THE ACMANT2 SOFTWARE PACKAGE

Peter Domonkos
Centre for Climate Change (C3), University Rovira i Virgili, Tortosa, Spain
e-mail: peter.domonkos@urv.cat

1. INTRODUCTION

Author has developed the statistical software package ACMANT2, which includes computer programs for the automatic homogenization of mean temperature (Tmean), daily maximum temperature (Tmax), daily minimum temperature (Tmin) and precipitation amounts (PP). The software treats either daily or monthly input, but the detection of inhomogeneities (IH) and the calculation of adjustment terms are always done on the annual or monthly scale, then daily data are adjusted with downscaling the monthly adjustment terms. This study will present the structure and the most important segments of ACMANT2 and will discuss why ACMANT was one of the best performing homogenization method in the international tests of the European project COST ES0601 (its popular name, HOME, will be referred hereafter). In the study, the latest version of ACMANT is referred to as ACMANT2, its previous version as ACMANT1, while its constant properties are often assigned to "ACMANT" without index.

Temperature and precipitation time series can be homogenized with ACMANT2 and the homogenization of these variables is done with very similar algorithms. In this study the description of temperature homogenization is provided only in detail, but the important differences between the algorithms for temperature homogenization and precipitation homogenization will be mentioned.

The organization of the study is as follows: In the next section, the motivation and the brief history of the development of ACMANT is presented. In the third section, the most important theoretical properties of ACMANT2 are shown. In section 4 some efficiency results are shown, while in section 5 the computer programs of the ACMANT2 package and their use are briefly described. The study is supplied with appendixes (AP) with the detailed descriptions of I) the differences between ACMANT1 and ACMANT2 in the homogenization of Tmean and Tmax of mid- or high latitudes; II) the specific rules of Tmin homogenization and homogenization of any temperature variable in tropical regions.

2. MOTIVATION AND BRIEF HISTORY OF THE DEVELOPMENT OF ACMANT

During HOME (2007-2011), the new method ACMANT for homogenizing monthly temperature series appeared and was found to be one of the most effective methods just in its first version, leaving relatively small residual errors in the homogenized series.

I have been dealing with testing the efficiency of homogenization methods since 2003 (Domonkos, 2008, 2011a, 2013a, etc.) and I often found large differences between the

efficiencies of different methods. Moreover, homogenization sometimes might worsen the quality of observational series (Domonkos, 2013a), therefore a new approach is needed in our general view around the task of time series homogenization. While the review study of Peterson et al. (1998) suggested that time series homogenization is generally recommendable anything is the method applied, the enhanced need for more reliable and more accurate observational data for climate change and climate variability studies forces us to select and use the best performing methods. The international tests with the HOME benchmark dataset (HBM) confirmed that the differences between method efficiencies are large (Venema *et al*., 2012) and based on these tests only five methods can be recommended for homogenizing monthly temperature and precipitation series, namely MASH (Szentimrey, 1999), PRODIGE (Caussinus and Mestre, 2004), ACMANT (Domonkos, 2011b, referred here D2011), USHCN (Menne and Williams, 2009) and the Craddock-test (Craddock, 1979). In the last stage of HOME, the HOMER method was created (Mestre *et al*., 2013) from the best performing segments of PRODIGE and ACMANT and incorporating the network wide joint segmentation method (Picard *et al*., 2011). After HOME, the climatologist community still has important tasks in continuing test experiments (Domonkos, 2013b), since the efficiencies measured by HOME are based on the use of a not very large benchmark dataset, i.e. 15 networks for testing each of temperature and precipitation homogenization methods (Venema *et al*., 2012). On the other hand, HOME recognized several weak points of the tested homogenization methods, fostering a new stage of the methodological developments.

In my opinion, the creation of an effective homogenization method must be based on three principles, namely: i) Consideration of the statistical properties of observational data (including its IHs), for which the homogenization method will be used; ii) Relying on the best results of earlier achievements; iii) Creating additional value with innovation and automation.

i) *Consideration of the statistical properties of observed temperature and precipitation series* - The mean frequency of detected IHs in European and North American climate records is around 5-6 per 100yr and per station (Auer *et al*., 2005; Menne *et al.* 2009; Venema *et al*., 2012), although it depends on network density and the examined climatic variable (Menne *et al.*, 2009) and on the homogenization method applied (Domonkos, 2011a). Note that IHs are modelled as a sudden shift in the mean and referred to as "break" throughout this study when no other specification is given. The true frequency of breaks is likely higher than their detected frequency, because small-size shifts and short-term biases often cannot be detected (Brohan *et al*., 2006; Menne *et al.*, 2009; Domonkos, 2011a, 2013a). Therefore, the true frequency of breaks in observational time series is expected to be at least equal but likely higher than 5 breaks per station and per 100yr, whilst other kinds of IHs (e.g. trend-like biases) may additionally occur in the series. The optimal way of homogenizing datasets with such IHs is the use of multiple break methods, i.e. methods in which the joint structure of IHs are searched and corrected directly, taking into account the mutual effects of individual IHs. The development of multiple break methods began in the last decade of the 20th century (apart from some basically subjective methods not considered in here) and now we have four methods of this kind: MASH, PRODIGE, HOMER and ACMANT. Considering that at present, climatologists apply approximately 40 methods for homogenizing temperature and precipitation series (Domonkos and Efthymiadis, 2013), multiple break methods compose a small cluster of the existing methods. It is striking that the cluster of the best performing methods in HOME tests is almost identical with the cluster of multiple break methods, showing that HOME tests justified the advantage of using multiple break methods. Consequently, the incorporation of multiple break techniques in ACMANT was a good decision.

Another positive feature of ACMANT related to this point is that the semi-sinusoid annual cycle of Tmean and Tmax biases is taken into account in the homogenization procedure. In mid- and high latitudes the annual cycle of insolation results in the annual cycle of biases caused by various IHs (Drogue *et al*., 2005; Domonkos and Štěpánek, 2009; Brunet *et al*., 2011), since bias-sizes are often related directly or indirectly to the duration and intensity of insolation. In ACMANT, the homogenization of Tmean and Tmax observed in mid- or high latitudes is performed with a bivariate detection where the two variables are the annual mean and the amplitude of summer – winter difference. The calculation of adjustment terms and some other routines of the computer program also consider the semi-sinusoid cycle of biases and I believe that these properties significantly contribute to the high efficiency of ACMANT.

ii. *Relying on the best results of earlier achievements* - ACMANT is based on the detection and correction method of PRODIGE (the name "ACMANT" came from "Adapted Caussinus - Mestre Algorithm for homogenising Networks of Temperature series). In testing detection parts of homogenization methods, PRODIGE showed the highest efficiency (Domonkos, 2011a, 2013a), while ANOVA is a correction method, with which even the results of other homogenization methods could be improved (Domonkos *et al*., 2011). The adaptation of routines, which once were effective in other homogenization methods, sets a good basis for the creation of new methods that could outperform the earlier methods.

iii. *Creating additional value with innovation and automation* - Three kinds of added value will be discussed here: a) Separating time scales in the detection of IHs; b) Exploitation of the partial regularity in the seasonal changes of biases caused by IHs; c) Automation of the homogenization procedure.

a) Although the main goal of time series homogenization is not the break detection (but the minimization of the residual non-climatic biases), break detection is included in every homogenization method, since the identification of break positions helps to eliminate the biases. On annual or multi-annual timescale, the spatial correlations and hence the signal to noise ratio is higher than on monthly scale. In addition, monthly data is often affected by annual cycle of bias, which is obviously absent in annual data. On the other hand, statistical samples are larger when time series are examined on monthly resolution, thus the advantages and drawbacks could seemingly be compensated by each-other. However, one main difficulty of the homogenization task is that both the number of breaks and their positions must be estimated from the sample, and the uncertainty can be reduced if breaks are searched at the time scale in which they manifest themselves best.

b) Tmean and Tmax data are often affected by seasonally varying biases and such variations can be modelled by sinusoid cycles. ACMANT and HOMER are the only methods, which exploit this feature of temperature data in their break detection algorithms.

c) ACMANT and HOMER have been developed from PRODIGE, but while PRODIGE and HOMER are semi-objective methods, ACMANT is a fully objective and fully automatic homogenization method. Note that we use the term "objective method" in the sense that the results do not depend on homogenizers. The objectivity and automation has four advantages: i) The use of automatic methods is advisable for large datasets and practically the only option for homogenizing datasets including 50 or more time series. ii) Testing of automatic methods is straightforward and tests are easily manageable even with huge test datasets. Owing to these tests the performance of automatic methods is more transparent than that of other methods. iii) A homogenization product of a fully objective method can be reconstructed at any time. iv) Automatic methods are easy-to-use for climatologists.

