



## **Homogenisation algorithm skill testing with synthetic global benchmarks for the International Surface Temperature Initiative**

Katherine Willet (1), Victor Venema (2), Claude Williams (3), Enric Aguilar (4), Ian Joliffe (5), Lisa Alexander (6), Lucie Vincent (7), Robert Lund (8), Matt Menne (3), Peter Thorne (9), Renate Auchmann (10), Rachel Warren (11), Stefan Bronniman (10), Thordis Thorarinsdotir (12), Steve Easterbrook (13), Colin Gallagher (8), Giuseppina Lopardo (14), Zeke Hausfather (15), and David Berry (16)

(1) Met Office Hadley Centre, United Kingdom, (2) University of Bonn, Meteorological Institute, Bonn, Germany, (3) National Climatic Data Center, Asheville, North Carolina, USA, (4) Center for Climate Change, C3, University Rovira i Virgili, Tarragona, Spain, (5) Exeter Climate Systems, University of Exeter, United Kingdom, (6) Climate Change Research Centre, University of New South Wales, Australia, (7) Climate Research Division, Environment Canada, Canada, (8) Department of Mathematical Sciences, Clemson University, USA, (9) Geography Department, Maynooth University, Ireland, (10) Oeschger Center for Climate Change Research & Institute of Geography, University of Bern, Switzerland, (11) College of Engineering, Mathematics and Physical Sciences, University of Exeter, United Kingdom, (12) Statistical Analysis, Pattern Recognition and Image Analysis (SAMBA), Norwegian Computing Centre, Oslo, Norway, (13) Department of Computer Science, University of Toronto, Canada, (14) Istituto Nazionale di Ricerca Metrologica (INRiM), Italy, (15) Berkeley Earth, Berkeley, CA, USA, (16) National Oceanographic Centre, Southampton, UK

Our surface temperature data are good enough to give us confidence that the world has warmed since 1880. However, they are not perfect - we cannot be precise in the amount of warming for the globe and especially for small regions or specific locations. Inhomogeneity (non-climate changes to the station record) is a major problem. While progress in detection of, and adjustment for inhomogeneities is continually advancing, monitoring effectiveness on large networks and gauging respective improvements in climate data quality is non-trivial.

There is currently no internationally recognised means of robustly assessing the effectiveness of homogenisation methods on real data - and thus, the inhomogeneity uncertainty in those data. Here I present the work of the International Surface Temperature Initiative (ISTI; [www.surface temperatures.org](http://www.surface temperatures.org)) Benchmarking working group. The aim is to quantify homogenisation algorithm skill on the global scale against realistic benchmarks. This involves the creation of synthetic worlds of surface temperature data, deliberate contamination of these with known errors and then assessment of the ability of homogenisation algorithms to detect and remove these errors. The ultimate aim is threefold: quantifying uncertainties in surface temperature data; enabling more meaningful product intercomparison; and improving homogenisation methods. There are five components work:

1. Create 30000 synthetic benchmark stations that look and feel like the real global temperature network, but do not contain any inhomogeneities: analog clean-worlds.
2. Design a set of error models which mimic the main types of inhomogeneities found in practice, and combined them with the analog clean-worlds to give analog error-worlds.
3. Engage with dataset creators to run their homogenisation algorithms blind on the analog error-world stations as they have done with the real data.
4. Design an assessment framework to gauge the degree to which analog error-worlds are returned to the original analog clean-worlds by homogenisation and the detection/adjustment skill of the homogenisation algorithms.
5. Present an assessment to the dataset creators of their method skill and estimated uncertainty remaining in the data due to inhomogeneity.