



A "Global Radiosonde and tracked-balloon Archive on Sixteen Pressure levels" (GRASP) going back to 1905 – Part 2: homogeneity adjustments for pilot balloon and radiosonde wind data

L. Ramella Pralungo and L. Haimberger

Department of Meteorology and Geophysics, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria

Correspondence to: L. Ramella Pralungo (lorenzo.ramella-pralungo@univie.ac.at)

Received: 11 April 2014 – Published in Earth Syst. Sci. Data Discuss.: 15 May 2014

Revised: 4 August 2014 – Accepted: 6 August 2014 – Published: 9 September 2014

Abstract. This paper describes the comprehensive homogenization of the “Global Radiosonde and tracked balloon Archive on Sixteen Pressure levels” (GRASP) wind records. Many of those records suffer from artificial shifts that need to be detected and adjusted before they are suitable for climate studies.

Time series of departures between observations and the National Atmospheric and Oceanic Administration 20th-century (NOAA-20CR) surface pressure only reanalysis have been calculated offline by first interpolating the observations to pressure levels and standard synoptic times, if needed, and then interpolating the gridded NOAA-20CR standard pressure level data horizontally to the observation locations. These difference time series are quite sensitive to breaks in the observation time series and can be used for both automatic detection and adjustment of the breaks.

Both wind speed and direction show a comparable number of breaks, roughly one break in three stations. More than a hundred artificial shifts in wind direction could be detected at several US stations in the period 1938/1955. From the 1960s onward the wind direction breaks are less frequent. Wind speed data are not affected as much by measurement biases, but one has to be aware of a large fair-weather sampling bias in early years, when high wind speeds were much less likely to be observed than after 1960, when radar tracking was already common practice. This bias has to be taken into account when calculating trends or monthly means from wind speed data.

Trends of both wind speed and direction look spatially more homogeneous after adjustment. With the exception of a widespread wind direction bias found in the early US network, no signs of pervasive measurement biases could be found. The adjustments can likely improve observation usage when applied during data assimilation. Alternatively they can serve as a basis for validating variational wind bias adjustment schemes. Certainly, they are expected to improve estimates of global wind trends.

All the homogeneity adjustments are available in the PANGAEA archive with associated doi:10.1594/PANGAEA.823617.

1 Introduction

From the 1900s tracked balloons, and from the 1940s also radiosondes, were practically the only upper-air wind observing system with global or regional coverage up to the beginning of the satellite era in the late 1970s. Despite the rapidly increasing amounts of satellite data since then (Dee et al.,

2011), balloons and radiosondes remain an essential component of the observing network. While the vertical extent of the records was limited to mostly below 400 hPa before 1940, it reached the stratosphere from the 1950s onward (Scherhag, 1962). Since the 1970s, 10 hPa are regularly reached by most balloons.

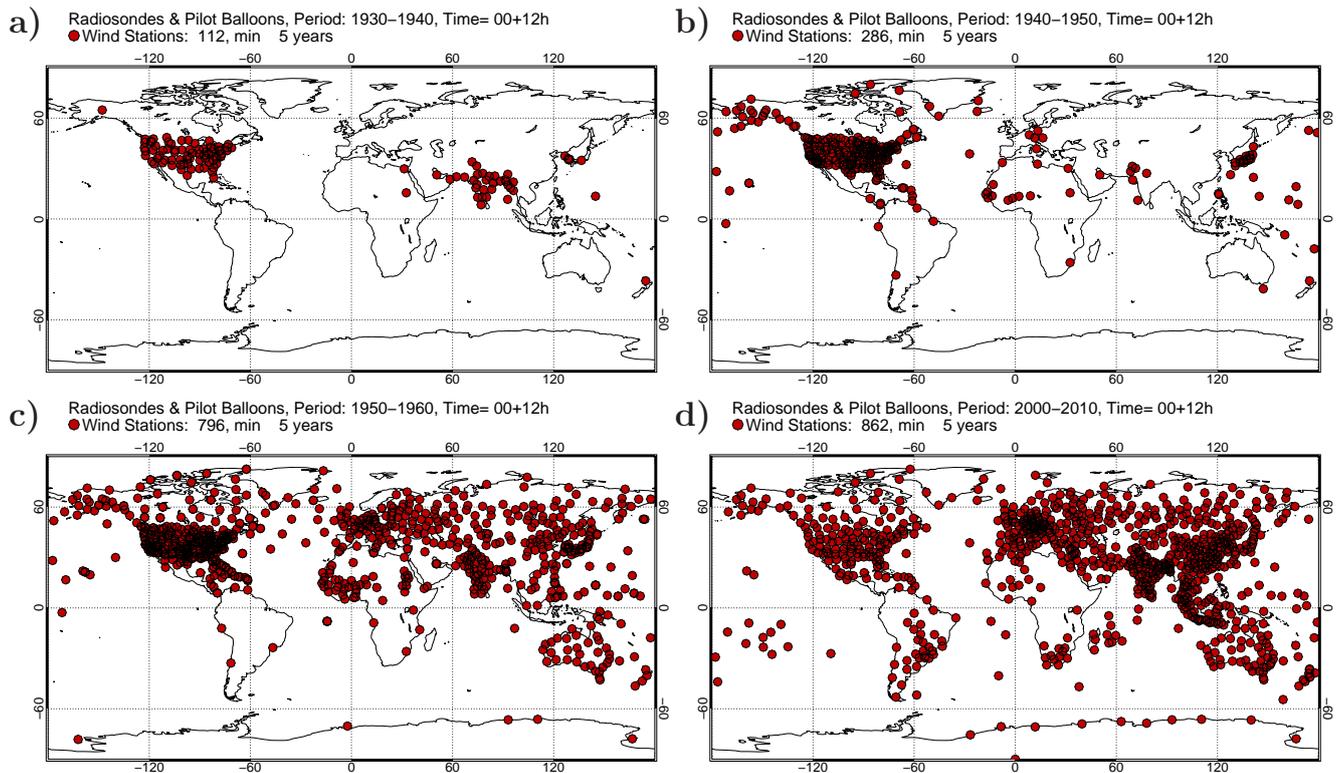


Figure 1. Maps of upper-air stations of the GRASP data set with at least 5 years of data per decade in the range of 850–500 hPa for different decades.

Early upper-air wind data have been used to reconstruct climate anomalies in the early 20th century, such as the Dust Bowl drought in the 1930s (Ewen et al., 2008b), and were instrumental in discovering the quasi-biennial oscillation (Graystone, 1959). More modern in situ upper-air wind data have helped attributing regional wind stilling in the Northern Hemisphere to increased surface roughness (Vautard et al., 2010). There is also increased interest in wind speed trends due to increasing installations of wind turbines. Allen and Sherwood (2008) have used wind data as a temperature proxy, applying the thermal wind equation to calculate temperature (gradient) trends. Interestingly, they found much larger warming trends than those estimated from radiosonde temperature measurements which are long known to be biased (Santer et al., 1999, 2005; Thorne et al., 2005).

Long and homogeneously observed time series are an essential source to diagnose the three-dimensional pattern of climate change. They are also precious input data for reanalysis efforts, particularly if they go back to before the satellite era (Kistler et al., 2001; Uppala et al., 2005; Ebita et al., 2011). Reanalyses have proved extremely fruitful for climate research and are essential input for studies in many disciplines (Hartmann et al., 2013). Reanalysis performance and homogeneity are highly dependent on the available input data and their quality. Any improvement there will increase the accuracy of the estimated climate state and will be directly

beneficial for future studies related to it. This fact triggered several efforts to digitize surface data and early upper-air data in many countries (Allan and Ansell, 2006; Allan et al., 2011; Brönnimann, 2003; Ewen et al., 2008a; Stickler et al., 2014). The present study builds upon these efforts. In part I of this study (Ramella-Pralungo et al., 2014) a new radiosonde and pilot balloon wind data set at 16 standard pressure levels, called GRASP, has been developed. Figure 1 shows maps of global coverage with wind data for different decades in this data set.

Upper-air wind time series from balloons have traditionally been assumed to be temporally relatively homogeneous compared to temperature or humidity. While biases are not as pervasive as for temperature, wind biases have been occasionally detected by monitoring the output of data assimilation systems (Hollingsworth et al., 1986). As an example, Fig. 2 shows the observed wind direction time series for the station Bismarck (North Dakota USA WMO ID 072764) at 00:00 UTC at different pressure levels. Already in these time series there are indications that, in around 1948, they are affected by artificial shifts in wind direction, caused by wrong north alignment of the station. Imperfect tracking of the horizontal motion, wrong height assignment or the inertness of the ascending balloon can also cause biases. These biases are hard to detect without a reliable reference if they are smaller

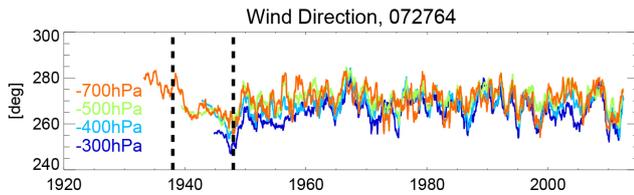


Figure 2. Time series of wind direction observations at station Bismarck (072764, North Dakota, USA), at the 700 (orange), 500 (green), 300 (light blue) and 200 (dark blue) hPa levels, respectively. A 365-day running mean has been applied to all time series. Note shifts in 1938 and 1948.

than in this example. That reference may be neighboring stations or, since recent times, reanalysis data.

Haimberger (2007) showed that background departure time series from reanalyses can be used effectively to detect and adjust breaks in radiosonde temperature time series. An automatic homogenization method (RADiosonde OBServation CORrection using REanalyses, RAOBCORE) showed some skill in adjusting the global radiosonde temperature data set. Subsequent refinements of the method also employed composites of neighboring radiosondes that allowed more accurate adjustment of breaks without sacrificing the independence of the adjusted radiosonde data from satellite data (Haimberger et al., 2008, 2012).

This method has been successfully applied to the global radiosonde wind data set back to 1958 as well (Gruber and Haimberger, 2008). While wind direction errors could be safely identified in their paper, there were concerns that wind speed background departures may not be independent enough from the assimilating model (ERA-40 (European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis) at this time (Uppala et al., 2005), which has some known inhomogeneities). This study also did not consider pilot balloon winds and was restricted to the period 1958–2002.

Since then, the availability of newly digitized data as well as the advent of the National Atmospheric and Oceanic Administration 20th-Century Reanalysis (NOAA-20CR; Compo et al., 2011) improved the prospects for a comprehensive homogenization of upper-air wind data back to the beginning of balloon observations. The NOAA-20CR is well suited as reference since it is independent of upper-air observations and since it has reasonably realistic temperature fields up to stratospheric levels. While some wind biases are evident in this reanalysis (see Fig. 3 of Part 1 and Stickler and Brönnimann, 2011), these are temporally quite stable, at least over the midlatitudes from 1950 onward, as will be shown below. It has been used for interpolating the global pilot balloon wind data set from geometric height to standard pressure levels in Part 1. The present paper describes how this combined radiosonde plus pilot balloon data set is homogenized using analysis departure information from the NOAA-20CR.

The next section describes the input data, and Sect. 3 outlines the homogenization method. Section 4 explains how sampling biases can translate into monthly mean biases and what is done to avoid that. Section 5 presents results, and conclusions are drawn in Sect. 6.

2 Input data

The GRASP data set as described in Part 1 is the main input to be homogenized. Its main sources are the ERA-Interim observation input and observation feedback data set (Dee et al., 2011), the ERA-40 observation input and feedback data set (Uppala et al., 2005), the Integrated Global Radiosonde Archive (IGRA) (Durre et al., 2006), updated until 2012 and supplemented with analysis departures from the NOAA-20CR, and the Comprehensive Historical Upper-Air Network (CHUAN; Stickler et al., 2010) including the ERA-CLIM (European Reanalysis of Global Climate Observations) Historical Upper-Air Data (Stickler et al., 2014).

