BUFR Template Replacement for the FM13 (SHIP) Code: Preliminary Comments and Questions

Scott Woodruff, Chair ETMC, DMPA/JCOMM
17 May 2006 (Scott.D.Woodruff@noaa.gov)

1) Previous ETMC assessment

In 2004, JCOMM's Expert Team on Marine Climatology (ETMC) made a limited review of NCEP's then-current version of BUFR used to store FM13 (and FM18), with findings that not all the features of the originally reported SHIP (and BUOY) data were accurately or completely preserved, partly because of transformations of the data into BUFR units (or into binary data). Findings from this general review, available under Agenda item 3.2 at:

http://icoads.noaa.gov/etmc/etmc1/

were to be provided to CBS by the WMO Secretariat (JCOMM, 2004, pp. 4-5). ETMC offered to continue this work if so requested, but I am not aware of any subsequent feedback from CBS.

2) Where will the SHIP BUFR reports be encoded?

JCOMM (2006) does not appear to clearly define whether the plan is to transmit BUFR from individual ships, or only to circulate BUFR over GTS from major centers (e.g., BUFR prepared centrally from FM13 messages, with the possible addition of extra metadata). This may not be fully decided yet, but it seems to me that it may be helpful to understand where BUFR will likely be encoded (and who will be decoding and using BUFR) because these factors might have a bearing on the template design. Is any clarification available at this stage?

3) Impact on VOS observational instructions from the transition to BUFR?

The documentation and table-driven aspects of BUFR appear to be optimized for machines, rather than humans, in many respects. If BUFR will be transmitted from individual ships, does it follow that national and international marine observing manuals should be rewritten e.g. to refer to the BUFR construct "[0 11 001]" rather than "dd" (wind direction)? Historically, it may be valuable to note that "dd" has been in use at least since 1913 in the International Synoptic Code, and I think there are very strong arguments to retain this continuity with long-term observing practices.

4) BUFR and software complexity

Partly because it attempts to represent many possible data types, BUFR is an extremely complex format, and highly dependent on software. If the plan is to encode and transmit BUFR from many different locations (e.g., NMCs, or from individual ships), several separate encoders (and decoders) may have to be developed and maintained (as BUFR evolves) in parallel (e.g., for SEAS versus TurboWin). In that case, there is very likely a greater possibility of software errors (simply since more computer code exists) or subtle differences (e.g., because of different rounding procedures or the use of computer hardware, when transforming to/from BUFR units).

5) Ensuring archival of the originally reported data

Concerns about format and software complexity relate to what I feel is a key requirement: ensuring access to all the originally reported data (e.g., temperatures in °C, rather than only as transformed to Kelvin in a binary representation). In the event the original data are not archived and the data are incorrectly transformed into non-original BUFR units (or into the binary representation), then the irreplaceable original data might be unrecoverable.

6) NCEP's solution to retaining the originally reported data

Probably with these software reliability and data continuity considerations in mind, NOAA/NWS/NCEP has wisely preserved the complete original reported FM13/18 character strings in their version of BUFR (e.g. see Table 1). This feature has allowed retrospective correction of a number of problems involving fields that were missing or inaccurately represented in BUFR, through access to the original character strings.

Possible solution: Could a similar approach be considered for the new BUFR template? Then, could any new field requirements (e.g., for metadata) be added via new BUFR fields?

7) Plans for independent validation of the template?

If the original message strings could be preserved in a fashion similar to the NCEP approach as part of the new BUFR template, then it is important to note that the accuracy and completeness of the representation of the data in BUFR could be fully validated by reprocessing the original message strings (preferably using an independent procedure), and then mechanically verifying that the results are identical (the approach used for the 2004 review; ref. item 1).

A further suggestion is that that an ongoing validation, e.g. by JCOMM, definitely be planned as part of development and implementation of the new BUFR template. Based on our experience with the 2004 review, CBS (perhaps because of an emphasis on provision of real-time data to NMCs) appears not to have adequately considered long-term climatological consistency and permanent data archival in the design of BUFR.