Note, however, that automatic methods are not competitors of subjective or semi-objective methods, since for the homogenization of small networks with the help of metadata, the use of subjective or semi-objective methods is preferred.

The ACMANT1 software was only for homogenizing monthly means of Tmean and Tmax observed in mid- or high latitudes. The full description of ACMANT1 is published in D2011. However, I have made new computer programs for homogenizing other variables and even the structure and content of the algorithm for homogenizing Tmean and Tmax have changed significantly since then. The two most important novelties in the general structure of ACMANT2 relative to its earlier version are as follows: i) In ACMANT2 the adjustment terms are always calculated by ANOVA, also in the phase of Pre-homogenization; ii) A new subroutine has been included, "Filtering of outlier period", that is always applied just after the common outlier filtering. The purpose of filtering of outlier periods is to remove large, short-term biases before homogenization, similarly as individual outlier values are removed. The decrease of the residual root mean square error in homogenized series can be expected from these changes.

The parameterization of ACMANT2 is based on tests with various large test datasets similar in climatic characteristics to the HBM, but varied in the properties of IHs included in them.

3. ACMANT2: STRUCTURE AND KEY THEORETICAL PROPERTIES

ACMANT2 is composed of four main segments, namely Introductory Operations, Pre-homogenization, Main Homogenization and Final Adjustments, and each main segment includes various routines (e.g. for break detection, outlier filtering, bias correction). Some routines are common for more than one main segment (Fig. 1), since the accuracy of certain operations increases with the improving homogeneity of the data during the procedure, and thus the repeated application of such routines improves the accuracy of the final results.

Fig. 1 shows the most important segments only, i.e. operations, like calculation of spatial correlations, building reference series, exclusion of detected breaks of insignificant size, etc. are not shown. On the other hand, routines marked with asterisk indicate that they are not included in precipitation homogenization. For homogenizing PP, most routines are applied in the same way as for temperature homogenization, after the row PP values are converted by a quasi-logarithmic transformation. However, the work on monthly scale is strongly reduced in PP homogenization due to the often large spatial and temporal irregularity in monthly PP totals (particularly in semiarid regions).

3.1. Relative homogenization based on reference series

ACMANT2 is a relative homogenization method, which means that the detection of IHs is performed by examining the differences between a candidate series and a reference series, and then any detected IH is assigned to the candidate series. The method of creating reference series from composite series mostly follows the rules of Peterson and Easterling (1994) with some differences in the details. Composite series are weighted according to the squared spatial correlations of monthly temperature anomalies and the first difference series

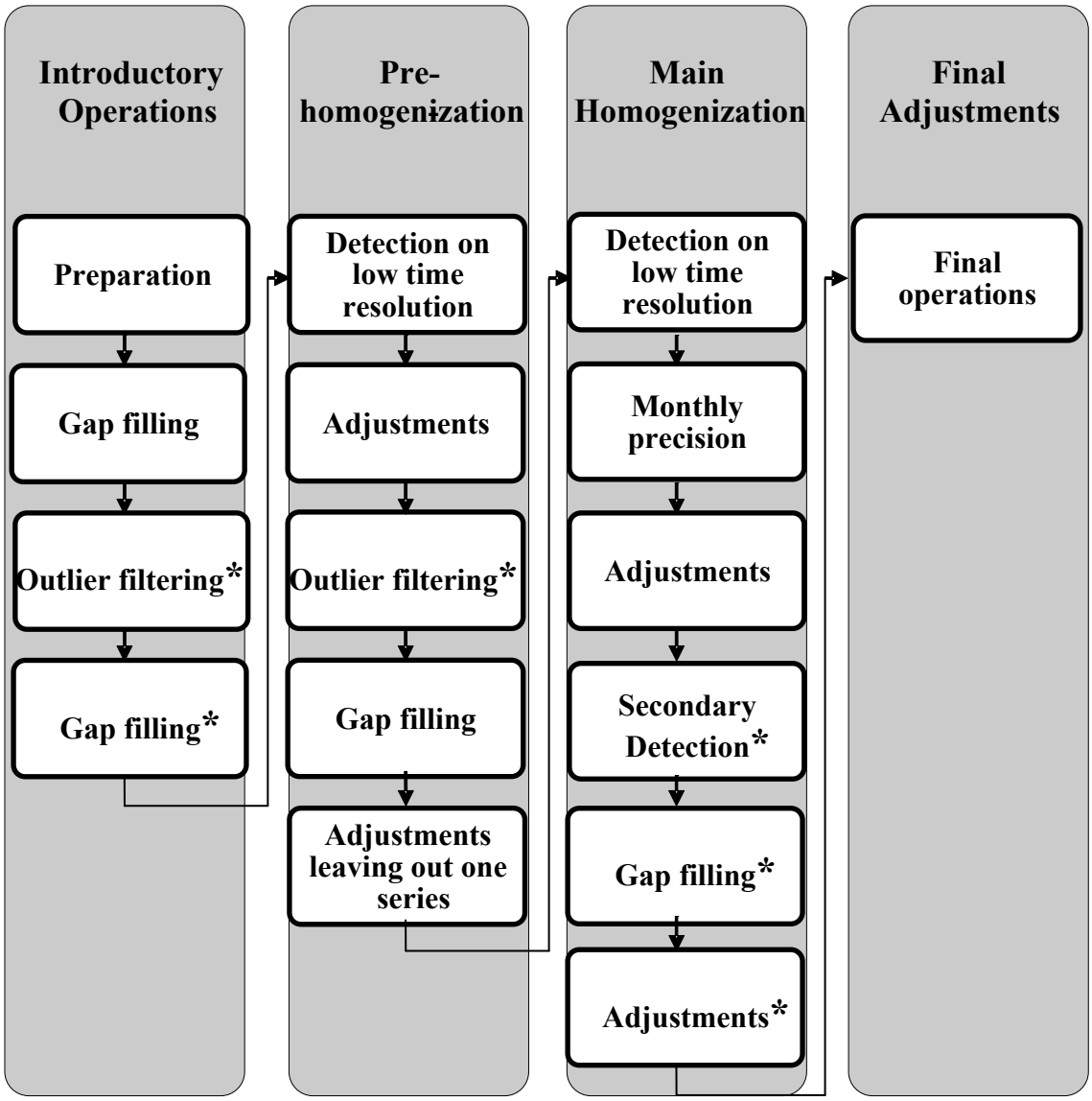

Fig. 1. Scheme of ACMANT2. The shown segments are common for all computer programs included in ACMANT2, except that the ones marked with (*) are included in temperature homogenization only.

(increment series) are used for calculating the correlations, in order to reduce the impact of IHs on the empirical correlations. Possible effects of IHs in the reference composites are not considered during the calculations of the spatial differences. Note, however, that in the phase of Main Homogenization the reference composites have been pre-homogenized (while the candidate series remains the raw, outlier filtered series), and in this way the IHs of the reference composites have markedly reduced impact on the final results. The parameterization for the calculation of reference series is shown in AP I-2.1.

A speciality of ACMANT2 is that the detection of IHs on low time resolution (i.e. detection of biases for at least 3 year long sections of the candidate series) has two main phases. The goal of the first phase within the Pre-homogenization segment is to remove or reduce the large-size biases of the time series, those that will be reference composites in the second phase, within the Main Homogenization stage. The first and second phases are performed in almost the same way, with a small change only in the parameterization. In the Pre-

homogenization phase, the future candidate series is excluded from the calculation of adjustment terms, and thus the multiple use of the same spatial relationship (including error term) is excluded.

The use of reference composites flexibly changes if the number of available reference composites is different for diverse sections of the candidate series (AP I-2.2), and all the reference composites of at least 0.4 spatial correlation with the candidate series are utilized, disregarding possible differences in the starting and ending dates of the series. Note that in ACMANT2, reference series are never used for calculating adjustment-terms. It is an important detail in which multiple break methods differ from more traditional homogenization methods (e.g. Standard Normal Homogeneity Test [SNHT] by Alexandersson and Moberg, 1997; RHTest by Wang, 2008, etc.).

