Monitoring European average temperature based on the E-OBS gridded data set

G. van der Schrier,1 E. J. M van den Besselaar,1 A. M. G. Klein Tank,1 and G. Verver1

Received 12 October 2012; revised 22 March 2013; accepted 24 April 2013; published 5 June 2013.

[1] A European average temperature with monthly resolution is constructed based on the E-OBS daily data set with near real-time updates for monitoring. Taken together, the European average temperature and the associated gridded daily maps of surface temperature from the E-OBS data set provide a detailed record of European climate variability and change since 1950. Both are based on validated station data provided by the European National Meteorological and Hydrological Services. A quantitative analysis of the uncertainty sources to the European average temperature indicates that the uncertainties due to urbanization, statistical interpolation, and the potential inhomogeneities in the input records to E-OBS dominate the total uncertainty estimate. In the aggregation of the interpolation uncertainty from the daily to the monthly level and then to a European averaged value, the effective sample size and the effective spatial degrees of freedom are estimated to account for spatial and temporal coherency in the uncertainty estimates. The European average temperature shows that 7 years of the top 10 warmest years are from the period starting as recent as the year 2000 and a clear upward trend in annual average temperatures over the last few decades is visible. The most recent year in the top 10 coldest years is 1987. It also shows that warming in Europe is accelerating compared to the warming over the global land masses and to a lesser extent compared to the Northern Hemisphere land masses over the period 1980–2010.


1. Introduction

[2] There is a wealth of evidence that global temperatures are on the rise [Trenberth et al., 2007]. This is true for Europe as well, where temperatures are not only rising faster than the global average but are also rising faster than expected from climate projections [van Oldenborgh et al., 2009]. With the warming of Europe, hot summers have occurred in the recent past, which were unprecedented in the instrumental record, like the 2003 summer with its emphasis on western Europe [Luterbacher et al., 2004], surpassed in extremity by the recent 2010 summer, which had its emphasis on eastern Europe [Barriopedro et al., 2011]. Also, autumn 2006 and winter 2007 have been exceptionally warm [Cattiaux et al., 2009; Luterbacher et al., 2007], while the 2010 winter has been cold but still warmer than what might be expected based on the persistently negative phase of the North Atlantic Oscillation [Cattiaux et al., 2010].

[3] Future climate simulations show that this trend is likely to persist, with hot summer extremes in south-central Europe, warming stronger than the mean and in northern Europe cold winter extremes warming stronger than mean temperatures [Fischer et al., 2012].

[4] The seemingly increase in frequency of extremely hot summers, and warm seasons in general, warrants reliable monitoring tools, which can place observed temperatures and trends in a historical perspective, and the associated daily fields can be used for describing extremes. This tool needs to be updated frequently enough to be of near real-time use for climate monitoring purposes. This paper describes the work leading to such a monitoring tool for European average temperature with a monthly resolution.

[5] The basis for the European average temperature will be the E-OBS gridded data set [Haylock et al., 2008]. This data set contains daily maps of gridded data spanning the period from 1 January 1950 to the present, at four different grid resolutions. The E-OBS data set was developed as part of the European Union Framework 6 ENSEMBLES project [van der Linden and Mitchell, 2009], with the aim to provide data for validation of Regional Climate Models (RCMs) and also for climate change studies.

[6] The data used in the gridding of E-OBS is from the European Climate Assessment & Dataset (ECA&D), [Klein Tank et al., 2002; Klok and Klein Tank, 2008]. ECA&D is a collection of daily station observations of currently 12 elements (of which five are gridded) and contains data from nearly 6600 European stations (status in June 2012) and is gradually expanding. ECA&D also contains data from the...

1Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands.

Corresponding author: G. van der Schrier, Royal Netherlands Meteorological Institute (KNMI), PO Box 201, 3730 AE De Bilt, The Netherlands. (schrier@knmi.nl)

©2013. American Geophysical Union. All Rights Reserved.
2169-897X/13/10.1002/jgrd.50444

5120
pre-1950 period. Data from the station network at ECA&D is updated on a monthly basis using data kindly provided by the National Meteorological and Hydrological Services (NMHSs), individual researchers affiliated with a university, global data centers like the National Climatic Data Center (NOAA, Asheville, USA) or the synoptic messages from the Global Telecommunication System [World Meteorological Organization, 2007]. ECA&D receives station data from 57 participants (mostly NMHSs) for 62 countries (status in June 2012).

[7] The maximum spatial extent of E-OBS covers the whole of Europe including Turkey, Northern Africa, and the Middle East. However, the station data availability at the southern and eastern edges of the domain is poor, in the early years in particular. For this reason we confine the calculations of a European average temperature to a substantially smaller area than E-OBS provides. The area over which the average is calculated are the land grid squares in the area between 12°W and 45°E, and 30°N and 75°N, supplemented by Iceland (25°W to 12°W and 63°N to 67°N). This area includes the Mediterranean and parts of Northern Africa (mainly Morocco, Algeria, and Tunisia). It excludes Svalbard and Greenland in the north and the Canary Islands, the Azores, and Madeira in the southwest. This choice is a compromise between quality, with simply too few data to extend the area further south or east, and the desire to cover the geographically defined Europe, reaching up to the Ural Mountains at 60°E, as much as possible.

[8] The paper is organized as follows. In section 2, some background is given to the validity of the station data, the gridding methodology, and the calculation of a European average temperature. Section 3 discusses and quantifies various uncertainty estimates as well as the aggregation of these uncertainties to a European average. Readers who are mainly interested in the actual European averaged temperature are encouraged to continue reading at sections 4 and 5 where the E-OBS-based European temperature is discussed and compared to other popular data sets. The study is summarized in section 6.

2. Method

2.1. Station Data Issues

[5] The station data provided by the NMHSs and other participants are not homogeneously distributed over Europe. Figure 1 shows the uneven station coverage, with the highest coverage in Germany, Slovenia, the Netherlands, Cyprus, and Switzerland. Germany has nearly 30 stations per 10,000 km² in ECA&D. The poorest coverage is found in Northern Africa and southeastern Europe with typically less than one station per 10,000 km².

[10] Although most data provided to ECA&D is validated by the responsible NMHS, they first go through a series of basic quality checks (QC) to remove obvious problems and outliers. These tests are described elsewhere [ECA&D Project Team, 2012]. No attempt is made to correct data flagged as suspect, and this data is not used in the gridding procedure.

