CHAPTER 2

BUFR TABLES

2.1 Introduction. BUFR employs 3 types of tables: BUFR tables, code tables and flag tables. It also provides for the use of local tables.

The tables in BUFR that contain information to describe, classify and define the contents of a BUFR message are called BUFR tables. There are 4 BUFR tables defined: Tables A, B, C and D. Code tables and flag tables are defined and discussed in paragraph 2.7. Local tables are defined and discussed in paragraph 2.8.

2.2 Table A - Data Category. Table A is referred to in Section 1 and provides a quick check for the type of data represented in the message. Of the 256 possible entries for Table A, 17 are currently defined:

Table 2-1. BUFR Table A - data category

Code Figure	Meaning
0	Surface data - land
1	Surface data - sea
2	Vertical soundings (other than satellite)
3	Vertical soundings (satellite)
4	Single level upper-air data (other than satellite)
5	Single level upper-air data (satellite)
6	Radar data
7	Synoptic data
8	Physical/chemical constituents
9	Dispersal and transport
10	Radiological data
11	BUFR tables, complete replacement or update
12	Surface data (satellite)
13-19	Reserved
20	Status information
21	Radiances
22-30	Reserved
31	Oceanographic data
32-100	Reserved
101	Image data
102-255	Reserved

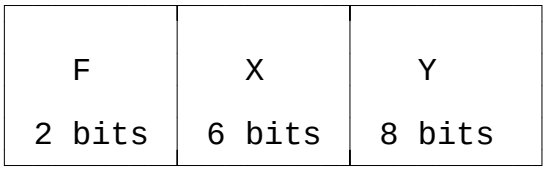
Placing one of the code figures from Table A (Table 2-1) in octet 9 of Section 1 is actually redundant. The descriptors used in

Section 3 of a message fully define the data in Section 4, regardless of the Table A code figure in Section 1. Decoding programs may, however, wish to reference Table A in order to obtain a general classification of the data available prior to actually decoding the information and/or passing it on to some subsequent application program.

2.3 Table B - Classification of Elements. Table B is referenced in Section 3 of a BUFR message and contains descriptions of parameters encoded in Section 4. Table B, as set forth in the WMO Manual On Codes, Volume 1, Part B (reference 4), is reproduced in its entirety in Appendix B. Table B entries consist of 6 entities:

- a descriptor consisting of the 3 parts F X and Y
- element name
- units: basic (SI) units for the element
- scale: factor (equal to 10 to the power [scale]) by which the element has been multiplied prior to encoding
- reference value: a number to be subtracted from the element, after scaling, (if any), and prior to encoding
- data width, in bits, the element requires for representation in Section 4

A Table B descriptor consists of 16 bits (2 octets) divided into 3 parts, F, X and Y.



F (2 bits) indicates the type of descriptor. In 2 bits there are 4 possibilities, 0, 1, 2 and 3. The numeric value of the 2 bit quantity F, indicates the type of descriptor.

F = 0 Element descriptor (Table B entry)
F = 1 Replication operator
F = 2 Operator descriptor (Table C entry)
F = 3 Sequence descriptor (Table D entry)

X (6 bits) indicates the class or category of descriptor. There are 64 possibilities, classes 00 to 63. Thus far, 28 classes have been defined.

Y (8 bits) indicates the entry within an X class. The 8 bits will yield 256 possibilities within each of the 64 classes. There are a varying number of entries within each of the 28 classes that are currently defined.

It is the F X Y descriptors in Section 3 that refer to data represented in Section 4. The 16 bits of F, X and Y are not to be treated as a 16 bit numeric value, but rather as 16 bits divided into 3 parts, where each part (F, X and Y) are in themselves 2, 6 and 8 bit numeric values. Some examples of descriptors with their corresponding bit settings:

Descriptor	F	X	Y	
0 01 001	00	000001	00000001	(Figure 1-4)
1 02 006	01	000010	00000110	
2 01 131	10	000001	10000011	
3 07 002	11	000111	00000010	

If the following descriptors were contained in Section 3,

0 01 001 0 01 002 0 02 001 0 04 001 0 04 002
0 04 003 0 04 004 0 04 005 0 05 002 0 06 002,

they would refer to the following extracts from BUFR Table B.

Table Reference F X Y	Element Name	Units	Scale	Reference Value	Data Width (Bits)
0 01 001	WMO block number	numeric	0	0	7
0 01 002	WMO station number	numeric	0	0	10
0 02 001	Type of station	code table	0	0	2
0 04 001	Year	Year	0	0	12
0 04 002	Month	Month	0	0	4
0 04 003	Day	Day	0	0	6
0 04 004	Hour	Hour	0	0	5
0 04 005	Minute	Minute	0	0	6
0 05 002	Latitude (coarse accuracy)	Degree	2	-9000	15
0 06 002	Longitude (coarse accuracy)	Degree	2	-18000	16

The element name is a plain language description of the element entry of the table.

