COMPARISON OF GTS DATA STREAMS

Revised version of document presented at ETMC-III by David Berry and Elizabeth Kent
Original available via http://icoads.noaa.gov/etmc/

Summary and Purpose of Document

This document provides information on a comparison of three GTS data streams: from the NOAA National Climatic Data Centre (NCDC), the Japan Meteorological Agency (JMA), and the European Centre for Medium-Range Weather Forecasts ECMWF, for December 2007. For this study, it proved impractical to compare GTS data from a fourth data stream, from the UK Met Office, due to translation of the original SHIP (FM13) data into BUFR.

DISCUSSION

1. Background

Reports from different GTS data streams have been compared for the month of December 2007. There were differences in the contents of the streams in terms of the reports present and absent in each data stream. Some similar reports (same position, time and callsign) were found to differ in some elements, the reasons for this are not always clear, but probably include: reports being corrected and sent again by the ship; and particular elements being added or excluded. The results of the study are described in Appendix A.

2. Further Work Required

Difficulties with the BUFR data stream archived by the UK Met Office meant that considerable time was expended before it was concluded that a comparison including this stream was not practical in the time available. A comparison of the GTS and delayed mode data stream has therefore not yet been carried out and would be extremely valuable. Further work may allow the Met Office data stream to be included in this comparison, if that was thought to be worthwhile.

Moreover, additional GTS data streams could potentially be included in a follow-up comparison, including the more heavily processed NCEP BUFR (plus attached original FM13 message string) data now used for ICOADS (http://icoads.noaa.gov/rt.html), which have been subject to a "dup-merge" processing in which exact duplicates were removed and partial duplicates blended to create more complete BUFR reports.

3. Recommendations

(a) When using GTS data in the climate record data streams from multiple centres should be merged, preferable from different regions. In doing this the number of observations in the record can be increased by 3 – 5 %.

(b) The BUFR template as used by the Met Office introduces differences to reports which make it hard to compare with the original FM13 data.

(c) The merged GTS data stream should be compared to the delayed mode data stream to determine whether any of the unresolved report differences can be understood using quality controlled data. Any "close match" report elements which differ and for which the reason cannot be determined should be flagged as suspect in the climate record.

APPENDIX A

RESULTS OF COMPARISON OF GTS STREAMS FOR DECEMBER 2007

David I. Berry and Elizabeth C. Kent

1 Introduction

This report is a comparison of 3 data streams extracted from the GTS for December 2007. The aims were:

1) To document the completeness of each stream and determine the potential advantages of merging data streams either for real time applications or for the climate record.

2) To determine the extent of any differences in the contents of reports in the different streams.

2 Data

Three different data streams containing surface weather reports in FM13 format (WMO 2009a,b) extracted from the Global Telecommunications System (GTS) during December 2007 by different weather centers have been compared. Each of the three streams contains both Voluntary Observing Ship (VOS) reports and buoy data in FM13 format. In this comparison only the VOS data has been used and the buoy data discarded.

The original intention was for the request to be made for data at a date in the future, unfortunately the request to the centers was delayed without the date being changed. This resulted in the data having to be retrieved from archive and one centre was therefore unable to provide data in FM13 format and provided BUFR formatted data (WMO 2001), significantly complicating the comparison. This did however have the advantage of demonstrating each center's archive recovery capabilities. The first stream contains GTS data archived by the US National Climatic Data Center (NCDC). This stream contains one weather report per line with the GTS bulletin headers removed and with all reports received during December 2007 in a single file. The second stream contains data extracted from the GTS by the European Centre for Medium-Range Weather Forecasts (ECMWF). The data in this stream is split into daily files, with all reports received on a given day included in the file for that day. These daily files contain the GTS bulletin headers followed by one or more weather reports split over multiple lines. The third stream contains data extracted from the GTS by the Japan Meteorological Agency (JMA). This stream is also split into daily files. Each line in the file contains a GTS bulletin, i.e. a bulletin header followed by all the weather reports contained within that bulletin. Examples of the data contained in each of these 3 streams is given below (Figures 1 - 3). Table 1 lists the variables available from the different data streams.

```
BBXX01120120070004 WCY7052 01004 99485 71268 41298 31016 10077 20070 40180
54000 70222 834// 22222 0//// 20101 328// 40905 5//// 6//// 80073 ICE /////

BBXX01120120070004 WDB3161 01004 99359 71273 41498 63426 10140 2011/ 40150
57020 70222 867// 22235 00169 20804 334// 41105 5//// 6//// 80120 ICE
/////

BBXX01120120070004 WDB3161 01004 99359 71273 41498 63426 10140 20110 40150
57020 70222 867// 22235 00169 20804 334// 41105 5//// 6//// 80120 ICE
/////

BBXX01120120070004 WDB9951 01004 99443 11512 41397 42916 10055 2131/ 40148
57010 70111 84150 22215 00050 20402 325// 41204 5//// 6//// 81020 ICE
/////

BBXX01120120070004 WDB9951 01004 99443 11512 41397 42916 10055 21310 40148
57010 70111 84150 22215 00050 20402 325// 41204 5//// 6//// 81020 ICE
/////
```

```
BBXX01120120070004 WDC4696 01004 99289 71385 41398 80717 10217 2014/ 40210
56012 70322 886// 22255 00205 20503 305// 40904 5//// 6//// 80172 ICE
/////

BBXX01120120070004 WDD3826 01004 99276 70796 41598 30709 10260 2023/ 40200
53010 70211 83218 22285 0//// 20302 33404 40606 5//// 6//// 80240 ICE
/////
```

