://doi.pangaea.de/10.1594/PANGAEA.925807).

Several studies have highlighted the importance of retrieving long meteorological series. Climate change is imposing new maximum and minimum temperature records in both Iberian countries at a wide range of temporal scales, from daily to decadal. It is important to contextualize these extreme events, and the associated climate tendencies using long-time series. For example, Portugal achieved an all-time record of 47.3°C during the extreme heatwave of 2003 (Trigo et al., 2006), while Lisbon attained its top rank temperature of 44°C in the recent heatwave of August 2018 (Sousa et al., 2019). In both cases, the possibility of comparing these new record values with longer homogenized time series starting in mid-19th century is relevant to assess their return periods and the likelihood of an anthropogenic fingerprint. Note that, the values included in our dataset are not homogenized. Even short historical series have been used to study past extreme events, as previously mentioned (e.g. Domínguez-Castro et al., 2013). Some projects published digitized data from meteorological observations made in Extremadura. Therefore, this database will help to better understand the climate of this region located at the southwest of Spain. Moreover, it can contribute to global studies on meteorological events that occurred in the Iberian Peninsula and Europe in the past or to be assimilated in reanalysis experiments.

## ACKNOWLEDGEMENTS

This research was supported by the Economy and Infrastructure Counselling of the Junta of Extremadura through project IB16127 and grants GR18097 (co-financed by the European Regional Development Fund) and by the Ministerio de Economía y Competitividad of the Spanish Government (CGL2017-87917-P and PID2019-108589RA-I00). The research of N. Bravo-Paredes has been supported by the predoctoral fellowship PRE2018-084897 from Agencia Estatal de Investigación (Ministerio de Ciencia, Innovación y Universidades) of the Spanish Government. J. Vaquero-Martínez thanks Junta de Extremadura and European Social Funds for the predoctoral fellowship PD18029. The authors acknowledge the help of the staff of the Observatory of Ebro and AEMet libraries where some historical observations included in this work were consulted.

## CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

## OPEN PRACTICES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The dataset is published on Pangaea: HYPERLINK "https://urldefense.com/v3/__https://doi.pangaea.de/10.1594/PANGAEA.928037__;!!N11eV2iwtfs!4qU2tpXGjOQFMqycOkP4rAxMF1xh-ymqbyAc6VtL3Z4zrIokjt6LI3VgkrWrIw$" https://doi.pangaea.de/10.1594/PANGAEA.928037. Learn more about the Open Practices badges from the Center for OpenScience: https://osf.io/tvyxz/wiki.

## ORCID

*José M. Vaquero* https://orcid.org/0000-0002-8754-1509
*Nieves Bravo-Paredes* https://orcid.org/0000-0002-4119-0395
*María Angeles Obregón* https://orcid.org/0000-0002-4645-5014
*Víctor M. S. Carrasco* https://orcid.org/0000-0001-9358-1219
*Maria Antonia Valente* https://orcid.org/0000-0001-8040-0829
*Ricardo M. Trigo* https://orcid.org/0000-0002-4183-9852
*Fernando Domínguez-Castro* https://orcid.org/0000-0003-3085-7040
*Javier Montero-Martín* https://orcid.org/0000-0001-7460-1405
*Javier Vaquero-Martínez* https://orcid.org/0000-0003-1741-3840
*Manuel Antón* https://orcid.org/0000-0002-0816-3758
*José Agustín García* https://orcid.org/0000-0001-5620-5660
*María Cruz Gallego* https://orcid.org/0000-0002-8591-0382

## REFERENCES

Alcoforado, M.J., Vaquero, J.M., Trigo, R.M. & Taborda, J.P. (2012) Early Portuguese meteorological measurements (18th century). *Climate of the Past*, 8(1), 353–371. https://doi.org/10.5194/cp-8-353-2012

Allan, R., Brohan, P., Compo, G.P., Stone, R., Luterbacher, J. & Brönnimann, S. (2011) The international atmospheric circulation reconstructions over the earth (ACRE) initiative. *Bulletin of the American Meteorological Society*, 92(11), 1421–1425. https://doi.org/10.1175/2011BAMS3218.1