The archive contains daily records at 00:00 and 12:00 UTC from 2924 wind stations with a WMO identifier, distributed all over the world, with data starting in 1905 at Lindenberg (10 393) for temperature and in Hamburg (10 148) for wind and with longer time series over the US starting in 1915 (e.g., Omaha 72 553). Figure 1 shows the wind network development and distribution of those stations that have at least 5 years of observations. Already in the 1930s, a network of stations was operating mainly in the US and in India. In the 1940s the distribution was already global, albeit sparse. The global coverage over land masses was constantly improved in the 1950s, particularly in the International Geophysical Year (IGY) 1958 when several new radiosonde stations also started operating in South America, southern Africa and the Pacific Islands. The network of radiosonde and pilot balloon observations in the 1950s is much denser than the radiosonde network only (compare with Fig. 10 in Part 1). Since the IGY, the radiosonde network has not changed much in size, peaking in the 1970s and since then slowly decreasing in density.

In addition to the observations at standard pressure levels, the GRASP data set contains the so-called *innovations* or *analysis departures*, i.e., observation minus analysis from the NOAA-20CR, collocated at each station location and each available observation time (00:00 and 12:00 UTC), and standard pressure levels. For homogenization the analyzed fields from the NOAA-20CR have been chosen again since they are independent of upper-air data and since NOAA-20CR goes back to 1872 so that analysis departures exist even for the earliest upper-air data. While background departures from ERA-40 and ERA-Interim have smaller variance, they are not completely independent of upper-air data, and they also do not reach back far enough. Evaluations of NOAA-20CR have found some biases in its wind field (Stickler et al., 2010) and several inhomogeneities have been detected by Ferguson and Villarini (2013), particularly in ensemble spread. Despite

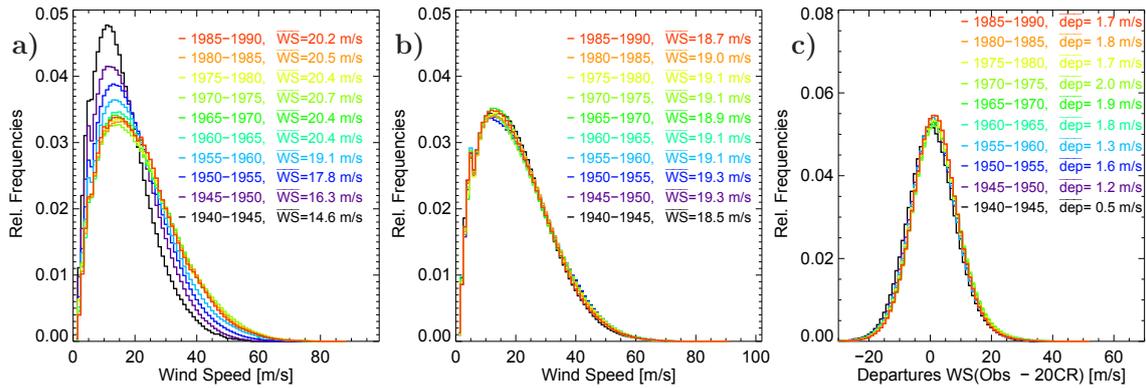


Figure 3. Histogram for several decades of (a) wind speed observations, (b) analyzed wind speeds and (c) innovations (obs-NOAA-20CR) at 300 hPa and 00:00 UTC for the same subset of stations as in Fig. 12. The histograms have been normalized so that the integral is 1.

these caveats one has to acknowledge that the NOAA-20CR reproduces the comprehensive global atmospheric circulation and also the surface temperatures in the 20th century quite well, given the reduced set of input data it assimilates. In many ways it is unique, and similar upcoming products such as ERA-20C (Poli et al., 2013) have yet to demonstrate significantly better temporal homogeneity than the NOAA-20CR.

3 Methods

The radiosonde wind homogenization system presented here is based on the methodology referred to as Radiosonde Observation Correction using Reanalysis (RAOBCORE). It was introduced by Haimberger (2007) and was originally applied only to temperature time series. Gruber and Haimberger (2008) demonstrated that the RAOBCORE's technique can be, when properly modified, applied also to wind data with promising results. In that pioneering study only radiosonde data collected during ERA-40 (Uppala et al., 2005), complemented with radiosonde data from the IGRA (Durre et al., 2006), were used, spanning the time interval 1958–2002.

The system used here, referred to as RAOBCORE 2.0, analyzes temperature, wind speed and wind direction simultaneously. In contrast to the original RAOBCORE, it uses NOAA 20th-Century Reanalysis data as reference. While the simultaneous treatment of temperature and wind would allow for combined break detection, we found that temperature and wind breaks rarely occur at the same time, the exception being perhaps station relocations. Therefore each variable is treated separately for break detection and adjustment. Metadata are an important source of information for homogenization purposes, and they can and have been used in the break detection procedure by assigning higher a priori probabilities to a break given particular events (Haimberger, 2007) or for applying certain adjustments (Luers and Eskridge, 1995). A new, promising version of upper-air metadata has recently become available (Tschudin and Schroeder, 2013). However,

the metadata have been collected mainly for temperature and humidity adjustment. For wind, the metadata information is rather sparse and most likely incomplete. At this stage, metadata have therefore not been used for wind homogenization.

3.1 Construction of reference time series

The NOAA-20CR zonal (U) and meridional (V) wind values used as reference are already stored in GRASP as departure time series, i.e., the NOAA-20CR values can be calculated simply by adding the departures to the observations. For homogenization purposes is often advantageous to represent wind information in polar coordinates (speed ff and direction dd). North alignment shifts in wind direction, which are a major source of biases affecting the whole vertical profile, are visible much better in wind direction time series than in U or V time series. Following the notation of Haimberger et al. (2012), we define the wind direction difference (τ) between an observation (obs) at time t_i (where i is index of day) at a given station and the NOAA-20CR analysis (an) interpolated to this as

$$\tau_{dd}(t_i) = \text{obs}_{dd}(t_i) - \text{an}_{dd}(t_i). \quad (1)$$

In cases where this difference is not in the range of $[-180^\circ, 180^\circ]$, we set

$$\begin{aligned} -\tau_{dd}(t_i) &= \tau_{dd}(t_i) - 360^\circ & \text{if } \tau_{dd}(t_i) > 180^\circ \\ -\tau_{dd}(t_i) &= 360^\circ + \tau_{dd}(t_i) & \text{if } \tau_{dd}(t_i) < -180^\circ. \end{aligned}$$

Wind direction departures are not necessarily Gaussian distributed, and care must be taken particularly at low wind speeds. Therefore, they are considered valid only if the wind speeds in both observations and in the reanalysis are higher than 1 m s^{-1} and if the wind direction differences are smaller than 90° . Now one can take the average over a time interval a (denoted by an overbar) which yields

$$\bar{\tau}_{dd}(a) = \overline{\text{obs}_{dd}(a)} - \overline{\text{an}_{dd}(a)}. \quad (2)$$

The wind speed departures Δff are defined as

$$\bar{\tau}_{ff}(a) = \overline{\text{obs}_{ff}(a)} - \overline{\text{an}_{ff}(a)}. \quad (3)$$

$\bar{\tau}_u(a)$ and $\bar{\tau}_v(a)$ can be defined in a similar way. All these averages exist at 00:00 and 12:00 UTC at the 16 standard pressure levels. The standard length of the intervals a, b is 2 years for break detection (to be able to separate two nearby breakpoints) and 8 years (for better accuracy) for break adjustment. It can be reduced to 200 days at the beginning and end of the time series and near gaps. At least 200 valid observations in each interval are required for valid averages.

3.2 Break detection

As discussed by Haimberger (2007), a variant of the standard normal homogeneity test (SNHT; Alexandersson, 1986) can be applied to the innovation time series $\tau_x(t_i)$ (where x can be u, v, dd, ff) in order to find possible break candidates.

The SNHT variant considered here calculates a test statistic Q_k for each potential breakpoint k . The intervals are then chosen as $a = [k - N/2, k]$ and $b = [k, k + N/2]$, with $N = 1460$ (2 years before, 2 years after a potential break) as default choice. We now define Q_k as

$$Q_k = N/2 \left\{ [\bar{\tau}_x(a) - \bar{\tau}_x(a, b)]^2 + [\bar{\tau}_x(b) - \bar{\tau}_x(a, b)]^2 \right\} / \sigma_x(a, b), \quad (4)$$

where $\sigma(a, b)$ is the standard deviation of the $\tau_x(t_i)$ over the interval $[a, b]$ and $\bar{\tau}_x(a, b)$ is the mean $\tau_x(t_i)$ over the whole interval $[a, b]$. The maximum number of missing data admitted in one subinterval of length $N/2$ is an important parameter. It is set to 650. With so many data allowed to be missing, special care is needed to ensure the equal sampling of the annual cycle in the interval before and after a potential breakpoint k , particularly in the case of data gaps. This is done by deleting data for a month that is missing in one interval also in the other interval. This simple measure minimizes false break detections at stations with strong annual cycles (Haimberger, 2007). This SNHT variant yields time series of Q_k for all k where there are enough data. The Q_k time series exist at each pressure level, and significant maxima are not reached at each level and not at the same k values. Thus one has to combine the breakpoint probability information from all levels to get unique breakpoints. The composite series of all the Q_k is obtained as a mean along all the pressure levels and times:

$$\langle Q_k(t_m) \rangle = \frac{1}{n_p} \frac{1}{n_t} \sum_{n_p, n_t} Q_k(p_l, t_m) \text{ with } p_l = 100, 150, \dots, 850 \text{ hPa.}$$

$$t_m = 00, 12 \text{ UTC}$$

Under the null hypothesis (homogeneous time series), the distribution of $\langle Q_k \rangle$ can be obtained from Monte Carlo simulations, taking possible autocorrelation into account. From these, a critical value Q_{crit} can be derived. It is a reasonable idea to set Q_{crit} not too high, since weaker breaks can be

eliminated later by applying robustness criteria (explained in the next section). Depending on the value set for Q_{crit} , the time series $\langle Q_k \rangle$ is converted into break “probabilities”, with range $[0, 1]$, in the same manner as Haimberger (2007), except that the a priori probability time series is flat (no use of metadata). If, on a given day, the maximum break probability exceeds 0.5, the date is recorded as a *possible* break. If it passes the robustness tests, it is adjusted as explained below. If there is a break, the $\langle Q_k \rangle$ are typically larger than Q_{crit} over a certain time interval (see, e.g., Fig. 6). The break location is set at the maximum $\langle Q_k \rangle$ value. Local $\langle Q_k \rangle$ maxima have to be separated by at least a year to be recognized as separate breakpoints.

3.3 Break adjustment

As mentioned, a common reason for breaks in wind direction is the wrong north alignment: in this case the wind direction error is expected to be constant in time and also in the vertical. The break size estimate at a given pressure level at 00:00 or 12:00 UTC at date k is defined as

$$\overline{\Delta \tau}_{dd}^k(a, b) = \bar{\tau}_{dd}^k(b) - \bar{\tau}_{dd}^k(a). \quad (5)$$

The intervals a, b are generally chosen to be longer (up to 8 years, as explained above) than the intervals used for break detection with SNHT, above. We consider a break significant if its magnitude is larger than 1.96 times the standard deviation of the $\bar{\tau}_{dd}$. We require that the vertical mean of the break sizes $\langle \overline{\Delta \tau}_{dd}^k(a, b) \rangle$ must be greater than 3° and that the break size estimates at individual pressure levels must not change sign in the vertical. At least four time series (two pressure levels at both 00:00 and 12:00 UTC times or four pressure levels at the same time) must be available in order for them to be accepted. In this case the break is adjusted at both 00:00 and 12:00 UTC. Since there are cases in which the device changes between the observation at 00:00 and 12:00 UTC (for example, pilot balloon at 00:00 UTC and radiosonde at 12:00 UTC), one may get different and independent break profiles at different times. For such cases, the significance test has been performed at both times separately. The breaks at station Bismarck as shown in Fig. 6 are typical examples. Others, like the Russian station Aktyubinsk (now known as Aktobe, but for historical reason here called Aktyubinsk, WMO ID 35229) or Marion Island (WMO ID 68994), highlighted by Gruber and Haimberger (2008), have been revisited for comparison purposes. The obtained mean break profiles agree well with those presented in Gruber and Haimberger (2008), although the variance of the estimates is larger because the scatter of the departures from surface data only reanalysis used as reference is larger when compared to background departures from full reanalyses such as ERA-Interim, particularly in remote regions.

