José A. Guijarro
State Meteorological Agency (AEMET)

24.1 Introduction ........................................................................................................501
24.2 Quality Control Procedures ..............................................................................502
24.3 Series Homogenization ......................................................................................503
   Causes and Consequences of Inhomogeneities • Homogeneity
   Test Algorithms • Homogenization Computer Packages • Application Example
24.4 Summary and Conclusions ..............................................................................510
References ..............................................................................................................512

AUTHOR
José A. Guijarro is a meteorologist in charge of the Mediterranean Meteorological Studies unit at the State Meteorological Agency (AEMET) office in the Balearic Islands (Spain). He worked on various posts (radiosounding operator, aeronautical forecaster), but most of his work has been devoted to different fields of climatology, which was also the subject of his PhD dissertation. He has published over 75 papers and some book chapters and has lectured in several international courses on applied and statistical climatology. He is currently a member of an Open Panel of Experts of the Commission for Climatology (World Meteorological Organization [WMO]) and leader of the Data Rescue Task Team of the Working Group on Climatology and Hydrology (WMO Regional Association VI).

PREFACE
This chapter addresses the importance of using homogenized and quality-controlled series in order to obtain reliable estimations of the trends and variability of any climatological element, including those relevant for hydrological studies. Currently used methodologies are discussed, focusing on available computer packages, and a case application example is presented to better illustrate the benefits of applying these procedures.

24.1 Introduction
Meteorological observatories, through their observational programs, collect the data that build the climatological series that can be seen as strands lengthening year after year along the operation life of the observatories. In any geographical area we have normally several meteorological stations, and when the
climatologist weaves their different strands, he or she obtains a carpet depicting the spatial and temporal variability of the involved climatic elements in that area. However, observations may be affected by many perturbations that introduce point errors or changes in the statistical properties of the series, hence compromising the reliability of any conclusion derived from their analysis. Apart from human errors in the observation or mechanization of manual instruments, observational practices may change along with time, and miscalibrations, instrumentation replacement, station relocations, and changes in the surrounding land use are common events in the history of any observatory, introducing inhomogeneities in their series. From a statistical point of view, a series would be inhomogeneous whenever their statistical properties change along with time, but this would include actual changes in the climate. Therefore, climatological series inhomogeneity is referred only to changes introduced by perturbations alien to the climate variability.

Concern about inhomogeneities existing in the climatological series is as old as the series themselves. Many methods have been proposed to detect and remove these unwanted perturbations from the series, and thorough reviews and advice about their use can be seen in Peterson et al. [10], Aguilar et al. [1], and Beaulieu et al. [3]. On the other hand, the Hungarian Meteoro-Hydrological Service has been celebrating a series of International Seminars on Homogenization and Quality Control of Climatological Databases that are also a good source of information about past and current efforts on this subject (see the more recent proceedings in WMO [15–17]), as well as the website of a COST Action [4] that emerged from one of these seminars.

The next sections will address the quality control approaches, and then homogenization algorithms and their implementation in publicly available computer packages will be discussed, to end with an example application of one of them.

### 24.2 Quality Control Procedures

Quality control of climatological data is normally first carried out at the institutions that collect, mechanize, and archive the observations and depends greatly on the observed element and the instrumentation used. For example, cooperative observers may be manually recording daily extreme temperatures with a Six–Bellani thermometer, and errors can originate by reading the wrong side of the index or when confusing the nearest 5°C thick tick mark of the scale. In contrast, common manual precipitation data errors are assigning yesterday’s 07–07 h total to today or accumulating the precipitation of several missing observations into one. Therefore, it is the institution managing the observing network the best suited to know all the peculiarities of the sites and their observing practices and to control the quality of their data accordingly.

Nevertheless, even when this first-stage quality control is properly applied, some errors will inevitably pass the filters or can be originated afterward, during the data transfer between different devices or archive systems, and therefore data users should perform additional quality controls whenever possible. These quality controls will be more effective when applied to original data, because derived data may smooth the errors making them more difficult to detect. For example, a 10°C error in a manual reading of a daily maximum temperature becomes a mere 0.33°C error in the monthly maximum average and a still less noticeable 0.17°C error in the monthly mean temperature.

