Random trend errors in climate station data due to inhomogeneities

Ralf Lindau | Victor Venema

Institute for Geosciences and Meteorology, University of Bonn, Bonn, Germany

Correspondence
Ralf Lindau, Institute for Geosciences and Meteorology, University of Bonn, Auf dem Hügel 20, D 53121 Bonn, Germany.
Email: rlindau@uni-bonn.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: LI 2830/1-1

Abstract
Inhomogeneities in station series are a large part of the uncertainty budget of long-term temperature trend estimates. This article introduces two analytical equations for the dependence of the station trend uncertainty on the statistical properties of the inhomogeneities. One equation is for inhomogeneities that act as random deviations from a fixed baseline, where the deviation levels are random and independent. The second equation is for inhomogeneities, which behave like Brownian Motion (BM), where not the levels themselves but the jumps between them are independent. It shows that BM-type breaks introduce much larger trend errors, growing linearly with the number of breaks. Using the information about type, strength, and frequency of the breaks for the United States and Germany, the random trend errors for these two countries are calculated for the period 1901–2000. An alternative and independent estimate is obtained by an empirical approach, exploiting the distance dependence of the trend variability for neighbouring station pairs. Both methods (empirical and analytical) find that the station trend uncertainty is larger in the United States (0.71 and 0.82°C per century, respectively) than in Germany (0.50 and 0.58°C per century, respectively). The good agreement of the analytical and the empirical estimate gives confidence in the methods to assess trend uncertainties, as well as in the method to determine the statistical properties of the break inhomogeneities.

KEYWORDS
break types, Brownian motion, climate stations, homogenization, inhomogeneities, random deviations, trend errors

1 | INTRODUCTION

Trends in climate records are affected by inhomogeneities, which are mostly caused by relocations of the stations or changes in the operational measuring techniques (Venema et al., 2019). These observational changes introduce a break signal into the climate record, which is modelled as a step function consisting of homogeneous subperiods interrupted by sudden jumps. The sizes of these breaks are expected to follow a normal distribution (Menne and Williams, 2005), which means that series also contain many small breaks, which are too small to be detected but do influence the homogenization and can still produce considerable trend errors. Based on detected breaks and expert judgement, the standard deviation of the jump size distribution is estimated to be about 0.8 K for European networks (Auer et al., 2007; Brunetti et al., 2006; Caussinus and Mestre, 2004; Della-Marta et al., 2004; Venema et al., 2012). In the United
States, jumps are thought to be somewhat smaller (Menne et al., 2009).

A few of these spurious jumps can be sufficient to perturb the actual trend considerably. Changes in the immediate or wider surroundings of the station may also produce sudden jumps or more gradual increases. The latter can be modelled by multiple small jumps (Venema et al., 2012), and this study will only consider jumps.

Homogenization algorithms are applied to detect these inhomogeneities and to eliminate their influence as much as possible in a correction step. In most cases, relative homogenization techniques are applied, where the difference time series of a considered candidate station with its neighbouring stations is analysed to reduce the dominating regional climate signal so that the inhomogeneity signal becomes better detectable (Conrad and Pollak, 1950). After the removal of the complicated climate signal, the difference time series is well modelled as a step function representing the inhomogeneities plus noise due to weather differences and observational errors. The signal-to-noise ratio (SNR) can then be defined as the square root of the variance of the break signal divided by the variance of the noise signal. Next to the above mentioned break sizes, the correlations between stations are thus also crucial. Lindau and Venema (2016) concentrated on the detection of breaks and showed that the uncertainty of the detected break positions can be approximated by a Brownian motion (BM) with drift, where the SNR defines the drift size. This fits qualitatively to earlier numerical results on position errors by Menne and Williams (2009) and to the relationship between the hit rate and the SNR by, for example, Wang (2008).

Lindau and Venema (2018a) showed that, for SNR < 1, problems may occur in the detection part. The problem is not simply that the detection itself may become difficult. More fundamental is that breakpoints explaining only the noise are erroneously assessed to be significant, because they also explain a large part of the break variance at the same time. In this way, completely false break positions could be the outcome.

Lindau and Venema (2018b) examined the correction part of homogenization algorithms and concentrated on the trends. They distinguished between biased and unbiased breaks, where the latter are characterized by an average jump size of zero, which on average do not change the trends. Again, the SNR was found to be crucial for the performance of the trend correction. They showed that the trend error of unbiased breaks could only be reduced if the number of breaks divided by the SNR is smaller than 6, even if the detection of breaks is perfect. Thus, for more than six breaks in each time series and SNR = 1, the trend error of the homogenized station series is larger than for the original data. Biased breaks additionally introduce a trend bias, which can be entirely corrected only if the break detection is perfect. However, in realistic cases, where the detection of the break positions is imperfect, the bias is only corrected partly.

Lindau and Venema (2019) pointed out that it is further important to distinguish between two types of inhomogeneities. On the one hand, there are random deviation (RD)-type breaks, where the levels between two adjacent breakpoints are RDs from a fixed baseline and, on the other hand, BM-type breaks, where the jumps between the levels are random so that the inhomogeneity signal has the characteristics of a BM. Consequently, BM levels have a memory and cannot be modelled as single random numbers but as the sum of n random numbers, where n is the number of all previous jumps. In their study, they presented a method to provide an average quantification of both break types in a given dataset. For RD-type breaks, the strength \( \sigma_\delta^2 \) and the frequency \( p_\delta \) could be separately given, whereas for BM-type breaks, only the combined parameter consisting of the product of strength and frequency \( p_\delta \sigma_\delta^2 \) could be derived. These break parameters were determined for the contiguous United States, where a mixture of RD- and BM-type breaks were detected and for Germany, where fewer and smaller breaks were found. An important difference was that German data contain only RD breaks and no BM breaks.

