CHAPTER 5

TABLE C DATA DESCRIPTION OPERATORS

5.1 Introduction. Table C data description operators (Table 5-1) are used when there is a need to redefine Table B attributes temporarily, such as the need to change the data width, scale or reference value of a Table B entry.

5.2 Changing Data Width, Scale, and Reference Value. If data from a DRIFTER observation (FM 18-IX Ext., Report of a drifting-buoy observation) were being encoded into BUFR, there are no Table B entries to correspond to latitude and longitude in thousandths of degrees. The Table B entries for latitude and longitude are high accuracy (hundred thousandths of a degree) and coarse accuracy (hundredths of a degree). There are several possible methods to handle the encoding of latitude and longitude for DRIFTER in thousandths of degrees. One method would be to choose the high accuracy Table B entries for latitude and longitude in hundred thousandths of degrees. There would be no loss of accuracy, but a lot of unused bits for each observation would be encoded in Section 4. The high accuracy latitude requires 25 bits for representation, high accuracy longitude 26 bits. To represent latitude and longitude to thousandths of degrees would require 18 and 19 bits respectively. If the extra bits from using high accuracy were not deemed a concern, this would be the easiest method. If, however, it were desirable to use only the bits required to represent latitude and longitude in thousandths of degrees, there are two ways for this to be accomplished. First, and the least desirable of any method, would be to create local descriptors for Table B with the appropriate scale and reference values for thousandths of degrees. This is the least desirable method because if the BUFR message were to be transmitted to another center, the receiving center would have to have available to their BUFR decoder program the correct definition of the local descriptors. The other method would be to use the Table C data description operators 2 01 Y to change the data width of the Table B descriptor for latitude and longitude, 2 02 Y to change the scale and 2 03 Y to change the reference values.

There is now a choice to be made between temporarily changing latitude and longitude from hundredths of degrees to thousandths, or, from changing them from hundred thousandths to thousandths. It doesn’t matter which is done, as the only difference between the choices will be the Y operand entries of the data description operators.

Table 5-1. BUFR Table C - data description operators

Table Reference F	Table Reference X	Operand	Operator Name	Operation Definition
2	01	Y	Change data width	Add (Y-128) bits to the data width for each data element in Table B, other than CCITT IA5 (character) data, code or flag tables
2	02	Y	Change scale	Multiply scale given for each non-code data elements in Table B by 10^(Y-128)
2	03	Y	Change reference values	Subsequent element descriptors define new reference values for corresponding Table B entries. Each new reference value is represented by Y bits in the Data Section. Definition of new reference values is concluded by encoding this operator with Y=255. Negative reference values shall be represented by a positive integer with the left-most bit (bit 1) set to 1
2	04	Y	Add associated field	Precede each data element with Y bits of information This operation associates a data field (e.g., quality control information) of Y bits with each data element.
2	05	Y	Signify character	Y characters (CCITT international Alphabet No. 5) are inserted as a data field of Y x 8 bits in length
2	06	Y	Signify data width for the immediately following local descriptor	Y bits of data are described by the immediately following descriptor

If it were decided to change the data width of latitude and longitude from hundredths to thousandths of degrees, it must first be determined how many bits are necessary to represent individually latitude and longitude in thousandths of a degree. The maximum value for latitude to be represented in the data in Section 4 must take into consideration the old reference value of -9000. The new reference value will be -90000 to accommodate thousandths of degrees. The maximum value of a reported latitude to be encoded into BUFR bits is 180000. This value is arrived at by a reported latitude of 90.000 North which must then be scaled to 10^3 (also to be changed from 10^2) to retain the desired precision, then subtracting the reference value of -90000, producing 180000. The number of bits to accommodate 180000_{10} is 18. To change the data width of the Table B entry for latitude (coarse accuracy) from 15 bits to 18 bits would require the Table C entry 2 01 131. The Y operand 131 is determined by the Operation Definition of adding Y-128 bits to the data width given for the element 0 05 002. The number 128 is the midpoint between 1 and 255 which is the range of values for the 8 bits of Y. Numbers between 1 and 127 will produce a negative value for changing data width, 129 to 255 a positive value.

The next step would be to change the scale from 10^2 to 10^3 in order to properly decode the reported latitude which will be encoded in Section 4 with 18 bits. The WMO BUFR definition for change scale, "Multiply scale given for each non-code data element in Table B by 10^(Y-128)," is referring to the result of 10^scale. For Table B entry 0 05 002, the scale is 2. In this case it is the resultant value 100 which is to be multiplied by 10^(Y-128), not the scale 2. Thus, the data description operator to change the scale for Table B entry 0 05 002 would be 2 02 129.

