Adapted Caussinus-Mestre Algorithm for Networks of Temperature series (ACMANT)

Peter Domonkos
Centre for Climate Change (C3), Geography Dept., University Rovira i Virgili, Campus Terres de l'Ebre, Tortosa - Spain
E-mail: *peter.domonkos@urv.cat*

Abstract

Any change in technical or environmental conditions of observations may result in bias from the precise values of observed climatic variables. The common name of these biases is inhomogeneity (IH). IHs usually appear in a form of sudden shift or gradual trends in the time series of any variable, and the timing of the shift indicates the date of change in the conditions of observation. The seasonal cycle of radiation intensity often causes marked seasonal cycle in the IHs of observed temperature time series, since a substantial portion of them has direct or indirect connection to radiation changes in the micro-environment of the thermometer. Therefore the magnitudes of temperature IHs tend to be larger in summer than in winter. A new homogenisation method (ACMANT) has recently been developed which treats in a special way the seasonal changes of IH-sizes in temperature time series. The ACMANT is a further development of the Caussinus-Mestre method, that is one of the most effective tool among the known homogenising methods. The ACMANT applies a bivariate test for searching the timings of IHs, the two variables are the annual mean temperature and the amplitude of seasonal temperature-cycle. The ACMANT contains several further innovations whose efficiencies are tested with the benchmark of the COST ES0601 project. The paper describes the properties and the operation of ACMANT and presents some verification results. The results show that the ACMANT has outstandingly high performance. The ACMANT is a recommended method for homogenising networks of monthly temperature time series that observed in mid- or high geographical latitudes, because the harmonic seasonal cycle of IH-size is valid for these time series only.

1. Introduction

For achieving reliable information about climate variability, large amount of observed time series of high quality is needed. One principal tool for improving data quality is the application of statistical homogenisation. In the recent decades large number of homogenisation methods have been developed and applied for homogenising climatic time series. In these methods large scale of statistical tools are applied for detecting change-points or trends of non-climatic origination (*i.e.* inhomogeneities [IH]) in observed meteorological time series. The main tools of IH detection are a) examination of accumulated anomalies [1], b) rank-order statistics [2,3], c) multiple linear regression [4,5], d) t-test based examinations [6,7], e) multiple analysis with Fisher-test [8], f)

fitting step-function [9]. Looking through reviewing articles about homogenisation methods [10-17] it can be seen that we know many details about homogenisation methods, but uncertainties still exist about their efficiencies, or with other words, about the practical usefulness of individual methods.

The recently developed ACMANT homogenisation method has some characteristics those are absolutely new, and may influence positively the effectiveness of homogenisation. The most important innovation in ACMANT is that it uses bi-variable test for detecting IHs in monthly series of temperatures. The two variables are the annual mean temperature and the amplitude of seasonal cycle. Considering that the sizes of IHs in temperature series often have seasonal cycle of considerable amplitude [18,19], the selected two variables often have

change-points in the same time. The chance of right detection increases when the change-points with common time-points are searched with unified test for the two variables. In ACMANT IHs with gradually changing deviation from the correct values are approached by series of change-points, similarly as in many other homogenisation methods.

The paper describes the operation of the whole homogenisation procedure, discusses its general properties, and gives information about its effectiveness. The organisation of the paper is as follow: The next section gives a general picture about the main properties of ACMANT, and describes the conditions of its practical application. Section 3 describes the main functions of ACMANT in detailed. In section 4 the operation of ACMANT is summarised and its steps are presented in the true sequence. In section 5 some verification results are presented. Finally, in section 6, the properties of ACMANT and the verification results are discussed, and conclusions are drawn. The paper has an appendix with the explanation of symbols applied.

2. Main Properties of ACMANT

1) A fully automated method.

2) The use of ACMANT is recommended for temperature series from mid- and high-latitudes, since its algorithm supposes quasi-harmonic annual cycle of considerable amplitude in IH-sizes.

3) Relative homogenisation method, thus it can be used for networks, and not for solely time series. Basically, the ACMANT applies the traditional comparison of candidate series – reference series pairs. Reference series are always built from minimum two component series. A speciality of ACMANT that if the time series of the network have observed values for different time periods, it may use different reference series for different sections of the same candidate series and the selection of reference series is automatic.

4) ACMANT incorporates the best detection and correction algorithms of known homogenisation methods, *i.e.* the detection part is based on the PRODIGE [9] method, while the final correction of time series is made with ANOVA [9].

5) Main novelties of ACMANT relative to PRODIGE: i) It applies the Caussinus – Mestre detection method jointly to two variables, *i.e.* to annual means, and summer-winter differences; ii) Fully automatic generation of reference series, iii) Pre-homogenisation for preparing composites of reference series of better quality than in the raw dataset, iv) Separated way of detection for long-term IHs and short-term IHs; v) After the calculation of correction-terms by ANOVA, those IHs that turn out to be insignificant are deleted from the list of IHs, and the ANOVA procedure is repeated with a reduced list of IHs.

6) Two types of IH-detection are included in ACMANT, *i.e.* the Main Detection is for long-term IHs (generally with at least 3 years duration), while the Secondary Detection is for large-size but short-term IHs. Both detection segments are developed from the Caussinus-Mestre detection algorithm.

7) The lengths of raw time series in a network can be different, and they may cover different time periods. Any candidate time series or a section of candidate series is subdued to homogenisation with ACMANT if at least two partner time series exist in the network that cover the period of the candidate series or its section and have at least 0.5 autocorrelation with the candidate series.

8) Time series often contain large number of missing values. If the ratio of missing values exceeds some preset thresholds in some section(s) of the time series that sections are classified non-applicable sections, while the section with acceptable ratio of available data is the applicable section. According to this, raw time series are generally split into three sections, *i.e.* one applicable section and two non-applicable sections before and after the applicable section. If the ratio of missing values is acceptably low in the tails of the time series, there are no non-applicable sections in the time series. The minimum ratio of available data is 25% in the first k years and last k years of the applicable section, for any k, $k = \{1,2,…15\}$, as well as the minimum ratio is 16,7% for any 30-year subsection of the applicable section. A time series must not contain more than one applicable section (i.e. the applicable section cannot be separated into two parts by a long pause of observation). The minimum length of applicable section is 30 years.

9) The ACMANT has own segments for filling missing data and for substituting outlier values. For this purpose spatial interpolation is applied. However, data of non-applicable sections are not treated, and they do not used at all during the homogenisation procedure.

10) Input data-field for ACMANT: Monthly temperature characteristics with monthly time resolution. The lengths of time series may be different, but the data-fields of each series are required to be converted into a common format (which format includes the same number of data for each temperature series) in a way that missing values are filled with -999.9. After the preparation only 4 parameters have to be introduced before application: a) Length of time series, b) First year of time series, c) Number of time series in the network, d) Identifier of network.

11) The result of homogenisation is a) Timings and sizes of inhomogeneities (IHs) for each series. b) Tim-

ings of outliers. c) Filled data-gaps caused by missing values or outliers inside the series. d) Homogenised time series. Sizes of IHs are characterised with two variables: a) shift in annual means, b) shift in the amplitude of seasonal cycle.

3. Main Functions of ACMANT

3.1. Basic definitions

Before the description of the operation of ACMANT, definitions of some basic statistical concepts are presented.

Time average of series **X** (denoted with upper stroke):

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (1)$$

Standard deviation of series **X**:

$$\sigma(X) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2} \qquad (2)$$

Time average of section $[j_1, j_2]$ of series **X**:

$$\overline{X_{[j1,j2]}} = \frac{1}{j_2 - j_1 + 1}\sum_{i=j_1}^{j_2} x_i \qquad (3)$$

Note that standard deviation for selected section of time series is defined and denoted with the same logic as section-average by Eqution (3).

Derivation of anomaly for station s, year j and month m:

$$a_{s,j,m} = x_{s,j,m} - \frac{1}{n'_{s,m}}\sum_{i=1}^{n_s} x_{s,i,m} \qquad (4)$$

Missing values are represented with $x_{s,i,m} = 0$, while n' stands for the number of available observed values in **X$_{s,m}$** in Eqution (4). Note that due to missing values the simple time-average cannot be used in Eqution (4).

ACMANT counts with anomalies during its whole procedure, only in the last step the climatic means (the second term in the right hand of Eqution (4) are added back to the homogenised anomaly series.

Spatial correlation:

$$r'_{g,s}$$

$$= \frac{\frac{1}{n'_{g,s}}\sum_{h=1}^{12n_{gs}} a_{g,h}\cdot a_{s,h} - \left(\frac{1}{n'_{g,s}}\sum_{h=hn_{gs}}^{hx_{gs}} a_{g,h}\right)\cdot\left(\frac{1}{n'_{g,s}}\sum_{h=hn_{gs}}^{hx_{gs}} a_{s,h}\right)}{\sigma\left(A_{g[hn_{gs},hx_{gs}]}\right)\sigma\left(A_{s[hn_{gs},hx_{gs}]}\right)}$$
$$(5)$$

In Eqution (5) g and s denote stations, h is the serial number of month from the beginning of time series (here time has one dimension only), hn_{gs}, hx_{gs} and n'_{gs} denote

the first month, the last months and total number of months, for which observed data exist in both series, respectively. Missing values of **A$_g$** and **A$_s$** are represented with 0 in Eqution (5).