Dunn et al. (2014) pointed out that there is a lack of small breaks sizes in the distribution of detected breaks, which was confirmed by Yosef et al. (2018) and Trewin (2013). Lindau and Venema (2019) derived the number of breaks without applying homogenization algorithms and found more than the double number of breaks compared with these traditional estimates. However, this is plausible because their estimate comprises also the large number of small size breaks, which are not detected by the homogenization algorithms.

In this article, we concentrate on the station trend errors as they are contained in the original data before any homogenization. Besides homogenization, quality control is an important issue in data processing, especially for the determination of extremes. However, Hunziker et al. (2018) showed that the error of mean temperature trends depends strongly on whether the data are homogenized or not, whereas the kind of applied quality control is less important. As the aim is to derive trend errors of the raw data neither homogenization nor specific quality control is applied.

In Section 2, we derive the expected trend uncertainty caused by the two inhomogeneity types (RD and BM) as a function of strength and frequency of the breaks. Using the derived parameters from Lindau and Venema (2019), an error estimate of the station trends in the United States and in Germany can be given; see Section 3. These estimates are then compared in Section 4 to the trend difference of neighbouring stations, which provide a second and independent measure of the trend uncertainty. Thus, an indirect method that derives the expected trend variance from the known breaks characteristics is compared with the results directly calculated from the data.

Please note that only the random part of the uncertainty is considered because both Lindau and Venema (2019) as well
as the empirical method applied in Section 4 use difference time series of neighbouring stations. This is necessary to eliminate the dominating climate signal, but it removes at the same time any possible mean trend bias caused by inhomogeneities.

2 | DERIVATION OF THE EXPECTED ERROR VARIANCE OF TRENDS

In this section, we determine the theoretical trend variance (caused by inhomogeneities), which is expected for both break types. These relationships are compared with numerical results based on simulated station series containing inhomogeneities. For this numerical part, we simulate the inhomogeneity signal by a step function of length \( l = 100 \) years containing a number of \( k \) breakpoints at random positions \( r_i \), \( i = 1 \) to \( k \). In these time series, breaks occur stochastically with a probability \( p \) so that the number of breaks \( k \) varies between the different time series. The breaks divide the time series into \( k + 1 \) homogeneous subperiods (HSPs) of length \( l_i \). Within a HSP, the step function remains at a constant level. For BM-type breaks, the jumps between the levels are random, independent, and normal distributed with zero mean. For RD-type breaks, the same is true but for the levels themselves.

2.1 | RD-type breaks

We created an ensemble of 10,000 series with a variance of the HSP levels of \( \sigma^2 = 0.5 \) K\(^2\). As each jump between two adjacent levels consists of two of such level deviations, the variance of the jumps is equal to 1 K\(^2\). Then, we calculated the trend of each time series with least square linear regression. Note that the expected mean trend of all step functions is zero; the crucial parameter is the variance of the obtained trends: it represents the uncertainty caused by the inserted inhomogeneities.

Figure 1 shows the obtained trend variance (given by the crosses) for 200 different break rates \( p \) between 0.1 and 20 breaks per century (cty). For break rates well below one, the variance increases steeply with a slope of about 1.2 K\(^2\)-cty\(^{-1}\) as given by the thin straight line. The maximum is reached at a rate of about 2–3 breaks cty\(^{-1}\).

To compare our numerical result with the theoretical expectation, we start with an approximation for the relative trend variance of entire networks \( f^2 \) given by Lindau and Venema (2018b, their eq. (17)):

\[
f^2 = \frac{\text{Var}_{RDNet}(IT)}{\sigma^2} = \frac{24k}{n(k+2)^2}.
\]  

(1)

It is given as the quotient of the variance of the temperature change \( \text{Var}_{RDNet}(IT) \) divided by the variance of the inhomogeneity levels \( \sigma^2 \). Here, \( T \) is the temperature trend and \( l \) the length of the series so that their product gives the temperature change. The relative trend variance \( f^2 \) depends on the number of breaks \( k \) in a station and the number stations \( n \) within a network. To convert the variance of the temperature change into the variance of the trend, we have to divide by the squared length of the time series \( l^2 \). Furthermore, for the station trends, the variance is \( n \) times larger, so the division by \( n \) in Equation (1) has to be omitted. Solving for the variance of the station trends, we obtain:

\[
\text{Var}_{RD}(T) = \frac{24k}{(k+2)^2} \sigma^2.
\]  

(2)

This theoretical approximation is given as thick continuous curve in Figure 1, which reaches its maximum at \( k = 2 \). That a maximum of variance must exist somewhere at moderate break rates is qualitatively plausible. RD breaks cause a random fluctuation of the HSP levels, which is comparable to the effect of noise, but on the longer time scale of the HSPs. Consequently, the variance of the trends converges with increasing number of breaks to zero, as more and more independent values are underlying. However, for zero breaks, the variance does not become maximal but on the contrary is equal to zero. Thus, a maximum must exist somewhere in between, which is obviously reached at \( k = 2 \).

