INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. (2010)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/joc.2192

The minimization of the *screen bias* from ancient Western Mediterranean air temperature records: an exploratory statistical analysis

Manola Brunet,[a,b]* Jesús Asin,[c] Javier Sigró,[a] Manuel Bañón,[d] Francisco García,[e] Enric Aguilar,[a] Juan Esteban Palenzuela,[f] Thomas C. Peterson[g] and Phil Jones[b]

[a] *Centre for Climate Change (C3), Department of Geography, University Rovira i Virgili, Av. Catalunya, 35, Tarragona 43071, Spain*
[b] *Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK*
[c] *Statistical Methods Department, University of Zaragoza, C/María de Luna, 3, 50015 Zaragoza, Spain*
[d] *Agencia Estatal de Meteorología, Alicante Meteorological Observatory, C/Regidor Ocaña, 26, 03011 Alicante, Spain*
[e] *Agencia Estatal de Meteorología, Galician Regional Meteorological Centre, C/Canceliña, 8-15011 La Coruña, Spain*
[f] *Agencia Estatal de Meteorología, Murcian Regional Meteorological Centre, Avda. de la Libertad, 11, Guadalupe, 30071 Murcia, Spain*
[g] *NOAA's National Climate Data Center, Asheville, NC, USA*

ABSTRACT: Here we present an exploratory statistical analysis aimed at the minimization of the 'screen bias' from affected ancient air temperature time series over the Western Mediterranean. Our approach lies in the statistical analysis of about 6 years of daily paired temperature observations taken using the ancient Montsouri shelter and the modern Stevenson screen for daily maximum (T_x) and minimum (T_n) temperature data recorded at two experimental sites: the meteorological gardens of La Coruña and Murcia, Spain (locations under the influence of the Oceanic/Atlantic/Galician and Mediterranean arid and semi-arid climate types, respectively), where ongoing field trials have been carried out. Descriptive statistical analysis of the paired series shows pre-sheltered temperatures tended to induce a strong warm bias in T_x data (of about 1 °C at the annual scale but with a clear seasonal cycle with higher values in summer and lower in winter), while T_n readings have a small (~0.2 °C, and sustained all year round) cold bias compared to the modern period. Statistical relationships between the screen bias and other related meteorological variables show the highest correlation coefficients between the 'screen bias' and T_x, T_n and the diurnal temperature range (DTR) recorded under the replicated ancient shelters at both locations and point to the reliability of these variables as potential predictors of the ΔT_x. We generate a parsimonious regression model based on the data from both experimental sites, which takes into account polynomial terms of lower order for the predictor variables (T_x and DTR recorded under the ancient shelter) and harmonic terms, in order to represent the seasonal cycle of the screen bias. The goodness-of-fit of the model is satisfactory, as it explains up to 51.7% of the additional ΔT_x variability. Copyright © 2010 Royal Meteorological Society

KEY WORDS screen bias; long temperature records; inter-screen comparison; statistical analysis; regression model; Western Mediterranean climates

Received 3 August 2009; Revised 29 December 2009; Accepted 17 May 2010

1. Introduction

1.1. The problem

It is well known that the exposure of thermometers, and the shelters that protect them, at land stations have changed over time with a wide variety of designs and exposures. These changes have compromised the homogeneity of century-long air temperature records and their adjustment is difficult to estimate by applying most state-of-the-art homogenization techniques, as the changeover to new screens took place at similar times in the past at most sites in the national meteorological networks. Relative homogeneity testing techniques in such cases are of little use. Instead, different approaches need to be designed to assess the scale of the problem and provide the likely adjustments that are needed.

In order to protect thermometers from direct or indirect radiation and wetting, different kinds of radiation screens were designed and installed. Since the second half of the 19th century, when a generalized concern for measuring 'the true air temperature' took place among meteorologists, different kinds of shelters and screens were constructed and used at meteorological gardens: shelters or open stands, free-standing screens and wall screens (see Sparks (1972) for a description of protectors used worldwide). Among the open stands, the so-called French or Montsouri shelter (MONT hereinafter) (Angot, 1903) and Glaisher stand (Figure 11 in Parker, 1994) were widely used. Particularly they were employed in South-western and Western Europe in the late 19th and early 20th centuries, depending on the country and

---

* Correspondence to: Manola Brunet, Centre on Climate Change (C3), University Rovira i Virgili, Av. Catalunya, 35, Tarragona 43002, Spain. E-mail: manola.brunet@urv.cat

meteorological networks (Sparks, 1972; Parker, 1994), until they were replaced by the new Stevenson screen (STEV from now on).

1.2. Previous attempts to identify and solve the problem

Contemporary intercomparison between temperature readings recorded under MONT shelter or Glaisher stand and the old or new STEV screens were carried out, as reported by Parker (1994), in order to assess the impact of the shelter used when measuring air temperature at different locations and climate types. For example, comparisons between MONT and STEV screens carried out at the meteorological gardens of Berlin during 1886–1887, Hamburg during 1910–1912, Pavlovsk (Russia) during 1898–1899 and in Paris for 50 years of paired data (based on Dettwiller (1978) and quoted by Parker (1995)) indicated that the MONT shelter when compared to the new STEV screen had introduced a strong warm bias in daily maximum temperature (T_x) readings with the highest values in summer (from 0.7 °C in Berlin to 1.1 °C in Hamburg) and the lowest in winter (from 0.1 °C in Pavlovsk to 0.3 °C in Berlin), as well as a very slight cold bias in daily minimum temperature (T_n) measurements, with steady differences across winter and summer (of about 0.2 °C). With regard to the difference between readings taken under a Glaisher stand and a STEV screen at Strathfield–Turgiss in 1869, at Greenwich during 1887–1889, at Croydon during 1877–1881 and Camden Square in London during 1881–1915, summarized by Parker (1994) for the UK and by Nordly *et al.* (1997) for the Swedish site at Karlstad during 7 years of paired observations, all showed a similar pattern of warm bias in T_x observations and cold bias in T_n readings, although the cold bias in T_n measurements is a bit larger than reported by the comparisons between MONT and STEV screens. From these early studies, it is clear that the *screen bias* associated with free-standing open stands has likely induced a warm (cold) bias, in the early period, in long T_x (T_n) records, whose magnitude depends on the latitude, on the time during the day/year and on the weather conditions when the measurements were taken (Parker, 1994). Therefore, the use of the unadjusted data in assessments of long-term temperature variability and change (for periods extending before about 1900) could lead to an underestimate of long-term trends.

More recent field tests of thermometer screens and comparisons between readings taken under ancient stands and modern screen have been carried out in order to estimate their influence on thermometric readings over different climates and world regions: over Nordic European countries (Andersson and Mattisson, 1991; Nordly *et al.*, 1996, 1997; Bergström and Moberg, 2002; Moberg *et al.*, 2002, 2003; Klingbjer and Moberg, 2003), over the United Kingdom (Butler *et al.*, 2005), over the Greater Alpine Region (Böhm *et al.*, 2009), over the Mediterranean Basin (Brunet *et al.*, 2004, 2006, 2007; Bañón *et al.*, 2008), over the United States (Chenoweth, 1992), over Australia (Richards *et al.*, 1992, 1993; Nicholls *et al.*, 1996) or on the global scale (Parker, 1994, 1995).

Although these studies and reviews identified, assessed or reviewed the sign and magnitude of the *screen bias* induced by the changeover of ancient exposures to modern screen, only a few have attempted to adjust the bias with respect to modern readings from the analysed time series. In addition, the need for enhancing the availability of homogenized climate records, particularly at the daily and sub-daily scales, is being highly recommended (e.g. Peterson *et al.*, 1998; Aguilar *et al.*, 2003; Brunet *et al.*, 2008; Böhm *et al.*, 2009). Both these studies and the World Meteorological Organization (WMO) guidelines clearly indicate the need to use high-quality instrumental climate datasets for more robustly assessing climate variability and change over different time and space scales. Along with other causes of inhomogeneities, these studies also describe, discuss or develop schemes to minimize the *screen bias* from the longest raw temperature records. Because the changeover to modern screen frequently took place at similar times in a network, the application of relative homogeneity tests are unable to identify this breakpoint in homogeneity, as most of land stations were likely to be in the same way affected in the earlier period. For instance, in Spain, the substitution of old shelters by the new STEV screens took place, first, in Madrid in 1893 and in the mid-1910s for most of the other stations that compose the Spanish historical air temperature network (see Table IX of Brunet *et al.* (2008) for specific details on the dates of the introduction of the STEV screen at each one of the stations that were used to develop the Spanish Daily Adjusted Temperature Series – SDATS – which incorporates part of the Spanish meteorological network since the 2nd half of the 19th century). According to the relevant contemporary literature and internal reports (see Table III of Brunet *et al.* (2008) for a list of these metadata sources) prior to the dates of the STEV screen's introduction in Spain, two kinds of open stands were used: the so-called French or MONT shelter and the Glaisher stand. No evidence of other kinds of thermometric exposures, such as window installations or north wall exposures, has emerged from the searches undertaken.

Instead of using relative homogeneity tests for detecting and minimizing this bias in the long affected records, it is necessary to use other approaches and homogenization methods. One way is to carry out a field experimental intercomparison of paired temperature readings recorded under the old and new screens and estimate relationships from the inter-screen difference series, in combination with other relevant variables. These relate the temperatures of the one screen (the old one) into those of the other (the new one), as highlighted by the WMO's guides on metadata and homogenization (Aguilar *et al.*, 2003) and on adjusting long daily temperature series (Brunet *et al.*, 2008). If additional variables (other than the temperature measurements themselves) are incorporated into the relationships, it is important to ensure that these variables were accurately measured in earlier times and their time series are homogenous. Additionally, the use of other meteorological variables for homogenizing temperature data also implies some restrictions, since the resulting adjusted data are no longer independent with respect to the other climate variable employed for estimating the adjustments. Therefore, the adjusted data are only usable for a limited number of applications, which cannot include any analysis for assessing the relationships between the homogenized temperature data and the variable used for estimating the adjustment.

Moberg *et al.* (2003) assessed the reliability of pre-1860 summer temperatures series in Stockholm and Uppsala (southern Sweden) since their station histories suggested earlier observations could be biased due to deficient radiation protection. The authors attempted to determine if independent support for the warm summers (i.e. as measured) could be gained from other related climate variables by means of stepwise multiple regression analysis. They investigated nine potential predictor variables (six air circulation indices, precipitation, air pressure and cloud amount) and concluded that the pre-1860 values are very likely positively biased as the observed decadal mean temperatures during 1780–1860 were above the upper limit of a 95% confidence interval that explains uncertainties both in the regression relationship and in the cloudiness records, a variable that was the most important one explaining summer temperature variance. This study clearly showed the importance of cloudiness, particularly in summer, but cloudiness measurements in Sweden before the 1950s are virtually impossible to homogenize with any confidence (Moberg *et al.*, 2003).