3.2. Filling the Gaps of Time Series

This operation can be considered as one step of the preparation of time series, because further segments of ACMANT require continuous data series. However, on the other hand, this operation is part of the ACMANT, and it is performed automatically.

The interpolation for a missing value of month h in the candidate series (**A$_g$**) relies on the same date values of surrounding stations. All the time series of minimum 0.4 spatial correlation with the candidate series are used in the interpolation, if they have observed value for month h. The interpolated value is a weighted average of the anomalies of the partner time series in h. For this interpolation, section-anomalies ($a_{[h1,h2]}$) are calculated for the symmetric window $(h_1,h_2]$ around h. The number of observed values used (n') is usually 100 or 101, i.e. the procedure search h_1 and h_2 for which $h_2 - h = h - h_1$, and $n' \varepsilon [100,101]$. However, when the window is 2x10 years wide, i.e. $h_2 - h_1 = 240$ and n' is still smaller than 100, $n' \varepsilon [30,99]$ is accepted. In case of $h_2 - h_1 = 240$, and $n' < 30$, the window-width widens further, and it stops when $n' \varepsilon [100,101]$, or when $h_2 - h_1 = 360$. Eqution (6) shows the calculation of section-anomalies for series s.

$$a_{s,h[h1h2]} = a_{s,h} - \frac{1}{n'}\sum_{i=h_1}^{h_2} a_{s,i} \qquad (6)$$

Missing values of **A$_s$** are represented with zero in Eqution (6).

The interpolated value is the weighted average anomaly of N^* surrounding series which is added to the period-average of **A$_g$**. The weights are the squared spatial correlations. When the sum of squared correlations is lower than a fixed threshold (0.64), zero anomaly $a_{g,h[h1,h2]} = 0$ is presumed with a certain or entire weight, according to Equtions (7) and (8).

$$a_{g,h} = \frac{1}{p}\sum_{s=1}^{N^*} r'^2_{g,s} a_{s,h[h1,h2]} + \frac{1}{n'}\sum_{i=h_1}^{h_2} a_{g,i}, \qquad (7)$$

$$p = \max\left(0.64, \sum_{s=1}^{N^*} r'^2_{g,s}\right) \qquad (8)$$

The gap-filling has to be accomplished before the other steps in ACMANT. However, the first estimations are not final, after some steps of pre-homogenisation the interpolation is repeated in the same way as it is presented here, but with the use of data of higher quality. There is one difference in the second round of the inter-

polation process, namely $n' < 100$ (more exactly: $n' \varepsilon$ [30,99]) for window $(h_1,h_2]$ is expected only when $h_2 - h_1 = 360$.

3.3. Constructing Relative Time Series

3.3.1. General Rules of Constructing Relative Time Series in ACMANT

ACMANT makes relative homogenisation which relies on the spatial comparison of time series. The way of this comparison basically follows the traditional rules introduced by [20], and applied later widely [21,23]. The relative time series are the arithmetical differences of the candidate series, and the so-called reference series (**F**) (Eqution 9).

$$T_{g,q} = A_{g[j_1,j_2]} - F_{g[j_1,j_2]} \qquad (9)$$

In Eqution (9) a section is defined for which the relative time series is calculated. This section cannot stretch out the ends of any reference-composites, or the ends of the candidate series.

Reference series are built from the composition of neighbouring series around the candidate series. The weights of individual composites depend on the spatial correlations between first difference (increment) series. In studies about homogenisation methods one can find different recommendations about the usefulness of first difference series, and about the optimal number of reference-components. For ACMANT the use of first difference series was selected (Eqution (10)) for evaluating spatial correlations, because possible large IHs might affect more seriously the spatial correlations of raw time series (**R'**), than the spatial correlations of first difference series (**R**).

$$\Delta a_{s,h} = a_{s,h+1} - a_{s,h} \quad \text{for} \quad \forall h \in (1, \; 12n_s - 1) \quad (10)$$

Spatial correlations for $\Delta \mathbf{A}$ are calculated according to Eqution (5) with the only difference from that, that at this stage there are no missing values in the series, thus instead of n' the length of the period examined must be applied. Denoting with $b_{g,s,h}$ the product of $\Delta a_{g,h}$ and $\Delta a_{s,h}$, the equation can be written in a simpler form (Eqution (11)).

$$r_{g,s} = \frac{\overline{B_{g,s\left[hn_{gs},hx_{gs}\right]}} - \overline{A_{g\left[hn_{gs},hx_{gs}\right]}} \cdot \overline{A_{s\left[hn_{gs},hx_{gs}\right]}}}{\sigma\left(A_{g\left[hn_{gs},hx_{gs}\right]}\right)\sigma\left(A_{s\left[hn_{gs},hx_{gs}\right]}\right)} \quad (11)$$

For determining the number of reference-components (S), minimum thresholds for acceptable spatial correlations are introduced. In the present version of ACMANT this threshold is relatively low, generally $r \geq 0.4$, although for at least two of the components it has to be at least 0.5. The application of the relatively low threshold

values relies on the outcome of some experiments according to the use of a large number of reference components usually results in good verification results, even if the spatial correlations of some components are relatively low. Its explanation is that the increase of reference components tends to reduce the mean effect of IHs and noise in the reference components.

The acceptable minimum of S is 2, in the reverse case homogenisation cannot be fulfilled with ACMANT. Note that one pre-selected time series is often excluded from the reference composites, it is because during the pre-homogenisation the candidate series for which the pre-homogenised reference composites will be used must not be taken into account in any form to avoid non-desired effects of a multiplied use of the same information for the same pieces of data in the homogenisation procedure. For this reason it may occur that the number of reference composites is not more than 1.

The composition of reference series for a predetermined section $[j_1,j_2]$ of series $\mathbf{A_g}$ is presented by Eqution (12).

$$F_{g[j_1,j_2]} = \frac{\sum_{s=1}^{S} r_{g,s}^2 A_{s[j_1,j_2]}}{w_{g[j_1,j_2]}} \qquad (12)$$

In Eqution (12) w stands for the total weight of the reference those composites that have acceptable spatial correlation with the candidate series, and available data in section $[j_1,j_2]$ as well (Eqution (13)).

$$w_{g[j_1,j_2]} = \sum_{s=1}^{S} r_{g,s}^2 \qquad (13)$$

3.3.2. Application of Homogenisation-Adjustment for Components of Relative Time Series

When relative time series are created, the candidate series are always raw or outlier-filtered time series and homogenisation-adjustments have not been applied earlier for them, while for reference-composites, homogenisation-adjustments are usually applied if adjustment factors are available for them at the contemporary phase of the procedure. Note that in the description of the algorithm (Sect. 4), certain deviations from these rules will be mentioned.

3.3.3. Constructing Different Relative Time Series for Different Sections of the Candidate Series

For sections of the candidate series distinct reference series are often built when the number of available series with adequately high spatial correlation is larger for a section, than for the entire series. In this way usually more than one relative time series are produced for one candidate series.

Considering that more than one relative time series

can be constructed for one candidate series, we apply an index (q) for denoting the individual reference series for the same candidate series, while index g will be omitted hereafter (since the candidate series is fixed in this examination). Generally Q reference series belong to one candidate series, they starts at year $y_{1,1}, y_{2,1}, \ldots y_{Q,1}$, while their last years are $y_{1,2}, y_{2,2}, \ldots y_{Q,2}$. Note that the minimum length of reference series ($y_{q,2} - y_{q,1} + 1$) is 30 years, and as reference series never extend over the borders of their candidate series, **Y** also marks the borders of relative time series.

The determination of the set of reference series has four main phases. In phase 1 three time series are determined, i) the one with the highest w (**F**$_{opt}$) ii) the one with the earliest $y_{q,1}$ ($y_{1,1}$) and iii) the one with the latest $y_{q,2}$. Note that **F**$_{opt}$ may have the earliest $y_{q,1}$ and/or the latest $y_{q,2}$, thus the final number of reference series originated from this phase can be lower than 3.

