Building Long Homogeneous Temperature Series across Europe: A New Approach for the Blending of Neighboring Series

ANTONELLO A. SQUINTU

*Koninklijk Nederlands Meteorologisch Instituut, De Bilt, and Meteorology and Air Quality Group,
Wageningen University, Wageningen, Netherlands*

GERARD VAN DER SCHRIER, ELSE J. M. VAN DEN BESSELAAR, AND RICHARD C. CORNES[a]

Koninklijk Nederlands Meteorologisch Instituut, De Bilt, Netherlands

ALBERT M. G. KLEIN TANK[b]

*Koninklijk Nederlands Meteorologisch Instituut, De Bilt, and Meteorology and Air Quality Group,
Wageningen University, Wageningen, Netherlands*

(Manuscript received 6 February 2019, in final form 18 October 2019)

ABSTRACT

Long and homogeneous series are a necessary requirement for reliable climate analysis. Relocation of measuring equipment from one station to another, such as from the city center to a rural area or a nearby airport, is one of the causes of discontinuities in these long series that may affect trend estimates. In this paper, an updated procedure for the composition of long series, by combining data from nearby stations, is introduced. It couples an evolution of the blending procedure already implemented within the European Climate Assessment and Dataset (ECA&D, which combines data from stations no more than 12.5 km apart from each other) with a duplicate removal, alongside the quantile matching homogenization procedure. The ECA&D contains approximately 3000 homogenized series for each temperature variable prior to the blending procedure, and approximately 820 of these are longer than 60 years; the process of blending increases the number of long series to more than 900. Three case studies illustrate the effects of the homogenization on single blended series, showing the effectiveness of separate adjustments on extreme and mean values (Geneva, Switzerland), on cases in which blending is complex (Rheinstetten, Germany), and on series that are completed by adding relevant portions of Global Telecommunications System synoptic data (Siauliai, Lithuania). A trend assessment on the whole European continent reveals the removal of negative and very large trends, demonstrating a stronger spatial consistency. The new blended and homogenized dataset will allow a more reliable use of temperature series for indices calculation and for the calculation of gridded datasets and it will be available online for users (https://www.ecad.eu).

1. Introduction

Long and high-resolution temperature series are fundamental in climatological studies for giving a historical perspective of average warming and climatic extremes (Peterson et al. 1998; Aguilar et al. 2003). An important requirement for these series is that they are homogeneous. The removal of nonclimatic or artificial signals, *inhomogeneities*, is essential since climatic trends may be contaminated by these signals as shown by, for example, Venema et al. (2013) and Squintu et al. (2019).

It is highly likely that the surroundings of long-running meteorological stations will change over time or that the location of the instruments and possibly also the measurement procedure will also change (Domonkos 2011; Kruger and Nxumalo 2017; Vincent et al. 2018). In addition, stations may be temporarily or permanently terminated, producing discontinuities in the dataset.

---

[a] Current affiliation: National Oceanography Centre, Southampton, United Kingdom.
[b] Current affiliation: Hadley Centre for Climate Change, Met Office, Exeter, United Kingdom.

Corresponding author: Antonello A. Squintu, antonello.squintu@knmi.nl

TABLE 1. Statistics with regard to length of series. In the case of OriHom, only the longest series for each station is considered.

	OriHom			NewHomBlend		
	TN	TX	TG	TN	TX	TG
No. of series	3199	3201	2710	3110	3114	2673
No. of series longer than 60 years	824	821	866	927	914	900
Avg length	46.41	46.00	52.31	48.70	48.11	53.98

A common procedure (ECA&D Project Team 2013; Menne et al. 2012; Kruger and Nxumalo 2017) that is used to build long and continuous time series is to fill gaps in a series using data from neighboring stations or to combine long segments of measurements. When data of stations from very different surroundings are combined, in particular from urban and rural (e.g., airport) areas, large inhomogeneities are introduced (Trewin and Trevitt 1996; Tuomenvirta 2001; Brunet et al. 2006; Böhm et al. 2010; Rahimzadeh and Nassaji Zavareh 2014; Vincent et al. 2018). The extension of series can be also performed with data related to sources that are not completely comparable. An example of this is the update of series to recent periods with Global Telecommunications System (GTS)-derived synoptic messages (''synops''; van den Besselaar et al. 2012). The presence of these discontinuities makes the combined data unreliable for accurate climatological analyses, since the amplitude of inhomogeneities can be as large as the climate change signal itself (Peterson et al. 1998; Caussinus and Mestre 2004; Begert et al. 2005; Della-Marta and Wanner 2006; Venema et al. 2013).

Several studies on singular cases or on small datasets have produced tailored homogeneity adjustments for composite series (Maugeri et al. 2002; Böhm et al. 2010; Yang et al. 2013; Dienst et al. 2017; Kruger and Nxumalo 2017; Delvaux et al. 2019; Nemec et al. 2013), and were able to account for the specific characteristics of the considered series and for metadata. In the case of large datasets, it is not possible to generalize a tailored approach for an individual case to the whole dataset due to the number of series involved (Dienst et al. 2017; Delvaux et al. 2019). For such datasets a fully automatic procedure is required that has to be able to handle the large spectrum of sometimes unknown circumstances (different surroundings, measurement techniques, protocols, and elevation) that each specific case presents. The procedure cannot rely on metadata as this information is often lacking

The aim of this study is to provide a new approach to the operational procedure used in the construction of composite series (blending) (ECA&D Project Team 2013) in the European Climate Assessment and Dataset (ECA&D; Klein Tank et al. 2002). The combination with a duplicate-removal procedure and with a quantile matching homogenization technique, analogous to the method presented by Squintu et al. (2019), allows the production of more reliable long temperature series, which are a fundamental tool for the estimation of indices and trends.

2. Data

The ECA&D collects data from thousands of stations spanning the European and Mediterranean domain. Temperature records play a key role, with about 3000 stored series per element (minimum, maximum, and mean) (for the status as of January 2019, see Table 1) that have been previously homogenized within the work presented in Squintu et al. (2019). Figure 1 shows the spatial distribution of homogenized nonblended (OriHom; see section 2a) series of average temperatures. Here the size of the dot represents the length of the series and the color coding the starting year. This dataset clearly has a high potential for generating longer series, especially in the denser areas, and a blending procedure was developed to construct these augmented series. The procedure works by replacing missing data in the original nonhomogenized series with measurements from surrounding stations and extends it to current times, when possible, with synops. A limitation of this approach is reflected in the production of duplicates of time series: in the case of a relocation, a new series is constructed joining donated data from the discontinued
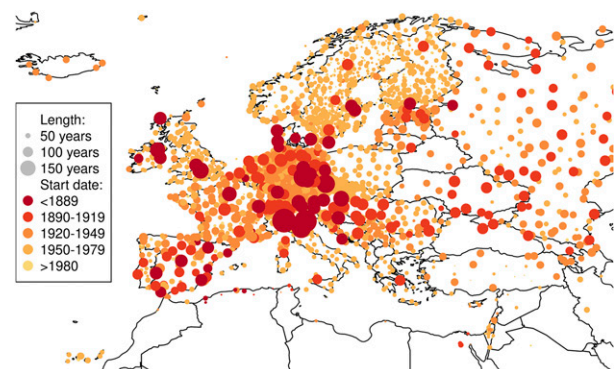

FIG. 1. Available TG homogenized series (OriHom; see section 2a) in ECA&D. The size of the dots indicates the length of the series, and color coding indicates the start year.

station to the receiving operational station *and vice versa*. Although the blended series are a fundamental tool for the development of the "E-OBS" gridded dataset (Haylock et al. 2008; van den Besselaar et al. 2011; Cornes et al. 2018), duplicating time series in the input to this gridded dataset overemphasizes the corresponding information. Thus an additional step to avoid any duplication is needed and has been included in this study.

