

Chapter 7

Magnetic Observatory Data and Metadata: Types and Availability

S Reay¹, D Herzog², S Alex³, E Kharin⁴, S McLean², M Nosé⁵, N Sergeyeva⁴

1. British Geological Survey, Edinburgh, United Kingdom
2. National Geophysical Data Center, Boulder CO, USA
3. Indian Institute of Geomagnetism, Mumbai, India
4. Geophysical Center, Russian Academy of Sciences, Moscow, Russia
5. Graduate School of Science, Kyoto University, Kyoto, Japan

This is an author created version of:

*Reay SJ, Herzog DC, Alex S, Kharin E, McLean S, Nosé M and Sergeyeva N (2011) Magnetic Observatory Data and Metadata: Types and Availability. In: Mandea M, Korte M (eds) *Geomagnetic Observations and Models, IAGA Special Sopron Book Series 5*, DOI 10.1007/978-90-481-9858-0_7 pp149-181*

The original publication is available at <http://www.springerlink.com>

Abstract The availability of magnetic observatory data has evolved rapidly with the transition of observatories from analogue photographic magnetograms to digital electronic recordings, and the advent of the internet for instant global access to information of every sort. Metadata (information about the data) is undergoing its own transformation in order to accompany the rapid and extensive dissemination of these data. This chapter describes the types of data historically and currently produced by geomagnetic observatories and introduces new data types such as one-second and quasi-absolute data recently discussed at the 11th IAGA Scientific Assembly in Sopron, Hungary. We review the availability of these data types from the World Data Centres, INTERMAGNET and other sources. Finally, we discuss developments in metadata describing the current efforts in the geomagnetism community to gather, store and distribute this information about the data to better assist scientific discovery.

7.1 Introduction

Magnetic observatories continuously measure the strength and direction of the Earth's magnetic field (Macmillan, 2007). The availability of data from these observatories has evolved greatly with the transition from the production of analogue photographic magnetograms to digital electronic

recordings, and the development of technological capabilities such as the Personal Computer (PC); CD- and DVD-ROMs for the inexpensive archiving of large quantities of data; and the internet for instant global access to information of every sort. Metadata (information about the data) are undergoing their own transformation in order to accompany the rapid and extensive dissemination of these various data sets. Metadata describes the content, quality, originator, and other characteristics of a data set, and supports the proper interpretation and use of the data.

For more than a century, magnetic observatory data were only available on printed materials as analogue traces on photographic paper known as magnetograms, products derived from those traces, such as temporal averages (hourly, daily, etc.) and magnetic activity indices. Metadata consisted of yearbooks in which information about the observatories was provided, and copies of these various data products were reproduced in these books. The launching of the International Geophysical Year (IGY) in 1957 brought with it the beginning of the World Data Centres (WDC) established as central repositories where scientists could go to obtain data from a number of locations worldwide.

In the late 1960's, magnetic observatories began converting to digital electronic equipment and a true revolution in data types and availability was underway. The individual traces could now be kept separate (traces often crossed each other and became

confused on magnetograms) and the standard observatory data product changed from hourly means to minute means. Most significantly, these data could now be collected, manipulated, archived and disseminated using computers.

The emergence of the PC in the 1980's, and the development of CD-ROMs for data storage saw the beginning of "small science" where individual scientists were able to have large quantities of data available to study for minimal cost. In 1990, for example, the U.S. Geological Survey (USGS) produced the first CD-ROM of one-minute magnetic observatory data for a 5-year period (1985-1989) from the USGS network of (then) 13 stations (Herzog and Lupica, 1992), and served as the model for the production of a series of CD-ROMs by INTERMAGNET containing a global set of observatory data.

During the 1980's and 1990's, significant advances were also made in the timeliness of data delivery through the use of satellites. Data Collection Platforms (DCPs), equipped with satellite transmission systems, began delivering data from the observatories to data processing centres, with delays of the order of tens of minutes. An organisation called INTERMAGNET was formed which began transferring data from observatories worldwide using satellites, and which developed Geomagnetic Information Nodes (GINs) where users could obtain preliminary data from the participating stations quickly by means of email requests.

But without question the greatest advancement in data availability has come about from the impact of the World Wide Web (WWW) as a data discovery and distribution tool for the internet. Beginning in 1992, when WWW access became more broadly available to the general public, the growth in use of the WWW for access to data has been astonishing. Today virtually every institution, public and private, has a website with the capability of providing access to their information and data. As magnetic observatories have progressed in their abilities to collect and process data faster and with higher cadences (one-second data, for example), these data have also become ever more readily available. This change has implications for traditional WDC roles of data discovery and delivery, but also for long-term data archive.

In this chapter we will look at some of these developments and advances in data and metadata types and how they have been made available. Using an historical perspective, and with the help of results presented at the 11th IAGA Scientific Assembly held

in Sopron, Hungary in 2009, we hope to provide an account of where we are today regarding these topics.

7.2 Data Types

7.2.1 Printed Media

From the earliest days of magnetic observatory operations, data were collected on paper records of one sort or another. Variations in the magnetic field were recorded as continuous traces on photographic paper. Observations of the absolute magnitudes of the field components were recorded and provided calibration reference data with which to convert the analogue trace amplitudes into magnetic field values. These observations could then be reduced to produce tables of results at various time-scales: hourly, daily, monthly, annual. In time, range data started to be presented and then activity measures, such as the *K*-index (1939) were developed. Magnetic records also included absolute measurements, information on the baseline used and indications of the most quiet and active days (from 1911).

These records were presented in the form of yearbooks. These yearbooks and photographic analogue traces are archived and are available from different sources and in a variety of formats. Many magnetogram traces are stored on microfilm and microfiche. In this digital age we are keen to preserve these, often, fragile documents and allow greater access for analysis of these historical records. Some have been scanned as images and some have been entered directly to form digitised electronic files.

7.2.1.1 Eye-Observations

The earliest records from magnetic observatories were from eye-observations made manually by observers at set times during each day. For example at Greenwich observatory observations were made at 2-hourly intervals except on "term-days" when measurements were every 5 minutes. At Colaba observatory observations were made hourly except during disturbed times when they were every 15 minutes or every 5 minutes during severely disturbed conditions. Fig 7.1 shows the 5-minute eye-observations for Colaba during the September 1859 "Carrington Storm". The laborious nature of these manual observations encouraged the development of automatic recording devices. In 1847 Brooke designed the automatic photographic magnetograph that was to form the standard technique for observatory operations for over a century (Brooke, 1847).

DISTURBANCE OBSERVATIONS, 1859.

| Date and Göttingen Mean Time. | DECLINATION. | | HORIZONTAL FORCE MAGNETOMETERS. | | | | VERTICAL FORCE MAGNETOMETER. | |
|-------------------------------------|--------------|--------|---------------------------------|---------------------------|----------------------------|----------------------------|---------------------------------|--------------|
| | Large. | Small. | Large. | Small. | Thermometers. | | 2 min. before Full Time. | Thermometer. |
| | | | | | 2 min. after Full Time. | 4 min. after Full Time. | | |
| H. M. | | | Sc. Read. Uncorrected. | Sc. Read. Uncorrected. | | | Sec. Read. Uncorrected. | |
| SEPT. 1ST. | 18.15 | 28.070 | 29.22 | 15.27 | 23.22 | 83.6 | 81.0 | 59.00 |
| | 18.30 | 56.131 | 56.42 | 11.52 | 21.20 | 84.0 | 81.2 | 58.35 |
| | 18.45 | 49.202 | 49.00 | — | — | 84.0 | 81.4 | 59.45 |
| | 19.00 | 49.202 | 48.95 | 3.20 | 10.22 | 84.1 | 81.6 | 58.75 |
| | 19.05 | 56.063 | 55.00 | 2.47 | 16.80 | 84.1 | 81.8 | 57.55 |
| | 19.10 | 50.368 | 49.22 | 3.88 | 17.20 | 84.1 | 81.8 | 56.40 |
| | 19.15 | 48.859 | 46.00 | 7.78 | 21.10 | 84.4 | 81.9 | 55.60 |
| | 19.20 | 36.548 | 35.00 | 13.00 | 26.40 | 84.5 | 81.9 | 54.86 |
| | 19.25 | 37.940 | 37.44 | 13.55 | 26.50 | 84.5 | 81.9 | 56.20 |
| | 19.30 | 31.707 | 31.24 | 13.45 | 26.20 | 84.6 | 82.0 | 55.75 |
| | 19.35 | 28.687 | 31.57 | 12.57 | 27.50 | 84.7 | 82.3 | 55.25 |
| | 19.40 | 27.727 | 28.70 | 12.10 | 22.75 | 84.7 | 82.5 | 57.62 |
| | 19.45 | 29.991 | 30.79 | 13.65 | 24.40 | 84.7 | 82.6 | 58.22 |
| | 19.50 | 32.657 | 32.62 | 14.45 | 26.20 | 84.8 | 82.7 | 58.02 |
| | 19.55 | 24.846 | 28.70 | 14.46 | 27.65 | 84.8 | 82.7 | 57.48 |
| | 20.00 | 30.196 | 23.99 | 16.54 | 28.80 | 84.9 | 83.0 | 58.55 |
| | 20.05 | 22.168 | 24.25 | 15.17 | 29.35 | 84.9 | 83.0 | 58.25 |
| | 20.10 | 24.914 | 23.73 | 16.28 | 28.55 | 84.9 | 83.0 | 59.28 |
| | 20.15 | 31.088 | 25.56 | 15.60 | 29.07 | 84.9 | 83.1 | 58.64 |
| | 20.20 | 22.101 | 22.63 | 14.45 | 30.13 | 84.9 | 83.2 | 58.15 |
| | 20.25 | 17.293 | 18.76 | 17.35 | 29.25 | 84.9 | 83.2 | 59.35 |
| | 20.30 | 18.379 | 13.01 | 17.50 | 30.90 | 85.0 | 83.5 | 58.00 |
| | 20.35 | 23.130 | 20.33 | 13.35 | 23.30 | 85.1 | 83.5 | 59.45 |
| | 20.40 | 19.767 | 19.02 | 13.45 | 25.20 | 85.1 | 83.6 | 59.55 |
| | 20.45 | 18.670 | 17.90 | 14.40 | 26.15 | 85.1 | 83.6 | 59.00 |
| | 20.50 | 21.071 | 19.81 | 15.45 | 27.88 | 85.0 | 83.6 | 59.60 |
| | 20.55 | 18.327 | 19.02 | 16.85 | 29.65 | 85.0 | 83.6 | 57.50 |
| | 21.00 | 22.787 | 17.72 | 13.40 | 26.25 | 85.2 | 83.7 | 59.85 |
| | 21.05 | 19.356 | 20.33 | 13.70 | 25.50 | 85.2 | 83.8 | 60.62 |
| | 21.10 | 20.522 | 20.85 | 12.75 | 23.60 | 85.3 | 83.8 | 60.40 |
| | 21.20 | 19.356 | 20.33 | 12.80 | 25.50 | 85.3 | 83.9 | 59.85 |
| | 21.30 | 18.944 | 19.81 | 14.05 | 25.40 | 85.3 | 84.0 | 59.60 |
| | 21.40 | 17.778 | 18.76 | 13.85 | 25.70 | 85.3 | 84.0 | 59.50 |
| | 21.50 | 14.554 | 18.50 | 14.10 | 26.00 | 85.3 | 84.0 | 59.40 |
| | 22.00 | 15.445 | 17.45 | 13.65 | 25.70 | 85.3 | 84.0 | 59.88 |
| | 22.10 | 16.955 | 17.72 | 13.50 | 25.45 | 85.3 | 84.0 | 60.15 |
| | 22.20 | 17.641 | 18.65 | 13.45 | 24.90 | 85.3 | 84.0 | 59.85 |
| | 22.30 | 17.572 | 18.76 | 13.48 | 25.15 | 85.4 | 84.0 | 59.95 |
| | 22.45 | 17.641 | 18.76 | 14.35 | 26.07 | 85.5 | 84.0 | 60.00 |
| | 23.00 | 17.778 | 19.02 | 13.93 | 25.60 | 85.4 | 83.9 | 60.25 |
| | 23.15 | 19.013 | 20.07 | 14.40 | 25.75 | 85.3 | 83.9 | 60.62 |
| | 23.30 | 21.758 | 22.68 | 14.53 | 26.85 | 85.2 | 83.8 | 60.65 |
| | 23.45 | 23.198 | 24.40 | 14.70 | 26.65 | 85.1 | 83.6 | 61.00 |
| SEPT. 2ND. | 0.00 | 24.639 | 25.04 | 14.64 | 26.68 | 85.0 | 83.5 | 61.00 |
| | 0.15 | 23.130 | 25.04 | 15.40 | 26.65 | 85.0 | 83.5 | 60.60 |
| | 0.30 | 23.473 | 26.08 | 14.92 | 25.50 | 84.9 | 83.4 | 60.30 |
| | 0.45 | 23.130 | 24.51 | 17.23 | 28.05 | 84.8 | 83.3 | 59.85 |
| | 1.00 | 19.836 | 21.90 | 17.43 | 28.60 | 84.6 | 83.2 | 60.20 |
| | 1.15 | 20.590 | 20.07 | 16.55 | 29.55 | 84.3 | 82.8 | 59.45 |
| | 1.30 | 17.298 | 20.91 | 14.97 | 28.00 | 84.1 | 82.8 | 59.75 |
| | 1.45 | 20.865 | 22.16 | 13.10 | 24.70 | 83.9 | 82.8 | 60.00 |
| | 2.00 | 19.836 | 20.07 | 12.57 | 25.00 | 83.6 | 82.4 | 60.50 |
| | 0.15 | 31.707 | 32.50 | 12.45 | 24.60 | 83.5 | 82.4 | 60.60 |
| | 0.30 | 24.363 | 21.38 | 12.10 | 23.45 | 83.3 | 82.1 | 60.60 |
| | 0.45 | 20.453 | 19.28 | 10.97 | 23.50 | 83.2 | 82.0 | 60.60 |
| | 3.00 | 19.767 | 21.64 | 10.50 | 22.00 | 83.0 | 82.0 | 60.20 |
| | 3.15 | 22.305 | 23.31 | 10.20 | 21.20 | 83.0 | 82.0 | 60.75 |
| | 3.30 | 24.159 | 25.82 | 9.87 | 21.30 | 83.0 | 82.0 | 60.65 |

43-1859.

Fig 7.1 Eye-observations made with Grubb's magnetometer at Colaba Observatory for 1-2 September 1859. The observation frequency increases from 15-minute to 5-minute observations as the severity of the storm increases (see Tsurutani et al., 2003). (Credit: Indian Institute of Geomagnetism)

7.2.1.2 Magnetograms

For more than a century, the changes in the Earth's magnetic field have been recorded at observatories as analogue traces on sheets of photographic paper. Under darkroom conditions, mirrors attached to magnets suspended from quartz filament fibres, known as *variometers*, reflect light from a source through a horizontal lens and onto a sheet of photographic paper generally mounted on a drum that completes a rotation once a day. Time marks on the hour, generated by a light flash, provide a temporal reference. These paper records are known as *normal magnetograms*. There are also two auxiliary types of magnetograms produced at some stations: *storm magnetograms*, with reduced sensitivity, provide better amplitude resolution during periods of

enhanced magnetic activity; and *rapid-run magnetograms* are recorded on drums that rotate more than once a day as the drum moved laterally, providing greater time resolution. In addition to the mirrors mounted on the magnets in the variometer, stationary mirrors are also used to reflect light onto the recording drum and produce a fixed (non-varying) trace known as a *baseline*. A schematic illustration of this arrangement can be seen in Fig 7.2a. An example of a *normal magnetogram*, from Kakioka, Japan can be seen in Fig 7.2b. Although still in use today at some observatories, these systems have become less common as the magnetic field is more often recorded using digital systems as described in section 7.2.2.

