



Regional-scale significance of climatic trends over Europe

Radan Huth (1,2,3) and Martin Dubrovský (2,3)

(1) Charles University, Faculty of Science, Dept. of Physical Geography and Geoecology, Praha 2, Czechia (huth@ufa.cas.cz), (2) Institute of Atmospheric Physics, Czech Academy of Sciences, Praha, Czechia, (3) Global Change Research Institute, Czech Academy of Sciences, Brno, Czechia

Detection of trends in climate elements and assessment of whether they are statistically significant are among the most important tasks in current climatology. The vast majority of studies only assesses the significance of trends at individual stations or grid points, that is, on a local level, providing no information on whether the occurrence of significant local trends is significant as a whole, that is, on a regional scale.

To fill this gap and provide a guidance on how to evaluate the regional-scale significance of trends, we compare five methods of assessing regional trend significance: (i) counts of trends of one sign regardless of their local (in)significance; (ii) counts of locally significant trends; (iii) multi-site Kendall test extended to compensate for spatial autocorrelation; (iv) Walker test, based on the smallest p-value (highest significance) of all local tests; (v) false detection rate, which can be considered a generalization of Walker test. The performance of the tests is compared on synthetic data, consisting of 10,000 realizations of time series on a regular rectangular grid with a given spatial and temporal autocorrelation and magnitude of trend, assuming first-order autoregressive process both in time and space. The data are generated by software tool SPAGETTA (SPAtial GEneraTor for Trend Analysis). Time series with no trend are used to create null distributions of test statistics and determine their critical values. Time series with non-zero trends are used to assess the type II error, that is, the probability of accepting the null hypothesis of no trend when it is false. We set the type II error of 5% as the limit of detectability of trends, and calculate the magnitude of a trend that can be detected with such an error for a wide range of spatial and temporal autocorrelations, grid sizes, and lengths of the series. The performance of tests is better (that is, trends of smaller magnitude are detected with type II error of 5%) for longer time series, larger grids, and lower autocorrelations. The multi-site Kendall and sign-counting trends perform best; the gap in the performance between them and the Walker and *fdr* tests gets narrower with increasing autocorrelations. The sign-counting test is, however, not applicable to cases with very high spatial autocorrelations because of its discrete nature.

The tests are applied to the detection of annual and seasonal temperature trends over Europe for various gridded datasets. Temperature change over period 1961-2010 is regionally significant annually and in all seasons except winter. On the example of winter, we demonstrate that even a larger number of locally significant trends (more than a half of all gridpoints) does not grant a regional significance if the spatial autocorrelation is high.