

OPERATIONS OF THE WORLD DATA CENTRE FOR GEOMAGNETISM, EDINBURGH

S J Reay, E Clarke, E Dawson, and S Macmillan*

**British Geological Survey, Murchison House, West Mains Road, Edinburgh, EH9 3LA, United Kingdom
Email: sjr@bgs.ac.uk*

ABSTRACT

The British Geological Survey has operated a World Data Centre for Geomagnetism since 1966. Geomagnetic time-series data from around 280 observatories worldwide at a number of time resolutions are held along with various magnetic survey, model, and activity index data. The operation of this data centre provides a valuable resource for the geomagnetic research community.

The operation of the WDC and details of the range of data held are presented. The quality control procedures that are applied to incoming data are described as is the work to collaborate with other data centres to distribute and improve the overall consistency of data held worldwide. The development of standards for metadata associated with datasets is demonstrated, and current efforts to digitally preserve the BGS analogue holdings of magnetograms and observatory yearbooks are described.

Keywords: Geomagnetism, World Data Centre, World Data System, Metadata, Digital data capture

1 INTRODUCTION

A World Data Centre (WDC) for Geomagnetism was established in the United Kingdom in 1966. This was operated by the Institute of Geological Science, which later became the British Geological Survey (BGS), in Herstmonceux, Sussex. The WDC moved to its current location in Edinburgh in 1977. BGS is part of the Natural Environment Research Council (NERC), a research centre funded by UK government. In the past this WDC focused its attention primarily on gathering data for use in global magnetic field modelling, particularly geomagnetic observatory annual means. The WDC for Geomagnetism, Copenhagen, then hosted by the Danish Meteorological Institute, gathered geomagnetic observatory one-minute and hourly mean values. In 2007 BGS agreed to take over responsibility for the digital datasets held at WDC Copenhagen. These were transferred to Edinburgh and a copy made of the data catalogue web pages to ensure the data were available in an identical form from the WDC Edinburgh website (<http://www.wdc.bgs.ac.uk/>).

The WDC operations in Edinburgh are carried out by the Geomagnetism Team of BGS. Several staff members are involved in WDC operations to various degrees although all have additional duties within the team. BGS staff have experience in processing and delivering geomagnetic data since the team operate eight geomagnetic observatories worldwide and conduct repeat station observations across the UK. They are also experienced in using geomagnetic data as the team's research scientists work in the field of global geomagnetic field modelling and space weather science.

In February 2012 the WDC for Geomagnetism, Edinburgh was formally accepted as a regular member of the newly established ICSU World Data System (WDS) (<http://www.icsu-wds.org/>).

2 DATA HOLDINGS

WDC for Geomagnetism, Edinburgh is an active and growing data centre keenly seeking out new data. An individually-tailored email is sent annually to more than 100 geomagnetic observatory operators requesting any new data and highlighting current gaps in the data holdings for their observatories. Reciprocal arrangements with other WDCs to share data directly submitted to us is in place and an agreement with INTERMAGNET (<http://www.intermagnet.org/>) has been established to collect data from this network and distribute via the WDC. BGS hope to increase links with other WDCs in the future and widen this data distribution network.

The WDC holds an extensive range of geomagnetic data including digital data from around 280 geomagnetic observatories worldwide containing one-minute mean values from 1969; hourly values from 1883; and annual

means from 1813 (Figure 1). Data from land, marine, satellite, and aeromagnetic surveys and repeat stations worldwide from 1900 onwards are available as well as charts and computations from main field models, including the World Magnetic Model (WMM) (<http://www.geomag.bgs.ac.uk/research/modelling/WorldMagneticModel.html>) and International Geomagnetic Reference Field (IGRF) (<http://www.ngdc.noaa.gov/AGA/vmod/igrf.html>) model. Further digital data include a complete set of the definitive magnetic activity indices (K , K_p , ap , A_p , aa , Aa , C_p and C_9) and solar activity indices (*International Sunspot Number* and $F10.7$). An archive of historical and analogue records is also maintained. These include geomagnetic data collected during NERC funded university projects; magnetograms from all UK observatories from 1850 (the digital capture of which is described in Section 5); and a library containing observatory yearbooks from around the world, expedition memoirs, original survey observations, and other miscellaneous items.