The approach for wind speed is slightly different, since there are no such strong physical constraints available as

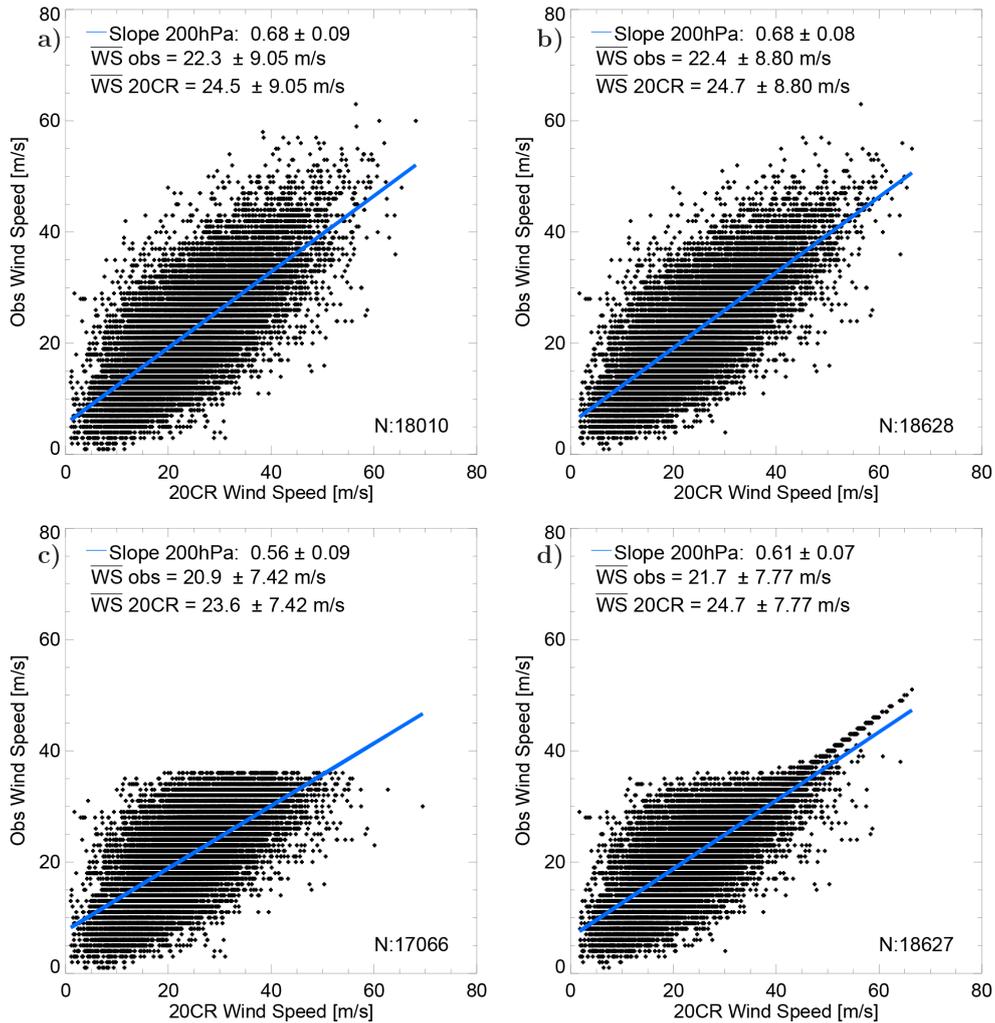


Figure 4. Scatterplots of observed wind speed vs. NOAA-20CR at station 72764 at 200 hPa with different treatment of high wind speeds during the period 1960–2010. **(a)** original observed wind speed; **(b)** observed wind speed, missing values replaced by scaled NOAA-20CR wind speeds; **(c)** observed wind speed, all values above 60% of the maximum observed wind speed set to “missing”; **(d)** as (c), but values set to missing in (c) are replaced with scaled NOAA-20CR wind speeds. Mean wind speeds as well as the linear relationship between observed and NOAA-20CR wind speeds are given in inset legends.

there are for wind direction. Larger biases may occur only at some pressure levels; thus, each single level is analyzed independently. In particular, it is well known that the largest inhomogeneities occur at high wind speeds, i.e., where the jet streams are located (7–12 km above the sea level for the polar jet and 10–16 km for the subtropical jet). A break at a given level is considered significant if its size is larger than 1.96 times the standard deviation of the $\bar{\tau}_{ff}$. If this criterion is fulfilled at least at two levels at a potential breakpoint, the wind speed profile is adjusted.

While the individual wind direction values are adjusted by a vertically constant value, the wind speeds at a given pressure level are adjusted by a factor λ such that high wind speed values are adjusted more than low wind speeds. Otherwise low wind speeds might become negative.

Before adjusting the observations, the NOAA-20CR wind speeds, which are known to be biased (Compo et al., 2011), have been adjusted by another constant factor λ_{20CR} , which has been derived from the mean wind speed quotient between the most recent part of the observation time series and the corresponding NOAA-20CR analysis time series:

$$\lambda_{20CR} = \overline{WS}_{obs} / \overline{WS}_{20CR}. \quad (6)$$

The λ_{20CR} factor is calculated for each pressure level and for each station using linear regression to reflect different biases in different climate zones. Only wind speeds larger than 1 m s^{-1} in the NOAA-20CR and after 1960 have been taken into account when calculating this factor. In general, the method delivers reasonable factors in the range of 0.9–1.2, assuming there are sufficient data available (at least

1000 values). One example is the ratio of the mean wind speeds given in Fig. 4a. Before 1960, the mean values of the observations in Fig. 3a are higher by this factor than the NOAA-20CR winds in Fig. 3b.

As has been described in Haimberger (2007) and Haimberger et al. (2012), the breakpoint adjustment procedure works backward in time, from the most recent to the earliest one; in this way, a progressively shorter section of the time series is adjusted. The default averaging interval for the break size estimation has been set to 8 years. If only a shorter time series is available (because of the next earlier break or because of a data gap), this parameter is reduced accordingly. The adjustment factors remain constant in amplitude between the breakpoints. After these procedures, the breaks due to changes in the measurement biases of the observed wind time series should be largely removed.

3.4 Calculation of unbiased monthly means

The early wind measurement systems used theodolites to track the ascending balloons, which worked only if there was good visibility and upper-level winds were not too strong. While, under fair weather conditions, the balloons could be tracked up to 200 hPa or higher, the balloons were sometimes lost below 2 km in stormy conditions. As a result, winds during disturbed weather conditions are underrepresented. While the single measurements are unbiased, their monthly mean is biased since only the low-wind part of the distribution is well sampled. Figure 3a shows histograms of observed wind speed for long records over the US at the 300 hPa level. Whereas there are very few measured wind speeds over 40 m s^{-1} in the 1940s, they are quite common from 1960 onward. When taking the mean, the 5-year averaged wind speeds increase from 14.6 m s^{-1} to more than 20 m s^{-1} in the above time frame. This increase is almost solely due to decreasing sampling bias. After the introduction of radar tracking in the 1960s, only very few observations were missing and the sampling bias vanished.

In contrast to the observations, the mean wind speed departures from the NOAA-20CR changed only slightly. Observed winds tend to be higher than NOAA-20CR winds by 0.5 m s^{-1} in the 1940s and 1.8 m s^{-1} from the 1960s onwards. The mean is positive since the NOAA-20CR has a predominantly low bias below 150 hPa in wind speed (Compo et al., 2011); see also Fig. 3 of Part 1 (U wind is a very good proxy for wind speed in the zonal mean). It increases slightly over time since the departures tend to be larger during high-wind conditions, which are undersampled in the early years. The increase is, therefore, not necessarily a sign of inhomogeneities in the measurements. The constant mean departures from the 1960s onward indicate that the low bias of NOAA-20CR is practically constant throughout these decades. Thus, it can be expected to be constant also earlier on, the exception being perhaps the period 1900–1940, when

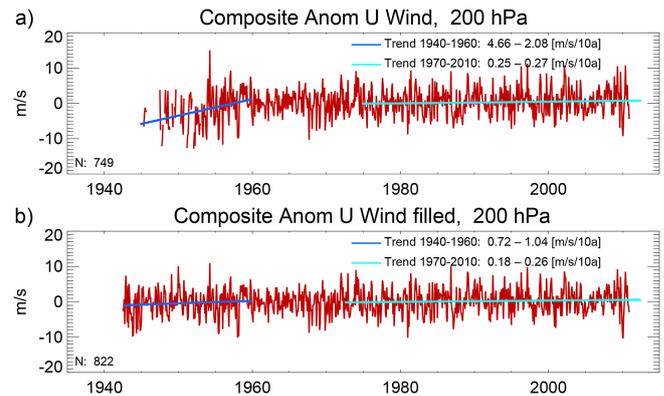


Figure 5. Monthly mean wind speed at 200 hPa, averaged over WMO stations with IDs between 70 000 and 75 000 (mostly USA and Canada), (a) without filling, (b) with missing data replaced with NOAA-20CR scaled wind values.

different background error variances have been used in the NOAA-20CR assimilation system (Compo et al., 2011).

There are several options to avoid a sampling bias in, e.g., monthly means. The simplest one is to calculate the means only if a very large fraction of observations are available (e.g., less than two observations missing per month). Such a strict requirement leads to very few monthly means being calculated in the early years. As a second option, one can substitute missing observations with analyzed winds from NOAA-20CR. We did this for months where at least 15 observations were available at a given observation time (00:00 or 12:00 UTC). The scaled NOAA-20CR winds (as explained before) are, of course, used for the substitution. Figure 4a, b show the effect of the filling at a station where only few data were missing. The filling procedure increases the observation count from 18 010 in (a) to 18 628 in (b). The mean wind speed is slightly increased through the filling in this case. To put more stress on the filling algorithm, we artificially set all high wind speed values to “missing” in panel (c), which reduces the mean wind speed by 1.4 m s^{-1} . Substitution of the values set to missing in (c) by scaled NOAA-20CR values reduces this artificial sampling bias to 0.6 m s^{-1} in panel (d).

A composite of monthly mean anomaly time series of stations with WMO numbers between 70 000 and 75 000 are shown in Fig. 5. The earliest part of the time series in (a) is affected by the sampling bias without filling. Complementing the daily data with scaled NOAA-20CR values before taking the monthly averages makes it possible to remove or at least reduce the spurious trends due to undersampling (lower panel). There may be room for further improvement in the treatment of the sampling bias, but the example demonstrates that the sampling bias cannot be ignored at all if one wants to consider time series going back before the 1960s.

