

INTERMAGNET Archive Format (IAF)

INTERMAGNET has published CDs and DVDs of geomagnetic observatory data since 1991. The CDs and DVDs contain a variety of metadata, including contact information and quality control records. The geomagnetic data on the CDs and DVDs is held in INTERMAGNET archive format. This format holds minute, hourly and daily mean values as well as K indices.

The data are coded as 32-bit (long integer) binary words, with 5888 words comprising a day- long record. Each file contains one month of day records (so files are variable length, from 28 to 31 records). Each day of data has a header and data section, the data being subdivided into minute means, hourly means, daily means and a set of K-indices.

Words 1 to 16 comprise the header section containing a mixture of text and numeric fields, including a 3-letter observatory identification (ID) code, the year concatenated with the day of the year, co-latitude, longitude, elevation, orientation, originating organization, a D-conversion factor, data quality, and instrumentation. The D-conversion factor is a fixed value used only in the graphics portion of the access software to allow Declination to be plotted in minutes of arc and equivalent nanoteslas (nT). It is given as $H/3438*10000$, where H is the annual mean value of the horizontal intensity. When XYZF or XYZG orientation is used the D-conversion factor should be set to 10000. ASCII values, such as the observatory ID and orientation, are also stored as 32-bit words, but are coded as the hexadecimal byte-string corresponding to the ASCII string. For example, the string "HDZF" is coded as the sequence "48 44 5A 46". Where a string is shorter than four bytes, it is padded to the **left** with spaces. For example, the string "ESK" is coded as the sequence "20 45 53 4B".

Words 17 to 5776 contain the minute mean values of the 4 geomagnetic elements (successively H,D,Z,F/G or X,Y,Z,F/G) for the day. The values are stored in tenth-units with an implied decimal point. Thus, an H value of 21305.6 is stored (in tenth-nT) as 213056 with a decimal point implied between the last and next-to-last digits.

Words 5777 to 5888 hold the hourly and daily mean values (the order of the elements and encoding method is the same as the minute mean data), K-indices and a small space for each contributing institution to use as they wish.

Missing data for minute, hour, and day values are stored as "999999". Missing K-index values are stored as "999".

Versions of INTERMAGNET Archive Format

Three versions of this format have been used:

Version Number	Years of data using	Description	Words affected by
----------------	---------------------	-------------	-------------------

	this version		version change
1.0	1991 - 2007	This is the original description of the format. It was not designated version 1.0 until 2007. Minor undocumented changes were made to how the header was used over the lifetime of this version.	N/A
1.1	2008	Addition of the publication date and encoding of the format version number in the header.	14 and 15
2.0	2009 onwards	Introduction of delta-F	6, 4337 to 5776, 5849 to 5872 and 5876

Detailed description of INTERMAGNET archive format

INTERMAGNET Archive Format-32 Bit Words:			
HEADER	Word	Description	Notes
	1	Station ID	eg. BOU
	2	Year Julian Day	eg. 1989001 ; January 1, 1989
	3	Co-Latitude	$(90^\circ - \text{Latitude}) * 1000$
	4	Longitude	East Longitude * 1000
	5	Elevation	elevation in metres above sea level
	6	Orientation (of the data in this file)	Valid values for version 1.0 and 1.1 are HDZF or XYZF. Valid values for version 2.0 are HDZG or XYZG – see note 1
	7	Origin	eg. USGS, GSC, BGS, EOPG etc.
	8	D Conversion	$H/3438*10000$ where H=annual mean of H For orientations of XYZF or XYZG set to 10000.
	9	Data Quality	IMAG
	10	Instrumentation	eg. RC (Ring Core), LC (Linear Core), etc.
	11	K-9 value in nT	e.g. 750
	12	Sample rate of original data, in ms	e.g. 125
	13	Sensor orientation	e.g. XYZF, HDZF, UVZ, DIF, etc.

		indicating which elements are actually measured	
	14	Publication date (version 1.1 onwards)	Only the year and month are held, in the form YYMM, e.g. "0806" (June 2008). This field is not used in version 1.0.
	15	Format version (version 1.1 onwards)	The left-most byte of the word is used to indicate the version number of the format. Although this field was not present in version 1.0, as a reserved field it should have been set to 0. Valid values for this field are: <ul style="list-style-type: none"> ▪ 0 – Format version 1.0 ▪ 1 – Format version 1.1 ▪ 2 – Format version 2.0
	16	RESERVED FOR FUTURE USE	
MINUTE MEAN DATA	17	1440 minute mean values - element 1 (one day)	
	...		
	1456		
	1457	1440 minute mean values - element 2 (one day)	
	...		
	2896		
	2897	1440 minute mean values - element 3 (one day)	
	...		
	4336		
	4337	1440 minute mean values - element 4 (one day) – see note 1	
...			
5776			
HOURLY MEAN DATA	5777	24 mean hourly values - element 1 (one day)	
	...		
	5800	24 mean hourly values - element 2 (one day)	
	5801		
	...		
	5824	24 mean hourly values - element 3 (one day)	
	5825		
...			