[11] In order to construct time series which are as long and as complete as possible, a “blending” step is made in which data, which has passed the QC, from nearby ECA stations are taken to infill gaps, i.e., days with data missing or flagged as suspect. If no valid data from nearby ECA stations are available, data from nearby synoptic stations, available via the GTS (Global Telecommunication System), which have passed the QC are taken to infill the gaps. The requirement for a nearby station to be used in this blending step is that it is no further away than 12.5 km and does not differ in altitude more than 25 m. Finally, if there is less than 10 years difference between the year of the last date of the series and the current date, the series are extended with synoptic messages from the GTS from nearby synoptic stations as well. This extends the time series to a date as recent as possible if the time series are not updated by the participant on a regular basis. Since daily average temperatures (i.e., true 24 h averages) are not available as synoptic messages from the GTS, a simple average of daily maximum and minimum temperatures is used as a proxy.

[12] Haylock et al. [2008] already pointed to the problem that the NMHSs in Europe do not adhere to one standard measuring interval and/or construction of daily mean values for temperature. The effect of this problem on the quality of E-OBS is hard to quantify. Moreover, synoptic messages for daily maximum and minimum temperatures are recorded over 12 h intervals rather than the 24 h intervals used by the NMHSs. In Europe and for maximum temperatures, this is between 06.00 UTC and 18.00 UTC; for European minimum temperatures, this is between 18.00 UTC and 06.00 UTC. This discrepancy in measuring intervals can have profound effects on the homogeneity of the blended series [van den Besselaar et al., 2012].

[13] Changes in routine observation practices or station relocations may have introduced inhomogeneities in time series, leading to trends of a non-climatic origin. In ECA&D, four homogeneity tests are implemented that flag, for specified periods, the data as useful, doubtful, or suspect depending on the number of homogeneity tests it passes [Wijngaard et al., 2003]. These tests are currently not applied to the series which are input to E-OBS, since the number of series which would pass the tests as “useful” would be too low to be able to construct meaningful daily gridded maps of temperature. Obviously, the mix of homogeneous and inhomogeneous data as input to the gridding routines is a handicap of E-OBS. In Section 3.3 the effects of this handicap on the validity of European average temperatures is quantified.

2.2. Gridding Methodology

2.2.1. Gridding Daily Data Versus Gridding Monthly Data

[14] A common approach to calculate regional average monthly temperatures is to aggregate station-based (sub)daily measurements to the monthly level, calculate a monthly climatology based on this data, and subsequently grid the monthly averages with respect to the climatology [e.g., Harris et al., 2013]. Here, we will diverge from this approach by using a gridded data set with daily gridded data, which is subsequently aggregated to the monthly level.

[15] The advantage of the first, more common approach is that a less dense network of stations is required since the decorrelation length of monthly temperature data is larger than that for daily data. It is also computationally less expensive.

[16] However, the principal reason for using a gridded daily data set as the basis for the monthly European average temperature is that other climate monitoring products
produced by the authors of this study, related to daily extremes, require the daily resolution of E-OBS. Having a common basis for all the climate monitoring products (i.e., the E-OBS gridded data set) motivated us to diverge from the more common approach for the calculation of regional averaged monthly temperatures.

The high-density station network underlying E-OBS allows for the gridding of daily average temperature before the monthly average is calculated. This can be quantified with the following: An analysis of the decorrelation length of daily data for daily average temperature for station De Bilt (the Netherlands) with the E-OBS gridded daily data (over the period 1950–2011) showed the smallest decorrelation isoline just including Wales, relating to a distance of approx. 700 km. A calculation of the distribution of distances between nearest stations that provide daily average temperature in the 1960 and the 2000 situations, shows median values shorter than 25 km. For the 1960 situation, 7.9% of the stations have a nearest station at more than 100 km, and this percentage drops to 0.6% for 200 km and 0.3% for 300 km. For the 2000 situation, these percentages are 4.7%, 0.3%, and 0.1%. This shows that the distance between stations is considerably less than the distance at which the 0.6 correlation isoline is found, validating the approach of first gridding daily data and then aggregating to the monthly level.

2.2.2. Methodology for Gridding Daily Data

Hofstra et al. [2008] compared several interpolation methods for daily temperature (and other elements) and concluded that the method which involves kriging using a geographically independent variogram is the most optimal interpolation method for the daily data set. This method was described in detail by Haylock et al. [2008]; a brief summary is given here. Note that the same program has been used successfully in southeastern South America to produce gridded daily maps of temperature [Tencer et al., 2011].

At its simplest, kriging is no more than a method of weighted averaging of the observations within some region [Webster and Oliver, 1990]. The weights are chosen such that kriging minimizes the variance of the observations around the interpolating surface. In that sense, kriging is a “simple” least squares problem. However, the underlying assumption is that the station data being interpolated have variances which do not change over Europe. This is not the case with stations spanning the whole of Europe with its many climate zones. The daily data therefore need to be made homogeneous across the region.

This problem was addressed by adopting a three-step methodology of interpolating the daily data: interpolating the monthly mean using thin-plate splines to define the underlying spatial structure of the data; kriging the daily anomalies with regard to the monthly mean; and applying the interpolated daily anomaly to the interpolated monthly mean to create the final result. This is similar to universal kriging [Journel and Huijbregts, 1978], where a polynomial is fit to the underlying spatial trend. In such a large and complex region as Europe, thin-plate splines are a more appropriate method for trend estimation than polynomials. For temperature, the difference between the daily values and the monthly mean are gridded. Monthly mean temperature data are gridded using thin-plate splines, and the anomalies are gridded using kriging, which turned out to be the most suited method [Hofstra et al., 2008]. For the gridding of the monthly mean temperature, a three-dimensional spline is used, taking into account the station elevation.

The observations were first interpolated to a high-resolution 0.1° × 0.1° rotated pole mastergrid and then averaged to produce the 0.5° × 0.5° regular latitude-longitude grid used in this study. The reason for doing this is that the interpolation methods were tuned to reproduce as accurately as possible a point observation whereas the regridding aims to produce grid square averages. For temperature, the benefits of this procedure may be small but it greatly improves the quality of gridded precipitation estimates (which is also part of the E-OBS data set). The period used is 1950 to the present.
The E-OBS gridded data set also includes an estimate of the interpolation uncertainty. This interpolation uncertainty relates partly to an uncertainty in the determination of the monthly means and partly to an uncertainty in the determination of the daily anomaly fields. Both sources of interpolation uncertainty are briefly discussed here.

The thin plate splines are used in the monthly interpolation which is a stochastic method that allows an estimate of the interpolation uncertainty. The ANUSPLIN package [Hutchinson, 1995] provides these monthly uncertainty estimates, a detailed description of which is beyond the scope of this paper. Regarding the daily uncertainty, the estimate provided by the kriging is not used since this estimate only reflects station density. An alternative method is used [Yamamoto, 2000], which is based on the notion that uncertainty is higher when data from neighboring stations are more variable (and conversely, the uncertainty is low when data from neighboring stations are very similar). Yamamoto [2000] estimates the “interpolation variance” and shows that this better reflects the true error than the kriging variance. This method is applied to every grid square for every day, giving the standard error for the daily anomaly. The estimates for the monthly and daily errors are added in quadrature to produce the final interpolation uncertainty.