Despite this encouraging performance of the filling method, we decided not to use it for homogenization

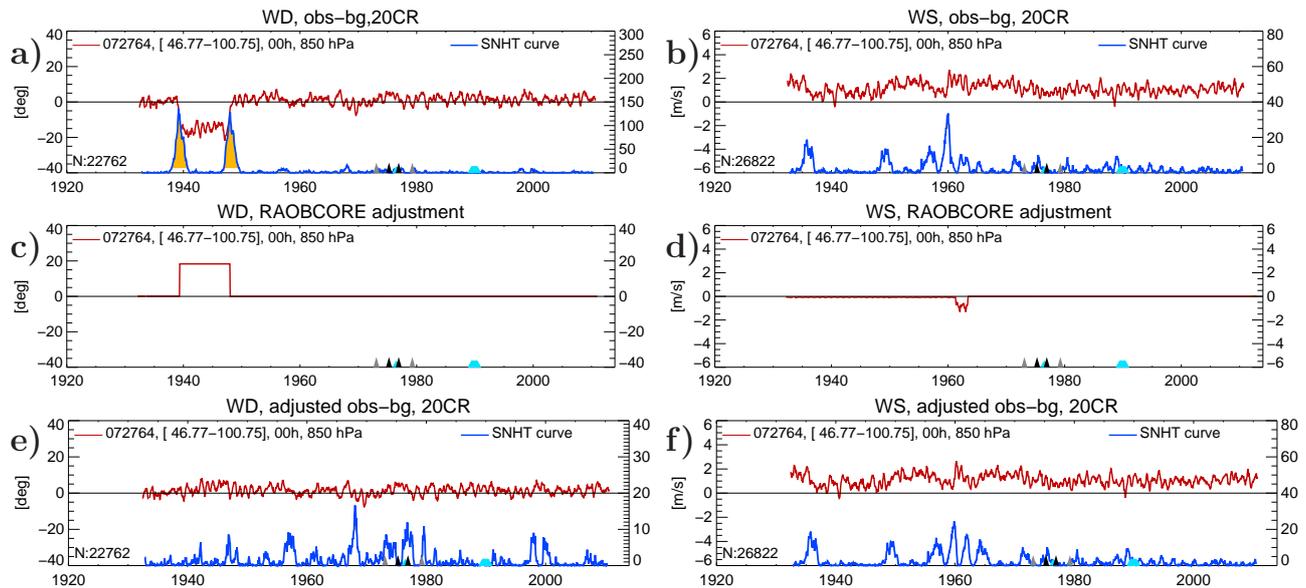


Figure 6. Upper row: (a) wind direction and (b) wind speed analysis departure (Obs-NOAA-20CR) time series at time 00:00 UTC at 850 hPa at station Bismarck (red curves). The blue curves (right axes) show the Q_k time series as defined in Eq. (4). Q_k values above $Q_{\text{crit}} = 10$ are statistically significant and are shaded. The small colored triangles on the x axes indicate changes in the radiosonde type. For dd , Q_{crit} has been set to 20. The mean of the wind speed differences is positive since NOAA-20CR wind speeds over the US are generally biased low. Metadata from IGRA are indicated as light blue trapezoids or triangles. Middle row: (c) wind direction and (d) wind speed adjustments calculated by RAOBCORE from analysis of the departure time series in upper row. Wind direction biases are constant in pressure and time between breakpoints. Wind speed adjustments occur only after 1960 at this station. They depend on observed wind speeds and are, therefore, not constant. Lower row: analysis departures after adding the adjustments in middle row to upper row. Note smaller right axis scales compared to the upper panel, indicating better homogeneity.

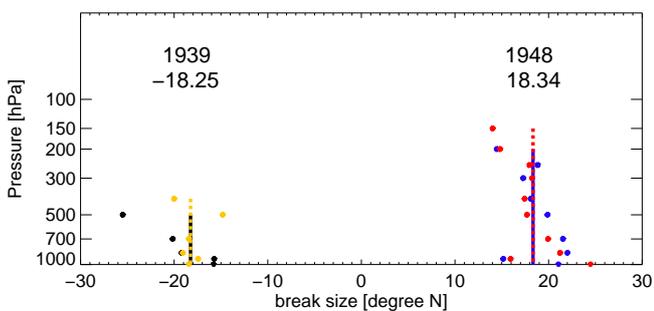


Figure 7. Vertical profiles of wind direction break size estimates at Bismarck (WMO ID 72764) for breaks in 1939 and 1948. The bullets are the estimated differences detected at each pressure level (black and blue at 00:00 UTC, yellow and red at 12:00 UTC). The solid and dotted lines are the respective vertically constant adjustment mean values for the whole vertical profiles

purposes. The most important reason is that the filling compromises the independence between observation and the reference. Thus, it would become harder to detect breaks when analyzing time series containing an appreciable amount of values from the reference series. Even if we did an adjustment, it would be underestimated for the observed values but overestimated for the filled values. This would re-

main a problem even if we had a more accurate reanalysis than the NOAA-20CR at our disposal. As such, one should use the filling approach only after homogenization.

In the following we display only results with data where the sampling bias is low so that the filling is essentially unnecessary. This is the case for the early period at lower pressure levels (< 500 hPa) and for the period after 1960.

Also, on PANGAEA, we only provide homogeneity adjustments (daily and monthly) but no filled values. It is important for the reader to understand that we removed the non-climatic shifts but that we only highlighted the presence of the sampling bias without actually removing it.

4 Results

4.1 Individual series analysis

As mentioned above, the main purpose of this paper is to find and possibly fix breaks, especially in the period before 1958, where other more sophisticated techniques (RICH (Radiosonde Innovation Composite Homogenization) and its versions (Haimberger et al., 2008), for example) are not applicable, mainly due to the low station density that makes the construction of reliable reference time series from neighboring stations rather difficult.

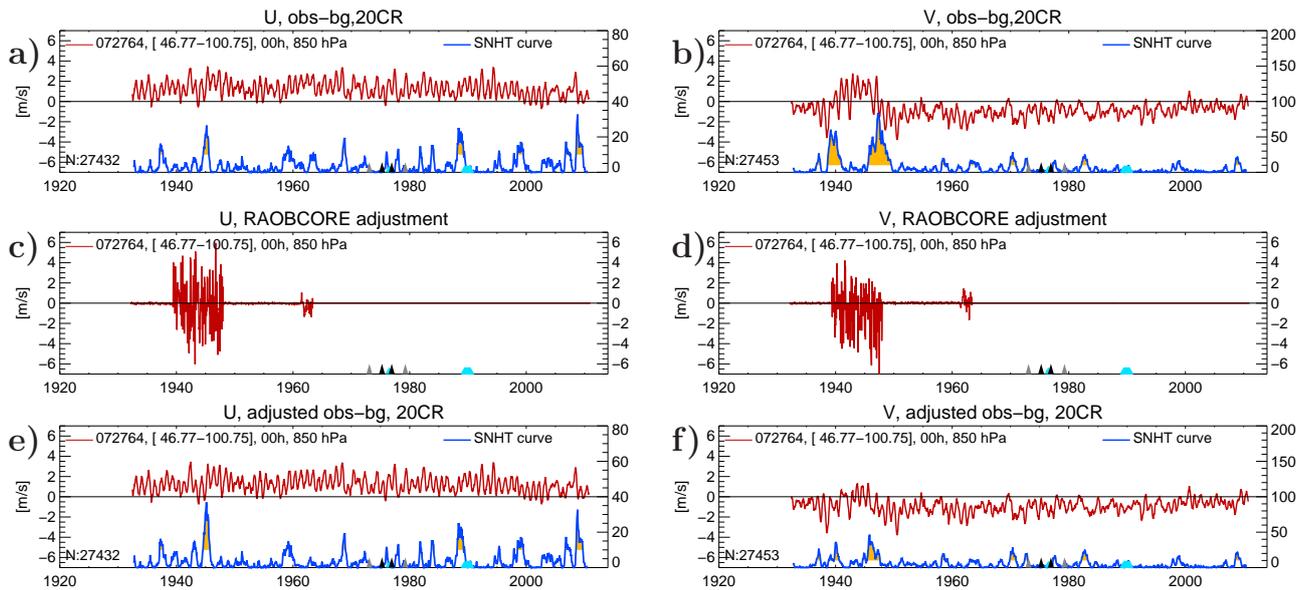


Figure 8. (a) U and (b) V wind innovations, (c, d) wind adjustments and (e, f) adjusted innovations at the 850 hPa level. Note that a positive adjustment of wind direction as shown in Fig. 6c translates to negative V adjustments, depending on the strength of the predominantly westerly winds.

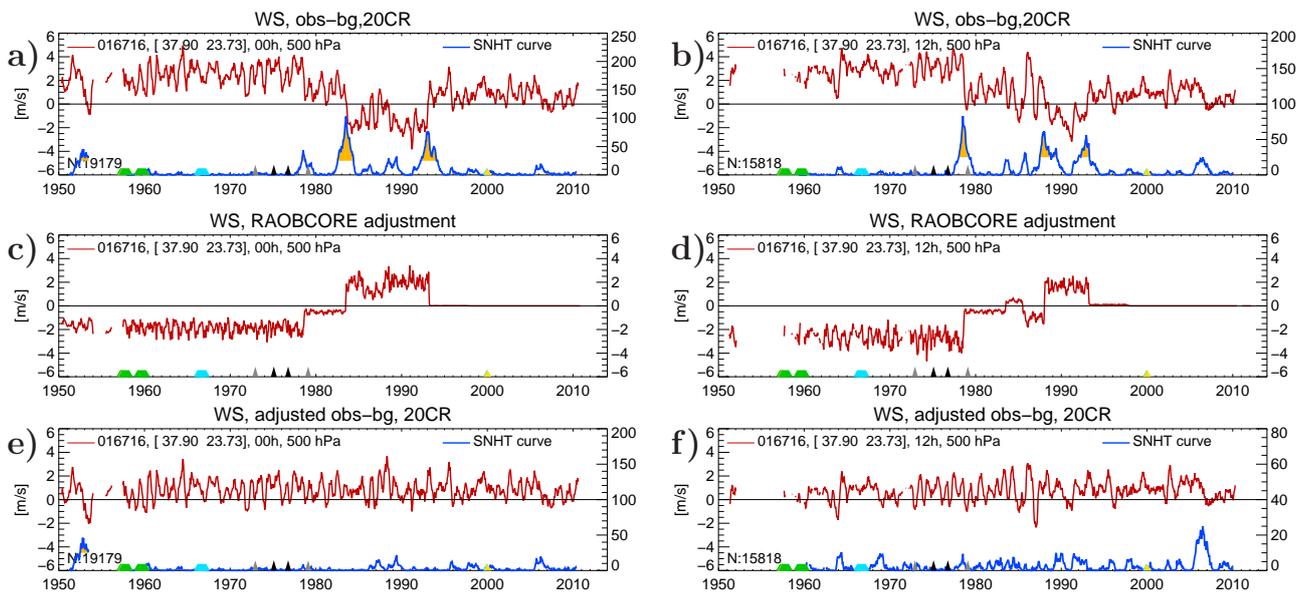


Figure 9. Wind speed innovations (upper row, red curves) at Athens, Greece, WMO ID 016716, at 300 hPa, at (a) 00:00 UTC and (b) 12:00 UTC. Blue curves are SNHT time series, (c, d) are corresponding adjustments and (e, f) are corresponding adjusted innovations.

