EU Surface Temperature for All Corners of Earth (EUSTACE): break-detection algorithm for a global air temperature dataset

Yuri Brugnara[1,2], Renate Auchmann[1,2], Stefan Brönnimann[1,2]
[1]Oeschger Centre for Climate Change Research – Bern, Switzerland
[2]Institute of Geography, University of Bern – Bern, Switzerland
Contact: yuri.brugnara@giub.unibe.ch

1. Abstract

EUSTACE is a Horizon 2020 project that involves 8 institutes in 4 countries. Its goal is to produce daily estimates of surface air temperature since 1850 across the globe for the first time, by combining surface and satellite data using novel statistical techniques.

Here we describe a fully automatic state-of-the-art break-detection algorithm that was developed for the global LSAT dataset used within EUSTACE. We evaluate the performance of the method using artificial benchmarks and present various statistics related to frequency and size of the inhomogeneities detected in the real data.

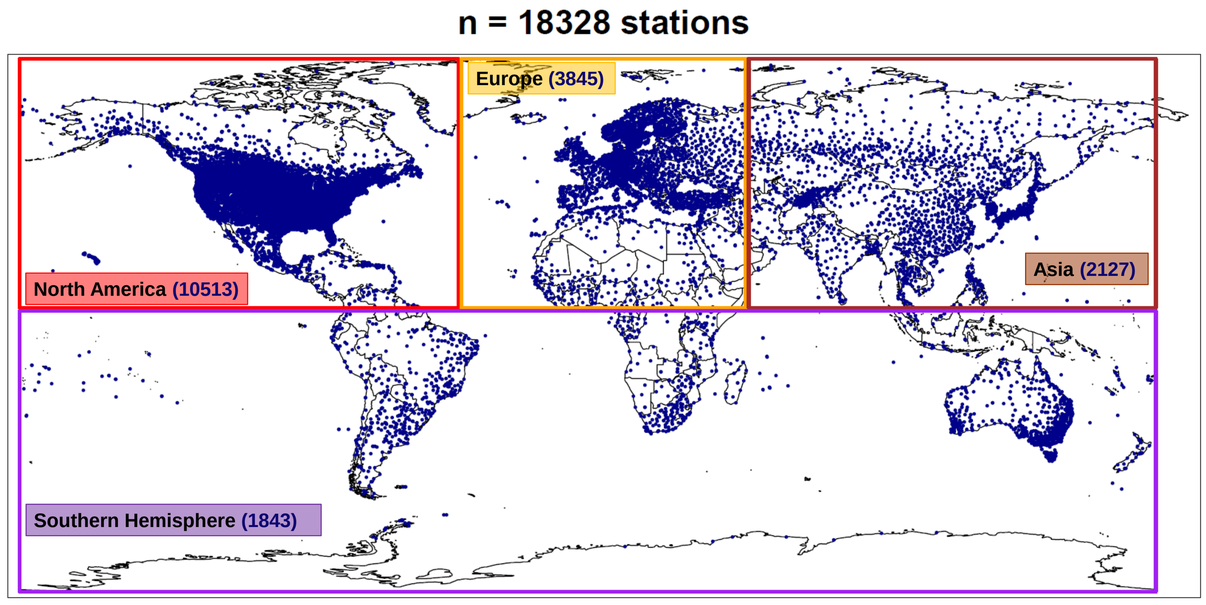


Figure 1 Map of the stations that have enough data and neighbours to perform the break detection (for mean temperature). Data source: ISTI (www.surfacetemperatures.org).

