Automated Quality Control for studying extremes in high resolution data: Creating HadISD

R.J.H. Dunn (1), K.M. Willett (1), P.W. Thorne (2,3), E.V. Woolley (1), D.E. Parker (1), I. Durre (2), A. Dai (4), and R.S. Vose (1)

(1) Met Office, Hadley Centre, Exeter, United Kingdom (robert.dunn@metoffice.gov.uk), (2) NOAA NCDC, Asheville, NC, USA, (3) CICS, NC State University, Asheville, NC, USA, (4) NCAR, Boulder, CO, USA

To meet the requirements of climate science in the 21st century we need high resolution, high quality, globally complete datasets in near real time. This demands a highly meticulous approach to data quality control, which must be done in an objective, reproducible and consistent manner. The Met Office Hadley Centre has quality controlled over 6000 stations with sub-daily, synoptic data using an automated suite of tests. We focus mainly on temperature, dewpoint temperature and sea-level pressure, but also includes cloud cover, wind speed and direction. These variables are key to characterising meteorological events with human impacts such as storms and heat waves. Our automated quality control suite addresses many known issues with observational data including individual and clustered outliers, repeated and frequently occurring values using inconsistency with neighbouring stations. We have validated these tests using real, well studied, extreme events such as Hurricane Katrina and the Australian heat waves in 2009. Here we introduce our new product which we plan to update on a yearly basis, describing our methodological choices and validation. Additionally, all software developed for the creation of this product is also now publicly available, the pros and cons of which will be briefly discussed.