Universal time variations in the Magnetosphere and the effect of CME arrival time: analysis of the February 2022 event that led to the loss of Starlink satellites

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Lockwood, M. ORCID logoORCID: https://orcid.org/0000-0002-7397-2172, Owens, M. J. ORCID logoORCID: https://orcid.org/0000-0003-2061-2453 and Barnard, L. A. ORCID logoORCID: https://orcid.org/0000-0001-9876-4612 (2023) Universal time variations in the Magnetosphere and the effect of CME arrival time: analysis of the February 2022 event that led to the loss of Starlink satellites. Journal of Geophysical Research: Space Physics, 128 (3). e2022JA031177. ISSN 2169-9402 doi: https://doi.org/10.1029/2022JA031177 Available at https://centaur.reading.ac.uk/110984/

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/10.1029/2022JA031177

Publisher: American Geophysical Union
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
Central Archive at the University of Reading
Reading’s research outputs online
Universal Time Variations in the Magnetosphere and the Effect of CME Arrival Time: Analysis of the February 2022 Event that Led to the Loss of Starlink Satellites

M. Lockwood1, M. J. Owens1, and L. A. Barnard1

1Department of Meteorology, University of Reading, Reading, UK

Abstract  We study Universal Time (UT) variations in the magnetospheric response to Coronal Mass Ejection (CME) impacts, using the example of the two CMEs that led to the destruction of 38 out of 49 Starlink satellites in early February 2022. We employ the Expanding-Contracting Polar Cap model to analyze the variation in the size of the ionospheric polar caps and an eccentric dipole model of the geomagnetic field and thereby quantify the UT variations caused by the inductive effect of the diurnal motions of the geomagnetic poles in a “geocentric-solar” frame of reference (i.e., a frame with an X axis that points from the center of the Earth to the center of the Sun). The results show that use of a quasi-steady convection model predicts a similar global power deposition into the thermosphere as that inferred here, but does not give the same division of that power between the northern and southern hemispheres. We demonstrate that, through the combined effects of the Russell-McPherron dipole-tilt mechanism on solar-wind magnetosphere coupling and of the diurnal polar cap motions in a geocentric-solar frame, the power deposited varies significantly with the arrival UT of the CMEs at Earth. We also show that in the events of early February 2022, both CMEs arrived at almost the optimum UT to cause maximum thermospheric heating.

Plain Language Summary  We use a recent well-publicized space-weather event as an example of a previously-overlooked aspect of the behavior of near-Earth space. The event took place in early February 2022, when 38 out of 49 Starlink satellites burned up in Earth’s atmosphere because two Coronal Mass Ejections (CMEs) emitted from the Sun hit the Earth and had a larger heating effect on the upper atmosphere than expected. The new element that we introduce is the effect of the eccentricity of Earth's magnetic field which is reflected in the offset of the magnetic pole from the geographic pole being considerably greater in the southern hemisphere than in the northern hemisphere. This introduces a daily variation into the response of Earth's magnetosphere to a given solar wind disturbance and we show that the effect would have been less severe during the February 2022 event had the CMEs arrived either earlier or later than they did.

1. Introduction

1.1. The Events of 3–4 February 2022

The impact of two Coronal Mass Ejections (CMEs) on Earth’s magnetosphere on 3 and 4 February 2022 caused the loss of 38 of 49 recently-launched SpaceX Starlink satellites due to enhanced upper atmosphere density during the resulting geomagnetic storm (Dang et al., 2022; Fang et al., 2022; Hapgood et al., 2022; Kataoka et al., 2022; Tsurutani et al., 2022; Y. Zhang et al., 2022). The full financial cost of this space weather event is not known but has been estimated to be upward of $20m for the lost satellites and $30m for the wasted launch capacity. A surprising element of this event is that the causal geomagnetic disturbance was moderately large but not extreme. The global geomagnetic activity Kp index reached a value of 5+ after the impact of each CME, a value exceeded 3.5% of the time since production of the Kp index began in 1932 (Bartels, 1949; Bartels et al., 1939). The level meant that the event was classified as a minor storm according to the scale used by NOAA’s Space Weather Prediction Center (SWPC). The 3-hourly global am geomagnetic index (Mayaud, 1972) indicates a slightly rarer event, reaching 84 nT after both CMEs - a level that is exceeded 2% of the time since the am index data series began in January 1959.

On 3 February 2022, at 18:13 UTC, the Starlink satellites had been launched by a SpaceX Falcon 9 rocket from NASA’s Kennedy Space Center in Florida into an orbit with perigee of 210 km. This was the 36th in a series of such launches since May 2019. However the geomagnetic storm, which at the time of launch had just begun,
subsequently raised the density of the atmosphere, increasing the drag on the satellites and induced re-entry before the planned deployment of electric thrusters could raise all but 11 of them to an operational orbit at altitude near 500 km. Consequently, 38 of the 49 satellites burnt up in the atmosphere on 7 February. SpaceX has responded to the loss by changing its launch procedures: the subsequent Starlink launch on 21 February used a higher initial orbit at 300 km altitude, but carried only 46 instead of 49 satellites. It should be noted that SpaceX’s adoption of a low initial orbit is good and responsible strategy: if a satellite fails initial checks it can be readily de-orbited from such a low altitude and so does not add to the accumulation of space junk. However, it is a strategy that places the satellites at risk from space weather-driven changes in atmospheric drag, which is what happened to the ill-fated satellites launched on 3 February. The event clearly demonstrates one of several reasons why it is important to link responsible procedures to minimize space junk with space weather forecasting (Hapgood et al., 2022).

Several recent papers have studied the science behind this event. Fang et al. (2022) have looked at the CMEs causing the event, their propagation from Sun to Earth and their forecasting using the standard tools of a 3-dimensional MHD model of the heliosphere, based on solar magnetograph data, with a CME “cone model”, which is inserted based on coronograph images of the event eruptions. It appears from this analysis that forecasts used by the SpaceX launch team underestimated the scale of thermospheric heating, and consequent density rises at the initial satellite orbit, that the event caused. For example, the standard empirical NRLMSISE-00 model (Picone et al., 2002) predicts a rise of only about 5% in neutral density at 210 km. Limb observations during the event indicate that the actual rise at 210 km was between 11% and 18% on the dusk side of the low-latitude Earth and 40%–59% on the dawn side (Y. Zhang et al., 2022). Dang et al. (2022) predict somewhat smaller rises in thermospheric densities than some other studies but show that the integrated effect on the satellites was still enough to cause re-entry from 210 km altitude. On the other hand, Kataoka et al. (2022) suggest the heating effect was greater and more widespread and the study by Tsurutani et al. (2022) of orbit changes from other satellites, such as Swarm, suggests the thermospheric density changes were large and the Starlink satellite orbits did not decay as fast as this would predict because they had been partially lifted from their initial orbit.