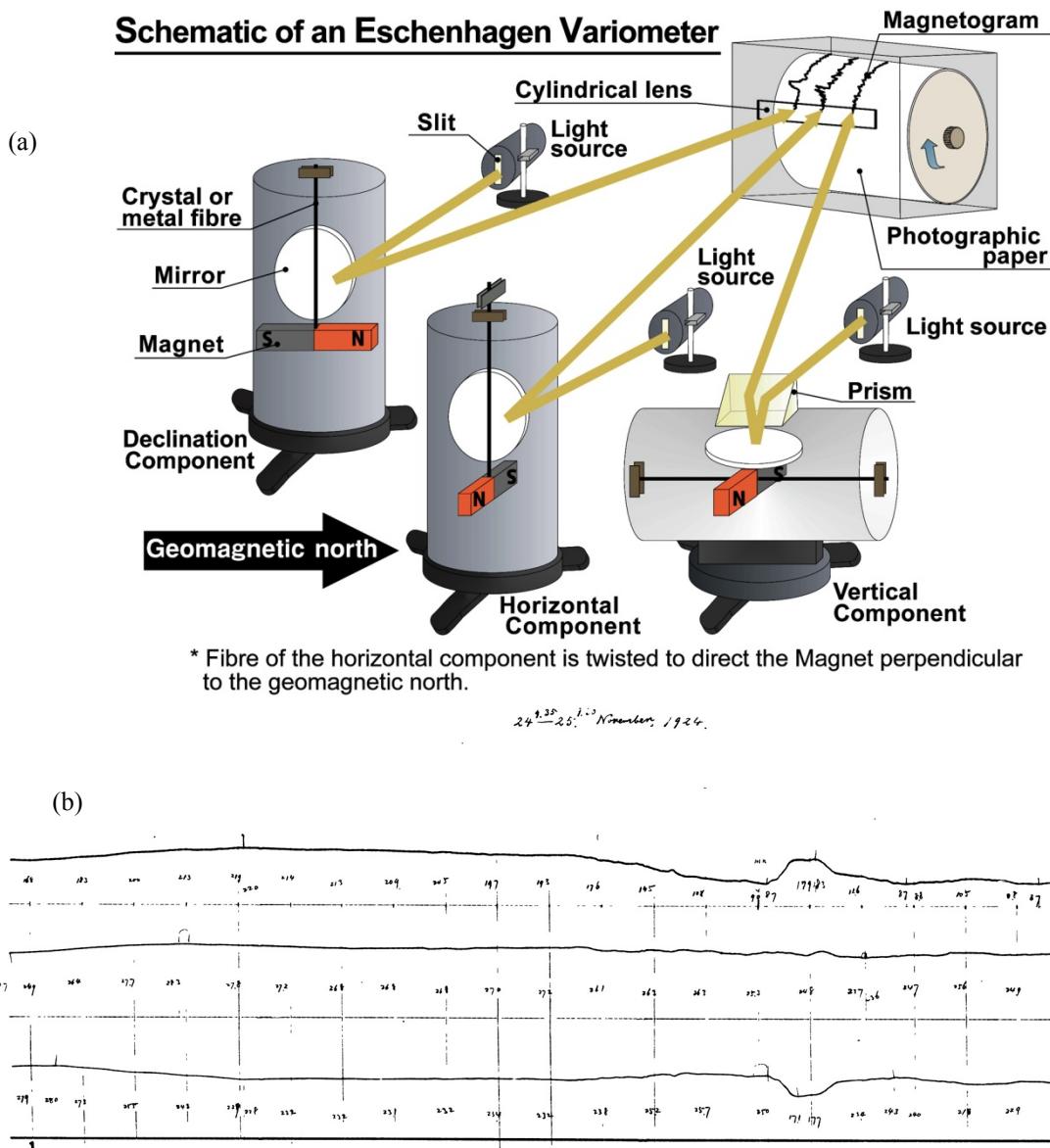


Fig. 7.2 (a) Schematic of an Eschenhagen type magnetic variometer (for details see Wienert, 1970) (b) Normal-run magnetogram recorded at Kakioka observatory on November 24, 1924. (Credit: WDC for Geomagnetism, Kyoto)

7.2.1.3 Calibration Data

In addition to the magnetogram traces continuously recording the variation in the magnetic field periodic measurements (generally weekly) of the absolute values of the magnetic field vector components were made using a variety of special instruments. Furthermore, special deflection magnets of known magnetic moment were used to obtain conversion factors, known as *scale values*, to convert the traces from units of distance to units of magnetic field strength. Knowing the absolute values of the field components for a given time interval, one could then combine these with the scale values to derive an absolute value for the baseline. With this, a calibrated value of the magnetic field at any point along the magnetogram trace could then be determined to convert the analogue traces into definitive magnetic values.

7.2.1.4 Hourly Values

After the magnetograms were developed, an observer would use a scaling glass to estimate the amplitude of each trace for each hour from its respective baseline. Using the calibration data, these could then be converted to hourly values of the magnetic field. In the earliest days of observatory operations these hourly values were often spot observations taken

every 60 minutes on the hour. Later, just after the turn of the 20th century, observatories began to estimate the average amplitudes over the whole hour to provide hourly mean values (HMVs). The first HMVs were calculated by Schmidt (1905) for Potsdam observatory with many others following suit shortly afterwards. These HMVs referred to the period of sixty minutes from the start of the hour, resulting in a mean centred on the following half-hour (i.e. a mean labelled as hour 10 is centred at 10h 30m). Convenient for hand-scaling between the hour marks on a magnetogram, this method continued into the digital era, thus ensuring homogeneity throughout the HMV data sets.

Until the advent of digital recording systems HMVs were the primary data product from magnetic observatories. Both the spot hourly values and the HMVs were usually compiled into monthly tables as shown in Fig 7.3. These tables were published in observatory yearbooks along with the information about the values that would enable distinction between spot and mean values. Unfortunately there were cases of insufficient or incorrect information being recorded, causing the potential for incorrect conclusions to be reached by researchers using these data sets and underlining the importance of clear and accurate metadata (see section 7.4).

| MAGNETOGRAF HOURS SCALINGS (UNIVERSAL TIME) | | | | | | | | | | | | | | U.S. DEPARTMENT OF COMMERCE COAST AND EARTHQUAKE SURVEY GEOMAGNETIC DIVISION | | | | | | | | | | OBSY. | YEAR | MONTH | ELEMENT | | | |
|--|----|-----|-----|-----|------|-----|-----|------|------|------|------|-------|-------|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|-------|------|-------|---------|------|------|-----|
| C | Q | S | G | No. | 01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 | 10 | 11 | 12 | N | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | SUM |
| <i>Values are in units of nT, and are averages for successive periods of one hour beginning at midnight. Hour 01 of local day is hour 12 of the previous universal day. Shaded areas indicate where corrections have been applied. Negative values are in red, with minus signs shown.</i> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| O | 1 | 01 | 280 | 258 | 257 | 247 | 240 | 238 | 240 | 235 | 237 | 247 | 257 | 260 | 01 | 276 | 287 | 288 | 289 | 300 | 316 | 320 | 307 | 275 | 249 | 229 | 229 | 6363 | | |
| O | 1 | 02 | 227 | 227 | 219 | 210 | 218 | 220 | 221 | 230 | 239 | 240 | 247 | 250 | 02 | 261 | 251 | 257 | 261 | 277 | 289 | 291 | 292 | 298 | 287 | 269 | 255 | 6036 | | |
| O | Q | 03 | 246 | 247 | 239 | 231 | 231 | 231 | 230 | 227 | 231 | 236 | 231 | 228 | 03 | 229 | 229 | 234 | 239 | 252 | 260 | 280 | 270 | 302 | 298 | 288 | 270 | 5975 | | |
| O | 1 | 04 | 243 | 232 | 239 | 235 | 228 | 222 | 213 | 203 | 218 | 210 | 218 | 219 | 04 | 214 | 219 | 221 | 226 | 249 | 261 | 285 | 292 | 291 | 283 | 272 | 274 | 5767 | | |
| O | Q | 05 | 267 | 257 | 237 | 239 | 230 | 230 | 230 | 232 | 232 | 232 | 234 | 237 | 05 | 239 | 238 | 232 | 230 | 240 | 262 | 288 | 280 | 271 | 252 | 242 | 258 | 5683 | | |
| O | Q | 06 | 266 | 266 | 255 | 241 | 245 | 253 | 252 | 251 | 240 | 245 | 229 | 249 | 06 | 220 | 240 | 238 | 235 | 261 | 281 | 280 | 270 | 282 | 280 | 262 | 230 | 6044 | | |
| O | 1 | 07 | 209 | 205 | 209 | 206 | 218 | 226 | 229 | 230 | 232 | 236 | 240 | 247 | 07 | 244 | 245 | 242 | 240 | 262 | 268 | 287 | 309 | 343 | 318 | 281 | 310 | 6061 | | |
| O | 1 | 08 | 297 | 250 | 245 | 229 | 238 | 241 | 230 | 210 | 157 | 182 | 162 | 190 | 08 | 198 | 202 | 217 | 220 | 250 | 269 | 266 | 252 | 250 | 262 | 252 | 255 | 5474 | | |
| O | 1 | 09 | 250 | 239 | 231 | 212 | 221 | 200 | 183 | 187 | 188 | 200 | 201 | 208 | 09 | 207 | 213 | 215 | 234 | 280 | 289 | 299 | 326 | 340 | 353 | 370 | 382 | 6042 | | |
| I | S | 10 | 392 | 362 | 305 | 203 | 028 | -210 | -258 | -138 | -0.5 | -0.68 | -0.32 | -0.49 | 10 | 058 | 058 | 047 | 024 | 050 | 079 | 107 | 111 | 130 | 110 | 071 | 089 | 1536 | | |
| I | S | 11 | 078 | 073 | -005 | 010 | 070 | 071 | 064 | 054 | 070 | 087 | 077 | 110 | 11 | 111 | 128 | 143 | 135 | 152 | 169 | 199 | 221 | 220 | 164 | 150 | 147 | 2709 | | |
| O | S | 12 | 132 | 125 | 111 | 092 | 093 | 092 | 107 | 140 | 150 | 154 | 150 | 142 | 12 | 178 | 183 | 172 | 177 | 173 | 200 | 193 | 178 | 169 | 123 | 123 | 3584 | | | |
| O | 13 | 128 | 152 | 157 | 140 | 129 | 126 | 154 | 148 | 149 | 149 | 189 | 189 | 187 | 13 | 162 | 170 | 171 | 173 | 181 | 196 | 211 | 220 | 210 | 179 | 181 | 208 | 4082 | | |
| O | 14 | 187 | 180 | 179 | 163 | 148 | 149 | 139 | 156 | 170 | 160 | 160 | 162 | 14 | 160 | 179 | 179 | 181 | 204 | 223 | 220 | 217 | 248 | 264 | 260 | 240 | 4530 | | | |
| O | 15 | 210 | 195 | 181 | 171 | 178 | 183 | 177 | 178 | 178 | 178 | 180 | 186 | 15 | 206 | 199 | 200 | 210 | 212 | 232 | 255 | 258 | 257 | 224 | 231 | 235 | 4916 | | | |
| O | Q | 16 | 220 | 207 | 190 | 185 | 189 | 188 | 182 | 182 | 182 | 191 | 191 | 190 | 16 | 191 | 195 | 192 | 192 | 211 | 218 | 222 | 220 | 230 | 249 | 261 | 249 | 4931 | | |
| O | 17 | 225 | 201 | 188 | 187 | 182 | 182 | 181 | 182 | 190 | 188 | 190 | 190 | 190 | 17 | 193 | 196 | 210 | 221 | 240 | 240 | 250 | 242 | 242 | 250 | 245 | 240 | 5054 | | |
| O | 18 | 231 | 219 | 211 | 205 | 208 | 211 | 212 | 221 | 229 | 233 | 239 | 239 | 221 | 18 | 200 | 210 | 210 | 210 | 225 | 226 | 226 | 226 | 250 | 250 | 230 | 198 | 176 | 5291 | |
| O | 19 | 202 | 208 | 209 | 200 | 198 | 199 | 199 | 202 | 208 | 211 | 219 | 213 | 211 | 19 | 211 | 207 | 210 | 210 | 219 | 239 | 257 | 260 | 249 | 221 | 202 | 201 | 5169 | | |
| O | S | 20 | 205 | 175 | 129 | 101 | 071 | 111 | 151 | 169 | 186 | 199 | 198 | 201 | 20 | 192 | 179 | 157 | 169 | 212 | 227 | 227 | 218 | 219 | 220 | 229 | 239 | 4386 | | |
| O | S | 21 | 222 | 121 | 140 | 165 | 170 | 170 | 173 | 165 | 180 | 187 | 170 | 183 | 21 | 195 | 197 | 181 | 181 | 182 | 192 | 192 | 211 | 228 | 220 | 222 | 220 | 213 | 4446 | |
| O | 22 | 228 | 219 | 200 | 191 | 193 | 205 | 205 | 212 | 210 | 215 | 219 | 221 | 22 | 23 | 231 | 241 | 230 | 236 | 244 | 239 | 228 | 210 | 198 | 198 | 198 | 5234 | | | |
| O | 23 | 205 | 207 | 191 | 178 | 179 | 163 | 153 | 162 | 171 | 184 | 209 | 208 | 23 | 216 | 206 | 203 | 197 | 207 | 228 | 239 | 237 | 221 | 210 | 200 | 212 | 4796 | | | |
| O | 24 | 199 | 189 | 192 | 187 | 187 | 193 | 193 | 198 | 200 | 208 | 210 | 211 | 210 | 24 | 218 | 222 | 225 | 226 | 230 | 258 | 262 | 267 | 257 | 251 | 262 | 249 | 5317 | | |
| O | 25 | 237 | 229 | 229 | 190 | 190 | 166 | 161 | 117 | 085 | 138 | 120 | 187 | 158 | 25 | 162 | 168 | 171 | 180 | 193 | 195 | 202 | 214 | 222 | 229 | 247 | 252 | 4405 | | |
| O | 26 | 241 | 219 | 188 | 171 | 170 | 170 | 192 | 190 | 190 | 179 | 179 | 179 | 179 | 26 | 190 | 190 | 194 | 189 | 202 | 212 | 228 | 221 | 221 | 212 | 206 | 198 | 4794 | | |
| O | 27 | 210 | 214 | 210 | 199 | 191 | 192 | 190 | 190 | 192 | 203 | 208 | 220 | 22 | 27 | 257 | 270 | 272 | 273 | 272 | 249 | 249 | 249 | 249 | 249 | 249 | 226 | 5466 | | |
| O | 28 | 239 | 240 | 232 | 228 | 227 | 220 | 205 | 206 | 210 | 219 | 237 | 248 | 28 | 250 | 228 | 222 | 238 | 229 | 255 | 240 | 239 | 229 | 203 | 193 | 182 | 5441 | | | |
| O | 29 | 178 | 170 | 178 | 169 | 180 | 193 | 192 | 190 | 183 | 230 | 221 | 200 | 20 | 29 | 194 | 192 | 191 | 208 | 221 | 250 | 259 | 256 | 250 | 233 | 219 | 209 | 4966 | | |
| O | Q | 30 | 207 | 208 | 220 | 221 | 211 | 210 | 207 | 210 | 212 | 210 | 212 | 220 | 31 | 221 | 220 | 222 | 228 | 228 | 239 | 249 | 250 | 252 | 262 | 258 | 250 | 262 | 5470 | |
| | | 31 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| SCALED BY <input type="checkbox"/> HO PERSONNEL <input checked="" type="checkbox"/> CHECKED BY <input type="checkbox"/> HO PERSONNEL | | | | | | | | | | | | | | () Interpolated [] Significant portion of record interpolated. □ No record; or no values available because of faulty record. ♦ Derived from _____ Mgs., converted to Normal Mgs. | | | | | | | | | | | | | | | | |
| MONTHLY SUM ¹⁴ 149993 | | | | | | | | | | | | | | MONTHLY MEAN 208 | | | | | | | | | | | | | | | | |
| DAYS WITH GAPS: / / | | | | | | | | | | | | | | USCGOM-DC 649-Pes | | | | | | | | | | | | | | | | |

Fig. 7.3 A table of hourly mean values from Honolulu observatory for the *H* component in June 1964. Each column shows the hour in UT and each row indicates the day of the month. These values are summed at the end of each row and a monthly sum and mean are produced. The table also indicates the five internationally quiet (Q) and disturbed (S) days.

7.2.1.5 Magnetic Activity Indices

In addition to the values of the magnetic field components themselves, efforts were made to develop indicators of the various levels of geomagnetic activity. A comprehensive guide to this subject is given in the chapter “Geomagnetic Indices” by Menvielle et al. (2010) within this book and in the classical text by Mayaud (1980). Here, we mention only a few of the most commonly used indices that were historically derived from and originally stored as analogue records.

The 3-hour K -index is a quantitative measure of local magnetic activity based upon the range of fluctuations in the observatory traces over 3 hours. It uses a semi-logarithmic scale from 0 to 9, with 0 indicating completely quiet conditions and 9, highly disturbed conditions. It is intended to measure geomagnetic disturbances outside the normal diurnal quiet time variations and so these had to be accounted for by the experienced hand-scaler. The largest range of either the horizontal component (H) or declination (D) trace (originally the vertical component (Z) was also used) is selected. K -indices can also be converted into eight 3-hour equivalent linear amplitudes, ak , and summed, these provide a local daily amplitude index, Ak .

The scaling of K -indices at a wide network of observatories enabled the derivation of various K based planetary indices adopted by IAGA: Kp from which ap is derived; am ; and aa . The planetary activity index, Kp , is derived from the standardised K -indices of 13, mostly mid-latitude, observatories (Bartels et al. 1939). The name Kp comes from the German phrase "planetarische Kennziffer", meaning planetary index. It is designed to monitor the influence of the auroral electrojet current system, and the magnetospheric ring current and field-aligned currents. As with Ak , ap is the 3-hourly equivalent planetary amplitude derived from Kp and Ap is the daily average of the eight values. Kp , ap and Ap extend back to 1932. Although much used by the external magnetic field communities, these indices are generally considered to have two main weaknesses: one is the poor global distribution of the observatory network used; and the second is that the length of the time series is insufficient for century-long studies. The first was solved by the design of the am -index (Mayaud, 1967), which uses a more extensive network of observatories and is available from 1959. The latter was solved by the design of a much simpler K based planetary index, the aa -index (Mayaud, 1972). Using only two near-antipodal observatories it could be extended back further in time to 1868.

