Comparison of homogenization packages applied to monthly series of temperature and precipitation: the MULTITEST project

J.A. Guijarro[1], J.A. López[1], E. Aguilar[2], P. Domonkos, V.K.C. Venema[3], J. Sigró[2] and M. Brunet[2]

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

9th Seminar for homogenization and quality control in climatological databases and 4th conference on spatial interpolation techniques in climatology and meteorology (Budapest, 3-7 April 2017)

Outline

Introduction

- The need to homogenize observational series before its use to assess climate variability emerged long time ago and many methodologies, some of them implemented in computer packages, have been developed since then.
- Action COST ES0601 *HOME* was very useful to promote discussion meetings of homogenization specialists and to inter-compare the performances of their methods and software developments.
- Yet, as many of these programs have been improved since the end of that Action, new benchmarking exercises are needed to compare their current performances.

Introduction

The results of a preliminary comparison are still shown at http://www.climatol.eu/DARE/testhomog.html, but the Spanish project MULTITEST (Multiple verification of automatic softwares homogenizing monthly temperature and precipitation series) aims at updating and improving those benchmarking experiments in various ways:

- More realistic temperature networks
- Inclusion of precipitation networks with different climatic characteristics (Temperate, Mediterranean and Monsoonal)
- More realistic inhomogeneities
- More tested homogenization methods

Yet only automatic procedures can be tested to achieve significant results with a reasonable effort!

Benchmarking methodology

- Data benchmarks are composed by 100 homogeneous series with 60 years of monthly values without any missing data. From them, 100 tests are made by:
  - Randomly sampling a subset of the series (true solution)
  - Applying inhomogeneities to them (problem series)
  - Homogenizating them (backward adjustent) by different methods (results)
  - Comparing the results with the true solutions, computing RMSE, trend differences, and other metrics
- Note that as these methods are applied in an automatic way, they are run with default settings, and their results may not be as optimal as when properly tuned to each problem network.

Methodology

Tested homogenization programs (those that we could run in completely automatic mode):

- Climatol 3.0 (Guijarro), with constant and variable corrections
- ACMANT 3.0 (Domonkos), versions for temperature and precipitation (sinusoidal and irregular seasonalities)
- MASH 3.03 (Szentimrey)
- RHTestV4 (Wang & Feng), absolute and relative, with or without quantile adjustment. (Average series were given as reference!)
- USHCN_v52d (Menne & Williams)
  (We could not compile the current version yet)
- HOMER 2.6 (Mestre *et al.*), with different iteration strategies

Temperatures: networks & inhomogeneities

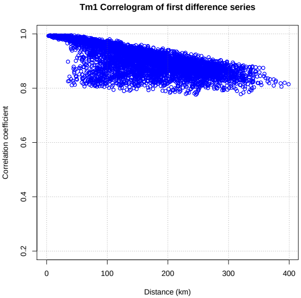
Generation of master networks Tm1, Tm2 and Tm3:

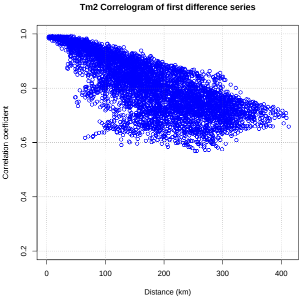
- 100 random points on a 4 x 3° lon-lat area
- Mean monthly homogenized temperatures from Valladolid (Duero basin, Spain) acting as seed series
- Closest point is assigned the same series plus white noise from $C \cdot \mathcal{N}(0, 1.5)$
- Coefficient $C = 0.18, 0.30, 0.65$ yield three master networks with decreasing correlation between stations, called Tm1, Tm2 and Tm3
- Series shifted to account for simulated elevation, 2°C/100yr trend added, and annual oscillation varied $\pm$20%

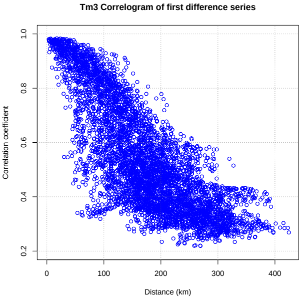
Inhomogeneities (mode 'rs'): Random number of shifts (5/100yr) with random size from $\mathcal{N}(0, 1)$ and sinusoidal seasonality of random amplitude from $\mathcal{N}(0, 0.7)$

