IDŐJÁRÁS
Quarterly Journal of the Hungarian Meteorological Service
Vol. 117, No. 1, January–March 2013, pp. 47-67

HOMER : a homogenization software – methods and applications

Olivier Mestre[1]*, **Peter Domonkos**[2], **Franck Picard**[3], **Ingeborg Auer**[4], **Stéphane Robin**[5,6], **Emilie Lebarbier**[5,6], **Reinhard Böhm**[4], **Enric Aguilar**[7], **Jose Guijarro**[8], **Gregor Vertachnik**[9], **Matija Klancar**[9], **Brigitte Dubuisson**[1], and **Petr Stepanek**[10]

[1]*Meteo-France, Direction de la Production, 42 avenue Coriolis, 31057 Toulouse cedex, France*

[2]*Center for Climate Change, Univ. Rovira i Virgili, Av. Remolins, 13-15, 43500-Tortosa, Spain*

[3]*UCB Lyon 1, UMR 5558, Villeurbanne, France*

[4]*Zentralanstalt für Meteorologie und Geodynamik, Wien, Austria*

[5]*AgroParisTech, UMR 518, Paris, France*

[6]*INRA, UMR 518, Paris, France*

[7]*Center for Climate Change, Univ. Rovira i Virgili, Tarragona, Spain*

[8]*Agencia Estatal de Meteorologia, Palma de Mallorca, Spain*

[9]*Environmental Agency of the Republic of Slovenia, Meteorology, Ljubljana, Slovenia;*

[10]*Czech Hydrometeorological Institute, Brno, Czech Republic*

**Corresponding author E-mail: Olivier.Mestre@meteo.fr*

(Manuscript received in final form October 24, 2012)

Abstract–Between 2007–2011, the European COST Action ES0601 called HOME project was devoted to evaluate the performance of homogenization methods used in climatology and produce a software that would be a synthesis of the best aspects of some of the most efficient methods. HOMER (HOMogenizaton softwarE in R) is a software for homogenizing essential climate variables at monthly and annual time scales. HOMER has been constructed exploiting the best characteristics of some other state-of-the-art homogenization methods, i.e., PRODIGE, ACMANT, CLIMATOL, and the recently developed joint-segmentation method (*cghseg*). HOMER is based on the methodology of optimal segmentation with dynamic programing, the application of a network-wide two-factor model both for detection and correction, and some new techniques in the coordination of detection processes from multiannual to monthly scales. HOMER also includes a tool to assess trend biases in urban temperature series (UBRIS). HOMER's approach to the final homogenization results is iterative. HOMER is an interactive method, that takes advantage of metadata. A practical application of HOMER is presented on temperature series of Wien, Austria and its surroundings.

Key-words: Homogenization, optimal segmentation, joint segmentation, ANOVA, temperature, precipitation, urban trend bias

1. Introduction

The accuracy of climatic observations is often affected by inhomogeneities due to changes in the technical or environmental conditions of the measurements (station relocations, changes of the type, height or sheltering of the instruments, etc., *Aguilar et al*., 2003, *Auer et al*., 2005). Most of such changes cause sudden shifts (change-points) in the series of local climatic data, while some others (particularly urban development) result in gradually increasing biases from the real macroclimatic characteristics. Correction of inhomogeneities before any climate variability analyses is highly desirable, and for this purpose, a large number of homogenization methods have been developed in the recent decades (*Peterson et al*., 1998; *Ducre-Robitaille et al*., 2003; *Beaulieu et al*., 2008; among others).

HOMER is a recently developed method for homogenizing monthly and annual temperature and precipitation data. It includes the best features of some other state-of-the-art methods, namely PRODIGE (*Caussinus* and *Mestre*, 2004), ACMANT (*Domonkos*, 2011), and *cghseg* a joint segmentation method that was developed originally by bio-statisticians in the context of DNA segmentation (*Picard et al*., 2011). PRODIGE and ACMANT have the same theoretical base regarding the optimal segmentation with dynamic programming DP (*Hawkins*, 2001), an information theory based formula for determining the number of segments in time series (hereafter: C&L criterion, *Caussinus* and *Lyazrhi*, 1997), and a network-wide unified correction model (ANOVA, *Caussinus* and *Mestre*, 2004). The results of blind test experiments conducted during COST Action ES0601 (*Venema et al*., 2012) validates these approaches, since PRODIGE and ACMANT rank among the best methods for homogenizing monthly and annual climate data (*cghseg* and HOMER were not tested during the HOME action). The joint segmentation is an extension of the optimal segmentation for finding network-wide optima by means of an iterative procedure, a modified BIC criterion being used for determining the number of changes (*Zhang* and *Siegmund*, 2007; *Picard et al*., 2011).

HOMER is an interactive semi-automatic method. In applying HOMER, users may choose between the *cghseg* detection results whose generation is fully automatic on the one hand, and a partly subjective pairwise comparison technique that is adapted from PRODIGE on the other hand. This freedom allows users to add subjective decisions based on metadata or research experiences. HOMER includes also some innovations of ACMANT in the coordination of working on different time scales. Basic quality control and network analysis are adapted from CLIMATOL (*Guijarro*, 2011).

Our paper is organized as follows: first, Section 2 describes the main models and procedures of HOMER. The methodology of characterizing urban trends (UBRIS) and the main properties of ACMANT are also presented there, together with a discussion. An application of HOMER on Wien temperature series is then shown in Section 3.

2. *HOMER main procedures*

In this section, we will focus on functions used during the homogenization process: statistical tools for pairwise detection (2.1), two factor model for joint detection and correction (2.2), UBRIS model for urban trend bias assessment (2.3), ACMANT functions (2.4). Usefulness of each task is discussed in 2.5, and a workflow of tasks is provided.

2.1. Detection of changes in pairwise series (univariate detection)

2.1.1. Model

Let Y be the annual or seasonal difference between two series. We model $Y_i, i=1,\ldots,n$ as a series of Gaussian variables of constant variance σ^2, but with varying mean μ from sub-period to subperiod. The number and positions of change-points are unknown.

Let k the number of changes and $\tau_1, \tau_2, \ldots, \tau_k$ their positions. We denote $K=\{\tau_1,\ldots,\tau_k\}$ the set of changes in the series. At most cases old data are adjusted relative to the modern data, and for simplicity $\tau_o=0$ is fixed at $\tau_{k+1}=n$. Further notations are:

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \,,$$

$$\overline{Y}_j = \frac{1}{n_j} \sum_{\tau_{j-1}+1}^{\tau_j} Y_i \ ,$$

where $n_j = \tau_j\text{-}\tau_{j-1} + 1 \ ; \ j = 1, \dots, k+1 \ ,$

$$\overline{Y}_{hm} = \frac{1}{m-h+1} \sum_{i=h+1}^{m} Y_i \ ,$$

$$W_{hm} = \sum_{i=h+1}^{m} \left(Y_i - \overline{Y}_{hm}\right)^2 .$$

Changes in the mean are $\mathrm{IE}[Y_i] = \nu_j$ for $\tau_{j-1} + 1 < i < \tau_j$

Maximum likelihood estimates of the ν_j's are straightforwardly given by $\hat{\nu}_j = \overline{Y}_j$. For a given number k, we wish to maximize the likelihood, which is equivalent to minimize deviance *D*:

$$D_k = \frac{\sum_{j-1}^{k+1} \sum_{\tau_{j-1}+1}^{\tau_j} \left(Y_i - \overline{Y}_j\right)^2}{\sigma^2} + 2n \log\left(\sqrt{2\pi}\,\sigma\right) \quad . \tag{1}$$