3.2. Inhomogeneity detection on low time resolution

ACMANT2 includes two markedly different types of break detection. One is for identifying long-standing biases whose characteristic time is longer than 2 years (referred as detection on low time resolution), while the other is for identifying temporarily existing, short-lived but large size biases lasting from 3 months to 24 months.

3.2.1. Fitting optimal step function (univariate detection)

Fitting optimal step function is a known technic for the detection of multiple breaks in time series (Hawkins, 1972; Caussinus and Mestre, 2004). Presuming that a time series contains K IHs and all of them are sudden shifts of the mean values (i.e. breaks), the time series is modelled by a step function of K +1 steps. The optimal step function can be found with the variance minimization of the data relative to the step function model. As true IHs are often sudden shifts (e.g. due to station relocation, instrumental change), this model is realistic. When gradually increasing bias (trend-like) IHs occur, the step function approach still provides fair (although slightly less accurate) results, transforming the trend into two or more steps.

Let the annual mean (E) and the section mean (upper stroke) for step k of variable x be defined by (1) and (2),

$$E(x_j) = \frac{\sum_{m=1}^{12} x_{j,m}}{12} \tag{1}$$

$$\overline{\mathbf{X_k}} = \frac{1}{j_k - j_{k-1}} \sum_{i=j_{k-1}+1}^{j_k} x_i \tag{2}$$

then the optimal step function for time series **Q** of length L, including K breaks is given by (3) and (4).

$$\min_{[j_1, j_2, \ldots j_K]} \left\{ \sum_{k=0}^{K} \sum_{i=j_k+1}^{j_{k+1}} (E(q)_i - \overline{\mathbf{E}(q)_\mathbf{k}})^2 \right\} \tag{3}$$

$$j_0 = 0, \qquad j_{K+1} = L \tag{4}$$

Note that when step function fitting is applied in ACMANT2, the minimum length of a step is at least 3 time units, i.e. 3 years or 3 months depending on the time step between two adjacent values (5).

$$j_{k+1} - j_k \geq 3 \qquad \forall k \in \{0 \leq k \leq K\} \tag{5}$$

3.2.2. Bivariate detection of breaks with fitting optimal step function

As it has been mentioned, biases in Tmean and Tmax series often have sinusoid annual cycle linked to the annual cycle of insolation. Therefore when a break occurs for station relocation or change in the instrumentation, etc., both the annual means and the amplitude of seasonal cycle can be affected. In the bivariate detection of Tmean and Tmax homogenization, breaks with common timings are searched for the annual mean (E) and for the amplitude of summer-winter difference (Z). Z is defined as:

$$Z(x_j) = \frac{1}{3.5} \sum_{m=1}^{12} c_m x_m \tag{6}$$

where m denotes calendar month and the monthly coefficients (c_m) are negative in winter and positive in summer:
$c_1 = c_{11} = c_{12} = -1$,
$c_2 = -0.5$,
$c_3 = c_4 = c_9 = c_{10} = 0$,
$c_5 = c_6 = c_7 = 1$,
$c_8 = 0.5$

Then the best fitting step function to series **Q** including K breaks is given by (7).

$$\min_{[j_1, j_2, \ldots j_K]} \left\{ \sum_{k=0}^{K} \sum_{i=j_k+1}^{j_{k+1}} (E(q)_i - \overline{\mathbf{E}(q)_\mathbf{k}})^2 + c_0^2 (Z(q)_i - \overline{\mathbf{Z}(q)_\mathbf{k}})^2 \right\} \tag{7}$$

$c_0 = 5^{-0.5}$ (empirical constant). Note that

$$\overline{\mathbf{Q}_\mathbf{k}} \equiv \overline{\mathbf{E}(q)_\mathbf{k}} \tag{8}$$

by definition, the longer form in (3) and (7) is included only for showing the identity of the treatment for *E* and *Z*. Note also that for presenting the average of whole time series the index (*k*) will be omitted.

3.2.3. Assessment of the number of breaks

The critical point of step-function fitting methods is the determination of *K*. In ACMANT2 a parameterized version of the Caussinus - Lyazrhi criterion (Caussinus and Lyazrhi, 1997) is applied. This criterion takes into account the reduction of variance due to the inclusion of breaks, but with balancing that with a penalty term due to the increasing number of steps, because the residual variance tends to decrease with the rising number of steps either the breaks between steps are significant or not. As a consequence, the rise of the number of steps will give better score only if the reduction of standard deviation overbalances the increase of the penalty term (9), (10).

$$\ln\left\{1-\frac{\sum_{k=0}^{K}(j_{k+1}-j_k)\cdot\left[(\overline{\mathbf{Q}_\mathbf{k}}-\overline{\mathbf{Q}})^2+c_0^2(\overline{\mathbf{Z}(q)_\mathbf{k}}-\overline{\mathbf{Z}(q)})^2\right]}{\sum_{i=1}^{L}(E(q)_i-\overline{\mathbf{Q}})^2+c_0^2(Z(q)_i-\overline{\mathbf{Z}(q)})^2}\right\}+S \tag{9}$$

$$S = p\frac{2K}{L-1}\ln(L) \tag{10}$$

The shown formula differs from the original one in one detail, i.e. the penalty term (*S*) here includes an empirical coefficient (*p*). The value of *p* is different for univariate and bivariate detection, as well as different in the Pre-homogenization than in the Main Homogenization. In the Main Homogenization $p = 1.4$ in univariate detection and $p = 1.0$ in bivariate detection. See more details in AP I-3.1 and AP II-1.

3.3. Detection of short-term biases

Short-term IHs can be modelled by a platform-shape bias from the true climatic values where the platform is the composition of a pair of breaks of the same shift-size but to the opposite directions (Fig. 2). Short-term IHs can be caused by temporal changes in the conditions of the observation. The frequency of the short-term biases can be much higher than their detected frequency, because the signal-to-noise ratio is relatively low for short sections of time series due to the limited sample size (Domonkos, 2013a). Experimental results indicate that the true frequency is really significantly higher than the detected frequency (Domonkos, 2011a; Rienzner and Gandolfi, 2011).

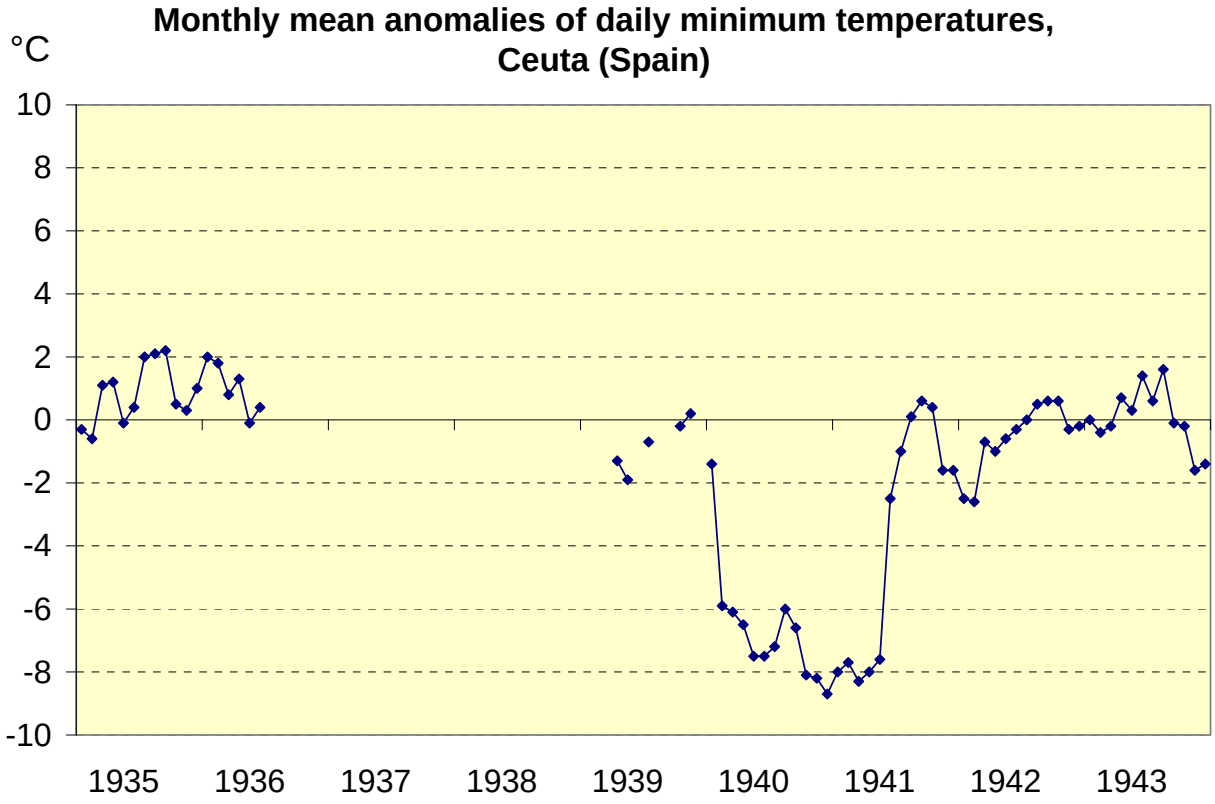

Fig. 2. Large, short-term, platform-shaped bias between 02-1940 and 05-1941 in the Tmin of Ceuta (Spain).

