UNIFIED MODEL DOCUMENTATION PAPER F3

FORMAT OF

ATMOSPHERIC, OCEANIC and WAVE DUMPS,

FIELDSFILES, ANCILLARY DATA SETS,

BOUNDARY DATA SETS AND

OBSERVATION FILES

for the

UNIFIED FORECAST / CLIMATE MODEL

D. ROBINSON

VERSION No 25 - Dated 23/11/98

This version matches version 4.5 of the Unified Model
Also covers pre-version 4.5 files

Numerical Weather Prediction
Meteorological Office
London Road
BRACKNELL
Berkshire
RG12 2SZ
United Kingdom

REVISION

Versions 3-20: see version 20.

Modification Record From Version 21		
Document version	**Author**	**Description.....**
21	D.M. Goddard	Definitions of headers updated and clarified Additional choices at vn3.4 for Fixed_length_header 3 and 5 New definition for Fixed_length_header Fixed_length_header 153
22	A.S. Lawless	REAL_CONSTANTS(37)&(38) added at vn3.3.
22.1	M.J. Hatton	20/10/95 Usage of PP header items 44 and 45 included
23	D. Robinson	25/07/97. UM Vn 4.0/4.1/4.2/4.3 (*) changes included. Integer Constants 22-24, Real Constants 22-28 and 30-38. PP Header, Item 45. Changes to LOOKUP(29/30) for Well-formed I/O Files. Changes to ocean dumps for MPP platforms. New section on Wave Dumps.
24	D. Robinson	18/05/98. Um Vn 4.4 changes included. Additions : Atmos - Integer Constants (22-23). Corrections : Atmos - Integer Constants (28) and Ocean dumps - Integer and Real Constants.
25	D. Robinson	23/11/98. Um Vn 4.5 changes included. All references to Pre Vn 2.8 files removed - See Version 24. Changes to FIXHD(15), Integer Constants (16-17), LOOKUP 21 and 39.

* - Vn 4.2 was not released for general use. Vn 4.3 is the first general release on the new T3E and contains new code to run the UM on a MPP platform.

Contents

1. Introduction

This note describes the format of the data files used within the unified forecast/climate model system. The structure of the files is accepted for both atmosphere and ocean applications for storing dumps, ancillary data, output data (fieldsfiles), boundary data and observations.

In designing this format the aim has been to avoid any features which would make the format inflexible with regard to changes in number of levels, number of field types or even the order in which fields are held, as well as being adaptable for different applications to which the UM might eventually be put. This has led to the construction of a file format which consists of a primary header record with pointers to a series of secondary header records and also records pointing to and describing the data areas as well as the data areas themselves.

Each field has an associated 64 (*) word header contained within a lookup table. The location of that field is obtained by a search on the appropriate element(s) (primarily that giving the PP code) of the lookup table. Full details are available online at PPP.DOC.TEXT(@FCODES). The lookup table describes the contents of that field, in a manner similar to that adopted by the fieldsfile. It is essentially the extended version of the pp header record and section 4 gives the format along with appropriate user defined extensions applicable to UM data sets.

Note that the adoption of the fieldsfile type lookup table implies a mixed integer/real array which has conventionally been decoded by EQUIVALENCING. This non standard convention is allowable within CONTROL routines only.

In the basic version of the model there is an assumption when searching the lookup that the levels for a given variable are contiguous. This can still be assumed but it is a programming restriction not a data set restriction. All authors of code which generate dumps, FIELDSfiles or ancillary data sets should obey the restriction. However, one cannot assume that the first occurrence of a particular field is the primary model data field (it might be a timeseries, zonal mean diagnostic etc), so a more comprehensive check on several LOOKUP words is required to uniquely identify a field. In practice, the primary model data fields will occur first in the dumps before other fields sharing the same primary PP code. Also in practice, there will be a well defined order in which fields are held within the dump. It would be too restrictive to lay down such an order precisely at this stage, but if we accept that prognostic variables will precede diagnostic variables which will precede timeseries then it will be a much simpler process to compare related dumps. FIELDSfiles will not have such a well defined order of fields. Ancillary data sets contain much smaller amounts of information with a conventional order of fields for each type of data set.

(*) *128 word header for observation files from version 3.3*

The constants included in section 3 of this version of the dump are not intended to be a comprehensive set as further constants may be required. It is important that we avoid proliferation of constants in the dump file, since if it becomes established that the DUMP is a repository for any sort of model 'Constant', there will be so many changes in content that an archive of dumps will rapidly become redundant. Constants which can easily be derived from other more basic constants should not be included. Tunable constants which are a function of the model and not the data should not be obtained from the dump since this will make rerunning of archived cases with the latest version of the model unnecessarily complicated, however it is permissible to put copies of such constants in the dump as a future reference source to define the run more exactly. In such instances it is important that any run-time variation (which will be taken from a namelist) to the values held in the initial file are copied into the output dump file. Such 'passive' constants are identified here by means of a '*'. Units for variables are given in UMDP no. 5.

Requests for additional constants fields should be channelled through the Unified Model Librarian. The examples given are specific to the basic atmospheric version of UM and a different set will be applicable to other applications.

