

Resolving issues concerning Eskdalemuir geomagnetic hourly values

S. Macmillan and E. Clarke

British Geological Survey, Murchison House, West Mains Road, Edinburgh, EH9 3LA, UK

Received: 24 January 2011 – Revised: 31 January 2011 – Accepted: 1 February 2011 – Published: 8 February 2011

Abstract. The hourly values of the geomagnetic field from 1911 to 1931 derived from measurements made at Eskdalemuir observatory in the UK, and available online from the World Data Centre for Geomagnetism at <http://www.wdc.bgs.ac.uk/>, have now been corrected. Previously they were 2-point averaged and transformed from the original north, east and vertical down values in the tables in the observatory yearbooks. This paper documents the course of events from discovering the post-processing done to the data to the final resolution of the problem. As it was through the development of a new index, the Inter-Hour Variability index, that this post-processing came to light, we provide a revised series of this index for Eskdalemuir and compare it with that from another European observatory. Conclusions of studies concerning long-term magnetic field variability and inferred solar variability, whilst not necessarily consistent with one another, are not obviously invalidated by the incorrect hourly values from Eskdalemuir. This series of events illustrates the challenges that lie ahead in removing any remaining errors and inconsistencies in the data holdings of different World Data Centres.

Keywords. Geomagnetism and paleomagnetism (Time variations, diurnal to secular; Time variations, secular and long term; General or miscellaneous)

1 Introduction

For the past few years it has been known that the Eskdalemuir hourly values from 1911 to 1931 were post-processed with unfortunate consequences. These data were formerly available online from the World Data Centre (WDC) for Geomagnetism at the Danish Meteorological Institute till the start of 2007 and then at the authors' institute. The post-processing was the application of a 2-point running mean and

transformation to declination, horizontal and vertical intensities (DHZ) from the original northerly, easterly and vertical intensities (XYZ) recorded in the tables of the observatory yearbooks. The reason for the post-processing was presumably to make these earlier hourly values compatible with the later values which were hourly mean DHZ values centred on the half hour (Eskdalemuir yearbooks, 1932 onwards). There is no record of when this post-processing was done. The 1911–1917 data were XYZ spot values on the hour (according to yearbook table headings but more on this later) and the 1918–1931 data were XYZ mean values centred on the hour (Eskdalemuir yearbooks, 1911–1931). This post-processing had the effect of lowering the spectral content for high frequency studies performed on the hourly value time series. However the significance of this only really came to light with the development of the Inter-Hour Variability index in the 2000s.

2 The discovery

The first indication of something unexpected in the early Eskdalemuir hourly values was from correspondence with Leif Svalgaard in 2004 where he presented ratios of the newly defined Inter-Hour Variability indices for the UK observatories. The IHV index is the sum of absolute differences of hourly means (or values) for a geomagnetic component from one hour to the next over the six hour interval around local midnight. The ratios for the annual means of the IHV indices for the horizontal component as recorded at the UK observatories at Eskdalemuir, Lerwick, Abinger and Hartland revealed a sharp change in the Eskdalemuir record at the end of 1931 (Fig. 1).

The creation of the IHV index was motivated by issues concerning the *aa* index (Svalgaard et al., 2003). The *aa* index has been critical for inferring long-term changes in the Sun (e.g. Lockwood et al., 1999) as it is the longest running index of magnetic activity available, dating back to 1868. However some records of observatory hourly values are also very long and this lead to the development and



Correspondence to: S. Macmillan
(smac@bgs.ac.uk)

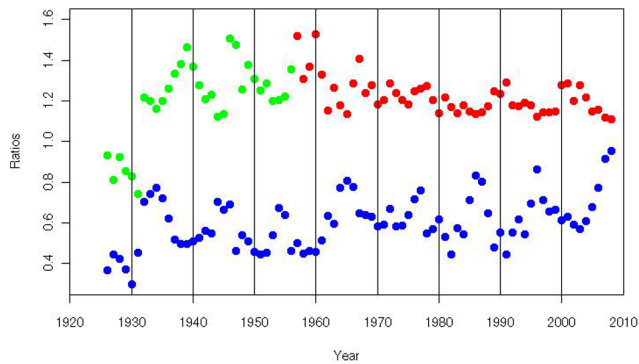


Fig. 1. Ratios of annual means of UK observatory IHV indices based on data available at the Edinburgh WDC before 9 August 2010. Green is ratio of Eskdalemuir to Abinger, red is ratio of Eskdalemuir to Hartland and blue is ratio of Eskdalemuir to Lerwick.

