UEA-CRU Responses to Questions 20.04.10

We wish to accurately record the exact details of raw instrument temperature station availability for each CRUTEM data set.

We now understand the following:

1. Raw instrument temperature station data is available from NMOs

2. Aggregated instrument temperature station data is available from WWR and GHCN

Questions:

A. Are there any other aggregated data sources ? We are unsure about the role and availability of NCAR as quoted in a UEA FoI response as being a source of data along with GHCN. We have visited the site and it does not seem to be set up as a global repository.

NCAR is a global data repository (many more variables than temperature). It is one of the biggest climate data repositories in the world. Its datasets (and tools) are introduced on this web site:

http://www.ncar.ucar.edu/tools/datasets/

You will see that they provide a number of different interfaces for identifying or searching for datasets, including the three listed in the central column of this webpage which are "The Community Data Portal", "EOL's Catalog of Field Expedition Data Sets", and "CISL Research Data Archive". If you click on the latter (i.e. "CISL Research Data Archive") you will reach this web page:

http://dss.ucar.edu/index.html

If you type "surface station temperature" in the "Dataset Search" box and click "Go" you will get a list of many such datasets. Perhaps the one of particular interest here is the one named "**World Monthly Surface Station Climatology, 1738-cont (ds570.0)**". The dataset description begins:

"This world monthly surface station climatology has data for over 4700 different stations (2600 in more recent years). Data for some stations goes as far back as the mid-1700's. See decadal coverage for more detail. Most of the data was obtained directly from the National Climatic Data Center (NCDC), …"

You could also get there by doing a google search on 'ncar dataset station temperature' and this webpage is the first one:

http://dss.ucar.edu/catalogs/free.html

which also lists the DS 570 dataset. Whichever way you get there, if you click on this dataset you will get to here:

http://dss.ucar.edu/datasets/ds570.0/

Dataset DS570 (above) is World Weather Records (WWR). It says it is updated annually, though it is actually slightly more frequently than that and it is about 6 months behind real time. Current last month is Oct 2009.

Earlier versions of this dataset (possibly with a different number) have been available in this format for decades. This dataset is the World Weather Record (WWR) dataset that we refer to in the 1986 papers. In the 1980s we got a version of this dataset on magnetic tape. Updates came this way as well for CRUTEM1.

B. Please can you provide full details as to how one would find and get hold of records through WWR? Our own researchers have only found *http://www.wmo.ch/pages///prog//wcp/wcdmp/wwr/index_en.html* *which is unhelpful.*

- *It gives only an email address to obtain a CD but gives no indication as to which station data is available.*
- *It links to NCDC in a way which does not seem to make access easy. It was unclear how one would use this link to obtain a set of global records.*
- *It is also unclear as to whether the WWR-NCDC link is related (or not) to the NCDC-GHCN link*

Dataset DS570 (above) is World Weather Records (WWR).

Dataset DS570.1 is a subset comprising WWR data for some specific decades (they are included in the main DS570.0 dataset). The link below shows the different decades referred to as WWR

http://dss.ucar.edu/datasets/ds570.1/

This gives World Weather Record (WWR) data for the 1961-70, 1971-80 and 1981-90 decades. These are the 10-year books of data that you saw in the CRU Library. The earlier volumes (pre-1921, 1920s, 1930s, 1940s and 1950s) are the original paper sources that comprise WWR (and hence DS570). These individual decades (the 1960s, 1970s and 1980s) within DS570.1 will have been incorporated within DS570.

Two other subsets are also available, but all are included in the main DS570.0 dataset:

http://dss.ucar.edu/datasets/ds570.2/

http://dss.ucar.edu/datasets/ds570.3/

WMO have limited resources to maintain their web site and therefore some links may not go directly to the desired pages (indeed some are out of date and no longer work). To reduce maintenance costs they may link to another organisation's home page rather than to a specific dataset page in case the latter later gets moved. WMO may have to link to a NOAA website which will be the one at NCDC, Asheville, because NOAA is an NMS whereas NCAR is not.