In ACMANT2, large-size, platform-shaped biases are detected in moving windows of monthly temperature anomalies. This new routine of ACMANT (such step is not included in ACMANT1) is named Filtering of outlier periods (AP I-4.1). This procedure has characteristics similar to filtering out outlier values, as well as to the minimization of standard deviation relative to platform-shape step functions. It is always performed just after the common Outlier filtering.

The detection of long-term biases might be affected by the existence of large-size short-term biases and vice versa. Therefore the detection and elimination of short-term biases is performed three times in ACMANT2, approaching step-by-step to the final solution. These operations are applied first in the Introductory Operations, then within the Pre-homogenization phase after the adjustment of long-term biases, and finally in the Main Homogenisation phase, after the adjustment of long-term biases. The way of the detection and correction of short term biases in the Main Homogenization differ from the Filtering of outlier periods, i.e. the routine Secondary Detection of ACMANT1 (D2011) has been kept with some little changes in the parameterization only (AP I-4.2 and AP II-4).

Note that in PP homogenization neither outlier filtering nor any kind of operation for filtering out short-term biases is performed.

3.4. Data adjustment

In ACMANT2, adjustment-terms are generally calculated with variance analysis (ANOVA, 3.4.1 – 3.4.2.). However, ANOVA determines temporal differences only, therefore one fix value (time series average, or a reference value of the time series which is considered to be unbiased) must be defined (3.4.3). For periods of very short-term biases, interpolation technique is applied instead of ANOVA (3.4.4).

3.4.1. The ANOVA model for the assessment of adjustment terms

The ANOVA procedure determines the minimum variance of anomalies relative to the climate signal of an examined region, relying on the timings of detected breaks in all the examined time series of the region. It is proven that ANOVA provides the optimum estimation of adjustment-terms when the spatial gradients of climate are temporally constant and the list of detected breaks is correct (Mestre, 2004; Caussinus and Mestre, 2004). Moreover, experiments showed that ANOVA performs better than conservative correction methods even when the list of detected breaks is partially correct only (Domonkos *et al.*, 2011).

For applying ANOVA in time series homogenization, a model is set up. In this model, the observed values are considered to be the sum of the climate signal (u), station effect (v) and noise (ε) for each time series and each time point (11).

$$\mathbf{X} = \mathbf{U} + \mathbf{V} + \boldsymbol{\varepsilon} \tag{11}$$

The spatial gradients of climate are temporally constant, which is approximately true for observational datasets when data for a specific climatic zone is examined. No other constraint is included for climate. Station effect means the sum of site effect (i.e. temporally constant difference relative to the climate signal) and the biases caused by IHs. In the model, all IHs are breaks, and their timings are known. This variance minimization can be solved with the construction of an equation system following the relationships in the model.

ANOVA searches the solution of (11) with the minimum variance of ε for the entire dataset. The minimum variance can be obtained by (12, 13).

$$Nu'_i + \sum_{s=1}^{N} v'_s = \sum_{s=1}^{N} x_{s,i} \quad \text{for every } i\text{: } i \in \left[j_{\min}, j_{\max} \right] \tag{12}$$

$$\sum_{i=j_k+1}^{j_{k+1}} u'_i + (j_{k+1} - j_k) v'_{s,k} = \sum_{i=j_k+1}^{j_{k+1}} x_{s,k} \quad \text{for every } s \text{ and } k \tag{13}$$

In (12) and (13) N denotes the number of station series, s the serial number of station series, $j_{\min}$ and $j_{\max}$ stand for the first and last years of the period, respectively, for which homogenisation is performed, while apostrophe denotes estimated variable.

When the period between adjacent breaks is very short or when simultaneous breaks occur in various station series, the efficiency of ANOVA is reduced. Moreover, if all time series have a detected break at the same time, then the equation system is undetermined. For these reasons, breaks of relatively small estimated size are sometimes deleted from the break list (AP I-5).

3.4.2. The use of ANOVA in ACMANT2

ANOVA is always applied using data without bias corrections, because the recursive application of ANOVA could multiply the errors of the estimated spatial relationships.

ANOVA can be applied separately for the variables under examinations, thus in the homogenization of Tmean and Tmax, ANOVA is applied separately for annual means ($E(x)$) and summer-winter differences ($Z(x)$), then the monthly adjustment-terms are derived from them. In the Pre-homogenisation, ANOVA is applied on data of annual resolution, while in the Main Homogenization the input data is monthly. In monthly resolution, **X** is examined directly, instead of **E**(x). However, **Z**(x) is a variable whose interpretation on monthly scale is not straightforward. Monthly values of **Z**(x) are defined for each month (h) of the series by using the data of the 12-month symmetric window around h (14).

$$Z(x)_{j,h} = \sum_{h'=h-5}^{h+5} c_{m(h')} \cdot x_{h'} + 0.5(c_{m(h-6)} x_{h-6} + c_{m(h+6)}) x_{h+6} \tag{14}$$

(Note: close to the endpoints of the series the extent of window is limited by the data availability.) Coefficients c_m are the same as in (6).

If break k has the timing $H(k)$ in monthly scale and α_k and β_k denote the estimated station effects for the homogeneous section of $[H(k\text{-}1)+1, H(k)]$ for **X** and **Z**(x), respectively, then the estimated station effect (v') is given by (15) for each month of the section.

$$v'_{i,m} = \alpha_k + c^*_m \beta_k \tag{15}$$

Coefficients c^*_m differ from c_m in a way that they provide the same summer – winter difference as c_m, but with a harmonic annual cycle of the coefficients.

When only one variable is examined as in the homogenization of Tmin, the determination of adjustment term is simplified to (16).

$$v'_{i,m} = \alpha_k \tag{16}$$

(16) shows that in the Tmin homogenization of ACMANT2 the monthly adjustment terms are independent from the season of the year.

3.4.3. Selection of reference period

In ACMANT2, the values of the last homogeneous section of the series are considered to be unbiased, and it is considered to be reference period in the adjustment of the other sections of the series. This assumption is rather general in time series homogenization, but it might have unfavourable consequences when the last homogeneous section is too short for acquiring the accurate estimates of its statistical properties or when it has characteristics atypical for the site due to instrument error or for any other reasons. If the statistical properties of the last homogeneous section do not reflect well the true climate, its use as reference period may cause biases in the mean climatic characteristics of the site, as well as in the spatial climatic gradients. Notwithstanding, this problem has no effect on the reliability of the temporal variability. Note that the possible application of adjustment terms varying according to the probability distribution function value (percentile) of the raw data (as for instance in Della-Marta and Wanner, 2006) would make the homogenization results more sensible to the choice of the reference period, since any bias of the empirical probability distribution from the true climate in the reference period would be exported to all the other sections of the time series. Therefore as long as ACMANT remains fully automatic, percentile dependent adjustments will not be included in it.

3.4.4. Adjustment of short-term biases

In the Introductory Operations and Pre-homogenization, the values within the period of detected short-term biases are always substituted with interpolated values. By contrast, in the Main Homogenization, biases shorter than 6 months are adjusted by interpolation only, while the longer ones are corrected by ANOVA.

4. EFFICIENCY OF ACMANT2 IN HOMOGENIZING THE HOME BENCHMARK

If we would like to compare the efficiencies of different homogenization methods, it is still the HBM is the best for this purpose, since the characteristics of the HBM are rather close to the characteristics of observational time series and the efficiencies with this dataset are known for several homogenization methods (Venema et al., 2012). As the biases of the HBM series have quasi sinusoid annual cycle, the program for homogenizing Tmean and Tmax can be tested with this dataset.