Naming convention

The current ECA&D blended series (from now on referred to as *OldBlend*) were generated from the original nonhomogenized series (*OriNonHom*) of ECA&D. The new blended series (*NewBlend*) are the result of the updated blending procedure described in the current study, which takes as input the homogenized original series (*OriHom*). These, as introduced by Squintu et al. (2019), are the result of two iterations of homogenization that remove large breaks. A third run of break detection revealed the presence of only minor inhomogeneities, which motivated not to further adjust so as to prevent overcorrection and preserve resemblance to the original data.

The homogenized version of the blended series will be referred to as *NewHomBlend*. The method described in this paper is applied to daily minimum temperature (TN), daily maximum temperature (TX), and daily average temperature (TG).

3. Methods

a. The modified blending procedure

The construction of a NewBlend is performed for each station individually. Nearby stations around a target station are identified in a radius of 12.5 km and with a maximum of 25 m of difference in altitude. These include the stations providing synops. This distance is the result of a pragmatic choice, based on the observation that mutual distances between stations affected by relocation tend to be below such threshold. All the OriHom data belonging to the target series and to the neighboring series, which can be more than one for each station, are involved in the blending.

The series with the most recent data of the target station is called the *basis* series, while other OriHom from the same and neighboring stations are defined as *donating* series since they donate their data to fill possible gaps in the basis and to extend it to both earlier times and to more recent times (with GTS data).

In this process, illustrated in Fig. 2, each day between the earliest and the latest date in the available series is considered individually. If the daily value is
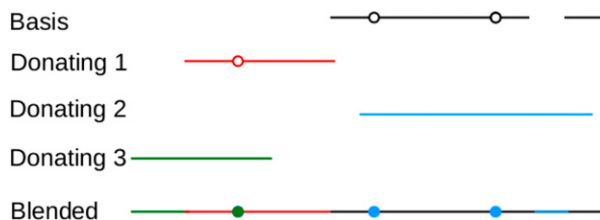

FIG. 2. Schematic diagram to show how data belonging to basis and donating series contribute to the creation of a long blended series. Lines indicate continuous segments, and open circles represent missing data. Missing data that have been infilled are indicated by filled circles having the color of the corresponding donor. Donating series are sorted according to their rank.

missing in the basis, the donating series are checked for availability for that date. The series that donates the datum is chosen according to a hierarchy based on distance from the basis; that is, if the closest donor has a value on that date, it is selected and otherwise the second closest is checked, and so on. In any case, priority is given to validated series over synop values; the latter are used only when no alternative data are available and updates from the data providers are lacking. Integration with synop data is performed only if the last validated datum lies in the 10 years preceding the current date (ECA&D Project Team 2013). In this step, the introduced data are not adjusted for elevation or changes in the surroundings.

Figure 2 illustrates that the blended series is composed of segments that come from different series. The length of such segments might vary from several years to single days. Furthermore, contributions from a specific donating series can be fragmented (filling more than one gap in the basis series), making the homogenization more complicated.

b. Avoiding duplicates

The described blending procedure determines the presence of identical segments of data due to the mutual exchange of data that can occur between two neighboring stations. The presence of duplicates has the effect of giving double weight to certain records in the construction of gridded datasets like E-OBS.

The avoidance of duplicates in the new method is illustrated by the example of the area of Berlin, Germany (Fig. 3). Initially all neighboring stations of Berlin-Tempelhof [ECA&D identifier (ID) 2759] are considered. For any of the neighboring stations, it is checked if additional surrounding stations can be included. The process is iterated until no more stations are in the radius of 12.5 km of any of the already selected stations, determining a cluster of stations that have donated data to each other. The stations in the
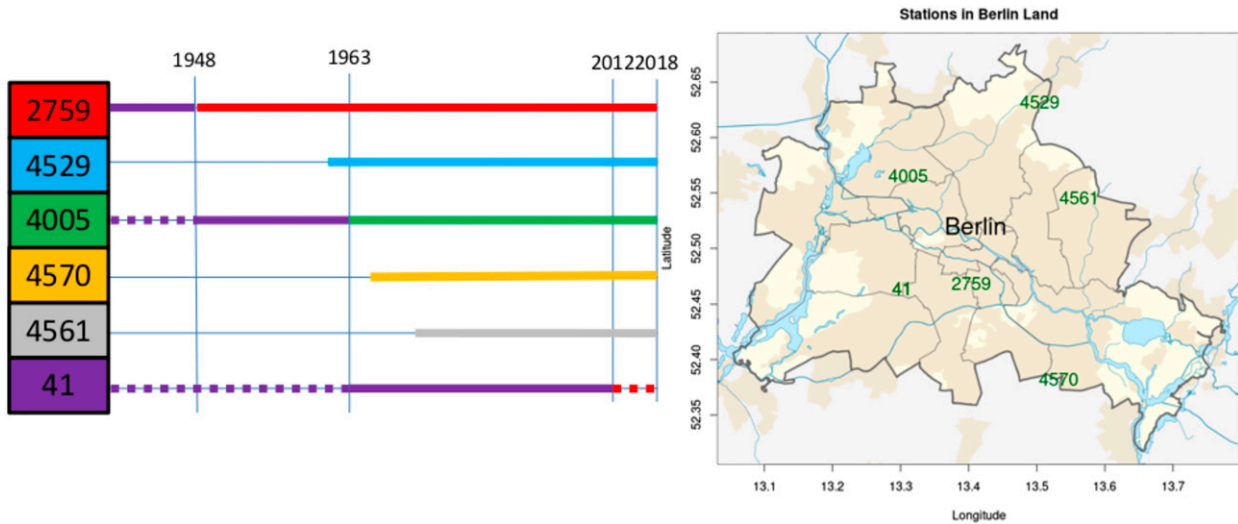
FIG. 3. Diagram and map representing the removal of duplicate records from blended series in the area of Berlin. Colors identify the series that contribute to the blended series. Dashed lines indicate data removed to avoid the presence of duplicates.

cluster are ranked, giving priority first to those with OriHom that are still updated and then to those with the longest duration. According to these criteria, Berlin-Tempelhof is ranked first. Among its donating series, the one that starts the earliest is the series of Berlin-Dahlem (ID: 41). This series is not updated anymore, and its data are used to extend Tempelhof and Berlin-Tegel (ID: 4005) to earlier times. At the same time the blending procedure found data with ID 2579 to complete the series with ID 41 to current times. The portion of 41 before 1948 is used in the blending of station 2759, thus the values of this time interval are removed from 4005 and 41, since they have lower rank. The same process is applied for the segment between 1948 and 1963, which is present in 4005 and thus removed from 41. Stations Berlin-Buch (ID: 4529), Berlin-Marzahn (ID: 4561), and Berlin-Schonefeld (ID: 4570) are part of the cluster, but their data cannot be used for the extension of 2579 nor can they profit from data from each other or from the lower-ranked station since their distance to other stations is larger than the preset threshold. As a result of this process, no daily measurement appears in more than one blended series.

c. Homogenization of the NewBlend series

In contrast to the OldBlend series, in which data from a neighboring station were simply inserted to fill gaps or to extend the record, the data of the NewBlend are adjusted to avoid the introduction of inhomogeneities. Although the generation of NewBlend series is applied to all elements available in the ECA&D database, the homogenization is applied to NewBlend

temperature series only and is based on the probability density functions of the respective OriHom series. Methods to homogenize time series with a daily resolution for other elements are currently in an experimental phase and are outside the scope of this paper.

