

## 0. Introduction

CMR.5 is a packed binary format designed as a compact alternative\* to LMR (Long Marine Reports), The National Climatic Data Center's TD-11 (Tape Deck-11), or other formats, containing some of the most frequently used variables. Each report has the internal structure given in Table D0-1. 192 bits was chosen as the minimum number of bits needed to represent the fields of interest, as well as being divisible by 16-, 32-, and 64-bit word sizes. 192 bits is also one-sixth the size of a 148-character TD-11 representation (given 8-bit character size).

It is assumed that the reader is familiar with techniques for transferring a binary block into memory and then extracting into INTEGER variables the bit strings whose lengths are given in Table D0-1. Refer to supp. H for more information. For a general discussion including the advantage in execution time and storage relative to traditional techniques see [3].

Compression was achieved by packing data represented as positive integers into fields whose lengths are specified in the *bits* column of Table D0-1. To accomplish this, a field's floating point *true value* (within the range of that column) was divided by the appropriate *units* (the smallest increment of the data that has been encoded). After rounding, the *base* was subtracted to produce a *coded* positive integer (within the range of that column), which was finally right-justified with zero fill in the field's position within the report. Using the sea surface temperature (field 9) *true value* 28.6 °C as an example,  $(28.6/0.1) - (-51) = 337$ .

Once a given field has been extracted into a *coded* value, the *true value* can be reconstructed by reversing the process:

$$\text{true value} = (\text{coded} + \text{base}) * \text{units}$$

The above *true value* example is reconstructed by  $(337 + (-51)) * 0.1 = 28.6$  °C. **NOTE: in each coded value, zero is reserved as an indicator of missing data.** Of course, none of BOX10, MONTH, BOX2, YEAR, X, or Y should ever be missing, although DAY and HOUR may be missing.

Explanations for each field in Table D0-1 are given under the corresponding headings that follow, where all information refers to the *true value* (unless explicit mention is made to the contrary), and some reference is made to TD-11 [5], [6], [7] or LMR (supp. F) documentation. The various indicators and flags show the reliability or precision of the data they refer to, and may be extant only if the data are also non-missing. Algorithms are expressed in FORTRAN.

---

\* CMR.5 supersedes CMR.4 (described in supp. E). The material in supp. E has been retained only for reference and includes details on translation from LMR (supp. F). The only omission from CMR.5 is the recorded wind speed. Because of rounding in the calculation of *coded* U and V, it can be only approximated by  $(U^2 + V^2)^{1/2}$ .

