**RMetS**

Royal Meteorological Society

# Homogenisation of temperature and precipitation time series with ACMANT3: method description and efficiency tests

P. Domonkos[a]* and J. Coll[b]

[a] *Centre for Climate Change, University of Rovira i Virgili, Tortosa, Spain*
[b] *Irish Climate Analysis and Research UnitS (ICARUS), Department of Geography, Maynooth University, Co Kildare, Ireland*

**ABSTRACT:** The development of Adapted Caussinus–Mestre Algorithm for homogenising Networks of Temperature series (ACMANT), one of the most successful homogenisation methods tested by the European project COST ES0601 (HOME) has been continued. The third generation of the software package 'ACMANT3' contains six programmes for homogenising temperature values or precipitation totals. These incorporate two models of the annual cycle of temperature biases and homogenisation either on a monthly or daily time scale. All ACMANT3 programmes are fully automatic and the method is therefore suitable for homogenising large datasets. This paper describes the theoretical background of ACMANT and the recent developments, which extend the capabilities, and hence, the application of the method. The most important novelties in ACMANT3 are: the ensemble pre-homogenisation with the exclusion of one potential reference composite in each ensemble member; the use of ordinary kriging for weighting reference composites; the assessment of seasonal cycle of temperature biases in case of irregular-shaped seasonal cycles. ACMANT3 also allows for homogenisation on the daily scale including for break timing assessment, gap filling and ANOVA application on the daily time scale. Examples of efficiency tests of monthly temperature homogenisation using artificially developed but realistic test datasets are presented. ACMANT3 can be characterized by improved efficiency in comparison with earlier ACMANT versions, high missing data tolerance and improved user friendliness. Discussion concerning when the use of an automatic homogenisation method is recommended is included, and some caveats in relation to how and when ACMANT3 should be applied are provided.

KEY WORDS    time series; homogenisation; ACMANT; spatial interpolation; surface air temperature; precipitation

*Received 12 February 2016; Revised 31 May 2016; Accepted 8 June 2016*

## 1. Introduction

For the analysis of climate change and climate variability the accuracy of observational data is of key importance (Williams *et al.,* 2012; Acquaotta and Fratianni, 2014). Although national meteorological services led by World Meteorological Organisation (WMO) initiatives foster the production of high quality and comparable climatic data temporally and spatially, technical changes in the measurement setup or observational practices often influence the usability of climate data records. Observational data can only be considered temporally homogeneous (hereafter: homogeneous) if temporal variations are exclusively influenced by weather and climate. In practice, several factors corrupt the homogeneity of climatic time series, these include: station relocations, changes of instrumentation, instrument position, site changes around the instrument, changes of the timing of reading instruments, etc. (Aguilar *et al.,* 2003; Menne *et al.,* 2009; Acquaotta *et al.,* 2016). A general observation is that long observational climatic records are seldom homogeneous, and that the quality of climatic records may also be affected by occasional

observational and data transfer errors. Certain data quality problems can be eliminated by general quality control (Durre *et al.,* 2010; Menne *et al.,* 2012), or with the analysis of the documents of the history of observations (so-called metadata, Bergstrom and Moberg, 2002; Prohom *et al.,* 2016). However, the statistical homogenisation of data provides additional quality control and allows for improved temporal and spatial comparisons between data for scientific purposes (Peterson *et al.,* 1998; Beaulieu *et al.,* 2008; Venema *et al.,* 2012). When observational networks are sufficiently dense, relative homogenisation (i.e. homogenisation methods including spatial comparisons of time series) can help to remove even relatively small non-climatic biases from the data. Therefore, the statistical methodology underpinning time series homogenisation is a widely studied topic of climatology (e.g. Series of Data Quality Control and Time Series Homogenisation, World Meteorological Organisation, 1996–2014).

Both the most common and frequent form of inhomogeneity in a climate time series is the sudden shift of the means, commonly referred to as a break or change-point. A set of breaks can be searched and corrected one-by-one in hierarchic structures (e.g. Alexandersson and Moberg, 1997), or jointly with appropriate mathematical tools. When time series include multiple breaks, joint treatments have theoretical advantages over hierarchic techniques

---

* Correspondence to: P. Domonkos, Centre for Climate Change, University of Rovira i Virgili, Av. Remolins 13-15, 43500 Tortosa, Spain.
E-mail: dpeterfree@gmail.com

(Szentimrey *et al.,* 2014; Lindau and Venema, 2016), as in hierarchic techniques early phase errors are delivered to the later steps of the homogenisation process. For purposes of this study 'multiple break method' means a method with joint detection of inhomogeneities, and one which incorporates the joint calculation of correction terms for adjusting inhomogeneities. As observed temperature time series include five to seven breaks per 100 years on average (Menne *et al.,* 2009; Venema *et al.,* 2012), or even more due to hidden short-term biases (Domonkos, 2011a; Rienzner and Gandolfi, 2011) multiple break methods are of key importance in providing high level solutions for homogenisation tasks particularly in relation to the homogenisation of temperature. Efficiency tests prove that multiple break methods generally outperform other homogenisation methods (Domonkos, 2011a,; Domonkos *et al.,* 2011; Venema *et al.,* 2012). Although some inhomogeneities result in gradually increasing biases instead of abrupt shifts of the means (e.g. urbanisation), this effect has little impact on the rank order of method efficiencies (Domonkos, 2011a).

The organisation of this paper is as follows: The development of multiple break methods and particularly the development of Adapted Caussinus–Mestre Algorithm for homogenising Networks of Temperature (ACMANT) is described in the next section; in Section 3, the novel features of ACMANT3 compared with earlier ACMANT versions are presented; some efficiency results are presented in Section 4; and the paper ends with a discussion and some recommendations in Section 5.

## 2. Development of multiple break methods and ACMANT

Although statistical break detections and corrections have been studied and applied for at least 90 years (Conrad, 1925), the theory and development of multiple break homogenisation appeared only in the 1990s coincident with the more widespread use of personal computers. At that time two approaches to multiple break homogenisation were established, namely Multiple Analysis of Series for Homogenisation (MASH, Szentimrey, 1996, 1999) and PRODIGE (Caussinus and Mestre, 1996, 2004). These two methods differ markedly from each other: MASH uses multiple reference series, selects the set of breaks with hypothesis testing, derives adjustments terms from the confidence intervals belonging to the hypothesis test results, and the MASH approach to the final solution is iterative. By contrast, PRODIGE uses pairwise comparisons, optimal step function fitting for break detection and the minimisation of residual variance (ANOVA) for the adjustments of inhomogeneities in a procedure without iteration. Both MASH and PRODIGE were among the most successful methods tested by the European project COST ES0601 (known as 'HOME', 2007–2011). During HOME, two new multiple break methods were created based on PRODIGE: one is the fully automatic ACMANT (Adapted

Caussinus–Mestre Algorithm for homogenising Networks of Temperature series, Domonkos, 2011b) and the other is Homogenization software in R (HOMER, Mestre *et al.,* 2013), the interactive homogenisation method officially recommended by HOME. Both HOMER and ACMANT provide additional functionality relative to the parent method PRODIGE. Therefore recently, HOMER and ACMANT have been applied more frequently than PRODIGE.

After the termination of HOME, the development of ACMANT has continued, and the second generation of ACMANT (ACMANT2) already incorporated methods for the precipitation homogenisation and for the treatment of daily data through downscaling the monthly homogenisation results to daily scale (Domonkos, 2014, 2015a). Most recently the development of ACMANT3 has improved the efficiency and user friendliness further, and some of these improvements are detailed below.