Allan, R., Endfield, G., Damodaran, V., Adamson, G., Hannaford, M., Carroll, F. et al. (2016) Toward integrated historical climate research: the example of Atmospheric Circulation Reconstructions over the Earth. *Wiley Interdisciplinary Reviews: Climate Change*, 7(2), 164–174. https://doi.org/10.1002/wcc.379

Anduaga Egaña, A. (2012) *Meteorología, ideología y sociedad en la España contemporánea*. Consejo Superior de Investigaciones Científicas.

Antón, M., Román, R., Sanchez-Lorenzo, A., Calbó, J. & Vaquero, J.M. (2017) Variability analysis of the reconstructed daily global solar radiation under all-sky and cloud-free conditions in Madrid during the period 1887–1950. *Atmospheric Research*, 191, 94–100. https://doi.org/10.1016/j.atmosres.2017.03.013

Antón, M., Vaquero, J.M. & Aparicio, A.J.P. (2014) The controversial early brightening in the first half of 20th century: a contribution

from pyrheliometer measurements in Madrid (Spain). *Global and Planetary Change*, 115, 71–75. https://doi.org/10.1016/j.gloplacha.2014.01.013
Aparicio, A.J.P., Antón, M., Gallego, M.C., Sanchez-Lorenzo, A. & Vaquero, J.M. (2019) Re-evaluation of trends in atmospheric column transparency from pyrheliometer measurements in Madrid (1910–1929). *Atmospheric Research*, 217, 165–171. https://doi.org/10.1016/j.atmosres.2018.11.003
Ashcroft, L., Coll, J.R., Gilabert, A., Domonkos, P., Aguilar, E., Sigró, J. et al. (2018) A rescued dataset of sub-daily meteorological observations for Europe and the southern Mediterranean region 1877–2012. *Earth System Science Data*, 10, 1613–1635. https://doi.org/10.5194/essd-10-1613-2018
Barriendos, M., Creus, J., González, M., López, P., Rivas, V. & Serrano, E. (2000) *La climatología histórica en España. Primeros resultados y perspectivas de la investigación. V Reunión Nacional de Climatología*, 15–56.
Barriendos, M., Gomez, B. & Peña, J. (1997) Old series of meteorological readings for Madrid and Barcelona (1780–1860). Documentary and observed characteristics. In: Martín-Vide, J. (Ed.) *Advances in Historical Climatology in Spain, Barcelona*. Oikos-Tau: pp. 157–172.
Barriendos, M., Martín-Vide, J., Peña, J.C. & Rodríguez, R. (2002) Daily meteorological observations in Cádiz – San Fernando. Analysis of the documentary sources and the instrumental data content (1786–1996). *Climatic Change*, 53, 151–170. https://doi.org/10.1023/A:1014991430122
Bravo-Paredes, N., Gallego, M.C., Antón, M., Núñez, M. & Vaquero, J.M. (2019) Analysis of actinometric measurements under all-sky and cloud-free conditions in Cáceres (Spain) for the period 1913–1920. *Tellus, Series B: Chemical and Physical Meteorology*, 71(1), 1–8. https://doi.org/10.1080/16000889.2019.1663597
Brombacher, W.G., Johnson, D.P. & Cross, J.L. (1960). *Mercury Barometers and Manometers*. Monograph 8. National Bureau of Standards.
Brönnimann, S. (2015) *Climatic Changes Since 1700*. Springer. ISBN 978-3-319-19042-6.
Brönnimann, S., Allan, R., Ashcroft, L., Baer, S., Barriendos, M., Brázdil, R. et al. (2019b) Unlocking pre-1850 instrumental meteorological records: A global inventory. *Bulletin of the American Meteorological Society*, 100(12), ES389–ES413. https://doi.org/10.1175/BAMS-D-19-0040.1
Brönnimann, S., Annis, J., Dann, W., Ewen, T., Grant, A.N., Griesser, T. et al. (2006) A guide for digitising manuscript climate data. *Climate of the Past*, 2(2), 137–144. https://doi.org/10.5194/cp-2-137-2006
Brönnimann, S., Brugnara, Y., Allan, R.J., Brunet, M., Compo, G.P., Crouthamel, R.I. et al. (2018) A roadmap to climate data rescue services. *Geoscience Data Journal*, 5, 28–29. https://doi.org/10.1002/gdj3.56
Brönnimann, S., Martius, O., Rohr, C., Bresch, D.N. & Lin, K.-H.-E. (2019a) Historical weather data for climate risk assessment. *Annals of the New York Academy of Sciences*, 1436(1), 121–137. https://doi.org/10.1111/nyas.13966
Brugnara, Y., Pfister, L., Villiger, L., Rohr, C., Isotta, F.A. & Brönnimann, S. (2020) Early instrumental meteorological observations in Switzerland: 1708–1873. *Earth System Science Data*, 12, 1179–1190. https://doi.org/10.5194/essd-12-1179-2020