Table D0-1  
CMR.5

#	Field	Description	True value	Units*	Base	Coded	Bits	
<u>Location</u>								
1	BOX10	10 ° box	1 ≤ 648**	1***	0	same	10	
2	MONTH		1 ≤ 12	1	0	same	4	
3	BOX2	2 ° box.	1 ≤ 16202	1	0	same	14	
4	YEAR		1800 ≤ 2054	1	1799	1 ≤ 255	8	
5	DAY		1 ≤ 31	1	0	same	5	
6	HOURL		0 ≤ 23	1	-1	1 ≤ 24	5	
7	X	lon (from BOX2	0 ≤ 2.0	0.1 °	-1	1 ≤ 21	5	
8	Y	lat SW corner)	0 ≤ 2.0	0.1 °	-1	1 ≤ 21	5	
							sub-total	56
<u>Temperature</u>								
9	S	sea surface temperature	-5.0 ≤ 40.0	0.1 ° C	-51	1 ≤ 451	9	
10	BI	bucket indicator	0 ≤ 2	1	-1	1 ≤ 3	2	
11	A	air temperature	-88.0 ≤ 58.0	0.1 ° C	-881	1 ≤ 1461	11	
12	DP	dew point depression	0 ≤ 70.0	0.1 ° C	-1	1 ≤ 701	10	
13	TI	temperature indicator	0 ≤ 5	1	-1	1 ≤ 6	3	
							sub-total	35
<u>Wind</u>								
14	U	eastward component	-102.2 ≤ 102.2	0.1 m s <sup>-1</sup>	-1023	1 ≤ 2045	11	
15	V	northward component	-102.2 ≤ 102.2	0.1 m s <sup>-1</sup>	-1023	1 ≤ 2045	11	
16	DI	direction indicator	0 ≤ 5	1	-1	1 ≤ 6	3	
17	WI	wind speed indicator	0 ≤ 1	1	-1	1 ≤ 2	2	
							sub-total	27
<u>Pressure and clouds</u>								
18	P	sea level pressure	870.0 ≤ 1074.6	0.1 mb	8699	1 ≤ 2047	11	
19	C	total cloud amount	0 ≤ 9	1	-1	1 ≤ 10	4	
20	NH	lower cloud amount	0 ≤ 9	1	-1	1 ≤ 10	4	
21	CL	low cloud type	0 ≤ 10	1	-1	1 ≤ 11	4	
22	H	cloud height	0 ≤ 10	1	-1	1 ≤ 11	4	
23	HI	cloud height indicator	0 ≤ 1	1	-1	1 ≤ 2	2	
24	CM	middle cloud type	0 ≤ 10	1	-1	1 ≤ 11	4	
25	CH	high cloud type	0 ≤ 10	1	-1	1 ≤ 11	4	
							sub-total	37
<u>Misc.</u>								
26	ST	ship type	0 ≤ 7	1	-1	1 ≤ 8	4	
27	PW	present weather	0 ≤ 99	1	-1	1 ≤ 100	7	
28	CD	card deck	0 ≤ 999	1	-1	1 ≤ 1000	10	
							sub-total	21
<u>Flags</u>								
29	LF	landlocked flag	0 ≤ 0	1	-1	1 ≤ 1	1	
30	SF	SST flag	0 ≤ 2	1	-1	1 ≤ 3	2	
31	AF	air temperature flag	0 ≤ 2	1	-1	1 ≤ 3	2	
32	RF	relative humidity flag	0 ≤ 2	1	-1	1 ≤ 3	2	
33	WF	wind flag	0 ≤ 2	1	-1	1 ≤ 3	2	
34	PF	pressure flag	0 ≤ 2	1	-1	1 ≤ 3	2	
							sub-total	11
35	CK	checksum	n/a	n/a	n/a	n/a	5	
							total	192

\* "Units" gives the smallest increment of the data that has been encoded. Thus a change of one unit in the integer coded value represents a change in the true value of one of the units shown.

\*\* m ≤ n denotes "from m through n inclusive."

\*\*\* Units of 1 are explained in the text.

## 1. Fields

### 1) BOX10 10° box

See supp. G for a description of the 10° box system, and supp. H for related software.

### 2) MONTH

1=January, 2=February, ..., 12=December.

### 3) BOX2 2° box

See supp. G for a description of the 2° box system, and supp. H for related software.

### 4) YEAR

The year can range from 1800 to 2054.

### 5) DAY

Day of the month.

### 6) HOUR

00 to 23 GMT.

### 7) X longitude

### 8) Y latitude

Position in tenths of a degree measured from the SW (lower-left) corner of the BOX2. Range is from 0 to 2.0 subject to the boundary constraints of a BOX2:

- a) Boxes in the NE quadrant have  $0 \leq X < 2.0$ ,  $0 \leq Y < 2.0$  (except if box E boundary is  $180^\circ$  E,  $0 \leq X \leq 2.0$ ).
- b) Boxes in the NW quadrant have  $0 < X \leq 2.0$ ,  $0 \leq Y < 2.0$  (except if box W boundary is  $180^\circ$  E,  $0 \leq X \leq 2.0$ ).
- c) Boxes in the SE quadrant have  $0 \leq X < 2.0$ ,  $0 < Y \leq 2.0$  (except if box E boundary is  $180^\circ$  E,  $0 \leq X \leq 2.0$ ).
- d) Boxes in the SW quadrant have  $0 < X \leq 2.0$ ,  $0 < Y \leq 2.0$  (except if box W boundary is  $180^\circ$  E,  $0 \leq X \leq 2.0$ ).
- e) Boxes 1 and 16202 have X and Y (by convention) equal 0 always.