For the daily uncertainty, the estimate is based on the concept that a higher uncertainty is expected at an interpolated point when the neighbors are more variable. When neighbors are similar, one would expect less uncertainty. The method of Yamamoto [2000] is applied to every grid point for every day to arrive at the standard error for the daily anomaly.

The final uncertainty at a grid point was calculated by combining the uncertainties from the monthly climatology and the daily anomaly in quadrature. Haylock et al. [2008] provides more information on the calculation of the uncertainties in E-OBS.

2.3. Calculation of a European Average Temperature

Based on E-OBS version 6.0 (spanning 1 Jan 1950–31 Dec 2011) with the grid resolution of 0.5° × 0.5°, released in April 2012, a European average temperature is calculated over the domain specified in the Introduction (section 1). The reason not to use the E-OBS data at the finer 0.25° × 0.25° resolution is that this would not add much more detail to a European average temperature, being based on the same master grid as the coarse resolution file. A more detailed land-sea mask and topography may produce a slight effect; along the coasts in the Mediterranean, a finer land-sea mask will better resemble the coast line and will result in a change in the total area of “land” grid squares, which will modify the European average temperature. Due to the long coastlines of countries and islands in the Mediterranean area, this effect will be strongest there. Moreover, the high temperatures—well above the European average—will tend to amplify this effect. A finer topography will change the average temperature over a mountainous area somewhat.

From a preliminary analysis of the European averaged absolute temperature, it turned out that a strong warming trend was evident, which was related to the gradual growth of the spatial extent of the data within the European domain, related to an increase in the number of stations over time. This growth was particularly strong in Northern Africa and the Middle East, which are the hottest areas in the domain considered and artificially increased the European average temperature.

The fraction of the grid squares, which has non-missing data is lowest for January 1950 at 0.83 and highest in the 1990s where it reaches nearly 0.96. In order to prevent this non-climatic trend in the European average temperature, the grid squares in the European domain which are identified as land points but do not have temperature data in a particular month are infilled by the 1961–1990 climatological value corresponding to this grid square. Climatological monthly means from the Climatic Research Unit’s CRUTS3.1 data [Harris et al., 2013] is used for this since in E-OBS, many of these grid squares have missing data in this period. The E-OBS data is aggregated to the monthly level first before the infilling with the CRUTS3.1 data, which is at a monthly resolution. The grid of the CRUTS3.1 data is similar to that of the E-OBS data used in this study. The explanation why the CRU TS 3.1 data covers areas for which the E-OBS data lacks the data is related to the monthly resolution of the CRU TS 3.1 data. More series may be available at that resolution for those areas and monthly temperatures have a much larger decorrelation length, allowing for the gridding of a much larger area by one station than would be the case for gridding daily data—as done in E-OBS.

The use of climatological monthly mean temperatures rather than the actual values for complementing E-OBS is motivated by the wish that all information on changes in the European average temperature is from the E-OBS data set. This gives explicit knowledge which stations contribute to the European temperature and allows for selectively leaving out stations might they fail to meet additional quality criteria.

Figure 2 shows a map with, for each grid square, the fraction of months with available E-OBS data for the 1950–2011 period. It shows that the largest part of Europe does not require infilling with the CRUTS3.1 data, except Iceland and a few grid squares in the UK where approx. 10% of the months are infilled. However, in the Middle East and Northern Africa, infilling is needed, where the Algerian Sahara, parts of Libya and Egypt, and Iraq have no E-OBS data.

The use of climatological data to fill the grid squares without data dampens variability and gives more conservative trend estimates. In an alternative data set, we exchanged the climatological data with the actual CRU TS 3.1 data. The difference between the alternative data and the current data is 0.016°C with a standard deviation of 0.052°C for annual values. There is a seasonal cycle in the difference series, and the bias and standard deviation for January are −0.005°C and 0.133°C, respectively. For August, these numbers are 0.033°C and 0.067°C. The biases are smaller than the uncertainties discussed in section 3. With the expectation that future versions of the European temperature series will have a better station coverage in Northern Africa, by, e.g., the MEDARE project now involved as a participant to ECA&D, we expect the effect of infilling to reduce. The uncertainty in the European temperature related to infilling is therefore not considered further.

The procedure to calculate the European temperature now summarizes to the following steps. Calculate a monthly climatology over the period 1961–1990 from the CRUTS3.1 data. A mask file is generated using the E-OBS elevation file.
and the E-OBS temperature data, aggregated to the monthly level, which will have “missing” at grid squares where both E-OBS data and elevation are missing (e.g., ocean), the value 0 where E-OBS elevation is not missing but E-OBS temperature data is, and the value 1 where both E-OBS elevation and temperature data are not missing. The CRUTS3.1 climatology data replaces the corresponding month for the E-OBS temperature data if the value of the mask for a grid square is 0. The resulting data file, now a blend of CRUTS3.1 data and E-OBS temperature data, is then averaged over the domain, where grid squares are weighted with the cosine of their latitude.

3. Uncertainty Estimates

In this section, the likely sources which may introduce an uncertainty to the European average temperature are discussed and quantified. Some of these sources turn out to dominate the uncertainty estimate. These sources are discussed in detail in this section. Other sources are of minor importance and ignored in the final estimate of the uncertainty. These are discussed in the Appendix A to this paper.

3.1. Urbanization Effects

There are several estimates of the urbanization effects. For the Netherlands, Brandsma et al. [2003] find an urbanization trend in station De Bilt of approx. 0.1°C per century while for Central London, Jones and Lister [2009] find that warming trends are not significantly different from those at rural sites. Estimates for larger areas vary likewise, with an increase of 0.05°C from 1900 to 1990, as suggested to be present in the HadCRUT data set [Folland et al., 2001], while Parker [2004] concludes that large-scale warming due to urban effects is too small to detect.

A motivated estimate of the urbanization effects requires detailed knowledge of the station surroundings and location. This type of information is available for several stations in ECA&D but too few to reliably estimate this effect. In the absence of a European estimate of urbanization effects of temperature trends, we adhere to the estimate adopted by Brohan et al. [2006], and use a one-sided estimate of urbanization effects on the European average temperatures of 0.05°C per century, starting in the year 1950 with a value of 0.025°C.

The bias error due to urbanization is assumed to have no seasonality, which is a simplification because of the weather dependency of this effect [Brandsma et al., 2003].

