Clim. Past Discuss., 6, 2517–2555, 2010 www.clim-past-discuss.net/6/2517/2010/ doi:10.5194/cpd-6-2517-2010 © Author(s) 2010. CC Attribution 3.0 License.



This discussion paper is/has been under review for the journal Climate of the Past (CP). Please refer to the corresponding final paper in CP if available.

# The construction of a Central Netherlands temperature

G. van der Schrier<sup>1</sup>, A. van Ulden<sup>2</sup>, and G. J. van Oldenborgh<sup>1</sup>

<sup>1</sup>KNMI, P.O. Box 201, 3730 AE De Bilt, The Netherlands <sup>2</sup>retired

Received: 27 October 2010 - Accepted: 29 October 2010 - Published: 10 November 2010

Correspondence to: G. van der Schrier (schrier@knmi.nl)

Published by Copernicus Publications on behalf of the European Geosciences Union.



### Abstract

The Central Netherlands Temperature (CNT) is a monthly daily mean temperature series constructed from homogenised time series from the centre of the Netherlands. The purpose of this series is to offer a homogeneous time series representative of a larger area to study large-scale temperature changes. It will also facilitate a comparison with

<sup>5</sup> area to study large-scale temperature changes. It will also facilitate a comparison with climate models, which resolve similar scales.

From 1906 onwards, temperature measurements in the Netherlands have been sufficiently standardised to construct a high-quality series. Long time series have been constructed by merging nearby stations, using the overlap to calibrate the differences.

These long time series and a few time series of only a few decades in length, have been subjected to a homogeneity analysis in which significant breaks and artificial trends have been corrected. Many of the detected breaks correspond to changes in the observations that are documented in the station metadata.

This version of the CNT, to which we attach the version number 1.1, is constructed as the unweighted average of four stations (De Bilt, Winterswijk/Hupsel, Oudenbosch/Gilze-Rijen and Gemert/Volkel) with the stations Eindhoven and Deelen added from 1951 and 1958 respectively onwards.

#### 1 Introduction

In the Netherlands, the earliest temperature observations were made at the end of the

- <sup>20</sup> 17th century. From 1706 onwards, systematic measurements were made and a continuous record, albeit constructed from several sources, exists (Labrijn, 1945). Because of the lack of standardisation in observation procedures, instruments and observations screens, the construction of a homogeneous record on the basis of these early instrumental records is difficult and is not attempted here.
- <sup>25</sup> In 1906 a climatological network had become operational in the Netherlands that employed a highly standardised observation practice and a type of Stevenson screens



at all stations but one. The exception was the station De Bilt. A set of main stations made observations on an hourly basis, while secondary stations in the climatological network made thrice daily measurements. Around 1950, a new synoptic network was installed. This was operated by the Weather Forecasting department of KNMI in paral-

Iel to the climatological network, which was operated by the Climate Division at KNMI. This situation persisted until around 1990, when the two networks were integrated to form a single, fully automated, observation network.

The locations of the observing stations relevant for this study, both the stations that ceased operation and the operational ones, are shown in Fig. 1.

The aim of this study is first to construct a set of homogeneous monthly averaged records for daily mean temperature at various locations. These records are either based on long continuous records from the KNMI network or, when these are not available, on combinations of two records from nearby stations to obtain time series as long as possible. Next, based on a selection of these homogeneous records, a Central Netherlands Temperature (CNT) record is assembled that is, by construction, representative for a larger area.

In a precursor of this study, van Ulden et al. (2009) based an earlier version of the CNT on the same monthly averaged station records but used a different method to homogenise these records and made different choices regarding construction of the

- 20 reference series, aggregation levels, window size etc. in the application of their method than what is done here. Despite these different approaches, we will show that the locations and sizes of most of the detected breaks in the station records are similar between the current study and that of van Ulden et al. (2009) and that consequently the differences between the CNT as presented in this study and the one presented by
- van Ulden et al. (2009) are very small.



#### 2 Construction of long records

At the secondary stations in the climatological network operating since the early 20th century, temperature readings were made at 08:00, 14:00 and 19:00 LT as well as the minimum and maximum temperatures reached in the time period 19:00-19:00 LT.

<sup>5</sup> Based on these measurements, van der Hoeven (1992) made accurate estimates of daily (00:00–00:00 LT) mean temperature. Note that only for the main stations in the climatological network, minimum and maximum temperatures reached between consecutive readings rather than those reached in a 24-h period, are recorded.

The approach of Van der Hoeven (1992) was to make five estimates of the daily mean temperature  $T_{24}$ :

$$T_{24}^{(i)} = T_i - C_i(t)(T_x - T_n), \quad i = 1, 2, 3, x, n$$

where subscripts 1, 2, 3 refer to temperature readings at 08:00, 14:00 and 19:00 LT and subscripts *x* and *n* to daily maximum and minimum temperatures. The coefficients  $C_i(t)$ were seasonally dependant and obtained from a comparison with 24-h temperature observations at De Bilt in the period 1961–1970. The seasonal dependence of the coefficients is introduced to account for the annual variations in the times of sunset and sunrise and are computed for the 36 decades of the year. Finally, the five estimates of the daily mean temperature were averaged to give the best estimate:

$$T_{24} = (T_1 + T_2 + T_3 + T_x + T_n - (C_1 + C_2 + C_3 + C_x + C_n)(T_x - T_n))/5.$$
 (2)

<sup>20</sup> Applying this approach to the De Bilt data, revealed that the value estimated from these five measurements agreed with the true 24-h mean to withing 0.006 °C.

From 1970 until the introduction of the automated weather systems, ten measurements a day were made at the secondary stations: eight at three-hourly intervals plus the minimum and maximum temperatures. Twenty-four hour averages were made by an unweighted average of these values.

The secondary stations involved in this study are indicated with a "G" in Table 1.

