EMS Annual Meeting Abstracts
Vol. 16, EMS2019-788, 2019

Relative statistical homogenization of observational networks with a low signal to noise ratio

Victor Venema and Ralf Lindau
University of Bonn, Meteorological institute, Bonn, Germany (victor.venema@grassroots.is)

Relative statistical homogenization compares a candidate station to its neighbouring reference stations. Relative methods assume that the reference stations experience the same large-scale climate signal, but have deviations due to inhomogeneities (typically modelled as a step function) and noise (typically modelled as white noise, sometimes as short range correlated noise).

Science has studied this problem in much detail for the case of one time series with one break. Climatologists also studied more realistic cases, with multiple station series and multiple breaks per series, but mostly for dense observational networks, which are rare globally and going back in time. Our recent work suggests that statistical homogenization in case of multiple stations and multiple breaks is especially hard when the signal to noise ratio is below one.

We studied the detection problem for a series (the difference of a pair of stations) consisting of a break signal and a noise signal. The signal to noise ratio of such a series is defined as the standard deviation of the break signal divided by the standard deviation of the noise signal.

The optimal break positions for a certain number of breaks are determined by searching for the combination that minimizes the unexplained variance. This variance and a penalty function is used to determine the number of breaks.

The fundamental problem in case of multiple breaks is that randomly inserted test breaks are already able to explain half of the break variance. The combination which maximizes the explained noise variance would be random with respect to the breaks and will thus have this noise variance plus half of the break variance. Because this combination explains more noise than expected in a pure noise signal, it will be regarded as statistically significant. The statistical test rightly detects that the series contains inhomogeneities, but when the signal to noise ratio is low the positions of the breaks will be mostly determined by the noise. The paper gives the full details and also shows that a cross-term of the break and the noise signal further increases the variance of the solution.

Studying the so-called ANOVA joint correction method, we found that when all breaks are correctly detected this correction will remove large-scale trend biases due to inhomogeneities. After correction the size of the random trend error is determined by the noise variance of the difference time series and not by the break variance. When there are errors in the break positions and thus especially when the signal to noise ratio is low, any biases in trends will only be partially corrected and in this case the break variance is also important.