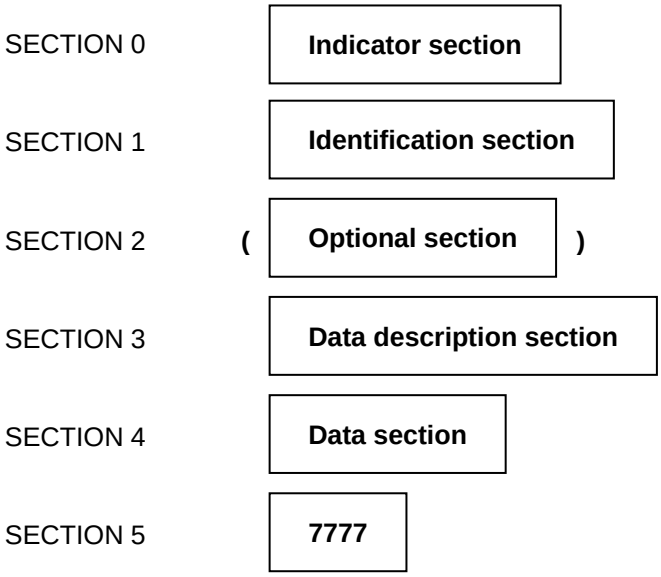
FM 94–XIV BUFR Binary universal form for the representation of meteorological data

REPRESENTATION FORM:



Notes:

(1) BUFR is the name of a binary code for the exchange and storage of data.

(2) The BUFR message consists of a continuous bit-stream made of a sequence of octets (1 octet = 8 bits).

(3) The terms “BUFR message” and “section” describe logical entities to assist BUFR definition.

(4) A BUFR message consists of one or more subsets of related meteorological data defined, described and represented by a single BUFR entity. For observational data, each data subset usually corresponds to one observation.

(5) The octets of a BUFR message are grouped in sections:

Section number	*Name*	*Contents*
0	Indicator section	“BUFR”, length of message, BUFR edition number
1	Identification section	Length of section, identification of the message
2	Optional section	Length of section and additional items for local use by automatic data processing centres
3	Data description section	Length of section, number of data subsets, data category flag, data compression flag and a collection of descriptors which define the form and content of individual data elements
4	Data section	Length of section and binary data
5	End section	7777

(6) It will be noted that the BUFR representation is not suitable for visual data recognition without computer interpretation.

(7) The representation of data by means of a series of bits is independent of any particular machine representation.

(8) Message and section lengths are expressed in octets. Section 0 has a fixed length of 8 octets; Section 5 has a fixed length of 4 octets. Sections 1, 2, 3 and 4 have a variable length which is included in the first three octets of each section.

(9) In the BUFR message, the bit length for “International Alphabet No. 5” is regarded as 8-bit, adding one bit “0” to the 7-bit of IA5 as the most significant bit.

REGULATIONS:

94.1 **General**

94.1.1 The BUFR form shall be used for the binary representation of meteorological data for exchange and storage. BUFR is particularly suitable for meteorological data that cannot be represented using FM 92 GRIB.

94.1.2 The beginning and the end of the code form shall be identified by 4 octets coded according to the International Alphabet No. 5 to represent, respectively, the indicators BUFR and 7777 in Indicator section 0 and End section 5. All other octets included in the code shall represent data in binary form.

Regulation for BUFR edition 3:

94.1.3 Each section included in the code form shall always contain an even number of octets. This rule shall be applied by appending bits set to zero to the section where necessary.

Regulation for BUFR edition 4:

94.1.3 Each section included in the code form shall always contain an integer multiple of 8 bits (octet). This rule shall be applied by appending bits set to zero to the section where necessary.

94.1.4 By convention, reserved values in sections 1 to 4 shall be set to zero.

94.1.5 Missing values shall be set to fields of all ones (e.g. each octet shall be set to 11111111 binary). This shall apply to code tables as well as data elements; flag tables shall be augmented to contain a missing indicator bit where this is deemed to be necessary. This regulation does not apply to the data description operator qualifiers in Class 31 of Table B.

94.1.6 The convention for representing missing data for compressed data within the binary Data section shall be to set the corresponding increments to fields of all ones.

