

Review of “HadISD: A quality controlled global synoptic report database for selected variables at long-term stations from 1973-2010”

Neal Lott, NOAA’s National Climatic Data Center

15 July 2012

General comments: Overall this is an excellent paper with no serious weaknesses or issues. More importantly, it reflects an outstanding research and development effort. The quality control techniques and algorithms appear to be sound and based on an iterative process of testing and improvement. The figures and tables are well organized and provide good examples of various QC checks.

Specific comments:

- In the introduction, reference is made to the ISD database being “non-trivial for the non-expert” to access. It should be noted that NCDC provides access to ISD via a GIS interface and with the ability for users to select desired parameters into a space or comma-delimited file. Providing this information would be useful for readers to know.
- The paper references ISD-Lite. Was this input source used in the effort? If so, there may be some questions regarding the ISD vs ISD-Lite that should be addressed. If not, then it’s a non-issue.
- In section 3, para 1, the paper states that a large number of stations report only rarely. This should be quantified in some way. Readers may infer that well over half of the stations have very little data.
- In section 3, para 2, the paper refers to the first 250 lines of each annual file. It might be useful to refer to the first 250 observations or the first 250 data records, so it’s obvious to the reader what this means.
- Regarding section 4, were the authors aware of NCDC’s online documentation regarding known problems in ISD? It would be interesting to know if the QC checks performed by the authors “caught” some of these problems.
- Regarding section 4.1.11, providing an example for a station here might be helpful. Also, several of the figures refer to the affect that station compositing may have on the QC results (eg, increasing the flagging rate) – very good to see that this was evaluated. For the QC checks in 4.1.11, could compositing have resulted in some artifacts in the data which would have affected the QC results? (NCDC recently ran a check which showed no occurrences of dew point higher than temperature in the ISD database – ie, for specific stations in each annual file.)
- Was consideration given to process the QC checks both before and after station compositing, at least for certain algorithms?