(1)

Stevenson huts were used in all stations until about 1990, with the exception of De Bilt until 1950. From around 1990 a new automated observing system was gradually introduced using small multi-plate thermometer screens. This transition has negligible effect on monthly mean temperatures (Brandsma and van der Meulen, 2008).

- <sup>5</sup> Most records in Table 1 were complete. Den Helder and Sittard had 9 missing months, Winterswijk had 1 missing month, Hoorn had 6 missing months and Eindhoven missed May and June 1952. These missing data were filled with data from alternative stations with a monthly adjustment to account for the any climatological difference (see Tables 1 and 2). The record from Vlissingen appeared to be too in-10 complete to be useful for this study. All records are tested for outliers, but none were found, with the exception of Eindhoven which is discussed in Sect. 6.12. Eight long
- records are constructed covering the period 1906–2008 by merging the records with records from nearby stations (see Table 3). The older parts of these merged records were adjusted to the recent parts using overlapping observation periods. The monthly adjustment factors were smoothed with a 5-point guasi-gaussian filter.

For the transition Winterswijk to Hupsel the overlapping period was only 10 months, which is too short for a reliable estimate of the adjustment factors. Therefore we used a 10 yr overlap of both time series with Deelen to determine the adjustments. The smoothed adjustment factors are shown in Fig. 2. We see in this figure that the adjustment factors are all negative, meaning that the recent stations are cooler than the older

stations. This may be related to an urban heat island effect for stations like Maastricht and Groningen, where observations were made in the centre of the city in the earlier period and at the airport later on. In Maastricht, the modern station at the airport is also located at a much higher and exposed position.

#### 25 3 Method

20

The approach taken to identify possible changepoints and estimate the size of the break is based largely on the two-phase regression technique suggested by Vincent



(1998). Potential discontinuities are detected on 40-yr sliding windows of the difference time series between the target series and a (homogeneous) reference series. The construction of the reference series is discussed in Sect. 4. Both the reference series and the target series have their seasonal cycles removed, following the approach by Menne

- <sup>5</sup> and Williams Jr. (2009). The seasonal cycle is calculated using the complete period over which there is data. Easterling and Peterson (1995) note that a windowing technique may obscure discontinuities which are close in time, but having a sliding window with increments of one year will (at least partially) eliminate this problem. In order to prevent this problem, homogenised records are put through the detection algorithm to detect possible inhomogeneities which were left undetected in the first run. For the
- Eindhoven record, this procedure yielded a breakpoint which remained undetected in the first run of the program.

The two-phase regression technique applies four different models. The first model determines whether the time series is homogeneous over the tested interval of time. If the time series fails to pass this test, three different models are used to estimate the location and size of potential step. Every year in the record is tested as a potential step, with the exception of the first and last three years.

To test if a series is inhomogeneous, a straight line is fitted to the data. The goodness of fit is quantified using the Durbin-Watson statistic, which is a test for the correlation of regression residuals (Wilks, 1995, Sect. 6.2.6). It tests the null-hypothesis that the residuals are serially independent against the alternative that they are consistent with a first-order autoregressive process. The threshold for the Durbin-Watson statistic relates to the 5% level where the null-hypothesis of zero serial correlation can either be rejected or where this statistic is indeterminate. Upper and lower bounds for the significance of the Durbin-Watson statistic are calculated using the NAG routine g01epf.

If the difference series is judged inhomogeneous, the location and size of the break are estimated using a simple two-phase regression model (Vincent, 1998)

 $X_t = \begin{cases} \mu_1 + \varepsilon_t, \ 1 \le t \le c\\ \mu_2 + \varepsilon_t, \ c < t \le n, \end{cases}$ 

(3)

where  $\mu_1$ ,  $\mu_2$  are mean values before and after the break and  $\varepsilon_t$  is the zero-mean independent random error with a constant variance  $\sigma_{\varepsilon}^2$ . The time *c* is called a changepoint if  $\mu_1 \neq \mu_2$ . The *F* statistic for a changepoint at time *c* is:

$$F_c = \frac{(\text{SSE}_{\text{red}} - \text{SSE}_{\text{full}})/1}{\text{SSE}_{\text{rm}}/(n-2)}$$

<sup>5</sup> where SSE<sub>full</sub> is the sum of squared errors of the "full" model (Eq. 3), which includes the break and SSE<sub>red</sub> is the sum of squared errors of the "reduced" model which assumes a constant mean.

Slightly more complex is the two-phase regression model with a common trend (Wang, 2003):

10 
$$X_t = \begin{cases} \mu_1 + \beta t + \varepsilon_t, \ 1 \le t \le c \\ \mu_2 + \beta t + \varepsilon_t, \ c < t \le n, \end{cases}$$

where  $\beta$  is the value of the trend. The *F* statistic for a changepoint at time *c* is:

$$F_c = \frac{(\text{SSE}_{\text{red}} - \text{SSE}_{\text{full}})/1}{\text{SSE}_{\text{full}}/(n-3)}.$$

The fourth model allows for a combination of a discontinuous trend and a step (Lund and Reeves, 2002):

$$X_{t} = \begin{cases} \mu_{1} + \beta_{1}t + \varepsilon_{t}, \ 1 \le t \le c \\ \mu_{2} + \beta_{2}t + \varepsilon_{t}, \ c < t \le n, \end{cases}$$

with  $\beta_1$ ,  $\beta_2$  values of the trend before and after the break. The *F* statistic for a changepoint at time *c* is:

$$F_c = \frac{(\text{SSE}_{\text{red}} - \text{SSE}_{\text{full}})/2}{\text{SSE}_{\text{full}}/(n-4)}.$$

Discussion Paper 6, 2517-2555, 2010 **Central Netherlands** temperature **Discussion** Paper G. van der Schrier et al. **Title Page** Abstract Introduction Conclusions References Figures Discussion Paper Tables Back Close Full Screen / Esc **Discussion** Paper Printer-friendly Version Interactive Discussion

(4)

(5)

(6)

(7)

(8)

Under the null hypothesis of no changepoints and assuming Gaussian errors  $\varepsilon_t$  in models (3, 5, 7), tables with the  $F_{max}$  percentiles are given by Jarusková (1996), Wang (2003) and Lund and Reeves (2002), respectively. The 95% significance level is used as threshold to determine if a break is significant or not.

