Theoretical and Applied Climatology (2019) 136:287–299
https://doi.org/10.1007/s00704-018-2488-3

ORIGINAL PAPER

Impact of missing data on the efficiency of homogenisation: experiments with ACMANTv3

Peter Domonkos[1] • John Coll[2]

Received: 21 September 2017 / Accepted: 18 April 2018 / Published online: 28 April 2018

Abstract

The impact of missing data on the efficiency of homogenisation with ACMANTv3 is examined with simulated monthly surface air temperature test datasets. The homogeneous database is derived from an earlier benchmarking of daily temperature data in the USA, and then outliers and inhomogeneities (IHs) are randomly inserted into the time series. Three inhomogeneous datasets are generated and used, one with relatively few and small IHs, another one with IHs of medium frequency and size, and a third one with large and frequent IHs. All of the inserted IHs are changes to the means. Most of the IHs are single sudden shifts or pair of shifts resulting in platform-shaped biases. Each test dataset consists of 158 time series of 100 years length, and their mean spatial correlation is 0.68–0.88. For examining the impacts of missing data, seven experiments are performed, in which 18 series are left complete, while variable quantities (10–70%) of the data of the other 140 series are removed.

The results show that data gaps have a greater impact on the monthly root mean squared error (RMSE) than the annual RMSE and trend bias. When data with a large ratio of gaps is homogenised, the reduction of the upper 5% of the monthly RMSE is the least successful, but even there, the efficiency remains positive. In terms of reducing the annual RMSE and trend bias, the efficiency is 54–91%. The inclusion of short and incomplete series with sufficient spatial correlation in all cases improves the efficiency of homogenisation with ACMANTv3.

1 Introduction

Modern climatology needs large observational climatic datasets of high quality (Thorne et al. 2017), and one issue is that some of these long climatic records are often affected by technical changes in the conditions of the climate observations, resulting in apparent temporal variations of climate. These changes are defined as inhomogeneities (IHs), and for the analysis of true climatic variability, their removal from the data is desirable in order to establish more accurate climate-based variations with other effects removed. Several factors can introduce IHs to climate time series (Aguilar et al. 2003; Auer et al. 2005). The most frequent form of IHs is a sudden change in the long-term means (break); this can occur due to station relocation, changes in the instrumentation or in observation practice, among other reasons. Urbanisation or other environmental changes in the surroundings of the observing station may introduce gradual, trend-shaped IHs (Jones and Lister 2010; Hausfather et al. 2013), while temporal deviations from the regular conditions of the observations may result in platform-shaped IHs (Rienzner and Gandolfi 2011; Domonkos 2011a).

The purpose of time series homogenisation is the removal of IHs from the data. As the data records of nearby observing stations often include a similar climate signal, the main tools of homogenisation include the exclusion of the common climate signal via spatial comparisons (referred to as relative homogenisation) and statistical tests. There are a large number of statistical homogenisation methods in use in climatology, and the use of statistical homogenisation methods is generally recommended (Peterson et al. 1998; Aguilar et al. 2003; Ribeiro et al. 2016). Relative homogenisation methods improve the homogeneity of data, whereas absolute homogenisation methods have the potential to make the data even more inhomogeneous (Acquaotta and Fratianni 2014; Hannart et al. 2014). Other tools for homogenisation are the documented information about changes in the conditions of

✉ Peter Domonkos
dpeterfree@gmail.com

[1] Unaffiliated, Tortosa, Spain

[2] Irish Climate Analysis and Research UnitS (ICARUS), Department of Geography, Maynooth University, Maynooth, Co Kildare, Ireland

observations (so-called metadata), and these are often used in tandem with statistical tests and a visual inspection of the data (Brunet et al. 2006; Mestre et al. 2013). Generally, the application of changes based on the statistical methods is given precedence over information from the metadata, even when the changes in observing networks are well documented. There are three main reasons for this: (i) Most pieces of metadata do not contain quantitative information about the size of the IH; (ii) Several IHs are unintentional (Thorne et al. 2016); hence, they are not well known and not included in metadata; (iii) The optimal use of metadata in homogenisation is subject to debate (Gubler et al. 2017). These issues notwithstanding the mix of tools applied in homogenisation practice depend primarily on the type of homogenisation task to be performed. In the case of homogenisation for data sparse regions or periods, the power of relative homogenisation is relatively low, and thus the information from metadata may be of key importance (Gimmi et al. 2007; Prohom et al. 2016). Whereas in the case of homogenising large and spatially dense datasets, automatic or semiautomatic statistical procedures are the more likely tools of choice (Menne and Williams 2009; Spinoni et al. 2015).

During the European project COST ES0601 (known also as "HOME", http://www.homogenisation.org), several statistical homogenisation methods were tested, and a surprisingly wide range of efficiencies were found in relation to both the test dataset properties and according to the homogenisation methods used (Venema et al. 2012). These indicate that it is important to learn more in relation to the method efficiencies, and arising from this to use homogenisation methods of proven high efficiency for climate data homogenisation.

In recently published efficiency tests of homogenisation methods, ACMANT (Domonkos 2011b; Domonkos and Coll 2017a) consistently produced the best results (Venema et al. 2012; Killick 2016; Guijarro et al. 2017). Additional advantages offered by ACMANT are that this method is fully automatic, easy-to-use and fast; therefore, it is easy to test it on large datasets.

Earlier tests have given relatively little information about the missing data effect on the efficiencies of homogenisation methods, despite the potential for missing data to affect climatological analyses by reducing the correlation coefficients of station pairs and hence affecting the performance of relative homogenisation methods (Hunziker et al. 2018). In this study, the missing data effect on the efficiency of ACMANTv3 will be analysed with the help of three monthly surface air temperature test datasets and with varied ratios of missing data. In our analyses, the temporal distribution of missing data follows a pattern frequently appearing in real observational datasets, namely an important ratio of the missing data is concentrated at the beginning of time series due to the varied starting date of observations.

2 Gap filling in statistical homogenisation procedures

Since data collected at all sites within the same climatic region is likely to be highly correlated and with similar patterns of temporal variability, it follows that a lot of missing data has the potential to affect these relationships even for stations in relatively close geographical proximity. One option to reduce the effects of missing data on the homogenisation results is to fill data gaps before homogenisation with interpolated values. However, any gap filling with interpolation before homogenisation results in an artificial multiplication of the same observational information within a region and also those of the IHs. Therefore, several experts advise not to fill data gaps before homogenisation (Auer et al. 2005, 2007; Guentchev et al. 2010). However, the use of composite reference series needs complete series at least for the period of the spatial comparison, since in case of incomplete reference composites, the reference series would have non-natural temporal variation in addition to IHs. Among modern homogenisation methods, the gap filling before homogenisation is obligatory in MASH (Szentimrey 1999), Climatol (Guijarro 2014) and ACMANT, and it is done automatically as part of the homogenisation procedure in the latter two methods, while the homogenisation can be done without gap filling in Pairwise Homogenisation Algorithm (PHA, known also as USHCN method, Menne and Williams 2009) and data gaps are filled at the end of the procedure in HOMER (Mestre et al. 2013). Homogenisation methods vary also in their ability to process datasets with larger ratios of missing data. Thus, ACMANT, Climatol and PHA, e.g., have a very high missing data tolerance, whereas the WMO Task Team on homogenisation guidelines on the tolerable missing data for the application of HOMER and MASH are 15 years and 30% respectively (WMO 2016).