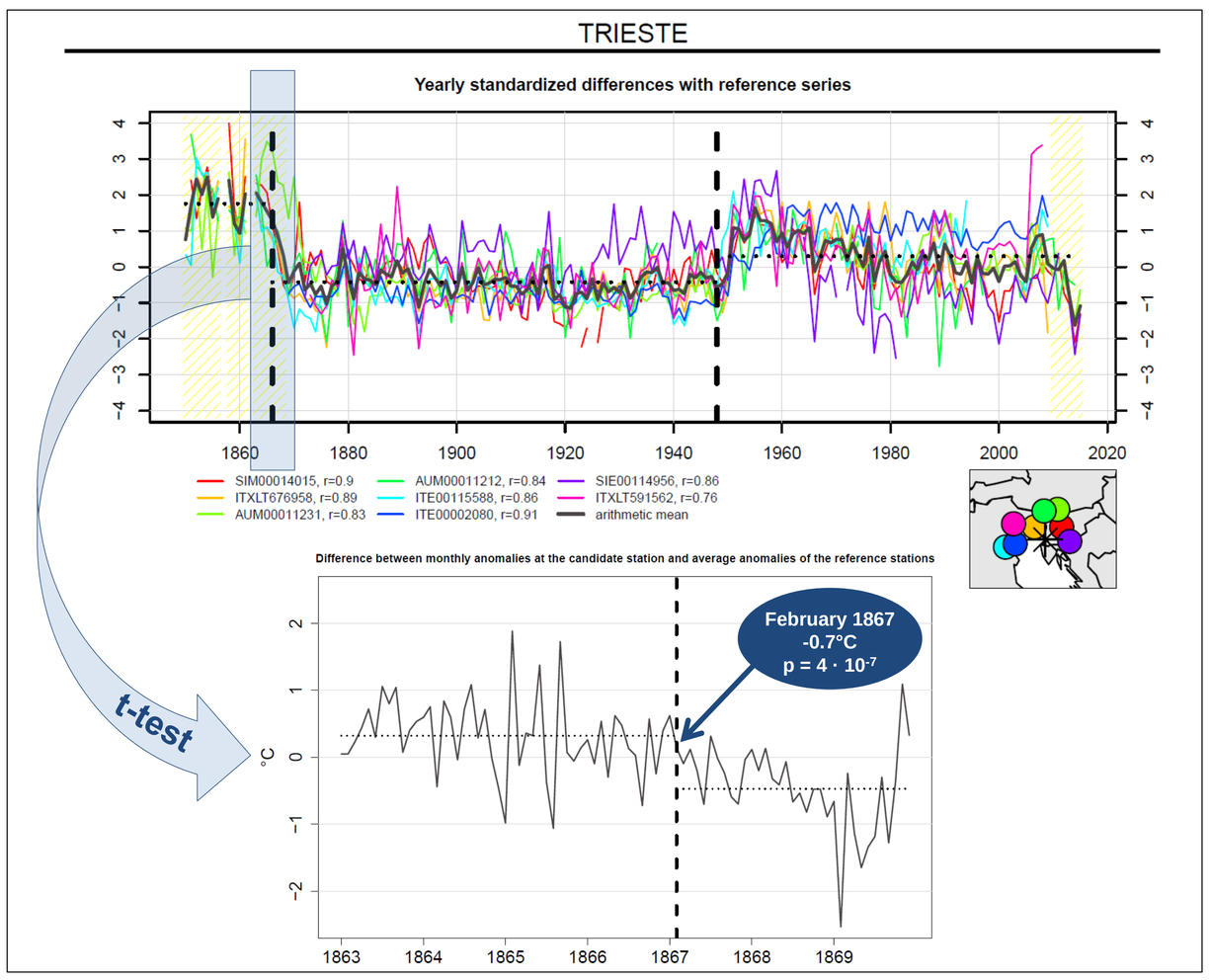


Figure 2 Break-detection algorithm applied to the mean temperature series of Trieste.

2. Automatic break detection

A breakpoint is an inhomogeneity in a time series caused by changes in the instrumentation, the observation procedure, and/or the environment. For a temperature series, a typical breakpoint is caused by the change of the radiation screen.

The breakpoint detection used within EUSTACE applies a fully automatic algorithm, adapted from Kuglitsch et al. (2012). After a selection of up to 8 reference series, which satisfy determinate requirements of distance, correlation and data availability, the detection is performed on yearly and semi-yearly means by combining 3 independent detection methods:

- **PRODIGE** (Caussinus and Mestre, 2004)
- **RHtests** (Wang et al., 2007; Wang, 2008)
- **GAHMDI** (Toreti et al., 2012)

A certain breakpoint is considered **significant if it is detected by at least 2 methods**. The detection is then further refined on monthly data by applying a Student's t-test on a 7-year window around the breakpoint, in order to detect the month when the breakpoint did more likely happen (i.e., where the probability of the null hypothesis has its minimum).

3. Validation & Analysis

The performance of the algorithm can be assessed using a benchmark dataset where breakpoints have been artificially added and are therefore known (surrogate series). Using the benchmark described by Venema et al. (2012) we obtain a **probability of detection** (i.e., number of hits divided by the sum of hits and misses**)** of about **60%** in the best case scenario represented by many high-correlated reference series. The **mean absolute detection error** is in the order of **6 months** (more than 40% of the hits have no detection error).

When we apply the algorithm to real data, we can look for systematic behaviours of the breakpoints. We find that on average the breakpoints cause a reduction of the observed **mean temperature**, i.e. **a reduction of the long-term trend**. This is found in every continent and it is particularly evident in the 1990s (worldwide introduction of automatic weather stations).

The evolution is less clear for the **daily temperature range** (Tmax – Tmin). A large mean change is found in North America in the 2000s, implying that a significant artificial negative trend has been introduced there. Recently, Oyler et al. (2015) documented a **network-wide instrumentation change in the western United States** in the mid-2000s that caused a systematic increase of minimum temperatures and that might explain our results, although we also find a systematic decrease of maximum temperatures.