The thermospheric modeling by Dang et al. (2022) used the National Center for Atmospheric Research (NCAR) Thermosphere Ionosphere Electrodynamics General Circulation Model (TIEGCM), a three-dimensional, self-consistent, physical model of the coupled ionosphere-thermosphere system (Richmond et al., 1992). For very quiet times particle precipitation seems to dominate magnetospheric energy deposition in the ionosphere-thermosphere system; however, for small storms, the Joule heating by the ionospheric E-region Pedersen currents associated with F-region convection (Kalafatoglu et al., 2018; X. Zhang et al., 2005) is slightly larger than that due to particle precipitation, and for large storms Joule heating is dominant (Hajra et al., 2022; Robinson & Zanetti, 2021; Wilson et al., 2006). To define the Joule heating input to the coupled thermosphere-ionosphere model, Dang et al. (2022) employed the empirical convection and Joule heating model of Weimer (2005) which uses the prevailing solar wind and IMF conditions and assumes a steady-state response of the magnetosphere with only directly-driven energy deposition.

1.2. Polar Cap Expansion and Contraction and Universal Time Effects

In this paper, we look at the magnetospheric behavior during the February 2022 events. In general, some of the energy extracted from the solar wind by the magnetosphere is directly deposited in the polar upper atmosphere by currents and precipitating particles, whereas a second component is stored in the geomagnetic tail and released after a delay (Baker et al., 1997; Blockx et al., 2009; Liu et al., 2018; Shukhtina et al., 2005). The energy is largely stored as magnetic energy of open flux in the tail and so the cycle of energy storage and release is reflected in the open magnetic flux, $F_{PC}$. However, $F_{PC}$ is not a perfect indicator of energy stored because the lobe magnetic energy density is proportional to the square of the magnetic field in the lobe, which can be increased in the near-Earth tail for a given $F_{PC}$ by the squeezing effect of enhanced solar wind dynamic pressure (Caan et al., 1973; Lockwood, McWilliams, et al., 2020; Lockwood, Owens, Barnard, Watt, et al., 2020). Nevertheless, because the ionospheric field $B_i$ is effectively constant, changes in the area of the region of ionospheric open flux (here termed the polar cap), $A_{PC} = F_{PC}/B_i$, indicate the energy storage in, and release from, the tail. The next section discusses the Expanding-Contracting Polar Cap (ECP) model and how the expansion and contraction of the polar cap relates to ionospheric F-region convection voltages (Cowley & Lockwood, 1992; Lockwood & Cowley, 2022), and the associated E-region Pedersen currents and hence energy dissipation in the ionosphere and
thermosphere by Joule heating. By using the ECPC model, in which the open flux rises and falls, we separate the energy deposition by the directly-driven system from that by the storage-release system, whereas if steady state is assumed (i.e., with constant $F_{pc}$) there is only directly-driven power dissipation and one is ignoring the existence of the storage-release system. There is not room here to review the literature on the relative importance of directly-driven and storage-release energy deposition in the ionosphere and thermosphere; however, we can say that both observations and global MHD modeling show that the storage-release system is certainly not negligible, is often dominant and cannot be ignored (Blockx et al., 2009; Liu et al., 2018; Shukhtina et al., 2005).

Another factor that we investigate in the present paper is the effect on cycles of energy storage and dissipation of motions of the magnetic poles in any “geocentric-solar” frame of reference, caused by Earth’s rotation (Lockwood et al., 2021). By geocentric-solar, we mean any frame with an X axis that points from the center of the Earth to the center of the Sun, such as GSE (Geocentric Solar Ecliptic), GSEQ (Geocentric Solar Equatorial) of GSM (Geocentric Solar Magnetospheric). These introduce Universal Time (UT) variations which are close to being in antiphase in the two hemispheres and which are larger in the southern hemisphere because the offset of the magnetic and rotational poles is greater. Lockwood and Milan (2023) have recently reviewed causes and observations of UT effects in the magnetosphere, as seen in averages from long-term datasets. These include global geomagnetic indices, auroral electrojet indices, partial ring current indices, transpolar voltage data and field-aligned current data from the AMPERE project, exploiting magnetometers on board the Iridium swarm of 70 satellites.

An often-discussed potential effect of pole motions on energy deposition is that of ionospheric conductivity variations. Enhanced conductivity, generated by solar EUV illumination, peaks when the polar cap is tipped toward the Sun whereas the pole-motion voltage effect peaks 6 hours earlier when the pole is tipping toward the Sun at its fastest rate. For the February period studied here, calculations of the mean conductivity in the northern polar cap and auroral ovals show almost no variation with UT as almost all of those regions is on the nightside (solar zenith angles, $\chi > 100^\circ$) all of the time. On the other hand, there is a considerable quasi-sinusoidal UT variation in conductivity for the southern polar cap and auroral oval as most of those regions is subjected to diurnal variations in zenith angles $\chi$ below 90° that induce major conductivity changes (Ridley et al., 2004).

However, there are a number of points that need to be considered about effects of variations in this solar-EUV-induced conductivity in the polar cap and auroral oval. Much of the energy dissipation during geomagnetic storms takes place in the auroral ovals, caused by the Pedersen currents that connect the Region 1 and Region 2 field-aligned currents and where conductivity is dominated by auroral precipitation rather than being generated by solar EUV (Carter et al., 2020). This greatly reduces the significance of the solar EUV generated conductivity to energy deposition during geomagnetic storms. In addition, UT effects are introduced into heating rates by the neutral wind velocity (Billett et al., 2018; Cai et al., 2014), but we note that regular diurnal effects due to neutral winds are less clear in disturbed times when the convection pattern is changing faster than the neutral wind can respond.