ACMANT3 is a complex software package incorporating six programmes, these are: temperature homogenisation with a sinusoid annual cycle of biases; temperature homogenisation with an irregular annual cycle of biases; precipitation homogenisation. Each of the preceding three has monthly and daily homogenisation versions (http://www.c3.urv.cat/data.html); and in total the six programmes incorporate 174 sub-routines. The software package also includes auxiliary files to support network construction. However, despite its complicated structure, ACMANT provides the fastest method implementation among all the available automatic homogenisation methods.

As both HOMER and ACMANT have been developed from PRODIGE, the two new multiple break methods have several similarities. Table 1 summarizes the main similarities and differences of the two methods. Note that although HOMER is only for monthly homogenisation, the joint use of HOMER and Spline Daily Homogenization (SPLIDHOM, Mestre *et al.,* 2011) can be applied for daily data homogenisation (www.homogenisation.org), and in Table 1 it is considered the daily homogenisation version of HOMER.

## 3. Methodological novelties of ACMANT3

The full scientific description of ACMANT2 has been published (Domonkos, 2014, 2015a), and therefore, only the new features of ACMANT3 in comparison with the earlier ACMANT versions are presented here. Appendix A details the mathematical formulations of some steps of the homogenisation connected to the content of this section. Note that some less important details are not shown due to the complexity of the methodology.

To aid interpretation of the present description of the methodology, it should be recalled that in ACMANT, the candidate series for homogenisation is compared with the weighted average of other available series in the same climatic area. This weighted average series is referred to as composite reference series and the contributor series as

Table 1. Similarities and differences between HOMER and ACMANT.

|  | HOMER | ACMANT3 |
|---|---|---|
| Temperature homogenisation | X | X |
| Precipitation homogenisation | X | X |
| Monthly homogenisation | X | X |
| Daily homogenisation | X | X |
| Pairwise comparison | X |  |
| Composite reference series |  | X |
| Optimal step function fitting | X | X |
| Derivatives of the C–L criterion | X | X |
| Network-wide joint segmentation | X |  |
| Bivariate detection for temperature (when applicable) | X | X |
| Bivariate detection for precipitation (when applicable) |  | X |
| Varied use of time resolution in break detection |  | X |
| Ensemble pre-homogenisation |  | X |
| Correction with ANOVA model | X | X |
| Iterations | X |  |
| Monthly precision of breaks | X | X |
| Daily precision of large magnitude breaks |  | X |
| Adjusting relative to the last homogeneous section | X | X |
| Quantile dependent daily correction terms | X |  |
| High missing data tolerance | X[a] | X[a] |
| Completion of missing data | X | X |
| Completion of data before break detection |  | X |
| Filtering of spatial outliers | X | X |
| Metadata use supported | X |  |
| Graphical results | X |  |
| Automatic execution |  | X |

[a]The missing data tolerance of ACMANT is higher than that of HOMER (results of authors' experiments, not shown).

reference composites, respectively. The difference (ratio) of the candidate and reference series is referred to as relative time series in the homogenisation of temperature (precipitation) series.

### 3.1. Ensemble pre-homogenisation

The aim of doing pre-homogenisation prior to the main homogenisation is to remove relatively large biases before the main homogenisation. This excludes the possibility of potentially large biases in the reference composites during the main homogenisation affecting the accuracy of the final homogenisation results. The concept for ACMANT1 and ACMANT2 was that during the pre-homogenisation of the reference series of the later candidate series, the later candidate series was excluded from that pre-homogenisation. The pre-homogenisation

routines employed by ACMANT3 are completely different and their core is the search for and application of ensemble minimums as adjustment terms. The underpinning concept is that the application of false adjustments can be excluded with high certainty by the use of ensemble minimums.

In the first phase, an ensemble of pre-homogenisation is produced, and this always excludes one potential reference component from the homogenization. The main steps of the pre-homogenisation are (1) creating relative time series; (2) break detection on an annual scale and; (3) calculation of adjustment terms with ANOVA (Caussinus and Mestre, 2004; Domonkos, 2014) are performed for each ensemble member. The adjustment terms indicated by the ensemble members are stored but not applied in this phase. In the second phase, the minimum absolute value of the stored adjustment terms is calculated for each time series and each year, and then, it is applied for obtaining the pre-homogenisation result. If the signs of the adjustment terms of ensemble members are mixed for a particular year of a particular time series, then the result adjustment term is zero. See also Appendix A1.4.

If univariate detection is applied, then the monthly adjustment terms of pre-homogenisation are constant within a particular year, while the shape of the annual cycle of adjustment terms is predefined when bivariate detection is applied (Domonkos, 2014). The pre-homogenisation is performed twice in any programme of ACMANT3. In the first execution the $c_p$ coefficient of the modified Caussinus–Lyazrhi (C–L) criterion (Appendix A1.3, formulas (8) and (9)) is elevated with 40%. The purpose of this distinction in parameter $c_p$ according to the stage of ACMANT3 procedure is to focus on the elimination of large inhomogeneities in the first pre-homogenisation.

### 3.2. Ordinary kriging for determining the weights of reference components

Ordinary kriging is a widely applied tool for the production of optimally interpolated values of a meteorological variable to a given location. Theoretically, ordinary kriging provides the optimal weights of reference series composites (Szentimrey, 2010), as the purpose of building composite reference series is to have another series beyond the candidate series (i.e. the reference series) whose climatic variability is the same as that of the candidate series. However, in practice, the unavoidable inaccuracy of the large number of parameter estimates incorporated in ordinary kriging has the potential to reduce the efficiency. The more traditional way is applying the squared spatial correlations of increment time series as weights of the reference composites (Peterson and Easterling, 1994).

If the number of reference composites is $N$, then the number of estimated parameters is $N$ in the traditional method, while for ordinary kriging it is $0.5 \times N \times (N-1)$ as ordinary kriging uses the entire covariance matrix of the candidate series and reference composites. According to tests (not shown) ordinary kriging performs better than the traditional weighting of reference composites if it is applied with some restrictions: (1) Ordinary kriging is

applied only when $N \geq 6$; (2) when an estimated weight is negative, the applied weight is 0, as a negative weight for a reference composite has no physical interpretation; (3) when an estimated weight would be higher than 0.4, the weights of the other reference composites are slightly elevated in order to reduce the overly strong influence of the most highly correlated reference composite. It is applied to reduce the likelihood of importing any potential inhomogeneity from the most closely correlated reference composite.

### 3.3. Irregular annual cycle of biases for temperature inhomogeneities

In the tropics and monsoon regions, the annual cycle of irradiation markedly differs from the semi-sinusoid cycle of the mid- and high-latitude regions. Therefore, the annual cycle of temperature biases is not sinusoid everywhere. In addition, biases of daily temperature minima are affected more by the frequency of clear sky and calm weather conditions than by the annual cycle of irradiation. In ACMANT2, the model annual cycle of biases was constant when the sinusoid model was not applicable. ACMANT3 provides season-dependent adjustments for these cases, in spite of the generally lower signal to noise ratio in the estimates for seasonal than annual characteristics. The assessment of the seasonal cycle of biases includes three main steps: (1) Optimal step function fitting (Appendix A1.3) is applied to the annual series of monthly temperatures for each calendar month, with coarser time resolution than in the default application, this to allow for the lower signal to noise ratio for monthly compared to annual values. Here, the minimum distance between two adjacent breaks is 5 years. (2) ANOVA is applied to the annual series of monthly values. In this step, the same ensemble procedure is included for each month as in the pre-homogenisation (Appendix A1.4). (3) Finally, monthly estimates of adjustment terms derived from the ensemble calculations are smoothed (Appendix A1.6).