Quality controls can be grouped in four types:

1. Physical plausibility, that is, checking whether any value is outside its expected range (e.g., negative precipitation). It is often adjusted to the local climate: 0°C can be a usual temperature in temperate zone winters but unexpected in a tropical coastal site.
2. Temporal increments: Limits may be imposed to the maximum and minimum change between two or more consecutive measures.
3. Internal consistency between two or more variables: Minimum temperature greater than the maximum, precipitation amounts when no precipitating meteor is reported, etc.
4. Spatial consistency: The series of nearby stations are expected to present a similar behavior, and hence their comparison is a powerful method for detecting errors. This is not applicable to sub-daily data unless sensors are close enough as to guarantee the series synchronization (time shifts could be applied, but this would complicate the method). A previous normalization of the series is advised in areas with complex orography where climatic variables may display different ranges of values in varied altitudes and topographic locations.

The two first controls are applied over single series, and the others require two at least (from the same observatory in the third case and from several in the fourth type of controls). More information about their implementation can be found in the literature, as in Shearman [13]; Peterson et al. [11]; Guijarro [7]; and Zahumenský [18].

Homogenization packages discussed in the following section only include very basic quality controls, such as the possibility to remove outliers greater than a prescribed threshold, although visual inspection of their graphic outputs may help in detecting other kind of problems in the series.

### 24.3 Series Homogenization

Series homogenization is the process by which any influence unrelated to climate and weather is removed from climatological data sets. In this section we will address the variety of these unwanted perturbations and the different strategies to eliminate them from the observational series.

#### 24.3.1 Causes and Consequences of Inhomogeneities

As advanced in the introduction, any change in the conditions of observation or in the environment of the observatory may have a significant impact on the measured data. The first ranges from changes of instrumentation (including accessories such as wind or radiation shields) or observer, observing schedules and formulas used to compute daily values from observations at fixed hours, to relocation of the instruments (both horizontal or vertical displacements) or the whole observatory. And the environment may change in a variety of scales, from growing trees or new buildings or other obstacles in the vicinity to general urbanization or other land use changes in the area (e.g., introducing irrigation practices or new crops). Some changes will produce abrupt variations in the series, while others will introduce gradual shifts. For example, a forest fire affecting the area will suddenly change the albedo and hydric balance to slowly return to the initial conditions as vegetation recovers.

The impact of these changes depends on the climatic element being considered. Trees and other obstacles will affect wind speed and rainfall, this last much modified by whirls in the rain gauge, but pressure will not be influenced, while any relocation implying changes in altitude will be clearly detected by the latter. Temperature is one of the more important elements and at the same time very sensitive to all kind of changes. Since a perfectly calibrated thermometer yields just its own temperature, we must grant that this is also the temperature of the air by allowing a perfect thermal equilibrium between both. This is accomplished by letting the air flow as freely as possible around the thermometer bulb, while it is concealed from any kind of radiation exchange (from the sun, direct or reflected by the ground, but also from infrared radiation imbalances with heated surfaces or the coldness of a cloudless sky). It is clear that both requirements are incompatible, but different thermometer screens have been developed in the past that try to approximate to these ideal conditions. In the first observatories, thermometers were merely mounted on a north wall, but then they were exposed to (mild but noticeable) sunshine in the early and late hours of the summer days. Therefore, they were shielded from the sun with simple shading panels. Afterward it was realized that the thermal inertia of the walls and their north orientation had a microclimatic influence, and thermometers begun to be installed either on terraces or open ground. But as the widest diurnal temperature range takes place
on the ground (due to its strong radiation exchanges), air temperature bears important vertical gradients, mainly on calm sunny days, and therefore the height at which the thermometer is placed has an important influence on the measured temperatures. That ended in thermometers being installed on the ground, at a fixed height (although not the same in all countries: 1.5, 1.8, 2 m, etc.). Even repainting a thermometer shield can change its albedo and have an (presumably small, but still noticeable) influence on its temperature.

Therefore, observational series may bear anomalous shifts or trends that will mislead any conclusions about their variability, unless provisions are made to correct them, as discussed in the following subsections.

