Time-of-day/time-of-year response functions of planetary geomagnetic indices

Mike Lockwood, Aude Chambodut, Ivan D Finch, Luke Barnard, Mathew J Owens, Carl Haines

▶ **To cite this version:**

Mike Lockwood, Aude Chambodut, Ivan D Finch, Luke Barnard, Mathew J Owens, et al.. Time-of-day/time-of-year response functions of planetary geomagnetic indices. Journal of Space Weather and Space Climate, 2019, 9, pp.A20. 10.1051/swsc/2019017 . hal-02369543

HAL Id: hal-02369543
https://hal.science/hal-02369543

Submitted on 19 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

J. Space Weather Space Clim. 2019, **9**, A20
© M. Lockwood et al., Published by EDP Sciences 2019
https://doi.org/10.1051/swsc/2019017

JSWSC

Available online at:
www.swsc-journal.org

RESEARCH ARTICLE OPEN ACCESS

Time-of-day/time-of-year response functions of planetary geomagnetic indices

Mike Lockwood[1,*], Aude Chambodut[2], Ivan D. Finch[3], Luke A. Barnard[1], Mathew J. Owens[1], and Carl Haines[1]

[1] Department of Meteorology, University of Reading, Whiteknights Campus, Earley Gate, PO Box 243, Reading RG6 6BB, UK
[2] Institut de Physique du Globe de Strasbourg, UMR7516, Université de Strasbourg/EOST, CNRS, 5 rue René Descartes, Strasbourg Cedex 67084, France
[3] Rutherford Appleton Laboratory, Science and Technology Facilities Council, Harwell Campus, Didcot OX11 0QX, UK

Received 11 January 2019 / Accepted 1 April 2019

Abstract – *Aims:* To elucidate differences between commonly-used mid-latitude geomagnetic indices and study quantitatively the differences in their responses to solar forcing as a function of Universal Time (*UT*), time-of-year (*F*), and solar-terrestrial activity level. To identify the strengths, weaknesses and applicability of each index and investigate ways to correct for any weaknesses without damaging their strengths.
Methods: We model how the location of a geomagnetic observatory influences its sensitivity to solar forcing. This modelling for a single station can then be applied to indices that employ analytic algorithms to combine data from different stations and thereby we derive the patterns of response of the indices as a function of *UT*, *F* and activity level. The model allows for effects of solar zenith angle on ionospheric conductivity and of the station's proximity to the midnight-sector auroral oval: it employs coefficients that are derived iteratively by comparing data from the current *aa* index stations (Hartland and Canberra) to simultaneous values of the *am* index, constructed from chains of stations in both hemispheres. This is done separately for eight overlapping bands of activity level, as quantified by the *am* index. Initial estimates were obtained by assuming the *am* response is independent of both *F* and *UT* and the coefficients so derived were then used to compute a corrected *F*-*UT* response pattern for *am*. This cycle was repeated until it resulted in changes in predicted values that were below the adopted uncertainty level (0.001%). The ideal response pattern of an index would be uniform and linear (i.e., independent of both *UT* and *F* and the same at all activity levels). We quantify the response uniformity using the percentage variation at any activity level, $V = 100\ (\sigma_S/\langle S\rangle)$, where S is the index's sensitivity at that activity level and σ_S is the standard deviation of S: both S and σ_S were computed using the eight *UT* ranges of the 3-hourly indices and 20 equal-width ranges of *F*. As an overall metric of index performance, we take an occurrence-weighted mean of V, V_{av}, over the eight activity-level bins. This metric would ideally be zero and a large value shows that the index compilation is introducing large spurious *UT* and/or *F* variations into the data. We also study index performance by comparisons with the *SME* and *SML* indices, compiled from a very large number of stations, and with an optimum solar wind "coupling function", derived from simultaneous interplanetary observations.
Results: It is shown that a station's response patterns depend strongly on the level of geomagnetic activity because at low activity levels the effect of solar zenith angle on ionospheric conductivity dominates over the effect of station proximity to the midnight-sector auroral oval, whereas the converse applies at high activity levels. The metric V_{av} for the two-station *aa* index is modelled to be 8.95%, whereas for the multi-station *am* index it is 0.65%. The *ap* (and hence *Kp*) index cannot be analyzed directly this way because its construction employs tabular conversions, but the very low V_{av} for *am* allows us to use $\langle ap\rangle/\langle am\rangle$ to evaluate the *UT*-*F* response patterns for *ap*. This yields $V_{av} = 11.20\%$ for *ap*. The same empirical test applied to the classical *aa* index and the new "homogenous" *aa* index, aa_H (derived from *aa* using the station sensitivity model), yields V_{av} of, respectively, 10.62% (i.e., slightly higher than the modelled

*Corresponding author: m.lockwood@reading.ac.uk

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

value) and 5.54%. The *ap* index value of V_{av} is shown to be high because it exaggerates the average semi-annual variation and has an annual variation giving a lower average response in northern hemisphere winter. It also contains a strong artefact *UT* variation. We derive an algorithm for correcting for this uneven response which gives a corrected *ap* value, ap_C, for which V_{av} is reduced to 1.78%. The unevenness of the *ap* response arises from the dominance of European stations in the network used and the fact that all data are referred to a European station (Niemegk). However, in other contexts, this is a strength of *ap*, because averaging similar data gives increased sensitivity and more accurate values on annual timescales, for which the *UT*-*F* response pattern is averaged out.

Keywords: indices / geomagnetism / substorm / space environment / space climate

1 Introduction

Geomagnetic indices are widely used to quantify the level of activity in the terrestrial magnetosphere-ionosphere-thermosphere system, and have sometimes been used without consideration being given to their response characteristics and whether or not they are appropriate to the application in question. In this paper, we study the response characteristics of three widely-used 3-hourly global geomagnetic indices compiled from observations by mid-latitude observatories; indices which primarily respond to the substorm current wedge (see review by Lockwood, 2013). These are the "mondial" (meaning "global") *am* index (and its components *an* and *as* in the northern and southern hemispheres, respectively), the "planetary" *ap* index (equivalent to *Kp*) and the *aa* index (and its hemispheric components, aa_N and aa_S). For each of these indices, we study the response as a function of the time-of-day (i.e., Universal Time, *UT*), the time of year (quantified by the fraction of the calendar year, *F*) and the level of geomagnetic activity.

Maps showing the stations currently contributing to these indices are presented in Figure 1. Lists of stations, their locations and the intervals over which each was used, are given for each index in Appendix A. The *am* index is designed to give good and even longitudinal coverage in both hemispheres, currently employing 14 stations in the northern hemisphere and 10 in the southern. There have been relatively minor changes to the network of stations deployed and the use of area-weighted averaging over longitude sectors has minimised the effect of these changes. It is available continuously from 1 January 1959 to the present day. The *ap* (and *Kp*) index is currently made from data from 11 stations in the northern hemisphere and 2 in the southern. It is available continuously since 1 January 1932, although the number and distribution of stations has varied considerably: initially there were 10 stations all in the northern hemisphere, the first southern hemisphere station being added in 1958. The *ap* and *Kp* indices have always been dominated by data from European stations. The *aa* index uses just two stations, one in each hemisphere, and although suppressing the resulting spurious annual variation in the response rather well, *aa* shows a large spurious diurnal variation (e.g., Lockwood et al., 2018b). The *aa* index is continuously available for the longest interval (1 January 1868 to the present day) and has been relatively homogeneous in its construction. It is this longevity that gives *aa* its importance. Recent work has shown that it is greatly improved by corrections to allow for the effects of secular change in the intrinsic geomagnetic field and the location-dependent sensitivity of the stations deployed, yielding the "homogeneous *aa* index" aa_H (Lockwood et al., 2018a, b). Note that above we discuss the 3-hourly indices *am*, *an*, *as*, *ap*, *aa*, aa_N, aa_S and aa_H: the same considerations apply to their respective daily means, *Am*, *An*, *As*, *Ap*, *Aa*, Aa_N, Aa_S and Aa_H, and their 8-point (24-hour) running (boxcar) means *Am**, *An**, *As**, *Ap**, *Aa**, Aa_N*, Aa_S* and Aa_H* (Allen, 1982).

Figure 2a shows an example of a 7-day interval of 3-hourly geomagnetic index data and Figure 2b shows the corresponding 24-hour (8-point) running means for the same interval. The interval is 27 October 2003 to 2 November and so includes the "Halloween storms" that generated considerable GIC events (Kappenman, 2005). This example is chosen here because it gave the highest daily means in the interval 1995–2017. The values given for the *am*, *ap*, *aa* and aa_H indices in Figure 2a (and their corresponding 8-point means, *Am**, *Ap**, *Aa**, and Aa_H* in Fig. 2b) are the published values but the *ap* and *Ap** values have been multiplied a factor $f = \langle am \rangle_{all} / \langle ap \rangle_{all}$, the ratio of overall means of *am* and *ap*, which allows for the difference in the scaling of *ap* and the other indices. Figure 2b shows the most clear-cut difference between the indices is that the scaled *ap* (and hence *Ap**) values during the storm are consistently larger than the corresponding values for *am*, *aa*, and aa_H whereas before and after they are proportionally much more similar. This demonstrates that differences between the responses of the indices can depend on the level of geomagnetic activity. Given the normalisation using *f*, this means there must be other times when $f \times ap$ is lower than *am*, *aa* and aa_H. This paper investigates how systematic this difference is with time-of year. In comparison, the *Aa**, Aa_H* and *Am** values are much more similar. The 3-hourly values shown in Figure 2a show considerable point-to-point variation in each index and point-to-point variation in the relationships between the various indices. There is no evidence for a systematic difference between the indices with *UT* (by which points in Fig. 2a are colour-coded to aid comparisons). Such differences would be convolved with random effects, such as the timing of perturbations within the 3-hour intervals over which the range indices are measured, and so are likely to emerge only in statistical surveys that also allow for the effects of time-of year and activity level.

1.1 *K* indices

K values, on which all these indices are based, were introduced by Bartels et al. (1939). They are derived for each magnetometer station from the range of variation observed in each of eight 3-hour intervals (0–3 *UT*, 3–6 *UT*, up to 21–24 *UT*). Originally scaled manually, *K*-value derivation was increasingly automated as magnetometer data recording moved from analogue to digital (Riddick & Stuart, 1984). The range of the irregular variations (i.e., after subtraction of the regular daily variation) in either of the horizontal components (*X* northward

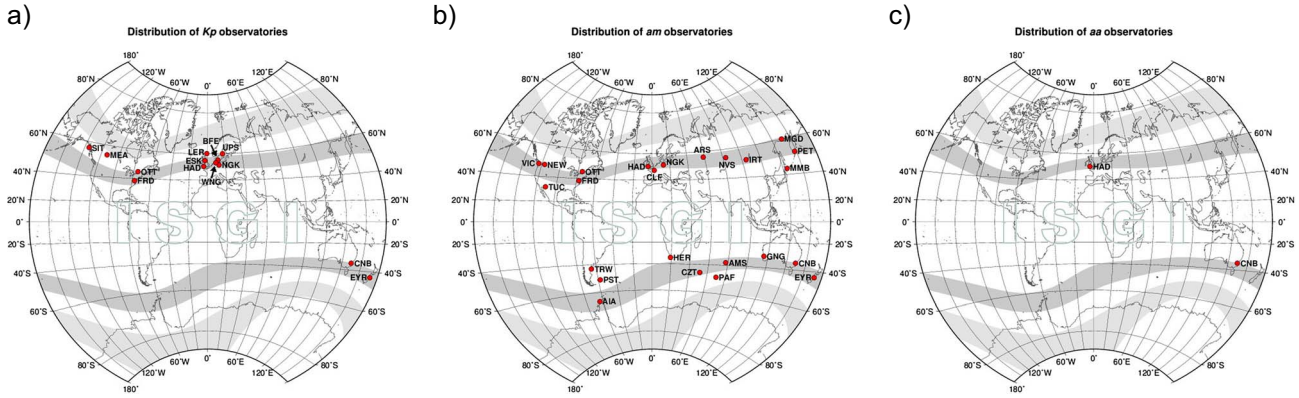

Fig. 1. Maps of networks of stations currently contributing to (a) the *Kp* (and hence *ap*) index, (b) the *am* index and (c) the *aa* index. In each map, the light grey bands are typical locations of the auroral oval and dark grey bands are ideal middle geomagnetic latitudes for stations to give a *K*-index value, being close enough to give a large signal, but far enough away that the response is monotonic because, for all but the very largest disturbances, the auroral oval approaches the station as the activity level increases. Details of these stations, and others used in the past, are given in Appendix A. Images courtesy of the International Service of Geomagnetic Indices (ISGI).

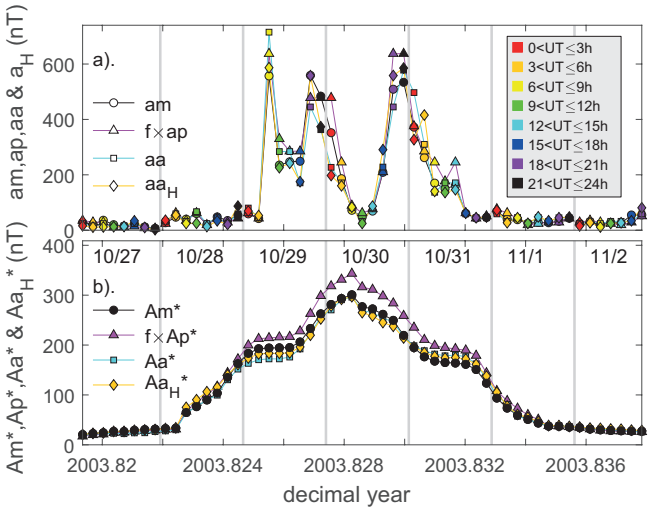

Fig. 2. Variations in geomagnetic range indices for 27 October 2003 to 2 November, showing the "Halloween storms": (a) 3-hourly values am, $f \times ap$, aa and aa_H (b) their 24-hour (8-point) running ("boxcar") means Am^*, $f \times Ap^*$, Aa^* and Aa_H^*. The ap and Ap^* values have been multiplied by $f = \langle am \rangle_{all} / \langle ap \rangle_{all}$, the ratio of overall means of am and ap for 1995–2017, to allow for the difference between the scaling of ap and that for other indices. Circles, triangles, squares and diamonds are for am (Am^*), $f \times ap$ ($f \times Ap^*$), aa (Aa^*), and aa_H (Aa_H^*), respectively. Points are colour-coded by the UT of observation in (a) and vertical grey lines are at $UT = 0$.

or Y eastward, whichever gives the larger value), ΔH_X or ΔH_Y, is ranked into one of 10 classes using quasi-logarithmic band limits that are specific to the observatory and to which a K value of 0–9 is assigned. The original idea of this procedure was that the scale of threshold values used to convert the continuous ΔH_X or ΔH_Y range values into the quantized K values is adjusted for each station to allow for its location and characteristics such that the K value is a standardized measure of the geomagnetic activity level, irrespective of where it is measured. In practice, the range limits for all K bands for a given station are all set by just one number, L, the lower limit of the $K = 9$ band, the lower limit for the $K = 0$ band being always set to zero (Menvielle & Berthelier, 1991). This is because the same relative scale is used at all stations with the thresholds for the K bands given in Table 1.

The very high correlations between the range indices and auroral electrojet indices such as AE and AL (e.g., Adebesin, 2016; Lockwood et al., 2019a) indicate that geomagnetic activity at mid-latitude observatories is dominated by the ionospheric segment of the substorm current wedge, i.e., the main westward auroral electrojet (e.g., Saba et al., 1997; Finch et al., 2008; Lockwood, 2013). (See further discussion in Sect. 1.5). As a result, the value of L used for a station is set by its closest proximity to the midnight Magnetic Local Time (MLT) sector of a nominal auroral oval which is where the range response of a station is greatest (Clauer & McPherron, 1974; Finch et al. 2008; Chambodut et al., 2013). The L value used is decreased with decreasing magnetic latitude because the range observed decreases with increasing distance from the auroral oval (e.g., Rostoker, 1972). In practice, this is quantified by the geocentric angle, δ, between the station and the point of closest approach of the nominal auroral oval (which occurs near midnight MLT) which is taken to be along a corrected geomagnetic latitude Λ_{CG} of 69°. Station K-indices are converted to a_K values (in nT) using a standard scale called "mid-class amplitudes", for which the range threshold for the $K = 9$ band is $L = 500$ nT: the conversion table for implementing this scale is referred to as K2aK and is given in the third column of Table 1. All the planetary and hemispheric indices discussed in this paper are based on these observatory K-values.

1.2 The *aa* index

The simplest of the indices that we study on this paper is the *aa* index, which was devised by Mayaud (1971, 1972, 1980) to give a continuous, well-calibrated and homogeneous record of geomagnetic activity that extends back to 1868. This index uses just two stations at similar geomagnetic latitudes, one in each hemisphere. The northern and southern hemisphere *aa* indices, aa_N and aa_S, are the a_K values from the station in that

Table 1. Bands of range values used to generate quantized *K*-indices for a station with a lower limit of the $K = 9$ band of L. $\Delta H_{X \text{ or } Y}$ is the range between extreme values in the 3-hour intervals of the northward or westward horizontal component, whichever is the larger. The right hand column gives the quantized a_K values ascribed to the *K*-levels using the "K2aK" or "mid-class amplitudes" scale.