If a fit of a model fails to meet the significance level using the Durbin-Watson statistic, it is not considered further.

A review of modern methods, including the methods used here, is given by Reeves et al. (2007) in which is concluded that the common trend two-phase regression model seems optimal for most time series.

- A hierarchy is used in determining which changepoint models (3, 5, 7) are used to estimate step sizes. If a difference series is inhomogeneous, models (3) and (5) are applied and information from model (7) is only used after a visual confirmation that a discontinuous trend is present. No distinction is made in information on the step size from models (3, 5), the estimate of the continuous trend from model (5) is not used to correct for this trend. The motivation for not correcting for a continuous trend is related
- to the construction of the reference series. Both the construction of the reference series and this motivation are discussed in Sect. 4. However, discontinuous trends (which are output from model 7), are corrected for.

#### 4 Reference time series

5

- In the absence of homogeneous time series, constructing a (near-)homogeneous reference series requires a special approach. Instead of using a tailored approach where an average of a small number of selected time series is used as reference series from the vicinity of the target record, we use the most dominant mode of variability from a Principal Component Analysis (PCA) which is based on all available long time series ries. The first mode of variability accounts for the maximum amount of joint variability
- of the variance-covariance matrix (Wilks, 1995) which is based, in its turn, on a selection of long station data homogeneously spread over the country. The principal mode



of variability will be a weighted average of the input series and contains a large fraction of the common variability of the series. This time series will have the warming trend common to all time series and, due to the averaging of all available long records, inhomogeneities in the individual records are damped. However, a reference series

- <sup>5</sup> constructed from time series scattered over the country will not reflect any regional signal. Other considerations are that the tailored approach is more labour-intensive and can hardly be automated, but the expectation is that a tailored approach will provide a reference series which captures more of the month-to-month variability which reduces the noise in the difference series. This should make it easier to detect smaller
- <sup>10</sup> breaks. The PCA-based method is a rather straightforward procedure, easily automated, but potentially less suited to homogenise a series which is not near the centre of the country. The decorrelation length of the interannual variability of monthly mean temperature around the Netherlands varies from about 1000 km in summer to 2000 km in winter, which is much larger than the size of the Netherlands, so regional effects are not expected to be very large.

Input to the PCA are the time series of the stations De Bilt, Groningen/Eelde, Winterswijk/Hupsel, Maastricht/Beek, Volkel/Gemert, Oudenbosch/Gilze-Rijen and Den Helder/De Kooy. This excludes the station Vlissingen in the south western corner of the Netherlands, which is too gappy and has seen too many relocations to be allowed in the reference time series.

20

Over the more recent period, from the 1950s onward, more stations are available to construct a reference series. In order not to introduce inhomogeneities in the reference series, we did not include these stations in the reference series. Moreover, the seven stations used for the reference series over the first part of the 20th century are scattered

<sup>25</sup> around the country and should be sufficiently able to pick up on large-scale variations of temperature.

When using a weighted sum of series as a reference series, a correction has to be made when one of the series which is input to the PCA analysis is adjusted. The reference series is written as



$$\sum_{i=1}^N c_i A_i(t),$$

5

where  $c_i$  is the weight associated with the *i*-th series and  $A_i(t)$  is the *i*-th series at time *t*. The sum of the weights  $c_i$  equals 1. When the adjustments to the *j*-th series itself needs to be calculated, a reference series excluding the *j*-th series would be required. The difference from which an adjustment is computed is:

$$\Delta T = A_j(t) - \frac{1}{1 - c_j} \sum_{i=1, i \neq j}^N c_i A_i(t).$$

By writing Eq. (9) as

$$\sum_{i=1}^{N} c_i A_i(t) = \left(\sum_{i=1, i \neq j}^{N} c_i A_i(t)\right) + c_j A_j(t),$$

we see that

10 
$$\Delta T = \frac{1}{1 - c_j} \left( A_j(t) - \sum_{i=1}^N c_i A_i(t) \right).$$
(12)

One can therefore just use the total reference series and multiply corrections by  $1/(1 - c_j)$  afterwards.

Peterson et al. (1998) note that the construction of a reference series by simply averaging series from surrounding stations has been done earlier by Potter (1981), <sup>15</sup> although he used an average of 18-stations for this. More specifically, Peterson and Easterling (1994) average the three best correlating series from the 5 nearest stations to build the reference series and a similar approach is taken by van Ulden et al. (2009). Using a PCA to determine the weights for the individual time series in a construction the reference series seems a new approach.



(9)

(10)

(11)

To test the merits of the PCA-based approach, the reference series used by van Ulden et al. (2009) and the PCA-based reference series are compared for a selection of stations. The RMS between the target series and either of the two reference series is calculated (correcting for any offset). It turns out that the RMS's are very similar (not shown), indicating that for this study involving stations that are much closer to each other than the decorrelation length, a PCA-based reference series does not give

higher noise-levels in the difference series compared to a tailored approach. The different models discussed in Sect. 3 are combined in the approach of this study which has merit in cases where the "true" regression model is unknown (Reeves et al.,

- 2007). Since the reference series used in this study only holds information associated 10 with country-wide spatial scales and is not specifically pin-pointed at a certain region, we expect that difference series may have a continuous trend throughout the record. This makes the use of model (5), which includes a step and a continuous trend, particularly suited to this approach.
- The principal mode explains 96.7% of the variability, and includes the warming trend. 15 The dominant mode of variability is a weighted average of the input series, with the weights shown in Table 4. The weights for the various stations are very similar, the relative difference between the extremes ((maximum-minimum)/maximum) is only 0.16. The largest weights are found in Maastricht/Beek and Winterswijk/Hupsel, the lowest

is found in Den Helder/De Kooy, located at the North Sea coast.

