LPM Data
Basic Guide

Document Version

Version 1.0 – Original
Version 1.1 – Added Output Path option and use of quotation marks to DGRange
Version 1.21 – Added DGMod status bit 5 update and reference to daily data

Associated Documents

1. Degum Technical Description
2. Degum Installation and Operation Guide

Author

David Judge, PSD BAS

British
Antarctic Survey

NATURAL ENVIRONMENT RESEARCH COUNCIL

Table of Contents

1. Introduction

This document details the location and file formats of the LPM data sets available from the BAS data repository. It is intended to act as a basic guide and reference source for scientists wishing to access the LPM data at BAS. This document is relevant for LPM data recorded from 2002 onwards, some LPM data exists for 2001 in the BAS DABS system.

This document also provides information about the applications that have been used to process the data, a format conversion utility and the location of more detailed documentation should it be required.

2. Station and Data File Naming Convention

Each LPM station is assigned a unique name that is a concatenation of the letter M, for magnetometer, followed by two numbers (separated by a hyphen) representing the station's location to the nearest whole degree of latitude south and longitude east. Westerly longitudes are always converted to easterly values.

For example, if an LPM station was located at position 82.775 S and 13.059 W, its station name would be:

M83-347

The file name for a data file acquired by a station in a particular year would begin with the station name followed by the year number, again separated by a hyphen. The year corresponds to the operating year of the station, it may begin or end before or after the calendar year. So, for data acquired in 2004 by the station in the above example, the file name would begin:

M83-347-2004

3. Directory Structure

From the root LPM data directory, *lpm*, the generic directory structure for an individual station is:

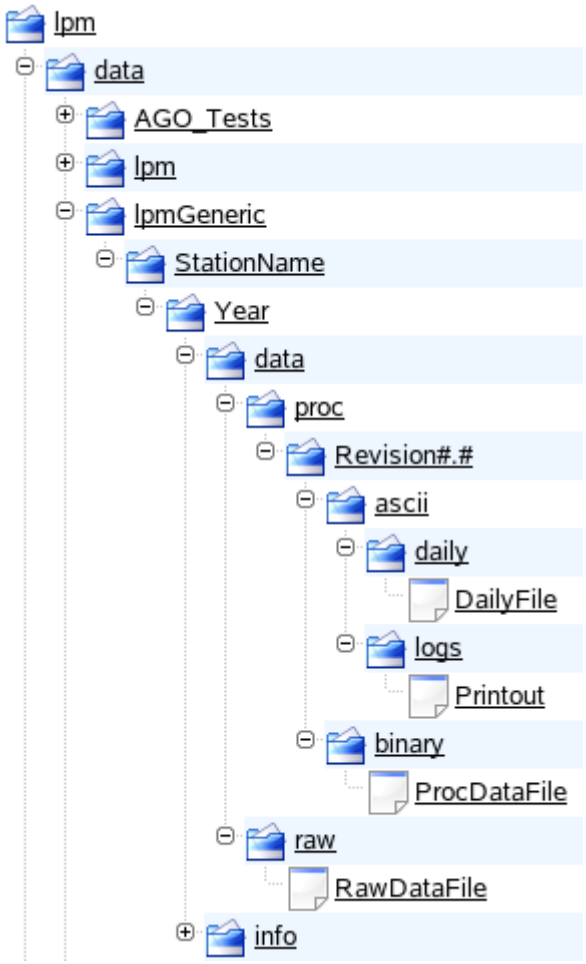

Fig. 1 Generic Directory Structure

where,

StationName and *Year* are the acquisition station name and year respectively, e.g. M83-347 and 2004.
Revision#.# is the revision of the processing software used to process the data.
DailyFile represents the daily ascii files
Printout represents the output files generated during processing
ProcDataFile represents the processed data files
RawDataFile represents the raw LPM data file

Here is an example directory tree, expanded for station M84-336, year 2004:

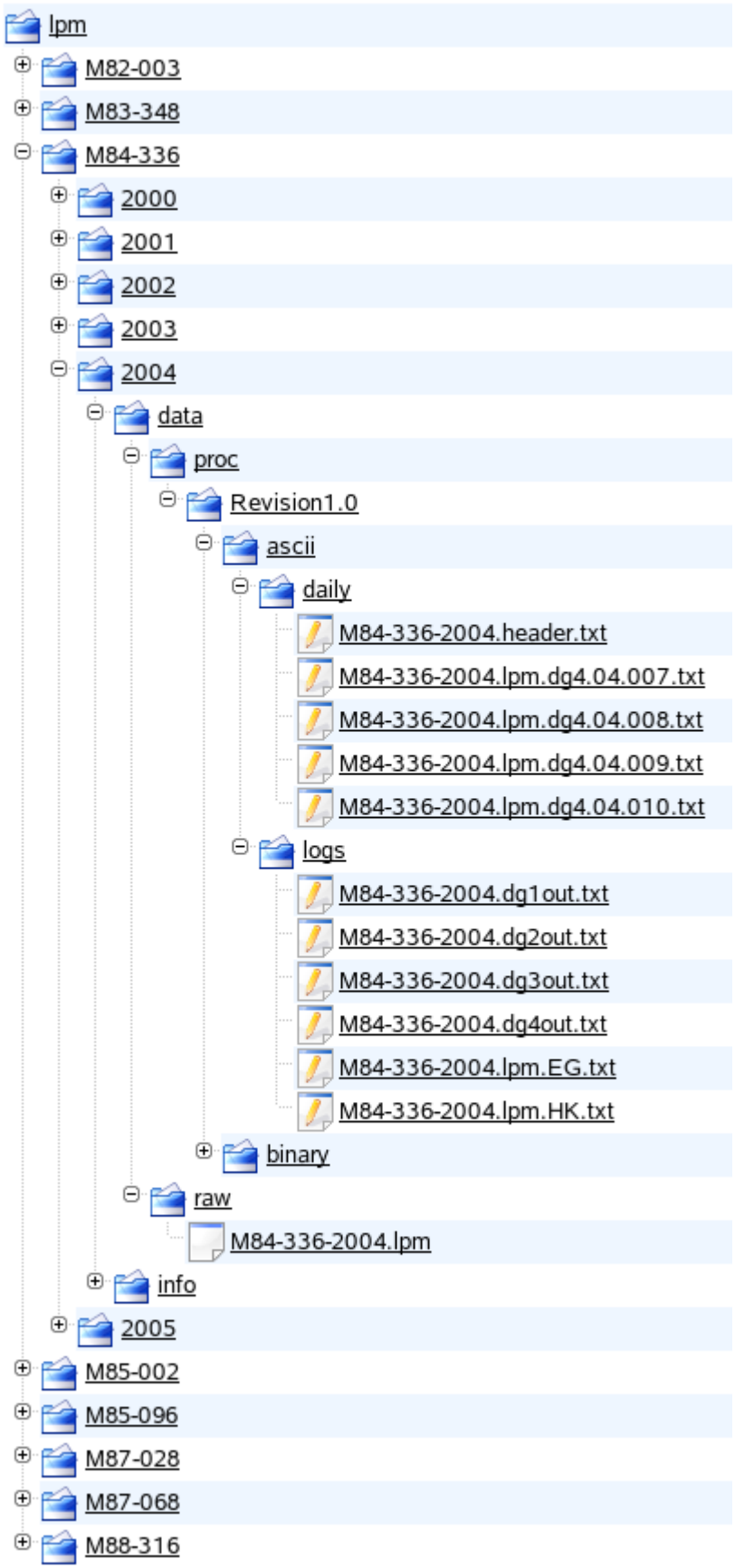

Fig. 2 Directory Tree

4. File Types

The *raw* folder contains the original raw LPM data file as recovered from the field. This data must be processed to be scientifically usable.

The *ascii/daily* folder contains all the data as daily text files (format described in section 9 of this document). Each daily file's name is of the form:

MLL-LLL-YYYY.lpm.dg4.YY.DDD.txt

Where,
MLL-LLL-YYYY.lpm.dg4 is the name of the source binary file from which the daily ascii file was generated.
YY.DDD is the two-digit year number and three-digit day number of the single day's data it contains.

e.g. M83-348-2003.lpm.dg4.03.140.txt

It is possible that up to three source files will supply a station's data for a single year. Hence it is possible for a daily folder to contain files with three different source file sections in their file name.

e.g.

M83-348-2002.lpm.dg4.03.015.txt
M83-348-2003.lpm.dg4.03.140.txt
M83-348-2004.lpm.dg4.03.352.txt

Also, a single day's data may be contained in two separate daily files if it was generated from two source files.

e.g.

M83-348-2003.lpm.dg4.03.341.txt
M83-348-2004.lpm.dg4.03.341.txt

No Degum processing header files are stored within the daily ascii files. Each daily folder does, however, contain an example of the headers from all the source files that contributed to that particular daily folder. Header information is stored in files of the format:

MLL-LLL-YYYY.header.txt

Where MLL-LLL-YYYY is derived from the name of the source binary file.