Another important measure of magnetic activity is the Disturbed storm time, Dst -index (Sugiura, 1964). This hourly index is based on the average value of H

measured hourly, as described in 7.2.1.4, at four near-equatorial geomagnetic observatories. It was originally introduced to measure the magnitude of the current which produces the symmetric disturbance field, but is often used to estimate the strength of magnetic storms caused largely by the magnetospheric ring current. This (mainly) westward flowing ring current causes H to be depressed from its normal level. Dst can be used as a measure of magnetic storm intensity because the strength of the surface magnetic field at low latitudes is inversely proportional to the energy content of the ring current.

7.2.1.6 Yearbooks

Every year the results from a magnetic observatory would be collated and published in the form of a yearbook. Along with photographic magnetometer traces yearbooks are one of the primary sources of printed geomagnetic observatory data. The type of data published in yearbooks has changed over the years as instrumentation has developed and methods were standardised, but all share commonality. Generally, yearbooks contain descriptive text describing the observatory; this may detail its location, housing, staff, instrumentation, and data processing methods. It also contains the final definitive measurements and results from the observatory. As an example, yearbooks from Greenwich observatory originally contained results called ‘indicators of the magnetometers’; firstly from eye-observations and then taken directly from the photographic magnetometer records. These were published alongside records of magnetic dip (these days normally referred to as inclination, I) and absolute measures of H . In later years hourly, daily, monthly and annual means were presented alongside K -indices, diurnal ranges, absolute measurements, baseline information and, sometimes, selective magnetograms and descriptive notes of magnetic disturbances. The results for many observatories were often published alongside meteorological observations.

It is worth noting when examining data from the earliest yearbooks that care must be taken regarding time recordings. Prior to the adoption of standard time-zones and Universal Time in 1884, time conventions at each observatory could vary considerably. Some used a local civil time where a day started at midnight whereas others used an astronomical time where a day started at noon or some other local convention (Boteler, 2006).

Presently, the tradition of publishing yearbooks is still practiced by many observatories worldwide. However since advent of digital distribution of results this practice has been discontinued at some observatories. Yearbooks continue to be the best source of valuable metadata on observatory practice,

instrumentation, data processing and quality control and the publication of these are encouraged by IAGA (Jankowski and Sucksdorff, 1996).

7.2.1.7 Conservation and Conversion of Printed Media

During classical, analogue operations magnetograms, HMV tables, *K*-index tables, and calibration data made up the core of data types produced at magnetic observatories, although other derived products such as sudden storm commencements (*ssc*) and magnetic pulsations (Matsushita and Campbell, 1967) were also generated at some stations. At some WDCs the various paper products were archived onto microfilm rolls and microfiche cards. WDCs periodically published individual and joint catalogues identifying available data at each of the geomagnetic data centres (e.g. Abston et al., 1985). Fig 7.4 shows the holdings of analogue and digital data at WDC-Kyoto and illustrates that the digital-age of magnetic observatories only began in the mid-1980s. There are vast amounts of data ‘locked away’ in analogue form that is difficult to access and analyse in this digital era. Special procedures are required to preserve these data and allow electronic access.

Much effort has been made over the years to digitise magnetogram traces (e.g. Curto et al., 1996), but this has proved to be a challenging task. During magnetically active times, generally of most interest to researchers, the traces for the different components often cross each other making it difficult to identify which is which. Nonetheless, at the 11th IAGA Scientific Assembly, the British Geological Survey (BGS) reported on a programme to digitise magnetograms from the UK observatories dating back to 1848 (Clarke et al., 2009). The primary goal

of the project is one of conservation, with the capture of the magnetograms as images (Fig 7.5), which will also enable increased access to the scientific community. The second aim is to develop a semi-automated procedure to convert the images into digital values with a greater time resolution than has previously been reported.

In addition, many of the WDCs discussed in section 7.3 of this chapter have been involved in a special project funded by an International Council for Science (ICSU) grant. This “Magnetogram Rescue Project” involved WDCs Kyoto, Moscow, and Mumbai and aimed to locate old magnetograms and convert them into digital images (Iyemori, 2005). The results of this project are available from these WDCs.

The observatory community has invested most of its efforts in this area to convert the tables of data in yearbooks and other bulletins (for example, HMVs, *K*-indices) into digital electronic values. Various attempts have been made at this, including the scanning techniques of Optical Character Recognition (OCR). The unreliability of this procedure, especially for HMV tables (generally due to the high variability of different typefaces and hand-written entries in the tables), caused these efforts to be directed more toward manual entry and checking (see for example, Nagarajan, 2008; Fouassier and Chulliat, 2009; Nevanlinna and Häkkinen, 2010). While the reliability of this manual procedure is very good, it is also labour intensive and can therefore be expensive. As a compromise, and in order to make the data accessible electronically, many data sets have been, and are being, scanned into images that will at least allow for computer manipulation by researchers.

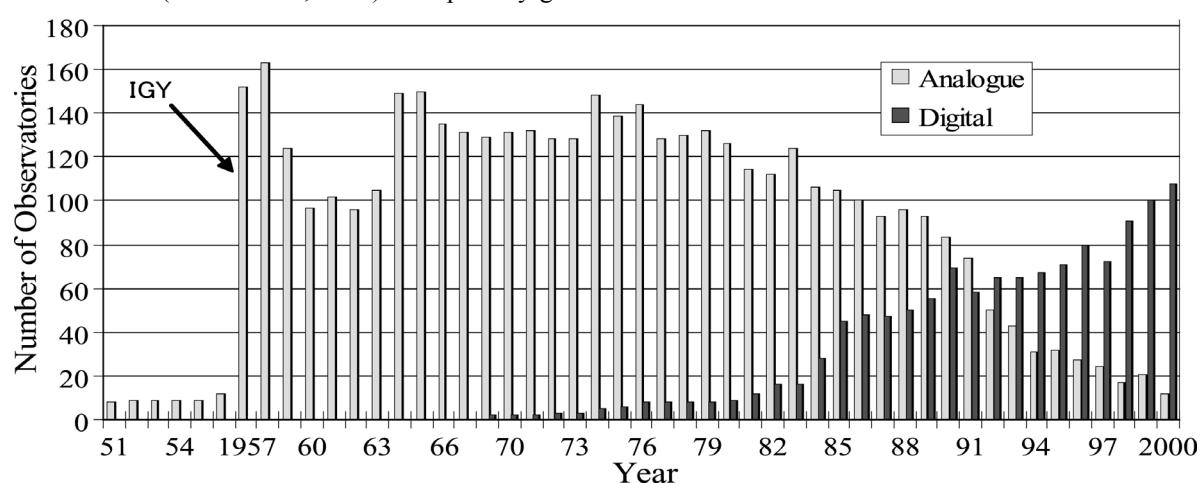


Fig 7.4 The number of analogue normal-run magnetograms and digital one-minute data held at the WDC-Kyoto. (Credit: WDC for Geomagnetism, Kyoto)



Fig. 7.5. Digital image capture of historical magnetograms (Copyright: BGS/NERC)

7.2.2 Electronic Media

The earliest digital electronic data from magnetic observatories were recorded in 1969 from two French stations, Dumont d'Urville (DRV) in Antarctica, and Port-aux-Français (PAF) on the Kerguelen Islands in the Indian Ocean. These stations used photoelectric feedback and Cesium vapour magnetometers to sample the northerly intensity (X), easterly intensity (Y) components, and total intensity (F) of the magnetic field at one-minute intervals. The introduction of the triaxial fluxgate magnetometer (see for example, Trigg et al., 1971), and use of analogue-to-digital converters in data loggers by the early 1970's, ushered in the true digital electronic age in magnetic observatory operations. These devices sampled the field multiple times per minute which were then filtered to produce one-minute values. This

also was the beginning of one-minute values becoming the standard published magnetic observatory results. Many institutes added proton precession scalar magnetometers at their stations which provided separate measurements of the absolute F , and also served as a reference to compare with the fluxgate vector data. Observatory data quality is discussed in the chapter "Improvements in geomagnetic observatory data quality" by Reda et al. (2010) within this book, where a full description of modern magnetic observatory instrumentation, operations and data processing is provided.

7.2.2.1 Minute Means

Various brands of linear core, three-component fluxgate magnetometers are installed at observatories throughout the world. Originally these instruments recorded an output voltage proportional to the

strength of the ambient magnetic field. Using the calibration data, the voltages were converted to magnetic field values. The output from a fluxgate magnetometer can be sampled at various frequencies; initially systems were set to sample every 10-seconds but 1-second sampling is now more common. Regular manual absolute observations are still required to derive baseline values for the fluxgate data. Combined, these provide definitive data (see for example, Turner et al. 2007). Early on, data were stored on magnetic tape and processed to derive 2.5-minute means. Over time these faded from use and one-minute means became the standard. These definitive digital one-minute data form the basis for easy computation of various time averages, including hourly, daily, monthly, and annual means.

As with the analogue records, the delay in the availability of early digital data could be considerable. This was due to several factors including the fact that recording tapes at many stations were left for weeks before they were retrieved. Even today, it requires several months of weekly absolute observations, to enable the production of reliable final baselines, before the final definitive data can be delivered to the WDCs.

With the many advantages that digital data provided, and the proliferation of low-cost computational capabilities, the demand for these one-minute data grew. New DCPs began allowing access remotely through modems, and requests for the data rapidly increased. Many researchers, for example in the space physics community, did not require definitive data as they are only interested in the variations in the magnetic field. The internet, with its email and FTP capabilities, allowed institutions to send preliminary data to users more quickly and efficiently. By 1991, INTERMAGNET had established GINs where a global collection of observatory one-minute data could be retrieved. Demand from industry for observatory data also grew. One example is its use, in near real-time, to improve the accuracy of bore-hole surveys in off-shore directional drilling operations carried out by oil and gas companies (Reay et al., 2005). Today, one-minute data are available from a multitude of locations through websites within minutes of the measurements being made.

However, three important points concerning the one-minute data should be emphasised. First, in the early 1990's magnetometers and data collection systems were developed that produced an output in magnetic units rather than voltage units (Narod, 2009). The field measured in this way may appear to be the correct value that one might expect for a given location, but it is not. There are correcting factors, such as zero-level offsets and orthogonality issues that must be corrected for with the use of baselines in the calibration data. With the widespread

proliferation of preliminary one-minute data, one could potentially mistake the preliminary data for the definitive data. The preliminary values can be as little as a few to as many as hundreds of nT apart from the definitive values. Nowadays INTERMAGNET clearly separates each type of data on their website and the IAGA-2002 data format explicitly states the type of data both in the filename and the headers. This gives another good example of the power of metadata.

Secondly, early on, the one-minute means were often computed using simple arithmetic mean algorithms, sometimes referred to as a "box-car" mean. However, when converting from analogue to digital, in order to preserve the full information in the analogue signal, it is necessary to sample at least twice the maximum frequency of the analogue signal. Sampling at lower rates introduces erroneous components due to *aliasing*. A numerical filter can minimise the effects of aliasing in the construction of one-minute values from one-second samples. A Gaussian filter is often used for this purpose, although other types meeting the same requirements are also used. INTERMAGNET now requires numerical filtering of one-minute data as a criterion for attaining INTERMAGNET Magnetic Observatory (IMO) certification and encourages the use of a Gaussian filter using 90 one-second samples. The USGS, for example, began applying a Gaussian filter to derive its one-minute data from 1995 and the BGS use a 61-point cosine filter, having adopted this method in 1997 when instruments capable of one-second sampling were first installed in the UK.

Finally, we should mention that minute means have not always been centred on the minute, but rather were often computed from the 60 seconds of values prior to the minute. It wasn't until the IAGA meeting in Canberra, Australia in 1979 that the decision was taken to encourage observatories to centre the minute values exactly on the minute (Resolution 12).

7.2.2.2 Hourly, Daily, Monthly, Annual Means

An on-going debate within the geomagnetic observatory community is the question of how HMVs should be computed when data are missing within the hour. The question centres on the accuracy of HMVs when the one-minute data for the hour are not complete. Over recent years a number of researchers have considered what may affect the accuracy of HMVs considering various factors including how many data points are missing, the distribution of these data gaps (random, continuous blocks), the level of magnetic activity and, the magnetic latitude of the observatory. Both Mandea (2002) and Schott and Linthe (2007) concluded that a reliable HMV could be computed if less than 10 % of the data were

missing. However Herzog (2009) found that during magnetically quiet times, including at a high-latitude station, up to two-thirds of the data could be missing without significant loss of accuracy. Newitt (2009) also found the 10% rule be an over-simplification leading both to the unnecessary rejection and unacceptable inclusion of certain data. He suggested a more statistical approach based on a set maximum permissible error and this was further considered in Marsal and Curto (2009) who studied the effect of a pre-established relative accuracy on HMVs. Love (2009) stated that the level of accuracy set for HMVs could not be better than the $<5\text{nT}$ level set for minute data by INTERMAGNET. In his study, he found that HMVs could satisfy this 5nT level of accuracy, on average, about 90% of the time if the “10 % rule” were applied. Whilst no common consensus yet exists on this issue, an IAGA task force is currently working towards determining a resolution (Hejda et al., 2009).

Questions have also arisen regarding how the daily means should be computed. Differences can arise depending upon whether one computes a daily mean from the average of the 1440 minutes of the day or from the average of the 24 hours of the day. Missing data during the day can give rise to these differences. INTERMAGNET guidance states that a daily mean is calculated from 24 hourly values and an hourly value is calculated from 60 one-minute values. Monthly and annual means are less sensitive to these differences, although this depends on the amount of missing data and the time distribution of the missing data. These have historically been computed both for all days and for magnetically quiet days using the daily mean values. At the WDC Edinburgh, a note is associated with any annual mean value derived from an incomplete data set. This allows these data to be down-weighted when used in global magnetic field modelling.

7.2.2.3 Digital Magnetic Activity Indices

Although production of digital one-minute data became the standard for observatory operations, the legacy of products produced from the analogue magnetograms, including magnetic indices, continued. The long time series of these products were important to researchers studying the magnetic field and development of “magnetic climatologies” requiring many decades of data.

The derivation of K -indices from digital data presented its own unique difficulties. At first the data were plotted and printed on paper thus enabling the index to be hand-scaled in the same way as traditionally done from photographic magnetograms. There was, however, a strong incentive to obtain K -indices automatically using computer algorithms. A comparison between the various proposed algorithms

was organised by an IAGA Working Group on geomagnetic indices, and four computer-derived methods were found to provide acceptable results when compared to the hand-scaled method. The four methods, discussed in Menvielle et al. (1995), were approved at the IUGG General Assembly in Vienna in 1991. Although many institutes gradually switched to digital derivation, some have continued to use the traditional hand-scaling method.

The convenience and flexibility of digital electronic media provided the opportunity to not only derive existing indices digitally, and in a more timely manner, but also to devise new magnetic activity indices, such as the AE -index and PC -index.

The AE -index is the auroral electrojet index that is derived from a set of 12 observatories located in the auroral zone latitudes of the northern hemisphere (Davis and Sugiura, 1966). Plots of the H -component for each station are superimposed together to form a collection of traces. This collection produces upper and lower bounds of the range of magnetic activity for the group. The maximum positive deviation of the H -traces produces an upper bound (AU), and the maximum negative deviation produces a lower bound (AL). AE is the difference between the two and provides a global measure of the magnetic activity caused by ionospheric currents within the auroral oval. It correlates well with magnetospheric substorms that can last for several hours. This index is attractive because it is easy to compute digitally, and can be used to provide a measure of magnetic activity over any time scale desired. It also has known limitations; for example during very large storms the auroral electrojet will move equatorward of AE stations and its strength will no longer be measured (e.g., Akasofu, 1981; Feldstein et al., 1997). This index was endorsed by IAGA in 1969 (Resolution 2).

The PC -index is the polar cap (dimensionless) index designed to measure geomagnetic activity over the polar cap regions using only a single station located near the geomagnetic North and South poles (Troshichev and Andrezen, 1985). These polar stations are Qaanaaq (previously Thule) in Greenland for the northern hemisphere and Vostok in Antarctica for the southern hemisphere. This index is used to describe the principal features of the solar wind and the Interplanetary Magnetic Field (IMF), and indicates the total energy input into the magnetosphere. This was recommended to IAGA in Birmingham (1999) and endorsed in Hanoi (2001). The endorsement was later suspended due to different procedures being used in the derivation of the index in each hemisphere (PCN and PCS). At Sopron, Stauning et al. (2009) outlined a unified procedure now in use for both indices.