While the US wind network was established already very early with up to 60 stations in the 1930s, it had some problems with measuring the wind direction. The raw observations show visible shifts, e.g., in the wind direction time series of station 072764 (Bismarck, North Dakota, USA) in Fig. 2. Shifts are visible at all the available pressure levels in the years 1938 and 1948. The shifts are much more clearly visible in Fig. 6a. It shows the wind direction innovations

$\tau_{dd}(t_i)$ time series with the NOAA-20CR winds as reference at 850 hPa for (a) wind direction and (b) wind speed. The SNHT test statistic reaches very high values for wind direction. The breaks are very similar for wind direction at other levels (Fig. 7). In these early years the ascents did not reach very high levels and already at 500 hPa the SNHT (blue line, right axis) can be calculated only after 1945. After those

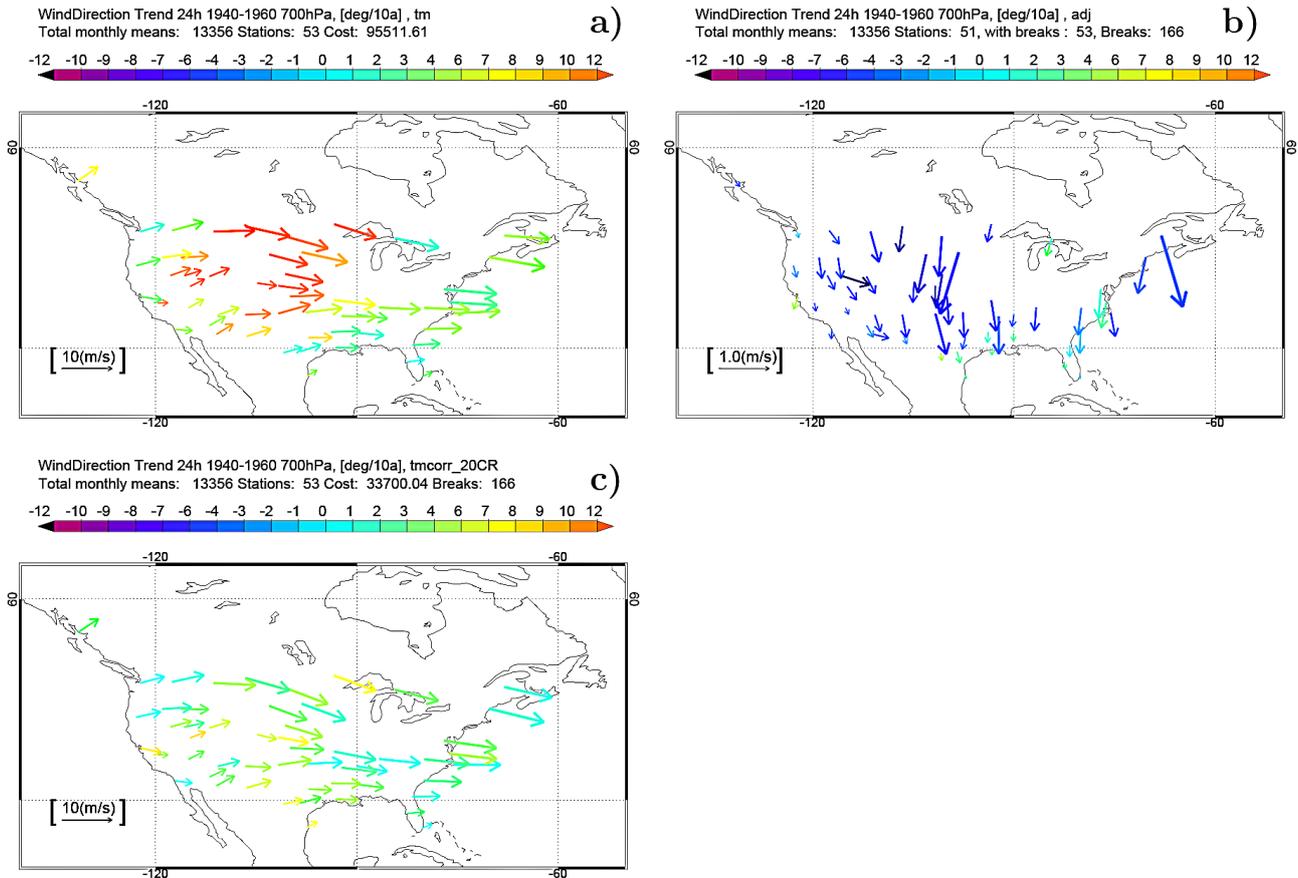


Figure 10. (a) Raw wind direction trends at 700 hPa at US stations (WMO numbers between 71 000 and 73 000) for the period 1940–1960 averaged over both 00:00 and 12:00 UTC (most stations reported twice daily). Arrows depict the mean wind speed and direction over this period. Arrow color indicates wind direction trend over the investigated period. Only stations with less than 1 year of missing values are shown. (b) Wind direction adjustments ($\Delta\bar{U}$, $\Delta\bar{V}$) as estimated by RAOBCORE 2.0. The arrows show size and direction of the adjustments; the colors show how much the wind direction trend has been changed due to the adjustments. Note scale difference in arrow length compared to (a). 174 shifts have been adjusted in total. (c) shows wind direction trends after adjustment. The cost function measuring spatial-trend heterogeneity (see Haimberger, 2007; Haimberger et al., 2012) has been reduced by a factor of 3 compared to (a).

breaks in early years, the wind direction departures seem quite homogeneous with no obvious shift.

Figure 6b shows the wind speed innovations $\tau_{ff}(t_i)$ and the SNHT time series. In the years 1930–1960, the time series are homogeneous. Weak shifts are visible in 1960, 1963 (for unknown reasons) and in 1973, which can be attributed to a station relocation according to S. Schroeder’s metadata database (Tschudin and Schroeder, 2013).

Panels 6c, d shows the different adjustments calculated with RAOBCORE 2.0. They reduce the detected jumps, as can be seen from the adjusted innovation time series in Fig. 6e, f, which are definitely more homogeneous than the raw series presented in Fig. 6a, b. No visible wind direction inhomogeneities remain. The inhomogeneities found may have contributed to the apparent strong wind anomalies found by Brönnimann et al. (2009) during the dust bowl drought. They would likely be weaker if homogenized data

had been used. This just shows how important it is to eliminate inhomogeneities before doing climate studies.

Wind speed adjustments are not constant in time, since they depend on wind speed itself. The breaks in the early 1960s are no longer visible in the adjusted innovation time series, only one shift in 1972 has not been adjusted. The adjustments for wind speed and direction can be split into U and V wind component adjustments, with a simple vectorial decomposition, and these can be applied in order to adjust U and V wind components. Due to the predominantly zonal wind direction, the wind direction adjustments are projected more onto the V wind components whereas wind speed adjustments affect mainly the U component. Figure 8 shows unadjusted innovations, adjustments and adjusted innovations of U and V for the 700 hPa level. Note that the V adjustments (panel d) have a negative mean because of the negative wind direction adjustment and the mostly westerly winds. The adjusted time series have no more shifts but also

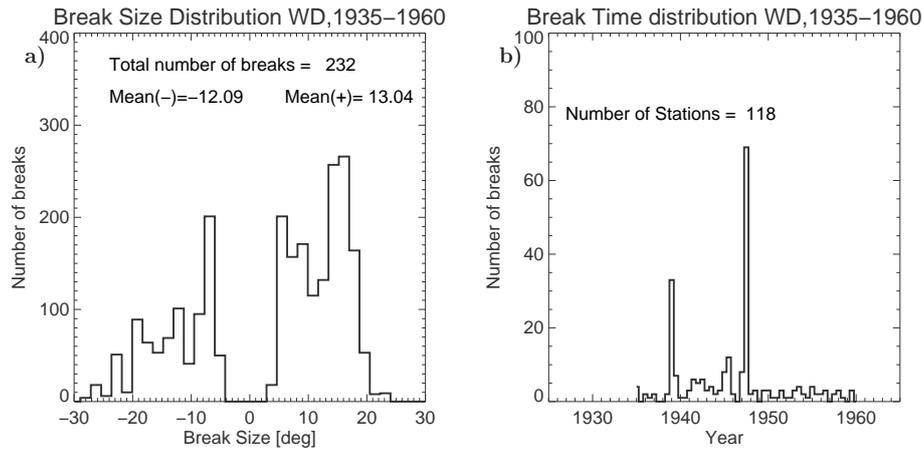


Figure 11. (a) size and (b) temporal distribution of wind direction breaks at North American stations (WMO numbers between 70 000 and 75 000) in the period 1935–1960; all available stations are analyzed. The means of positive and negative breaks are indicated.

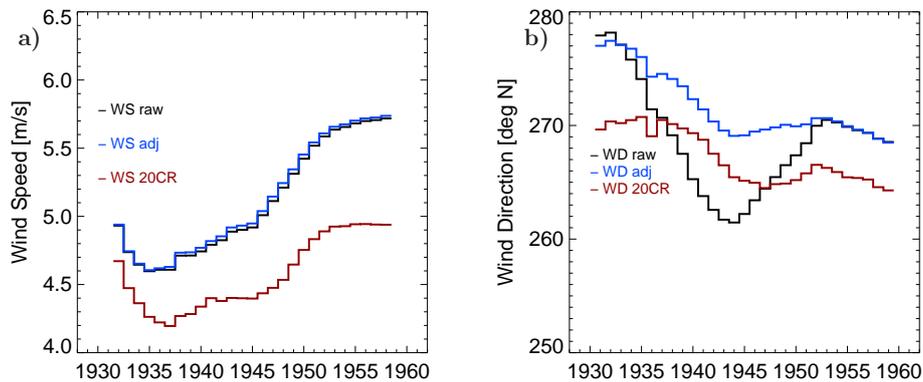


Figure 12. Unadjusted (raw), adjusted and NOAA-20CR wind speeds (a) and directions (b) in the period 1940–1960 at 700 hPa, over a selected group of stations with long and complete time series (WMO numbers: 70261, 72206, 72208, 72250, 72261, 72293, 72317, 72363, 72365, 72405, 72469, 72493, 72518, 72528, 72562, 72572, 72597, 72632, 72681, 72698, 72764, 74626, 72231, 72334, 72374). Running mean: 8 years.

have reduced variance. Generally it has been found that wind speed breaks over the US are relatively rare and relatively weak. Already the unadjusted innovations often look homogeneous, which indicates good quality of both observations and NOAA-20CR.

Wind speed breaks are weak at Bismarck, but at other places they can be quite strong, e.g., in Athens (016716, Greece). This station has already been studied by Gruber and Haimberger (2008). Figure 9 shows the innovation time series at 300 hPa at 00:00 and 12:00 UTC. In 1993, 1987, 1983 and 1979, four massive breaks are detected by the SNHT. The resulting adjustments calculated by RAOBCORE 2.0 and the adjusted innovation time series are shown in the middle and lower panels of Fig. 9. All breaks already detected by Gruber and Haimberger (2008) could be detected again, although a less accurate reference (NOAA-20CR instead of ERA-40) was used. Also, the wind direction series at Athens seems to suffer from two inhomogeneities, but these are not removed

since they are not vertically constant (close to linear growth) and they change sign. As such they are not attributable to wrong north alignment.

4.2 Regional and global trends

The overall beneficial impact of the applied homogeneity adjustments is best visible in maps of trends of both wind direction and speed before and after the data have been treated with RAOBCORE 2.0. The raw, the adjusted and adjustment time series are available at daily resolution, but for trend calculations monthly averages have been used, taking particular care of the sampling problem as noted above. The trend calculation is easy for wind speed (scalar variable), but more challenging for wind direction, which is derived from U and V trends, as explained in Appendix A.

While maps of trends do not reveal all inhomogeneities, they are rather sensitive to inhomogeneities and the comparison with trends at neighboring stations helps to understand

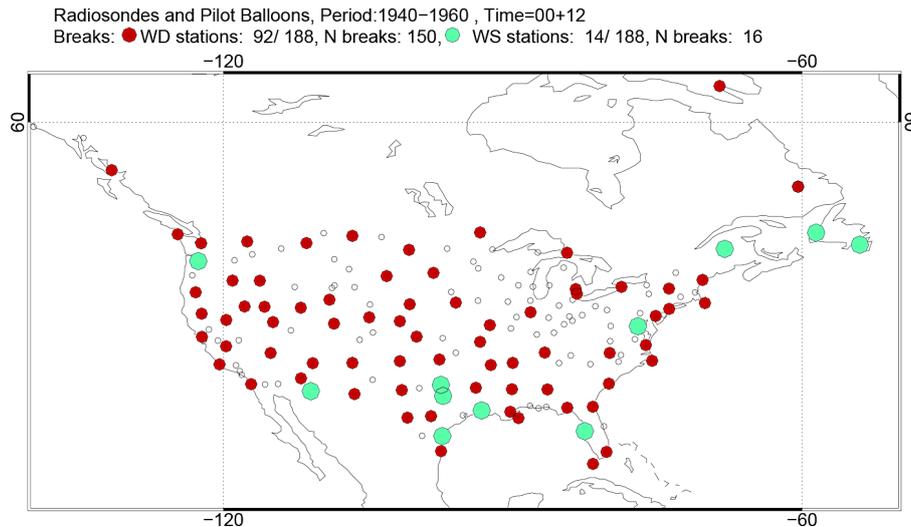


Figure 13. Stations with breaks in wind direction (red bullets) and speed (green bullets) for the period 1940–1960, as well as stations with no breaks (open circles). 188 stations with at least 10 years of observations were analyzed. For wind direction, 92 stations show 150 breaks in this period, whereas only stations with 16 wind speed breaks were found.

where the raw data are affected by them. To estimate the spatial consistency of the trends, the *cost* function introduced by Haimberger (2007) has been used. In this way, smaller improvements in spatial homogeneity can also be objectively measured.