analyses of IHV indices for inferring long-term changes in the Sun and near-Earth space (Svalgaard et al., 2004; Mursula et al., 2004; Mursula and Martini, 2006; Svalgaard and Cliver, 2007a, b). The problem with the Eskdalemuir IHV series was known at this stage and attempts were made to correct it (Martini and Mursula, 2006). Another point to note at this stage was that the World Centres were acknowledged collectively for the supply of the hourly values and the individual WDC supplying the data was not identified. This turns out to be important. In the meantime the long-term solution to the problem seemed to lie with either undoing the post-processing or redigitising the original hourly value tables in the observatory yearbooks.

3 Undoing the post-processing

In theory, it should be possible to undo such relatively simple post-processing by obtaining an initial value from the yearbooks and applying it to each series of post-processed data using

$$E_{i+1} = 2E'_i - E_i$$

where E_i is an original hourly value of an element of the Earth's magnetic field for hour i and E'_i is a 2-point averaged hourly value of an element of the Earth's magnetic field for hour i . Knowledge of E_0 (the value at the start of the time series) should be all that is required to regenerate the original series of data. However it turns out that the transforming from XYZ to DHZ and the limitations of the WDC format significantly complicates matters. We explain this now in more detail.

The digital storage and exchange format that is widely used for hourly values is the WDC format, described at <http://www.wdc.bgs.ac.uk/catalog/master.html>. This is the natural extension of the way that hourly values were presented in the older yearbooks. The use of a baseline value,

one line of data per element per day and 1 nT resolution is common to both the yearbook tables and the WDC-format files. The unit for declination of the magnetic field in the WDC-format file is tenth arc-minute, and the resolution is 1 tenth arc-minute.

There are 4 options of how the post-processing was done:

1. The XYZ values were transformed to DHZ, then 2-point averaged and rounded to 1 nT.
2. The XYZ values were transformed to DHZ, rounded to nearest 1 tenth arc-minute and 1 nT (the resolution of the WDC format), then 2-point averaged and rounded to 1 nT. This option is possible if WDC-format files of unaveraged DHZ existed at an intermediate stage.
3. The XYZ values were 2-point averaged, then transformed to DHZ and rounded to nearest 1 tenth arc-minute and 1 nT.
4. The XYZ values were 2-point averaged, rounded to nearest 1 nT, then transformed to DHZ and rounded (again) to nearest 1 tenth arc-minute and 1 nT. This option is possible if WDC-format files of averaged XYZ existed at an intermediate stage.

Within each of these options are choices for rounding 0.5 to the resolution of the WDC-format files which is 1. Whilst the default rounding in the IEEE 754-2008 floating point arithmetic standard is to round to even, its implementation is dependent on computer platform and programming language. Rounding to zero and rounding up are other possibilities. Also unknown is whether the same rounding was applied to each component. A further complication is that the Y component in the yearbooks assumes west is positive, but it is not known whether this was the case in any intermediate WDC-format files.

Because of these uncertainties in the post-processing, and also because any resulting values from attempts to undo the post-processing would require comparison with the original data sources, the yearbook tables were scanned and digitised. It should be noted that the original data are the analogue magnetograms and tables of absolute observations and whilst there is a programme underway in BGS to digitise these for long-term preservation, it was not deemed practical to use these to obtain new series of absolute hourly values within a reasonable timescale.

4 Experiences with scanning and digitising yearbook tables

In support of a study by Martini and Mursula (2006), work began to redigitise the original hourly values directly from the monthly tables in the Eskdalemuir observatory yearbooks. Originally this involved making scanned copies of only the January tables for each year as described in Martini