3.2. Interpolation Uncertainty

An estimate of the interpolation uncertainty on a daily basis is provided along with the E-OBS gridded data set. On a daily basis and on the 0.5° × 0.5° grid, these uncertainties can be quite large, reaching up to 2–2.5°C at the edges of the grid. High station density areas have estimates of interpolation uncertainties well below 0.5°C, while the largest areas in Europe have estimates below 1°C. Nevertheless, uncertainties associated with the interpolation are up to a magnitude larger than other uncertainty sources for the daily gridded data set (discussed in the Appendix A), which has already been noted by Haylock et al. [2008, section 3.4]. For the aggregation of the uncertainty from a daily level to a monthly level and for the aggregation of grid values to a European average value, it needs to be established to what extent the daily uncertainties can be considered independent in time and if these uncertainties have a spatial coherence.

Since the E-OBS data set is based on a combination of thin-plate splines and kriging for the gridding methodology, a spatial coherence in the uncertainty estimates seems likely. These gridding methods relate the grid square value on a particular time to neighboring grid squares, introducing a spatial correlation. These methods also tend to dampen the amplitude of extreme events [Hofstra et al., 2010], suggesting a temporal correlation in the uncertainty estimate as well.
In the following, the approach is taken to aggregate the daily interpolation uncertainty maps to the monthly level first, and the monthly maps are aggregated to a monthly European average value. The time and space averaging could be reversed, except for practical reasons not favored, since the spatial averaging involves a Principal Component Analysis, which is computationally costly when using maps of daily data (discussed in section 3.2.2).

### 3.2.1. Aggregation from Daily to Monthly Uncertainty Estimates

When daily uncertainties can be assumed to be independent (serially uncorrelated, white noise), the aggregation follows the simple rule

\[
\sigma_m = \sigma \sqrt{\frac{1}{M}},
\]

where \(\sigma_m\) is the aggregated (monthly) uncertainty estimate for month \(m\), \(\sigma\) is the daily interpolation uncertainty estimate, and \(M\) is the number of days in the aggregation interval. In the case of uncertainties which are correlated, an effective sample size (ESS) \(M'\) needs to be substituted for \(M\). A simple estimate for \(M'\) is [Wilks, 1995, equation 5.12]:

\[
M' = M \frac{1 - \rho_1}{1 + \rho_1},
\]

where \(\rho_1\) is the lag 1 autocorrelation.

Next to this simple estimate of the ESS, other (more complex) estimates are available in the literature. Thiébaux and Zwiers [1984] compare a few of these estimates and test their performance. It is shown in that study that there is a considerable spread in the estimate of ESS by the various methods, indicating that estimating ESS is less than trivial.

In order to make an estimate of the ESS, we examine for a set of stations over Europe the difference time series between the actual daily station data and the E-OBS value for the corresponding grid square. This quantity will be used as a measure for the “real error”. The selected stations are located in the European region defined earlier and have (blended) data spanning over at least the period E-OBS has data (01 Jan. 1950 to 31 Dec. 2011) with no gaps, resulting in a set of 614 stations (status June 2012). In order to remove any bias in the difference series related to differences in the actual station elevation and the elevation used in the E-OBS gridding, the station data is adjusted to the E-OBS elevation using a constant lapse-rate of 0.6°C/100 m. A more appropriate way of calculating the real error would be to average the data from all stations within this grid square. However, only a minority of grid squares has more than one station within their area, and we therefore expect that this will only give a minor improvement.

The assumption implicitly made is that the real error is dominated by deficiencies introduced by the interpolation.

Figure 3 shows the ESS for the selection of stations as a function of the month (red dots). For each month, the median and standard deviation (green and blue lines) are calculated. Figure 3 also shows that there is considerable variation in ESS over the stations in Europe. For January, the ESS is relatively low in Iceland, from the south side of the Alpine region to the Adriatic coast of the Balkan and Turkey. In July, the ESS is low in Iceland, Turkey, and the most northwestern parts of Russia (not shown). Apart from this, there are stations with low ESS scattered over Europe surrounded by stations with (much) higher ESS values. Low ESS values suggest a high lag 1 temporal correlation in the difference time series between the grid square and corresponding stations series. Note that some of the spread in the ESS estimates may be linked to the sampling variability.

In the aggregation from daily to monthly maps, the averaged value of the ESS is used for all grid squares for the sake of simplicity.

A possible explanation for stations with a low ESS is that the relative coarseness of the grid (0.5 ° × 0.5 °) in combination with a high-density station network showing steep gradients might introduce a bias in the difference series of the “real error”, leading to high lag 1 autocorrelations. This may be the case in winter on the Adriatic coast and in the Alpine region, where the gradients associated with the coast or mountains are not sufficiently resolved in E-OBS.

### 3.2.2. Aggregation from Monthly Grid Values to Monthly European Averages

Aggregation of the interpolation uncertainty over Europe follows, analogously to (1), the rule

\[
\sigma_m^\text{E} = \frac{\sigma_m}{\sqrt{N_{\text{eff}}}},
\]

where \(N_{\text{eff}}\) is the estimate of the “Effective Spatial Degrees Of Freedom” (ESDOF), \(\sigma_m\) is the spatial average over the domain of the uncertainty \(\sigma_m\), and \(\sigma_m^{\text{E}}\) is the European averaged value of the uncertainty for month \(m\). The ESDOF is the effective number of independent points. This number will be considerably less than the number of grid squares due to the spatial coherence of geophysical variables. One approach to estimate ESDOF is the method proposed by Bretherton et al. [1999] based on a Principal Component Analysis.

The method uses the \(N \times N\) covariance matrix of the variances of the stations, with \(N\) stations. The formula for the effective spatial degrees of freedom \(N_{\text{eff}}^*\) is [Bretherton et al., 1999]:
where $\lambda_k$ is the $k$th eigenvalue of the $N \times N$ covariance matrix. In the case of equal $\lambda_k$ (in which no eigenvector of the covariance matrix dominates), $N_{\text{eff}} = N$ and in the other extreme case (one eigenvector dominates all), $N_{\text{eff}} = 1$, so $N \geq N_{\text{eff}} \geq 1$. Bretherton et al. [1999] offer some further interpretations of $N_{\text{eff}}$.

Figure 4 gives ESDOF as a function of the month for the real error field, based on a comparison between 614 station records and the corresponding E-OBS grid square (red). The green line gives the 5 point Gaussian smoother.

\[ N_{\text{eff}}^* = \left( \frac{\sum_{k=1}^{N} \lambda_k^2}{\sum_{k=1}^{N} \lambda_k^2} \right)^2, \]

(4)

between 0.080°C and 0.082°C. The monthly averages show the largest values in the winter months and the smallest values in June to September. This will be related to having a higher estimate of the ESDOF in the warm months than in the winter months. Interpolation uncertainty for January and February starts from approx. 0.14°C in the 1950s to decrease to approx. 0.12°C at the end of the record. The 1950 value for January is the highest at 0.16°C.