The *ascii/logs* folder contains all the text printout files produced during processing. Files ending in dg#out.txt are the processing printout files. Files ending EG.txt or HK.txt contain non-magnetometer information extracted from the raw LPM data files during processing. Further details and the format description for the EG and HK files can be found in the associated document, *Degum_Technical_Description.doc*

The *binary* folder contains the processed magnetometer data in binary format. The file ending .dg4 is the final processed file, all other files are intermediate files created during previous processing stages. Further details can be found in the associated document, *Degum_Technical_Description.doc*

5. LPM Instrumentation Versions

Scientists should be aware that not all of the data available here have been acquired by standard instruments or are operated by BAS personnel.

Stations M65-297, M66-294 and M67-292-2002 / 2003 were operated by BAS but utilised a non standard sensitivity and sampling rate.

Stations M68-041, M69-041, M70-039, M70-044, M74-043 and M77-040 are operated by the Japanese National Institute of Polar Research.

These data should be used with the understanding that they may be non-standard. Should further information be required about these stations, the request should be directed to the appropriate organisation.

6. Binary Data Format

The final processed data, the .dg4 file, is in binary format. A brief description of the data format is given here, please refer to the associated documents for a full description.

The data format consists of a header, of no fixed size, containing a record of the processing history followed by a single array of all the data samples. Each sample is stored in a fixed format of 30 bytes in total.

The first four bytes in the file represent a 32-bit integer giving the number of character strings that form the processing history header. Each string is of indeterminate length but is terminated by a new line character (\n).

The sample structure contains, in order, Time (8 bytes), Uncertainty in time (2 bytes), Status (2 bytes), Data (12 bytes) and Accumulated Processing Changes (6 bytes)

Structure	Variable	Size(bytes)	Description
.dg4	History Sample(1)...Sample(N)	[Variable]	Final processed LPM data
History	NumStrings	4	Number of history strings
	String(1)\n String(2)\n ... String(NumStrings)\n	[Variable]	History character string terminated by '\n' character
Sample	Time	8	UT Time encoded to seconds * 10
	Uncertainty	2	Uncertainty in time (seconds * 10)
	Status	2	Status flags
	Data	12	XYZ components of magnetometer data. 4 bytes each in (nT * 10) Units
	Delta	6	Accumulated changes. 2 bytes for each component each in (nT * 10) Units

Table 1 .dg4 file format description

The sample time is stored as a 64-bit integer. It represents the time, in seconds, since 1/1/2000, at a resolution to 0.1 seconds so has been scaled by 10.

The time uncertainty is stored as a 16-bit unsigned integer, again scaled by 10 to give a resolution of 0.1 seconds.

The two status bytes store additional information about the sample as 16 bit fields as detailed below:

Bit	Description
0	Sample acquired during a Normal loop
1	Sample acquired during a GPS Fixing loop
2	A GPS transition sample (first sample acquired after GPS switched on or off)
3	Sample required an ADC Zero fix during processing
4	Sample required a Despike fix during processing
5	One or more components have been replaced post-processing
6	Sample is within an anomalously short block
7	Sample is within a block affected by a data gap
8	Set if the uncertainty in time overflows 16 bits
9	Set if the accumulated changes overflow 16 bits
10	Unassigned
11	Unassigned
12	Unassigned
13	Unassigned
14	Unassigned
15	Unassigned

Table 2 Status bit description

Any data with any combinations of status bits set is considered good, although data with status bits 6 and 7 set should not be used without an understanding of the implications. The magnetometer data is stored in 12 bytes, a 4 byte signed integer for each component in H, D, Z order. The data units are nT at a resolution to 0.1 nT so have been scaled by 10. The data have been rotated on to the standard orthogonal reference frame where the H component is in the direction towards the magnetic North Pole, the D component is positive in the eastward direction and the Z component is in the vertical direction (negative in the southern hemisphere).

The accumulated changes between the raw recorded field data and the final processed data for each sample are stored in the final 6 bytes, 2 byte signed integers for each component. Units are nT at a resolution to 0.1 nT so have been scaled by 10.

7. Replaced Data

Due to instrument malfunction, some of the data acquired is erroneous and is not considered to be scientifically usable. This data has been erased from the processed binary files (and hence the derived daily files) and replaced with a dummy value of 99999.9 nT.