2. List of data set components

The full list of components of the data set format is given below. This looks a formidable list but the option should be available of having empty records if they are not required. Note that an empty record is defined as having a negative start address. The dimensioning of that missing record will be 1 to avoid problems with zero dimension arrays, but there is no requirement to read that single word.

2.1 FIXED_LENGTH_HEADER
A primary record of 256 integer words whose main purposes are to locate the secondary records and identify the type of dump. There should be nothing in this header which is specific to any one application. All the remaining records have a variable number of words can be configured differently for different applications.

2.2 INTEGER_CONSTANTS
An integer array containing model variables which further identify the data, or model variables which need to passed between runs or tunable model variables which although fixed might usefully be held in the dump to identify the run for future reference. This record should not contain any arrays.

2.3 REAL_CONSTANTS
As 2.2 above but for real variables.

2.4 LEVEL_DEPENDENT_CONSTANTS
Within the code this is anticipated to be a two dimensional array containing variables for which values are required at every level. It should not be assumed that the first dimension of the array is the number of levels. Since we need the flexibility of holding constants which might apply to different sets of levels (eg all levels, wet levels, soil levels etc). Real or integer values may be specified.

2.5 ROW_DEPENDENT_CONSTANTS
As 2.4 above but for variables which are required for each row

2.6 COLUMN_DEPENDENT_CONSTANTS
As 2.4 above but for variables which are required for each column (ie point along a row)

2.7 FIELDS_OF_CONSTANTS
As 2.4 above but for variables which are required for each grid point in the field

2.8 EXTRA_CONSTANTS
An unstructured one dimensional array containing other essential 'constants' that do not conveniently fit into 2.2 to 2.7 above.

2.9 TEMP_HISTORYFILE
A fixed length record of 2048 words containing a copy of the temporary

HISTORY COMMON block. It helps to identify the run and aid setting up of repeat runs. **This was set to zero length from UM Vn 3.4.**

2.10 COMPRESS_FIELD_INDICES

Can be used in place of an integer array of 1/0 should we wish to eliminate the need to store land points (where storage is at a premium and the additional complexity of unpacking with this indexing system is an acceptable alternative to using more storage (initially not required for atmosphere use . It is anticipated that three indices will be required and full implementation details are in Ocean Model documentation.

2.11 LOOKUP

A 64 or 128 word header table (usually one header per 2D field) describing and pointing to the model's real and integer data fields including timeseries. This is essentially an adaption of the pp headers. Full details are given in section 4.

2.12 DATA

Real, integer or logical data fields or observations including timeseries.

3 Details of the data set contents

3.1 FIXED_LENGTH_HEADER

1. Data Set Format Version Number.
   - 15 Data set format : pre-Vn 3.1.

   Since Vn 3.1, the use of Word 1 in the fixed header has been discontinued and is set to -32768 (IMDI). Any format change is now connected with a new version of the UM.

2. Indicator for Sub-Model :
   - 1 Atmosphere
   - 2 Ocean
   - 4 Wave **Vn4.1 onwards**

3. Indicator for vert coord type :
   - 1 hybrid
   - 2 sigma
   - 3 pressure
   - 4 depth
   - 5 Charney-Phillips on radius levels needed for dynamics grid, used by variational assimilation tangent linear dumps. **Vn3.4 onwards**
   - 6 Wave model direction and frequency pseudo-levels. **Vn4.1 Onwards**

4. Indicator for horiz grid type :
   - 0 Global
   - 1 Northern Hemisphere
   - 2 Southern Hemisphere
   - 3 LAM(No Wrap)
   - 4 LAM(Wrap) for equatorial lat-long LAM add 100

5. Indicator for dataset type :
   - 1 Instantaneous dump
   - 2 Mean dump
   - 3 FIELDSfile
   - 4 Ancillary dataset
   - 5 Boundary dataset
   - 6 AC observation file
   - 7 VAR observation file
   - 8 Cx file (model columns at observation location
   - 9 Covariance File
   - 10 OPS Obstore File

6. Run identifier :
   - 0 Undefined
   - 1 Main
   - 2 Update

   or programmer code (>100)

7. Experiment Number eg identify version of model (1) or as for climate model?
8. Indicator for calendar :
   - 1 standard gregorian calendar
   - 2 360 day calendar
9. Indicator for grid staggering eg (1-5 for Arakawa a-e grids)
10. Indicator for ancillary data sets showing times at which data is provided:
    - 0 Single time
    - 1 Time series
    - 2 Periodic time series
11. Projection Number
12. Model version number x 100 + release number. Eg Vn2.4 is stored as 204
14. Type of Observation File 1/2/3/4 Atmosphere/Ocean/SST/Wave

15. Type of arithmetic operation last done in fieldop. **From Vn 4.3**
    - 100 Addition
    - 200 Subtraction
    - 300 Multiplication
    - 400 Division by Integer

21. Year }
22. Month. }
23. Day of month. }
24. Hour. }
25. Minute. }
26. Second. }
27. Day number }

Dumps : Initial data time
Observation file : Validity time
Ancillary File, Boundary dataset or FIELDSFILE : First Validity Time

28. Year. }
29. Month. }
30. Day of month. }
31. Hour. }
32. Minute. }
33. Second. }
34. Day number }

Dumps : Validity Time of instantaneous fields (Start time of mean fields)
Observation file : Validity Time
Ancillary Files, Boundary datasets, FIELDSFILE : Last validity time.