The calculation of adjustments is based on the quantile matching method documented by Squintu et al. (2019) that is based upon earlier work of Trewin (2013). This consists of the comparison of the distributions of temperatures of the basis and of the donating OriHom. Data related to each month are considered separately over a period of at least 5 years from both basis and donating series. The records are divided into bins, each corresponding to the quantiles from 5 to 95. Sequences of values related to the quantiles are obtained, one for the basis and one for the donating series. These are subtracted from each other in order to obtain an estimate of the adjustments. Since a climate change signal may be present in the difference of these quantiles, a comparison against parallel segments extracted from surrounding stations, which are used as references, is performed to remove this climate-related signal. More details are provided in the appendix of this paper, in Squintu et al. (2019), and in Trewin (2013). The reference series are selected among all the available OriHom in a coordinate box. This box always spans 6° in latitude, whereas the range of longitude varies according to the latitude in order to cover the same direction in the west–east direction that the box covers in the north–south direction. Each reference allows the calculation of an estimate of the adjustments. As a last step,

FIG. 4. Map of Geneva (provided through the courtesy of Google Maps) with the locations of the stations Parc de l'Observatoire (south) and Cointrin Airport (north) connected by the white line.

considering the set of estimates, the median is taken as the final correcting factor. The adjustments are then applied to each datum that appears in the blended series and that does not belong to the basis, according to its donating series, month, and quantile.

4. Results

The procedure described above is applied to all the OriHom temperature series of ECA&D. Three case studies are reported in this section. The first (Geneva, Switzerland) illustrates the improvements on to the extremes indices, the second (Rheinstetten, Germany) represents an example of a series with discontinuous blending contributions from donating series, and the third (Siauliai, Lithuania) explains how the GTS data have been treated and adjusted. Following these examples, an analysis of the effects on the whole dataset is reported.

a. Case study: Geneva

Among the several examples of stations that have been relocated, one of particular interest is the series of TN from Geneva. The station in Parc de l'Observatoire was decommissioned in 1961. This station was likely influenced by nearby buildings and trees and by the urban heat island effect (see position on the map in Fig. 4). Simultaneously, the station in the open fields of the Cointrin Airport was established at a distance of 5.8 km from and 15 m higher than the old station. In the blending step, these two OriHom are joined and an adjustment is required to avoid the introduction of an inhomogeneity.

As expected (Tuomenvirta 2001; Böhm et al. 2010; Yang et al. 2013), the (indirect) removal of the urban heat island and, with lower impact, the elevation difference introduces a cooling factor to the NewBlend. A clear steplike pattern (black lines in Fig. 5) is observed
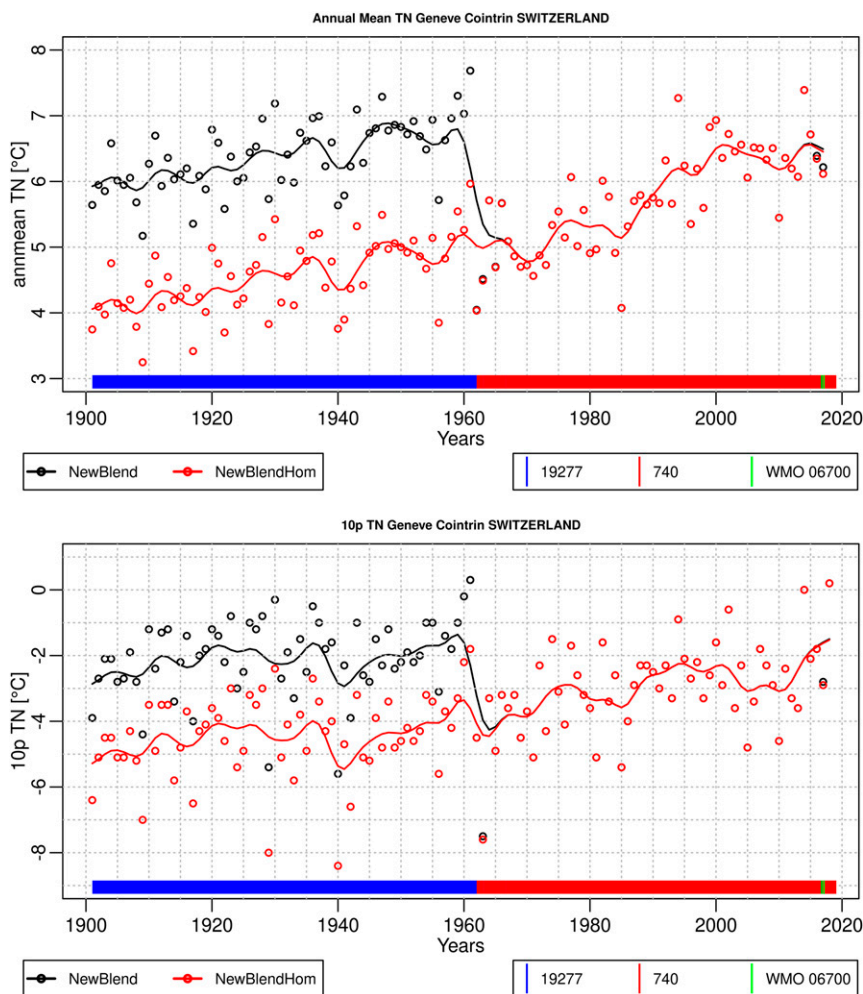
FIG. 5. (top) Annual (top) means and (bottom) 10th percentiles of minimum temperatures of the blended series of Geneva. Open circles are yearly values, and the line is a Gaussian weighted running mean. Black items represent the NewBlend; red items represent the NewHomBlend. Contributing segments are reported in the sequences of bars (one per day) on the bottom of the plots: the blue sequence indicates the period covered by Geneva Observatoire (identifier 19277), red indicates Geneva Cointrin (identifier 740), and green indicates sparse contributions of the closest synop series (WMO identifier 06700).

when the series are joined without adjustment, whereas the red lines show the blending of the series using the homogenization adjustments. The average effect of the homogenization on annual means of TN (Fig. 5, top panel) is $-1.3°C$; for the annual 10th percentile (Fig. 5, bottom panel) the adjustments are stronger, $-1.9°C$. This shows that corrections based on mean values, in this case, underestimate adjustments for the more extreme values. This example illustrates how the combination of blending and homogenization generates a long and homogeneous series where the nonclimatic effects of site location have not been adjusted.

A final check has been performed to validate the results of this process against the homogenized temperatures provided by MeteoSwiss. These are monthly values of TG; thus the average of the 12 values has been compared with the average of the 365 daily values of the TG NewHomBlend series of Geneva. The strong agreement between the two series (see Fig. 6) is confirmed by the low RMSE of 0.15°C.

b. Case study: Rheinstetten

The meteorological station at Rheinstetten Airport, Germany, was in operation between January 1948 and 1984 and again from 2009 onward. Data from this station are blended with data from the station in Karlsruhe, Germany, which is located in a park approximately 8 km away from Rheinstetten (Fig. 7)
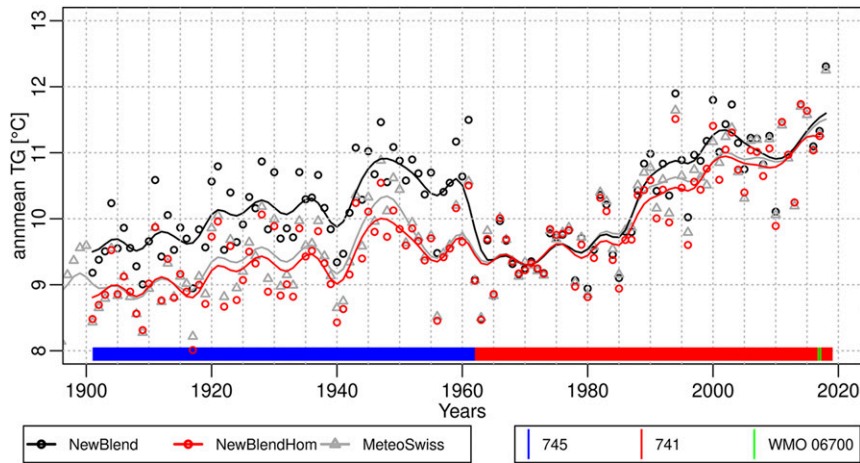
FIG. 6. Annual means of mean temperatures of the blended series of Geneva. The color code is as in Fig. 5 with the addition of the annual average of the monthly homogenized values provided by MeteoSwiss (gray lines and open triangles).

