Derivation of InterDiurnal Variability Index (IDV), Revisited

Leif Svalgaard[1]

[1]Cyprus Hall, W.W. Hansen Experimental Physics Laboratory, Stanford University, Stanford, CA 94305, USA

Key Points:

- = enter point 1 here =
- = enter point 2 here =
- = enter point 3 here =

Corresponding author: Leif Svalgaard, leif@leif.org

Abstract
Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

1 Introduction

It has long been thought (e.g. *Bartels* [1932]: "Observations of terrestrial-magnetic activity yield therefore not only information about geophysical influences of such solar phenomena that may be traced in astrophysical observations, but supplement these direct observations themselves") that the Earth itself could be used as an instrument to quantitatively gauge solar activity and solar wind properties. An unsuccessful, but seminal and influential, early attempt at this by *Lockwood et al.* [1999] has been superseded by the, now generally accepted, work of Svalgaard and colleagues [*Svalgaard et al.*, 2003; *Le Sager and Svalgaard*, 2004; *Svalgaard and Cliver*, 2005, 2007a,b, 2010; *Cliver and Herbst*, 2018] who introduced new geomagnetic indices (IDV, IHV, and PCP) with different dependencies (B, BV^2, and BV) on solar wind magnetic field, B, and solar wind speed, V, allowing both properties to be derived and cross-checked from this over-determined set [*Svalgaard and Cliver*, 2010; *Svalgaard*, 2014; *Owens et al.*, 2016]. They are now well-constrained (at better than the 10% level) back to the beginning of systematic observations [*Gauss and Weber*, 1837] of the variations of the geomagnetic field in the 1830s.

We first give an overview of the IDV-index: its definition and normalization, its physical interpretation, its relationship with the Heliospheric magnetic field (HMF) and with the sunspot group number. Then we describe the automatic procedure we have developed for construction of the index using modern data stored in World Data Centers (WDCs) and the INTERMAGNET repository. And we end with connecting the IDV-index so generated to the previously published series to obtain a unified index from 1835 to the present and the HMF strength inferred from the adopted series.

2 The IDV Index

The InterDiurnal Variability (IDV) index for a given geomagnetic observatory (a "station") is defined [*Svalgaard and Cliver*, 2005, 2010] as the average difference without regard to sign, from one day to the next, between hourly mean values of the Horizontal Component, H, measured one hour after local midnight. The average should be taken over a suitably long interval of time, such as one year, to eliminate various seasonal complications. IDV has the useful property of being independent of solar wind speed and is found to be highly correlated with the near-Earth heliospheric magnetic field (HMF) strength, B. Thus once IDV is determined, solar wind B can be inferred as well.

2.1 Choice of Time Interval

Svalgaard and Cliver [2005] emphasized that IDV is a modern version of the u-measure [*Joos et al.*, 1952; *Bartels*, 1932] building on ideas of more than a century ago [*Broun*, 1861; *Moos*, 1910; *van*

Bemmelen, 1903; *Schmidt*, 1918]. *Kertz* [1958], *Mayaud* [1980], and *Svalgaard* [2005] suggested using only nighttime values to avoid contamination by the regular diurnal variation. *Svalgaard and Cliver* [2005] used their IDV index (IDV05) augmented with the *u*-measure before AD 1890 to reconstruct the HMF strength for the years 1872-2004; later extending (IDV09) with many more modern stations, while using the *u*-measure for the interval 1835-1871 where it was calculated from the variability through the day [*Svalgaard and Cliver*, 2010].

Lockwood and colleagues [*Lockwood et al.*, 2013, 2014] have suggested to reduce the influence of noise in the early 19th century data by averaging the 24 individual time series of IDV calculated for each of the 24 hours of the day (IDV(1d)). In spite of the day-to-day variability of the (semi-regular) diurnal variation of the geomagnetic field, the 'IDV signature' is strong enough to reduce this extra variance to a tolerably small second-order effect.

The tension between using night hours only versus variability throughout the day was resolved in *Svalgaard* [2014]. Before 1872 there were no daily mean data available for any magnetic observatory, so to calculate the *u*-measure - "more for illustration than for actual use" - Bartels turned to use the monthly averages of the 'summed ranges' (designated *s*, being the daily sum of unsigned deviations of instantaneous values on each hour from the daily mean) supplied by *Moos* [1910] (Table 261) as the main contributor to a proxy for the *u*-measure.

2.2 The Summed Ranges

Moos [1910] noted that: "perhaps the sum of all ordinates of the inequality [deviation] without regard to signs which gives the average ordinate in 24 hours, may be considered as a more appropriate factor representing the variation due to disturbing effects. The daily range, or preferably the summed ranges, figures of the diurnal inequality of each day would probably serve as the most appropriate data for this purpose...". On any given day, the variation consists of a regular pattern (Figure 1) although varying a bit from day to day with superposed 'noise' from geomagnetic activity, thus increasing the variance; this increase is what we are interested in.

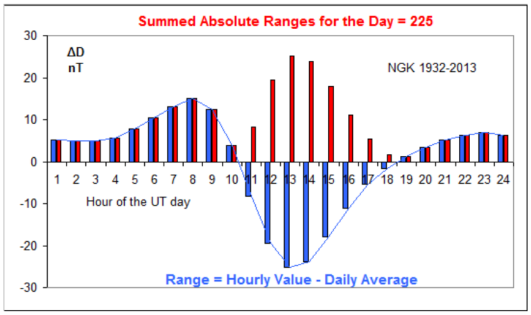

Figure 1. Average diurnal variation of Declination (expressed in force units, nT) at Niemegk (NGK). The signed deviations [blue bars determined every hour - either from an instantaneous value on the hour or from the hourly mean] from the daily mean are converted to unsigned departures (red bars) which are then summed over the day giving the Summed Ranges for each day (denoted s(D) or s(H) for the horizontal force).

Svalgaard [2014] (Appendix A) showed that using *daily s*-values computed from *hourly* values for the several 19th-century stations that have now become available (especially Helsinki [*Nevanlinna*, 2004] and several other Russian stations from the epoch of the 'Magnetic Crusade') allows determination of IDV with no ill effects. This is particularly important for early stations where the Horizontal component, H, is often very noisy and temperature-sensitive, while the Declination, D, is well-observed (or at times even the only component observed). Combining the different methods of computing IDV for 19th-century stations, *Svalgaard* [2014] re-evaluated the early IDV normalized to IDV09 (dubbing it IDV14), with which the series by *Lockwood et al.* [2014] (also scaled to IDV09) satisfactorily agree. IDV is a robust index.

The IDV index for a given station is calculated as the average (usually over a year) unsigned differences between the hourly value (mean or instantaneous on the hour) for the hour following local solar midnight of the horizontal component of the geomagnetic field. We shall denote that quantity by IDVn ['n' for 'night'] in what follows. We emphasize that IDVn for a given station is local to

the station, depending on corrected geomagnetic latitude and local underground conductivity and, in some cases, distance from the (electrically conducting) sea. We normalize s(D) to s(H) because IDVn is calculated for the horizontal force, H. The procedure is shown schematically in Figure 2.

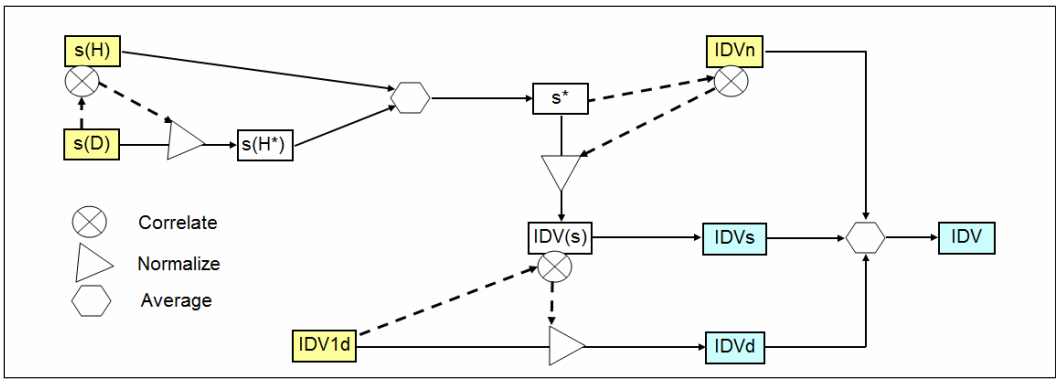

Figure 2. The observed [yellow boxes] series of Summed Ranges over a day s(H) and s(D) are correlated [circle with cross]. The fit is used to normalize [triangle] s(D) to the scale of s(H) [s(H*)] which when averaged with s(H) [hexagon] yields the composite series s*, that is correlated [circle with cross] with IDVn [other yellow box]. That fit is used to normalize s* [triangle] to the scale of IDVn to give us an alternative series IDVs [light blue box IDVs]. IDV(1d) can be normalized to the scale of IDVs [light blue box IDVd]. Finally, all three versions can be averaged. Perhaps with suitable weight factors.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

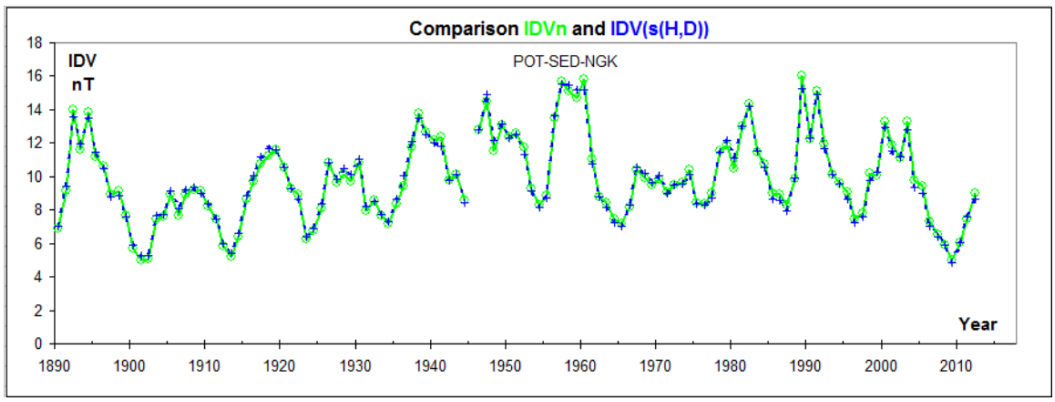

Figure 3. IDVn for POT-SED-NGK [green line] compared to IDV computed from s(H,D) [blue dashed line]. Because the two curves are so close to at times be indistinguishable, each yearly value is also marked with a symbol: green circle for IDVn and blue plus sign for IDV(s(H,D)).