35. Year. }
36. Month. }
37. Day of month. }
38. Hour. }
39. Minute. }
40. Second. }
41. Day number. }

Dumps and observation files : Wall clock time file generated.
Ancillary File and Boundary dataset : Interval between validity times of data. Set to zero if such data prsent for one time only.

100. start of **INTEGER_CONSTANTS**
101. dimension of integer constants array

105. start of **REAL_CONSTANTS**
106. dimension of real constants array

110. start of **LEVEL_DEPENDENT_CONSTANTS**
111. first dimension of level dep constants array
112. second dimension of same

115. start of **ROW_DEPENDENT_CONSTANTS**
116. first dimension of row dep constants array
117. second dimension of same

120. start of **COLUMN_DEPENDENT_CONSTANTS**
121. first dimension of col dep constants array
122. second dimension of same

125. start of **FIELDS_OF_CONSTANTS**
126. first dimension of fields of constants array
127. second dimension of same

130. start of **EXTRA_CONSTANTS**
131. length of this array.

135. start of **TEMP_HISTORYFILE**
136. length of this array.

140. start of **COMPRESSED_FIELD_INDEX1**
141. length of this array
142. start of **COMPRESSED_FIELD_INDEX2**
143. length of this array
144. start of **COMPRESSED_FIELD_INDEX3**
145. length of this array

150. start of **LOOKUP** table
151. first dimension of this lookup table (64 or 128)
152. second dimension of this lookup table
153. number of prognostic fields in dump. **vn 3.4 onwards**
(Only used for instantaneous dumps FIXED_LENGTH_HEADER(5) = 1)

160. start of **DATA**
161. dimension of **DATA**
162. Maximum length of all fields in **DATA**
(Only used by observation file so far)

3.2 INTEGER_CONSTANTS
Location and dimensions defined by words 100-101 of FIXED_LENGTH_HEADER
Separately defined for ocean and atmosphere use

Full list of words reserved for **Atmosphere** use

1. Number of timesteps since start of run
2. Meaning interval for the mean fields (hours)
3. Number of instantaneous dumps used to generate mean field
4. User defined if non-contiguous period used for meaning. (0) if contiguous otherwise no. of hours between neighbouring contiguous. sections of means
5. User defined if non-contiguous period used for meaning. (0) if contiguous otherwise no. of hours between end of one contiguous. section and start of next
6. Number of points E-W (x direction or 1st dimension)
7. Number of points N-S (y direction or 2nd dimension)
8. Number of levels (P_LEVELS) (numbered away from the earths surface)
9. Number of wet levels (Q_LEVELS) (ditto)
10. Number of soil levels (DS_LEVELS)
11. Number of cloud levels (CLOUD_LEVELS)
12. Number of tracer levels
13. Number of boundary layer levels
14. Number of 'passive' tracers to be advected excluding moisture
15. Number of different field types in dump
16. Not used. **Pre Vn 4.5** : Number of Soil variables.
17. Not used. **Pre Vn 4.5** : Number of Vegetation variables.
18. Number of radiation variables
19. NORTHERN_FILTERED_P_ROW
20. SOUTHERN_FILTERED_P_ROW
21. Value of integer missing data indicator (-32768)

22-24 : Pre Vn 4.0

22. Start level for gravity wave drag
23. Start level for vertical diffusion
24. Final level for vertical diffusion

22-23 : Vn 4.0 onwards

22. Not used
23. Not used

22-23 : Vn 4.4 onwards

22. Calling period (days) for TRIFFID vegetation model otherwise not used.
23. Number of atmosphere timesteps since last call to TRIFFID otherwise not used.

24. Not used - **Vn 4.0 onwards**.

25. Number of land points
26. Number of ozone levels
27. Number of levels at which tracers advected
28. Atmos Dump : Number of soil hydrology (moisture) levels. **Vn 4.1 onwards**
    Atmos Obs files : Number of observations in observation file.
29. Number of data values in observation file.
30. Time window (minutes) of observations before observation file time.
31. Time window (minutes) of observations after observation file time.
32. Number of AC Observation Types in observation file.

For atmosphere dumps : Dimension is 29. Words 1-28 in the above list are used, Word 29 is not used.

For atmosphere fieldsfiles and ancillary data sets : Dimension is 15 and the following words are used :-

3. Number of different times for which data is present in data set. This is used to determine the periodicity of the data. e.g. 12 for twelve monthly values.
6. Number of points E-W (x direction or 1st dimension)
7. Number of points N-S (y direction or 2nd dimension)
8. Number of vertical levels in model providing or receiving data.
15. Number of different field types present in data.

For atmosphere boundary data sets : Dimension is 15 and the following words are used :-

3. Number of different times for which data is present in data set.
6-9. As in full list. The values refer to the model to receive the data.
15. Number of different field types present in data. (5) ***prior to version 3.3 this information was stored in position 13.***

For **atmosphere observation files** : Dimension is 32.