Gap filling may be performed at either a daily or monthly time scale, but gap filling on a daily scale is more problematic due to the larger temporal and spatial variability of daily climatic data. Tests conducted by Kemp et al. (1983) showed that spatial interpolation is the best method for the gap filling of temperature series, even on a daily scale. However, the use of more climatic variables or information based on categorising the prevailing synoptic situation may improve the accuracy of spatial interpolation on a daily scale (Huth and Nemesova 1995; Oyler et al. 2015).

In this study, we focus on the treatments with monthly time series. For the gap filling of monthly series, spatial interpolation is clearly the best tool, but there is no consensus on the number of data from the neighbourhood which should be used (Domonkos and Coll 2017b, hereafter DC2017b). In applying any interpolation procedures, data of the closest stations and/or the data most highly correlated with the candidate station should generally be more highly weighted than the other data of the neighbourhood. Szentimrey (2010) recommends kriging as the theoretically most correct interpolation method. However, kriging uses a large number of parameters, which in practice may affect accuracy. For instance, the tests of Borges et al. (2013, 2014) showed that a regression based on geographical latitude, longitude and altitude and on the residual inverse distance to be the most accurate interpolation method for filling the gaps of seasonal precipitation series in Brazil. In ACMANTv3, the squared spatial correlations are used as weights in the interpolation of any climatic variable, as these correlations can be considered the combined information of the distance and other geographical characteristics.

Spatial correlations can vary seasonally between stations; hence, routines taking account seasonal variation would be useful for the calculation of interpolated values. However, in many instances, the common period of observations between the candidate series and the time series from neighbouring stations is simply too short to properly select and weight the data of the neighbourhood if the estimations of the spatial relationships are derived from annual data (Costa and Soares 2009). Therefore, we recommend the use of deseasonalised monthly data for the spatial interpolation, which although excluding any consideration of the seasonal variation among the spatial relationships, it provides a sufficient amount of data for the calculations in the vast majority of practical gap filling tasks. The gap filling with ACMANTv3 is described in Sect. 3.3 and Appendix 1.

3 Methods

3.1 Test datasets

In DC2017b, the effect of the number of partner series on the efficiency of homogenisation with the help of five test datasets (denoted there with A, B, C, D, E) was examined, and three of the five datasets (A, B, C) are used here with little modification.

The origin of the homogeneous base of our test datasets is the US benchmark dataset for testing the efficiency of daily temperature homogenisation methods (Willett et al. 2014). The data of "Wyoming 2" are used here, and both the source data and our homogeneous base consist of 158 time series. The original dataset contains daily mean surface air temperatures of 42 years long series (1970–2011), but 100 years long monthly mean temperature series (1901–2000) were derived from that with a simple method described in DC2017b, and these monthly data are used in this study. After the transformation to monthly series, IHs and outliers are randomly inserted into them. The sequence of IHs can be characterised as a limited random walk, since the cumulative impact of IHs (i.e. the cumulated bias from the true climate) is maximised during the dataset generation. Three datasets were generated (A, B and C): In dataset A, the frequency of IHs is low, and the mean size of biases is relatively small. In datasets B (C), the frequency of IHs is higher (the highest) and the mean size of biases is higher (the highest). The systematic trend bias caused by IHs is small in A and C, but pronounced in B. The only differences here from the versions used by DC2017b are that (a) all the datasets are 100 years long; (b) in dataset B, the frequency of platform-shaped IHs is 5 per 100 years; and (c) in dataset C, the size of platform-shaped IHs is elevated by 30% compared to the other IH types. More details of the test datasets are provided in DC2017b for the interested reader.

The mean of spatial correlations (r) is 0.88 in A and B, while it is 0.68 for C (Fig. 1).

3.2 Missing data fields

Each test dataset is examined with 7 kinds of mean missing data ratio, but at least 18 series from the 158 are always left complete, and these complete series offer a sufficient number of comparable synchronous data for the homogenisation of any series, except for a few cases of low spatial correlations in dataset C. This is since in the present analysis we do not want to confound the problem of the lack of synchronous data with that of using interpolated data. The availability of complete series does not reduce importantly the use of interpolated data, as partner series are selected firstly according to their spatial correlations and not according to the completeness of the series (Sect. 3.4 and Appendix 2).

Eighteen time series were arbitrary selected from the 158, and this subset remained unchanged in all experiments. Regarding the other 140 series of the datasets, they include 10, 20,…,70 percentage missing values on average in the 7 experiments conducted. The missing data ratio is then varied for time series within the same experiment (Table 1). In each experiment and for each time series, the monthly missing

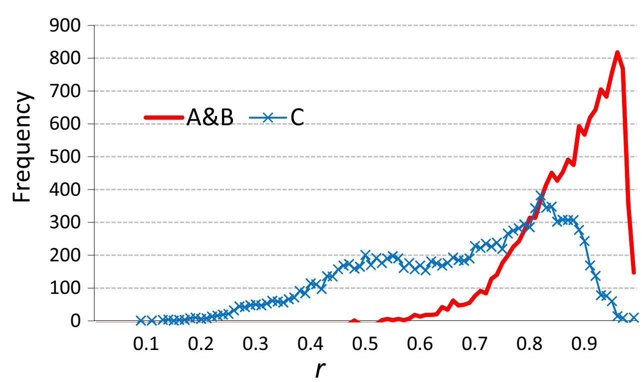

Fig. 1 Frequency distribution of spatial correlations in the three test datasets (A, B and C) in arbitrary units

Table 1 Missing data ratios in time series in the 7 experiments

Exp. 1		Exp. 2		Exp. 3		Exp. 4		Exp. 5		Exp. 6		Exp. 7	
N	Ratio	*N*	Ratio	*N*	Ratio	*N*	Ratio	*N*	Ratio	*N*	Ratio	*N*	Ratio
18	0	18	0	18	0	18	0	18	0	18	0	18	0
55	0	55	0	15	0	15	0	30	20	15	20	15	40
40	10	40	20	40	10	30	20	40	40	30	40	15	60
30	20	30	40	30	30	50	40	40	60	40	60	55	70
15	30	15	60	15	40	30	60	30	80	55	80	55	80
				25	50	15	80						
				15	70								

Exp. experiment, *N* number of time series, *ratio* missing data ratio in percentage

values are positioned within the 100-year period according to the following rules: (i) 50% of the missing values are entered at the beginning of the period; 20% of the missing values are entered at the end of the period; 20% of the missing values are positioned in data gaps of entire years, but otherwise the dates are selected randomly; and 10% of the missing values are positioned fully randomly. For instance, the positions of 80% missing values within 1901–2000 are as follows: the data segments for the periods 1901–1940 and 1985–2000 are completely missing, the data of a further 16 years are completely missing between 1941 and 1984, and there are a further 96 monthly missing values in the remaining 28 years. Although in real observational time series, the extent of missing data at the beginning and end sections of time series is highly varied; the distribution of missing data in our datasets is representative of a typical pattern of the temporal evolution of data availability from stations. Thus, there is an initial sharp increase in data availability, followed by a relatively stable, low proportion of missing values, and then latterly a tendency for growing ratio of missing values to be present. This pattern has been observed to be the case for data networks in the USA (Menne et al. 2009), in China (Hua et al. 2017), in Europe (Klok and Klein Tank 2009; Coll et al. 2018) and also in global datasets (Thorne et al. 2016).