Figure 1: Example data from NCDC GTS data stream

```
346

SNVD01 KWBC 042300 RRJ

BBXX

44029 04231 99425 70706 46/30 /3110 11016 49966 51029 92304 22200

 00076 10502 70010 333 91213 555 11127 22133=

44141 04231 99430 70580 46/// /2714 10074 49869 52005 22200 00098

 11413 70063 333 91218=

44138 04231 99443 70536 46/// /2301 10070 49890 57004 22200 00076

 11410 70051 333 91203=

44137 04231 99422 70620 46/// /2517 10045 49863 52006 22200 00091

 11316 70078 333 91222=

46232 04231 99326 71174 46/// ///// 1//// 92322 22200 00158 11204

 20902 328// 41203 70019=

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=

46237 04231 99378 71226 46/// ///// 1//// 92321 22200 00120=

42362 04231 99278 70907 46/// ///// 10160 20090 40176 56010 92300

 22200 00250=

42362 04231 99278 70907 46/// ///// 10160 20090 40176 92315 22200

 00250=

44040 04231 99410 70736 46/// /3110 10003 21065 40043 92319 22200

 10301 70007 333 91214 555 11118 22124=
```

Figure 2: Example data from ECMWF GTS data stream

```
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 BBXX 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 336060
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=
```

Figure 3: Example data from JMA GTS data stream

The data received from ECMWF appeared to be the least processed stream, for example containing many more duplicated reports than the other streams. The NCDC stream appeared to be the most processed and the JMA stream intermediate to the others.

Table 1: FM13 elements matched in the comparison of the GTS streams

Section	Group	Element	Description
0	D….D	DDDD	Ship's callsign
	$YYGGi_w$	YY	Day of the month
		GG	Time of observation to nearest hour
		i_w	Wind speed indicator
	$99L_aL_aL_a$	$L_aL_aL_a$	Latitude in degrees and tenths
	$Q_cL_oL_oL_oL_o$	Q_c	Quadrant of the globe
		$L_oL_oL_oL_o$	Longitude in degrees and tenths
1	i_ri_xhVV	i_R	Indicator figure for precipitation group
		i_X	Indicator for type of station
		H	Height of base of lowest cloud
		VV	Horizontal visibility at surface
	Nddff	N	Total amount of cloud in eighths (oktas)
		dd	True direction, in tens of degrees, from which the surface wind is blowing
		ff	Speed of surface wind
	00fff	fff	Wind speed in units indicated by i_W of 99 units or more
	$1s_nTTT$	s_n	Sign of air temperature
		TTT	Air temperature in whole degrees and tenths
	$2s_nT_dT_dT_d$	s_n	Sign of dew point temperature
		$T_dT_dT_d$	Dew point temperature in whole degrees and tenths
	$3P_0P_0P_0P_0$	$P_0P_0P_0P_0$	Station-level pressure in hPa (omitting thousands figure)
	$4a_3hhh$	a_3	Standard isobaric surface for which the geopotential is reported
		hhh	Geopotential of an agreed standard isobaric surface given by a_3, in standard geopotential metres, omitting the thousands digit.
	4PPPP	PPPP	Sea level pressure in hPa (omitting thousands figure)
	5appp	a	Characteristic of pressure tendency during the three hours preceding the time of observation
		ppp	Amount of barometric tendency (i.e. net change in barometer reading) in the three hours preceding the observation expressed in tenths of a millibar.
	$6RRRt_R$	RRR	Amount of precipitation which has fallen during the period preceding the time of observation, as indicated by t_R.