To complete the necessary changes for Table B, the reference value also needs to be modified from -9000 to -90000. Here again it must be determined how many bits are necessary to accommodate the new value, as the new reference value itself is encoded into Section 4. The number of bits to accommodate 90000 (positive value) is 17. It is, however, necessary to indicate this is to be a negative value which will require an additional bit. To indicate a new reference value as negative, the left most bit of the reference value encoded into Section 4 is set to 1. The sequence of operators needed to redefine or change a reference value is:

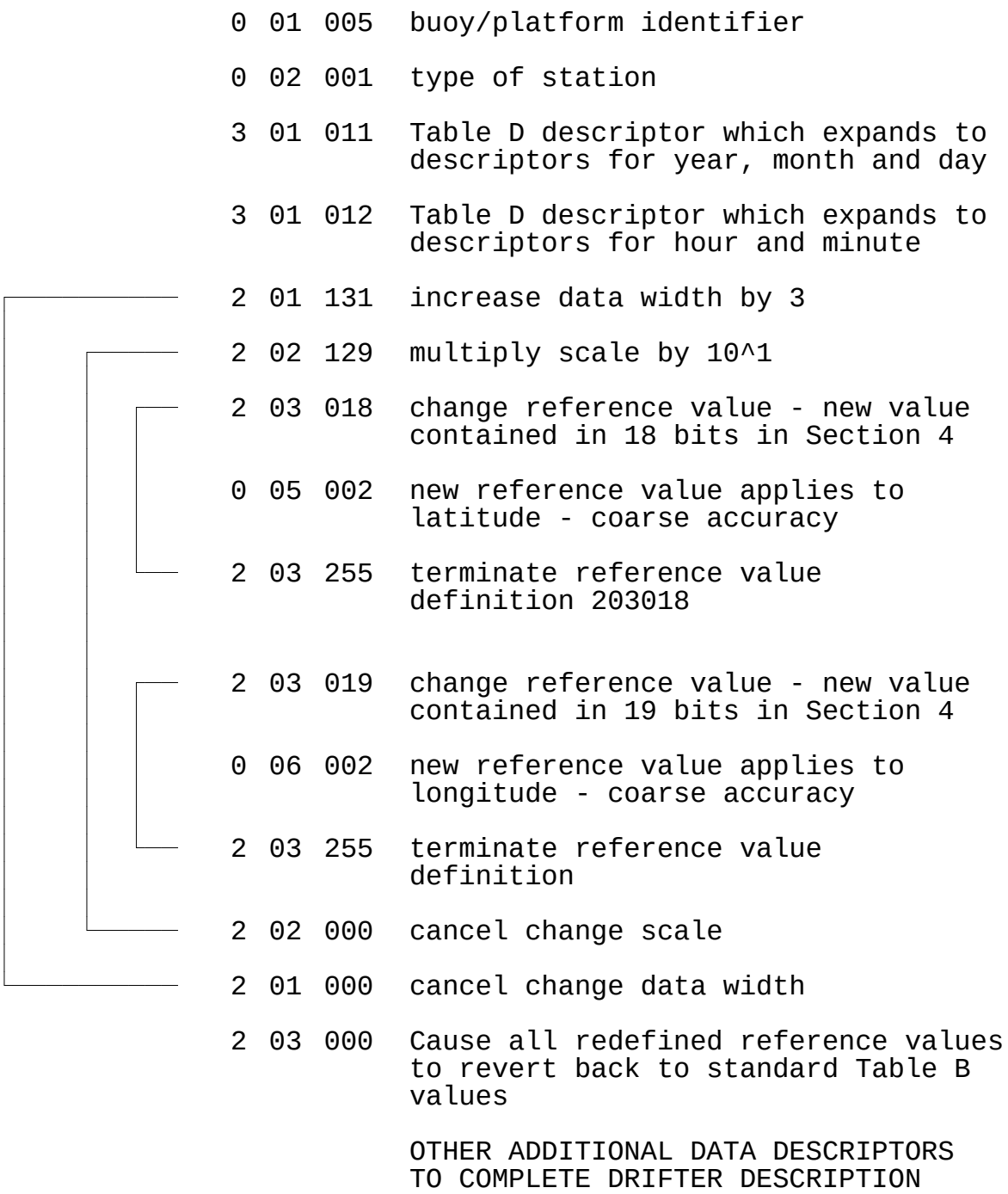
1) the 2 03 018 "change reference values operator," which announces a change and states how many bits are set aside for the new reference value in the data section (18 in this example)

2) one or more regular (F=0) data descriptors to indicate which variable(s) are to have new reference values. There are, of course, as many 18-bit values in the data as there are data descriptors following the 2 03 018 descriptor.

