Break Position Errors in Climate Records

Ralf Lindau and Victor Venema

University of Bonn, Germany

Long instrumental climate records suffer from inhomogeneities due to, e.g., relocations of the stations or changes in instrumentation, which may introduce sudden jumps into the time series. These inhomogeneities may have influences in both directions: masking true or introducing spurious trends. Homogenization algorithms use the difference time series of neighboring stations to identify breaks. Modern multiple break point methods search for the optimum segmentation, which is characterized by minimum internal variance within the segments and maximum external variance between the segment means.

We analyze the accuracy of these homogenization methods and concentrate on the uncertainty in the position of the break. Due to unavoidable random noise in the difference time series, the segmentation method may find a slightly shifted break position, which attains a higher external variance than the true one. Not only direct neighbors of the true break needs to be considered, but all neighbors; that one with the largest external variance will be chosen as erroneous optimum. The variances of shifted segmentations are describable by a sum over a successively expanded sequence of a normal distributed random variable minus a term, which grows linearly with the length of the sequence. Such a process is known as Brownian motion with drift. Thus, the probability distribution of break position deviations can be largely described by the time of the maximum of a Brownian motion with drift, where the jump height to noise ratio defines the drift size.