Section	Group	Element	Description
		t_R	Period of observation for RRR
	$7wwW_1W_2$	ww	Present weather
		W_1	Past weather 1
		W_2	Past weather 2
	$8N_hC_LC_MC_H$	N_h	Amount of low cloud present or, if no low cloud present, all the medium cloud present
		C_L	Type of low cloud
		C_M	Type of medium cloud
		C_H	Type of high cloud
	9GGgg	GG	Time of observation in hours UTC (if different from GG in Section 0)
		gg	Time of observation in minutes UTC (if different from GG in Section 0)
2	$222D_sv_s$	D_s	Ship's course made good during the three hours preceding the time of observation
		v_s	Ship's average speed made good during the three hours preceding the time of observation
	$0s_sT_wT_wT_w$	s_s	Sign and type of sea surface temperature measurement
		$T_wT_wT_w$	Sea surface temperature in degrees and tenths
	$1P_{wa}P_{wa}H_{wa}H_{wa}$	$P_{wa}P_{wa}$	Period of waves (wave recorder) in seconds
		$H_{wa}H_{wa}$	Height of waves (wave recorder) in ½ metres
	$2P_wP_wH_wH_w$	P_wP_w	Period of sea waves in seconds
		H_wH_w	Height of sea waves in units of ½ metres
	$3d_{w1}d_{w1}d_{w2}d_{w2}$	$d_{w1}d_{w1}$	Direction, in tens of degrees, from which the first swell waves are coming
		$d_{w2}d_{w2}$	Direction, in tens of degrees, from which the second swell waves are coming
	$4P_{w1}P_{w1}H_{w1}H_{w1}$	$P_{w1}P_{w1}$	Period of first swell waves in seconds
		$H_{w1}H_{w1}$	Height of first swell waves in ½ metres
	$4P_{w2}P_{w2}H_{w2}H_{w2}$	$P_{w2}P_{w2}$	Period of second swell waves in seconds
		$H_{w2}H_{w2}$	Height of second swell waves in ½ metres
	$6I_sE_sE_sR_s$	I_s	Type of ice accretion
		E_sE_s	Thickness of ice accretion in cm
		R_s	Rate of ice accretion
	$70H_{wa}H_{wa}H_{wa}$	$H_{wa}H_{wa}H_{wa}$	Wave height measured in units of 0.1 metre
	$8s_wT_bT_bT_b$	s_w	Indicator for the sign and type of wet bulb temperature reported
		$T_bT_bT_b$	Wet bulb temperature in degrees and tenths
	ICE + $c_iS_ib_iD_iz_i$	c_i	Concentration or arrangement of sea ice
		S_i	Stage of development
		b_i	Ice of land origin
		D_i	Bearing of ice edge
		z_i	Ice situation and trend of conditions over the preceding 3 hours

3 Merging of Data Streams

The different data streams have been merged by first indexing the VOS weather reports contained in the NCDC data stream based on the location and time of the weather reports. The VOS reports contained in the remaining 2 data streams have then been added to this index, checking for reports already indexed and where a match on location and time is found the report categorized into one of four categories (Table 2).

Table 2: Categories for merging process

Category	Description
New report	Same location and time but with different FM13 elements
Duplicate report	Same location, time and FM13 elements (including callsign)
Masked report	Same location, time and FM13 elements but with different callsign
Close match	Same location, time and callsign but with FM13 elements that differ

4 Results

4.1 Number of reports

In total, excluding duplicate reports, 111374 VOS weather reports were indexed from the three different data streams. 108041 of these reports were found to be unique (including some observations masked by Japan and USA from mid-December 2007) and 3333 further reports identified as close matches to those unique observations. In nearly two-thirds of these reports the only difference between the close matches is the inclusion of the dew point temperature in one and not the other. Where this occurs, the report without the dew point tends to be present in all three streams whilst the report with the dew point included is only present in the NCDC stream. In a few cases, the report with dew point occurs in a stream other than the NCDC stream. When these close matches are excluded from the comparison the number of reports indexed changes to 108769 with 107972 unique reports and 797 close matches. Table 3 lists the number of unique reports in each stream, excluding the close matches caused by the inclusion of the dew point, and the number of reports missing from each stream.

Overall, the ECMWF data stream contains the greatest number of reports and has the fewest missing, with 105097 of the 108769 reports indexed present. Both the NCDC and JMA streams have ~2000 fewer reports, with 103297 and 103285 reports respectively. When the location of the reports missing from each stream is plotted (Figure 4) strong similarities can be seen between the NCDC and JMA streams. In both cases, the same 10 ships account for over 80 % of the missing reports, with the reports from several of these ships missing completely from the NCDC and JMA streams. The originating centre for the reports from these 10 ships is set to either Offenbach (EDZW) or OSLO (ENMI) in the bulletin headers in the ECMWF stream, however, other reports from these centres are in the JMA data stream. As a result it is unclear why the reports from these 10 ships are missing from the JMA and NCDC streams. The reports missing from the ECMWF data stream are located over the major shipping lanes with no apparent pattern to the missing data. In contrast to the JMA and NCDC streams, the 10 ships with the most reports missing from the ECMWF stream account for only 17 % of the missing data. Figures 5 and 6 show histograms of the reporting hour and day of the missing reports for the different streams respectively. The majority of the reports missing from the ECMWF stream are made on the synoptic hour whilst the majority of the reports missing from the JMA and NCDC data streams are made at non-standard reporting hours.