1-5. Not used.
6-9. As in full list. These values refer to the model being used to assimilate the observations.
10-27 Not used.
28-32 Used as in full list above.

3.3 REAL_CONSTANTS
Location and dimensions defined by words 105-106 of FIXED_LENGTH_HEADER
Separately defined for ocean and atmosphere use
Constants marked (*) are 'passive' - see introduction.

Full list of words reserved for **Atmosphere** use

1. EW (x) grid spacing in degrees
2. NS (y) grid spacing in degrees
3. Latitude of first PTR row in degrees (latitudes in range 90 to -90)
4. Longitude of first PTR point on row in degrees(longitudes in range 0-360)
5. Real latitude of 'pseudo' N pole in degrees
6. Real longitude of 'pseudo' N pole in degrees
7. Not used
8. ATMOS YEAR copied to FIXED_LENGTH_HEADER(28)
9. ATMOS DAY copied to FIXED_LENGTH_HEADER(34)
10. ATMOS HOUR copied to FIXED_LENGTH_HEADER(31)
11. ATMOS MINUTE copied to FIXED_LENGTH_HEADER(32)
12. ATMOS SECOND more precise value to be copied to FIXED_LENGTH_HEADER(33)

13-17 Not used
18. Global mean diabatic flux
19. MASS
20. ENERGY
21. ENERGY_DRIFT

22-38 : Pre Vn 4.0

22. (*)LATITUDE_BAND
23. (*)VERTICAL_DIFFUSION
24. (*)SNOW_MASKING_DEPTH
25. (*)CO_2
26. (*)C
27. (*)KAY_GWAVE
28. (*)NU BASIC
29. VALUE OF REAL MISSING DATA INDICATOR
    - Version 3.2 onwards : -2^{-30}
    - Pre-Version 3.2 : -32768.0
30. ATMOS_STEP (converted from history file integer)
31. (*)FILTERING SAFETY FACTOR
32. (*)CO2START
33. (*)CO2RATE
34. (*)CO2END
35. (*)CLOUD WATER THRESHOLD OVER LAND FOR CONVERSION TO PRECIPITATION
36. (*)CLOUD WATER THRESHOLD OVER SEA FOR CONVERSION TO PRECIPITATION
37. (*)WIND_LIMIT for half timestep dynamics
38. (*)DIV_LIMIT Divergence limit for half timestep dynamics

22-38 : Vn4.0 onwards

22-28. Not used
29. As pre Vn 4.0
30-38. Not used.

In atmosphere dumps : Dimension is 38 for Version 3.3 (36 for Versions 2.8 - 3.2, 34 for pre-version 2.8) dumps. Words 1-38 (1-36, 1-34) in the full list are used.

In **atmosphere fieldsfile, ancillary data sets** and **boundary data sets** :
Dimension is 6 and the first 6 values in the full list above are used.

In atmosphere observation files : Dimension is 34 and words 1-6 and 29 are used as above.

3.4 LEVEL_DEPENDENT_CONSTANTS
Location and dimensions defined by words 110-112 of FIXED_LENGTH_HEADER Separately defined for ocean and atmosphere use

For **atmosphere data sets** except observation files :

The first dimension is the number of model levels and the second dimension refers to :-

Pre Vn4.0

		Type
1.	AK	Real
2.	BK	Real
3.	DELTA_AK	Real
4.	DELTA_BK	Real
5.	THETA_REF	Real
6.	(*) Diffusion coefficients for u and v	Real
7.	(*) Diffusion coefficients for q	Real
8.	Diffusion exponent for θ, u and v	Integer
9.	Diffusion exponent for q	Integer
10.	(*) Divergence damping coeffs. - assimilation	Real
11.	(*) Divergence damping coeffs. - forecast	Real
12.	RHCRIT Critical RH for cloud amount calculations	Real
13.	SOIL_THICKNESS(first DS_LEVELS only *)	Real

4.0 onwards

1.	AK	Real
2.	BK	Real
3.	DELTA_AK	Real
4.	DELTA_BK	Real
5.	THETA_REF	Real
6.	SOIL_THICKNESS(first DS_LEVELS only *)	Real

* DS_LEVELS = No of soil levels, stored in Integer Constants - Word 10

For atmosphere dumps the second dimension is 6 (13 before Vn4.0) corresponding to the full list above.

For atmosphere fieldsfiles, ancillary data sets and boundary data sets the second dimension is 4 and the first 4 values above are used. In the case of FIELDSfiles, these refer to the levels of the model providing the results, not the output levels. In ancillary and boundary data sets they refer to the model levels for which the data are present (model level of field in LOOKUP(33) for ancillary data sets). These constants are not included for ancillary and boundary data sets containing single level data.

For atmosphere observation files this array stores the levels of the observation data for each observation type.

The first dimension will be the observation type with the maximum number of level information plus two and the data is arranged as follows :-

1. Observation level type
   - 0 Single level, no level in data
   - 1 Single level, level included in data.
   - 2 Model levels.
   - 3 Pressure levels.
   - 4 Pressure levels corresponding to layer boundaries.
   - 10 A-value of hybrid levels.
   - 11 B-value of hybrid levels.
2. Number of observation levels or layers.
3- Observation levels. Any levels not used are padded with the real missing data indicator. Note that for N layers there will be N+1 pressure levels corresponding to the layer boundaries.