2.3 Normalizing IDV

The basic tool for normalization is regressing raw IDV for NGK (i.e. calculated without applying any normalization) against the raw IDV for a station, see upper panel of Figure 4. The regression line is always almost linear with fair degree of homoscedacity and we fit it to a power-law ($NGK = a\ OBS^{\,b}$) with coefficient of determination (calculated by fitting the logarithms), R^2, in excess of 0.85 for more than 83% of the stations. The parameters a and b are then used to normalize the station 'OBS' to NGK. In (the very rare) cases where the correlation is deemed to be of too low significance, equation (1) can be used. It is immediately clear (see lower panel of Figure 4) that IDV can be determined from a single (good) station (NGK). Adding many more (ideally 'all') stations (making the series maximally 'inhomogeneous') serves three purposes: (1) ensuring that the 'master' station (NGK in this case) is not 'drifting' with respect to all the other stations with which it overlaps, (2) making the whole series as an average over many stations maximally homogeneous (i.e. reduced to the same scale), and (3) making IDV a *global* index as the additional stations can be (and are) widely distributed in latitude and, more importantly, in longitude. Here we subscribe to the view that 'more is better' as it also allows to estimate the spread of values (and thus the error bar) about the resulting average series without being biased by selection effects.

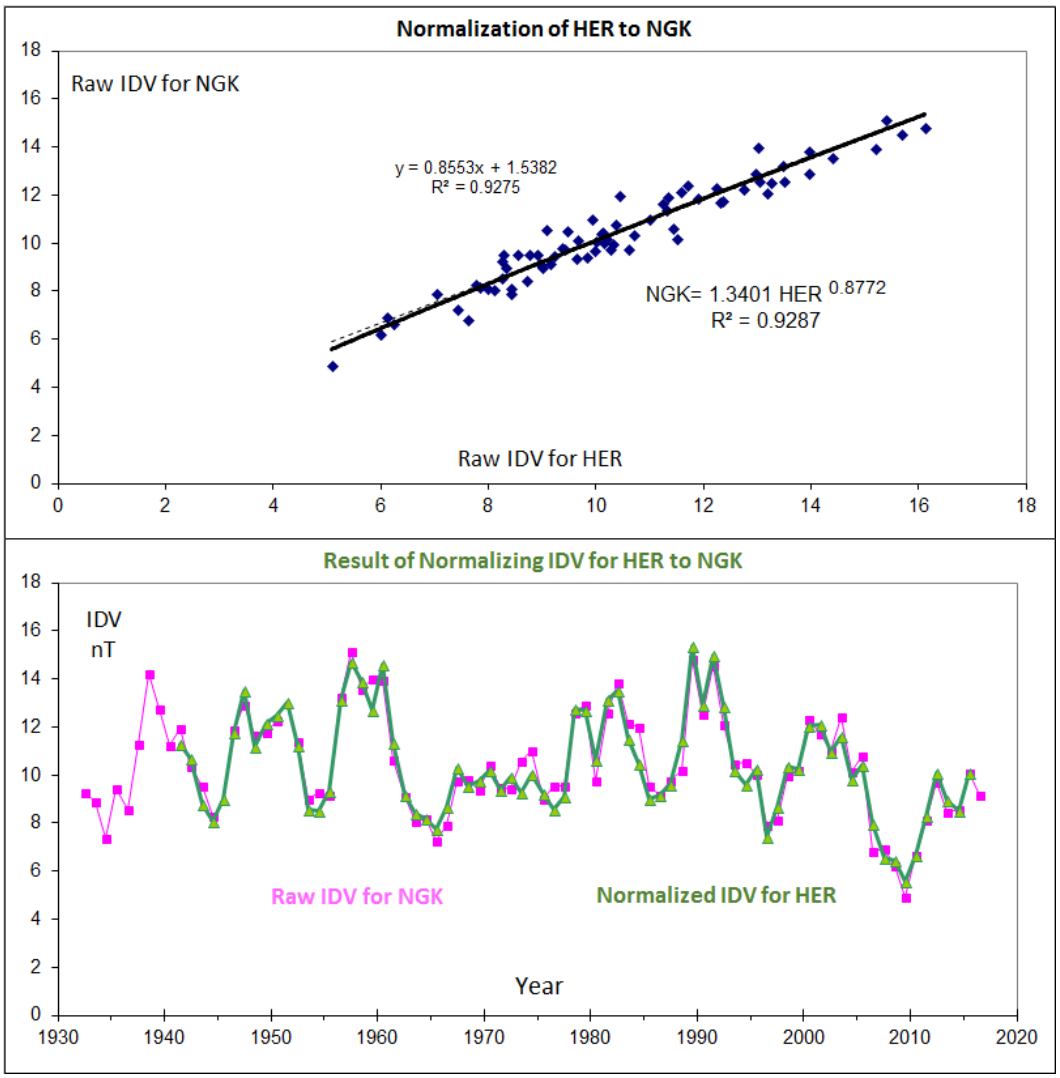

Figure 4. Upper panel: . Lower panel: .

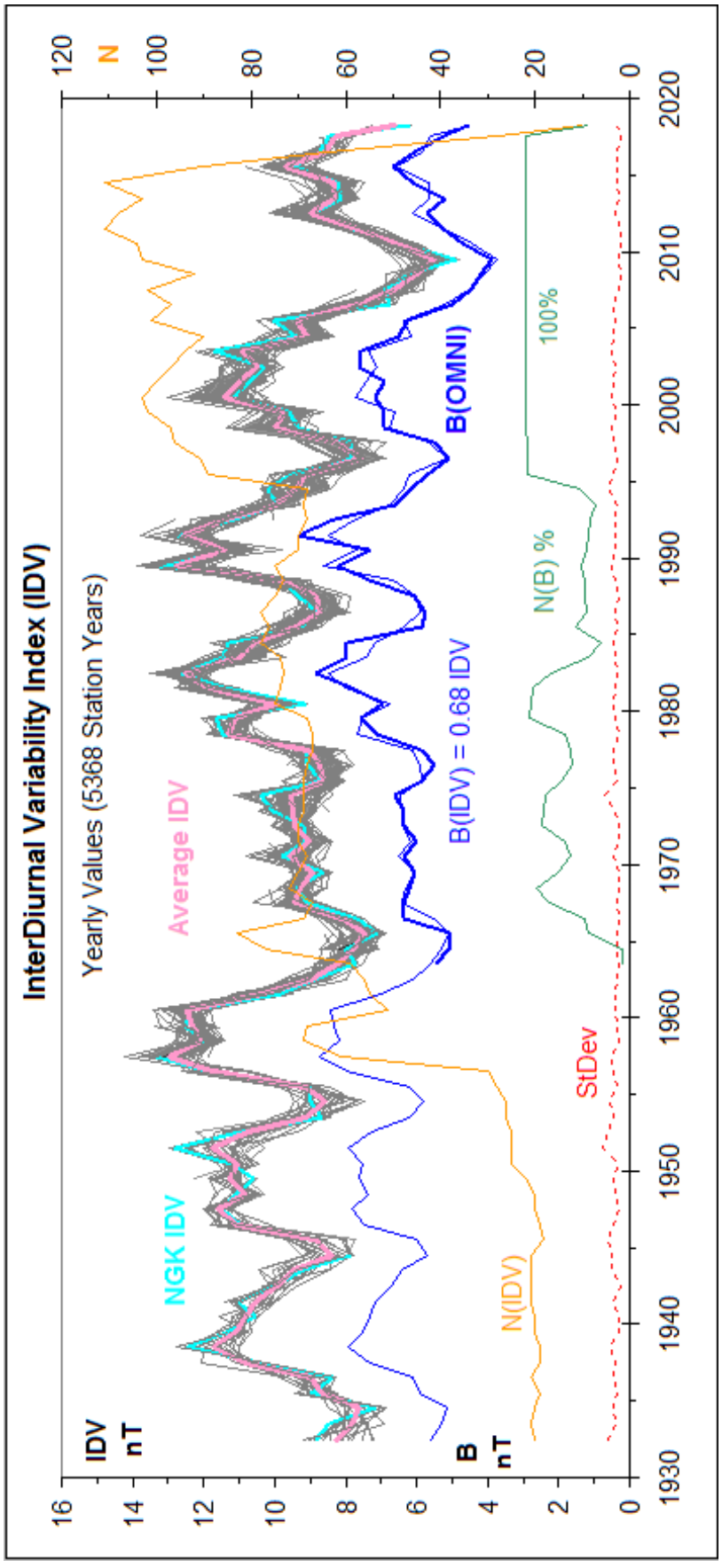

Figure 5. Variation of yearly averages of IDV for all stations (grey curves) normalized to NGK (light-blue curve). The number of stations in each average is shown by the orange curve (right-hand scale). The Standard Deviation of the ensemble is shown by the red dashed curve at the bottom of the panel. The IDV index is a good proxy for the Heliospheric Magnetic Field strength (B = 0.68 IDV_{32}, light blue curve). The observed HMF B is shown by the heavy blue curve. As shown by the green curve, the coverage of observations of B was a times poor and erratic before 1996, but for most years, the missing data did not seem to significantly degrade the relationship between IDV and B.

2.4 Latitude Dependence

IDV is ordered in Corrected Geomagnetic Latitude (β). IDV is smallest at $|\beta| = 45°$, increasing slightly towards lower latitudes, and increases above $|\beta|$ of ≈50° and increases dramatically above $|\beta|$ of ≈55°, see upper panel of Figure 6. At higher latitudes, the magnetic effects of the auroral electrojets begin to overwhelm the effect due to the ring-current, which is the physical quantity measured primarily by IDV. We therefore generally only include stations with $|\beta|$ not higher than 52°, see Figure 7. Empirically, the dependence on latitude for a given Station 'A' is somewhat weaker than the 'theoretical' $1/\cos(\beta)$ dependence that Bartels assumed for the u-measure (and used today for the D_{st} index), namely:

$$IDV(normalized\ to\ NGK) = \frac{IDV(Station\ A)}{1.26\ cos^{0.6}(\beta\ (Station\ A))} \tag{1}$$

Physically, it would have made more sense to normalize to the equator, but we retain the historical choice of NGK (originally its nearby predecessor station Potsdam (POT)).