Table 3: Number of reports and number missing from each data stream

Data Stream	Number of reports indexed	Number missing
NCDC	103297	5472
ECMWF	105097	3672
JMA	103285	5484

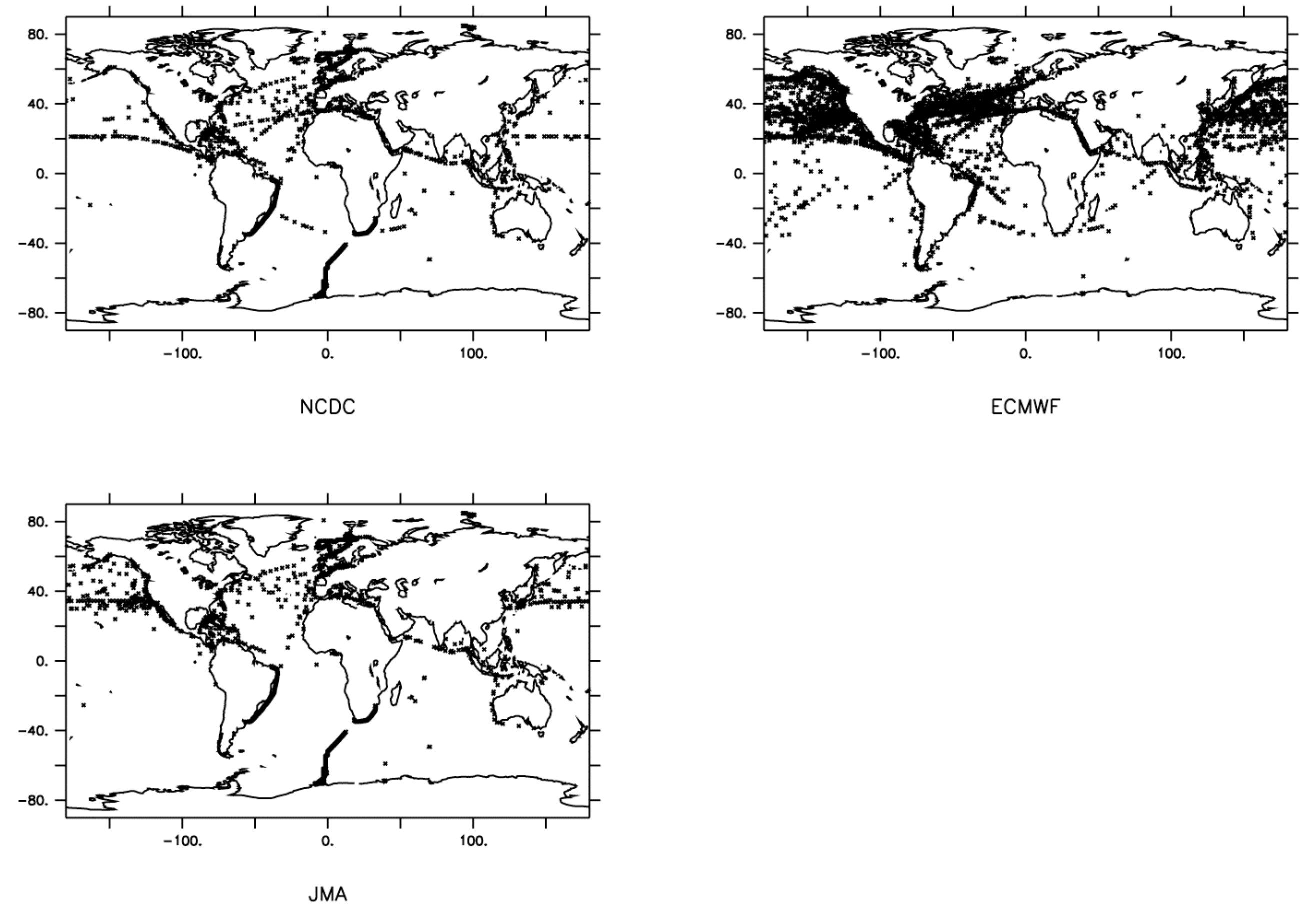

