Geophysical Research Abstracts,
Vol. 11, EGU2009-3943, 2009
EGU General Assembly 2009

Penalized Maximal F Test for Detecting Change Points of Temperature and Wind Speed Data Series

L. Cao, X. Liu, and Q. Li

Climatic Data Center, National Meteorological Information Center, China Meteorological Administration, Beijing, 100081, China

The homogeneity of the climate record continues to receive considerable attention. Time series are commonly contaminated by non-climatic discontinuities that result from station relocations, observation time changes, and station specific trends related to environmental changes in the proximity of the observation sites. Several statistical methods have been proposed for detecting undocumented shifts. Wang Xiaolan et al. proposed the penalized maximal F test (PMFT) for detecting undocumented mean shifts that are not accompanied by any sudden change in the linear trend of time series. This method is based on the penalized maximal F test, which are embedded in a recursive testing algorithm, with the lag-1 autocorrelation (if any) of the time series being empirically accounted for. In this research the PMFT method is used for detecting the shifts of long time series of temperature and wind speed data series over China. The monthly average temperature data of about tens of meteorological observing stations and the annual average wind speed data of 753 meteorological observing stations have been detected. The results show that this homogeneity detection method works well for these two meteorological data series over China.

1. Results of monthly average temperature data series

To create a reference time series is sometimes very useful for homogeneity detection, while it is difficult to get a good reference time series especially for the hundred-year long temperature data with a lot of year data lost over China. The PMFT method are used without building a references series to detecting the change points of the monthly average maximum temperature of twelve meteorological stations and the monthly average minimum temperature of twenty-nine stations. The results show that this method is fit for the homogeneous detection and we needn't interpolating the data and building a reference time series before the detection. Although discontinuities in temperature time series can be caused by any number of changes in, for example, sensor type, and even the observation schedule, station relocations are the likely cause of the majority of abrupt shifts identified in the temperature series evaluated here.

2. Results of annual average wind speed data series

52 of the data series are too short to be detected among 753 meteorological observing stations. A total of 356 change points over 271 stations are detected of the annual average wind speed time series, which accounts for 38.7% of the evaluated stations. The homogeneous data series are of 231 stations and another 199 stations are not significant that can be considered as homogeneous, which accounts for 61.3% of the evaluated stations. It is found that the data of 61.3% stations are homogeneity among the detected 701 stations, which shows that the homogeneity of the annual average wind speed is good. The change points of the annual average wind speed range from 1 to 2. The changes of instrument and location are the main reason for the non-homogeneity, while the change of the type of the observation instrument for the wind speed is the most important reason for the non-homogeneity of the annual average wind speed over China. The environment change seems not so remarkable, because the relocation and the instrument change may take place at the same time to conceal the effect of the environment change.

All of the works we have done are the preliminary experiments of using this method. Although we get some results, there are still a lot of works need to do because the wind speed data are so special and the probability distribution are not the exact Gaussian distribution. At the same time the data of the wind speed are affected mostly by the topography and the barriers aside of the observation fields. The use of reference series can help to diminish departure from Gaussian distribution. We will do more experiments on detecting of the wind speed data. What the

important thing is that rely on most detail metadata information to help the work of homogeneity detection.