### 9) S sea surface temperature

### 10) BI bucket indicator

Temperature S in tenths of a degree Celsius. BI shows the method by which S was taken:

0 = unknown

1 = bucket

2 = implied bucket (an HSST source or any match thereof)

NOTE: BI values 0 and 1 are unreliable at least for U.S. recruited ships (i.e., country code 0K or 02) until starting on 1 May 1973, or perhaps earlier (see *COADS Release 1*, and for country codes see [6]).

- 11) A        air temperature  
12) DP       dew point depression

Temperatures A and DP in tenths of a degree Celsius. Let DPT denote dew point temperature. Dew point depression is defined as

$$DP = A - DPT$$

- 13) TI        temperature indicator

Shows the precision and units that S, A, and DP were recorded in or later translated to (see supp. I):

- 0 = degrees Celsius and tenths
- 1 = whole degrees Celsius
- 2 = half degrees Celsius
- 3 = degrees Fahrenheit and tenths
- 4 = whole degrees Fahrenheit
- 5 = half degrees Fahrenheit

- 14) U        vector wind eastward component  
15) V        vector wind northward component

U and V were computed to tenths of a meter per second, using the wind direction in degrees (D) and wind speed in tenths of a meter per second (W) as follows:

$$\begin{aligned} \text{ANG} &= D * (3.14159265359 / 180.) \\ U &= - W * \text{SIN}(\text{ANG}) \\ V &= - W * \text{COS}(\text{ANG}) \end{aligned}$$

(Supp. F describes how the original compass reading was translated into whole degrees.)

- 16) DI        direction indicator

DI shows the compass (and approximate precision) used to report the direction contributing to U and V:

- 0 = 36-point compass
- 1 = 32-point compass
- 2 = 16 of 36-point compass
- 3 = 16 of 32-point compass
- 4 = 8-point compass
- 5 = 360-point compass

- 17) WI        wind speed indicator

WI shows if the wind speed was:

0 = estimated (or unknown method of observation)  
1 = measured

18) P        sea level pressure  
In tenths of a millibar.

- 19) C        total cloud amount
- 20) NH      lower cloud amount
- 21) CL      low cloud type
- 22) H       cloud height
- 23) HI      cloud height indicator
- 24) CM      middle cloud type
- 25) CH      high cloud type

Except for HI, the cloud fields 19)-25) have possible codes 0 to 9 as given by TD-11, or a 10 corresponding to the minus sign given therein for CL, H, CM, and CH. Alternately, see supp. F for these definitions. HI shows if H was:

0 = estimated  
1 = measured

26) ST       ship type

The type of observing vessel was obtained according to supp. I, and the unreliability of this field is discussed in *COADS Release 1*.

- 0 = U.S. Navy or "deck" log, or unknown
- 1 = merchant ship or foreign military
- 2 = ocean station vessel -- off station or station proximity unknown
- 3 = ocean station vessel -- on station
- 4 = lightship
- 5 = buoy
- 6 = research ship
- 7 = expendable or mechanical bathythermograph (XBT or MBT)

27) PW       present weather

Codes 0 to 99 as given by TD-11 or supp. F.

28) CD       card deck

Number of the source card deck the report came from, as assigned by NCDC and described in supp. F.

- 29) LF landlocked flag
- 30) SF sea surface temperature (S) flag
- 31) AF air temperature (A) flag
- 32) RF relative humidity (R) flag
- 33) WF wind (W, U, V) flag
- 34) PF pressure (P) flag

The flags 29)-34) show whether the variables they refer to were trimmed (i.e., excluded from the summaries but retained in CMR) as apparent statistical outliers or for other reasons.