#### **Quality check** 5

5

25

In order to assess the quality of the various records used in this study, a running standard deviation of the difference of the annual average of each series with the reference series is shown over 41 yr sliding windows (Fig. 3). The standard deviations vary considerably with time. Figure 3a shows that Sittard has a maximum in the first few decades of the record, which is in part related to a warm bias (not shown). Gemert has a very pronounced peak around 1950, which is related to a very significant break



in that period (discussed in Sect. 6.6). The running standard deviation for the series composed of Oudenbosch and Gilze Rijen has a maximum in the late 1960s which is related to a break in that period (discussed in Sect. 6.4).

De Bilt, which is situated in the center of the Netherlands, has low standard deviations for the whole observation period. This relates to the fact that the reference series reflects climatic conditions of the central part of the Netherlands best.

Figure 3b shows that Soesterberg, despite its central location and Twenthe show high noise levels.

Around 1990, the noise levels of all stations (except for Beek) are significantly reduced (not shown). This is probably related to the transition to an automated network and improved observation practices.

#### 6 Detected breaks and trends

# 6.1 De Bilt

10

The tests indicate breaks in annual averaged values near 1918–1920, near 1950, near 1968–1969 and near 1976. There is no information from the metadata which could explain the breaks near 1918–1920 and 1976. The break at 1950 is clearly related with a string of modifications which happened in the early 1950s. On 17 May 1950, the large thermometer screen (the so-called "Pagoda") was replaced by a Stevenson screen. On 16 September 1950, a relocation westwards had taken place, subsequently
20 followed by a 300 m relocation southward on 27 August 1951.

The break at 1968–1969 is most probably related to the uprooting of the nearest part of the neighbouring orchard (which was turned into a parking lot) at the end of March 1967 and/or the construction of a new 25 m high KNMI office which started in November 1967 and had reached its highest point on 2 January 1969.



The only metadata which could possible relate to the detected break at 1919–1920 are routine replacements of thermometers. The detected break near 1918–1920 is left uncorrected due to the absence of metadata indicating possible causes for this jump. After correcting for the 1950 and 1968 jumps, a new run through the program did detect the possible inhomogeneity at 1976 but it failed to exceed the 95% significance

detect the possible inhomogeneity at 1976 but it failed to exceed the 95% significance level.

Brandsma (KNMI, personal communication) homogenised the De Bilt record using more physical methods rather than the statistical approach used here. He corrected for the replacement of the screen to the Stevenson screen and its transition (September

- 10 1950), a relocation of the Stevenson screen (August 1951), the lowering of Stevenson screen from 2.2 m to 1.5 m (June 1961) and the transition of artificial ventilated Stevenson screen to KNMI round-plated screen (June 1993). Finally, Brandsma corrected for a warming trend of 0.11 °C per century caused by urban warming. The correction factors Brandsma used for the changes around 1950 and the correction factors used in
- this study are shown in Fig. 4 and are very similar. However, the use of smoothed adjustment factors in this study rather than the original, more noisy monthly break-values, makes that the adjustments for the De Bilt record associated with this break are more conservative than those by Brandsma. The additional correction for the break around 1968 and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction of a possible warming trend makes the two records different and the lack of a correction differe
- <sup>20</sup> different. The RMS difference between annual values of the De Bilt record obtained using the adjustments detected statistically in this study and the one homogenised using Brandsma's method is 0.1 °C.

#### 6.2 Den Helder/De Kooy

No inhomogeneities are detected.

<b>Discussion</b> Pa	6, 2517–2	<b>2</b> D 555, 2010				
per	Central Ne tempe	Central Netherlands temperature				
Discus	G. van der S	G. van der Schrier et al.				
sion Pa	Title Page					
aper	Abstract	Introduction				
—	Conclusions	References				
Discu	Tables	Figures				
ssior	I	►I				
) Pap	•	•				
)er	Back	Close				
– Dia	Full Screen / Esc					
scussi	Printer-friendly Version					
on Pa	Interactive Discussion					
aper		BY				

## 6.3 Groningen/Eelde

No inhomogeneities are detected in the annually averaged values. However, evaluating the months separately gives possible inhomogeneities in the years 1952–1953, 1973 and 1996. The 1952–1953 break has an amplitude of roughly equal size for July and

<sup>5</sup> December, but of opposite sign. The metadata indicate that on 28 February 1952 the thermograph was corrected with 1°C and that it has been replaced on 2 June 1953. The evidence from the metadata is judged too scanty to justify a correction for this break.

On 1 May 1973 the measurement field was relocated to the west side of the runway. This move coincides with changing measurement practice from manual to the use of electronic equipment. The Groningen/Eelde series is adjusted for this break.

No metadata indicating a possible cause for the 1996 break could be found. Apparently, some suspicion at the KNMI staff of the time must have existed, since on 4 October 1996, a comparison is made between the temperature sensors and a calibrated sensor. No deviations were found though. The series is not corrected for this detected break.

#### 6.4 Oudenbosch/Gilze Rijen

Breaks in the Oudenbosch/Gilze Rijen series are detected around 1946–1948, 1966–1967, 1971 and near 1984.

- <sup>20</sup> The metadata indicates that corrections on the minimum and maximum recording thermometers changed frequently and significantly around the 1946–1947 period, but these should not affect the daily averaged temperatures. No metadata has been found which might explain the break in daily averaged temperature around this time. This break is left uncorrected.
- The break around 1966–1967 is likely associated to two changes: the construction of a paved road on the southeast side of the terrain and the related uprooting of high trees which made the surroundings more open and the move of the instrument field



25

1400 m southward, further away from trees in a more exposed setting in July 1966. The series is adjusted for this break.

The 1971–1972 break has only a detectable break for March, August and September, not in the other months and not in the annual mean values. The only metadata indicating a possible cause for this break was the reinstalment of the thermometer screen following 1971 "new style" specifications at 5 October 1970. Furthermore, KNMI noted in their logs around this time that a garage was erected in the vicinity of the location. No corrections are made for this break.