The units of Table B entries refer to the format of how the data in Section 4 is represented. The data may be numeric as in the case of a WMO block number or character data as in the case of an aircraft identifier. When data is in character form, the character representation is always according to the CCITT International Alphabet No. 5. The units may also refer to a code or flag table, where the code or flag table is described in the WMO Manual On

Codes using as the code or flag table number the same number as the F X Y descriptor. For example, from page I-Bi--123 of the manual:

0 02 001
Type of station

Code Figure	
1	Automatic station
2	Manned station
3	Reserved
4	Missing value

Other units are in Standard International (SI) units, such as meters or degrees Kelvin.

The scale refers to the power of 10 by which the element in Section 4 has been multiplied in order to retain the desired precision in the transmitted data. For example, the units of latitude are whole degrees in Table B. This is not precise enough, however, for most uses. Therefore, the elements are to be multiplied by 100 (10^2) so that the transmitted precision will be centidegrees, a more useful precision. On the other hand, the (SI) unit of pressure in Table B is Pascals, a rather small unit that would result in unnecessarily precise numbers being transmitted. The BUFR Table B calls for pressure to be divided by 10 (10^-1) resulting in a transmitted unit of 10ths of hPa, or tenths of millibars, a more reasonable precision for meteorological usage. These precisions can be changed on the fly, so to speak, if the table values are not appropriate in special cases. This is done through the use of "data description operators" - see paragraph 2.4 below.

The reference value is a value that is to be subtracted from the data after multiplication by the scale factor, if any, before encoding into Section 4 in order to produce, in all cases, a positive value. In the case of latitude and longitude, south latitude and west longitude are negative before applying the reference value. If, for example, a position of 35.50 degrees south latitude were being encoded, multiplying -35.50 by 100 (scale of 2) would produce -3550. Subtracting the reference value -9000 would give 5450 that would be encoded in Section 4. To obtain the original value in decoding Section 4, adding back the -9000 reference value to 5450 would result in -3550, then dividing by the scale (100) would obtain -35.50.

The data width of Table B entries is a count of how many bits the largest possible value of an individual data item of Section 4 occupies.

In those instances where a Table B descriptor defines an element of data in Section 4, where that element is missing for a given

subset, then all bits for that element will be set to 1's in Section 4.

Obviously, without an up-to-date Table B, a decoder program would not be able to determine the form or content of data appearing in Section 4.

2.3.1 Data Replication. A special descriptor called the replication operator (F = 1) is used to define a range of subsequent descriptors, together with a replication factor. This enables the appropriate descriptors to be considered to be repeated a number of times. In general for data replication, X indicates the number of immediately following descriptors that are to be replicated as a repeated set, and Y indicates the total number of replications. This, of course, implies, that the same pattern will be found in Section 4, the data section. This ability to describe a repeated pattern in the data by a single set of descriptors contributes to the efficiency of BUFR.

As an example, consider that the following sequence appears in Section 3:

1 02 006 0 07 004 0 01 003

The meaning of 1 02 006 is that the next 2 descriptors are repeated 6 times, or the equivalent set of descriptors:

0 07 004 0 01 003 0 07 004 0 01 003 0 07 004 0 01 003
0 07 004 0 01 003 0 07 004 0 01 003 0 07 004 0 01 003

A special form of the replication operator allows the replication factor to be stored with the data in Section 4, rather than with the descriptor in Section 3. This special form is called delayed replication. It is indicated by Y = 0. It allows the data to be described in a general way, with the number of replications being different from subset to subset. Since the data now contains an additional data element, the actual replication count, a descriptor must be added to Section 3 to account for, and describe, this (special) data element. The appropriate descriptor is found in Class 31. Special note: the 0 31 YYY (delayed replication factor) descriptor follows immediately after the 1 X 000 (delayed replication) descriptor but is NOT included in the count (X) of the following descriptors to be replicated.

Another form of delayed replication enables both the data description and the corresponding data item or items to be repeated. Entries in Class 31 of Table B (see Appendix B) are used in association with the delayed replication operator to enable this to be done.

2.4 Table C - Data Description Operators. Table C (F = 2) data description operators (Chapter 5) are used when there is a need to redefine Table B attributes temporarily, such as to change data width, scale or reference value of a Table B entry. Table C is also used to add associated fields such as quality control information, indicate characters as data items, and signify data width of local descriptors.