Figure 4: Location of reports missing from each data stream

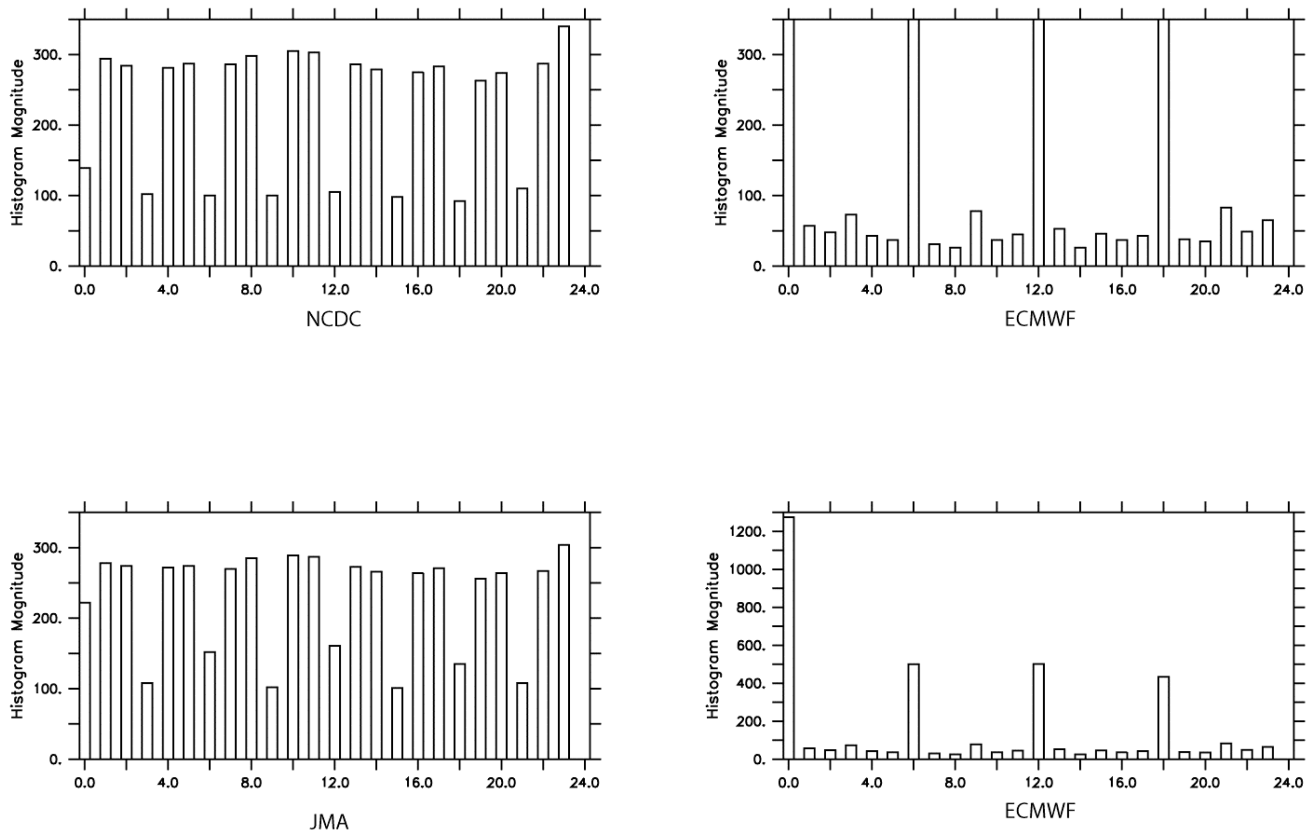

Figure 5: Histograms of reporting hour for reports missing from each stream (data from ECMWF plotted twice on different scales).