and 4 m of elevation lower. This urban station has an uninterrupted record running from 1876 to 2008, allowing the series of Rheinstetten to be extended farther back in time and to fill the gap between 1985 and 2008. Figure 8 shows how in both periods the temperature in the Karlsruhe station appears to be warmer (as expected for a station in an urban environment) than the rural station. The applied adjustments
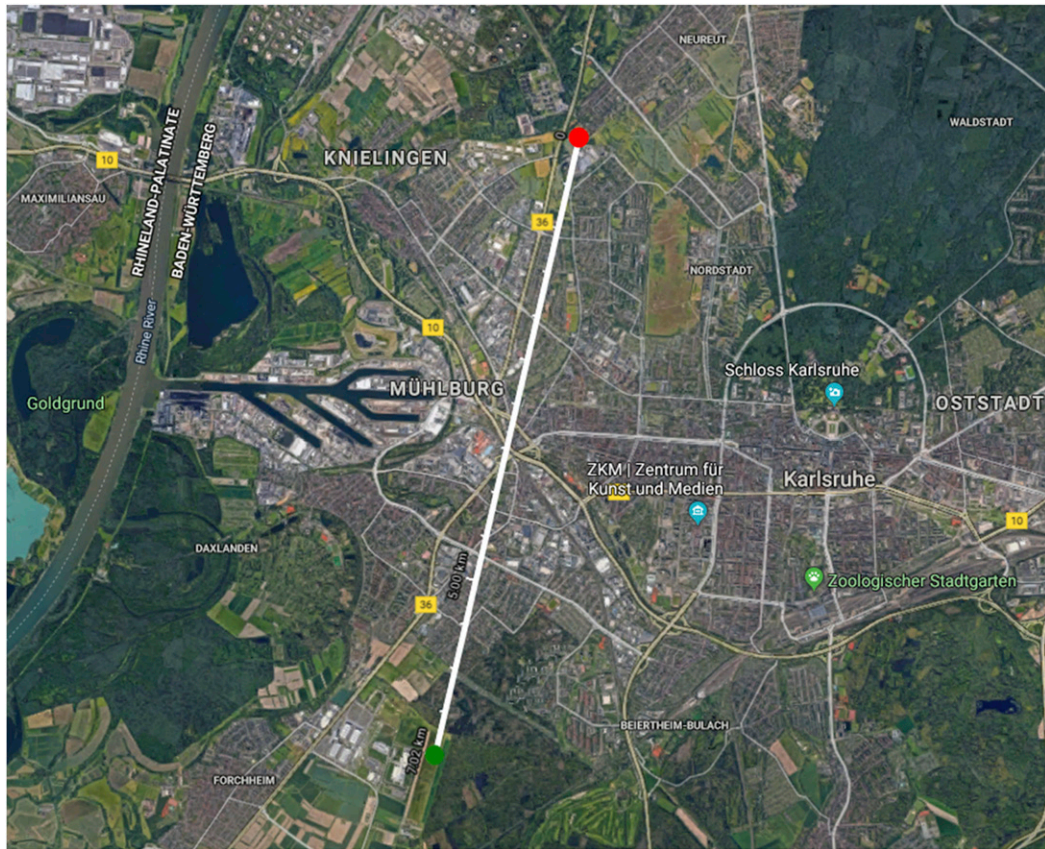

FIG. 7. Map of Karlsruhe and Rheinstetten (provided through the courtesy of Google Maps) with the locations of the stations Karlsruhe (north) and Rheinstetten airstrip (south), with the white line connecting the two stations.
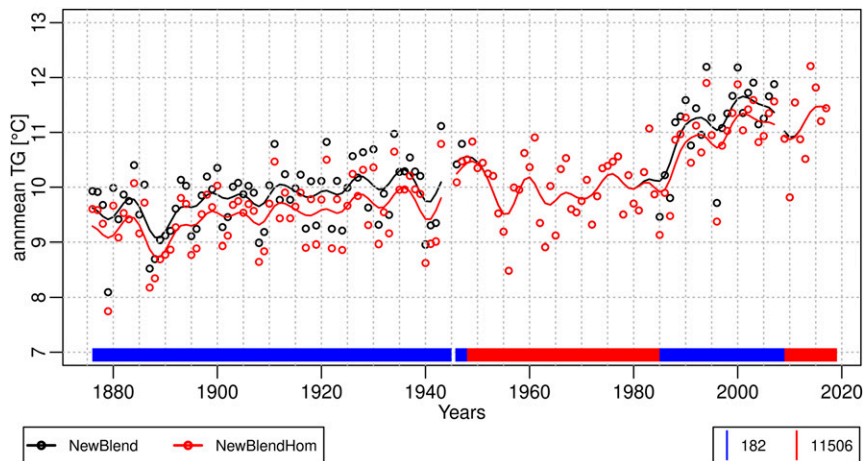
FIG. 8. Annual mean of the blended TG series of Rheinstetten. The color code is as in Fig. 5. The red bar indicates data from Rheinstetten (11 506); the blue bar indicates data from Karlsruhe (182).

convert these data to values that are more consistent with the basis.

The structure of this blended series is a representative, but rare, example with multiple relevant contributions from the same donating series. More frequently it is observed that multiple contributions from the same donating series consist of a main segment and several individual or short sequences of donated data that fill sparse missing values in the basis. During the homogenization step each segment is adjusted according to the statistical features of the whole donating series. Thus the quantile adjustments are the same for all segments belonging to the same donor. Calculating separate adjustments for each segment would require arbitrary choices: determining which criteria to follow for splitting the donating series would make the process subjective. Furthermore, such a decision would imply the development of a more complicated algorithm that would be applied to a very small portion of the dataset (about 0.3%).

This motivated us to apply as adjusting factors to blended data before 1948 and between 1984 and 2000 values that are drawn from the same set of quantile-based adjustments. Since the temperature values in the record before 1948 are generally lower than those in the latter period, the adjustments for the early period will sample the lower quantiles more than those from the more recent segment.

c. *Extending the series with homogenized synop data: Siauliai*

Synop data are used for a maximum of 10 years to extend to present day those series that are not regularly updated (ECA&D Project Team 2013).

The daily TN and TX values are based on 12-h periods (from 0600 to 1800 UTC for TX and from 1800 to 0600 UTC for TN). This implies that synop values for TX are equal to or underestimate the 24-h-based TX values. Conversely, synop values for TN are equal to or overestimate the 24-h-based TN values. Daily mean values (TG) are calculated as a simple average of TX and TN. Therefore, this system is likely to introduce systematic biases on the last portion of the series and thus to introduce inhomogeneities (van den Besselaar et al. 2012). The quantile matching approach is used to adjust these biases, making synoptic data more consistent with data from the surrounding stations.