On the basis of a short field trial, Butler *et al.* (2005) assessed and corrected thermometric readings at Armagh Observatory (Northern Ireland) accounting for changes in exposure by means of comparing about 1 year (December 2003 to October 2004) of paired observations recorded in a light metal box located in a north window and in a STEV screen placed in the garden. However, the authors concluded that the readings in both settings were in remarkable agreement, as only a slight systematic warm bias of ∼0.2 °C was detected. Consequently, the authors corrected old temperatures recorded in the light metal box to the STEV screen, a correction that was not seasonally dependent. Given the shortness of the comparison and the more cloudy nature of the climate of Ireland, the magnitude may be biased low.

Over the Greater Alpine Region (GAR), Böhm *et al.* (2009) employed recent dual observations taken during 8 years under the old exposure (thermometers placed on north-facing walls or windows across the GAR) and modern STEV screen at Kremsmünster observatory (Austria). The authors used the orientation angle of the buildings (sites across the GAR in the range from NE to NW), along with the difference series between both unscreened and screened exposures at Kremsmünster and estimated adjustments for minimizing the pre-1870 biased temperature diurnal cycle under unscreened exposures with respect to modern screens at each GAR station. Their results indicated that summer averaged (April to September) mean temperatures are biased warm by about 0.4 °C before 1870, with winters (October to March) staying much the same, with the effects on monthly temperature averages largest in June (ranging from 0.21 to 0.93 °C, depending on the station). They also corrected for the slight cooling (up to 0.3 °C for some sites in February). The Böhm *et al.* (2009) study also had to adjust for changes in the way monthly mean temperatures have been calculated across the GAR. Surprisingly, Böhm *et al.* (2009) found that the best method for estimating the true daily and hence monthly average temperature (the mean of each hourly reading) is one of the methods that is most affected by the screen bias.

1.3. The screen bias over the Mediterranean climate

For Mediterranean climates, it has been documented (i.e. Nicholls *et al.*, 1996; Brunet *et al.*, 2004; Bañón *et al.*, 2008) that pre-sheltered temperatures tended to induce a strong warm bias in T_x records, which shows a clear seasonal cycle (higher bias in summer than in winter) compared to current observing practices while T_n readings indicate a small and steady cold bias across the year. When combined this leads to the derived daily mean temperature (T_m) series calculated for the STEV screen being lower than those estimated for the open stand. These results emerged, first, from a long-running (60 years) comparison at Adelaide (Australia), and they support the fact that earlier (for Australia pre-1915) observations are biased warm relative to modern exposures (Nicholls *et al.*, 1996). Similar results are highlighted by Brunet *et al.* (2004) and Bañón *et al.* (2008) through comparisons of dual temperature observations recorded under a replicated MONT shelter and the standard STEV screen installed at La Coruña and Murcia meteorological gardens in the framework of the Spanish-funded project SCREEN (*assessment and minimization of 'screen bias' incorporated into the longest Spanish air temperature records by time changing thermometric exposures throughout dual temperature observation*. CICYT research project: Ref. No. REN2002-0091/CLI). The same authors applied the Spanish data to a preliminary empirical minimization, before undertaking the homogeneity testing of the records, in order to adjust the biased measurements on a monthly basis to conform to modern observations. The first attempt consisted of subtracting from the affected monthly T_x raw values the median of the daily differences series recorded under ancient and new screens (Brunet *et al.*, 2006), since the authors only had 1 year of pair-wise time series. Later, the same authors (Brunet *et al.*, 2007) reassessed their screen bias minimization exercise, after collecting two complete years of simultaneous daily temperature observations. On this occasion, the authors generated two linear regression models, one for each location, using MONT monthly averaged T_x data to predict STEV monthly T_x values. This process enabled them to derive corrections at other stations across Spain, assuming that the two locations were characteristic of the various climate types across the country.

Thus, there is an accumulation of evidence that points to the need to adjust for the screen-biased temperatures. Long-term trend assessments are compromised, either for the area-averaged or for the single-site trend analysis. These problems reduce the data quality and the consistency of the results based on unadjusted data and they affect both the estimation of trends over large and small spatial scales (from global to local) (Parker, 1995). Although the different sign of this bias for T_x and T_n partially cancels in the derived T_m series, the stronger magnitude of the bias in T_x than in T_n series and its seasonal dependence, particularly over subtropical climates, must still be giving negatively biased long-term trends even for T_m series. The opposite sign of the bias will be magnified in DTR.

Following the Brunet *et al.* (2007) study, here we discuss a new and more robust empirical modelling strategy for adjusting screen-biased data, particularly applicable for minimizing the T_x bias from long western Mediterranean temperature records. The approach is based on the analysis of about 6 years of paired observations: the 5-year period (2003–2007) for model estimation and about 1 year (2008) for its validation, using the data recorded at the meteorological gardens of La Coruña and Murcia (Spain).

This article is organized as follows: details of the field trial, the data employed and the results from the quality controls of the data are addressed in Section 2. Section 3 shows results from the exploratory statistical analysis of the screen bias, its relationships with other related meteorological variables and the varying monthly shape of the relationships shown by the chosen predictors. The estimation of the regression model developed for minimizing the screen bias along with the results of the goodness-of-fit of the screen bias of both daily extreme variables (T_x and T_n) is presented in Section 4. Finally, in Section 5, we discuss the advantages of the application of the generated model to adjust screen-biased data on a daily basis, along with the reliability of its application to detect and minimize the screen bias from the affected raw time series at the two experimental sites of La Coruña and Murcia, finishing with some concluding remarks.

2. Experimental approach, data used and results from applying data quality control (QC)

In the framework of the Spanish-funded research project SCREEN, the 19th century MONT shelters were rebuilt, according to the details given in contemporary publications (e.g. Rico Sinobas, 1857; ICM, 1893; Angot, 1903), and installed in April 2003 at the meteorological garden of La Coruña (43°22′02″N/08°25′10″W at 63 m asl) located in the Spanish North-western Atlantic coast and in March 2003 at Murcia (37°58′59″N/01°07′14″W at 57 m asl) situated in a low-land area about 40 km west of the Spanish Southeastern Mediterranean coast. These two locations are representative of two different climatic types and subtypes: the Oceanic/Atlantic/Galician and the Mediterranean arid and semi-arid types, respectively, according to Martín Vide and Olcina's (2001) Spanish climate classification. Both field experiments were and are managed by meteorologists from Agencia Estatal de Meteorología (AEMET, Spanish Met Office) at the regional centres of Galicia (La Coruña station) and Murcia (Guadalupe, Murcia station), where the field trials are still in operation and will continue for the foreseeable future. Figure 1 shows the locations of the experimental sites and details of the reconstructed ancient MONT shelters at both meteorological gardens. Figure 1a shows a location map of the two experimental sites. Figure 1b shows a side view of the MONT shelter together with two STEV screens at the experimental site of Murcia (the closest to the MONT shelter is the screen used in this field trial and the other the official screen used at this meteorological garden), and Figure 1c gives a general view of the installation at La Coruña garden. The instruments inside both types of screens (MONT and STEV) are identical and calibrated sensors connected to two Campbell CR510-2M dataloggers and placed at 1.60 m above the ground at both experimental sites. The nearness of the two screens, where 10-min readings are being recorded, and the identical micro- and topo-climatic conditions surrounding both shelters and observatories at both sites ensure identical measurement conditions. Therefore, the differences found between readings taken in the MONT and STEV protectors will be associated with the influence induced by the type of shelter.

The 10-min observations were, first, subjected to several quality controls. For the field experiment at Murcia, we used 10-min data from March 2003 to June 2008, and in the case of La Coruña, 10-min data from April 2003 to September 2008. We scrutinized the 278 640 ten-minute paired values recorded under MONT and STEV shelters at Murcia; meanwhile for La Coruña 283

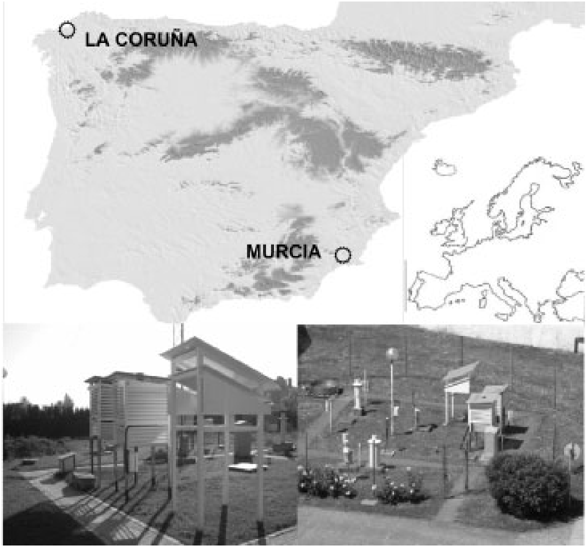

Figure 1. Locations of both experimental sites (a), side view of the MONT and STEV protectors at Murcia (b) and general view of the experimental site at La Coruña (c).

680 paired values were examined. These 10-min values were subjected to the following tests:

- Gross error checks:
  - o Calendar date consistency: no. of readings per day, month and year (aimed at eliminating duplicate readings).
  - o Aberrant values (T_x and T_n values $>50\,^\circ$C and $<-30\,^\circ$C).
- Tolerance tests, such as four or more successive identical 10-min values and readings beyond ±4 standard deviation (σ).
- Temporal coherency: visual inspection of the daily cycle and time of recording of T_x and T_n daily values.

From the 144 ten-minute daily values (from 0 to 24 h), the highest 10-min averaged reading was chosen as the daily T_x for that day and the lowest averaged value as the daily T_n, conforming daily T_x and T_n paired time series for both shelters. From the paired T_x and T_n series at both locations, we calculated the difference (MONT minus STEV) series or the *screen bias* (ΔT_x and ΔT_n) series. These 10-min averages plus the two screen bias series were then subjected to other quality controls. First, we examined those values exceeding $\pm 4\sigma$ from their means. Then, we looked at the coherence of the annual cycle by visually inspecting the monthly and annually resolved boxplots of the two daily paired time series at both locations. For the screen bias series, we labelled as potentially erroneous data those values larger than $\pm 1.70\,^\circ$C. Finally, consecutive daily T_x and T_n values exceeding $15\,^\circ$C were also scrutinized.

After undertaking the QC exercise, from a potential of 278 640 ten-minute paired values for T_x and T_n at Murcia about 10% had to be rejected from further analysis. Most of these values were discarded due to the malfunctioning of the datalogger or for the absence of one of the paired readings (26 731 values). Only 1301 (0.47%) labelled values were spotted as erroneous values and, then, were also not used in further analysis. For La Coruña, from a potential of 283 680 values about 12% of them were discarded, most (11.6%) were also associated with the datalogger misreading or the absence of the paired measurement and only 0.27% (776 values) did not pass the QCs. Therefore, from a potential of 1935 (1970) daily values of the difference series (ΔT_x and ΔT_n) at Murcia (La Coruña), 94.5% (90.4%) of the data were used for Murcia (La Coruña) in the current analysis.