4.2.1 Wind trends over the US in the early period 1940–1960

In 1940 the US already had a well established upper-air network, whereas observations were very rare in other regions. Pilot balloons and radiosondes were launched twice daily, typically reaching the 500 hPa level.

As our first example, we show wind direction trends at 700 hPa in the period 1940–1960 over the US in Fig. 10a. The arrows indicate the mean wind speed over the considered period and point to the mean wind direction. The color scale tells us about the wind direction trend for the investigated period. The figure shows suspiciously strong trends (more than 10° over 10 years, mainly located over the central US; station Bismarck, North Dakota, 072764, belongs to this group, see Fig. 6). Applying RAOBCORE2.0 and using the NOAA-20CR as a reference, around 150 wind direction breaks were detected in over 45 station records in the period 1935–1958. The applied adjustments are shown in Fig. 10b, where the arrows represent the adjustment vectors. For the period 1940–1960, the adjustments are such that winds change from westerly to more northwesterly in most cases (since the breaks in the opposite direction mostly occur before 1940), stressing that all the stations are affected by a similar north-alignment-related bias. The colors show how the adjustments affect the wind direction trends. The adjusted trends in Fig. 10c look much more homogeneous and the cost function has also been

reduced by a factor of 3 due to the improved spatial and temporal homogeneity.

The nature of the breaks affecting the USA in the period 1935–1960 is highlighted in Fig. 11, which shows that the break distribution is clearly bimodal with absolute values of break sizes definitely greater than 0. This indicates that the adjustments are applied only where the break significance has been carefully verified. Moreover, the break time distribution shows that most breaks occurred in 1939 and 1948. S. Schroeder’s metadata database (Tschudin and Schroeder, 2013) indicates a change from theodolite pilot balloons to first-generation radiosondes in 1939, as was the case at Bismarck. In many cases there are metadata events indicating measurement system changes in this database in the above cited years. However, the information does not reveal what exactly caused these changes in the north alignment. According to our estimates, this problem was corrected more than a decade later. If one investigates the break distributions shown in Fig. 11 for the years 1938–1941 and separately for 1947–1951, one obtains two tight and approximately Gaussian distributions, with means of -12° and $+13^\circ$, respectively.

In order to get an idea of how the observed mean wind speed and direction looked like before and after the break analysis, a subset of central US stations was averaged year by year for both wind speed and direction. As is well visible in Fig. 12, wind directions have been strongly adjusted in the period 1935–1948, whereas wind speed shows only slight changes due to the homogenization. The sampling problem does not play a major role for this figure, since, at 700 hPa, most series involved were almost complete. Thus, and because it is visible also in the NOAA-20CR, the increase in wind speed during this period is trusted. Some in-stationarity in wind direction is visible as well. However, this is also

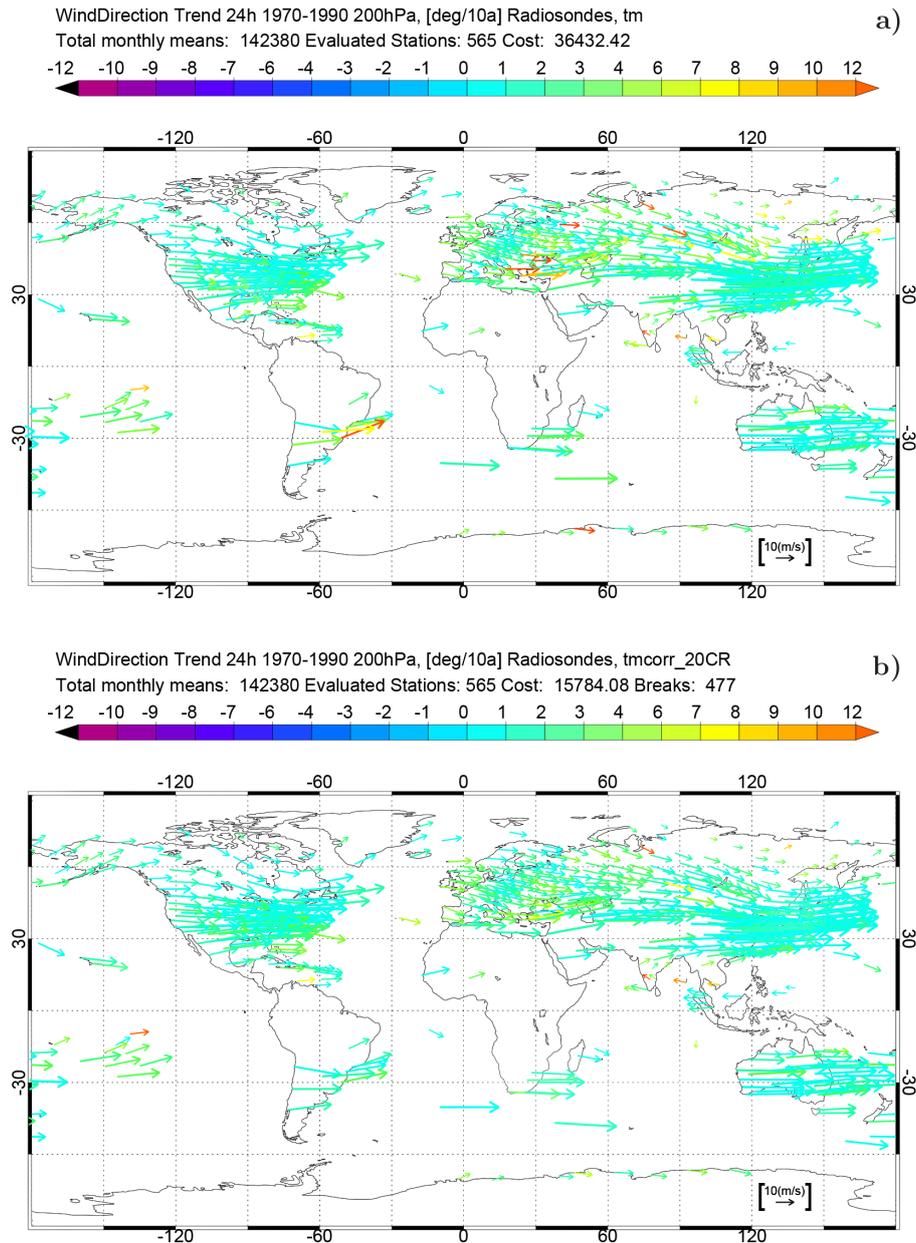


Figure 14. Global observed wind direction trends at 200hPa for the period 1970–1990, (a) before adjustment, (b) after adjustment with RAOBCORE. 565 WMO stations with less than three missing years in the inquired period have been used. The length, direction and color of the arrows have the same meaning as in Fig. 10.

found in the NOAA-20CR. The offset by several degrees is less than the measurement increment of wind direction ($5\text{--}10^\circ$).

Figure 13 shows that 92 out of 188 US stations reporting wind observations in the period 1940–1960 with at least 10 years of observations had breaks in wind direction (150 breaks in total). Only 14 stations were affected by 16 wind speed breaks.

4.2.2 Global trends in more recent times

The upper-air wind observation coverage is fairly global from 1950 onward, with a significant improvement in 1958, the International Geophysical Year. The inclusion of pilot balloon data is essential to get acceptable coverage in the Tropics and Southern Hemisphere in the early years. Inhomogeneities do, however, also occur in later periods. Figure 14 shows global wind direction trend maps at 200 hPa for the period 1970–1990. Quite a few stations show suspiciously strong trends compared to their neighbors. 477

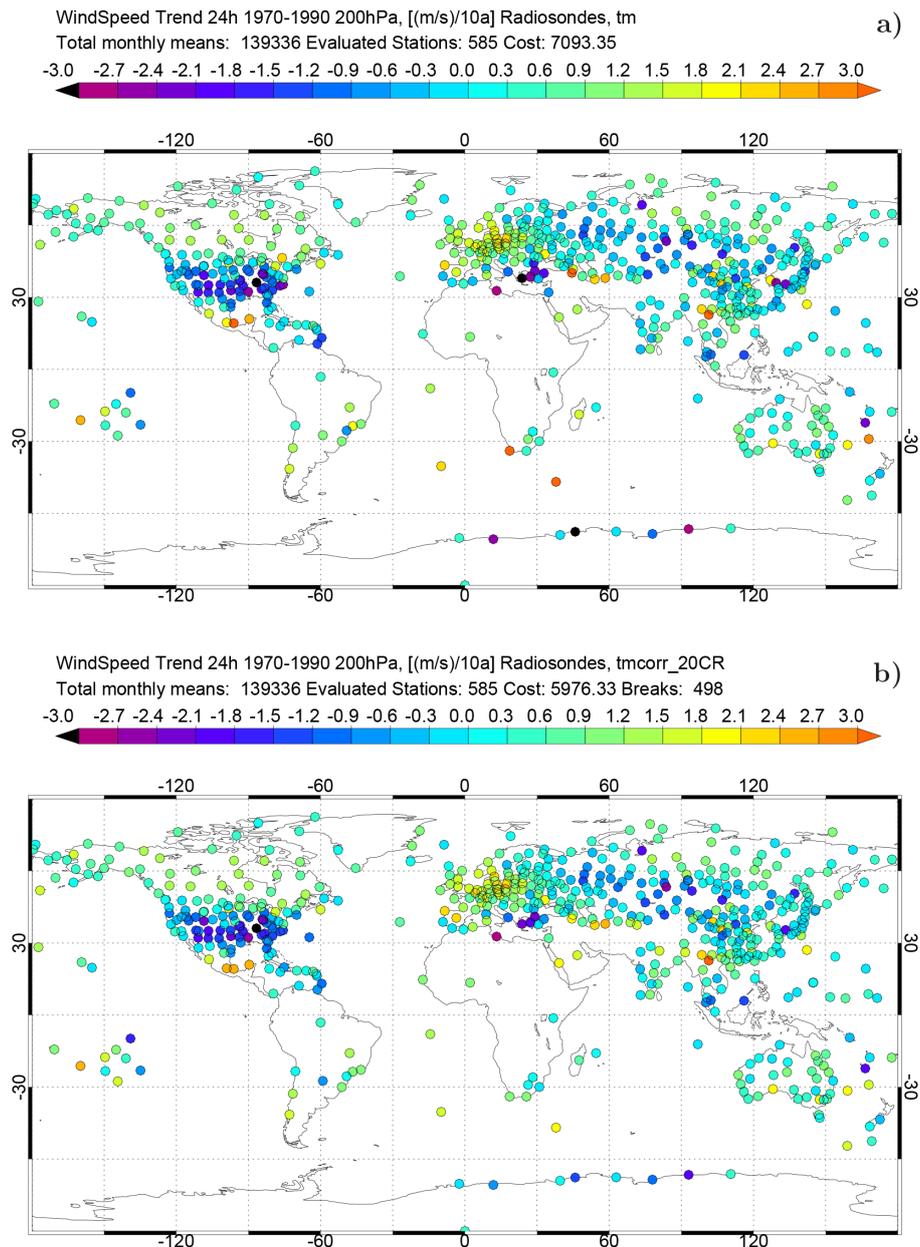


Figure 15. Global observed wind speed trends at 200hPa for the period 1970–1990, (a) before adjustment, (b) after adjustment with RAOBCORE. 585 WMO stations with less than three missing years in the inquired period have been used. Cost values in legend indicate moderate reduction of trend heterogeneity.

breaks have been detected and adjusted by RAOBCORE and the cost function is reduced to about 40 %, which is substantial. The adjusted trends in Fig. 14 show only very few remaining outliers.