The flag LF has (NOTE: distinct from the usual *coded* 0 for missing) one possible extant *true value*:

0 = landlocked

indicating that the 2° box is landlocked according to the landlocked file described in supp G. In this case, the other flags (SF, AF, RF, WF, PF) are all missing and any data were automatically trimmed because trimmed summaries were generated only for reports with LF missing.

If LF is missing, the other flags may still be missing or else carry one of the following values:

$$\begin{aligned}
 0 &= g - 2.8\sigma_1 \leq a_i \leq g + 2.8\sigma_5 \text{ (not trimmed)} \\
 1 &= g - 3.5\sigma_1 \leq a_i \leq g + 3.5\sigma_5 \text{ (not trimmed)} \\
 2 &= a_i < g - 3.5\sigma_1 \text{ or } a_i > g + 3.5\sigma_5 \text{ (trimmed)}
 \end{aligned}$$

where  $a_i$  is an individual observation of the variable under scrutiny,  $g$  is the smoothed median, and  $\sigma_1$  and  $\sigma_5$  are the smoothed lower and upper median deviation. The computation and format of these smoothed limits is described in supp. C.

If LF is missing, other flags set to missing indicate either that the smoothed limits are missing, and thus the variables referred to were automatically trimmed, or that the variables are missing. Thus missing flags must be evaluated in conjunction with the actual variables (the special case of RF is discussed in the following).

Assignment of the flags SF, AF, and PF was accomplished as follows:

- a) If the 2° box was landlocked, the flag was set to missing.
- b) If the smoothed limits were missing, the flag was set to missing.
- c) If the variable was missing, the flag was set to missing.
- d) If the variable fell within the narrow interval ( $g - 2.8\sigma_1 \leq a_i \leq g + 2.8\sigma_5$ ), the flag was set to 0 and the variable was included in the summaries.
- e) If the variable fell within the wide interval ( $g - 3.5\sigma_1 \leq a_i \leq g + 3.5\sigma_5$ ), the flag was set to 1 and the variable was included in the summaries.
- f) Otherwise, the flag was set to 2 and the variable was trimmed from the summaries because it fell outside the wide interval.

Assignment of WF depended jointly on U and V, but followed the same basic rules. U was tested against its limits first, and then V was tested against its limits, retaining the maximum flag value found in the separate tests for U and V. In case either the limits for U or V were

missing, or both, the flag was set to missing. As a result, all three wind variables (U, V) and W (wind speed) were included in the summaries only if WF had a value 0 or 1.

Assignment of RF was a special case only because relative humidity does not appear directly, and depends on A (air temperature) and DP (dew point depression). To handle this dependence, AF was assigned first. Only if AF then had a value 0 or 1 and DP was extant could R be computed; otherwise R was considered missing.

### 35) CK      checksum

A checksum was computed and stored with each report as a measure of reliability during storage and transmission. The checksum is computed by

- 1) Summing *coded* values of all other fields in the report besides the checksum.
- 2) Obtaining the modulo ( $2^5 - 1$ ) of the sum.

Repeating this calculation for every unpacked report, and then verifying that the checksum so obtained agrees with the *coded* checksum stored in the report, is strongly encouraged. For example, supposing that the *coded* values of the preceding 34 fields are available in an array FIELD, the checksum CK is computed and verified against the stored checksum CKS in FORTRAN as follows:

```
INTEGER CK,J,FIELD(34),CKS
CK = 0
DO 500 J = 1,34
500 CK = CK + FIELD(J)
CK = MOD(CK, 31)
IF (CK .NE. CKS) THEN
  PRINT *,'ERROR. CK = ',CK,' .NE. CKS = ',CKS
  STOP
ENDIF
```

Note that using modulus  $2^5 - 1$  takes into account every bit of CK, versus chopping at the sixth bit using modulus  $2^5$ .