The approximation given in Equation (2) is valid for \( k \) being a constant number of breaks for every time series. Here, we consider, however, stochastically occurring breaks so that the number of breaks will vary according to the
Poisson distribution; this reduces the height of the peak occurring in the numerical result. Thus, to obtain the results for mean break rates, a weighted average with the according Poisson weights $w$ is necessary.

$$w(k) = \frac{\lambda^k e^{-\lambda}}{k!},$$

(3)

where $\lambda$ denotes the average and $k$ the actual number of breaks. The thin curve in Figure 1 shows the result of this (weighted) averaging process. It leads to a reduction especially around the maximum at $p = 2 \text{ cty}^{-1}$ so that the average curve matches the data much better. At a mean rate of $5 \text{ cty}^{-1}$, the thick and thin curve cross, as the curvature (second derivative) changes its sign so that the averaging process leads for larger break rates to a slight increase compared with the original curve. Apart from these details, the approximation given in Equation (2) fits for break rates larger than about $4 \text{ cty}^{-1}$ reasonably well to the modelled data.

### 2.2 BM-type breaks

To create BM-type breaks, we choose a variance of the jumps between two adjacent HSPs of $\sigma_\beta^2 = 1 \text{ K}^2$. Formally, $\sigma_\beta^2$ is twice as large as $\sigma_\delta^2$. In this way, the size of the jumps is equal for both break types because two deviations of the RD signal produce one jump.

Figure 2 shows that the trend variance caused by BM-type breaks quickly becomes much larger than the corresponding variance for RD-type breaks. It grows linearly with the number of breaks instead of decreasing as we have found for RD breaks. Lindau and Venema (2019) already expected a linear variance increase of the BM trend variance and gave an approximation based on the following consideration. The slope of the connection line between the first and the last value of a BM is a rough estimate for the trend. The last value is equal to the sum of $k$ random numbers so that its variance is $k\sigma_\beta^2$. As the first value of a BM is constant and zero, the variance of the temperature change is also $k\sigma_\beta^2$.

However, for the linear trends, we empirically found $1.2 k\sigma_\beta^2$, thus a 20% larger variance. That a discrepancy between approximation and truth exists is not surprising but the intriguing factor of 1.2 asks for an explanation. This is given in detail in the Appendix, where the factor 6/5 is confirmed. Thus, also the theory shows:

$$\text{Var}_{BM}(T) = \frac{6}{5} \frac{p_\beta \sigma_\beta^2}{l},$$

(4)

where $p_\beta$ denotes the frequency of BM-type breaks and $l$ the length of the considered time series. Thus, for BM-type breaks, the variance of the trend grows linearly with the break rate and the slope is $6/5 \text{ K}^2 \text{ cty}^{-1}$. The same slope was found for RD breaks at very small break rates (please compare the thick straight line in Figure 1). This shows that the two break types are difficult to distinguish for small break rates. This is plausible, as it is impossible to assign a single break in a time series to one of the two break types. At least two breaks are necessary to test the independence of consecutive jump heights as it is necessary for the proof of BM-type breaks. For small break rates, this is important to bear in mind.

### 3 INDIRECT CALCULATION OF TREND UNCERTAINTY BY USING THE KNOWN BREAK PARAMETERS

Lindau and Venema (2019) were able to separate the effects of RD- and BM-type breaks by considering the variance of the difference time series $\text{Var}(D(L))$ between two neighbouring climate stations for different time lags $L$. RD breaks produce a saturating exponential signal, whereas BM breaks cause a linear increase of the variance as a function of lag. The noise reflecting the weather differences between two stations affects a lag independent offset that depends only on the noise variance $\sigma_\varepsilon^2$. Summing these three terms, the variance of the difference is a function of time lag $L$ and given by:

$$\text{Var}(D(L)) = 2p_\beta \sigma_\beta^2 L + 4\sigma_\delta^2 (1 - e^{-p_\delta L}) + 4\sigma_\varepsilon^2.$$

(5)

Fitting Equation (5) to the data, Lindau and Venema (2019) obtained break parameters for the United States and
German stations, which are given in Table 1. In the United States, a mean rate of $p_\beta = 17.1$ RD breaks per century is found with a size of $\sigma_\delta^2 = 0.1175 \text{ K}^2$. When we insert these values into Equation (2), we obtain the trend variance caused by RD breaks:

$$\text{Var}_{\text{RD}}(T) = \frac{24k}{(k+2)^2} \sigma_\delta^2 = \frac{24 \times 17.1 \times 0.1175}{19.1^2} \text{ K}^2 \text{ cty}^2 = 0.132 \text{ K}^2 \text{ cty}^2.$$  \hspace{1em} (6a)

The combined BM parameter is $0.45 \text{ K}^2 \text{ cty}^{-1}$. Inserted in Equation (4), we get:

$$\text{Var}_{\text{BM}}(T) = 1.2 \frac{p_\beta \sigma_\beta^2}{l} = 1.2 \frac{0.45 \text{ K}^2 \text{ cty}}{1 \text{ cty}^2} = 0.540 \text{ K}^2 \text{ cty}^2.$$  \hspace{1em} (6b)

As the two break signals are assumed to be independent, their variances can be summed up to obtain the total error variance ($0.672 \text{ K}^2 \text{ cty}^{-2}$). The square root yields the standard deviation of the trend in the United States due to inhomogeneities:

$$\sigma_{US}(T) = 0.82 \text{ K \ cty}.$$  \hspace{1em} (7)