Wind speed trend patterns are already relatively smooth without homogenization, as shown in Fig. 15 at 200 hPa. Isolated outliers such as Athens (see Fig. 9) or stations in Turkey, Algeria, the Republic of South Africa, California, Mexico and Antarctica are detected and in most cases adjusted. The cost function is reduced only slightly, by 15 %

(Fig. 15). One could have achieved a lower cost function value by setting wind speed break thresholds lower, but we decided to be more conservative. Some strong trend features are also present in the NOAA-20CR (which has a cost value of 3900) as well, such as the strong deceleration of winds over the US, so we do not expect the cost function to become much lower.

Figure 16a shows the size distribution of wind direction breaks over the whole investigated period (1920–2010), whereas (b) shows the temporal distribution of the breaks.

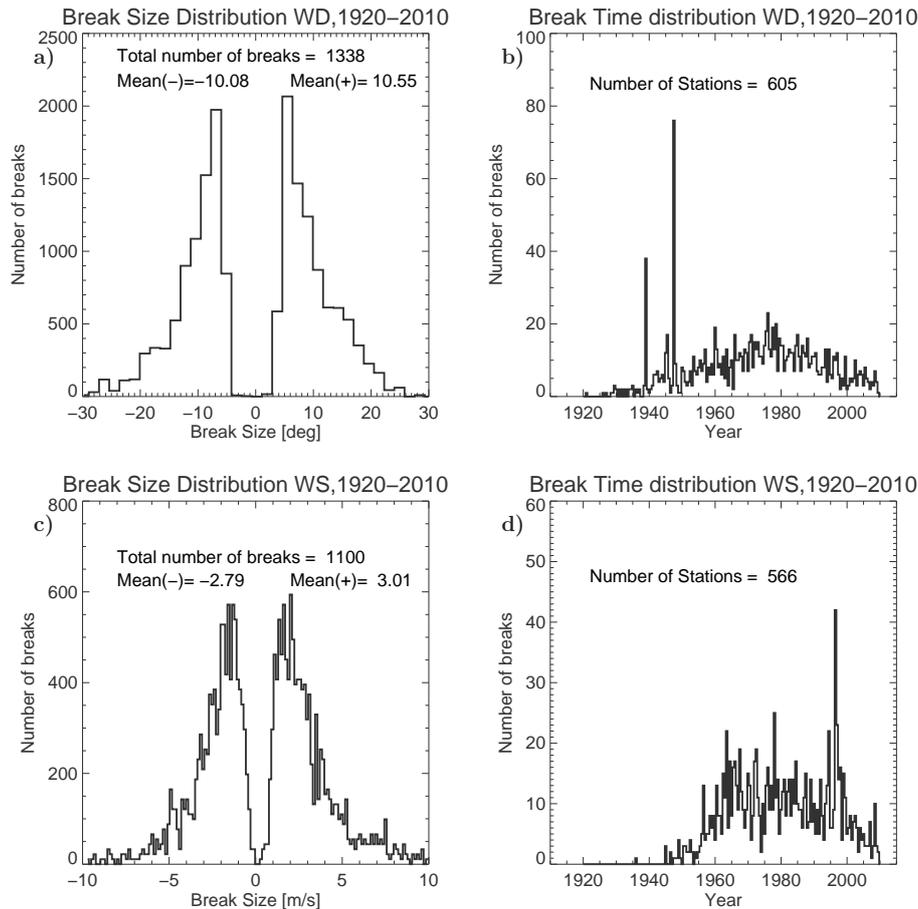


Figure 16. (a) Wind direction break size and (b) time distributions for the whole period investigated (1900–2010), for all available stations. (c) shows corresponding wind speed break size and (d) break time distributions.

The two peaks around 1937/8 and 1947/8 stem from the problems of the early US stations. Afterwards the temporal distribution is smooth. The size distribution shows that really only larger wind direction breaks have been adjusted. Wind speed break sizes in Fig. 16a again show a bimodal break size distribution. Regarding the time distribution in (b), we observe sizeable break numbers only after 1945, most probably because, before, the sparse (only in the USA) soundings did not reach high levels where the wind speed is higher and problems with speed measurements are more likely. The prominent peak in 1998 can be attributed mainly to stations with breaks in the Middle East, northern Africa and South Asia (together ≈ 25 stations). We found no indication of a break in the NOAA-20CR when we compared difference series of ERA-Interim and NOAA-20CR wind speeds (not shown). A total of 1100 wind speed breaks at 566 stations (out of 2924 stations with more than 1 year of data) were detected. In addition, 1338 wind direction breaks at 605 stations have been found.

5 Conclusions

This paper documents our efforts to homogenize the global wind radiosonde and pilot balloon archive described in Part 1 (Ramella-Pralungo et al., 2014), using the NOAA-20CR (Compo et al., 2011) as reference. Since this reanalysis was produced using only surface data it is independent of upper-air data. The already well-known RAOBCORE method (Haimberger, 2007; Gruber and Haimberger, 2008) has been extended and reinforced such that it is able to treat temperature (results are not shown in this work) and wind together.

In contrast to Gruber and Haimberger (2008), we analyzed innovation time series of wind speed and direction only (not U and V) for break detection, since the measurement instruments report speed and direction and the breaks are related to biases linked to the instrument itself. For wind direction we looked specifically for north alignment errors. Regarding wind speed, one expects breaks at those levels where the mean wind speed is particularly large, i.e., near jet streams. When a break was detected, the adjustments were made by

applying a constant factor estimated from comparing the wind speeds and wind speed innovations before and after the break. Combining information from wind speed and direction, the U and V wind components could also be adjusted.

Since the NOAA-20CR was used as reference and since a much more comprehensive input data set was used, several inhomogeneities in the period before 1958 could be detected and adjusted, the most prominent being a pervasive wind direction bias of up to 15° over the central US between 1938 and 1948. After 1960, inhomogeneities are relatively rare compared to temperature inhomogeneities. Fewer breaks have been detected and adjusted compared to Gruber and Haimberger (2008). The reasons are, on the one hand, inhomogeneities in the reference used then (ERA-40, which has some shifts, e.g., in 1972, 1974, 1976, 1979, 1986, mostly related to the satellite observing system; Uppala et al., 2005). Therefore, we did not consider the breaks detected by Gruber and Haimberger (2008) as metadata for the present adjustment procedure. On the other hand, the ERA-40 background departures were much smaller than the NOAA-20CR analysis departures used here. For example σ_u in the period 1979–2001 at 200 hPa at station 35 229 (Aktyubinsk) is 8.1 m s^{-1} for NOAA-20CR analysis departures, 5.9 m s^{-1} for ERA-40 background departures and 4.1 m s^{-1} for ERA-Interim departures. This is simply because upper-air observations have been assimilated in the latter two reanalyses. It is quite likely that some of the smaller shifts that could be detected with ERA-40 could not be detected with our adjustment system simply because of the higher noise level, which makes the SNHT less sensitive (σ is in the denominator of Eq. 4).

The adjustments could be used as input in a reanalysis bias correction scheme or they could be used as a reference to test a future variational wind bias adjustment scheme for radiosonde wind biases with the basic algorithm similar to the one used for adjusting satellite radiances (Dee and Uppala, 2009). In particular, wind direction biases seem good candidates for variational adjustment since wind direction is well constrained in a state-of-the-art multivariate data assimilation system and expected model biases in wind direction are much smaller than the biases found in some of the observation records.

The wind data set developed is the most comprehensive homogenized data set of its kind. It is an ideal basis for estimating the tropical zonal mean warming maximum through exploiting the thermal wind relationship, as pioneered by Allen and Sherwood (2008). The large number of pilot balloon records should allow us to extend these estimates back to 1950, and, because of the more complete data, higher accuracy can be expected compared to their study. It may be very interesting as well to check whether the global mean kinetic energy has increased in the past decades (Bengtsson et al., 2004; Kung, 1966).

Although the data are now homogenized, one should be aware that observed wind speeds have a sampling bias. In the early period, high wind speeds are underrepresented because of measurement limitations. This effect still has to be taken into account when calculating monthly means and trends.

Appendix A: Wind direction trend

While estimating trends of scalar variables is straightforward, it is more challenging for wind variables, especially wind direction. The wind vector in polar coordinates can be expressed as $ff e^{i(dd \cdot \pi/180)}$, which is a nonlinear product. Calculating the wind direction trends just from regressing dd values over time will not always work, amongst other things because of its periodicity. To avoid such problems we split the wind vector into its orthogonal U and V components and regress those independently over time. Then we transform them back to dd and ff . The detailed steps are as follows:

1. For the selected time window (normally chosen to be 20 years), $\overline{U}_{\text{raw}}$, $\overline{V}_{\text{raw}}$, $\overline{U}_{\text{adj}}$ and $\overline{V}_{\text{adj}}$ are calculated using unadjusted (raw) and adjusted (adj) data (where the mean is taken over the selected time window).
2. From the U_{raw} , V_{raw} and U_{adj} , V_{adj} , time series, the linear trends are calculated. These are multiplied with the length of the time interval to yield differences $\Delta \overline{U}_{\text{raw}}$, $\Delta \overline{V}_{\text{raw}}$, $\Delta \overline{U}_{\text{adj}}$ and $\Delta \overline{V}_{\text{adj}}$.
3. Using the standard inner product definition, the wind direction difference Δdd can be calculated as

$$\Delta dd = \text{acos} \left(\frac{\left(\frac{\overline{U}}{\overline{V}} \right) \cdot \left(\frac{\overline{U} + \Delta \overline{U}}{\overline{V} + \Delta \overline{V}} \right)}{ff(ff + \Delta ff)} \right) \quad (\text{A1})$$

for raw and adj observations.

Dividing Δdd by the length of the time interval used for calculating ΔU , ΔV yields the wind direction trend as it is shown color-coded, e.g., in Fig. 14.

Wind speed trends are calculated similarly, by calculating first the wind speed difference

$$\Delta \overline{ff} = \sqrt{(U + \Delta U)^2 + (V + \Delta V)^2} - \sqrt{U^2 + V^2} \quad (\text{A2})$$

and then dividing it by the time interval.

Appendix B: NetCDF file structure

This Appendix only describes the format of the netCDF adjustment files. The format of the merged archive, which is very similar, has already been described in Part 1 of this paper. The netCDF adjustment files contain the time series with daily resolution (00:00 and 12:00 UTC ascents) at 16 standard pressure levels (10, 20, 30, 50, 70, 100, 150, 200, 250, 300, 400, 500, 700, 850, 925 and 1000 hPa); one file for each variable (temperature, U wind and V wind components) and station. The file names are self-explanatory, such as

123456_corr_t_X.nc for daily adjustments

123456_corr_monthly_t_X.nc for monthly adjustments.

The first digit is either 0 or 1:

- 0: the station has been identified as a WMO station
- 1: the station does not have a WMO identifier.

The next five digits are the WMO station identification number if the first is a 0; otherwise they are the progressive number with which the station has been saved in the respective archive (The Comprehensive Historical Upper-Air Network (CHUAN version 1.7) or the ERA-CLIM historical upper air data sets (ECUD), since only those two archives contain unknown stations). The X refers to the variable reported in the file, and it can be

- T → temperature
- U → U wind component
- V → V wind component
- WS → wind speed
- WD → wind direction.