2.1.2. Dynamic programming

The naive way to minimize deviance *D* is to consider every combination of the position of the change-points. But the number of hypotheses rises very fast with *n*, the length of the series, and *k*, the number change-points. When detection is performed for change-points in a normal sample, a DP algorithm can be used (*Lavielle*, 1998; *Hawkins*, 1972, 2001; etc.). Computation time then becomes only linear in k and quadratic in *n*. It is based on a recursion between optimal *k* and *k*–1 solutions. DP allows us to find an optimal solution without computing all possibilities. For *k* changes, the problem is to minimize:

$$Q = \sum_{j=1}^{k+1} \sum_{i=\tau_{j-1}+1}^{\tau_j} \left(Y_i - \overline{Y}_j\right)^2 = \sum_{j=1}^{k+1} W_{\tau_{j-1}\ \tau_j} \ . \tag{2}$$

The solution is given by the following recursion:

- $F_{1,m} = W_{0m}$ for $m = 1, n$,

- for each $r = 2, \dots k + 1$, let us compute
  $F_{r,m} = \mathrm{MIN}_{0<h<m}\lfloor F_{r-1,h} + W_{h,m}\rfloor$ for $m = 1, n$,

- for each $F_{r,m}$ value, let us keep in table $H_{r,m}$ the *h* value that corresponds to the minimum of $F_{r,m}$,

- the change-point estimates are given by: $\tau_{k+1} = n$, and for $r = k, k-1, \dots, 1$ we get $\hat{\tau}_r = H_{r+1,\tau_{r+1}}$.

2.1.3. Selecting the number of changes.

The fit of the change-point model increases monotonously with k ($Q = 0$ for $k = n$). The model selection is guided finding the most parsimonious model that gives a "good" explanation of data vector *Y*. Several penalized likelihood criteria can be found in the literature. In the latest version of HOMER, we take the advantage of the uniseg procedure from the R package which uses the modified BIC criterion of *cghseg*. As in *Schwarz*'s BIC (1978), *Zhang* and *Siegmund* approach this problem by deriving an asymptotic approximation of the Bayes factor, using a uniform prior on change-points location (among other hypotheses).

The procedure is as follows: for each value of *k*, DP allows us to select the optimal position for the *k* change-points $\{0, \hat{\tau}_1, \hat{\tau}_2, \dots, \hat{\tau}_k, n\}$. For each k value, *MBIC*(*Y*; *k*) is computed:

$$MBI(Y;k) = \left(\frac{n-k+1}{2}\right)\log\left[1 - \frac{\sum_{j=1}^{k+1} n_j\left(\bar{Y}_j - \bar{Y}\right)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}\right] + \log\left[\frac{\Gamma\left(\frac{n-k+1}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)}\right]$$
$$+\frac{k}{2}\log\left[\sum_{i=1}^{n}(Y_i - \bar{Y})^2\right] - \frac{1}{2}\log\left[\sum_{i=1}^{n}\left(\bar{Y}_i - \bar{Y}\right)^2\right] \qquad (3)$$
$$-\frac{1}{2}\sum_{j=1}^{k+1}\log(n_j) + \left(\frac{1}{2} - k\right)\log(n) \ ,$$

where Γ denotes the Gamma function. The model selection consists in selecting the number of change-points k that minimizes *MBIC*:

$$\text{select } k^* \text{ such that } k^* = \text{Argmin}_k(MBIC_k(Y;k)). \tag{4}$$

This criterion is more complex than the classical BIC or C&L criteria used in PRODIGE, but does not require any user-chosen shrinkage parameters like in *Tibshirani* (1996), *Birgé* and *Massart* (2001) or *Gu* and *Wang* (2003). The first term in Eq. (3) corresponds to a likelihood ratio term, the subsequent ones are the penalty. One has to note that the penalty depends on n, k, but also on the closeness of the changes via the sum of $\log(n_j)$ term: close change-points are more penalized. Simulations (not shown here) show that *MBIC* criterion is slightly less powerful than the C&L, but less sensitive to small autocorrelation that still might be present in the pairwise comparisons.

Standard deviation of the residuals is then estimated by:

$$\hat{\sigma}^* = \sqrt{\frac{1}{n-k^*}\sum_{j=1}^{k^*+1}\sum_{i=\hat{\tau}_{j-1}+1}^{\hat{\tau}}\left(Y_j - \bar{Y}_j\right)^2} = \sqrt{\frac{1}{n-k^*}\sum_{j=1}^{k^*+1} W_{\hat{\tau}_{j-1}\hat{\tau}_j}} \ . \tag{5}$$

We will see in practice that this estimation of noise is very useful, since detection power is directly related to the signal (i.e., amplitude of changes) to noise ratio. Smaller values of noise ensure more accurate detection.

2.2. ANOVA: a two-factor model for joint-detection and correction

2.2.1. Model

Let us consider p series belonging to the same climate area in such a way that all the series are affected by the same climatic conditions at the same time. This assumption is realistic when considering monthly or annual observations of the same geographical region. We assume that each series of observations is the sum of a climatic effect, a station effect, and random white noise. This is a simple two-factor analysis of variance model without interaction, and we will denote it by ANOVA in the following.

Let X be a matrix of n observations X_{ij} on p series where $i=1,...,n$ is the time index and $j=1,...,p$ is the station index. Let k_j be the number of change-points, let $\tau_{1,j}, \tau_{2,j}, \ldots, \tau_{k_j,j}$ be the positions of these k_j change-points. Let $K_j = (\tau_{1,j}, \ldots, \tau_{k_j,j})$ be the set of change-points for series j. To simplify the notation, we set again $\tau_{o,j}$=0, and $\tau_{k_{j+1},j} = n$, so that K_j becomes $K_j = \left(0, \tau_{1,j}, \dots, \tau_{k_j,j}, n\right)$.

The station effect is constant if the series is homogeneous. If not, the station effect is constant between two shifts. In the following, level denotes a homogeneous sub-period between two discontinuities of a given series. For a series j with k_j breaks, let L_{jh} be the hth level (h=1,...,k_j+1), thus L_{jh} is the interval: $\left[\tau_{h-1,j} + 1, \tau_{h_j,j}\right]$. Note that the level h for the observation X_{ij} depends both on time i and station j: when necessary it will be written $h(i,j)$.

Let μ_i be the climate effect at time i and ν_{jh} the station effect of station j for level L_{jh}. If there are no outliers, the data are described by the linear model:

$$\mathrm{IE}\left(X_{ij}\right) = \mu_1 + \nu_{jh(i,j)} \quad , \quad \mathrm{Var}(X) = \sigma^2 I_{np} \quad . \tag{6}$$

One parameter of the model can be freely chosen and it is done with introducing the condition$\sum_{i=1}^{n} \mu_i = 0$, so that μ_i are defined as climate anomalies.

The number of independent parameters of the model without discontinuities is n+p–1.

Examples:

- No break in series 1: $\mathrm{IE}(X_{i1}) = \mu_i + \nu_1 \quad ,$
- One break at i_o for series 2: $\begin{cases} \mathrm{IE}(X_{i2}) = \mu_i + \nu_{21} \ , \ \text{for } i \leq i_0 \\ \mathrm{IE}(X_{i2}) = \mu_i + \nu_{22} \ , \ \text{for } i > i_0 \end{cases}$.