7.2.2.4 One-Second Data

With advances in magnetometers and other hardware systems and circuitry, it was inevitable that observatories would begin collecting and storing one-second data. Currently data reported from observatories are most commonly one-second values recorded from a magnetometer sampling at 1 Hz. Unlike one-minute and hourly mean data these are spot values and the accuracy of the timing (to the millisecond) is not well controlled. Providing true one-second mean data (obtained by filtering higher frequency data) is a challenging endeavour requiring sophisticated magnetometer specifications, timing accuracy, sampling restrictions, physical and numerical filtering processes, data processing and storage techniques. The greatest obstacle to overcome is instrumental noise which at 1Hz is typically $10 \text{ pT}/\sqrt{\text{Hz}}$, greater than geomagnetic signal at this frequency.

At the IAGA conference in Sopron, several authors reported on issues regarding developments regarding one-second data. Korepanov et al. (2009) presented the results of tests with a new magnetometer, developed at the Lviv Center of Institute for Space Research in the Ukraine, which meets the standards for one-second data set by INTERMAGNET. Their magnetometer accomplished a 1 pT resolution matching the standard proposed by INTERMAGNET. The sensitivity threshold is close to 1 pT, and the damping of higher frequency ambient noise, especially that of power line harmonics has been achieved. Dourbes and Conrad observatories are currently using this particular magnetometer. Worthington et al. (2009a) reported on an effort to estimate the noise levels and timing accuracy of the fluxgate magnetometers and A/D converters for the one-second data being collected at the USGS, compared to the standards proposed by INTERMAGNET. The results showed that the USGS observatories studied have noise levels of 0.01-0.02 nT and the timing is close to the 10 millisecond accuracy proposed by INTERMAGNET in 2008.

Others observatories around the world are investigating this issue and developing new techniques and instrumentation to provide one-second data (for example Chulliat et al., 2009a; Shanahan and Turbitt, 2009; Worthington et al., 2009b). Whilst this is a new development for most, the Japan Meteorological Agency has been recording one-second data at Kakioka observatory for about 25 years and at Memambetsu and Kanoya observatories for over 10 years (Minamoto, 2009).

A difficulty with one-second data is the question of what format to use for distribution. The IAGA-2002 format (see Appendix 1) allows for the possibility of one-second data, but only provides for a

resolution of 0.01 nT or 10 pT, and this is insufficient for the INTERMAGNET proposal to record one-second data to 1pT resolution. Nonetheless, it was decided at the INTERMAGNET meeting in Sopron (2009), that one-second preliminary data would be distributed from the website. IMO's will be requested to submit all one-second data that they hold. Initially the IAGA-2002 data format is to be used for these data, despite the resolution problem. Several possibilities for a more satisfactory format were discussed at the Sopron meeting, including: the development of a new format using XML; the development of a new format using the CDF (Common Data Format) [1] currently used in space physics and magnetic satellites communities; or modifying the current IAGA-2002 format.

7.2.2.5 Quasi-Definitive Data

The significant time delay between the availability of preliminary data and definitive data from observatories prompted the proposal of a new data type at the 11th IAGA Scientific Assembly in Sopron. Preliminary data are accessible almost immediately in many cases, but the definitive data can take a year or more to become available. Field modellers have a need for what has been termed “quasi-definitive data,” defined as “data corrected using temporary baselines shortly after their acquisition and very near to being the final data of the observatory” (Chulliat et al., 2009b). Such baseline corrected observatory data are useful to modellers to quickly detect geomagnetic jerks (Chulliat et al., 2010), to test IGRF-11 candidate models (Chulliat, 2009c; Finlay et al., 2010) and for the validation of level-2 products within the context of the upcoming Swarm mission (Friis-Christensen et al., 2006), expected to be launched in 2012. The appeal of this data to modellers is apparent, but one concern with this proposal is the notion of the values being “very near” to that of the final definitive data. One standard proposed at the INTERMAGNET meeting following IAGA was that the data should be within 5 nT of the final definitive values, although verification will only be possible following production of the definitive data. Recent studies suggest observatories should be able to meet this requirement, but confirmation of the accuracy achieved will always be after the fact.

One study presented in Sopron reported on work carried out at the BGS and demonstrated that quasi-definitive data can be produced in near real-time (Baillie et al., 2009). BGS use piecewise polynomials to estimate baselines, and these are extrapolated to provide the daily baseline values used to construct quasi-definitive data in near real-time. For the five BGS INTERMAGNET observatories, comparisons were made between the X , Y and vertical (Z) component quasi-definitive hourly mean values,

available on a next day basis, and the final definitive hourly mean values from 2000 to 2007. The distribution of differences showed that for all observatories the quasi-definitive data were within 5nT of the final data 94 - 99% of the time. Peltier and Chulliat (2010) also recently proposed a method for producing quasi-definitive data every month and demonstrated that the difference between quasi-definitive and definitive data was less than 1 nT, well within INTERMAGNET standard, for nine observatories having very different baseline characteristics. At Sopron, IAGA encouraged magnetic observatories to produce baseline-corrected quasi-definitive data shortly after their acquisition (Resolution 5) (IAGA, 2009).

7.3 Data Availability

7.3.1 World Data Centres for Geomagnetism

During the IGY, the ICSU created a number of WDCs that were designed to collect, catalogue, archive and distribute geophysical and solar data sets from a few centralised locations (ICSU Panel on WDCs, 1996). There was, and continues to be, a clear need for a few central repositories where scientists can go to obtain data in a consistent format and avoid catastrophic and irreversible losses of collected data (Mande and Papitashvili, 2009). Originally, WDCs were set up in the United States (WDC-A), Russia (WDC-B), Europe (WDC-C1) and Asia-Oceania

(WDC-C2). This has expanded to other regions, notably China (WDC-D) in 1988. WDCs were originally established to manage and preserve data from the physical sciences, such as geomagnetism, oceanography, and meteorology. Since these beginnings WDCs have expanded greatly in both the disciplines they support and the ways in which they meet the scientific community's requirements. Today, WDCs are more broadly environmental and include such disciplines as biodiversity and ecology, soils, and land processes. There are now more than 50 WDCs located in 12 different countries, as shown in Fig 7.6. Geomagnetic observatory data, geomagnetic indices, survey data, models and other sources are maintained at several of these WDCs.

WDCs are funded and maintained by their host countries on behalf of the international scientific community. The WDCs operate under the guidelines of an ICSU panel, and from the outset have championed full and open access with data made available free of charge, or for the nominal costs of reproduction. WDCs are encouraged to exchange data with each other and there are informal arrangements between data centres holding geomagnetic data to exchange new data sets received. WDCs are also expected to assure a reasonable standard of data quality and documentation. The ultimate responsibility however for data quality lies with the data contributor and not the WDC (ICSU Panel on WDCs, 1996). Recently Korte et al. (2007)

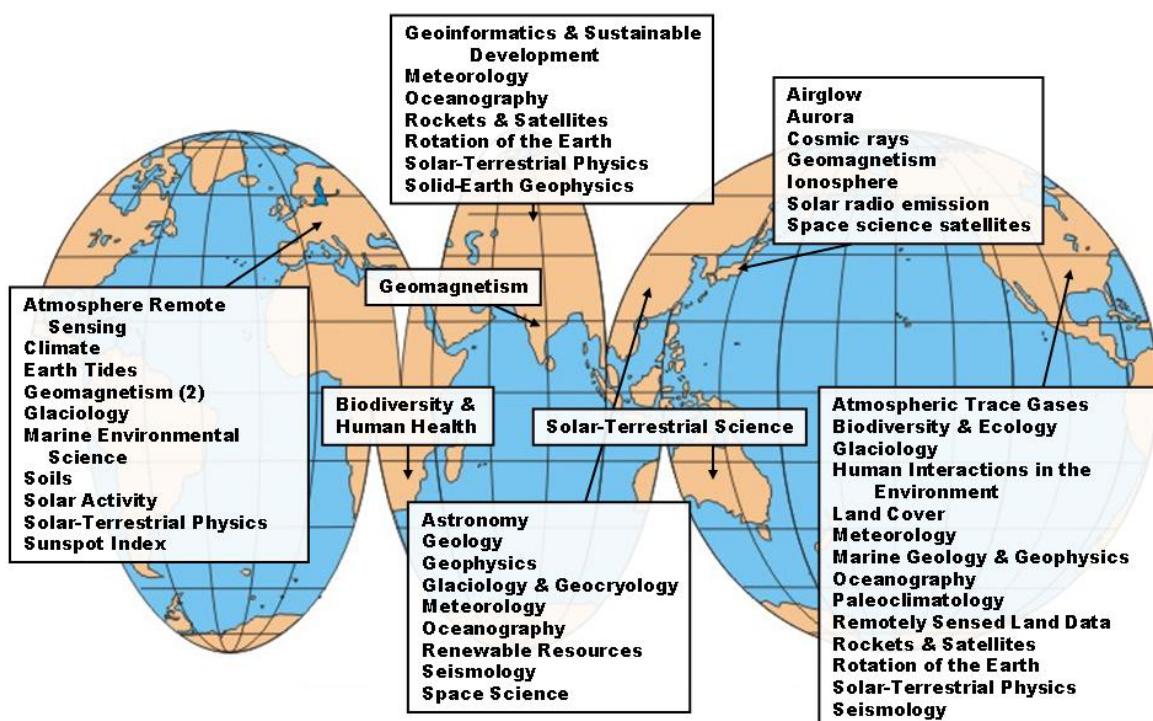


Fig. 7.6 Locations and scope of the various World Data Centres worldwide. (Redrawn from original courtesy of Prof B. Minster, past Chair of the ICSU Panel on WDC)

identified a number of inconstancies in the hourly mean data sets held at the (then) WDC for Geomagnetism in Copenhagen. It is preferable that data producers re-examine their data sets and update the WDC database accordingly. However in some cases WDCs have, and do, make changes to the data held within. For example, Dawson et al. (2009) reported in Sopron on WDC for Geomagnetism, Edinburgh's efforts to correct simple typographical and formatting errors to the data sets. In this case it is not necessary to return the data to the individual observatory to correct. Improving the metadata associated with data sets, noting any errors, and if changes are made, is perhaps the first step to improving the quality of the data held.

7.3.1.1 World Data Centre for Geomagnetism, Edinburgh

The WDC for Geomagnetism, Edinburgh [2] was established in 1966 at the Institute of Geological Sciences in Sussex, which later became the BGS, and moved to its current location in Edinburgh in 1977. BGS, which is part of the Natural Environment Research Council (NERC), a publicly-funded agency, had concentrated on gathering data primarily for use in global magnetic field modelling - mainly annual mean data from the worldwide observatory network. Whereas, the WDC in Copenhagen, which was hosted by the Danish Meteorological Institute (DMI), gathered one-minute and hourly observatory data. WDC-Copenhagen provided access to these data sets online via a 'Data Catalogue' website. In 2007, BGS agreed to take over responsibility for these data sets and the operation of this 'Data Catalogue' website.

Data now held at the WDC-Edinburgh include digital data from geomagnetic observatories worldwide (Fig. 7.7) including one-minute mean values from 1969; hourly values from 1883; and annual means from 1813. Also available are uncorrected one-second data from 2000 onwards from UK observatories and digital versions of some yearbooks. Data from land, marine and aeromagnetic surveys and repeat stations worldwide from 1900 onwards are available, as well as charts and computations of main field models, including the World Magnetic Model (WMM) and International Geomagnetic Reference Field (IGRF) model. Further digital data include definitive magnetic activity indices (K , K_p , ap , Ap , aa , Aa , Cp and $C9$); estimated real-time planetary indices (ap_{est} , Ap_{est} , aa_{est} and Aa_{est}) and solar activity indices (SSN and $F10.7$) [3]. In addition its analogue data holdings include shipborne declination data from 1590 onwards; archived magnetograms for several UK stations from 1850; and a library containing yearbooks, expedition memoirs, original survey observations and similar

items. Data are available online, by anonymous FTP or by request.

7.3.1.2 World Data Center for Geomagnetism, Kyoto

In 1957, the Faculty of Science, Kyoto University was assigned to establish the WDC-C2 for Geomagnetism [4] since it was well-known for its achievements in geomagnetic research. This led to the foundation of the World Data Archive for Geomagnetism in the Kyoto University Library in December of the same year. In 1976, the Solar Terrestrial Physics Subcommittee of Science Council of Japan adopted the resolution that the existing data archive should be managed by a new research institution. As a result, on April 18, 1977, the Data Analysis Center for Geomagnetism and Space Magnetism (DAGSM) was established as a new institution in the Faculty of Science, Kyoto University and an associate professor was assigned to it. For over 30 years, DAGSM has been operating the WDC for Geomagnetism, Kyoto and providing a leading data service.

WDC Kyoto holds geomagnetic field data in form of normal-run magnetograms, rapid-run magnetograms, hourly digital values, one-minute digital values, and one-second digital values. Kyoto also derives geomagnetic indices, including the AE -index and the Dst -index, both of which are IAGA sanctioned indices, as well as the SYM/ASY indices (Iyemori, 1990). Recent advances in computing have made it possible for this data centre to collect and display geomagnetic field data in real-time from more than 30 observatories. Magnetograms have been copied onto microfilm or microfiche and are available back to the early 1900's. Converted digital images of old magnetograms resulting from the ICSU "Magnetogram Rescue Project" are also available from the Kyoto WDC [5].

7.3.1.3 World Data Center for Geomagnetism, Copenhagen

The WDC for Geomagnetism, Copenhagen is maintained by the National Space Institute at the Technical University of Denmark (DTU Space) [6] since 2010. It was established at the DMI for the IGY in 1957. Its Geomagnetic Data Master Catalogue for geomagnetic observatory hourly means and one-minute means moved to WDC Edinburgh in 2007.

WDC-Copenhagen currently holds digital data for the PCN -index derived from Qaanaaq geomagnetic observatory, and a mirror of the GFZ Potsdam ftp-server, home of the Kp -index. It has substantial analogue holdings and catalogues of these holdings, in the form of searchable PDF files, are available online [7]. This includes approximately 4300 years of magnetograms from about 280 geomagnetic

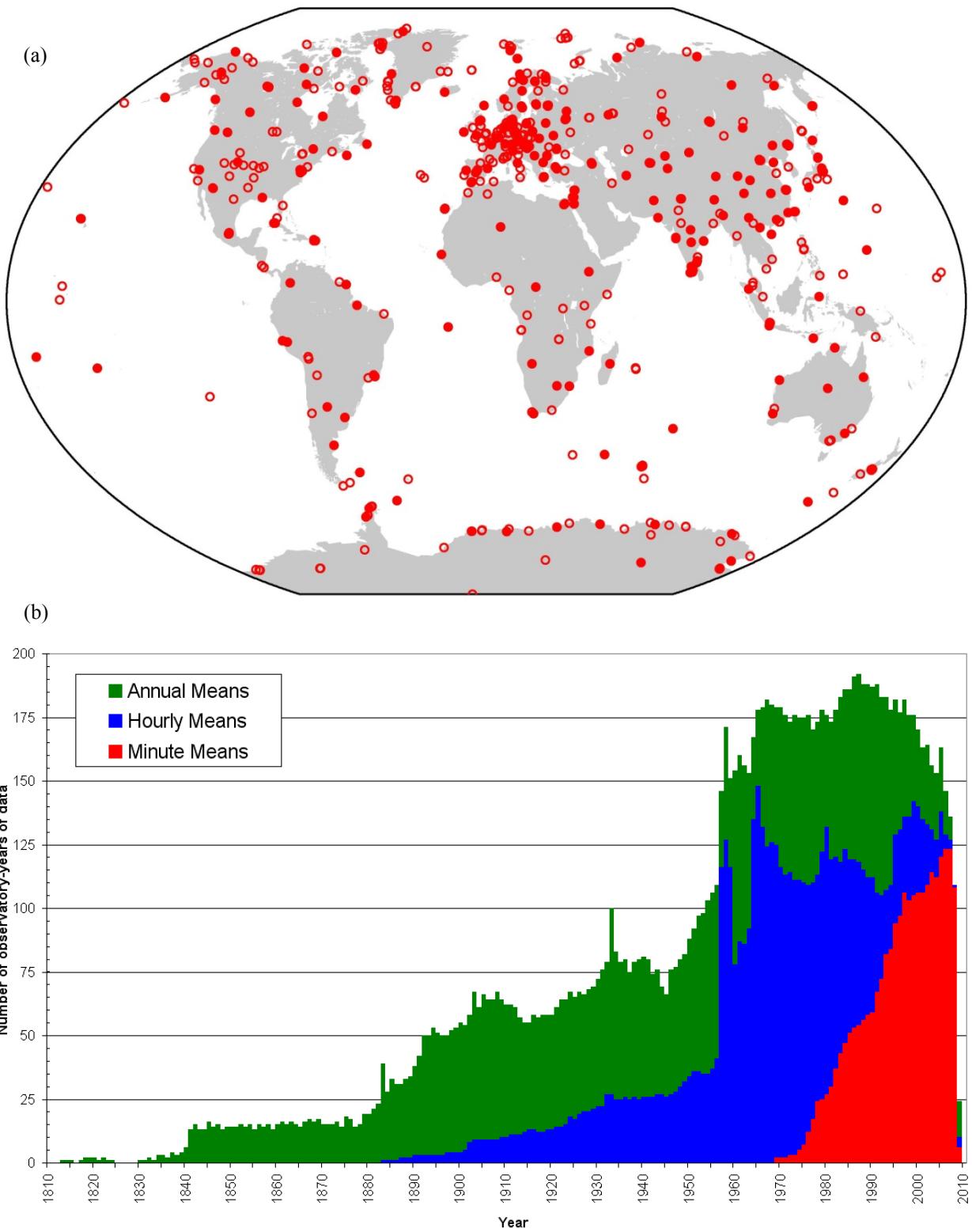


Fig. 7.7 **(a)** Locations of geomagnetic observatories worldwide, with data held in the Edinburgh WDC. Solid dots indicate operational observatories, open dots are sites that are currently closed **(b)** The number of data holdings at the World Data Centre (Edinburgh) for various time resolutions. (Copyright BGS/NERC)

observatories preserved on microfilm or microfiche; about 3200 years of tabulated hourly means from about 230 observatories; nearly 790 years of quick-run or pulsation magnetograms from about 90 observatories; approximately 1400 records of special events from about 90 different observatories; nearly 3300 years of K -indices from about 180 observatories; and on the order of 100 years of Q -index for about 22 observatories. The WDC-Copenhagen holdings also include data from 406 observatories mainly from the IGY onwards, also data from the Second Polar Year, 1932-1933. Most of the data are stored on microfilm, microfiche, and in printed publications.