In phase 2 potential reference series are examined whose starting year ($y_{q,1}$) is after $y_{1,1}$, but earlier than $y_{opt,1}$. Obviously, when $y_{opt,1} - y_{1,1} < 2$, this phase is omitted. During this examination each possible $y_{q,1}$ is examined, proceeding step-by-step from $y_{1,1} + 1$ until $y_{opt,1} - 1$. Two parameters (p_1 and p_2) are monitored during this examination (Equtions (14) and (15)).

$$p_1 = y_{q,1} - y_{q-1,1} \qquad (14)$$

$$p_2 = \frac{w_q}{w_{q-1}} \qquad (15)$$

A new reference series is selected when a) $p_2 \geq 1.3$, or b) $p_1 \geq 5$ and $p_2 \geq 1.1$, or c) $p_1 \geq 10$ and $p_2 \geq 1.03$, or d) $p_1 = 30$. Once a reference series is selected, the examinations are continued recursively by examining the potential starting years between $y_{q,1}$ and $y_{opt,1}$. It follows from the written rules that a new reference series is selected with 30 years distance in the starting years at latest. However, when the selection is based on condition d) ($p_1 = 30$), two special cases need more detailed description,: i) It may occur that the available reference composites for $y_{q-1,1} + 30$ are the same as for $y_{q-1,1}$. In this case, although there is no new reference series, the procedure continues with the virtual $y_{q,1}$ that equals with $y_{q-1,1} + 30$. ii) Although w usually increases between $y_{1,1}$ and $y_{opt,1}$, p_2 of lower than 1 may occur for a 30-year subsection. In this case, the maximal p_2 is searched for between $y_{q-1,1} + 1$ and $y_{q-1,1} + 30$, and the corresponding reference series is selected. Especially, it may occur (very rarely for observed climatic time series) that even the maximal p_2 does not satisfy the minimum conditions for creating reference series, and a discrete subsection of $[y_{q-1,1} + 30, y_{opt,1}]$ cannot be subdued to homogenisation, meanwhile other sections before and after that subsection can be.

In phase 3 the symmetric procedure is applied for potential reference series with ending years between $y_{opt,2}$ and $y_{q(latest),2}$ than in phase 2, proceeding backwards, step-by-step from $y_{q(latest),2} - 1$ until $y_{opt,2} + 1$.

In phase 4 the selected reference series are ordered according to w, and multiple selections of the same reference series are excluded.

3.3.4. Unified Relative Time Series

We introduce the concept of unified relative time series (**T**$^+$). It will be used for calculating temporary corrections and filtering outliers in the section of pre-homogenisation. **T**$^+$ is a combination of **T$_q$** series. The reasoning of its introduction is empirical, *i.e.* adjustment factors that are derived from different relative time series, often result in relatively large artificial biases in the low frequency variability of the adjusted time series, even if the individual estimations of change-point effects are relatively good. The concept of the unified relative time series exploits the fact, that a relative time series can be modified with an arbitrary constant without any effect on the estimation of adjustment factors.

The set of relative time series is ordered according to the decreasing values of w belonging to the relevant reference series, then the **T** series are examined one-by-one following this order. The values of **T**$^+$ for year j are determined when the condition of **T$_q$** includes $t_{q,j}$ is satisfied first. In the calculation of values for **T**$^+$ the relevant values of **T$_q$** are usually adjusted with the mean difference between **T**$^+$ and **T$_q$** according to Equtions (16)-(18). When j lies before the section(s) for which **T**$^+$ has values determined from the previously examined **T** series, Eqution (16) is applied, while Eqution (17) is applied in the opposite case.

$$t_j^+ = t_{q,j} + \frac{1}{p} \sum_{i=Y_1^*}^{Y_1^* + p - 1} \left(t_i^+ - t_{q,i} \right) \qquad (16)$$

$$t_j^+ = t_{q,j} + \frac{1}{p} \sum_{i=Y_2^* - p + 1}^{Y_2^*} \left(t_i^+ - t_{q,i} \right) \qquad (17)$$

where

$$\begin{aligned} Y_1^* &= \min \left\{ y_{1,1}, y_{2,1}, \cdots, y_{q-1,1} \right\} \\ Y_2^* &= \max \left\{ y_{1,2}, y_{2,2}, \cdots, y_{q-1,2} \right\} \end{aligned} \qquad (18)$$

Notes:

1. Month-indexes for t and t^+ are omitted from Eqs. 16 and 17, first for simplicity, and secondly because often some annual characteristics are used in the homogenisation procedure instead of monthly values (see Sect. 3.4.1).

2. Value p is preset before the calculation of unified

time series. The value is relatively low ($p = 5$) in the beginning of the homogenisation procedure, but with the advance of the homogenisation it is higher (up to 30). The philosophy of limiting p stems from the fact that unadjusted change-points may affect the apparent difference of $\mathbf{T^+}$ and $\mathbf{T_q}$, therefore sections that are relatively close to the newly determined values are used only. For the very same reason p is lower in the beginning of the homogenisation procedure, and with the decreasing influence of unadjusted change-points p increases.

3. The p applied can be lower than its predefined value when the number of available value-pairs of $\mathbf{T^+}$ and $\mathbf{T_q}$ is lower than the predefined value of p. When the number of applicable p is lower than 3, adjustment is not made, *i.e.* in this case $t_j^+ = t_{q,j}$. It is always the case for $q = 1$.

4. It is seldom, but may occur that previously determined values of $\mathbf{T^+}$ exist on both sides of the newly determined values. Introducing Y_1' for the starting year of the right section and Y_2' for the ending year of the left section, and substituting Y_1^* and Y_2^* in Equtions (16) and (17) with them, the estimations of the adjustment factors included in the referred formulas are calculated first and they are denoted with E_1 and E_2, respectively. Then a linear transition between E_2 and E_1 provides the adjustment factors for the values between Y_2' and Y_1' (Eqution (19)).

$$ t_j^+ = t_{q,j} + E_2 \frac{Y_1' - j}{Y_1' - Y_2'} + E_1 \frac{j - Y_2'}{Y_1' - Y_2'} \qquad (19) $$

3.4. Detecting IHs with the Main Detection

3.4.1. Detection Process

In the Main Detection the timings and sizes of IHs are searched with fitting step-functions to two variables, namely to the annual means (**TM**) and summer-winter differences (**TD**) in relative time series (Equtions (20) and (21)). In the following description the index q of relative time series is omitted.

$$ tm_j = \frac{\sum_{m=1}^{12} t_{j,m}}{12} \qquad (20) $$

$$ td_j = \frac{t_{j,5} + t_{j,6} + t_{j,7} + 0.5t_{j,8} - t_{j,11} - t_{j,12} - t_{j,1} - 0.5t_{j,2}}{3.5} \qquad (21) $$

In Eqution (21) the second index represents calendar month.

Solutions with common timings of steps are considered only, and the minimum sum of squared errors is searched in a similar way as it is described by [24], [9], etc. (Equtions (22) and (23)).

$$ \min_{[j_1, j_2, \ldots j_{K'}]} \left\{ \sum_{k=0}^{K'} \sum_{i=j_k+1}^{j_{k+1}} \left(tm_i - \overline{TM_k} \right)^2 + c_0^2 \left(td_i - \overline{TD_k} \right)^2 \right\} \qquad (22) $$

$$ j_0 = 0 , \quad j_{K'+1} = L \qquad (23) $$

L denotes the length of the period examined, in years, c_0 is constant, while the number of change-points for the selected period is K'. k characterises not only the serial number of change-points, but the serial number of sections between adjacent change-points too (Eqution (24)).

$$ \overline{TM_k} = \overline{TM_{[j_k+1, j_{k+1}]}} \qquad (24) $$

c_0 represents the estimated significance of changes of TD in comparison to that of TM in relative time series. In its estimation not only the mean sizes, but the signal/noise rate also had to be taken into account. In the present application $c_0 = 2^{-0.5}$.

For selecting the most appropriate K', the Caussinus – Lyazrhi criterion [25] is applied (Equtions (25) and (26)).

$$ \ln \left\{ 1 - \frac{\sum_{k=0}^{K'} (j_{k+1} - j_k) \cdot \left[(\overline{TM_k} - \overline{TM})^2 + c_0^2 (\overline{TD_k} - \overline{TD})^2 \right]}{\sum_{i=1}^{L} (tm_i - \overline{TM})^2 + c_0^2 (td_i - \overline{TD})^2} \right\} $$
$$ + G \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (25) $$

$$ G = \frac{2K'}{L-1} \ln(L) \qquad (26) $$

The Main Detection differs in two points from the classic Caussinus-Mestre detection method:

1) Step functions are fitted to two variables,

2) The minimum distance between two change-points is 3 time-units:

$$ j_{k+1} - j_k \geq 3 \qquad \forall k \in \{0 \leq k \leq K'\} \qquad (27) $$

3.4.2. Selection of Relative Time Series

Q different relative time series ($\mathbf{T_1}, \mathbf{T_2}, \ldots \mathbf{T_Q}$) are derived for one candidate series, and they are ordered according to the sum of squared correlations of the reference-composites ($w_1 > w_2 > \ldots w_Q$). Each of them is used in the Main Detection, but frequently some sections of the relative time series are used only. The algorithm of the relative time series selection is as follows:

1. First the $\mathbf{T_1}$ series is used with its whole length. In this step section $[y_{1,1}, y_{1,2}]$ of the candidate series is homogenised.

