1 Input data

The study area comprises the whole Spanish territory (Figure S1) including the Iberian Peninsula (IP), the Balearic Islands (BAL) and the Canary Islands (CAN), covering a total area of 504,660 km^2 (IP: 492,175 km^2; BAL: 4,992 km^2; CAN: 7,493 km^2). The altitudes range from 0 m a.s.l. at the coast to maximums of 3,404 m a.s.l. (Aneto summit) in the IP, 358 m a.s.l. (Toro Mountain summit) in BAL and 3,718 m a.s.l. (Teide summit) in CAN.
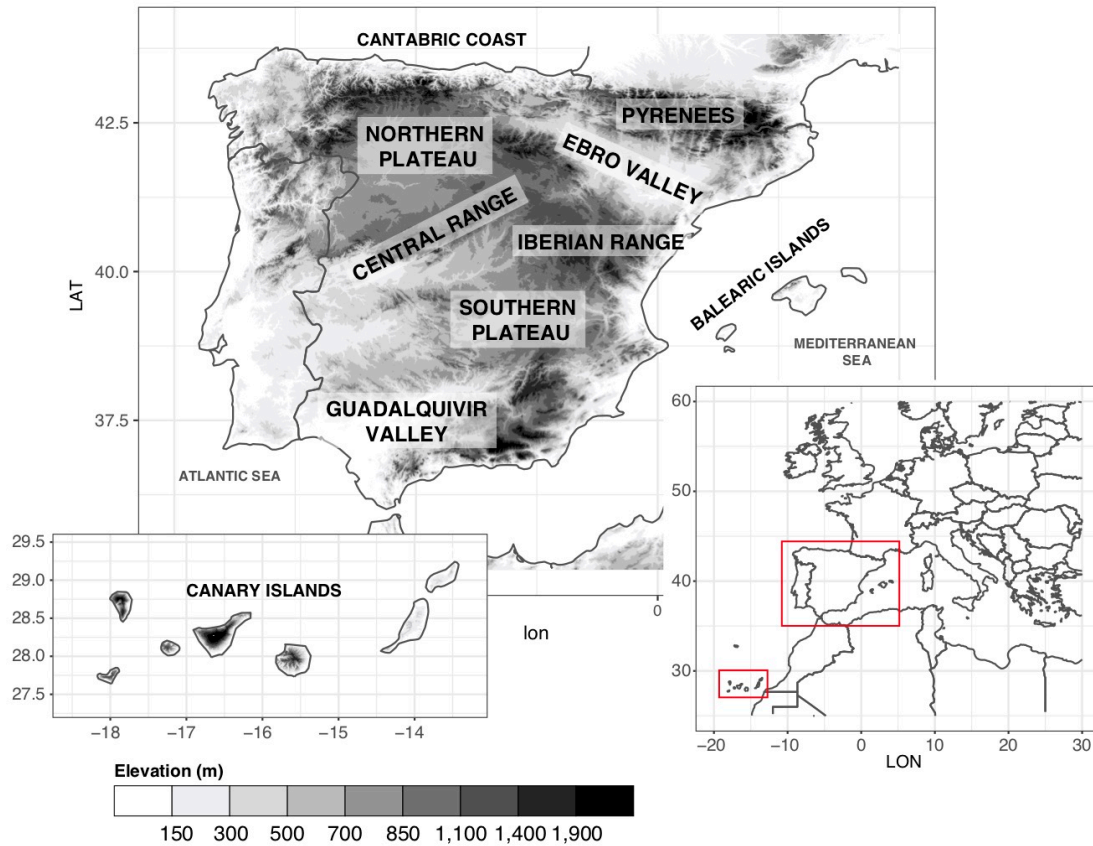
Figure S1: Geographical context of the study area and names of places mentioned in the text.

Most of the input data sourced from the Spanish Meteorological Agency (Aemet) (IP: 90.8%; BAL: 90.1%; CAN: 90.7%) and from the Spanish Ministry of Agriculture and Environment (MAGRAMA) (IP: 8.2%; BAL: 8.9%; CAN: 8.3%) using the 1901-2014 period for IP and 1971-2014 period for BAL and CAN. More than 5,000 stations were used as original observations with more than 33 millions of raw data (Table S1).