94.1.7 When a local reference value for a set of element values for compressed data is represented as all ones, this shall imply that all values in the set are missing.

94.2 **Section 0 – Indicator section**

Section 0 shall be 8 octets long. Octets 1 to 4 shall be character coded according to the International Alphabet No. 5 as BUFR. The remainder of the section shall contain the length of the entire BUFR message (including the Indicator section) expressed in binary form over octets 5 to 7 (i.e. 24 bits), followed by the BUFR edition number, in binary, in octet 8.

94.3 **Section 1 – Identification section**

94.3.1 The length of the section, in units of octets, shall be expressed in binary form over the group of the first three octets of the section.

94.3.2 Octet 8 of the section shall be used to indicate the inclusion or the omission of section 2.

94.4 **Section 2 – Optional section**

94.4.1 Regulation 94.3.1 shall apply.

94.4.2 Octet 5 and subsequent octets shall contain additional items as may be defined within each centre for its own use.

94.5 Section 3 – Data description section

94.5.1 Regulation 94.3.1 shall apply.

94.5.2 Octets 5 and 6 of the section shall be used as a 16-bit number to indicate the number of data subsets within the BUFR message. Octet 7 shall be used to indicate whether observed data or other data are reported, and whether data are compressed or not. Octet 8 and subsequent octets shall contain a collection of descriptors which define the form and content of individual data elements in the Data section. A “data subset” shall be defined as the subset of data described by one single application of this collection of descriptors.

94.5.3 *Data description syntax for* BUFR

94.5.3.1 Data description shall consist of one or more descriptors. Each descriptor shall occupy 2 octets and contain 3 parts: F (2 bits), X (6 bits) and Y (8 bits).

94.5.3.2 If F = 0, the descriptor shall be called an “element descriptor”. An element descriptor shall define a single data item by reference to Table B.

Notes:

(1) X denotes the Table B class, Y denotes the element within that class. The corresponding data item is depicted according to the definition contained in Table B, unless otherwise modified.

(2) The definition(s) of one or more data item(s) may be modified by means of data description operators.

94.5.3.3 Element descriptors corresponding to the following classes in Table B shall remain in effect until superseded by redefinition:

Class	
01	Identification
02	Instrumentation
03	Reserved
04	Location (time)
05	Location (horizontal – 1)
06	Location (horizontal – 2)
07	Location (vertical)
08	Significance qualifiers
09	Reserved

Note: Redefinition is effected by the occurrence of element descriptors which contradict the preceding element descriptors from these classes. If two or more elements from the same class do not contradict one another, they all apply.

94.5.3.4 The consecutive occurrence of two identical element descriptors or identical sets of element descriptors from Classes 04 to 07 inclusive shall denote a range of values bounded by the corresponding element values. This enables the definition of layers and simple time periods.

94.5.3.5 The definition of line, areas, volumes and more complex time attributes shall be accomplished using descriptors from Classes 04 to 07 in association with suitable descriptors from Class 08.

94.5.3.6 The consecutive occurrence of two or more non-identical element descriptors from Classes 04 to 07 inclusive shall infer that all such elements remain in effect until redefined, unless such elements define an increment.

94.5.3.7 Data items defined by element descriptors in Class 10 or above shall not behave as coordinates with respect to subsequent data.

94.5.3.8 Increments: Any occurrence of an element descriptor from Classes 04 to 07 which defines an increment shall indicate that the location corresponding to that class shall be incremented by the corresponding data value. In the case of successive increments from the same class, this means that each increment shall apply in a cumulative manner, with all preceding increments remaining in effect.

Displacements: In contrast, any displacement descriptor from Classes 04 to 07 does not redefine the location corresponding to that class, but shall define only a transient displaced location from the location corresponding to that class. In the case of successive displacements from the same class, this means that each displacement shall apply independently and in a non-cumulative manner to the location corresponding to that class.

94.5.3.9 If a BUFR message is made up of more than one subset, each subset shall be treated as though it was the first subset encountered.

94.5.4 ***The replication operation***

94.5.4.1 If F = 1, the descriptor shall be called a “replication descriptor”. For this case, X shall indicate the number of descriptors to be repeated, and Y the total number of occurrences (replications) of the repeated subsequence.