2.5 Table D - Lists of Common Sequences. Table D (F = 3) contains descriptors which describe additional descriptors. A single descriptor used in Section 3 with F = 3 is a pointer to a Table D entry which contains other descriptors. If the Table D descriptor 3 01 001 were used in Section 3, the expansion of that descriptor is two Table B descriptors, 0 01 001 and 0 01 002.

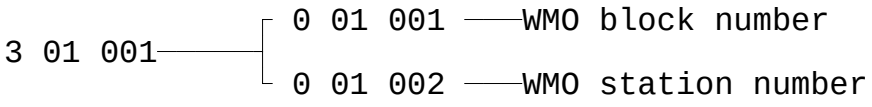
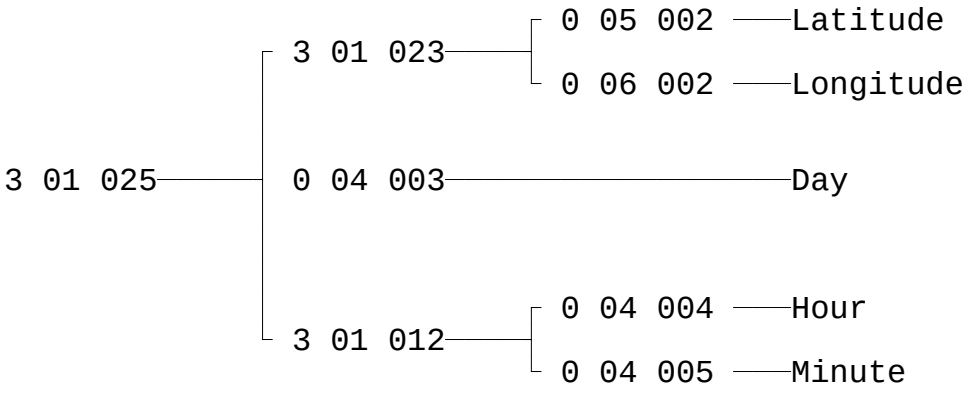


Table D descriptors may also refer to an expansion list of descriptors that contain additional Table D descriptors. The descriptor 3 01 025 expands to 3 01 023, 0 04 003 and 3 01 012. In the expansion, 3 01 023 additionally expands to 0 05 002 and 0 06 002. The remaining descriptor 3 01 012 expands to 0 04 004 and 0 04 005. Thus, the single Table D descriptor 3 01 025 expands to a total of 5 separate Table B entries.



The order of the data in Section 4 is then according to the following sequence of Table B entries: 0 05 002, 0 06 002, 0 04 003, 0 04 004, 0 04 005.

There are currently 19 categories of common sequences defined in BUFR Table D (Table 2-2 below).

Table 2-2. BUFR Table D list of common sequences

F	X	Category of Sequences
3	00	BUFR table entries sequences
3	01	Location and identification sequences
3	02	Meteorological sequences common to surface data
3	03	Meteorological sequences common to vertical sounding data
3	04	Meteorological sequences common to satellite observations
3	05	Reserved
3	06	Meteorological or oceanographic sequences common to oceanographic observations
3	07	Surface report sequences (land)
3	08	Surface report sequences (sea)
3	09	Vertical sounding sequences (conventional data)
3	10	Vertical sounding sequences (satellite data)
3	11	Single level report sequences (conventional data)
3	12	Single level report sequences (satellite data)
3	13	Sequences common to image data
3	14	Reserved
3	15	Oceanographic report sequences
3	16	Synoptic feature sequences
3	18	Radiological report sequences
3	21	Radar report sequences

Any BUFR message may be encoded without using Table D. The data description contained within Section 3 can be accomplished entirely by using only element descriptors of Table B and operator descriptors of Table C. To do so, however would involve considerable overhead in terms of the length of the Section 3 data description. The use of Table D is another major contributor to the efficiency of BUFR.

2.6 Message Layout. Figure 2-1 illustrates how the single descriptor 3 07 002 expands into 2 more Table D descriptors, 3 01 032 and 3 02 011. The descriptor 3 01 032 further expands into 5 more descriptors 3 01 001, 0 02 001, 3 01 011, 3 01 012 and 3 01 024. As is shown in Figure 2-1, descriptors in Table D may themselves refer to Table D, provided no circularity results on repeated expansion. Completion of the expansion process leads to a total of 31 Table B descriptors. The 16 bits in Section 3 taken by the descriptor 3 07 002 results in a savings of 480 bits (30 x 16 bits) over what the 31 Table B descriptors would occupy in bits.