The metadata for the years around 1984 only indicate routine maintenance and replacements to the instruments, the most profound being a replacement of the thermograph on 19 March 1983 due to a bended pen arm and adjustments to the thermograph on 6 October 1983 and 29 November 1984 by 1.0 °C and 0.5 °C respectively. Thermographs were replaced approximately once a year, which makes it improbable that these defects relate to the observed breaks in the record. However, this latter break was large and robust enough to justify adjustment in the absence of metadata.

### 6.5 Winterswijk/Hupsel

Large breaks in the Winterswijk/Hupsel record are detected at 1940 and 1950. Much smaller breaks are detected near 1960 and 1970–1972. The break at 1940 is possibly related to a relocation at 12 March 1940 to a more ideally located site. The new site is open, facing the observer's house to the north and a meadow to the south, but the thermometer screen is placed between two shrubs. On 27 February 1950, the thermometer screen was relocated 5 m eastwards away from the shrubs. The growth of the shrubs might have introduced an artificial trend in the data. The annually averaged difference record does show a trend over this 10-year period, but it is small (approx. 0.18 °C in 10 years) and has been left uncorrected.

The breaks near 1960 and 1970–1972 have been left uncorrected due to absence of metadata which could be related to the break.



Interestingly, a break is detected near 1984–1985. Based on the F-statistic, this break is not significant and is not corrected for. However, it seems to be related to a modest relocation of the station some 50 m in SW direction on 27 March 1985.

#### 6.6 Gemert/Volkel

A large break and a discontinuous trend are detected around 1950. The break and discontinuous trend are obvious from a visual inspection of the difference record (Fig. 5). This is clearly associated with the reinstallment of the station on 27 September 1949. Preceding this reinstallment, reports have been made (in June and July of 1949) indicating that the site did not meet regulations regarding the surroundings. The clearing was too small for proper ventilation. Additionally, the height of the thermometer screen was not according to regulations.

Having identified the combination of a step and a discontinuous trend for 1950, estimates of the adjustments from the other model, Eq. (7), are to be used. It turns out that the estimates of the size of the trend for the 1950 break, calculated for each sliding

- <sup>15</sup> window, show much variability. This is probably related to the relatively high year-toyear variability of the difference series in relation to the modest trend. Because of this, the few estimates (< 10%) of a negative trend were not used in the final estimate of the trend, nor were estimates used which showed trends of > 0.5 °C/10 year (7 instances). The record is corrected for the break and the discontinuous trend.
- <sup>20</sup> There is some discussion on the validity of the measurements from Volkel airbase. When the transition between the Volkel and Gemert records is set at 1980, then the homogenisation procedure detects a break and discontinues trend at this year, with a negative trend in the difference series after 1980. No break nor discontinuous trend is detected when the transition between the series is set at 1990. The poor quality of
- the measurements from the 1980s of the Volkel air base (not shown) may be related to the spurious trend. The introduction of automated measurements in the early 1990s will have improved the quality of the data. With the nearest tree line at 245 m of the thermometer screen in N-NW direction, with trees of heights between 18–20 m, the



situation at the observation site is in line with WMO regulations for a meteorological observation site.

No corrections are made to the Volkel record which is blended to the Gemert record from 1990 onwards.

#### 5 6.7 Maastricht/Beek

The Maastricht/Beek record shows a break in 1931 in the annually averaged values, but fails to show significant values in an analysis for each month separately. The Maastricht observations were made on top of a tower (at approx. 20 m above street level) on a building in the centre of the city. From 1 July 1951, parallel measurements were made at the outskirts of the city in a garden area which show considerable differences with the Maastricht observations. The suboptimal setting of the Maastricht station and the relocation to a new site some 65 m higher in altitude may be the principal reasons for having the largest adjustment factors associated with the move to a more ideally located setting (Fig. 2).

#### 15 6.8 Twenthe

Twenthe shows a break in 1969 in the annually averaged values, but significant values fail to show in an analysis for each month separately. Twenthe is a military air base and no metadata exists which might substantiate this break. No adjustments are made to this record.

#### 20 6.9 Hoorn

Hoorn shows a break which is barely significant in 1948. Initially, the observation site was located on a terrain for agricultural use. On 1 November 1946, the observation terrain was relocated to the gardens of the local slaughterhouse, facing nearby buildings in southeast to southwest directions. On 21 November 1947, measurements ceased



and on 28 April 1948, the station was relocated back to its original terrain. No adjustments are made for this break.

Another break is detected around 1970–1973. The metadata indicate that a school was build at a distance of approximately 20 m from the observation site in the early 1970s and a relocation in SW direction of 15 m was effectuated on 2 July 1971, but it is unclear if this rather modest relocation could explain the break to warmer conditions. No correction for this break are made to the Hoorn record.

### 6.10 Schiphol

5

25

The Schiphol (Amsterdam International Airport) record shows a significant break in 1981. The Schiphol metadata is not very clear about the cause for this break. Presumably it is related to a relocation of the measurement field from the vicinity of the main buildings to the outer edge of the Schiphol area, near a runway. The KNMI archives holds a map for the situation around 1960 and one for 1986 from which a change in position of the measurement field is evident but a more precise timing of the relocation cannot be given. Given the rapid growth of the airport, this period is likely to have seen more than one relocation of the instrumentation. The record is adjusted for this break.

#### 6.11 Deelen

In the period 1954–1957, measurements were taken during weekdays only at Deelen airbase, reliable monthly means could only be constructed for January 1958 onwards. Breaks are detected near 1962 and around 1984–1985. However, the metadata from this military airport provided no leads to what might have caused these breaks. The breaks are large enough to warrant adjustment.