K value	Range band	a_K (nT)
0	$0 \leq \Delta H_{X \text{ or } Y} < 0.01L$	2.3
1	$0.01L \leq \Delta H_{X \text{ or } Y} < 0.02L$	7.3
2	$0.02L \leq \Delta H_{X \text{ or } Y} < 0.04L$	15
3	$0.04L \leq \Delta H_{X \text{ or } Y} < 0.08L$	30
4	$0.08L \leq \Delta H_{X \text{ or } Y} < 0.14L$	55
5	$0.14L \leq \Delta H_{X \text{ or } Y} < 0.24L$	95
6	$0.24L \leq \Delta H_{X \text{ or } Y} < 0.40L$	160
7	$0.40L \leq \Delta H_{X \text{ or } Y} < 0.66L$	265
8	$0.66L \leq \Delta H_{X \text{ or } Y} < L$	415
9	$\Delta H_{X \text{ or } Y} \geq L$	667

hemisphere multiplied by a station scaling factor. For the "classic" *aa* index [the official *aa* index generated by École et Observatoire des Sciences de la Terre (EOST), and available from International Service of Geomagnetic Indices (ISGI), at http://isgi.unistra.fr/ and data centres around the world] the station scaling factors are constant with time; however, recently Lockwood et al. (2018a) have shown that anomalies in the secular variations of aa_N and aa_S are removed if time-dependent factors, calculated from models of the global intrinsic ("main") field, are used. The hemispheric *aa* indices are then averaged together to give $aa = (aa_N + aa_S)/2$. This averaging, to a large extent, gives cancellation of the seasonal variation in the geomagnetic response to solar forcing that is found in aa_N and aa_S individually. Lockwood et al. (2018b) have demonstrated how good cancellation is effectively achieved for the annual variation but that, because the best available stations are roughly 10 h apart in local time (instead of the ideal 12 h), the diurnal variations at the two stations do not cancel as completely as the annual variation. Recently, Lockwood et al. (2018b) have used a model of the stations' response (and made the correction to allow for the effects of the secular change in the geomagnetic field) to generate the "homogeneous" version of the *aa* index, aa_H, so named because it largely eliminates well-known hemispheric asymmetries between the mean values and distributions of aa_N and aa_S.

1.3 The *am* index

The stations used to compile the *am* index (Mayaud, 1980) are situated at sub-auroral latitudes close to corrected geomagnetic latitude $\Lambda_{CG} = 50°$. They are grouped into longitude sectors, with five such groups in the Northern hemisphere, and four in the Southern. The *K* indices for stations in a longitude sector are averaged together and the result is converted into a sector a_K value using the standard K2aK scale. Weighted averages of these sector a_K values are then generated in each hemisphere giving *an* and *as*, the weighting factors accounting for the differences in the longitude extents of the sectors. The index *am* is equal to $(an + as)/2$. Note that, like *aa*, *am* is compiled using only mathematical operations. We here employ all available *am* data up to the end of 2017 and that after the end of 2014 these data are classed as "provisional" which means they have passed initial quality checks and can be used, but not yet been through the final review that defines them as "definitive". We here apply additional checks to the data for 2015–2017 by testing for, and removing, any outliers in the scatter plots (more than 3σ from the mean) with the SuperMAG *SME* index or the Auroral Electrojet *AE* index (see Sect. 1.5).

1.4 The *ap (Kp)* index

The *ap* index is currently made using *K* index data from 11 northern and two southern hemisphere stations between corrected geomagnetic latitudes $|\Lambda_{CG}|$ between 44° and 60°. The *K* indices are first converted into standardised K_S values to account for the time-of-year and *UT* response characteristics of the observing site and to, as far as possible, normalise them to the values seen simultaneously by Niemegk, which was chosen as the reference station. The *K*p index is the arithmetic mean of the 3-hour standardized K_S values for the observatories employed. The 3-hourly *K*p values are converted into *ap* values using a standard table that is constructed such that *ap* may be regarded as the range of the most-disturbed of the two horizontal field components, expressed in units of 2 nT, at a station at dipole latitude of 50°.

The standardization from *K* to K_S is achieved using conversion tables for each observatory that were defined for the original stations by Bartels (1949, 1957). These give a multiplication factor K_S/K that depends on the station location, UT, time-of-year, *F*, and the *K* value and hence application of these factors is a non-linear operation. The present conversion tables used are for three seasons and many were generated using selected data from 1943 to 1948 only. The three seasons are: (1) the months around winter solstice (January, February, November and December); (2) the months around the equinoxes (March, April, September and October); and (3) the months around summer solstice (May, June, July and August). The network of stations used to compile the *ap* and *Kp* indices has varied, and the intervals over which each station was used are given in Appendix A.

In theory, if the *K* to K_S conversions were always ideal, the distribution of *ap* stations would be of no consequence as the various K_S values would be all be different measurements of the same thing. However, to some extent the tables will fall short of the ideal so the index response pattern will also depend, to some extent, on the distribution of stations. Even if the *K* to K_S conversions were ideal, converting the data to what Niemegk would have seen causes the index to take on the *UT-F* response characteristics of the Niemegk site. Note that we here quote *ap* (and hence ap_C, Ap_C* and $[Ap_C*]_{max}$) in the most widely used form – namely as an index without units: the standard *ap* values are an index in units of 2 nT and hence the values in units of nT would be double those given in this paper (Menvielle & Berthelier, 1991).

1.5 Relationship of mid-latitude *K*-based indices to auroral electrojet indices

In this section, we compare the mid-latitude range indices (*am*, *ap* and *aa*) with the *SME* and *SML* indices from the northern hemisphere SuperMAG magnetometers (Newell & Gjerloev, 2011) and the standard *AE* and *AL* auroral electrojet indices from the ring of 12 northern hemisphere stations at

auroral latitudes (Davis & Sugiura, 1966). The *AE* indices are based on 1-minute data on the quiet-day-subtracted horizontal magnetic component from the 12 stations. The upper and lower envelopes of the distribution of values at any one time define *AU* and *AL*, respectively, and *AE* is the width of the envelope between these maximum and minimum values, $AE = AU - AL$. The *AL* values are negative and driven by the westward auroral electrojet that is the ionospheric segment of the substorm current wedge and so are usually recorded by a station in the midnight MLT sector and *AL* responds mainly to the enhancement of the electrojet during substorm expansion phases (see review by Lockwood, 2013). On the other hand, *AU* responds primarily to the eastward electrojet associated with westward convective flow in the afternoon MLT sector and is enhanced during both the growth and expansion phases of substorms. The SuperMAG *SME* and *SML* indices are computed in exactly the same way as *AE* and *AL* but using 1-minute data from a much greater number of stations, which varies between 93 and 118 over the interval studied in this section (1996–2017). These indices capture better the extreme deflections which define them. The SuperMAG network is global but has an excess of northern hemisphere stations that are particularly clustered in the European and American longitude sectors and *SME* and *SML* are generated from northern hemisphere stations only. However, because *SME* and *SML*, like *AE* and *AL*, use extreme rather than average values, the clustering is not an issue and the network is extensive enough for these indices to be considered almost free of spurious diurnal variations. However, the use of only northern hemisphere stations will mean these indices, like *AE* and *AL*, will have an annual variation due to seasonal effects. The *am*, *ap* and *aa* indices are all based on the range of variation in 3-hour intervals, so it makes sense to compare them with the largest *SME* value (SME_{max}) and the minimum *SML* value (SML_{min}) in the same 3-hour intervals. However, as shown by Table 2, strong correlations are also found with the average *SME* and *SML* for the coincident 3-hour intervals ($\langle SME\rangle_{\tau=3h}$ and $\langle SML\rangle_{\tau=3h}$), although not quite as strong as for SME_{max} and SML_{min}. Correlation coefficients are typically between 0.8 and 0.9 and always stronger for the 24-hour smoothed values than the 3-hourly values and for the SuperMAG indices than for the traditional *AE* indices. In most cases, the *am* index yields the best correlations, but those for *aa* are very similar: correlations for *ap* are invariably lower. Note that the correlations for *AL* are slightly weaker than those for *AE*, as are those for *SML* compared to *SME*. This indicates that the mid-latitude range indices do respond to the directly driven system, as detected by *AU* and *SMU*, as well as the strong influence of the substorm unloading system, as detected by *AL* and *SML*.

Figure 3 illustrates examples of the relationships of (a) *am*, (b) *ap*, and (c) *aa* to SME_{max} for the years 1996–2017 (inclusive). The general features are the same in corresponding plots for $\langle SME\rangle_{\tau=3h}$, $-SML_{min}$, $-\langle SML\rangle_{\tau=3h}$, AE_{max}, $\langle AE\rangle_{\tau=3h}$, $-AL_{min}$, or $-\langle AL\rangle_{\tau=3h}$. In each panel a scatter plot of the three-hourly values is given by grey dots and of their 8-point running means (24-hour averages) by orange dots. Note the grey dots show the quantization of both *aa* and *ap* levels, whereas *am* (like *Am**, *Ap** and *Aa**) is essentially continuous. Black dots are means in 1-percentile bins of the mid-latitude range index along the *x*-axis (giving 614 samples in each bin) and the horizontal and vertical error bars are plus and minus one standard deviation for that bin. The mauve lines are fourth-order polynomial fits to the 3-hourly values. The RMS deviation of the observed 3-hourly SME_{max} values from the fitted polynomial values, Δ_{RMS} are given in each panel, as are the linear correlation coefficients *r* and *r** for the 3-hourly values and the 8-point running means, respectively. Note that the polynomial fit order $m = 4$ that was used was the largest that could be employed (desirable as it preserves the non-linearity of the variation) that gave a good fit to the tail of the index distribution; the latter criterion being tested by evaluating the $\Delta_{RMS}^2/(n - m - 1)$ value for the $n = 614$ samples in the top percentile of the index dataset and checking it was not significantly larger (at the $1 - \sigma$ level) than the minimum value for any *m*. The cyan lines are linear fits to the 3-hourly values for quiet-times ($SME_{max} < 750$ nT), and are plotted to gauge the deviations from linearity of the data for larger SME_{max}. Figure 3 shows that all three mid-latitude range indices have a non-linear response, with values being overestimates, relative to SME_{max} values, at the highest activity levels. However, the tendency is much stronger for *ap* than for *am* and *aa*: Figure 3a and c shows that the deviation from linearity is only present in the top 1% of samples for *am* and *aa*; however, for *ap* the deviation is persistent for the top 35% and is significant (greater than 1 standard deviation) for the top 17% of *ap* samples. The scatter (quantified by the RMS deviation of samples from the fourth-order polynomial fit, Δ_{RMS}) is smallest for *am* but largest for *aa*, but the linear correlations *r* and *r** for *aa* are higher than for *ap* because the polynomial fit deviates from linearity much less for *aa* than for *ap*. Indeed, for 24-hour running means, the correlation for *Aa** is even very slightly higher than that for *Am**.

Table 2. Linear correlation coefficients between mid-latitude range indices and the SuperMAG and auroral electrojet indices: *r* and *r** are for 3-hourly values and the 8-point running means, respectively. The data used are for 1996–2017 (inclusive). This yields 61368 3-hourly samples (*am*, *ap* and *aa*) and 6136124 running-mean samples (*Am**, *Ap** and *Aa**). For all correlations, the large number of samples ensures that the correlation significance level, derived by comparison with the AR1 red noise model, is 100% to within at least three decimal places for all cases. The maximum *SME* and minimum *SML* in each 3-hour intervals is SME_{max} and SML_{min} respectively.

	Interval (h)		*am*	*ap*	*aa*
SME_{max}	3	*r*	0.840	0.806	0.832
$\langle SME\rangle_{\tau=3h}$	3	*r*	0.784	0.741	0.772
AE_{max}	3	*r*	0.824	0.767	0.815
$\langle AE\rangle_{\tau=3h}$	3	*r*	0.784	0.741	0.772
SML_{min}	3	*r*	−0.823	−0.777	−0.807
$\langle SML\rangle_{\tau=3h}$	3	*r*	−0.792	−0.748	−0.779
AL_{min}	3	*r*	−0.822	−0.766	−0.810
$\langle AL\rangle_{\tau=3h}$	3	*r*	−0.792	−0.748	−0.779
			*Am**	*Ap**	*Aa**
SME_{max}*	24	*r**	0.908	0.881	0.913
$\langle SME\rangle_{\tau=24h}$	24	*r**	0.866	0.826	0.864
AE_{max}*	24	*r**	0.891	0.853	0.901
$\langle AE\rangle_{\tau=24h}$	24	*r**	0.856	0.826	0.864
SML_{min}*	24	*r**	−0.890	−0.867	−0.867
$\langle SML\rangle_{\tau=24h}$	24	*r**	−0.848	−0.847	−0.848
AL_{min}*	24	*r**	−0.900	−0.865	−0.865
$\langle AL\rangle_{\tau=24h}$	24	*r**	−0.875	−0.847	−0.848

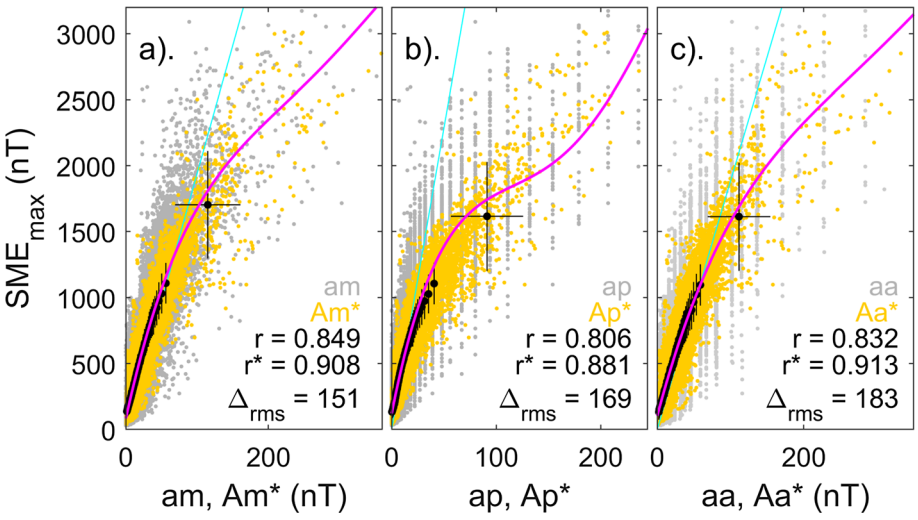

Fig. 3. Scatter plots for 1996–2017 (inclusive) of the mid-latitude range indices with the maximum of the *SME* index, SME_{max}, seen in the same 3-hour intervals by the SuperMAG global magnetometer network. (a) (grey) 3-hourly SME_{max} values as a function of 3-hourly *am* and (orange) 24-hour running means of SME_{max} as a function of corresponding running means of *am*, *Am**. Black dots are means in 1-percentile ranges of *Am** (giving 614 samples in each bin) and the horizontal and vertical error bars are ±1 standard deviation. The mauve line is a fourth-order polynomial fit to the 3-hourly values. The RMS deviation of the observed 3-hourly SME_{max} values from the fitted polynomial value for the corresponding *am*, Δ_{RMS} is given, as is the correlation coefficient *r* between 3-hourly SME_{max} and *am* values. (b) The same as (a) for 3-hourly *ap* and its 24-hour running mean, *Ap**. (c) The same as (a) for 3-hourly *aa* and its 24-hour running mean, *Aa**. In each panel the cyan line is a linear fit to the 3-hourly values for $SME_{max} < 750$ nT, and is plotted to gauge the deviation from linearity of the data for larger SME_{max}.

1.6 Relationship of mid-latitude *K*-based indices to interplanetary conditions

In Section 2 we employ the concept of the sensitivity of station to "solar forcing". This sensitivity depends only on the station's location, all other site factors (such as instrumentation sensitivity and ground conductivity) being accounted for by the instrument calibration. This sensitivity gives the station's response to all solar forcings, including photon and particle ionization and heating effects (Aksnes et al., 2002; Ieda et al., 2014) as well as the enhanced electric fields and the associated expansion/contraction of the auroral oval due to solar wind-magnetosphere coupling (Lockwood et al., 1990; Cowley & Lockwood, 1992; Milan et al., 2012). With this definition, not all of this solar forcing comes from the solar wind, because of the effects of solar EUV and X-ray ionizing and heating radiations on ionospheric conductivities. Hence "solar wind forcing" is a part of, but not all of, "solar forcing".

To quantify solar wind forcing, a number of "coupling functions" have been proposed as predictors and explainers of geomagnetic disturbance (see review in Lockwood et al., 2019a). These are combinations of parameters characterizing the near-Earth planetary environment, combined with various coefficients and exponents that are free fit parameters that are derived empirically to get the best fit to the observed geomagnetic activity response. The recent work by Lockwood et al. (2019a) highlights a serious problem with most previous coupling function studies that have generally neglected the effect of gaps in the interplanetary data series on the grounds that they occur at random and so their effects will average out. Lockwood et al. use Monte-Carlo analysis by inserting synthetic data gaps at random into near continuous data to show that this is far from being a valid assumption and that the effect of data gaps is to add considerable noise into solar wind – geomagnetic activity correlation studies. This is true irrespective of the method used to deal with the data gaps (for example interpolation, piecewise removal of geomagnetic data, or simply ignoring their effects). This raises the potential for "overfitting" which is a serious problem in multiple regression analysis of geophysical time series that have internal noise: it is a recognized pitfall in areas where quasi-chaotic behaviors give large internal noise such as climate science (e.g., Knutti et al., 2006) and population growth (e.g., Knape & de Valpine, 2011) but had not been considered in solar-wind/magnetosphere coupling studies. Overfitting occurs if a fit has too many degrees of freedom which allows it to fit to the noise in the training subset, and hence is not robust in general. Including all of the factors with their own weighting factors and/or exponents can result in extremely good fits that can, nevertheless, give details that are statistically meaningless as each additional fit parameter reduces the statistical significance of the correlation. Such fits can have limited, and in extreme cases, zero predictive capability because they have fitted noise rather than the signal. The addition of noise by neglecting the effect of gaps in interplanetary data (which before 1995 were common and often of long duration) means that overfitting is a serious problem. Another complication is that the relative performance of different coupling functions depends strongly on the data averaging timescale and the averaging timescale used to generate the best fit exponents (Finch & Lockwood, 2007).