In Germany, they found only RD-type breaks with a frequency of 4.1 breaks $\text{ cty}^{-1}$. Their size is $0.1275 \text{ K}^2$:

$$\text{Var}_{\text{RD}}(T) = \frac{24k}{(k+2)^2} \sigma_\delta^2 = \frac{24 \times 4.1 \times 0.1275}{6.1^2} \text{ K}^2 \text{ cty}^2 = 0.337 \text{ K}^2 \text{ cty}^2.$$  \hspace{1em} (8)

This is considerably larger than the error variance for RD breaks in the United States (compare Equations (6a) and (8)). The main reason is that the break frequency in Germany is about four times lower. This means fewer independent breaks, which increases the random scatter. That fewer breaks causes larger errors may seem counterintuitive, but is indeed true for the unbiased (zero mean) breaks assumed here. Applying the square root to Equation (8) yields again the standard deviation, now for Germany:

$$\sigma_{GER}(T) = 0.58 \text{ K \ cty}.$$  \hspace{1em} (9)

Thus, the overall effect of inhomogeneities is larger in the United States caused by the strong impact of BM-type breaks, which are not found in Germany.

### 4 DIRECT CALCULATION OF THE TREND UNCERTAINTY BY USING THE OBSERVED TREND VARIABILITY

In this section, trends are directly calculated from real observations. The variance of the trend difference between two stations is considered as a function of station distance. The effect of the inhomogeneities is independent of distance and appears as an offset readable at zero distance, whereas with increasing distance between two stations, the variance grows due to real differences in the regional climate.

We used monthly data of the International Surface Temperature Initiative (Rennie et al., 2014) and extracted time series from the United States and Germany with at least 80 years of data in the entire period 1901–2000 and at least 20 years in the baseline period 1961–1990. Then we subtracted the monthly mean for the years 1961 to 1990 to remove the variance of the annual cycle. From this data, least square linear trends are calculated for the period 1901–2000. In the next step, pairs of these station trends are built and sorted into 12 distance classes from 0 to 100 until 1,100 to 1,200 km. All possible pairs of stations are built so that each station occurs $n$ times when $n$ partner stations are found. Within these classes, two parameters are calculated:

The variance of the difference (i.e., the trend differences of pairs of stations) and the sum of the variance (i.e., the variance of all contributing station pairs). The latter fluctuates weakly with distance (Figure 3, uppermost thin curve), but this is only caused by the slightly different selection of the underlying stations; a station will not occur in a distance class, if no partner is found in that specific distance. However, essentially the sum of the variance is a constant describing the (double) variance of the trends in the whole area considered, for example, contiguous United States.

The variance of the trend difference (bold curves) consists of two components: A constant offset describing the variance caused by the inhomogeneities and a second component caused by the different climate zones, which is growing with increasing distance. Thus, the inhomogeneity part, that is, the uncertainty of trends caused by the breaks, can be read at the (marginally extrapolated) axis intercept of the lower curve.

For the upper fat curve representing the United States, the intercept is equal to about $1 \text{ (K cty}^{-1})^2$. From this value, the $SD$ for a single station can be obtained by dividing by
2 (because two stations are contributing) and computing the square root, which yields 0.71 K cty$^{-1}$.

The difference between the thick and the thin curves at zero distance, being for the United States equal to about 0.6 K$^2$ cty$^{-2}$, can be attributed to real climatic differences of the trends. This additional component nearly vanishes for distances larger than 1,200 km. Thus, with 0.6 K$^2$ cty$^{-2}$, we have additionally an estimate for the magnitude of real trend variations on large spatial scales.

An analogous test, but on finer scale for distances below 120 km, shows that both curves remain rather constant (Figure 4). Only for the shortest distance class of 5 km, the lower curve decreases abruptly. However, not only the trend correlation changes but also the correlation of the time series themselves jumps from 0.935 to 0.984. This suggests that undetected identical data segments from different data sources are the reason rather than real climate signals.

The procedure is repeated for Germany where much less data are available so that the two curves are characterized by larger scatter. Nevertheless, clear differences compared with the U.S. results are obvious. The first is that the thick and thin curves are nearly equal and more or less constant over all distances classes. This indicates that climatic differences in the trends play almost no role, although over similar distances there is an effect in the United States. The trends of all stations throughout Germany are similar, only the inhomogeneities impose a constant variance offset. A second difference with the U.S. results is that this level is much lower with about 0.5 K$^2$ cty$^{-2}$. The square root of the half of this value yields 0.5 K cty$^{-1}$. However, due to the sparse underlying data, this value has a larger confidence interval.

In Germany, a reliable analysis of the variance for short distances similar to Figure 4 is difficult, as only few pairs are available for these distance classes. Consequently, the result is dominated by large scatter and therefore not shown.

For U.S. data, the trend variability caused by inhomogeneities is shown to be as large as 1 K$^2$ cty$^{-2}$ for nearby station pairs, which corresponds to a $SD$ of about 0.71 K cty$^{-1}$ for a single station. For Germany, the corresponding value is about 0.5 K cty$^{-1}$. These values are slightly lower, but nevertheless in good agreement with the indirectly derived estimates given in Equations (7) and (9).