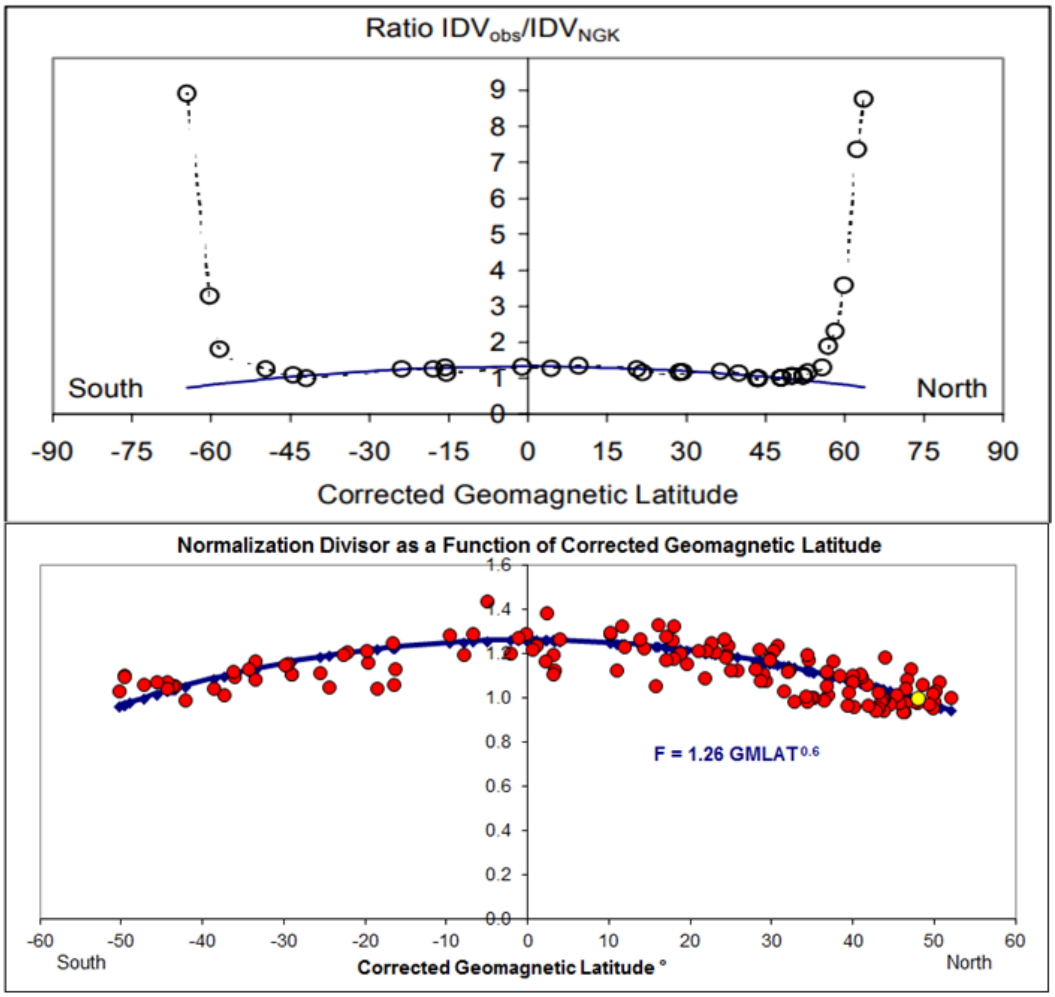

Figure 6. Upper panel: Mean ratios between yearly average raw IDV for the 34 observatories used for IDV05 and yearly average IDV for NGK over the interval 1965-2003 as a function of corrected geomagnetic latitude (β), after *Svalgaard and Cliver* [2005]. Lower panel: The variation of the empirically-derived normalization divisor for IDV18 (over the interval 1932-2018) as a function of β. The yellow dot is for NGK.

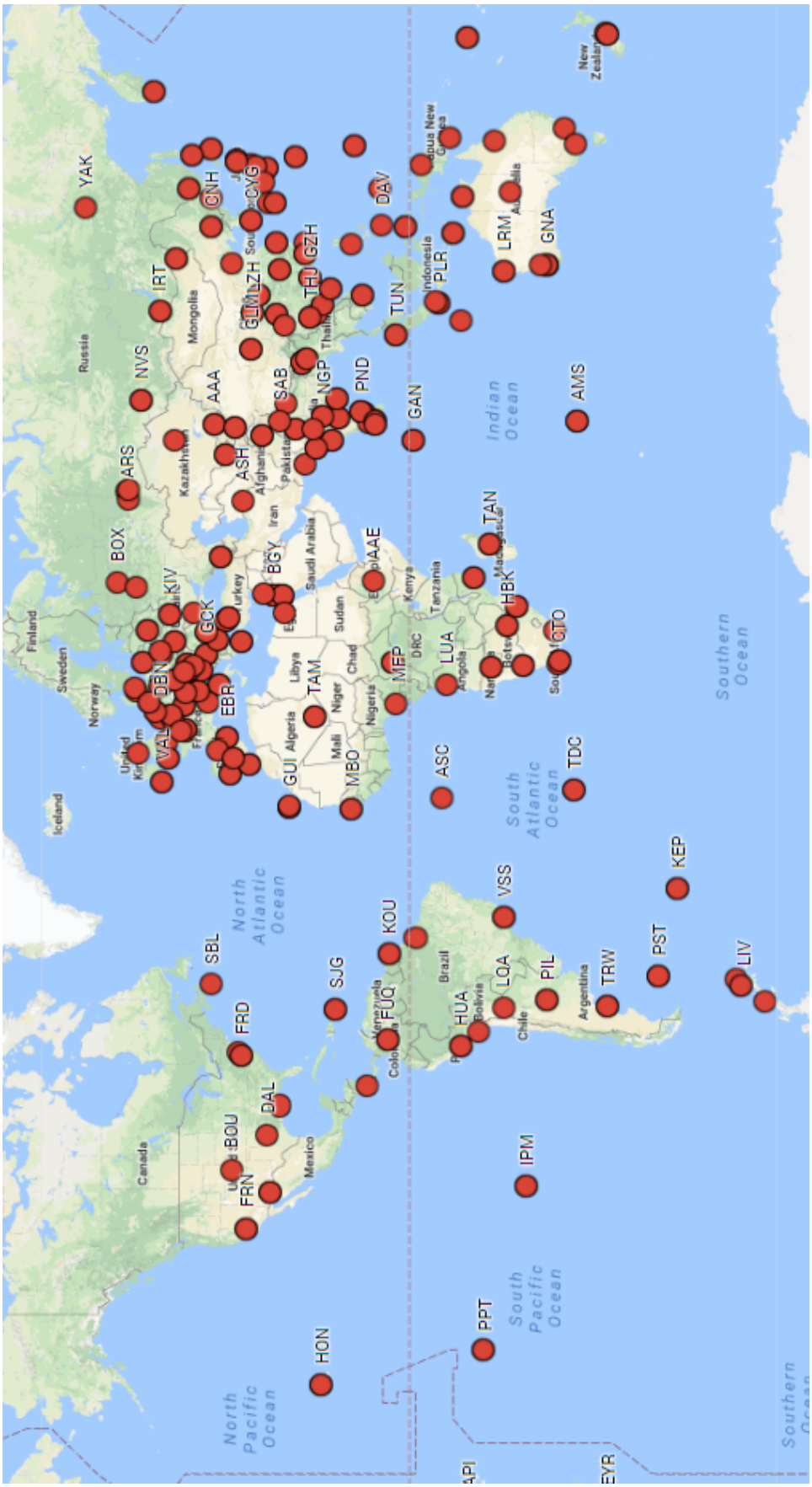

Figure 7. Upper panel: Mean ratios between yearly average IDV for the 34 observatories used for IDV05 and yearly average IDV for NGK over the interval 1965-2003 as a function of corrected geomagnetic latitude (β), after *Svalgaard and Cliver* [2005]. Lower panel: The variation of the empirically-derived normalization divisor for IDV18 (over the interval 1932-2018) as a function of β. The yellow dot is for NGK.

3 IDV is a Ring-Current Index

It was noted long ago [*Broun*, 1861; *van Bemmelen*, 1903] that the intensity of the geomagnetic elements, especially the horizontal component H, is changing sharply during a strong magnetic disturbance ('storm' ['ungewitter'] as von Humboldt called it), before slowly returning to normal during the following days. This Post-Perturbation (the 'Nachstörung') is a worldwide phenomenon: a vector directed parallel to the Earth's magnetic axis, decreasing in intensity from the equator to the poles and changing with time as a damped oscillator. It is remarkable that this important discovery was ignored and forgotten for four decades until its rediscovery at the turn of the 20th century when it formed the basis for the concept behind the u-measure. But even then, the u-measure was severely criticized until the invention of the D_{st} index by Sugiura in the 1960s [*Sugiura*, 1964]. This index has been maintained to the present day and is extensively used in magnetospheric research as a measure (albeit imperfect) of the Ring-Current in the Van Allen Belts.

3.1 IDV Measures the Negative Part of the D_{st} Index

The fact that positive and negative values of D_{st} are due to different physical processes (controlled roughly by solar wind pressure and magnetic reconnection, respectively) makes a simple yearly average of D_{st} somewhat suspect as a physical quantity. If we consider only negative values of D_{st} in the average, we are focusing on the single physical phenomenon an index is meant to represent. Indeed, *Svalgaard and Cliver* [2005] reported that IDV is closely correlated with the negative part of the D_{st} index. Figure 8 shows how closely the *negative* part of the D_{st} index (maintained by the WDC in Kyoto) matches the IDV index. The much simpler to derive IDV index is an excellent proxy for the negative part of the D_{st} index: $-D_{st} = 3.33\ IDV - 12.1$.