As a result of these considerations, we here use just one coupling function, P_α, that has just a single free fit parameter, the coupling exponent, α, which estimates the power input into the magnetosphere, using the theory of Vasyliunas et al. (1982). This theory is based on the fact that the dominant energy flux in the solar wind is the bulk flow kinetic energy flux of the

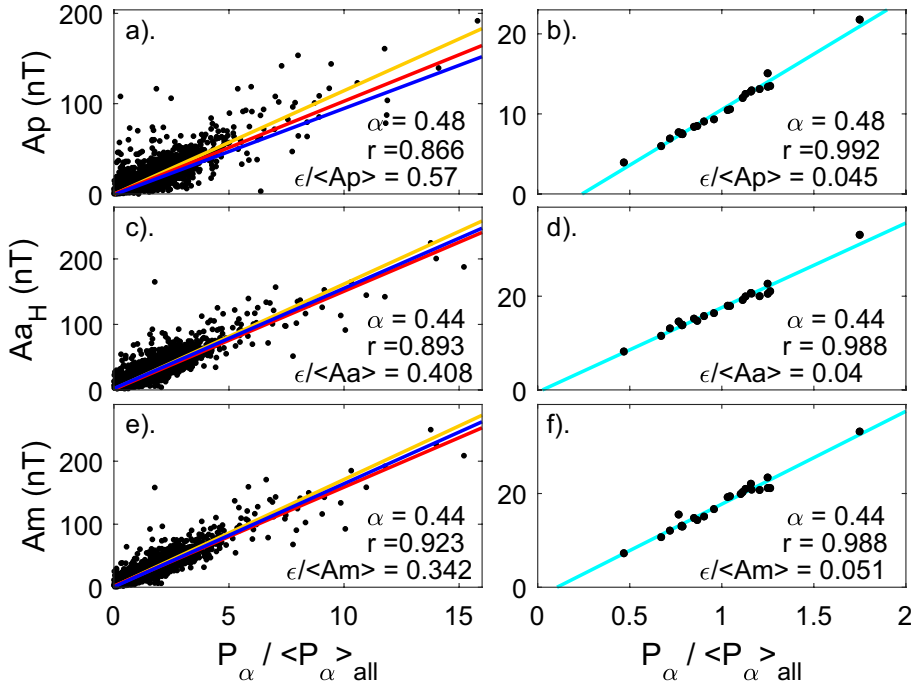

Fig. 4. Scatter plots of geomagnetic indices as a function of normalized power input into the magnetosphere computed from near-Earth solar wind observation, $P_\alpha/\langle P_\alpha\rangle_{all}$, where the average is over the full period considered (1995–2017, inclusive). The left hand panels are for daily means, the right hand panels for annual means. The top panels are for the *ap* index, the middle for *aa* index, the bottom for *am* index. For the daily data, linear regression fits are shown for: (red line) 91 days around the June solstice; (blue line) 91 days around the December solstice; and (orange line) 91 days around either equinox). For annual means the cyan lines are linear regression fits for all data. The number of valid daily P_α data points is $N = 8375$ (an availability of 99.7%) and for annual means is $N = 23$. The best-fit coupling exponent used to generate P_α is $\alpha = 0.44$ for *am* and *aa* and $\alpha = 0.48$ for *ap*. The linear correlation coefficients, *r*, and the Root Mean Square (RMS) linear fit residual ε (as a ratio of the overall mean value of the index) are given in each panel.

particles (and not the Poynting flux assumed by the much-used but flawed epsilon parameter). The fraction of this energy flux that is converted into Poynting flux by currents flowing in the bow shock, magnetosheath and tail magnetopause is taken to be the (necessarily dimensionless) factor $M_A^{-2\alpha}$, where M_A is the Alfvén Mach number of the solar wind and α is an unknown coupling exponent. The value of α could vary from zero (which would mean all incident power could enter the magnetosphere if the orientation of the interplanetary magnetic field, IMF, were favorable) and a large positive value (which would mean a vanishingly small fraction could enter the magnetosphere). Typically, values of α between 0.35 and 0.5 have been derived from the fits to data which, for a typical M_A of 10, means that 5–10% of solar wind power is available to enter the magnetosphere. The coupling function P_α contains terms in solar wind mean ion mass, number density and velocity and the IMF orientation, but the exponents for each are all self-consistent, all being set by the single fit parameter α and the theory. (Note that, in some studies the exponent of the IMF orientation factor has been treated as a separate fit variable, but we employ the procedure recommended by Vasyliunas et al. that allow it to be computed self-consistently from the data once the optimum α has been determined). Finch & Lockwood (2007) found that the optimum coupling exponent α used in generating P_α depended on averaging timescale, a result that implied that there was at least one other mechanism at work and that P_α was failing to capture all of the relevant physics. However, Lockwood et al. (2019a) have shown that this variation in α was an artefact caused by data gaps and that when steps are taken to minimize the effect of such data gaps, α is effectively constant on all timescales from 1 min up to 1 year.

Figure 4 shows scatter plots of *ap* (top), the new homogeneous *aa* index, aa_H (Lockwood et al., 2018a, b) (middle), and *am* (bottom) as a function of the power input into the magnetosphere, P_α (normalized by dividing by its average for the whole interval, $\langle P_\alpha\rangle_{all}$, in order to cancel some constants). The left hand panels are for daily means and the right hand panels for annual means. We use daily means because at shorter timescales the lag introduced by substorm growth phases becomes a significant factor (Lockwood et al., 2019a). For *aa*, aa_H and *am* the best fit coupling exponent (at all timescales) is $\alpha = 0.44$, whereas for *ap* it is $\alpha = 0.48$ (Lockwood et al., 2019a). The correlation coefficients for the daily means *Ap*, Aa_H and *Am* with daily means of $P_\alpha/\langle P_\alpha\rangle_{all}$ are very high, being 0.866, 0.893 and 0.923, respectively. The root-mean-square (RMS) fit residual for a linear regression of all the data, ε, was also computed, giving $\varepsilon/\langle Ap\rangle_{all} = 0.570$ for *Ap*, $\varepsilon/\langle Aa\rangle_{all} = 0.401$ for *Aa* and $\varepsilon/\langle Am\rangle_{all} = 0.342$ for *Am*. Hence both these metrics give best agreement for *am*, and worst agreement for *ap*. We tested the significance of the difference between the correlations using the Meng-*Z* test (which allows for intercorrelations between the datasets (Meng et al., 1992)) and found that the *p*-value for the null hypothesis that they are actually the same was undetectably small. Hence the agreement with the

coupling function is significantly better for *am* than for *aa* and that for *aa* is significantly better than that for *ap* on this 1-day averaging timescale. For the traditional *aa* index, the correlation was 0.883, which is slightly lower than for aa_H but the Meng-*Z* test gives that the *p*-value for the null hypothesis that these two correlations are the same is only 5×10^{-5}. Thus the improvement of aa_H over *aa* is small but statistically highly significant.

For the annual means shown in the right hand panels, the ranking order of the correlations is reversed. In this case, the correlations for *ap*, aa_H (and *aa*) and *am* are 0.992, 0.988 and 0.987. These correlations are exceptionally high and differences are small. The Meng-*Z* test gives *p*-values against the null hypothesis that they are the same of 0.314 for *ap* and aa_H and of 0.371 for *ap* and *am*. Hence the difference between the correlations for *ap* and aa_H is just significant at the $1 - \sigma$ level but it is not quite significant at the $1 - \sigma$ level for *ap* and *am*. The RMS residuals (as a ratio of the overall mean value) are lowest for *aa*, but higher for *am* than for *ap*. In conclusion, both metrics show that *am* outperforms *ap* in daily averages but is out-performed by *ap* in annual averages.

The linear regression lines shown indicate why this is the case. For annual means (right-hand panels) the cyan lines are linear regression fits to all data but for the daily-averaged data (left-hand panels) the data have been linearly regressed in three subsets: (1) 91 days around the June solstice (giving the red line); (2) 91 days around the December solstice (blue line); and (3) 91 days around either equinox (orange line). In Figure 4a the regression lines for the three seasons are different, with the *Ap* values for equinox at a given $P_\alpha/\langle P_\alpha \rangle_{all}$ being larger, whereas for northern hemisphere winter they are smaller. For *Aa* and *Am* these differences are much smaller. The annual and semianual variations in the response of *aa* and *am* are real and in a later publication we will show that they are highly significant features of magnetospheric behaviour, however they are both exaggerated in *ap*. On the other hand, when we average out these 6 and 12 month periodicities by taking annual means, *ap* performs better than the other two indices.

Figure 5 explains the seasonal variations of the correlations shown in the left hand panels of Figure 4 by showing the semi-annual variations in the mean values of the geomagnetic indices and their best fit of the corresponding variation in P_α. The power input P_α shows a clear semi-annual variation with peaks at the equinoxes, the September peak being somewhat more pronounced than the March one. There is a slight difference between the solstices with the June minimum being slightly deeper than the December one. This variation results almost completely from the $\sin^4(\theta/2)$ term in P_α (where θ is the IMF clock angle in the Geocentric Solar Magnetospheric, GSM, frame of reference) and is caused by the Russell-McPherron effect of the Earth's dipole tilt on IMF orientation in the GSM frame (Russell & McPherron, 1973). Figures 4 and 5 use daily means to avoid complications associated with the lag of the geomagnetic response due to the variable length of substorm growth phases. Lockwood et al. (2016) show that for averaging timescales greater than 12 h the *F-UT* response pattern due to the Rusell-McPherron effect becomes "axial" in form, i.e., showing the equinoctial peaks but no *UT* variation. Note that on shorter timescales the *F-UT* pattern of geomagnetic response becomes "equinoctial" in form (best demonstrated by the *am* index, but can also be detected in aa_H, Lockwood et al., 2018b) whereas

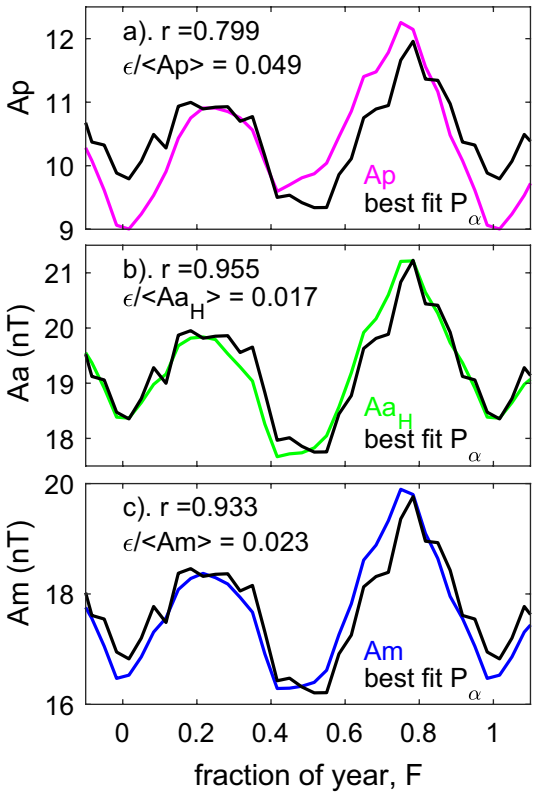

Fig. 5. The annual and semi-annual variations in geomagnetic indices and estimated power input into the magnetosphere, P_α, for coincident data from 1995 to 2017, inclusive. In each panel the coloured line shows mean values of daily means of the geomagnetic index in 30 equal-width bins of time-of-year, *F*, smoothed with a 3-point running mean. The black line is the best-fit variation of the near-continuous P_α data for the same interval processed the same way. (a) is for the *Ap* index; (b) is for the Aa_H index, and (c) is for the *Am* index. In each panel, two goodness of fit metrics are given: the correlation coefficient *r* and the Root Mean Square (RMS) fit residual, ε, as a ratio of the overall mean value of the index.

P_α shows the Russell-McPherron form (Lockwood et al., 2016) which confirms that the geomagnetic indices are not responding to the "directly-driven" system. This difference in the *F-UT* patterns is therefore associated with the "storage-release" system to which mid-latitude range indices primarily respond (Finch et al., 2008; Lockwood, 2013). Hence the origins of the equinoctial pattern must be associated with the variable lags cause by the durations of substorm growth phases.

Figure 5 shows that the index that most closely reflects the semi-annual variation in P_α is Aa_H, for which the fits are, surprisingly, slightly better than for *Am*. The fits are not as good for *Ap*, which exaggerates the September equinox peak and has a deep minimum in December. Hence Figure 5 explains the seasonal differences in the regression fits in Figure 4 for *Ap*, and why they are smaller for Aa_H nor *Am*.

2 Analysis

Finch (2008) employed the concept of the location-dependent magnetometer station sensitivity, *s*, defined simply

for any given type of single-station geomagnetic activity measure g by

$$s = \frac{g}{I_S} \tag{1}$$

where I_S is a measure of the input solar forcing (which includes the effects of both induced currents in near-Earth space driven by solar wind-magnetosphere coupling and of conductivity due to ionizing EUV and X-ray radiations from the Sun or particle precipitations). Finch defined s to be a function of only the instrument co-ordinates because instrument and local site characteristics are accounted for by other inter-calibration procedures. By taking ratios of g seen simultaneously at many pairs of different stations, the I_S factor is cancelled and the ratios of the station sensitivities is known. Note that this concept is the same as that introduced by Bartels (1949) and that is still used today in the compilation of the *ap* (and *kp*) index: this is because Bartels' look-up tables used to convert K data from a given station (XXX) into what Niemegk (NGK) would have seen (the K_S value) are tables of average empirical values of s_{NGK}/s_{XXX} (as a function of UT, F and activity level). If the data from different stations are combined into a geomagnetic index using linear mathematics, then the sensitivities are similarly combined. For example, if the g data from N stations are averaged together with weighting functions ω to give a planetary index G,

$$G = \sum_{i=1}^{N} \frac{\omega_i g_i}{N} = \sum_{i=1}^{N} \frac{\omega_i s_i I_S}{N} = I_S \sum_{i=1}^{N} \frac{\omega_i s_i}{N} = I_S S \tag{2}$$

where S is the sensitivity of the index which is the weighted mean of the station sensitivities, s_i. From comparisons of the ratios for many pairs of stations, Finch (2008) derived a functional form for computing the sensitivity of a station as a function of its geographic coordinates, date, time-of-year and time-of-day:

$$s = \frac{\{1 + A\ \cos^{0.7}(\chi)\}\left\{B\ \cos\left[(T - T^*)\left(\frac{2\pi}{24}\right)\right] + 1\right\}}{m} \tag{3}$$

where

$$T^* = 1.5 \sin\left(2\pi[F_1 - F]\right) - 0.5 \tag{4}$$

A and B are constants, χ is the solar zenith angle (a function of location, UT and F), T is the MLT of the station (in hours – also a function of location, UT and F and, on long timescales, of the intrinsic geomagnetic field), F is the fraction of the year and $F = F_1$ at the spring equinox (taken to be 100/365.25 for the northern hemisphere and 283/365.25 for the southern hemisphere). Lastly, m is a normalising factor that ensures that the average value of s, over all UT and all times-of-year (F) and activity levels, is unity for a given station and year: it is used to retain calibrations that allow for instrument characteristics and local site effects.

The Biot-Savart law states that the field disturbance at an observatory O, ΔB_o, is proportional to the integral over space of J_{np}/r^2 for all points P, where J_{np} is the current density at P normal to the line OP and r is the distance OP. The inverse-square dependence on r means that there can be a range of contributions to the observed signal at a mid-latitude station from large variations in J_{np} in the auroral oval (i.e., at larger r) to smaller fluctuations in J_{np} more local to the observatory (at smaller r). It is well known that substorm electric fields can penetrate the shielding provided by the region-2 field-aligned currents (that connect to the ring current) and so give "bays" in mid-latitude magnetometer records (e.g., Caan et al., 1978; Kikuchi et al., 2000). Hence the conductivities of the ionosphere over the observatory can influence the local J_{np} and hence ΔB_o. On the other hand, the large and variable currents flowing along the midnight-sector auroral oval during substorms will also have an effect that will depend strongly on how close this sector of the auroral oval is to the observatory. The first term on the right of equation (3) allows for the effect of solar zenith angle χ on the ionospheric conductivity over the station due to solar EUV and X-ray radiation and thus depends on the station's geographic coordinates, the UT and the time-of-year, F. If the Sun is below the horizon, χ is set to (π/2): hence the coefficient A controls the extent to which the effect of dayside conductivity at a given χ is enhanced over residual nightside values. Note that there are small changes to the precise formulation of Finch (2008), who used a $\cos^{0.5}(\chi)$ dependence, as predicted by Chapman production-layer theory and as also used in a great many prior applications. However, Ieda et al. (2014) have shown that a conductivity dependence on $\cos^{0.7}(\chi)$ fits better with observations and is also predicted by theory when the upward gradient of the neutral atmospheric scale height is accounted for. Using the conductivity over the observatory is an approximate parameterisation as there will be contributions to the total ΔB_o that arise from currents that are between the observatory and the auroral oval.

The second term on the right-hand-side of equation (3) is the station's sensitivity due to its distance from the location of peak response, which is at an MLT of T^* in the midnight sector. The sine term in equation (4) is used to model the known earlier onset of enhanced substorm activity in summer. Equation (4) yields T^* of 1 h MLT and 22 h MLT for the winter and summer solstices, respectively. This is based on the survey of mid-latitude station responses to substorm expansion phases by Finch (2008) and agrees well with the results of Liou et al. (2001), who found substorm onset was typically at T = 22 h in summer but 23.5 h in winter. Similar behaviour was deduced by Wang et al. (2007). We note that we are most interested in the MLT where auroral electrojet currents have peak effect on mid-latitude K indices: this is close to, but not the same as, the MLT of substorm onset (Clauer & McPherron, 1974; Chu et al., 2014).

In this paper, the solar zenith angle at a given station is computed as a function of time (year, fraction of year, F, and UT) using an ephemeris that gives the solar declination at that time. The MLT for a given UT is computed for that date using the IGRF-15 model of the geomagnetic field (Thébault et al., 2015).

Finch (2008) assumed that the factors A and B were constants and had considerable success in modelling the average response of different stations and indices. However, there are reasons to also think that the relative importance of the two terms in equation (3) might change systematically with the level of geomagnetic activity. Firstly, particle precipitation fluxes are higher during enhanced activity over a wide range of locations (including mid-latitudes (e.g., Shiokawa et al., 2005)), which could mean that photon-induced conductivity is less important,

and hence the dependence on $\cos^{0.7}(\chi)$ is weaker: as a result, the factor A would be reduced at higher activity levels. Secondly, the auroral oval expands equatorward when activity is enhanced, making the second factor in equation (3) (associated with the spatial proximity of the auroral electrojet) more important. The factor B sets the amplitude of the diurnal variation seen by the station because of the variation in its proximity to the peak of the substorm current wedge. For these reasons the factors A and B are here both treated as functions of geomagnetic activity level.