5 | DISCUSSION AND CONCLUSIONS

We have developed an analytical and an empirical method to estimate the uncertainty of secular stations trends due to inhomogeneities. The empirical method found for contiguous United States and Germany a random trend uncertainty of about 0.71 and 0.5 K cty$^{-1}$, respectively. The alternative method to use the analytically derived Equations (2) and (4) together with the statistical properties from Lindau and Venema (2019) yielded slightly larger values with 0.82 and 0.58 K cty$^{-1}$, respectively.

Trends are certainly less sensitive to missing data compared with, for example, extremes. However, the effect of missing data will be included in the trend uncertainty, slightly increasing the uncertainty for the empirical method.
The analytical method, however, is based on lag differences. Data gaps make the correct estimation of the parameters more difficult but will not lead to systematically larger trend uncertainties. Consequently, missing data can make the estimates of the empirical method somewhat higher than those of the analytical estimate. The effect of missing data seems to be limited, as the analytical estimates are even slightly larger than the empirically obtained trend uncertainties.

The random trend error is a large part of the station trend error budget but for some regions considerable network-wide trend biases have been found (Begert et al., 2005; Brunetti et al., 2006; Böhmer et al., 2010; Menne et al., 2010), which are smaller but comparable in size and should thus be part of a complete trend uncertainty budget.

The large-scale bias introduced by the breaks is difficult to determine directly from the data. In single-station series, this component cannot be separated from the true climate trend; and in difference time series, actually used to eliminate the climate signal, this component is lost, too. Thus, we are restrained to indirect methods that compare the trends before and after applying a homogenization algorithm to the data. How accurately these are able to estimate the full trend bias should be tested in the future.

We can compare our estimates of the trend errors with those in the inhomogeneities of the HOME benchmark based on a community consensus of typical European inhomogeneities, which were 1.2°C cyt−1 (Venema et al., 2012). However, the breaks in the HOME benchmark were √2 times too large; taking this into account, the station trend errors would have been 0.85 K cyt−1. In the present study, we found errors between 0.5 and 0.82 K cyt−1, depending on country and method. However, these comprise only random errors and no biases so that they may underestimate the full uncertainty. Taking this into account, these values seem to match reasonably.

Random errors between 0.5 and 0.82 K cyt−1, as we found them for individual stations, will cause random errors also in the network mean trend. Assuming 10 independent stations in a network, the error will be reduced by a factor of √10 to 0.16 to 0.26 K cyt−1. This is still not a bias but solely the consequence of the underlying random errors. Such an uncertainty is smaller than the observed climate trends but not negligible. Furthermore, the error estimate is a lower threshold for the uncertainty of network mean trends. Any dependency between the break signals of the stations will further increase it. The same is true for possibly occurring trend biases, which are not considered in this study.

The above estimate is valid for the situation before homogenization. However, Lindau and Venema (2018b) showed for realistic scenarios (but assuming a perfect detection of the breaks) that the correction step is responsible for the production of dependent solutions for the stations. Consequently, similar trend errors have to be expected also in homogenized data.

Even ignoring network-wide trend biases, inhomogeneities can be very complicated: the break frequency does not have to be constant nor do the break positions need to be independent in time or between stations. RD and BM are idealizations. Clearly BM must be an approximation as its variance grows unbounded in time. Thus, that the analytical and the empirical estimates found about the same station trend uncertainties gives some confidence that the description of inhomogeneities as a superposition of a RD and a BM signal is sufficiently realistic and that complications are at least not too important for trend errors due to inhomogeneities.

ORCID

Ralf Lindau https://orcid.org/0000-0003-4963-4402

REFERENCES

Auer, I., Böhmer, R., Jurkovic, A., Lipa, W., Orlik, A., Potzmann, R., Schöner, W., Ungersböck, M., Matulla, C., Brifka, K., Jones, P., Efthymiadis, D., Brunetti, M., Nanni, T., Maugeri, M., Mercalli, L., Mestre, O., Moisselin, J.M., Begert, M., Müller-Westermeier, G., Kveton, V., Bochnicek, O., Statny, P., Lapin, M., Szalai, S., Szentimrey, T., Cegnar, T., Dolinar, M., Gajic-Capka, M., Zaninovic, K., Majstorovic, Z., and Nieplova, E. (2007) HISTALP—historical instrumental climatological surface rime series of the greater Alpine region. International Journal of Climatology, 27, 17–46. https://doi.org/10.1002/joc.1377.

Begert, M., Schlegel, T. and Kirchhofer, W. (2005) Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000. International Journal of Climatology, 25, 65–80. https://doi.org/10.1002/joc.1118.

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

Brunetti, M., Maugeri, M., Monti, F. and Nanni, T. (2006) Temperature and precipitation variability in Italy in the last two centuries from homogenized instrumental time series. International Journal of Climatology, 26, 345–381. https://doi.org/10.1002/joc.1251.

Causinus, H. and Mestre, O. (2004) Detection and correction of artificial shifts in climate series. Applied Statistics, 53(3), 405–425. https://doi.org/10.1111/j.1467-9876.2004.00535.x.

Conrad, V. and Pollak, C. (1950) Methods in Climatology. Cambridge, MA: Harvard University Press, p. 459.

Della-Marta, P.M., Collins, D. and Braganza, K. (2004) Updating Australia’s high quality annual temperature dataset. Australian Meteorological Magazine, 53, 277–292.

Dunn, R.J., Willett, K.M., Morice, C.P. and Parker, D.E. (2014) Pairwise homogeneity assessment of HadISD. Climate of the Past, 10(4), 1501–1522.