2. When the first q relative time series has already been used, $\mathbf{T_{q+1}}$ is applied for sections that a) lie within $[y_{q+1,1}, y_{q+1,2}]$ and b) have not been homogenised with the first q relative time series.

3. When $\mathbf{T_{q+1}}$ is applied, the tails of the sections homogenised with the previous $\mathbf{T}$ series are often over-

lapped. It means that when $a_{g,e1}$ belongs to a section that has been homogenised previously, but $a_{g,e1-1}$ has not, and $e_1 > y_{q+1,1}$, a d_1-year long section after e_1-1 is subdued again homogenisation. The usual length of this overlap is 9 years, but an overlapping section is not allowed to extend over a) IHs detected in previous steps, b) ends of $\mathbf{T_{q+1}}$ (i.e. $y_{q+1,1}$ and $y_{q+1,2}$). Let the first IH after e_1 be denoted by k_1 then the length of overlap is determined by Eqution (28).

$$ d_1 = \min\left\{9, k_1 - e_1 + 1, y_{q+1,2} - e_1 + 1\right\} \qquad (28) $$

For tails that lie in the other ends of the previously homogenised sections (i.e. whose last point is e_2), the overlap (d_2) is calculated with the same logic (Eqution (29)).

$$ d_2 = \min\left\{9, e_2 - k_2, e_2 - y_{q+1,1} + 1\right\} \qquad (29) $$

In Eqution (29) k_2 stands for the first IH before e_2.

3.5. Secondary Detection

In the Secondary Detection short-term IHs are searched in relative time series adjusted according to the results of the Main Detection. This operation is performed only when the maximum of accumulated anomalies in adjusted relative time series exceeds some predefined thresholds. Adjusted relative time series and the timing of the maximum of accumulated anomalies are denoted with $\mathbf{T^*}$ and H^*, respectively. $\mathbf{T^*}$ is examined in its monthly resolution, and the section that is selected around H^* is not allowed to be longer than 60 months.

This section has two sub-sections, namely the search of the maximum of accumulated anomalies and the detection of IHs around H^*.

3.5.1. Search of the Maximum of Accumulated Anomalies
5-month and 10-month moving averages are calculated for the normalised anomalies of $\mathbf{T_q^*}$. All the relative time series of the candidate series are used ($q = 1,2,\dots Q$). In eqs. (30) and (31) the calculation of accumulated anomalies is shown for 5-month and 10-month periods, respectively.

$$ \mathrm{MA}5\left(b_{q,h}\right) = \frac{1}{5}\sum_{i=h-2}^{h+2}\frac{t_{q,h}^* - \overline{\boldsymbol{T}_q^*}}{\sigma(\boldsymbol{T}_q^*)} \qquad (30) $$

$$ \mathrm{MA}10\left(b_{q,h}\right) = \frac{1}{10}\sum_{i=h-5}^{h+4}\frac{t_{q,h}^* - \overline{\boldsymbol{T}_q^*}}{\sigma(\boldsymbol{T}_q^*)} \qquad (31) $$

Note that in Equtions (30) and (31) i is not allowed to be lower than 1 or higher than the number of months (nm) in $\mathbf{T_q^*}$, therefore the conditions of $3 \leq h \leq nm$-2 and $6 \leq h$

$\leq nm$-4 have to be satisfied for Equtions (30) and (31), respectively.

After having the accumulated anomalies (MA5(b) and MA10(b)) for each $\mathbf{T_q^*}$, their maximums are determined. The maximums are calculated without sorting the accumulated anomalies according to q, and in this way only two maximal values are obtained, one for MA5 and another for MA10. In the present version, Secondary Detection is made only when max(MA5(b)) $\geq$ 2.0 or max(MA10(b)) $\geq$ 1.4. When exists a H^* for which the former relation is satisfied, the timing of the maximum of 5-month anomalies is used, while the timing of the maximum of 10-month anomalies is used when only the latter relation is satisfied.

3.5.2. Detection of IHs around the Maximum of Accumulated Anomalies

a) Selection of relative time series
A window is edited around H^* and for that operation available data are needed in both sides of H^*. Usually the $\mathbf{T_q^*}$ of the highest w is used for which at least 20 data are available in each side of H^*. When non of the $\mathbf{T_q^*}$ series meet with this condition (because H^* is close to one end of the candidate series), all the $\mathbf{T_q^*}$ series are examined again, with less strict conditions. In the second round the expected minimum of available data (in each side of H^*) is 10, and in the third round (if that is necessary) the minimum threshold is 2 only.

b) Edition of window around H^*
Usually a symmetric window $[h_1,h_2]$ of 60 months length is edited around H^*, but the window must not overlap IHs detected in earlier steps or any end of the $\mathbf{T_q}$ series, therefore it can be narrower than 60 months. Let suppose that K_1 change-points has been detected in the earlier steps of the homogenisation procedure for section $[1,H^*]$ of the series, then the borders of the window are determined by Equtions (32) and (33).

$$ h_1 = \max\left\{H^* - 29, H_{K_1} + 1, 1\right\} \qquad (32) $$

$$ h_2 = \min\left\{H^* + 30, H_{K_1+1}, nm\right\} \qquad (33) $$

c) Detection of IHs in windows
The base of the detection process is again the Caussinus – Mestre method. However, in the Secondary Detection the constant sections of step-function are substituted with harmonic functions of annual cycle for sections of longer than 9 months, and some other modifications also exist relative to the original Caussinus – Mestre method. The list of differences from the original method is as follows.

1) Constant sections are substituted with harmonic function of annual cycle for sections longer than nine

months;

2) The minimum distance between two change-points is 3 time-units, *i.e.* 3 months in this case, but Eqution (26) is applicable here otherwise;

3) Maximum two change-points are allowed to be detected in a window $[h_1,h_2]$. ($K' = 0, 1$ or 2)

Describing more detailed point i), Eqution (34) is applied for fitting harmonic function of annual cycle.

$$u'_h = \alpha'_k + \beta'_k \left(c_A \sin \frac{2\pi m}{12} + c_B \cos \frac{2\pi m}{12} \right), \quad (34)$$
$$h \in [H_k + 1, H_{k+1}]$$

c_A and c_B are general constants in the procedure, because they characterise the relation between the phase of the annual cycle and calendar month. Their values are set to have the modus of the harmonic functions at the solstices ($c_A = -0.1031$, $c_B = -0.9947$). Between two adjacent change-points, α_k' and β_k' are also constants, and their most proper values are searched with iteration. For finding the best fitting model (K', as well as the most applicable timings of change-points), the function fitting of Eqution (34) must be applied for each subperiod of the examined window, then the model with the lowest Caussinus – Lyazrhi score is selected. This operation contains large number of calculations, but by applying an economic algorithm, the computation time is kept fairly short.

In case of $\beta' = 0$ the function is constant and this fact shows that the step-function is a special case of the function family applied here.

3.6. Calculation of Adjustment Factors

3.6.1. ANOVA
The interaction of IHs of individual time series makes it necessary to calculate the precise adjustment-terms by an equation system. In ACMANT the ANOVA method is applied. In brief, the application of ANOVA to calculate adjustment-terms for correcting IHs, is based on the division of observed time series to three components, namely to climate effect, stations effect (IH) and noise. In [9] it is proved that the optimal estimation of adjustment-terms is provided when the noise term is set to be zero in the equation system, thus the ANOVA provides the optimal estimation of adjustment-terms. The referred study also contains the detailed description of the application of ANOVA in the homogenisation of climatic time series.

In ACMANT the ANOVA operates within an automatic procedure, therefore a special attention is needed to treat cases when the equation system has no deterministic solution. It could occur when all the time series of a network (those that are comparable in the homogenisa-

tion procedure) have a change-point in the same time. In that case the behaviour of sections before and after the common change-point is independent. To avoid this case the smallest IH is cancelled if all time series show IH at the same time.

In ACMANT the ANOVA is applied separately for two annual variables (i.e. for **TM** and **TD**), then the monthly correction-terms are derived from them (see Sect. 3.6.3). The ANOVA is applied after the Main Detection, then it is repeated after the Secondary Detection, and if the number of IHs is reduced at the end of the procedure relying on posterior tests, the application of ANOVA is repeated again. However, ANOVA is not applied during the pre-homogenisation, because it is not a tool for making step-by-step improvements in the dataset.

3.6.2. Homogenisation-Adjustments during the Pre-Homogenisation
After having the timings of IHs identified, adjustment-terms for executing homogenisation can be deduced directly from relative time series. Although a part of the detected biases can be caused by the impreciseness of reference series, adjustment-terms are always applied with its full content to the candidate series.