Lockwood et al. (2018b) quantified the factors A and B by assuming that the large number of stations used to derive the *am* index, and their even longitudinal spacing, results in the sensitivity of *am* index, S_{am}, being always unity, independent of both *UT* and time-of-year, F. We here use a more exact and iterative procedure, but get results which are very similar to those found by Lockwood et al. (2018b). In general

$$a_K/am = s_K/S_{am} \quad (5)$$

where the station of sensitivity s_K gives a K-index value that transforms to a_K using the standard K2aK scale. We here quantify the geomagnetic activity level using the *am* index and divide it into eight (generally overlapping) activity ranges: $0 \le am < 10$ nT, $10 \le am < 20$ nT, $20 \le am < 40$ nT, $30 \le am < 50$ nT, $40 \le am < 60$ nT, $50 \le am < 90$ nT, $60 \le am < 110$ nT, and $am \ge 70$ nT for which the years 1959–2017 give N_b = 58183, 51083, 40894, 22691, 13157, 8302, 10869, and 6060 samples, respectively, and the mean *am* values are 5.32, 13.96, 27.50, 37.71, 47.76, 63.56, 73.90, and 109.14 nT. The distribution of *am* values and these band limits are shown in Figure 6.

In the iterative procedure we adopt, we make the initial assumption of uniform S_{am} ($S_{am1}(UT, F) = 1$). This gives initial estimates of $s_K(UT, F)$ for each of the eight *am* ranges for the Canberra and Hartland *aa* stations studied (we here denote these initial values as s_{K1}). We then obtain initial A and B estimates (A_1 and B_1) for each of the *am* bins, using equations (2)–(4) by fitting modelled sensitivity ratios s_K/S_{am} to the a_K/am ratio values in the relevant *am* subset, using the fitting procedure in the *UT-F* parameter space described by Lockwood et al. (2018b). The modelled sensitivities were computed for 24 UT values (1 h apart) and 365 F values (daily) and their ratios then averaged into the same bins as for the observational data (namely, 3 h width in *UT* and 1/20 width in F). From these initial A_1 and B_1 values we can use equations (3) and (4) to compute the corresponding sensitivity value for each of the *am* stations, and then use equation (2) to re-compute the sensitivity of the network of stations used in compiling the index, S_{am}, giving a new estimate S_{am2}(UT, F). Using these values in equation (5) gives revised (first iteration) values of $s_K(UT, F)$, values (s_{K2}). This loop was repeated until the RMS deviation of the modelled s_K/S_{am} values for CNB and HAD from the observed a_K/am ratios converged on a constant, minimum value (to within an adopted uncertainty level of 0.001%). This iterative procedure yielded A values of 0.6116, 0.2727, 0.2083, 0.2010, 0.2001, 0.2001, 0.2003, and 0.2000 for the eight *am* bins (in order of increasing $\langle am \rangle$) and B values of 0.2890, 0.3293, 0.3631, 0.3786, 0.3711, 0.3163, 0.2800 and 0.2797. This iteration allows us to compute the *am* index sensitivity values, $S_{am}(UT, F)$, self-consistently rather than assuming it is constant at unity (the assumption that was employed by Lockwood et al., 2018b).

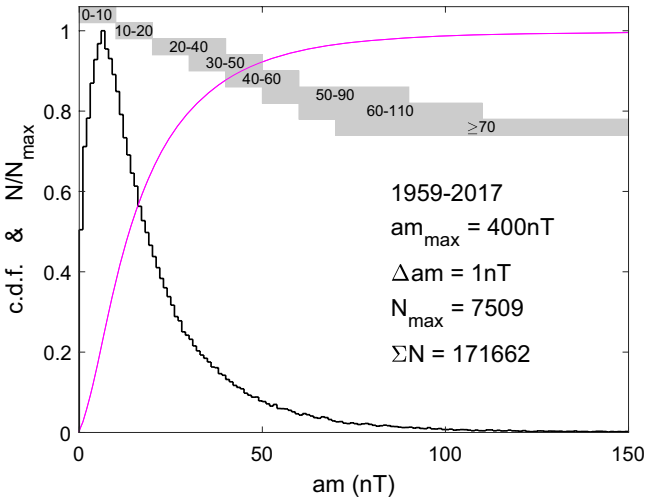

Fig. 6. Cumulative probability distribution (c.d.f, mauve line) and histogram of number of *am* samples in bins Δam = 1 nT wide (N, shown by the black line as N/N_{max}, where N_{max} is the maximum value of N) for all *am* data in the years 1959–2017 (inclusive). The grey bars give the eight overlapping *am* bands employed in this paper: $0 \le am < 10$ nT, $10 \le am < 20$ nT, $20 \le am < 40$ nT, $30 \le am < 50$ nT, $40 \le am < 60$ nT, $50 \le am < 90$ nT, $60 \le am < 110$ nT, and $am \ge 70$ nT which contain a numbers of samples N_b of 58183, 51083, 40894, 22691, 13157, 8302, 10869, and 6060, respectively, and for which the mean *am* values are 5.32, 13.96, 27.50, 37.71, 47.76, 63.56, 73.90, and 109.14 nT.

The right-hand columns of Figures 7 and 8 show the *UT-F* plots of the observed ratios $\langle a_K \rangle / \langle am \rangle$ for the two current *aa* stations, respectively Hartland and Canberra. The rows are for the eight *am* activity level bins shown in Figure 6, with the highest activity at the top and the lowest at the bottom. The left-hand column of both figures gives the *UT-F* plots of the modelled sensitivity of the *am* index, S_{am}, and the middle panels in both figures give the corresponding plots of the best-fit modelled sensitivity, s_K/S_{am}.

Figures 7 and 8 both show that the local ionospheric conductivity term is more important at low geomagnetic activity with a strong peak around the minimum of the solar zenith angle χ (at F = 0.5 and 14 *UT* for Hartland and F = 0 and F = 1 and 4 *UT* for Canberra). This is reflected in the large values A for the lowest activity levels. This peak becomes increasingly less pronounced with increasing *am* activity level (i.e., A decreases) and at the largest *am* the pattern is determined mainly by the distance of the station from the midnight auroral oval.

3 Response functions of geomagnetic indices

From the derived values of A and B for each *am* activity level bin, we can use equations (2)–(4) to compute the (*F-UT*) response pattern of any mid-latitude range indices that is compiled using a mathematical algorithm. Rather than show

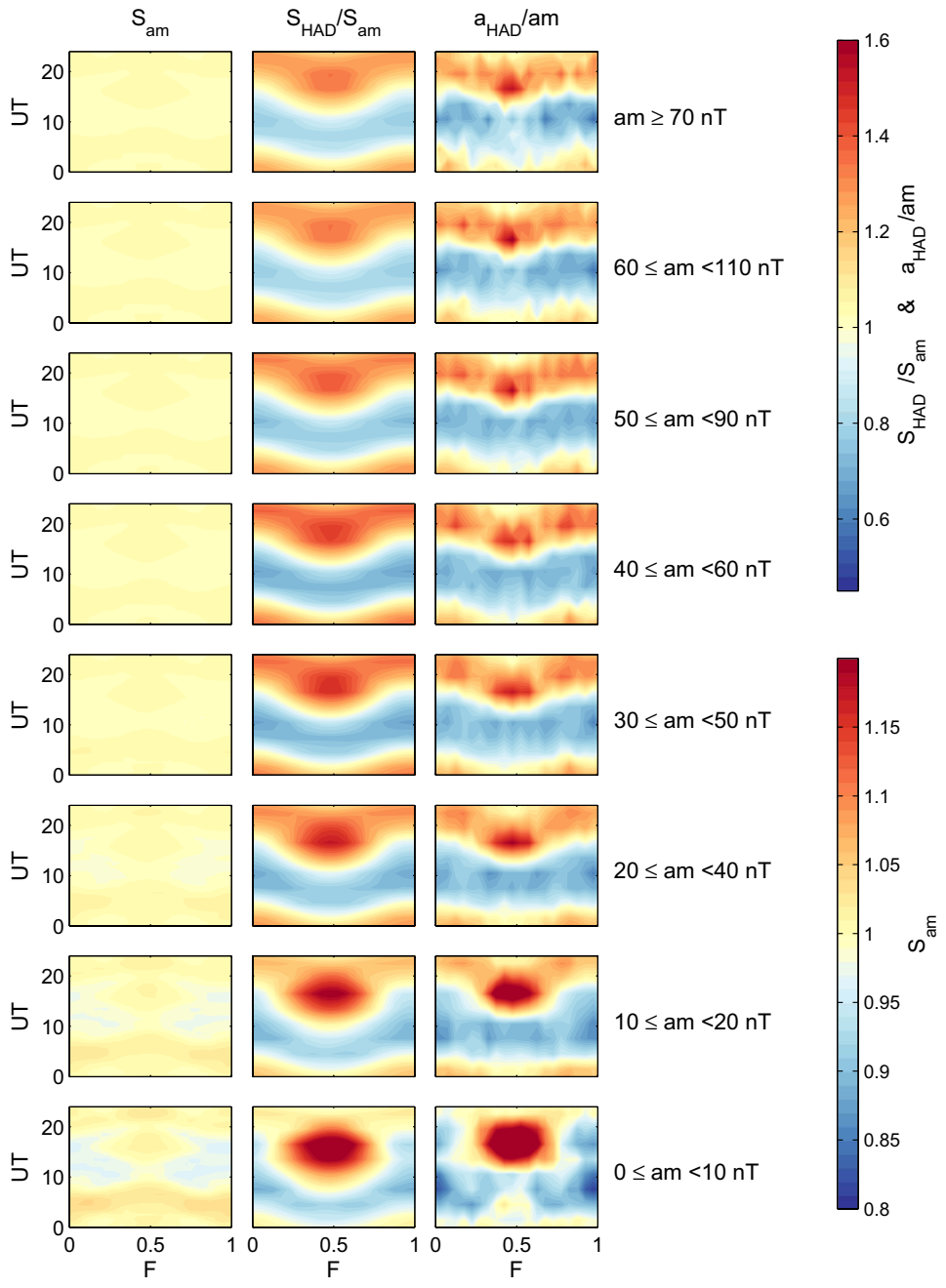

Fig. 7. Analysis of the sensitivity of the Hartland (HAD) station. Time-of-day (*UT*)/time-of-year (*F*) plots of: (left column) the modelled sensitivity for the *am* index, S_{am}, for the current stations and sector weighting functions; (middle column) modelled values of the ratio s_{HAD}/S_{am} where s_{HAD} is the sensitivity of the Hartland magnetometer station for measuring its a_K values, a_{HAD}; and (right column) means of the observed values of the ratio $\langle a_{HAD}\rangle/\langle am\rangle = s_{HAD}/S_{am}$. All data are for eight *UT* bins 3 h wide and 20 *F* bins 18.25 days wide over the years 1959–2017 (inclusive). The panels are for *am* ranges (from top to bottom) of: $am \geq 70$ nT; $60 \leq am < 110$ nT; $50 \leq am < 90$ nT; $40 \leq am < 60$ nT; $30 \leq am < 50$ nT; $20 \leq am < 40$ nT; $10 \leq am < 20$ nT; and $0 \leq am < 10$ nT shown in Figure 6. The modelled values are based on the mean *am* in each band which equals, respectively, 109.14, 75.94, 63.56, 47.76, 37.71 27.50, 13.96, and 5.32 nT. Modelled sensitivities are computed at points 1 h apart in *UT* and 1/365 apart in *F* and then averaged into the same sized *UT-F* bins (3 h by 0.05) as used for the observations. Note that the left-hand plots are colour-contoured using the 0.8–1.2 scale given by the lower colour bar while the modelled and observed s_{HAD}/S_{am} sensitivity ratios both use the 0.5–1.6 scale given by the upper colour bar. In all plots unity values are coloured yellow.

results for all eight *am* activity bins in every case, we here show just two illustrative ones, representative of low and moderately high activity levels. We choose $10 \leq am < 20$ nT for low activity and $60 \leq am < 110$ nT for high activity. We avoid the lowest *am* bin because the lowest row of Figure 8 shows that the fit is not always as good as it is for other panels: this is a sensitivity effect associated with the lowest activity level that can be detected by a single station, compared to that for the

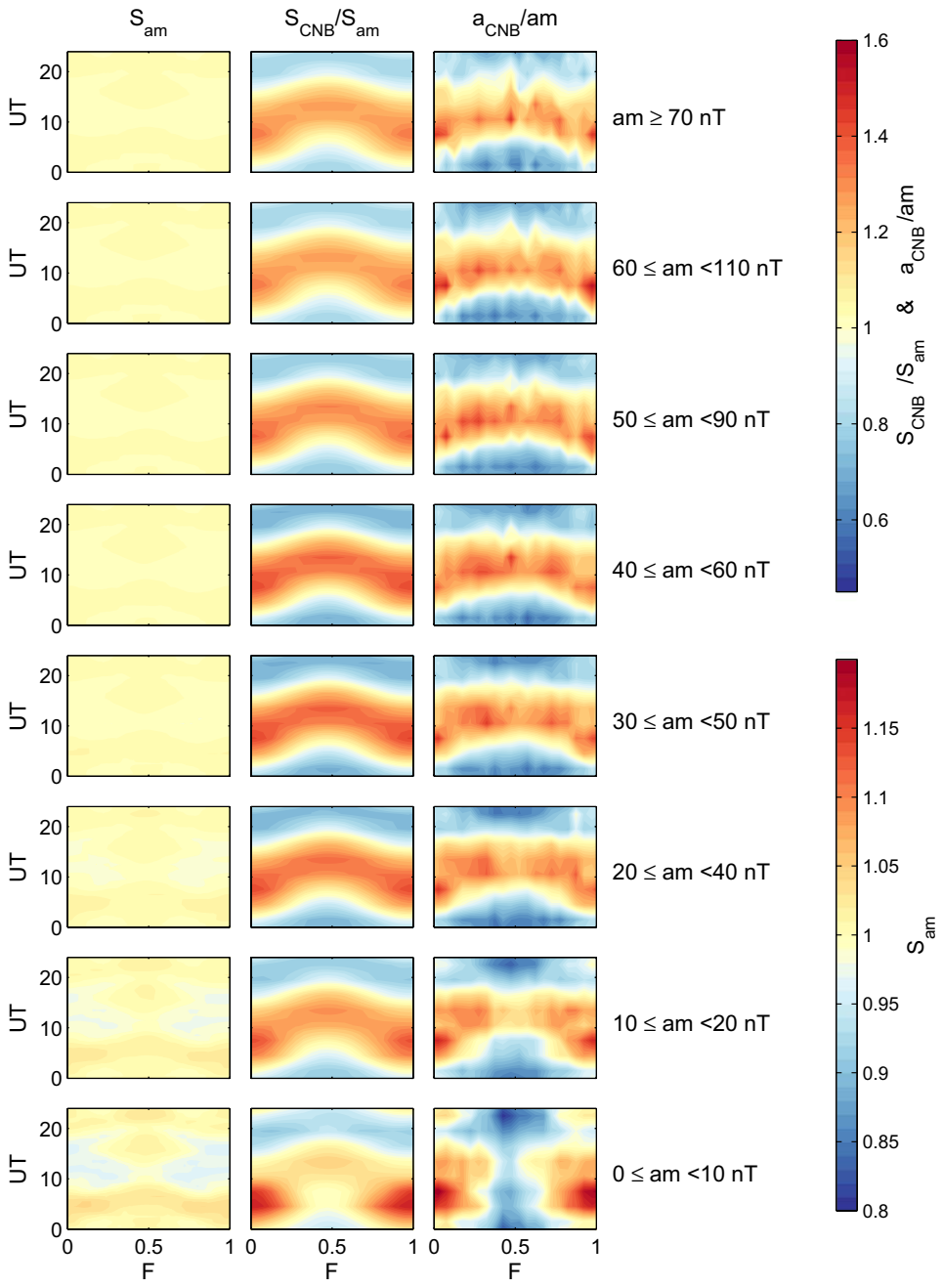

Fig. 8. The same as Figure 7 for the Canberra (CNB) station, giving *UT-F* plots of: (left column) the modelled sensitivity for the *am* index, S_{am}, for the current stations and sector weighting functions; (middle column) modelled values of the ratio s_{CNB}/S_{am} where s_{CNB} is the sensitivity of the Canberra magnetometer station for measuring its a_K values, a_{CNB}; and (right column) observed values of the ratio $a_{CNB}/am = s_{CNB}/S_{am}$.

am or *ap* index; that limit being lower for the indices, by virtue of the averaging of data from a number of stations. This means that there are times (as in the case of the bottom row of Fig. 8) when the station is not detecting any activity yet the *am* index is: in Figure 8 (for Canberra in the southern hemisphere) this is particularly true in mid-winter ($F = 0.5$). We avoid the largest activity bin ($am \geq 70$ nT) because it is based on the smallest number of samples. Instead we use the second largest and the second smallest activity bins as examples.

In all cases, the response functions are computed from the fitting procedure described in Section 2: the modelled sensitivities were computed for 24 different *UT*s (1 h apart) and 365 *F* values (daily) and for plotting of the *UT-F* patterns these were then averaged into the same bins as for the observational data (namely, eight 3-hour bins in *UT* and 20 18.25-day bins of *F*). We also made 3-hourly means of the sensitivity (for the same UT bins as the observations, i.e., 0–3 UT, 3–6 UT, etc.) and identified the maximum and minimum 3-hourly value of each *UT-F* pattern (S_{max} and S_{min}, respectively) as well as the mean ($\langle S \rangle$ and standard deviation (σ_S) of the 160 values (eight UT values and 20 *F* values) in each pattern. As well as taking the maximum percentage deviations of from the mean (S_{max} and

Table 3. Uniformity of average time-of-day/time-of-year response, $S_{av}(UT, F)$, of the various mid-latitude geomagnetic range indices.

Index	Rank	Start date	Method	F bin size	V_{av}	Largest positive % deviation of $S_{av}(F, UT)$ from the mean	Largest negative % deviation of $S_{av}(F, UT)$ from the mean
am	1	1959	Model	1 day	0.65	1.4	−1.2
ap_C	2	1932	Data	0.05 year	1.78	5.0	−4.6
aa_H	3	1868	Data	0.05 year	5.54	12.6	−12.3
aa	4	1868	Data	0.05 year	10.62	40.8	−17.0
aa	–	1868	Model	1 day	8.95	29.0	−16.8
ap	5	1932	Data	0.05 year	11.20	25.4	−21.6

S_{min}), we quantify the percentage standard deviation $V = (100\sigma_S)/(\langle S \rangle)$ for each am activity band. To evaluate the average behaviour, patterns were made for each of the eight *am* bands defined in Figure 6 and a weighted mean of the V values taken:

$$V_{av} = \sum_{b=1}^{8} \left(\frac{N_b}{\sum N_b} \right) V = 100 \cdot \sum_{b=1}^{8} \left(\frac{N_b}{\sum N_b} \right) \frac{\sigma_S}{\langle S \rangle} \quad (6)$$

where N_b is the numbers of *am* observations in the band (given earlier).