Table D has been limited to lists of descriptors likely to be most frequently used. Table D was not designed to be comprehensive of all sequences likely to be encountered. To do so would require an excessively large Table D and would considerably reduce flexibility when encoding minor differences in reporting practices. More

flexibility is retained if the Data Description Section contains several descriptors.

A complete layout of a BUFR message containing just 1 surface observation is illustrated in Figure 2-2. As indicated in octets 5-7 of Section 1, there are a total of 78 octets in the message, or 624 bits. Of the 624 bits, 267 are for the actual parameters of data (Figure 2-1) and the remaining 357 bits are BUFR overhead. BUFR overhead in this context is the number of bits that are not actual surface data. In this example there are more bits used for the overhead than for the surface data.

Figure 2-3 is a complete layout of a BUFR message containing the maximum number of 448 subsets to fit within the 15000 octet limit. This message would contain 14996 octets or 119968 bits. Of these 119968 bits, 119616 are data and 352 bits are BUFR overhead. The 5 bit difference in overhead from Figure 2-2 (357 bits) and Figure 2-3 (352 bits) is due to the number of bits set to 0 at the end of Section 4 in order to complete the section at the end of an even numbered octet. For 1 subset of 267 bits, 5 additional bits are needed to complete the octet. For 448 subsets, or 119616 bits, no additional bits are needed to complete the last octet.

2.6.1 Comparison of BUFR and Character Code Bit Counts. The surface observations illustrated in Figures 2-1 to 2-3 are the equivalent of the following parameters in the WMO code form FM 12-IX Ext. SYNOP:

$YYGGi_w$ IIiii $i_R i_X$hVV Nddff $1s_nTTT$ $2s_nT_dT_dT_d$ $3P_oP_oP_oP_o$
$4PPPP$ 5appp $7wwW_1W_2$ $8N_hC_LC_MC_H$

Data encoded in this form would consist of 55 characters plus 10 spaces between each group of 5 characters for a total of 65 characters. For transmission purposes these 65 characters would require a total number of 520 bits (65 X 8 bits per character). A complete BUFR message with 1 observation (Figure 2-2) requires 78 octets or 624 bits, 104 more than the corresponding character representation. Of these 624 bits, 267 are taken by the surface observation and 357 as BUFR overhead. If, however, 448 observations in character form were transmitted, the total number of bits would be 232960 (520 X 448). The corresponding BUFR representation (Figure 2-3) would require 14996 octets, or 119968 bits, a savings of 112992 bits over the character representation. The 112992 bits is equivalent to 217 observations in character form or 423 observations in BUFR, not counting the BUFR overhead. While these numbers may be viewed in different ways, the real significance is that BUFR is far more efficient, in terms of number of bits to represent a set of meteorological observations, than character forms.