There is also a need to be consistent when evaluating the roles of ionospheric conductivity and flux transport, a need that is imposed by Maxwell’s equation $\nabla \cdot \mathbf{B} = 0$. The point being that this fundamental equation of electromagnetism (the nonexistence of magnetic monopoles) demands that the open flux in the two hemispheres must be identical because it is generated and lost by magnetic reconnections which affect both hemispheres (as opposed to reconnection between field in one lobe and a northward interplanetary field which causes a circulation in that lobe and polar cap but does not change $F_{pc}$). In addition, when averaged over sufficient time, the antisunward magnetic flux transport rate of open flux in both ionospheric polar caps (i.e., the transpolar voltages) must be the same as that of the parts of the open field lines that are in interplanetary space. The latter cannot be influenced by ionospheric conductivity because the antisunward flow there is supersonic and super-Alfvénic. The same is true for most of the tail magnetosheath (Li et al., 2020). Several numerical simulations show that increased polar cap conductivity reduces transpolar voltages (e.g., Borovsky et al., 2009). This is expected as field-perpendicular conductivities (both Hall and Pedersen) arise from collisions between ions and electrons and neutral atoms and ion-neutral collisions also give frictional drag on the motion of F-region plasma and frozen-in magnetic field (Ridley et al., 2004). As discussed by Tanaka (2007) and (for an isolated flux tube) by Southwood (1987), this is the “line-tying” concept introduced by Atkinson (1967) and Atkinson (1978) to explain the origin of field-aligned currents and how they transfer momentum and energy down into the ionosphere. Because the interplanetary segments of open field lines, outside the bow shock in the “Stern gap”, and indeed in the tail magnetosheath, are flowing supersonically and super-Alfvénically away from the ionospheric polar cap that they are connected to,
they can have no information about the state of the ionosphere and so are not influenced in any way by the slowing of their field line foot points. Hence the reduction in transpolar voltage associated with enhanced polar cap conductivity must give induction effects in the field of the relevant tail lobe between the ionosphere and the tail magnetopause and hence a rise in the energy stored in that field. This means that enhanced conductivity is really influencing the balance between energy stored in the tail (and later released) and energy directly deposited in the ionosphere. Some (we will define it to be a fraction $f_L$) of the energy stored by the inductive field changes that decouple the ionospheric transpolar voltage and the Stern gap voltage, may be in the tail and antisunward of the reconnection X-line in the cross-tail current sheet that closes open flux: this cross-tail energy will be lost to the near-Earth magnetosphere and the ionosphere/thermosphere and returned to the solar wind. On the other hand, the remainder (a fraction $1 - f_L$) of the energy stored by the inductive field changes will be sunward of the tail reconnection X-line and that stored energy is deposited by Joule and particle heating via the storage-release system and in the ECPC, associated with the nightside reconnection voltage.

Because the resistance to motion of open field lines is in the ionosphere, almost all of any induced field changes and extra energy storage will be in the near-Earth magnetosphere and $f_L$ will be very small. This means that if transpolar voltage in a given polar cap is reduced by enhanced polar cap conductivity (resulting in the associated directly-driven energy deposition in the ionosphere and thermosphere being reduced), after a delay (typically a substorm growth phase duration) the lost energy is deposited by the flows and associated Joule heating accompanying the enhanced nightside reconnection as part of the storage-release system. Hence “saturation” effects associated with enhanced ionospheric conductivity which limit the transpolar voltage and Joule dissipation (X. Zhang et al., 2005) do not cause a reduction in total flux transport seen in the ionosphere, but they do spread it out over a longer time interval via the storage/release system and this will have a corresponding effect on the variation in Joule heating. Hence if we were to include modulation of the directly-driven transpolar voltage by ionospheric EUV-generated conductivity effects, we would also need to modulate the storage-release system in an appropriate way which requires knowledge of the tail lobe field changes.

In theory, this can be achieved using a full global MHD model of the magnetosphere, but here we do not include EUV-induced conductivity effects on transpolar voltage in the summer (southern) hemisphere because to modulate the directly-driven system without including a matching modulation of the storage-release system would be an inconsistent analysis. In addition, we argue that for the dominant energy deposition in the auroral oval, conductivity induced by particle precipitation dominates over that due to solar EUV photoionization.

2. Theory and Methods

2.1. Polar Cap Expansion and Contraction

Two key parameters in the analysis presented in this paper are the dawn-dusk diameters of the ionospheric polar caps, $d_{PC}$, and the voltage placed across them by the solar wind flow, the transpolar voltage, $\Phi_{PC}$ (also referred to as the cross-cap potential drop and synonymous with the polar cap flux transport rate). We investigate the variation of both using the Expanding-Contracting Polar Cap (ECPC) model (Cowley & Lockwood, 1992; Lockwood & Cowley, 2022). Continuity of open flux is the core equation of the ECPC model and is equivalent to Faraday’s law (in integral form) applied to the open-closed field line boundary:

$$\frac{dF_{PC}}{dt} = \Phi_D - \Phi_N$$

where $F_{PC}$ is the magnetospheric open flux, $\Phi_D$ is the reconnection voltage in the dayside magnetopause (the rate of production of open flux) and $\Phi_N$ is the reconnection voltage in the nightside, cross-tail current sheet (the rate of loss of open flux).

We here adopt the major simplifying assumption that the ionospheric open-flux polar caps remain circular. This is certainly not valid all of the time, particularly for strongly and prolonged northward IMF when $F_{PC}$ is well below its average value of about 0.4 GWb. In such cases, a “horse-collar” auroral form is often seen, indicating a teardrop-shaped open flux region (Bower et al., 2022; Elphinstone et al., 1993; Hones et al., 1989; Imber et al., 2006; Milan et al., 2020; Wang et al., 2022). Nevertheless, the assumption of a circular polar cap has been successfully used many times with the ECPC model (e.g., Lockwood et al., 1990; Milan et al., 2017), including an analysis of a full year of data (Milan et al., 2021) and has two major advantages for the present study. Firstly the transpolar voltage is given by (Connor et al., 2014; Lockwood, 1991; Milan et al., 2017, 2021)
where $\Phi_v$ is the sum of the voltages induced by all viscous-like (non-reconnection) mechanisms. Secondly, this assumption allows us to relate the flux $F_{PC}$ and diameter $d_{PC}$ of the polar cap. We use Equation 4 of Milan et al. (2021), which they employ to derive the open flux $F_{PC}$ by integration of the vertical ionospheric field $B_i$. Inside the polar cap from the latitude of the region 1 field-aligned currents using a model geomagnetic field with an offset of the circular polar cap center from the magnetic pole toward the nightside of $4^\circ$. These authors find the values $F_{PC}$ are insensitive to this offset in the range $1^\circ$–$10^\circ$. We remove the latitudinal offset of $\Delta^\circ = 4^\circ$ between the open-closed boundary and the region 1 currents that Milan et al. employed and express the latitudinal radius in terms of the open flux polar cap diameter, an arc length, $d_{PC}$, at an altitude of 400 km:

$$F_{PC} = (3.259 \times 10^{-5})d_{PC}^2 + 23.53d_{PC} - (2.7 \times 10^7)$$

where $F_{PC}$ is in Wb and $d_{PC}$ is in m.

Note that subtle changes in the shape of the open polar cap are a key part of understanding the pattern of ionospheric convection, as predicted by the ECPC model (Cowley & Lockwood, 1992; Lockwood & Cowley, 2022; Lockwood et al., 1990; Lockwood, Lanchester, et al., 2006; Lockwood & Morley, 2004; Tulegenov et al., 2023), particularly in determining the pattern of flow following a burst in either the magnetopause or the tail reconnection voltages. However, the distortions to the boundary, and the flow patterns associated with them, propagate around the boundary (Morley & Lockwood, 2005). These transient features, and others associated with filamentary field-aligned currents caused by dynamic pressure pulse impacts on the magnetosphere (Lühr et al., 1996), are therefore not part of the present paper because of the simplifying assumption of a circular polar cap.

When using the ECPC it is important to understand the importance of the timescale $\tau$ over which the data are averaged. If a large $\tau$ is used, this is averaging over many cycles of expansion and contraction of the polar cap and $d_{PC}/d\tau$ tends to zero. From Equation 1 this steady-state condition means that $\langle \Phi_{pol} \rangle = \langle \Phi_{p} \rangle$. Only for small $\tau$ do we see the full expansion and contraction of the polar cap. Increasing $\tau$ causes the analysis to tend toward steady state. An important timescale in this is $\tau \sim 1$ hr, which is close to the average duration of the substorm cycle. In such cycles, although much of the open flux generated in the growth phase (by enhanced magnetopause reconnection voltage $\Phi_p$) can be lost in the subsequent expansion and recovery phases by enhanced reconnection voltage in the cross-tail current sheet, $\Phi_v$, one substorm cycle does not generally return the polar cap flux to the value it had at the start of the growth phase. It can take a string of weakening substorm expansions to achieve that (Lockwood & McWilliams, 2021b). Conversely, some substorms deposit more energy than was stored in the growth phase leaving the stored tail energy at a lower level that it was at the start of that growth phase (Baker et al., 1997). Furthermore there are steady convection intervals (Lockwood et al., 2009) and intervals of driven convection (Milan et al., 2021) which can last considerably longer than an hour. Later in this paper, we demonstrate that the first CME in the February 2022 event is an example of how large, persistent and increasing $\Phi_p$ prevents $\Phi_v$ establishing a steady-state, despite several substorm expansions, even on averaging timescales of 1–2 days. In general, the voltage $\Phi_p$ is constantly changing because of the variability in the solar wind parameters (Lockwood, 2022; Lockwood, Bentley, et al., 2019; Lockwood & Cowley, 2022). As a result, although steady state is a good approximation for $\tau$ of several days, we can still detect the effects of non-steady behavior at $\tau = 1$ hr, although they will be reduced in magnitude by the averaging.

### 2.2. Determination of the Magnetopause Reconnection Voltage, $\Phi_p$

Lockwood and McWilliams (2021b) used a dataset of 25 years of hourly-averaged data ($\tau = 1$ hr) to show that the optimum solar wind coupling function depends on which magnetospheric response index it is aimed at predicting. In particular, they showed that the coupling function that best predicts transpolar voltage $\Phi_{PC}$ is considerably different from those that best predict geomagnetic activity indices. Best practice in deriving these coupling functions was discussed by these authors and by Lockwood (2022). In particular, the results of regression and correlation analysis tends to be weighted toward the means of the distributions and the fit often underestimates the full range and extreme values of the observations (Lockwood, Rouillard, et al., 2006; Sivadas & Sibeck, 2022). The method used by Lockwood and McWilliams (2021b) and Lockwood (2022) avoids this by fitting to averages in bins that cover the full range, meaning that the weighting is equal across the whole range of the data and not dominated by the larger number of data points close to the mean. There are also pitfalls over
time resolution (Laundal et al., 2020; Lockwood, 2022). Figure 1 demonstrates the best-fit coupling function to the transpolar voltage observed by the SuperDARN coherent radar network and shows that another major pitfall, overfitting, has been avoided because the fit is essentially the same for the independent test half of the dataset as for the fit data subset that was used in the derivation of the coupling function. Overfitting is a problem that is particularly facilitated by the presence of datagaps in the interplanetary data which were a serious but neglected problem for coupling function studies using data from before the advent of the near-continuous data from the ACE and Wind interplanetary monitors in 1995. The best way to handle data gaps has been discussed by Finch and Lockwood (2007) and Lockwood, Bentley, et al. (2019). In both cases shown in Figure 1, an optimum lag of 20 min of $\Phi_{PC}$ behind the interplanetary coupling function is employed (see Figure 6 of Lockwood & McWilliams, 2021b).

