Geophysical Research Abstracts
Vol. 19, EGU2017-13972, 2017
EGU General Assembly 2017

Automatic benchmarking of homogenization packages applied to synthetic monthly series within the frame of the MULTITEST project

José A. Guijarro (1), José A. López (1), Enric Aguilar (2), Peter Domonkos (), Victor Venema (3), Javier Sigró (2), and Manola Brunet (2)

(1) AEMET (State Meteorological Agency), Spain, (2) Universitat Rovira i Virgili, Spain, (3) Meteorological Institute, University of Bonn, Germany

After the successful inter-comparison of homogenization methods carried out in the COST Action ES0601 (HOME), many methods kept improving their algorithms, suggesting the need of performing new inter-comparison exercises. However, manual applications of the methodologies to a large number of testing networks cannot be afforded without involving the work of many researchers over an extended time. The alternative is to make the comparisons as automatic as possible, as in the MULTITEST project, which, funded by the Spanish Ministry of Economy and Competitiveness, tests homogenization methods by applying them to a large number of synthetic networks of monthly temperature and precipitation.

One hundred networks of 10 series were sampled from different master networks containing 100 series of 720 values (60 years times 12 months). Three master temperature networks were built with different degree of cross-correlations between the series in order to simulate conditions of different station densities or climatic heterogeneity. Also three master synthetic networks were developed for precipitation, this time mimicking the characteristics of three different climates: Atlantic temperate, Mediterranean and monsoonal.

Inhomogeneities were introduced in every network sampled from the master networks, and all publicly available homogenization methods that we could run in an automatic way were applied to them: ACMANT 3.0, Climatol 3.0, MASH 3.03, RHTestV4, USHCN v52d and HOMER 2.6. Most of them were tested with different settings, and their comparative results can be inspected in box-plot graphics of Root Mean Squared Errors and trend biases computed between the homogenized data and their original homogeneous series.

In a first stage, inhomogeneities were applied to the synthetic homogeneous series with five different settings with increasing difficulty and realism: i) big shifts in half of the series; ii) the same with a strong seasonality; iii) short term platforms and local trends; iv) random number of shifts with random size and location in all series; and v) the same plus seasonality of random amplitude. The shifts were additive for temperature and multiplicative for precipitation.

The second stage is dedicated to study the impact of the number of series in the networks, seasonalities other than sinusoidal, and the occurrence of simultaneous shifts in a high number of series.

Finally, tests will be performed on a longer and more realistic benchmark, with varying number of missing data along time, similar to that used in the COST Action ES0601.

These inter-comparisons will be valuable both to the users and to the developers of the tested packages, who can see how their algorithms behave under varied climate conditions.