The $_t$ refers to *time series*, as this is the form in which the data have been stored. The following variables list specifies data type (float or int) and the name of the variable and provides some explanations:

- *integer* → **station** → station ID, the same as the first six digits in the file name;
- *float* → **lat** → station latitude, in degrees north, range $[-90^\circ, 90^\circ]$;
- *float* → **lon** → station longitude, in degrees east, range $[-180^\circ, 180^\circ]$;
- *float* → **alt** → station altitude, in m.a.s.l., range $[-400, 8000]$ m;
- *integer* → **press(pressure_levels)** → array with the 16 standard pressure levels;
- *integer* → **obs_time(obs_time)** → array with launch times (00:00 and 12:00) UTC;
- *integer* → **dates(n_g_values)** → progressive day index from 1900-01-01 with Gregorian calendar, range $[19000101, 20230316]$;
- *float* → **X_AN_raso_corrections(obs_time, pressure_levels, n_g_values)** → this variable contains the actual adjustments; X could be either **T**, **U**, **V**, **WS**, **WD**, as for the file names.

The arrays have dimensions

$$\text{obs}(\text{obs_time}, \text{pressure_levels}, \text{n_g_values})$$

. In this way, we use the minimum number of days in order to map the time series. Note that for wind speed and direction only the adjustment files are available. If one

wants to also have the raw data in terms of wind speed and wind direction, one has to calculate those from U and V .

With the same conventions, we also have

- *float* → *X_AN_break_uncertainties(obs_time, pressure_levels, n_g_values)* → the standard errors of the break size estimates
- *float* → *X_AN_breaks(obs_time, pressure_levels, n_g_values)* → the break profiles; these are identical with adjustments for temperature but are different from adjustments for wind direction.

The monthly files have one additional variable:

- *integer* → *dates(n_g_values)* → Number of data used for producing the monthly mean

Both daily and monthly files have the following global attributes:

- *conventions* = “CF-1.4” → NetCDF files conventions;
- *title* → project title;
- *institution* → file owner “University of Vienna”;
- *history* → when the file has been produced;
- *data type* → “radiosonde measurement bias corrections”;
- *source* → source archive;
- *references:* → www.univie.ac.at/theoret-met/research/raobcore.

Routines for reading the archived NetCDF files are available in FORTRAN and IDL and can be downloaded from [doi:10.1594/PANGAEA.823609](https://doi.org/10.1594/PANGAEA.823609).

Acknowledgements. This work has been funded by projects P21772-N22 and P25260-N29 of the Austrian Fonds zur Förderung der wissenschaftlichen Forschung (FWF), as well as by the EU 7th Framework Programme collaborative project ERA-CLIM (Grant No. 265229). The authors thank the collaborators within ERA-CLIM, most notably Hans Hersbach, Alexander Stickler and Stefan Brönnimann, for their support and constructive comments. The comments of two anonymous reviewers, also regarding the format of the adjustment files, substantially improved the manuscript.

Edited by: G. König-Langlo

References

- Alexandersson, H.: A homogeneity test applied to precipitation data, *Int. J. Climatol.*, 6, 661–675, 1986.
- Allan, R., Brohan, P., Compo, G. P., Stone, R., Luterbacher, J., and Brönnimann, S.: The international atmospheric circulation reconstructions over the earth (ACRE) initiative, *B. Am. Meteorol. Soc.*, 92, 1421–1425, doi:10.1175/2011BAMS3218.1, 2011.
- Allan, R. J. and Ansell, T. J.: A new globally complete monthly historical mean sea level pressure data set (HadSLP2): 1850–2004, *J. Climate*, 19, 5816–5842, 2006.
- Allen, R. J. and Sherwood, S. C.: Warming maximum in the tropical upper troposphere deduced from thermal winds, *Nat. Geosci.*, 1, 399–403, 2008.
- Bengtsson, L., Hagemann, S., and Hodges, K. I.: Can climate trends be calculated from reanalysis data?, *J. Geophys. Res.*, 109, D11111–D11119, 2004.
- Brönnimann, S.: A historical upper air-data set for the 1939–44 period, *Int. J. Climatol.*, 23, 769–791, 2003.
- Brönnimann, S., Stickler, A., Griesser, T., Ewen, T., Grant, A. N., Fischer, A. M., Schraner, M., Peter, T., Rozanov, E., and Ross, T.: Exceptional atmospheric circulation during the “Dust Bowl”, *36, L08802*, doi:10.1029/2009GL037612, 2009.
- Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Å., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The twentieth century reanalysis project, *Q. J. Roy. Meteor. Soc.*, 137A, 1–28, doi:10.1002/qj.776, 2011.
- Dee, D. P. and Uppala, S. M.: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis, *Q. J. Roy. Meteor. Soc.*, 135, 1830–1841, 2009.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kallberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553–597, doi:10.1002/qj.828, 2011.
- Durre, I., Vose, R., and Wuertz, D. B.: Overview of the Integrated Global Radiosonde Archive, *J. Climate*, 19, 53–68, 2006.
- Ebita, A., Kobayashi, S., Ota, Y., Moriya, M., Kumabe, R., Onogi, K., Harada, Y., Yasui, S., Miyaoka, K., Takahashi, K., Kama-hori, H., Kobayashi, C., Endo, H., Soma, M., Oikawa, Y., and Ishimizu, T.: The Japanese 55-year reanalysis “JRA-55”: An interim report, *SOLA*, 7, 149–152, doi:10.2151/sola.2011-038, 2011.
- Ewen, T., Brönnimann, S., and Annis, J. L.: An extended Pacific North American index from upper air historical data back to 1922, *J. Climate*, 21, 1295–1308, 2008a.
- Ewen, T., Grant, A., and Brönnimann, S.: A monthly upperair data set for north america back to 1922, *Mon. Weather Rev.*, 136, 1792–1805, 2008b.
- Ferguson, C. R. and Villarini, G.: An evaluation of the statistical homogeneity of the twentieth century reanalysis, *Clim. Dynam.*, 42, 2841–2866, doi:10.1007/s00382-013-1996-1, 2013.
- Graystone, P.: Meteorological office discussion on tropical meteorology, *Met. Magazine*, 88, 117ff., 1959.
- Gruber, C. and Haimberger, L.: On the homogeneity of radiosonde wind time series, *Meteorol. Z.*, 17, 631–643, 2008.
- Haimberger, L.: Homogenization of radiosonde temperature time series using innovation statistics, *J. Climate*, 20, 1377–1403, 2007.
- Haimberger, L., Gruber, C., Sperka, S., and Tavolato, C.: Homogenization of the global radiosonde temperature and wind dataset using innovation statistics from reanalyses, in: *Proceedings of the Third WCRP International Conference on Reanalyses*, 5 pp., http://wcrp.ipsl.jussieu.fr/Workshops/Reanalysis2008/Documents/G4-411_ea.pdf, 2008.
- Haimberger, L., Tavolato, C., and Sperka, S.: Homogenization of the global radiosonde temperature dataset through combined comparison with reanalysis background series and neighboring stations, *J. Climate*, 25, 8108–8131, 2012.
- Hartmann, D. L., Klein Tank, A. M. G., and Rusticucci, M. (Eds.): *IPCC 2013: The Physical Science Basis, Chapter Observations: Atmosphere and Surface*, 159–254, IPCC, 2013.
- Hollingsworth, A., Shaw, D. B., Lönnberg, P., Illari, L., and Simmons, A. J.: Monitoring of observation and analysis quality by a data assimilation system, *Mon. Weather Rev.*, 114, 861–879, 1986.
- Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R., and Fiorino, M.: The NCEP/NCAR 50-year Reanalysis: Monthly Mean CDROM and Documentation, *B. Am. Meteorol. Soc.*, 82, 247–267, 2001.
- Kung, E. C.: Large-scale balance of kinetic energy in the atmosphere, *Mon. Weather Rev.*, 94, 627–640, 1966.
- Luers, J. K. and Eskridge, R. E.: Temperature corrections for the VIZ and Vaisala radiosondes, *J. Appl. Meteorol.*, 34, 1241–1253, 1995.
- Poli, P., Hersbach, H., Tan, D., Dee, D., Thépaut, J. N., Simmons, A., Peubey, C., Laloyaux, P., Komori, T., Berrisford, P., Dragani, R., Tremolet, Y., Holm, E., Bonavita, M., Isaksen, L., and Fisher, M.: The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20C), ECMWF, 2013.
- Ramella Pralungo, L., Haimberger, L., Stickler, A., and Brönnimann, S.: A global radiosonde and tracked balloon archive on 16 pressure levels (GRASP) back to 1905 – Part 1: Merging and

- interpolation to 00:00 and 12:00 GMT, *Earth Syst. Sci. Data*, 6, 185–200, doi:10.5194/essd-6-185-2014, 2014.
- Santer, B. D., Hnilo, J., Wigley, T. M. L., Boyle, J. S., Doutriaux, C., Fiorino, M., Parker, D. E., and Taylor, K. E.: Uncertainties in observationally based estimates of temperature change in the free atmosphere, *J. Geophys. Res.*, 104, 6305–6333, 1999.
- Santer, B. D., Wigley, T. M. L., Mears, C., Wentz, F. J., Klein, S. A., Seidel, D. J., Taylor, K. E., Thorne, P. W., Wehner, M. F., Geckler, P. J., Boyle, J. S., Collins, W. D., Dixon, K. W., Doutriaux, C., Free, M., Fu, Q., Hansen, J. E., Jones, G. S., Ruedy, R., Karl, T. R., Lanzante, J. R., Meehl, G. A., Ramaswamy, V., Russell, G., and Schmidt, G. A.: Amplification of surface temperature trends and variability in the tropical atmosphere, *Science*, 309, 1551–1556, 2005.
- Scherhag, R.: Langperiodische Schwankungen der Stratosphärenzirkulation, *Beitr. Phys. Atmos.*, 35, 245–251, 1962.
- Stickler, A. and Brönnimann, S.: Significant bias of the ncep/ncar and twentieth century reanalyses relative to pilot balloon observations over the west african monsoon region (1940–57), *Q. J. Roy. Meteor. Soc.*, 137, 1400–1416, doi:10.1002/qj.854, 2011.
- Stickler, A., Grant, A. N., Ewen, T., Ross, T. F., Vose, R. S., Comeaux, J., Bessemoulin, P., Jylhä, K., Adam, W. K., Jeannet, P., Nagurny, A., Sterin, A. M., Allan, R., Compo, G. P., Griesser, T., and Brönnimann, S.: The comprehensive historical upper air network (CHUAN), *B. Am. Meteorol. Soc.*, 91, 741–751, doi:10.1175/2009BAMS2852.1, 2010.
- Stickler, A., Brönnimann, S., Jourdain, S., Roucaute, E., Sterin, A., Nikolaev, D., Valente, M. A., Wartenburger, R., Hersbach, H., Ramella-Pralungo, L., and Dee, D.: Description of the ERA-CLIM historical upper-air data, *Earth Syst. Sci. Data*, 6, 29–48, doi:10.5194/essd-6-29-2014, 2014.
- Thorne, P., Parker, D. E., Tett, S. F. B., Jones, P. D., McCarthy, M., Coleman, H., and Brohan, P.: Revisiting radiosonde upper-air temperatures from 1958 to 2002, *J. Geophys. Res.*, 110, D18105, doi:10.1029/2004JD005753, 2005.
- Tschudin, M. E. and Schroeder, S. R.: Time constant estimates for radiosonde temperature sensors, *J. Atmos. Ocean. Tech.*, 30, 40–56, doi:10.1175/JTECH-D-11-00181.1, 2013.
- Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., da Costa Bechtold, V., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, I., Janssen, P. A. E. M., Jenne, R., McNally, A. P., Mahfouf, J. F., Morcrette, J. J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 Re-analysis, *Q. J. Roy. Meteor. Soc.*, 131, 2961–3012, 2005.
- Vautard, R., Cattiaux, J., Yiou, P., Thepaut, J. N., and Ciais, P.: Northern hemisphere atmospheric stilling partly attributed to an increase in surface roughness, *Nat. Geosci.*, 3, 756–761, 2010.