Table 1. Output from NCEP BUFRLIB routine UFDUMP. This is a dump of the contents of one BUFR report (for ship call sign DBLK) received in December 2005. The original FM13 report character string is attached (portion highlighted) via 14 (i.e., the value of RAWRPT) replications of field RRSTG (BUFR code 058008). This highlighting indicates some fields added by NCEP as part of the GTS processing or from their QC of the data, some of which could also be considered for the template, depending on where the BUFR data will be encoded (e.g., bulletin header information can prove valuable for some QC-related applications). (Note: The raw computer output has been edited to resolve some minor formatting problems.)

```
Wed May 17 13:08:17 MDT 2006

 MESSAGE TYPE NC001001

004001  YEAR            2005.0  YEAR             YEAR
004002  MNTH              12.0  MONTH            MONTH
004003  DAYS              31.0  DAY              DAY
004004  HOUR              21.0  HOUR             HOUR
004005  MINU               0.0  MINUTE           MINUTE
001198  RPID              DBLK  ( 8)CCITT IA5    REPORT IDENTIFIER
        <ID1SQ>            1.0
001203  SHPC8             DBLK  ( 8)CCITT IA5    SHIP CALL SIGN (8 CHARACTERS)
        <ID2SQ>            0.0
        <ID3SQ>            0.0
005002  CLAT            -69.30  DEGREES          LATITUDE (COARSE ACCURACY)
006002  CLON             -5.90  DEGREES          LONGITUDE (COARSE ACCURACY)
007001  SELV               0.0  METERS           HEIGHT OF STATION
033215  CORN               0.0  CODE TABLE       CORRECTED REPORT INDICATOR
        {BID}              1.0
035195  SEQNUM             795  ( 4)CCITT IA5    CHANNEL SEQUENCE NUMBER
035021  BUHD            SIVJ30  ( 6)CCITT IA5    BULLETIN...MONITORED (TTAAii)
035023  BORG              KWBC  ( 4)CCITT IA5    BULLETIN...MONITORED (CCCC)
035022  BULTIM          312100  ( 6)CCITT IA5    BULLETIN...MONITORED (YYGGgg)
035194  BBB             MISSING ( 6)CCITT IA5    BULLETIN...MONITORED (BBB)
        {RCPTIM}           1.0
008202  RCTS               0.0  CODE TABLE       RECEIPT TIME SIGNIFICANCE
004200  RCYR            2005.0  YEAR             YEAR   - TIME OF RECEIPT
004201  RCMO              12.0  MONTH            MONTH  - TIME OF RECEIPT
004202  RCDY              31.0  DAY              DAY    - TIME OF RECEIPT
004203  RCHR              21.0  HOUR             HOUR   - TIME OF RECEIPT
004204  RCMI              12.0  MINUTE           MINUTE - TIME OF RECEIPT
002193  ITSO               0.0  CODE TABLE       IND TYPE OF STN OP Pst/Pre WX
002001  TOST               1.0  CODE TABLE       TYPE OF STATION
013194  INPC               4.0  CODE TABLE       IND INCLUSION/OMISSION OF PREC
020001  HOVI           50000.0  METERS           HORIZONTAL VISIBILITY
002002  TIWM               8.0  FLAG TABLE       TYPE OF INSTR...WIND MEAS
033195  QMWN            MISSING CODE TABLE       SDMEDIT/QUIPS QCmark FOR WIND
011001  WDIR             230.0  DEGREES TRUE     WIND DIRECTION
011002  WSPD               9.0  METERS/SECOND    WIND SPEED
        <WNDSQ2>           0.0
033193  QMAT            MISSING CODE TABLE       SDMEDIT/QUIPS QCmark FOR TEMP
012101  TMDB            272.05  DEGREES KELVIN   TEMPERATURE/DRY BULB TEMP
033194  QMDD            MISSING CODE TABLE       SDMEDIT/QUIPS QCmark FOR MOIST
012103  TMDP            269.85  DEGREES KELVIN   DEW POINT TEMPERATURE
002038  MSST            MISSING CODE TABLE       METHOD OF SEA SURFACE TEMP
033218  QMST            MISSING CODE TABLE       SDMEDIT/QUIPS QCmark FOR SST
```

```
022043  SST1            272.15  DEGREES KELVIN  SEA TEMPERATURE
        <TMPSQ2>           0.0
        <TMPSQ3>           0.0
033207  QMPR           MISSING  CODE TABLE      SDMEDIT/QUIPS QCmark FOR PRESS