The mauve line in Figure 1 is an average fit to many years of data, which means that steady-state applies to the fitted value so that, from Equation 2, $\langle \Phi_D \rangle = \langle \Phi_N \rangle = \langle \Phi_{PC} - \Phi_V \rangle$. The best-fit is given by:

$$\langle \Phi_{PC} \rangle = \langle \Phi_D \rangle + \langle \Phi_V \rangle = s_T \left[ B_1 \rho_{sw} V_{sw} \sin^d (\theta/2) \right] + i_T$$

where $B_1 = \left( B_1^2 + B_2^2 \right)^{1/2}$ is the IMF transverse to the $X$ direction of the GSM frame (in nT); $\rho_{sw}$ is the solar wind mass density (in kg cm$^{-3}$); $V_{sw}$ is the $X$-component of the solar wind velocity (in km s$^{-1}$); $\theta = \tan^{-1}(B_Y/B_Z)$ is the IMF clock angle in the GSM frame and $B_Y$ and $B_Z$ are the $Y$ and $Z$ IMF components in that frame. The best-fit constants are $a = 0.6554, b = 0.0522, c = 0.6676, d = 2.5, i_T = 13.5$ kV, and $s_T = 8.4075$ km s$^{-1}$ for the above parameter units and $\Phi_{PC}$ in kV. The above parameters give an root-mean-square error (RMSE) of $\Delta = 12.863$ kV for the fit data set (Figure 1a) and of $\Delta = 10.347$ kV for the test data set (Figure 1b). Given that the viscous-like voltage is not predicted to have the dependence on the IMF clock angle, we here take $\Phi_V$ to be equal to $i_T = 13.5$ kV, a value

Figure 1. Data density plots (2-dimensional histograms) of observed and predicted hourly means of transpolar voltage, $\Phi_{PC}$. The observations are from the survey of 25 years’ data from the northern hemisphere coherent-scatter SuperDARN radar network by Lockwood and McWilliams (2021b). The predicted values are made using the interplanetary data in the Omni2 dataset (King & Papitashvili, 2005), using the procedure described by Lockwood and McWilliams (2021a) to generate a coupling function for transpolar voltage (see Equation 4 of text). (a) is for the fit data subset (2012–2019, inclusive) and (b) for the independent test data subset (1995–2011, inclusive). In both panels the fraction of valid samples, $n/\Sigma n$, is color-coded on a logarithmic scale as a function of observed and predicted transpolar voltage, in bins of size 1 kV by 1 kV with $n \geq 3$ samples. The total number of samples, $\Sigma n$, the correlation coefficient, $r$, and the r.m.s. deviation of fitted from observed values, $\Delta$, are given. The mauve line are perfect agreement of observed and predicted $\Phi_{PC}$.
Equation 5 is derived from hourly means. We studied the relationship between the 1-min values of $\Phi_D$ computed using this equation and a variety of one-minute geomagnetic indices ($SML$, $SMU$, $AL$, $AU$, $SMR$, and $SYM-H$) at the optimum lag and compared it with that for 1-hr running means of the same data. The correlation is naturally lower for the 1-min data (e.g., for $-SML$ 0.66 as opposed to 0.80) in each case as the scatter is greater, but the best-fit polynomial is almost identical for the two time resolutions in all cases. Hence we here use Equation 5 to generate 1-min values of $\Phi_D$. We employ a lag of $\delta t = 5$ min between the interplanetary observations (which are propagated from L1 to the bow shock) and $\Phi_D$ that being the lag derived by Lockwood and McWilliams (2021b) between the hourly means of $\Phi_{PC}$ and running hourly means of the coupling function from IMF data when auroral electrojet activity was quiet ($-AL \leq 100$ nT).

A different analysis of the same dataset reveals that much of the scatter in Figure 1 is caused by a mechanism other than dayside magnetopause reconnection. The key point to note about Figure 2 is that the contours slope diagonally across the plot. This means that a given predicted $\Phi_D$ the transpolar voltage $\Phi_{PC}$ increases with increasing $-SML$. This is true in all regions of this parameter space except when both $\Phi_D$ and $-SML$ are very large when the contours become vertical showing that $\Phi_{PC}$ is then a function of $\Phi_D$ only. Thus even using hourly data we can detect an influence on transpolar voltage which depends on the auroral electrojet activity level. This is therefore a separate influence on the transpolar voltage, as predicted to be supplied by $\Phi_D$ in the ECPC model. The same picture emerges from the Active Magnetosphere and Planetary Electrodynamics Response Experiment (AMPERE) analysis of data on field-aligned currents from magnetometers on board more than 70 Iridium satellites in circular low-Earth orbit (altitude 780 km) in six orbit planes, which give 12 cuts at different MLTs in each orbit through the auroral oval (Anderson et al., 2014; Milan et al., 2015): from Chree analysis of these data, Milan et al. (2018) show that the field-aligned currents that bring convection circulation of the magnetosphere down to the ionosphere are enhanced in response both to a coupling function that quantifies dayside magnetopause reconnection and also to the $-AL$ index.

2.3. Determination of the Reconnection Voltage in the Cross-Tail Current Sheet, $\Phi_N$ The problem in applying the ECPC is that the nightside reconnection voltage $\Phi_N$ is hard to quantify (Walach et al., 2017; Øye, 2018). Several studies have used Equation 1 to infer it from the rate of change of open polar cap flux $dF_{PC}/dt$ and the value of $\Phi_D$ deduced from interplanetary parameters using a coupling function of the type given by Equation 5. For example, Grocott et al. (2002) used $dF_{PC}/dt$ deduced from the bite-out in the nightside polar cap in a substorm expansion phase observed by magnetometers, radars, and imagers. The same basis was used by Milan et al. (2007), Øye (2018), and Milan et al. (2021) using the circular polar cap assumption and looking at the change in radius of the polar cap inferred from global auroral images or the locations of field-aligned current sheets deduced from the AMPERE program. This method is not useful here where we wish to use $\Phi_N$ to predict the variation of $F_{PC}$ The way forward was first established by Holzer et al. (1986) who used the $AL$ nightside auroral electrojet index as a proxy for $\Phi_N$ to study two polar cap expansions and contractions using Equation 1. The validity of this approach has been confirmed by several studies of larger datasets, including the statistical survey of 25 years of data by Lockwood and McWilliams (2021b) who found that transpolar voltage
was a function of $\Phi_D$ with an independent influence related to $AL$, something we have demonstrated here with Figure 2 using the SuperMAG $SML$ index in place of $AL$. This confirms $SML$ can be used as a basis for the quantification of $\Phi_N$.