### 3.4. Daily precision of large-size breaks

A step function with exactly one break is fitted to daily data of relative time series in a section, which includes both the preceding and subsequent 6 months of the pre-estimated timing of the break. Univariate detection is applied, and the break is accepted in a narrower window only, which includes the preceding and subsequent 2 months of the pre-estimated timing. This operation is done only when the pre-estimated break magnitude is not lower than the 75% of the empirical standard deviation of daily data, this also considers either the entire window (1 year) or the narrower window (4 months). When this operation is omitted for a break of relatively small magnitude, the default timing of the last calendar day of the pre-estimated month of the break is retained.

### 3.5. Treatment of daily data when monthly outliers are detected

A monthly outlier or an outlier value of the mean of a few months period may indicate that on daily scale either a platform-shaped inhomogeneity exists, or the daily values are very scattered with significant mean bias. For identifying the former case (which is likely the more frequent), a step function with two steps is fitted to the daily data of the relative time series spanning the period that starts 4 months before the outlier and ends 4 months after the outlier ('wide window'). The two breaks are searched across a narrower window, specified for 1 month before and after the outlier period with a minimum platform length of 10 days, and univariate detection is applied. No platform-shaped inhomogeneity is identified when (1) there are less than 4 months with observed monthly data of the candidate series among the 8 months before and after the narrow window but within the wide window; or (2) the standard deviation of the relative time series values within the candidate platform-shaped inhomogeneity is at least twice as high as for all the values within the wide window. If a platform-shaped bias cannot be identified, then all the daily data of the outlier period are treated as missing data.

### 3.6. ANOVA on daily scale

In the late phase of daily homogenisation, the break times are provided at daily resolution and ANOVA is then applied to the daily resolution data to provide direct daily adjustment terms. The equation system for the practical solution of the minimisation of the residual variance with ANOVA (Domonkos, 2014, 2015a) can be applied on any time scale.

### 3.7. Selective exclusion of years with too few observed data from the homogenisation

ACMANT's default operational mode is to first infill data gaps with interpolated values, and all the routines of the software use continuous time series. However, when the number of synchronously available observed data in the network is very low, the reliability of interpolated values is correspondingly low, and hence, can affect the efficiency of homogenisation. In ACMANT3, years with too few observed values are excluded from most steps of the homogenisation procedure and are only included in the final interpolations for completing time series and for applying final adjustments to correct inhomogeneities.

Typically, years with intact observed data from less than three stations are excluded. A yearly observation is classified as intact when the year has at least nine observed monthly values (at least 9 months with at least 75% complete observed daily values) in the case of monthly (daily) input data. Whereas in the case of precipitation, only one missing daily data value excludes the entire month from contributing positively to the evaluation of the intact or non-intact character of the yearly observation.

### 3.8. Elevated missing data tolerance

Time series must comprise at least ∼10 years observed data, more precisely 114 monthly observations and a series

may have more than a 100-year data gap. (See more requirements for data input in the ACMANT3 Manual, http://www.c3.urv.cat/data.html.)

As relative time series often have different lengths (Domonkos, 2011b, 2014), long data gaps would result in some reference series including virtual composites, i.e. composites without their own observational data. In the present parameterisation scheme of ACMANT3, a reference composite can be included if it has at least 12 observed monthly values falling within the corresponding period of the target relative series. This threshold is low, since experience with ACMANT3 shows that the inclusion of a very few additional observed values compensates well for the uncertainty associated with the inclusion of interpolated values.

### 3.9. Modifications in the interpolation of monthly values for substituting missing data

The interpolation for a missing value of a candidate series is provided by the weighted average of the anomalies in other series of the network ('partner series') applied to the missing data points, and these are then adjusted to the climatic mean of the candidate series (Appendix A1.5). Two changes are incorporated compared to earlier ACMANT versions: (1) Only the homogenized periods of partner series are taken into account other than at the early phase of the homogenisation procedure, or when the interpolation is for a missing value out of the homogenized period of the candidate series; (2) changes to the parameterisation scheme.

### 3.10. Gap filling of daily data before homogenisation

The advantage of gap filling on data at daily scale before homogenisation is that temporally fragmented observed values can be incorporated into the homogenisation. The interpolation of daily values is performed in the same way as that of the monthly values in monthly homogenisation, other than for some changes in the parameterisation scheme applied. This interpolation is repeated several times during a homogenisation procedure, as after the adjustments for biases caused by inhomogeneities, more accurate interpolated values can be provided for infilling data gaps.

### 3.11. Completion of time series

For ACMANT3, even if the input data comprises time series of various lengths, the homogenized output series are (optionally) completed with data for the same time period. These are generally from the earliest year with available observational data in the network through until the latest year with observational data in the network.

## 4. Efficiency tests

For measuring the efficiency of homogenisation methods, artificially generated test datasets are needed, in which the positions, shapes and magnitudes of inhomogeneities

Table 2. Summary statistics for the proportion of the test time series completeness and the extent of missing data. Completeness and exterior missing data are proportional to the entire period examined (100 years), while interior missing data to the periods of observations. All values are in percentage units.

|  | Completeness | Exterior missing data | Interior missing data |
|---|---|---|---|
| Dataset (A) | 89.0 | 10.0 | 1.1 |
| Dataset (B) | 50.1 | 39.2 | 17.7 |

are known. In good test datasets, both the climatic and inhomogeneity characteristics should be realistic and the dataset should be of sufficient size to make confident statistical estimations (Domonkos, 2013). In this section, the efficiencies of the earlier and new ACMANT versions are compared for the homogenisation of monthly temperatures with regular annual cycle of biases using two test datasets. The homogeneous dataset was taken from the HOME benchmark, its large, 200 network sized version is used (Venema *et al.,* 2012). For the exercise here, climatic trends, additional noise and various kinds of inhomogeneities have been added (Appendix A2). The two test datasets differ only in the number and completeness of their time series, and in the spatial correlations between the test series. Each dataset consists of 200 networks of 100-year long time series. In dataset (A), the spatial correlations are around 0.9, and each network includes 10 time series with a low missing data ratio. In dataset (B), the mean spatial correlation is 0.8, each network contains 15 time series, and the missing data ratio is elevated. Table 2 summarizes the missing data characteristics of datasets (A) and (B) where missing data due to the shortage of the period of observations are referred to external missing data, and those within the period of observations internal missing data, respectively.

It can be seen that although in dataset (B) networks include 15 time series, the total number of observed values is lower with 16% in dataset (B) than in dataset (A). Homogenisation efficiency measures are presented here via the comparison of the residual errors with the raw data errors. The examined efficiency measures are: (1) monthly root mean squared error (RMSE), (2) annual RMSE, (3) absolute trend bias of individual time series, (4) absolute network mean trend bias, (5) systematic network mean trend bias. Figures 1 and 2 present the raw data errors and the errors after the ACMANT homogenisation with three ACMANT versions from the earliest to the most recent: ACMANT1 (AC1), ACMANT2 (AC2) and ACMANT3 (AC3).