3. Exploratory statistical analysis of the screen bias at both experimental sites

An exploratory statistical analysis is presented first, in order to document the most prominent characteristics of the empirically estimated *screen bias* for both daily extreme parameters at both experimental sites. Bidimensional relationships between *screen bias* (ΔT_x and ΔT_n) series and other potentially explanatory variables are discussed afterwards.

Table I. Percentage of T_x and T_n bias observations exceeding several temperature thresholds during the assessed period.

	La Coruña (%)	Murcia (%)
T_x bias $>0\,^\circ$C	95.6	100
T_x bias $>1\,^\circ$C	34.9	77.0
T_x bias $>1.5\,^\circ$C	10.6	27.5
T_x bias $>2\,^\circ$C	1.7	1.4
T_n bias $<0\,^\circ$C	98.1	93.1
T_n bias $<-0.1\,^\circ$C	81.3	74.2
T_n bias $<-0.25\,^\circ$C	26.3	40.3
T_n bias $<-0.5\,^\circ$C	3.0	8.4
T_n bias $<-1\,^\circ$C	0.5	0.2

The annual evolution of the statistical distribution of ΔT_x and ΔT_n series for La Coruña and Murcia are shown in Figure 2 by means of monthly boxplots, and in Table I, the frequency distribution of the data over/under several temperature limits is provided. Figure 3 shows the mean value of the ΔT_x and ΔT_n over the year at both sites, smoothed by a lowess curve (Cleveland, 1979). From a study of that figures and table, several characteristics should be highlighted:

1. The ΔT_x is predominantly positive across the year at both locations, showing the warm bias of the old stand compared to the modern screen. The bias is on average slightly larger in magnitude and less variable at Murcia than at La Coruña, which points to the role of the different climate types as one of the factors influencing the screen bias. However, comparing the annual cycles of the ΔT_x at both locations, it is also clear that there are stronger similarities than differences between these annual patterns. This reinforces the marked effect that the usage of open stands could have induced in the affected long T_x records for any meteorological station located in Mediterranean and Atlantic climatic types. In short, a positive bias characterizes almost all ($\sim$96%) daily differences in Murcia (La Coruña) and occurrences higher than $1.5\,^\circ$C are found in a 27.5% (10.6%) of the cases for Murcia (La Coruña).
2. The ΔT_n is mostly negative (98.1% and 93.1% of the cases in La Coruña and Murcia, respectively), but of much lower magnitude and variability than the T_x bias. About 81% and 74% of the negative differences are lower than $0.1\,^\circ$C and only 3% and 8.4% of T_n biases are beyond the 0 and $-0.5\,^\circ$C limits for La Coruña and Murcia, respectively.
3. A discernible seasonal cycle is detected only for the ΔT_x series, with higher positive differences during summer than in winter. For ΔT_n, the annual cycle is negligible in Murcia and does not appear in La Coruña, at all. Also, the variability reflected by the interquartile range of the differences does not show seasonal cycle in any of the two variables and two locations.

Relationships between the screen bias and other potential explanatory variables are explored and described

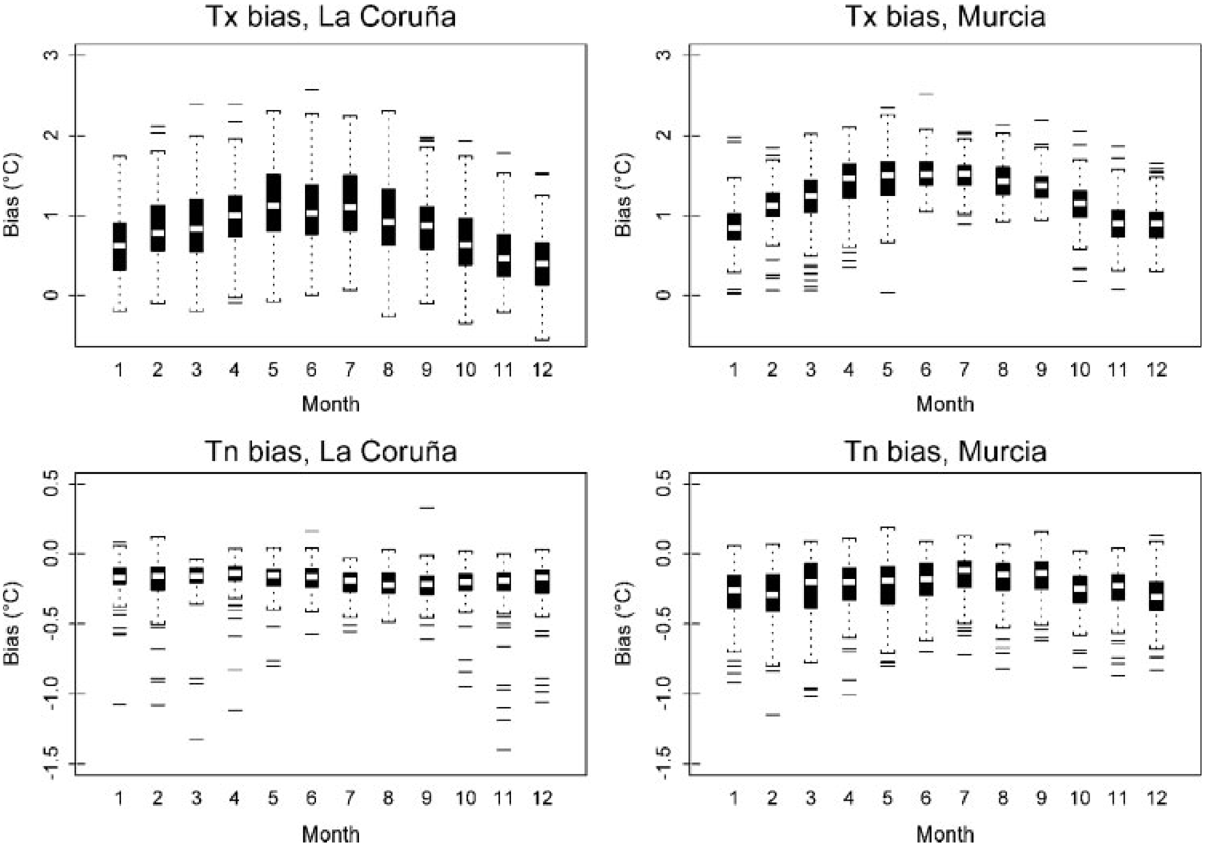

Figure 2. Monthly boxplots of the T_x and T_n screen bias series for La Coruña and Murcia. Note the difference in scale between the T_x and T_n plots.

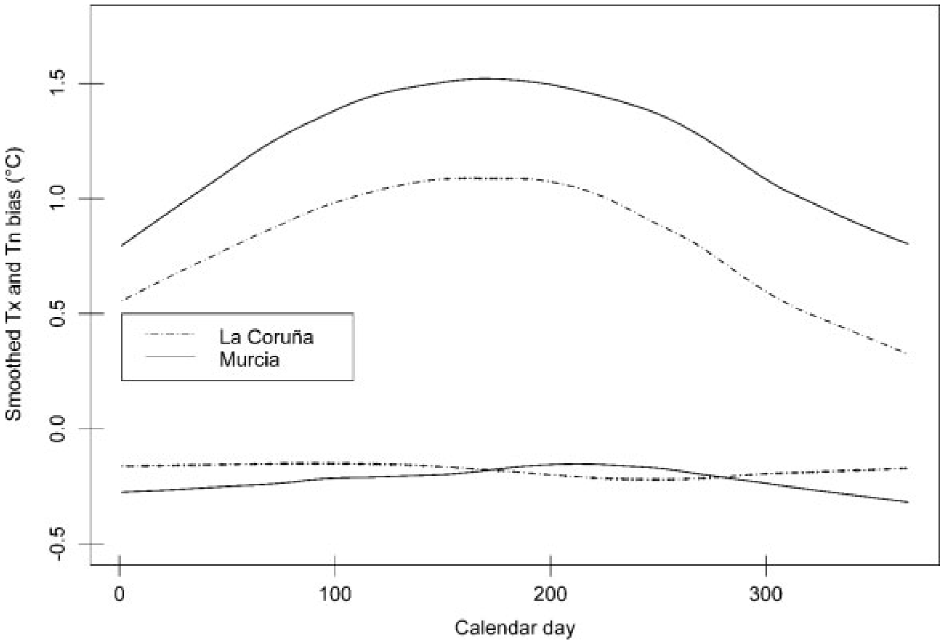

Figure 3. Annual evolution of T_x (upper curves) and T_n (lower curves) screen bias series, smoothed with Lowess filter (Cleveland, 1979) for La Coruña and Murcia.

here. The aim of this analysis is to examine the relationships between both ΔT_x and ΔT_n and those variables that are potential predictors for the definition of a statistical regression model with which to minimize the *screen bias* after its application to other climatically related and affected T_x series. The correlation analysis between the ΔT_x and ΔT_n series and other related climate variables records, such as MONT and STEV daily T_x, T_n and

Table II. Pearson r (Spearman Rho in brackets) correlation coefficients between daily ΔT_x and ΔT_n series and several potentially explanatory climatic variables recorded at both experimental sites.

