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

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Lockwood, M., Chambodut, A., Finch, I. D., Barnard, L. A., Owens, M. J. and Haines, C. (2019) Time-of-day/time-of-year response functions of planetary geomagnetic indices. Journal of Space Weather and Space Climate, 9. A20. ISSN 2115-7251 doi: https://doi.org/10.1051/swsc/2019017 Available at https://centaur.reading.ac.uk/83659/

It is advisable to refer to the publisher’s version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1051/swsc/2019017

Publisher: EDP Sciences

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

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

Mike Lockwood¹,* , Aude Chambodut², Ivan D. Finch³, Luke A. Barnard¹, Mathew J. Owens¹, and Carl Haines¹

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, \( V = 100 \left( \frac{\sigma_a}{S} \right) \), where \( S \) is the index’s sensitivity at that activity level and \( \sigma_a \) is the standard deviation of \( S \). Both \( S \) and \( \sigma_s \) 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 \( V \) for the two-station aa index, 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 \( V_{av} \approx 11.20\% \) for ap. The same empirical test applied to the classical aa index and the new “homogenous” aa index, \( aa_{11} \) (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 $V_{St}$ 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 $V_{St}$ 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 $aah$ 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 $f = (am)_{i}/(ap)_{j}$, 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 $aah$ 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 $aah$. 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 ($K$ northward
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 hemisphere, respectively, for each hour, and these are multiplied together to give the $aa$ index.

1.2 The $aa$ index

The $aa$ index is defined as the product of the northern and southern hemisphere $aa$ indices:

$$aa = aa_N \times aa_S$$

where $aa_N$ and $aa_S$ are the northern and southern hemisphere $aa$ indices, respectively.

The $aa$ index is calculated for each hour and is used to calculate the $Kp$ index. The $Kp$ index is a measure of geomagnetic activity and is used to predict the impact of geomagnetic storms on satellite operations. The $Kp$ index is calculated by taking the average of the northern and southern hemisphere $aa$ indices:

$$Kp = \frac{aa_N + aa_S}{2}$$

The $Kp$ index is expressed as a whole number between 0 and 9, with 0 being calm and 9 being extremely active. The $Kp$ index is used to predict the impact of geomagnetic storms on satellite operations and is used to issue geomagnetic storm warnings.
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 $L$. $\Delta H_N$ or $\gamma$ 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         | $0 \leq \Delta H_N < 0.01L$ | 2.3 |
| 1         | $0.01L \leq \Delta H_N < 0.02L$ | 7.3 |
| 2         | $0.02L \leq \Delta H_N < 0.04L$ | 15 |
| 3         | $0.04L \leq \Delta H_N < 0.08L$ | 30 |
| 4         | $0.08L \leq \Delta H_N < 0.14L$ | 55 |
| 5         | $0.14L \leq \Delta H_N < 0.24L$ | 95 |
| 6         | $0.24L \leq \Delta H_N < 0.40L$ | 160 |
| 7         | $0.40L \leq \Delta H_N < 0.66L$ | 265 |
| 8         | $0.66L \leq \Delta H_N < L$ | 415 |
| 9         | $\Delta H_N \geq L$ | 667 |

The $ap$ index is currently made using $K$ index data from 11 northern and two southern hemisphere stations between corrected geomagnetic latitudes $[\text{ACcl}]$ 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 $Kp$ index is the arithmetic mean of the 3-hour standardized $K_S$ values for the observatories employed. The 3-hourly $Kp$ 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 $K/S$ 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 $[Ap_C^*]_{\text{max}}$) 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 & Sugiuira, 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 deficiencies 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\(_{\text{max}}\)) and the minimum SML value (SML\(_{\text{min}}\)) 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 \( (\langle \text{SME} \rangle_{3h} \text{ and } \langle \text{SML} \rangle_{3h} ) \), although not quite as strong as for SME\(_{\text{max}}\) and SML\(_{\text{min}}\). 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.