The mauve line in part (a) of Figure 3 is a polynomial fit to all the hourly means of $\Phi_{PC}$ from the survey of Lockwood and McWilliams (2021b) as a function of the simultaneous sigma hourly means of $\cdot SML$. There is considerable spread in the data which we expect for two known reasons: firstly $\Phi_D$ contributes to $\Phi_{PC}$ as well as $\Phi_N$, and secondly there are ionospheric conductivity and other seasonal effects in the relationship between $\Phi_{PC}$ and the geomagnetic $SML$ index. Part (b) of Figure 3 shows the equivalent plot for strongly northward IMF ($B_Z\geq 10 \text{nT}$) when we can take $\Phi_D$ to be zero. The mauve line is the same as in part (a) and remains a good fit to the average data, which covers a smaller range of both $\Phi_{PC}$ and $SML$, as expected for northward IMF. However, although the scatter is reduced because $\Phi_D$ is not a factor, it is still considerable.

To estimate the contribution to $\Phi_{PC}$ of processes associated with the substorm current wedge, we adopt the fitted form to $\Phi_{PC}$ as a function of $SML$ (given at the top of Figure 2) but scale it with a factor $k(F)$, that is a function of time of year, $F$, to allow for the seasonal effects in these northern hemisphere radar and magnetometer data. This gives us an estimate of the tail reconnection rate associated with auroral electrojet activity, as quantified by $SML$:

$$\left[\Phi_{N1} \text{in kV}\right] = k(F) \times f(SML) = k(F) \times \left[ -3.59y^4 + 134y^3 - 187y^2 + 15 \right]$$

where $y = [-\cdot SML \text{ in nT}]$.

However, the scatter in Figure 3b shows that $SML$ does not uniquely define $\Phi_N$. The same conclusion can be drawn from the survey of 1-year of data by Milan et al. (2021) in which the ratio $\PhiB/SML$ appears to differ, on average, for different phases of magnetospheric behavior. On the other hand, Milan et al. (2007) show that almost all $\Phi_N$ occurs soon after bursts in $\Phi_D$ and a base-level of $\Phi_D$ between events was not detected. However, we note that the viscous-like voltage $\Phi_D$ may well actually be due to low-level continuing $\Phi_D$ because ongoing unbalanced nightside reconnection has the ability to mimic all the ionospheric flows that have been attributed to viscous-like interaction.

Compared to the scatter in Figure 1, that in Figure 3 is large and so we here take steps to reduce it. Specifically, we add to the highly variable loss rate that is predicted by $SML$ alone, which we term $\Phi_{N2}$, a second quasi-steady
loss rate that depends only on the magnetic shear across the cross-tail current sheet (and hence related to the open flux \(F_{PC}\)), but which does not register in SML: we call this second loss rate \(\Phi_{N2}\), and is defined such that the total loss rate is

\[
\Phi_N = \Phi_{N1} + \Phi_{N2}
\]

(7)

We take this additional loss of \(F_{PC}\) to be linear with a time constant \(\tau_N\) so

\[
\Phi_{N2} = F_{PC}(\delta t)/\tau_N
\]

(8)

where \(\delta t\) is the lag to allow for the propagation of open flux into the tail after its generation (at rate \(\Phi_D\)). We here use \(\delta t = 15\) min, the derived optimum lag between the predicted \(\Phi_D\) and observed \(\Phi_{PC}\) in the study by (Lockwood & McWilliams, 2021a, —see their Figure 6). We repeated our analysis for \(\delta t = 10\) min and \(\delta t = 30\) min and found only small changes to our results. Note that there is a potential improvement we could make to Equation 8 in that the rate could be made a function of solar wind dynamic pressure \(p_{SW}\) and the dipole tilt angle \(\delta\) as there are observations that indicate that, together, they influence the rate of nightside reconnection by squeezing the tail (Lockwood et al., 1973; Finch et al., 2008; Hubert et al., 2009; Karlsson et al., 2000; Kokubun et al., 1977; Lockwood, 2013; Lockwood et al., 2021; Lockwood, McWilliams, et al., 2020; Lockwood et al., 2020; Schielle & Siscoe, 1970; Yue et al., 2010). This would not influence the equation for the voltage \(\Phi_N\) as any raise in -SML is accounted for; however, it could influence \(\Phi_{N1}\) if \(\Phi_N\) were increased by more than is predicted from the rise in -SML. However, variations in \(p_{SW}\) have also been proposed to influence \(\Phi_D\), field aligned currents and Joule heating (Boudouridis, 2005; Lee, 2004; Lukianova, 2003; Palmroth, 2004; Stauning & Troshichev, 2008); hence, it might be possible to include solar wind dynamic pressure effects in one area and not another. To be sure that we are not inconsistent in the present paper we do not include the effects of \(p_{SW}\) variations in the present analysis. However, we do highlight times when changes in \(p_{SW}\) may have had an effect.

The reason for proposing a second component of \(\Phi_N\) that does not depend on SML is that without it we found that during the events studied \(F_{PC}\) grows to either unrealistic values (exceeding 2.5 GWb) and/or falls to below zero for any \(k(F)\). This is because in the delays between enhanced \(\Phi_D\) and the SML response, positive (\(\Phi_D - \Phi_N\)) can be very large if \(\Phi_N\) is only based on SML and so \(F_{PC}\) can grow unrealistically large. Conversely there are intervals when \(\Phi_D\) has declined (i.e., the IMF has turned northward) but SML remains high and these can give large negative (\(\Phi_D - \Phi_N\)) and if they persist for long enough the estimated \(F_{PC}\) turns negative, which is unphysical. We found that, for the event studied here, quantifying \(\Phi_N\) from SML alone, there was no scaling factor which could prevent \(F_{PC}\) at certain times, becoming negative and from ever becoming unrealistically large (and in fact, usually both occurred).

To compute the scaling factor \(k(F)\), for the time of year \(F\) of the events (early February) we make use of the fact that averaged over a long enough period \(\tau\), the means \((\Phi_D)\) and \((\Phi_N)\) become equal and hence steady-state applies (and by Equation 2, the polar cap flux is constant). Choosing the interval \(\tau\) needed, however, is a compromise between two factors: if \(\tau\) is too large the seasonal variation between \(\Phi_N\) and SML becomes a factor but if is too short (less than a couple of days) then steady-state is not achieved.