Some further characteristics of the model:

a) Estimation can be performed with missing data with the following conditions: there should be at least one non-missing value per year on the whole network (estimation of the μ's) and one non-missing value between two breaks for each subperiod on each series.

b) Climate signal is treated as a fixed parameter so that no assumption is made about the shape of this signal.

c) Conditionally to the climate signal, the disturbances are considered independent.

d) Local variabilities are very similar, which leads to the expression of Var(X).

Note that conditions c) and d) are approximately true within the same climatic region. Small spatial autocorrelation may be observed in the residuals.

So far, this model has been used in PRODIGE and ACMANT mainly for correction purposes – although *Caussinus* and *Mestre* (2004) propose some clue to use it for detection. It has been shown that the inclusion of ANOVA correction improves significantly the results of other methods participated in HOME blind test experiments (*Domonkos et al*., 2012b). Using HOME benchmark and the set of break-points detected using for example standard normal homogeneity test (SNHT), correcting the inhomogeneities by ANOVA allowed a much better homogenization than the standard SNHT correction method. We will see below that this model can be used for detection as well, allowing for joint detection of a whole set of series.

2.2.2. Joint-detection

The change-point model Eq. (6) can theoretically be used for joint detection of the changes on the whole set of series. However, due to the introduction of factor μ, the classical DP algorithm cannot be applied (*Caussinus* and *Mestre*, 2004) and until recently, joint segmentation was considered computationally intractable. Adapted algorithms allow us to solve this problem in a reasonable computing time. *Picard et al*. (2011) rely on two "computational tricks". The first one solves the problems caused by segmentation of multiple series. Let us set all μ_i = 0. Since DP complexity is quadratic with the size of the data, just considering segmentation of the ν factor may become problematic when considering multiple series. *Picard et al*. (2011) propose a "two-stage" DP algorithm that significantly reduces the computation time. Briefly, the first stage consists in finding all optimal solutions for each ν_j factor separately, from $k = 1$ to $kmax_j$. The second stage uses outputs from the first stage to optimally allocate the number of segments to each factor $\nu_1,...,\nu_p$, in order to maximize the overall fit. The model selection is provided by a multivariate version of *Zhang* and *Siegmund* criterion derived in *Picard et al*. (2011).

The second strategy consists in iteratively estimating μ_i and the segmentation of factor ν: at step (s+1), μ_i is estimated by:

$$\hat{\mu}_i^{(s+1)} = \frac{1}{p}\sum_{j=1}^{p} Y_{ij} - \hat{\nu}_{jh(i,j)}^{(s)} \quad , \tag{7}$$

where the segmentation of factor ν is updated using two-stage DP on $X_{ij} - \hat{\mu}_i^{(s+1)}$.

2.2.3. Correction and reconstitution of missing data

Once segmentation has been achieved, correction can be computed. Estimates $\hat{\nu}_{jh(i,j)}$ are used in the following way: let L_{jk_j} be the last level of series j, and $\hat{\nu}_{jk_j}$ the corresponding estimation of the station effect. Then, for every $X_{ij} \in L_{jh}$ $\left(1 \leq h \leq k_j + 1\right)$, corrected X_{ij} (denoted by X_{ij}^*) is given by:

$$X_{ij}^* = X_{ij} - \hat{\nu}_{jh(i,j)} + \hat{\nu}_{j,k_j+1} \;. \tag{8}$$

Note that the model allows the imputation of missing data and the correction of outliers. For any missing data or outlier (i,j), the imputation is naturally given by $\hat{X}_{ij} = \hat{\mu}_i + \hat{\nu}_{jh(i,j)}$. Since the two-factor model takes into account the change-points in the series, this allows an unbiased reconstitution of missing values, contrary to classical regression or interpolation methods.

2.3. Characterization of urban trends: UBRIS

UBRIS (urban bias remaining in series) procedure allows us to characterize artificial trends – in most cases related to urbanization, which are sometimes present in the climate series. UBRIS works jointly analyzing time series with potential artificial trends (“urban”) and without potential artificial trends (“rural”). This is an improvement compared to traditional urban trend characterization, where rural and urban series are homogenized separately, before being compared (*Peterson*, 2003 for example). This requires a large set of both rural and urban series, which may be problematic on earlier periods for example.

UBRIS relies on an extension of model Eq. (6). Let us assume that the $j < m < p$ series are free of urban trends, and that for $m \leq j \leq p$, an additional trend may affect the series.

$$\mathrm{IE}\left(X_{ij}\right) = \mu_i + \nu_{jh(i,j)} \;, \qquad \text{for} \quad 1 \leq j < m < p \;\;,$$

$$\mathrm{IE}\left(X_{ij}\right) = \mu_i + \nu_{jh(i,j)} + \beta_j i \;\;, \quad for \quad m \leq j \leq p \;, \tag{9}$$

$$\mathrm{Var}(X) = \sigma^2 I_{np} \;\;.$$

Practically, UBRIS model is slightly more complicated than Eq. (9), since trend may not affect the whole period of the series. For computation, at least one series has to be free of trend, otherwise there is no unique solution when estimating climate factor μ and trend term β. Estimation is performed via ordinary least squares. Standard student t-test allows us to test significance of the trends (β_j). UBRIS ensures a posterior estimation of those additional trends.

Prior to UBRIS analysis, HOMER has to be run in order to detect abrupt changes.

UBRIS relies on the knowledge of climatologists who decide *a priori* which series may or may not be affected by urban trends. This human expertise is important. If series corrupted by artificial trends enter the "rural" group, they will bias the estimates of climate factor μ and trend term β.

2.4. ACMANT

ACMANT (adapted caussinus mestre algorithm for homogenizing networks of monthly temperature data, *Domonkos*, 2011) was developed from PRODIGE during the HOME period. However, in contrast with PRODIGE and HOMER, ACMANT is fully automatic and it applies reference series built from composites for time series comparisons. The other main novelties of ACMANT are i) it applies pre-homogenization in a way that the double use of the same spatial connection is excluded, ii) it coordinates the operations on different time scales (from multiannual to monthly) in a unique way.

2.4.1. ACMANT bivariate detection

Observed temperature data often have inhomogeneities with significant seasonal cycles in the resulted bias (*Drogue et al*., 2005; *Brunet et al*., 2011; etc.). Therefore, change-points are searched by fitting step-functions to two annual characteristics, i.e., to annual means (Y) and to the range of the seasonal cycle (R) in relative time series, that is, candidate series minus reference series. In HOMER, the reference series are the climate signals (μ coefficients in ANOVA model) or, with other words, the reference series for ACMANT detection are always pre-homogenized. Adapting notations of Section 2.1. to R series, ACMANT detection procedure aims at minimizing:

$$Q_{YR} = \sum_{j=1}^{k+1} \sum_{i=\tau_{j-1}+1}^{\tau_j} \left(Y_i - \bar{Y}_j\right)^2 + \frac{1}{2}\left(R_j - \bar{R}_j\right)^2 \ . \tag{10}$$

The ½ factor in Eq. (10) was chosen empirically. Solutions with common timings of change-points on Y and R are considered only, so that the standard DP algorithm applies the cost function Q_{YR}. In order to set the number of changes, the C&L criterion is used both in original ACMANT and in its adaptation to HOMER:

$$C_0(Y, R) = 0 \quad \text{and}$$

$$C_k(Y,R) = \log\left[1 - \frac{\sum_{j=1}^{k+1} n_j \left[\left(\bar{Y}_j - \bar{Y}\right)^2 + \frac{1}{2}\left(\bar{R}_j - \bar{R}\right)^2\right]}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2 + \frac{1}{2}(R_i - \bar{R})^2}\right] + \frac{2k}{n-1}\ln(n) \quad . \quad (11)$$

The selection rule is: select k^* such that k^*=Argmin$_k$(C_k(Y)). In many cases, this procedure will allow us to detect changes hardly noticeable in annual means.