7.3.1.4 World Data Centre for Geomagnetism, Mumbai

The Indian Institute of Geomagnetism (IIG), Navi Mumbai houses the WDC for Geomagnetism, Mumbai [8]. The centre is situated on the campus of old Colaba Magnetic Observatory, Bombay, which operated from 1841 to 1905. India's participation in the study of the earth's magnetism dates back to the period 1834-1841, when the country joined the Gottingen Magnetic Union. Full-fledged functioning of the WDC for Geomagnetism started at the IIG in 1991 under the recommendation of ICSU and the WDC Panel.

The centre holdings include magnetograms, geomagnetic hourly values, yearbooks of various observatories, digital data on magnetic tape, diskettes and CD-ROMs, microfilms and microfiche of the magnetograms, solar geophysical data bulletins containing solar rotation numbers; Kp , Ap , Cp and Zürich sunspot numbers, and indices such as Dst , AE , AU , AL , and aa . The centre is equipped with ample facilities for data retrieval, and these data are available in machine readable form upon written request or personal visit to the WDC.

7.3.1.5 World Data Center for Solar-Terrestrial Physics, Moscow

The WDC system in Russia consists of two facilities. The WDC-B was one of the original data centres established in the USSR by the Academy of Sciences of the USSR in 1957 to support the IGY. In 1971, the WDC for Solid Earth Physics (WDC for SEP) in Moscow was established as a part of WDC-B, and is maintained by the Geophysical Center of the Russian Academy of Sciences (RAS). WDC for SEP has concentrated on gathering data about the main magnetic field and its secular variations. A considerable part of these data are stored as publications on paper and microfilm, but some are also available in digital electronic form from their website [9]. Data stored at the WDC for SEP, Moscow include digital annual means values from

1813 onwards; publications of original land, marine and aeromagnetic survey observations; catalogues of measured values of the geomagnetic field elements, secular variation data, and charts of the geomagnetic field elements and its anomalies.

The other facility is the WDC for Solar-Terrestrial Physics (STP), also in Moscow [10]. The WDC for STP holds analogue data (mainly on microfilm and microfiche) from 1957 for the worldwide network of 264 magnetic stations as normal-run magnetograms, rapid-run magnetograms, hourly-, minute- and daily-mean values, as well as global magnetic activity indices and various publications. All standard geomagnetic data stored in the WDC for STP, Moscow, are accessible on the Moscow SPIDR (Space Physics Interactive Data Resource) website and mirrored worldwide (see section 7.3.5.1).

Additional data in non-standard formats are available on the website of the WDC for STP. Digital electronic data from the WDC for STP website include data from magnetic observatories in Russia and the Former Soviet Union; hourly-mean values for 38 observatories mainly from the IGY onwards; one-minute values from 41 observatories mainly from 1983 onwards; global magnetic activity indices (aa , Kp , Ap , AE , Dst , Pc); digital images of magnetograms beginning from 1957; ssc from 1868 onwards; and a catalogue of geomagnetic $Pc1$ pulsations at the Borok and Mirny observatories from 1957-1992. Participation in the ICSU "Rescue of Magnetograms Project" resulted in digital images of magnetograms from nine observatories covering over 100 observatory-years worth of data.

7.3.1.6 World Data Center for Solar-Terrestrial Physics, Boulder

WDC-A was one of the original data centres established in the United States by the Coast & Geodetic Survey (C&GS) in 1957 to support management of the full suite of IGY data. The National Geophysical Data Center (NGDC) was created in 1965 to assume responsibility for long-term management of the geophysical data, including the WDCs for Geomagnetism, Gravity, and Seismology. NGDC moved to Boulder, CO from Washington, DC in 1972, assuming responsibility for the WDC-A for Solar Activity hosted by the University of Colorado. Currently, NGDC operates two WDCs; the WDC for Solar-Terrestrial Physics (WDC-STP) [11] and the WDC for Geophysics and Marine Geology (WDC-GMG) [12].

The WDC-STP holds data sets relating to solar activity, the space environment, ionosphere, thermosphere, geomagnetism, and cosmic rays. Geomagnetic observations at magnetic observatories are maintained as one-minute, hourly, daily, monthly, and annual mean values. In addition to the

observatory data, the WDC-STP also manages magnetic variation, repeat observation, and land survey data, satellite data and global indices of magnetic activity. The WDC-GMG manages magnetic data collected at sea, aeromagnetic data, and develops and distributes magnetic field models, including the joint US/UK WMM and IAGA IGRF. The marine and airborne magnetic data are available through the GEODAS (Geophysical Data System) online system at [13]. Digital data are available online through their websites or by FTP. The WDC-STP also supports the SPIDR data portal that allows internet users to access, browse, display, and analyse STP data.

7.3.1.7 World Data Centre for Solar-Terrestrial Science, Sydney

The WDC for Solar-Terrestrial Science [14] is operated by IPS (Ionospheric Prediction Service) Radio and Space Services, a program within the Bureau of Meteorology of the Australian Government. The centre was established in 2000 and operates in Sydney, Australia. It contains ionospheric, magnetometer, spectrograph, cosmic ray data and solar images that are available for download via FTP. Currently WDC-Sydney holds magnetometer data for ten locations in Australia and two in Pakistan.

7.3.1.8 World Data Center for Geophysics, Beijing

The WDC for Geophysics [15] is operated by the Institute of Geology and Geophysics, Chinese Academy of Sciences. It was established in 1988 and is based in Beijing, China. With the support of the institute, Chinese Academy of Sciences and Ministry of Science and Technology, they have established online access to scientific data, such as geomagnetic data, gravity data, geoelectric field data and seismic wave data. This WDC holds one-minute and *K*-index data for Beijing Ming Tombs (BMT) observatory. Much effort has also been made to rescue historical magnetograms from 1877-1962 from Sheshan magnetic observatory. This collection has now been microfilmed and some magnetograms have been scanned and made available from their website (Peng, 2007).

7.3.2 INTERMAGNET

INTERMAGNET stands for *International Real-time Magnetic Observatory Network*. Its objective, to establish a “global network of cooperating digital magnetic observatories, adopting modern standard specifications for measuring and recording equipment, in order to facilitate data exchange and the production of geomagnetic data products in close to real-time” (St-Louis et al. 2008) was achieved and

further improvements continue to be the goal. The origin of the organisation and current operation are described in detail by Kerridge (2001) and Rasson (2007). Since its inception in 1986 INTERMAGNET has become the *de facto* organisation for setting standards for the highest quality observatory operations. Data types are defined as *Reported* (the raw data without any corrections), *Adjusted* (data that have some modifications made to it), and *Definitive* (the final published data), and these codes are used in the headers of the adopted data exchange formats. Instrument standards, such as accuracy and resolution were also established. Observatories that meet these standards and requirements are known as INTERMAGNET Magnetic Observatories, or IMOs.

In 1991, the USGS proposed that INTERMAGNET produce a CD-ROM of definitive one-minute data from IMOs, based upon a model the USGS had published (Herzog and Lupica 1992). Since then, INTERMAGNET has produced CD-ROMs, and now DVD-ROMs, annually. For 2007, the DVD-ROM contains data for 104 observatories from 41 countries. In 1997 a website was created [16] and this is now an important portal for the availability of both definitive and preliminary worldwide observatory data.

Following the 11th IAGA Scientific Assembly in Sopron, INTERMAGNET management resolved to make both one-second and quasi-definitive data available from all IMOs that are able to provide it (S Flower, personal communication, 2009). It was decided to restore all archived preliminary data to the website. The ability to provide commercial protection of data by time-limited distribution from the website will be introduced. The ‘gold standard’ for near real-time data delivery was agreed to be at 2-3 minutes delay and a discussion session was held to try to identify techniques that might be used to implement this standard.

7.3.3 World Data System

In 2006 ICSU established an ad-hoc Strategic Committee on Information and Data (SCID) to advise ICSU on the future organisation and direction of its activities in relation to scientific data and information. The SCID considered input from three ICSU Interdisciplinary Bodies: the WDC, the Federation of Astronomical and Geophysical Data analysis Services (FAGS) and the Committee on Data for Science and Technology (CODATA).

Despite the success of the WDCs, the review highlighted some basic problems that needed to be addressed. There was no real “system” in the sense of a managed or coordinated effort between the WDCs. While the ICSU WDC panel provided guidance, the WDCs have, for the most part, operated as individual institutions rather than a single entity, and the

concept of data interoperability was only minimally implemented. There are also large disparities between WDCs in the developed countries and those in developing countries, no central comprehensive directory or catalogue of data available at the WDCs and an increasingly urgent problem with old data holdings being at risk of decay, lost, or discarded (ICSU 2008).

In its review the SCID also considered the operation of the FAGS and CODATA. The FAGS [17] was formed by ICSU in 1956 as part of IGY. Their principal purpose has been to encourage the analysis of long-term data sets and produce data products for the scientific community. The review by SCID identified similar issues to the WDCs with the services being very much a product of history rather than by design. CODATA were established by ICSU in 1966. It provides a cross-disciplinary focus on scientific data and aims to improve the quality and accessibility of scientific data and the methods by which those data are acquired, managed, and analysed.

Following the review by the SCID, at the 2008 ICSU General Assembly, it was agreed to abolish the ICSU Panel on WDCs and create a World Data System (WDS) incorporating former WDCs with the FAGS, and other centres and services [18]. Those WDCs and FAGS currently in existence must now apply to ICSU to re-establish membership as a part of the WDS. The WDS will provide a coordinated and strategic response to the data needs of the global scientific community with international initiatives such as the International Polar Year (IPY).

The WDS will transition from a set of stand-alone WDCs, and individual Services, to a common globally interoperable distributed data system, while incorporating new scientific data activities. The new system will have advanced interconnections between data management components for disciplinary and multidisciplinary scientific data applications. WDS will have a broader disciplinary and geographic base than the previous ICSU bodies and will strive to become a worldwide ‘community of excellence’ for scientific data. Eventually the WDS plan is to implement a Global Data System of Systems (GDSS) that should be interoperable with current and emerging data systems such as Global Earth Observations System of Systems (GEOSS) [19] (Minster et al., 2009).

7.3.4 The International Service of Geomagnetic Indices

The International Service of Geomagnetic Indices (ISGI) is a collaborative international service established by IAGA and falling within the FAGS. ISGI collects, derives, validates, maintains and distributes time-series of geomagnetic indices and

other geomagnetic data products. The service is also responsible for the publication and distribution of any IAGA Bulletins containing information on geomagnetic indices.

The main ISGI operations are carried out at the Laboratoire Atmosphères, Milieux, Observations Spatiales (LATMOS), Paris (official derivation of *aa*, *am*, *Km*). The other ISGI collaborating institutes are: GeoForschungsZentrum (GFZ), Potsdam (official derivation of *Kp*, *Ap* and international quiet and disturbed days); Observatorio del Ebro, Roquetes (official derivation of *ssc* and solar flare effects (*sfe*)); and WDC-C2 for Geomagnetism, Kyoto (official derivation of *Dst* and *AE*). Data availability for the latter is described in more detail in section 7.3.1.2 and further information on the other institutes follows.

7.3.4.1 LATMOS, France

Established in 2009 following a reorganisation of scientific organisations in France, LATMOS has now replaced Centre d'étude des Environnements Terrestre et Planétaires (CETP) as the main ISGI host. As well as having the primary responsibility for the derivation of the *aa*, *am* and *Km* indices, LATMOS also maintains the ISGI database of the IAGA endorsed indices and operates the ISGI publication office and the main ISGI website [20]. As well as providing online access to the ISGI data, the website contains valuable information (metadata) on the derivation and meaning of the indices. The definitive indices available online are: *aa*, *Kpa* (from 1868); *am*, *Kpm* (from 1959); *Dst* (from 1957); *AU*, *AE* and *AL* (from 1957); *ssc* (from 1868); international quietest days (from 1932); and *Kp* (from 1932). Monthly bulletins are also produced, which contain various provisional indices and include graphs of musical diagrams of *Km* and *aa*. These are circulated by post as well as being made available on the website.

7.3.4.2 GeoForschungsZentrum, Germany

Helmholtz-Center Potsdam, the German Research Center for Geosciences or GFZ is the national German research centre for the Earth sciences. The Adolf-Schmidt-Observatory Niemegk, belonging to GFZ, derives and maintains the *Kp*, *ap* and *Ap* indices as well as the lesser known planetary indices, *Cp* and *C9*. Definitive indices are produced twice a month and quick look versions are provided in near-real time. As well as the planetary indices, GFZ's ISGI responsibilities include the selection of the international quietest days (*Q-days*) and most disturbed days (*D-days*) of each month, which are classified using the *Kp* indices. These, the aforementioned planetary indices and *ssc* data provided by Observatorio del Ebro are all available

on the GFZ website [21] in various formats, including ASCII tables, histograms and Bartels music diagrams. They are also available to download directly from the GFZ anonymous FTP site.

7.3.4.3 Observatorio del Ebro, Spain

The Ebro Observatory is a research institute founded by the Society of Jesus in 1904 to study the Sun-Earth relationships. It is a non-profit organisation associated with the Spanish Research Council (CSIC). Since 1975, Ebro Observatory was entrusted by IAGA to host the International Service on Rapid Magnetic Variations. It has responsibility for collecting, creating and publishing the official IAGA lists of rapid variations, which consist of *ssc* and *sfe* records. Annual lists of *ssc* data are available online from 1868 and *sfe* data from 1995 [22]. Monthly updates are provided to LATMOS and GFZ for further distribution.

7.3.5 Other Data Resources

7.3.5.1 Space Physics Interactive Data Resource

The SPIDR [23] is a data portal designed to allow a solar terrestrial physics customer to intelligently access and manage historical space physics data for integration with environment models and space weather forecasts. SPIDR is a distributed network of synchronous databases, web-portals and web-services. There are SPIDR sites in Boulder, Paris, Nagoya, Sydney, Beijing, Kiev, and Cape Town. SPIDR databases include geomagnetic one-minute, hourly means and annual means, as well as global geomagnetic and solar indices.

7.3.5.2 Observatory Operator's Websites

Many institutions operating magnetic observatories in the modern era have their own websites. This is an additional resource providing access to geomagnetic observatory data and data products to the user community. Since the advent of digital data distribution in the 1990's, observatories might not necessarily submit their data to a WDC as recommended by IAGA. It is recognised (Mandea and Papitashvili, 2009) that the traditional "push data" approach (to a WDC for example) is gradually being replaced by a "pull data" approach via data mining techniques such as virtual observatories (VO).

7.3.5.3 Variometer Networks

In addition to the data resources mentioned thus far, where full-field definitive observatory data and activity indices can be obtained, there are also a large number of variometer stations located around the globe providing additional data resources. Unlike magnetic observatories, variometer stations record

only the short-term variations in the magnetic field and not the absolute level. These instruments complement the network of magnetic observatories and these measurements are used especially in the study of solar-terrestrial science.

There is a large amount of data (past and present) from many different magnetometer chains or networks around the world. A worldwide collaborative effort to pull these together was initiated and a single source of ground based magnetometer data is now available at SuperMAG [24]. This data portal website provides access to variometer data from more than 200 stations worldwide. Data are available from 1997 and can be viewed and downloaded online as linear plots and maps.