3.3. Effects of Using Non-Homogeneous Input Data

[53] In the collection of station data as input to E-OBS, no distinction is made between series which are homogeneous and which are not. It is possible to make a quantitative esti-
Figure 6. Difference in monthly mean temperatures using all available series and only those series that have passed the homogeneity tests as “useful” or “doubtful” (in red) with the annually averaged mean (green) and standard deviation (blue).

estimate of the effects of including non-homogeneous data to the E-OBS input data by pre-selecting only those series which are tested as homogeneous in the gridding script.

ECA&D includes four simple homogeneity tests [Wijngaard et al., 2003]. For temperature, the annual averages of DTR (Diurnal Temperature Range; maximum-minimum temperature) and vDTR (absolute day-to-day difference of DTR) are used as input to these tests. The use of derived annual variables avoids autocorrelation problems with testing daily series.

All four tests suppose under the null hypothesis that in the series of the testing variable, the values are independent with the same distribution. Under the alternative hypothesis, the tests assume that a step-wise shift in the mean (a break) is present or that the series is not randomly distributed (depending on the test).

The test results are condensed into a single flag: “useful” (zero or one tests reject the null hypothesis at the 1% level), “doubtful” (two tests reject the null hypothesis at the 1% level) and “suspect” (three or four tests reject the null hypothesis at the 1% level).

The category is calculated separately for DTR and vDTR. If the results are different, the least favorable category is assigned to the temperature series.

The homogeneity tests are conducted using data from the 1950–2011 period or a shorter period if the series do not completely overlap with this period.

Note that the four tests are not completely independent. Also, the use of stand-alone data rather than difference data, with a homogeneous reference series in the homogeneity tests, is a possible concern in the application of these tests.

By using data labeled as “useful” or “doubtful” only, alternative gridded data sets for daily average temperature are produced. Based on these, monthly and annually averaged European temperatures are calculated and compared to the European temperature based on all available series. For this comparison to be valid, the extent of the grids of the two data sets has to be made equal. The extent of the grid of the alternative data set is smaller than that of the data set based on all available series. Correcting this prevents the differences in the European average temperatures to relate in part to the use of different amounts of climatological data being used as supplement to cover the European domain. The percentage of the series which have been labeled as “suspect” is 65.6%, and a total of 2257 series are not used in the construction of this alternative E-OBS data set (retaining 1182 series).

Based on this analysis, annual values of the European average temperature are, on average, too warm by 0.051°C compared to the alternative homogeneous data set (with a 1 σ of 0.015°C). The highest difference between the annually averaged European temperatures using all available series and the one using the homogeneous subset of this series is found in the 1950s and the lowest in the 2000s (not shown). Figure 6 shows that there is a strong seasonal cycle in this difference, with a warm bias in summer (up to 0.123°C) and a cold bias in winter (up to −0.047°C). The standard deviation is smallest in spring and autumn and largest in winter and summer.

Averaged maps of the difference between the gridded data based on all series and based on the homogeneous series only, show that in January, the cold bias is mainly found in northwestern Spain, eastern Turkey, Sweden, and northern Scandinavia. In July, the warm bias is found in Spain, Morocco, eastern Algeria, and eastern Turkey (Figure 7). These large areas with warm or cold biases are mainly in regions with a low station density. This illustrates that single series impact on a large spatial scale in low station density areas.

Due to the vast number of series discarded in the homogeneous E-OBS data set, the estimate of the interpolation uncertainty will also be larger. This analysis does not correct this side effect, making the estimate of the uncertainty related to the use of non-homogeneous data an upper limit estimate.

Due to the spatial averaging of temperatures to the European level, homogeneity issues that are of random nature will be averaged out. A possible candidate of a systematic effect is the moving of stations in the 1950s from a site in or close to a city, to nearby airports, resulting in an overall cooling [van der Schrier et al., 2011]. Another effect may be transition from a network of manually operated stations to an automated network and a change (often associated with this transition) in the calculation of daily average temperature, but it is unclear if this will produce a non-random effect.

3.4. Combining Uncertainties

The biases due to urbanization are one sided: temperatures may be too high because of urbanization, but they will not be too low [Brohan et al., 2006]. This means that in accounting for this, the uncertainty margin at the “cold” side of the graph (the lower uncertainty bound) will increase, while at the “warm” side (the upper uncertainty bound), the uncertainty margin will not change. Similarly, the warm bias in the summer average temperatures related to the lack of homogenization in E-OBS will result in an
increase in the uncertainty margin at the “cold” side of the graph. In the winter months, where a cold bias is present, an increase in the uncertainty margin at the “warm” side of the graph is added.

[66] The upper and lower uncertainties in the monthly means of the European average temperature are therefore calculated separately by adding the associated uncertainties of the lack of homogenization, the interpolation and the urbanization in quadrature.

[67] Annual values of the uncertainty bounds are calculated as simple averages of the monthly values.

[68] Figure 8 gives the monthly and annual European average temperatures with the uncertainty estimates. At monthly resolution, the month-to-month variability is larger than the uncertainty estimates related to the lack of homogenization, the interpolation uncertainty and the bias-related uncertainty.

4. Comparison With Other Temperature Data Sets

4.1. GISTEMP, GHCN-M, and CRUTEM4v

[69] The annual averaged European-averaged (EurAvg) temperature based on the E-OBS data set, with respect to its 1961–1990 climatology, is compared with other data sets that can be used to estimate the temperature changes over Europe but are more frequently used to construct estimates of the global mean temperature (UEA/CRU’s CRUTEM4v [Jones et al., 2012], NOAA/NCDC’s GHCN-M (version 3)}
similar for all data sets. That the total area, over which EurAvg is calculated, is applied over the region de-mention versions. The mask of the E-OBS data has been "nearest neighbor approach. This makes sure that the coarse resolution data set is preserved in the other estimates, while the GISTEMP estimate is at the CRUTEM4v estimate is most of the time higher than the other estimates, while there is only 1 year where the E-OBS estimate is larger than the other three estimates.

Figure 9. European average temperature with respect to the 1961–1990 climatology as calculated by E-OBS, GISTEMP (blue), CRUTEM4v (purple), and GHCN-M (yellow). The lower plot gives the E-OBS uncertainty bounds. Figure 9 also shows that the EurAvg temperatures on the basis of these data sets have been calculated by going through the following three steps. The grid squares with missing data have been filled with zeroes (i.e., climatology) and the data sets have been regridded to the 0.5° × 0.5° E-OBS grid using a simple nearest neighbor approach. This makes sure that the “checkerboard” pattern in temperature of the original coarse resolution data set is preserved in the finer resolution versions. The mask of the E-OBS data has been applied over the region defined in section 1, to make sure that the total area, over which EurAvg is calculated, is similar for all data sets.

For the GISTEMP and GHCN-M data sets, an adjust-ment is calculated for the anomalous EurAvg temperatures to define this data set with respect to the 1961–1990 normal rather than their original climatologies (1951–1980 and 1971–2000, respectively).