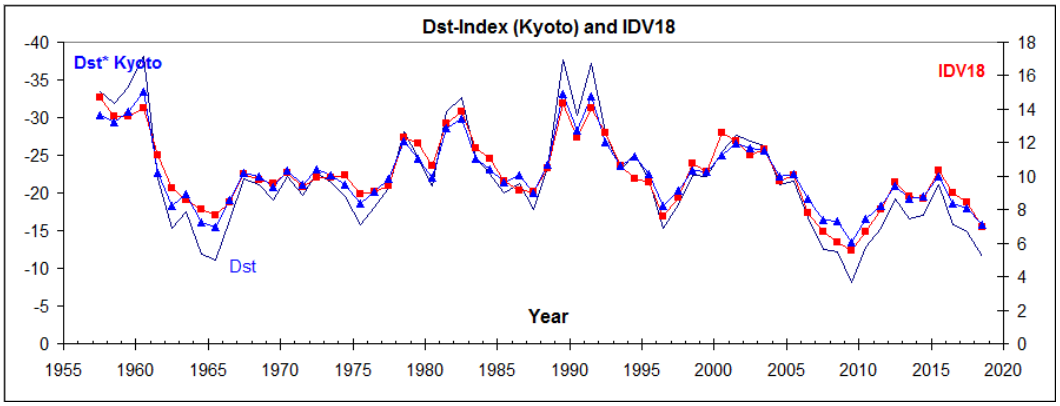

Figure 8. The D_{st}-index derived at the Kyoto WDC (thin blue line Dst) scaled (blue triangles Dst* Kyoto) to IDV (as determined in this article, red squares) shows an excellent match (R^2=0.86) between the two indices.

3.2 Is There a 'Correct' D_{st} Index?

The 'correct' construction (and perhaps physical interpretation) of D_{st} has been under debate for some time. *Love* [2007], using a wider set of stations, derived a century-long D_{st} series (J. J. Love, personal communication), extending back to 1905. He used modern signal-processing methods to separate the various [epi-]cyclic contributions of the variations [*Love and Gannon*, 2009]. The removal of the solar-quiet variation is done through time- and frequency-domain band-stop filtering, selectively removing specific Fourier terms approximating stationary periodic variation driven by the Earth's rotation, the Moon's orbit, the Earth's orbit around the Sun. The resulting non-stationary disturbance time series are weighted by observatory-site geomagnetic latitude and then averaged together across longitudes. *Karinen and Mursula* [2005] (based on *Cliver et al.* [2001]) argued

that the 'quiet-time' level of D_{st} was incorrectly determined (with a seasonally variable 'non-storm' component) and proposed that D_{st} be suitably corrected (to their D_{cx}).

Svalgaard [2005] re-examined the complicated and unsound procedure used for the standard (Kyoto) derivation of D_{st} and proposed a simpler alternative (covering 1929-2005) not suffering from the known deficiencies of the standard procedure. There are three issues to be considered:

(1) Selecting observatories at low latitude in both hemispheres, yet still away from the equatorial electrojet. He identified 14 stations (7 in the Northern Hemisphere - SJG, HON, KAK, SSH, MBO, TKT, ABG; 7 in the Southern Hemisphere - HER(CTO), API, TAN, PIL, PPT, TRW, VSS) in the low-latitude band suitable for derivation of D_{st} and with long-term coverage.

(2) Removing the Main Field. A common practice is to calculate an average 'quiet' field using the 'five quiet days' per month, except that some of these days may not be quiet at all; they just happen to be the least disturbed during that month. He chose instead to use only days where *no* 3-hour interval had an *aa*-index value exceeding 12. The next step is to compute the yearly averages of those day numbers within the year and of the geomagnetic component field value for all these 'truly' quiet days within the year. A 2nd-order polynomial fit to these yearly pairs of numbers for five years centered on the year within which we wish to derive the main field is then used to interpolate the main field for any given day within that year. Occasionally, (instrumental) discontinuities must be identified and manually corrected for.

And (3) Removing the Solar-Diurnal Variation. The daily variation is complex and varies with season (solar zenith angle) and phase of the sunspot cycle (EUV flux). Index-builders have tried to describe this complicated variation by a combination of linear terms and a smoothed 2D Fourier expansion as function of time of year and time of day (e.g. *Karinen and Mursula* [2005]). This is not entirely satisfactory, as the (un-modeled) day-to-day variation of the daily variation is as large as the variation itself. The regular solar-diurnal variation is effectively absent during the night-hours, so one can bypass the problem by only using night hours and calculate (to be dubbed) D_{sv} as the observed value minus the interpolated main field with no empirical adjustments of any kind. In the average of the two hemispheres, any seasonal variations cancel out in a natural way.

3.3 On a Yearly Timescale, it Doesn't Matter

All of these series overlap with the Kyoto-series and can be scaled to match during the intervals of overlap. As we are interested in the long-term variation of geomagnetic activity, we compute yearly averages of the negative hourly values of the each series and plot them together in Figure 9.

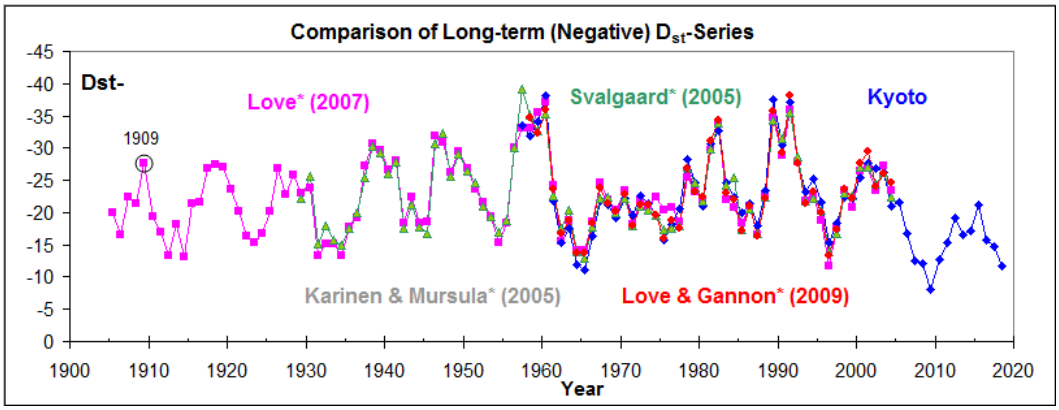

Figure 9. Comparison of the negative part of four long-term D_{st} series constructed by the authors noted on the figure using very different methods, all scaled to the level of the 'official' Kyoto series (blue symbols). Note: *the K&M 2005 curve is in preparation.*

From this we conclude that the debate about what the best method for constructing D_{st} might be, is moot on this timescale as they all substantially agree and that the variation of IDV is a satisfactory proxy for the long-term variation of negative D_{st}, as we saw in Figure 8.

4 Heliospheric Magnetic Field Strength

The OMNI dataset [*King and Papitashvili*, 2005] (http://omniweb.gsfc.nasa.gov/) collates the near-Earth solar wind measurements from numerous spacecraft. The record begins in late 1963 and continues to the present day, having grown to about 350,000 hourly values. In the present article, we correlate the IDV index with all available Heliospheric (solar wind) magnetic hourly data. Cutting the error bars of a purported relationship by a factor of two will require 200 years of additional data, so the result of the present work will be definitive for quite some time to come.

A serious concern is the stability of the long-term calibration of a dataset spanning several solar cycles and derived from many (22 for OMNI at the time of writing) not always overlapping data sources. On average, the root-mean-square value of B is $rms(B) = B/2.76$ for the years 1977-2018. We can invert that relation and calculate B from $rms(B)$: $B_{calc} = 2.76\ rms(B$,OMNI). As Figure 10 shows, the so calculated values of B (purple diamonds) match reasonably well B derived from IDV14 (pink triangles) and also given by the OMNI dataset (green circles), *except* for the years before 1977 where the values are about 22% too high (purple dashed curve). Correcting for this, brings B_{calc} down to the same level as the observed values from OMNI and those derived from IDV (that already agreed well). Because of this agreement we have confidence in the values for B itself, in spite of the problem with the *rms* (likely caused by changing observing cadence or by the way the *rms* is calculated). This may be a case where the geomagnetic data help to validate the spacecraft data.

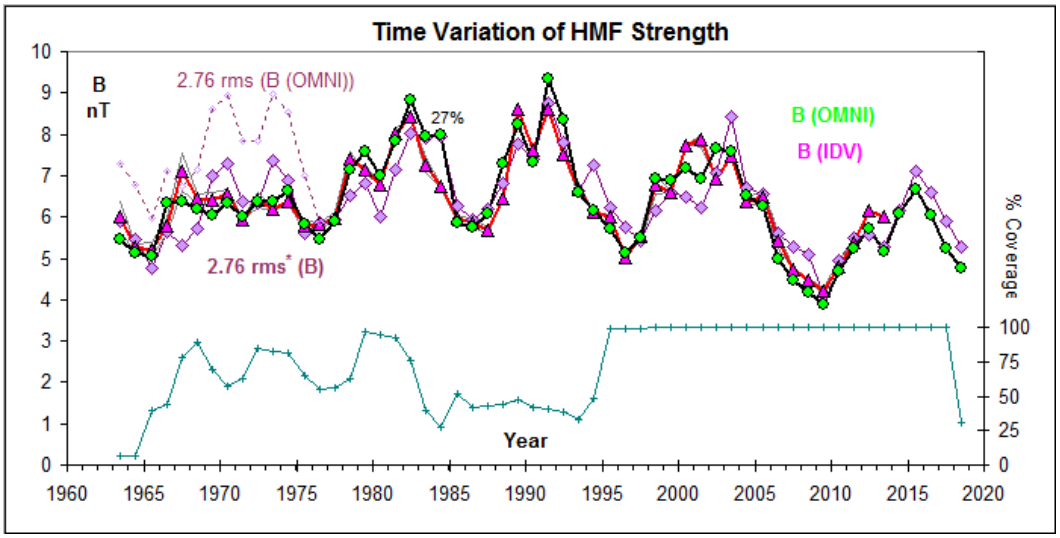

Figure 10. Variation of yearly averages of HMF B (from OMNI: green circles; from IDV: pink triangles). Average B_{calc} derived from $B_{calc} = 2.76\ rms(B)$ is shown by purple diamonds (see text for correction of the OMNI value for $rms(B)$, purple dashed curve). At bottom: percentage coverage. For 1984, the coverage was only 27%, possibly explaining the large disagreement.