In the annual averaged values for Deelen and to a lesser extent in the February monthly means, a discontinuous trend is detected with a break in 1977. Before 1977, a distinct upwards trend is detected indicating that Deelen warms faster than



the reference record, after that year the warming trends are similar. Again, there is no indication what might be the reason for the discontinuous trend and it is left unadjusted.

#### 6.12 Eindhoven

A curiously low value for annual averaged temperature for Eindhoven airport is observed for 1952, which is 0.95 °C lower than the reference. A comparison with adjusted records for Oudenbosch/Gilze Rijen and Gemert/Volkel indicated that the monthly averages for the month May to July were up to 3 °C lower than surrounding stations. The monthly averages of 1952 for these months were replaced by an average of the corresponding months of the adjusted records of Oudenbosch/Gilze Rijen and Gemert/Volkel. The metadata indicated that observations from 1 May 1952 onwards were made by airforce personnel rather than civil servants from the aviation authority,

and reports of KNMI inspectors of the mid-1950s complained of many false readings. Breaks are detected at 1969–1970 and 1986–1988. The exact timing of the latter break is vague, non-significant breaks are also reported for 1985, but strangely enough

<sup>15</sup> none for 1987. The metadata provided no information on the possible origin of the first break. The relocation to a new terrain on 3 July 1984 may be related to the latter break. There is some indication of a break in the month of May only, around 1980–1981;

only 3 sliding windows indicated this break. No metadata had been found which might account for this break and it is left uncorrected.

After corrections for the 1969–1970 and 1986–1988 breaks, the Eindhoven series is put through the break detection script again and this yields a break in 1980–1981 and a newly detected break in 1958, which was apparently not large enough (in terms of the *F*-statistic) to be reported in the uncorrected series. This break must be related to a station relocation at 17 October 1958. Before this date, the site did not meet KNMI specifications. This break is corrected for.



#### 6.13 Rotterdam

No breaks are detected at Rotterdam.

#### 7 The Central Netherlands Temperature

#### 7.1 Definition

<sup>5</sup> The Central Netherlands Temperature (CNT) record is based on homogenised monthly means of daily averaged temperatures from a selection of series from the central part of the Netherlands. These series are from De Bilt, Winterswijk/Hupsel, Oudenbosch/Gilze-Rijen and Gemert/Volkel. The record from Eindhoven is included from 1951 onwards and Deelen is included from 1958 onwards. The CNT is a simple unweighted average of these records. Monthly adjustments were applied to the CNT prior to the inclusion of the Eindhoven record in 1951 and to the CNT record from 1951 to the inclusion of Deelen in 1958 to account for the transition from 4 to 5 to 6 stations.

15

Gaussian filter, similar to the adjustments in Sect. 2 and are small at  $O(0.01 \degree C)$ . The long records from the coastal station Den Helder/De Kooy, the series from Groningen/Eelde and Leeuwarden in the north and Maastricht/Beek and Sittard/Beek in the south of the Netherlands have not been included since they are too far at the outer extremes of The Netherlands and are therefore less representative of the Central

These adjustments are calculated over the 1961–2008 period, smoothed by a 5-point

Netherlands. The principal motivation not to include the records from the airports of
 Rotterdam and Amsterdam (Schiphol) is that these stations are relatively close to the sea and that they may be influenced by the large cities in their vicinity and in the case of Schiphol the rapid development of the airport itself.

Other long records in The Netherlands have either been discontinued (Hoorn, Soesterberg) or are too gappy (Vlissingen) to be included in the CNT. The running



standard deviations discussed in Sect. 5 and Fig. 3 indicate that the records from Twenthe and Sittard are too noisy to be included.

Although the CNT is constructed from stations roughly in the area centred in the central-south east of the Netherlands, a correlation analysis between winter (DJF) and

- <sup>5</sup> summer (JJA) averages of the CNT and similar averages of the E-OBS gridded dataset based on averaged temperature from surrounding stations from the European Climate Assessment & Dataset (Haylock et al., 2008) (Fig. 6a,b), shows that the CNT is representative of a much larger area. Correlations remain high for stations in the Netherlands, including those in the north and along the coast and high correlations are found
- <sup>10</sup> in Belgium and Germany as well. Monthly mean values show smaller correlations, especially in summer.

#### 7.2 Comparison with an earlier version

Preceding this study, van Ulden et al. (2009) constructed an earlier version of the Central Netherlands Temperature record. This earlier version, to which we attach the ver-

- <sup>15</sup> sion number 1.0, is based on the same selection of series as the current version 1.1. However, the homogenisation procedure between the van Ulden et al. (2009) study and the current study is different. These differences relate to the construction of the reference series (see Sect. 4) but also to the break and spurious trends detection algorithms. For the construction of CNT1.0, homogeneity tests based on Easterling and
- Peterson (1995) for the detection of breaks were used and a method based on that of Alexandersson and Moberg (1997) is used to detect spurious trends. In contrast to Easterling and Peterson (1995), van Ulden et al. (2009) used moving windows, both for breaks and trends. In both tests, the critical significance levels derived from Alexandersson and Moberg (1997) are used.
- Because of the differences between the current approach and that of van Ulden et al. (2009), differences in the homogenised versions of the underlying station data can be expected. Figure 7 shows the adjustments made to the records in comparison with the adjustments made in version 1.0. Below we will argue that the differences between



CNT1.1 and CNT1.0 are very modest, despite the differences in approaches, which adds to the robustness of the CNT.

Figure 8 shows the RMS between CNT1.1 and CNT1.0 as a function of the month. The RMS has been determined over the timeperiod 1906–2008. This figure shows that

- the RMS is near 0.03 °C, except for late spring and for July to September. In the first period, the RMS rises to nearly 0.04 °C, in the second to approximately 0.07 °C. The rise in April-May can be attributed to differences in the homogenisation of the Eindhoven record (Fig. 9). The current study has an additional correction for a break in 1958. The rise in the period July to September is mainly attributed to the Winterswijk/Hupsel,
- <sup>10</sup> Deelen and Gemert/Volkel series (Fig. 9). In the homogenisation of the first record (Sect. 6.5) a small break in 1960 is detected, but we have not corrected for this break due to the absence of metadata related to this break. However, van Ulden et al. (2009) does correct for this break and the largest amplitude of the correction can be found in the months June–September. In the Deelen record, breaks are detected which are cor-
- <sup>15</sup> rected for in the current study, but not in the series used to construct CNT1.0. Finally, the adjustments to the Gemert/Volkel record are slightly different in the current study compared to that of van Ulden et al. (2009).