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\(_{\text{max}}\) and SML\(_{\text{min}}\), respectively.

| Interval (h) | am | ap | aa |
|-------------|----|----|----|
| SME\(_{\text{max}}\) | 3  | 0.840 | 0.806 | 0.832 |
| \( \langle \text{SME} \rangle_{3h} \) | 3  | 0.784 | 0.741 | 0.772 |
| AE\(_{\text{max}}\) | 3  | 0.824 | 0.767 | 0.815 |
| \( \langle \text{AE} \rangle_{3h} \) | 3  | 0.784 | 0.741 | 0.772 |
| SML\(_{\text{min}}\) | 3  | -0.823 | -0.777 | -0.807 |
| \( \langle \text{SML} \rangle_{3h} \) | 3  | -0.792 | -0.748 | -0.779 |
| AL\(_{\text{min}}\) | 3  | -0.822 | -0.766 | -0.810 |
| \( \langle \text{AL} \rangle_{3h} \) | 3  | -0.792 | -0.748 | -0.779 |
| Am\(*\) | 24 | 0.908 | 0.881 | 0.913 |
| Ap\(*\) | 24 | 0.866 | 0.826 | 0.864 |
| Aa\(*\) | 24 | 0.891 | 0.853 | 0.901 |
| SME\(_{\text{max}}^*\) | 24 | 0.856 | 0.826 | 0.864 |
| \( \langle \text{SME} \rangle_{24h} \) | 24 | -0.890 | -0.867 | -0.867 |
| AE\(_{\text{max}}^*\) | 24 | -0.848 | -0.847 | -0.848 |
| \( \langle \text{AE} \rangle_{24h} \) | 24 | -0.900 | -0.865 | -0.865 |
| SML\(_{\text{min}}^*\) | 24 | -0.875 | -0.847 | -0.848 |
| \( \langle \text{SML} \rangle_{24h} \) | 24 | -0.884 | -0.847 | -0.848 |
| AL\(_{\text{min}}^*\) | 24 | -0.848 | -0.847 | -0.848 |
| \( \langle \text{AL} \rangle_{24h} \) | 24 | -0.900 | -0.865 | -0.865 |

Figure 3 illustrates examples of the relationships of (a) am, (b) ap, and (c) aa to SME\(_{\text{max}}\) for the years 1996–2017 (inclusive). The general features are the same in corresponding plots for \( \langle \text{SME} \rangle_{3h} \), SML\(_{\text{min}}\), \( \langle \text{SML} \rangle_{3h} \), AE\(_{\text{max}}\), \( \langle \text{AE} \rangle_{3h} \), AL\(_{\text{min}}\) or \( \langle \text{AL} \rangle_{3h} \). 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\(_{\text{max}}\) values from the fitted polynomial values, \( \Delta_{\text{RMS}} \), is 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 \( \Delta_{\text{RMS}}^2 / (n - m - 1) \) 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\(_{\text{max}} \leq 750 \text{nT} \)) and are plotted to gauge the deviations from linearity of the data for larger SME\(_{\text{max}}\). Figure 3 shows that all three mid-latitude range indices have a non-linear response, with values being overestimates, relative to SME\(_{\text{max}}\) 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, \( \Delta_{\text{RMS}} \)) 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\(*\).
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 to 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, \( P_x \), that has just a single free fit parameter, the coupling exponent, \( x \), 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. 3. Scatter plots for 1996–2017 (inclusive) of the mid-latitude range indices with the maximum of the SME index, SME\(_{\text{max}}\), seen in the same 3-hour intervals by the SuperMAG global magnetometer network. (a) (grey) 3-hourly SME\(_{\text{max}}\) values as a function of 3-hourly \( am \) and (orange) 24-hour running means of SME\(_{\text{max}}\) as a function of corresponding running means of \( am \), SME\(_{\text{max}}\). 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\(_{\text{max}}\) values from the fitted polynomial value for the corresponding \( am, \Delta_{\text{RMS}} \) is given, as is the correlation coefficient \( r \) between 3-hourly SME\(_{\text{max}}\) and \( am \) values. (b) The same as (a) for 3-hourly \( ap \) and its 24-hour running mean, SME\(_{\text{max}}\). (c) The same as (a) for 3-hourly \( aa \) and its 24-hour running mean, SME\(_{\text{max}}\). In each panel the cyan line is a linear fit to the 3-hourly values for SME\(_{\text{max}} < 750\) nT, and is plotted to gauge the deviation from linearity of the data for larger SME\(_{\text{max}}\).
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 $\alpha$ is an unknown coupling exponent. The value of $\alpha$ 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 all incident power could enter the magnetosphere if the orientation, but the exponents for each are all self-consistent, all being set by the single fit parameter $\alpha$ 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 $\alpha$ has been determined). Finch & Lockwood (2007) found that the optimum coupling exponent $\alpha$ used in generating $P_a$ depends on averaging timescale, a result that implied that there was at least one other mechanism at work and that $P_a$ was failing to capture all of the relevant physics. However, Lockwood et al. (2019a) have shown that this variation in $\alpha$ was an artefact caused by data gaps and that when steps are taken to minimize the effect of such data gaps, $\alpha$ 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, $P_a$ (normalized by dividing by its average for the whole interval, $\langle P_a\rangle_{all}$, 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 $P_a/\langle P_a\rangle_{all}$ 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, $\epsilon$, was also computed, giving $\epsilon/\langle Ap \rangle_{all} = 0.570$ for $Ap$, $\epsilon/\langle Aa \rangle_{all} = 0.401$ for $Aa$ and $\epsilon/\langle Am \rangle_{all} = 0.342$ 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