We also computed the average response pattern for the index, $S_{av}(UT, F)$ as the similarly weighted means of the eight patterns for the different activity levels

$$S_{av}(UT, F) = \sum_{b=1}^{8} \left(\frac{N_b}{\sum N_b} \right) S(UT, F) \quad (7)$$

Table 3 summarises the results by comparing the largest positive and negative deviations of $S_{av}(UT, F)$ for the tested indices as percentages of the mean, along with the metric V_{av} for various geomagnetic indices: results are given for both the modelling described above and from an empirical comparison with the *am* index.

3.1 Modelled response functions of the *an*, *as* and *am* geomagnetic indices

Figure 9 compares the modelled response patterns (the *UT-F* plots of index sensitivity) of the hemispheric *an* and *as* indices to that for *am* = (*an* + *as*)/2. These are all evaluated for the spatial distribution and weighting functions of stations for a selected example year which is 2014. At this time the IAGA codes of the stations in use are MGD, MMB, PET, POD, ARS, IRT, NVS, CLF, HAD, NGK, FRD, OTT, NEW, TUC, and VIC in the northern hemisphere and CNB, EYR, AMS, GNG, CZT, HER, PAF, AIA, PST, and TRW in the southern hemisphere (see Appendix A for the corresponding observatory names, locations and the intervals over which they were used to construct a given index). The computation procedure for deriving *am* and the longitude-sector weighting functions is described at http://isgi.unistra.fr/Documents/am_LWFs_example.pdf.

At both the moderately high and the low activity level examples shown (in the left and right and columns of Fig. 9, respectively) and in both hemispheres, there is a clear seasonal variation with enhanced index sensitivity in summer, S_{an} being largest around $F = 0.5$ and S_{as} being largest around $F = 0$ (which is the same as $F = 1$). If the longitudinal distributions of stations were ideal, the contours would all be vertical in these plots as there would be no *UT* variation. This is not quite the case, as S_{an} is slightly but consistently larger around 18 UT, and slightly lower around 04 UT at all times of year and all activity levels. S_{as} shows the converse behaviour, being similarly larger around 04UT and lower around 18 UT. Because many features in the S_{an} pattern are the converse of those in the S_{as} pattern, they are averaged out in *am* and the S_{am} patterns shown in the bottom panels of Figure 9 are much more uniform, especially for high geomagnetic activity. For 3-hourly values, the largest value is $S_{max} = 1.0284$ for the lower activity range with a minimum of $S_{min} = 0.9741$ and hence the largest percentage deviations from $\langle S_{am} \rangle = 1$, are 100 $(S_{max} - \langle S_{am} \rangle)/\langle S_{am} \rangle = 2.8\%$ and 100 $(S_{min} - \langle S_{am} \rangle)/\langle S_{am} \rangle = -2.6\%$, respectively. Hence, in this case (for which $\langle am \rangle = 14$ nT) the time-of-year/time-of-day response pattern for *am* is uniform to within maximum deviations of ±2.8%. For the higher activity range ($\langle am \rangle = 74$ nT) the corresponding extrema are much smaller, being +0.21% and −0.25%. The patterns are complex and general features very weak but include very slightly stronger annual variation at 10–15 UT and a band of very slightly higher S_{am} around 5 UT. For weighted means of the average *UT-F* response pattern over all activity levels, $S_{av}(UT, F)$ as defined by equation (7), the extrema are +1.4 and −1.2%, as given by the top row in Table 3. The other metric that we use to quantify the flatness of the average *UT-F* response pattern is V_{av} defined by equation (6). Table 3 shows that the modelled V_{av} value for the *am* index is 0.65%.

3.2 Modelled and observed response functions of the *aa* geomagnetic index

Figure 10 shows the *F–UT* response patterns for *aa*, $S_{aa}(UT, F)$ for the same high- and low-activity ranges of *am* used in Figure 9. Note that the colour scale range is considerably expanded in Figure 10 compared to Figure 9 because S_{aa} shows much greater deviations from unity than S_{am}, as is to be expected because *aa* is derived from data from just two stations. The response pattern will depend on which pair of stations is employed to generate *aa*. In Figure 10 various years are chosen as examples of long-lived *aa* station combinations: 2010 (contributing stations CNB and HAD), 1970 (TOO and HAD), 1930 (TOO and ABN) and 1890 (MEL and GRW). Table 4 shows that the minimum of the 3-hourly *aa* sensitivity, S_{aa}, is reasonably constant for these years being −18.5% in 1890 and −16.7% in 2010 for low activity and −18.9% in 1890 and

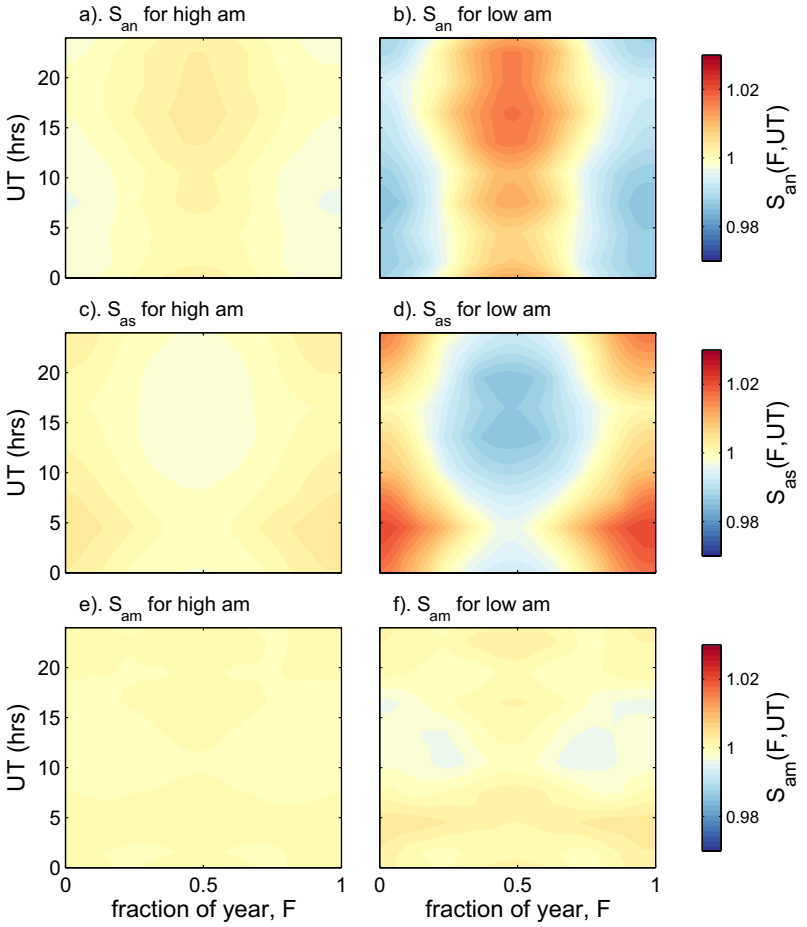

Fig. 9. Time-of-day (*UT*)/time-of-year (*F*) plots of the modelled sensitivity of (top row) the northern-hemisphere *an* index, S_{an}; (middle row) the southern-hemisphere *as* index, S_{as}; and (bottom row) the global index *am* = (*an* + *as*)/2, S_{am}. The left hand plots are for relatively high geomagnetic activity (defined as *am* = 74 nT, the mean of the 60 ≤ *am* < 110 nT band) and the right hand plots are for relatively low geomagnetic activity (defined as *am* = 14 nT, the mean of the 10 ≤ *am* < 20 nT band). All plots are for the stations and longitudinal sector weighting functions used in 2014.

16.1% in 2010 for high activity. The corresponding maximum values are almost constant at 12.8% for low activity and 12.0% for high activity. As shown in the next section, these extreme deviations from unity are actually smaller than those for *Kp* (*ap*) because, although *aa* is compiled using just two stations, those stations have been chosen to give as much cancellation of the stations' diurnal and annual response sensitivity variations as possible (being in opposite hemispheres and about 10 h apart in local time).

The "checkerboard" response pattern for *aa* seen in Figure 10 is present for both low and high activity and is found in the *aa* data, as demonstrated by Figure 11. This plot shows the sensitivity of the *aa* index $S_{aa} = (\langle aa\rangle/\langle am\rangle)S_{am}$, where the ratio $\langle aa\rangle/\langle am\rangle$ is taken from data for all available years (1959–2017). Note that the modelled sensitivity of the *am* index, S_{am}, is very close to unity at all *F* and UT and so the patterns for $\langle aa\rangle/\langle am\rangle$ are almost identical to those shown. The left-hand and right-hand panels are for high and low geomagnetic activity respectively. To get enough samples for this plot, high activity is defined in Figure 11 as *am* ≥ 40 nT both when averaging the observed $\langle aa\rangle$ and $\langle am\rangle$ values and when calculating the model sensitivity of *am*, S_{am}.

Table 4 shows that the peak deviations from unity are actually very similar in the two activity-level cases. Note also that the response patterns for the hemispheric *aa* indices aa_N and aa_S are given, for the current pairing of *aa* stations, by Figures 7 and 8.

Figure 10 shows that there are no strong differences in the modelled *aa* sensitivity patterns for the years studied. This means that the effects of station changes and of secular drift in the magnetic field on the response pattern for *aa* have been small. There is an important point to clarify here about the effect of secular change in the intrinsic geomagnetic field. All the *UT-F* response patterns in this paper have been normalised to unity. This means that any changes because of the drift in the average geographic latitude of the auroral oval will not be included. These effects have been studied by Lockwood et al. (2018a) and a new "homogeneous" *aa* index, aa_H, presented that makes allowance for this drift. The patterns presented in Figure 10 are the time-of-year/time-of-day variations around the annual mean, and although the annual mean estimates will have varied because of secular drift, the patterns will not have changed much at all. However, there is a small secondary effect of the secular change in the field that is included in the patterns

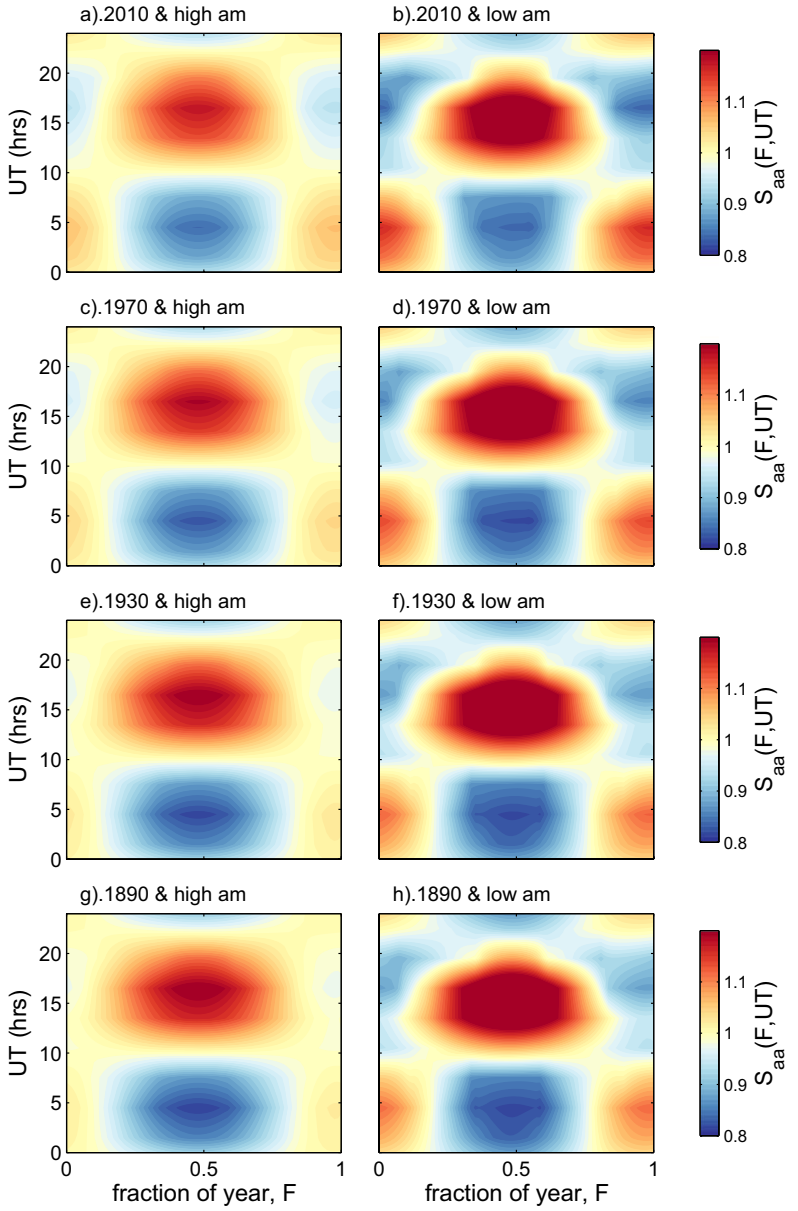

Fig. 10. Time-of-day (*UT*)/time-of-year (*F*) plots of the modelled sensitivity of the *aa* index, S_{aa}, for various years. The left-hand and right-hand columns are for relatively high and low geomagnetic activity (defined as for Fig. 9), respectively. Plots are for: (a) and (b) 2010; (c) and (d) 1970; (e) and (f) 1930 and (g) and (h) 1890.

Table 4. Maximum and minimum percentage deviations of modelled 3-hourly index sensitivities, *S*, from unity for selected years and middle and low geomagnetic activity levels.

Index	*S*	Year	Largest percent deviations of sensitivity from mean			
			$10 \leq am < 20$ nT		$60 \leq am < 110$ nT	
			Min	Max	Min	Max
am	S_{am}	2014	−2.59%	+2.84%	−0.49%	+0.49%
aa	S_{aa}	2010	−16.7%	+12.7%	−16.1%	+11.8%
aa	S_{aa}	1970	−18.5%	+12.8%	−17.7%	+11.9%
aa	S_{aa}	1930	−18.4%	+12.9%	−18.6%	+12.1%
aa	S_{aa}	1890	−18.5%	+12.9%	−18.9%	+12.1%

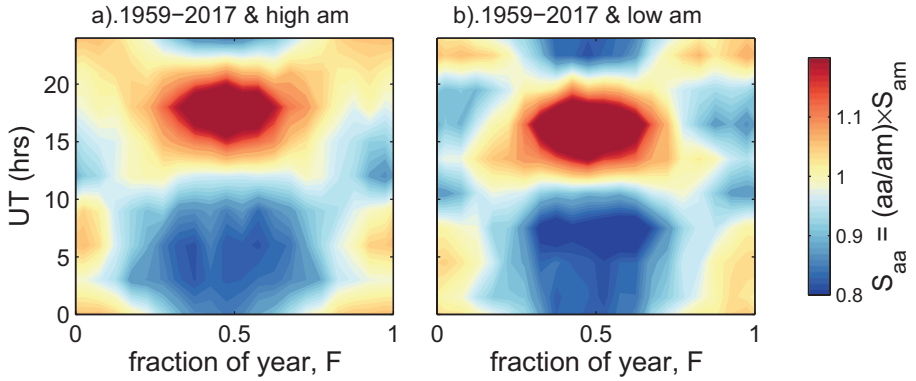

Fig. 11. Time-of-day (*UT*)/time-of-year (*F*) plots of the observed sensitivity of the *aa* index, $S_{aa} = (\langle aa\rangle/\langle am\rangle) \times S_{am}$, for the years 1959–2017. The left-hand and right-hand columns are for high and low geomagnetic activity respectively. The low activity range is $10 \leq aa < 20$ nT as used in Figures 7 and 8, but to get sufficient samples, high activity is here defined as $am \geq 40$ nT both when averaging the observed $\langle aa\rangle$ and $\langle am\rangle$ values and when calculating the model sensitivity of *am*, S_{am}.

presented and this is in the *UT* at which midnight MLT occurs, which has an effect via the term *T* in equation (3). The fact that there is no detectable effect in Figure 10 shows that this effect on the response pattern is very small.

The fourth row of Table 3 gives the the largest percentage deviations of the average pattern for the *aa* index and the V_{av} metric, both modelled and from data (by comparison with simultaneous *am* data).

3.3 Response function of the *ap* (*kp*) geomagnetic index

Our model cannot be employed to analyse the response pattern of the *ap* and *Kp* indices. This is because equation (2) requires the index to be compiled by linear mathematics so that the solar forcing term can be isolated and cancelled. In the compilation of *Kp*, the K_S/K factor used is a complex function of *K* and hence the process is achieved by a look-up table, rather than an analytic function and is also non-linear as the look-up table used depends on the activity level. Hence we cannot use the same model analysis as applied above to *am* and *aa*. However, we can apply the data-based approach (comparison with the *am* index) that was used in the previous section for *aa*. As for *aa*, we can estimate the response function for *ap* using the equation

$$S_{ap}(UT,F) = \left[\frac{\langle ap\rangle}{\langle am\rangle}\right]_{(UT,F)} \times S_{am}(UT,F) \qquad (8)$$

Figure 12 compares the *am* and *ap* index response patterns over the full interval over which we have both, namely 1959–2017 (inclusive). The middle column of Figure 12 presents the *UT*-*F* patterns of the ratio $\langle ap\rangle/\langle am\rangle$ (again for eight 3-hourly UT values, and twenty bins in *F* that are $\Delta F = 0.05$ wide, and the eight *am* activity level bins shown in Fig. 6 and as employed in Figs. 7 and 8). The left hand column gives the modelled *am* sensitivity patterns, $S_{am}(UT, F)$. Using these patterns, we can estimate the *ap* sensitivity S_{ap} using equation (8). The results are shown in the right-hand column of Figure 12. Note that the patterns for S_{ap} are very similar indeed to those for $(\langle ap\rangle/\langle am\rangle)$ in the middle column because S_{am} is so close to unity at all *UT* and *F*. Figure 12 shows a consistent pattern in $S_{ap}(UT, F)$ with response at 0–8 UT being greater in northern hemisphere winter but that at all other UT being greater in northern hemisphere summer. At low activity levels, the 0–8 UT variation tends to dominate but the 8–24 UT variation increasingly dominates with increasing activity level.