The second dimension is the number of observation types plus two for the A and B values of the hybrid levels if included.

3.5 ROW_DEPENDENT_CONSTANTS
Location and dimensions defined by words 115-117 of FIXED_LENGTH_HEADER
Separately defined for ocean and atmosphere use
first dimension is over rows
for atmosphere, second dimension refers to :
1. FILTER_WAVE_NUMBER_P_ROWS
2. FILTER_WAVE_NUMBER_U_ROWS
3. Number of E-W sweeps per row used in tracer advection stability calculations.

Used in atmosphere and ocean dumps only.

3.6 COLUMN_DEPENDENT_CONSTANTS
Location and dimensions defined by words 120-122 of FIXED_LENGTH_HEADER
Separately defined for ocean and atmosphere use
first dimension is over columns
No specific atmosphere use at present

3.7 FIELDS_OF_CONSTANTS
Location and dimensions defined by words 125-127 of FIXED_LENGTH_HEADER
Separately defined for ocean and atmosphere use
first dimension is over fields
No specific atmosphere use at present

3.8 EXTRA_CONSTANTS
Location and dimensions defined by words 130-131 of FIXED_LENGTH_HEADER
Separately defined for ocean and atmosphere use. No specific atmosphere use at present

3.9 TEMP_HISTORYFILE
Location and dimension defined by words 135-136 of FIXED_LENGTH_HEADER. Copy of temporary history common block. Used by atmosphere and ocean dumps only. Not used from UM Vn 3.4.

3.10 COMPRESSED_FIELD_INDEX
Locations and dimensions defined starting at word 140 of FIXED_LENGTH_HEADER
See ocean model documentation for details

NOTE: The complexities of this indexing system for allowing this option of data compression, mean that it is unlikely to be adopted for atmosphere use where possibly only a single land sea compression mask might be required. (This scheme is only viable where a separate mask is required for each level). If only simple 0/1 bitmasks are required, these are held in DATA as integer strings of 0/1. The lookup contains an indicator for compressed fields and a pointer to that part of the model DATA area containing the integer string required for the expansion. Further details in section 4.

3.11 LOOKUP table
Location and dimensions defined by words 150-152 of FIXED_LENGTH_HEADER. Full details given in section 4.

3.12 DATA
Array for all data, pointed to by words 160-161 of FIXED_LENGTH_HEADER and decoded LOOKUP. For observation files word 162 contains the maximum length of observational data which will be the observation type with the highest number of data values. The number of observations varies from file to file and this value enables the correct amount of workspace to be allocated in the UM I/0 routines.

NOTE: For atmosphere use: the format of the dump is sufficiently flexible to include any field in any order {although coding consideration imply that multilevel fields (and all soil fields and all veg fields) are contiguous}. A multitude of diagnostic fields of both 1 and 2 dimensions can be included if desired, but strictly a dump should only contain those fields necessary for restarting.

4. The LOOKUP table

For Dumps, FIELDSFILES, Boundary data sets and ancillary data sets,each entry in the lookup table is the 64 word PP header currently used. For Observation files a 128 word header is used see below section 5 for details

Word No.	Name	Contents		
1-45	**INTEGERS**			
1	LBYR	Year	}	Validity time for instantaneous
2	LBMON	Month	}	field. Accumulations/Means :
3	LBDAT	Day of month	}---	start of A/M period. **(*)**
4	LBHR	Hour	}	Start time of partial
5	LBMIN	Minute	}	timeseries if data record is a
6	LBDAY	Day number	}	timeseries
7	LBYRD	Year	}	Data time for instantaneous field.
8	LBMOND	Month	}	Accumulations/Means : End of A/M
9	LBDATD	Day of month	}---	period. **(*)**
10	LBHRD	Hour	}	End time of partial timeseries if
11	LBMIND	Minute	}	data record is a timeseries
12	LBDAYD	Day number	}	

* If using FIELDCOS with OPER=T in namelist, Words 1-6 become the validity time and Words 7-12 become the data time for Accumulations and Means in the converted file.

13	LBTIM	Time indicator
14	LBFT	Forecast period (hours)
15	LBLREC	Length of data record

Dumps : LBLREC is the original number of data points.

PP-files : LBLREC is the number of words needed to hold the information.

Timeseries : LBLREC is the length of the timeseries. Note that number of individual points sampled is LBLREC/LBUSER(3).

For PP-files LBLREC may be machine dependent. For unpacked data it makes no difference, but for packed data it does. On a HP workstation, or a HDS or indeed any 32 bit word machine the value of LBLREC will be twice the value it has on a CRAY, a 64 bit word machine, because it takes twice as many 32 bit words to hold the same amount of data as 64 bit words.

16	LBCODE	Grid type code

For rotated grids: Grid Type = 100 + (grid type for un-rotated grid).

17	LBHEM	Hemisphere indicator

LBHEM is set to 99 for a ‘rim’ data set as used for lateral boundary data sets.

18	LBROW	Number of rows in grid

LBROW is set to the width of the boundary zone for lateral boundary data sets.