	Murcia		La Coruña	
	ΔT_x	ΔT_n	ΔT_x	ΔT_n
MONT T_x	**0.541 (0.524)**	**0.127 (0.148)**	**0.334 (0.345)**	−0.036 **(−0.180)**
MONT T_n	**0.469 (0.490)**	**0.259 (0.271)**	**0.175 (0.190)**	−0.023 **(−0.144)**
MONT DTR	**0.248 (0.124)**	**−0.224 (−0.202)**	**0.336 (0.361)**	−0.030 **(−0.105)**
STEV T_x	**0.503 (0.490)**	**0.127 (0.148)**	**0.221 (0.245)**	−0.034 **(−0.172)**
STEV T_n	**0.471 (0.492)**	**0.231 (0.244)**	**0.176 (0.191)**	0.014 **(−0.170)**
STEV DTR	**0.150** (0.037)	**−0.183 (−0.163)**	**0.128 (0.172)**	0.038 (−0.039)
Daily sunshine	**0.382 (0.345)**	*0.064* **(0.130)**	**0.296 (0.310)**	*0.068* (0.012)
Pressure 00	**−0.157 (−0.185)**	**0.078 (0.064)**	**−0.108 (−0.141)**	**0.097** (*0.062)*
Pressure 07	**−0.145 (−0.181)**	**0.075 (0.068)**	**−0.071 (−0.106)**	**0.085** (0.053)
Pressure 13	**−0.173 (−0.196)**	**0.072 (0.067)**	−0.025 (*−0.063*)	**0.070** (0.040)
Pressure 18	**−0.205 (−0.224)**	*0.051* (0.045)	−0.007 (−0.045)	0.053 (−0.023)
Precipitation	**−0.207 (−0.140)**	**0.075** (0.044)	**−0.191 (−0.104)**	**−0.133 (−0.103)**
Daily radiation	**0.237 0.199)**	0.043 (0.069)		
Cloudiness 07	**−0.156 (−0.131)**	−0.008 (−0.039)	−0.016 (−0.051)	*−0.061* (−0.022)
Cloudiness 13	**−0.088 (−0.058)**	*−0.061* ((**−0.082)**	**−0.113 (−0.160)**	**−0.077** (−0.037)
Cloudiness 18	**−0.110 (−0.074)**	−0.025 (−0.038)	−0.042 (*−0.068*)	**−0.082** (−0.047)
Daily cloudiness	**−0.139 (−0.091)**	−0.036 (*−0.065*)	*−0.066* (**−0.128**)	**−0.085** (−0.046)
Daily wind speed	**0.099 (0.197)**	0.034 (**0.93**)	**−0.165 (−0.157)**	**−0.102** (*−0 .071*)

Bold (italic) indicates significance at the 1% (5%) confidence level.

diurnal temperature range (DTR) series, daily sunshine, hourly air pressure, daily precipitation amount, daily radiation (only available for the Murcia site), daily and hourly cloudiness amount and daily averaged wind speed series, have been carried out and their results are shown in Table II for both experimental sites. This gives both the calculated Pearson r and Spearman Rho (in brackets) correlation coefficients, together with their statistical significance at the 1% (bold) and 5% (italic) confidence levels.

As expected and shown in Table II, the best and highest relationships between the ΔT_x and the other related climate variables are with T_x series taken under the MONT shelter, and to a slightly lesser extent under the STEV screen at both sites, although relationships with MONT DTR series, particularly at La Coruña and also at Murcia, are remarkable. There are also significant relationships between ΔT_x and daily sunshine duration series at both sites, although lower than for DTR at La Coruña. Other relationships are much weaker, although some are statistically significant, but with little physical meaning. The results of the correlation analysis imply the potential of MONT T_x and DTR series to explain a satisfactory fraction of the ΔT_x variability.

We have also explored the monthly shape of the relationships between the ΔT_x and MONT T_x and DTR series in order to identify possible changes in the type and shape of the functional relationships throughout the year. Figure 4 shows the shape of these relationships smoothed by a lowess filter for April (upper plots) and September (lower plots). From inspection, there is a clear change in the shape of these relationships throughout the year. The ΔT_x and MONT T_x relationships at both locations indicate a shift in the shape between April and September (right top and bottom panels in Figure 4); meanwhile for ΔT_x and MONT DTR relationships, April shows the greatest variation (left top and bottom panels in Figure 4). This finding indicates the need to take into account in the regression model the interaction of the annual cycles of the predictor variables (MONT T_x and DTR) as well. Besides, and not less important, there is also a remarkably good agreement in the linear shape of these monthly relationships when DTR values are in the same range at both experimental sites; whereas for the higher values of DTR, the behaviour is not linear. The common behaviour of the monthly shapes of the relationships between ΔT_x and MONT T_x together with similar DTR effects on the ΔT_x at both locations leads us to consider one unique functional relationship between the response and these explanatory variables and, at the same time, it justifies the generation of a statistical model where the covariates and relationships with ΔT_x are expressed by one unique equation.

Therefore, relying first on the best relationships found between the ΔT_x and the related MONT temperature series at both locations and, second, on the previously mentioned caution of using other variables for adjusting temperature series, together with the rare availability of homogenous century-long time series for other distinct variables than temperature itself, the screen bias minimization approach presented and discussed in this study will make use of the MONT T_x and DTR series at both locations to generate a general regression model that can be applied to adjust the biased data from climatically related stations to both experimental sites.

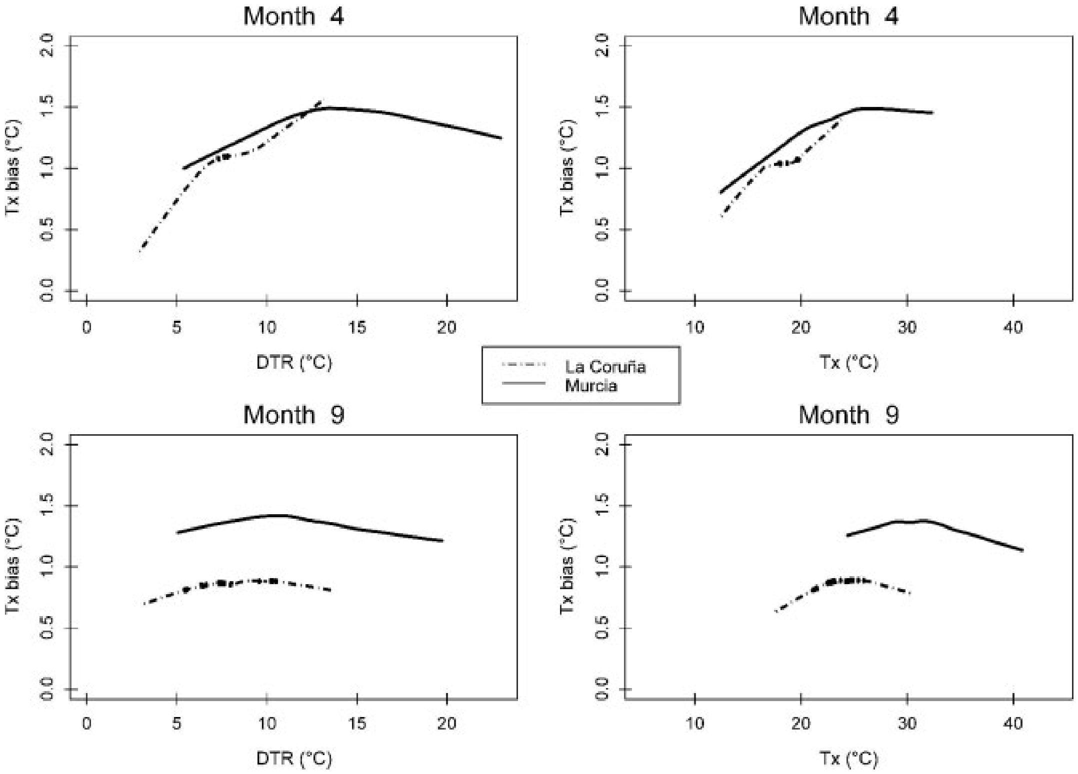

Figure 4. Lowess smoothed mean value of ΔT_x *versus* DTR (left panels) and T_x (right panels) for April and September showing the shape of these relationships.

As evident in Table II, the relationships between ΔT_n and the other climate variables are much weaker than those for ΔT_x. The relationships are barely significant and close to zero at La Coruña and are very small in magnitude at Murcia for most of the related variables, including the temperature series taken under both MONT and STEV shelters and the daily averaged wind speed at both sites. Only for ΔT_n in Murcia, there is a weak relationship to MONT and STEV T_n data and inversely related to the MONT DTR values. The monthly shapes of the ΔT_n relationships with its covariates (MONT T_n and DTR) also show weaker relationships than for ΔT_x and its covariates and there is little change throughout the year between ΔT_n and its covariates at La Coruña, while only Murcia indicates a slight change between April and September (not shown).

Although we will describe the results of the same approach as ΔT_x, for ΔT_n minimization using a single regression model based only on Murcia data, it is not advisable to adjust the slightly affected T_n data with the estimated model as it does not explain any additional variance than the mean of the ΔT_n. This aspect will be further discussed in Section 5.

4. The estimation of regression models for minimizing the screen bias targeting the affected ancient Western Mediterranean records

After carrying out the exploratory statistical analysis, here we describe the estimation of the general regression model to minimize ΔT_x from the biased old Western Mediterranean records. Our model takes into account data from both experimental sites and the potential predictor variables are time of the year, MONT T_x, MONT DTR, and some transformations. Our regression model is shown in Equation (1),

$$\hat{Y}(t) = \alpha_0 + \sum_{i=1}^{3} \left(\begin{array}{c} \beta_i + \sum\limits_{j=1}^{3} \beta_{ij}^{s} \sin(2\pi jt/366) \\ +\beta_{ij}^{c} \cos(2\pi jt/366) \end{array} \right) \mathrm{DTR}^{i} + \sum_{i=1}^{3} \left(\begin{array}{c} \gamma_i + \sum\limits_{j=1}^{3} \gamma_{ij}^{s} \sin(2\pi jt/366) \\ +\gamma_{ij}^{c} \cos(2\pi jt/366) \end{array} \right) T_x^{i} \quad (1)$$

where $\hat{Y}(t)$ is the expected value of the screen bias for a day t in the year, which depends upon an expression that includes the linear and quadratic terms of MONT T_x (represented by γ coefficient) and MONT DTR (represented by β coefficient) and whose coefficients for each day t are given by an expression that includes interactions with Fourier harmonics up to order three, which represent the change of the relationships throughout the year. The model includes the possibility that non-linear relationships are also captured by using polynomial expressions for both predictors.

The same general model can be applied for both the minimization of ΔT_x and ΔT_n through substituting the term T_x by T_n, but as has already been stated, the application of the model to adjust ΔT_n is no better than employing a constant ΔT_n correction factor.

The modelling process is aimed at generating a parsimonious model that includes only those statistically significant effects, in order to avoid the development of unnecessary complex models. The significance of each term is determined using an F test. When considering their interaction with harmonic terms, the F test can express the joint significance of the two terms corresponding to the sine and cosine components. Therefore, if the power of a covariate is significant, we include in the model the power of lower order. Also, as ΔT_x variability is higher at La Coruña than at Murcia, we developed a weighted regression equation, where the La Coruña and Murcia's responses have associated a weight of a 1/2 and 1, respectively.

Finally, we have also estimated two single models based on each experimental site's data for minimizing the ΔT_x, in order to compare both the goodness of the fit and the results estimated by applying either the general model or the single-site-based model, estimated for each location (La Coruña and Murcia), separately.

4.1. Results for maximum temperature *screen bias* (ΔT_x)

Table III shows the results for the selected general model, which includes 12 parameters related to terms with significant effect and gives the estimated coefficients, the p-values of the t-test, and the lower and upper limits at the 95% confidence interval. We have also calculated same coefficients for the regression models separately estimated for La Coruña and Murcia (not shown). A comparison between results from the site-based models and those shown in Table III indicates that the general model has similar signs and magnitudes of coefficients as those estimated by the site-based models. The structures of the covariates showing significant terms coincide between the general and the individual models, as well as the shapes of the polynomial terms (not shown).