The latest version of the 10-year book (the 1990s) is available as a CD, which we obtained from NCDC. These 1990s data (still referred to as WWR by some) will have been incorporated into DS570 above and also into GHCN. We incorporated these data in

CRUTEM2 (Jones and Moberg, 2003). This might be the CD that WMO is referring to. Copies of this CD will have been sent to all NMSs at the time.

Another link that can be got to from the NCDC site (with some difficulty) is this one

http://www.ngdc.noaa.gov/wdc/

which describes the system of World Data Centers and there are links to World Data Centres (WDC) around the world. There are, in fact, three WDCs for meteorology (NCDC, China and Russia) and they provide another way of accessing climate station data from around the world. The NCDC link takes you to:

http://www.ncdc.noaa.gov/oa/wdc/index.php

which has links to various data sets (e.g. "Climate of the World").

It is a little difficult to navigate around these and it does require a bit of knowledge about what you're looking for and what it might be called. The WDC also provides a link to another large database called the Global Change Master Directory (GCMD)

http://gcmd.gsfc.nasa.gov/KeywordSearch/Home.do?Portal=wdc&MetadataType=0

C. GHCN data is clear and easy to obtain today (we have done this). Can you confirm the period for which this was available, and any periods when it was restricted. For example we understand it restricted data in 2003/4 due to an WMO resolution?

GHCN has been available on an NCDC web site since the late 1990s. I suspect this has been the case since web access became possible. We obtained versions by ftp in the late 1990s. The first reference to GHCN seems to be a paper from 1997. Before GHCN, NCDC used to make WWR data available. WWR is also the basic source of GHCN.

GHCN access was restricted to some European countries during the period 2002-2004. I cannot recall the exact dates or the countries involved. The countries included the UK and some other European countries. These European countries (their NMSs) complained to WMO and NOAA that NCDC/NOAA (through GHCN) were giving free access to their national data (when it was in GHCN). The issue was resolved about a year later. It was easy to bypass the restrictions as all you had to do was to ask an American or Canadian colleague to send you the current copy of GHCN. The WWR dataset was also freely available at this time, but it seems that most people had forgotten about WWR being available at NCAR. The access issue was a disagreement between some European NMS and NOAA about data availability and related to the charging for national climate data by some European NMSs. To my knowledge it wasn't related to WMO Resolution 40.

As an aside, WMO Res. 40 is a list of climate data that NMSs should make freely available. Most have to be asked to do this and it can take ages. WMO has no power to enforce it but they have passed it.

D. Assuming the list in '2' above is complete, it is important that we know exactly what fraction of CRUTEM station data would have been available from them at what times. Please can you detail this, i.e

- *CRUTEM1986: At time of publication X% available from WWR, Y% available from GHCN.*
- *CRUTEM(1): At time of publication X% available from WWR, Y% available from GHCN.*
- *CRUTEM2: At time of publication X% available from WWR, Y% available from GHCN.*

- *CRUTEM3: At time of publication X% available from WWR, Y% available from GHCN.*

Please indicate if there was any significant change to availability after publication date. In any case where coverage is < 100% please state where the remaining data could be obtained.

It isn't always a simple case of saying that X% was from WWR and Y% was available from GHCN.

CRUTEM1986 – Apart from the data we had digitised during the 1980s (see TR documents and papers), the rest would have come from WWR (so the then equivalent of DS570). None would have come from GHCN as GHCN did not exist then. All the station temperature data we used (including the data we digitized and the adjustments we made) have been available here for a very long time:

http://cdiac.ornl.gov/ftp/ndp020/

In particular, see the dataset description here:

http://cdiac.ornl.gov/ftp/ndp020/ndp020r1.txt

and the NH and SH station data at these locations:

http://cdiac.ornl.gov/ftp/ndp020/jonesnh.dat

http://cdiac.ornl.gov/ftp/ndp020/jonessh.dat

The Antarctic station temperature data were, and still are, also available here:

ftp://cdiac.ornl.gov/pub/ndp032/old/

which was updated later here:

ftp://cdiac.ornl.gov/pub/ndp032/

(again, see the ndpNNN.txt files for the dataset descriptions).