19	LBNPT	Number of points per row
20	LBEXT	Length of extra data

21 LBPACK Packing method indicator **Vn 2.8 onwards.**

The LBPACK code describes packing and compression of fields in memory and on disk. The integer value is the combination of a series of codes as follows :-

LBPACK = $N_5N_4N_3N_2N_1$
(eg LBPACK=1210; N_5=0, N_4=1, N_3=2, $N_2$1, N_1=0)

where

N1 Packing
- 0 data not packed.
- 1 data packed using WGDOS method. **(*)**
- 2 data packed using CRAY 32 bit method.
- 3 data compressed using the GRIB method.

N2 Data Compression
- 0 data not compressed.
- 1 data compressed using the N3rd group of compressed field index arrays in the dump.
- 2 data compressed with the N3rd bit mask.

N3 Compression
If N_2=1, then
N_3 is the number of the group of compressed field index arrays used
If N_2=2, then
N_3 = 1 use land mask
N_3 = 2 use sea mask

N4 Number Format
- 0 Native (to the target machine) format
- 1 IBM format
- 2 CRAY format
- 3 IEEE format
- 4 GRIB format
- 5 Standard VAX format

N5 Reserved for future use (=0)

* - Note that the WGDOS packing method would not usually be adopted because it would not allow results to be replicated on a restart

Changes for 4.5
1. N4 : This has been reset from 2 to 3 for fieldsfiles.

22 LBREL Header release number **From Vn 3.3**
23 LBFC Field code
24 LBCFC Second field code
25 LBPROC Processing code
26 LBVC Vertical co-ordinate type
27 LBRVC Co-ordinate type for reference level
28 LEXP Experiment number

29	LBEGIN	Disk address/Start Record Files with well-formed records : Disk address (Word Number) relative to start of file. **From Vn 4.3** Fieldsfile direct access data sets : Start record For all other files including those not set up with well-formed records : set to zero
30	LBNREC	Disk length/No of records. Files with well-formed records : Length on disk which is no. of words. The length includes rounding up to an exact number of sectors. Sector length currently in use on T3E is 512 words. Fieldsfile direct access data sets : No of records For all other files including those not set up with well-formed records : set to zero
31	LBPROJ	Met.O.8 projection number
32	LBTYP	Met.O.8 field type
33	LBLEV	Met.O.8 level code Set to 7777 for a multi-level field, as occurs in lateral boundary data sets.
34	LBRSVD(1)	Reserved for future PP-package use
35	LBRSVD(2)	Reserved for future PP-package use
36	LBRSVD(3)	Reserved for future PP-package use
37	LBRSVD(4)	Reserved for future PP-package use Words 34-37 are reserved for PP-package use and should not be used by programmers.
38	LBSRCE	=1111 to indicate following apply to UM
	If LBSRCE=1111, then Words 39-45 = LBUSER area for UM use	
39	LBUSER(1)	Data Type **Vn 2.8 onwards** 1 - Real field 2 - Integer Field 3 - Logical Field -1 - Real Timeseries -2 - Integer Timeseries -3 - Logical Timeseries **Changes for 4.5** 1. For fieldsfiles, the data type for the Land-Sea mask has been reset from 1 to 3. Note that CUMF will not compare a 4.4 and 4.5 mask.
40	LBUSER(2)	Start address in DATA
41	LBUSER(3)	No of sampling periods for Time Series
42	LBUSER(4)	Stash section number and Item number Held in the form : Section Number*1000 + Item Number
43	LBUSER(5)	Stash Pseudo Dimension Used to specify level eg. Radiation Band or Assimilation Group.

44	LBUSER(6)	Free space for users. Used by external program FIELDCOS to denote length of unpacked field size. Contains number of times FIELDOP has been applied to a dump of fieldsfile. **From Vn 4.3.**
45	LBUSER(7)	Internal Model Number. Vn 4.0 onwards. 1 - Atmosphere 2 - Ocean 3 - Slab 4 - Wave (Vn 4.1 onwards) Pre Vn 4.0 - free space for users.
46-64	**REAL**	
46	BRSVD(1)	Reserved for future PP-package use
47	BRSVD(2)	Reserved for future PP-package use
48	BRSVD(3)	Reserved for future PP-package use
49	BRSVD(4)	Reserved for future PP-package use Words 46-49 are reserved for PP-package use and should not be used by programmers.
50	BDATUM	Datum value
51	BACC	(Packed fields) Packing accuracy
52	BLEV	Level. For hybrid levels, B-value of level.
53	BRLEV	Reference level
54	BHLEV	For hybrid levels, A-value of level.
55	BHRLEV	For hybrid levels, A-value of ref. level.
56	BPLAT	Real latitude of ‘pseudo’ N pole
57	BPLON	Real longitude of ‘pseudo’ N pole The latitude and longitude of the ’pseudo’ north pole need only be coded for rotated grids.
58	BGOR	Grid orientation
59	BZY	Zeroth latitude
60	BDY	Latitude interval
61	BZX	Zeroth Longitude
62	BDX	Longitude interval
63	BMDI	Missing data indicator
64	BMKS	M.K.S scaling factor

5. Atmosphere Observation Files

An observation file has the following components :-

Fixed Length Header, Integer Constants, Real Constants, Level dependent constants, Lookup table and a Data section.