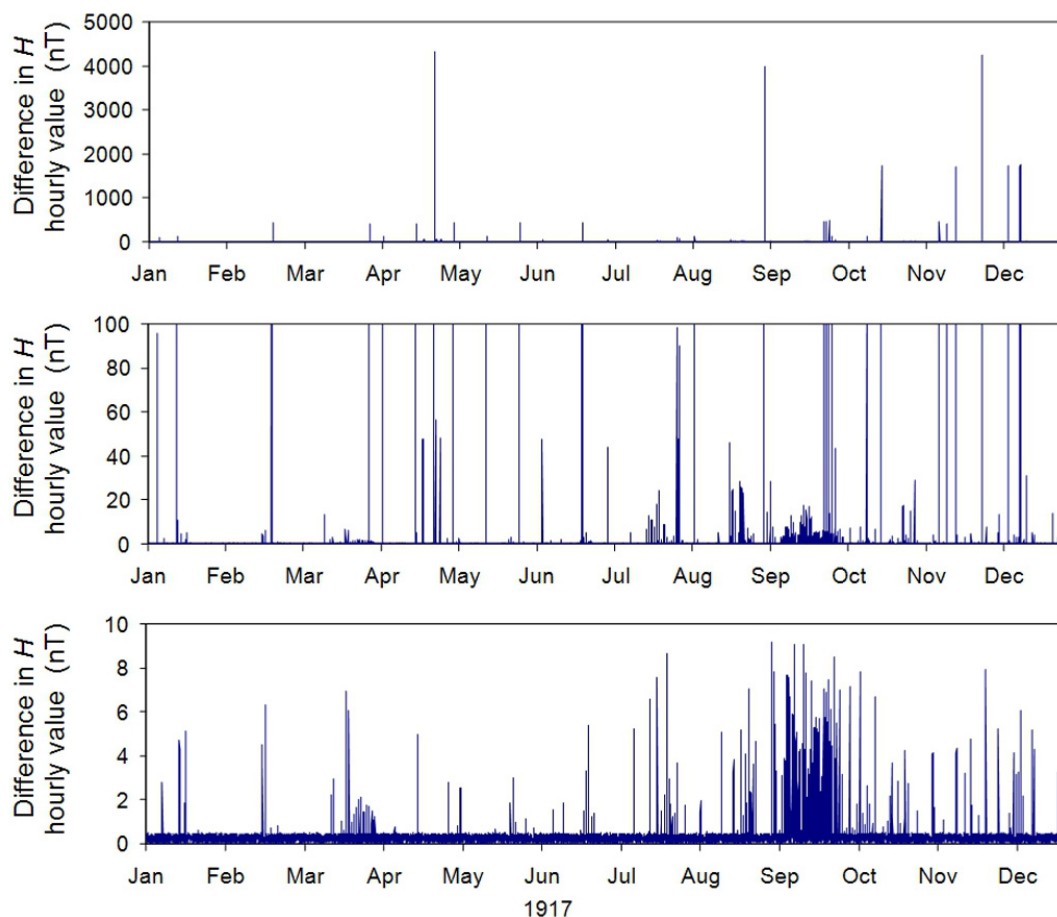


Fig. 3. Errors in the hourly mean values for 1917 from manual data entry carried out in 2007. The top panel shows full range of errors, the middle panel reduces the range to 100 nT to show more detail and the bottom panel shows the remaining errors following the corrections described in the text.

was the intention to complete the processing by removing all the remaining errors. This task has now been superseded by the discoveries and work carried out in the remainder of this paper.

5 Comparison of WDC holdings

Until recently it had been widely assumed that the WDC-format files for Eskdalemuir held at different WDCs contained the same data. However it was noted in Love et al. (2010) that the files held at Edinburgh and Kyoto WDCs were different and that the Kyoto files appeared to contain unaveraged XYZ values. A detailed comparison of the files was made by the authors, including monthly plots of differences, using the 4 options of post-processing listed in Sect. 3.

It was found that option 1 with rounding up for 0.5, resulted in the least differences between the 2 data series, particularly for Z . Table 1 lists the annual means and standard deviations of the differences in DHZ for this option. Although Table 1 shows that the overall differences are not al-

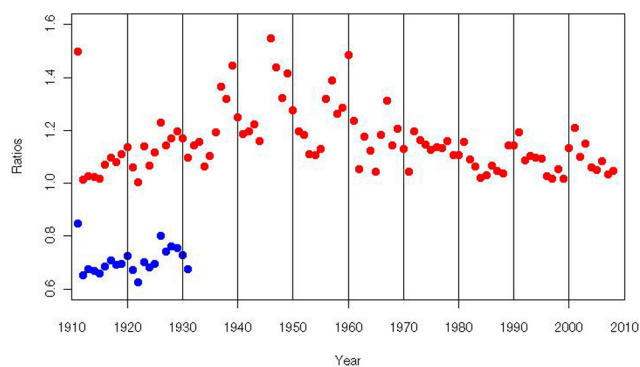


Fig. 4. Ratios of annual means of Eskdalemuir and Niemegk IHV indices based on data available at the Edinburgh WDC before (blue) and after (red) 9 August 2010.

ways zero with this option, this, along with the monthly plots of differences, was strong support for Love's claim that the data in the Kyoto WDC-format files were not post-processed.