Regarding the trends, Table 5 shows that the differences between CNT1.0 and CNT1.1 are very small.

#### 20 8 Conclusions

Climate models compute meteorological variables at a typical scale of 100 km. Local effects caused by vegetation, small lakes or small variations in altitude, are not resolved by the models. In order to make a sensible comparison between model output and observations, the latter need to be defined at a spatial scale similar to the model results.

<sup>25</sup> The Central Netherlands Temperature record (CNT) has been designed to meet this demand. Additionally, the CNT is expected to be of interest for climate research, being based on high-quality homogenised records and representative for a larger area.



The CNT is based on a selection of homogeneous monthly averaged records for daily mean temperature from the KNMI network. Long records have been constructed by blending data from nearby station, using the overlap period to calibrate the differences. This resulted in nine records which start in the early 1900s. Using seven of these records in a Principal Component Analysis, a weighted average of these seven series is obtained which contains a large fraction of the common variability of the series. This time series contains the warming trend common to all time series and, due to the averaging of all available long records, inhomogeneities in the individual records are damped. The weighted average is used as reference series to homogenise the available records.

Based on an assessment of the noise levels of each difference record, the location of the record and whether or not the station is still operational, a selection of four records is made which span the period from 1906 onwards. Two additional records are included from 1951 and 1958 onwards. The CNT is constructed as an unweighted average of these records with adjustments added to account for the additional records.

The CNT series, to which we attach the version number 1.1 and the ten homogenised station series are updated monthly and can be downloaded from the KNMI web site, www.knmi.nl.

Acknowledgement. The authors would like to thank Theo Brandsma and the Royal Netherlands Meteorological Institute for making available the homogenised De Bilt record from Brandsma (http://www.knmi.nl/klimatologie/onderzoeksgegevens/homogeen\_260/tg\_hom\_ mnd260.txt). We thank Jan Huizinga for his assistance in uncovering the metadata from the KNMI archives.



#### References

10

- Alexandersson, H. and Moberg, A.: Homogenization of Swedish temperature data, Part I: homogeneity test for linear trends, Int. J. Climatol., 17, 25–34, doi:10.1002/joc.1097, 1997. 2537
- <sup>5</sup> Brandsma, T. and van der Meulen, J. P.: Thermometer screen intercomparison in De Bilt (the Netherlands) – Part II: description and modeling of mean temperature differences and extremes, Int. J. Climatol., 28, 389–400, doi:10.1002/joc.1524, 2008. 2521
  - Easterling, D. R. and Peterson, T. C.: A new method for detecting undocumented discontinuities in climatological time series, Int. J. Climatol., 15, 369–377, doi:10.1002/joc.3370150403, 1995. 2522, 2537
- Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded daat set of surface temperature and precipitation for 1950–2006, J. Geophys. Res.-Atmos., 113, D20119, doi:10.1029/2008JD010201, 2008. 2537, 2552
- <sup>15</sup> Jarusková, D.: Change-point detection in meteorological measurement, Mon. Weather Rev., 124, 1535–1543, 1996. 2524
  - Labrijn, A.: The climate of the Netherlands during the last two and a half centuries, Tech. Rep. KNMI No. 102, Royal Netherlands Meteorological Institute, mededelingen en Verhandelingen, 1945. 2518
- Lund, R. and Reeves, J.: Detection of undocumented changepoints: a revision of the two-phase regression model, J. Climate, 15, 2547–2554, 2002. 2523, 2524
  - Menne, M. J. and Williams Jr., C. N.: Homogenization of temperature series via pairwise comparison, J. Climate, 22, 1700–1717, doi:10.1175/2008JCLI2263.1, 2009. 2522
  - Peterson, T. C. and Easterling, D. R.: Creation of homogeneous composite climatological ref-
- erence series, Int. J. Climatol., 14, 671–680, 1994. 2526
- Peterson, T. C., Easterling, D. R., Karl, T. R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., Vincent, L., Heino, R., Toumenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Forland, E., Hanssen-Bauer, I., Alexandersson, H., Jones, P. D., and Parker, D.: Homogeneity adjustments of in situ atmospheric climate data: a review, Int. J. Climatol., 18, 1493–1517, 1998. 2526
  - Potter, K. W.: Illustration of a new test for detecting a shift in mean in precipitation series, Mon. Weather Rev., 109, 2040–2045, 1981. 2526



Reeves, J., Chen, J., Wang, X. L., Lund, R., and Lu, Q.: A review and comparison of changepoint detection techniques for climate data, J. Appl. Meteorol. Clim., 46, 900–915, 2007. 2524, 2527

van der Hoeven, P. C. T.: Etmaaltemperatuur en Dagextremen uit Termijnwaarnemingen (in Dutch), memorandum KD 92-08, KNMI, 1992. 2520

5

10

van Ulden, A., van Oldenborgh, G. J., and van der Schrier, G.: The Construction of a Central Netherlands temperature, Tech. Rep. WR 2009-03, KNMI, 2009. 2519, 2526, 2527, 2537, 2538, 2553

Vincent, L. A.: A technique for the identification of inhomogeneities in Canadian temperature series, J. Climate, 11, 1094–1104, 1998. 2521, 2522

- Wang, X. L.: Comments on "Detection of undocumented changepoints: a revision of the twophase regression model", J. Climate, 16, 3383–3385, 2003. 2523, 2524
- Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Academic Press, New York, 1995. 2522, 2524



**Table 1.** Climatological records analysed in this study. (H-records based on 24 hourly observations; G-records based on 5 observations per day until 1970, 10 observations per day from 1971 onwards.)