Table S1. Summary of all the used information of maximum and minimum temperature in the Spanish Iberian Peninsula (PEN), the Balearic (BAL) and Canary Islands (CAN). Number of months: months with at least 1 daily record in TMAX or TMIN; Number of complete months: months with records in all days; Number of daily pairs: days with record of TMAX and TMIN. The figures are referred to 1901-2014 period in PEN and 1971-2014 in BAL and CAN.

	PEN	BAL	CAN
Number of stations	5,056	124	340
Number of months	1,133,100	21,875	51,210
Number of complete months	1,057,941	18,530	39,831
Number of daily pairs	33,972,660	639,134	1,485,160

2 Results of quality control

The *iQC* results disaggregated by territorial unit (PEN: Peninsular Spain; BAL: Balearic Islands; CAN: Canary Islands) are presented in absolute figures and in percentage of the total number of available days and months in the original dataset (Table S2). The number of removed days corresponding to those out of range ($TMAX >= 50 \mid TMAX <= -30 \mid TMIN <= -35 \mid TMIN >= 40$) and to complete months ($n < 3 \mid sdTMAX == 0 \mid sdTMIN == 0 \mid meanTDIF == 0 \mid sdTDIF == 0$) were very similar.

Table S2. Number of data and percentage of total removed by the initial quality control (*iQC*) in the Spanish Iberian Peninsula (PEN), the Balearic (BAL) and Canary Islands (CAN).

	PEN	BAL	CAN
Number of removed days *(out of range)*	52 (<0.01%)	7 (<0.01%)	0 (0.00%)
Number of removed months	1,179 (0.10%)	34 (0.16%)	136 (0.27%)
Number of removed days *(complete months)*	3,965 (0.01%)	57 (0.01%)	286 (0.02%)

Once the data series passed the *iQC*, daily reference values (*dRV*) were computed for each station based on a general monthly reference (rmRV) and a local monthly reference (fmRV). Figure S2 shows the differences between them on an example month (July 2005). Results show that they only differ in some decimal points, being these differences always lower than 1 ºC.
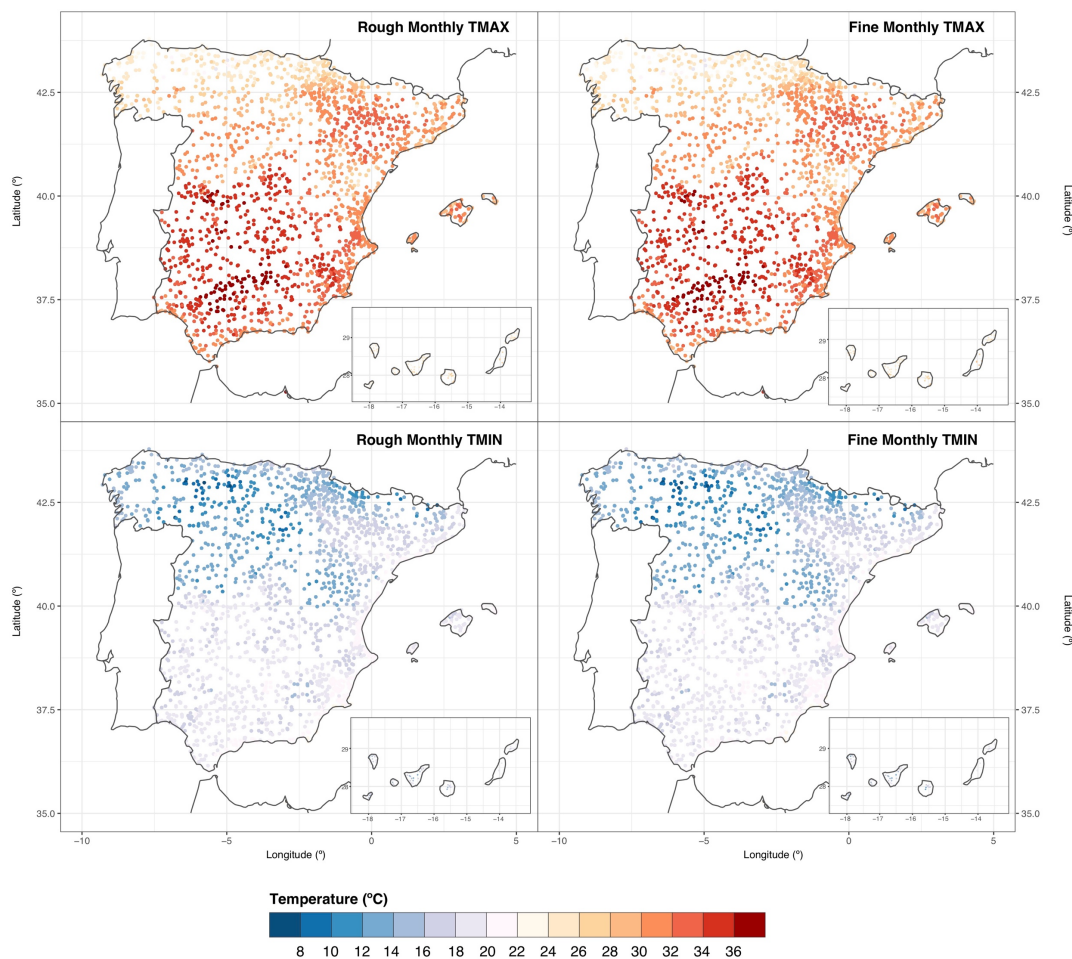
Figure S2. Rough Monthly RV (left) and Fine Monthly RV (right) for maximum (top) and minimum (bottom) temperature of an example month (July 2005).