3.3 Homogenisation and gap filling with ACMANTv3

We use ACMANTv3 in all of the experiments. ACMANT is a modern, multiple break homogenisation method in the sense that it detects jointly the breaks of time series and calculates jointly the correction terms for all the IHs of a climatic network whose data are homogenised together. While the first version of ACMANT (Domonkos 2011b) was usable only for the homogenisation of monthly temperatures of mid- and high-latitudes, the most recent version (Domonkos and Coll 2017a) is usable for temperature and precipitation homogenisation of any climatic region and for data at both daily or monthly time resolution.

The accuracy of homogenisation has also been improved with the release of newer versions. For instance, efficiency improving novelties in ACMANTv3 compared to earlier versions are the calculation of monthly temperature adjustment terms with irregular seasonality for temperature minimums in general and for any surface air temperature characteristic of tropical or monsoon regions. Other new features include an ensemble pre-homogenisation and the application of ordinary kriging for setting the weights of the reference composites.

Other than for extreme cases, the missing data tolerance of ACMANTv3 is practically unlimited, and for the interested reader, more details are in the Manual (http://www.c3.urv.cat/softdata.php).

The gap filling for a missing monthly value of a candidate series is performed by using the weighted average of observed values synchronous with the missing value, but the values are tuned to a section mean value characterising both the common effect of local climate and the possible inhomogeneity effect around the missing value in the candidate series (Appendix 1). In the best scenario, the interpolated value retains the same bias from the true climate as the temporally closest observed values of the candidate series. To achieve this, relatively narrow windows are favoured in the estimation of the mean difference between the mean temperature of the candidate station and that of the other stations. The optimisation of sample size at the proximity of data gaps was examined by Tardivo and Berti (2012) in a partly similar interpolation task. In our case, the optimisation of sample size was part of the development of ACMANTv3.

The mean annual cycle is removed from all-time series before gap filling. Gap filling routines are performed three times within a homogenisation procedure, first with the use of raw data, and then in the second and third stages of the procedure with the use of pre-homogenised data. Any observed value synchronous with the missing data is used in the interpolation when the spatial correlation between the time series of the observed value and the candidate series is at least 0.4. More details of the gap filling procedures in ACMANTv3 are provided in Appendix 1.

3.4 Automatic networking

Relative homogenisation can be performed within networks of the same climatic region, and to enable this, large datasets are often split into smaller subsets. In addition, DC2017b showed that the optimal number of partner series for homogenising a candidate series with ACMANTv3 is about 30, and that the accuracy of homogenisation cannot be improved with the inclusion of a larger number of partner series, even in situations where the spatial density of data and a similarity of climate would allow this. While some homogenisation software packages include the construction of climatic networks, e.g. Climatol and PHA, the construction of a suitable climatic network is a separate preparatory task for homogenisation with ACMANT.

In DC2017b, distinct networks were constructed for each candidate series, with the inclusion of the other time series most closely correlated with the candidate series as the reference series in each case. However, the task encountered here is somewhat more complicated, as missing data can markedly lessen the number of the truly comparable observed data. Accordingly, we introduce the concept of “number of effective partners”, which indicates the temporally and spatially usable observed values of partner series for any given month of the candidate series. During the automatic network construction applied here, the increase of the number of effective partners and the inclusion of highly correlated partner series are favoured, while any excess in network size is penalised (more details are in Appendix 2).

Figures 2 and 3 show the distribution of the number of partner series and that of the effective partners in the experiment with a 70% missing data ratio for datasets A and C. When the spatial correlations are higher (dataset A), the number of partner series has little variation (i.e. it varies between 42 and 50), while in dataset C, some candidate series have much less partner series due to insufficient spatial correlations. The difference of the spatial correlations between the two datasets has consequences also for the number of effective partners. In the case of dataset A, each date of each candidate series has at least 17 effective partners, as the 18 complete series of the dataset are well correlated. There is a high peak at 18, since for the starting and ending sections of many series just the values of the 18 complete series act as effective partners. By contrast, with dataset C, the number of effective partners is sometimes much smaller than for dataset A. Nevertheless, even for dataset C and for a 70% mean missing data ratio with the exception of the 18 complete series, the number of effective partners is at least 10 in 96.0% of the cases, and at least 15 in 91.8% of the cases.

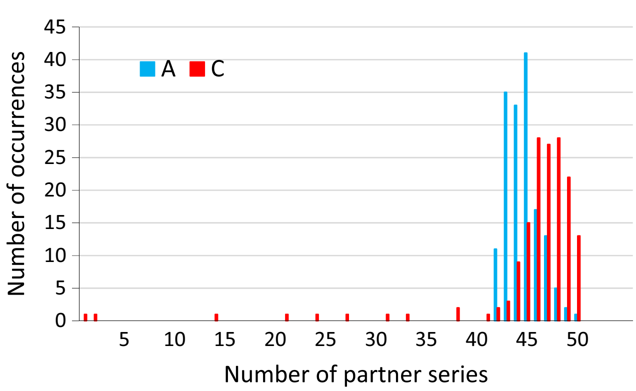

Fig. 2 Number of partner series for the candidate series of datasets A and C

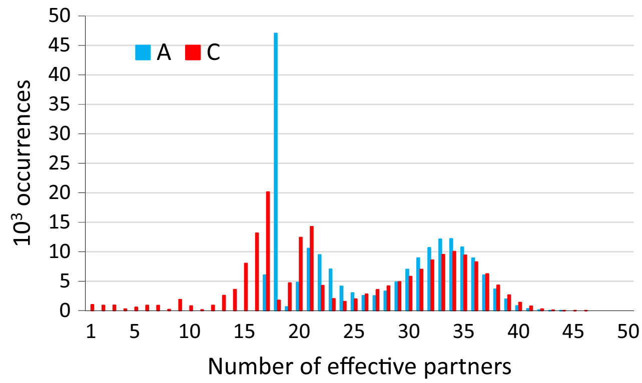

Fig. 3 The distribution of the number of effective partners for the months of candidate series in datasets A and C

3.5 Efficiency measures

The efficiency of homogenisation is evaluated with the ratio of the removal of raw data error. Three kinds of errors are monitored, namely the centred root mean squared error (RMSE) for monthly values, the RMSE for annual values, and the bias of the linear trend for time series between the first and the last observed values in them. These error terms are calculated in the same way as in the closing study of HOME (Venema et al. 2012).