The values of S_{am} are very uniform and close to unity (note that the S_{am} scale is the lower one in Figure 12 and covers a smaller dynamic range than for the middle and right columns by a factor of six). This is particularly true for moderate and high activity. However, even at low activity, the differences between the pattern of $\langle ap\rangle/\langle am\rangle$ and the corresponding pattern of $(\langle ap\rangle/\langle am\rangle)S_{am}$ cannot be detected. The *ap* index is often treated as homogeneous but, in reality, the network of stations used has changed considerably even over the interval 1959–2017. Given the variability that this introduces, the use of the iteratively-derived factor S_{am} is probably not justified and in the remainder of this section we assume S_{am} is unity and just look at the ratio $(\langle ap\rangle/\langle am\rangle)$. This has the advantage of making the analysis purely empirical (instead of the empirical-model mixture involved in S_{ap} estimates).

Figure 13 studies the relationship between the *ap* and *am* indices. The grey dots in Figure 13a are 3-hourly values (*ap* and *am*) and the cyan points are 8-point (1 day) running means (*Ap** and *Am**). The black dots are means (with error bars that are $\pm 1\sigma$) taken in bins of *am* that are $\Delta am = 10$ nT wide (only bins with six or more samples are shown). The mauve line shows *am* = *ap* and the plot shows that *ap* is consistently smaller than *am*. In addition, the difference grows with activity levels so that *Ap** values are increasingly smaller than *Am** values for larger activity. Lockwood et al. (2019b) point out that this is a significant factor when geomagnetic storms during the recent grand solar maximum, as defined using the *Ap** data, are compared to those seen during 1868–1932 in the *Aa** data or to the estimates for during the Carrington storm (that are often quantified by a proxy equivalent for *Aa**). The overall mean of *am* is larger than that for *ap* by a factor $\eta = 1.5339$.

Note that to a large extent taking 24-hour running means reduces the effect of the *UT* variation in the sensitivity of any one 3-hourly index; however, it does not remove it completely.

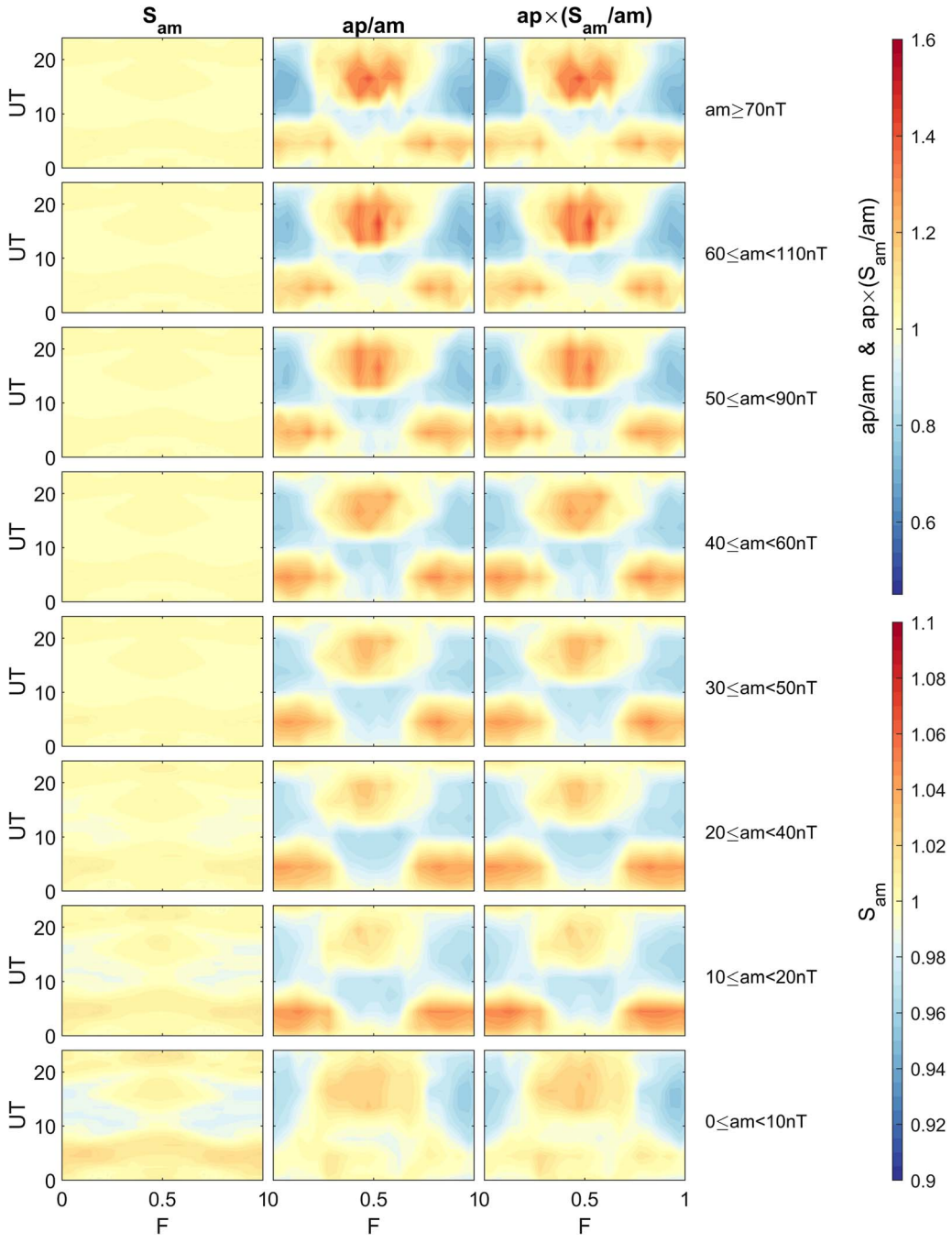

Fig. 12. Time-of-day (*UT*)/time-of-year (*F*) plots of: (left-hand column) the sensitivity of the *am* index, S_{am}, for 1988 (the mid-point of the interval 1959–2017 for which *am* data are available); (middle column) observed $\langle ap\rangle/\langle am\rangle$ and (right column) $S_{ap} = (\langle ap\rangle/\langle am\rangle)S_{am}$. Data are for 1959–2017 (inclusive) and the *am* ranges defined in Figure 6.

This is because the 3-hourly index value will generally have some variation within each 24-hour interval and the phasing of this variation, relative to the *UT* variations in the index sensitivities, will influence the *Ap** and *Am** values and influence them differently.

Figure 13b and c compares the annual variations, by taking mean values in 20 bins of the fraction of the year, *F*, that are $\Delta F = 0.05$ wide (18.25 days). The black dots and line show means of *am*, the mauve dots and line means of $\eta \times ap$. It can be seen that the annual variations are very similar, but that the well-known semiannual variation (Cortie, 1912, Chapman & Bartels, 1940; Russell & McPherron, 1973; Cliver et al., 2002; Le Mouël et al., 2004) is proportionally slightly larger, on average, in *ap* than in *am*. The average variation with *F* of the ratio of $\langle ap/am\rangle$ is shown in Figure 13c. Figure 13b and c show that *ap* also has an annual variation as it underestimates geomagnetic activity during northern hemisphere winter, which is perhaps not surprising given the dominance of northern

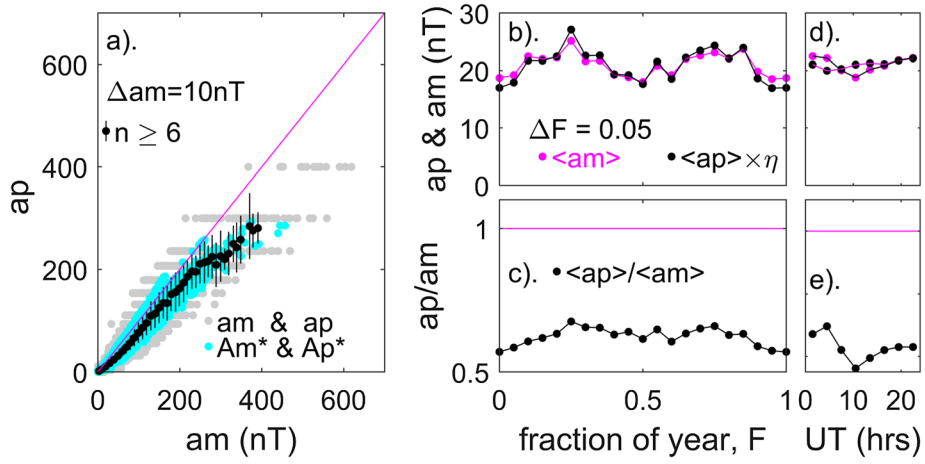

Fig. 13. (a) Grey points shows *ap* index values as a function of simultaneous *am* values for 1959–2017 (inclusive). Cyan points show the scatter plot of the corresponding 24-hour (8-point) running means, *Am** and *Ap**. The black points are means of *ap* (with error bars of plus and minus one standard deviation) as a function of means of *am* in *am* bins of width Δam = 10 nT (only means for bins containing six or more samples are shown). The mauve line is the ideal case for which *ap* = *am*. (b) and (c) Annual variations shown by means in fraction-of-year (*F*) bins of width ΔF = 0.05. (b) shows (in black) the annual variations of $\langle am \rangle$ and (in mauve) $\langle ap \rangle \times \eta$, where η is the ratio $\langle am \rangle_{all}/\langle ap \rangle_{all}$ for means taken over the whole dataset. (c) shows the annual variation of the ratio of *ap*/*am*. (d) and (e) Diurnal variations shown by means for the eight UT values of both indices using the same color coding as in (b) and (c).

hemisphere stations and the lower ionospheric conductivities in winter. Figure 13d and e presents the same analysis for diurnal variations. Figure 13d shows that there is a slight diurnal variation in *am* and a larger one in *ap*. The diurnal variation in the ratio of the two indices is shown in Figure 13e.

The average $S_{ap}(UT, F)$ pattern gives percentage deviations of 11.2%, 25.4% and −21.6% for σ_S, S_{max} and S_{min}, respectively. Table 3 shows this is the least uniform of the indices tested in this paper. In Section 4 we develop and test an empirical correction for this.

3.4 Response function of the homogeneous *aa* index, aa_H

For comparison, it is useful to apply to the new homogenized *aa* index, aa_H, the same tests as used in the last section to study the constancy of the *ap* index. It is not instructive to test this index against the sensitivity model because the model is used to correct *aa* and give aa_H. Figure 14 compares with the *am* index and so corresponds to Figure 12. The left-hand column again gives the $S_{am}(UT, F)$ patterns, the middle column $(\langle aa_H \rangle/\langle am \rangle)_{(UT,F)}$ and the right hand column $S_{aaH} = (\langle aa_H \rangle/\langle am \rangle)_{(UT,F)} \times S_{am}(UT, F)$. The fluctuations around unity in the middle and right column are of smaller amplitude than for *ap* and in character are more of a regular *UT* variation (at all *F*), which means that averaging over a calendar day or making 8-point running means (to give Aa_H and Aa_H*, respectively) will further reduce the effect of the non-uniformity of $S_{aaH}(UT, F)$.

Figure 15 corresponds to Figure 13 for the aa_H index. The agreement of aa_H and *am* (grey dots) and of Aa_H* and *Am** (cyan dots) in part (a) are both good and linear. In parts (b) and (c), the annual variations of $\langle aa_H \rangle$ and $\langle am \rangle$ with *F* are very similar (with aa_H being just very slightly lower values round *F* = 0.5 and very slightly higher around *F* = 0 and *F* = 1). Parts (d) and (e), show there is a *UT* variation in average aa_H that is somewhat greater than that in *am*. The mean $S_{aaH}(UT, F)$ pattern gives percentage deviations of 11.20%, 25.4% and −21.6% for σ_S, S_{max} and S_{min}, respectively. Table 3 shows that the average pattern gives percent deviations of 5.54%, +12.6% and −12.3% for σ_S, S_{max} and S_{min}, respectively for aa_H which is considerably better than for the uncorrected *aa* data and therefore also better than for *ap*.

4 A corrected *ap* index, ap_C

In this section, we present a method for empirically correcting the *ap* index to allow for its non-uniform response function. We do this using the same basic principle as introduced by Bartels and still used today to compile *ap*, namely activity-dependent look-up tables to correct the time-of-day/time-of-year response. A difference is that we are applying it to an index and not the data from a single station. The tables are provided in the form of arrays of values that can be interpolated to give the multiplicative correction factor required for a given *ap*, *UT* and *F*. We recommend use of Piecewise Cubic Hermite Interpolating Polynomial, (PCHIP) interpolation because it maintains smooth changes in gradient through the data points. Tests show that, although in general that it can sometimes give excessive oscillation between points, in this case it gives more accurate value than a linear interpolation).

Activity level in previous sections was quantified using *am* index bands, but that cannot be applied here; instead, *ap* must be used as the aim is to correct *ap* even at times when *am* was not measured. Figure 16 studies the distribution of *ap* values. The blue line gives the c.d.f of the (quantised) 3-hourly *ap* values and the mauve line the (almost continuous) *Ap** values. The vertical grey and white bands divide the *Ap** distribution into 20 percentiles, each containing N_{20} = 8612 samples. The separation of these percentiles is smaller than the separation of the *ap* levels below 22 and so all these levels contain

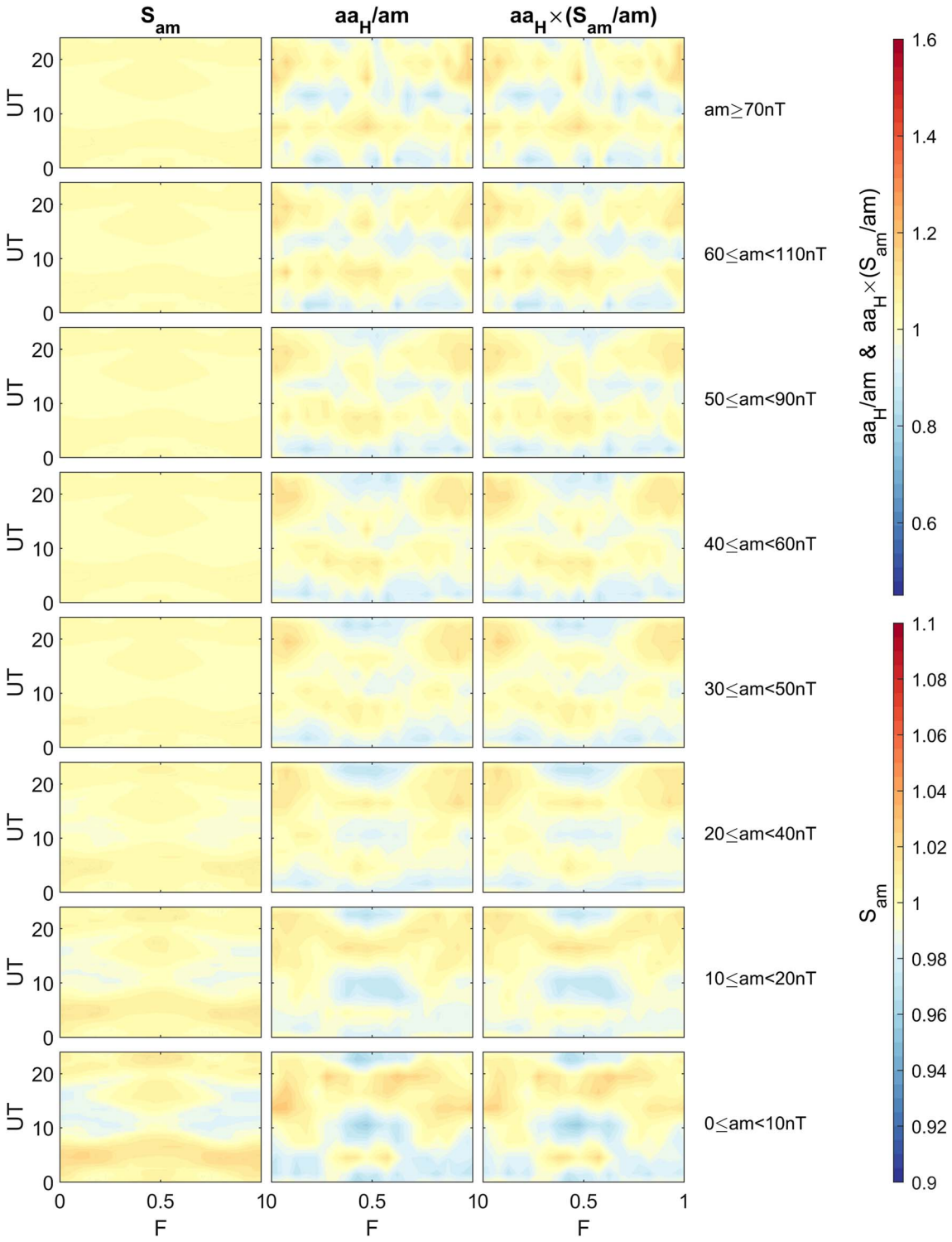

Fig. 14. The same as Figure 12 for the homogenized *aa* index, aa_H. Data are for 1959–2017 (inclusive) and the *am* ranges defined in Figure 6.

considerably more than N_{20} samples. The ap = 22 level containts very close to N_{20} samples level and the ap = 27 and ap = 32 levels combined give slightly more than N_{20} samples, but we have to combine all ap levels greater than 39 to get more than N_{20} samples in the tail of the distribution. Hence we derive corrections for ap levels of 2, 3, 4, 5, 6, 7, 9, 12, 15, 18, 22, 27 combined with 32, and 39 and greater. We also deal with each of the eight *UT*s of the ap index separately. For each of these *UT-ap* combinations we then compute the empirical correction factor $f_c(UT, F, ap) = \langle ap \rangle / \langle am \rangle$ in 20 equal-width bins in F. These can be used to convert an ap value for a general F and UT using PCHIP interpolation between the relevant 20 $f_c(F)$

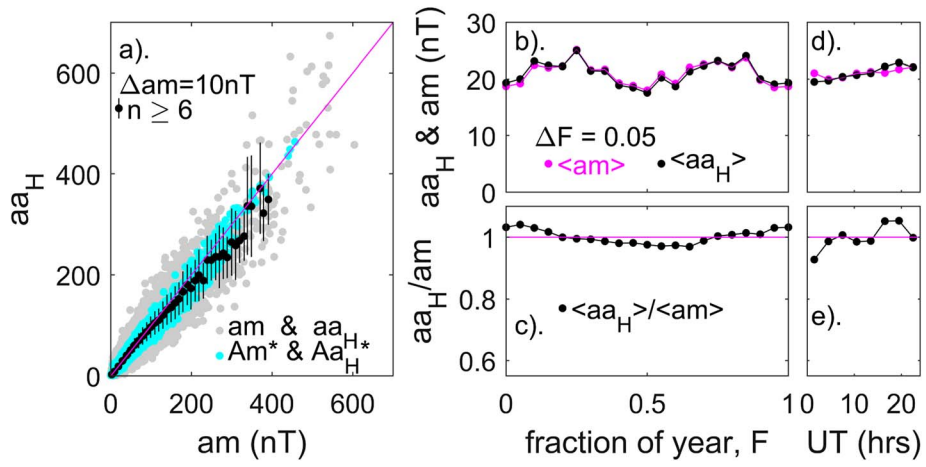

Fig. 15. Same as Figure 13 for the homogenized *aa* index, aa_H. Data are for 1959–2017 (inclusive).