Higher differences than the *iQC* were shown between territories in the results of the deep quality control (*dQC*) (Table S3). In regard of daily similarities, more data were removed in the Canary Islands than the rest of the areas (about three times more than PEN and almost five times more than Balearic Islands in TMAX). The number of removed data regarding the daily differences was more similar between zones, only in PEN was higher than 1% (in TMAX).

Table S3. Number of data and percentage of total removed by the deep quality control (*dQC*) in the Spanish Iberian Peninsula (PEN), the Balearic (BAL) and Canary Islands (CAN).

		PEN	BAL	CAN
Daily similarities (number of removed months and days)				
TMAX	Number of removed months	14,181 (1.25%)	144 (0.66%)	1,316 (2.57%)
	Number of removed days	411,880 (1.21%)	3,521 (0.55%)	35,101 (2.36%)
TMIN	Number of removed months	17,706 (1.56%)	204 (0.93%)	1,582 (3.09%)
	Number of removed days	517,790 (1.52%)	5,383 (0.84%)	43,262 (2.91%)
Daily differences (number of removed days)				
TMAX	Number of removed days	534,363 (1.57%)	6,262 (0.98%)	10,650 (0.72%)
TMIN	Number of removed days	286,696 (0.84%)	4,066 (0.64%)	9,042 (0.61%)