Hunziker, S., Brönnimann, S., Calle, J., Moreno, I., Andrade, M., Ticona, L., Huerta, A. and Lavado-Casimiro, W. (2018) Effects of undetected data quality issues on climatological analyses. Climate of the Past, 14, 1–20. https://doi.org/10.5194/cp-14-1-2018.

Lindau, R. and Venema, V. (2019) A new method to study inhomogeneities in climate records: Brownian motion or random deviations?
International Journal of Climatology, 39, 4769–4783. https://doi.org/10.1002/joc.6105.

Lindau, R. and Venema, V.K.C. (2016) The uncertainty of break positions detected by homogenization algorithms in climate records. International Journal of Climatology, 36(2), 576–589. https://doi.org/10.1002/joc.4366.

Lindau, R. and Venema, V.K.C. (2018a) The joint influence of break and noise variance on the break detection capability in time series homogenization. Advances in Statistical Climatology, Meteorology and Oceanography, 4, 1–18. https://doi.org/10.5194/ascmo-4-1-2018.

Lindau, R. and Venema, V.K.C. (2018b) On the reduction of trend errors by the ANOVA joint correction scheme used in homogenization of climate station records. International Journal of Climatology, 38, 5525–5527. https://doi.org/10.1002/joc.5728.

Menne, M.J., C.N. Williams Jr. and M.A. Palecki (2010) On the reliability of the U.S. surface temperature record. Journal of Geophysical Research, 115: D11108. doi:https://doi.org/10.1029/2009JD013094

Menne, M.J. and Williams, C.N., Jr. (2005) Detection of undocumented changepoints using multiple test statistics and composite reference series. Journal of Climate, 18, 4271–4286. doi:https://doi.org/10.1175/JCLI3524.1.

Menne, M.J. and Williams, C.N., Jr. (2009) Homogenization of temperature series via pairwise comparisons. Journal of Climate, 22(7), 1700–1717. https://doi.org/10.1175/2008JCLI2263.1.

Menne, M.J., Williams, C.N., Jr. and Vose, R.S. (2009) The U.S. historical climatology network monthly temperature data, version 2. Bulletin of the American Meteorological Society, 90(7), 993–1007. https://doi.org/10.1175/2008BAMS2613.1.

Rennie, J., Lawrimore, J., Gleason, B., Thorne, P., Morice, C., Menne, M., Williams, C., Gambi de Almeida, W., Christy, J., Flannery, M., Ishihara, M., Kamiguchi, K., Klein Tank, A., Mhanda, A., Lister, D., Razuvaev, V., Renom, M., Rusticucci, M., Trewin, B., Worley, S., Venema, V., Angel, W., Brunet, M., Dattore, B., Diamond, H., Lazzara, M., Le Blancq, F., Luterbacher, J., Mächel, M., Revadekar, J., Vose, R. and Yin, X. (2014) The International Surface Temperature Initiative: Global land surface databank: monthly temperature data version 1 release 3. https://doi.org/10.5194/cp-8-89-2012.

Trewin, B. (2013) A daily homogenized temperature data set for Australia. International Journal of Climatology, 33(6), 1510–1529. https://doi.org/10.1002/joc.3530.

Venema, V., Mestre, O., Aguilar, E., Auer, I., Guijarro, J.A., Domonkos, P., Vertacnik, G., Szentimrey, T., Stepanek, P., Zahradnicek, P., Viare, J., Müller-Westermeier, G., Lakatos, M., Williams, C.N., Menne, M.J., Lindau, R., Rasol, D., Rustemeier, E., Kolokythas, K., Marinova, T., Andresen, L., Acquaotta, F., Fratianni, S., Cheval, S., Klanac, M., Brunetti, M., Gruber, C., Prohom Duran, M., Likso, T., Esteban, P. and Brandsma, T. (2012) Benchmarking homogenization algorithms for monthly data. Climate of the Past, 8, 89–115. https://doi.org/10.5194/cp-8-89-2012.

Yosef, Y., Aguilar, E., and Alpert, P. (2018) Detecting and adjusting artificial biases of long-term temperature records in Israel. International Journal of Climatology, 38(8), 3273–3289.

How to cite this article: Lindau R, Venema V. Random trend errors in climate station data due to inhomogeneities. Int J Climatol. 2020;40:2393–2402. https://doi.org/10.1002/joc.6340

APPENDIX: VARIANCE OF THE TRENDS OF BROWNIAN MOTION STEP FUNCTIONS

We consider a set of Brownian motion step functions $y_i$ of length $n$ containing a number of breakpoints at random timings $\tau_i$, $i = 1$ to $k$. Breaks occur stochastically with a probability $p$ so that the number of breaks $k$ varies between the different time series. The breaks divide the time series into $k + 1$ homogeneous subperiods (HSPs) of length $l_i$. Within a HSP, the step function remains at a constant level $z_i$. The levels $z_i$ form a Brownian motion in time with:

\[ z_i = \sum_{j=0}^{k} x_i, x_i \sim N(0, \sigma_p^2), \text{for } 1 \leq i \leq k, x_0 = 0. \]  

(A1)

The increments $x_i$ are random, independent, and normal distributed with zero mean.