It is important to stress that synops are appended and corrected, but not used as references; their statistical features are not transferred in any case to the surrounding series. An example of the correction of synops, for the TN and TX series of Siauliai is shown in Fig. 9. As expected, the corrections applied to synops have low amplitude and different sign for maximum and minimum temperature. While this behavior has been observed on almost all the most relevant contributions of GTS data, the histograms of adjustments applied to synops (Fig. 10) show that the negative adjustments dominate for minimum temperatures, while for maximum temperatures the distribution is centered around zero. This difference may be due to the fact that the probability of recording maximum temperature outside the 0600–1800 interval (local time) is lower than the probability of having minimum temperature occurring outside the 1800–0600 interval (van den Besselaar et al. 2012). Second, the location of the station providing the synops
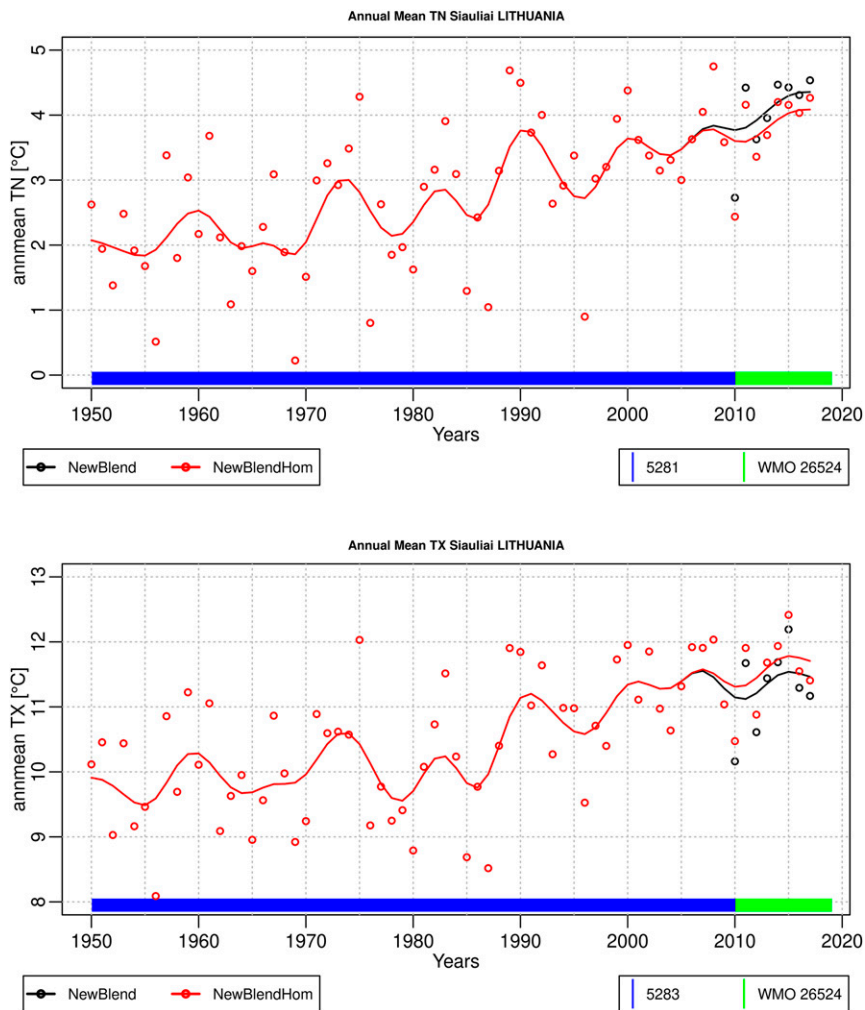
FIG. 9. Example of homogenization of GTS contributions to the series of Siauliai for (top) TN and (bottom) TX. The color code is as in Fig. 5.

series does not always coincide with the basis series, a fact that might interfere with the systematic biases introduced above.

d. Application to the whole dataset: Statistics

Table 1 describes the effects on the whole dataset. Considering the average lengths of OriHom (taking the longest for each station) and NewHomBlend, no considerable changes are observed (around +2 years). The transplant from a series to another together with the removal of duplicates explains the low increase. The most relevant change is observed in the distribution of the length of the series. This is illustrated in the histogram of Fig. 11, which highlights the shift toward larger values (and thus longer series) that is confirmed by the increase (~+16%) of very long series (more than 60 years); see Table 1. At the same time stations with a very low amount of data slightly

increase their number. These series have donated most of their data to their neighbors.

e. Effects of modified blending on trends

The results presented in sections 4a–4c indicate the power of the coupled blending and homogenization on individual series. Nonetheless, the correction of anomalous trends and the pursuit of a stronger spatial consistency of the trends is one of the primary tasks of the blending procedure.

Figure 12 compares the results for NewBlend and NewHomBlend. The trends displayed in those plots are for the annual 90th percentile of TX and are calculated using Sen's slope method (Sen 1968), which is more robust than linear regression; the significance of these trends is estimated using the Mann–Kendall test. Only the trends of those series with at least 80 years of data between 1911 and
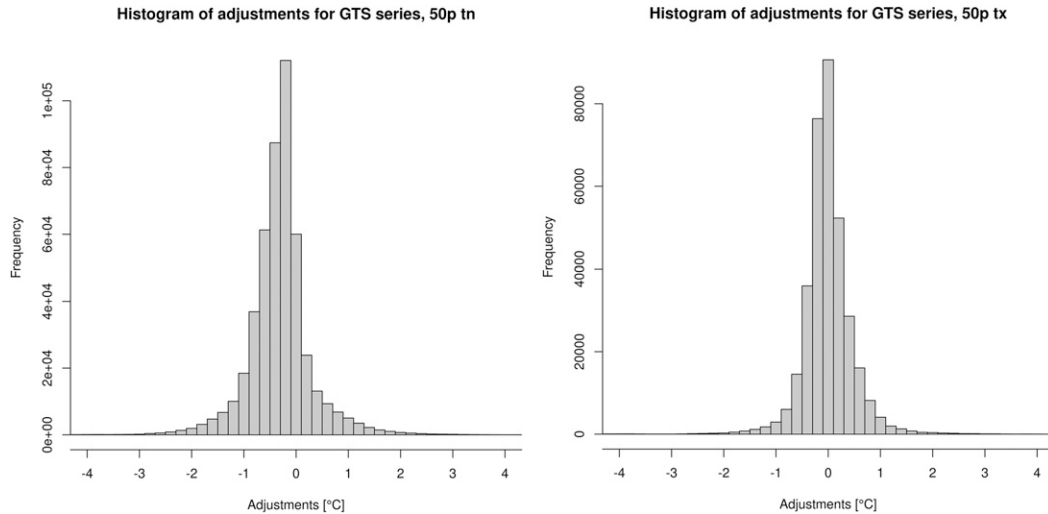
FIG. 10. Histograms of adjustments applied to the 50th percentiles of GTS series for (left) TN and (right) TX.

2010 are considered so as to focus on long term phenomena.

In the top panel of Fig. 12 a lack of order is evident in the distribution of trends, with a few apparent outliers: for example, Girona (Spain) with 0.70°C $(10\,\text{yr})^{-1}$, Milan (Italy) with -0.20°C $(10\,\text{yr})^{-1}$, and Uccle (Belgium) with -0.16°C $(10\,\text{yr})^{-1}$. These is a clear sign of the presence of gross inhomogeneities. At the same time, less-evident anomalies alter the spatial consistency, since the amplitude of such signals is comparable to those of the climatic trends. The disappearance of the extremely high and extremely low trends and the more consistent spatial patterns indicate the improved quality of the NewBlendHom dataset.

The map in the bottom panel of Fig. 12 shows the difference between the trends of NewHomBlend and NewBlend for each station. The difference of two trends is considered significant if the 95% confidence intervals of the two subtraction terms do not overlap. The absence of any pattern in such a plot indicates that the modified blending approach is neutral and does not favor an increase (decrease) of temperature trends over Europe, which would correspond to a predominance of orange/red (blue/purple) circles in the bottom panel of Fig. 12.

The median of the trends over Europe of NewBlend and NewHomBlend do not show relevant changes [from 0.14°C $(10\,\text{yr})^{-1}$ to 0.15°C $(10\,\text{yr})^{-1}$ for the 50th percentile of the three elements]; that is, when looking at the whole distribution, the differences in sign of the corrections compensate each other. Nevertheless, changes in trends of the individual series are not negligible (Tuomenvirta 2001). The specific features of each area are filtered out by the summarizing

calculation (as mean, median, and percentiles) because of the heterogeneity (Donat and Alexander 2012; Li et al. 2016) of the dataset and its nonuniform density.

f. Assessment of trends

The geographical patterns of temperature trends can now be assessed without being affected by the identified and introduced inhomogeneities. The quantile matching method provides different adjustments for the lower and the higher quantiles, and hence trends in the 10th and 90th percentiles of temperature are evaluated in this section. The maps in Fig. 13 show how the calculated trends vary over Europe. Winter (December, January, and February) minimum temperatures (Fig. 13, left column) show large trends in eastern Europe and in the
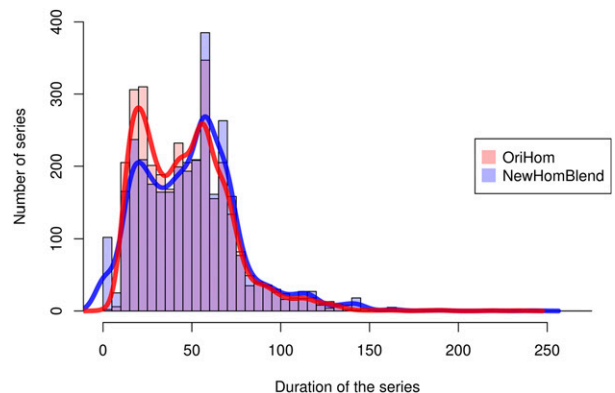

FIG. 11. Histogram and density function of the length of series before (OriHom) and after (NewHomBlend) the blending and the homogenization of the blended series. Red and blue items represent the original and final datasets, respectively.