Dataset (B) cannot be homogenized with AC1. In the other homogenisation results, three test statistics are shown for each method version and the associated efficiency measure: mean error, value of percentile 0.95 and maximal error, other than for the systematic network mean error where only the mean error is presented.
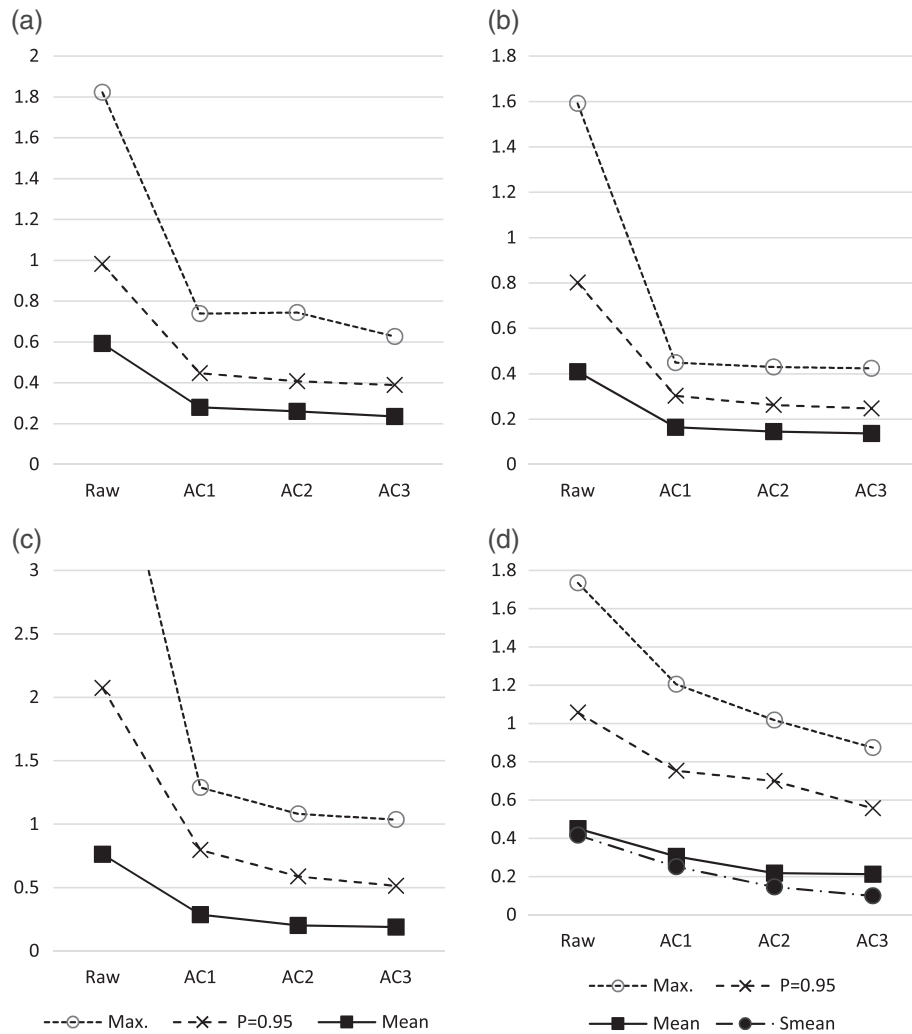
Figure 1. Errors of raw data and those of ACMANT homogenisation products for dataset (A). (a) Monthly RMSE, (b) annual RMSE, (c) trend bias for individual series, (d) network mean trend bias. Smean means systematic trend bias.

The main characteristics of the results are as follows:

1 The statistics are always better for the ACMANT homogenisation results than for the raw data with the exception of maximum monthly RMSE in dataset (B). The largest reduction of raw data errors (∼75%) are associated with the individual trend bias and systematic network mean trend bias in dataset (A). However, it is also worth noting that the improvement is strongly related to the statistical characteristics of the inhomogeneities in the data.

2 Although inhomogeneities are inserted into datasets (A) and (B) according to the same rules, the raw data errors are slightly lower in dataset (B) than in dataset (A). This is because the mean length of periods with observed data is shorter in dataset (B) than in dataset (A).

3 Residual homogenisation errors for dataset (B) are consistently higher than for those of dataset (A), but the differences are generally not very large; e.g. the increment is smaller than 50% for the mean errors (except for the systematic network mean bias) and the 0.95 percentile

values (with the exception of monthly RMSE). The difference between the two datasets is more marked in relation to the maximal residual errors having occurred in the homogenisation of the 200 networks. Considering that in dataset (B) both the amount of observed data and the spatial correlations are lower, and that the internal missing data ratio is higher than in dataset (A), the efficiency of the homogenisation of dataset (B) with ACMANT2 or ACMANT3 is acceptable.

4 The comparison of residual errors between ACMANT versions shows a slight but consistent improvement of efficiency in the transition from older versions to newer versions based on each efficiency measure and the test statistic applied. Relatively large improvement can be seen in the residual trend biases, particularly at the 0.95 percentile and in the systematic network mean trend bias as well.

Figure 3 shows the accuracy of the interpolation for missing monthly temperatures in datasets A and B in function of the number of the partner series used. If no partner series could be used then the empirical climatic
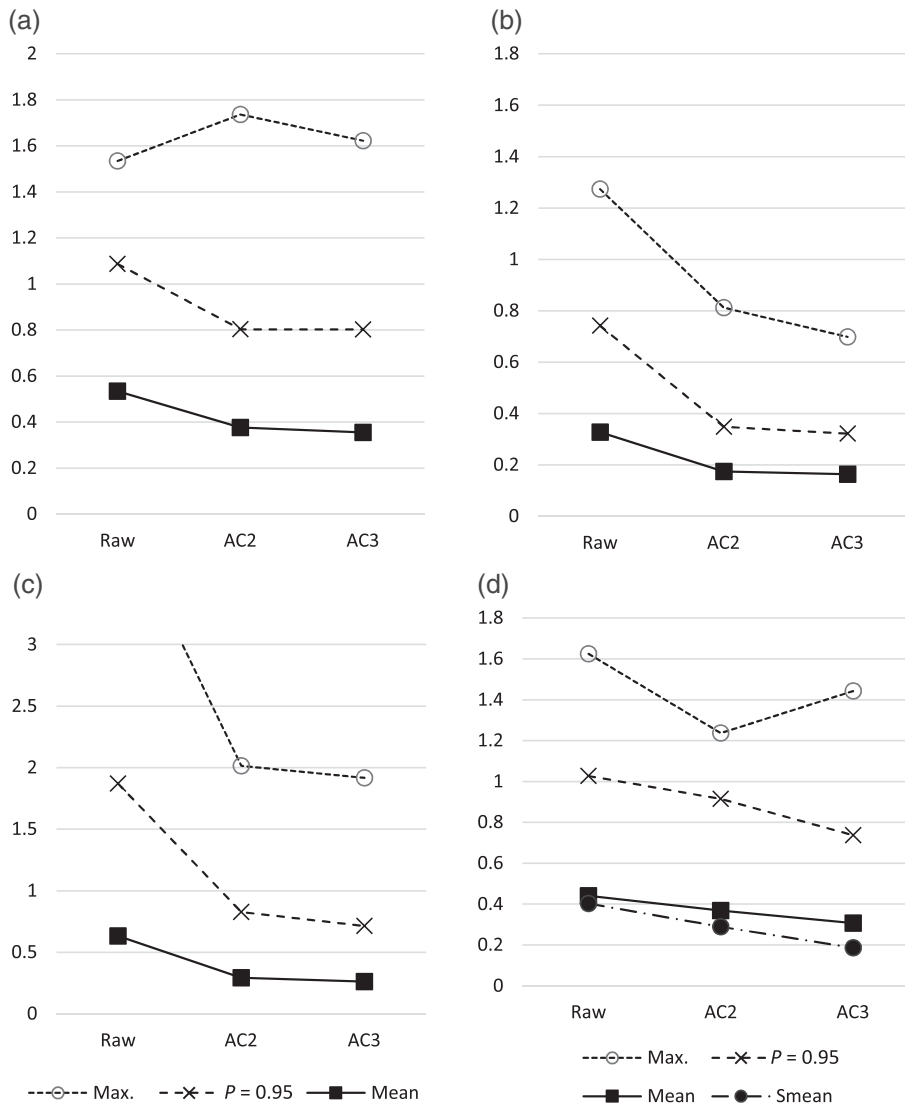
(a)