The comparison is given in Figure 9 and shows that for many years, the global data sets are close to the E-OBS estimate and within the uncertainty bounds. There are no years that all three estimates are outside the E-OBS uncertainty bounds. Figure 9 also shows that the divergence between the estimates of the global data sets is largest in the first half of the 1950s and increases from the 1990s onward, which is related to having the same 1961–1990 climatological period. From the 1990s, the CRUTEM4v estimate is most of the time higher than the other estimates, while the GISTEMP estimate is at the low side. In the 1950–2011 period, there are 15 years where the E-OBS estimate is larger than the other estimates, while there is only 1 year where the E-OBS estimate is lower than the other three estimates.

Figure 10. European average temperature with respect to the 1961–1990 climatology as calculated by E-OBS, the 20th Century (green), ERA-40 (orange), and ERA Interim (blue) reanalysis. The lower plot gives the E-OBS uncertainty margins in grey boxes and the difference between the reanalysis data and E-OBS.

4.2. 20th Century, ERA40, and ERA-Interim Reanalysis

The annual averaged EurAvg temperature is compared to European averages of 2 m temperatures as calculated in popular reanalysis data sets. These are the 20th Century Reanalysis [Compo et al., 2011], which starts in January 1871 and is based on a reanalysis using surface pressure data only. The spatial resolution of this reanalysis data set is 2° × 2°. The ERA-40 reanalysis [Uppala et al., 2005] spans the period September 1957 to August 2002 at a spatial resolution of 1.5° × 1.5° and the ERA-Interim reanalysis [Dee et al., 2011] spans the period January 1979 to December 2011 (at the time of writing in June 2012) with 0.7° × 0.7° resolution.

The European average temperature is calculated for the reanalysis data sets similarly as in section 4.1, and Figure 10 shows the results, with respect to the 1961–1990 climatology. This hides the bias in the 20th Century Reanalysis; its annual average temperatures over the 1961–1990 period is 11.01°C, while that of the ERA-40 and ERA-Interim reanalysis are 10.14°C and 10.16°C, respectively.

Figure 10 shows that the EurAvg temperature and the reanalysis temperatures generally agree, with the exception of 1992 for which all three reanalysis temperatures are below the E-OBS uncertainty estimate. Interesting is that 1989 is the warmest year in the 20th Century Reanalysis (outside the E-OBS uncertainty) and that from 1994 onward, the temperatures with respect to the 1961–1990 climatology in this reanalysis have been structurally lower than those in E-OBS and the other reanalysis products. In fact, the difference increases in time and even falls outside the E-OBS uncertainty from 1994 onward with the exception of only 2 years. Apparently, the warming over Europe evident in E-OBS and the other reanalysis products over the last decade fails to be captured in the 20th Century Reanalysis.

A comparison between monthly values of the E-OBS data and reanalysis data from ERA40 and ERAInterim shows that E-OBS deviates most strongly from the reanalysis data in
June 1992 (with the reanalysis data being colder). A relatively large area with large differences is found in the desert of Algeria where temperatures in E-OBS are 1°C–2°C higher. Next to this larger region, clusters of a few grid squares with differences of up to 5°C are found south of the Alps, between the Black and Caspian Sea, in Iraq, Israel, and southern Iceland. Given the location of most differences, data scarcity is suspected as the reason for the deviation. Note that E-OBS deviates less strongly from the global data sets in 1992 than it does for the reanalysis data (Figure 9). Apparently, the reanalysis data sets produce a temperature anomaly for June 1992, which is dynamically consistent with the data sets assimilated in the models, but which are not evident in the in situ data sets.

Figure 11 shows the difference in trends between the 20th Century Reanalysis and E-OBS over the period 1979–2008. This figure shows that the largest differences between these two are in central and eastern Europe and northern Scandinavia. A provisional explanation for the divergence in estimates of the warming in Eastern Europe is related to the use of observed Sea Surface Temperatures (SST) as boundary conditions for the 20th Century Reanalysis and the use of pressure observations only in the assimilation scheme. Apparently, the further away from the sea and ocean, the less the warming signal present in the SSTs impacts on the surface temperatures. The absence of any land-based temperature measurements in the assimilation scheme of the 20th Century Reanalysis means that no temperature adjustment is made.

Note that Figure 11 also reveals some of the problems associated with the use of non-homogeneous data in E-OBS. There are isolated spots, for instance in Romania, with a large discrepancy between the E-OBS-based trends and the trends from the 20th Century Reanalysis. These issues are dealt with in newer releases of the E-OBS data set.

Simmons et al. [2004] compares the trends in 2m temperature observed from the ERA-40 reanalysis with the CRUTEM2v data set (which is an earlier version of the CRUTEM4v data set used in section 4.1) and shows that ERA-40 is close to the CRU data from 1967 onward for Europe, although the 1979–2001 temperature trend is underestimated in the reanalysis. The discrepancy between ERA-40 and CRUTEM2v before 1967 is related to a limited availability of surface observations for ERA-40 for several European countries and a warm bias in the reanalysis. Figure 10 does reproduce a persistent bias between E-OBS and ERA-40 for the 1960s over the domain used in the current study (which is different from the one used by Simmons et al. [2004]). The ERA-40 temperature is higher than the E-OBS estimate for the years before 1967 and the ERA-40 temperature in the 1960s is above the E-OBS uncertainty estimate for 5 years.

4.3. CRUTS3.1

The E-OBS based EurAvg temperature is compared to the CRUTS3.1 data set [Harris et al., 2013]. This latter data set is defined on the same grid as the E-OBS data but with a different land-sea mask; the CRU data has markedly more data at the land-sea boundary. For Europe, this means that the CRU data has much more grid squares in the Mediterranean, which would result in a warm bias in the estimate of the actual European temperatures if left unadjusted. To make estimates of the E-OBS and CRUTS3.1 data sets comparable, the E-OBS land-sea mask is applied to the CRU data. Figure 12 shows that the resulting estimates of the European average temperature are similar with CRUTS3.1 temperatures generally higher than the E-OBS temperatures. For only 3 years, the CRUTS3.1 temperature estimate is below the E-OBS value (but within the uncertainty). For 45 years, the CRUTS3.1 temperature estimate is on average lower than the CRU estimate with 0.119°C over the common period 1950–2009.

A comparison between the EurAvg temperature and the CRUTS3.1 data for the seasons shows that for winter (DJF), the variability in the European temperatures...
The warmest year in the record is 2007 (11.043°C) above climatology. The E-OBS estimate is 0.093°C warmer than the CRU estimate. In spring, summer, and autumn, the E-OBS estimate is cooler than CRU with 0.149°C, 0.238°C, and 0.158°C, respectively.