The unified relative time series ($\mathbf{T}^+$) are used for calculating temporary adjustments. Let suppose that for IH $k*$ Equtions (35) and (36) show the shifts in TM and TD.

$$\alpha_{k*} = \overline{TM^+_{k*+1}} - \overline{TM^+_{k*}} \quad (35)$$

$$\beta_{k*} = \overline{TD^+_{k*+1}} - \overline{TD^+_{k*}} \quad (36)$$

If the number of detected IHs is K, the cumulated effect of IHs on the candidate series in year i, $i \leq k*$ is characterised by Equtions (37) and (38).

$$\alpha_i = \sum_{k=k*}^{K} \left(\overline{TM_{k+1}} - \overline{TM_k} \right) \quad (37)$$

$$\beta_i = \sum_{k=k*}^{K} \left(\overline{TD_{k+1}} - \overline{TD_k} \right) \quad (38)$$

3.6.3. Derivation of Monthly Correction-Terms
If IH $k*$ has the timing $H(k*)$ in monthly scale and $\mathbf{U}$ denotes the adjustment-term, it can be given for any year i and month m within the period $[H(k*-1)+1,H(k*)]$ by Equation (39). Corrections from β are distributed among the calendar months in a way that the annual cycle is a harmonic function with its extreme values in the solstices, and the degree of changes in *td* satisfies Equation (21). These conditions determine the monthly constants (c_m) in Equation (39).

$$u_{i,m} = \sum_{k=k*}^{K} \alpha_k + c_m \beta_k \quad (39)$$

3.7. Outlier-Filtering

Outlier filtering is applied two times in ACMANT, namely for raw time series first, then after some steps of the pre-homogenisation it is applied again. The applied method is the same in both cases, only one little difference will be mentioned below.

Two operations are performed in this step.

1) Anomalies with higher than 4 standard deviations are filtered according to Eqution (40).

$$\left| t_{s,q,h} - \overline{T_{s,q}} \right| > 4\sigma\left(T_{s,q}\right) \qquad (40)$$

For each series s and month h always the $\mathbf{T_{s,q}}$ of the highest w is selected which contains h. Note that in the first outlier-filtering the mean of $\mathbf{T_{s,q}}$ is considered to be 0, because for differences of raw anomaly series the expected value is 0.

2) If in a 10-month long period, more than one outliers of the same sign occur according to i), a confirmation is needed, because the accumulation of seeming outliers might be caused by large long-term variability. Therefore in this case a second operation is accomplished, in which potential outliers are examined with the statistical properties of the time-neighbourhood. Nineteen-month long windows are used for this purpose. The potential outlier is confirmed when the deviation from the window-average is larger than 3.5 standard deviation of the values within the window, but is not considered to be outlier in the reverse case (Equation (41)). For this operation, again the relative time series of the highest w, but with available data around h is selected.

$$\left| t_{q,h} - \overline{T_{q[h-9,h+9]}} \right| > 3.5\sigma\left(T_{q[h-9,h+9]}\right) \qquad (41)$$

Note that when the number of available data is less than 9 in one side of the window, this calculation is made in a narrower window. In this case, if another $\mathbf{T_{s,q}}$ series (of lower w) contains data for a full-size window, the calculation is repeated with that data, and the results are overwritten by that.

4. The Algorithm of ACMANT Homogenisation Method

After the main functions of ACMANT have been described, in this section the operation of the whole procedure is presented.

In the homogenisation procedure raw time series are converted several times (by interpolation, outlier-filtering, pre-homogenisation), thus they go through several stages before achieving their final form. During the procedure some earlier stages are preserved, and reused. To make distinctions among different stages of the same series stage-codes are introduced. The code of raw time series is 0, and if the series does not have sufficient spatial correlations with other time series it might remain unchanged during the homogenisation procedure of the network. However, the gap-filling is done even for these series, thus the code 0 indicates raw but continuous time series. The meaning of the codes:

0 – raw
1 – outlier-filtered
2 – two times outlier-filtered
3 – pre-homogenised without outlier filtering
4 – pre-homogenised and outlier-filtered
4a – one time series is excluded during the pre-homogenisation
5 – pre-homogenised and two times outlier filtered
5a – one time series is excluded during the pre-homogenisation
6 – homogenised

The algorithm of ACMANT
Part I: Preparation
   1. Calculation of anomalies (Eqution (4)).
   2. Calculation of spatial correlations (Eqution (5)).
   3. Filling of data gaps in each time series (Sect. 3.2). The results are the series of code 0.
   4. Calculation of first difference series (Eqution (10)) and their spatial correlations (Eqution (11)).
   5. Creation of relative time series from series-0 (Sects. 3.3.1 and 3.3.3).
   6. Outlier-filtering for each time series (Sect. 3.7), the results are series-1.
   7. Spatial correlations are recalculated from series-1.
   8. Interpolation for substituting outliers for each time series (Sect. 3.2). When an outlier was detected in a series g in the same year and month at which interpolation occurred in another series ($s \neq g$) at step 3, its value is re-interpolated at this step.
   Part II: Pre-homogenisation
   1. Calculation of first difference series and their spatial correlations from series-1.
   2. Creation of relative time series (Sects. 3.3.1, 3.3.3) and unified relative time series (in Sect. 3.3.4 in Eqs. 16 and 17 $p = 5$). Type of candidate series: 1. Type of reference components: 1.
   3. Use of the Main Detection (Sect. 3.4) for ranking time series according to the degree of inhomogeneous character. Let the timings of IH (k) be denoted by j_k, the mean estimated bias of $[j_k+1,j_{k+1}]$ by u_k', then an indicator (z) of the degree of inhomogeneous character for section $[j_{k1}+1,j_{k2}]$ is calculated by Eqution (42).

$$z = \frac{\left(\sum_{k=k_1}^{k_2}\left(j_{k+1} - j_k\right)\cdot u'\right)^2}{\left(j_{k2+1} - j_{k1}\right)^{1.2}} \qquad (42)$$

Maximums of z values are searched examining each possible $k_1 - k_2$ pair $(0 \leq k_1 \leq k_2 \leq K)$, then time series are ordered according to their maximal z values. The actual form of Eqution 42 was chosen after some experiments regarding the efficiency achievable. – In this step multiple relative time series are used for determining the timings of IHs (as usual), but the $\mathbf{T}^+$ is used for calculating indicator z.

Steps 4 - 7. compose a block that is accomplished for each series in the order determined in step 3.

4. Calculation of relative time series. Type of candidate series: 1. Type of reference composites: 4a when the composite has already bean pre-homogenised, 1 otherwise.

5. Main Detection.

6. Calculation of relative time series and $\mathbf{T}^+$ with the exclusion of one of the possible composites, $p = 10$. Type of candidate series: 1. Type of reference composites: 4 when the composite has already been homogenised, 1 otherwise.

7. Homogeneity-adjustments (Sects. 3.6.2 and 3.6.3). a) Input: series-0, adjustment-terms: from $\mathbf{T}^+$ of step 3, type of results series: 3; b) Input: series-1, adjustment-terms: from $\mathbf{T}^+$ of step 3, type of results series: 4; c) Input: series-1, adjustment-terms: from $\mathbf{T}^+$ of step 6, type of results series: 4a.

8. Calculation of first difference series and their spatial correlations from series-4.

9. Calculation of relative time series. This step prepares the repetition of outlier-filtering, and for this reason the candidate series are without previous outlier-filtering. Type of candidate series: 3. Type of reference composites: 4.

10. Outlier-filtering. The results are series-5.

11. Calculation of spatial correlations from series-5.

12. Recalculation of interpolated values using series-4. The primer results are series-5. Then from the interpolated values the homogenisation-adjustments are subtracted, and with these values the outliers in series-0 are substituted. The results are series-2.

13. Calculation of relative time series and $\mathbf{T}^+$ with the exclusion of one of the possible composites, $p = 30$. Type of candidate series: 2. Type of reference composites: 5.

14. Homogenisation-adjustments (Sects. 3.6.2 and 3.6.3). Input: series-2, adjustment-terms: from $\mathbf{T}^+$ of step 13, type of result series: 5a.

Part III: Homogenisation

Note: In this part unified relative time series are not used.

1. Calculation of relative time series. Type of candidate series: 2. Type of reference composites: 5a.

2. Main Detection.

3. Exclusion of one IH, if there occurs a common timing of IHs for all time series. Then the least significant IH is selected, based on the calculations of indicator z^* around the timings of the common IH (k), Eqution (43).

$$z^* = \left(j_{k+1} - j_{k-1} \right) \cdot \left(\alpha_k^2 + c_0^2 \beta_k^2 \right) \qquad (43)$$

In Eqution (43) α_k and β_k are determined with Equtions (35) and (36). The smallest z^* indicates the least significant IH, then that is excluded from the list of detected IHs.

