Results of the ETMC GTS comparison:
Dec. 2007 data

David Berry and Elizabeth Kent
Outline

- Purpose of the comparison
- The data streams
- Results
- Conclusions
Purpose of the comparison

- Observations exchanged on the GTS form part of the climate record
- Some observations are available only in real time and are not submitted through the delayed-mode system
- Questions include:
  - Do the GTS-derived data streams at different centres contain the same reports?
  - Do the reports common to each data stream have any differences in content?
  - Is there any benefit to using multiple GTS data streams?
  - How does the GTS stream compare with the delayed mode? (not addressed in detail here)
The WMO requested that JMA, ECMWF, NCDC and the Met Office provided GTS data for December 2007 for comparison. Unfortunately the request was sent out after Dec. 2007 meaning that the data had to be retrieved from archives.

- This did have the benefit that we were able to additionally examine the archival at the different centres.
- But it did present additional technical difficulties.
- Complicated by the masking of callsigns.
JMA

- multiple reports in one bulletin, header on same line, "=" denotes end of report

```plaintext
SMWF01 ENMI 200000 BBXX LDWR 20001 99660 10023 41398 82414 10072 20049 40245 58004 70222 887// 22200 04078 10908 3/// 4/// 5/// 70040=

SMVD22 KWBC 200000 RRD BEXX 46028 20001 99357 71219 46/// ////// 10122 40243 55000 92350 22200 00125 11307 70033= 46069 20001 99336 71202 46/// /3405 10130 40219 55000 92350 22200 00137 11306 20501 329/// 41306 70032 333 91207 555 11056 22058 32303 43308 62349 338053 337053 335057 33650 332064 336062= 46042 20001 99368 71224 46/// /1503 10112 40224 56009 92350 22200 00127 11406 20401 329/// 41406 70029 333 91204 555 11028 22029 32328 41405 62349 147027 140032 141031 147027 150024 151020= 46089 20001 99459 71258 46/// /2310 10089 30009 40009 54000 92350 22200 11009 20605 323/// 41007 70044 333 91213 555 11108 22115 32307 42 213 62349 234099 229087 227093 229089 224101 227085= 46053 20001 99342 71199 46/// /2802 10139 40221 56004 92350 22200 00142 11303 328/// 41303 70017 333 91203 555 11021 22021 62349 276020 287024 279030 267030 261027 244026=
```
One report per line, each with own header

<table>
<thead>
<tr>
<th>Code</th>
<th>WCY7052</th>
<th>01004</th>
<th>99485</th>
<th>71268</th>
<th>41298</th>
<th>31016</th>
<th>10077</th>
<th>20070</th>
<th>40180</th>
<th>54000</th>
<th>70222</th>
<th>834 // 22222</th>
<th>0 // 20101</th>
<th>328 // 40905</th>
<th>5 // 6 // 80073</th>
<th>ICE // // //</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code</td>
<td>WDB9951</td>
<td>01004</td>
<td>99443</td>
<td>11512</td>
<td>41397</td>
<td>42916</td>
<td>10055</td>
<td>2131</td>
<td>40148</td>
<td>57010</td>
<td>70111</td>
<td>84150</td>
<td>22215</td>
<td>00050</td>
<td>20402</td>
<td>325 // 41204</td>
</tr>
<tr>
<td>Code</td>
<td>WDB9951</td>
<td>01004</td>
<td>99443</td>
<td>11512</td>
<td>41397</td>
<td>42916</td>
<td>10055</td>
<td>2131</td>
<td>40148</td>
<td>57010</td>
<td>70111</td>
<td>84150</td>
<td>22215</td>
<td>00050</td>
<td>20402</td>
<td>325 // 41204</td>
</tr>
<tr>
<td>346</td>
<td>SNVD001 KWBC 042300 RRJ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>---------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44029 04231 99425 70706 46/30 /3110 11016 49966 51029 92304 22200 00076 10502 70010 333 91213 555 11127 22133=</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44138 04231 99443 70536 46:/// /2301 10070 49890 57004 22200 00076 11410 70051 333 91203=</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44137 04231 99422 70620 46:/// /2517 10045 49863 52006 22200 00091 11316 70078 333 91222=</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46232 04231 99326 71174 46:/// ///## 1/// 92322 22200 00158 11204 20902 328//= 41203 70019=</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46088 04231 99483 71232 46:/// /3601 10098 20092 40197 92320 22200 00084 333 91202 555 11016 22018 32252 41005 62319 004013 020021 344025 074019 280014 272017=</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Other data streams - not yet compared

- **Met Office**
  - Archived in BUFR with some translation of fields, therefore more difficult to identify matches

- **GCC**
  - Delayed mode data

- **ICOADS**
  - merge of NCEP GTS, GCC and other data sources
Number of reports (NCDC, JMA & ECMWF)

- Once duplicates had been excluded 111374 VOS reports remained
- 3.3k additional reports were close matches
  - 2.5k cases only difference is addition of dewpoint
    - dewpoint mostly present in NCDC only
    - otherwise identical report with missing dewpoint in all 3 streams
  - 0.6k reports with other elements present/absent
    - often wave elements and cloud information
  - Some of the remaining close matches seemed to be reports corrected/changed by the ship and resubmitted
  - A few reports are just different and which is correct is unclear
Missing reports

NCDC

ECMWF

JMA
Missing reports

Histograms for different sources:
- NCDC
- ECMWF
- JMA

www.noc.soton.ac.uk
Missing reports
Missing reports

- NCDC: 5472 (5%)
- ECMWF: 3672 (3%)
- JMA: 5484 (5%)
Results

- There are differences between the streams
  - Some are more complete than others - in this case ECMWF had more reports than NCDC or JMA
  - Some streams are fairly similar - in this case NCDC and JMA

- Gains from combining streams vary
  - ECMWF & (JMA or NCDC) gives ~3k more than ECMWF alone
  - JMA & NCDC gives ~300 extra reports
Results

- Missed reports seemed to be most strongly related to reporting hour e.g.
  - ECMWF missed > 1200 report at 0000
  - ECMWF missed ~ 500 reports at 0600/1200/1800
  - ECMWF missed < 100 reports at other hours
  - NCDC missed ~ 100 reports 0000/0300/0600 etc.
  - NCDC missed ~ 300 reports at other hours
  - JMA similar to NCDC but with more missed 0600 cf 0300 etc.

- Reasons unclear but may be due to cut-offs at different centres
Results

- Some reports have missing elements that are present in otherwise matching reports
  - mostly dewpoint, but also waves and clouds
- Some reports are flagged in header as being changed by the centre
  - in these cases we couldn't find any close matching reports
- Some reports are present in the streams with both present/absent callsigns
- Some matching reports have substantial differences
  - Some seem to be resent by ship
  - A few have differences which are not easy to explain
NCEP GTS & Delayed Mode Data

- 700: UK Met. Office VOSClim GTS BUFR Data
- 874: Shipboard Environmental (Data) Acquisition System (SEAS)
- 926: International Maritime Meteorological (IMM) Data
- 927: International Marine (US- or foreign-keyed ship data)
Conclusions

- From this study a more comprehensive GTS data archive would be obtained by merging streams.
- More information on exactly how centres handle GTS data would be extremely helpful: e.g.
  - why do some reports have missing elements?
  - why do some centres miss more reports at certain hours?
- Comparisons with delayed mode data may help resolve some of the unexplained differences.
- Need to feed back results to centres.