Fig. 4. Scatter plots of geomagnetic indices as a function of normalized power input into the magnetosphere computed from near-Earth solar wind observation, $P_a/\langle P_a\rangle_{all}$, 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 $P_a$ 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 $P_a$ 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 $\epsilon$ (as a ratio of the overall mean value of the index) are given in each panel.
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 \( aaq \) but the Meng-Z test gives that the \( p \)-value for the null hypothesis that these two correlations are the same is only \( 5 \times 10^{-2} \). Thus the improvement of \( aaq \) 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 \( ap, aaq \) (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 \( aaq \) and of 0.371 for \( ap \) and \( am \). Hence the difference between the correlations for \( ap \) and \( aaq \) 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 \( P_{20} / P_{z} \) being larger, whereas for northern hemisphere winter they are smaller. For \( Aa \) and \( Am \) these differences are much smaller. The annual and semiannual 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 semiannual variations in the mean values of the geomagnetic indices and their best fit of the corresponding variation in \( P_{z} \). The power input \( P_{z} \) shows a clear semi-annual variation with peaks at the solstices and 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 \( \sin^{2} (\theta / 2) \) term in \( P_{z} \) (where \( \theta \) 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 Russell-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 \( aaq \), Lockwood et al., 2018b) whereas

\[ P_{z} \]

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 \( P_{z} \) is \( aaq \), 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 \( aaq \) than for \( Am \).

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

\[
s = \frac{g}{I_S}
\]

(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 \( \omega \) to give a planetary index \( G \),

\[
G = \sum_{i=1}^{N} \frac{\omega_i g_i}{N} = \sum_{i=1}^{N} \frac{\omega_i I_S}{N} = I_S \sum_{i=1}^{N} \frac{\omega_i S_i}{N} = I_S S
\]

(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:

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

(3)

where

\[
T^* = 1.5 \sin \left( 2\pi \left[ F_1 - F \right] \right) - 0.5
\]

(4)

\( A \) and \( B \) are constants, \( \chi \) 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 \( O \), \( \Delta B_o \), 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 \( \Delta B_o \). 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 \( \chi \) 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, \( \chi \) is set to \((\pi/2)\); hence the coefficient \( A \) controls the extent to which the effect of dayside conductivity at a given \( \chi \) is enhanced over residual nightside values. Note that there are small changes to the precise formulation of Finch (2008), who used a \( \cos^{-0.7}(\chi) \) 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 \( \Delta B_o \), 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 Liu 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 (Czader & 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., Shikiwaka 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

$$a_k/am = s_k/S_{am}$$

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: $0 \leq am < 10$ nT, $10 \leq am < 20$ nT, $20 \leq am < 40$ nT, $30 \leq am < 50$ nT, $40 \leq am < 60$ nT, $50 \leq am < 90$ nT, $60 \leq am < 110$ nT, and $am \geq 70$ nT for which the years 1959–2017 give $N_k = 58183, 51083, 40894, 22691, 13157, 8302, 10869, and 6060$ 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 $S_{am}(UT, F) = 1$. 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 $s_k^i$). 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 $S_{am}(UT, F)$. Using these values in equation (5) gives revised (first iteration) values of $s_k(UT, F)$, values ($s_k^i$). 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, 0.7277, 0.2083, 0.2030, 0.2010, 0.2001, 0.2003, and 0.2000 for the eight $am$ bins (in order of increasing $am$) 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).