The effect of the T_x data for both the general and single-site models is of quadratic order, interacting with the Fourier harmonic of first order, whereas the DTR's effect is cubic with a linear term that varies across the year. The DTR and T_x effects over the fitted values of the ΔT_x are shown in Figure 5 for the calendar days 30, 120, 210 and 300, corresponding to late January, April, July and October. Notice DTR effect is up to three times larger than T_x effect on the adjusted ΔT_x values. Besides, the effect of DTR is different in the four cases/days, although similar increasing shapes of the DTR values up to the range of 10 °C and a stabilization of the shape for the upper values are evident from Figure 5 top panels. There are even larger differences in the shape of T_x effect among the four cases (Figure 5 bottom panels), as it is almost linear on the calendar day 300 (late October) and not linear at all on day 120 (late April). The comparison of these results with the ones shown in Figure 4 is interesting, as it proves that the statistical model represents well the relationships with the covariates. In this regard, the variation that the general model captures is reflected in the fitted values of ΔT_x. In the case, for instance, that T_x is equal to 20 °C and DTR equals to 10 °C fitted adjustments of 1.04, 1.39, 1.28 and 0.93 °C values will be returned for the calendar days 30, 120, 210 and 300 depicted in Figure 5, respectively.

Table IV gives the R^2 and the residual standard deviation (s) values for ΔT_x estimated for each single experimental site and for both sites together for both the general and the site-based models during the estimation (2003–2007) and the validation (2008) periods. The goodness of the adjusted general model is also proven by the 49.7% (30.3%) of ΔT_x-explained variability for Murcia (La Coruña) and 51.7% for both sites together during the estimation period. Although the explained variability is slightly better at each individual site when estimated by the site-based models than by the general model, the joint explained variability for both sites with the general model is better than that explained for each site by the site-based models. This result also reinforces the simplicity of choosing and applying the general regression model estimated to minimize the biased T_x time series recorded under similar climate types. On

Table III. Estimated coefficients, p-values and the lower and upper limits of the 95% confidence interval for the ΔT_x from the general model of Equation (1), which includes 12 parameters related to terms with significant effect on the screen bias estimates.

Term	Estimated coefficient	p-value t-test	Lower limit at 95% confidence interval	Upper limit at 95% confidence interval
Intercept	−0.7611	0	−1.0184	−0.5037
MONT T_x	0.0311	0.007	0.0085	0.0537
Sin 1: MONT T_x	0.0123	0	0.0076	0.017
Cos 1: MONT T_x	−0.0085	0.008	−0.0148	−0.0022
MONT T_{x^2}	−0.0005	0.065	−0.001	0
Sin 1: MONT T_{x^2}	−0.0005	0	−0.0006	−0.0003
Cos 1: MONT T_{x^2}	0.0002	0.127	0	0.0004
MONT DTR	0.3015	0	0.2626	0.3404
Sin 1: MONT DTR	0.0102	0	0.0049	0.0154
Cos 1: MONT DTR	−0.0107	0	−0.0165	−0.005
MONT DTR^2	−0.0193	0	−0.0228	−0.0158
MONT DTR^3	0.0004	0	0.0003	0.0005

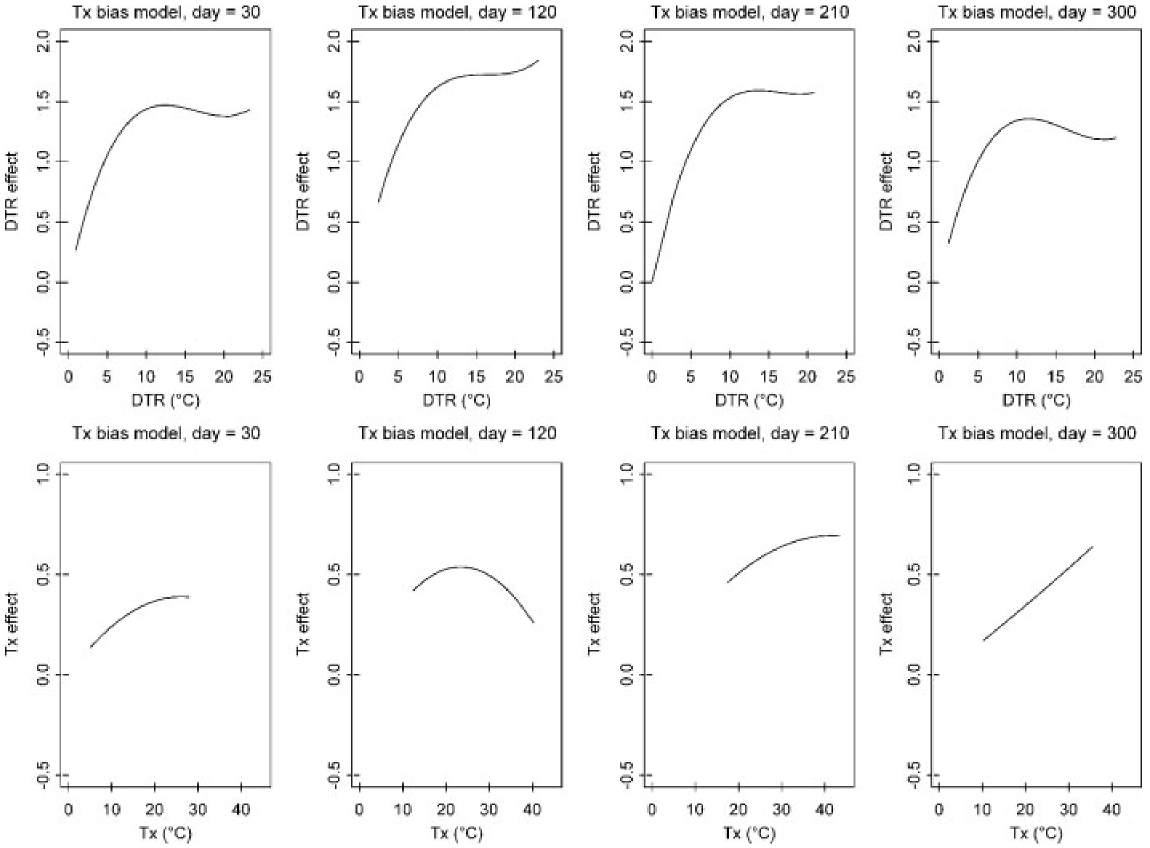

Figure 5. Effects of DTR (top panels) and T_x (bottom panels) in the selected model for the ΔT_x assessed for the calendar days 30, 120, 210 and 300.

Table IV. R^2 and residual standard deviation (s) estimated by the ΔT_x general and individual models for each single experimental site and for both sites together during the estimation (2003–2007) and the validation (2008) periods.

	Estimation period 2003–2007		Validation period 2008	
	R^2 (%)	s (°C)	R^2 (%)	s (°C)
General model				
Both	51.7	–	38.7	–
Murcia	49.7	0.258	43.6	0.264
La Coruña	30.3	0.424	29.9	0.408
Site-based models				
Murcia	51.0	0.255	42.3	0.265
La Coruña	38.7	0.408	38.9	0.384

the other hand, the better goodness-of-fit at the Murcia experimental site is related to both the more marked annual cycle and the lower variability shown by Murcia's ΔT_x when compared with those estimated for La Coruña (Figure 2).

Finally, Figure 6 compares the ΔT_x monthly mean fitted values returned by the general model (Equation (1)) for both the estimation (2003–2007) and the validation (2008) periods. As shown in this figure, the degree of the reproducibility of the fitted ΔT_x values is reasonably good when compared with the estimation period (left plot in Figure 6) and the validation period (right plot in Figure 6) values. It also indicates the goodness of the generated regression model based on a unique equation that explains the ΔT_x responses at both sites. Moreover, the general model also has an advantage over the individual models by incorporating the slightly different ΔT_x patterns related to the different climate types where the experimental sites are located. Therefore, it integrates both responses in one unique model and, then, other locations with similarly affected records within the two climate types can be confidently adjusted by applying the regression Equation (1).

4.2. Results for minimum temperature *screen bias* (ΔT_n)

As stated in the exploratory statistical analysis (Section 3) describing the screen bias on minimum temperatures (ΔT_n), there is a markedly different behaviour than the one characterizing the ΔT_x. First, its magnitude is notably lower than for ΔT_x. Second, it hardly shows any seasonality at the Murcia site and there is no seasonality at all at La Coruña (Figures 2 and 3). Table V, as Table III, shows the estimated coefficients, the p-valuesof the t-test and the lower/upper 95% confidence intervals, but for the ΔT_n and its covariates. Although some terms indicate significant seasonal changes in the effect of the covariates and their interaction with Fourier harmonics on the ΔT_n across the year, they are much lower and less marked than in the case of ΔT_x.

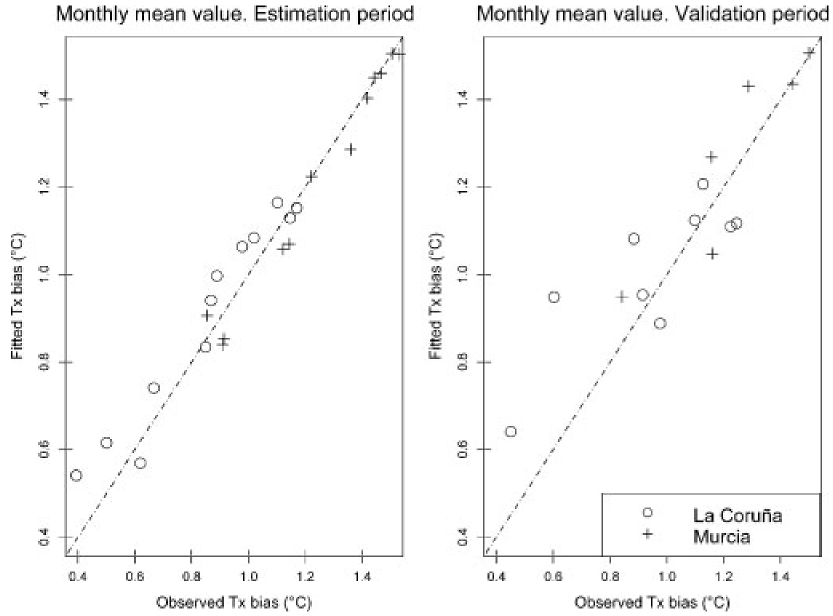

Figure 6. Monthly mean values of the observed and fitted ΔT_{x} for La Coruña and Murcia for the estimation period (left panel) and the validation period (right panel).

Table V. As Table III, but for ΔT_{n}.