Fig. 3 shows the efficiency of ACMANT2 in comparison with the efficiencies of several other methods. It can be seen that the efficiency of ACMANT2 is slightly lower than that of ACMANT1. The slight decrease may have 3 reasons:
i) The parameterization of ACMANT1 can be overfitted to HBM, since in the development of ACMANT1 I used the HBM.
ii) In the HBM the number of short-term biases is unrealistically low, and thus the positive effect of the inclusion of filtering of outlier periods in ACMANT2 does not appear in the tests with this dataset. Note here that short-term biases are not inserted to HBM, thus short-term biases in HBM are present only in the rare cases of their accidental formation from randomly placed breaks.

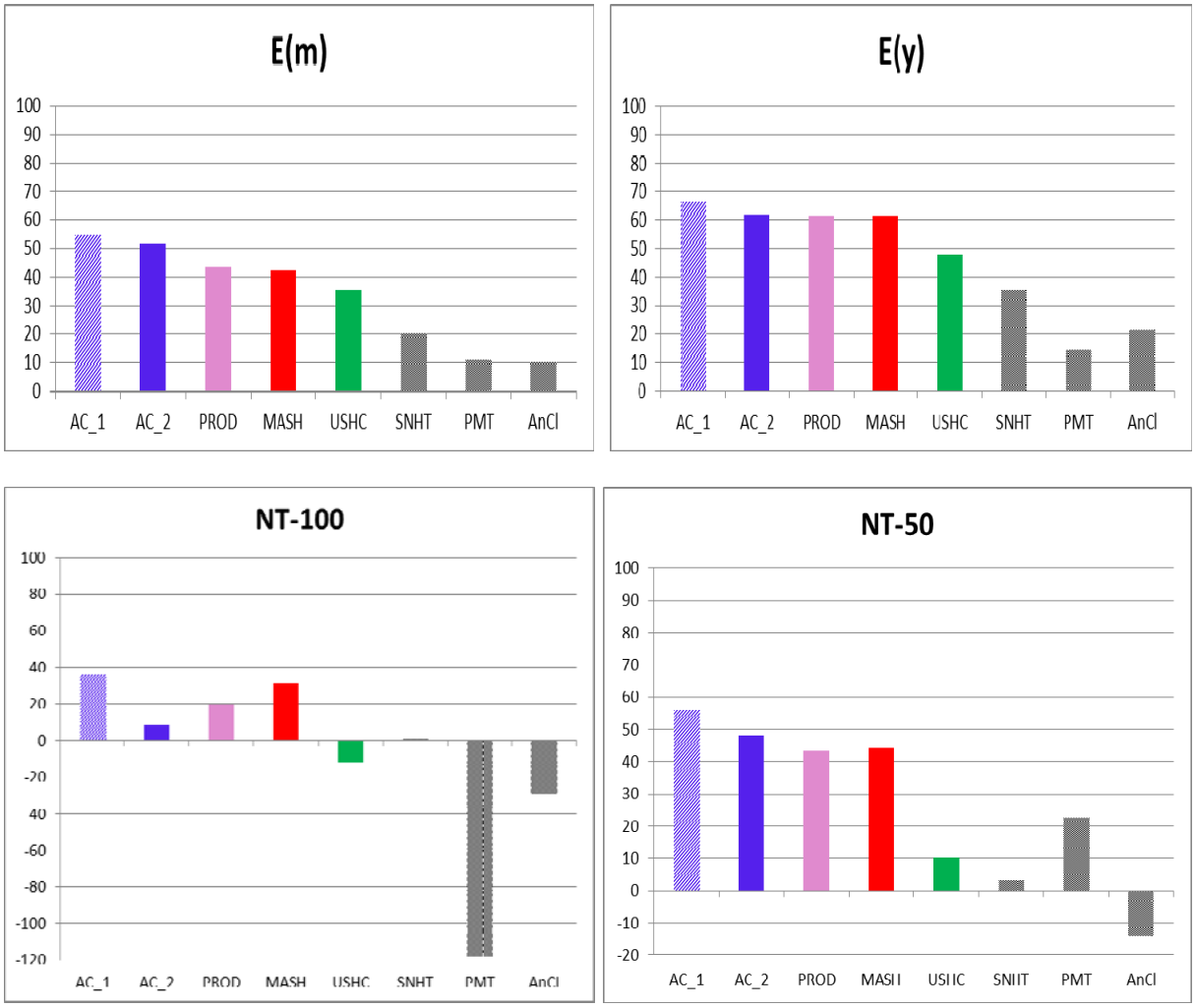

Fig. 3. Efficiency (%) in reducing RMSE of HBM with various homogenization methods. E(m) – RMSE of monthly values, E(Y) – RMSE of annual values, NT-100 – RMSE of network mean trends for the whole period (100 years) examined, NT-50 – RMSE of network mean trends over the last 50 years, AC_1 – ACMANT1, AC_2 – ACMANT2, PROD – PRODIGE, PMT – penalized maximum t-test of RHTest, AnCl – AnClim (Štěpánek et al., 2009).

iii. Random fluctuation of the results due to the small sample size. Note here that the sample size of HBM (15 networks) is obviously very small for the assessment of the efficiency in reducing network mean trend errors.

Fig. 3. shows that even with the few percentages drop relative to its earlier version, ACMANT is one of the most effective homogenization methods, and considering the fully automatic methods tested by HOME, still ACMANT shows the highest efficiency. Note however, that several other homogenization methods have also been developed since the HOME tests, thus new comparative tests with a large new benchmark of realistic time series properties are needed to see more clearly the rank order of the efficiencies.

5. USE OF ACMANT2 SOFTWARE

5.1. Some notes on the use of the software

The software package has a manual available in web (www. c3.urv.cat/data.html) together with the software. Therefore only some important points of the use are described here.

The software package includes 6 different computer programs which can be chosen according to the characteristics of the input raw data. Three programs are for homogenizing daily data, while the other three programs treat monthly data only. The three programs differ according to the variable treated: One program is for the homogenization of Tmean or Tmax, another one is for Tmin and the third one is for PP. The software package contains also some auxiliary files, their function and use is described in the Manual.

Anything is the input variable, some rules are common for the homogenization with ACMANT2. At least 4 time series with adequate spatial correlations (AP-I-2.1) are needed. The length of the input series may vary between 10 years and 200 years. The rules of input data preparation shown in the Manual must be followed accurately, otherwise the selected program will not run or will stop with some error message.

5.2. Selection of the appropriate program

The selection of the appropriate program seems to be straightforward, since the kind of the input variable (i.e. Tmin, Tmax, Tmean or PP) and its time resolution (daily or monthly) determine which program matches best. Yet there is a gap in this simple matching between variables and programs: The program including the bivariate detection for annual mean and summer – winter difference is proposed to use for Tmean and Tmax from the mid- or high latitudes only. As quasi sinusoid annual cycle of biases is not expected in temperature data of the tropical belt and in monsoon regions, the program with bivariate detection is not recommended to use there. The recommended matching between variables and programs is shown in Table 1.

Table 1. Recommended matching between input data type and programs of ACMANT2 software

	Program		
Input variable and region	T_{mean}&T_{max}	T_{min}	PP
T_{mean} or T_{max} in mid or high latitudes	X		
T_{mean} or T_{max} in tropical or monsoonal regions		X	
T_{min} anywhere		X	
PP anywhere			X

5.3. Options offered for users

Although ACMANT2 is fully automatic, there some options beyond the choice of the appropriate program are offered for the users at the initiation of the homogenization procedure. Four kinds of options are asked from the users: i) in temperature homogenization: the programs can be run with or without outlier filtering; ii) in precipitation homogenization: the program can be run with dividing the year into snowy and rainy seasons or without such division; iii) in homogenization of daily data: inputting both daily and monthly data or monthly raw data is constructed from daily data by the program; iv) output data format.

i) The programs of temperature homogenization can be run with or without outlier filtering. The proposed mode is the inclusion of outlier filtering, but there is one exception: Users may check manually the detected outliers by ACMANT, and this check might find that some detected outliers should be considered true extreme values instead of outliers. In such a case the proposed continuation of the homogenization procedure is as follows: a) Justified outliers must be shown with missing data code in the raw data; b) Accepted extreme values must be left unchanged in the raw data; c) Repeating the run of ACMANT2 with the use of the modified input data and without outlier filtering.