5. Monitoring the European Temperature

Seven years of the top 10 warmest years (Table 1) are from the period starting as recent as the year 2000. The earliest year of the top 10 warmest years is 1989. In the top 10 coldest years, the most recent year is 1987 and 5 years of this top 10 precede or are in the 1960s (Table 2).

Figure 13 shows the annual European average temperatures as anomalies with respect to the 1961–1990 climatology, including uncertainty estimates, based on the approach put forward in previous sections. A clear upward trend in annual average temperatures over the last few decades is visible. The most recent year with an average temperature below the 1961–1990 climatology was 1996 (at −0.004°C below climatology). The warmest year in the record is 2007 (11.043°C); the coldest year is 1956 (8.753°C).

5.1. Comparison of Trends With the Global and Northern Hemisphere Average Temperatures

The annual European average temperatures as anomalies with respect to the 1961–1990 climatology, including uncertainty estimates, based on the approach put forward in previous sections. A clear upward trend in annual average temperatures over the last few decades is visible. The most recent year with an average temperature below the 1961–1990 climatology was 1996 (at −0.004°C below climatology). The warmest year in the record is 2007 (11.043°C); the coldest year is 1956 (8.753°C).

Figure 13 shows the annual European average temperatures with respect to the 1961–1990 climatology as calculated by E-OBS with the uncertainty bounds in grey.

is very large and the two data sets agree fairly good. In winter, the E-OBS estimate is 0.093°C warmer than the CRU estimate. In spring, summer, and autumn, the E-OBS estimate is cooler than CRU with 0.149°C, 0.238°C, and 0.158°C, respectively.

Figure 13 shows the annual European average temperatures as anomalies with respect to the 1961–1990 climatology, including uncertainty estimates, based on the approach put forward in previous sections. A clear upward trend in annual average temperatures over the last few decades is visible. The most recent year with an average temperature below the 1961–1990 climatology was 1996 (at −0.004°C below climatology). The warmest year in the record is 2007 (11.043°C); the coldest year is 1956 (8.753°C).

Figure 14 shows the European average temperatures stratified by season. In winter, an increase in temperature can be distinguished, but the year-to-year variability is strong. This is confirmed in Table 1 where the top 10 warmest winters include years as recent as 2007 and 2008 but also the earlier years 1955 and 1961. Visual inspection indicates that the trend toward higher temperatures is clearer for the other three seasons, with summer showing the most striking increase. For summer, the year 2010 is the warmest on record, exceeding the number two (2003) with 0.45°C. The top 10 warmest summers all fall in the most recent 13 years. There is nearly a 3°C difference between the summer of 2010 and that of 1978, the coldest in the record.

Temperature averaged over Europe strongly depends on how Europe is defined. An alternative to the current domain relates more to the politically defined Europe and is the area bounded by 10°W–30°E and 35°N–75°N. This area, excluding Iceland, Russia, and most of Northern Africa, is also used in, e.g., the Annual Bulletin on the Climate in WMO Region VI as published by the Deutscher Wetter Dienst.

Figure 15 shows the European temperature as defined in this study and the alternative, defined on the smaller domain. It shows that the temperature defined on the smaller grid is significantly lower, by approx. 1°C. The ranking of warm years and seasons changes as well, with 2011 being the warmest year and 2007 the second-warmest year on record for the smaller domain, while 2007 and 2011 are the warmest and fourth-warmest years over the larger domain. Similarly, the summers of 2003 and 2010 are the warmest and second-warmest year, respectively over the smaller domain, while these years are reversed in the ranking over the larger domain.

The trends of the curves in Figure 15 are only slightly different. For the trend over the 1950–2010 period, the large and small domains give values of 0.179 and 0.178°C/10 years, respectively. For the 1980–2010 period, these values are 0.414 and 0.433°C/10 years.

Table 1. List of Years With the Highest Annual and Seasonal Averaged European Temperatures

|------|------|-------|--------|-------|--------|-------|--------|-------|--------|-------|
temperature the smallest. Figure 16 shows that the global land temperature is increasing less rapidly than the Northern Hemisphere and European averages. The difference in trends between the latter two averages for the last three decades (Table 3) is mainly related to the rather low values of the European temperatures in the 1980s. Comparing the trend values over the 1980–2010 period shows that Europe warms approx. 1.2 times faster than the Northern Hemisphere land and approx. 1.6 times faster than the global land.

### 6. Summary

[90] This study presents a European average temperature with monthly and annual resolution spanning the period 1950–present. The area over which this average is calculated stretches from Northern Africa (30°N) to northern Scandinavia (75°N) and includes Iceland and Eastern Europe up to 45°E. The European average temperature is based on E-OBS v6.0 at the 0.5° resolution, combined with CRUTS3.1 data to

---

**Table 2. List of Years With the Lowest Annual and Seasonal Averaged European Temperatures**

|------|------|--------|-------|--------|-------|--------|-------|--------|-------|

---

**Figure 14.** European average temperatures with respect to the 1961–1990 climatology as calculated by E-OBS including the error margins. The upper panels show winter (DJF) and spring (MAM), the lower panels show summer (JJA) and autumn (SON).
provide the 1961–1990 climatology for those grid squares for which E-OBS has insufficient data to calculate the climatology.

Quantitative estimates have been made for the various uncertainties associated with collecting, sampling, and gridding station based data as well as non-climatic trends due to urbanization and changes in thermometer screens. One aspect of the E-OBS data is that not all station data on which it is based are homogeneous. An estimate of the uncertainty in the European average temperature associated with this has been made.

Of these uncertainties, the lack of homogenization, the interpolation, and the possible bias due to urbanization dominate. The calculation of monthly averages of these uncertainties explicitly takes the effective sample size in account, which adjusts for the effects of autocorrelations in the uncertainty estimates. In the aggregation to a European average, the effective spatial degrees of freedom are used to adjust for spatial coherency in the uncertainty estimate.

A comparison between the E-OBS based European average temperature and estimates of this quantity using popular global gridded data sets (NASA/GISS’s GISTEMP, NOAA/NCDC’s GHCN-M (version 3), and UEA/CRU’s CRUTEM4v) shows that the divergence between these estimates and E-OBS is largest in the 1950s. Also, in this period the CRUTEM4v estimate is most of the time higher than the other two estimates, while the GISTEMP estimate is at the low side. In the 1950–2011 period, there are 15 years where the E-OBS estimate is larger than the other estimates, while there is only 1 year where the E-OBS estimate is lower than the other three estimates.

European temperatures and temperature trends as produced by the ERA-40 and ERA-Interim reanalysis data sets are similar as to what the E-OBS data gives. A comparison with the 20th Century Reanalysis shows that this data set has 1989 as the warmest year since 1950 (outside the E-OBS uncertainty), but fails to reproduce the warming of the past 15–20 years which is observed in Europe.