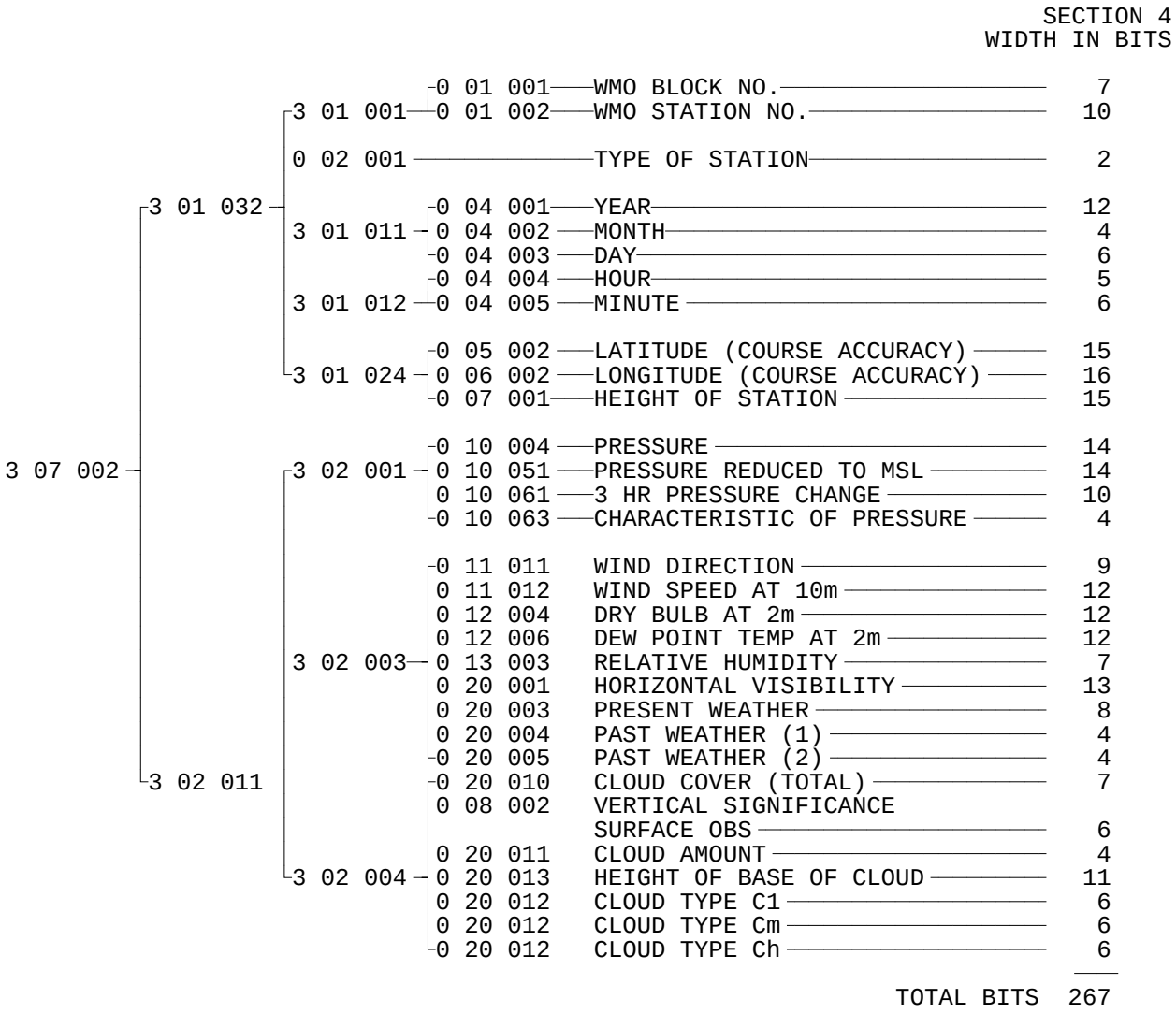


Figure 2-1. Example of surface observations sequence using Table D descriptor 3 07 002

	Section Octet No.	Octet in Message	Encoded Value	Description
Section 0 (indicator section)	1-4	1-4	BUFR	encoded international CCITT Alphabet No. 5
	5-7	5-7	78	total length of message (octets)
	8	8	2	BUFR edition number
Section 1 (identification section)	1-3	9-11	18	length of section (octets)
	4	12	0	BUFR master table
	5-6	13-14	58	originating center (U.S. Navy - FNOC)
	7	15	0	update sequence number
	8	16	0	indicator that Section 2 not included
	9	17	0	Table A - surface land data
	10	18	0	BUFR message sub-type
	11	19	2	version number of master tables
	12	20	0	version number of local tables
	13	21	92	year of century
	14	22	4	month
	15	23	18	day
	16	24	0	hour
	17	25	0	minute
	18	26	0	reserved for local use by ADP centers (also needed to complete even number of octets for section
Section 3 (Data description section)	1-3	27-29	10	length of section (octets)
	4	30	0	reserved
	5-6	31-32	1	number of data subsets
	7	33	bit 1=1	flag indicating observed data
	8-9	34-35	3 07 002	Table D descriptor for surface land in F X Y format
	10	36	0	need to complete section with an even number of octets
Section 4 (Data section)	1-3	37-39	38	length of section (octets)
	4	40	0	reserved
	5-38	41-74	data	continuous bit stream of data for 1 observations, 267 bits plus 5 bits to end on even octet (see Figure 2-1 for expansion)
Section 5 (End section)	1-4	75-78	7777	encoded CCITT International Alphabet No. 5

Figure 2-2. BUFR message of 1 surface observation using Table D descriptor 3 07 002

	Section Octet No.	Octet in Message	Encoded Value	Description
Section 0 (indicator section)	1-4	1-4	BUFR	encoded international CCITT Alphabet No. 5
	5-7	5-7	14996	total length of message (octets)
	8	8	2	BUFR edition number
Section 1 (identification section)	1-3	9-11	18	length of section (octets)
	4	12	0	BUFR master table
	5-6	13-14	58	originating center (U.S. Navy - FNOC)
	7	15	0	update sequence number
	8	16	0	indicator that Section 2 not included
	9	17	0	Table A - surface land data
	10	18	0	BUFR message sub-type
	11	19	2	version number of master table
	12	20	0	version number of local tables
	13	21	92	year of century
	14	22	4	month
	15	23	18	day
	16	24	0	hour
	17	25	0	minute
	18	26	0	reserved for local use by ADP centers (also needed to complete even number of octets for section
Section 3 (Data description section)	1-3	27-29	10	length of section (octets)
	4	30	0	reserved
	5-6	31-32	448	number of data subsets
	7	33	bit 1=1	flag indicating observed data
	8-9	34-35	3 07 002	Table D descriptor for surface land in F X Y format
	10	36	0	need to complete section with an even number of octets
Section 4 (Data section)	1-3	37-39	14956	length of section (octets)
	4	40	0	reserved
	5-14956	41-14992	data	continuous bit stream of data for 448 observations, 267 bits per observation with no added bits to end on an even octet
Section 5 (End section)	1-4	14993-14996	7777	encoded CCITT International Alphabet No. 5