The lookup table for an observation file is different from the other data sets as its first length is 128 compared to 64 in the other data sets. The first 64 words is the PP Header as described in section 4 and Words 65-128 contains extra information on the observation data.

Word 65-75 is reserved for values specific to each observation type. Those in use are

65	AC Observation Type Number
66	Number of Observations
67	Number of Data Values per observation

68-75 reserved for future use

76-128 reserved for the data descriptors to describe the data contents and is used by the ocean observation files and maybe in the atmosphere observation files in future.

The total number of data values for an observation type is Word 66 multiplied by

Word 67. The total number of observations and data values in an observation file are given by Word 28 and 29 respectively of the Integer Constants array.

Observation Data

The observation data is contiguous by observation type and is ordered according to the headers in the lookup table. For each observation type, the data is contiguous by data value. The data is stored in 32 bits on disk and 64 bits in memory. Full details on the data contents for each observation type is given in UMDP P3, Chapter 14.

5.1 Observation files for VAR and OPS

Details on the file formats for the new schemes can be found at :-

VAR :
http://fr0400:80/~frva/var0/documentation/views/Var_Comp/latest/Doc/VTDP8.html

OPS : http://fr1700/~opsrc/ops0/documentation/latest/OTDP3.html

6. Ocean dumps

6.1 MPP and non-MPP platforms.

Ocean dumps are set up differently depending on the type of platform the ocean model is to be run on. For non-MPP model runs (including a single PE non-mpp run on mpp platforms), an ocean dump may contain compressed model fields. For MPP model runs, the ocean model can only handle uncompressed fields. Any dumps containing compressed fields must be converted to contain uncompressed fields for a MPP model run.

6.2 Contents

Many components of the ocean dump are adequately described in sections 1-5 (atmospheric) above. Those items which differ in content are mentioned in the following description.

FIXED_LENGTH_HEADER:
Words 1 to 256: as atmospheric except ;

Pre Vn 4.4

Word 11: indicator for boundary conditions: cyclic east-west (1 if cyclic, 0 otherwise) + symmetric at southern row (2 if symmetric, 0 otherwise).
Word 12: Indicator for type of model dynamics used: barotropic stream function (0), Killworth code for free surface (1), no barotropic currents (2).

INT_CONST: (Length - 29 words)
Word 1: number of tracer timesteps since start of run (ITT)
Words 2-7: as atmospheric.
Word 8 : number of vertical levels.
Words 9-10: spare
Word 11: number of points in a compressed array (number of sea points)
Word 12: as atmospheric.
Word 13: spare
Word 14: number of ’passive’ tracers (including salinity) (NT)
Word 15: as atmospheric.
Words 16 to 20: spare
Word 21: as atmospheric.
Words 22-29: spare

REAL_CONST: (Length - 36 words)
Words 1 to 6: as atmospheric.
Word 7: latitude (in degrees) of Southern wall (1st u,v row) (SWLDEG)
Word 8: longitude (in degrees) of Western boundary (1st u,v column) (WEDGEDEG)

Words 9 to 19: spare
Word 20: total kinetic energy (EKTOT)
Word 21: external energy change (PLICEX)
Word 22: internal energy change (PLICIN)
Word 23: work done by buoyancy (BUOY)
Words 24 to 28: spare
Word 29: as atmospheric.
Word 30: area of surface of model basin (AREA)
Word 31: total volume of model basin (VOLUME)
Words 32 to 36: spare

LEV_DEP_CONST:
Words 1 to KM: thicknesses (in metres) of model layers

ROW_DEP_CONST:
Words 1 to JM: north-south grid spacing (in degrees latitude)

COL_DEP_CONST:
Words 1 to IM: east-west grid spacing (in degrees longitude)

FIELDS_CONST:
Words 1 to IM*JM: FKMP(IM,JM)=number of ocean levels at tracer grid points

EXTRA_CONST:
Word 1: NISLE=number of islands
Word 2: ISEG(1)=number of segments for first island
Words 3 to ISEG(1)+2: ISIS(1,N),N=1,ISEG(1)=start columns of island segments for first island
Words ISEG(1)+3 to 2*ISEG(1)+2: IEIS(1,N),N=1,ISEG(1)=end columns of island segments for first island
Words 2*ISEG(1)+3 to 3*ISEG(1)+2: JSIS(1,N),N=1,ISEG(1)=start rows of island segments for first island
Words 3*ISEG(1)+3 to 4*ISEG(1)+2: JEIS(1,N),N=1,ISEG(1)=end rows of island segments for first island
Word 4*ISEG(1)+3: ISEG(2)=number of segments for second island
Words 4*ISEG(1)+4 to ISEG(2)+4*ISEG(1)+3: ISIS(2,N),N=1,ISEG(2)=start columns of island segments for second island

etc

TEMP_HIST_FILE:
Words 1 to 2048: copy of temporary history common block

COMP_FIELD_INDEX:
Words 1 to n_seg: elements of the array index_comp(n_seg)
Words n_seg+1 to 2*n_seg: elements of the array index_exp(n_seg)
Words 2*n_seg+1 to 2*n_seg+rows*levels: elements of the array index(rows,levels)

Set all compression indices to 1 for dumps with uncompressed fields.