Temperatures: Correlograms

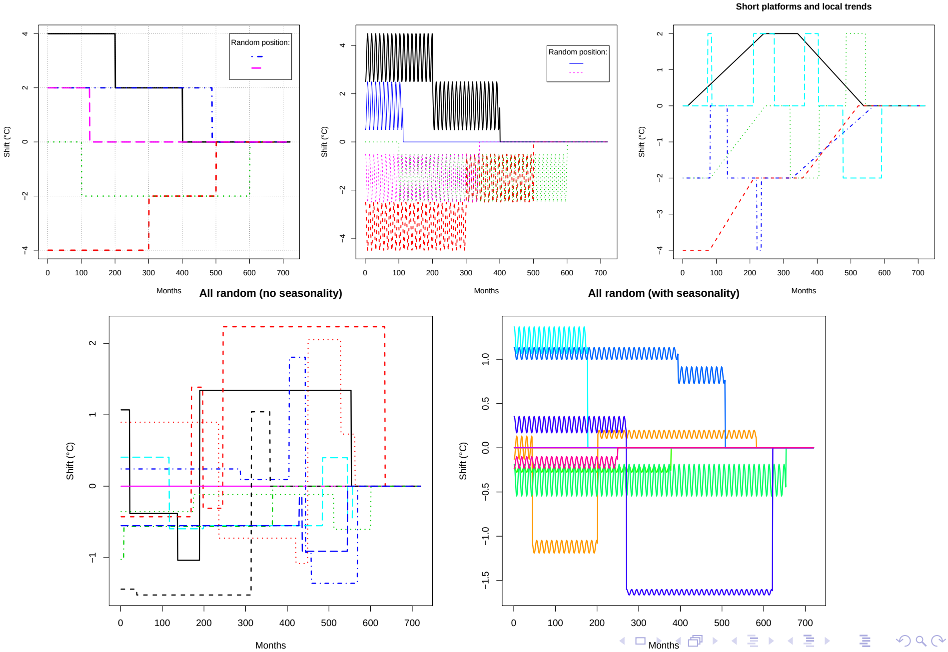
Correlograms of the first differences of the temperature networks Tm1, Tm2 and Tm3:



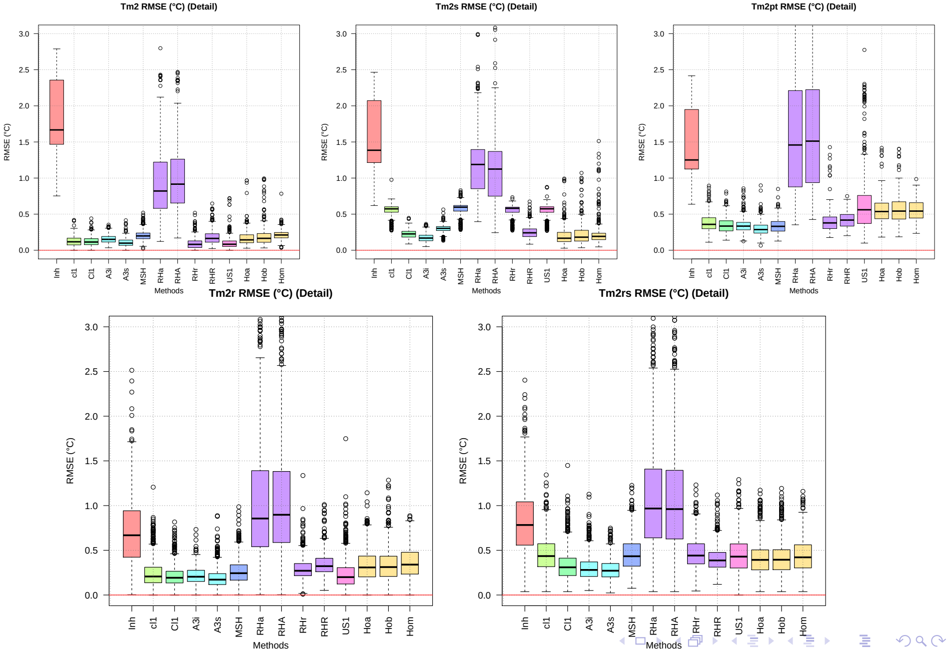




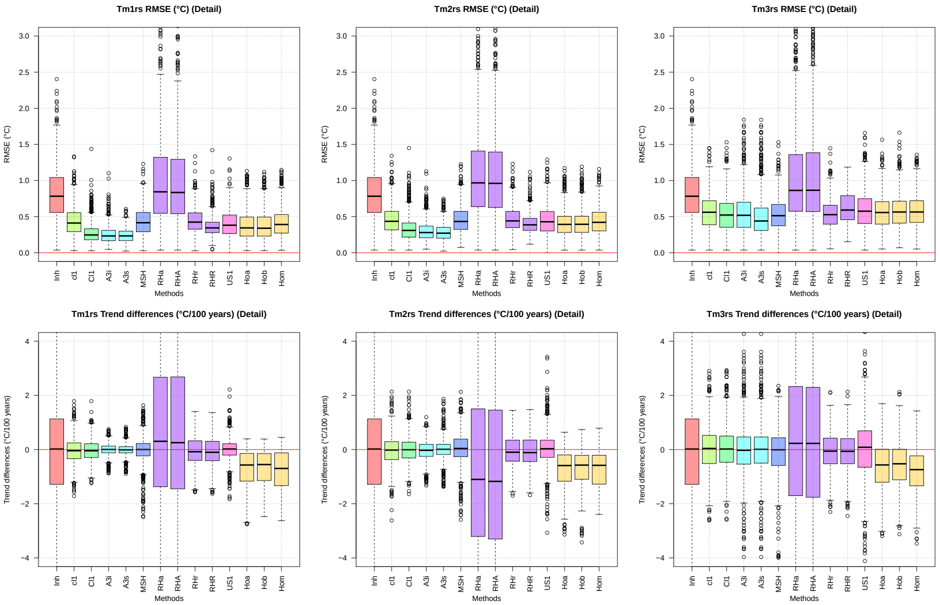
Temperatures: 5 experiments



Temperatures: Tm2 RMSE results



Temperatures: RMSE and trend diff. in mode 'rs'



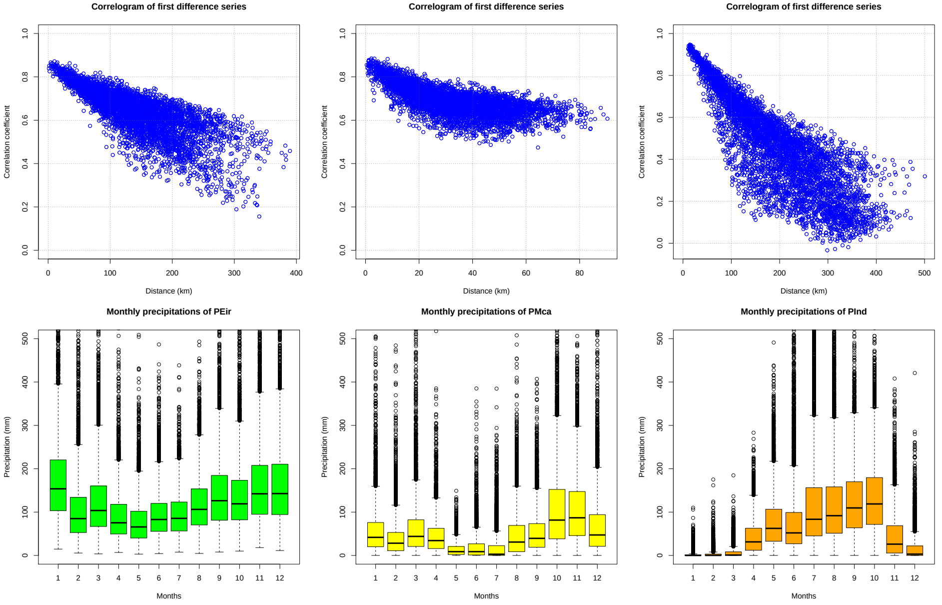
Precipitations

Three monthly precipitation networks were built simulating three different climates: Atlantic temperate (PEir), Mediterranean (PMca) and Monsoonal (PInd).