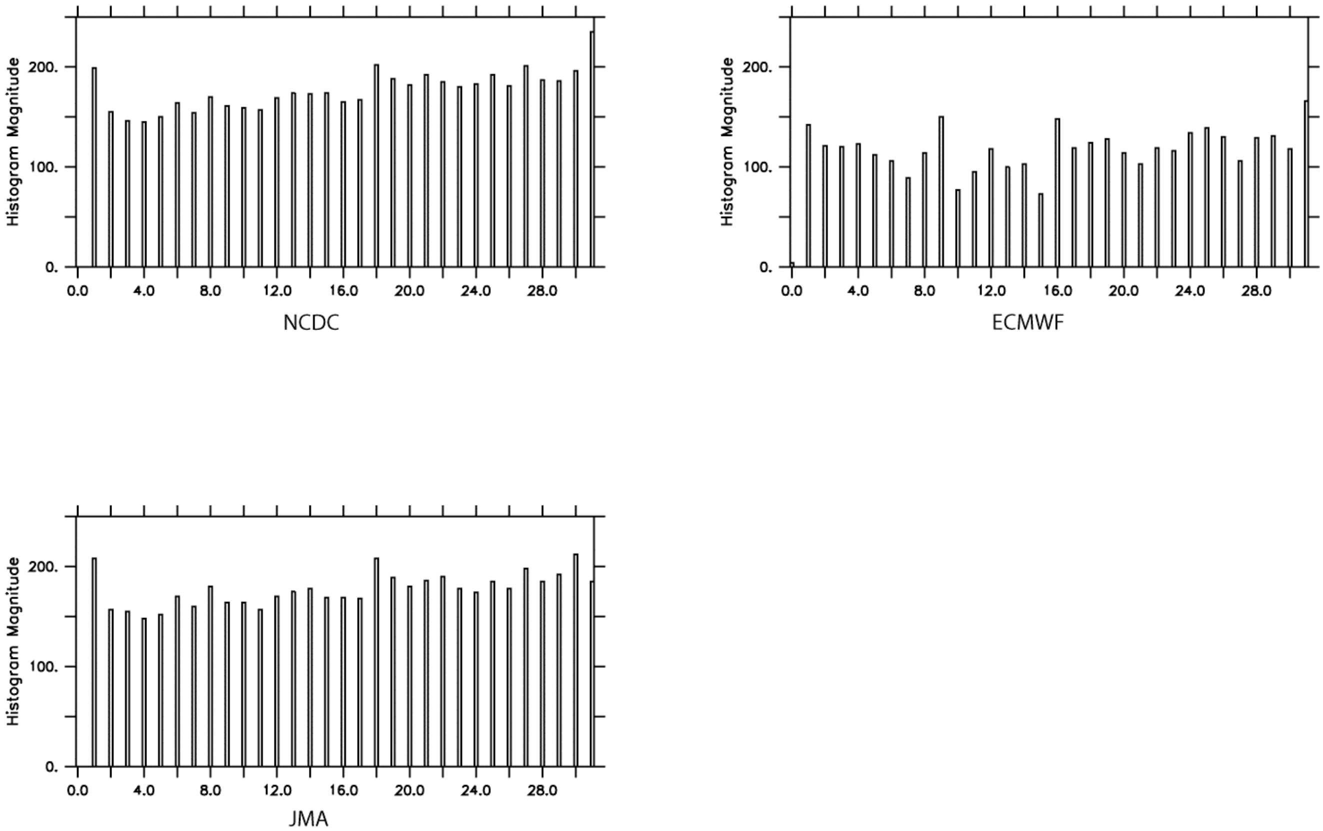

Figure 6: Histograms of day of month for reports missing from each stream

4.2 Analysis of differences between close matches

In total, 3333 close matches were found during the indexing and comparison of the different data streams. As noted above, the majority of these were identified as the dew point being set in one report and not the other. The origin of these reports, i.e. from the originating centre contained in the bulletin header, cannot be identified as the majority occur in the NCDC stream and no headers are present to identify the originating centre. When these reports are excluded the number of close matches decreases to 797, the majority of which are caused by FM13 elements either set to the missing flag or absent in one report and not the other. In the majority of these cases it appears an incomplete report is sent by the VOS followed by a more complete report. If the close matches with only missing elements are excluded the number of close matches decreases further to 177.

The cause of these remaining close matches ranges from an incomplete report being sent by a ship followed by a more complete report, but with slight changes to the other fields, to close matches with substantial differences between reports that cannot be explained. An example of a close match due to an incomplete report being transmitted by a ship follow by a more complete report can be seen in Figure 7. In this example, the wind speed has changed from 14 to 13 knots and the air temperature changed from 17.6 to 17.7°C between the reports being made. Both reports appear in all three data streams.

```
BBXX KRGB 29004 99326 71373 43/// /3114 10176 2//// 40311 5//// 7////
8//// 22255 00184 2//// 3//// 4//// 5//// 6//// ICE /////=

BBXX KRGB 29004 99326 71373 43398 33113 10177 2015/ 40311 56005 7////
8213/ 22255 00184 20804 335// 41205 5//// 6//// 80161 ICE /////=
```

Figure 7: Example of weather reports identified as a close match. Elements that differ have been underlined for clarity.

Figure 8 shows another example where the cause of the differences between the close matches can be identified. In this case, the first two reports form the close match with the second report containing the same time, location and date as the first report but with different data. The report made by the same ship an hour later (i.e. the third one in the figure) also contains the same data as the second report, suggesting that the second report was initially transmitted with time and location left unchanged from the report at 0500h. This was then corrected and retransmitted by the ship.

```
BBXX PBGH 23054 99140 70615 416// 40914 10270 40090 54000 7//// 844//
22243 20301 3//// 4//// 5////=
BBXX PBGH 23054 99140 70615 416// 41325 10270 40085 58005 7//// 844//
22243 20301 3//// 4//// 5////=
BBXX PBGH 23064 99138 70614 416// 41325 10270 40085 58005 7//// 844//
22243 20301 3//// 4//// 5////=
```

Figure 8: Example reports where the time and location information have been incorrectly repeated from the previous report and transmitted followed by the transmission of the correct report.