Note: Where a replication operation includes delayed replication(s) within the scope of its replication, the replication (or repetition) factor descriptor(s) from Class 31 shall be counted for X, except the one (if any) located immediately after the replication description for which X is being calculated, as in the following example:

106000 031001 008002 103000 031001 005002 006002 010002.

94.5.4.2 A value of Y = 0 associated with the replication descriptor shall indicate delayed replication. In this case, the replication data description operator shall be completed by the next element descriptor, which shall define a data item indicating the number of replications. This descriptor may also indicate (by its value of Y) that the following datum is to be replicated together with the following descriptor.

94.5.4.3 Time or location increment descriptors, from Classes 04 to 07 inclusive, may be associated with replication descriptors in the following way: when an increment descriptor immediately precedes a replication descriptor, or is separated from it by one or more operator descriptors from Table C, this shall infer that all such increments be applied for each replication; the application of the increments shall have effect from the beginning of each defined replication, including the first.

94.5.5 ***Further operations on element and sequence descriptors***

94.5.5.1 If F = 2, the descriptor shall be called an “operator descriptor”. An operator descriptor shall define an operation by reference to Table C.

Notes:

(1) X denotes the value corresponding to an operator defined within Table C.

(2) Y contains a value to be used as an operand in completing the defined operation.

94.5.5.2 When the Y operand of any operator descriptor, or a count associated with it, refers to a specific number of descriptors preceding the operator, this shall infer that those preceding descriptors are all from Table B or C, i.e. all references to Table D descriptors shall have been completely resolved. Any forward reference to descriptors shall infer that the descriptors are enumerated as they are found in the original record, i.e. Table D descriptors are not expanded.

94.5.5.3 A data present bit-map shall be defined as a set of N one bit values corresponding to N data entities described by N element descriptors (including element descriptors for delayed replication, if present); the data description of a data present bit-map is comprised of a replication operator followed by the element descriptor for the data present indicator.

Notes:

(1) Where an operator descriptor requires a data present bit-map of length N to complete the operator definition, the N consecutive element descriptors which correspond to the N data entities to which the N bit values refer shall end with the element descriptor which immediately precedes the first such operator, or with the element descriptor which immediately precedes the first occurrence of such an operator following the occurrence of a cancel backward reference operator.

(2) All references to previously defined element descriptors effected through the application of operators which are qualified by data present bit-maps shall refer to the element descriptors concerned including any modifications resulting from change data width, change reference value, and change scale factor.

(3) The define data present bit-map for re-use operator enables a data present bit-map to be defined and later re-used; the definition of a data present bit-map shall remain defined until the occurrence of a cancel defined data present bit-map operator or a cancel backward data reference operator.

(4) Where an operator descriptor is qualified by a data present bit-map of length N there shall be defined a number of values of the type indicated by that operator together with subsequent appropriate element descriptors; the number of values defined shall correspond to the number of bits set to zero in the data present bit-map; the description of each data item shall be obtained by substituting the appropriate element descriptors, modified by the operator, at each subsequent occurrence of a marker operator.

94.5.6 ***Indirect reference to descriptors***

94.5.6.1 If F = 3, the descriptor shall be called a “sequence descriptor”. A sequence descriptor shall define a list of element descriptors, replication descriptors, operator descriptors and/or sequence descriptors by reference to Table D.

Note: X denotes the Table D category, Y denotes the entry within the category. Table D entries contain lists of commonly associated descriptors for convenience.

94.5.6.2 A sequence descriptor shall be equivalent to the corresponding list of descriptors in Table D.

Note: If a sequence descriptor is included within the scope of a replication descriptor 1 X Y, the number of descriptors to be repeated shall be modified if the sequence descriptor is replaced by the corresponding list of descriptors from Table D.

94.5.7 ***Unit rules***

94.5.7.1 The unit of an element descriptor, when not defined as a code table, flag table or CCITT5, should be based on the International System of Units (SI), established by the eleventh General Conference on Weights and Measures in 1960, and extended at the 1980 Conference. Alternatively, in exceptional cases, consideration may be given to other standard common units used by the data producer and the users, where a convincing case can be made that those units are more appropriate for the intended purpose of the descriptor. In such cases, priority shall be given to units contained in WMO Common Table C-6 or, in the case of descriptors for aviation products, ICAO Annex 5.