When we look at the correlations between observations and estimates of removed data (daily similarities stage of *dQC*), we note that the patterns of correlation frequencies are similar in TMAX and TMIN in peninsular Spain (Figure S3a and b) and in Canary Islands (Figure S3e and f) (always higher in TMIN in both cases), but different in Balearic Islands (Figure S3c and d). Nevertheless, the average correlation in almost all cases is negative near to zero (except in BAL TMAX, Figure S3c).
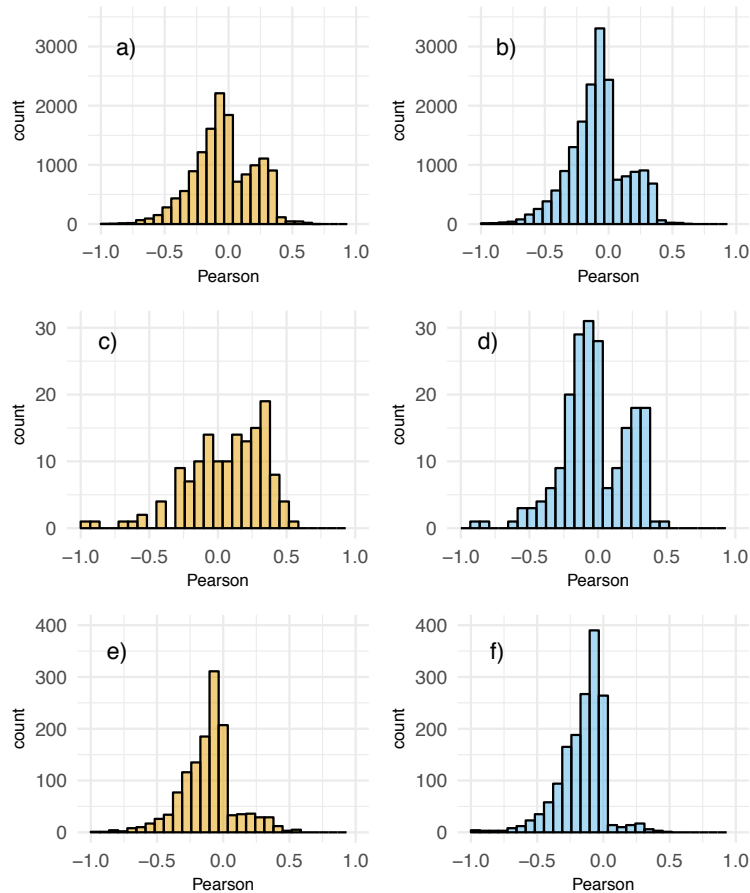
Figure S3. Correlation frequencies between observations and estimates of removed data in quality control process. Maximum (orange) and minimum (blue) temperatures are shown for peninsular Spain (a and b), Balearic Islands (c and d) and Canary Islands (e and f). Dashed orange and blue lines represent mean correlation of maximum and minimum temperature, respectively.