2.4.2. Month of change specification

Another feature of ACMANT that has been included in HOMER is its procedure for finding the most likely month of a change-point. If the precise month of the change is not known, since detection is mainly performed on annual indices, the default is to validate the break at the end of the year. At the end of the homogenization procedure, a more precise detection is made, using the monthly series serially (that is, the sequence of January, February, March, etc, for each year). Both candidate monthly series and reference series (computed from monthly μ factors) are deseasonalized; when analyzing change τ_j, standard DP algorithm is run on series of differences on interval [τ_{j-1}, τ_{j+1}]. Algorithm allows us to change the position of the change in a range of +/–2 years (in the original ACMANT the range is +/–12 months). Alternatively, the monthly precision can be determined by metadata. In HOMER, a flag marks whether a detected break is validated by metadata or not.

2.5. Discussion

The different methods contributing to the operation of HOMER have their own strengths and weaknesses. PRODIGE relies on a pairwise strategy for detection of the changes. A candidate series is compared to its neighbors in the same climatic area by computing series of differences. These difference series are then tested for discontinuities. On such a difference series without metadata, the detected changes may have been caused by the candidate or the neighbor. But, if a detected change-point remains constant throughout the set of comparisons of a candidate station with its neighbors, it can be attributed to this candidate station: this is called “attribution phase”. There are two advantages in this approach. First, we avoid creating composite reference series averaging non-homogeneous series. Second, detection relies on an efficient univariate detection procedure whose level and power are well controlled. But, because of the randomness of the difference series, the change-points of weak amplitude will lead to less accurate detection and sometimes no detection at all for some comparisons (in particular in the case of simultaneous breaks). At most cases, however, the induced ambiguity can be removed by considering the whole set of comparisons and using the metadata archives of the climate stations when available, as well as the knowledge of climatologists. This break-points detection phase has been considered the main drawback of PRODIGE, since it has to be performed manually, a process which may be tedious and time consuming, thus very difficult to apply to a large dataset and requiring a high level of regional climate knowledge and homogenization expertise.

To overcome the detection problem, an alternative approach is obtained by using the overall two-factor model, that allows the analysis and correction of a whole set of series (Section 2.2.). The *multiseg* (*cghseg* package) function determines the proper number of change-points using the MBIC criterion. This detection process with DP is quick and automatic. However model selection in a multivariate framework is a complex task, and the power of this procedure is sometimes lower than expected. In HOMER, function *multiseg* allows the automate attribution of the changes to a large extent, and in some cases the pairwise detection allows us to put into evidence changes that were not detected by *multiseg*.

ACMANT helps finding changes with a strong seasonal behavior in temperature series. In many cases, changes in observation conditions (location, sheltering, etc.) may have effects of opposing signs regarding the seasons, for example a positive effect in summer and a negative effect in winter. Such inhomogeneities are often hardly detectable on annual means, but clearly detectable with the ACMANT bivariate detection. A useful additional feature of ACMANT is the detection with monthly preciseness. The structure of HOMER has built in a way that it intends to exploit optimally the positive characteristics of the contributing methods. The tasks flow chart of HOMER is given in *Fig. 1*.

Detection is an iterative process. The initial detection phase usually reveals the most obvious changes which are corrected. Analyzing the result of this correction allows us to create an updated set of detected changes on a network. The joint detection is accompanied by the pairwise detection for allowing the use of metadata and for checking the results. The ACMANT detection follows the first cycle of detection and correction, since ACMANT detection needs pre-homogenized reference series. Note that correction is always performed on the initial data, simply by updating the set of the validated change-points before running ANOVA.

The process ends, whenever pairwise, joint-detection, and ACMANT bivariate detection find no additional changes on corrected series. In practice, the user may tolerate some pairwise comparisons still exhibiting unattributed isolated breaks, probably due to 1st kind errors.

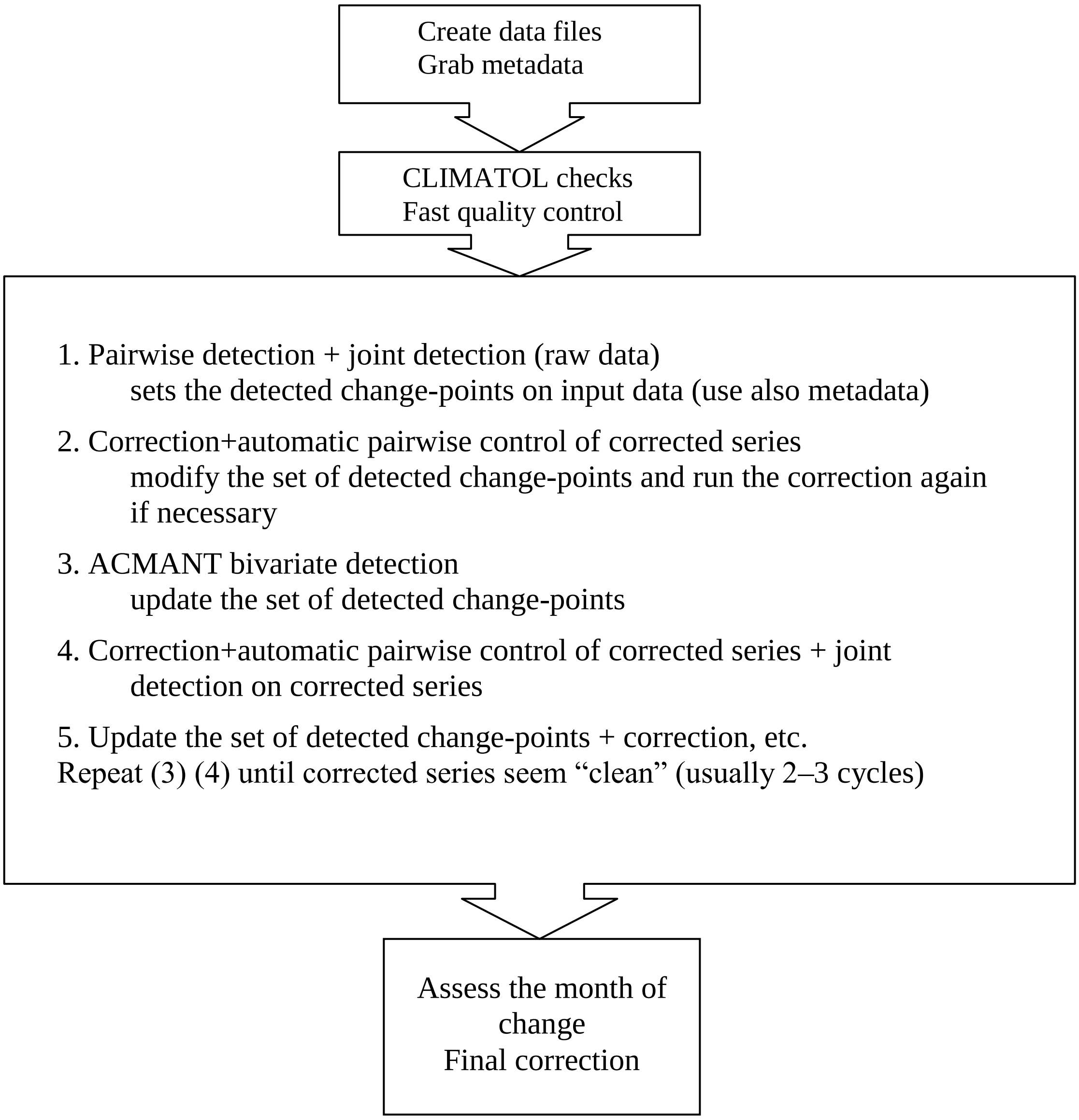

Fig. 1. Tasks flow chart of HOMER.

