Kurzfassungen der Meteorologentagung DACH
Bonn, Deutschland, 20.–24. September 2010
DACH2010-317

First results from the COST-HOME monthly benchmark dataset with temperature and precipitation data for testing homogenisation algorithms

V.K.C. Venema (1), O. Mestre (2), and the COST-HOME Team

(1) University of Bonn, Meteorological institute, Bonn, Germany (victor.venema@uni-bonn.de), (2) Meteo France, Ecole Nationale de la Meteorologie, Toulouse, France (olivier.mestre@meteo.fr)

Climate records are known to be inhomogeneous, i.e. they contain changes that are not due to a change in the regional climate. For a regional network of climate stations, the variability can be decomposed in regional climate variability, which occurs at all stations, and noise plus inhomogeneities, which usually occur at only one station. Such inhomogeneities can be detected by comparing neighbouring stations (relative statistical homogenisation). A jump in a single station could be sudden climate change, but a jump in the difference time series of two nearby stations is not related to a change in the regional climate. Removing these inhomogeneities is important, since the jumps in the records may have the same order of magnitude as trends and climate variability.

Typical causes of inhomogeneities are a change in measurement location or of the surrounding, the type of instrument shelters used, typical measurement heights and in the instrumentation itself. There are also changes in measurement procedures such as the way the daily mean temperature is computed, the observation times, and also maintenance procedures and timing.

As part of the COST Action HOME (Advances in homogenisation methods of climate series: an integrated approach) a dataset was generated that serves as a benchmark for homogenisation algorithms. Members of the Action and third parties have been invited to homogenise this dataset. Based upon a survey among homogenisation experts we chose to work with monthly values for temperature and precipitation.

Surrogate and synthetic data was generated to represent homogeneous climate data. To this data known inhomogeneities are added: outliers, as well as break inhomogeneities and local trends. Furthermore, missing data is simulated and a global trend is added.

The participants have returned around 26 contributions. Some fully automatic algorithms were applied, but most homogenisation methods need human input. For well-known algorithms, MASH, PRODIGE, SNHT, multiple contributions were returned. This allowed us to study the importance of the implementation and the operator for homogenisation, which was found to be an important factor. Furthermore, algorithms designed for multiple break detection are clearly better. We analysed both the root mean square error and the difference in the trends between the original data and the homogenised data. Interestingly, these two error measures can lead to a very different ranking for some algorithms.

For more information on the COST Action on homogenisation see:
http://www.homogenisation.org