The differences between observations and estimates were very similar in the three areas both in TMAX (Figure S4b, e and h) and TMIN (Figure S5b, e and h). It is noteworthy that the mean difference was always zero or near to zero, which means that the removed data are very different from their corresponding observations. The display of the anomalies (Figures S4 and S5c, f and i) showed a clustering of the data in two groups: negative and positive anomalies in both spatial (Zs) and temporal (Zt) dimensions. However, most of the data were concentrated near to zero. The interpretation of this clustering is that the majority of the data in the plot was removed due to a combination of (negative or positive) anomalies between -10 and 10. Thus, these anomalies were not very extreme but enough to be as different as be rejected. Some of the removed data in the plot of anomalies with low Zs or Zt should not be present. Though, they derive from situations in which other daily values of one month are removed because different criteria and then, that month has less than three daily values. Subsequently, these values are removed fulfilling the *iQC* criteria.
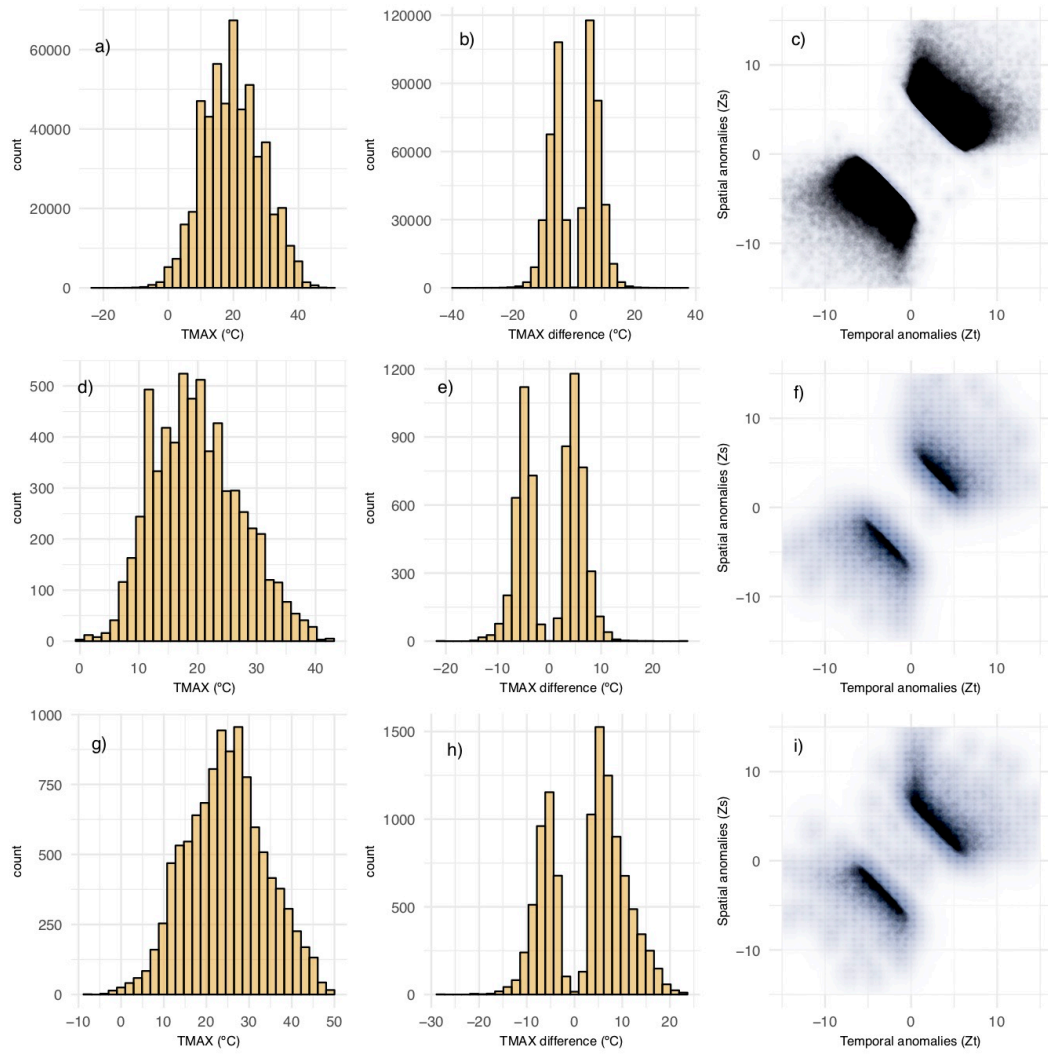
Figure S4. Daily maximum temperature data removed by quality control process in peninsular Spain (a, b and c), Balearic Islands (d, e and f) and Canary Islands (g, h and i). Left column: removed data by magnitude; central column: removed data by differences between observations and estimates; and right column: temporal anomalies (Zs) vs spatial anomalies (Zs). Dashed vertical lines in histograms represent mean values.
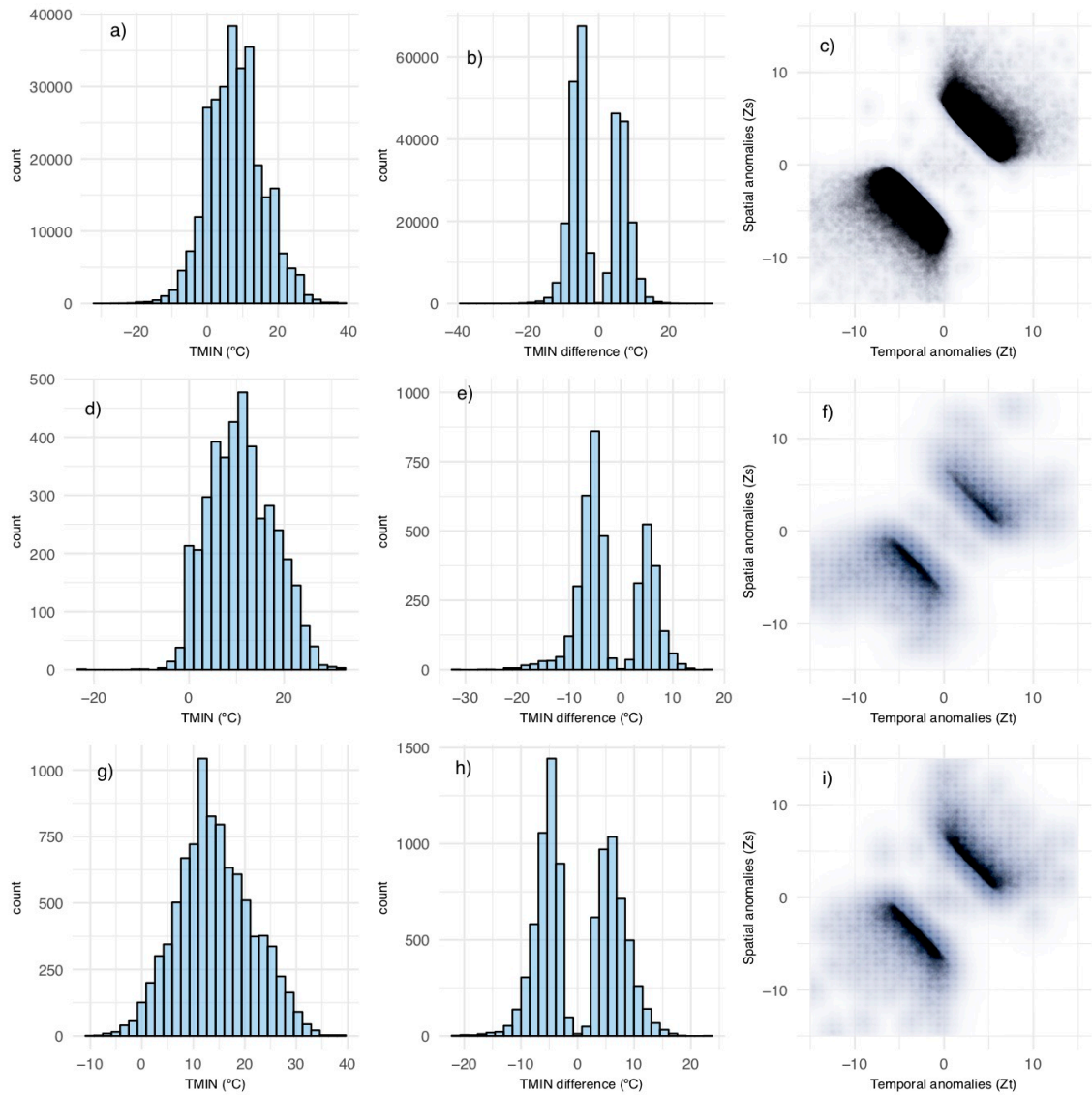
Figure S5. Daily minimum temperature data removed by quality control process in peninsular Spain (a, b and c), Balearic Islands (d, e and f) and Canary Islands (g, h and i). Left column: removed data by magnitude; central column: removed data by differences between observations and estimates; and right column: temporal anomalies (Zs) vs spatial anomalies (Zs) of removed data. Dashed vertical lines in histograms represent mean values.