4. Refining timings of IHs in monthly time scale. In the Main detection annual characteristics are used only, so the timings of IHs cannot be assessed with monthly preciseness from that. Here, the relative time series ($\mathbf{T}$) of step 1 are used in monthly resolution. – The first estimation for the timing of IH k is taken from step 2, and it is the December of the year (j) detected. Then two-phase harmonic functions are fitted in a 48-month wide window centred at December of year j. This fitting is made in the same way as for the Secondary Detection (Equation (34)), but here the number of change-points is fixed to be 1 within the window. Further limitation is that the timing of the change-point is searched up to 12 months distance from the centre of the window. – For the calculations of this step always the $\mathbf{T_q}$ of the highest w is selected that contains the 48-month window around j_k.

5. Calculation of correction-terms with ANOVA. Input field: series-2 and timings of IHs from steps 2-4. This step consists of three operations: a) ANOVA for correction-terms in annual means ($\mathbf{TM}$) (Sect. 3.6.1); b) ANOVA for correction-terms in summer-winter differences ($\mathbf{TD}$) (Sect. 3.6.1); c) Calculation of monthly correction-terms (Sect. 3.6.3). – In ACMANT the input variables are introduced to ANOVA in monthly resolution, as at this phase of the procedure the timings of IHs are available with monthly preciseness. Problems do not occur in calculating α, since the calendar months are nearly evenly distributed between any two adjacent change-points. By contrast, the summer-winter difference is a characteristic which has no values in monthly time-scale. The problem is tackled by applying moving weighted averages of monthly anomalies, providing monthly values in this way for ANOVA calculations (Equation (44)).

$$td'_{j,m} = \sum_{H=H^*-5}^{H^*+5} c'_{m(H)} \cdot a_H + 0.5 \left(c'_{m(H-6)} + c'_{m(H+6)} \right) a_H \quad (44)$$

Examining Eqution (44) it can be seen that $\mathbf{TD'}$ characterises the summer-winter difference, and it has monthly values. Values for c_m' are set in harmony with Equation (21), ($c_1' = -1/3.5$, $c_2' = -0.5/3.5$, etc.).

6. Application of homogenisation-adjustments on relative time series of step 1. Each $\mathbf{T_{g,q}}$ ($q = 1,2,\ldots Q$) is ad-

justed with the adjustment-factors determined by step 5 for series g. The results are $\mathbf{T}^*$.

Steps 7 and 8 compose a block whose steps are fulfilled cyclically as long as inner indicators show the necessity of the continuation of the cycle.

7. Calculation of the maximums of accumulated and normalised anomalies in adjusted relative time series. (Sect. 3.5.1). If one of the maximums exceeds the pre-defined thresholds, step 8 follows, otherwise the procedure continues with step 9.

8. Secondary Detection (Sect. 3.5.2). Step 7 follows, but before that the section of time series that has already subdued to Secondary Detection, is excluded from further examinations of accumulated anomalies.

9. Calculation of correction terms with ANOVA. The only difference relative to step 5 is that here the list of IHs is supplied with the result of the Secondary Detection.

10. Application of correction-terms on series-2. The results are series-6.

Part IV. Final adjustments

Some IHs might turn out to be insignificant after the ANOVA calculations. They are removed, and the ANOVA is repeated with the rest of the IHs. These steps are fulfilled cyclically as long as there is no insignificant IH after the ANOVA calculations.

1. The significance of each IH (k) is controlled with t-test. The size of IH is characterised by the sum of absolute values of α_k and $c_0\beta_k$ where α_k and β_k are determined with Equtions (35) and (36). Series-6 are used. The periods to which the shift is assigned are presented by Equtions (45) and (46).

$$l_1 = H_k - H_{k-1} \qquad (45)$$

$$l_2 = H_{k+1} - H_k \qquad (46)$$

The standard deviation for both periods is considered to be the same as the standard deviation of the relative time series in which k was detected. Let the sum of l_1 and l_2 be denoted by l, then statistic p^* can be given by Equation (47).

$$p^* = \frac{\left(|\alpha_k| + c_0|\beta_k|\right)\sqrt{l_1 l_2 (l-2)}}{l\sigma} \qquad (47)$$

Note that p^* differs from regular t-statistics of l-2 freedom because it includes β, while σ characterises the standard deviation of α only. – In the present application, thresholds for selecting significant IHs equal to the regular thresholds (depending on l) for the 0.05 significance level. If at least 1 IH is excluded, step 2 follows, otherwise step 4.

2. Calculation of correction terms with ANOVA. The only difference relative to step 5 of Part III is, that here the list of IHs is altered relative to the previous calcula-

tions.

3. Application of correction-terms on series-2. The results are series-6. Step 1 follows.

4. Long-term means of monthly temperatures are added to the homogenised anomalies (Eqution (48)). a^* marks homogenised anomalies and $\mathbf{V}$ stands for homogenised temperature series.

$$v_{s,j,m} = a^*_{s,j,m} + \frac{1}{n'_{s,m}}\sum_{i=1}^{n_s} x_{s,i,m} \qquad (48)$$

5. Verification

5.1. Test-Datasets

In the COST HOME action (COST ES0601) a benchmark dataset with known artificial IHs was built by a group of experts, and was announced [26] for the climatological community, for comparing the results of different homogenisation methods. This benchmark consists of 40 networks of 100-year long series of simulated monthly temperatures, 40 networks of simulated precipitation data and some networks with real observed data. For verification purpose only the simulated temperature data are used in this study, since in datasets of real observations the properties of inhomogeneities are never known perfectly, thus exact verification cannot be performed for them.

Half of the simulated networks was generated with white noise background (synthetic data), while in the other half of the networks the background noise mimics the low-frequency variability of time series better (surrogated data). The benchmark contains networks of 15 time series (25%), networks of 9 time series (25%) and networks of 5 time series (50%). The spatial correlations are high (often higher than 0.8) what is typical for temperature networks of the last 100-150 years observations in Europe and the US. The frequency of IHs varies widely, but most networks contain a few large IHs and a lot of small IHs. See more details in [26].

In this study another test dataset is also used. It consists of 100 surrogated and 100 synthetic networks of monthly temperature data. This dataset (referred as benchmark-2 hereafter) was created exactly in the same way as the benchmark dataset, the only difference is that in benchmark-2 all the networks contain 9 time series (50%) or 15 time series (50%).

5.2. Evaluation of efficiency

The root mean squared error (RMSE) caused by IHs or imperfect homogenisation was calculated for some statistical characteristics of raw time series (W_R) and for that of homogenised time series (W_H). Then the efficiencies

were calculated by Eqution (49).

$$Eff = \frac{W_R - W_H}{W_R} \qquad (49)$$

The verified characteristics are: a) RMSE of monthly biases, b) RMSE of biases in annual means, and c) RMSE of slope-biases in the linear trend for the entire period of time series.

5.3. Efficiency Results

In January of 2010 the first evaluation of efficiencies was presented for the participants of COST event, and some of the results have been published [27]. It was found that the performance of ACMANT is one among the most effective homogenisation methods. However, the author has found problems with the reliability of climatic trend estimation of that version (ACMANTv0), particularly when the method was used for homogenising small networks ($N = 5$). After the causes of the bias have been identified, some modifications were made in ACMANT. The present version (ACMANTv1, February, 2011) that is described in this study contains several modifications relative to ACMANTv0. In the ACMA- NTv0 neither ANOVA, nor unified relative time series were used for calculating correction-terms, but those terms were calculated using the same relative time series as in which the IHs were detected. The lack of harmonisation among individual assessments may cause accumulated biases in the estimation of long term trends. Another change is that in ACMANTv1 the modus of the annual cycle coincide with solstices, why in ACMA- NTv0 they were in mid-January and mid-July. However, in the benchmark-homogenisation I still used the seasonal cycle of the earlier version, because it provides the best results for the benchmark.

The ACMANTv1 produces better results than the ACMANTv0 not only in the trend estimation, but also in the reconstructions of other statistical properties of time series. For the 40 simulated networks of the benchmark the efficiency of trend-slope estimation increased from 0.400 to 0.745, that of the annual mean temperatures increased from 0.516 to 0.661, and that of the monthly mean temperatures increased from 0.434 to 0.553.

As since February 2010 I know the true positions of IHs in the benchmark, another test dataset, the benchmark-2 was used for the validation of ACMANTv1. Surprisingly, the results for benchmark-2 are slightly even better, than the results for the former benchmark (**Figure 1**). When the values for **Figure 1** were calculated, the small networks ($N = 5$) of the benchmark were excluded, for using test-datasets of the same statistical properties. It can be seen that (i) for relatively large net-

works even the ACMANTv0 performed quite well, (ii) the ACMANTv1 has always better efficiency than the ACMANTv0, (iii) the blind-test results are slightly even better than the results achieved in the benchmark, (iv) the results are better when synthetic data are homogenised in comparison with the homogenisation of surrogated networks, (v) the efficiency in reducing the monthly RMSE is slightly poorer than the reduction of the annual RMSE.