3. *Case study*

3.1. Homogenization using HOMER

A set of 13 series from Wien, Austria and its surroundings is provided by Zentralanstalt für Meteorologie und Geodynamik (ZAMG). Stations marked with (r) are considered rural: Fuchsenbigl(r), Gross-Enzersdorf(r), Klosterneuburg, Langenlebern(r), Schwechat, Wien-Innere-Stadt, Wien-Laaerberg, Wien-Mariabrunn(r), Rosenhügel, Rathauspark, Stadlau, Wien-Unterlaa, Wien-Hohe-Warte (*Fig. 2*).

Fig. 2. Map of Wien series

Let us take Stadlau as an example: results of pairwise detection are given in *Fig. 3*. A quick examination of pairwise comparisons puts into evidence changes in 1969 or 1970, 1984, 2001 or 2002, and potential additional changes in 1953, and 1979.

The second step consists in running *cghseg* joint-detection (*multiseg* function). Combining pairwise and joint detection allows a quick attribution of the changes: 1954, 1969, 1984, and 2002 (*Fig. 4*). Note also the good agreement in the amplitudes of the changes detected in pairwise comparisons (triangles are black for breaks detected on pairwise annual series, blue for winter, and red for summer) and joint detection (green ⊕). However, the automatic joint-detection is not perfect. On Wien series, *multiseg* tends to detect a change around 1985-1987, which is not supported at all by pairwise comparisons, and thus, it is rejected manually by the user (large red cross in the same year). During estimation of μ and segmentation ν, *multiseg* iterative algorithm has wrongly attributed a climatic feature to the ν factor. Furthermore, the rather obvious change in 1979 (when considering pairwise comparisons) was not detected by *multiseg*. User has to validate it manually using the graphical user interface. When clicking on the window, the user adds red crosses to remove or validate breaks. The *y* axis is not important, only the date (*x* axis) is taken into account. Clicking on a date selected by *multiseg* (symbol ⊕ is present) removes the corresponding date, while clicking elsewhere validates a new change-point. Metadata allow us to validate changes in 1980 (relocation of the weather station) and 2002 (changes in instrumentation). There are also sufficient statistical clues to validate the other changes, even if metadata are lacking.

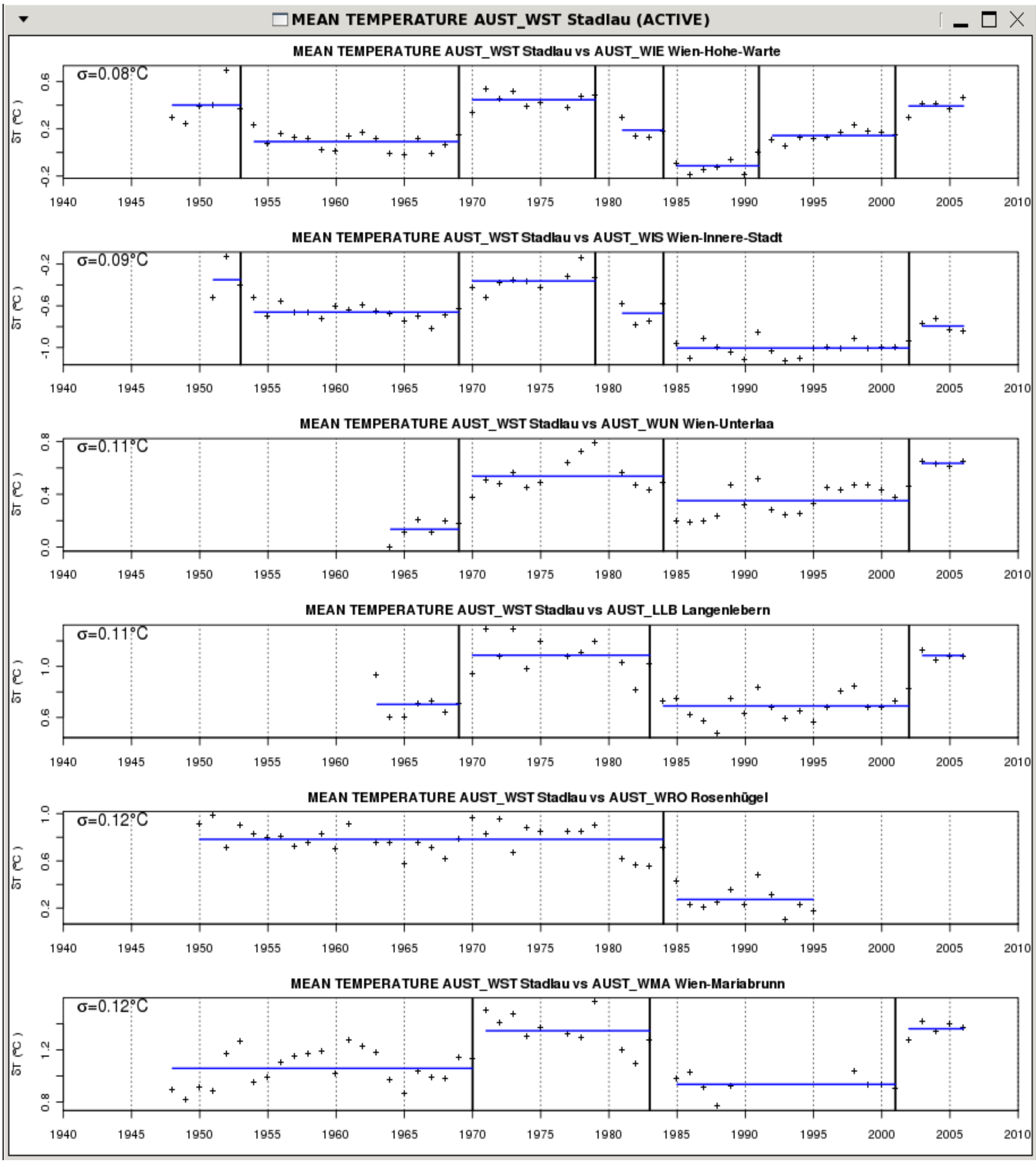

Fig. 3. Screen capture of HOMER outputs: Stadlau series compared to its neighbours. Pairwise comparison are sorted according to the increasing values of the noise standard deviation (upper left corner of each plot), computed using Eq. (5). For clarity reasons, only 6 comparisons with the smallest noise are shown.

After a correction step, ACMANT bivariate detection confirms the selected changes on Stadlau series (not shown). The raw and corrected Stadlau series after the final correction are shown in *Fig. 5* (upper panel for the raw, lower panel for the corrected series).

