BINARY REPRESENTATION OF DATA IN THE UK MET. OFFICE

P.J. Smith and W.H. Lyne
Meteorological Office
Bracknell, UK

1. INTRODUCTION

The Meteorological Office uses two binary representations for grid-point data, FM 92 GRIB and an internally developed format which, for the purpose of this paper, shall be called MOBNRY. Both code forms have their advantages and disadvantages and these, together with their applications, are described below.

The use of BUFR within the Office is described in another presentation at this meeting.

2. APPLICATIONS

2.1 FM 92 GRIB

This code form is used for the dissemination of numerical forecast products over the GTS. A comprehensive range of analysis and forecast fields is available out to T+72 from the global model, with a more limited set to T+120. Most of these products are on a 5 deg. longitude by 2.5 deg. latitude grid, with each individual bulletin covering a quadrant of the globe.

A similar range of products is available from the regional model to T+36, but on a finer 1.25 by 1.25 degree grid.

A summary of these products and grids is given in table 1, with further information available from the Forecast Products Branch of the Meteorological Office. Details of the global grids will be issued as a supplement to Volume B of WMO publication No. 9.

2.2 MOBNRY This code form was developed internally by the Meteorological Office for archiving analyses and forecasts from the operational models, both global and regional. Data are archived at the model grid points and on standard pressure levels. This packed archive was introduced for the atmospheric models on the 18th December 1985 and for forecast wave model data on 1st May 1987. An earlier archive exists for a limited number of atmospheric model fields from May 1983, and hindcast wave data have been archived from 1978 but in a different format.

Most standard surface and upper air fields are archived as analyses, with up to thirteen pressure levels for upper air data. A more limited set of pressure levels is used for archiving forecast data. Vertical velocity is also archived, together with surface fluxes of sensible and latent heat, surface stress and accumulated precipitation. Some radiation and cloud data are archived, but not all are currently available from the global model. Further details may be obtained from the Forecast Products Branch.

3. COMPARISON OF CODE FORMS

FM 92 GRIB is documented in the WMO Manual on Codes and its description will not be repeated here. The MOBNRY code is similar in its method of packing the data, but differs in that fields are grouped together by run type.

Thus a MOBNRY archive for an individual run of a model consists of three sections:

HEADER SECTION:- 10 words containing data time and details of run.

LOOKUP TABLE:- 21 half-words for each field containing its specification and required precision.

DATA BLOCKS:- Blocks of 20,000 words containing the packed fields of data.

(Note that a word consists of four bytes or octets.)

Within the DATA BLOCKS each field is stored a row at a time with a one word header preceding each row of data constructed as follows:
Octets 1,2 & 3:- Scaled minimum for row.
Octet 4:- No. of bits in which each data item is held.
If the leading bit of octet 4 is set then the data field is sparse and the first 6 words of the data field contains a bit map (c.f. the Bit Map Block of GRIB). The data themselves are coded as non-negative scaled differences from the scaled minimum for the row.

A data item Y is unpacked according to the formula:

$$Y = (R + X)*2^{P}$$

where R is the scaled minimum (a 24 bit integer), X the coded difference, and P the precision. This should be compared with the GRIB representation:

$$Y = R + (X)*2^{P}$$

where R is now an unscaled reference value coded as a 32 bit floating point number. (P is often referred to as the scale factor rather than the precision in descriptions of GRIB).

4. DISCUSSION

GRIB enjoys the advantages of being an approved WMO code form in which the contents of the Product Definition Block and the Grid Definition Block enable the data to be self defining. Apart from this, the actual methods of packing the data in the two code forms is very similar, the main difference being the choice of a separate reference value for each row of data in MOBNRY.

This difference can lead to significant increases in packing density, especially for those fields which have a large

latitudinal variation such as stratospheric height fields. A more appropriate reference value enables the data to be packed in a smaller number of bits to achieve the required precision. The same advantage accrues for those rows which have only a limited variation along the rows, the main example being near polar rows for models with latitude - longitude grids.

When packing IBM 32-bit floating point numbers, a compression to 25% of the original size is typical with MOBNRY whereas 33% is more representative of GRIB.

The current specification of FM 92 GRIB mentions that complex packing of grid-point data is to be developed. It is suggested that one form of this could be developed incorporating the row by row nature of MOBNRY, thus achieving a further significant increase in compression.

During 1990 the introduction of higher resolution forecast models is expected to result in an almost threefold increase in output data. One of the packing methods described here will be used to compress these data and reduce the demands on on-line storage.

TABLE 1 Products currently available in GRIB

Fields from both global and regional models

Surface: Sea level pressure, 10m wind components

Upper air: Height, wind components, temperature at 850,700,500,400,300,250,200,150,100 hPa

Relative humidity at 850,700,500 hPa

Additional fields available from regional model

Surface: 1.5m temperature, accumulated precipitation

Upper air: Wind comps., temp., rel. humidity at 950 hPa

Additional fields available from global model

Upper air: Height, wind components, temperature at 50 hPa

Maximum wind and pressure level

Tropopause pressure and temperature

Regional model resolution: 1.25 x 1.25 deg. 32.5-75N, 70W-35E

Global model resolution: 2.5 deg. latitude x 5 deg. longitude

Data times: 0000 and 1200 UTC

Forecast times

Regional model: T+0 to 36 (3 hourly steps)

Global model: T+0 to 48 (6 hourly steps),T+60

T+72,96,120 (Sea level pressure and 500 hPa height only)