(b)

(c)

(d)

---○--- Max.    --✕-- $P = 0.95$    —■— Mean

---○--- Max.    --✕-- $P = 0.95$

—■— Mean    —●— Smean
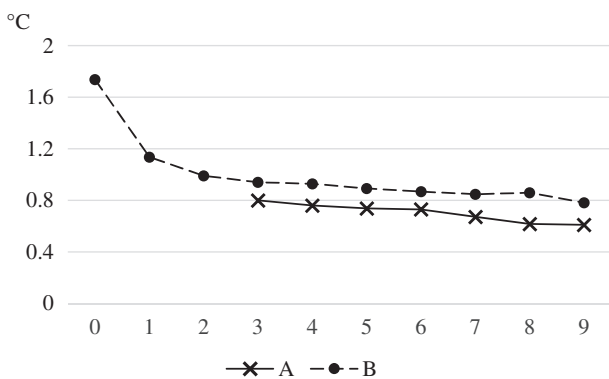
Figure 2. The same as Figure 1, but for dataset (B).

°C

Figure 3. RMSE of interpolated monthly values of the homogenized datasets A and B. In the horizontal axis, the number of partner series used is shown.

mean substitutes the missing data, thus, the other pieces of the results can be compared with the error of this simple substitution occurred in dataset B. The results show that the accuracy depends on the number of partner

series and the spatial correlations between time series. The RMSE of interpolated values is mostly smaller with 45–65% than the RMSE of the empirical climatic mean. Note that the accuracy of annual values is markedly higher than that of the monthly values even when each monthly value of the year is interpolated. For such years, the mean residual annual RMSE is 0.29 °C in dataset A and 0.34 °C in dataset B. These results indicate that the accuracy of annual values from interpolated monthly values tends to be slightly higher than the accuracy of the annual values from observed monthly values before homogenisation.

## 5. Discussion and recommendations

Approximately 20 years have passed since the multiple break theory was proposed. The underpinning principle of this theory is simple and is not limited to the topic of time series homogenisation: when a physical problem includes mutually dependent factors, adequate

mathematical methods must include the joint treatment of these factors. Although the frequency of the application of HOMER has increased exponentially in the recent years (e.g. Freitas *et al.,* 2013; Mamara *et al.,* 2014; Noone *et al.,* 2016) due to the recommendation of HOME, the overall frequency of using multiple break methods is still low, and there are indications that the general understanding of the multiple break homogenisation concept is still missing. Although HOMER was the first recommended method of HOME, we do not yet have clear evidence whether MASH, HOMER or ACMANT have the highest efficiency when they are applied to particular homogenisation tasks. Considering the similarities of the theoretical background of HOMER and ACMANT, it seems reasonable to assume that the efficiency of these two methods is likely to be similar. Therefore, the choice between HOMER and ACMANT for particular homogenisation tasks should be based on the dataset characteristics rather than on the efficiency order (even if this were known). The use of ACMANT is particularly recommended for (1) datasets with little or no metadata; (2) datasets from dense networks with large numbers of time series and where there are high spatial correlations; (3) very large datasets (>∼200 time series) for which the use of automatic methods is the most feasible and easily managed solution.

One important purpose of homogenisation is to deliver regional and global mean temperature trend estimates which are more accurate (Rohde *et al.,* 2013; Rennie *et al.,* 2014; Venema *et al.,* 2015). The efficiency tests presented here provide firm indications that ACMANT3 can considerably reduce initial regional trend biases at any spatial scale, although the efficiency achieved depends both on the spatial density and the extent of the intact record of the observational data. Further research is needed in this important and emerging area, for both the development and testing of statistical methods (Domonkos and Guijarro, 2015) and alongside an analysis of the causes of possible systematic biases in temperature records, with parallel measurements (http://www.surfacetemperatures.org/databank/parallel_measurements).

Certainly in the case of ACMANT, there is still room for further development and refinement. For instance, the ACMANT3 daily homogenisation programmes do not yet contain varied adjustment terms according to the percentiles of the probability distribution function, while such procedures are included in some other homogenisation methods, e.g. Della-Marta and Wanner (2006); Kuglitsch *et al.* (2009); Mestre *et al.* (2011). The analysis of the frequency changes of the 0 values in precipitation time series can also be incorporated in automatic homogenisation as it is shown by Wang *et al.* (2010). Nevertheless it is considered that the main phases in the development of ACMANT as a fully automatic homogenisation method for temperature and precipitation homogenisation have been completed.

Beyond the methodological improvements, ACMANT3 is also more user-friendly. Of particular note, the software package includes auxiliary files supporting input data preparation and the joint use of ACMANT3 with the software 'Rclimdex-extraQC' for the common quality control of daily temperature and precipitation data.

The use of ACMANT is recommended with the following caveats:

1. A general quality control of data before the application of ACMANT is necessary. Date order errors or accidental mixing of station series might cause serious errors in the final results, and may even impede the execution of the programmes. Physical outliers due to data transcription errors (e.g. 100 °C instead of 10.0 °C) might result in serious biases in the calculation of climatic means affecting the final results of homogenisation. At current development, the inner quality control routine of ACMANT for filtering spatial outliers is not sufficient without the previous filtering of physically implausible values. Frequency of zero precipitation events should be checked before using ACMANT, as zeros are sometimes erroneously shown instead of missing data code in climate records.

2. Synchronous breaks might affect seriously the efficiency of homogenisation, as the concept of relative homogenisation is that breaks can be identified from the differences of the candidate and reference time series. The potential danger is obvious when the synchronous break is present in at least in half of the time series of a network and tests indicate (not shown) that the efficiency declines even when the ratio of affected time series is much lower than 50%. However, synchronous breaks are often the consequences of new protocols in observing networks, and hence, they are often well documented. It is recommended to apply adjustments for known synchronous breaks before using ACMANT, even if the precise break magnitudes are not known.

3. If a break magnitude is known from parallel measurements, then it is the best to apply the known correction term before using ACMANT.

4. The use of raw climate records is preferred as the input for ACMANT rather than products of previous homogenisations or other kinds of secondary data products. In the optimal case, the ACMANT input does not contain interpolated values or adjustments based on spatial comparisons other than the adjustments applied for synchronous breaks if those are reasonable.

5. The maximum number of time series that can be homogenized within one network is 99 using the ACMANT homogenisation methods, and the use of very large networks is not recommended. The optimal size of networks is usually around 20–30 time series, although when the potential reference series are shorter than the candidate series or where they are often incomplete, then the optimum number of time series can be much higher. For large and dense networks, it is recommended to apply automatic networking following the suggestions of Domonkos (2015b).