010004  PRES           MISSING  PASCALS         PRESSURE
010051  PMSL           98870.0  PASCALS         PRESSURE REDUCED TO MSL
010063  CHPT               7.0  CODE TABLE      CHARACTERISTIC OF PRESS TEND
010061  3HPC            -200.0  PASCALS         3 HOUR PRESSURE CHANGE
010062  24PC           MISSING  PASCALS         24 HOUR PRESSURE CHANGE
013021  TP06           MISSING  KG/METER**2     TOTAL PRECIP PAST 6 HOURS
        <PCPSQ2>           0.0
        <PCPSQ3>           0.0
020010  TOCC               0.0  %               CLOUD COVER (TOTAL)
020201  HBLCS              9.0  CODE TABLE      Ht ABOVE Sfc BASE LOWEST CLOUD
        {CLDSQ1}           3.0
008002  VSSO               7.0  CODE TABLE      VERTICAL SIG (SFC OBSERVATION)
020011  CLAM           MISSING  CODE TABLE      CLOUD AMOUNT
020012  CLTP              30.0  CODE TABLE      CLOUD TYPE
020013  HOCB           MISSING  METERS          HEIGHT OF BASE OF CLOUD
008002  VSSO               8.0  CODE TABLE      VERTICAL SIG (SFC OBSERVATION)
020011  CLAM           MISSING  CODE TABLE      CLOUD AMOUNT
020012  CLTP              20.0  CODE TABLE      CLOUD TYPE
020013  HOCB           MISSING  METERS          HEIGHT OF BASE OF CLOUD
008002  VSSO               9.0  CODE TABLE      VERTICAL SIG (SFC OBSERVATION)
020011  CLAM           MISSING  CODE TABLE      CLOUD AMOUNT
020012  CLTP              10.0  CODE TABLE      CLOUD TYPE
020013  HOCB           MISSING  METERS          HEIGHT OF BASE OF CLOUD
        <PPWSQ1>           1.0
020003  PRWE               2.0  CODE TABLE      PRESENT WEATHER
020004  PSW1               0.0  CODE TABLE      PAST WEATHER (1)
020005  PSW2               0.0  CODE TABLE      PAST WEATHER (2)
        <WAVSQ1>           0.0
        <WAVSQ2>           1.0
022012  POWW               2.0  SECONDS         PERIOD OF WIND WAVES
022022  HOWW               0.5  METERS          HEIGHT OF WIND WAVES
        {WAVSQ3}           1.0
022003  DOSW              80.0  DEGREES TRUE    DIRECTION OF SWELL WAVES
022013  POSW               7.0  SECONDS         PERIOD OF SWELL WAVES
022023  HOSW               1.5  METERS          HEIGHT OF SWELL WAVES
        <MPLSQ1>           1.0
001193  TDMP               4.0  CODE TABLE      TRUE DIR SHIP PAST 3 Hr
001200  ASMP               2.0  CODE TABLE      AVG SPD SHIP PAST 3 Hr
        <ICESQ1>           0.0
        <RPSEC3>           0.0
        {RAWRPT}          14.0
058008  RRSTG         DBLK 312  ( 8)CCITT IA5   RAW REPORT STRING
058008  RRSTG         11 99693  ( 8)CCITT IA5   RAW REPORT STRING
058008  RRSTG          50059 4  ( 8)CCITT IA5   RAW REPORT STRING
058008  RRSTG         1999 023  ( 8)CCITT IA5   RAW REPORT STRING
058008  RRSTG         09 11011  ( 8)CCITT IA5   RAW REPORT STRING
058008  RRSTG          21033 4  ( 8)CCITT IA5   RAW REPORT STRING
058008  RRSTG         9887 570  ( 8)CCITT IA5   RAW REPORT STRING
058008  RRSTG         20 70200  ( 8)CCITT IA5   RAW REPORT STRING
058008  RRSTG          80000 2  ( 8)CCITT IA5   RAW REPORT STRING
058008  RRSTG         2242 070  ( 8)CCITT IA5   RAW REPORT STRING
058008  RRSTG         10 20201  ( 8)CCITT IA5   RAW REPORT STRING
058008  RRSTG          308// 4  ( 8)CCITT IA5   RAW REPORT STRING
058008  RRSTG         0703 ICE  ( 8)CCITT IA5   RAW REPORT STRING
058008  RRSTG            64740  ( 8)CCITT IA5   RAW REPORT STRING

 >>> END OF SUBSET <<<
```

References

JCOMM, 2004: Expert Team on Marine Climatology, First Session, Gdynia, Poland, 7-10 July 2004, Final Report. JCOMM Meeting Report No. 32, 55 pp.

JCOMM, 2006: JCOMM/OCG Workshop to Establish a Pilot Project for the Collection of Real-time Metadata regarding Sea Surface Temperature and Water Temperature Profile Data, Reading, UK, 28-29 March 2006, Draft Final Report (received 24 April 2006).