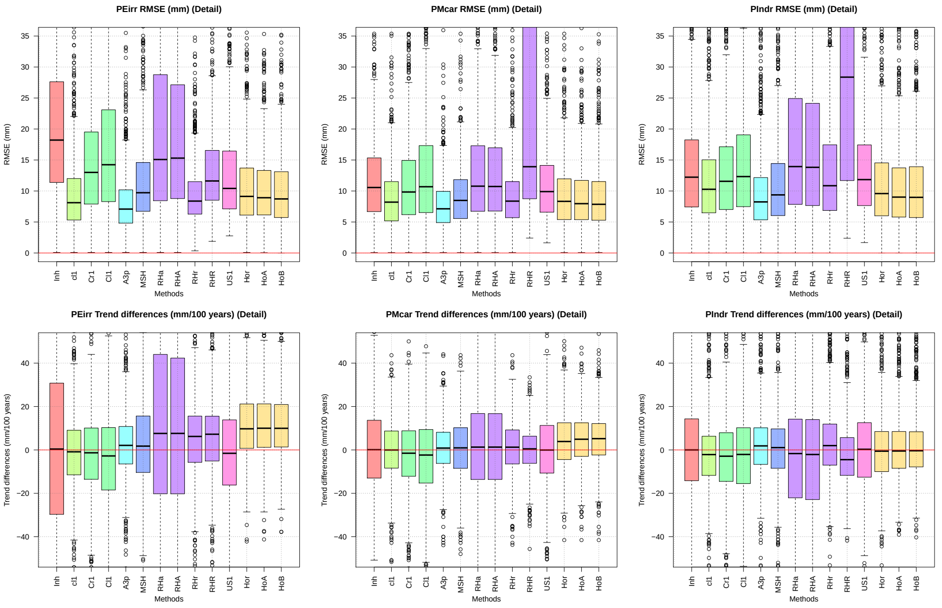
Real series from Ireland, Majorca and SW India (gridded) were respectively used to derive variograms, gamma coefficients and frequency of zeroes, which were used to compute their synthetic series by means of the R package gstat, preserving the spatial correlation structure.

A random number of shifts (5/100yr) were introduced as **factors** drawn from $\mathcal{N}(1, 0.2)$ (in mode 'r': no seasonal perturbation were applied to these factors)

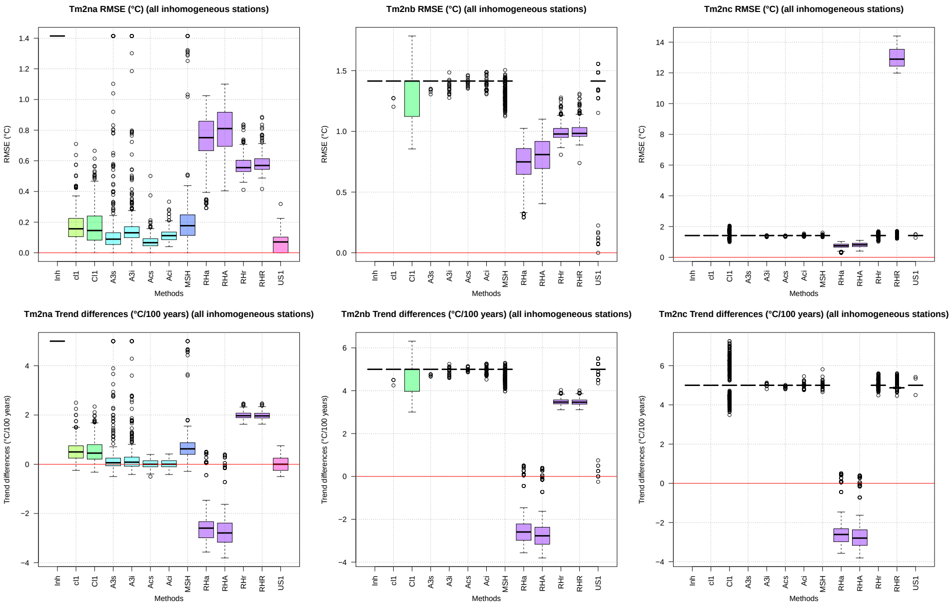
Precipitation characteristics



Precipitations: RMSE and trends ('r')