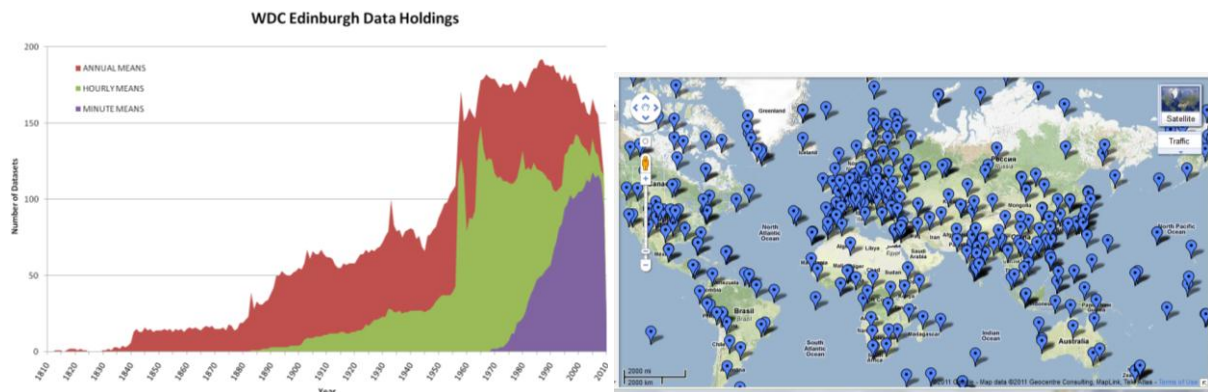


Figure 1. Data holdings of the WDC: Left – the number of observatory data holdings by time resolution and year. Right – the locations of current & past observatories for which we hold annual, hourly, or minute data. (Google map is © Google 2011, map data © Europa Technologies, MapIT, Tele Atlas)

3 QUALITY CONTROL WITHIN OUR WDC

All data submitted to the WDC are subjected to quality assurance checks before ingestion into the database. When problems are encountered, the BGS staff work with the data originators to try to resolve the issues and improve the dataset. It is known that errors do exist within the historical data holdings; in some cases data quality could be improved or there are inconsistencies in the data held at the various WDCs for Geomagnetism. This is of great concern to the scientific community and is the main focus of BGS's current WDC activities.

The problem of disagreement between datasets held at different WDCs is not a straightforward one. Geomagnetic observatory data are commonly published annually for the preceding calendar year. Finalised data are classed as 'definitive'. However, data may then have corrections applied to produce a second, improved 'definitive' dataset. In the past, subsequent corrections may have been made by the data originator and not passed on to all data centres. Also, some data centres may have made their own corrections to the data and not documented or distributed the changes. The corrections may take the form of a change to the absolute ('baseline') level of the data, the removal of obviously erroneous data points (such as cases where there is evidence of significant environmental noise), or the correction of simple typographical and formatting errors. In all cases, because of the lack of version information within the geomagnetism community's file format standards, the history of these changes may be lost.

Since 2007 BGS has worked to improve the quality of the data to aid the scientific community. Simple typographical or formatting errors within the hourly observatory data holdings were sought and corrected (Dawson et al., 2009). Errors of this type, such as the use of an incorrect "missing sample flag" value, can result in gross effects in any subsequent data analysis. Furthermore, holdings of hourly and minute mean values have been compared with those held in other WDCs for Geomagnetism in order to identify gaps in the respective holdings.

BGS are working with the WDC for Geomagnetism, Kyoto in particular to harmonise the data holdings. In theory, the datasets common to different WDCs should be identical (given that they are holding 'definitive')

data). However, in a recent analysis by Dawson et al. (2011), it was discovered that almost 20% of the datasets have some level of disagreement between them. The WDCs at Edinburgh and Kyoto have agreed to work together to resolve these issues. This is not a simple undertaking; it is not always clear why some data are in disagreement, and without records of the corrections made, the task of identifying the authoritative ‘definitive’ dataset is made more difficult. This work will need to be carried out in partnership with the observatory operators who will have the final say on which version should be considered definitive. However in some cases, where observatories are no longer operational or the host institute no longer exists, absolute certainty may not be possible.