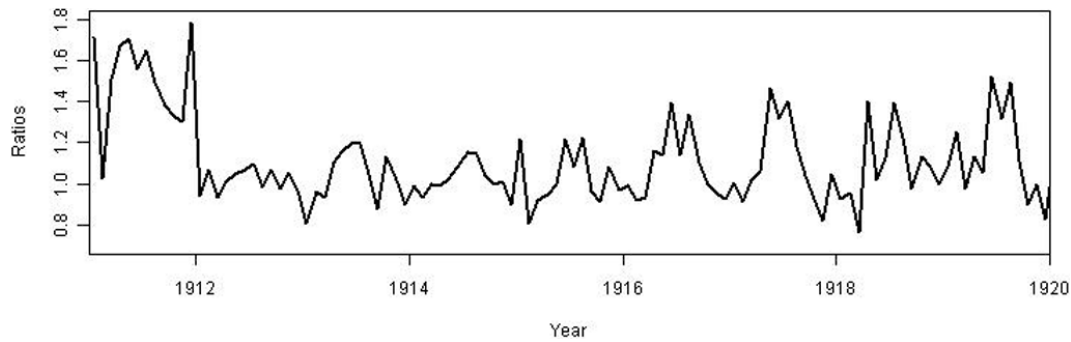


Fig. 5. Ratios of monthly means of Eskdalemuir and Niemegk IHV indices based on data available at the Edinburgh WDC after 9 August 2010.

Table 1. Annual statistics of differences between post-processed Kyoto WDC hourly values (option 1 in text) and Edinburgh WDC hourly values prior to 9 August 2010.

Year	Mean of differences in D (0.1 arc-minutes)	Standard deviation of differences in D (0.1 arc-minutes)	Mean of differences in H (nT)	Standard deviation of differences in H (nT)	Mean of differences in Z (nT)	Standard deviation of differences in Z (nT)
1911	0.0	0.1	0.0	0.1	0.0	0.0
1912	0.0	0.1	0.0	0.1	–	–
1913	0.0	0.1	0.0	0.1	0.0	0.0
1914	0.0	0.1	0.0	0.1	0.0	0.0
1915	0.0	0.1	0.0	0.1	0.0	0.0
1916	0.0	0.1	0.0	0.1	0.0	0.0
1917	1.8	87.4	–0.2	9.8	0.0	0.0
1918	0.0	0.1	0.0	0.1	0.0	0.0
1919	0.0	0.1	0.0	0.1	0.0	0.0
1920	0.0	0.1	0.0	0.1	0.0	0.0
1921	0.0	0.1	0.0	0.1	0.0	0.0
1922	0.0	0.1	0.0	0.1	0.0	0.0
1923	0.0	0.1	0.0	0.1	0.0	0.0
1924	0.0	0.1	0.0	0.1	0.0	0.0
1925	2.8	51.9	–0.4	6.8	0.0	0.0
1926	0.0	0.1	0.0	0.1	0.0	0.0
1927	0.0	0.1	0.0	0.1	0.0	0.0
1928	0.0	0.1	0.0	0.1	0.0	0.0
1929	0.0	0.1	0.0	0.1	0.0	0.0
1930	0.0	0.1	0.0	0.1	0.0	0.0
1931	0.0	0.1	0.0	0.1	0.0	0.0

Before replacing the Edinburgh files with those from Kyoto the cause of the large differences in 1917 and 1925 (Table 1) had to be found. For this the monthly plots of differences, yearbook scans and digitised values were used. For 1917 it was the July values in Y and Z in the Kyoto files that were wrong. For 1925 it was the January values in Y in the Kyoto files that were wrong and the December values in Y and Z in the Edinburgh files that contained additional errors over and above the post-processing.

Further checking of the Kyoto files revealed missing chunks of data in 1915 and 1921. Again the yearbook scans

and digitised values were used to complete the files. A final quality-control measure was to select one day at random for each year from 1911 to 1931 and to check that the values in the WDC-format files agreed with those from the yearbooks. The corrected Kyoto files were transmitted back to Kyoto and the post-processed Edinburgh files available online were replaced with the correct files on 9 August 2010.