The connection between raw data error (f_{raw}), error of the homogenised data (f_{hom}) and efficiency (E) is shown by (1)

$$E = \frac{f_{raw} - f_{hom}}{f_{raw}} \quad (1)$$

If the homogenisation is perfect then $E = 1$, if it is neutral then $E = 0$ and when the homogenised data have a bigger error than the raw data, E is negative. The efficiency can be calculated for any error term and for any characteristic of the error distribution. In this study, the arithmetic average of errors and the 0.95 value of the

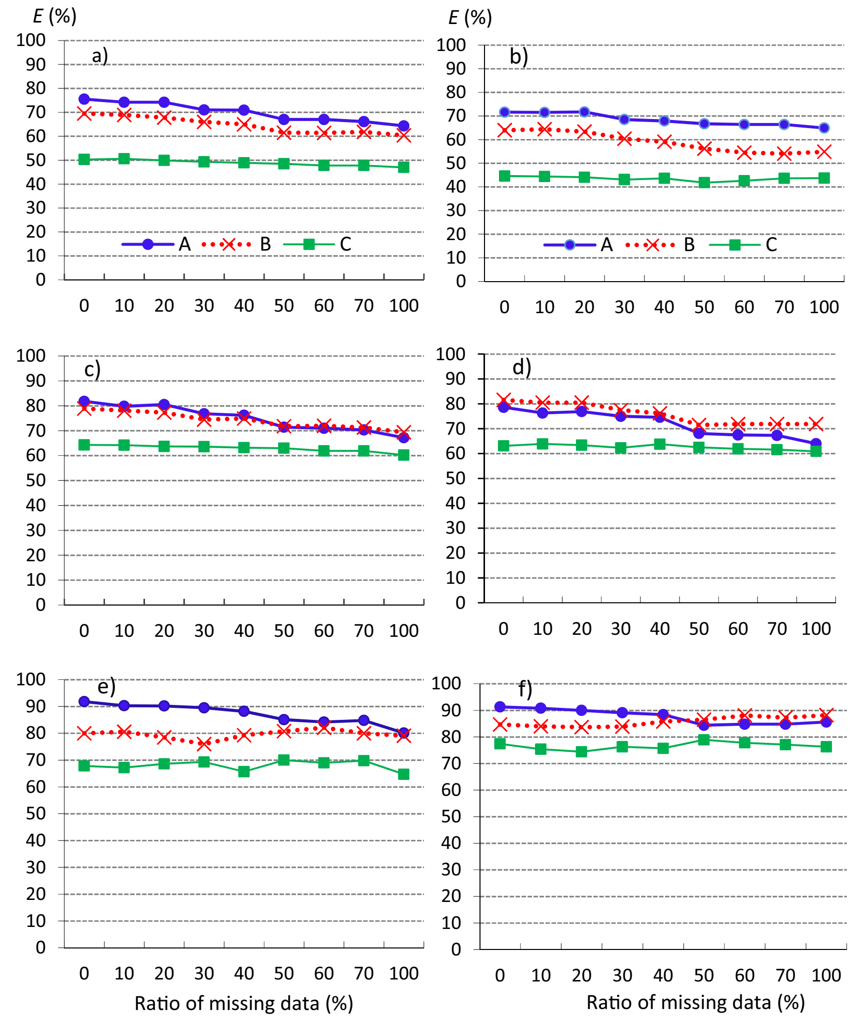

Fig. 4 Efficiency of ACMANTv3 in reducing the errors of the 18 complete series. Upper panels (a, b): monthly RMSE, middle (c, d): annual RMSE, bottom (e, f): linear trend bias for the entire period. Left panels (a, c, e): mean error, right panels (b, d, f): P95 of errors. 100% missing data means that the 18 complete series are homogenised without the use of the data of other series

probability distribution function (P95) will be shown, the latter is chosen for examining the efficiency in reducing the threshold for the largest 5% of errors.

4 Results

4.1 Effect of missing data in the partner series on the homogenisation results of the complete series

Figure 4a, b shows the effect of missing data on the monthly RMSE reduction of the 18 complete series. It can be seen that the inclusion of partner series improves the efficiency even when the missing data ratio is high in them, except for P95 in datasets B and C. In the latter two cases, the use of complete series gives slightly better results than some of the experiments where incomplete series are included, but these differences are very small. The efficiency is 50–80% for datasets A and B and 40–50% for dataset C. Surprisingly, the results of dataset C, particularly those for P95, have a low dependence on the missing data ratio of the partner series.

Figure 4c, d shows similar results for the annual RMSE to the results for monthly RMSE, except that the efficiencies here always reach or exceed 60%. The efficiencies for dataset B are often higher than those for dataset A, despite the fact that the IH structure in dataset B is more complex than that in dataset A. The inclusion of partner series with any ratio of missing values impacts positively the results in all the datasets, and this applies both for the means and for the thresholds of the largest errors. The efficiencies are much less dependent on the ratio of missing data in the results of dataset C than in the results of the other datasets.

The efficiency of removing trend bias is generally higher than that of removing RMSE (Fig. 4e, f). The inclusion of partner series with any ratio of missing values reduces the mean residual trend bias in dataset A and C, while the impact is less clear for dataset B and for the P95 values. There does not seem to be any clear impact of the missing data ratio either, except in the results of dataset A.

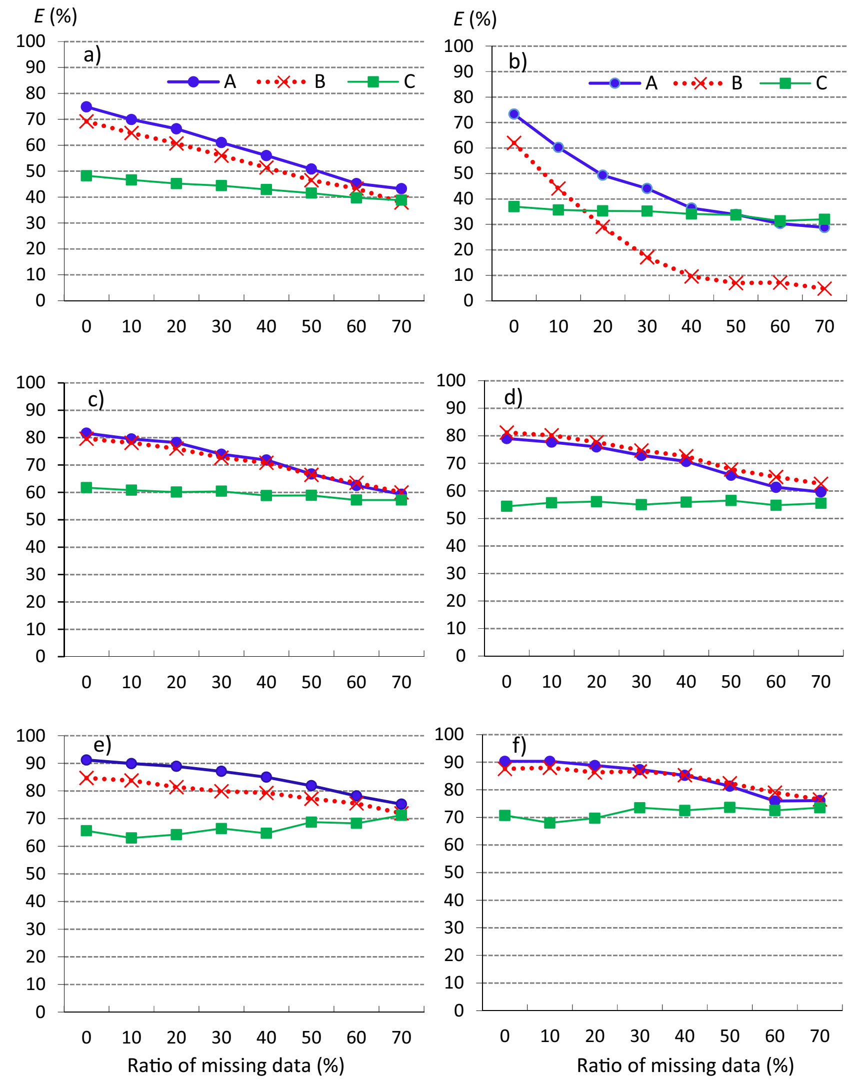

Fig. 5 Efficiency of ACMANTv3 in reducing errors in all of the time series. The explanation of the panels is the same as for Fig. 4

4.2 Overall effect of missing data on the homogenisation results

Figure 5a, b presents the results of the mean reduction of monthly RMSE for all of the 158 series. The efficiencies are markedly lower here than for those presented in Fig. 4, and demonstrate that the removal of monthly RMSE is less successful in time series with data gaps than in complete series. The order of efficiencies according to test datasets is A, B and C for the mean errors, with a sharp decline being observed in the differences according to the test dataset used when the missing data ratio increases. Regarding the P95 values, the efficiencies decline rapidly associated with the rising missing data ratio in dataset A and B, whereas they remain moderately low in dataset C. The results indicate that the largest 5% of monthly RMSE for dataset B is not much smaller in the homogenised data than in the raw data when the missing data ratio is larger than 20%.