Term	Estimated coefficient	p-value t-test	Lower limit at 95% confidence interval	Upper limit at 95% confidence interval
Intercept	−0.056	0.045	−0.1106	−0.0013
MONT T_{n}	−0.0246	0	−0.0317	−0.0175
Sin 1: MONT T_{n}	0.0048	0.005	0.0014	0.0082
Cos 1: MONT T_{n}	−0.0009	0.593	−0.0042	0.0024
MONT $T_{\mathrm{n}2}$	0.0011	0	0.0008	0.0014
Sin 1: MONT $T_{\mathrm{n}2}$	−0.0001	0.208	−0.0003	0.0001
Cos 1: MONT $T_{\mathrm{n}2}$	0.0003	0.006	0.0001	0.0005
MONT DTR	0.0036	0.261	−0.0027	0.0098
Sin 1: MONT DTR	−0.0026	0.002	−0.0043	−0.001
Cos 1: MONT DTR	−0.0071	0	−0.0091	−0.0051
MONT DTR^2	−0.0006	0	−0.0009	−0.0003

In Figure 7, the MONT DTR (top panels) and T_{n} (bottom panels) effects on the ΔT_{n} are shown for the days 30, 120, 210 and 300, as done for T_{x}. First, it is evident that both effects are quadratic, but of much lower values than the ones found for the ΔT_{x} as shown in Figure 5 (notice the big change in the intervals of the y-axis in Figure 7 compared to Figure 5). Although some weak variation can be identified, the effects of the covariates hardly change throughout the year. Here we can see that the DTR effect for calendar day 30, which indicates that with higher values of DTR there are higher negative ΔT_{n} values, is different from the effect on calendar day 210, which shows no identifiable change within lower and higher values of the DTR. Still there is less variation when looking at the MONT T_{n} effect on the fitted ΔT_{n}.

In this context, the fitted model for ΔT_{n} at both experimental sites only represents the mean value of the response at La Coruña, without explaining any of the variability, whereas only 14% of the ΔT_{n} variability is explained for Murcia (see Table VI results). Therefore, the application of the regression model in order to minimize the ΔT_{n} from likely biased T_{n} records is neither meaningful nor recommended. Instead of it, the application of any other constant correction factors is advisable.

Table VI. As Table IV, but for ΔT_{n}.

	Estimation period 2003–2007		Validation period 2008	
	R^2 (%)	s (°C)	R^2 (%)	s (°C)
Both	9.6	–	0	–
Murcia	14	0.17	7	0.191
La Coruña	0	0.146	0	0.191

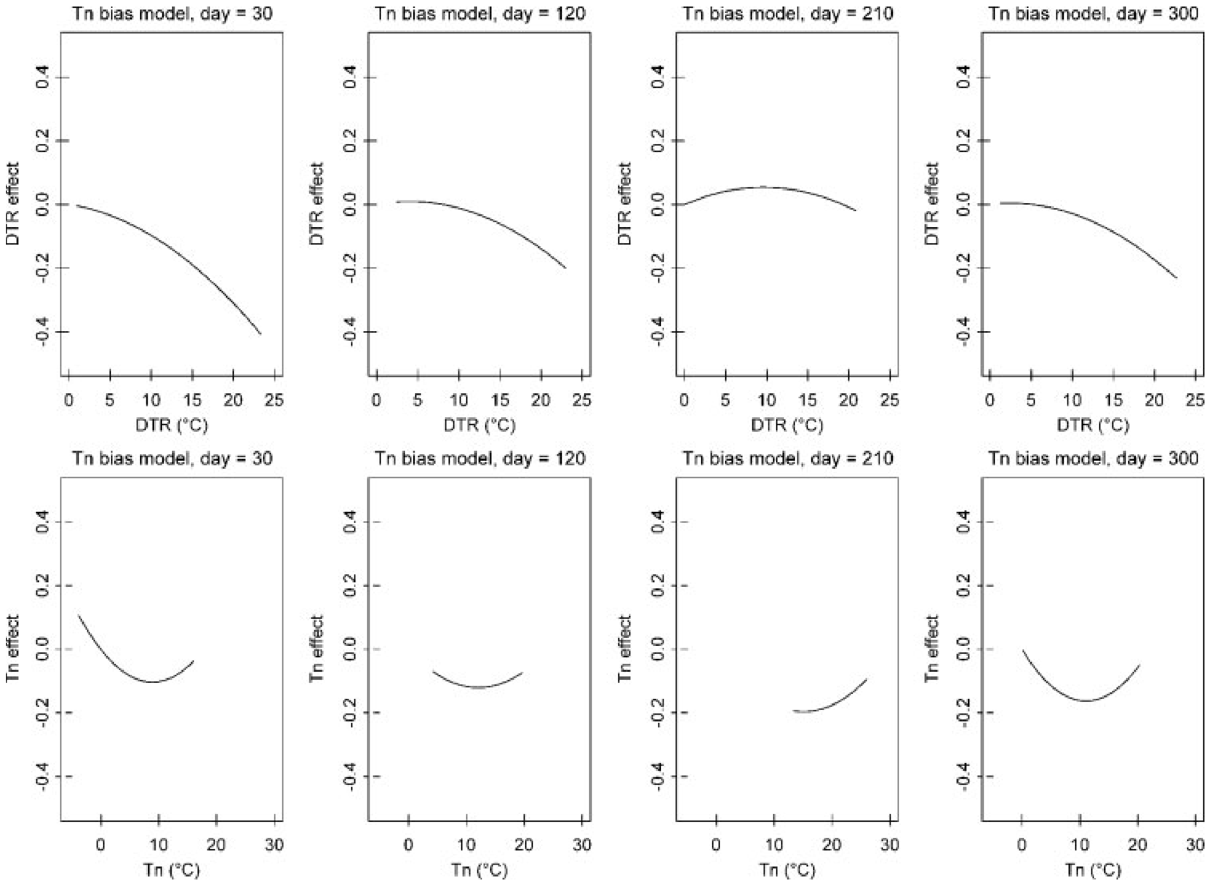

Figure 7. As Figure 5, but for the effects of DTR (top panels) and T_n (bottom panels) on ΔT_n. Note, however, the change in scale of this plot compared with the effects for T_x shown in Figure 5 due to the smaller magnitude of T_n bias.

4.3. An application of the regression model to minimize the screen bias from the historical La Coruña and Murcia time series

Here we show and discuss the results of the application of the general model to the likely biased T_x original historical time series at both stations (La Coruña and Murcia) and compare the adjustments applied with those experimentally estimated at both sites. This can be considered as additional validation.

According to the available metadata, temperatures were measured in the past under free-standing open stands at Murcia and La Coruña from 1863 to 1911 and from 1882 to 1911, respectively, when the old stands were substituted by the standard Stevenson screen during the same year at both stations.

We first assess possible problems of extrapolation when applying the weighted regression equation to the likely affected 19th and early 20th centuries time series at both stations. As the regression equation (1) is a function of MONT T_x and DTR data, we explore the range of values of these covariates during both the experimental (2003–2007) and the historical records periods at both locations. We identify and compare the highest and lowest values of the daily MONT T_x and DTR series at both the experimental and the instrumental periods. For Murcia, during the 1863–1911 period only 11 days recorded T_x values out of the range of the observed during the 6-year experimental period (2003–2007): five exceeded the highest value (43.4 °C) and six below the lowest value (5.2 °C), whereas for the DTR series only 36 days were beyond the DTR range of the experimental period: 29 exceeded the highest DTR (23.6 °C) and 7 the lowest value (1.1 °C). For La Coruña and during the targeted 1882–1911 period, only 16 values were out of the range of the lowest monthly T_x value (7.2 °C) and none exceeded the highest value (35.3 °C), meanwhile for the DTR series just 33 days exceeded the highest DTR (17.5 °C) recorded during the experimental period and, of course, none exceeded the lowest monthly DTR value (0°) observed at this experimental site. Therefore, these results prove the application of the general model to adjust the likely biased 19th and 20th centuries T_x data at both locations requires little data extrapolation, which gives confidence on the applicability of the generated model to other historical similarly screen-biased time series.

Figure 8 shows histograms given the fitted ΔT_x adjustments returned by the application of the general model to Murcia (top panels) and La Coruña (bottom panels) data, in order to compare the goodness-of-fit over the experimental period (left panels) and the historical period (right panels). As evident from this figure, each station has a similar shape of distribution between both periods: the experimental one (2003–2007) and the historical one (i.e.

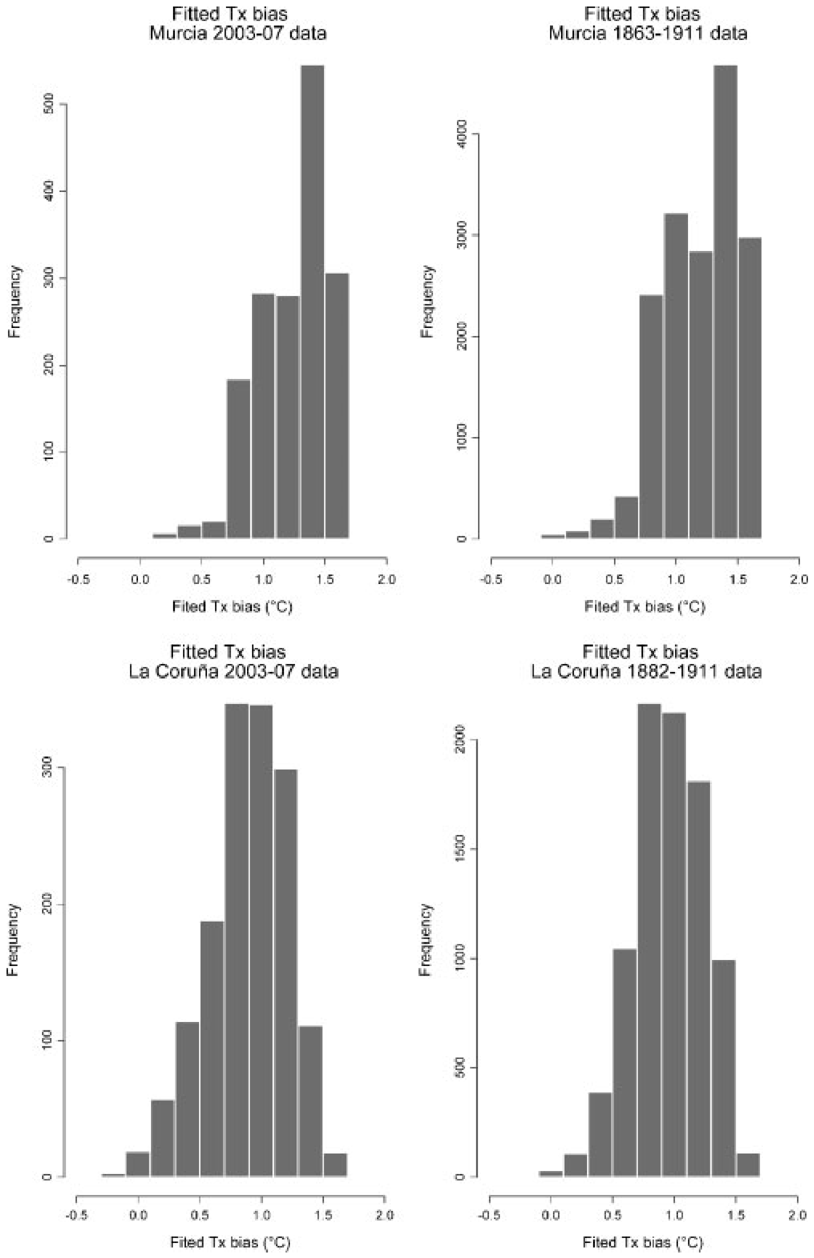

Figure 8. Histogram showing the distribution of the fitted ΔT_{x} at Murcia for the experimental period (top left panel) and the historical period (top right panel). Same information for La Coruña for the experimental period (bottom left panel) and the historical period (bottom right panel).