ii) If in a part of the year the dominant form of the precipitation is snow, the starting and ending months of the snowy season must be introduced at the initialization of the PP homogenization. It is because the IHs of snow data often markedly differ from the IHs of rain, due to the technical problems of catching snow precipitation and converting it into water amount comparable with rain precipitation. To manage this problem, PP homogenization in ACMANT2 includes two different modes, one mode is similar to the univariate homogenization of Tmin (when there is no snowy season in the region), while the other mode is bivariate homogenization in which the two variables are the total amount of the PP for the rainy season and that is for the snowy season. The program selects the appropriate mode automatically, after the requested parameters are introduced by the user.

iii) The detection of IHs and the calculation of adjustment terms are always based on monthly data. If the input is daily data, ACMANT2 develops the monthly dataset, does the homogenization and finally adjusts both the monthly and daily data. As input data may include missing values or outliers, the characteristics of the developed monthly dataset might depend on the treatment of these quality problems of the initial dataset. The development of monthly dataset in ACMANT2 is automatic, but there is an option for the users to introduce their own developed monthly dataset (together with daily data). This option is recommended to use in the case when the user has a monthly dataset which has been developed with the meticulous check of the possible quality problems in the daily data.

iv) The program offers that the output will consist of a default output package, but the user may select his choices if he wants. The goal of leaving free options in the form of the output package is to provide the opportunities of a) having homogenized dataset that is immediately applicable as input in extreme index calculation softwares, b) having the dataset with or without infilling the missing values, c) providing supplementary information about the spatial correlations and other characteristics related to the homogenization procedure.

6. CONCLUSIONS

ACMANT2 homogenization software has recently been developed. This software has been prepared for the automatic homogenization of observational temperature and precipitation datasets. This paper together with D2011, provides the whole description of the temperature homogenization with ACMANT2. The high efficiency of temperature homogenization with ACMANT2 is illustrated with the homogenization of the HOME benchmark dataset. Due to its high efficiency, author recommends the use of the software in each case when the size of the dataset and the spatial correlations allow the use of automatic homogenization method. The use of ACMANT2 is particularly recommended for very large datasets when it is hard to use non-automatic methods.

Acknowledgements: The research was supported by the European project UERRA FP7-SPACE-2013-1.

REFERENCES

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

Auer I, Böhm R, Jurkovic A, Orlik A, Potzmann R, Schöner W, Ungersböck M, Brunetti M, Nanni T, Maugeri M, Briffa K, Jones PD, Efthymiadis D, Mestre O, Moisselin J-M, Begert M, Brazdil R, Bochnicek O, Cegnar T, Gajic-Capka M, Zaninovic K, Majstorovic Z, Szalai S, Szentimrey T, Mercalli L. 2005. A new instrumental precipitation dataset for the greater Alpine region for the period 1800–2002. *Int. J. Climatol.* **25**: 139–166.

Brohan P, Kennedy JJ, Harris I, Tett SFB, Jones PD. 2006. Uncertainty estimates in regional and global observed temperature changes: a new data set from 1850. *Journal of Geophysical Research* **111**: D12106, doi: 10.1029/2005JD006548.

Brunet M, Asin J, Sigró J, Bañon M, García F, Aguilar E, Palenzuela JE, Peterson TC, Jones P. 2011. The minimization of the screen bias from ancient Western Mediterranean air temperature records: an exploratory statistical analysis, *Int. J. Climatol*. **31**: 1879–1895.

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

Caussinus H, Mestre O. 2004. Detection and correction of artificial shifts in climate series. *J. Roy. Stat. Soc*. **C 53**: 405-425.

Craddock JM. 1979. Methods of comparing annual rainfall records for climatic purposes. *Weather*, **34**: 332-346.

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

Domonkos P. 2008: Testing of homogenisation methods: purposes, tools and problems of implementation. *Proceedings of the 5th Seminar for Homogenisation and Quality Control in Climatological Databases*. (Ed. Lakatos, M., Szentimrey, T., Bihari, Z. and Szalai, S.), WCDMP-No. 71, WMO/TD-NO. 1493, 126-145.

Domonkos P. 2011a. Efficiency evaluation for detecting inhomogeneities by objective homogenisation methods. *Theor. Appl. Climatol.* **105**: 455-467.

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. 2013a. Efficiencies of inhomogeneity-detection algorithms: comparison of different detection methods and efficiency measures. *Journal of Climatology* pp15, doi:10.1155/2013/390945.

Domonkos, P. 2013b: Measuring performances of homogenization methods. Időjárás 117, 91-112.

Domonkos P, Efthymiadis D. 2013. After HOME: Progress in the practical application of homogenization methods. 11th ECAM conference, 09-13 Sept 2013, Reading, England, EMS2013-99.

Domonkos P, Štěpánek P. 2009. Statistical characteristics of detectable inhomogeneities in observed meteorological time series. *Studia Geoph. Geod*. **53**: 239-260, doi: 10.007/ s11200-009-0015-9.

Domonkos P, Venema V, Mestre O. 2011. Efficiencies of homogenisation methods: our present knowledge and its limitation. In: Seventh Seminar for Homogenisation and Quality Control in Climatological Databases (Eds. Lakatos M, Szentimrey T, Vincze E.), WCDMP-78, WMO, Geneva, 11-24.
Drogue G, Mestre O, Hoffmann L, Iffly J-F, Pfister L. 2005. Recent warming in a small region with semi-oceanic climate, 1949-1998: what is the ground truth? *Theor. Appl. Climatol*., **81**: 1-10.
Hawkins DM. 1972. On the choice of segments in piecewise approximation. *J. Inst. Math. Appl*. **9**: 250–256, doi:10.1093/imamat/9.2.250
Menne M, Williams Jr CN. 2009. Homogenization of temperature series via pairwise comparisons. *J. Clim*. **22**: 1700-1717.
Menne MJ, Williams Jr CN, Vose RS. 2009. The U.S. Historical Climatology Network monthly temperature data, version 2. *Bull. Amer. Meteor. Soc*., **90**: 993-1007.
Mestre O. 2004: Correcting climate series using ANOVA technique. In: Fourth Seminar for Homogenization of Surface Climatological Data (Eds. Szalai S, Szentimrey T) WCDMP 56, WMO-TD 1236, WMO, Geneva, 93-96.
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, *Időjárás* **117**: 47–67.
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, Böhm R, Gullett D, Vincent L, Heino R, Tuomenvirta H, Mestre O, Szentimrey T, Salinger J, Førland 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.
Picard F, Lebarbier E, Hoebeke M, Rigaill G, Thiam B, Robin S. 2011. Joint segmentation, calling and normalization of multiple CGH profiles. *Biostatistics*, **12**: 413-428. doi:10.1093/biostatistics/kxq076
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.
Štěpánek P, Zahradnícek P, Skalák P. 2009. Data quality control and homogenization of the air temperature and precipitation series in the Czech Republic in the period 1961–2007. *Adv. Sci. Res*., **3**: 23–26.
Szentimrey T. 1999: Multiple Analysis of Series for Homogenization (MASH). In: Second Seminar for Homogenization of Surface Climatological Data (Eds. Szalai S, Szentimrey T, Szinell Cs.) WCDMP 41, WMO-TD 962, WMO, Geneva, 27-46.
Venema V, Mestre O, Aguilar E, Auer I, Guijarro JA, Domonkos P, Vertacnik G, Szentimrey T, Štěpánek P, Zahradnícek 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. *Climate of the Past* **8**: 89-115, doi:10.5194/cp-8-89-2012.
Wang XLL. 2008. Accounting for autocorrelation in detecting mean shifts in climate data series using the penalized maximal t or F test. *J. Appl. Meteor. Climatol*. **47**: 2423–2444.
www. c3.urv.cat/data.html: Software package and Manual of ACMANT2.

APPENDIX I. CHANGES FROM ACMANT1 IN HOMOGENIZING MONTHLY MEANS OF DAILY Tmean OR Tmax.

I-1. Replacing missing data and outliers with interpolated values

In ACMANT2 this routine is applied three times, while in ACMANT1 it was run only twice. There is no change of the parameterization for the first two running of the routine. In the third running, the parameterization is the same as for the second routine.

I-2. Construction of relative time series

Let the candidate series and reference series be denoted by **A** and **F**, respectively, then relative time series (**Q**) are defined as their difference series: **Q = A – F**.