### 24.3.2 Homogeneity Test Algorithms

The first approaches to investigating the homogeneity of meteorological series consisted in expert judgment on visual inspection of the plotted series. It soon became evident that absolute homogeneity was incompatible with long-term climatic variability, and then cumulative values of the problem series were compared with those of an assumed homogeneous reference (double-mass analysis). Another common approach is to compute differences or quotients between the problem series and the reference, allowing objective detection methods to be applied on them. The history of the observatories provides invaluable information about possible changes, hence helping in deciding which series have suffered them and which can serve as references because of their continuity. But these histories (*metadata*) are often incomplete, if not absent, and therefore no series can be fully trusted of being homogeneous, making difficult to ascribe an inhomogeneity to the problem or to the reference station. Two main strategies have been adopted to overcome this difficulty: to compare the problem series with a *composite reference series* made of several neighbor or well-correlated stations and to make *multiple comparisons* with all the potential references, looking for repeating inhomogeneities at the same time step. Composite references, on their side, can be built in different ways, but generally a number of stations in the area of the problem station are chosen based on proximity and/or correlation. Data of the reference series can be normalized to avoid too different ranges, and the composite series can be a plain mean of all the candidates or an average weighted by distance or correlation. Other methodologies compute the reference series with a multiple regression model that estimates the problem data as a function of other available data in the study area.

When a composite reference series has been obtained, differences with the problem series are computed when the variable has a near-normal frequency distribution as, for example, temperature. But quotients are sometimes preferred with elements with a zero lower limit (the case of precipitation or wind speed), frequently presenting an L-shape biased distribution. Nevertheless, quotients become problematic in arid areas, where months with very low or null precipitation are not uncommon. In these cases differences can be better than quotients, and a previous transformation of the data (e.g., through a cubic root) can be used to approximate the frequency distribution to the normal. When both the problem and the reference series are homogeneous, the difference (or quotient) series should be normally distributed, and therefore statistical methods to test this normality can be applied, and this hypothesis can be rejected when the results are unexpected at a chosen significant level.

These tests can be parametric, as the classical t-test, the standard normal homogeneity test (SNHT) [2], the maximum likelihood ratio, or the serial correlation Durbin–Watson test, or nonparametric, as the Wilcoxon–Mann–Whitney [6], and can be applied on the whole series or on running windows (to improve the detection of multiple break points). Other algorithms compute trends before and after a moving time step, as is the case of two-phase regression [5], looking for significant differences between both models. Most of these algorithms provide indication of the location of the inhomogeneous shifts, and it is advisable to compare them with the metadata that can help in confirming or better adjusting the date of the shift.
The following step is to correct the detected inhomogeneities. Several statistical techniques can be used, from computation of the mean differences before and after the break to analysis of variance (ANOVA) or direct computation of the missing values after splitting the series into their homogeneous segments.

The sensitivity of these detection algorithms relies on a good signal-to-noise ratio: with series with a high variability, only big shifts can be detected reliably. Some climatic elements are more variable than others, and daily data have a much higher variability than monthly, seasonal, or annual means or totals. As a result, less inhomogeneities are normally detected on precipitation than in temperature series, and homogenization techniques are usually not applied to daily data but rather to monthly series, whose correction terms are afterward transferred (with some smoothing interpolation scheme) to the daily series.

### 24.3.3 Homogenization Computer Packages

We have seen that there exist a variety of algorithms among which climatologists can choose the most appropriate for their research, depending on the climatic element, the density of the measuring stations, and the geographical characteristics of the study area. When dealing with a reduced amount of series, these algorithms may be applied manually with the help of spreadsheets or other elementary computing aids. Otherwise they must be laboriously implemented in more complex computer programs. Fortunately several climatologists that have already made this task offer their developments to other users by preparing and documenting ready to use computer packages and distributing them in the Internet or by other means.

Tables 24.1 through 24.3 summarize the characteristics of the currently available homogenization packages, most of them in active stage of development to refine their procedures or to include new capabilities (updates to this information can be followed at http://www.climatol.eu/DARE/index.html). Although developed for the same purpose, they differ in a variety of ways that allow users to choose them according to their preferences, beginning by the operating system on which they run, or whether they are closed or open source. Interactivity or automatism is also an important feature: some users may love the possibility to fully interact with every step of the process through a sophisticated graphical user interface (GUI) and many algorithms to choose (as is the case in AnClim), while others will prefer the ability to run the software in a fully automatic way, either from a terminal or from a script. Most packages are able to compute or select the reference series by themselves, but that is not the case of RHTestV3 that, working on single series, leaves the provision of the references entirely to the user. HOMER is even more interactive in its design, since it is addressed to trained users that can make use of their expertise.