The efficiency in reducing annual RMSE (Fig. 5c, d) is generally much higher than in reducing monthly RMSE. Similar to the case of monthly RMSE reduction, the decline of efficiency with increasing missing data ratio is more pronounced than for the complete series (Fig. 4c, d) in datasets A and B, while the missing data ratio is nearly neutral to the efficiencies in dataset C. When the missing data ratio is high, the difference of efficiencies according to test datasets is small.

The efficiency in the removal of trend bias (Fig. 5e, f) is generally higher than in the removal of annual or monthly RMSE, and it is always higher than 60% in the experiments performed. While the efficiencies clearly decline with increasing missing data ratio for datasets A and B, a slight opposite tendency is observed in the results for dataset C. When the missing data ratio is high, the difference of efficiencies according to test datasets is small, similar to the efficiencies in reducing annual or monthly RMSE.

5 Discussion

5.1 The impact of using short and incomplete series in homogenising time series

The results show that the inclusion of time series even with large missing data fields in them almost always improves the efficiency, and no significant decrease of efficiency for this reason was observed during the experiments. All these indicate that in spite of at least 18 complete series being included in each experiment, the inclusion of the available short and incomplete time series is always advisable in homogenising with ACMANTv3, as long as the spatial correlations remain satisfactory (i.e. they are not lower than 0.4). Note that in several other homogenisation methods, short and incomplete time series, as well as series with moderately high spatial correlations are often excluded from network construction and reference series selection, and hence the number of partner series is sometimes heavily restricted (DC2017b and references therein).

In homogenising complete candidate series with partly incomplete series, the efficiency decreases slightly with an increasing missing data ratio, but the efficiency generally remains higher than if the incomplete series was excluded. In homogenising incomplete series, the decrease of efficiency with increasing missing data ratio is faster than is the case for complete series, particularly in the reduction of monthly RMSE. Nevertheless, the efficiency has remained positive in all the experiments, and it is always above 50% for annual RMSE and trend bias.

5.2 Differences between the monthly RMSE and annual RMSE results

The reduction of trend bias and annual RMSE is generally more successful than the reduction of monthly RMSE. This feature is not specific to ACMANT; rather, it is the case for statistical homogenisation methods generally (Venema et al. 2012). It is remarkable that in the experiments presented here, the decline of efficiency with an increasing missing data ratio is much more rapid for monthly RMSE than for annual RMSE. As a consequence, the uniform application of annual correction terms to each calendar month can be more effective than the derivation and application of month-specific correction terms. However, the raw monthly RMSE is larger than the raw annual RMSE (Table 2). We have not undertaken experiments applying seasonally uniform downscaling of annual correction terms, but gross estimations about the relation of the efficiencies can be obtained by relying on the results presented in Fig. 5 and the data in Table 2. According to these, the use of month-specific correction terms is the most efficient for reducing the mean monthly RMSE in datasets A and B, but it seems slightly less effective for dataset C where the annual cycle of biases is faint and irregularly shaped. By contrast, the downscaling of annual correction terms applied in a seasonally consistent fashion is the most effective for reducing the largest few percentages of monthly RMSE in series with missing data, and its advantage in this respect is notable.

Table 2 Ratio of monthly and annual RMSE for the raw and complete datasets

	Ratio of means	Ratio of P95
Dataset A	1.628	1.567
Dataset B	1.819	1.453
Dataset C	1.268	1.118

The reduction of monthly RMSE is impacted more by the presence of missing data than that of the annual RMSE, as the interpolation errors of monthly values act as additional noise on a monthly scale; hence, their impact is attenuated on longer time scales. The signal to noise ratio generally influences the efficiency of homogenisation, as it is shown for the estimation of break positions (Lindau and Venema 2016), for the accuracy of monthly homogenisation results (Venema et al. 2012) and for the estimation of daily correction terms (Mestre et al. 2011). The results presented here emphasise that although several statistical techniques are known for the estimation of correction terms with high temporal resolution, their practical benefit depends on several factors, first of all on the signal to noise ratio.

5.3 Variation of results according to test datasets and missing data ratio

In dataset B, the characteristic IH size is larger than in dataset A, which favours the signal to noise ratio in dataset B, but the higher frequency and more complicated structure of IHs in dataset B makes the homogenisation of this dataset more difficult. These favourable and unfavourable factors apparently largely neutralise their common effect, as the efficiencies for dataset A and dataset B are similar with slightly better results for dataset A. The characteristic IH size is even larger in dataset C than in dataset B, but there the spatial correlations are lower and also the 15 per 100 years frequency of short-term IHs acts as a larger noise component; hence, the efficiencies are generally smaller for dataset C.

Interestingly, the efficiencies for different datasets seem to converge with an increasing missing data ratio, as efficiencies decline faster with an increasing missing data ratio for datasets A and B than for dataset C. It should also be noted that such convergence does not appear in the efficiencies for complete time series. Often, there is no significant decline of efficiency for dataset C with an increasing missing data ratio, and exceptionally, the efficiency slightly increases with a higher missing data ratio in Fig. 5e, f. We cannot explain the reason for this paradox, perhaps the shortening of the periods with observed data in experiments with large missing data fields reduces the frequency of some inhomogeneity problems in dataset C.

The homogenisation results in a much smaller improvement in the largest monthly RMSE of dataset B than in any other efficiency characteristic examined, and this case is also exceptional in the sense that here the results are poorer for dataset B than for dataset C. A possible explanation of the change of the rank order here is that the estimation of a sinusoid shaped annual cycle of correction terms is more sensitive to the errors associated with monthly data than for that of the irregular shaped annual cycle, as the latter includes ensemble calculations and smoothing (Domonkos and Coll 2017a).

5.4 Representativeness of the results

The results of the study characterise the efficiency of ACMANTv3 in homogenising monthly temperature series from data dense areas of northern, continental mid-latitude regions, but they are also at least indicative for the efficiency of a large number of other kinds of homogenisation tasks.

In each experiment presented here, at least 18 series are complete, and this situation provides good conditions for gap filling and relative homogenisation, anything is the missing data ratio in the other series, although note that in dataset C not all of the 18 complete series are spatially correlated with each other. For sparser data in less dense networks, the efficiencies would likely be lower.

Seventy percent of the missing values are positioned at the beginning or end of the time series (outer missing values); hence, they shorten the effective length of time series, while the rest of the missing values are positioned between observed values (inner missing values). The effects of outer missing data and inner missing data are not the same on the efficiency of ACMANTv3, as in most steps of the homogenisation, only the inner data gaps are filled with interpolated values. Considering the ratios of inner data gaps relative to the effective lengths of time series with missing data in the seven experiments, they are 3, 7, 11, 17, 23, 31 and 41%. A larger ratio of inner data gaps than for those in the experiments performed here would likely enhance the missing data effect on the efficiency.