1863–1911 for Murcia and 1882–1911 for La Coruña). Only 25 out of 10 958 (128 out of 17 897) daily adjusted biases in the historical period for La Coruña (Murcia) are beyond the range of the adjusted biases during the experimental period. This also confirms the credibility of the fitted ΔT_{x} adjustments for these two stations.

Finally, Figure 9 shows three-dimensional images of the surface shapes of the daily fitted ΔT_{x} values (vertical axes) estimated by the general model in order to adjust, and then to minimize, the *screen bias* from the historical records of Murcia (top panel) and La Coruña (bottom panel), together with its functional relationships with T_{x} and DTR data (horizontal axes). Both panels reveal the different shapes of the ΔT_{x} adjustments required to minimize the screen bias from both records, which indicate the performance of the empirically estimated regression model as it differentially captures and distinguishes the required ΔT_{x} adjustments for each historical record.

These results imply that it is possible to generate adequate adjustments in order to minimize ΔT_{x} from the affected time series of La Coruña and Murcia and, by extension, we are confident in the reliability of its application to adjust other affected and climatically related time series, as discussed in the last section.

5. Discussion and concluding remarks

The ΔT_{x} fitted general model is satisfactory from several points of view. First, the inclusion of the whole range of observed ΔT_{x} values at both locations reinforces the

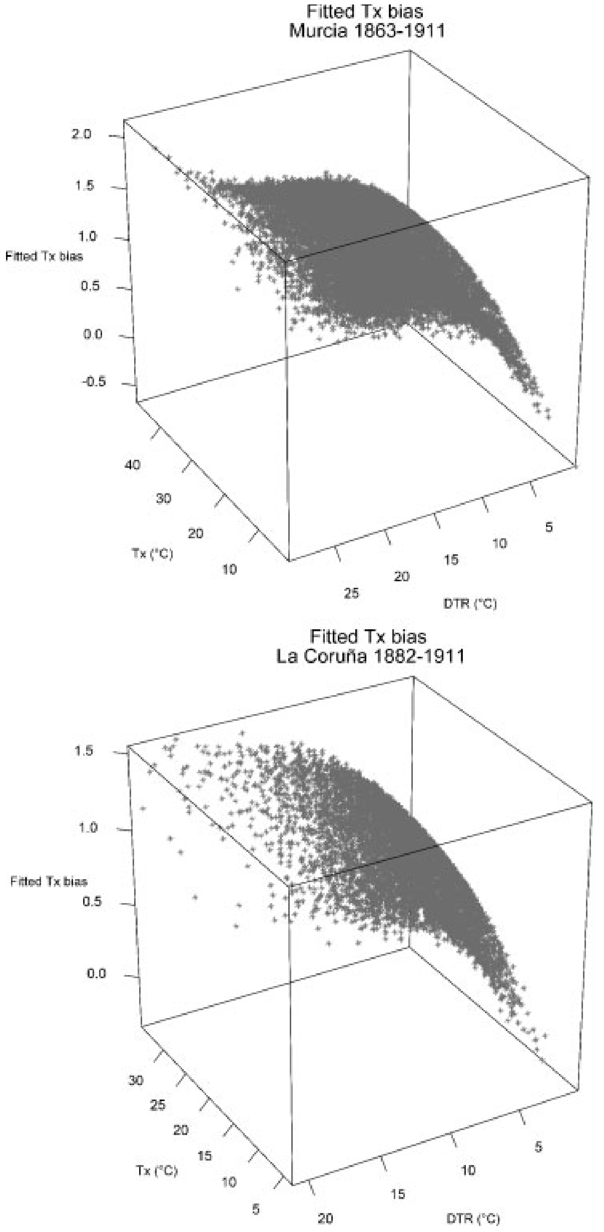

Figure 9. Three-dimensional plots giving the surface shape of the ΔT_x adjustments fitted by the regression equation (1) at Murcia (upper plot) and La Coruña (lower plot) for the periods of 1863–1911 and 1882–1911, respectively.

reliability of its application to other affected time series recorded at other locations representative of climate types similar to the Oceanic/Atlantic/Galician and the Mediterranean arid and semi-arid types, without overestimating the adjustments. In this regard, the general model, with respect to the site-based models, has the advantage of including the whole range of experimentally observed ΔT_x at the two locations (La Coruña and Murcia) without a significant loss of fit. As already mentioned, the structure of the covariates, the shape of the polynomial terms and their coefficients are coincident, of the same sign, and of similar magnitudes among the three models developed. This point gives us confidence on the application of the generated regression equation to minimize the screen bias from other affected and climatically related time series.

Second, the developed model is aimed at adjusting the ΔT_x on a daily basis, which will improve the Brunet *et al.* (2007) adjustments made on a monthly basis. In this regard, the residual standard deviation of the fitted adjustments estimated by the general model is 0.26 °C (0.42 °C) at Murcia (La Coruña), which is larger at La Coruña than at Murcia due to the slightly different ΔT_x magnitudes, variability and annual patterns at both locations (Table IV and Figure 2). However, when the adjustments are just estimated on a monthly basis, the associated error ranges are: 0.28 and 0.45 °C for both locations, which also proves the better adjustments produced by the regression model discussed here.

Third, the flexibility of the model enables us to fit different daily adjustments that are dependent on the day of the year and the actual temperature to be adjusted. In Figure 10 (top panels), the results of the fitting by applying the general model to adjust ΔT_x at the daily (small circles) and monthly (thick slashes) scales are compared for both La Coruña (right top panel) and Murcia (left top panel). From the upper panels of this figure is evident that better fitted values are resolved at the daily scale than at the monthly resolution. The first fit preserves the observed daily ΔT_x variability at both locations, which can only be gained by from a model that distinguishes different situations that cannot be reproduced by monthly adjustment approaches. In the bottom panels of Figure 10, we compare the ΔT_x adjustments from the model application (triangles) and the ΔT_x experimentally observed (small circles) at Murcia (left bottom panel) and La Coruña (right bottom panel), showing again the good agreement. This also highlights the robustness of the procedure developed in this study as the adjustments preserve the different ΔT_x correction patterns (magnitude, amplitude and variability of the adjustments) required for each experimental site without any overestimation of the minimization of the screen bias. This result improves the application of earlier 'easy-correction methods' (e.g. like that applied in the Brunet *et al.*, 2006, approach), contributing to enhancing long-term daily data quality and homogeneity. However, it is also worth developing a similarly robust regression model for adjusting ΔT_x from records only available on a monthly basis, which also takes into account similar covariates and interactions with Fourier harmonics to incorporate the annual cycle. For many regions, unfortunately, there are still lots of potentially screen-biased records that are currently only available on a monthly basis. This would deserve further attention.

Fourth, we have estimated an equation for minimizing the ΔT_x that is not connected to any parameter of any specific observatory, which makes the application of the model easier for any other climatically related station with biased records.

And fifth, the fitted equation to minimize ΔT_x from other climatically related and affected temperature records is simple, as it only contains polynomial terms

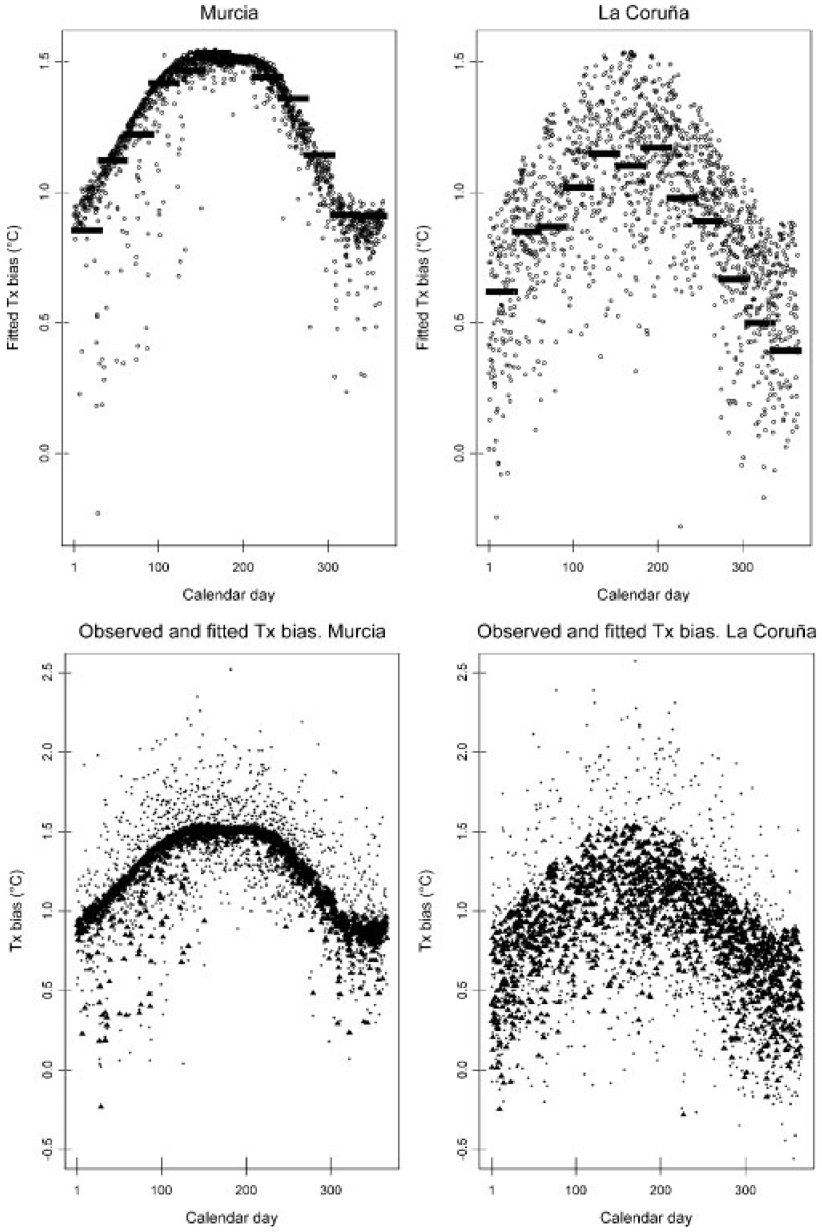

Figure 10. The daily fitted ΔT_x adjustments (small circles) and estimated monthly based corrections (thick slashes) at La Coruña (left upper panel) and Murcia (right upper panel), together with the fitted (triangles) and observed (small circles) ΔT_x values at La Coruña (right lower panel) and Murcia (left lower panel).

of lower order in the temperature variables and harmonic terms in order to represent the seasonal cycle. In addition, the equation estimated from the data recorded at the two experimental sites also has an advantage over the integration of the two slightly different ΔT_x responses, which are climatically dependent. This will also ensure the corrections applied to any biased time series recorded under both climate types (Oceanic/Atlantic/Galician and Mediterranean arid and semi-arid) could not be overestimated as their potentially biased data will not be out of range of their covariates. The opposite, an underestimation of the adjustments, could become true (see bottom panels of Figure 10 where the magnitude of the underestimation is clearly shown). Thus, the application of the current regression model will not ensure the total adjustment of the ΔT_x from other biased records, but it will ensure, at least, its partial minimization.