I-2.1. Parameterization

Minimum length of **A**: 10 years
Minimum length of **F**: 10 years (Note: it implies that each **F** must have at least 10 year common section with **A**).
Minimum number of reference composites: 4
Maximum number of reference composites: not limited
Minimum number of months for which observed values are available both in **A** and in a reference composite: 50
Minimum threshold of spatial correlation (r_0): it depends on the number of reference composites (J):

if $J = 4$ then $r_0 = 0.6$
if $J = 5$ then $r_0 = 0.48$
if $J > 5$ then $r_0 = 0.4$

Further restrictions: If the sum of the accumulated weight (w) of reference composite j reaches 4.0 without reference composites with length (n) of shorter than 20 years ($n_j < 20$), then the composites of $n_j < 20$ are excluded. Similarly, if $w \geq 7.0$, then composites of $n_j < 30$ are excluded.

I-2.2. Constructing different relative time series for different sections of the candidate series

(a) First the longest section of **A** that can be homogenized is paired with at least one **F** series in a way that the result **Q** series are the possible longest.

(b) In the second step, for each section of **A**, the **Q** series with the maximal accumulated weight of reference composites is created.

(c) If after (a) and (b) the number of **Q** series for a specific **A** would exceed 80, then phase (b) is repeated in a modified way, namely sections with the highest overlap are not distinguished, thus the number of **Q** series is reduced. It is managed in a way that the time lapse between similar sections is monitored and a parameter indicating the threshold degree of time lapse between sections is increasing gradually from 0 as long as the number of created **Q** series remains under 80.

(d) If the accumulated weight of the reference composites for two **Q** series differs with at least 5%, then both series are retained, irrespectively to the degree of time lapse between sections. This rule overwrite rule (c), but the total number of retained **Q** series cannot be higher than 80.

I-2.3. Unified relative time series

In ACMANT2 it is not used.

I-2.4. Selection of relative time series

There are often more than one relative time series are available for the examination of a specific section of the candidate series. As a principal rule, the **Q** with the highest accumulated weight (w) of the reference composites is used only, but w is modified (w*) according to the length of **Q** (17), since the use of relatively long **Q** series is preferred.

$$w^* = w \log(6 n_{\mathbf{Q}}) \tag{17}$$

Only in the Secondary Detection, all the **Q** series are used in the check of the maximal accumulated anomalies. The rules of the harmonization of section-examinations in the detection on low time resolution are unchanged (D2011).

I-3. Detection on low time resolution

I-3.1. Pre-homogenization

(i) Ranking of time series according to the inhomogeneous character is not applied in ACMANT2.

(ii) Pre-detection is done according to the rules of the detection on low time resolution (sect. 3.2). In the penalty-term of the Caussinus – Lyazrhi criterion an empirical coefficient (p) is applied (eq. 10). In the Pre-detection, this coefficient depends on the accumulated weight of the reference components (18).

$$p = p_1 - \left(\frac{p_2 \log(10w - 7)}{\log(10w)} \right)^{p_3} \tag{18}$$

$$p_1 = 2.0; \quad p_2 = 1.05; \quad p_3 = 17.0$$

(iii) Parameter c_0 (vs. Eq. 7):
In ACMANT2 $c_0 = 5^{-0.5}$ both in the Pre-homogenization and in the Main Detection.

I-3.2. Monthly precision of breaks

The width of the window, in which break is searched in the monthly precision (step III/4 in the algorithm of D2011) is 29 months (it was 25 months in ACMANT1).

I-4. Inhomogeneity detection on short time-scale

I-4.1. Filtering of outlier periods

Outlier periods could also be referred to as short-term inhomogeneities, since their model is a short-term, platform-like bias from the correct values. In this model the bias is constant for the outlier period. Both the detection and the adjustment of outlier periods are more similar to outlier filtering than to the detection and adjustment of long-term biases.

Filtering of outlier periods is applied for 2 - 27 month long periods, always after the routine of common outlier filtering. In switched off outlier filtering mode, the minimum duration of outlier periods is 5 months.

In searching outlier-periods, relative time series (**Q**) are used on monthly scale and the values are transformed to standard anomalies (**B**). Further denotations: l – length of outlier period, h_1 and h_2 – starting and ending months (respectively) of the outlier-period in the first estimation, l_A and l_B are lengths of outer sections of (h_1, h_2) before that and after that, respectively, int and sgn – integer part and sign of arithmetic expression, respectively, mod – function of modulo, λ statistic of significance.

The detection of outlier periods is a step-by-step procedure, since only one outlier-period is identified in a particular step, i.e. the one with the highest λ (19). The mean value of a potential outlier-period is compared with the mean value of the adjacent outer sections in both sides of the potential outlier-period (20). Once an outlier-period has been selected, its values are adjusted to make it possible searching the next most significant outlier-period. Here, temporal adjustments are applied, which are valid during the operations of this routine only. The temporal adjustments eliminate the difference between the means of the outlier-period and its outer sections, and thus the next most significant outlier period can be selected in the next round. The procedure stops when $\lambda < 30$.

The identification of an outlier-period comprises two phases. In the first phase (i), the most significant outlier-period of the time series is selected and a first estimate is made for its position. In the second phase (ii) the starting and ending months of the outlier-period are determined.

Phase (i): the outlier-period with the maximal λ is searched for each h_1, h_2 pairs ($2 \leq h_2 - h_1 < 27$) of standardised relative time series.

$$\lambda = l'^{0.75}\, d^2 \tag{19}$$

where d (magnitude-characteristic) and l' (duration-characteristic) are determined by Eqs. (20) and (21):

$$d = \overline{b_{h_1 - l_A}, b_{h_1 - 1}} + \overline{b_{h_2 + 1}, b_{h_2 + l_B}} - \overline{b_{h_1}, b_{h_2}} \tag{20}$$

$$l' = \mathrm{int}(\max\left\{ l - \frac{0.75}{3.5} \sum_{[h_1, h_2]} c_m, 1 \right\}) \tag{21}$$

Further conditions are that

$$\mathrm{sgn}(\overline{b_{h1}, b_{h_2}} - \overline{b_{h_1 - l_A}, b_{h_1 - 1}}) = \mathrm{sgn}(\overline{b_{h_1}, b_{h_2}} - \overline{b_{h_2 + 1}, b_{h_2 + l_B}}) \tag{22}$$

$$\mathrm{mod}(l_A, 12) = 0 \qquad \mathrm{mod}(l_B, 12) = 0 \tag{23}$$

The usual length of the outer periods is 24 months in both sides of the potential outlier-period. However, if an outlier-period is close to an endpoint of **B**, l_B or l_A can be 12 or even 0. The two outer periods together must contain at least 36 months for providing statistical sample of adequate size for the calculations. For the fulfilment of this condition, if $l_B = 0$ then $l_A = 36$ and if $l_A = 0$ then $l_B = 36$. In (21), the sum of c_m within the outlier period is included in order to take into account the seasonal imbalance of the period. It is necessary, because biases due to breaks seasonally vary, and thus a long-standing bias with enhanced seasonal cycle could be detected as short-term outlier-period when a seasonal peak of the long-term bias and random noise accidentally add up. The sum of c_m is an indicator of the seasonal imbalance and it is normalised with the absolute value of sum of c_m over a half year (i.e. 3.5, see the denominator of the coefficient). The 0.75 in the counter is an empirical constant.

Phase (ii): The first and last months of the outlier-period are re-estimated with fitting optimal step-function in window $[b_{h1-lB}, b_{h2+lA}]$. For longer than 9 month sections harmonic functions are fitted instead of constant values and from this point of view the procedure is the same as the break detection part of Secondary Detection (D2011). Differing from Secondary Detection, solutions with exactly two breaks are accepted only, and the first and second breaks are expected in the periods $[h_1 - 14, h_1 - 1]$ and $[h_2, h_2 + 13]$, respectively. So that, the final duration of an outlier-period is equal or greater than the pre-estimated duration. If h_1 or h_2 coincides with one endpoint of **B**, then one only break is searched, since the other endpoint of the outlier-period is defined by the endpoint of **B**.

I-4.2. Secondary detection

The parameterized penalty term of the Caussinus – Lyazrhi criterion is applied (eqs. 9-10), and here $p = 1.8$. There is no other change relative to ACMANT1.