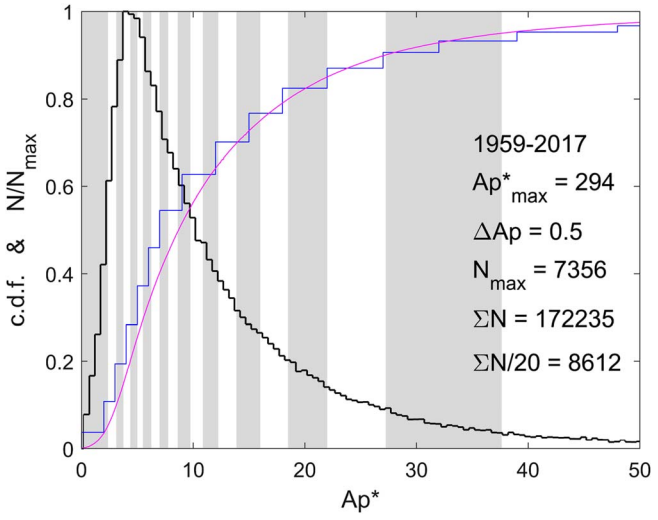

Fig. 16. The distributions of *ap* and *Ap** values. The mauve and blue lines are the c.d.f.s of, respectively, *Ap** and *ap* values for 1959–2017, inclusive. The black line is the histogram of N/N_{max}, where N is the number of *Ap** samples in bins 0.5 wide and N_{max} is the maximum value of N. The vertical white and grey bands divide the distribution of *Ap** into the 20-percentiles, each containing $\Sigma N/20 = 8612$ samples.

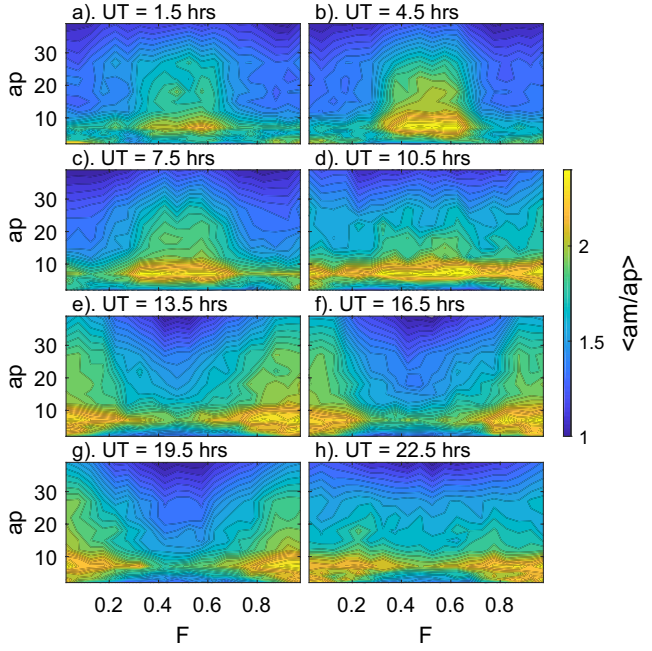

Fig. 17. Plots of fits to the ratio of *am*/*ap* as a function of time of year F and *ap* value for the 8 *UT*'s of the *ap* and *am* index samples, derived as described in the text so that they can be used to generate the corrected *ap* index, ap_C from *ap*.

values. Figure 17 presents contour plots of daily $f_c(ap, F)$ values for the eight *UT* ranges.

This allows us to turn every *ap* value into a corrected value $ap_C = ap \times f_c(UT, F, ap)$. Figure 18 compares the ap_C and *am* values, in the same format as Figures 13 and 15. Part (a) shows that there is very good agreement between ap_C and *am* and between Ap_C* and *Am**. Parts (b) and (c) shows that the average annual variations of ap_C and *am* are very similar and parts (d) and (e) even their UT variations are well matched. The mean $S_{apC}(UT,F)$ pattern gives percentage deviations of 1.78%, 5.0%, and −4.6% for σ_S, S_{max} and S_{min}, respectively. Table 3 shows this is a major improvement compared to *ap*.

The coefficients needed to implement the correction to *ap* are given in the Supporting Information file attached to this paper. The metadata given in the header to that file also gives some MATLAB program code that converts *ap* into ap_C.

5 Conclusions and recommendations

5.1 Conclusions

We have presented analysis of the time-of-day (*UT*)/time-of-year (*F*) response patterns of various planetary geomagnetic

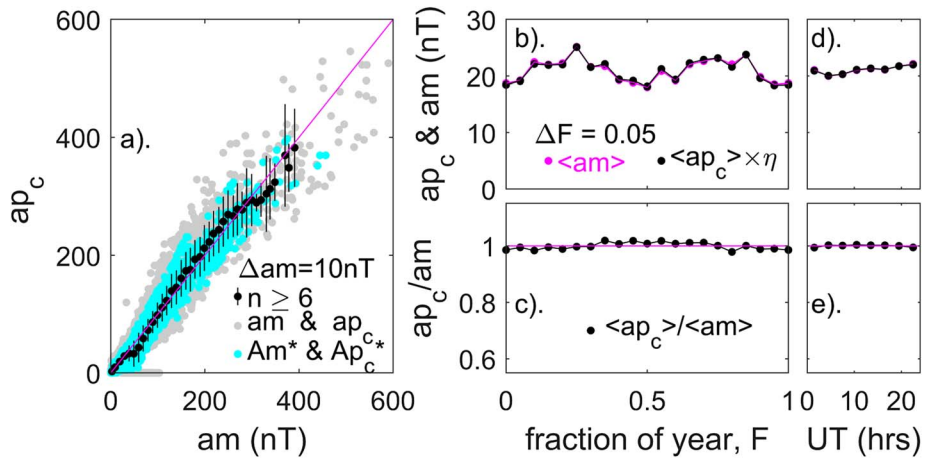

Fig. 18. Same as Figures 13 and 15 for the corrected *ap* index, ap_C.

indices. In general, responses depend on the level of geomagnetic activity because the effects of EUV/X-ray generated solar conductivity local to the observatory dominate the response at low geomagnetic activity, but the effects of the great circle distance between the observatory and the midnight sector auroral oval dominate at high geomagnetic activity.

The diurnal variations in the hemispheric *an* and *as* indices are small because of the use of station groupings that are as even in longitudinal separation as possible. On averaging to give the *am* index, the seasonal variations cancel to a very large extent and so the index sensitivity, S_{am}, is close to unity at all *F* and *UT*. This is valid for three-hourly values to within extrema of ±2.9% at low geomagnetic activity, falling to just ±0.5% at high geomagnetic activity. As an overall average, the modelled (*UT*-*F*) response pattern of *am* is constant to within a standard deviation of 0.65% with extreme deviations of +1.4% and −1.2%.

The *ap* (*Kp*) index station distribution is not uniform but the use of the *K* to *K*s conversion tables makes allowance for this. These tables were constructed using data from a limited number of years when solar and geomagnetic activity were at moderate levels only. Hence it not that surprising that they may not be ideal over all the 1957–2017 interval tested here. The general similarity of the average annual and diurnal variations of *ap* and *am* in Figure 13 indicates that the tables are performing well in that they do not introduce major errors. The right hand column of Figure 11b, however, shows that there are spurious time-of-year variations in the response of *ap* that depend strongly on *UT*. For $0 < UT < 10$ h there are peaks at almost all activity levels around $F = 1$ (northern hemisphere winter), whereas at all other *UT* the peak is around $F = 0.5$ (northern hemisphere summer). Averaging over all *UT* gives peaks in the sensitivity at the equinox at that can be seen in Figure 13c, which reveals that that *ap* exaggerates slightly the semiannual variation in geomagnetic activity. There is also a spurious net annual variation with values lower around $F = 1$ than $F = 0.5$. Note that Figure 4a is also consistent with this, showing the greatest average response in *ap* to solar wind forcing at equinox and a greater response around the June solstice than around the December solstice. Figure 13d shows there is also a persistent net diurnal variation in the *ap* response with a minimum at 9–12 *UT*. These average values however hide the fact that the *ap* response function is a complex and variable function of *UT*, *F* and activity level, as shown by Figure 11. We note that these spurious diurnal and annual variations in *ap* arise because this index employs concentrations of stations in certain regions (particularly Europe). But this, and making the data from all stations mimic the Niemegk reference station by converting to K_S, does also have advantages in noise suppression because one is averaging different estimates of the same thing and, when averaged over a whole year of continuous data (for which $\langle S_{ap} \rangle_{\tau=1yr} = 1$), the spurious variations with *UT* and *F* are largely averaged out – hence the *ap* index almost certainly provides the most reliable and accurate estimate on annual timescales. This is reflected in the response to solar wind forcing shown in the right hand panels of Figure 4 which reveals that *ap* performs slightly, but significantly, best on annual averaging timescales. However, we also note that the uneven pattern of S_{ap} will introduce some random sampling noise, depending on the time of year and *UT* at which the quasi-randomly-occurring largest interplanetary disturbances happen to hit Earth's magnetosphere.

The sensitivity modelling indicates that the *aa* index *response* pattern has remained very constant over time and is comparable, and actually slightly better than, that derived empirically for *ap*. This is because the southern hemisphere data is given equal weight in *aa*, whereas it has much lower weight in *ap*.

Our analysis shows that the new "homogeneous" *aa* index aa_H (Lockwood et al., 2018a, b) performs considerably better than *aa*, having a flatter *UT*-*F* response at all activity levels. Naturally, being based on just one station in each hemisphere it cannot match the performance of *am*, but it has the advantage (for studies of long-term change or requiring large sample numbers) of extending back to 1868 whereas *am* extends back to only 1959.

We have also presented a method for correcting the *ap* index for its uneven response pattern. The correction presented here uses all *ap* and *am* data since 1959, despite the fact that the network of stations used to generate the *am* index has undergone (relatively minor) changes in that interval and that used to generate the *ap* index has undergone considerable changes. This means that, in effect, we are accepting the tabular *K*-to-K_S conversions inherent in *ap* for the purposes in deriving a correction factor to *ap*, $f_c(ap, F, UT)$, but this is one step removed from accepting them for the purposes of compiling

Table 5. Standard deviations and maximum and minimum percentage deviations from unity of observed 3-hourly index sensitivities, S, estimated assuming the *am* index is ideal for 1959–2017, for low, middle and high geomagnetic activity levels. The bottom two rows give the factor by which the corrected index is improved in terms of the uniformity of its response.

Index	$am < 8$ nT			$8 \leq am < 30$ nT			$am \geq 30$ nT		
	Lowest 20%			Inter 20%			Largest 20%		
	σ_S (%)	Max (%)	Min (%)	σ_S (%)	Max (%)	Min (%)	σ_S (%)	Max (%)	Min (%)
ap	10.18	20.2	−22.8	12.21	29.10	−17.40	12.88	28.40	−24.52
ap_C	3.07	8.10	−9.40	2.27	5.52	−5.73	3.85	10.62	−13.73
aa	12.74	39.80	−20.80	11.44	40.20	−20.02	9.70	34.91	−18.30
aa_H	2.37	8.11	−5.80	1.41	4.94	−3.05	2.51	6.31	−8.49
	Ratios of values for uncorrected indices to those for corrected								
For ap_C	3.32	2.49	2.43	9.34	5.27	3.04	3.35	2.67	1.79
For aa_H	5.38	4.91	3.59	8.11	9.96	6.56	3.86	5.53	2.16

ap itself. Hence this is a first order correction. It would be possible to make a more detailed correction and analyse the response and associated $f_c(ap, F, UT)$ factors for each of the various combinations of stations used to construct *ap* after 1959, as given by Appendix A. However, this would introduce other uncertainties caused by the reduction in the number of samples in each case. In addition, this would still not account for the several changes to the *ap* network made before 1959. The optimum correction to *ap* would be to correct the *K* indices for each station individually using the station sensitivity model and then average them together as this could be done in a consistent way for all the data (since 1932). However, this is a large task and well beyond the scope of this paper.

Table 5 quantifies the improvements to the *aa* and *ap* indices made by aa_H and ap_C, respectively, using *am* as the calibration standard. Note that in the case of aa_H, the correction is made through application of the station sensitivity model, whereas for ap_C it is purely empirical. The most revealing division of activity levels is to consider the lowest 20% of samples (as determined by *am*), the highest 20% and the remainder in the inter 20-percentile range. The bottom two rows of Table 5 give the factor improvement brought about in the corrected indices: ideally these values should be infinite as *am* is used to correct the data and then used to test the corrected index. However, this is not going to be the case for a variety of reasons: in the case of aa_H, the main reason is that by averaging many stations *am* will always suppress geophysical and instrumental noise more effectively than *aa* which is based on just two stations; in the case of ap_C, the main reason is the effect of changes in the *ap* observation network. It can be seen that improvements, in terms of lowering σ_S (flattening the response function) are always in excess of three and that for extrema the improvement in S_{max} is always better than that in S_{min}. In both cases the improvement for the bulk of the distribution is always better than for the lower 20% and the upper 20%.

Because taking running means over 24 h averages out the UT dependence, the improvements in Aa_H* and in Ap_C* are not as great as for the 3-hourly values aa_H and ap_C. Nevertheless, the ranking order of storm days in these indices can be considerably different from those for *Aa** and in *Ap**. The occurrence of storm days, as quantified by Aa_H* since 1868 will be the subject of another paper. Lockwood et al. (2019b) have studied the behaviour and ranking order of extreme events since 1932, as quantified using the Ap_C* index.

5.2 Recommendations

The major differences between the indices studied in this paper arise from the geographic distribution of stations used to compile them. The *aa* index uses just two stations in order to give a long data sequence. The *am* index has been designed to make the stations' distribution as even as possible and so give the most even time-of day/time-of-year response that is allowed by the availability of land that is suitable for housing a magnetometer. On the other hand, the *ap* index is, and has always been, dominated by European stations and is constructed in a way that re-calibrates all data to a European station (Niemegk): we have shown that because it is averaging more data that is similar (or is made similar), this gives it better sensitivity and more accurate annual means but an un-even time-of-day/time-of-year response. Lockwood et al. (2018a, b) have presented a model-based means for correcting the *aa* index, to give the "homogeneous" *aa* index aa_H, that allows for both long-term secular changes and the unevenness of the time-of-day/time-of-year response pattern caused by the use of just two stations. In this paper, we have used empirical comparisons with the *am* index to generate a corrected *ap* index, ap_C, that allows for the unevenness of the time-of-day/time-of-year response pattern of *ap* identified in this paper. Because of the differences between these corrections, their potential applications are different. The corrections to *aa* are based on physical model and employ parameters that can be projected into and measured in the future (the stations' geographic and geomagnetic coordinates and the solar declination angle): hence we recommend that it is used for both re-evaluation of past studies and for studies that extend into future. On the other hand, the *ap* corrections (to give ap_C) are empirical and based on comparison with past *am* index data and so will become increasingly unreliable with time for future data. Hence ap_C will be useful in re-evaluation of some past work for which evenness of the time-of-year/time-of-day response to solar wind forcing is important (for example studies involving the ranking order of the largest geomagnetic storms, Lockwood et al., 2019b, c) but it would not make sense to recommend its use for future work with that requirement, because as long as the *am* index is available it is the better option in this context. That is not to say, that the *ap* index is not of value in other contexts. For example, in annual means of the data, the time-of-year/time-of-day response is averaged out and so long term trends are well monitored and for such

applications keeping the same index (and keeping its compilation as homogeneous as possible) is important. We note in that *ap* extends back 27 years further into the past than *am* and so for studies where the sub-annual response is important, ap_C can be used to extend the *am* data sequence back a further 27 years to 1932.

Looking to any potential development of the networks of stations, it is important that any changes made enhance, rather than undermine, the strengths of the index in question. From the above, we would argue that in the case of *aa* the strength is longevity. In the case of *ap* it is part longevity and part sensitivity brought about by averaging many nearby stations and calibrating all station to Niemegk. For longevity, homogeneity of compilation is by far the most important consideration, meaning that both past and future station changes are undesirable. In the case of *aa*, the station sensitivity model can be used to minimise the effects of station changes, as does the scaling to Niemegk in the case of *ap*. Hence moving *ap* stations to make the geographic distribution more even (and make *ap* more like *am*) would not be desirable change as it would undermine the advantages that *ap* has. For *am*, the choice of stations and the method of their combination have been shown to be good in that it gives a very flat time-of-day/time-of-year response pattern, which is the index's most important strength. This is true at all levels of geomagnetic activity, except the very quietest which suggests a sensitivity issue. In terms of improving the network distribution for *am*, the major limitation is the availability of accessible land on which a magnetometer can be placed and operated. However, there may be advantages for some applications in increasing the numbers of stations in all the longitude sectors to give increased averaging out of local site and instrumentation effects and so increase sensitivity. However, so as not to disrupt the evenness of the response, one would want to make similar, matching, improvements in all sectors.

Supplementary material

Supplementary material is available at https://www.swsc-journal.org/10.1051/swsc/2019017/olm.