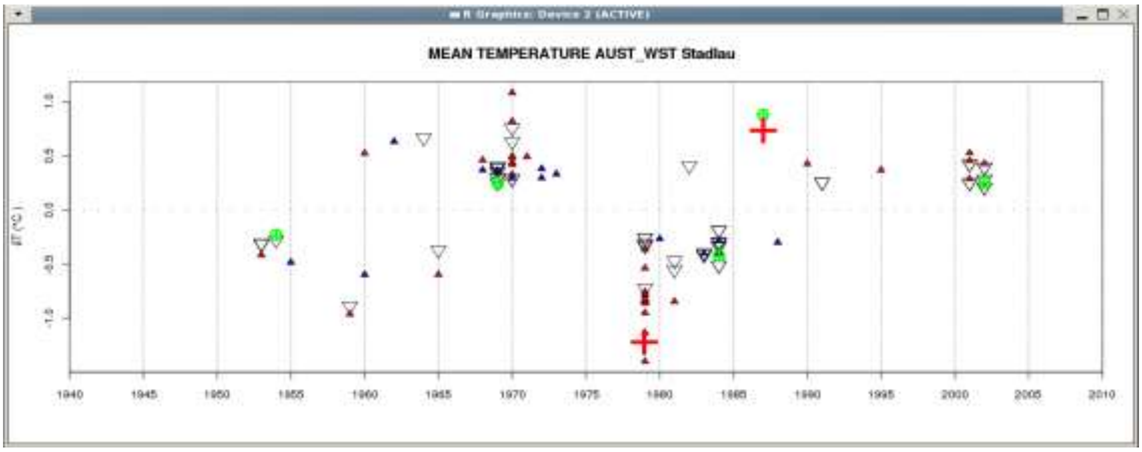

Fig. 4. Screen capture of HOMER outputs: date (*x* axis) and amplitude (*y* axis) of change-points detected on the whole set of pairwise comparisons: annual comparisons (black), winter (blue) and summer (red) triangles. Joint detection results are pointed as green ⊕ symbols. Red crosses mark user's interventions.

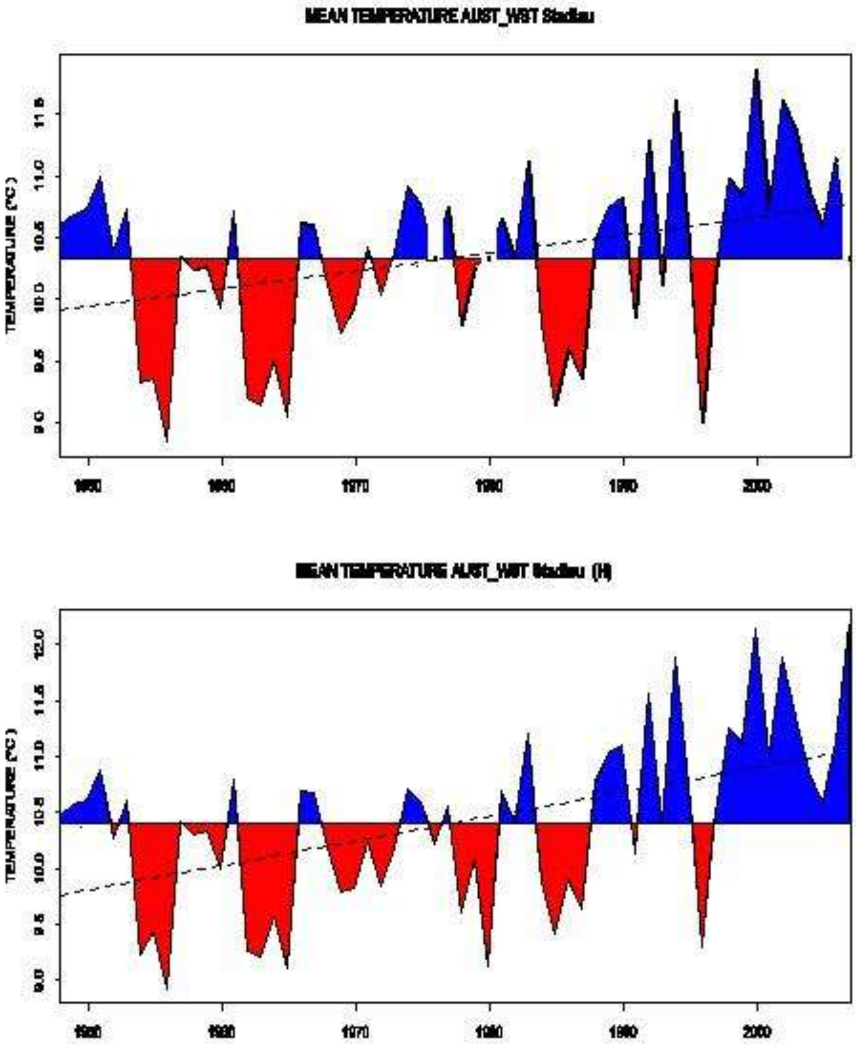

Fig. 5. Raw (up) and corrected (down) series of Stadlau.

Pairwise comparison of corrected series is characteristic of a good homogenization (*Fig. 6*).