4 METADATA

The issues caused by the lack of documentation, with regard to quality control procedures carried out on data held at the WDCs, demonstrate the importance of metadata. There have been discussions between the WDCs and among the wider geomagnetic community to address this matter and establish a metadata standard for geomagnetic observatory data (e.g., Fischman et al., 2009 or Reay et al., 2011). Current metadata standards for geospatial data may act as a guide in establishing a standard, but, as geomagnetic observatory data are time-series and every aspect of metadata associated with an observatory (up to and including its location) could change with time, it is difficult to determine what to assign a metadata record to.

Within the WDC for Geomagnetism, Edinburgh the decision was made to focus on metadata gathering, collecting useful information that can be re-formatted according to an agreed metadata standard at a later date. As part of the annual ‘call-for-data’, basic metadata, such as observatory name, location, dates of operation, contact information, instruments used, and other similar details, are requested. Historically, observatory yearbooks were produced in a reasonably similar format around the world and can be considered the unofficial metadata standard for observatory data. The continued production of yearbooks is actively encouraged, and electronic copies of these are made available for download via the WDC website wherever possible. These annual reports thoroughly describe the operations carried out at an observatory for a given year and usually contain all the information, such as processing history, instrumentation, etc. that would be required for a complete metadata record.

The metadata records collated to date are available online for users to examine. These records are being populated with information on known QC issues associated with the data, including notes of the version history of the data if it has been modified. To help populate these records, feedback from both the providers of the original datasets and the WDC users are welcome.

5 PRESERVING HISTORICAL DATA

In the middle of the 19th century regular systematic measurements of the Earth’s magnetic field began in earnest. Analogue data, in various forms, were produced until the late 20th century when the majority of observatories introduced digital data recording. BGS are custodians of the original records and results from historical UK observatories. These consist of magnetogram records on photographic paper between 1848 - 1987 and a range of yearbooks produced to summarise the results and describe the operations. The operational timeline of each observatory is shown in Figure 2. These records provide a unique and continuous record of magnetic variations across the UK. Records for historic British colonial observatories are also held and copies of various yearbooks from observatories worldwide are stored as part of WDC operations.

In the past it was realised that some of these records were in a poor physical state caused by storage in unsuitable conditions. There was also a risk of further deterioration. In 2008 a programme commenced to create digital back-up copies of each paper magnetogram and make these available online (Clarke et al., 2009). The records were also moved to a secure, environmentally controlled archive room conforming to BS5454 (the British Standard for the preservation of archival material). The digital capture was carried out using a 21-megapixel camera at a fixed focal length. The images are stored in a database and an online magnetogram archive is available from <http://www.bgs.ac.uk/data/Magnetograms/home.html>. This work is ongoing; as of December 2011, approximately 90% of the 472-observatory years of data have been digitally captured.

A collection of 378 yearbooks from current and historical UK and British colonial observatories are held plus many articles and survey data records. Work to digitise the historical observatory yearbooks, in conjunction with the digitisation of the magnetograms from the same observatories, was started early in 2011. Some articles

related to observatory results and operations were also included. The yearbooks contain the published results for each observatory and information about the observatory operations and measurements - a critical source of metadata. A 'Bookeye' scanner was used to capture good quality images of each page. This work is now complete and the electronic copies are available from http://www.geomag.bgs.ac.uk/data_service/data/yearbooks/yearbooks.html.