SuperMAG relies on acquiring data from a variety of sources including magnetometer networks such as IMAGE [25] (International Monitor for Auroral Geomagnetic Effects) and CARISMA [26] (Canadian Array for Realtime Investigations of Magnetic Activity). IMAGE is a network of 31 magnetometer stations operated by 10 institutions from Estonia, Finland, Germany, Norway, Poland, Russia and Sweden. CARISMA operate an array of 25 stations throughout Canada. INTERMAGNET also contributes data to SuperMAG since magnetic observatories can double as variometer stations. There are dozens of similar magnetometer networks located across the globe and are too numerous to mention here. Most of these are collaborative ventures in their own right, relying in turn on the individual institutes operating the stations.

7.4 Metadata and Metadata Standards

Metadata are an increasingly important aspect of geomagnetic data provision, both for the WDCs and for other data providers. Metadata are information about data; they are the "who, what, when, where, why and how" of a data set. They describe the content, quality, originator, and other characteristics of a data set that help users understand the nature of that data set and how to use it. As time passes, and personnel, instrumentation, and data processing procedures change, it is important that current and future generations have sufficient information about the data to enable them to independently understand and use them. The objective is to provide documentation that ensures that the values contained in a data set have the necessary reference information, and that even non-experts in a discipline will be able to use the data properly.

As Buneman (2005) notes, the web has radically changed the way scientific research is carried out, with more rapid and varied access to data, but this

has produced new issues in maintaining the scientific record. As data can readily be copied and altered we need some means of verifying authorship and data provenance for researchers. For data producers too the issue of data citation is a growing concern. Wayne (2005a; 2005b) demonstrates the many benefits of metadata but also some of the obstacles to its production. Adhering to a metadata standard can be complex and time-consuming and there may be little immediate, tangible benefit to data producers. However the potential benefit is great.

The subject of metadata was raised at the IAGA conference in Sopron, and at the INTERMAGNET meeting that followed. Reay et al. (2009) highlighted the need for good data provenance to assist scientific examination of geomagnetic data sets. Better metadata would lead to improvements in the curation of data at the WDCs; addressing any inconsistencies seen in data sets and providing a clear ‘paper-trail’ of any transformations or corrections to data. It would also provide clear quality assurance to researchers, assisting data selection for global field models. In addition; common, well developed metadata improves the “discoverability” of data through the many metadata clearinghouses.

Svalgaard and Cliver (2007) discuss the current issues of data quality within WDCs and note that metadata about, especially, historical data is sorely lacking. They also point out the current difficulties in feeding back any corrections into the WDCs. If comprehensive metadata records were associated with a dataset a researcher could trace all the data processing steps applied to a data set, more easily identifying any inconsistencies. This would also provide a means for archivists to document any subsequent corrections made.

Consider another scenario: say, if you wanted to gather all geomagnetic data recorded by a fluxgate magnetometer, spanning 1990-2000, located within N30-50 latitude range. Currently this would require a significant amount of investigation to determine which observatories this applied to and where to get this data. If this simple metadata was recorded and stored in a central database in a known metadata standard this type of query would take seconds.

At the Sopron assembly, IAGA recognised the importance of metadata preservation in supporting geophysical studies (Resolution 7) and encouraged relevant agencies to support the generation, preservation and dissemination of metadata to ensure the future usability of these data for interdisciplinary studies. A session on metadata has been proposed for the next IUGG General Assembly in 2011.

INTERMAGNET has assigned a representative to work with groups in Japan, the UK, the US, and elsewhere on the issue of station-level metadata requirements and proposals. INTERMAGNET will notify the IMO's that a proposed standard would be

forthcoming, and they should prepare to incorporate a new element into the observatory operations. It was also suggested that observatories place more effort on generating yearbooks with the omission of magnetogram plots if that would facilitate their production. The WDCs are also considering this issue and a wiki has been established to further coordinate discussions on this matter.

There are many issues to be considered in the development of a metadata standard for magnetic observatories. In the following sections we discuss some of the types of metadata currently available, or in development, that might provide a solution to capturing this important information resource.

7.4.1 Magnetic Observatory Metadata

7.4.1.1 Yearbooks

Yearbooks provided the earliest form of metadata for geomagnetic observatories, and are still being produced by many institutions today. These contain a permanent record of the status of the observatory for a given year and generally contain information about the station location, instrumentation, contact persons and record significant events or changes to the observatory throughout the year. These are published alongside the actual data products reported from an observatory.

Traditionally these were produced as printed books and these are stored in libraries, institutes and WDCs across the globe. Nowadays yearbooks are often produced in electronic PDF format that can be found from the websites of some institutions, WDCs, and INTERMAGNET [27]. Additionally some efforts are being made to scan historical yearbooks allowing electronic access to these important records.

Whilst the metadata held in yearbooks are often precisely what we wish to capture in a metadata standard, the format that this information is held is not currently standardised and is not suitable for data discovery applications. Principally, observatory operators should be encouraged to record this metadata, in whatever form, and then those within the geomagnetism community engaged in data curation must consider how to transfer this into an established metadata format. Improvements in data mining may be a way forward in this respect.

7.4.1.2 INTERMAGNET

The best source for digital observatory metadata, at least for those observatories within the organisation, is INTERMAGNET. Each year INTERMAGNET publish annual definitive observatory results on a CD/DVD-ROM. Alongside the data a *Readme* file is supplied for each observatory detailing basic information on location, instrumentation, sampling, filtering, contacts etc. This ASCII text file is in a

standardised format with mandatory and optional fields. As these metadata records are directly related to a single published dataset they also provide temporal metadata that is challenging to capture otherwise.

Beyond this published metadata there is much implicit metadata associated with classifying an observatory as an IMO. If an observatory meets INTERMAGNET standards you can infer much information about its data processing, quality assurance, data distribution and instrumentation standards. However to a non-expert most of this information is not readily apparent: to access this type of metadata you would need to refer to the Technical Reference manual (St-Louis et al. 2008) held separate to the data.

7.4.1.3 Metadata Standards

Despite the wealth of metadata, in yearbooks and from INTERMAGNET, there is currently no metadata standard that can adequately describe data from magnetic observatories. With metadata, standardisation is important because it provides a common set of formats, terminology and definitions that facilitates the use and exchange of metadata digitally. It provides uniformity and consistency over time, enabling the development of tools and resources for metadata exchange and manipulation. One might say that metadata standards are to metadata what data formats are to data.

Considering what form a metadata standard for magnetic observatory data may take, we can suggest some of the basic information that would be required. These include, but are not limited to:

- *Contact Information* - The name of the responsible institute that produces the data, including addresses and personal contacts.
- *Data Description* - A description of the type of data, including a narrative summary of the nature of the data and its possible applications and a list of entities and attributes relevant to the data.
- *Station Description* - A description of the station in question, including coordinates, elevation, and possibly photographs and maps.
- *Instrumentation* - A description of the types of instruments in use at the observatory, and the nature of the data that are retrieved from them.
- *Data Processing* - A description of the processes and methodology used to process the data from instrument recordings to the final definitive values. This should include a description of data formats and how the numerical values are to be interpreted.
- *Data Quality* - A general assessment of the quality of the data set considering completeness, accuracy, quality control methods and so on.

- *Data Distribution* - A description of how and from where the data may be acquired, particularly through the WDCs.

To help us towards a metadata standard for geomagnetic observatory data we must look toward those standards currently in use for other, similar, datasets. Metadata standards established for geospatial data may act as such a guide. While these standards address many of the basic requirements of our metadata (for data discoverability, distribution, etc.) they lack the scope to deal with parameters that change with time. For an observatory nothing is fixed for all time: instrumentation, processing techniques, responsible institutions, staff contacts, data types, data quality, and even the observatory location can change. How to address this issue is one the greatest challenges facing metadata standardisation.

7.4.2 FGDC Standard

One of the first standards produced for geospatial data was developed by the Federal Geographic Data Committee (FGDC) in the US [28], which was designed primarily as the template for storing and distributing geographic spatial metadata. The FGDC standard consists of 7 information sections:

1. *Identification* - Basic information about the data set and originator.
2. *Data Quality* - General assessment of the quality of the data set.
3. *Spatial Data Organisation* - Used to represent data set spatial information.
4. *Spatial Reference* - Reference frame for data set coordinates.
5. *Entity and Attribute* - Details about the information content.
6. *Distribution* - Distributor information and how to obtain the data set.
7. *Metadata Reference* - Metadata responsible parties and current status.

This standard is complex, requiring the data provider to fill in many metadata fields to comply. These are often in a terminology that is unclear to non-experts in the standard. The FGDC standard was developed primarily for data products that have a single publication date, such as maps, and consequently has limitations with regard to parameters that change in time - an important factor for observatory metadata. One would have to write a separate record for each change that occurs.

The NGDC has produced a modified standard for observatory data based on FGDC standard, omitting sections that are not applicable to magnetic observatories (Fischman et al., 2009). Within SPIDR, authorised institute personnel can now enter metadata for their own observatories using this modified

FGDC template. SPIDR provides online forms corresponding to the various sections of the modified FGDC standard, from which observatory-level metadata records in the form of XML files are generated. While this goes some way to capture the required metadata it is still overly complex for observatory data providers to complete, does not solve the problem of tracking observatory histories, and has not been well received by the geomagnetic community as yet.

7.4.3 ISO-19115 Standard

In 2003 the International Standards Organisation (ISO) Technical Committee 211 on Geographic Information/Geomatics (ISO/TC 211) released the metadata standard ISO-19115. This combined various elements of national geospatial metadata standard including the FGDC (USA), CEN/TC287 (Netherlands), and ANZLIC (Australia & New Zealand), as well as contributions from other members of the Open Geospatial Consortium [29].

The FGDC mandatory fields and ISO core metadata fields are quite similar, although ISO-19115 offers the possibility of more detail, and includes special coverage of raster and imagery information. While the FGDC standard has 7 informational sections, ISO-19115 has 14 comprising more than 400 individual metadata elements. As with the FGDC standard, most of these are optional and only a few of the available elements are likely to be used. While this ISO-19115 standard offers greater complexity and detail, it has the same issues as the FGDC standard with no capacity to track time-changing parameters at observatories.

7.4.4 SPASE Data Model

In 2006 NASA funded five new VOs to cater for different aspects of solar system science. A VO is a complex distributed environment with a goal to provide a single point of discovery of data and related resources (Merka et al., 2008). Groups, including the Space Physics Archive Search and Extract (SPASE)

```
<instrumentTypeUsageInstance>
  <beginDate>1975-05-09</beginDate>
  <vectorMagnetometerType>EDA Fluxgate magnetometer</vectorMagnetometerType>
  <endDate>1981-12-31</endDate>
</instrumentTypeUsageInstance>
```

The angle brackets surround a tag name, with the backslash used to indicate the end of an XML element. In this example *instrumentTypeUsageInstance* is an element that identifies a period of time when a particular type of instrument was in use. The *vectorMagnetometerType* element describes the type of instrument in use and the time-specific elements *beginDate* and *endDate* define the boundaries of the period when this

consortium [30], are defining metadata standards to aid in archiving and sharing of information resources from VOs.

The SPASE data model provides enough detail to allow a scientist to understand the content of the solar-terrestrial data products together with essential retrieval and contact information (King et al., 2010). The SPASE metadata schema divides the solar-terrestrial environment into a limited set of 12 resource types. Most commonly used are Display Data, Numerical Data, Granule, Instrument, Observatory and Person. These resources are interconnected to create a network of resources and the relationship between these resources fully describes a dataset. VOs can harvest these resource descriptions and allow users to find data.

As the SPASE data model manages the interrelationship between resources, such as observatory and instrument, it may prove to be better placed to manage the issue of time-varying information in magnetic observatory data.

7.4.5 XML

XML stands for eXtensible Markup Language and was designed to store and transmit structured data in plain text. It allows users to define their own vocabulary using tags to mark-up data in a similar way to how HTML is used to mark-up web pages. XML has become the de facto standard for the exchange of information, in part, because of this flexibility. A user can develop any set of properties to define the elements of a data set and describe that data set using corresponding tags. This is known as creating an *XML schema*. Data and metadata can be stored in a document conforming to that XML schema, and be exchanged and read by both computers and humans.

For example, consider an XML document used to represent an instance of an instrument type in operation at a magnetic observatory:

instrument was in use. Note also that XML elements may be nested to reflect the structure of the data.

By constructing an XML schema that fully encapsulates magnetic observatory metadata we begin to define a type of metadata standard. This standard would be able to describe time-varying aspects of an observatory's operation by defining appropriate elements, which is an advantage over using current established standards. By using XML

as a base, we can define the metadata and the method of storage and exchange at the same time.

XML is primarily used to exchange information between computers. XSLT (Extensible Stylesheet Language: Transformations) is a tool that allows XML to be transformed into various formats such as plain text, HTML or into XML of a different schema which may be more appropriate for the end-user [31]. For example the SPIDR system stores metadata in XML files and uses XSLT to transform this information into HTML for display to the user.

7.4.6 Databases

Another possibility being pursued at the NGDC, and reported on in Sopron, is the construction of a database schema to store the relevant observatory parameters with time histories of these parameters as fields in the database tables (Mabie et al., 2009). For example, changes in instrumentation from one time period to another (start time/stop time) can be stored in the database tables, and retrieved when a user requests metadata on that instrumentation within a given time-frame. The advantage of this approach is that web pages can be developed that provide selectable metadata categories (e.g. contact information, changes to instrumentation, etc.) to accompany data sets requested by users. If a user requests data from a certain observatory for a given interval, it can be designed so that the user may also request the metadata on instrumentation, institution address, data quality, or any other metadata category deemed of interest to the user. A menu of choices could be designed so that the user could select the kind of metadata of importance to them. The user would not have to search for the metadata separately, but could download it as an accompanying file with the data itself.

Whilst the XML and database approach may allow for the complete capture of observatory metadata this is still one step away from an established international metadata standard. Translation algorithms will be required to convert metadata held in a local form or within a magnetic observatory database/XML schema into an accepted standard when one is established.

7.5 Metadata Distribution

Once a metadata standard has been established there are many other issues to be addressed. We must consider how this metadata is stored and managed and where it can be accessed from. Should multiple sites distribute the metadata, and how would we handle consistency between sites? How would we collect and ingest metadata records? Although observatory operators will have the required information it may require a third-party to translate

this information into the established metadata standard.

Beyond establishing a metadata standard these and many other issues must be addressed in time. For now, we will look at some selected examples of how metadata records for various geophysical data are distributed. All these metadata portals require a manual process to obtain the metadata separate from acquiring the data itself. They all lack the desirable feature of obtaining metadata simultaneously with data.

7.5.1 SPIDR VO (USA)

The SPIDR is a distributed network of synchronised web-accessible databases, developed by the NGDC, which provides users with access to current and historical STP data and metadata. SPIDR allows users to login as a guest or register with a username and password. After selecting the *geomagnetic view* option, the user can then select *view/modify stations metadata*. After selecting an observatory the user is presented with the modified FGDC metadata record for that observatory.

Each institution can assign a person responsible for maintaining and updating the metadata associated with the data from their own observatories. That person registers with SPIDR, requesting permission to access and edit the metadata records for those observatories.

As this system is based on the modified FGDC standard it has the same drawbacks regarding observatory metadata as discussed in section 7.4.2.

7.5.2 GeoMIND (Europe)

GeoMIND (*Geophysical Multilingual Internet-Driven Information Center*) [32] is an internet site dedicated to information about geophysical systems in Europe. It is a consortium of 12 organisations from 9 different countries, and provides metadata on geophysical data holdings from these countries in a consistent and seamless way, in whatever language the user selects. The website allows a user to identify a data set and its general content, and then be informed about how to obtain the data itself. This system allows for searching and editing of metadata records, as well as the importing and exporting of records in a variety of formats.

Whilst geomagnetic monitoring is included as a geophysical record within GeoMIND it is not currently well populated with metadata. The metadata are available on the web in the form of XML files conforming to ISO-19115. Only basic contact, citation and data distribution metadata are currently covered.

7.5.3 GEOMET (Australia)

GEOMET is the name of a database developed by Australian Geosciences that holds metadata records in the ANZLIC standard. ANZLIC (Australia New Zealand Land Information Council) is the metadata standard adopted by Australia and New Zealand that provides a schema required for describing geographic information, products, and services. It is similar to the FGDC standard in the parameters it defines, giving information about the identification, geographical boundaries, the quality, the spatial and temporal limits, spatial reference, and distribution of digital geographic data. ANZLIC assigns a unique identifier for each metadata record, which includes the usual metadata parameters such as an abstract describing the data set in general terms; geographic bounding coordinates; data quality; and contact information. ANZLIC has been made compatible with the ISO-19115 standard for geospatial metadata. Again, as this is similar to the FGDC and ISO-19115 standards, it will suffer from the same issues concerning geomagnetic observatory data.

7.5.4 IUGONET (Japan)

IUGONET (Inter-university Upper atmosphere Global Observation NETwork) [33] is a cooperative program in Japan between the National Institute of Polar Research, Tohoku University, Nagoya University, Kyoto University, and Kyushu University to expand the global radar, magnetometer and optical observation equipment and ground-based network in order to enhance each agency's effective use of various observational data sets.