In this particular case it will not be necessary to have separate Data Description operators to modify longitude data width and change of scale. The increase in number of bits for data width to accommodate longitude to thousandths of degrees is also 3. The change of scale also remains the same. There will, however, be a required change of reference value from -18000 to -180000. By following the same steps as when changing the latitude Table reference value, the Data Description operator for changing the longitude reference value would be 2 03 019 followed by the data descriptor 0 06 002, followed by the descriptor 2 03 255 to indicate the end of the list of descriptors for which reference values are being changed.

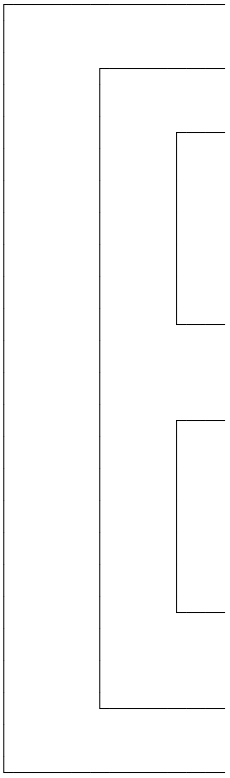
Once Data Description operators 2 01 Y, 2 02 Y and 2 03 Y have been used in Section 3, they remain in effect for the rest of whatever follows in the Section 3 data descriptions. To cancel operator 2 01, and 2 02, the additional entries 2 01 000 and 2 02 000 must be included in Section 3. To cancel the reference value change indicated by the operator 2 03 018, there must be included in Section 3 an operator 2 03 000.

The data description operators encoded into Section 3 for DRIFTER observations would then be as shown on the following page.



The order for cancellation of nested Data Description operators follows the above pattern; that is, the last change defined is the first canceled.

If instead of changing latitude and longitude from hundredths to thousandths, it were to be changed from hundred thousandths to thousandths the following descriptions would be used:



0	01	005	buoy/platform identifier
0	02	001	type of station
3	01	011	Table D descriptor which expands to descriptors for year, month and day
3	01	012	Table D descriptor which expands to descriptors for hour and minute
2	**01**	**121**	**decrease data width by 7**
2	**02**	**127**	**multiply scale by -1**
2	03	018	change reference value - new value contained in 18 bits in Section 4
0	**05**	**001**	**new reference value applies to latitude - high accuracy**
2	03	255	terminate reference value definition 203018
2	03	019	change reference value - new value contained in 19 bits in Section 4
0	**06**	**001**	**new reference value applies to longitude - high accuracy**
2	03	255	terminate reference value definition
2	02	000	cancel change scale
2	01	000	cancel change data width
2	03	000	Cause all redefined reference values to revert back to standard Table B values
			OTHER ADDITIONAL DATA DESCRIPTORS TO COMPLETE DRIFTER DESCRIPTION

Which would be the better of the methods? Again, use of local descriptors to define latitude and longitude is not a good idea as their use may cause a BUFR message to be undecodable in some other center. Of the two other methods, (1) using high accuracy latitude and longitude, or (2) using Data Description operators to change latitude and longitude definitions to thousandths of degrees, each will produce the same results. In terms of number of bits saved by changing to thousandths of degrees over high accuracy, a DRIFTER observation containing data equivalent to the DRIFTER code (FM 18-IX Ext. Sections 0 through Section 2) would require 214 bits per observation using high accuracy latitude and longitude. If latitude and longitude were changed by Data Description operators to thousandths of degrees then the observation would require 200 bits per observation, or a savings of 14 bits per observation - hardly worth the effort!

The preceding example does not imply that changing data width, scale and reference values should not be done, but it does point out that to do so to lower the number of bits within the data section for a given parameter is probably not that beneficial. In those instances where none of the possible Table B entries provide enough significance for new technologies, then the flexibility is provided within BUFR to handle those situations. If, for example, satellites were to measure latitude and longitude to millionths of degrees, then, to maintain significance of those measurements would require changing data width, scale and reference values, at least until (or if) there is a new Table B entry.