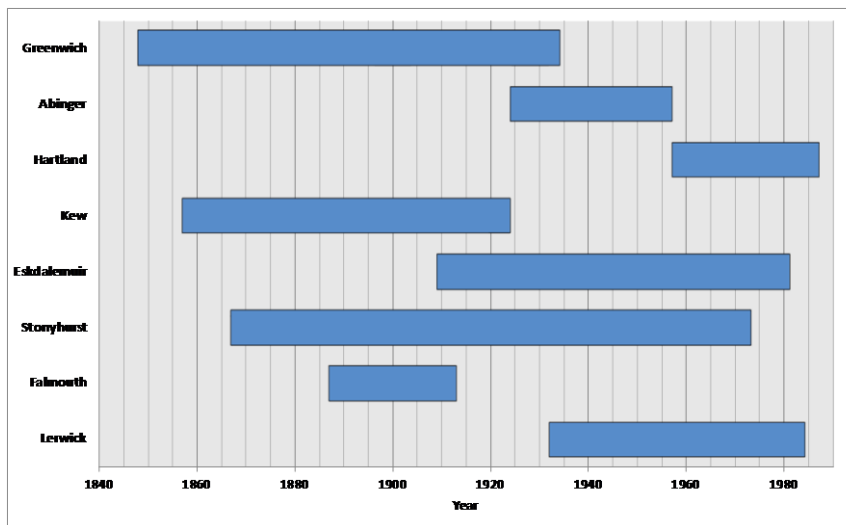


Figure 2. The temporal coverage of the analogue UK magnetogram records for eight UK observatories.

6 FUTURE DEVELOPMENTS

This paper has discussed the status of the World Data Centre for Geomagnetism, Edinburgh and the work currently undertaken there. New data will continue to be ingested, and in the future the collection of additional data types, such as one-second magnetic observatory data, will be considered. The aim to make progress on quality control issues in collaboration with the WDC for Geomagnetism, Kyoto is high on the agenda. Datasets will also be better documented by increasing the quality and range of metadata held. The programme to digitise BGS holdings of analogue UK magnetograms will continue until all records are captured and available online. Furthermore, a catalogue of all other analogue records held will be completed, with unique and at-risk documents identified for digitisation. Finally, the current WDC online interface will be replaced to improve how users can access and analyse data in the future. Web services technologies, such as that described in Dawson et al (2012), will be used to better deliver data to the scientific community.

7 ACKNOWLEDGEMENTS

This paper is published with the permission of the Director of the British Geological Survey (Natural Environment Research Council).

8 REFERENCES

Clarke, E., Flower, S., Humphries, T., McIntosh, R., McTaggart, F., McIntyre, B., Owenson, N., Henderson, K., Mann, E., MacKenzie, K., Piper, S., Wilson, L., & Gillanders, R. (2009) The digitization of observatory magnetograms. Poster presented at: 11th IAGA Scientific Assembly, 23-30 Aug 2009, Sopron, Hungary. Retrieved from the WWW, November 30, 2012: <http://nora.nerc.ac.uk/12494/>

Dawson, E., Lowndes, J., & Reddy, P. (2012) The British Geological Survey's new geomagnetic data web service. Manuscript submitted for publication to Data Science Journal (*Proceedings of 1st ICSU World Data System Conference*).

Dawson, E., Reay, S., Macmillan, S., Flower, S., & Shanahan, T. (2009) Quality Control Procedures at the

World Data for Geomagnetism (Edinburgh). Poster presented at: 11th IAGA Scientific Assembly, 23-30 Aug 2009, Sopron, Hungary. Retrieved from the WWW, November 30, 2012: <http://nora.nerc.ac.uk/11740/>

Dawson, E., Macmillan, S., Humphries, T., & Beggan, C. (2011) A comparison of data holdings at World Data Centres for Geomagnetism in Edinburgh and Kyoto. Poster presented at: IUGG XXV General Assembly, 28 June - 7 July 2011, Melbourne, Australia. Retrieved from the WWW, November 30, 2012: <http://nora.nerc.ac.uk/15365/>

Fischman, D., Denig, W.F., & 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, pp. 82-85.

Reay, S.J., Herzog, D.C., Alex, S., Kharin, E., McLean, S., Nosé, M., & Sergeyeva, N. (2011) Magnetic Observatory Data and Metadata: Types and Availability. In: Manda, M. & Korte, M. (eds) *Geomagnetic Observations and Models*, IAGA Special Sopron Book Series 5, DOI 10.1007/978-90-481-9858-0_7 pp. 149-181.

(Article history: Available online 31 January 2013)