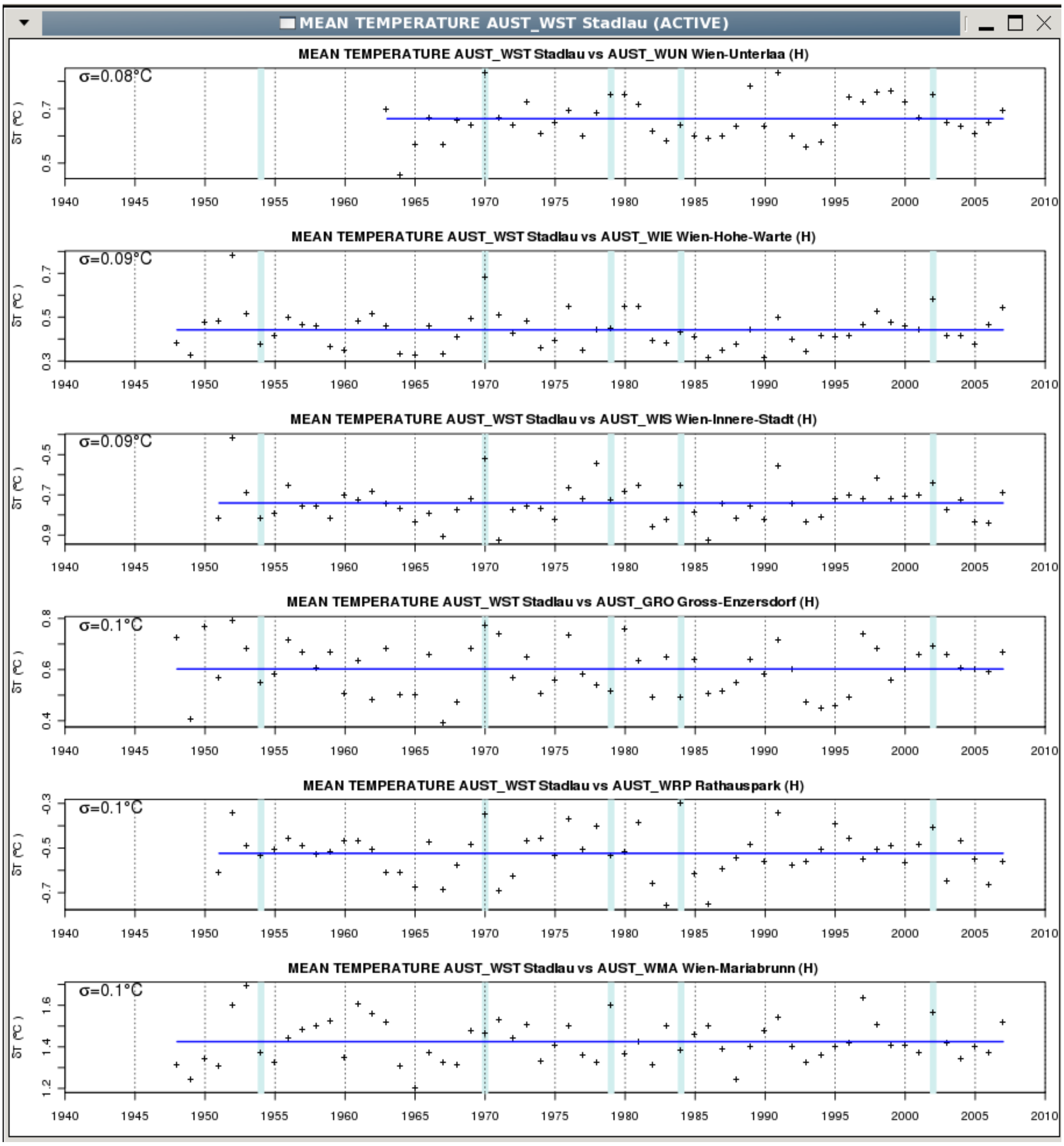

Fig. 6. The same as *Fig. 3*, but for the corrected Stadlau series compared to its corrected neighbors. The list of pairwise comparisons changed a little bit, since estimates of noise standard deviation slightly varied.

Another example of the effect of correction is shown for Rathauspark series (*Fig. 7* upper panel for the raw, lower panel for the corrected series).

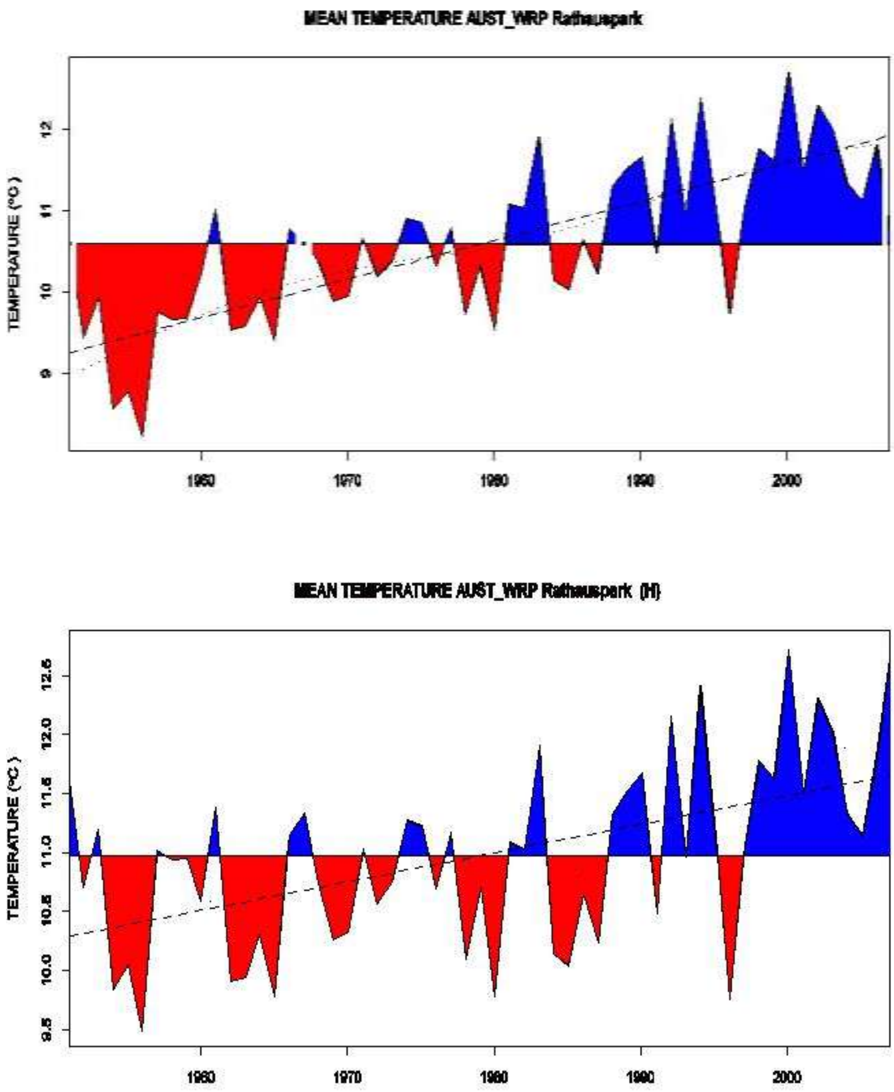

Fig. 7. Raw (up) and corrected (down) series of Rathauspark.

3.2. UBRIS characterization of urban trends

Running UBRIS allowed us to estimate jointly the effect of abrupt changes and the potentially significant urban trends on Wien series. UBRIS procedure is run in the following way: a first estimation allows us to put into evidence some urban series having no additional trend (large *p* values of the Student *t*-test for corresponding β). Those series are included into the rural set, and trends are re-estimated. At the end of the process, central temperature series exhibit no significant urban trends at level 0.05. Only suburban series (Wien Laaerberg, +0.10°C/decade, Rosenhügel +0.08°C/decade) exhibit significant positive trends (with student *t*-test *p* values lower than 10e–4). Corrected series of Laaerberg, with and without urban trend, is shown in *Fig. 8*.

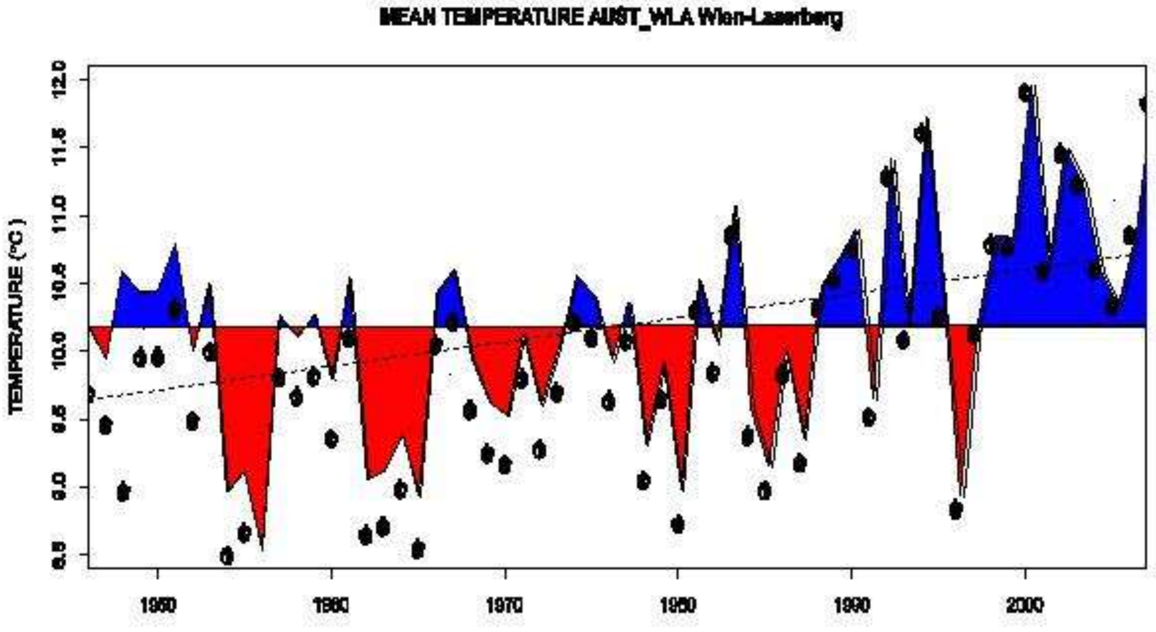

Fig. 8. Homogenized series of annual mean temperature of Laaerberg, with urban trend (series of ⊕ symbols) and removed urban trend(solid line).