This example also shows that, when changing data width, scale and reference values, a single Table D descriptor cannot be used in Section 3. The reason is that changing data width and scale apply to all descriptors in Table B until the change data width and/or change scale is canceled. Since the descriptor to be affected may be deep within the Table D expansion process, there is no way to include the Data Descriptor operators in that expansion. A change in reference value, however, can be accomplished while still using a single Table D entry. This is possible because, after the entry for change reference value, 2 03 YYY, there must also be included the Table B descriptor or multiple descriptors that are to have new reference values.

5.2.1 Changing Reference Value Only. The Table B entries for geopotential, 0 07 003 and 0 10 003 have a reference value of -400 which is too restrictive for very low pressure systems. The Table C-Data Description operator 2 03 YYY can be placed as the first descriptor in Section 3, followed by the Table B descriptor(s) to which it applies. Placing 2 03 010, followed by 0 10 003 before the Table D descriptor means that each time data is encountered in Section 4 for 0 10 003, the new reference value indicated by the count of 10 bits specified by YYY applies. Within 10 bits the limit of the new reference value as a negative number is -511. The descriptor to conclude the list of descriptors for which new reference values are supplied follows immediately, and is followed

in turn by the Table D descriptor. See Figure 5-1 where the order of the Section 3 descriptors is:

2 03 010 0 10 003 2 03 255 3 09 008

The Section 4 data will be in the order indicated by that figure.

5.3 Add Associated Field. The Data Description operator 2 04 Y permits the inclusion of quality control information of Y bits attached to each following data element. The additional YYY bits of the associated field appear in the data section as prefixes to the actual data elements. The Add Associated Field operator, whenever used, must be immediately followed by the Class 31 Data Description Operator Qualifier 0 31 021 (data width 6 bits; see code table below) to indicate the meaning of the associated fields.

0 31 021
Associated field significance

Code Figure		
0	Reserved	
1	1 bit indicator of quality	0 = good 1 = suspect or bad
2	2 bit indicator of quality	0 = good 1 = slightly suspect 2 = highly suspect 3 = bad
3-6	Reserved	
7	Percentage confidence	
8-20	Reserved	
21	1 bit indicator of correction	0 = original value 1 = substituted/corrected value
22-62	Reserved for local use	
63	Missing value	

If quality control information were to be added to a single parameter such as pressure, Table B descriptor 0 07 004, the following sequence would appear in Section 3:

2 04 007 0 31 021 0 07 004 2 04 000

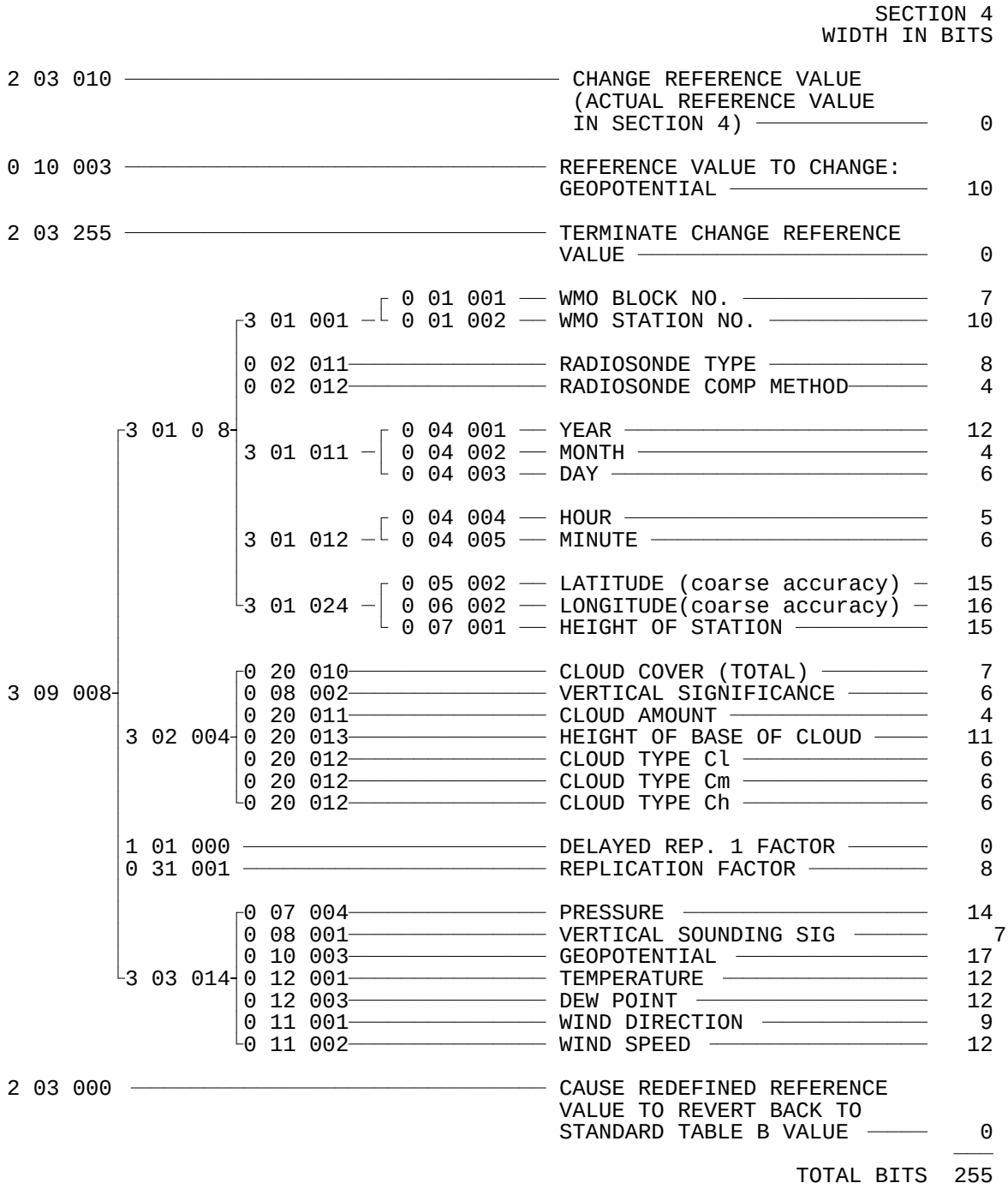


Figure 5-1. Change reference value of geopotential

The meaning of the above sequence is:

2 04 007 - indicator that 7 bits of data precede all following Table B entries

0 31 021 - code table entry for the meaning of the 7 bits preceding the Table B entry

0 07 004 - Table B entry for pressure

2 04 000 - cancellation of the Add Associated Field operator

The Section 4 data width for this sequence is 27 bits. The operators 2 04 007 and 2 04 000 do not occupy any bits within Section 4. The 27 bits are taken by 0 31 021 (6 bits) and 0 07 004 (21 bits, 7 bits of associated field plus 14 bits of pressure value)

When multiple Table B entries are preceded by 2 04 YYY as in:

2 04 007 0 31 021 0 07 004 0 31 021 0 10 003 2 04 000

the Add Associated Field operator 2 04 007 and the Data Description Operator Qualifier 0 31 021 both apply to the Table B descriptors 0 07 004 and 0 10 003. The Section 4 data width for the sequence is then:

2 04 007	0	bits
0 31 021	6	
0 07 004	21	(7 associated bits plus bits 14 data)
0 31 021	6	(change meaning of associated field)
0 10 003	24	(7 associated bits plus 17 bits data)
2 04 000	0	

Note that the associated fields are not prefixed onto the data described by 0 31 YYY descriptors. **This is a general rule: none of the Table C operators are applied to any of the Table B, Class 31 descriptors.**

If quality control information were to be added to the following sequence of parameters as described by the Table D descriptor 3 03 014 the following will result:

			SECTION 4 WIDTH IN BITS
3 03 014	0 07 004	PRESSURE	14
	0 08 001	VERTICAL SOUNDING SIG	7
	0 10 003	GEOPOTENTIAL	17
	0 12 001	TEMPERATURE	12
	0 12 003	DEW POINT	12
	0 11 001	WIND DIRECTION	9
	0 11 002	WIND SPEED	12
			83

By placing in Section 3 the operators 2 04 YYY and 0 31 021 immediately preceding 3 03 014, and the cancellation operator 2 04 000 following 3 03 014, the following sequence would be produced:

		SECTION 4 WIDTH IN BITS
2 04 007	ADD ASSOCIATED FIELD	0
0 31 021	ASSOCIATED FIELD SIG	6
	ASSOCIATED FIELD	7
0 07 004	PRESSURE	14
	ASSOCIATED FIELD	7
0 08 001	VERTICAL SOUNDING SIG	7
	ASSOCIATED FIELD	7
0 10 003	GEOPOTENTIAL	17
	ASSOCIATED FIELD	7
0 12 001	TEMPERATURE	12
	ASSOCIATED FIELD	7
0 12 003	DEW POINT	12
	ASSOCIATED FIELD	7
0 11 001	WIND DIRECTION	9
	ASSOCIATED FIELD	7
0 11 002	WIND SPEED	12
2 04 000	CANCEL ADD ASSOCIATED FIELD	0
		138

Adding associated fields to a data sequence that is described by a Table D descriptor means the associated fields are placed before all data items in the sequence. If quality control information were to be applied only to the pressure and geopotential parameters, the Table D descriptor could not be used but instead each individual parameter would have to be listed in Section 3.