It is clear that the applicability of any regression model becomes limited by extrapolation problems (i.e. cases in which the covariate values can be out of the range of the used in our statistical approach), this problem will likely be minimized when applying the equation to the biased time series measured under both climate types, and also the differences in the daily and annual temperature cycles of the covariates, as well as in the ΔT_x daily and seasonal cycles, are much less pronounced than their similarities.

The minimization of ΔT_n by applying the equation developed here is more questionable for a number of reasons. First, ΔT_n is of a very low magnitude and does not exhibit any seasonality at all, particularly at

La Coruña. Second, it shows weak relationships with the potential predictor variables (MONT T_n and DTR data) and other related climate variables, which together with the first point makes the application of a unique regression model that takes into account polynomial terms of lower order less general. Fewer harmonic terms are necessary for representing the seasonal cycle, which is almost absent in ΔT_n. Third, for minimizing ΔT_n from the affected time series, any other 'easy-correction method' based on subtracting the bias at a monthly scale and only taking into account the minimization of monthly means of ΔT_n would be advisable.

Finally, and based on the above reasoning, the authors are persuaded that the application of the model discussed here to other affected records, for similar climatic regions, will contribute to improve the state-of-the-art in homogenization techniques and to enhance the availability of high-quality and homogeneous temperature records with which to undertake more robust analysis of climate variability and change. In this regard, the approach presented here enables homogenization of the biased Spanish temperature records all the way back to the late 1700s enhancing both currently available long temperature time series (as those included in the Spanish Daily Adjusted Temperature Series developed by Brunet *et al.* (2006)) an also Spanish records that are being or will be developed by different research groups. The use of more reliable temperature time series will improve climate change detection and prediction assessments over Spain.

Acknowledgements

We would like to thank the authorities and staff of the AEMET at both regional centres (Galicia and Murcia regions) and meteorological gardens (La Coruña and Murcia-Guadalupe) for their kindness by allowing the authors to install and operate the MONT replicas in the ongoing experimental project once the research project grant that made the field experiment possible is over. We also thank Susan Osborne (NOAA's NCDC) for her review. This work was supported by: the EU-funded EMULATE project (EVK2-CT-2002-00161), the Spanish-funded SCREEN project (Comisión Interministerial de Ciencia y Tecnología: REN2002-0091/CLI), the CLICAL and ACCEBRO projects (Comisión Interministerial de Ciencia y Tecnología: CGL2006-13327-C04-03/CLI and CGL2006-02485/CLI) and the CAFID-EXPI project (Comisión Interministerial de Ciencia y Tecnología: CGL2007-65546-C03-02). Phil Jones has been supported by the US Department of Energy (grant DE-FG02-98ER62601)

References

Aguilar E, Auer I, Brunet M, Peterson TC, Wieringa J. 2003. *Guidelines on Climate Metadata and Homogenization*. World Meteorological Organization (WMO) TD No. 1186/World Climate Data Monitoring Program (WCDMP) No. 53, Geneva; 52.

Andersson T, Mattisson I. 1991. *A Field Test of Thermometer Screens*. SMHI RMK No. 62, Swedish Meteorological Institute: Norrköping; 40.

Angot A. 1903. *Instructions Météorologiques*, 4th edn. Gauthier-Villars, Imp-Libraire: Paris, 49–50.

Bañón M, Brunet M, Sigró J, Aguilar E, García F, Palenzuela JE, Asín J. 2008. Comparación de las temperaturas obtenidas en las garitas Montsouri y Stevenson. Resultados provisionales del proyecto SCREEN In *Agua y cambio climático XXX Jornadas de la Asociación Española de Meteorología, IX Encuentro hispano-luso de Meteorología, XII Congreso Latinoamericano e Ibérico de Meteorología, Zaragoza, 5–7 May 2008*, Publicaciones de la Asociación Española de Meteorología: Zaragoza; 13.

Bergström H, Moberg A. 2002. Daily air temperature and pressure series for Uppsala (1722–1998). *Climatic Change* **53**: 213–252.

Böhm R, Jones PD, Hiebl J, Frank D, Brunetti M, Maugeri M. 2009. The early instrumental warm-bias: a solution for long central European temperature series 1760–2007. *Climatic Change* DOI: 10.1007/s10584-009 9649-4.

Brunet M, Bañón M, García F, Aguilar E, Saladié O, Sigró J, Asín J, López D. 2004. *Una aproximación experimental tendente a la minimización del sesgo artificial asociado al tipo de garita meteorológica a través de la observación dual de la temperatura del aire*. In *Actas de las XXVIII Jornadas Científicas de la AME, Badajoz, 11–13 February 2004*, Pub. Asociación Meteorología Española y Ministerio de Medio Ambiente: La Meteorología y el Clima Atlánticos, Badajoz: 13.

Brunet M, Saladie O, Jones PD, Sigro J, Aguilar E, Moberg A, Walther A, Lister D, Lopez D, Almarza C. 2006. The development of a new daily adjusted temperature dataset for Spain (1850–2003). *International Journal of Climatology* **26**: 1777–1802.

Brunet M, Jones PD, Sigro J, Saladie O, Aguilar E, Moberg A, Della-Marta PM, Lister D, Walther A, Lopez D. 2007. Temporal and spatial temperature variability and change over Spain during 1850–2005. *Journal of Geophysical Research – Atmosphere* **112**: D12117. DOI: 10.1029/2006JD008249.

Brunet M, Saladié O, Jones PD, Sigró J, Aguilar E, Moberg A, Lister D, Walther A, Almarza C. 2008. *A Case-Study/Guidance on the Development of Long-Term Daily Adjusted Temperature Datasets*. World Meteorological Organization (WMO) WMO-TD-1425/WCDMP-66, Geneva; 43.

Butler CJ, García Suárez AM, Coughlin ADS, Morrell C. 2005. Air temperatures at Armagh Observatory, Northern Ireland, from 1796 to 2002. *International Journal of Climatology* **25**(8): 1055–1079. DOI: 10.1002/joc.1148.

Chenoweth M. 1992. A possible discontinuity in the US historical temperature record. *Journal of Climate* **5**: 1172–1179.

Cleveland WS. 1979. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* **74**: 829–836.

Dettwiller J. 1978. Evolution séculaire de la température de l'air à Paris. *La Météorologie, VI Ser.* **13**: 95–130.

ICM (Instituto Central Meteorológico). 1893. *Treinta años de observaciones meteorológicas: exposición y resumen de las efectuadas en el Observatorio de Madrid desde el 1° de enero de 1860 al 31 de diciembre de 1889*, Establecimiento Tipográfico de los Sucesores de Cuesta: Madrid, Spain; 237.

Klingbjer P, Moberg A. 2003. A composite monthly temperature record from Tornedalen in northern Sweden, 1802–2002. *International Journal of Climatology* **23**(12): 1465–1494. DOI: 10.1002/joc.946.

Martín Vide J, Olcina J. 2001. *Climas y tiempos de España,*, Alianza Editorial: Madrid, Spain.

Moberg A, Bergström H, Ruiz Krigsman J, Svanered O. 2002. Daily air temperature and pressure series for Stokoholm (1756–1998). *Climatic Change* **53**: 171–212.

Moberg A, Alexandersson H, Bergström H, Jones PD. 2003. Were southern Swedish summer temperatures before 1860 as warm as measured? *International Journal of Climatology* **23**(12): 1495–1521. DOI: 10.1002/joc.945.

Nicholls N, Tapp R, Burrows K, Richards D. 1996. Historical thermometer exposures in Australia. *International Journal of Climatology* **16**: 705–710.

Nordly PØ, Alexandersson H, Frich P, Førland EJ, Heino R, Jónsson T, Steffensen P, Tuomenvirta H, Tveito OE. 1996. *The Effect of Radiation Screens on Nordic Temperature Measurements*. DNMI-klima, Report No. 4/96, Norwegian Meteorological Istitute, Oslo; 56.

Nordly PØ, Alexandersson H, Frich P, Førland EJ, Heino R, Jónsson T, Tuomenvirta H, Tveito OE. 1997. The effect of radiation screens on Nordic time series of mean temperature. *International Journal of Climatology* **17**: 1667–1681.

Parker DE. 1994. Effects of changing exposure of thermometers at land stations. *International Journal of Climatology* **14**: 1–31.

Parker DE. 1995. Maximum and minimum temperatures: a backward and a forward look. *Atmospheric Research* **37**: 3–9.

Peterson TC, Easterling DR, Karl TR, Groisman P, Nicholls N, Plummer N, Torok S, Auer I, Boehm R, Gullett D, Vincent L, Heino R, Tuomenvirta H, Mestre O, Szentimrey T, Salinger J, Førland EJ, Hanssen-Bauer I, Alexandersson H, Jones P, Parker D. 1998. Homogeneity adjustments of in situ atmospheric climate data: a review. *International Journal of Climatology* **18**(13): 1493–1517. DOI: 10.1002/(SICI)1097-0088(19981115)18:131493::AID-JOC3293.0.CO;2-T.

Richards D, Wilson G, Sheng How K, Kang S, Tan A, Chueng S. 1992. Comparison of Temperature Measures. Report for Bureau of Meteorology Research Centre, Swinburne University of Technology, November 1992; 176.

Richards D, Nind G, Ramchand K, Keyhoe P. 1993. *Differences in Temperature Recordings between the Glaisher Stand and Stevenson Screen*. Report for Bureau of Meteorology Research Centre, Swinburne University of Technology, October 1993; 123.

Rico Sinobas M. 1857. *Resumen de los trabajos meteorológicos correspondientes al año 1854 verificados en el Real Observatorio de Madrid*. Imprenta de Eursebio Aguado, Madrid, Spain; 125.

Sparks WR. 1972. *The effect of thermometer screen design on the observed temperature*. WMO No. 315, Geneva, 106.