Another important aspect is the tolerance to missing data in the series to be homogenized. When homogenization was applied by manual means, investigators restrained themselves to use only the longest and/or more complete series available for their study. But now that automatic methods allow processing hundreds or even thousands of series, a high tolerance to missing data permits a denser network of stations to be homogenized. The benefits are double: on one side, each problem station will have nearer and better correlated reference data, improving the quality of the results. On the other, stations with short functioning periods will have their series homogenized and their missing data estimated, making them useful for many applications.

Most packages that are primarily devised to operate on monthly data can be applied to daily data as well. This is advisable for the detection of isolated errors in the form of outliers, but not for diagnosing nor correcting inhomogeneities. Current homogenization strategies for daily data relay on different methods of adjusting the monthly correction terms or factors to them. Some packages are programmed in R, a popular statistical computer environment [12], hence allowing its use on virtually any platform or operating system. USHCNv2 is also provided in source code, but in Fortran, and requires a nontrivial
### Table 24.1 Homogenization Packages: Software Types and Availability

<table>
<thead>
<tr>
<th>Package</th>
<th>Version</th>
<th>License</th>
<th>Open Source</th>
<th>Operating System</th>
<th>Program Type</th>
<th>Primary Operation</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAHMDI/HOMAD</td>
<td>?</td>
<td>GPL</td>
<td>Yes</td>
<td>(Most) Linux</td>
<td>R source</td>
<td>Automatic Interactive</td>
<td>Mail to <a href="mailto:andrea.toreti@giub.unibe.ch">andrea.toreti@giub.unibe.ch</a></td>
</tr>
<tr>
<td>HOMER</td>
<td>?</td>
<td>?</td>
<td>Yes</td>
<td>(Most) R source</td>
<td>Interactive</td>
<td></td>
<td>Mail to <a href="mailto:olivier.mestre@meteo.fr">olivier.mestre@meteo.fr</a></td>
</tr>
<tr>
<td>MASH</td>
<td>3.03</td>
<td>Freeware</td>
<td>No</td>
<td>DOS/Windows</td>
<td>Binary</td>
<td>Automatic (and interactive)</td>
<td>Mail to <a href="mailto:szentimrey.t@met.hu">szentimrey.t@met.hu</a></td>
</tr>
<tr>
<td>ReDistribution test</td>
<td>?</td>
<td>Freeware</td>
<td>Yes</td>
<td>(Most) R source</td>
<td>Interactive</td>
<td></td>
<td>Mail to <a href="mailto:predrag.petrovic@hidmet.gov.rs">predrag.petrovic@hidmet.gov.rs</a></td>
</tr>
</tbody>
</table>

Note: An interrogation sign marks unknown or doubtful items.