The right-hand columns of Figures 7 and 8 show the $UT-F$ plots of the observed ratios $am/am$ 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 $\chi$ (at $F = 0.5$ and 14 $UT$ for Hartland and $F = 0$ and 4 $UT$ for Canberra). This is reflected in the large values $A$ for the lowest activity levels. This peak becomes increasing 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...
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
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 UTs (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 ($S_{am}$ and $S_{min}$, respectively) as well as the mean $\langle S \rangle$ and standard deviation ($\sigma_S$) 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

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 $a_{CNB}/am = S_{CNB}/S_{am}$.
Table 3. Uniformity of average time-of-day/time-of-year response, $S_{av}(UT, F)$, of the various mid-latitude geomagnetic ranges.

| 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                                           |
| $aa_d$ | 3    | 1868       | Data   | 0.05 year   | 5.54    | 12.6                                            | −12.3                                          |
| $aa$  | 4    | 1932       | Data   | 0.05 year   | 10.62   | 40.8                                            | −17.0                                          |
| $ap$  | 5    | 1932       | Data   | 0.05 year   | 8.95    | 29.0                                            | −16.8                                          |

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

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

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

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

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 columns of Fig. 9, respectively) and in both hemispheres, there is a clear seasonal variation with enhanced index sensitivity in summer, $S_{am}$ being largest around $F = 0.5$ and $S_{an}$ 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. $S_{am}$ shows the converse behaviour, being similarly larger around 04 UT and lower around 18 UT. Because many features in the $am$ pattern are the converse of those in the $S_{an}$ 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 $S_{an}$ 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, $S_{min}$ 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
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 $Kp$ 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 $am \geq 40$ nT both when averaging the observed $\langle aa \rangle$ and $\langle am \rangle$ values and when calculating the model sensitivity of $am$, $S_{am}$.

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, $aa_{eq}$, 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 | Year | 10 \( \leq \ am < 20 \) nT | 60 \( \leq \ am < 110 \) nT |
|-------|------|--------------------------|--------------------------|
|       | \( S_{am} \) | Min | Max | Min | Max |
| \( am \) | 2014 | \( -2.59\% \)  | \( +2.84\% \) | \( -0.49\% \)  | \( +0.49\% \) |
| \( aa \) | 2010 | \( -16.7\% \) | \( +12.7\% \) | \( -16.1\% \) | \( +11.8\% \) |
| \( aa \) | 1970 | \( -18.5\% \) | \( +12.8\% \) | \( -17.7\% \) | \( +11.9\% \) |
| \( aa \) | 1930 | \( -18.4\% \) | \( +12.9\% \) | \( -18.6\% \) | \( +12.1\% \) |
| \( aa \) | 1890 | \( -18.5\% \) | \( +12.9\% \) | \( -18.9\% \) | \( +12.1\% \) |
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 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 $K_p/K$ 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

$$S_{ap}(UT, F) = \left(\frac{<ap>/<am>}{\langle UT,F\rangle}\right) \times S_{am}(UT, F) \quad (8)$$

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 $<ap>/<am>$ (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 $<(ap)/(am)>$ 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 $(ap)/(am)$ and the corresponding pattern of $(<ap>/<am>)S_{am}$ 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 $(ap)/(am)$). This has the advantage of making the analysis purely empirical (instead of the empirical-model mixture involved in $S_{am}$ 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.
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 ΔF = 0.05 wide (18.25 days). The black dots and line show means of am, the mauve dots and line means of η × 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 (ap/am) 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. 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 (ap)/(am) and (right column) S_{ap} = ((ap)/(am))S_{am}. Data are for 1959–2017 (inclusive) and the am ranges defined in Figure 6.
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 $S_{am}(UT, F)$ pattern gives percentage deviations of $11.2\%$, $25.4\%$ and $-21.6\%$ for $\sigma_S$, $S_{max}$ 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 $S_{aa_H} = (\langle aa_H\rangle/\langle am\rangle)_{UT,F} \times S_{am}(UT, F)$. 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 $S_{aa_H}(UT, F)$.