Whilst the cause of some of the close matches are identifiable and possible to correct the cause of other close matches are less clear and require the delayed mode data to choose the 'correct' observation. Such an example is given in Figure 9 – in this case there are 18 differences between the close matches coupled with a number of missing elements. The differences between the two reports are too large to be explained by two reports being made close together in time and it is unclear which one is correct. In this example, large errors could be introduced into the climate record if the incorrect report is kept and the correct report discarded – highlighting the need for the delayed mode data in order to identify the correct report.

```
BBXX DHDM 18124 99203 10387 41498 13617 10280 20267 40110 57030 7//00
81110 22275 04280 20508 301// 40404 5//// 80270=
BBXX DHDM 18124 99203 10387 41998 03520 10270 20206 40105 56/// 70122
80000 22275 04280 20502 334// 40603 5//// 80225=
```

Figure 9: Example close matches where the cause of the differences cannot be identified

In the examples given above, both the original report and the close match appear in all three data streams. Little additional information is provided by the bulletin headers in these examples other than the originating centre – the BBB indicator is usually either absent or the same for both the original report and the close match. This suggests the differences are originating in the reports when they are transmitted from the VOS. It should be noted a small number (e.g. 453 in the JMA data stream) do have the BBB indicator set to CCx in the bulletin header, indicating a correction to a previously transmitted bulletin. For these cases, the reports in the correction generally do not have a close match to any of the other reports in the GTS streams. Overall, the number of reports that form a close match with another report only form a small subset of the total number of reports, 177 out of ~108000 (or less than 0.2 % of the total number of reports).

4.3 Benefits from merging data streams

Table 4 lists the number of reports that would result from merging one or more data streams (union) and that are shared (intersection) between the different data streams. The greatest number of weather reports from merging two data streams is given by merging the ECMWF and NCDC streams, with 108496 reports in the merged data stream. This is an increase of 5199 compared to the NCDC stream and an increase of 3399 reports compared to the ECMWF data stream. When the ECMWF stream is merged with the JMA stream a similar increase in the number of reports is found, increasing by 5206 compared to the JMA stream and 3394 compared to the ECMWF stream. When the JMA and NCDC streams are merged the merged stream contains 103580 reports, increases of 283 and 295 relative to the NCDC and JMA streams respectively. When all three streams are merged the total number of reports in the merged data stream increases to 108769. In the case of the streams examined in this study the number of reports in the JMA and NCDC streams would increase by ~ 5% when merged with the ECMWF stream. The number of reports in the ECMWF stream would be increased by ~ 3% when merged with one of the other two streams.

Table 4: Number of reports in the union and intersection of the different data streams

Data Stream	Union	Intersection
NCDC or ECMWF or JMA	108769	99881
NCDC or ECMWF	108496	99898
ECMWF or JMA	108491	99891
NCDC or JMA	103580	103002

Summary

The results show significant differences between the GTS streams. These differences are generally in the reports included in each stream, with some reports present in one stream and not another, rather than changes to the data contained in individual reports. Where close matches between the reports exist both the first report and corresponding close match tend to exist in all the streams examined. In some cases it is possible to determine the cause of the close matches but in others delayed mode data may be required to determine which report is correct.

Conclusions

(a) When using GTS data in the climate record data streams from multiple centres should be merged, preferable from different regions. In doing this the number of observations in the record can be increased by 3 – 5 %.

(b) The BUFR template as used by the Met Office introduces differences to reports which make it hard to compare with the original FM13 data.

(c) The merged GTS data stream should be compared to the delayed mode data stream to determine whether any of the unresolved report differences can be understood using quality controlled data. Any "close match" report elements which differ and for which the reason cannot be determined should be flagged as suspect in the climate record.

References

WMO, 2009a: *Manual on Codes*. International Codes Vol. I.1 (Annex II to Technical Regulations), Part A – Alphanumeric Codes. WMO–No.306 (2009 Ed.).

WMO, 2009b: *Manual on Codes*. International Codes, Vol. I.2 (Annex II to Technical Regulations), Part B–Binary Codes and Part C–Common Features to Binary and Alphanumeric Codes. WMO–No.306 (2009 Ed.).