The author cannot explain why the efficiency results are better for the benchmark-2 than for the benchmark. The opposite relation should be expected, since the ACMANTv1 contains some semi-empirical parameters which were assumed based on experiments with the benchmark. Perhaps the 20 networks of the benchmark that were used for **Figure 1** frequently contain difficult coincidences of IH and noise-terms by accident, in comparison with the mean properties of the much larger benchmark-2. Anyhow, the achieved performance is outstandingly high. The author does not know other homogenisation method with similar or better efficiency.

6. Discussion and conclusion

The development of ACMANT relies on previous analyses of efficiencies of homogenisation methods [28], [29]. Notwithstanding, the application of homogenisation methods on such a complete dataset as the benchmark of the COST HOME offered and offers further opportunities to reveal the efficiency characteristics of methods. This study restricts to discuss characteristics whose impacts to the efficiency of ACMANT are unambiguous. Seven favourable characteristics and three still existing shortcomings are discussed:

1) IH-sizes of temperature series usually have seasonal cycle. It is caused by the fact that most of the IHs have relation to the change of radiation-sheltering or other radiation effects, and these effects are larger in summer than in winter. Thus, following the annual cycle of radiation intensity, most of temperature IH sizes have quasi-harmonic cycle. The benchmark of COST HOME contains these seasonal changes of IHs rather realistically. The ACMANT applies a bi-variable test for detect IHs, one variable is the annual mean (**TM**), and the other is the summer - winter difference (**TD**). This way of detection is very effective, because a) The two variables often have change-points with the same timings, thus the signal/noise ratio is better in a unified tests; b) The use of summer-winter difference is more effective, than the use of seasonal or monthly means, because the signal/noise ratio is relatively low for the summer - winter difference (**TD** is calculated from the values of eight monthly means, and the noise decreases with the increase of the number of averaged values); c) Distinct analyses of sea-
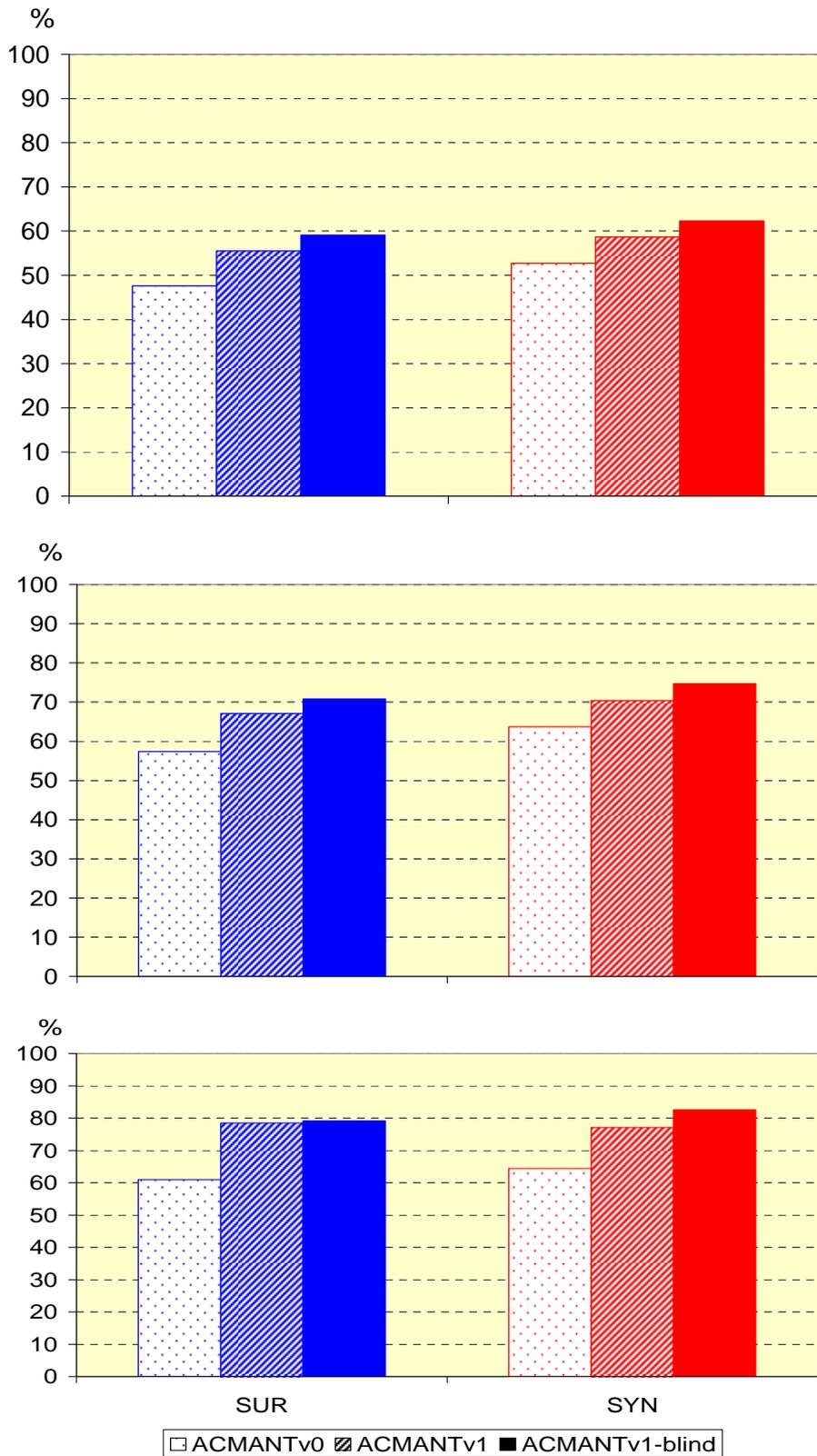
Figure 1. Efficiency in reducing the RMSE of monthly biases caused by IHs or imperfect homogenisation. Results for ACMANTv0, ACMANTv1-benchmark and ACMANTv1- benchmark-2 are presented by dotted, striped and filled fields, respectively. Upper part: RMSE of monthly values; mid-part: RMSE of annual means, bottom: RMSE of trend-slopes.

sonal and monthly series may result in set of detected change-points with poor seasonal coherence. In that case, after having the primer results, a harmonisation is necessary to obtain a realistic description of IH properties. Considering that that harmonisation needs further estimations loaded with uncertainties, it might cause biases in the result. In contrast, when **TM** and **TD** are examined, such a harmonisation is unnecessary, because the seasonal cycles of IHs can be derived directly from the detection result.

2) One base of the development is the Caussinus - Mestre detection method. Verification results show that the Caussinus - Mestre detection method is one of the best among the existing methods [29].

3) According to some tests fulfilled with the Caussinus - Mestre detection method, its performance has turned out to be better, when detection of IHs with shorter than 3 time units is not allowed. It is likely, because noise can produce IH-shaped changes in the very short time-scale more frequently than in longer time-scale. The ACMANT applies this experience.

4) In ACMANT the ANOVA is applied for calculating corrections which method provides the optimal estimations of correction-terms.

5) The pre-homogenisation of reference components is favourable when information specific for the connection between the candidate series and its references is not utilised in that step. The pre-homogenisation included in ACMANT improves the certainty of the estimation of number of IHs and that of the timings of IHs.

6) When the assessment of IH-positions are relatively confident at least for large-size IHs, the applied method for finding the timings of IHs in monthly time-scale gives improvement. Note that according to some experiments with the COST-benchmark, the application of the same technique in an earlier phase of the procedure for assessing timings with monthly preciseness did not result in any improvement of efficiency.

7) The calculation of correction-terms with ANOVA may show that some IHs which seemed to be significant in the detection process, do not have significance in the final results. With the exclusion of the insignificant IHs (their percentage was approximately 10% when the COST-benchmark was used) a better estimation of correction-terms can be provided basing on the significant IHs only.

Three shortcomings still wait for the application of further developments.

1) The aim of the Secondary Detection is to identify short-term but large-size IHs. However, as it operates after the Main Detection, a) the results of the Main Detection have errors due to unfiltered large-size, short-term IHs, b) the results of the Secondary Detection are affected by the errors of earlier steps.

2) The ACMANT is a purely statistical procedure, now it is impossible to use metadata (documental information of technical or environmental changes in the observation) information within ACMANT. However, the author thinks that the potential usability of metadata is limited when the spatial density and spatial correlations of data is appropriate to perform statistical homogenisation, since metadata do usually not contain quantitative information about the degree of local effects.

3) The present method is applicable for monthly temperatures from mid- or high-latitudes, and is not applicable for other climatic variables. The ACMANT contains innovations whose application would likely be useful for homogenising other variables than monthly temperatures, therefore further developments of homogenisation methods are needed.

A task for the future is, to apply efficiency-tests of wider properties of time series and data networks than which are provided by the COST-benchmark. The parameterisation of ACMANT has to be checked or modified relying on further tests.