### Table 24.2 Homogenization Packages: Data Types and Detection Methods

<table>
<thead>
<tr>
<th>Package</th>
<th>GUI</th>
<th>Time Resolution</th>
<th>Input Format</th>
<th>Metadata Use</th>
<th>Detection Method</th>
<th>Ref. Series Selection</th>
<th>Detection Statistic</th>
<th>Climatic Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACMANT</td>
<td>No</td>
<td>Monthly</td>
<td>Text</td>
<td>No</td>
<td>Reference</td>
<td>Correlation</td>
<td>Causinus–Lyazrhi</td>
<td>Temperature</td>
</tr>
<tr>
<td>AnClim/ProClimDB</td>
<td>Yes</td>
<td>Any</td>
<td>Text DBF</td>
<td>Yes</td>
<td>Ref. and pairwise</td>
<td>Correlation and distance</td>
<td>Several</td>
<td>Any</td>
</tr>
<tr>
<td>Climatol</td>
<td>No</td>
<td>Monthly</td>
<td>Text</td>
<td>No</td>
<td>Reference</td>
<td>Distance</td>
<td>SNHT</td>
<td>Any</td>
</tr>
<tr>
<td>GAHMDI/HOMAD</td>
<td>No</td>
<td>Monthly Daily</td>
<td>Text</td>
<td>Yes</td>
<td>Pairwise</td>
<td>Correlation</td>
<td>New method</td>
<td>Any Temperature</td>
</tr>
<tr>
<td>HOMER</td>
<td>No</td>
<td>Monthly</td>
<td>Text</td>
<td>Yes</td>
<td>Pairwise</td>
<td>Correlation</td>
<td>Penalized likelihood</td>
<td>Temperature</td>
</tr>
<tr>
<td>MASH</td>
<td>No</td>
<td>Monthly daily</td>
<td>Text</td>
<td>Yes</td>
<td>Multiple references</td>
<td>Correlation</td>
<td>Multiple linear regression (MLR) and hypothesis test</td>
<td>Any</td>
</tr>
<tr>
<td>ReDistribution Test</td>
<td>No</td>
<td>Subdaily</td>
<td>Text</td>
<td>No</td>
<td>Distribution</td>
<td>None</td>
<td>SNHT-like</td>
<td>Wind speed and direction</td>
</tr>
<tr>
<td>RHTestV3</td>
<td>Yes</td>
<td>Monthly daily</td>
<td>Text</td>
<td>Yes</td>
<td>Reference</td>
<td>Correlation</td>
<td>Penalized max. t and F tests</td>
<td>Any</td>
</tr>
<tr>
<td>USHCNv2</td>
<td>No</td>
<td>Monthly</td>
<td>Text</td>
<td>Yes</td>
<td>Pairwise</td>
<td>Correlation</td>
<td>MLR</td>
<td>Temperature</td>
</tr>
</tbody>
</table>
Apart from preferences about the operating system, interactivity, or whether a GUI is available, potential users will be interested in the performance of these packages in homogenizing data sets. The most thorough comparison exercise was made during the COST Action ES0601, its results being summarized in Venema et al. [14], but as this field is in active development, many packages have improved since those Action results and more intercomparisons are needed with the new versions. It is very difficult to repeat that exercise in the same way, which included time-consuming manual methods, but automatic benchmarking experiments are currently under way, although restrained to fully automatic runs of the packages. The first results of a comparison on ten simulated temperature series with one or two shifts in the mean of ±2°C in five of them show a general good and similar performance of ACMANT, Climatol, RHTestV3 relative with constant correction, and HOMER. All of them corrected quite well the anomalous trends of the unhomogenized series. RHTestV3 with quantile matching gave worse results, although not as bad as absolute homogenization, which should be avoided whenever possible [9]. The referred packages were those easier to automatize, but others will be included in the future and more complex homogenization problems will be tried as well (see http://www.climatol.eu/DARE/testhomog.html).

### 24.3.4 Application Example

For a better illustration of the use of these packages, we can proceed with a case study in which, from a set of raw monthly precipitation data, we want to obtain average values and trends after a homogenization process. The data come from 37 real stations located in a 115 × 65 km rectangular area centered at 41.3°N, 1.7°E (a coastal area including Barcelona, Spain), and they cover a period of 60 years (1951–2010).

As is usually the case, these series are far from being complete, as can be seen in Figure 24.1. Therefore, and due to the strong variability of the precipitation, missing data must be estimated before computing consistent averages of those stations. This task will be accomplished with the Climatol
package that has a high tolerance to missing data because reference stations are chosen by proximity and not by correlation (which cannot be reliably computed without a minimum number of simultaneous observations).

On the other hand, the area has a typical Mediterranean climate, with a marked minimum in July, although null precipitations may happen eventually in any month. Therefore, monthly precipitations exhibit a strong L-shape frequency distribution (Figure 24.2a), with values ranging from 0 to 730 mm. Climatol offers two possibilities in these situations, either to normalize the data using rate to normals (instead of the default standardization with the mean and the standard deviation) or to apply a root transformation to the data in order to approximate their distribution to the normal. A cubic root transformation has been chosen in our example, producing a near-normal distribution (Figure 24.2b) if we

**FIGURE 24.1** Data availability in the example area. (a) Data coverage in each station (lighter gray indicate less data in that year). (b) Number of data in every time step for the whole area.

**FIGURE 24.2** Frequency distributions of the precipitation data used. (a) Original. (b) After a cubic root transformation.
leave apart the zero-precipitation months. To apply this package to our case, we need to have installed the free R statistical environment and the Climatol contributed package (see the R help for installation instructions).

The second step is either to read the data directly from our archives or databases with our own R procedures or to save them in the Climatol plain text input formats, which is the option followed here. Only two files are needed in this case: one for the station coordinates and the other for the data. If we use “Prec” as short for precipitation, then the name of the first file must be “Prec_1951–2010.est,” and it lists the X(°), Y(°), Z(m) coordinates, the codes, and the names of the stations, as shown in these first lines:

1.054 41.074 19 “0013” “CAMBRILS”
1.052 41.141 123 “0015I” “RIUDOMS (CAMARA AGRARIA)”
1.179 41.15 68 “0016A” “REUS (AEROPORT)”
1.145 41.112 53 “0017” “VILASECA DE SOLCINA”
...