To date, the following data contains, to some extent, replaced samples:

Station	Year	Replaced Data
M67-292	2004	All Z samples and H, D samples from 04/11/04 to 13/02/05
M67-292	2005	All H, D and Z samples
M88-316	2005	All H and D samples

Table 3 Replaced erroneous data

The original data is still available, in binary form. It is located in the *binary* folder for the above stations as files ending in *.dg4.org*

8. Conversion to Ascii Format

A utility, *DGRange*, that can be used to convert all, or a portion, of a binary data file to ascii format is provided as part of the suite of processing applications (*Degum*). The source code can be downloaded from the BAS website.

A brief description of how to use the *DGRange* utility to obtain an ascii representation of the data follows, please refer to the associated documentation for a full description of how to install, compile and use the *Degum* suite of processing applications and utilities.

DGRange is a non-processing utility application that allows the user to extract whole-day portions of data from a *Degum* file and export it to a new file. Various options and file formats may be used but only the conversion of .dg4 files to ascii format will be discussed here.

Optional parameters are provided to specify the ascii file field delimiting character and if the export of the file processing history is required.

The utility can also be used to just report the time range of a supplied Degum file.

Command Line Parameters
DGRange usage is:

```
DGRange [-t['ch']][-h['ch']][-s][-pOutput Path][-rDay Month Year Day Month Year]] DGFile
```

where *DGFile* is the name of the input file.

The options are:

`-t['ch']` Export in text format.
The number and content of fields will vary with the type of input file (see below)
The optional *ch* character will be used as the field delimiter. Default is ' ' (space)

`-h['ch']` Export processing history information.
This option only applies if exporting in text format or reporting the input file range
The optional *ch* character will be used as the comment character. Default is ' ' (space)

`-s` Split a multi-day export range into separate output files.

`-pOutput Path`
Supply alternative path for the output file. Output files will be placed in the input directory if this option is not supplied. Do not use the '~' character for /home paths.

`-rDay Month Year Day Month Year`
The range of data to export. If not specified, the whole input file will be exported.
First *Day*, *Month* and *Year* (DMY) specify the start date, the second triplet specify the end date, both inclusive
Only whole days can be exported. Duplicate DMY triplets will export a single day's data
DMY parameters are numerical and integers. Years are 4 digit.
All 6 DMY values must be specified

A command with no parameters, except the input file name, will report the range of the input file only.

If the `-t` option is used without the `-r` option then the whole input file will be exported as a text file.

For some parameters, such as single characters or path names containing spaces, it may be necessary to wrap the string or character in quotation characters. Single or double quotes should be used – depending on operating system – but in general, use single quotes for linux/unix systems and double quotes for Windows/DOS systems.

Examples

`DGRange M66-294.dg4`	Reports the range of file *M66-294.dg4*
`DGRange -t -r14 3 2003 15 3 2003 M84-336.dg4`	Exports data from 14/3/03 to 15/3/03 of *M84-336.dg4* file in text format with space delimiter and no processing history
`DGRange -t, -h; -r1 6 2002 1 6 2002 M78-337.dg4`	Exports a comma delimited text file with colon commented processing history for data from the single day 1/6/02 of file *M78-337.dg4*
`DGRange -t -s -r1 6 2002 3 6 2002 M78-337.dg4`	Exports, to separate daily text files, data from each day between 1/6/02 and 3/6/02 of file *M78-337.dg4*
`DGRange -t –p/home/lpm -r1 6 2002 3 6 2002 M78-337.dg4`	Exports data from 1/6/2002 to 3/6/2002 of *M78-337.dg4* in text format. The output is directed to the folder: */home/lpm*
`DGRange -t –p"C:\lpm data" -r1 6 2002 3 6 2002 M78-337.dg4`	Exports data from 1/6/2002 to 3/6/2002 of *M78-337.dg4* in text format. The output is directed to the folder: *C:\lpm data*

9. Ascii Data Format

The ascii output format will contain, if requested, the processing history as a series of character strings each starting with the comment character, if specified, followed by multiple rows of data fields – one row per sample. The content of each field is summarised in Table 3 below:

Field #	Description
1	UT Time (seconds since 00:00 01/01/2000)
2	Seconds since start of first sample in file
3	Year
4	Month
5	Day
6	Hour
7	Minute
8	Second
9	Uncertainty in time (seconds)
10	Status (see table 2 in section 6)
11	Data H component (nT units)
12	Data D component (nT units)
13	Data Z component (nT units)
14	Delta H (nT units)
15	Delta D (nT units)
16	Delta Z (nT units)

Table 3 DGRange ascii output format description