TX ann NewBlend trends of 90p 1911-2010

TX ann NewHomBlend trends of 90p 1911-2010

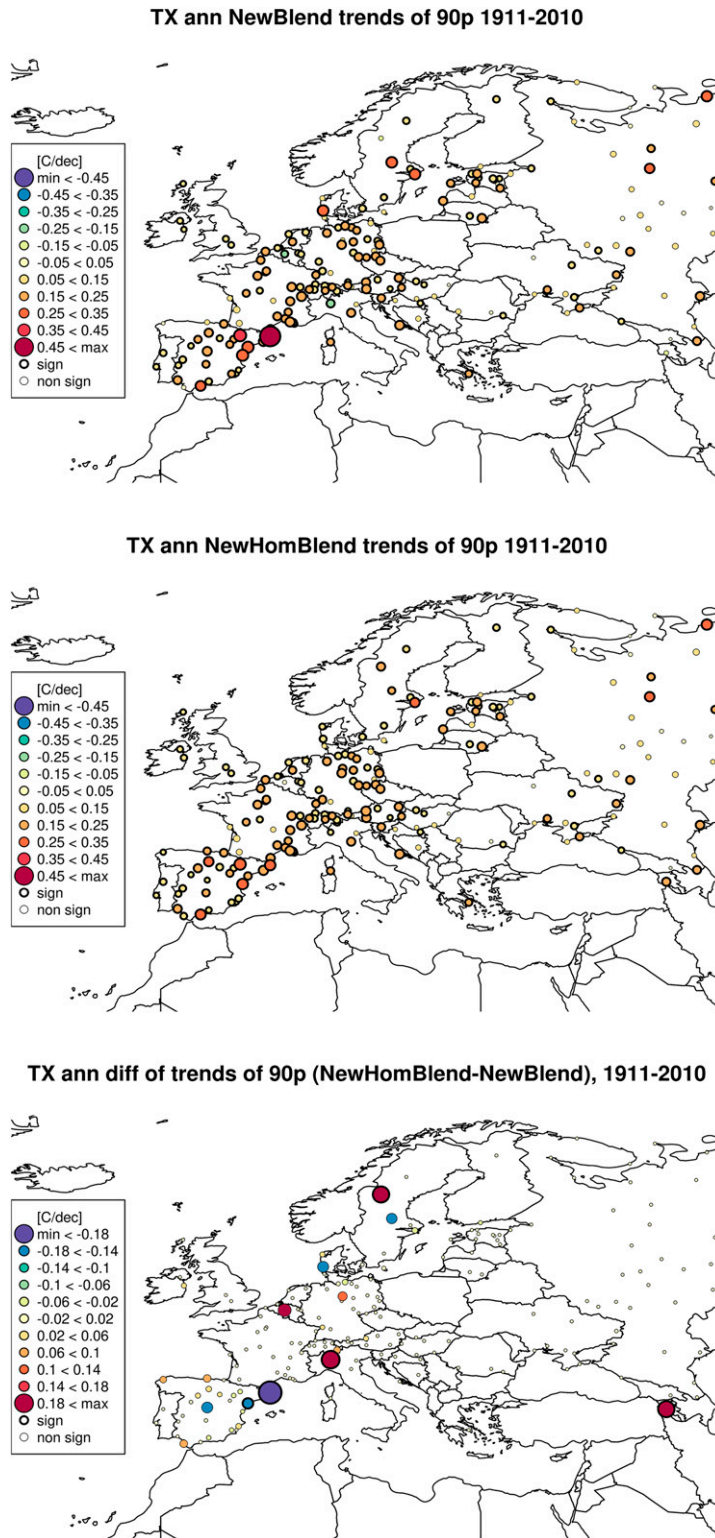**TX ann diff of trends of 90p (NewHomBlend-NewBlend), 1911-2010**

FIG. 12. Observed trends on 90th percentile of annual TX on the (top) nonhomogenized (NewBlend) and (bottom) homogenized (NewHomBlend) blended series. Also shown is (bottom) the difference (NewHomBlend − NewBlend) of these trends for each station.
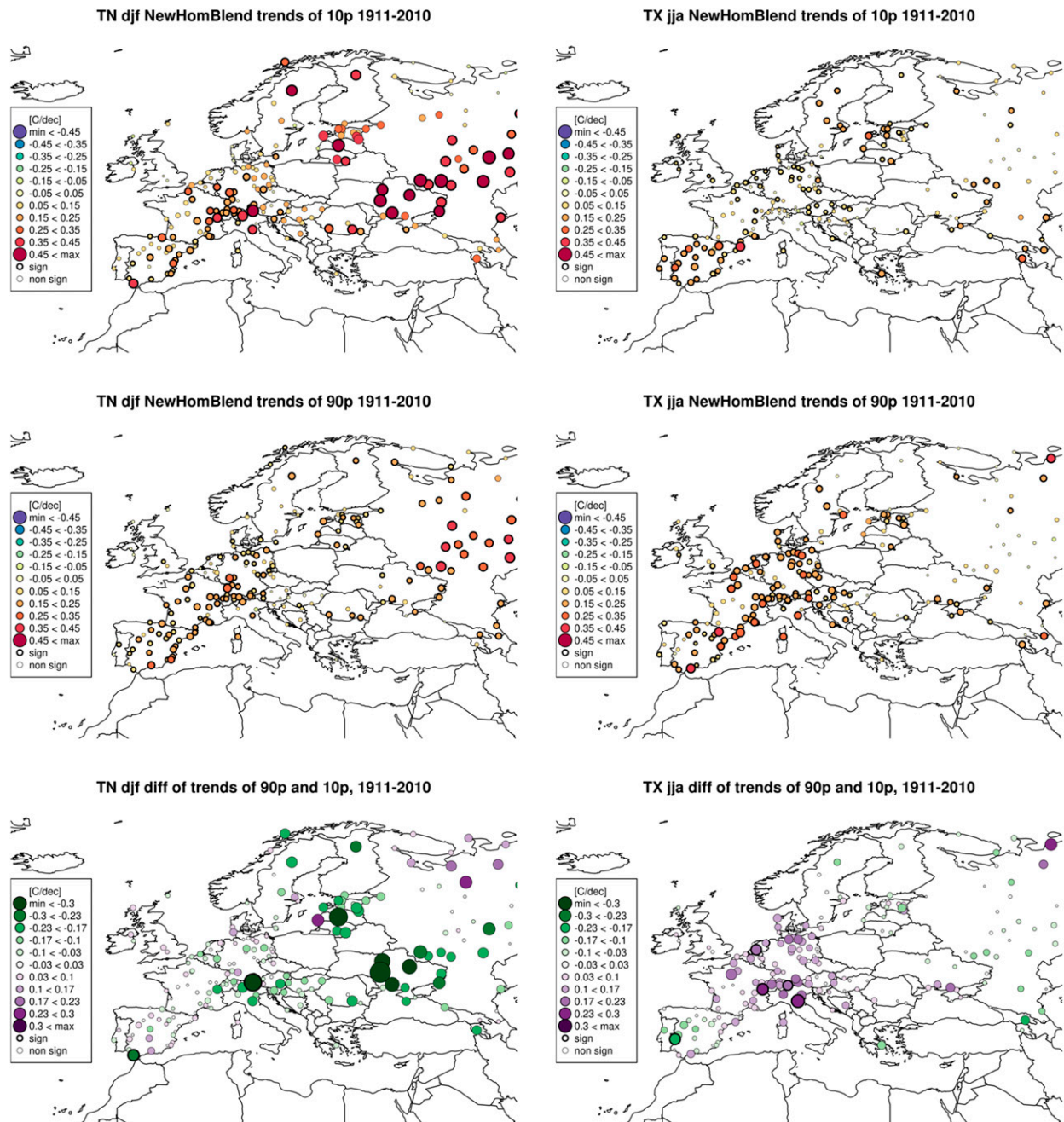
FIG. 13. Maps of trends of NewHomblend of (left) winter TN and (right) summer TX for (top) 10th and (center) 90th percentiles for the 1911–2010 period. (bottom) The difference between the trends for 90th and 10th percentiles: green and purple represent narrowing and broadening of the distribution, respectively.

