Geophysical Research Abstracts,
Vol. 11, EGU2009-12020, 2009
EGU General Assembly 2009

Climatological series shift tests comparison on running windows

J.A. Guijarro
AEMET (Meteorology State Agency), Balearic Islands Office, Palma de Mallorca, Spain (jaguijarro@inm.es)

The detection and correction of inhomogeneities in the climate series is of paramount importance for avoiding misleading conclusions in the study of climate variations. Several comparisons of detection methods have been undertaken so far, but their results are influenced by the type (shifts and/or local trends), number and position of the simulated inhomogeneities. During the development of an automated homogenization function for the "Climatol" R contributed package, the chosen approach for the detection of multiple change points was the application of the shift tests to windows running along the whole series of anomalies (differences between the problem station and a synthetic reference series computed from neighbor stations).

To select the bests tests to be applied in this way, 500 Monte Carlo simulations have been done for the ideal case of a 600 normally distributed terms (equivalent to a monthly climatic series of 50 years), with a single shift in the middle of magnitudes from 0 to 2 standard deviations (in steps of 0.2 s). The compared tests have been: 1) Classical t test; 2) Standard Normal Homogeneity Test; 3) Two Phase Regression; 4) Wilcoxon-Mann-Whitney (similar to Wilcoxon Rank Sums); and 5) Durbin-Watson test (lag-1 serial correlation). The criterion for qualifying the performance of each test was the ability to detect shifts without false alarms and to locate them at the correct point.

Results indicate that, under these precise simulated conditions, the best test is the classical t test, followed very closely by Alexandersson's SNHT, and by Wilcoxon-Mann-Whitney, while Two Phase Regression and Durbin-Watson performances are very poor.

As this tests must be applied many times in iterative runs during the homogenization of a climatological database, computing efficiency is also important, and therefore the time used by each of the tests has also been compared. Tests adjusting regression models (Two Phase Regression and Durbin-Watson) are the most time consuming (using the R lm function). The R implementation of the t test is much faster, but at the same time it is also much slower than SNHT, probably due to its higher complexity and the involved computation of p-values and other statistical parameters. For the case where computer time is critical, an alternative to SNHT is provided, giving identical results with 20% higher speed.