Station	Record ID	Period	Missing data	Filled with
De Bilt	D001 H	1901–1970		
Den Helder	D002 H	1906–1970	Sep 1944– May 1945	Hoorn
Vlissingen	D003 H	1906–1970	1918–1930, 1944–1945	Excluded
				from analysis
Eelde	D004 H	1946–1970		
Beek	D005 H	1946–1970		
Groningen	D006 H	1906–1951		
Maastricht	D007 H	1906–1952		
De Kooy	D009 H	1961–1970		
Winterswijk	D020 G	1906–1990	Nov 1944, Oct 1988	De Bilt
Hoorn	D029 G	1906–1990	Nov 1947–Apr 1948	Den Helder
Oudenbosch	D032 G	1906–1992		
Gemert	D033 G	1906–1990		
Sittard	D145 G	1906–1948	Apr–Aug 1940,	Maastricht
			Nov 1944–Feb 1945	
Gilze-Rijen	D132 G	1953–1970		
Twenthe	D146 G	1947–1970		



Station	Record ID	Period	Missing data	Filled with
Den Helder	235	1971–Jul 1972		
De Kooy	235	Aug 1972–2008		
Schiphol	240	1951-2008		
De Bilt	260	1951–2008		
Soesterberg	265	1953–2007		
Leeuwarden	270	1951–2008		
Deelen	275	1958–2008		
Eelde	280	1951–2008		
Hupsel	283	1990–2008		
Twenthe	290	1971–2008		
Rotterdam	344	1957–2008		
Gilze-Rijen	350	1971–2008		
Eindhoven	370	1951–2008	May, Jun 1952	Gemert and
				Oudenbosch
Volkel	375	1953–2008		
Beek	380	1971–2008		

 Table 2. Synoptic records based on 24 hourly observations.



Station	Record IDs	Period	Transition	Overlap
De Bilt	D001+260	1901–2008	Jan 1971	
Den Helder/De Kooy	D002+235	1906–2008	Aug 1972	1961–1970
Groningen/Eelde	D004/006+280	1906–2008	Jan 1946	1946–1951
Maastricht/Beek	D005/007+380	1906–2008	Jan 1946	1946–1952
Winterswijk/Hupsel	D020+283	1906–2008	Jan 1991	Mar–Dec 1990
Hoorn	D029+270	1906–1990	No suitable	
			follow-up	
Oudenbosch/Gilze-Rijen	D032+350	1906–2008	Jan 1993	1988–1992
Gemert/Volkel	D033+375	1906–2008	Jan 1991	1990
Sittard/Beek	D145+380	1906–2008	Jan 1946	1946–1948
Twenthe	D146+290	1946–2008	Jan 1971	
Schiphol	240	1951–2008		
Soesterberg	265	1953–2008		
Deelen	275	1958–2008		
Rotterdam	344	1957–2008		
Eindhoven	370	1951–2008		

Table 3. Long composite records used for break and trend analysis.



Table 4. Loading of the seven long records on the first principal component.

Station	Loading
De Bilt	0.143
Den Helder/De Kooy	0.128
Groningen/Eelde	0.144
Maastricht/Beek	0.150
Winterswijk/Hupsel	0.149
Oudenbosch/Gilze-Rijen	0.141
Gemert/Volkel	0.145



<b>Discussion</b> Pape	6, 2517–2	<b>PD</b> 2555, 2010			
	Central Netherlands temperature				
Discu	G. van der S	G. van der Schrier et al.			
ssion F	Title Page				
aper	Abstract	Introduction			
	Conclusions	References			
Discu	Tables	Figures			
ISSIO	I	۶I			
n Pap	•				
oer	Back	Close			
	Full Scre	een / Esc			
Discuss	Printer-friendly Version				
ion F	Interactive Discussion				
Paper	C	<b>O</b> BY			

**Table 5.** Trends (°C/10 yr) in CNT1.0 and CNT1.1 for winter (DJF), spring (MAM), summer (JJA) and autumn (SON), calculated over the periods 1906–2008, 1950–2008 and 1975–2008.

	1906–2008		1950-2008		1975-2008	
	CNT1.0	CNT1.1	CNT1.0	CNT1.1	CNT1.0	CNT1.1
DJF	0.111	0.114	0.334	0.332	0.544	0.540
MAM	0.126	0.130	0.308	0.314	0.655	0.662
JJA	0.154	0.148	0.241	0.240	0.381	0.391
SON	0.127	0.123	0.149	0.149	0.284	0.277



Fig. 1. Map of the Netherlands with the station locations and station types.





**Fig. 2.** Adjustment factors used to blend series from operational stations to those from stations which ceased operation (red = Den Helder/De Kooy, green = Gemert/Volkel, blue = Groningen/Eelde, pink = Maastricht/Beek, magenta = Oudenbosch/Gilze Rijen, yellow = Winterswijk/Hupsel, black = Sittard/Beek) [°C].







**Fig. 3.** Running standard deviations over 41 yr windows of annual averages of the difference between target series and reference series. Upper panel shows the long records, lower panel shows the shorter records.









**Fig. 5.** Trends and breaks in the annual mean temperature difference between the Gemert/ Volkel series and the reference series [°C].













**Fig. 7.** Break and trend corrections for various stations for this study and those of van Ulden et al. (2009), the latter marked with (WR). Shown are corrections for Winterswijk/Hupsel **(a)**, Oudenbosch/Gilze Rijen **(b)**, Gemert/Volkel **(c)** and Eindhoven **(d)**, Maastricht/Beek **(e)** and Schiphol **(f)**.



Fig. 8. Root Mean Square error of the difference between CNT1.1 and CNT1.0.





**Fig. 9.** Root Mean Square error of the difference between the homogenised stations records which construct CNT1.1 and CNT1.0.