As the *dQC* process is iterative, the detection, flagging and removing of suspect data is repeated until no suspect data is detected. The number of iterations depends on the original number and quality of the data, being more likely the need of more iterations when the dataset is higher. In case of Spain (Figure S6), PEN required between 12 and 14 iterations to complete the quality control process, while BAL and CAN used less than 10 (7 and 8, respectively).
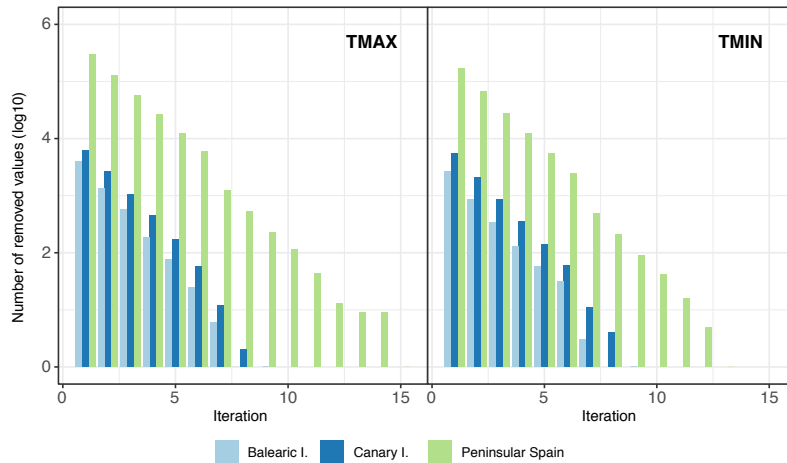
Figure S6. Number of removed values (in logarithmic scale) by iteration in *dQC* process in maximum (left) and minimum (right) temperature.