The ACMANT3 software package together with its manual is freely accessible from http://www.c3.urv.cat/data.html.

## Appendix A1: Mathematical formulations

### A1.1. Explanation of symbols

$A$ – time series of deseasonalized observed values
$B$ – preliminary adjustment term
$c, c_p$ – parameter
$d$ – internal distance
$e$ – external distance
$F$ – composite reference series
$g$ – index of candidate series
$h$ – length of period
$h'$ – number of available data in a period
$h_1, h_2$ – starting and ending points of period in any time scale
$i, i*$ – time point
$j$ – year
$j_1, j_2$ – starting and ending years
$k$ – serial number of break/step
$K$ – total number of breaks
$m$ – calendar month
$n$ – number of years in time series
$N$ – total number of stations in network or of reference series
$N'$ – total number of usable reference series at a particular step
$P$ – penalty term
$Q$ – relative time series
$r$ – spatial correlation
$s$ – serial number of reference series
$T$ – matrix
$U$ – operator
$w$ – weight
$W$ – accumulated weight
$x$ – serial number of ensemble homogenisation
$Z, Z'$ – adjustment term
upper stroke – section mean or mean of entire time series

### A1.2. Reference series and relative time series

Reference series ($F$) are built for candidate series ($G$) are composed from the other time series ($A_s$) of the same climatic network. The reference series often covers only a section $[h_1, h_2]$ of the candidate series. Reference series usually contain entire years, but the reference series with daily resolution for operations on daily scale are the exceptions.

$$F_{g[h_1,h_2]} = \sum_{s=1}^{N} w_g A_{s[h_1,h_2]}{}^g \qquad (A1)$$

The weights of the reference composites are determined by ordinary kriging (Szentimrey, 2010) if $N > 5$, while by the spatial correlations with the candidate series (Domonkos, 2011b) in the reverse case.

Break detection and outlier detection are performed on relative time series ($Q$).

$$Q_g = A_{g[h_1,h_2]} - F_{g[h_1,h_2]} \qquad (A2)$$

### A1.3. Determination of the number of breaks and step function fitting with predetermined number of steps

Optimal step function fitting with $K$ steps:

The task of optimal step function fitting as a model of time series is identical with the minimisation of the variance of internal distances (variation of observational data within constant sections of the model) and maximisation of the variance of external distances (variation of the modelled values, Lindau and Venema, 2013). The step function has $K + 1$ constant sections ($k = 0, 1, 2, \ldots K$).

Internal distance ($d$):
$$d(U)_i = U(q)_i - \overline{U(q)_k} \text{ where } i \in k \qquad (A3)$$
External distance ($e$):
$$e(U)_i = \overline{U(q)_k} - \overline{U(Q)} \text{ where } i \in k \qquad (A4)$$
$U$ is operator, most frequently (but not always) the generator of time average.

(1) Univariate detection

$$\min_{[h_1, h_2, \ldots h_K]} \left\{ \sum_{k=0}^{K} \sum_{i=h_k+1}^{h_{k+1}} \left( d(U)_{k,i} \right)^2 \right\} \qquad (A5)$$

(2) Bivariate detection

$$\min_{[j_1, j_2, \ldots j_K]} \left\{ \sum_{k=0}^{K} \sum_{i=j_k+1}^{j_{k+1}} \left( d(U_1)_{k,i}^2 + c \left( d(U_2)_{k,i} \right)^2 \right) \right\} \qquad (A6)$$

The bivariate detection is always performed on series with annual resolution. $c$ is empirical constant ($c = 0.2$) in temperature homogenisation, while it is the squared ratio of snowy months in proportion to the rainy months in precipitation homogenisation.

For univariate detection on annual scale and for bivariate detections:

$$j_{k+1} - j_k \geq j* \text{ for every } k \in \{0 \leq k \leq K\} \qquad (A7)$$

In most operations $j* = 3$, but in the break detection for annual temperature series of particular calendar months $j* = 5$.

In determining the optimal number of steps, the modified Caussinus–Lyazrhi criterion (C–L) is used, in which expression (A8) is minimized. This criterion is used only in annual scale detections.

$$\ln\left\{1 - \frac{\sum\limits_{k=0}^{K} (j_{k+1} - j_k) \cdot \left[\left(e\,(U_1)_k\right)^2 + c\,\left(e\,(U_2)_k\right)^2\right]}{\sum\limits_{i=1}^{h} \left\{\left(U_1\,(q)_i - \overline{\mathbf{U_1\,(Q)}}\right)^2 + c\,\left(U_2\,(q)_i - \overline{\mathbf{U_2\,(Q)}}\right)^2\right\}}\right\} + P \tag{A8}$$

$$P = c_p \frac{2K}{h-1} \ln(h) \tag{A9}$$

Note (1) that (A8) and (A9) differ only by the inclusion of coefficient $c_p$ in the penalty term from the original C–L (Caussinus and Lyazrhi, 1997); (2) When $c = 0$, the formula is usable for univariate detection.

Usually $c_p = 1.4$ in univariate detection and $c_p = 1.0$ in bivariate detection, but in the first pre-homogenisation the coefficient is elevated with 40%.

### A1.4. Taking the minimum correction terms of ensemble homogenisation

Phase 1: Let us assume we have $N$ time series in a given network where the length of the studied period is $n$ years. For any time series $s$, the possible highest number of ensemble members is $N - 1$, as it is the maximum possible number of reference components. Due to the different lengths of time series or low spatial correlations the real numbers of ensemble members ($N'$) for a particular station and particular year can be lower, so that $N'_{s,j} \le N - 1$ for any $s$ and $j$. For each ensemble member, the matching adjustment terms (**T**) are calculated with ANOVA on annual scale and these preliminary results are stored. **T** is a three dimensional matrix with the dimensions of station serial number ($s$), year ($j$), and the serial number of ensemble experiment ($x$).

Phase 2: The final adjustment in a pre-homogenisation procedure (**Z**) for any station $s$ and year $j$ is:

$$z_{s,j} = \begin{cases} \min\limits_{x=\{1,N'_{s,j}\}} |t_{s,j,x}| & \text{if } sign\,(t_{s,j,x}) = c \text{ for every } x \\ 0 & \text{if } sign\,(t_{s,j,x}) \neq c \text{ for any } x \in \{1, N'_{s,j}\} \end{cases} \tag{A10}$$

$$c = 1 \quad or \quad c = -1 \tag{A11}$$

### A1.5. Interpolation for substituting missing values

$$a_{g,i*} = \frac{1}{W} \sum_{s=1}^{N} c\,(h)\,r_{g,s}^2\,a_{s,h(s)[j1,j2]} + \frac{1}{h'} \sum_{i=j_1(g)}^{j_2(g)} a_{g,i} \tag{A12}$$

$$W = \max\left\{0.4, \sum_{s=1}^{N} c\,(h)\,r_{g,s}^2\right\} \tag{A13}$$

The lengths of the periods ($h$) and the number of data utilized ($h'$) for the calculation of empirical climatic mean around the date for which the interpolation is performed ($i*$) depend on the data availability in the partner series. However, periods taken into account always include entire years (i.e. they start with 01 January and end with 31 December). The weights are set subjectively, and they depend on factors influencing the usability of values in partner series in contributing to the accurate estimation of the target value: (1) spatial correlation ($r_{g,s}$); (2) data availability around $i*$; (3) whether data are pre-homogenized or not; (4) the phase of the homogenisation (as the interpolation is repeatedly performed in ACMANT, with data of increasing quality); (5) daily or monthly interpolation is applied.