The correlation $r$ between data $y$ and time steps $t$ defines the trend: The linear trend $T$ is given by the slope of the regression line where the data are treated as independent and the time as dependent parameter:

\[ T = r \frac{\sigma_y}{\sigma_t} = \frac{\text{cov}(y,t)}{\text{var}(t)} = \frac{\sum t y_i - \frac{1}{n} \sum t \sum y_i}{\sum t^2 - \frac{1}{n} \sum t^2}, \]  

(A2)

where $\sigma_y$ and $\sigma_t$ denote the standard deviations of data and time, respectively. The covariance between data and time is denoted by $\text{Cov}(y,t)$ and the variance of the time by $\text{Var}(t)$.

The sum over the time steps and sum over the squared time steps, which occur in Equation (A2), are given by:

\[ \sum_{t=1}^{n} t = \frac{1}{2} (n+1) n, \]  

(A3)

\[ \sum_{t=1}^{n} t^2 = \frac{1}{6} (2n+1)(n+1) n. \]  

(A4)

The sum over the data $y_i$ is:
In a strict sense, the breakpoints $\tau_i$ are only defined for $i > 0$. However, as the beginning of the first HSP is identical to the beginning of the entire time series, a reasonable definition is:

$$\tau_0 = 1.$$ (A6)

The sum over the product of data and time is:

$$\sum_{i=1}^{n} t y_i = \sum_{i=1}^{n} x_0 t + \sum_{i=1}^{n} t x_0 + \cdots + \sum_{i=1}^{n} (\tau_1 - 1) x_0$$

$$+ \tau_1 (x_0 + x_1) + (\tau_1 + 1) (x_0 + x_1) + \cdots$$

$$+ (\tau_2 - 1) (x_0 + x_1) + \cdots.$$ (A7)

In Equation (A7), the random number $x_0$ occurs 1+2+...+$n$ times. The random number $x_1$ appears only from the year $\tau_1$ on (and analogously for the subsequent HSPs). Thus, it follows:

$$\sum_{i=1}^{n} t y_i = \sum_{i=1}^{n} x_0 \sum_{i=1}^{n} t + \sum_{i=1}^{n} t x_0 + \cdots + \sum_{i=1}^{n} (\tau_1 - 1) x_0$$

$$+ \sum_{i=1}^{n} \sum_{f=\tau_1}^{\tau_i} x f \sum_{i=1}^{n} x_i$$ (A8)

The sum over subsequent natural numbers is given in Equation (A3). However, to solve the last summation in Equation (A8), the sum until values smaller than the lower boundary $\tau_1$ has to be subtracted:

$$\sum_{i=1}^{n} t y_i = \sum_{i=1}^{n} x_i \left( \sum_{i=1}^{n} t - \sum_{i=1}^{\tau_i-1} t \right) = \sum_{i=1}^{n} x_i \left( \frac{(n+1)n}{2} - \tau_i (\tau_i - 1) \right).$$ (A9)

The numerator of the last expression in Equation (A2) is equal to the covariance multiplied by $n - 1$. We denote this expression by $C$ and insert Equations (A9), (A3), and (A5):

$$C = \sum_{i=0}^{k} \left( (n+1)n - \tau_i (\tau_i - 1) - (n+1)(n - \tau_i + 1) x_i \right),$$ (A10)

which can be reduced to:

$$C = \frac{1}{2} \sum_{i=0}^{k} \left( n (\tau_i - 1) - (\tau_i - 1)^2 \right) x_i = \frac{1}{2} \sum_{i=1}^{k} \left( n (\tau_i - 1) - (\tau_i - 1)^2 \right) x_i,$$ (A11)

The second equal sign in Equation (A11) is valid, because $\tau_0 = 1$ (Equation (6)) so that the sum vanishes for $i = 0$ and the lower boundary can be set to 1.

The denominator of the last expression in Equation (A2) is equal to the variance multiplied by $n - 1$. We denote this expression by $V$, insert Equation (A3) and (A4) and obtain:

$$V = \frac{1}{6} (2n+1)(n+1)n - \frac{1}{n} \left( \frac{(n+1)n}{2} \right)^2.$$ (A12)

$V$ is a constant for all time series and depends only on the length $n$. It can be rewritten to:

$$V = \frac{n}{12} (n^2 - 1).$$ (A13)

Now the trend $T$ for each time series could be obtained by dividing $C$ by $V$. However, we are interested in the variance of the trend of many of such time series. As $V$ is a constant depending only on $n$, it is sufficient to calculate the variance of $C$ and divide just by $V^2$. Furthermore, the expected value of $C$ is zero so that the variance is equal to the mean squared $C$, denoted by square brackets.

$$\text{var}(T) = \frac{\text{var}(C)}{V^2} = \frac{C^2}{V^2}.$$ (A14)

Equation (A11) shows that $C$ is obtained by multiplying a time-depending coefficient $a_i$ by a random number $x_i$, summing up this product over all HSPs $i$ and dividing by 2. Thus, denoting:

$$a_i = n (\tau_i - 1) - (\tau_i - 1)^2,$$ (A15)

we can rewrite Equation (A11) in a concise form:

$$C = \frac{1}{2} \sum_{i=1}^{k} a_i x_i.$$ (A16)

For the mean squared $C$, it follows:

$$C^2 = \left[ \frac{1}{2} \sum_{i=1}^{k} a_i x_i \right]^2 = \frac{1}{4} \sum_{i=1}^{k} \left[ a_i^2 \right] \left[ x_i^2 \right]$$

$$= \frac{\sigma_p^2}{4} \sum_{i=1}^{k} \left[ a_i^2 \right] = \frac{\sigma_p^2}{4} \sum_{i=1}^{k} a_i^2.$$ (A17)

Equation (A17) is valid for two reasons. First, all mixed terms that arise when a sum is squared becomes zero because $x_i$ is a random variable. The second is that the mean squared $x_i$ is always equal to $\sigma_p^2$ independently from the HSP. In the last transformation, the order of summation and averaging is again reinverted.