3 Observations-estimates comparison

Most of the correlations between daily observations and estimates in the final (reconstructed) dataset were near to 0.8 (Figure S7). However, the Canary Islands showed slightly lower correlations (near to 0.75), probably due to the great differences between observatories in raw data. This was also noted in daily precipitation (Serrano-Notivoli et al., 2017a) showing that the great variations in orography and the high differences in the climate between sides of same islands, have a key influence on the differences between data series of observatories, having an impact on climatic reconstructions.
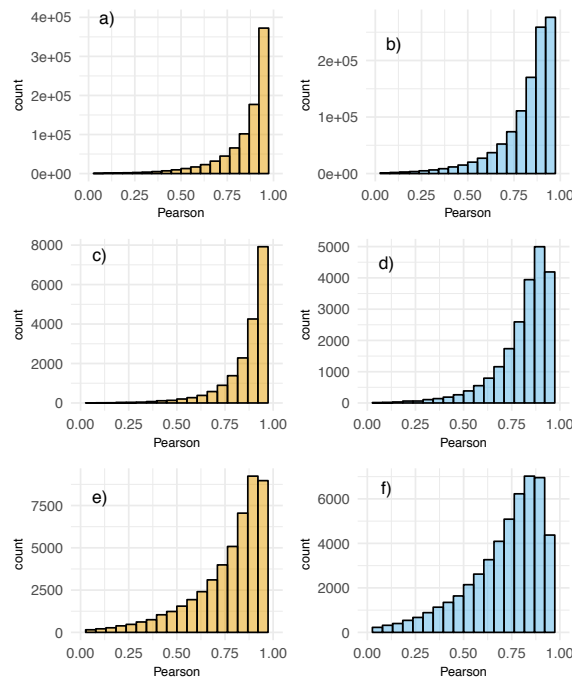
Figure S7. Correlation frequencies between observations and estimates of data in the final dataset. Maximum (orange) and minimum (blue) temperatures are shown for peninsular Spain (a and b), Balearic Islands (c and d) and Canary Islands (e and f). Dashed orange and blue lines represent mean correlation of maximum and minimum temperature, respectively.