In order to find an interval around the events of interest we searched for two times that were of order 1 week apart for which both the predicted \(\Phi_D\) (from Equation 1) and observed -SML were very low (below 1 kV and 20 nT, respectively) and followed intervals of at least 1 day when \(\Phi_D\) and -SML had remained low. Two such times around the events of interest were \(UT = 11\) hr of day-of-year (d-o-y) 38 (30 January 2022) and \(UT = 5\) hr on d-o-y 38 (7 February 2022), giving an interval of duration \(\tau\) of 7.75 days. At these times, images of the auroral oval (see below) gave similar and small values of \(F_{PC}\) of 0.267 ± 0.014 GWb. Therefore the difference \(\Delta F_{PC} = 0 ± 0.028\) GWb and the uncertainty in \(\Delta F_{PC}/\Delta t = \Delta F_{PC}/\tau\) is 41.8 Wbs⁻¹, in other words 0.042 kV. This is the level to which the averages of \(\Phi_D\) and \(\Phi_N\) can be taken as equal for this interval. We will refer to this as the “calibration interval”. From Equations 1 and 6–8

\[
\frac{dF_{PC}}{dt} = \Phi_D - k(F) \times f(S M L) - F_{PC}(\delta t)/\tau_N
\]

(9)

Equation 9 has two unknowns, \(k(F)\) and \(\tau_N\) - but for given values of these two parameters we can compute the variation of \(F_{PC}\) from the known variations of \(\Phi_D\) and SML. We start these computations for a wide range of initial values of \(F_{PC}\) at the start of doy 1 of 2022 and we find that this initial condition has no effect on the variation after d-o-y 30 (30 January 2022). For a given \(\tau_N\), we iterate \(k(F)\) until \(F_{PC}\) at the end of the calibration interval
equals that its start. This means that the integral of \( \Phi_D \) over the interval equals the integral of \( \Phi_N \) and steady state is achieved over this timescale.

If \( \tau_N \) is too small \( F_{\text{PC}} \) becomes negative. If \( \tau_N \) is too large we find \( F_{\text{PC}} \) reaches peaks larger than the largest values that have been detected, which are near 1 GWb (Boakes et al., 2009; Milan et al., 2021). We find that this \( F_{\text{PC}} \) is reached in both the peaks in the event studied here for \( \tau_N = 6.8 \text{ min} \) Section 4 provides evidence that it is an appropriate value for \( \tau_N \).

2.4. Universal Time Effect: Polar Cap Motions

A series of 4 papers Lockwood, Owens, Barnard, Haines, et al. (2020); Lockwood, McWilliams, et al. (2020); Lockwood, Owens, Barnard, Watt, et al. (2020); Lockwood et al. (2021) have investigated the semi-annual and Universal Time (UT) variations in the magnetosphere and in geomagnetic activity. The last of this series brings the conclusions together and successfully models the UT (and annual) variations of observed hemispheric geomagnetic indices by introducing a key component that had been hitherto overlooked. The interplanetary electric field is measured by spacecraft in geocentric-solar frames, such as GSM, GSE or GSEQ, that is, frames that are fixed in relation to the center of the Earth with an X-axis between the centers of the Earth and the Sun. Because Earth’s magnetic poles are offset from the rotational poles, the magnetic poles and the ionospheric polar caps move in the GSE and GSEQ frames in a diurnal circle generating almost circular loci (Lockwood et al., 2021). For GSM (for an eccentric dipole field but the GSE-to-GSM rotation angle determined using a geocentric dipole), the motions are ellipses but with the major axis in the X-direction of length equal to the diameter in the same direction of the loci for GSEQ and GSE. In this paper we are concerned with the component of motion in the X-direction which is the same for GSE, GSEQ and GSM frames.

Figure 4 illustrates these pole motions. Note that, by convention, the north/south pole of a magnet is where field lines diverge/converge and so, for the current polarity of the geomagnetic field, the magnetic pole in the southern hemisphere is, by that convention, a ‘north’ magnetic pole and vice-versa: in this paper, we refer to a magnetic pole by the hemisphere of the Earth that it is in and not by the magnet convention. At around 05 hr UT the north magnetic pole (i.e., the magnetic pole in the northern hemisphere) is pointed away from the Sun and the south magnetic field is pointed toward it, as shown in Figure 4a. At around 17 hr UT the north/south magnetic pole is pointed toward/away from the Sun (Figure 4c). Between these two times, around 11 hr UT the north/south magnetic pole is moving toward/away from the Sun (Figure 4b) and at around 23 hr UT the north/south magnetic pole is moving away from/toward the Sun (Figure 4d). The sunward component of the motion of the north/south magnetic pole at ionospheric F-region altitudes in any geocentric-solar frame is \( V_{\text{SP}} \) and \( V_{\text{SP}} \), respectively. These motions can be seen in the right-hand panels which show the pole loci at an altitude of 800 km with a dot denoting where the pole is at the UT in question. The sunward/antisunward velocities of the poles in the GSEQ frame are almost independent of the time-of-year and the larger offset of the pole in the Southern hemisphere from the rotational axis makes the radius of the near-circular orbits roughly twice as large as in the north and so the amplitude of the diurnal variation in \( V_{\text{SP}} \) is roughly twice that in \( V_{\text{NP}} \). Note also that the longitude separation for the eccentric field poles is smaller than the 180° that it would be for a geocentric field and so the variations in \( V_{\text{SP}} \) and \( V_{\text{SP}} \) are not in exact antiphase.