Brunet, M. & Jones, P. (2011) Data rescue initiatives: Bringing historical climate data into the 21st century. *Climate Research*, 47(1–2), 29–40. https://doi.org/10.3354/cr00960
Brunet, M., Saladié, O., Jones, P., Sigró, J., Aguilar, E., Moberg, A. et al. (2006) The development of a new dataset of Spanish daily adjusted temperatures series (SDATS) (1850–2003). *International Journal of Climatology*, 26, 1777–1802. https://doi.org/10.1002/joc.1338
Camuffo, D., Della Valle, A., Bertolin, C. & Santorelli, E. (2017) Temperature observations in Bologna, Italy, from 1715 to 1815: a comparison with other contemporary series and an overview of three centuries of changing climate. *Climatic Change*, 142, 7–22. https://doi.org/10.1007/s10584-017-1931-2, 2017
Cowan, T., Hegerl, G.C., Colfescu, I., Bollasina, M.A., Purich, A. & Boschat, G. (2017) Factors contributing to record-breaking heat waves over the great plains during the 1930s Dust Bowl. *Journal of Climate*, 30, 2437–2461. https://doi.org/10.1175/JCLI-D-16-0436.1
Domínguez-Castro, F., Ramos, A.M., García-Herrera, R. & Trigo, R.M. (2015) Iberian extreme precipitation 1855/1856: An analysis from early instrumental observations and documentary sources. *International Journal of Climatology*, 35(1), 142–153. https://doi.org/10.1002/joc.3973
Domínguez-Castro, F., Trigo, R.M. & Vaquero, J.M. (2013) The first meteorological measurements in the Iberian Peninsula: Evaluating the storm of November 1724. *Climatic Change*, 118(2), 443–455. https://doi.org/10.1007/s10584-012-0628-9
Domínguez-Castro, F., Vaquero, J.M., Gallego, M.C., Farrona, A.M.M., Antuña-Marrero, J.C., Cevallos, E.E. et al. (2017) Early meteorological records from Latin-America and the Caribbean during the 18th and 19$^{th}$. *Scientific Data*, 4, 170169. https://doi.org/10.1038/sdata.2017.169
Domínguez-Castro, F., Vaquero, J.M., Rodrigo, F.S., Farrona, A.M.M., Gallego, M.C., García-Herrera, R. et al. (2014a) Early Spanish meteorological records (1750–1850). *International Journal of Climatology*, 34(3), 593–603. https://doi.org/10.1002/joc.3709
Domínguez-Castro, J.C., Vaquero, J.M., Gallego, M.C. & García-Herrera, R. (2014b) Climatic potential of Islamic Chronicles in Iberia: Extreme droughts (AD 711–1010). *The Holocene*, 24(3), 370–374. https://doi.org/10.1177/0959683613518591
Donat, M.G., King, A.D., Overpeck, J.T., Alexander, L.V., Durre, I. & Karoly, D.J. (2016) Extraordinary heat during the 1930s US Dust Bowl and associated large-scale conditions. *Climate Dynamics*, 46, 413–426. https://doi.org/10.1007/s00382-015-2590-5
Dörries, M. (2003) Global science: the eruption of Krakatau. *Endeavour*, 27, 113–116. https://doi.org/10.1016/S0160-9327(03)00107-8
Fernández-Fernández, M.I., Gallego, M.C., Domínguez-Castro, F., Trigo, R.M., García, J., Vaquero, J.M. et al. (2014) The climate in Zafra from 1750 to 1840: history and description of weather observations. *Climatic Change*, 126(1), 107–118. https://doi.org/10.1007/s10584-014-1201-5
Fernández-Fernández, M.I., Gallego, M.C., Domínguez-Castro, F., Trigo, R.M. & Vaquero, J.M. (2015) The climate in Zafra from 1750 to 1840: Precipitation. *Climatic Change*, 129, 267–280. https://doi.org/10.1007/s10584-014-1315-9
Fernández-Fernández, M.I., Gallego, M.C., Domínguez-Castro, F., Trigo, R.M. & Vaquero, J.M. (2017) The climate in Zafra from 1750 to 1840: Temperature indexes from documentary sources. *Climatic Change*, 141, 671–684. https://doi.org/10.1007/s10584-017-1910-7
Fragoso, M., Trigo, R.M., Zêzere, J.L. & Valente, M.A. (2010) The exceptional rainfall event in Lisbon on 18 February 2008. *Weather*, 65(2), 31–35. https://doi.org/10.1002/wea.513
García-Garrido, J.P., Gallego, M.C., Palacios, T., Trigo, R.M. & Vaquero, J.M. (2020) Heavy rainfall and landslide event in January 1831 at the Pedregoso Mountains (Cabeza Del Buey, SW Spain). *Atmosphere*, 11, 544. https://doi.org/10.3390/atmos11050544