5 Automatic Production of IDV

While the earlier IDV-series were 'hand-crafted' from various (sometimes hard-to-get and proprietary) data sources with judicious omission of 'obvious' outliers and a somewhat subjective selection (by experts) and 'cleansing' of data (necessary with noisy and data-poor material), the present article aims at presenting a fully automatic and flexible derivation procedure that will allow anyone (in particular an eventual hosting institution) to produce an IDV-series using high-quality and readily available public data since 1932. The resulting IDV18 series and its underlying procedure will be submitted to IAGA for consideration as an official IAGA-supported and endorsed dataset.

The process (run by a combination of configuration files, scripts, and software programs, Figure 12) works on each observatory in turn by first downloading all the yearly data from a WDC (1) and/or the INTERMAGNET repository (2) using the FTP-protocol (3 and 4). As we use hourly values, we elect to use the data in the time-honored WDC-format, mirroring the way the values were printed in the observatory yearbooks (http://www.wdc.bgs.ac.uk/catalog/format.html). The INTERMAGNET data is converted (5) to the WDC-format (program GETINT). For testing, safekeeping, and 'What-if' exploration purposes the data files may be stored (6 and 9) in a local archive (7) from which it may be later retrieved (8). The program GETWDC (11) sanity-checks the WDC-files (10), omitting years with too little coverage, resolving Y2K-issues, and the like, and outputs all the data for the observatory for all the years as a simple textual file (12) which could, of course, also be produced in other ways, serving as a common interface format to the processing following.

In the next step, the program GETMAG (13) extracts (14), initially to separate files, the time-stamped H, X, and Y components (the latter two needed for calculation of H if the observatory reported X (North) and Y (East) components). Again, at this step, various sanity-checks are performed. Any errors found by this and other programs are written to a special error-file (not shown) for later consideration. The component-files are then concatenated (15) into a file (16) to be sorted (17) in time-order, bringing together (18) data values pertaining to the same hours. The GETCMP program (19) extracts H (or computes it from X and Y, if needed) from this file and retrieves (20) from the configuration and control file 'STATIONS' the UT-hour to use for calculation of IDV and outputs to a result file (21) the one value per day for that UT-time.

In the final step, the program GETIDV (22) uses the control file (20) to guide the calculation of IDV from the result file (21). Of special interest is the column in the control file marked 'I' (Figure 11). Letters 'C' or 'O' indicate INTERMAGNET stations that are either 'C'urrent or 'O'ld. The script can be run to only include stations that are current or past INTERMAGNET stations, as well as with several other selection criteria.

Obs.	Lat	Long	CGMlat	h	Fact	Exp	QY	IA
AAA	43.3	76.9	38.2	19	1.1207	0.9045	1Y	C
AAE	9.0	38.8	-1.3	21	1.0072	0.8919	1Y	O
ABN	51.2	-0.4	49.0	0	1.7780	0.7637	1Y	. AP
AIA	-65.2	-64.3	-49.8	4	1.1114	0.9109	1Y	CM
ALH	29.1	82.0	22.9	19	0.8000	1.0000	0N	.
AML	-43.2	172.7	-49.8	12	1.0194	0.9521	1Y	.M P
ANN	11.4	79.7	3.1	19	1.3323	0.7983	1Y	.
HER	-34.4	19.2	-42.3	23	1.3401	0.8772	2Y	CM
NGK	52.1	12.7	48.0	23	1.0000	1.0000	1N	CM

Figure 11. Sample lines from the STATIONS control file. Each line gives for an observatory (denoted by its IAGA code) its position, hour to use, normalization parameters, quality (Q), whether to normalize (Y or N), INTERMAGNET status (I), and membership in other activity indices (aM, aA, and/or aP).

The final output consists of three files: a file (25) with IDV values for each month and year from the first year we have data to the last year with data, a file (23) that contains values (in Comma Separated Value-format) that can be directly used to populate a table in EXCEL, and the same, but in the newer XML-format (24). The XMLCTL program (27) adds the necessary header (26) and trailer (28) to the XML file. The station list file (20, 29) can be used as input to the program GETKML (30) to produce a layer KML-file (31) for Google Maps to show the location of the stations. The file formats and other documentation can be found in the Installation Package in the Supplementary Information. In addition, all data and programs (including their source code) are similarly available.

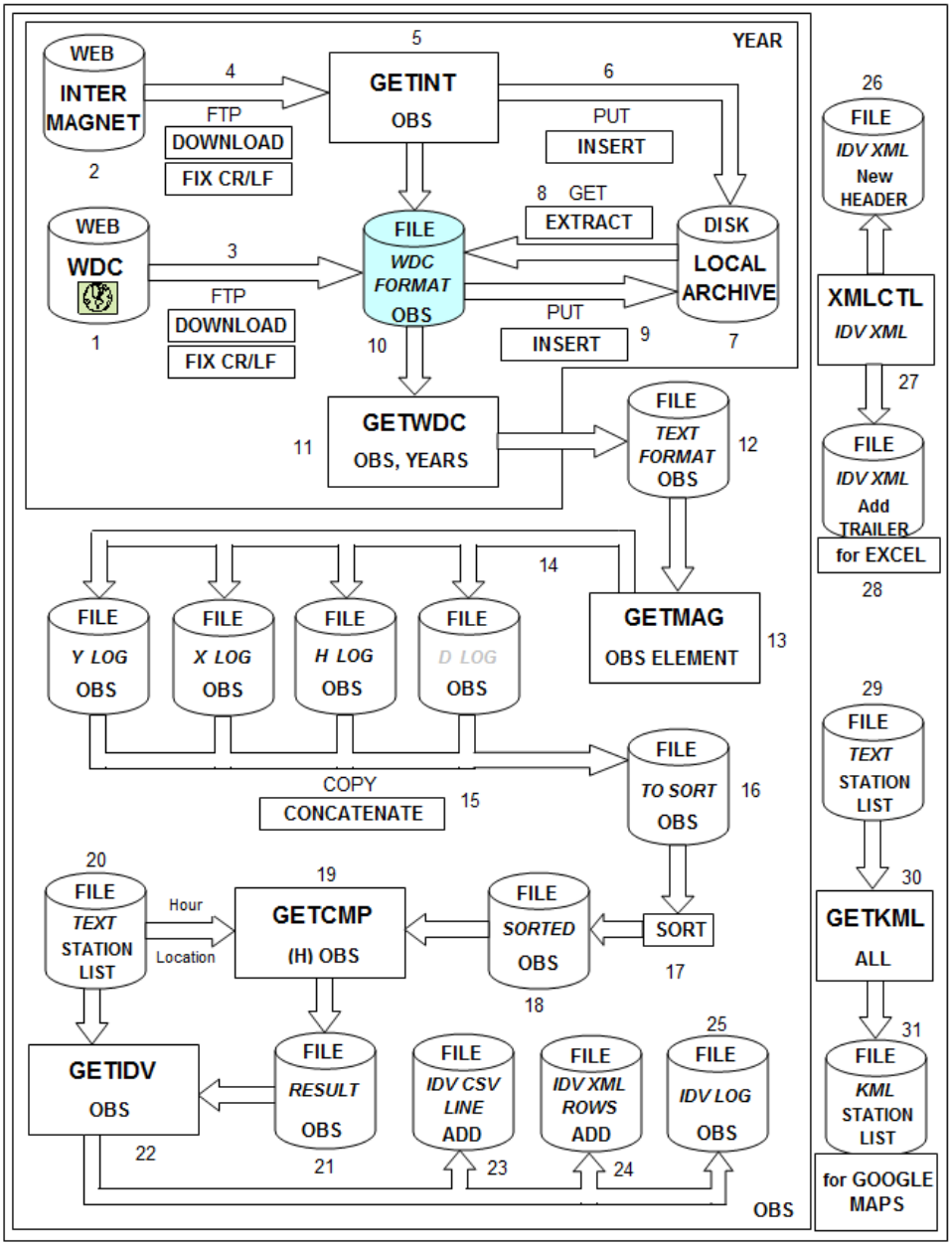

Figure 12. Flowchart of the processing steps for the automatic production of the IDV-index. The chart shows the data files and their flow through programs and scripts. Numbers next to symbols refer to explanatory notes in the text.

The INTERMAGNET Archive offers two access methods: (1) interactive access from a Web-browser (Figure 13) and (2) Unattended FTP access. Processing step #4 (in Figure 12) was a shorthand for the actual, more elaborate, processing flow depicted in Figure 14.

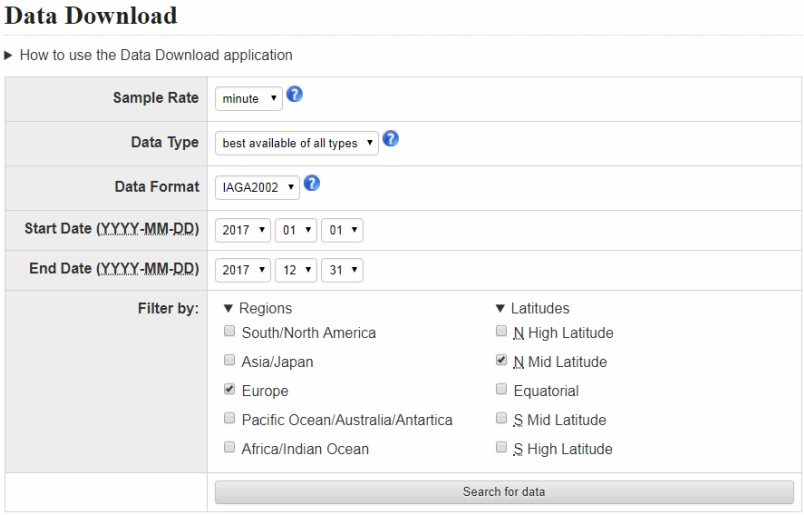

Figure 13. Interactive download form. The user (32) can select a begin-date and an end-date, but at most one year. A region and/or a latitude range must also be specified. Clicking on 'Search for data' brings up a form for selection of specific observatories.