2 04 007	ADD ASSOCIATED FIELD	0
0 31 021	ASSOCIATED FIELD SIG	6
	ASSOCIATED FIELD	7
0 07 004	PRESSURE	14
2 04 000	CANCEL ADD ASSOCIATED FIELD	0
0 08 001	VERTICAL SOUNDING SIG	7
2 04 007	ADD ASSOCIATED FIELD	0
0 31 021	ASSOCIATED FIELD SIG	6
	ASSOCIATED FIELD	7
0 10 003	GEOPOTENTIAL	17
2 04 000	CANCEL ADD ASSOCIATED FIELD	0
0 12 001	TEMPERATURE	12
0 12 003	DEW POINT	12
0 11 001	WIND DIRECTION	9
0 11 002	WIND SPEED	12
		109

If quality control information were to be added to the pressure and geopotential parameters in the TEMP observations as described in

Figure 3-1, the following adjustments would have to be made. The single Table D descriptor 3 09 008 could no longer be used as the expansion includes the additional Table D descriptor 3 03 014 which further expands to those parameters where quality control information would need to be inserted. The actual order of the Section 3 descriptors would now be (as illustrated in Figure 5-2):

```
 3 01 038   3 02 004   1 13 000   0 31 001   2 04 007   0 31 021
 0 07 004   2 04 000   0 08 001   2 04 007   0 31 021   0 10 003
 2 04 000   0 12 001   0 12 003   0 11 001   0 11 002
```

5.4 Encoding Character Data. There may be occasions when it is necessary to encode character data into BUFR. An observation encoded into BUFR that originated from the character code FM 13-IX Ext. SHIP, for example, has within that code form the optional inclusion of plain language. If this character information were carried over for encoding into BUFR, the Data Description operator 2 05 Y would be used in Section 3 to indicate the inclusion of character data in Section 4 of the BUFR message. The Y operand of the Data Descriptor indicates the number of characters, encoded CCITT International Alphabet No. 5, inserted as a data field in Section 4.

The following parameters from the FM 13-IX Ext. SHIP code form:

$$\left(\begin{bmatrix} 6I_S E_S E_S R_S \\ \text{or ICING +} \\ \text{plain language} \end{bmatrix} \right)$$

described by BUFR descriptors would be:

```
0 20 033   cause of ice accretion
0 20 031   ice deposit (thickness)
0 20 032   rate of ice accretion
```

It would have to be determined in advance how many characters would be allowed for the plain language. If only the word ICING were to be placed in Section 4, the Data Descriptor 2 05 005 would be used. If it were determined that ICING plus 25 additional characters, including spaces, were to be described then the descriptor would be 2 05 030. The data descriptors and data width in Section 4 would then be:

		data width in bits
0 20 033	cause of ice accretion	4
0 20 031	ice deposit (thickness)	7
0 20 032	rate of ice accretion	3
2 05 030	character information	240