I-5. Exclusion of detected breaks

I-5.1 Causes of possible exclusions

Detected breaks may be deleted for the following reasons:

i) If all the time series have break at the same time, the equation system of ANOVA is non-determined.

ii) A large number of simultaneous breaks reduce the reliability of the results, since the basic theory of the statistical homogenization is that the break is individual and the other station series are free from break at the same time. If the number of simultaneous breaks approaches to or exceeds the half of the number of time series within network, then there is a high risk that the correct series will be adjusted instead of the biased series.

iii) The inclusion of very close breaks (in time) or breaks with insignificant shift-size reduces the accuracy, since the random error of v' in eq. (13) increases with the shortening of the homogeneous period.

iv) When a break is detected with bivariate detection, it is possible that the shift-size is significant only in one of the variables examined.

v) The calculated shift sizes by ANOVA may indicate that a detected break is not significant statistically, in spite of it seemed to be significant during the detection phase.

In accordance with i), ii) and iii), the number of simultaneous breaks is limited in ACMANT2 even when all breaks seem to be significant, and, on the other hand, breaks with non-significant shift sizes are always deleted. The technical solution of the exclusion of breaks is as follows.

I-5.2. Significance of breaks

For the possible exclusion of one or more breaks, a significance order must be determined. For breaks which are detected in different phases (by different routines) of ACMANT2, the phase determines the order of significance, i.e. breaks detected in later phases of the procedure are considered to be more significant than those that detected in earlier phases, independently from other characteristics of the breaks. As a consequence, breaks of Secondary Detection are more significant than the breaks of Main Detection, the results of Main Detection overwrite the results of Pre-detection and the breaks of Main Detection are more significant than the breaks detected by Filtering of outlier periods.

Sometimes the rank order of significance must be determined for breaks detected by the same routine. As the assessment of significance of break k is limited to the examination of the period between j_{k-1} and j_{k+1} (which includes 1 detected break), the single break model can be applied for these assessments and thus the use of t-test and its modified versions are appropriate here. t-test is applied with the simplification that the standard deviation (σ) is

considered to be constant for a given series, it is because the signal-to-noise ratio is generally too low to estimate specific σ values for individual homogeneous sections with sufficient confidence. With this simplification, the calculation of t-statistic (τ) is given by (24).

$$\tau = \frac{|d|\sqrt{l_1 l_2 (l-2)}}{l\sigma} \tag{24}$$

In (24), $l_1 = j_k - j_{k-1}$, $l_2 = j_{k+1} - j_k$, $l = l_1 + l_2$ and d denotes the shift size. For determining the order of significances only, instead of absolute significances, τ^* (25) can be used instead of τ.

$$\tau^* = \frac{l_1 l_2 d^2}{l} \tag{25}$$

I-5.3. Reduction of the number of synchronous breaks

In ACMANT2, the number of synchronous breaks is not allowed to reach the 50% of the number of time series which are homogenized together at the section including the synchronous break. This rule is valid both for Pre-Homogenization and Main Homogenization. In Pre-homogenization the number of breaks is checked separately for variables E and Z, while in Main Homogenization the break-list is always identical for E and Z until the final filtering of insignificant breaks, for technical reasons. For the limitation of synchronous breaks the least significant breaks are deleted from the break-list when it is necessary. The rank order of significance is determined by the origin of the break or with the calculation of τ^*. If the timings of two breaks have 1 month difference only, then they are considered to be synchronous. In the Main Homogenization, a combination of the shift-sizes of E and Z is considered in d^2 (26).

$$d^2 = d(E)^2 + d(Z)^2 \tag{26}$$

I-5.4. Exclusion of breaks due to too short section of homogeneous period

The minimum distance between two adjacent breaks of the break-list is 5 months. The origin of the breaks is considered for determining and excluding the less significant break, if it is necessary for complying with the rule. Note that Filtering of outlier periods and Secondary Detection can detect shorter IHs than 5 months, but the biases due to such IHs are treated with interpolation and the breaks bordering such IHs are never included in break lists.

I-5.5. Exclusion of insignificant breaks

The significance of breaks is checked by *t*-test, separately for *E* and *Z*, both in Pre-homogenization and Main Homogenization. In Main Homogenization, breaks with τ indicating insignificant shift-size at the 0.05 level are considered to be insignificant. If two adjacent breaks are insignificant, then only one break (with the smaller τ) is excluded in one particular step, thereafter the check of significance is repeated applying the reduced break-list. If more than two sequent breaks are insignificant, then the first and last insignificant breaks are excluded, thereafter the check of significance is repeated.

In Pre-homogenization, only one break can be excluded in one specific step, i.e. the one with the lowest statistical significance. Here, modified *t*-statistics are calculated, i.e. $d\tau$ is examined instead of τ and the threshold statistic is lower (higher) for *E* (*Z*) than the relevant threshold of τ by the multiplier 0.842 (1.786). These modifications are based on test experiments.

I-6. Adjustments before the homogenized section

Relative homogenization often cannot be applied for early sections of time series when the density of observing network is inadequate. However, breaks of the homogenized section might be responsible for biases both within the homogenized section and before that. Thus adjustments can be applied for early sections of time series, even when break detection was not performed for them. Such adjustments will improve the data accuracy when the impact of detected breaks is not overwritten by some undetected breaks in the early sections.

I-6.1. Concepts of treated section and homogenized section

For treated sections the ratio of missing data and the length of data gaps are limited. Observed data out of the treated section do not take part in any calculation and they remain unchanged during the homogenization procedure. The homogenized section can be shorter than the treated section when the number of comparable time series and their spatial correlations are inadequate to create reliable reference series for some periods of the treated section. Each input time series has 0 or 1 treated section including 0 or 1 homogenized section (see their deduction in the Manual). Periods of the treated section out of the homogenized section are not subjected to break detection and outlier filtering, but the data of such periods may take part in gap filling, and they might be adjusted as well.

I-6.2. Deduction of the adjustment terms for data before the homogenized section

In ACMANT2 the mean estimated bias of the first 30 years of the homogenized section is considered to be “persistent bias” and the relevant adjustment is applied for the treated section before the homogenized section. If the length of the homogenized section is shorter than 30 years, then the persistent bias is zero by definition. The change of the adjustment-terms close to the beginning of the homogenized section is gradual to avoid creating seeming breaks due

to the rapid alteration of adjustment terms. For this reason, adjustment terms change linearly in the 3 years before the homogenized section (from the adjustment terms of the first year of the homogenized section to those of the persistent bias). From the fourth years before the homogenized period the adjustment terms of the persistent bias will be applied backwards until the beginning of the treated section.

I-7. Deduction of daily adjustment terms

Monthly adjustment terms equal with the daily adjustment term in a middle day of months, more precisely: on 15 January, 14 February, and on 16 (15) of other months in non-leap years (leap years). These days are named middle days. For any other day of the year the adjustment term is calculated with linear interpolation between the adjustment terms of the two closest middle days.

APPENDIX II. DIFFERENCES IN HOMOGENIZING Tmin RELATIVE TO HOMOGENIZING Tmax OR Tmean

II-1. Detection on low time resolution

In homogenizing Tmin (or homogenizing temperatures of tropical or monsoonal regions) always the univariate detection (formulas 3, 4 and 5) is applied.

In the Main Detection, the coefficient in the modified Caussinus – Lyazrhi criterion (eq. 10) $p = 1.4$.

In the Pre-homogenization eq. (18) is applied for determining p, but with modified parameterization:

$$p_1 = 2.6; \quad p_2 = 1.3; \quad p_3 = 20.0$$

II-2. Monthly precision

Two-phase step function is fitted instead of harmonic functions. There is no change here in the parameterization.

II-3. Filtering of outlier periods

Eqs. (19) and (21) are simplified here to (27), as the modification of l due to seasonal accumulation of biases is not applicable here.

$$\lambda = l^{0.75} d^2 \tag{27}$$

Another change is that in phase ii of this routine always step functions are fitted (and never harmonic functions).

There is no change here in the parameterization.

II-4. Secondary Detection

In searching the most likely break positions around the maximum of accumulated anomalies, always step functions are fitted.

Changes in the parameterization:
Threshold for 5-month accumulated anomalies: 2.15
Threshold for 10-month accumulated anomalies: 1.5
p-coefficient of the modified Caussinus – Lyazrhi criterion: 2.0.

II-5 Calculation of adjustment terms

The biases of *E* only are calculated. In the model of Tmin homogenization the annual cycle of biases is zero, therefore the adjustment term is constant within homogeneous subperiods.