Alpine region, especially for the lower percentiles. In the bottom panels of Fig. 13 changes in the shape of the distribution have been inspected by subtracting trends of 90th and 10th percentiles from each other. For winter minimum temperature, a narrowing of the probability distribution across the eastern Mediterranean (Alps, Balkans, and Ukraine) and Baltic area is observed. This

might be related to a decrease in the snow coverage of the areas during winter that disproportionally affects the *cold* tail of the distribution. At the same time Atlantic regions and Arctic Russia show a slight increase in the distribution width. In contrast, summer (June, July, and August) maximum temperatures (Fig. 13, right column) have larger trends across southern Europe. While this

increase has almost the same amplitude over the Iberian Peninsula, the results for central Europe and northern Italy are affected by steeper trends for the 90th percentile, which implies that the distribution of summer maximum temperature is becoming wider in these latter areas. These are the areas where the increase in intensity and duration of summer heat waves has had the highest impact, resulting in a larger trend on the *warm* tail of the distribution. These results are strongly related with the increase of frequency of extremely warm events, confirmed by several works on heat waves over the Mediterranean area (Della-Marta et al. 2007; Simolo et al. 2010; Yosef et al. 2018).

5. Summary and discussion

In this study, the homogenization procedure for temperature series documented in Squintu et al. (2019) is coupled with the blending procedure that is used operationally in the ECA&D (ECA&D Project Team 2013). The aim of the blending is to produce series that are as long and as complete as possible, facilitating climatological assessments. Long and homogeneous temperature records are constructed and trends in these records are compared against trends based on records that are constructed using the traditional method.

The most common situation that results in inhomogeneities is the relocation of a station from the city to the airport or to the countryside, which is expected to introduce a cooling signal in the resulting blended series. Furthermore, several other factors might also affect the series, including changes in the instruments and changes in the way the daily value is determined. Another specific issue that has been addressed is the extension of the series with synoptic messages. The inhomogeneity introduced when combining validated data with these synoptic data (van den Besselaar et al. 2012) is also corrected.

The approach is illustrated using examples of three locations in Europe. A trend assessment comparing the new and the traditional method shows significant changes (variation in trends on 90th percentile of annual TX above 5% or below −5%) in 22% of the stations (Fig. 12).

With this new development, more reliable trend assessments on the extreme parts of the temperature distribution can be made as unrealistic outliers are removed. One example demonstrating this is the difference in trends in the 10th and 90th percentiles of seasonal temperature records. Dramatic differences between eastern (warm trends in winter TN) and western Europe (warm trends in summer TX) have been observed. In addition, the difference between trends in the 90th and 10th percentile of seasonal temperature has revealed a narrowing

of winter TN distribution over southeast Europe and a widening of summer TX distribution over central Europe. This highlights the importance of conducting regional analyses of climate change impacts.

Considering the results of the modified blending procedure for temperature, additional actions are needed in the future for its further development. One of the most ambitious ideas is the use of measurements of other meteorological variables (such as solar radiation, humidity, or wind speed) to better interpret the differences among the stations that donate data to the blended series. The availability of a dataset with such long and homogeneous series will be a starting point for further studies on the characteristics of the warming climate of Europe.

The new blended and homogenized series will be used as the basis for a future version of E-OBS (Cornes et al. 2018) and should serve as a reliable benchmark for comparisons against computed trends in climate models. All original data as provided by the ECA&D data providers, as well as the blended homogenized data, will be made available where permission to do so is given.

APPENDIX

Quantile Matching: Brief Description of the Procedure

The quantile matching procedure is described in detail in Squintu et al. (2019), which has been developed from previous work by Trewin (2013). As discussed in the main text the homogenization procedure applied to the NewBlend series takes as input the OriHom series that have contributed to the creation of the blended series itself. The latest ending series (excluding GTS) that belongs to the target station is called *basis* **B**, while all the other series (from the same and from neighboring stations) will be referred to as *donating* series. The procedure makes use of a set of OriHom series located in the surrounding (within a coordinate box whose width varies according to latitude, see main text).

Adjustments are calculated considering individually each donating series, D_i. For each series the first step is the definition of the reference list. At least three

reference series, having at least five years of overlap and a correlation of 0.75 or higher with the basis and the considered donor, are required to proceed with the homogenization of the donated data; in the case of high availability, the most correlated reference series are selected. For each reference $\mathbf{R}_j$, $\mathbf{R}_j^B$ and $\mathbf{R}_j^{D_i}$ are defined as the portions of the reference that overlap with the basis and the donating series, respectively.

Adjustment calculation is performed for each month separately, including adjacent months in order to reduce the noise. The absolute temperature measurements related to the selected months of $\mathbf{B}$, $\mathbf{D}_i$, $\mathbf{R}_j^B$, and $\mathbf{R}_j^{D_i}$ are sorted in ascending order. The 5th, 10th, . . . , 90th, and 95th quantiles are selected, generating four quantile sequences ($\mathbf{b}_{q,m}$, $\mathbf{d}_{i,q,m}$, $\mathbf{r}_{j,q,m}^B$, and $\mathbf{r}_{j,q,m}^{D_i}$). The adjustments for $\mathbf{D}_i$, related to reference $\mathbf{R}_j$ for the month m and the quantile q are then calculated as

$$\mathbf{a}_{i,j,q,m} = (\mathbf{b}_{q,m} - \mathbf{d}_{i,q,m}) - (\mathbf{r}_{j,q,m}^B - \mathbf{r}_{j,q,m}^{D_i}). \quad (A1)$$

This process is iterated for each reference $j = 1, \ldots, r$.

At this point each value v that has been donated by $\mathbf{D}_i$, knowing its month, is adjusted by looking, for each $\mathbf{R}_j$, at the quantile $\tilde{q}_j$ to which it belongs in the overlapping period between $\mathbf{D}_i$ and $\mathbf{R}_j$. Thus r estimates of the adjusted value are obtained:

$$\tilde{v}_j = v + \mathbf{a}_{j,\tilde{q}_j,m}. \quad (A2)$$

The final adjusted value is then calculated taking the median of the estimations:

$$\overline{v} = \text{median}_j(\tilde{v}_j). \quad (A3)$$

This process is iterated for each donating series that contributes to the blended series.

REFERENCES

Aguilar, E., I. Auer, M. Brunet, T. C. Peterson, and J. Wieringa, 2003: Guidance on metadata and homogenization. WMO Rep. WCDMP 53, WMO-TD 1186, 55 pp., https://library.wmo.int/doc_num.php?explnum_id=9252.

Begert, M., T. Schlegel, and W. Kirchhofer, 2005: Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000. *Int. J. Climatol.*, **25**, 65–80, https://doi.org/10.1002/joc.1118.