With interactive access, the user (32) specifies the observatory and the start- and end-date on the Web form (33). The Browser returns a list of compressed data-files in the venerable ZIP-format, one for each day. The data-files encode their properties in the very *name* of the file (34), e.g. tuc20170101dmin.min.gz.zip, but are renamed (35) to retain only the essential part as obsyyyymmdd (e.g. tuc20170101.zip). The zip-file (36) contains either the data itself as text or yet another compressed file in the open gz-format (originally written to avoid patent-encumbered compression algorithms). The program GETZIP (37) determines (38) what the format is and (39) hands the file over to a script to process either (40) the zipped yearly text file (GETZ with helper-script GETZM for each month) or to the script GETI for processing of gz-files (41).

Unpacking of the data is done with either the UNZIP program (44) or the 7z program (42). In both cases, the CR/LF issue must be fixed (43 and 45) to yield the set of daily text files (46) for the observatory for the year in question. That set is processed by the GETINT program (47) to produce a yearly file (48) of the data in the WDC format ready for continued processing (49) as before (step #10 in Figure 12).

With FTP-access (50), the process is simpler: the yearly gz-files (53) are directly downloaded (51) by the script GETD (52) and passed (54) to the gz-unpacker (42) to be processed as for the interactive case. There are the usual issues about the permanence, maintenance, and reliability of the INTERMAGNET FTP-access. The scripts use a hard-coded internet URL for FTP service. Here is an example:

```
set user= anonymous
set pwd=leif@leif.org
echo user %user% %pwd%>down.ftp
echo cd intermagnet/minute/definitive/IAGA2002/year/month>>down.ftp
echo binary>> down.ftp
echo prompt>> down.ftp
echo verbose>> down.ftp
echo mget obsyearmonth*.* >>down.ftp
echo bye>>down.ftp
echo.>> down.ftp
echo downloading from remote machine
ftp -v -n -s:down.ftp ftp.seismo.nrcan.gc.ca
```

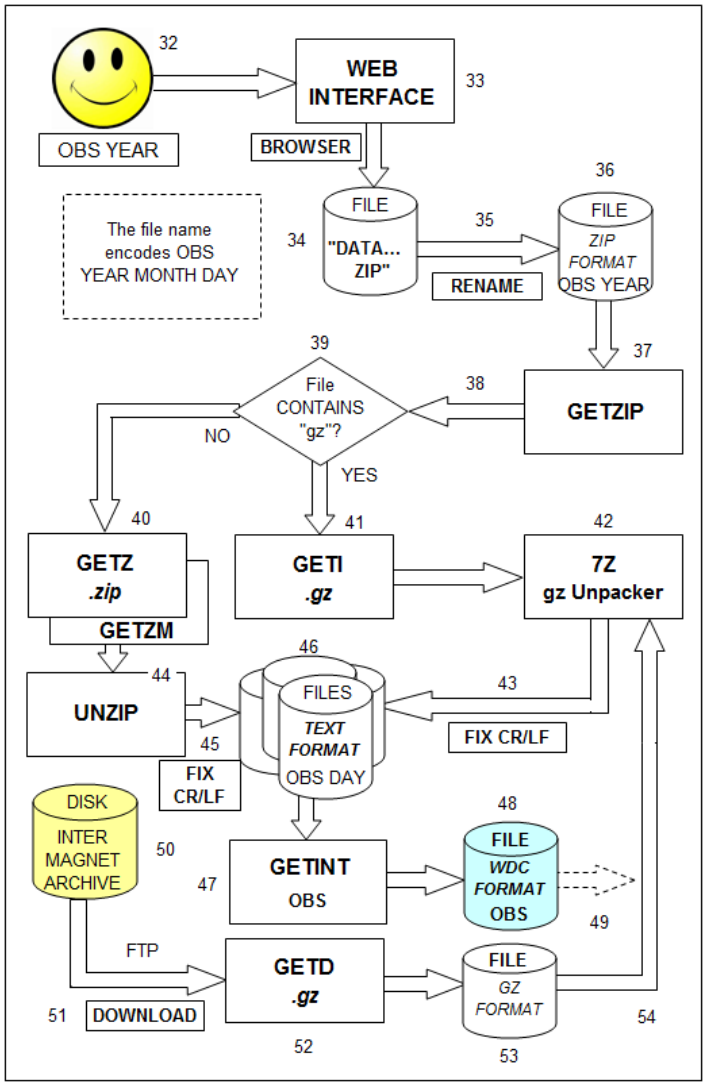

Figure 14. Flowchart of the processing steps for the automatic production of the IDV-index using data files from the INTERMAGNET Archive. The chart shows the data files and their flow through programs and scripts. Numbers next to symbols refer to explanatory notes in the text.

At the time of writing, the procedures run in a DOS box under 32-bit versions of Microsoft Windows XP, 7, Vista, and 10. If need be, they can run under 64-bit systems in a DOS-emulator, e.g. DOSBox (https://www.dosbox.com/information.php) or vDos (https://www.vdos.info/index.html), but a more cost-effective option is simply to dedicate an older (retired) machine with a 32-bit operating system to the job. The scripts are ordinary DOS batch files, and the well-structured and modular programs are written in a (portable) minimal subset of COBOL using only a handful of verbs (move, compute(add, subtract), perform(exit), if(else), display(accept), call). Because the programs basically read, write, and reformat numbers as textual character-based data, easy-to-understand COBOL is an efficient tool for this. All file-access is purely sequential and done through an abstraction layer hiding the (non-portable) intricacies (such as file name extensions). The philosophy behind these choices is akin to that behind the use of FORTRAN for compute-intensive problems in Particle Physics and Weather/Climate-modeling.

6 Capping IDV at an Upper Limit

Earlier versions of IDV imposed a limit on the daily value of IDV of 75 nT in order not to give undue weight to exceptionally large values during great storms. With an automatic process to construct a new IDV series (typical run time of the order of one hour using all stations) it is now trivial to experiment with the limit value and find the cap that gives the best correlation with the observed HMF B.

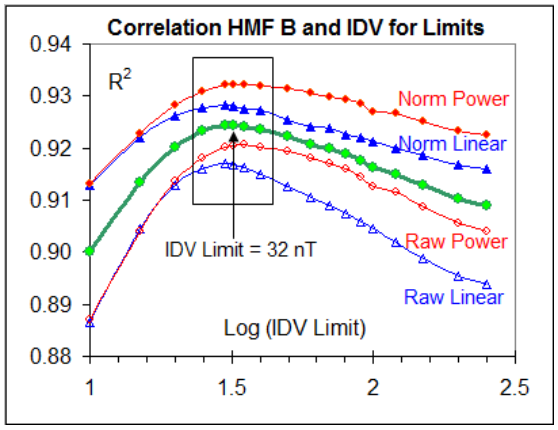

Figure 15. The Coefficient of Determination (R^2) for the correlation between yearly values of the HMF B at Earth for years 1964-2018 and the IDV-index capped at an upper limit blah blah pical run time of the order of one hour using all stations) it is now trivial to experiment with the limit value and find the cap that gives the best c

7 More to come...

In all its Merciless Glory ...

8 Adopted IDV18

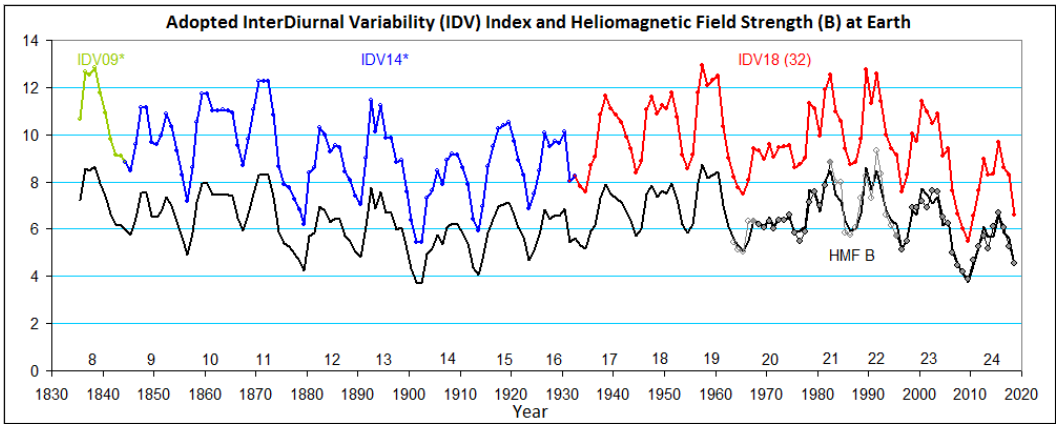

Figure 16. Variation of yearly averages of HMF B (from OMNI: green circles; from IDV: pink triangles). Average B derived from $B = 2.76\ rms(B)$ is shown by purple diamonds (see text for correction of the OMNI value for $rms(B)$, purple dashed curve). At bottom: percentage coverage. For 1984, the coverage was only 27%, possibly explaining the large disagreement.

It is often claimed (e.g. *Russell et al.* [2010]) that the recent minimum between solar cycles 23 and 24 was highly unusual (perhaps even unprecedented!) in duration and depth. With the extended perspective provided by almost two centuries of IDV (and hence of B) we see that the minima between cycles 11 and 12 and, especially, between 13 and 14 (Figure 17) were equally deep and 'late'. "There is nothing new under the sun". We are simply returning to conditions of a century ago (and, of course, of many other such past low activity periods). In addition, it is clear that the latter half of the 20th century has not been exceptionally active ("The most active in 10,000 years" as is sometimes erroneously claimed) compared to the mid-19th century before it.