94.5.7.2 In cases where an element descriptor is defined as a code table that references values requiring units, Regulation 94.5.7.1 shall apply.

94.6 **Section 4 – Data section**

94.6.1 Regulation 94.3.1 shall apply.

94.6.2 Reported values shall be coded using the number of bits for each parameter indicated by reference to the sequence descriptors, replication descriptors, operator descriptors, element descriptors and associated tables.

94.6.3 Values shall be coded in the order indicated by the sequence descriptors, replication descriptors, operator descriptors and element descriptors.

Notes:

(1) Where more than one data subset is included in a single BUFR message without data compression:

(i) The first set of data values shall be in the order defined by the data description, and shall represent the first data subset;

(ii) Subsequent sets of data values shall also be in the order defined by the data description, representing subsequent data subsets.

(2) Where more than one data subset is included in a single BUFR message, data compression may be used as follows:

(i) Values for each data element are grouped into sets, and the sets shall be in the order defined by the data description; the first value in each set shall represent a minimum value for the set; for character data the first value in the set shall be set to all bits zero; however, if the character data values in all subsets are identical, the first value shall represent the character string; this value is termed a "local reference value", R^o, with respect to the subsequent set of data;

(ii) Local reference values shall be coded according to Regulation 94.6.2;

(iii) If all values of an element are missing, R^o shall be coded with all bits set to 1s;

(iv) The local reference value shall be followed by a 6-bit quantity specifying the number of bits for each increment or for character data, specifying the number of octets needed for representing the character string in the data subsets. However, if the character data values in all subsets are identical, sub-note vii shall apply;

(v) Integer values (V), other than character values and missing values, will then be obtained as:

$$V = R + R^o + I$$

where R = table reference value

R^o = local reference value

I = increment;

Actual data values (V_a) will be then obtained by:

$$V_a = V \times 10^{-S}$$

where S = table scale value

(vi) Missing values will be denoted by setting all bits of the corresponding I to 1s;

(vii) Data elements all having the same value throughout a set shall be signified by coding the number of bits required for storing I as zero; in such cases, the increments shall be omitted;

(viii) When operators qualified by a data present bit-map are present, it is required that the length and contents of the bit-map shall be identical for each data subset if data compression is to be used;

(ix) When delayed replication is present, it is required that the number of replications shall be identical for each data subset if data compression is to be used. In such cases, sub-note vii shall apply when coding the number of replications.

94.7 **Section 5 – End section**

The End section shall always be 4 octets long, character coded according to the International Alphabet No. 5 as 7777.

Editorial note: click following links to respective chapters in separate files.

SPECIFICATIONS OF OCTET CONTENTS

BUFR TABLES, CODE TABLES AND FLAG TABLES

FM 94 BUFR refers to three types of tables: BUFR tables, code tables and flag tables.

BUFR tables

Tables containing information used to describe, classify and define the contents of a BUFR message are called BUFR tables. Four BUFR tables are defined: Tables A, B, C and D. Entry numbering shall be the same in BUFR tables and CREX tables (see definition of FM 95 CREX in Part C, Common Features to Binary and Alphanumeric Codes) for the same entity represented. Table B entries shall be listed in the common BUFR/CREX Table B. Table D common sequences shall not be defined in both BUFR Table D and CREX Table D, unless otherwise a conversion between both Tables D is not simple, that is, the conversion is not completed by simple replacement of part “F” of each descriptor. A new BUFR Table D sequence shall be assigned a number not used by any CREX Table D sequence. Similarly, if a CREX Table D sequence is not defined in BUFR Table D, it shall be assigned a number not used by any BUFR sequence.

BUFR Table A - *Data category* (non-operational entries)

BUFR/CREX Table B - *Classification of elements* (non-operational entries)

BUFR Table C - *Data description operators* (non-operational entries)

BUFR Table D - *List of common sequences* (non-operational entries)

Code tables and flag tables (non-operational entries)

BUFR Table B defines some elements by means of code tables or flag tables. Within this general description are included code tables referenced by code figures, and flag tables where each bit is set to 0 or 1 to indicate a false or true value with respect to a specific criterion. The concept of a flag table is especially useful where combinations of criteria are represented. Within BUFR, all 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.