It is very likely that our results also fairly characterise the efficiency of daily temperature homogenisation with ACMANTv3, as although the interpolation is performed on a daily scale, most of the steps of the homogenisation are still performed at an annual or monthly scale in daily homogenisation with ACMANTv3.

The results are indicative also as to the likely efficiency of temperature homogenisation for tropical regions. Although the spatial structures of climate are different in the tropics compared with those of the mid-latitudes, these differences are likely to have little effect on the efficiencies. Similarly, the results are transferable to the likely efficiencies of precipitation homogenisation with ACMANTv3, as the solution of the homogenisation task is similar for precipitation and temperature (Domonkos 2015).

Regarding the missing data effect on efficiencies using other homogenisation methods, the representativeness of the results presented here are uncertain as other homogenisation methods use other gap filling techniques, and the organisation of gap filling—homogenisation sequence(s) can be different from those of ACMANTv3. It would be useful to see the results of similar experiments with other high-performance homogenisation methods such as Climatol or MASH, but these comparisons are beyond the scope of this study.

6 Conclusions

The impact of missing data on the efficiency of ACMANTv3 in homogenising monthly air surface temperature series has been examined using three realistic test datasets. Each test dataset consisted of 158 time series of 100 years length, from which 18 series were left complete in all the experiments. The mean missing data ratio for the other 140 series was varied between 10 and 70%. The efficiency was measured via the reduction of the monthly and annual RMSE and the reduction of trend bias by homogenisation. The main conclusions are as follows.

- All the results showed positive efficiency, and the reduction of the raw data error is mostly larger than 50%. Short and incomplete series can be safely homogenised with ACMANTv3 if the spatial density of data satisfies the minimum conditions set for the method.
- Although data gaps cause some decrease in the efficiency, the inclusion of short and incomplete series as partner series has an overall positive impact on the efficiency of homogenisation.
- For time series with missing data, the reduction of trend bias and annual RMSE is much more successful than the reduction of monthly RMSE. For reducing the largest monthly RMSE, the use of uniform monthly correction terms is the most efficient.
- The results of the study provide an indication of the efficiency of ACMANTv3 in relation to various homogenisation tasks.

As ACMANT has most often been ranked at the first place in recent international efficiency tests of homogenisation methods, we recommend the use of ACMANTv3, first of all to the homogenisation of large datasets, and to the homogenisation of data sets of any size when metadata is not provided.

Acknowledgements The authors thank Kate Willett and her colleagues for giving open access to the temperature database they developed.

Funding information The second author was funded by the Irish Environmental Protection Agency under project 2012-CCRP-FS.11.

Appendix 1

I Gap filling within ACMANTv3

i) Concepts and definitions

The dataset consists of N monthly time series of n years length, but the series are incomplete. Series s ($s = 1,2,\ldots N$) can be presented as

$$\mathbf{X_s} = x_{s,1}, x_{s,2}, \ldots x_{s,h} \ldots x_{s,12n} \tag{A1}$$

without indicating possible data gaps, h stands for the serial number of month from the beginning of the time series. The relation between h, the serial number of year from the beginning of time series i and the serial number of calendar month m is.

$$h = 12(i-1) + m \quad i \in \{1, 2, \ldots n\} \tag{A2}$$

We will denote the cluster of observed values (distinguishing from missing values) of series s with J_s, and its sub-cluster for month m with $J_{s,m}$, and the number of elements in there with K_s and $K_{s,m}$, respectively.

Before any other operation with the data, the seasonal cycle is removed by extracting the monthly climatic normal ($U_{s,m}$) from the observed values, then the deseasonalised series are denoted with $\mathbf{A_s}$ (Eqs. A3–A5).

$$\mathbf{A_s} = a_{s,1}, a_{s,2}, \ldots a_{s,h} \ldots a_{s,12n} \tag{A3}$$

$$a_{s,h} = x_{s,h} - U_{s,m} \tag{A4}$$

$$U_{s,m} = \frac{1}{K_{s,m}} \sum\nolimits_{J_{s,m}} x_{s,h} \tag{A5}$$

Missing data of a candidate series ($\mathbf{A_g}$) will be filled with interpolation using the synchronous values of partner series ($\mathbf{A_s}$). In the selection of partner series and for weighting their contribution, spatial correlations are considered.

The spatial correlation between series g and series s ($r_{g,s}$) is defined by the Pearson correlation coefficient. The sample size for its calculation includes each h for which both series have observed values. When the sample size is lower than 50, the correlation is zero by our definition.

ii) Gap filling

The method of gap filling has remained similar to that in the first version of ACMANT (Domonkos 2011b), but some details have been changed since then.

The interpolation for a missing value of month $h0$ in the candidate series relies on the synchronous observed values of surrounding stations, but the values of the partner series are tuned to a section mean value characterising the common effect of local climate and possible inhomogeneity effects at the timing of the missing value in the candidate series. For this purpose, window $[h_1, h_2]$ around $h0$ is constructed. This window must be wide enough to have sufficient data for the reliable estimation of the difference between the section means for the candidate series $\mathbf{A_g}$ and its partner series $\mathbf{A_s}$, but narrow enough to exclude the effects of temporally distant IHs.

The window width can be regulated by parameterising it directly, or via the minimum number of the value pairs for series g and s. In practice, the window width is varied according to the frequency of missing data around $h0$ in the time series participating in the interpolation, hence $h_2 - h_1$ is functions of both h and s. Table 3 shows various sets of conditions for the window constructions in terms of the half window width (L) and the number of observed value pairs within the window (k). Moving down in Table 3 the conditions soften, and always the strictest conditions allowed by data availability are selected for the window construction.

All the series with $r_{g,s} \geq 0.4$ are considered as partner series, if they have observed value for month $h0$. Eq. (A6) shows the tuning of value $a_{s,h0}$ to the candidate series.

$$a'_{s,h0} = a_{s,h0} + \frac{1}{k} \sum\nolimits_{h=h_1}^{h_2} \left(a_{g,h} - a_{s,h}\right) \tag{A6}$$

When at least one of the two series do not have observed data (h is not $\in J_g \cap J_s$), then $a_{s,h} = a_{g,h}$ in (6) by definition. The interpolated value will be the weighted average of the tuned values of N^* partner series ($N^* \leq N - 1$). The weights are the squared spatial correlations corrected by coefficient c depending on the window width (Table 3).

Table 3 Connections between half window width (L), number of observed value pairs for the candidate series and its partner series (n') and coefficient of weight correction (c) in the construction of window around the timing of the missing data (h). Always the strictest conditions allowed by data availability are applied

First round			Later rounds		
L (years)	n'	c	L (years)	n'	c
< 5	100	1.0	< 12	100	1.0
5	30≤	1.0	12	30≤	1.0
< 12	30	0.8	< 25	30	0.9
No limit	30	0.5	No limit	30	0.7

When the sum of the corrected weights (p) is lower than 0.4, zero anomaly ($a_{g,h}=0$) is presumed for the missing value with a certain or entire weight, according to Eqs. (A7) and (A8).

$$a_{g,h0}=\frac{1}{p}\sum_{s=1}^{N^*}c_s r_{g,s}^2 a'_{s,h0} \quad \text{(A7)}$$

$$p=\max\left(0.4,\sum_{s=1}^{N^*}c_s r_{g,s}^2\right) \quad \text{(A8)}$$

Note that the optimal sample for the calculation of U and r (Huang et al. 1996; Tardivo 2015) may also differ from the sample including all available data and used in this study. However, this difference from the optimal parameterisation likely has a minor effect on the accuracy of interpolation.