These are all available through the Global Change Master Directory (GCMD) and the Carbon Dioxide Information Analysis Center (CDIAC), and have data through 1991 (1999 for the updated Antarctic set). So in terms of X and Y, Y is zero and X would be about 90% from WWR. The only data that were not in WWR were data CRU digitized (or gained access to in digital form) and these data were documented in the TR publications. The extra 10% of data was available in the various CDIAC sources (with the other 90%).

After publication, the real-time data came from two sources (CLIMAT and MCDW). CLIMAT is the data exchanged between NMSs that we obtained from the Met Office. From the early 1980s we received CLIMAT as hard copy print outs, which we input using UEA's system for typing in data and programs from coding sheets. This changed at some point to floppy disks in the 1990s. About 2000 we gained access to the data by email and then through a web site. Monthly Climatic Data for the World (MCDW) came to us as a publication from NCDC. This was added manually until it became available from NCDC in a digital format and we accessed it by ftp. CLIMAT and MCDW come from the same data source (the NMSs). MCDW was printed later, so included more back data and also included more station data from the US.

MCDW is available in Met Services libraries around the world. We have copies in the CRU library. It has been available at this ftp site since 1995

ftp://ftp1.ncdc.noaa.gov/pub/data/mcdw/

The first month there is September 1994 (and this was put up on 8 Aug 1995, file called ssm9409.fin). The latest month there is currently ssm0912.fin .

CLIMAT is available on the http://www.hadobs.org/ web site from Jan 2000 onwards

For CRUTEM1, the same data as CRUTEM1986 were used together with an extra 1088 stations that came from the availability of more digital data (no more were digitized by CRU). These digital sources were extra data from WWR and the rest would have come from NCDC (essentially pre-cursors of GHCN that we would have got by ftp). Both would now be in GHCN. The WWR volume data for the 1980s was not fully published until the late 1990s.

In terms of X and Y, Y is zero. All the data would have been available from either the GCMD, the pre-cursors of GHCN at NCDC, WWR or from CLIMAT/MCDW.

Updates to CRUTEM1 came from the CLIMAT and MCDW sources

CRUTEM2 and CRUTEM3. CRUTEM2 included the WWR data for 1991-2000 which were also included into GHCN. In our response to Issue 4 (from March 1) we estimated that there are 430 usable stations (i.e. those that have 1961-1990 normals) that are not in GHCN. This means that about 10% of the data are not available through GHCN (430/4348). So Y is ~90%. Some of these 430 would have been available in the GCMD source. Others would have been available at the NMS web sites of a number of countries given in the Tables in Jones and Moberg (2003). These include Australia, New Zealand, Canada and a number of European project websites such as HISTALP and IMPROVE.

Updates to CRUTEM2 and CRUTEM3 came from the CLIMAT and the MCDW sources, which by the time of CRUTEM2 were available from ftp or web sites. We also add real-time data from Australia into CRUTEM2 and CRUTEM3 (but not every month). These data would have been available on the Bureau of Meteorology web site in Melbourne, where we get them from.

Additional Points about the Minutes from the March 4, 2010 meeting

The minutes are fine, except for two aspects

#4 bullet point 2. This should read ‘Following the 2003 update (CRUTEM2), multiple versions were found to exist for some stations….’

So change “Follow” to “Following” and ‘data sets’ to ‘versions’

#15 The additional figure removed data from China, Taiwan, Hong Kong and Macau. It did not exclude data from Korea, so that sentence needs to be corrected.