Figure 2-3. BUFR message of 448 surface observations using Table D descriptor 3 07 002

2.7 Code Tables and Flag Tables. Since some meteorological parameters are qualitative or semi-qualitative, they are best represented with reference to a code table.

2.7.1 Code Tables. BUFR code tables and flag tables refer to elements defined within BUFR Table B. They are numbered according to the X and Y values of the corresponding Table B reference. For example, the Table B entry 0 01 003, WMO Region number, geographical area, indicates in the Unit column that this is a BUFR code table, the number of that code table being 0 01 003.

Many of the code tables included in the BUFR specification are similar to existing WMO code tables for representing character data. Attachment II of the WMO Manual on Codes, Volume 1, Part B is a list of the code tables associated with BUFR Table B and the existing specifications and code tables of the WMO Manual on Codes, Volume 1, Part A.

There is no one-to-one BUFR code table relationship to the character code tables. The character Code Table 3333, Quadrant of the Globe, for example, has no meaning in BUFR, as all points on the globe in BUFR are completely expressed as latitude and longitude values.

2.7.2 Flag tables. In a flag table, each bit indicates an item of significance. A bit set to 1 indicates an item is included, or is true, while a bit set to 0 indicates omission, or false. In any flag table, when all bits are set it is an indication of a missing value. Flag tables additionally enable combinations to be identified. In all flag tables within the BUFR specification, bits are numbered from 1 to N from most significant to least significant within a data width of N bits, i.e., from left (bit 1) to right (bit N). For example, from page I-Bi--135 of the manual:

0 08 001
Vertical sounding significance

Bit No.	
1	Surface
2	Standard level
3	Tropopause level
4	Maximum wind level
5	Significant level, temperature
6	Significant level, wind
All 7	Missing value

2.7.3 Flags. Flags, without reference to a flag table, are also used within Sections 1 and 3 of a BUFR message. In Section 1, octet 8, if bit 1 = 0 this is an indication that the optional Section 2 is not contained within the message. If bit 1 = 1, then Section 2 is included.

```
Section 1                          Section 1
 Octet 8                            Octet 8

00000000                           10000000
|
└ Section 2 not included           └ Section 2 included
```

Similarly, the two flag bits in Section 3, octet 7 have these meanings:

```
Section 3                          Section 3
 Octet 7                            Octet 7
00000000                           11000000
||
|└ non-compressed data              |└ compressed data
|
└ other data                       └ observed data
```

2.8 Local Tables. Since a data processing center may need to represent data conforming to a local requirement and such local data may not be defined within Table B, specific areas of Table B and D are reserved for local use (Figure 2-4). These areas are defined as descriptor part Y entries 192 to 255 inclusive of all descriptor part X classes. Also, centers defining descriptor part X classes for local use should restrict their use to the range 48 to 63 inclusive.