These results are consistent with those obtained by *Böhm* (1998), who used a more traditional homogenization technique, and analyzed the trends of the series of differences of central series minus mean of the rural series. Note that those conclusions may not apply to other cities, since Wien population is remarkably stable since 1950 for example. UBRIS model should be run on each case study.

Additionally, Klosterneuburg series (not shown here) exhibits a remarkable feature, a highly significant decreasing trend for summer months (−0.02°C/decade). This site should be investigated for a potential shadowing effect.

4. Conclusion and perspectives

This paper presents a set of homogenization procedures integrated in the new software package HOMER (available at www.homogenization.org). This package was built relying on the results of the 4-year long COST-HOME project, so it implements the most significant findings achieved by its different working groups. The evolution of PRODIGE, combined with ACMANT and CLIMATOL procedures and supported by the R-package *cghseg* into HOMER provides a state-of-the-art homogenization tool for monthly to annual data applicable to most essential climate variables. However, HOMER shall not be considered as an automatic method, since manual input is still required in order to control the homogenization process.

HOMER is recommended by the COST Action ES0601, together with *Craddock* (1979), MASH (*Szentimrey*, 2007), USHCN (*Menne* and *Williams*, 2005), ACMANT (2011) software that got valuable results during COST benchmark experiments (*Venema et al.*, 2012).

The addition of UBRIS procedures adds value to the package since artificial trends have remained a problematic issue in homogenization.

Further development planned in this work is using a generalized least squares estimation for the correction model, in order to take into account the spatial dependency of the residuals. Although this technique is expected to have a weak effect on the correction estimates themselves, it may provide more accurate confidence intervals. A Bayesian criterion for automatic attribution of changes detected in pairwise comparison is also in development.

Acknowledgements–HOMER has been developed with support of the European Union, through the COST Action ES0601 – Advances in Homogenization Methods of Climate Series: an Integrated Approach (HOME).

References

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

Auer, I., Böhm, R., Jurkovic´, A., Orlik, A., Potzmann, R., Schöner, W., Ungersböck, M., Brunetti, M., Nanni, T., Maugeri, M., Briffa, K., Jones, P., Efthymiadis, d., Mestre, O., Moisselin, J.M., Begert, M., Bradzil, R., Bochnicek, O., Cegnar, T., Gagić-Ĉapka, M., Zaninović, K., Majstorović, Z., Szalai, S.,Szentimrey, T. and *Mercalli M.*, 2005: A new instrumental precipitation dataset for the greater Alpine region for the period 1800–2002. *Int. J. Climatol. 25*, 139–166.

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

Birgé, L. and *Massart, P.*, 2001: Gaussian model selection. *J. Eur. Math. Soc. 3*, 203–268.

Böhm, R. 1998: Urban bias in temperature time series – a case study for the city of Vienna, Austria. *Climatic Change, 38*, 113–128.

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

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

Caussinus, H. and *Mestre, O.*, 2004: Detection and correction of artificial shifts in climate series. *J. Roy. Stat. Soc. Series C53*, 405–425.

Craddock, J.M., 1979: Methods of comparing annual rainfall records for climatic purposes. *Weather 34*, 332–346.
Domonkos, P., 2011: Adapted Caussinus-Mestre Algorithm for Networks of Temperature series (ACMANT). *Int. J. Geosci. 2*, 293–309.
Domonkos, P., Venema, V., Auer, I., Mestre, O., and *Brunetti, M.*, 2012a: The historical pathway towards more accurate homogenization. *Adv. Sci. Res. 8*, 45–52.
Domonkos, P., Venema, V., and *Mestre, O.*, 2012b: Efficiencies of homogenization methods: our present knowledge and its limitation. *Proceedings of the 7th Seminar for Homogenization and Quality Control in Climatological Databases* in press.
Drogue, G., Mestre, O., Hoffmann, L., Iffly, J-F., and *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.
Ducré-Robitaille, J-F., Vincent, L.A., and *Boulet, G.*, 2003: Comparison of techniques for detection of discontinuities in temperature series. *Int. J. Climatol. 23*, 1087–1101.
Gu, C. and *Wang, J.*, 2003: Penalized likelihood density estimation: Direct cross-validation and scalable approximation. *Statistica Sinica 13*, 811–826.
Guijarro, J.A., 2011: User's guide to CLIMATOL. http://www.meteobal.com/climatol/climatol-guide.pdf
Hawkins, D.M., 1972: On the choice of segments in piecewise approximation. *J. Inst. Math. Appl. 9*, 250–256.
Hawkins, D.M., 2001: Fitting multiple change-points to data. *Comput. Statist. Data Anal. 37*, 323–341.
Lavielle, M., 1998: Optimal segmentation of random processes. *IEEE Trans. on Signal Proc. 46*, 1365–1373.
Menne, M. J. and *Williams, C.N.Jr.*, 2005: Detection of undocumented changepoints using multiple test statistics and composite reference series. *J. Climate 18*, 4271–4286.
Peterson, T.C., Easterling, D.R., Karl, T.R., 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, E. J., Hanssen-Bauer, I., Alexandersson, H., Jones, P., and *Parker, D.*, 1998: Homogeneity adjustments of in situ atmospheric climate data: a review. *Int. J. Climatol. 18*, 1493–1517.
Peterson, T.C., 2003: Assessment of Urban Versus Rural In Situ Surface Temperatures in the Contiguous United States: No Difference Found. *J. Climate 16*, 2941–2959.
Picard, F., Lebarbier, E., Hoebeke, M., Rigaill, G., Thiam, B. and *Robin, S.*, 2011: Joint segmentation, calling and normalization of multiple CGH profiles. *Biostatistics 12*, 413–428.
Schwartz, G., 1978: Estimating the dimension of a model. *Ann. Statist. 6*, 2745–2756.
Szentimrey, T., 2007: Manual of homogenization software MASHv3.02, Hungarian Meteorological Service.
Tibshirani, R., 1996: Regression shrinkage and selection via the lasso. *JRSS Series B 58*, 267–288.
Venema, V., Mestre, O., Aguilar, E., Auer, I., Guijarro, J.A., Domonkos, P., Vertacnik, G., Szentimrey, T., Štěpánek, P., Zahradnicek, P., Viarre, J., Müller-Westermeier, G., Lakatos, M., Williams, C.N., 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, M.P., Likso, T., Esteban, P. and *Brandsma,T.*, 2012: Benchmarking monthly homogenization algorithms, *Climate of the Past 8*, 89–115.
Zhang, N.R. and *Siegmund, D.O.*, 2007: A Modified Bayes Information Criterion with Applications to the Analysis of Comparative Genomic Hybridization Data. *Biometrics 63*, 22–32.