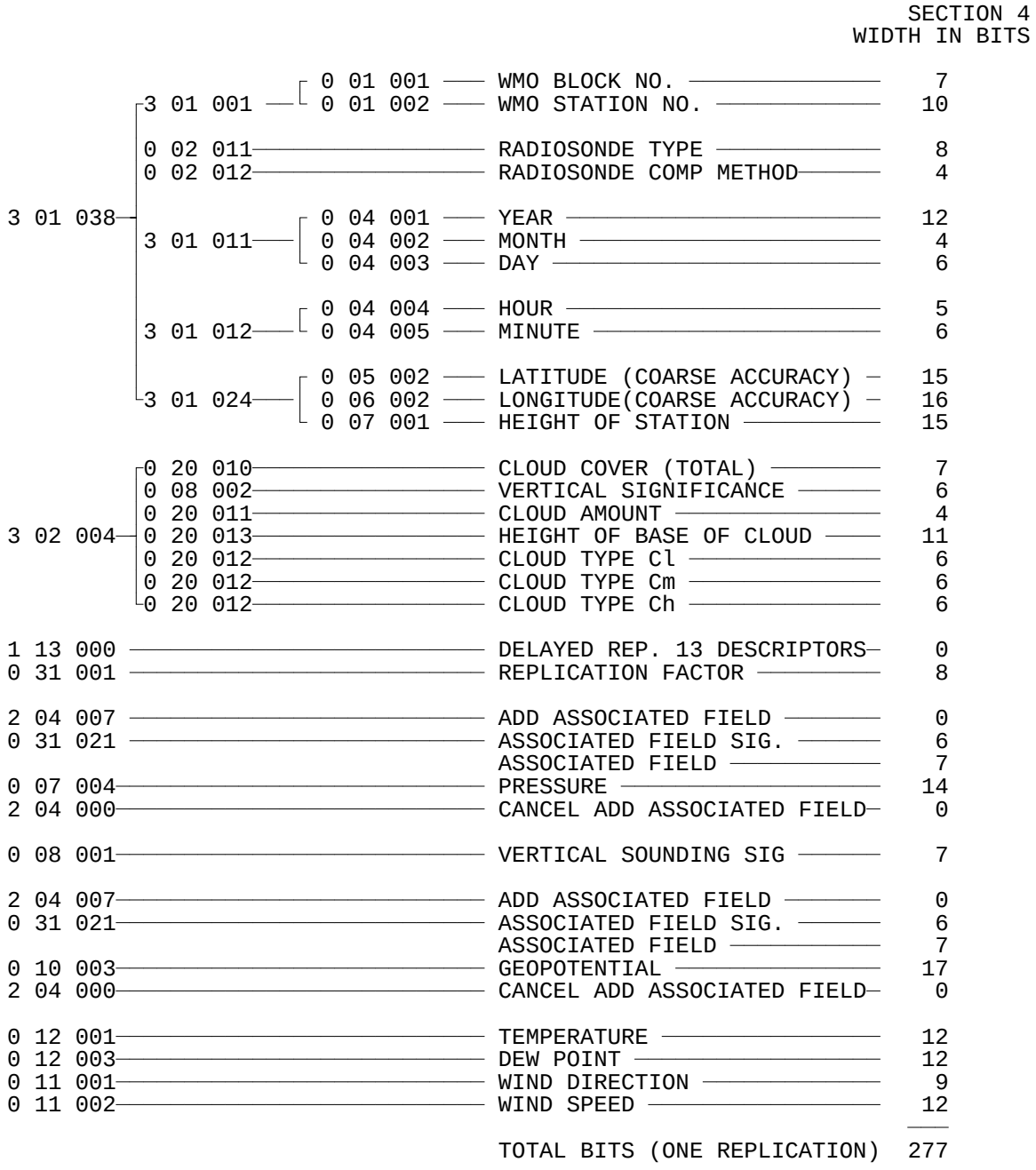


Figure 5-2. Example of TEMP observations sequence using delayed replication and quality control information.

Since an observation in FM 13-IX EXT. SHIP code would have either the parameters for ice reported, or ICING + plain language, but not both, then if there were no plain language the character information would be set to spaces. If the ICING + plain language were reported, then the data for descriptors 0 20 033, 0 20 031 and 0 20 032 would be set to missing, all bits set. Since Section 3 indicates a count of how many subsets (observations) are included in Section 4, the above descriptors apply to all subsets, even if an individual observation does not contain any icing information. In that case the entire set of icing data for an observation would be set to missing and spaces.

5.5 Signifying Length of Local Descriptors. Local Descriptors were provided in BUFR to enable a data processing center the capability of describing information of any type within BUFR for the center's internal use (Figure 2-4). There does exist, however, the possibility that once data is described in BUFR it may be necessary to transmit a BUFR message to another center, where the BUFR message would contain local information. Since a receiver of the BUFR message may or not know the meaning of the local descriptor, it could be impossible to be able to decode the message, as the receiver would not know the data width in Section 4 of the local information (Figure 2-5). While it could be argued that BUFR messages containing local information should never be transmitted to another center, it may require a separate set of software to remove local information before the message is ready for transmission. To overcome this situation the Data Description operator 2 06 Y was developed to allow local information to be contained within a transmitted message and to give information to the receiver that indicates the length in bits of the local data. The meaning of the Data Description operator 2 06 Y is that the following local descriptor is describing Y bits of data in Section 4. Knowing the width in bits of data in Section 4 then allows the receiver of the message to bypass that number of bits and permits proper decoding of Section 4. For example, see Figure 5-3.