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 $Ap^*$ (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 $S_{aa_H}(UT, F)$ pattern gives percentage deviations of $11.2\%$, $25.4\%$ and $-21.6\%$ for $\sigma_S$, $S_{max}$ and $S_{min}$ respectively. Table 3 shows that the average pattern gives percent deviations of $5.54\%$, $+12.6\%$ and $-12.3\%$ for $\sigma_S$, $S_{max}$ 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...
considerably more than $N_{20}$ samples. The $ap = 22$ level contains 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 UTs of the $ap$ index separately. For each of these $UT$-$ap$ combinations we then compute the empirical correction factor $f_c(UT, F, ap) = (ap)/(am)$ 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. 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.
values. Figure 17 presents contour plots of daily $f_c(\text{ap, } F)$ values for the eight UT ranges. This allows us to turn every $\text{ap}$ value into a corrected value $\text{ap}_C = \text{ap} \times f_c(\text{UT, } F, \text{ap})$. Figure 18 compares the $\text{ap}_C$ and $\text{am}$ values, in the same format as Figures 13 and 15. Part (a) shows that there is very good agreement between $\text{ap}_C$ and $\text{am}$ and between $\text{Ap}^*$ and $\text{Am}^*$. Parts (b) and (c) shows that the average annual variations of $\text{ap}_C$ and $\text{am}$ are very similar and parts (d) and (e) even their UT variations are well matched. The mean $S_{\text{ap}_C}(\text{UT},F)$ pattern gives percentage deviations of 1.78%, 5.0%, and −4.6% for $\sigma_S$, $S_{\text{max}}$ and $S_{\text{min}}$, respectively. Table 3 shows this is a major improvement compared to $\text{ap}$.

The coefficients needed to implement the correction to $\text{ap}$ 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 $\text{ap}$ into $\text{ap}_C$.

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...
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 am 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 $\langle S_{ap} \rangle = 1$), 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 $aa_{w0}$ (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, $f_{ap}(F, UT)$, but this is one step removed from accepting them for the purposes of compiling
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 $f_2(ap, F, UT)$ 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 $\sigma_{S}$ (flattening the response function) are always in excess of three and that for extrema the improvement in $\Delta_{min}$ is always better than that in $\Delta_{max}$. 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 $Aq_{H}^*$ 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 $Aq_{H}^*$ 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.

### 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 | $< 8 \text{ nT}$ | $8 \leq am < 30 \text{ nT}$ | $\geq 30 \text{ nT}$ |
|-------|---------------|----------------|----------------|
|       | Lowest 20%    | Inter 20%      | Largest 20%    |
| $am$  | $\sigma_{S}$ (%) | Max (%) | Min (%) | $\sigma_{S}$ (%) | Max (%) | Min (%) |
| ap    | 10.18         | 20.2          | -22.8          | 12.21         | 29.10 | -17.40 |
| $ap_{C}$ | 3.07         | 8.10          | -9.40          | 2.27          | 5.52  | -5.73  |
| $aa$  | 12.74         | 39.80         | -20.80         | 11.44         | 40.20 | -20.02 |
| $aa_{H}$ | 2.37         | 8.11          | -5.80          | 1.41          | 4.94  | -3.05  |
| $ap_{C}$ | 3.32         | 2.49          | 2.43           | 9.34          | 5.27  | 3.04   |
| $aa_{H}$ | 5.38         | 4.91          | 3.59           | 8.11          | 9.96  | 6.56   |
| Ratios of values for uncorrected indices to those for corrected |

#### 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-day/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, \( ap \) 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://isgiunistra.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. Shikokawa; 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 Elfiyia Zesta; 210 Chain, PI K. Yumoto; SAMNET, PI Faridhem 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; ICESSAR; RAPIDMAG; PENGUIN; British Antarctic 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. Chamboudat; Data obtained in cooperation with Geoscience Australia, PI Marina Costelloe; SuperMAG, PI Jesper W. Gjerloev. SuperMAG data are available from http://supermag.jhuapl.edu/indexes/?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 Ecole 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 studentship 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, Vondra RR, et al. 2002. Instantaneous ionospheric global conductance maps during an isolated substorm. Ann Geophys 20(8): 1181–1191. DOI: 10.5194/angeo-201181-2002. 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 \( K \) 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.

Chamboudat A, Marchaudon A, Mervielle M, El-Lendami F, Lathuillere C. 2013. The \( K \)-derived MLT sector geomagnetic fields. J Geophys Res 118: 7605–7621. DOI: 10.1002/jgre.18760. 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.
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      |       |
| Vysoyka Dubrava             | SVD       | 1958       | 1972     | 56.733        | 61.067        | 2      | Replaced by ARS |
| Arvit                       | 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       | 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 |
| Eyrewill                    | EYR       | 1978       |          | –43.423       | 172.355       | 6      |       |
| Martin de Vivies (Amsterdam Isl.) | AMS | 1981     |          | –37.800       | 77.570        | 7      |       |
| Gwangara                    | GNA       | 1958       | 2013     | –31.780       | 115.947       | 7      | Replaced by GNG |
| Gingin                      | GNG       | 2011       |          | –31.356       | 115.915       | 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       | 294.817       | 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  |                        |
| Spero         | 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   | 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.