Summarising, the ACMANT is a new homogenisation method which has been developed for homogenising mid- and high-latitude temperature series of observing networks. It has outstanding efficiency among statistical homogenisation methods. Its use is particularly recommended for homogenising networks comprising large number of time series with sufficient spatial correlations, since the ACMANT is a fully automatic method.

The executable file of ACMANTv1 together with its user-guide is freely downloadable from http://www.c3.urv.cat/members/pdomonkos.html.

7. Acknowledgements

8. References

[1] T. A. Buishand, "Some methods for testing the homogeneity of rainfall records," J. Hydrology, Vol 58, 1982, pp. 11-27.

[2] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," Ann. Math. Stat., Vol 18, 1947, pp. 50-60.

[3] J. R. Lanzante, "Resistant, robust and non-parametric techniques for the analysis of climate data: theory and examples, including applications to historical radiosonde stationdata," Int. J. Climatol., Vol 16, 1996, pp. 1197-1226.

[4]   A. R. Solow, "Testing for climate change: an application of the two-phase regression model," J. Clim. Appl. Meteor., Vol 26, 1987, pp. 1401-1405.

[5]   L. A. Vincent, "A technique for the identification of inhomogeneities in Canadian temperature series," J. Climate, Vol 11, 1998, pp. 1094-1104.

[6]   H. Alexandersson, "A homogeneity test applied to precipitation data," J. Climatol, Vol 6, 1986, pp. 661-675.

[7]   X. L. Wang, Q. H. Wen and Y. Wu, "Penalized maximal t test for detecting undocumented mean change in climate data series," J. Appl. Meteor. Climatol., Vol 46/6, 2007, pp. 916-931. doi:10.1175/JAM2504.1

[8]   T. Szentimrey, "Multiple Analysis of Series for Homogenization (MASH)," Second Seminar for Homogenization of Surface Climatological Data, WCDMP 41, WMO-TD 962, WMO, Geneva, 1999, pp. 27-46.

[9]   H. Caussinus and O. Mestre, "Detection and correction of artificial shifts in climate series," J. Roy. Stat. Soc. Series C, vol 53, 2004, pp. 405-425.

[10] T. C. Peterson and 20 co-authors, "Homogeneity adjustments of in situ atmospheric climate data: a review," Int. J. Climatol, Vol 18, 1998, pp. 1493-1517.

[11] E. Aguilar, I. Auer, M. Brunet, T. C. Peterson and J. Wieringa, "WMO Guidelines on climate metadata and homogenization," WCDMP-No. 53, WMO-TD No 1186, WMO, Geneva, 2003.

[12] J-F. Ducré-Robitaille, L. A. Vincent and G. Boulet, "Comparison of techniques for detection of discontinuities in temperature series," Int. J. Climatol, Vol 23, 2003, pp. 1087-1101. doi: 10.1002/joc.924

[13] I. Auer, and 24 coauthors, "A new instrumental precipitation dataset for the greater Alpine region for the period 1800-2002," Int. J. Climatol, Vol 25, 2005, pp. 139-166. doi: 10.1002/joc.1135

[14] M. J. Menne and C. N. Williams Jr, "Detection of undocumented changepoints using multiple test statistics and composite reference series," J. Climate, Vol 18, 2005, pp. 4271-4286. doi: 10.1175/JCLI3524.1

[15] M. J. Menne and C. N. Williams Jr, "Homogenization of temperature series via pairwise comparisons," J. Climate, Vol 22, 2009, pp. 1700-1717. doi: 10.1175/2008JCLI2263.1.

[16] M. Brunet, M, O. Saladié, P. Jones, J. Sigró, E. Aguilar, A. Moberg, D. Lister, A. Walther, D. Lopez and C. Almarza, "The development of a new dataset of Spanish daily adjusted temperature series (SDATS) (1850-2003)," Int. J. Climatol, Vol 26, 2006, pp. 1777-1802. doi: 10.1002/joc.1338

[17] M. Brunet, O. Saladié, P. Jones, J. Sigró, E. Aguilar, A. Moberg, D. Lister, A. Walther and C. Almarza, "A case-study/guidance on the development of long-term daily adjusted temperature datasets," WC-DMP-66/WMO-TD-1425, 2008.

[18] P. Domonkos, "Application of objective homogenization methods: Inhomogeneities in time series of temperature and precipitation," Időjárás, Vol 110, 2006, pp. 63-87.

[19] M. Brunet, J. Asin, J. Sigró, M. Bañon, F. García, E. Aguilar, J. E. Palenzuela, T. C. Peterson and P. Jones, P," The minimization of the screen bias from ancient WesternMediterranean air temperature records: an exploratory statistical analysis," Int. J. Climatol., early view 2011, doi:10.1002/joc.2192

[20] T. C. Peterson and D. R. Easterling, "Creation of homogeneous composite climatological reference series," Int. J. Climatol., Vol 14, 1994, pp. 671-679.

[21] H. Alexandersson, and A. Moberg, "Homogenization of Swedish temperature data. Part I: Homogeneity test for linear trends," Int. J. Climatol, Vol 17, 1997, pp. 25-34.

[22] M. Begert, T. Schlegel and W. Krichhofer, "Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000," Int. J. Climatol, Vol 25, 2005, pp. 65-80. doi: 10.1002/joc.1118

[23] A. T. DeGaetano, "Attributes of several methods for detecting discontinuities in mean temperature series," J. Climate, Vol 19, 2006, pp. 838-853. doi: 10.1175/JCLI3662.1

[24] D. M. Hawkins, "On the choice of segments in piecewise approximation," J. Inst. Math. Appl., Vol 9, 1972, pp. 250-256.

[25] H. Caussinus and F. Lyazrhi, "Choosing a linear model with a random number of change-points and outliers," Ann. Inst. Statist. Math, Vol 49/4, 1997, pp. 761-775.

[26] COST HOME home-page, http://www.meteo.uni-bonn.de/venema/themes/homogenisation/

[27] V. Venema, O. Mestre and the COST HOME Team, "Benchmark database," EGU General Assembly, Vienna, Austria, 3-7 May, 2010, EGU CL4.6-13357.

[28] P. Domonkos, "Testing of homogenisation methods: purposes, tools and problems of implementation," Proceedings of the 5th Seminar and Quality Control in Climatological

[29] Databases, (Ed. M. Lakatos, T. Szentimrey, Z. Bihari and S. Szalai), WCDMP-No. 71, WMO/TD-No. 1493, 2008, pp. 126-145.

[30] P. Domonkos, "Efficiency evaluation for detecting inhomogeneities by objective homogenisation methods," Theor Appl Climatol, early view 2011. doi: 10.1007/s00704-011-0399-7

Appendix: Explanation of Symbols and Acronyms

The list below contains the explanation of symbols those are used throughout the paper. Note that one can find further explanations in the text about contemporary use of variables, as well as about some constants and parameters whose role varies in the paper. Bold capital letter marks vector or matrix variable.

A   - anomaly
Δ**A**   - first difference (increment) of anomalies
A*   - adjusted anomaly
B   - normalised anomaly
F   - reference series
g - index of candidate series
G - penalty term
h - serial number of month in a series
hn - first month of a series
hx - last month of a series
H - timing of change-point in months
H^* - timing of selected change-point (in months)
j - year
k - serial number of change-point or that of section in fitted function
K - total number of detected change-points for a time series
K' - number of detected change-points for a section of a time series
l,L - length of series in a selected examination
m - calendar month
n - whole length of time series in years
n' - number of observed values in time series
N - number of stations
nm - length of relative time series in months
p^* - statistic of t-test
q - serial number of relative time series
Q - number of reference series for a specified candidate series
R   - spatial correlation between first difference series
R'   - spatial correlation
s - station

S - number of components of reference series
T   - relative time series
T$^+$ - unified relative time series
T* - corrected relative time series
TM - annual mean
TD - summer-winter difference
U   - correction-term
u' - summarised impact / correction-term belongs to some selected IHs
V   - homogenised temperature
w - sum of the squared correlations of reference components
W_R - error-term for raw time series
W_H - error-term for homogenised time series
X   - raw time series
Y$_1$ - first year of relative time series
Y$_2$ - last year of relative time series
z, z^* - indicator of significance of IHs
α - shift in annual means at an IH
β - shift in summer-winter differences at an IH
σ - standard deviation

ACMANT – Homogenisation method developed by the author: Adapted Caussinus-Mestre Algorithm for homogenising Networks of Temperature series.
ACMANTv0 – The version of ACMANT that took part in the blind test experiment of the COST HOME.
ACMANTv1 – The version of ACMANT that is described in this paper.
ANOVA – Equation-system based calculation method of correction-terms for homogenising time series.
COST HOME / COST ES0601 – International scientific action on the development and testing of homogenisation methods. It is sponsored by the European Union.
IH – Inhomogeneity: technical-born bias from the true climate in the series of observed data.
MA – moving average
PRODIGE – Homogenisation method that was developed by Caussinus and Mestre (2004).
RMSE – Root mean squared error.