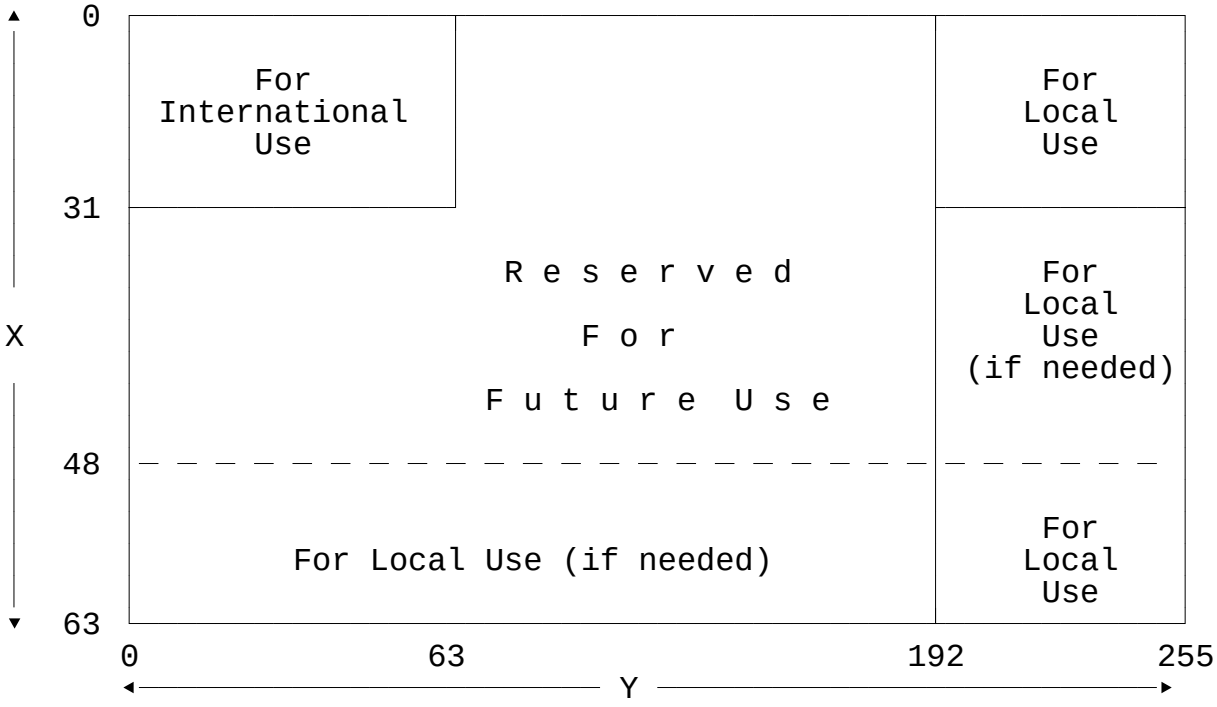


Figure 2-4. Table reservations

If a data processing center had multiple sources of data receipt, for example, it may be necessary to indicate the source of an observation by encoding the circuit from which the data was received. A local Table B descriptor such as 0 54 192 could be used which may be a code table specifying circuits of transmission. The Table B entry could be:

Table Reference	Element Name	Units	Scale	Reference Value	Data Width (Bits)
0 54 192	Circuit	code table	0	0	3

The corresponding local code table could be:

0 54 192
Circuit designators for data receipt

Code Figure	Circuit
0	GTS
1	AWN
2	AUTODIN
3	ANTARCTIC
4-7	Reserved

Using the same Table D descriptor, 3 07 002, as in Figure 2-1, adding the local descriptor 0 54 192 would produce the expansion as in Figure 2-5. The following modifications would have to be made to the BUFR message if the local descriptor 0 54 192 were to be included in a message (Figure 2-6):

Section 0, octets 5-7, the total length of the message, increases from 14996 octets to 14998 octets.

Section 1, octet no. 12 (octet 20 within the message) would have the version number of the local tables in use.

Section 3, octets 1-3, the encoded value would increase from 10 octets to 12 octets. If one descriptor were being added, the length of the section increases by 2 in order to keep the section an even number of octets. Octets 5-6, number of data subsets decreases from 448 to 443. The number of data subsets have been reduced to keep the total message length under the 15000 octet maximum.

Also in Section 3, the descriptors will occupy octets 8-11 vice octets 8-9 to accommodate the added descriptor.