The data are stored in a file named “Prec_1951–2010.dat,” also as plain text, containing the data of all the stations in the same order as in the previous file and in chronological order within each station, from January 1951 to December 2010 (including special codes or values for any missing data). The first lines of this file, corresponding to the years 1951–1953 of Cambrils, are (NA stands for “not available,” the standard missing data code in R) as follows:

42 16.6 NA 59.9 56.2 27.8 11.9 45.7 146.2 122.5 59 36
21 21 47 42 34 49 31 3 NA 104.7 7.5 9.1
0 2.1 40.1 13 20 140.2 13 37 58 62.5 28.7 64.4
...

After having prepared these two files, we start R in our computer and issue these two commands to load the Climatol library in memory and homogenize our example data (for detailed explanations on the command parameters, see the user’s guide [8]):

```r
library(climatol)
homogen(“Prec”, 1951, 2010, rtrans=3, vmin=0, gp=4)
```

After several minutes (depending on the processor speed) we get four output files, with the same base name as the input files and with extensions:

- **esh**: Output coordinates and names of the stations, with supplementary information
- **dah**: Output homogenized data
- **txt**: Log of all console output during the process
- **pdf**: Collection of diagnostic graphics (123 pages in our example)

All output files are plain text files except the PDF. These graphics show us the course of the homogenization process, in which, after normalizing all series, anomalies computed as differences between each one and an average of their closest neighbors are subject to the reliable and well-known standard normal homogeneity test, and the more outstanding inhomogeneities result in splitting the series at that point in an iterative process, until no inhomogeneity remains greater than a prescribed threshold. Figures 24.3 and 24.4 show one of those anomaly graphics and the reconstruction of both series resulting from splitting the original. Eight splits were made in our example exercise, two of them in the same station.

The homogenized output data in “Prec_1951–2010.dah” has the same text format as the input data, which is not much user friendly because of the lack of any time or station reference apart from the position of each data item in the file. Fortunately Climatol has a post-processing function to help in
obtaining various indexes from the homogenized data. If we are interested in the trends of the last 50 years (1961–2010), we issue the following command:


which yields monthly and annual trends in mm/century, computed as regression coefficients with time, of the 37 stations (choosing the longest period with original data when they have been split). In order to assess the effect of the homogenization, we can ask Climatol to compute the missing data without splitting any series and then get a new set of trends without homogenization. Figure 24.5 displays the frequency histograms of both trend sets, and it is evident that trends are more consistent and less spread after homogenization. Not only the trend range is reduced as a result of correcting the inhomogeneous series, but all trends are affected through the missing data interpolation process: the mean trend changes from −40.7 to −21.1 mm/century, and the more robust median, from −36.9 to −22.1 mm/century. The same homogen() function can be used to compute 1961–1990 normals or any empirical percentile for, for example, categorize the forthcoming new monthly values for climate monitoring purposes, and these categories (dry, normal, wet, etc.) can vary significantly if the percentiles are computed with inhomogeneous data.

24.4 Summary and Conclusions

This chapter has presented the variety of unwanted perturbations that can affect climatological series in such a way as to induce to incorrect evaluations about their trend and variability, hence compromising the quality of the decisions made by the user.
Different techniques to correct these inhomogeneities have been listed, but as their application to a big number of series can be very time consuming, attention has been focused on different implementations in the form of computer packages that are freely available. Their characteristics have been summarized in synoptic tables, and an example application to a real case has been illustrated, showing the higher consistency of the trends by comparison to the values obtained from the raw series.

The same may happen with other statistical characteristics of the series, and although the homogenized series obtained by different methods may differ, any good homogenization methodology will be worth being applied instead of using the original inhomogeneous data.

Ideally, the best situation would be to have homogeneous observational series, making unnecessary the use of these techniques. Unfortunately this is often not the case, but network managers can help in minimizing the problem by trying to keep observation conditions unchanged as much as possible, logging detailed and easily available records of unavoidable changes to provide good metadata to the users and avoiding by all means any change affecting a whole network at the same time, since these would invalidate the use of even the best relative homogenization methods.
References