Hawkins, E., Burt, S., Brohan, P., Lockwood, M., Richardson, H., Roy, M. & et al. (2019) Hourly weather observations from the Scottish Highlands (1883–1904) rescued by volunteer citizen scientists. *Geoscience Data Journal*, 6, 160–173. https://doi.org/10.1002/gdj3.79

Neukom, R., Steiger, N., Gómez-Navarro, J.J., Wang, J. & Werner, J.P. (2019) No evidence for globally coherent warm and cold periods over the preindustrial Common Era. *Nature*, 571, 550–554. https://doi.org/10.1038/s41586-019-1401-2

Obregón, M.A., Gallego, M.A., Antón, M. & Vaquero, J.M. (2020) Sunshine duration data in San Fernando (South of Spain) during 1880s: The impact of Krakatoa volcanic eruption. *Geoscience Data Journal*, 7(2), 185–191. https://doi.org/10.1002/gdj3.101

Rodrigo, F.S. (2002) Changes in climate variability and seasonal rainfall extremes: A case study from San Fernando (Spain), 1821–2000. *Theoretical and Applied Climatology*, 72(3–4), 193–207. https://doi.org/10.1007/s007040200020

Rodríguez, R., Barriendos, M., Jones, P.D., Martín-Vide, J. & Peña, J.C. (2001) Long pressure series for Barcelona (Spain). Daily reconstruction and monthly homogenization. *International Journal of Climatology*, 21(13), 1693–1704. https://doi.org/10.1002/joc.696

Sousa, P., Barriopedro, D., Ramos, A.M., García-Herrera, R., Espirito-Santo, F. & Trigo, R.M. (2019) Saharan air intrusions as a relevant mechanism for Iberian heatwaves: The record breaking events of August 2018 and June 2019. *Weather and Climate Extremes*, 26, 100224. https://doi.org/10.1016/j.wace.2019.100224

Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J. et al. (Eds.) (2014): *Climate change 2013: The physical science basis*. Cambridge University Press.

Tan, L.S., Burton, S., Crouthamel, R., van Engelen, A., Hutchinson, R., Nicodemus, L. et al. (2004) *Guidelines on Climate Data Rescue*. World Meteorological Organization, 1210.