LOOKUP:
contains 64-word pp-headers (as atmospheric) for contiguous fields in DATA

DATA:
Fields are stored in the order of their STASH codes.
Compressed **or** uncompressed forms of:
T (field code 601, STASH code 101)
S (602,102)
U (701,121)
V (702,122)
are followed by:
stream-function fields at present timestep (611,130)
stream-function fields at previous timestep (612,131)
stream-function tendencies at present timestep (613,132)
stream-function tendencies at previous timestep (614,133)
zonal windstress (721,150)
meridional windstress (722,151)
wind mixing energy (627,152)
solar heat flux (625,161)
surface heat flux (626,162)
precipitation minus evaporation (629,165)

6.3 Data compression technique for ocean variables.

This compression technique may only be used in ocean dumps when running non-MPP **models.**

Data compression of three-dimensional ocean variables is designed to permit ready access to the fields either as horizontal layers or as zonal slices. Data are stored in the dump in an array compressed (n_elements) and are written to the array expanded (nx, ny, nz). Data are grouped in segments of contiguous data, of which there are n_seg.

Unpacking of the data is achieved through the algorithm:

```
set level to the level to be retrieved
set row to the row to be retrieved

if ( row = ny ) then
    row_next = 1
    level_next = level + 1
else
    row_next = row + 1
    level_next = level
end if

if ( level_next > nz ) then
    num_seg = n_seg - index_to_rows(row,level) + 1
else
    num_seg = index_to_rows(row_next,level_next) - index_to_rows(row,level)
end if

for seg = 1 to num_seg

    seg_pos = index(row,level) + seg -1
    if ( seg_pos < n_seg ) then
        len_seg = index_compressed(seg_pos+1) - index_compressed(seg_pos)
    else
        len_seg = n_elements - index_compressed(seg) + 1
    end if

    for count = 1 to len_seg
        ipoint_exp= index_expanded(seg_pos) + count - 1
        ipoint_com = index_compressed(seg_pos) + count  - 1
        x_pos = ipoint_exp - (level -1)*number_of_points_in_layer -
                             (row - 1)*number_of_points_in_row
        expanded(x_pos, row, level) = compressed(ipoint_com)
    end
end
```

7. Wave Model Dumps.

7.1 Contents.

FIXED_LENGTH_HEADER:

Words 1 to 162: as atmospheric (but note below).

2. Indicator for Sub-Model. Set to 4.
3. Indicator for vert coord type. Set to 6.

INTEGER CONSTANTS Dimension = 15

Words 1-7 as atmosphere then:

8. Number of levels = frequencies, NFREQ
9. Number of directions NTHETA
10. Number of sea DATA POINTS
11-14 spare
15. Number of different field types in dump

REAL_CONSTANTS Dimension = 29

Words 1-12 and 29 as atmosphere then

13 THETA at centre of direction bin 1 (radians)
14 Δ(THETA) (positive ANTICLOCKWISE - in radians)
15 C0 (WAM frequency spacing constant) =1.1

16-28 spare

LEVEL_DEPENDENT_CONSTANTS Second dimension is 1

Frequency dependent - values of Fcentre for each bin

Fcentre 1 to nfreq (lev_dep_consts(i,1) i=1,nfreq)

ROW_DEPENDENT_CONSTANTS - not used.
COLUMN_DEPENDENT_CONSTANTS - not used.
FIELDS_OF_CONSTANTS - not used.
EXTRA_CONSTANTS - not used.
TEMP_HISTORYFILE - not used.
COMPRESSED_FIELD_INDEX - not used.

LOOKUP table

1-45 INTEGER - as for atmosphere except

26 LBVC Vertical co-ordinate type FREQUENCY for waves (set value =2)

44 LBUSER(6) for wave energy WAVE MODEL DIRECTION BIN NUMBER

46-64 REAL - as for atmosphere except

52 BLEV for WAVE ENERGY: frequency value of bin
54 BHLEV for WAVE ENERGY: direction value of bin

DATA
Array for all data, pointed to by words 160-161 of FIXED_LENGTH_HEADER and decoded LOOKUP.

Fields are stored in the order of their STASH codes. For wave model, store in order of usefulness - then STASH codes should be allocated in this order.

All data except LS mask to be stored at sea data points only. Compress using LAND_SEA MASK

ORDER OF FIELDS IN DUMP

a) ENERGY for direction=1,ntheta (for each level = 1-nfreq). PPCODE 351

b) Land Sea Mask : LOGICAL LAND=T PPCODE 38

c) Water depth (m) PPCODE 2

d) Winds NWP 10m (or level1) U component PPCODE 56
V component PPCODE 57

e) Winds U* from NWP if coupled / from WAM if not coupled.
U component PPCODE 61
V component PPCODE 62

f) ICE concentration PPCODE 37

g) Wave stress τ_w at this timestep X component PPCODE 364

h) Wave stress τ_w at this timestep Y component PPCODE 365

END of list