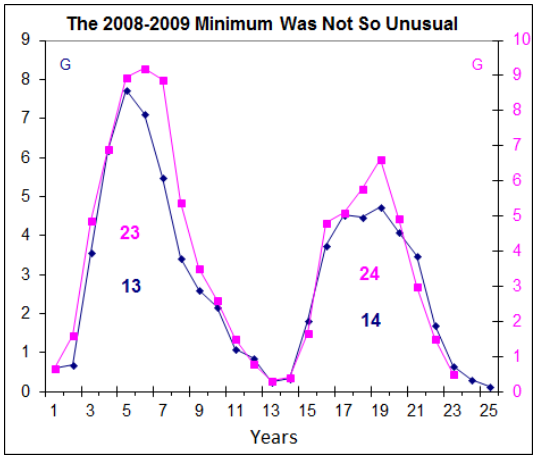

Figure 17. HMF (not unusual... more more more more more more

9 IDV and Sunspot Group Numbers

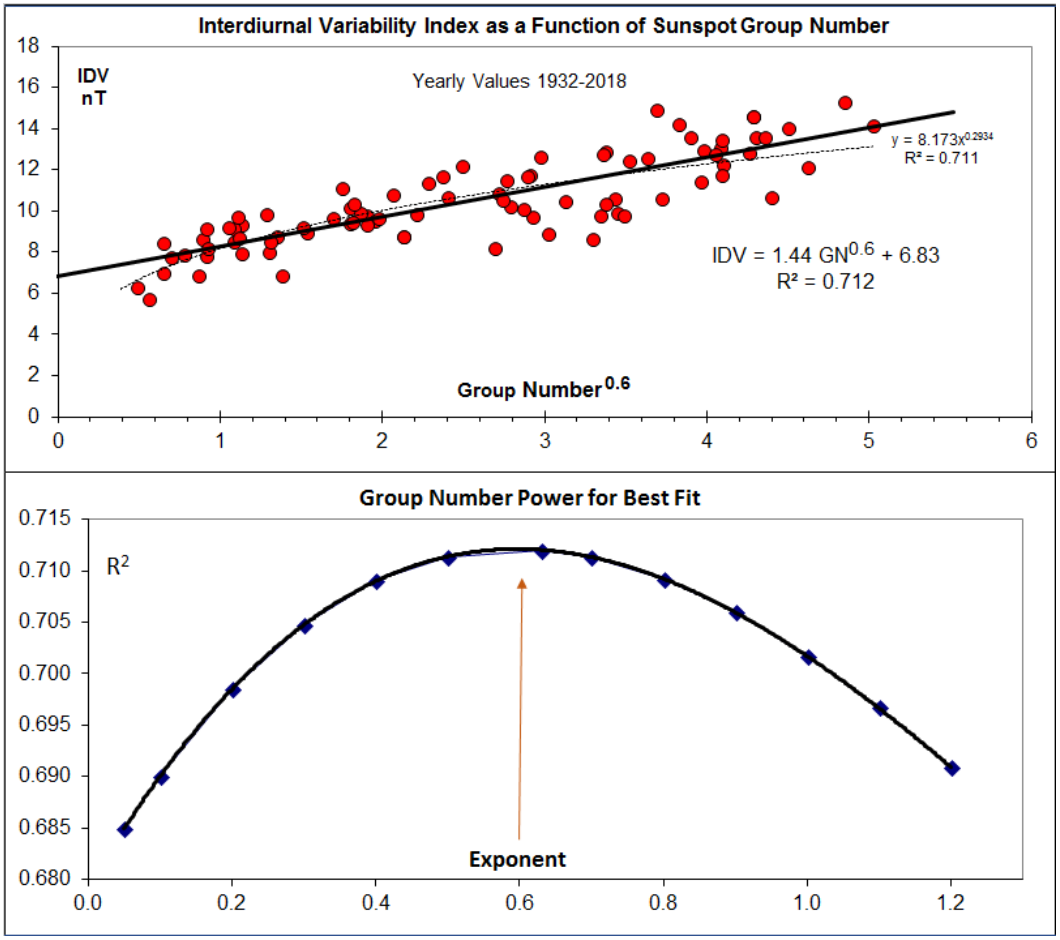

Figure 18. Variation of yearly averages of HMF B (from OMNI: green circles; from IDV: pink triangles). Average B derived from $B = 2.76\ rms(B)$ is shown by purple diamonds (see text for correction of the OMNI value for $rms(B)$, purple dashed curve). At bottom: percentage coverage. For 1984, the coverage was only 27%, possibly explaining the large disagreement.

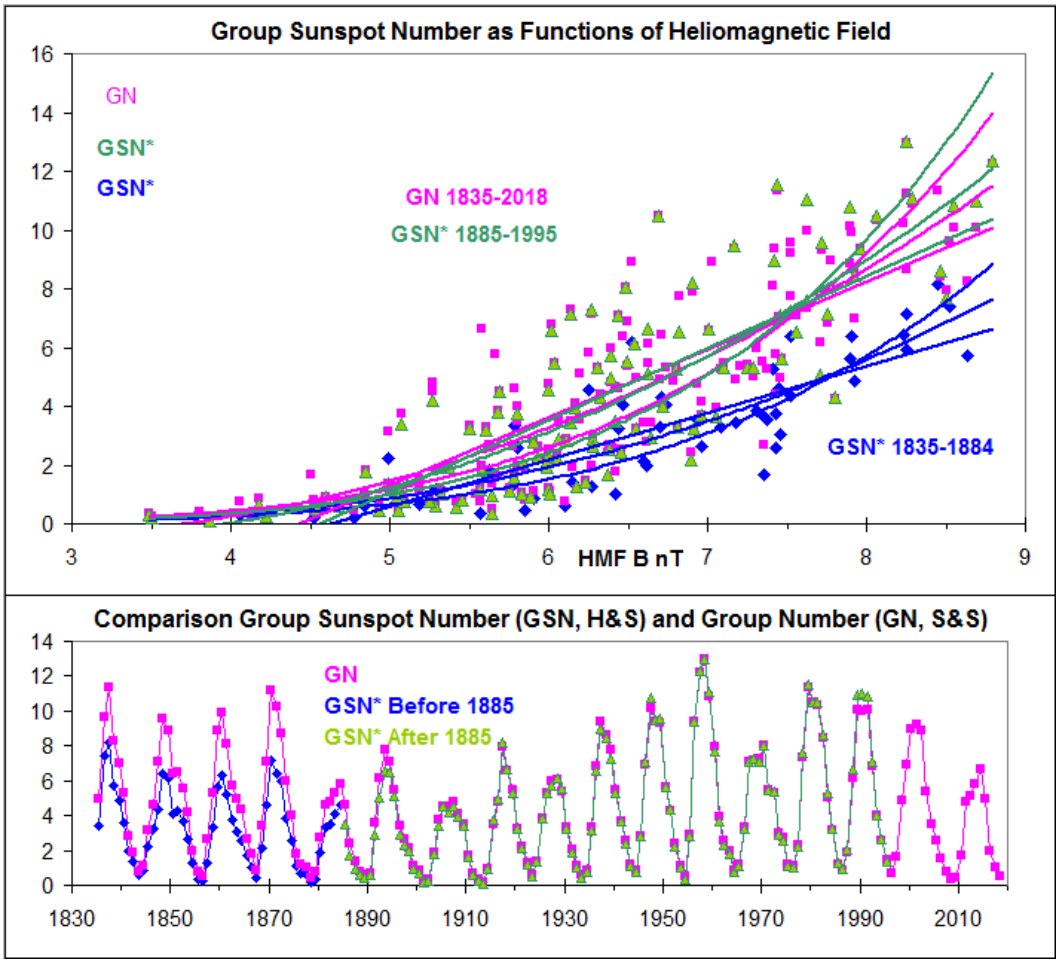

Figure 19. Variation of yearly averages of HMF B (from OMNI: green circles; from IDV: pink triangles). Average B derived from $B = 2.76\ rms(B)$ is shown by purple diamonds (see text for correction of the OMNI value for $rms(B)$, purple dashed curve). At bottom: percentage coverage. For 1984, the coverage was only 27%, possibly explaining the large disagreement.

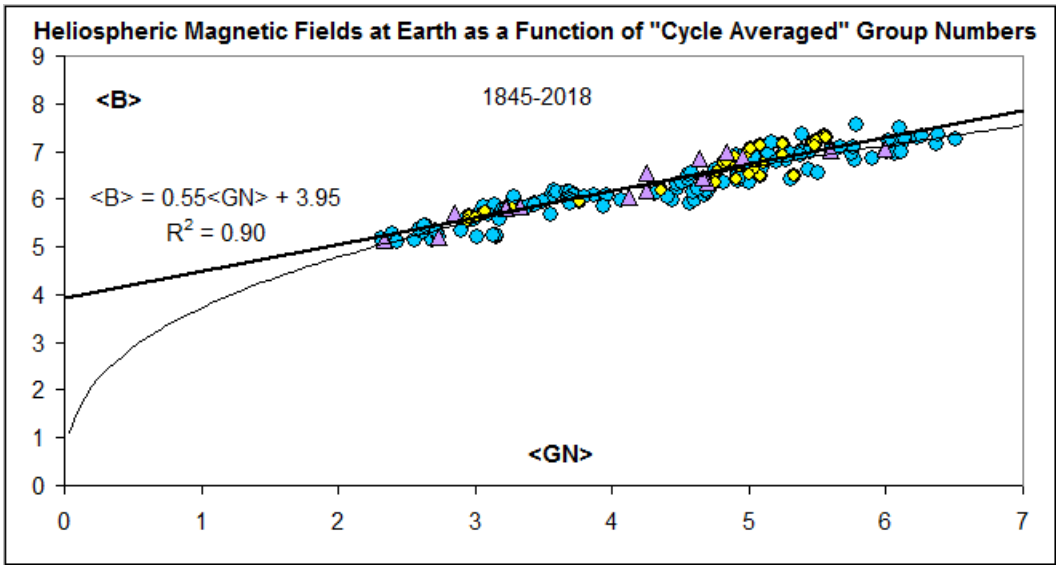

Figure 20. Variation of cycle-averaged Heliospheric Magnetic Field strength <B> [at Earth] as a function of the cycle-averaged sunspot Group Number <GN> for 1845-2018. Purple triangles show single averages for each cycle. Blue dots show running 11-year averages, while yellow diamonds show running 11-year averages for the years before 1885. The thin curve shows a power-law fit to the data points and as such is forced to go through the origin. The thick line is a least-square fit to the purple data points.

