



Automatic homogenization of large data sets of long-term precipitation series

E. Rustemeier, A. Kapala, V. Venema, and C. Simmer

University of Bonn, Meteorological Institute, Bonn, Germany (elke.rustemeier@uni-bonn.de)

Data from long-term in-situ precipitation observations are known to be inhomogeneous. The reasons for this are numerous such as relocation of measuring stations or modifications of the instrumental set up. For investigations based on these data, such as robust trend analysis, reliable data is essential.

This contribution focuses on the automatic homogenization of monthly precipitation totals and, if possible, supported by monthly indices (e.g., number of precipitation days, intensity, etc.). The automation of the methods allows an objective correction of large data sets in an acceptable time frame.

The data base consists of more than 2000 precipitation series located in Germany covering 50 years of data. Additionally, data from 118 stations of daily data are included, which have been digitized during the KLIDADIGI project by the Deutscher Wetterdienst and the Meteorological Institute of the University of Bonn covering 100 years of data.

The homogenization process consists of three steps: 1) Selections of networks of comparable precipitation behavior to estimate the climate signal and to construct reference series; 2) Detection of breakpoints; 3) Trend correction. Detection and correction are based on the Caussinus-Mestre approach (Caussinus and Mestre 2004).

Since the neighbor stations cannot be expected to be homogeneous, the algorithms are applied iteratively. The software has been successfully validated on the artificial time series of the COST Action HOME. In addition to its application for data homogenization, automation allows a sensitivity study of the uncertainties introduced by estimating the climate signal from neighbor stations.

References:

H. Caussinus and O. Mestre. Detection and correction of artificial shifts in climate series. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 53(3):405–425, 2004.