It follows from (A12) and (A13) that if no data of partner series is available, then the missing data of the candidate series will be substituted with the climatic mean.

### A1.6. Smoothing of monthly adjustment terms

The connection between the preliminary adjustment terms ($b$) and final adjustment terms ($z'$) is shown by (A14).

$$z'_{j,m} = 0.3b_{j,m-1} + 0.4b_{j,m} + 0.3b_{j,m+1} \quad m = \{1, 2, \dots 12\} \tag{A14}$$

$$b_{j,0} = b_{j-1,12}; \qquad b_{j,13} = b_{j+1,1} \tag{A15}$$

### Appendix A2: Additional noise, climatic trends and inhomogeneities in test datasets (A) and (B).

Additional noise is applied only to dataset (B), while the other operations are applied uniformly to datasets (A) and (B).

### A2.1. Definitions

1. Central series: The series which has the highest spatial correlations on average with the other series of the same network, is named the central series.
2. Limit bias: Biases of multiple inhomogeneities of the same time series may be accumulated until the parameter of subjectively defined limit values, they are named limit biases.
3. Platform: Platform-shaped inhomogeneity, pair of shifts of the same sign and opposite directions.

### A2.2. Additional noise

A monthly series of normally distributed red noise with 0 mean and 0.15 autocorrelation is added. The variance is a function of the spatial correlation with the central series, and the parameters of this function are empirically

determined for obtaining the target mean spatial correlation in networks.

### A2.3. Additional climatic trends

Each network receives one additional climatic trend. The trend magnitude is a random variable of uniform distribution between 0 and 2 °C. The length of the trend section varies between 1 and 100 years.

The climatic trend is added first to the central series, then with little variation to the other time series. The variation is random, proportioned to the network mean change magnitude, and its possible highest value depends on the spatial correlation with the central series. This variation is generally small, and seldom exceeds 10% of the magnitude of the network mean change.

### A2.4. Inserted inhomogeneities

1 The positions and properties of inhomogeneities are determined with the use of a random number generator. As a consequence, although all kinds of inhomogeneities fall with equal probability to any time series, the number and sizes of inhomogeneities randomly vary between networks and time series.
2 Forms of inhomogeneities: Sudden shift, trend and short-term platform. All inhomogeneities are changes in the means.
3 Mean frequency of inhomogeneities: sudden shift with 4/100 year, trend with 1/100 year, platform with 3/100 year.
4 Magnitude of inhomogeneities: Random variable with 0 mean and normal distribution. The standard deviation is larger in the first 50 years (0.8 °C) than in the second half of the series (0.5 °C). The magnitude distribution is the same for sudden shifts and trends, but all magnitudes are elevated with 30% for platforms. Accumulated biases are not allowed to exceed the limit biases, which are −1 and +3 °C in the first 50 years while −1.25 and +1.25 °C in the second half of the series. Biases due to platforms may exceed with 30% the limit bias. Limit bias values are not symmetrical around zero in the first 50 years for producing significant network mean trend biases. When the inclusion of a randomly generated inhomogeneity would cause the exceedance of a limit bias, then that inhomogeneity is dropped and a new inhomogeneity of the same shape and with the same timing is generated. No other constraint is applied in the temporal sequence of the inserted inhomogeneities.
5 Length of trends: Variable of uniform distribution between 5 and 100 years.
6 Length of platforms: Varies between 1 and 120 months, the frequency quadratically decreases with increasing length. Thus, most platform-shaped biases are very short lived and in the special case of 1 month duration, outlier value is formed. Note that most of the short-term changes cannot be detected for the low signal-to-noise ratio.

7 Seasonal cycle of biases: 25% of the inhomogeneities are without seasonal variation of bias. In the other 75%, the form of the variation is semi-sinusoid with modes in July and December (so that the spring half period lasts 7 months, while the autumn half period lasts 5 months). The magnitude is a random variable of uniform distribution. The peak-to-peak amplitude (i.e. the difference between July bias and December bias) varies between −1 and +4 °C in the first 50 years and between −1.5 and +1.5 °C in the second 50 years. There are limit biases also for the seasonal differences of accumulated biases. The limits are −1 and +4 °C in the first 50 years, while −2 and +2 °C in the second half of the time series.

### References

Acquaotta F, Fratianni S. 2014. The importance of the quality and reliability of the historical time series for the study of climate change. *Rev. Bras. Climatol.* **10**: 20–38.

Acquaotta F, Fratianni S, Venema V. 2016. Assessment of parallel precipitation measurements networks in Piedmont, Italy. *Int. J. Climatol.* , doi: 10.1002/joc.4606.

Aguilar E, Auer I, Brunet M, Peterson TC, Wieringa J. 2003. *Guidelines on climate metadata and homogenization. WCDMP-53, WMO-TD 1186*. Geneva, Switzerland: World Meteorological Organization.

Alexandersson H, Moberg A. 1997. Homogenization of Swedish temperature data. 1. Homogeneity test for linear trends. *Int. J. Climatol.* **17**: 25–34.

Beaulieu C, Seidou O, Ouarda TBMJ, Zhang X, Boulet G, Yagouti A. 2008. Intercomparison of homogenization techniques for precipitation data. *Water Resour. Res.* **44**: W02425, doi: 10.1029/2006WR005615.

Bergstrom H, Moberg A. 2002. Daily air temperature and pressure series for Uppsala (1722–1998). *Clim. Change* **53**: 213–252.

Caussinus H, Lyazrhi F. 1997. Choosing a linear model with a random number of change-points and outliers. *Ann. Inst. Stat. Math.* **49**(4): 761–775.

Caussinus H, Mestre O. 1996. New mathematical tools and methodologies for relative homogeneity testing. *First Seminar for Homogenization of Surface Climatological Data*, Budapest, Hungary, 63–82.

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

Conrad V. 1925. Homogenitätsbestimmung meteorologischer Beobachtungsreihen. *Meteorol. Z.* 482–485.

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

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

Domonkos P. 2011b. Adapted Caussinus-Mestre Algorithm for Networks of Temperature series (ACMANT). *Int. J. Geosci.* **2**: 293–309, doi: 10.4236/ijg.2011.23032.

Domonkos P. 2013. Measuring performances of homogenization methods. *Idojaras Q. J. Hung. Meteorol. Serv.* **117**: 91–112.

Domonkos P 2014. The ACMANT2 software package. *Eighth Seminar for Homogenization and Quality Control in Climatological Databases*. WMO-WCDMP-84: 46–72.

Domonkos P. 2015a. Homogenization of precipitation time series with ACMANT. *Theor. Appl. Climatol.* **122**: 303–314, doi: 10.1007/s00704-014-1298-5.

Domonkos P. 2015b. Automatic networking for the homogenization of large climatic datasets. *15th Annual Meeting of the European Meteorological Society Sofia (Bulgaria), EMS2015-84*.

Domonkos P, Guijarro JA. 2015. Efficiency tests for automatic homogenization methods of monthly temperature and precipitation series. *10th EUMETNET Data Management Workshop*, Oct 28–30, St. Gallen, Switzerland.

Domonkos P, Venema V, Mestre O. 2011. Efficiencies of homogenisation methods: our present knowledge and its limitation. *Seventh Seminar for Homogenisation and Quality Control in Climatological Databases*, WMO-WCDMP-78, 19–32.