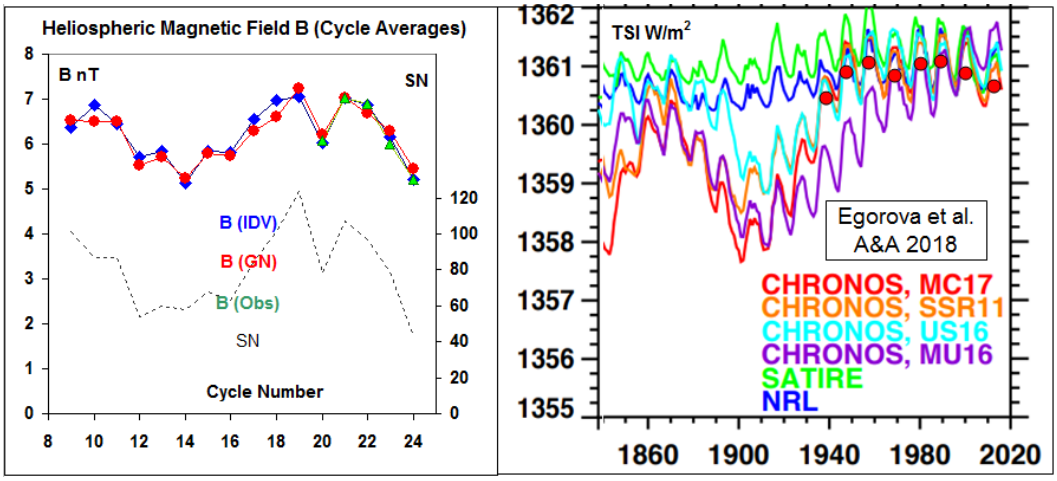

Figure 21. Left: Solar cycle (from minimum to minimum) Variation of cycle averages of HMF B (from IDV: blue diamonds; from Group Number: red circles; Observed in situ from OMNI: green triangles). The average Sunspot Number (SN) is shown for reference by the thin dashed line. Right: Reconstructions of Total Solar Irradiance (adapted from *Egorova et al.* [2018] based an assumed background level given by 22-yr running averages of cosmic ray modulation potentials). The red circles show the variation for times for which the modulation potential was actually based on observations.

10 Discussion and Conclusion

For me.

Acknowledgments
The research presented in this article rely on the data collected at magnetic observatories worldwide, and we thank the national institutions that support them. We thank the WDCs in Edinburgh (UK) and Kyoto (Japan) for their effort in curating hourly values of the geomagnetic elements and making them available through their excellent websites. We also recognize the role of the INTERMAGNET program in promoting high standards of magnetic observatory practice. We acknowledge use of NASA/GSFC's Space Physics Data Facility's OMNIWeb (ftp) service, and the OMNI data. LS thanks Stanford University for continuing support.

References

Bartels, J., Terrestrial-magnetic activity and its relations to solar phenomena, *Terr. Mag. Atm. Electr.*, **37**(1), 1, doi:10.1029/TE037i001p00001.

Broun, J. A., On the Horizontal Force of the Earth's Magnetism, *Edinburgh Roy. Soc. Trans.*, **22**, 511.

Cliver, E. W., Y. Kamide, A. G. Ling, and N. Yokoyama (2001), Semiannual variation of the geomagnetic Dst index: Evidence for a dominant nonstorm component, *J. Geophys. Res.*, **106**(A10), 21297-21304, doi:10.1029/2000JA000358.

Cliver, E. W., and K. Herbst, Evolution of the Sunspot Number and Solar Wind B Time Series, *Space Science Rev.*, **214**(2), #56, doi:10.1007/s11214-018-0487-4.

Egorova, T., W. Schmutz, E. Rozanov, A. I. Shapiro, I. Usoskin, J. Beer, R. V. Tagirov, and T. Peter (2018), Revised historical solar irradiance forcing, *Astron. & Astroph.*, **615**, id.A85, 10 pp. doi:10.1051/0004-6361/201731199.

Gauss, C. F., and W. Weber (1837), *Resultate aus den Beobachtungen des magnetischen Vereins im Jahre 1836*, 150 pp., Dieterich, Göttingen.

Joos, G., J. Bartels, and P. Ten Bruggencate (1952), Landolt-Börnstein: Zahlenwerte und funktionen aus physik, chemie, astronomie, geophysik und technik, in *Astronomie und Geophysik*, **18**, Springer, New York.

Karinen, A., and K. Mursula (2005), A new reconstruction of the Dst index for 1932-2002, *Ann. Geophys.*, **23**(2), 475-485, doi:10.5194/angeo-23- 475-2005.

Kertz, W. (1958), Ein neues mass für die feldstärke des erdmagnetischen äquatorialen Ringstroms, *Abh. Akad. Wiss. Göttingen Math. Phys.*, **2**, 1-83.

King, J. H., and N. E. Papitashvili (2005), Solar wind spatial scales in and comparisons of hourly Wind and ACE plasma and magnetic field data, *J. Geophys. Res.*, **110**, A02104, doi:10.1029/2004JA010649.

Le Sager, P., and L. Svalgaard (2004), No increase of the interplanetary electric field since 1926, *J. Geophys. Res.*, **109**, A07106, doi:10.1029/2004JA010411.

Lockwood, M., R. Stamper, and M. N. Wild (1999), A doubling of the Sun's coronal magnetic field during the past 100 years, *Nature*, **399**(6735), 437-439, doi: 10.1038/20867.

Lockwood, M., L. Barnard, H. Nevanlinna, M. J. Owens, R. G. Harrison, A. P. Rouillard, and C. J. Davis (2013), Reconstruction of geomagnetic activity and near-Earth interplanetary conditions over the past 167 yr - Part 1: A new geomagnetic data composite, *Ann. Geophys.*, **31**, 1957-1977, doi:10.5194/angeo-31-1957-2013.

Lockwood, M., H. Nevanlinna, M. Vokhmyanin, D. Ponyavin, S. Sokolov, L. Barnard, M. J. Owens, R. G. Harrison, A. P. Rouillard, and C. J. Scott (2014), Reconstruction of geomagnetic activity and near-Earth interplanetary conditions over the past 167 yr - Part 3: Improved representation of solar cycle 11, *Ann. Geophys.*, **32**, 367-381, doi:10.5194/angeo-32-367-2014, 2014.

Love, J. J. (2007), A continuous long-term record of magnetic-storm occurrence and intensity, *Eos Trans. AGU*, **88**(23), Joint. Assembl. Suppl., Abstract SH54B-03 (http://www.leif.org/research/AGU%20Spring%202007%20SH54B-03.pdf).

Love, J. J, and J. L. Gannon (2009), Revised Dst and the epicycles of magnetic disturbance: 1958-2007, *Ann. Geophys.*, **28**(8), 3101-3131, doi:10.5194/angeo-27-3101-2009

Mayaud, P.-N. (1967), Calcul preliminaire d'indices Km, Kn et Ks ou Am, An, et As, mesures de l'activité magnétique a l'échelle mondiale et dans les hémispheres Nord et Sud, *Ann. Géophys.*,

23, 585-617.

Mayaud, P.-N. (1980), Derivation, Meaning, and Use of Geomagnetic Indices, *Geophys. Monogr. Ser.*, **22**, 154 pp., AGU, Washington D. C.

Moos, N. A. F. (1910), The Phenomenon and its Discussion, in *Colaba Magnetic Data, 1846 to 1905*, 782 pp., Cent. Gov. Press, Bombay, India.

Nevanlinna, H. (2004), Results of the Helsinki magnetic observatory 1844-1912, *Ann. Geophys.*, **22**, 1691-1704, doi:10.5194/angeo-22-1691-2004.

Owens, M. J., E. W. Cliver, K. G. McCracken, J. Beer, L. Barnard, M. Lockwood, A. P. Rouillard, D. Passos, P. Riley, I. G. Usoskin, and Y.-M. Wang (2016), Near-Earth heliospheric magnetic field intensity since 1750: 1. Sunspot and geomagnetic reconstructions, *J. Geophys. Res. Space Physics*, **121**, 6048-6063, doi:10.1002/2016JA022529.

Russell, C. T., J. G. Luhmann, and L. K. Jian (2010), How unprecedented a solar minimum?, *Rev. Geophys.*, **48**, doi:10.1029/2009RG000316.

Schmidt, A. (1918), Über den Verlauf der Nachstörung, *Archiv des erdmagnetismus*, Heft **3**, 12, Georg Reimer, Berlin.

Schmidt, A. (1926), Über die Abhängigkeit der täglichen Variation vom Zustande der Sonne, *Archiv des erdmagnetismus*, Heft **4**, 4, Georg Reimer, Berlin.

Svalgaard, L. (1977), Geomagnetic activity: Dependence on solar wind parameters, in *Skylab Workshop on Coronal Holes*, edited by J. B. Zirker, pp. 371-441, Colo. Assoc. Univ. Press, Boulder, CO.

Svalgaard, L. (2005), Re-derivation of D_{st}, *Eos Trans. AGU*, **86**(52), Fall Meet. Suppl., Abstract SA12A-04.

Svalgaard, L. (2014), Correction of errors in scale values for magnetic elements for Helsinki, *Ann. Geophys.*, **32**, 633-641, doi:10.5194/angeo-32-633-2014.

Svalgaard, L., and E. W. Cliver (2005), The IDV index: Its derivation and use in inferring long-term variations of the interplanetary magnetic field strength, *J. Geophys. Res.*, **110**, A12, doi:10.1029/2005JA011203.

Svalgaard, L., and E. W. Cliver (2007a), 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.

Svalgaard, L., and E. W. Cliver (2007b), Long-term geomagnetic indices and their use in inferring solar wind parameters in the past, *Adv. Space Res.*, **40**, 1112-1120, doi:10.1016/j.asr.2007.06.066.

Svalgaard, L., and E. W. Cliver (2010), Heliospheric magnetic field 1835-2009, *J. Geophys. Res.*, **115**, A09111, doi:10.1029/2009JA015069.

Svalgaard, L., E. W. Cliver, and P. LeSager (2003), Determination of interplanetary magnetic field strength, solar wind speed and EUV irradiance, 1890-2003, in *Proceedings of ISCS 2003 Symposium: Solar Variability as an Input to the Earth's Environment*, edited by A. Wilson, pp. 15-23, Eur. Space Agency, Paris, Tatranska Lomnica, Slovakia.

Sugiura, M. (1964), Hourly values of equatorial Dst for the IGY, *Annals of the International Geophysical Year*, **35**, 945-948.

van Bemmelen, W. (1903), The diurnal field of magnetic disturbance, *Terr. Mag. Atm. Electr.*, **8**(4), 153-174, doi:/10.1029/TE008i004p00153.