The operator 2 06 Y can only be used when it precedes a local descriptor with F = 0. While it is within the rules of BUFR to create local descriptors with F = 3 (sequence descriptor), the Data Description operator 2 06 Y cannot be used to bypass whatever number of bits are being described by a sequence descriptor. Since a sequence descriptor expands to other descriptors and in the expansion process other local descriptors or delayed replication may be encountered, there is no way of knowing in advance how many total bits are covered by a sequence descriptor.

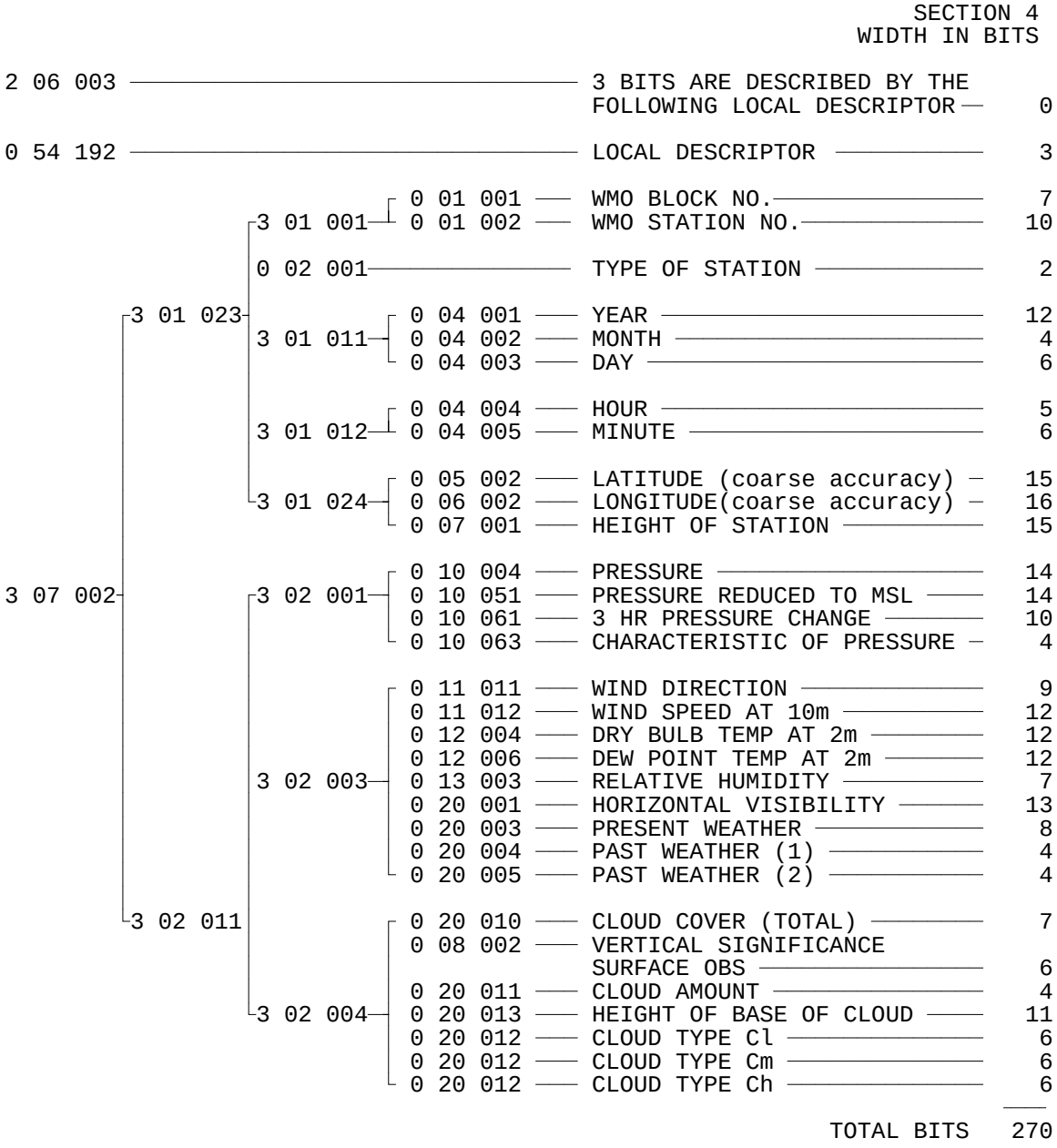


Figure 5-3. Example of surface observations with local descriptor and data descriptor operator 2 06 Y