Durre I, Menne MJ, Gleason BE, Houston TG, Vose RS. 2010. Robust automated quality control of daily surface observations. *J. Appl. Meteorol. Climatol.* **49**: 1615–1633, doi: 10.1175/2010 JAMC2375.1.

Freitas L, Pereira MG, Caramelo L, Mendes M, Nunes LF. 2013. Homogeneity of monthly air temperature in Portugal with HOMER and MASH. *Idojaras Q. J. Hung. Meteorol. Serv.* **117**: 69–90.

Kuglitsch FG, Toreti A, Xoplaki E, Della-Marta PM, Luterbacher J, Wanner H. 2009. Homogenization of daily maximum temperature series in the Mediterranean. *J. Geophys. Res.* **114**: D15108, doi: 10.1029/2008JD011606.

Lindau R, Venema V. 2013. On the multiple breakpoint problem and the number of significant breaks in homogenization of climate records. *Idojaras Q. J. Hung. Meteorol. Serv.* **117**: 1–34.

Lindau R, Venema V. 2016. The uncertainty of break positions detected by homogenization algorithms in climate records. *Int. J. Climatol.* **36**: 576–589, doi: 10.1002/joc.4366.

Mamara A, Argiriou AA, Anadranistakis M. 2014. Detection and correction of inhomogeneities in Greek climate temperature series. *Int. J. Climatol.* **34**: 3024–3043, doi: 10.1002/joc.3888.

Menne MJ, Williams CN Jr, Vose RS. 2009. The U.S. Historical Climatology Network monthly temperature data, version 2. *Bull. Am. Meteorol. Soc.* **90**: 993–1007, doi: 10.1175/2008BAMS2613.1.

Menne MJ, Durre I, Vose RS, Gleason BE, Houston TG. 2012. An overview of the Global Historical Climatology Network-daily database. *J. Atmos. Oceanic Tech.* **29**: 897–910, doi: 10.1175/JTECH-D-11-00103.1.

Mestre O, Gruber C, Prieur C, Caussinus H, Jourdain S. 2011. SPLIDHOM: a method for homogenization of daily temperature observations. *J. Appl. Meteorol. Climatol.* **50**: 2343–2358, doi: 10.1175/2011JAMC2641.1.

Mestre O, Domonkos P, Picard F, Auer I, Robin S, Lebarbier E, Böhm R, Aguilar E, Guijarro J, Vertacnik G, Klancar M, Dubuisson B, Štěpánek P. 2013. HOMER: homogenization software in R – methods and applications. *Idojaras Q. J. Hung. Meteorol. Serv.* **117**: 47–67.

Noone S, Murphy C, Coll J, Matthews T, Mullan D, Wilby RL, Walsh S. 2016. Homogenization and analysis of an expanded long-term monthly rainfall network for the Island of Ireland (1850–2010). *Int. J. Climatol.* **36**: 2837–2853, doi: 10.1002/joc.4522.

Peterson TC, Easterling DR. 1994. Creation of homogeneous composite climatological reference series. *Int. J. Climatol.* **14**: 671–679.

Peterson TC, Easterling DR, Karl TR, Groisman P, Nicholls N, Plummer N, Torok S, Auer I, Boehm R, Gullett D, Vincent L, Heino R, Tuomenvirta H, Mestre O, Szentimrey T, Salinger J, Forland EJ, Hanssen-Bauer I, Alexandersson H, Jones P, Parker D. 1998. Homogeneity adjustments of in situ atmospheric climate data: a review. *Int. J. Climatol.* **18**: 1493–1517.

Prohom M, Barriendos M, Sanchez-Lorenzo A. 2016. Reconstruction and homogenization of the longest instrumental precipitation series in the Iberian Peninsula (Barcelona, 1786–2014). *Int. J. Climatol.* **36**: 3072–3087, doi: 10.1002/joc.4537.

Rennie JJ, Lawrimore JH, Gleason BE, Thorne PW, Morice CP, Menne MJ, Williams CN, Gambi de Almeida W, Christy JR, Flannery M, Ishihara M, Kamiguchi K, Klein-Tank AMG, Mhanda A, Lister DH, Razuvaev V, Renom M, Rusticucci M, Tandy J, Worley SJ, Venema V, Angel W, Brunet M, Dattore B, Diamond H, Lazzara MA, Le Blancq F, Luterbacher J, Mächel H, Revadekar J, Vose RS, Yin X. 2014. The international surface temperature initiative global land surface databank: monthly temperature data release description and methods. *Geosci. Data J.* **1/2**: 75–102, doi: 10.1002/gdj3.8.

Rienzner M, Gandolfi C. 2011. A composite statistical method for the detection of multiple undocumented abrupt changes in the mean value within a time series. *Int. J. Climatol.* **31**: 742–755, doi: 10.1002/joc.2113.

Rohde R, Muller R, Jacobsen R, Perlmutter S, Rosenfeld A, Wurtele J, Curry J, Wickham C, Mosher S. 2013. Berkeley Earth temperature averaging process. *Geoinform. Geostat.* **1**: 2, doi: 10.4172/gigs.1000103.

Szentimrey T 1996. Statistical procedure for joint homogenization of climatic time series. *First Seminar for Homogenization of Surface Climatological Data*, Budapest, Hungary, 47–62.

Szentimrey T. 1999. Multiple Analysis of Series for Homogenization (MASH). *Second Seminar for Homogenization of Surface Climatological Data*, WMO-WCDMP-41, 27–46.

Szentimrey T 2010. Methodological questions of series comparison. *Sixth Seminar for Homogenization and Quality Control in Climatological Databases*, WMO-WCDMP-76, 1–7.

Szentimrey T, Lakatos M, Bihari Z. 2014. Mathematical questions of homogenization and quality control. *Eighth Seminar for Homogenization and Quality Control in Climatological Databases*, WMO-WCDMP-84, 5–22.

Venema V, Mestre O, Aguilar E, Auer I, Guijarro JA, Domonkos P, Vertacnik G, Szentimrey T, Štěpánek P, Zahradnicek P, Viarre J, Müller-Westermeier G, Lakatos M, Williams CN, Menne M, Lindau R, Rasol D, Rustemeier E, Kolokythas K, Marinova T, Andresen L, Acquaotta F, Fratianni S, Cheval S, Klancar M, Brunetti M, Gruber C, Duran MP, Likso T, Esteban P, Brandsma T. 2012. Benchmarking monthly homogenization algorithms. *Clim. Past* **8**: 89–115, doi: 10.5194/cp-8-89-2012.

Venema V, Jones P, Lindau R, Osborn T. 2015. Is the global mean land surface temperature trend too low? *15th Annual Meeting of the European Meteorological Society Sofia (Bulgaria), EMS2015-557*.

Wang XL, Chen H, Wu Y, Feng Y, Pu Q. 2010. New techniques for detection and adjustment of shifts in daily precipitation data series. *J. Appl. Meteorol. Climatol.* **49**: 2416–2436, doi: 10.1175/2010JAMC2376.1.

Williams CN, Menne MJ, Thorne P. 2012. Benchmarking the performance of pairwise homogenization of surface temperatures in the United States. *J. Geophys. Res.* **117**: D05116, doi: 10.1029/2011JD016761.

World Meteorological Organisation 1996–2014. Series of homogenization and quality control in climatological databases. *World Climate Data Monitoring Program*. WMO-WCDMP 41, 56, 71, 76, 78, 84, Geneva, Switzerland.