The E-OBS data, the associated interpolation uncertainty, and the European average temperature are available via the internet at http://www.ecad.eu.

Appendix A

A1. Station Errors

A1.1. Measurement Error

The random error in a single thermometer reading has been estimated at 0.2°C [Folland et al., 2001], which we will use as an estimate of the measurement error in the daily maximum (Tx) and minimum (Tn) temperatures. Daily averaged temperatures are based at least on two measurements (Tx and Tn), which would make the error in this measurement smaller (by a factor \(1/\sqrt{2}\)). Here we will focus on an estimate valid for daily maximum or minimum temperatures, too. For a monthly average, the measurement error in the monthly aggregated Tx or Tn temperature is at most 0.2/\(\sqrt{30} \approx 0.037°C\). This error will be uncorrelated with the value for any other station or the value for any other month. Aggregating this error to a European average will then make this error very small. In January 1950, there were approx. 600 stations with daily mean temperatures available for E-OBS, which is the smallest number over the timespan E-OBS covers. Assuming independence, the European averaged value will be

\[0.037°C/\sqrt{600} \approx 0.0015°C;\]

so small that this error will not be considered further.

A1.2. Calculation and Reporting Error

Station data used for E-OBS are quality checked before inclusion in the database but not adjusted. Data
are flagged when they fail the quality test and discarded for further use. However, it is possible that in reporting the data to ECA&D, errors are introduced, like reporting 29.1°C instead of 19.1°C. Many of these errors will be identified by the quality control. Those that remain, are likely to be small and because they are also uncorrelated in time and space, will have a negligible effect on the large-scale average. Following Brohan et al. [2006], this error will not be considered further.

A2. Sampling Error

[98] The sampling error will depend on the number of stations in a grid square, on the position of the stations and on the actual variability of climate in that grid square. Jones et al. [1997] give a method for calculating sampling error:

\[ SE^2 = \frac{\sigma_i^2 (1 - \rho)}{1 + (n - 1)\rho}, \]

where \( \sigma_i^2 \) is the mean station standard deviation for grid square \( i \), \( n \) is the number of stations in the grid square and \( \rho \) is the average inter-site correlation of stations within the grid square.

Note that stations in a grid square may have different baselines although they are highly correlated, for instance, if the elevation of stations is different within one grid square. Prior to application of (5), it is assumed that temperatures are reduced to the height associated with the grid square.

[99] Sampling errors will be large when the grid squares are large, like the \( 5^\circ \times 5^\circ \) used in the Brohan et al. [2006] study. For finer grids, like the E-OBS grid, the correlation between stations within a grid square is near 1, making the sampling error (equation 5) small. Furthermore, the number of stations in a \( 0.5^\circ \times 0.5^\circ \approx 50 \text{ km} \times 50 \text{ km} \) grid square is usually 1 or 0, except in the densest sampled areas. The station density in ECA&D varies from 1 station per 330 km\(^2\) for Germany, to 1 station per 51,400 km\(^2\) for Bosnia-Herzegovina (excluded in this overview are very small countries like Gibraltar and countries which are only partially in the European domain like Algeria).

[100] Evaluating the sampling error (equation 5) for the 410 grid squares, which have more than one station, shows that the average error over these grid squares varies between 0.165°C for July to 0.236°C for January. Taking these values to be representative of all grid squares and assuming that the sampling errors have little spatial correlation, the European averaged sampling error will be between 0.002°C and 0.003°C considering that the European averaged temperature is calculated over 5618 grid squares. These errors are an order of magnitude smaller than the bias errors discussed in section 3 and are therefore neglected.

A3. Thermometer Exposure Changes

[101] Most of Europe’s NMHSs have adopted the use of a Stevenson screen (one of which is an aspirated screen). For averaged monthly values, differences between the screens and the reference screen, the multi-plate unaspirated screen used at KNMI, were smaller than 0.05°C [Brandsma and van der Meulen, 2008, their Figure 1] with a few screens giving higher values throughout the year and a few giving lower values. An exception is the aspirated Young screen which showed a deviation of approx. –0.2°C in summer. It was also found that the differences in recorded temperatures increased with decreasing cloud cover and wind strength, which may be an additional problem given the “stilling” effect in wind speed observed in some studies [Vautard et al., 2010], although this effect is hard to quantify. The problem related to a change in the ventilation of the thermometer screens is not considered further.

[102] The metadata in ECA&D is not sufficiently detailed to relate the thermometer readings to the type of screen used for every European country, but ECA&D includes some pictures of the stations from recent date, provided by the NMHSs. Pictures of the stations are provided by 11 European countries in the domain considered here, of which three are now using a round-plated screen and eight (a variant of) the Stevenson screen. Given the observation that most countries in this (limited) survey are still using Stevenson screens, adopting an estimate of the error associated with the thermometer exposure change of 0.05°C as suggested by Brandsma and van der Meulen [2008], valid for the whole of Europe would be strongly exaggerated. It is more likely that only a few European countries will have seen a screen replacement which affected the temperatures to this degree, making a European averaged estimate for this error much smaller by an order of magnitude than other errors.

[103] The transition from a manual to an automated network may have introduced inhomogeneities as well. ECA&D metadata for a subset of the Israeli stations indicate that this transition (on 1 January 2005) produced differences in daily minimum and maximum temperatures in the order of up to a few 10th of a degree (information provided by Israel Meteorological Service). We expect that this effect will not be that large in the daily average temperatures because of compensating errors. On the other hand, for the Netherlands from around 1990, a new automated observing system, using small multi-plate thermometer screens, was gradually introduced replacing manual measurements from instruments housed in wooden Stevenson screens. This transition has had negligible effect on monthly mean temperatures [Brandsma and van der Meulen, 2008], but the noise levels (in terms of variability of daily differences with a reference series) of most stations were significantly reduced for some stations [van der Schrier et al., 2011]. A quantitative assessment of the effects of the transition to an automated network is not included in the E-OBS uncertainty estimate.

Acknowledgments. We acknowledge the E-OBS data set from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu). The three reviewers and the Associate Editor are thanked for their constructive comments on a previous version of this study. In some analysis, the KNMI Climate Explorer (elệmexp.knmi.nl) has been used. The GISS Surface Temperature Analysis are made available by the NASA Goddard Institute for Space Studies (http://data.giss.nasa.gov/gis temp/). The NCDC grided GHCN temperature anomalies are made available by the National Climatic Data Center (http://www.ncdc.noaa.gov/temp-and-precip/ghcn-gridded-products.php). The CRUTEM4v variance adjusted land air temperature anomalies are...
References


van der Linden, P., and J. Mitchell, eds. (2009) ENSEMBLES: Climate Change and its Impacts: Summary of Research and Results from the ENSEMBLES project, Met. Office Hadley Centre, Exeter, UK.