6 Revised IHV indices

As further evidence of the reliability of the updated files, we compute IHV indices for a number of observatories and compare ratios of their annual means. In Fig. 4 it can be seen that the revised Eskdalemuir IHV index series no longer shows a large step at end of 1931 when compared with that from Niemegk. In order to investigate the anomalous value for 1911 we “zoom in” and use ratios of monthly means of IHV indices. In Fig. 5 we can see there is a step at the end of 1911. This is when the hourly values at Eskdalemuir changed from being spot values to mean values (as noted by Martini and Mursula, 2006, the headings of the yearbook tables of hourly values are not clear in this respect, but a note in the text of the yearbook for 1912 spotted by Love et al., 2010, states that the values are mean values centred on the hour). Changing from spot to mean values also happens in other observatory series (in Niemegk it happened before the start of the Eskdalemuir series) and whilst an adjustment to the resulting IHV series can be estimated for each observatory (Svalgaard and Cliver, 2007a), the ultimate remedy may lie in making high-quality scans of the original magnetograms and yearbooks and producing hourly mean values where previously only hourly spot values are available. A programme to do this is currently underway for the UK observatories. In the meantime extra information has been added to the metadata for Eskdalemuir held at <http://www.wdc.bgs.ac.uk/>. This states that 1911 hourly values are spot values centred on the hour and that 1912–1931 are mean values centred on the hour.

7 Conclusions

The significance of this work is that Eskdalemuir data are widely used in studies concerning long-term trends in magnetic field variability and inferred solar variability (for example Mursula et al., 2004; Svalgaard and Cliver, 2007b). For that reason inconsistencies and errors in the Eskdalemuir time series need to be identified and carefully corrected. Fortunately, because Eskdalemuir is but one of several observatories used in these studies, the conclusions on solar variability are unlikely to need reconsidering. However it was the detailed analyses undertaken in these solar variability studies that brought the problem reported here to light. For this we are grateful and we believe that future research using long series of hourly mean values can now be made with high confidence following the extensive analysis described here. Work is underway to improve the metadata, including time-dependent metadata, held for each observatory at the World Data Centres.

Acknowledgements. The authors thank Andrew Snedden for his help in scanning and carrying out OCR tests on the hourly value tables in the Eskdalemuir yearbooks. We also thank Kyoto WDC for supplying their Eskdalemuir WDC hourly value files. This paper is published with the permission of the Executive Director of the British Geological Survey (Natural Environment Research Council).

Topical Editor R. Nakamura thanks one anonymous referee for her/his help in evaluating this paper.

References

- Lockwood, M., Stamper, R., and Wild, M. N.: A doubling of the Sun's coronal magnetic field during the past 100 years, *Nature*, 399, 437–439, 1999.
- Love, J. J., Tsai, V. C., and Gannon, J. L.: Averaging and sampling for magnetic-observatory hourly data, *Ann. Geophys.*, 28, 2079–2096, doi:10.5194/angeo-28-2079-2010, 2010.
- Martini, D. and Mursula, K.: Correcting the geomagnetic IHV index of the Eskdalemuir observatory, *Ann. Geophys.*, 24, 3411–3419, doi:10.5194/angeo-24-3411-2006, 2006.
- Mursula, K. and Martini, D.: Centennial increase in geomagnetic activity: latitudinal differences and global estimates, *J. Geophys. Res.*, 111, A08209, doi:10.1029/2005JA011549, 2006.
- Mursula, K., Martini, D., and Karinen, A.: Did open solar magnetic field increase during the last 100 years? A reanalysis of geomagnetic activity, *Solar Phys.*, 224, 85–94, 2004.
- Svalgaard, L. and Cliver, E. W.: 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, 2007a.
- Svalgaard, L. and Cliver, E. W.: Long-term geomagnetic indices and their use in inferring solar wind parameters in the past, *Adv. Space Res.*, 40, 1112–1120, 2007b.
- Svalgaard, L., Cliver, E. W., and Le Sager, P.: Determination of interplanetary field strength, solar wind speed, and UV irradiance, 1890–present, in: *ESA-SP-535*, edited by: Wilson, A., ESTEC, The Netherlands, 15–25, 2003.
- Svalgaard, L., Cliver, E. W., and Le Sager, P.: IHV: a new long-term geomagnetic index, *Adv. Space Res.*, 34, 436–439, 2004.