The gap filling is performed three times within the homogenisation, first with the use of raw data, then with the use of pre-homogenised data in the second and third stages of the procedure.

Appendix 2

II Automatic networking

Appropriately constructed networks for homogenisation with ACMANT have three positive attributes: (i) The candidate series have a sufficient number of highly correlated partner series; (ii) Each section of the candidate series is covered with a sufficient number of synchronous observed data of the partner series; (iii) There is no unnecessary excess of the network size. The algorithm presented here is structured to give solutions with these positive attributes.

The number of partner series and the number of effective partners (see its definition in Sect. 3.4) are denoted with M and F, respectively. The spatial correlations used here (r^*) are not the same as those which are used for the interpolation, namely r^* is calculated from the first difference (increment) series of the deseasonalised monthly temperatures, and following from how it was introduced to time series homogenisation by Peterson and Easterling (1994).

For the homogenisation of each candidate series, one distinct network is constructed. First, the most highly correlated partner series are selected up to 30 series. When the number of potential partner series with $r^*\geq 0.4$. is higher than 30, the following steps are performed recursively.

i) Possible improvements in F by the inclusion of any further partner series (s) are considered using score S1 for dates of $F<10$ (Eqs. A9, A10).

$$\mathrm{S1}(s)=\sum_{h=1}^{12n}5r_s^{*4}\left(12-F_{s,h}^*\right)^3 \quad \text{(A9)}$$

$$F^*=\begin{Bmatrix} F \text{ if } F<10 \\ 12 \text{ if } F\geq 10\end{Bmatrix} \quad \text{for every } s \text{ and } h \quad \text{(A10)}$$

ii) Possible improvements in F by the inclusion of any further partner series are considered using score S2 for decadal sections of the candidate series where $F<10$ in at least 25% of the decade. Months belonging to at least one of such decades are denoted with m in Eqs. A11-A12.

$$\mathrm{S2}(s)=\sum_m 5r_s^{*4}\left(20-F_{s,m}^{**}\right)^2 \quad \text{(A11)}$$

$$F^{**}=\begin{Bmatrix} F \text{ if } F<20 \\ 20 \text{ if } F\geq 20\end{Bmatrix} \quad \text{for every } s \text{ and } m \quad \text{(A12)}$$

iii) The exceedance of M above 30 is penalised with score S3 (Eq. A13).

$$\mathrm{S3}=(M-30)^2 \quad \text{(A13)}$$

iv) Summarised score S is calculated for each s (Eq. A14).

$$\mathrm{S}(s)=\mathrm{S1}(s)+\mathrm{S2}(s)-\mathrm{S3} \quad \text{(A14)}$$

Then the series with the highest S is selected, and the procedure continues with step i.

If $\mathrm{S}\leq 0$ for all s, no further partner series is selected, and the procedure terminates.

The development of this algorithm is based on subjective decisions, but the important elements of the procedure can be reasoned well. Frequent occurrence of low F within a relatively short period is considered more destructive to the efficiency of homogenisation than its sporadic occurrences; therefore, higher minimum threshold of F is applied in S2 than in S1. It is more important to raise the smallest F values (if it is possible) than to raise a large number of F values, that is why the second factors of (A9) and (A11) are assigned higher powers. When more series are comparably useful in raising F, it is important to give preference to the one with relatively high correlation, therefore the power of r^* is raised by 4. Note that some parameter values are close to those of the networking in PHA (Menne and Williams 2009), as in PHA the 40 best correlating partner series are taken at the first step, and the correlation threshold is 0.5.

References

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

Aguilar E, Auer I, Brunet M, Peterson TC, Wieringa J (2003) Guidelines on climate metadata and homogenization. World Meteorological Organization (WMO)-TD No. 1186, WCDMP No. 53, Geneva, Switzerland, 55 pp

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, Brazdil R, Bochnicek O, Cegnar T, Gajic-Capka M, Zaninovic K, Majstorovicp Z, Szalai S, Szentimrey T, Mercalli L (2005) A new instrumental precipitation dataset for the Greater Alpine Region for the period 1800–2002. Int J Climatol 25:139–166. https://doi.org/10.1002/joc.1135

Auer I, Böhm R, Jurkovic A, Lipa W, Orlik A, Potzmann R, Schöner W, Ungersboeck M, Matulla C, Briffa K, Jones P, Efthymiadis D, Brunetti M, Nanni T, Maugeri M, Mercalli L, Mestre O, Moisseline JM, Begert M, Muller-Westermeier G, Kveton V, Bochnicek O, Stastny P, Lapin M, Szalai S, Szentimrey T, Cegnar T, Dolinar M, Gajic-Capka M, Zaninovic K, Majstorovic Z, Nieplova E (2007) HISTALP—historical instrumental climatological surface time series of the Greater Alpine Region HISTALP. Int J Climatol 27:17–46. https://doi.org/10.1002/joc.1377

Borges P, Franke J, Tanaka M, Weiss H, Bernhofer C (2013) Spatial interpolation of climatological information: comparison of methods for the development of precipitation distribution in Distrito Federal, Brazil. Atmos Clim Sci 3(2):208–217. https://doi.org/10.4236/acs.2013.32022

Borges P, Franke J, Santos Silva FD, Weiss H, Bernhofer C (2014) Differences between two climatological periods (2001–2010 vs. 1971–2000) and trend analysis of temperature and precipitation in Central Brazil. Theor Appl Climatol 116:191–202. https://doi.org/10.1007/s00704-013-0947-4

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

Coll J, Curley M, Walsh S, and Sweeney J (2018) HOMERUN: relative homogenisation of the Irish precipitation network. EPA Research Report 2012-CCRP-FS.11 Report No. 242. Environmental Protection Agency, Wexford, pp32

Costa AC, Soares A (2009) Homogenization of climate data: review and new perspectives using geostatistics. Math Geosci 41(3):291–305

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

Domonkos P (2011b) Adapted Caussinus-Mestre algorithm for networks of temperature series (ACMANT). Int J Geosci 2:293–309. https://doi.org/10.4236/ijg.2011.23032

Domonkos P (2015) Homogenization of precipitation time series with ACMANT. Theor Appl Climatol 122:303–314. https://doi.org/10.1007/s00704-014-1298-5

Domonkos P, Coll J (2017a) Homogenisation of temperature and precipitation time series with ACMANT3: method description and efficiency tests. Int J Climatol 37:1910–1921. https://doi.org/10.1002/joc.4822

Domonkos P, Coll J (2017b) Time series homogenisation of large observational datasets: the impact of the number of partner series on the efficiency. Clim Res 74:31–42. https://doi.org/10.3354/cr01488

Gimmi U, Luterbacher J, Pfister C, Wanner H (2007) A method to reconstruct long precipitation series using systematic descriptive observations in weather diaries: the example of the precipitation series for Bern, Switzerland (1760–2003). Theor Appl Climatol 87:185–199. https://doi.org/10.1007/s00704-005-0193-5