Acknowledgements. The authors are grateful to the staff of The ISGI, France and collaborating institutes for the compilation and databasing of the *am* index which were downloaded from http://isgi.unistra.fr/data_download.php. We also thank the staff of Geoscience Australia, Canberra for the southern hemisphere *aa*-station *K*-index data, and at British Geological Survey (BGS), Edinburgh for the northern hemisphere *aa*-station *K*-index data. For the SuperMAG indices data we gratefully acknowledge: Intermagnet; USGS, Jeffrey J. Love; CARISMA, PI Ian Mann; CANMOS; The S-RAMP Database, PI K. Yumoto and Dr. K. Shiokawa; The SPIDR database; AARI, PI Oleg Troshichev; The MACCS program, PI M. Engebretson, Geomagnetism Unit of the Geological Survey of Canada; GIMA; MEASURE, UCLA IGPP and Florida Institute of Technology; SAMBA, PI Eftyhia Zesta; 210 Chain, PI K. Yumoto; SAMNET, PI Farideh Honary; The institutes who maintain the IMAGE magnetometer array, PI Eija Tanskanen; PENGUIN; AUTUMN, PI Martin Connors; DTU Space, PI Dr. Rico Behlke; South Pole and McMurdo Magnetometer, PI's Louis J. Lanzarotti and Alan T. Weatherwax; ICESTAR; RAPIDMAG; PENGUIn; British Artarctic Survey; McMac, PI Dr. Peter Chi; BGS, PI Dr. Susan Macmillan; Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation (IZMIRAN); GFZ, PI Dr. Juergen Matzka; MFGI, PI B. Heilig; IGFPAS, PI J. Reda; University of L'Aquila, PI M. Vellante; BCMT, V. Lesur and A. Chambodut; Data obtained in cooperation with Geoscience Australia, PI Marina Costelloe; SuperMAG, PI Jesper W. Gjerloev. SuperMAG data are available from http://supermag.jhuapl.edu/indices/?layers=SME.UL. The work at University of Reading (UoR) is supported by the SWIGS NERC Directed Highlight Topic Grant number NE/P016928/1 with some additional support from STFC consolidated grant number ST/M000885/1. The work at École et Observatoire des Sciences de la Terre (EOST) is supported by CNES, France. Initial work for this paper was carried out by IDF as part of his PhD studies at Southampton University, where he was a part-time student: we are grateful to Rutherford Appleton Laboratory and to Southampton University for supporting that PhD work. CH is supported on a NERC PhD studentsship as part of the SCENARIO Doctoral Training Partnership. The editor thanks Lauri Holappa and an anonymous referee for their assistance in evaluating this paper.

References

Adebesin BO. 2016. Investigation into the linear relationship between the *AE*, *Dst* and *ap* indices during different magnetic and solar activity conditions. *Acta Geod Geophys* **51(2)**: 315–331. DOI: 10.1007/s40328-015-0128-2.

Allen JH. 1982. Some commonly used magnetic activity indices: Their derivation, meaning, and use. In: *Proceedings of a Workshop on Satellite Drag, March 18–19, 1982, Boulder, Colorado*, Joselyn JAC (Ed.), pp. 114–134.

Aksnes A, Stadsnes J, Bjordal J, Østgaard N, Vondrak RR, et al. 2002. Instantaneous ionospheric global conductance maps during an isolated substorm. *Ann Geophys* **20(8)**: 1181–1191. DOI: 10.5194/angeo-20-1181-2002.

Bartels J, Heck NH, Johnston HF. 1939. The three-hour-range index measuring geomagnetic activity. *Terr Magn Atmos Electr* **44(4)**: 411–454. DOI: 10.1029/TE044i004p00411.

Bartels J. 1949. The standardized index Ks and the planetary index Kp. *IATME Bull* **12b**: 97.

Bartels J. 1957. The geomagnetic measures for the time-variations of solar corpuscular radiation, described for use in correlation studies in other geophysical fields. *Ann Intern Geophys Year* **4**: 227–236.

Caan MN, McPherron RL, Russell CT. 1978. The statistical magnetic signature of magnetospheric substorms. *Planet Space Sci* **26(3)**: 269–279. DOI: 10.1016/0032-0633(78)90092-2.

Chambodut A, Marchaudon A, Menvielle M, El-Lemdani F, Lathuillere C. 2013. The K-derived MLT sector geomagnetic indices. *Geophys Res Lett* **40**: 4808–4812. DOI: 10.1002/grl.50947.

Chapman S, Bartels J. 1940. *Geomagnetism*, vol. **2**, Oxford University Clarendon Press, Oxford and London, UK. ISBN 978-5-88199-480-8.

Chu X, Hsu T-S, McPherron RL, Angelopoulos V, Pu Z, Weygand JJ, Khurana K, Connors M, Kissinger J, Zhang H, Amm O. 2014. Development and validation of inversion technique for substorm current wedge using ground magnetic field data. *J Geophys Res Space Phys* **119**: 1909–1924. DOI: 10.1002/2013JA019185.

Clauer CR, McPherron RL. 1974. Mapping the local time-universal time development of magnetospheric substorms using mid-latitude magnetic observations. *J Geophys Res* **79(19)**: 2811–2820. DOI: 10.1029/JA079i019p02811.

Cliver EW, Kamide Y, Ling AG. 2002. The semiannual variation of geomagnetic activity: phases and profiles for 130 years of aa data. *J Atmos Sol-Terr Phys* **64**: 47–53. DOI: 10.1016/s1364-6826(01)00093-1.

Cortie AL. 1912. Sunspots and terrestrial magnetic phenomena, 1898–1911. *Mon Not R Astron Soc* **73**: 52–60. DOI: 10.1093/mnras/73.1.52.

Cowley SWH, Lockwood M. 1992. Excitation and decay of solar-wind driven flows in the magnetosphere-ionosphere system. *Ann Geophys* **10**: 103–115.

Davis TN, Sugiura M. 1966. Auroral electrojet activity index AE and its universal time variations. *J Geophys Res* **71(3)**: 785–801. DOI: 10.1029/JZ071i003p00785.

Finch ID. 2008. The use of geomagnetic activity observations in studies of solar wind- magnetosphere coupling and centennial solar change, *Ph.D. Thesis*, Southampton University, Southampton, UK.

Finch ID, Lockwood M. 2007. Solar wind-magnetosphere coupling functions on timescales of 1 day to 1 year. *Ann Geophys* **25**: 495–506. DOI: 10.5194/angeo-25-495-2007.

Finch ID, Lockwood M, Rouillard AP. 2008. The effects of solar wind magnetosphere coupling recorded at different geomagnetic latitudes: separation of directly-driven and storage/release systems. *Geophys Res Lett* **35**: L21105. DOI: 10.1029/2008GL035399.

Ieda A, Oyama S, Vanhamäki H, Fujii R, Nakamizo A, Amm O, Hori T, Takeda M, Ueno G, Yoshikawa A, Redmon RJ, Denig WF, Kamide Y, Nishitani N. 2014. Approximate forms of daytime ionospheric conductance. *J Geophys Res Space Phys* **119**: 10397–10415. DOI: 10.1002/2014JA020665.

Kappenman JG. 2005. An overview of the impulsive geomagnetic field disturbances and power grid, impacts associated with the violent Sun-Earth connection events of 29–31 October 2003 and a comparative evaluation with other contemporary storms. *Space Weather* **3**: S08C01. DOI: 10.1029/2004SW000128.

Kikuchi T, Lühr H, Schlegel K, Tachihara H, Shinohara M, Kitamura T-I. 2000. Penetration of auroral electric fields to the equator during a substorm. *J Geophys Res* **105(A10)**: 23251–23261. DOI: 10.1029/2000JA900016.

Knape J, de Valpine P. 2011. Effects of weather and climate on the dynamics of animal population time series. *Proc R Soc London B: Biol Sci* **278(18)**: 985–992. DOI: 10.1098/rspb.2010.1333.

Knutti R, Meehl GA, Allen MR, Stainforth DA. 2006. Constraining climate sensitivity from the seasonal cycle in surface temperature. *J Clim* **19(17)**: 4224–4233. DOI: 10.1175/JCLI3865.1.

Liou K, Newell PT, Sibeck DG, Meng C-I, Brittnacher M, Parks G. 2001. Observations of IMF and seasonal effects in the location of auroral substorm onset. *J Geophys Res* **106(A4)**: 5799–5810. DOI: 10.1029/2000ja003001.

Le Mouël J-L, Blanter E, Chulliat A, Shnirman M. 2004. On the semiannual and annual variations of geomagnetic activity and components. *Ann Geophys* **22**: 3583–3588. DOI: 10.5194/angeo-22-3583-2004.

Lockwood M. 2013. Reconstruction and prediction of variations in the open solar magnetic flux and interplanetary conditions. *Living Rev Solar Phys* **10**: 4. DOI: 10.12942/lrsp-2013-4.

Lockwood M, Cowley SWH, Freeman MP. 1990. The excitation of plasma convection in the high latitude ionosphere. *J Geophys Res* **95**: 7961–7971. DOI: 10.1029/JA095iA06p07961.

Lockwood M, Owens MJ, Barnard LA, Bentley S, Scott CJ, Watt CE. 2016. On the origins and timescales of geoeffective IMF. *Space Weather* **14(406)**: 432. DOI: 10.1002/2016SW001375.

Lockwood M, Chambodut A, Barnard LA, Owens MJ, Clarke E, et al. 2018a. A homogeneous aa index: 1. Secular variation. *J Space Weather Space Clim* **8**: A53. DOI: 10.1051/swsc/2018038.

Lockwood M, Finch ID, Chambodut A, Barnard LA, Owens MJ, Clarke E. 2018b. A homogeneous aa index: 2. Hemispheric asymmetries and the equinoctial variation. *J Space Weather Space Clim* **8**: A58. Doi: 10.1051/swsc/2018044.

Lockwood M, Bentley S, Owens MJ, Barnard LA, Scott CJ, Watt CE, Allanson O. 2019a. The development of a space climatology: 1. Solar-wind magnetosphere coupling as a function of timescale and the effect of data gaps. *Space Weather* **17**: 133–156. DOI: 10.1029/2018SW001856.

Lockwood M, Bentley S, Owens MJ, Barnard LA, Scott CJ, Watt CE, Allanson O, Freeman MP. 2019b. The development of a space climatology: The distribution of power input into the magnetosphere on a 3-hourly timescale. *Space Weather* **17**: 157–179. DOI: 10.1029/2018SW002016.

Lockwood M, Bentley S, Owens MJ, Barnard LA, Scott CJ, Watt CE, Allanson O, Freeman MP. 2019c. The development of a space climatology: 3. The evolution of distributions of space weather parameters with timescale. *Space Weather* **17**: 180–209. DOI: 10.1029/2018SW002017.

Mayaud P-N. 1971. Une mesure planétaire d'activité magnetique, basée sur deux observatoires antipodaux. *Ann Geophys* **27**: 67–70.

Mayaud P-N. 1972. The *aa* indices: A 100-year series characterizing the magnetic activity. *J Geophys Res* **77**: 6870–6874. DOI: 10.1029/JA077i034p06870.

Mayaud P-N. 1980. *Derivation, meaning and use of geomagnetic indices*, Geophysical monograph, vol. **22**, American Geophysical Union, Washington, DC. DOI: 10.1029/GM022.

Meng X-I, Rosenthal R, Rubin DB. 1992. Comparing correlated correlation coefficients. *Psychol Bull* **111(1)**: 172–175. DOI: 10.1037//0033-2909.111.1.172.

Menvielle M, Berthelier A. 1991. The K-derived planetary indices: Description and availability. *Rev Geophys* **29(3)**: 415–432. DOI: 10.1029/91RG00994.

Milan SE, Gosling JS, Hubert B. 2012. Relationship between interplanetary parameters and the magnetopause reconnection rate quantified from observations of the expanding polar cap. *J Geophys Res* **117**: A03226. DOI: 10.1029/2011JA017082.

Newell PT, Gjerloev JW. 2011. Evaluation of SuperMAG auroral electrojet indices as indicators of substorms and auroral power. *J Geophys Res* **116**: A12211. DOI: 10.1029/2011JA016779.

Riddick JC, Stuart WF. 1984. The generation of K indices from digitally recorded magnetic data. *Geophys Surv* **6(3/4)**: 439–456. DOI: 10.1007/BF01465559.

Rostoker G. 1972. Geomagnetic indices. *Rev Geophys* **10(4)**: 935–950. DOI:10.1029/RG010i004p00935.

Russell CT, McPherron RL. 1973. Semiannual variation of geomagnetic activity. *J Geophys Res* **78**: 82–108. DOI: 10.1029/JA078i001p00092.

Saba MMF, Gonzalez WD, clua de Gonzalez AL. 1997. Relationships between the AE, ap and Dst indices near solar minimum (1974) and at solar maximum (1979). *Ann Geophys* **15**: 1265. DOI: 10.1007/s00585-997-1265-x.

Shiokawa K, Ogawa T, Kamide Y. 2005. Low-latitude auroras observed in Japan: 1999–2004. *J Geophys Res* **110**: A05202. DOI: 10.1029/2004JA010706.

Thébault E, Finlay CC, Beggan CD, Alken P, Aubert J, et al. 2015. International geomagnetic reference field: the 12th generation. *Earth Planets Space* **67**: 79. DOI: 10.1186/s40623-015-0228-9.

Vasyliunas VM, Kan JR, Siscoe GL, Akasofu S-I. 1982. Scaling relations governing magnetospheric energy transfer. *Planet Space Sci* **30(4)**: 359–365. DOI: 10.1016/0032-0633(82)90041-1.

Wang H, Lühr H, Ma SY, Frey HU. 2007. Interhemispheric comparison of average substorm onset locations: evidence for deviation from conjugacy. *Ann Geophys* **25**: 989–999. DOI: 10.5194/angeo-25-989-2007.

Appendix A

Stations used to compile range-based global indices.

Table A1. *am* index magnetometer stations.

Name	IAGA code	Start date	End date	Latitude (°N)	Longitude (°E)	Sector	Notes
Magadan	MGD	1966	–	60.117	150.017	1	
Memambetsu	MMB	1958	–	43.910	144.189	1	
Paratunka/Petropavlovsk	PET	1969	–	52.971	158.248	1	
Bor (Podkamenaya Tunguska)	POD	1969	–	61.600	90.010	2	
Tomsk	TMK	1958	1970	56.470	84.930	2	
Vysokaya Dubrava	SVD	1958	1972	56.733	61.067	2	Replaced by ARS
Arti	ARS	1973	–	56.433	58.567	2	
Irkutsk	IRT	1958	–	52.167	104.450	2	
Novosibirsk (Klyuchi)	NVS	1967	–	54.850	83.230	2	
Chambon La Foret	CLF	1958	–	48.025	2.260	3	
Hartland	HAD	1958	–	50.995	355.516	3	
Witteveen	WIT	1958	1988	52.813	6.668	3	Replaced by NGK
Niemegk	NGK	1989	–	52.072	12.675	3	
Fredericksburg	FRD	1958	–	38.210	282.633	4	
Ottawa	OTT	1968	–	45.403	284.448	4	
Newport	NEW	1966	–	48.267	242.883	5	
Tucson	TUC	1958	–	32.170	249.270	5	
Victoria	VIC	1958	–	48.520	236.580	5	
Toolangi	TOO	1958	1979	−37.533	145.467	6	Replaced by CNB
Canberra	CNB	1980	–	−35.315	149.363	6	
Lauder	LAU	1979	1985	−45.033	169.683	6	
Amberley	AML	1958	1977	−43.152	172.722	6	Replaced by EYR
Eyrewell	EYR	1978	–	−43.423	172.355	6	
Martin de Vivies (Amsterdam Isl.)	AMS	1981	–	−37.800	77.570	7	
Gnangara	GNA	1958	2013	−31.780	115.947	7	Replaced by GNG
Gingin	GNG	2011	–	−31.356	115.715	7	
Port Alfred	CZT	1974	–	−46.431	51.867	8	
Hermanus	HER	1958	–	−34.425	19.225	8	
Port-Aux-Francais	PAF	1958	–	−49.353	70.262	8	
Faraday Islands (Argentine Isl.)	AIA	1958	–	−65.245	295.742	9	
South Georgia (Grytviken)	SGE	1983	1992	−54.282	323.507	9	
Port Stanley	PST	1994	-	−51.704	302.107	9	
Trelew	TRW	1958	-	−43.267	294.617	9	

Table A2. *ap* (*kp*) index magnetometer stations.

Name	IAGA code	Start date	End date	Latitude (°N)	Longitude (°E)	Sector	Notes
Lerwick	LER	1932	–	60.138	358.817	European	
Meanook	MEA	1932	–	54.616	246.653	N. American	
Sitka	SIT	1932	–	57.076	224.670	N. American	
Eskdalemuir	ESK	1932	–	55.314	356.794	European	
Lovö	LOV	1954	–	59.344	17.824	European	
Brorfelde	BFE	1932	–	55.625	11.672	European	
Agincourt	AGN	1932	1969	43.783	280.733	N. American	Replaced by OTT
Ottawa	OTT	1968	–	45.403	284.448	N. American	
Rude Skov	RSV	1932	1984	55.850	12.450	European	
Cheltenham	CLH	1932	1957	38.733	283.158	N. American	Replaced by FRD
Fredericksburg	FRD	1958	–	38.210	282.633	N. American	
Abinger	ABN	1932	1956	51.185	359.613	European	Replaced by HAD
Hartland	HAD	1957	–	50.995	355.516	European	
Witteveen	WIT	1932	1988	52.813	6.668	European	Replaced by NGK
Niemegk	NGK	1989	–	52.070	12.680	European	
Wingst	WNG	1938	–	53.743	9.073	European	
Toolangi	TOO	1958	1979	−37.533	145.467	Australasian	Replaced by CNB
Canberra	CNB	1980	–	−35.315	149.363	Australasian	
Amberley	AML	1958	1977	−43.152	172.722	Australasian	Replaced by EYR
Eyrewell	EYR	1978	–	−43.423	172.355	Australasian	

Table A3. *aa* index magnetometer stations.

Name	IAGA code	Start date	End date	Latitude (°N)	Longitude (°E)	Sector	Notes
Greenwich	GRW	1868	1925	51.477	0.000	N	Replaced by ABN
Abinger	ABN	1926	1956	51.185	359.613	N	Replaced by HAD
Hartland	HAD	1957	–	50.995	355.516	N	
Melbourne	MEL	1868	1919	−37.830	144.975	S	Replaced by TOO
Toolangi	TOO	1958	1979	−37.533	145.467	S	Replaced by CNB
Canberra	CNB	1980	–	−35.315	149.363	S	

Cite this article as: Lockwood M, Chambodut A, Finch ID, Barnard LA, Owens MJ, et al. 2019. Time-of-day/time-of-year response functions of planetary geomagnetic indices. *J. Space Weather Space Clim.* **9**, A20.