Revised version of document presented at ETMC-III

APPENDIX B

EXAMPLE OF LETTER SENT TO MET SERVICES REQUESTING DATA

World Meteorological Organization
Organisation météorologique mondiale

Temps Climat Eau
Weather Climate Water

Secretariat
7 bis, avenue de la Paix
Case postale 2300
CH 1211 Genève 2
Suisse
Tel.: +41 (0) 22 730 81 11
Fax: +41 (0) 22 730 81 81
E-mail: wmo@wmo.int
Website: www.wmo.int

Our ref.: 11025-08/OBS/WIGOS/OSD/MAR/
ETMC-SHIP

Prof. John Mitchell
Permanent Representative of
the United Kingdom of Great Britain and
Northern Ireland with WMO
Met Office
FitzRoy Road
EX1 3PB EXETER
United Kingdom of Great Britain and
Northern Ireland

GENEVA, 25 April 2008

Dear Professor Mitchell,

As you will recall, the Marine Climatological Summaries Scheme (MCSS) established in 1963 (Res. 35, Cg-IV), has its primary objectives in the international exchange, quality control and archival of delayed-mode marine climatological data, in support of global climate studies and the provision of a range of marine climatological services. Eight WMO Member States and Territories (now represented by Germany; Hong Kong, China; India; Japan; Netherlands; Russian Federation; United Kingdom and United States of America) were designated as Responsible Members (RMs) to gather and process the data, including also, data from other Contributing Members (CMs) worldwide; and to regularly publish Marine Climatological Summaries (MCS) for representative areas, in chart and/or tabular forms. Two Global Data Collecting Centres (GCCs) were established in 1993 in Germany and the United Kingdom to facilitate and enhance the flow and quality control of the data (Res. 11, CMM-XI, and Res. 10, EC-XLV). Eventually, all data are to be archived in the appropriate archives, including the International Comprehensive Ocean-Atmosphere Data Set (ICOADS).

As you will also recall, the Expert Team on Marine Climatology (ETMC) was established by the WMO/IOC Joint Technical Commission for Oceanography and Marine Meteorology (JCOMM) at its First Session (JCOMM-I, Akureyri, Iceland, June 2001) and renewed at its Second Session (JCOMM-II, Halifax, Canada, September 2005), and has engaged in an exercise to propose a restructuring of the MCSS.

At its second Session in Geneva, Switzerland (26-27 March 2007), while discussing Voluntary Observing Ship (VOS) data quality and exchange, the Team noted that the data inserted on GTS and the data received by operational centres did not always match, in part because of decisions made at the operational centres before storing or archiving the data (reports deleted or changed; data excluded). The Team suggested that in order to validate the data collection process of the MCSS, a detailed intercomparison

cc: Mr Robert Keeley, Coordinator of the JCOMM Data Management Programme Area
Mr Scott D. Woodruff, Chairperson of the JCOMM ETMC
Elizabeth Kent, ETMC Member

survey should be conducted between FM 13 SHIP sample data sets received at key operational centres such as the Met Office. The UK National Oceanography Centre, Southampton (NOCS) has volunteered to undertake such a survey provided the required sample data are submitted to it. The survey will help understand how the data reported by ships make their way into national and international archives, and why the data streams are different (e.g. differences due to the way the data are extracted from the GTS, from later QC, or from forecast cut-off times). As climate studies rely more and more on the real-time data for the climate archives, the need to understand the mechanisms of data flow to the archives becomes more important.

I am, therefore, inviting you to consider delivering appropriate sample data to the NOCS. As you know, all available *in situ* observations from the marine environment, including FM-13 SHIP data are regarded as essential with regard to the WMO Resolution 40 (Cg-XII) and ought to be provided on a free and unrestricted basis. Information regarding the data of interest for this intercomparison are provided in the Annex.

I sincerely hope that you will find this proposal favourable, as it will greatly contribute to increase the consistency and quality of Marine Climatological Summaries which in turn are essential for climate studies.

I look forward to your continuous and active participation in marine climatology activities.

Yours sincerely,

(J. Lengoasa)
for the Secretary-General

ANNEX

INFORMATION REGARDING THE SAMPLE OF GTS DATA TO BE SUBMITTED TO THE UK NATIONAL OCEANOGRAPHY CENTRE, SOUTHAMPTON (NOCS) FOR THE EVALUATION SURVEY FOR VALIDATING THE MCSS DATA COLLECTION PROCESS

What data	All data from the GTS in format below for the observations made during the period of December 2007
Period	December 2007
Format	FM-13 SHIP, including the GTS bulletin headers
Media	Data can be submitted to NOCS via ftp or CD-Rom
Contact point	Dr Elizabeth Kent Ocean Observing and Climate (254/25) National Oceanography Centre, Southampton European Way Southampton, SO14 3ZH United Kingdom Tel: +44 (0) 23 8059 6646 Fax: +44 (0) 23 8059 6400 E-mail: eck@noc.soton.ac.uk