	5848	
	5849	24 mean hourly values - element 4 (one day) – see note 1
	...	
	5872	
DAILY MEAN DATA	5873	1 daily mean value - element 1
	5874	1 daily mean value - element 2
	5875	1 daily mean value - element 3
	5876	1 daily mean value - element 4 – see note 1
K INDICES	5877	8 K Indices
	...	
	5884	Reserved for each contributing institution
	5885	
	...	
	5888	

Note 1 – How Total Field is recorded

In version 1.0 the fourth element is a total field value from either a scalar (independent) instrument or a total field value calculated from the main observatory instrument. INTERMAGNET has a list of which observatories supplied which type of total field value between 1991 and 2007 – this list is available on request.

In version 1.1 the fourth element is a total field value from a scalar (independent) instrument. Observatories not recording an independent total field value must set this element to the missing data value.

From version 2.0 the way data is encoded into the fourth element has changed significantly. The fourth element now holds ΔF , which is defined as

$$\Delta F = F(v) - F(s).$$

Where $F(v)$ represents the total field value calculated from the main observatory instrument ('vector F ') and $F(s)$ represents the total field from an independent instrument ('scalar F '). Both $F(v)$ and $F(s)$ must be corrected to the location in the observatory where geomagnetic absolute observations are made. When $F(s)$ is missing or both $F(s)$ and $F(v)$ are missing, set ΔF to the missing value. When $F(v)$ only is missing, set ΔF to $-F(s)$.

- Minute mean values (words 4337 to 5776), record ΔF as described above.
- Hourly mean values (words 5849 to 5872) and daily mean values (word 5876), record a missing value even if data is available – this is done because the 4th component in the data is a quality check for minute mean data – this quality check is meaningless for hourly means and daily means.

From version 2.0, word 6 should contain either "HDZG" or "XYZG" – the 'G' in the fourth element being the code for ΔF .

Storage Requirements

Each 1-day record requires 23,552 bytes, so a month-file for January would require 730,112 bytes of storage. A year of observatory data requires almost 8.6 Megabytes (Mb) of storage. The storage capacity of a CD-ROM is about 640 Mb. A single sided, single layer DVD holds about 4.7 Gb, a single sided, double layer DVD about 8.5Gb.

Directory Structure

The files on the INTERMAGNET CD-ROM/DVD are set up in a particular directory structure. The root directory contains a "README.TXT" file, which is an ASCII file describing the CD-ROM/DVD and where to obtain information about it, the software, and documentation; CDs also hold a "README.EXE" file, which is an executable version of the README.TXT file that allows the user to scroll back and forth through the information.

On the 1991 CD-ROM there are also two sub-directories. One is labelled "XTRAS", and the other "MAG1991". The XTRAS directory contains one file labelled "STRUCTUR.DAT", and another "PRNSTRUC.EXE". The STRUCTUR.DAT file provides a schematic of the data structure for the records on the CD-ROM/DVD and the PRNSTRUC.EXE file enables the user to obtain a printout of that record structure.

The MAG1991 directory contains a sub-directory for each observatory identified by its 3-letter ID code. In addition, there are sub-directories labelled "1991MAPS", "CTRY_INF", and "OBSY_INF". The 1991MAPS directory contains the *.PCX files that are the map images of each country for use in the access software. These are labelled by a 3-letter country ID with the PCX extension, and one labelled "ALL.PCX" for the "All Countries" option. The CTRY_INF directory contains a "CTRYLIST.IDX" file that is used internally, *.PCX files for each country (and one for ALL) that are the images used to show the flag and organizational Logo for the different countries, and the README files that pertain to each country's geomagnetism program (including a README for the ALL option). The OBSY_INF subdirectory contains a "91OBSYDAT.DBF" file that is used internally in the software.

The individual sub-directories (e.g. BFE for Brorfelde, TUC for Tucson, etc.) contain the 12 months of data labelled with the 3-letter ID, 2-character year, 3-letter month abbreviation, and a "BIN" extension indicating they are binary files. For example, "BFE91AUG.BIN" is a file of 31 sequential day-records for Brorfelde, for 1991, for August. In addition, there are the "README.XXX" files for the individual observatory, where the XXX indicates the 3-letter observatory ID.

This sub-directory may also contain a file labelled as XXXYRK.DKA, where the XXX is the 3-letter observatory ID, the YR is the 2-character year value and the K indicates a K-Index file. Originally the DKA extension was used to indicate that the data were generated from a digital algorithm in an ASCII format, however subsequently these files have been used to hold both digitally derived and hand-scaled K indices. Since the 2005 CD-ROM the DKA files have been created by INTERMAGNET using data from the binary IAF file (before 2005 these files were provided by the observatories). These ASCII K-Index files are used, even though the data are in the binary records, because they are much faster to access than paging through the binary records on the CD-ROM/DVD.