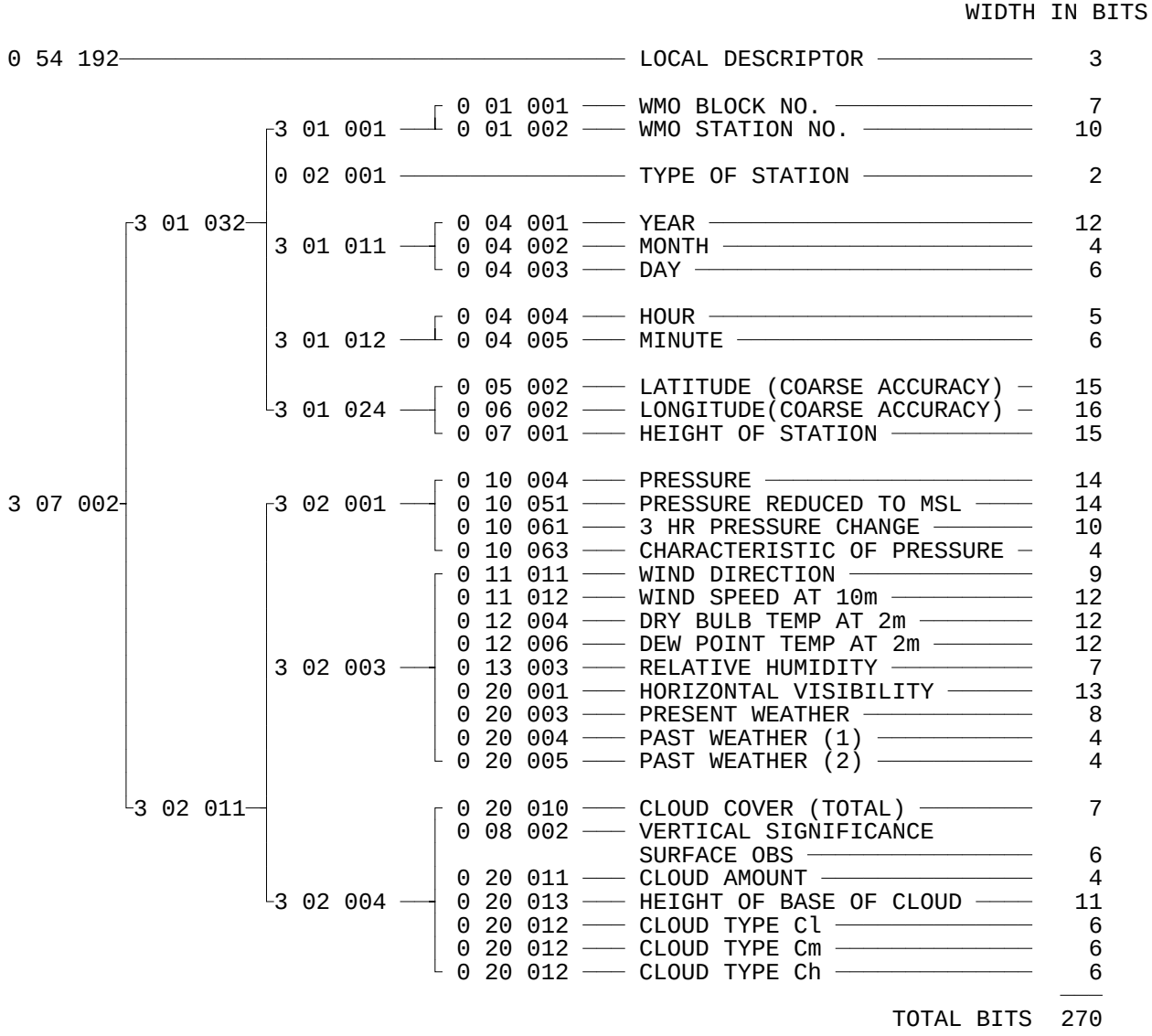


Figure 2-5. Example of surface observations sequence using Table D descriptor 3 07 002 and a local descriptor.

	Section Octet No.	Octet in Message	Encoded Value	Description
Section 0 (indicator section)	1-4	1-4	BUFR	encoded international CCITT Alphabet No. 5
	5-7	5-7	14998	total length of message (octets)
	8	8	2	BUFR edition number
Section 1 (identification section)	1-3	9-11	18	length of section (octets)
	4	12	0	BUFR master table
	5-6	13-14	58	originating center (U.S. Navy - FNOC)
	7	15	0	update sequence number
	8	16	0	indicator that Section 2 not included
	9	17	0	Table A - surface land data
	10	18	0	BUFR message sub-type
	11	19	2	version number of master tables
	12	20	1	version number of local tables
	13	21	92	year of century
	14	22	4	month
	15	23	18	day
	16	24	0	hour
	17	25	0	minute
	18	26	0	reserved for local use by ADP centers (also need to complete even number of octets
Section 3 (Data description section)	1-3	27-29	12	length of section (octets)
	4	30	0	reserved
	5-6	31-32	443	number of data subsets
	7	33	BIT 1=1	flag indicating observed data
	8-11	34-37	0 54 192 3 07 002	local and Table D descriptors in F X Y format
	10	38	0	need to complete section with an even number of octets
Section 4 (Data section)	1-3	39-41	14956	length of section (octets)
	4	42	0	reserved
	5-14956	43-14994	data	continuous bit stream of data for 443 observations, 270 bits per observation plus 6 bits to end on even octet
Section 5 (End section)	1-4	14995-14998	7777	encoded CCITT international Alphabet No. 5

Figure 2-6. BUFR message of 443 surface observations using 2 descriptors, local descriptor 0 54 192 and Table B descriptor 3 07 002.

Note that in Section 4, octets 1-3, the encoded value for length of section remains the same at 14956 octets. The number of bits needed for 448 subsets without a local descriptor is 119616 (448 X 267), or exactly 14952 octets. For 443 subsets with 3 bits added to each subset for the local information, 119610 bits are needed (443 X 270). Adding 6 bits to complete the octet brings the total bit count for all 443 subsets to 119616, the same number of bits as 448 subsets without the added local information.