Böhm, R., P. D. Jones, J. Hiebl, D. Frank, M. Brunetti, and M. Maugeri, 2010: The early instrumental warm-bias: A solution for long central European temperature series 1760–2007. *Climatic Change*, **101**, 41–67, https://doi.org/10.1007/s10584-009-9649-4.

Brunet, M., and Coauthors, 2006: The development of a new dataset of Spanish Daily Ajusted Temperature Series (SDATS) (1850–2003). *Int. J. Climatol.*, **26**, 1777–1802, https://doi.org/10.1002/joc.1338.

Caussinus, H., and O. Mestre, 2004: Detection and correction of artificial shifts in climate series. *J. Roy. Stat. Soc.*, **53C**, 405–425, https://doi.org/10.1111/j.1467-9876.2004.05155.x.

Cornes, R. C., G. van der Schrier, E. J. M. van den Besselaar, and P. D. Jones, 2018: An ensemble version of the E-OBS temperature and precipitation data sets. *J. Geophys. Res. Atmos.*, **123**, 9391–9409, https://doi.org/10.1029/2017JD028200.

Della-Marta, P. M., and H. Wanner, 2006: A method of homogenizing the extremes and mean of daily temperature measurements. *J. Climate*, **19**, 4179–4197, https://doi.org/10.1175/JCLI3855.1.

——, M. R. Haylock, J. Luterbacher, and H. Wanner, 2007: Doubled length of western European summer heat waves since 1880. *J. Geophys. Res.*, **112**, D15103, https://doi.org/10.1029/2007JD008510.

Delvaux, C., R. Ingels, V. Vrábeĺ, M. Journée, and C. Bertrand, 2019: Quality control and homogenization of the Belgian historical temperature data. *Int. J. Climatol.*, **39**, 157–171, https://doi.org/10.1002/JOC.5792.

Dienst, M., J. Lindén, E. Engström, and J. Esper, 2017: Removing the relocation bias from the 155-year Haparanda temperature record in northern Europe. *Int. J. Climatol.*, **37**, 4015–4026, https://doi.org/10.1002/joc.4981.

Domonkos, P., 2011: Efficiency evaluation for detecting inhomogeneities by objective homogenisation methods. *Theor. Appl. Climatol.*, **105**, 455–467, https://doi.org/10.1007/s00704-011-0399-7.

Donat, M. G., and L. V. Alexander, 2012: The shifting probability distribution of global daytime and night-time temperatures. *Geophys. Res. Lett.*, **39**, L14707, https://doi.org/10.1029/2012GL052459.

ECA&D Project Team, 2013: EUMETNET/ECSN optional programme: "European Climate Assessment & Dataset (ECA&D)" algorithm theoretical basis document (ATBD), version 10.7. Royal Netherlands Meteorological Institute Doc., 46 pp., https://www.ecad.eu/documents/atbd.pdf.

Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New, 2008: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *J. Geophys. Res.*, **113**, D20119, https://doi.org/10.1029/2008JD010201.

Klein Tank, A. M. G., and Coauthors, 2002: Daily dataset of 20th-century surface air temperature and precipitation series for the European climate assessment. *Int. J. Climatol.*, **22**, 1441–1453, https://doi.org/10.1002/joc.773.

Kruger, A. C., and M. Nxumalo, 2017: Surface temperature trends from homogenized time series in South Africa: 1931–2015. *Int. J. Climatol.*, **37**, 2364–2377, https://doi.org/10.1002/joc.4851.

Li, Z., L. Cao, Y. Zhu, and Z. Yan, 2016: Comparison of two homogenized datasets of daily maximum/mean/minimum temperature in China during 1960–2013. *J. Meteor. Res.*, **30**, 53–66, https://doi.org/10.1007/s13351-016-5054-x.

Maugeri, M., L. Buffoni, B. Delmonte, and A. Fassina, 2002: Daily Milan temperature and pressure series (1763-1998): Completing and homogenising the data. *Climatic Change*, **53**, 119–149, https://doi.org/10.1023/A:1014923027396.

Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston, 2012: An overview of the global historical climatology network-daily database. *J. Atmos. Oceanic Technol.*, **29**, 897–910, https://doi.org/10.1175/JTECH-D-11-00103.1.

Nemec, J., C. Gruber, B. Chimani, and I. Auer, 2013: Trends in extreme temperature indices in Austria based on a new homogenised dataset. *Int. J. Climatol.*, **33**, 1538–1550, https://doi.org/10.1002/joc.3532.

Peterson, T. C., and Coauthors, 1998: Homogeneity adjustments of in situ atmospheric climate data: A review. *Int. J. Climatol.*, **18**, 1493–1517, https://doi.org/10.1002/(SICI)1097-0088(19981115)18:13<1493::AID-JOC329>3.0.CO;2-T.

Rahimzadeh, F., and M. Nassaji Zavareh, 2014: Effects of adjustment for non-climatic discontinuities on determination of temperature trends and variability over Iran. *Int. J. Climatol.*, **34**, 2079–2096, https://doi.org/10.1002/joc.3823.

Sen, P. K., 1968: Estimates of the regression coefficient based on Kendall's tau. *J. Amer. Stat. Assoc.*, **63**, 1379–1389, https://doi.org/10.1080/01621459.1968.10480934.

Simolo, C., M. Brunetti, M. Maugeri, T. Nanni, and A. Speranza, 2010: Understanding climate change–induced variations in daily temperature distributions over Italy. *J. Geophys. Res.*, **115**, D22110, https://doi.org/10.1029/2010JD014088.

Squintu, A. A., G. van der Schrier, Y. Brugnara, and A. Klein Tank, 2019: Homogenization of daily temperature series in the European Climate Assessment & Dataset. *Int. J. Climatol.*, **39**, 1243–1261, https://doi.org/10.1002/joc.5874.

Trewin, B., 2013: A daily homogenized temperature data set for Australia. *Int. J. Climatol.*, **33**, 1510–1529, https://doi.org/10.1002/joc.3530.

——, and E. Trevitt, 1996: The development of composite temperature records. *Int. J. Climatol.*, **16**, 1227–1242, https://doi.org/10.1002/(SICI)1097-0088(199611)16:11<1227::AID-JOC82>3.0.CO;2-P.

Tuomenvirta, H., 2001: Homogeneity adjustments of temperature and precipitation series—Finnish and Nordic data. *Int. J. Climatol.*, **21**, 495–506, https://doi.org/10.1002/joc.616.

van den Besselaar, E. J. M., M. R. Haylock, G. van der Schrier, and A. M. G. Klein Tank, 2011: A European daily high-resolution observational gridded data set of sea level pressure. *J. Geophys. Res.*, **116**, D11110, https://doi.org/10.1029/2010JD015468.

——, A. M. G. Klein Tank, G. Van der Schrier, and P. D. Jones, 2012: Synoptic messages to extend climate data records. *J. Geophys. Res.*, **117**, D07101, https://doi.org/10.1029/2011JD016687.

Venema, V. K., and Coauthors, 2013: Benchmarking homogenization algorithms for monthly data. *AIP Conf. Proc.*, **1552**, 1060–1065, https://doi.org/10.1063/1.4819690.

Vincent, L. A., E. J. Milewska, X. L. Wang, and M. M. Hartwell, 2018: Uncertainty in homogenized daily temperatures and derived indices of extremes illustrated using parallel observations in Canada. *Int. J. Climatol.*, **38**, 692–707, https://doi.org/10.1002/joc.5203.

Yang, P., G. Ren, and W. Liu, 2013: Spatial and temporal characteristics of Beijing urban heat island intensity. *J. Appl. Meteor. Climatol.*, **52**, 1803–1816, https://doi.org/10.1175/JAMC-D-12-0125.1.

Yosef, Y., E. Aguilar, and P. Alpert, 2018: Detecting and adjusting artificial biases of long-term temperature records in Israel. *Int. J. Climatol.*, **38**, 3273–3289, https://doi.org/10.1002/joc.5500.