Thus, the average of the sum over $a_i^2$ is the crucial term. We reinsert Equation (A15):
\[
\left[ \sum_{i=1}^{k} a_i^2 \right] = \left[ \frac{1}{2} \left( 6 - \frac{1}{2} \right) \right] \left( n(\tau_i - 1) - (\tau_i - 1)^2 \right)^2 .
\]  
(A18)

We expand the square expression and separate the sums. As \( n \) is a constant, it can be written in front of the sums:

\[
\left[ \sum_{i=1}^{k} a_i^2 \right] = \left[ \sum_{i=1}^{k} (\tau_i - 1)^4 \right] - 2n \left[ \sum_{i=1}^{k} (\tau_i - 1)^3 \right] + n^2 \left[ \sum_{i=1}^{k} (\tau_i - 1)^2 \right] .
\]  
(A19)

In Equation (A19), the year before the break happens \((\tau_i - 1)\) is the crucial number. We consider the first summation on the right-hand side in more detail. For a given year \( \tau_i \), the expression \((\tau_i - 1)^4\) is trivially a fixed value. The mean sum of these forth power break dates can be alternatively obtained by multiplying this fixed value with its mean occurring frequency \( p \) and summing up the resulting product over all years. As \( p \) is a constant, it can be written in front of the summation sign:

\[
\left[ \sum_{i=1}^{k} (\tau_i - 1)^4 \right] = p \sum_{i=1}^{n} (t - 1)^4 .
\]  
(A20)

Analogous considerations can be made for the other powers of \( \tau_i - 1 \). It follows from Equation (A19):

\[
\left[ \sum_{i=1}^{k} a_i^2 \right] = p \sum_{i=1}^{n} (t - 1)^4 - 2pn \sum_{i=1}^{n} (t - 1)^3 + pn^2 \sum_{i=1}^{n} (t - 1)^2 .
\]  
(A21)

The sum of the second power of \( t \) is already given in Equation (A4) and also the results for higher powers are commonly known. However, here we need the sums of the powers of \( t - 1 \), so that some modifications are necessary. With \( s = t - 1 \), the lower boundary is shifted from 1 to 0 and the upper boundary from \( n \) to \( n - 1 \). The inclusion of zero has no consequence for the lower boundary because all powers of \( s \) are zero for \( s = 0 \); only the change of upper limit has to be taken into account:

\[
\sum_{i=1}^{n} (t - 1)^2 = \sum_{s=0}^{n-1} s^2 = \frac{1}{6} (2(n - 1) + 1)n(n - 1) = \frac{1}{6} (2n - 1)n(n - 1) ,
\]  
(A22a)

\[
\sum_{i=1}^{n} (t - 1)^3 = \sum_{s=0}^{n-1} s^3 = \left( \frac{1}{2} n(n - 1) \right)^2 .
\]  
(A22b)

\[
\sum_{i=1}^{n} (t - 1)^4 = \sum_{s=0}^{n-1} s^4 = \frac{1}{30} (2(n - 1) + 1)n(n - 1) \left( 3(n - 1)^2 + 3(n - 1) - 1 \right),
\]  
(A22c)

Then, Equations (A22a)–(A22c) are inserted into Equation (A21):

\[
\frac{1}{p} \left[ \sum_{i=1}^{k} a_i^2 \right] = \frac{1}{30} (2n - 1)n(n - 1) \left( 3n^2 - 3n - 1 \right) - 2n \left( \frac{1}{2} n(n - 1) \right)^2
\]  
(A23)

\[
+ \frac{6}{(2n - 1)n^3} .
\]

Some algebraic reformulations:

\[
\frac{30}{n(n - 1)p} \sum_{i=1}^{k} a_i^2 = (2n - 1) \left( (3n^2 - 3n - 1) + 5n^2 \right) - 15n^3(n - 1) ,
\]  
(A24)

lead to:

\[
\frac{30}{n(n - 1)p} \sum_{i=1}^{k} a_i^2 = n^3 + n^2 + n + 1 = (n^2 + 1)(n + 1) .
\]  
(A25)

And it follows:

\[
\sum_{i=1}^{k} a_i^2 = \frac{p}{30} (n^2 + 1) (n^2 - 1)n .
\]  
(A26)

We insert Equation (A26) into Equation (A17) and get:

\[
[C^2] = \frac{p \sigma_p^2}{120} (n^2 + 1)(n^2 - 1)n .
\]  
(A27)

Finally we insert Equation (A27) and Equation (A13) into Equation (A14):

\[
\frac{\text{var}(T)}{V^2} = \frac{6p \sigma_p^2}{5} \frac{n^2 + 1}{n^2 - 1} .
\]  
(A28)

The typical length of a yearly climate series is \( n = 100 \) so that the last fraction in Equation (A28) can be neglected:

\[
\frac{\text{var}(T)}{V^2} = \frac{6p \sigma_p^2}{5} .
\]  
(A29)