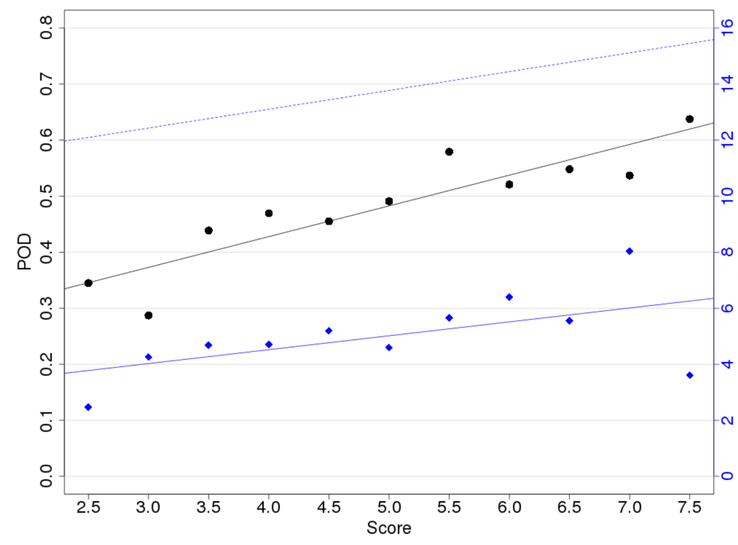
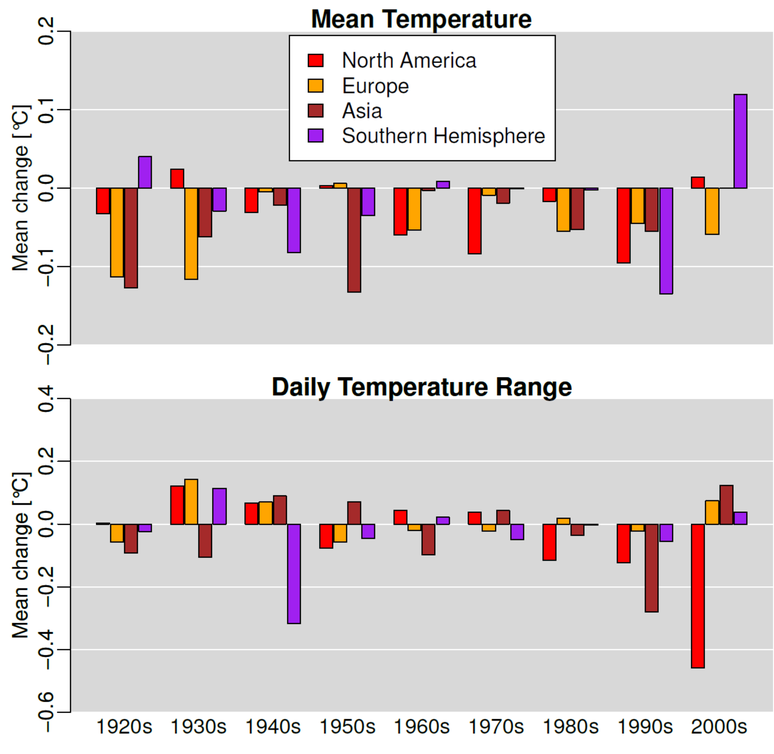


Figure 3 Probability of detection (POD; black points) and mean absolute error (MAE; blue points) as a function of the detection score (proportional to the number of reference series). The dashed blue line represents a regression of the average absolute errors obtained by selecting the month randomly within ± 1 year of the breakpoint detected with annual resolution. 66 unique surrogate series were used to produce these results.

Figure 4 Mean changes caused on average by the detected breakpoint in the four macro-areas defined in Fig. 1, divided by decade. Negative values indicate that the observed temperature has decreased after the breakpoint.



Acknowledgements

We thank Andrea Toreti (EC Joint Research Centre) and Yang Feng (Environment Canada) for providing software for their break-detection methods. Temperature series were retrieved from the International Surface Temperature Initiative's website.

References

H Caussinus and O Mestre. Detection and correction of artificial shifts in climate series. *J. R. Stat. Soc.*, 53: 405–425, 2004.

F G Kuglitsch, et al. Break detection of annual Swiss temperature series. *J. Geophys. Res.*, 117: D13, 2012.

J W Oyler, S Z Dobrowski, A P Ballantyne, A E Klene, and S W Running. Artificial amplification of warming trends across the moountains of the western United States. Geophys. Res. Lett. 42: 153-161, 2015.

A Toreti, F G Kuglitsch, E Xoplaki, and J Luterbacher. A novel approach for the detection of inhomogeneities affecting climate time series. *J. Appl. Meteorol. Clim.*, 51: 317–326, 2012.

V K Venema, et al. Benchmarking homogenization algorithms for monthly data. *Clim. Past*, 8: 89-115, 2012.

X L Wang. Accounting for autocorrelation in detecting mean shifts in climate data series using the penalized maximal t or f test. Journal of Applied Meteorology and Climatology, 47: 2423–2444, 2008.

X L Wang, Q H Wen, and Y Wu. Penalized maximal t test for detecting undocumented mean change in climate data series. *J. Appl. Meteorol. Clim.*, 46: 916–931, 2007.

EUSTACE has received funding from the European Union's Horizon 2020 Programme for Research and Innovation, under Grant Agreement no 640171

Official website: **www.eustaceproject.eu**