Tapiador, F. (2020) *The Geography of Spain: A Complete Synthesis*. Springer.

Trigo, R.M., Pereira, J.M.C., Pereira, M.G., Mota, B., Calado, M.T., DaCamara, C.C. & et al. (2006) The exceptional fire season of summer 2003 in Portugal. *International Journal of Climatology*, 26(13): 1741-1757, 15 2006.

Trigo, R.M., Vaquero, J.M., Alcoforado, M., Barriendos, M., Taborda, J., Garcia-Herrera, R. & et al. (2009) Iberia in 1816, the year without a summer. *International Journal of Climatology*, 29, 99–115. https://doi.org/10.1002/joc.1693

Trigo, R.M., Varino, F., Ramos, A.M., Valente, M.A., Zêzere, J.L., Vaquero, J.M. et al. (2014) The record precipitation and flood event in Iberia in December 1876: Description and synoptic analysis. *Frontiers in Earth Science*, 2, 1–15. https://doi.org/10.3389/feart.2014.00003

Vaquero, J.M. & Gallego, M.C. (2000) Una red meteorológica privada en el Badajoz decimonónico. *Revista Española de Física*, 14(4), 58–60.

**How to cite this article:** Vaquero, J.M., Bravo-Paredes, N., Obregón, M.A., Carrasco, V.M.S., Valente, M.A., Trigo, R.M., et al (2022) Recovery of early meteorological records from Extremadura region (SW Iberia): The 'CliPastExtrem' (v1.0) database. *Geoscience Data Journal*, 9, 207–220. https://doi.org/10.1002/gdj3.131

## APPENDIX 1

## Documentary sources consulted

- Observatorio Central Meteorológico (España). 1914. Resumen de las observaciones meteorológicas efectuadas en la península y algunas de sus islas adyacentes durante los años 1911 y 1912. Observatorio Central Meteorológico.
- Observatorio Central Meteorológico (España). 1915–1932. Resumen de las observaciones meteorológicas efectuadas en las estaciones del Servicio Meteorológico Español durante el año 1913–1926. Observatorio Central Meteorológico.
- Observatorio Central Meteorológico (España). 1932–1950. Resumen de las observaciones efectuadas durante el año 1927–1937. Observatorio Central Meteorológico.
- Annaes do Observatório do Infante Dom Luiz em Lisboa. 1863–1939. Vols. 1–77.
- Observatorio Astronómico de Madrid (España). 1884–1904. Resumen de las observaciones efectuadas en la Península y algunas de sus islas adyacentes durante los años 1876–1901. Observatorio Astronómico de Madrid.
- Anuario del Servicio Meteorológico Español.
- Monografía de las aguas y baños minerales de Alange. 1829, 1838, 1847, 1848, 1850, 1861, 1865, 1866, 1868, 1872, 1876–1887, 1889–1893, 1895–1901. –Manuscripts preserved in the Library of the Universidad Complutense (Madrid, Spain).
- Monografía de las aguas y baños minerales de Montemayor. 1839–1842, 1844, 1845, 1860–1866, 1876, 1879, 1880–1885, 1887, 1889–1892, 1894–1901, 1916. Manuscripts preserved in the Library of the Universidad Complutense (Madrid, Spain).
- Boletín del Observatorio Meteorológico del Colegio de Nuestra Señora del Recuerdo. Madrid. Published data of the meteorological station of the "Colegio de San José" (Villafranca de los Barros, Badajoz, Spain).
- Revista de Extremadura. 1899–1902. Observaciones meteorológicas tomadas en Cáceres referentes a los años 1897–1901. Tomo I–IV. Revista de Extremadura.
- Gazeta de Madrid. Colección histórica. 1863–1901. https://www.boe.es/buscar/gazeta.php?
- Consultas y Decretos del Duque de Feria. Archivo Municipal de Zafra, Badajoz (Spain).
- Logbooks of Meteorological Observations. Library of Territorial headquarters of the AEMET in Badajoz.