Gubler S, Hunziker S, Begert M, Croci-Maspoli M, Konzelmann T, Brönnimann S, Schwierz C, Oria C, Rosas G (2017) The influence of station density on climate data homogenization. Int J Climatol 37: 4670–4683. https://doi.org/10.1002/joc.5114

Guentchev G, Barsugli JJ, Eischeid J (2010) Homogeneity of gridded precipitation datasets for the Colorado River Basin. J Appl Meteorol Climatol 49:2404–2415. https://doi.org/10.1175/2010JAMC2484.1

Guijarro JA (2014) User's Guide to Climatol. http://www.meteobal.com/climatol/climatol-guide.pdf

Guijarro JA, López JA, Aguilar E, Domonkos P, Venema V, Sigró J, Brunet M (2017) Comparison of homogenization packages applied to monthly series of temperature and precipitation: the MULTITEST project. Ninth Seminar for Homogenization and Quality Control in Climatological Databases (Ed Szentimrey T, Lakatos M, Hoffmann L) WMO WCDMP-85:46–62

Hannart A, Mestre O, Naveau P (2014) An automatized homogenization procedure via pairwise comparisons with application to Argentinean temperature series. Int J Climatol 34:3528–3545. https://doi.org/10.1002/joc.3925

Hausfather Z, Menne MJ, Williams CN, Masters T, Broberg R, Jones D (2013) Quantifying the effect of urbanization on U.S. Historical Climatology Network temperature records. J Geophys Res Atmos 118:481–494. https://doi.org/10.1029/2012JD018509

Hua W, Shen SSP, Weithmann A, Wang H (2017) Estimation of sampling error uncertainties in observed surface air temperature change in China. Theor Appl Climatol 129:1133–1144. https://doi.org/10.1007/s00704-016-1836-4

Huang J, van den Dool HM, Barnston AG (1996) Long-lead seasonal temperature prediction using optimal climate normals. J Clim 9: 809–817

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

Huth R, Nemesova I (1995) Estimation of missing daily temperatures: can a weather categorization improve its accuracy? J Clim 8:1901–1916

Jones PD, Lister DH (2010) The urban heat island in Central London and urban-related warming trends in Central London since 1900. Weather 64:323–327. https://doi.org/10.1002/wea.432

Kemp WP, Burnell DG, Everson DO, Thomson AJ (1983) Estimating missing daily maximum and minimum temperatures. J Clim Appl Meteorol 22:1587–1593

Killick RE (2016) Benchmarking the performance of homogenisation algorithms on daily temperature data. PhD thesis, University of Exeter, UK. https://ore.exeter.ac.uk/repository/handle/10871/23095

Klok EJ, Klein Tank AMG (2009) Updated and extended European dataset of daily climate observations. Int J Climatol 29:1182–1191. https://doi.org/10.1002/joc.1779

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

Menne MJ, Williams CN (2009) Homogenization of temperature series via pairwise comparisons. J Clim 22:1700–1717. https://doi.org/10.1175/2008JCLI2263.1

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

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

Mestre O, Domonkos P, Picard F, Auer I, Robin S, Lebarbier E, Böhm R, Aguilar E, Guijarro J, Vertacnik G, Klancar M, Dubuisson B, Štěpánek P (2013) HOMER: homogenization software in R—

methods and applications. Idojaras Q J Hung Meteorol Serv 117:47–67

Oyler JW, Ballantyne A, Jencso K, Sweet M, Runninga SW (2015) Creating a topoclimatic daily air temperature dataset for the conterminous United States using homogenized station data and remotely sensed land skin temperature. Int J Climatol 35:2258–2279. https://doi.org/10.1002/joc.4127

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

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

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

Ribeiro S, Caineta J, Costa AC (2016) Review and discussion of homogenisation methods for climate data. Phys Chem Earth 94:167–179. https://doi.org/10.1016/j.pce.2015.08.007

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

Spinoni J, Szalai S, Szentimrey T, Lakatos M, Bihari Z, Andrea Nagy A, Németh Á, Kovács T, Mihic D, Dacic M, Petrovic P, Kržič A, Hiebl J, Auer I, Milkovic J, Štepánek P, Zahradnícek P, Kilar P, Limanowka D, Pyrc R, Cheval S, Birsan M-V, Dumitrescu A, Deak G, Matei M, Antolovic I, Nejedlík P, Štastný P, Kajaba P, Bochnícek O, Galo D, Mikulová K, Nabyvanets Y, Skrynyk O, Krakovska S, Gnatiuk N, Tolasz R, Antofie T, Vogt J (2015) Climate of the Carpathian region in the period 1961–2010: climatologies and trends of 10 variables. Int J Climatol 35:1322–1341. https://doi.org/10.1002/joc.4059

Szentimrey T (1999) Multiple analysis of series for homogenization (MASH). In: Szalai S, Szentimrey T, Szinell Cs (eds) Proc 2nd Seminar for Homo-genization of Surface Climatological Data. WMO WCDMP 41, pp 27–46

Szentimrey T (2010) Methodological questions of series comparison. In: Lakatos M, Szentimrey T, Bihari Z, Szalai S (eds) 6th Seminar for Homogenization and Quality Control in Climatological Databases. WMO WCDMP-76, pp 1–7

Tardivo G (2015) Spatial and time correlation of thermometers and pluviometers in a weather network database. Theor Appl Climatol 120:19–28. https://doi.org/10.1007/s00704-014-1148-5

Tardivo G, Berti A (2012) A dynamic method for gap filling in daily temperature datasets. J Appl Meteorol Climatol 51:1079–1086. https://doi.org/10.1175/JAMC-D-11-0117.1

Thorne PW, Menne MJ, Williams CN, Rennie JJ, Lawrimore JH, Vose RS, Peterson TC, Durre I, Davy R, Esau I, Klein-Tank AMG, Merlone A (2016) Reassessing changes in diurnal temperature range: a new data set and characterization of data biases. J Geophys Res Atmos 121:5115–5137. https://doi.org/10.1002/2015JD024583

Thorne PW, Allan R, Ashcroft L, Brohan P, Dunn R, Menne M, Pearce P, Picas J, Willett K, Benoy M, Bronnimann S, Canziani P, Coll J, Crouthamel R, Compo G, Cuppett D, Curley M, Duffy C, Gillespie I, Guijarro J, Jourdain S, Kent E, Kubota H, Legg T, Li Q, Matsumoto J, Murphy C, Rayner N, Rennie J, Rustemeier E, Slivinski L, Slonosky V, Squintu A, Tinz B, Valente M, Walsh S, Wang X, Westcott N, Wood K, Woodruff S, Worley S (2017) Towards an integrated set of surface meteorological observations for climate science and applications. Bull Am Meteorol Soc 98: 2689–2702. https://doi.org/10.1175/BAMS-D-16-0165.1

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

Willett KM, Williams CN, Jolliffe I, Lund R, Alexander L, Brönniman S, Vincent LA, Easterbrook S, Venema V, Berry D, Warren R, Lopardo G, Auchmann R, Aguilar E, Menne M, Gallagher C, Hausfather Z, Thorarinsdottir T, Thorne PW (2014) A framework for benchmarking of homogenisation algorithm performance on the global scale. Geosci Instrum Method Data Syst 3:187–200. https://doi.org/10.5194/gi-3-187-2014

WMO (2016). Web site of the Task Team on HOMOGENIZATION (OPACE2, WMO Commission for Climatology). http://www.climatol.eu/DARE. Accessed Aug 2017