Because of the diversity of the data from upper atmosphere observations, the program is developing a database for the metadata, which will include observation times, location, equipment, data formats, etc. to be made available on the internet. The metadata will make the data understandable to all researchers, and promote interdisciplinary research by a variety of different agencies, including those in fields other than that of upper atmosphere science. IUGONET will handle metadata in XML in the IUGONET format based on SPASE. Figure 7.8 shows an example of such metadata for the HF radar.

```

<?xml version="1.0" encoding="UTF-8"?>
<Spase lang="en" xmlns="http://www.spase-group.org/data/schema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.iugonet.org/data/resouces/iugonet-
1.0.0.xsd">
  <Version>1.0.0</Version>
  <NumericalData>
    <ResourceID>
      spase://IUGONET/NumericalData/SuperDARN/HOK/common_erg_cdf
    </ResourceID>
    <ResourceHeader>
      <ResourceName>
        SuperDARN Hokkaido HF radar, common mode data distributed by
        ERG-SC
      </ResourceName>
      <ReleaseDate>2009-04-01T00:00:00</ReleaseDate>
      <ExpirationDate>2199-12-31T23:59:59</ExpirationDate>
      <Description>
        Common mode data generated by SuperDARN Hokkaido HF radar. Data
        files are distributed in the CDF format through ERG-SC
      </Description>
      <Contact>
        <PersonID>spase://IUGONET/Person/Nozomu.Nishitani</PersonID>
        <Role>PrincipalInvestigator</Role>
      </Contact>
      <Contact>
        <PersonID>spase://IUGONET/Person/Kanako.Seki</PersonID>
        <Role>DataProducer</Role>
      </Contact>
      <Contact>
        <PersonID>spase://IUGONET/Person/Tomoaki.Hori</PersonID>
        <Role>MetadataContact</Role>
      </Contact>
    </ResourceHeader>
  .....

```

Fig. 7.8 An example of metadata archived by the metadata database of IUGONET. All metadata are described as XML in the IUGONET format based on the SPASE model

7.5.5 GeoNetwork (Open Source)

The GeoNetwork project [34] is an open-source and standards-based application for managing and delivering spatially referenced information and other resources via the internet. The software is intended to provide access to a wide variety of data, metadata, and other kinds of information, obtained from multiple disciplines, and organised and documented in standardised ways. It provides the ability to manage and administer access to geospatial databases containing data and related metadata. It provides support for ISO-19115, FGDC and Dublin Core metadata standards. It also incorporates various services such as access to ESRI-based servers running the ArcIMS protocol. There are many agencies that use the GeoNetwork open source system to manage their metadata, including the Group on Earth Observations [35].

7.6 Conclusion

In this chapter we have reviewed the range of geomagnetic observatory data and indices that are available to the scientific community. The types and availability of this data are summarised in Table 1. We have seen that the type of data recorded and supplied by observatory operators is not fixed for all time. Extracting information from historical printed data products are of interest to those examining trends in long time-series data and current efforts worldwide to digitally capture this important information resource should be lauded. Furthermore, exciting developments in the supply of one-second and quasi-absolute data should have a significant impact on the range and type of scientific analysis that can be achieved.

Curation of these important data sets for our and future generations is of utmost importance. This task has been carried out with great care for over 50 years by the WDCs providing long-term data stewardship and free and open access to all. This is an interesting

| | | Data Sources | | | | | | | | | | | |
|--------------------------|---------------------------------------|--------------|----------|-------------|-------------|-------------|------------|-----------------|-----------------|----------------|----------------|-----------|---------------|
| | | ISGI-Ebro | ISGI-GFZ | ISGI-LATMOS | INTERMAGNET | WDC-Beijing | WDC-Sydney | WDC-GMG Boulder | WDC-STP Boulder | WDC-STP Moscow | WDC-Copenhagen | WDC-Kyoto | WDC-Edinburgh |
| Digital Observatory | One-Second | ■ | ■ | - | - | - | - | - | - | - | - | - | - |
| | One-Minute | ■ | ■ | - | - | ■ | - | ■ | - | ■ | - | ■ | ■ |
| | Hourly Values | ■ | ■ | - | ■ | ■ | - | ■ | - | - | - | ■ | ■ |
| | Annual Means | ■ | - | - | - | - | ■ | ■ | - | - | - | - | - |
| Indices | Planetary (<i>K</i> -derived) | ■ | ■ | ■ | ■ | ■ | - | ■ | - | - | - | ■ | ■ |
| | Storm (<i>Dst</i> , <i>ASY/SYM</i>) | - | ■ | - | ■ | ■ | - | ■ | - | - | - | ■ | - |
| | Auroral (<i>AE</i>) | - | ■ | - | ■ | ■ | - | ■ | - | - | - | ■ | - |
| | Polar (<i>PC</i>) | - | - | ■ | - | ■ | - | - | - | - | - | - | - |
| | <i>SSC/SI</i> | ■ | - | - | - | ■ | - | - | - | - | - | ■ | ■ |
| | <i>SFE</i> | ■ | - | - | - | - | - | - | - | - | - | - | ■ |
| | Solar | ■ | - | - | ■ | - | ■ | - | ■ | - | - | - | - |
| | Other | ■ | ■ | - | - | ■ | - | ■ | - | ■ | - | - | - |
| | Repeat | ■ | - | - | - | - | - | ■ | - | - | - | ■ | - |
| Survey Data | Land/Air/Marine | ■ | - | - | - | - | ■ | ■ | - | - | - | - | - |
| | Global Magnetic Field Models | ■ | ■ | - | - | - | ■ | - | ■ | - | - | - | - |
| Digital Metadata Records | | ■ | ■ | - | - | - | ■ | - | ■ | - | ■ | ■ | - |

Table 1 A Summary of the type of geomagnetic data available and where this data is available from

time for data centres as we move towards a new World Data System to replace WDCs and provide a data service for the 21st century. INTERMAGNET and ISGI also play an important role in data availability, both in encouraging the establishment and improvement of standards at observatories and in the standardisation and distribution of data sets.

As data are now recorded and distributed electronically, important information about these data sets can become disassociated from the actual data. These metadata are vital to gain a full understanding of the data and any treatment applied to it. IAGA have recently recognised the importance of recording metadata and there are numerous efforts within the geomagnetism community to try and establish a way of capturing and distributing this important resource. This is one of the most pressing and challenging tasks in geomagnetic data curation today; but if we can achieve consensus and standardisation the scientific community will reap the benefits.

Acknowledgments The authors wish to thank David M. Clark, Dr. Hans-Joachim Linthe, Dr. Jürgen Matzka and Dr. Susan Macmillan for their valued contributions. We would like to thank Ellen Clarke for her critical review and many helpful additions to the manuscript. We would also like to thank Prof. Mioara Mandea and Dr. Monika Korte for the opportunity to prepare this chapter. This paper is published with the permission of the Director of BGS (NERC).

Glossary

ANZLIC - Australia New Zealand Land Information Council
BGS - British Geological Survey
CARISMA - Canadian Array for Realtime Investigations of Magnetic Activity
CDF - Common Data Format
CETP - Centre d'étude des Environnements Terrestre et Planétaires
CODATA - Committee on Data for Science and Technology
CSIC - Consejo Superior de Investigaciones Científicas (Spanish Research Council)
DAGSM - Data Analysis Center for Geomagnetism and Space Magnetism
DCP - Data Collection Platform
DMI - Danish Meteorological Institute
ERSI - Environmental Systems Research Institute
FAGS - Federation of Astronomical and Geophysical data analysis Service
FGDC - Federal Geographic Data Committee
FTP - File Transfer Protocol
GEODAS - Geophysical Data System
GIN - Geomagnetic Information Nodes
GFZ - GeoForschungsZentrum
HMV - Hourly Mean Value

IAGA - International Association of Geomagnetism and Aeronomy
ICSU - International Council for Science
IGRF - International Geomagnetic Reference Field
IGY - International Geophysical Year
IIG - Indian Institute of Geomagnetism
IPY - International Polar Year
IMAGE - International Monitor for Auroral Geomagnetic Effects
IMO - INTERMAGNET Observatory
INTERMAGNET - International Real-time Magnetic Observatory Network
ISGI - International Service of Geomagnetic Indices
ISO - International Standards Organisation
IUGONET - Inter-university Upper atmosphere Global Observation NETwork
LATMOS - Laboratoire Atmosphères, Milieux, Observations Spatiales
NERC - Natural Environment Research Council
NGDC - National Geophysical Data Center
OCR - Optical Character Recognition
SEP - Solid-Earth Physics
sfe - Solar Flare Effects
SPASE - Space Physics Archive Search and Extract
SPIDR - Space Physics Interactive Data Resource
ssc - Sudden Storm Commencements
STP - Solar-Terrestrial Physics
USGS - United States Geological Survey
VO - Virtual Observatory
WDC - World Data Centre
WDS - World Data System
WMM - World Magnetic Model
WWW - World Wide Web
XML - eXtensible Markup Language
XSLT - eXtensible Stylesheet Language Transformations

Appendix 1: Data File Formats

With the advent of digital data records it was soon recognised that a common data format for storing and presenting geomagnetic observatory was required. Without data standards it would be impossible for users of worldwide observatory data to make use of the vast collection of data available from many different institutions and countries.

Over time different data formats have been established with different motivations behind each approach. The most prevalent standard for one-minute data is currently IAGA-2002 whereas WDC Exchange Format is still often used for hourly mean data.

World Data Centre Exchange Format

In the 1980s a common file format for dissemination of geomagnetic data was developed. This WDC Exchange Format [36] was a continuation from the

format for punched card and was designed to make maximum use of the limited RAM and disk capacity of computers at that time. It is not a convenient format for the user and there is no space for any metadata. For hourly means each value is decomposed into a tabular base value (shared by all hourly values for a given day), and a tabular value (Fig 7.9). For each day the tabular base and the 24 values are presented with a daily mean value at the end of the row. For minute means each value in an hour is displayed along one horizontal row with an hourly mean value given at the end.

INTERMAGNET GIN Dissemination Formats

INTERMAGNET defined a number of format standards for the dissemination of data from their GINs [37]. There are standards for one-minute (IMFV1.22), hourly (IHVF1.01) and daily (IDFV1.01) data. For example, for one-minute means, data are organised on a day-file basis. One file contains 24 one-hour blocks, each containing 60 minutes worth of values. Blocks are padded with 9's if incomplete. Information is encoded in ASCII and there is only very limited metadata included e.g. noting if data are reported, adjusted or definitive.

```
ESK DEC0109 335 00 HDZF R EDI 03473568 000000 RRRRRRRRRRRRRRRR
174920 -022201 463428 495316 174931 -022205 463427 495320
174937 -022207 463427 495321 174928 -022201 463427 495318
174925 -022201 463427 495317 174931 -022209 463427 495320
...
ESK DEC0109 335 01 HDZF R EDI 03473568 000000 RRRRRRRRRRRRRRRR
174923 -022227 463424 495313 174922 -022228 463424 495314
174921 -022230 463424 495314 174924 -022232 463424 495313
174924 -022232 463424 495313 174924 -022232 463423 495314
...
ESK DEC0109 335 02 HDZF R EDI 03473568 000000 RRRRRRRRRRRRRRRR
174928 -022221 463420 495312 174926 -022223 463419 495312
174926 -022225 463420 495311 174924 -022223 463420 495310
174923 -022223 463420 495311 174924 -022221 463421 495311
```

IAGA-2002 Format

The IAGA ASCII Exchange Format [38] was adopted in 2001. It is used as a data exchange format for geomagnetic data (samples and means) from observatories and variometer stations at cadences from milliseconds up to and including monthly means. This ability to accommodate data of different cadence is a major strength of this file format. It is also flexible and can account for files containing different durations of data (e.g. a month file of one-minute data) with its well-defined file naming scheme. The file names themselves provide critical metadata and within each file there are twelve mandatory file header records for further metadata. These include the type of data (definitive, provisional, variation), if the data are instantaneous or a mean, and how the values are centred. It also has an optional, variable length comment field for more descriptive metadata associated with the dataset. The data records are presented in a fixed four-column ASCII format which is convenient for the user to read and manipulate.

| Format | IAGA-2002 | | | | |
|---|-----------------------------------|----------|----------|----------|----------|
| Source of Data | Danish Meteorological Institute | | | | |
| Station Name | Narsarsuaq | | | | |
| IAGA CODE | NAQ | | | | |
| Geodetic Latitude | 61.160 | | | | |
| Geodetic Longitude | 314.560 | | | | |
| Elevation | 4 | | | | |
| Reported | XYZF | | | | |
| Sensor Orientation | DIF | | | | |
| Digital Sampling | 0.01 seconds | | | | |
| Data Interval Type | Filtered 1-minute (00:30 - 01:29) | | | | |
| Data Type | Definitive | | | | |
| # This area is where the data source or distributor can include | | | | | |
| # any additional information needed for proper use of data. | | | | | |
| DATE | TIME | DOY | NAQX | NAQY | NAQZ |
| 2001-03-13 00:00:00.000 | 072 | 10800.11 | -6100.23 | 53381.51 | 54801.12 |
| 2001-03-13 00:01:00.000 | 072 | 10800.31 | -6100.20 | 53381.51 | 54801.12 |
| 2001-03-13 00:02:00.000 | 072 | 10801.11 | -6101.23 | 99999.00 | 54801.12 |
| 2001-03-13 00:03:00.000 | 072 | 10803.12 | -6100.23 | 99999.00 | 54801.12 |

Tabular bases D, H & Z mean values for hour 1, day 1, month 1, 2008 are 345 deg 47.0 min, 12575 nT and 50161 nT

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|------------|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| LRV0801D01 | 20 | 345 | 470 | 436 | 429 | 418 | 425 | 447 | 459 | 447 | 437 | 434 | 432 | 423 | 413 | 411 | 404 | 412 | 415 | 418 | 423 | 427 | 433 | 440 | 462 | 441 | 432 | | |
| LRV0801D02 | 20 | 345 | 440 | 437 | 436 | 435 | 430 | 435 | 428 | 433 | 436 | 437 | 436 | 432 | 418 | 406 | 403 | 406 | 415 | 412 | 415 | 423 | 429 | 434 | 439 | 443 | 427 | | |
| ... | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| LRV0801H01 | 20 | 120 | 575 | 571 | 584 | 582 | 589 | 592 | 595 | 595 | 593 | 591 | 589 | 589 | 589 | 589 | 591 | 593 | 592 | 593 | 592 | 591 | 590 | 589 | 589 | 588 | 587 | 588 | |
| LRV0801H02 | 20 | 120 | 592 | 586 | 585 | 588 | 590 | 592 | 594 | 596 | 595 | 596 | 593 | 590 | 588 | 587 | 588 | 591 | 594 | 597 | 595 | 592 | 594 | 591 | 591 | 588 | 591 | 591 | 591 |
| ... | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| LRV0801Z01 | 20 | 500 | 161 | 144 | 148 | 166 | 179 | 180 | 177 | 176 | 179 | 180 | 181 | 182 | 183 | 183 | 185 | 186 | 186 | 185 | 184 | 185 | 188 | 189 | 182 | 178 | | | |
| LRV0801Z02 | 20 | 500 | 182 | 181 | 182 | 184 | 183 | 182 | 183 | 182 | 181 | 182 | 182 | 183 | 183 | 184 | 185 | 185 | 185 | 188 | 191 | 189 | 189 | 186 | 184 | 184 | | | |

Fig 7.9 WDC hourly mean value format, with the tabular base values highlighted

Appendix 2: Internet Links

- [1] <http://cdf.gsfc.nasa.gov>
- [2] <http://www.wdc.bgs.ac.uk>
- [3] http://www.geomag.bgs.ac.uk/data_service/home.html
- [4] <http://wdc.kugi.kyoto-u.ac.jp>
- [5] <http://wdc.kugi.kyoto-u.ac.jp/film>
- [6] <http://www.space.dtu.dk/English.aspx>
- [7] http://www.space.dtu.dk/English/Research/Scientific_data_and_models/World_Data_Center_for_Geomagnetism.aspx
- [8] <http://www.wdciiig.res.in>
- [9] <http://www.wdcb.ru/sep>
- [10] <http://www.wdcb.ru/stp/index.en.html>
- [11] <http://www.ngdc.noaa.gov/stp/WDC/wdcstp.html>
- [12] <http://www.ngdc.noaa.gov/mgg/wdc/wdcgmg.html>
- [13] <http://www.ngdc.noaa.gov/mgg/geodas/trackline.html>
- [14] http://www.ips.gov.au/World_Data_Centre
- [15] <http://gp.wdc.cn>
- [16] <http://www.intermagnet.org>
- [17] <http://www.icsu-fags.org>
- [18] <http://wds.geolinks.org>
- [19] http://www.earthobservations.org/documents/geo_brochure.pdf
- [20] <http://isgi.latmos.ipsl.fr>
- [21] http://www-app3.gfz-potsdam.de/kp_index/index.html
- [22] http://www.obsebre.es/php/geomagnetisme/variacion_ap.php
- [23] <http://spidr.ngdc.noaa.gov/spidr>
- [24] <http://supermag.jhuapl.edu>
- [25] <http://www.geo.fmi.fi/image>
- [26] <http://bluebird.phys.ualberta.ca/carisma>
- [27] http://www.intermagnet.org/Yearbooks_e.php
- [28] <http://www.ngdc.gov/metadata>
- [29] <http://www.opengeospatial.org>
- [30] <http://www.spase-group.org>
- [31] <http://www.w3schools.com/xsl/default.asp>
- [32] http://www.geomind.eu/portal/md_search.jsf
- [33] <http://www.iugonet.org/en>
- [34] <http://geonetwork-opensource.org>
- [35] http://www.geoportal.org/web/guest/geo_home
- [36] <http://www.wdc.bgs.ac.uk/catalog/format.html>
- [37] http://www.intermagnet.org/FormatData_e.php
- [38] <http://www.ngdc.noaa.gov/IAGA/vdat/iagaformat.html>