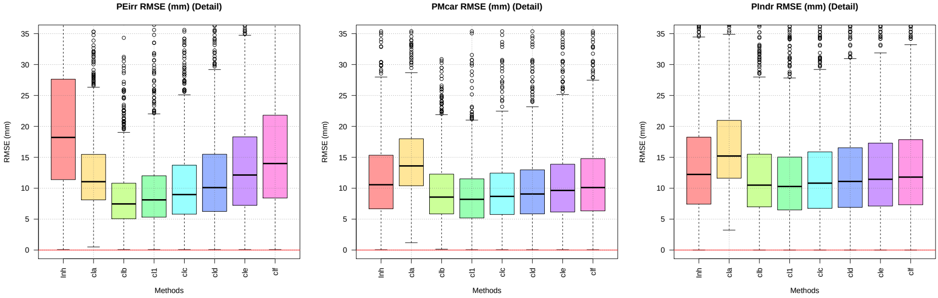


One only simultaneous shift in 40, 70 and 100% Tm2r



Precip.: Climatol RMSE vs SNHT thresholds

RMSE obtained by Climatol (with rate normalization) on the precipitation tests with thresholds of SNHT = 5, 10, 15, 20, 25, 35 and 50 :



Conclusions

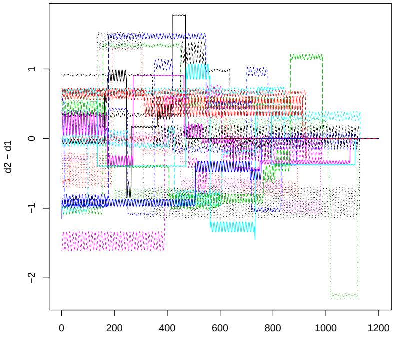
- The performance of the methods can vary depending on the characteristics of both the network and the inhomogeneities
- Hence the importance of showing results from different networks that can be representative of different real climates
- Unrealistically designed experiments also help in detecting the strengths and weaknesses of the methods
- Precipitation appears as probably being the most difficult variable to homogenize (many zeroes and a very biased PDF)
- The graphics displaying the results of the tests, as well as other characteristics of the software packages shown in http://www.climatol.eu/tt-hom/index.html, will facilitate the user to choose the method that better suits his needs

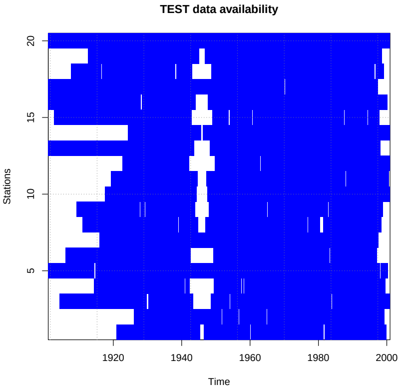
Future work

Ongoing and future work includes:

- Test the influence of non sinusoidal seasonalities in the shifts
- Test missing data tolerance of new packages
- Try longer series with missing data mimicking those in the HOME benchmark
- Put all results and scripts in a web page to allow reproducibility

Example network with missing data





Out of the scope of the MULTITEST project:

Why not take advantage of the implemented comparison scripts to test the performances of the methods with daily data?

It would only require:

- Choose homogeneous daily networks
  (Those developed by R. Killick?)
- Adapt the scripts to these networks
- Test the methods

(Ideally developers would provide automatic scripts that read the problem network and yield its homogenized version. In this way, they could test different settings of their programs.)

Acknowledgements

Project MULTITEST (CGL2014-52901-P) is funded by the Spanish Ministry of Economy and Competitiveness.

Thanks to Met Éireann for for providing the Irish monthly precipitation series that served as model to synthesize the network of Atlantic Temperate precipitations. Mallorca monthly precipitations were taken from AEMET data bases, and monthly precipitations from SW India, gridded at 0.5° resolution, were obtained from the Global Precipitation Climate Center (GPCC).

Many thanks for the advice received on the settings needed to run MASH (Tamás Szentimrey) and HOMER (Gregor Vertacnik, John Coll and Stefanie Gubler).