The structure of the final dataset, after the iterative quality control and the estimate of TMAX and TMIN for gap filling, showed very low differences between observations and estimates (Figure S8a, d and g). Most of these differences were near to zero, being the TMAX in Balearic Islands slightly broader than the rest of the areas. This means that the maximum temperature in this region is more difficult to estimate due to higher differences in original values. The spatial (Zs) and temporal (Zt) anomalies in this case (Figure S8b, c, e, f, h and i) are the inverse of those in the removed data. The great majority of the Zs and Zt in peninsular Spain were concentrated between -1 and 1, while this figure is higher in the islands from -2 to 2. Nevertheless, the extreme anomalies are extended to absolute values not higher than 4 in the IP and 6 in BAL and CAN.
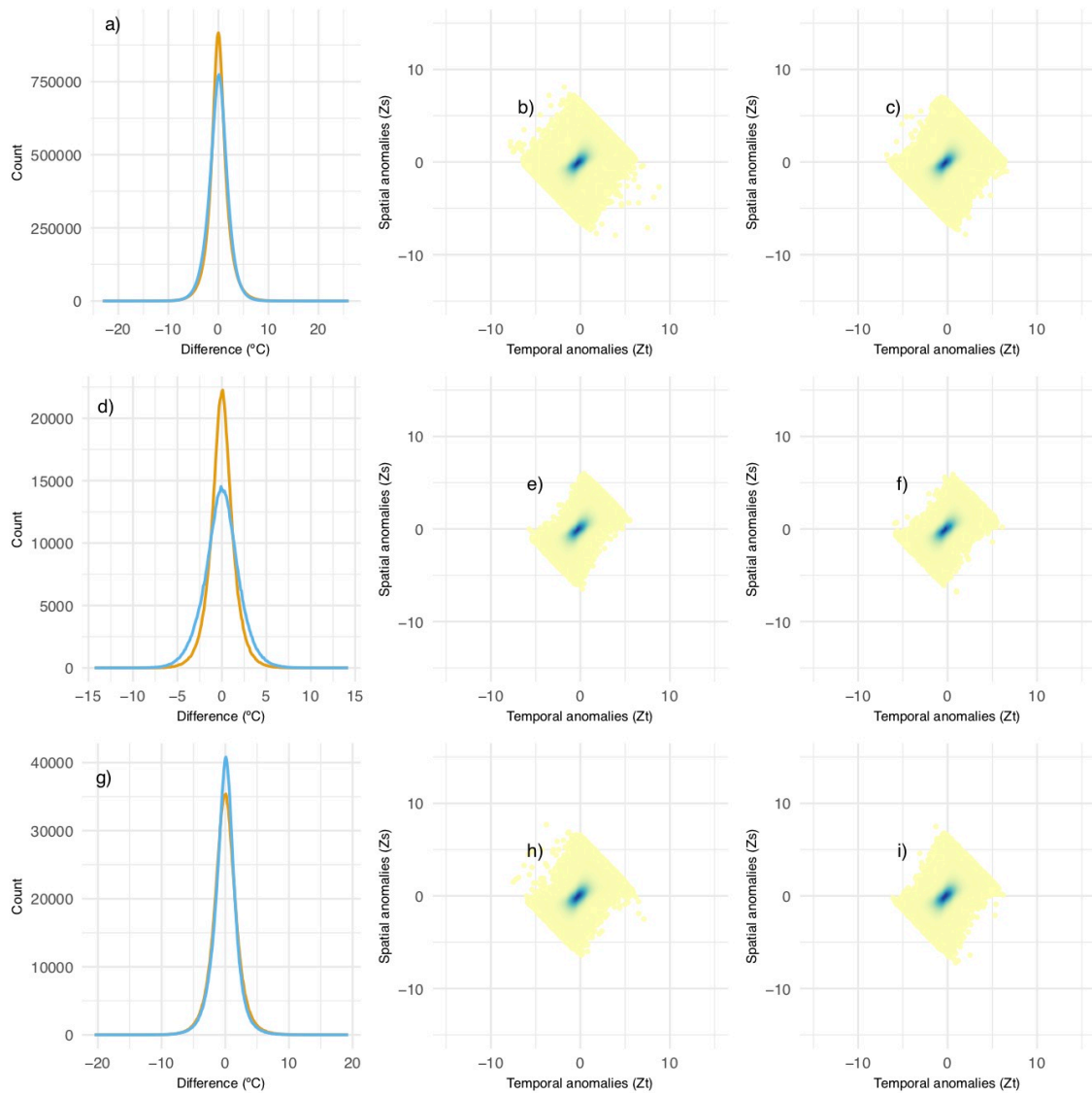
Figure S8. Structure of the final dataset in peninsular Spain (a, b and c), Balearic Islands (d, e and f) and Canary Islands (g, h and i). Left column: temperature differences between observations and estimates; central and right columns: temporal anomalies (Zs) vs spatial anomalies (Zs) of final dataset.