References

- Abston CC, Papitashvili NE and Papitashvili VO (1985) Combined International Catalog of Geomagnetic Data, Report UAG-92, U.S. National Geophysical Data Center
- Akasofu S-I (1981) Relationships between the AE and Dst indices during geomagnetic storms. *J. Geophys. Res.*, 86:4820–4822
- Baillie O, Clarke E, Flower S, Reay S, and Turbitt C (2009) Reporting quasi-definitive observatory data in near real time. 11th IAGA Scientific Assembly, Abstract Book 94 pp
- Bartels J, Heck NH, and Johnston HF (1939) The three-hour-range index measuring geomagnetic activity. *J. Geophys. Res.* 44:411-454
- Boteler DH (2006) Comment on time conventions in the recordings of 1859. *Adv. Space Res.* 38:301–303
- Brooke C (1847) Description of apparatus for automatic registration of magnetometers and other meteorological instruments by photography. *Philosophical Transactions of the Royal Society of London* 137: 69–77
- Buneman P (2005) What the Web Has Done for Scientific Data, and What it Hasn't. In *Advances in Web-Age Information Management: 6th International Conference, Web Age Information Management 2005*, Hangzhou, China, October 2005 Proceedings. 1-7 pp
- Chulliat A (2009c) Evaluation of IGRF-11 candidate models, secular variation. From IGRF-11: Progress reports, candidate models, evaluations and test models http://www.ngdc.noaa.gov/IAGA/vmod/EVALS/Eval_Chulliat.pdf
- Chulliat A, Peltier A, Truong F, and Fouassier D (2009b) Proposal for a new observatory data product: quasi-definitive data. 11th IAGA Scientific Assembly, Abstract Book 94 pp
- Chulliat A, Savary J, Telali A, Lalanne X (2009a) Acquisition of 1-Second Data in IPGP Magnetic Observatories. *Proceedings of the XIIIth IAGA Workshop on Geomagnetic Observatory Instruments, Data Acquisition, and Processing: U.S. Geological Survey Open-File Report 2009-1226*, 54-59 pp
- Chulliat A, Thébault E and Hulot G (2010) Core field acceleration pulse as a common cause of the 2003 and 2007 geomagnetic jerks. *Geophys. Res. Lett.* 37:L07301, doi:10.1029/2009GL042019
- Clarke E, Flower S, Humphries T, McIntosh R, McTaggart F, McIntyre B, Owenson N, Henderson K, Mann E, MacKenzie K, Piper S, Wilson L, and Gillanders R (2009) The digitization of observatory magnetograms. 11th IAGA Scientific Assembly, Program Abstract 43 pp
- Curto JJ, Sanclement E and Torta JM (1996) Automatic measurement of magnetic records on photographic paper. *Computers & Geosciences* 22:359-368
- Davis TN, and Sugiura M (1966) Auroral electrojet activity index AE and its universal time variations. *J Geophys. Res.* 71:785
- Dawson E, Reay S, Macmillan S, Flower S, Shanahan T (2009) Quality Control Procedures At The World Data For Geomagnetism (Edinburgh). 11th IAGA Scientific Assembly, Program Abstract 44 pp
- Feldstein YI, Grafe A, Gromova LI and Popov VA (1997) Auroral electrojets during geomagnetic storms. *J. Geophys. Res.* 102:14223–14236, doi:10.1029/97JA00577
- Fischman D, Denig WF and Herzog D (2009) A Proposed Metadata Implementation for Magnetic Observatories. *Proceedings of the XIIIth IAGA Workshop on Geomagnetic Observatory Instruments, Data Acquisition, and Processing: U.S. Geological Survey Open-File Report 2009-1226*, 82-85 pp
- Finlay CC, Maus S, Beggan C, Hamoudi M, Lesur V, Lowes FJ, Olsen N and Thébault E (2010) Evaluation of candidate geomagnetic field models for IGRF-11. *Earth Planets Space* 62 (10). 787-804
- Fouassier D and Chulliat A (2009) Extending backwards to 1883 the French magnetic hourly data series. *Proceedings of the XIIIth IAGA Workshop on Geomagnetic Observatory Instruments, Data Acquisition, and Processing: U.S. Geological Survey Open-File Report 2009-1226*, 86-94 pp
- Friis-Christensen E, Lühr H and Hulot G (2006) Swarm: A constellation to study the Earth's magnetic field. *Earth Planets and Space*, 58:351-358

- Hejda P, Herzog D, Linthe HJ, Mandea M, Schott J-J, and Svalgaard L (2009), On the derivation of hourly mean values with incomplete data. 11th IAGA Scientific Assembly, Abstract Book 94 pp
- Herzog DC and Lupica CW (1992) Magnetic observatory data on CD-ROM. *Eos, Transactions, American Geophysical Union* 73:236
- Herzog DC (2009) The effects of missing data on mean hourly values. Proceedings of the XIIIth IAGA Workshop on Geomagnetic Observatory Instruments, Data Acquisition, and Processing: U.S. Geological Survey Open-File Report 2009-1226, 116-126 pp
- IAGA (2009) IAGA News 46, http://www.iugg.org/IAGA/iaga_pages/pubs_prods/Newsletters/IAGA_News_46.pdf
- ICSU Panel on World Data Centers (1996) Guide to the World Data Center System. <http://www.ngdc.noaa.gov/wdc/guide/wdguide.html>
- ICSU (2008) Ad hoc Strategic Committee on Information and Data. Final Report to the ICSU Committee on Scientific Planning and Review
- Iyemori T (1990) Storm-time magnetospheric currents inferred from mid-latitude geomagnetic field variations. *J Geomag Geoelectr* 42:1249-1265
- Iyemori T, Nose M, McCreadie H, Odagi Y, Takeda M, Kamei T and Yagi M (2005) Digitization of Old Analogue Geomagnetic Data. Proceedings of PV-2005: Ensuring Long-term Preservation and Adding Value to Scientific and Technical data, Royal Society, Edinburgh, UK. <http://www.ukoln.ac.uk/events/pv-2005/pv-2005-final-poster-papers/014-poster.doc>
- Jankowski J and Sucksdorff C (1996) Guide for Magnetic Measurements and Observatory Practice. International Association of Geomagnetism and Aeronomy, Warsaw, Poland
- Kerridge DJ (2001) INTERMAGNET: Worldwide near-real-time geomagnetic observatory data, Proc. ESA Space Weather Workshop, ESTEC, Noordwijk, The Netherlands, http://www.esa-spaceweather.net/spweather/workshops/SPW_W3/PROCEEDINGS_W3/ESTEC_Intermagnet.pdf
- King T, Thieman J and Roberts DA (2010) SPASE 2.0: a standard data model for space physics. *Earth Sci Infor*, doi:10.1007/s12145-010-0053-4
- Korte M, Mandea M and Olsen N (2007) Worldwide Observatory Hourly Means from 1995 to 2003: Investigation of Their Quality. In Reda J (ed) XII IAGA Workshop on Geomagnetic Observatory Instruments, Data Acquisition and Processing Belsk, 19-24 June 2006, Publications of Institute of Geophysics Polish Academy of Sciences, C-99 (398)
- Korepanov V, Marusenkov A, and Rasson J (2009) 1-second INTERMAGNET Standard Magnetometer – 1-year operation. 11th IAGA Scientific Assembly, Abstract Book 94 pp
- Love JJ (2009) Missing data and the accuracy of magnetic-observatory hour means. *Ann. Geophys.* 27:3601-3610
- Macmillan, S (2007) Observatories: an overview. In: Gubbins, D and Herrero-Bervera E (eds.) Encyclopedia of Geomagnetism and Paleomagnetism. Netherlands, Springer, 708-711 pp.
- Mabie J, Denig W, McLean S (2009) Geomagnetic Data Management at the NGDC. 11th IAGA Scientific Assembly, Program Abstract 114 pp
- Marsal S and Curto JJ (2009) A new approach to the hourly mean computation problem when dealing with missing data. *Earth Planets Space* 61:945-956
- Mandea M (2002) 60, 59, 58 ... How many minutes for a reliable hourly mean? Proceedings of the Xth IAGA Workshop on Geomagnetic Instruments, Data Acquisition, and Processing, Hermanus Magnetic Observatory, Hermanus, South Africa, 112–120 pp
- Mandea M and Papitashvili V (2009) Worldwide Geomagnetic Data Collection and Management, *Eos, Transactions, American Geophysical Union*, 90:409-410
- Matsushita S, Campbell WH (eds.) (1967) Physics of geomagnetic phenomena (2 vols.). New York-London: Academic Press
- Mayaud PN (1967) Calcul préliminaire d'indices Km, Kn et Ks ou Am, An, et As, mesures de l'activité magnétique à l'échelle mondiale et dans les hémisphères Nord et Sud. *Ann. Géophys.* 23:585-617
- Mayaud PN (1972) The aa index: a 100-year series characterizing the geomagnetic activity, *J. Geophys. Res.* 77:6870-6874
- Mayaud PN (1980) Derivation, Meaning, and Use of Geomagnetic Indices, AGU Geophys Monogr 22
- Menvielle M, Iyemori T, Marchaudon A and Nose M (2011) Geomagnetic indices. In: Mandea M, Korte M (eds) Geomagnetic Observations and Models, IAGA Special Sopron Book Series 5, DOI 10.1007/978-90-481-9858-0_7 pp 183-228
- Menvielle M, Papitashvili N, Hakkinen L, and Sucksdorff C (1995) Computer production of K indices: review and comparison of methods. *Geophys J Int* 123:866 – 886
- Merka J, Narock TW and Szabo A (2008) Navigating through SPASE to heliospheric and magnetospheric data. *Earth Sci. Inform.* 1:35-42
- Minamoto Y (2009) Ongoing Geomagnetic field 1-second value measurement by JMA. Proceedings of the XIIIth IAGA Workshop on Geomagnetic Observatory Instruments, Data Acquisition, and Processing: U.S. Geological Survey Open-File Report 2009-1226, 190-193 pp
- Minster JH, Capitaine N, Clark DM and Mokrane M (2009) Initial Progress in Developing the New ICSU World Data System. American Geophysical Union, Fall Meeting 2009, abstract #IN23A-1059
- Narod B (2009) Private communication. Further information can be obtained from Narod Geophysics Ltd., #310, 2475 York Ave., Vancouver, B.C., Canada, V6K-1C9. Tel/fax: 604-732-9083
- Nevanlinna H and Häkkinen L (2010) Results of Russian geomagnetic observatories in the 19th century: magnetic activity, 1841–1862, *Ann. Geophys.*, 28:917-926
- Nagarajan N (2008) Retrieval and Distribution of Digital Archives of Photographic Records of Magnetic Variations at Ngri, India. Abstract for XIIIth IAGA Workshop on Geomagnetic Observatory Instruments, Data Acquisition, and Processing http://mines.conference-services.net/resources/328/1322/pdf/IAGA2008_0017.pdf
- Newitt LR (2009) The Effects of Missing Data on the Computation of Hourly Mean Values and Ranges. Proceedings of the XIIIth IAGA Workshop on Geomagnetic Observatory Instruments, Data Acquisition, and Processing: U.S. Geological Survey Open-File Report 2009-1226, 194-201 pp
- Peltier A and Chulliat A (2010) On the feasibility of promptly producing quasi-definitive magnetic observatory data. *Earth Planets Space*, 62(2):e5-e8, doi:10.5047/eps.2010.02.002

- Peng F, Shen X, Tang K, Zhang J, Huang Q, Xu Y, Yue B, and Yang D (2007) Data-Sharing Work of the World Data Center for Geophysics, Beijing. *Data Science Journal*, 6:404-407
- Rasson JL (2007) Observatories, INTERMAGNET. In: Gubbins D and Herrero-Bervera E (eds) *Encyclopedia of Geomagnetism and Paleomagnetism*, 715–717 pp, Springer-Verlag, New York
- Reay SJ, Allen W, Baillie O, Bowe J, Clarke E, Lesur V and Macmillan S (2005) Space weather effects on drilling accuracy in the North Sea. *Ann Geophys* 23:3081–3088
- Reay S, Dawson E, Flower S, Macmillan S, and Herzog D (2009) Towards a metadata standard for geomagnetic observatory data. 11th IAGA Scientific Assembly, Abstract Book 94 pp
- Reda J, Fouassier D, Isac A, Linthe H-J, Matzka J, Turbitt CW (2011) Improvements in geomagnetic observatory data quality. In: Mandea M, Korte M (eds) *Geomagnetic Observations and Models*, IAGA Special Sopron Book Series 5, DOI 10.1007/978-90-481-9858-0_7 pp 127–148
- St-Louis BJ, Sauter EA, Trigg DF and Coles RL (2008) INTERMAGNET Technical Reference Manual, version 4.4
http://www.intermagnet.org/publications/im_manual.pdf
- Schmidt, A (1905) Ergebnisse der magnetischen Beobachtungen in Potsdam, Veröffentl. des Preuss. Meteorol. Inst., Berlin
- Schott J-J and Linthe HJ (2007) The hourly mean computation problem revisited. Proceedings of the XIIth IAGA Workshop on Geomagnetic Instruments, Data Acquisition, and Processing, Institute of Geophysics, Polish Acad Sci, Warszawa, 135–143 pp
- Shanahan TJG and Turbitt CW (2009) Evaluating the Noise for a Commonly Used Fluxgate Magnetometer – for 1-second Data. Proceedings of the XIIIth IAGA Workshop on Geomagnetic Observatory Instruments, Data Acquisition, and Processing: U.S. Geological Survey Open-File Report 2009-1226, 239-245 pp
- Stauning P, Troshichev O, Janzhura A (2009) The polar cap (PC) index. Present (unified) and past index calculations 11th IAGA Scientific Assembly, Abstract Book 99 pp
- Sugiura M (1964) Hourly values of equatorial Dst for IGY. *Annals of the International Geophysical Year*, vol. 35, Pergamon Press, Oxford 945-948 pp
- Svalgaard L and Cliver EW (2007) Interhourly variability index of geomagnetic activity and its use in deriving the long-term variation of solar wind speed, *J. Geophys. Res.*, 112:A10111 doi:10.1029/2007JA012437
- Trigg DF, Serson PH, and Camfield PA (1971) A solid-state electrical recording magnetometer. Department of Energy, Mines, and Resources 41 Ottawa, Canada.
- Troshichev OA and Andrezen VG (1985) The relationship between interplanetary quantities and magnetic activity in the southern polar cap. *Planet Space Sci*, 33:415–419
- Tsurutani BT, Gonzalez WD, Lakhina GS and Alex S (2003) The extreme magnetic storm of 1–2 September 1859. *J. Geophys. Res.* 108(A7):1268 doi:10.1029/2002JA009504.
- Turner GM, Rasson JL, Reeves CV (2007) Observation and measurement techniques, in *Treatise on Geophysics*, vol. 5, ed. by G. Schubert, Elsevier, Amsterdam 93-146 pp
- Wayne L (2005a) Metadata in Action. In: *Proceedings of the 2005 GISPlanet Conference Proceedings* <http://www.fgdc.gov/metadata/documents/MetadataInAction.doc> (Accessed 28th July 2009)
- Wayne L (2005b) Institutionalize Metadata Before It Institutionalizes You, *Federal Geographic Data Committee* http://www.fgdc.gov/metadata/documents/InstitutionalizeMeta_Nov2005.doc (Accessed 28th July 2009)
- Wienert KA (1970) Notes of geomagnetic observatory and survey practice. UNESCO Earth Sciences 5
- Worthington EW, Sauter EA and Love JJ (2009b) Analysis of USGS One-Second Data. *Proceedings of the XIIIth IAGA Workshop on Geomagnetic Observatory Instruments, Data Acquisition, and Processing*: U.S. Geological Survey Open-File Report 2009-1226, 262–266 pp
- Worthington EW, White T, Sauter EA, and Stewart DC (2009a) Evaluation of 1-Hz data from USGS geomagnetic observatories. 11th IAGA Scientific Assembly, Abstract Book 94 pp