Both global images (Stubbs et al., 2005) and geomagnetic field modeling (Tsyganenko, 2019) of the auroral oval show that the polar cap moves with its magnetic pole with very little change in shape. There are a number of observations relevant to these diurnal motions of the polar cap caused by Earth dipole tilt. Newell and Meng (1989) surveyed 3 years’ data from the DMSF (Defense Meteorological Satellite Program) F7 satel- lite and showed that the region of solar wind precipitation in the cusp region migrated in geomagnetic latitude by about 0.06° for each 1° shift in dipole tilt angle. That means that (94%) of the motion of the magnetic pole in a geocentric-solar frame is reflected in the cusp location and only 6% in the geomagnetic frame. The key point is that the cusp boundary and the magnetic pole location are defined in the same geomagnetic latitude frame and the fact that their separation changes very little, means that their separation also changes very little when both are transformed into a geocentric-solar frame, no matter what geomagnetic latitude frame was used to measure their separation. The cusp precipitation is on newly-opened field lines generated by the voltage \( \Phi_p \) (see review by Smith & Lockwood, 1996) and hence this dipole-tilt induced motion of the dayside open-closed boundary (OCB) in a geocentric-solar frame mainly reflects that in the magnetic pole. Similarly on the nightside, Vorobjev and Yagodkina (2010) showed that the magnetic latitude of the poleward edge of the nightside northern-hemisphere
Figure 4.
auroral oval, as detected in DMSP satellite data from 1986, had a sinusoidal diurnal variation in amplitude near 2°, whereas the offset of the rotational northern eccentric axial pole at that time was about 8°. Hence in a geocentric-solar frame only about 75% of the motion in the magnetic pole is reflected in this boundary. However, this boundary is generally equatorward of the nightside OCB and this is likely to make this percentage a poor estimate of the real value that would apply to the nightside OCB. The magnetic field tracing needed by global MHD model of the magnetosphere means that the OCB can be identified in simulations and Kabin et al. (2004) found that magnetic latitude shifts in the noon OCB were 1.3° and −0.9° for dipole tilts of +35° and −35°, that is, 3.9% and 2.7%, respectively. The corresponding shifts in the midnight OCB were 0.8° and −0.5° (2.3% and 1.5%, respectively). Hence these simulations show the nightside OCB moves even more closely with the magnetic pole than the dayside OCB. The fact that the OCB is largely moving with the geomagnetic poles shows that closed field lines outside the open field line region are also taking part in this diurnal wobble caused by the pole motion. Oznovich et al. (1993) showed that during low auroral activity, the auroral oval as a whole was shifted by 1° in geomagnetic coordinates for every 10° change in the dipole tilt angle (90%). Being at large longitudinal separations (if not exactly the 180° for a geocentric dipole model) the motion of the auroral ovals induced by the magnetic pole motions would be close to, but not exactly, in antiphase in a geocentric-solar frame with the southern pole moving antisunward when the northern is moving sunward, and vice-versa. This was directly observed by (Stubbs et al., 2005), using full and simultaneous auroral images of the northern and southern auroral ovals made by the IMAGE and Polar satellites.

In the northern hemisphere the sunward speed of motion $V_{NP}$ is a sine wave of amplitude 57 ms$^{-1}$, but the larger offset of the southern magnetic pole from the rotational pole makes the corresponding amplitude of the speeds there, $V_{SP}$, about 135 ms$^{-1}$. These values were computed using the eccentric dipole field model of the geomagnetic field of Koochak and Fraser-Smith (2017) by Lockwood et al. (2021). This eccentric dipole model uses the first 8 coefficients of the IGRF model that define a spherical harmonic expansion of the magnetic scalar potential; this is compared to the first three used to define a centered dipole. This is a standard way of describing an eccentric dipole (Bartels, 1936). The speeds of these motions of the polar cap in a geocentric-solar frame are very much smaller than that of the solar wind. However, in the ionosphere the magnetic field $B$ is approximately $5 \times 10^{-5} T$, which is much greater than the flow-perpendicular field in interplanetary space (which is typically 5 nT). The diurnal motion toward and away from the Sun induces an electric field across the northern polar cap in a geocentric-solar frame of $V_{NP}B$ and a voltage in that frame across the polar cap of $\phi_{MN} = V_{NP}Bd_{PC}$ where $d_{PC}$ is the polar cap diameter. Using a typical polar cap angular radius of 15° gives $d_{PC} \approx 3.6 \times 10^{6} m$ and a sinusoidal diurnal voltage variation of amplitude 10 kV. In the southern hemisphere, the larger offset of the poles means that this amplitude is 24 kV. Another important factor is that the pole-motion voltages are applied consistently, rising and falling again sinusoidally over a 12-hr period, whereas rapid variations on IMF orientation mean that the voltages associated with solar wind flow are only consistently applied over intervals of several hours during rare, large CME events. Note also the importance here of the eccentric dipole model. Use of a geocentric dipole makes every effect on the North Pole motion equal and opposite to that of the south pole. Hence for a geocentric dipole, although Earth’s rotation alternately causes a given effect in one hemisphere and then the other, the global effect (the sum of the two) is always zero. This is not the case for an eccentric dipole.

These effects of pole motions, like the conductivity-induced changes discussed earlier, change the balance between directly-deposited energy and energy stored in the tail and then deposited via the storage-release system. Let us consider the Northern polar cap: when it is moving sunward (with a velocity $V_{NP} > 0$ that peaks at around 12 UT) the antisunward convection in the polar cap in a geocentric-solar frame will be reduced by the motion of the cap as a whole (i.e., by $V_{NP}$) and the voltage across the cap in that frame will be reduced by $\phi_{MN} = V_{NP}Bd_{PC}$ even if the reconnection voltages $\phi_p$ and $\phi_a$ are unchanged. The convection pattern perturbations for sunward/antisunward motions of the polar cap as a whole (and dawnward/duskward motions) were sketched by (Lockwood, 1991). This means the directly-deposited Joule heating is reduced. In a geocentric-solar frame, the flux transfer rate over the

Figure 4. The left-hand column gives schematics of pole motions in a geocentric-solar frame, based on the global MHD model predictions in Lockwood, Owens, Barnard, Watt, et al. (2020). The right hand panels show the loci of the axial poles, at an altitude of 800 km and over 24 hr, in the geocentric-solar GSEQ frame, predicted using the eccentric dipole model of Koochak and Fraser-Smith (2017) for the year 2018. The schematics show the magnetosphere in the noon-midnight (XZ) plane at (a) 05 hr UT, (b) 11 hr UT, (c) 17 hr UT, and (d) 23 hr UT at equinox. The mauve arrow shows the magnetic moment of an eccentric dipole. The dot-dashed lines are the magnetopause and the dashed lines the bow shock. Outside the bow shock the solar wind flows in the −X direction at speed $V_{SW}$. The pairs of panels in the right hand column give the pole loci in the GSEQ frame for the northern and southern hemispheres and four times of year: (green) March equinox, (mauve) June solstice, (orange) September equinox and (blue) December solstice. The colored dots on each locus shows the pole position at the UT of the corresponding schematic on the left-hand side.
northern ionospheric polar cap is reduced to \( \Phi_{PC} - \phi_{MN} \), but, given that the voltage across the region of open field lines in interplanetary space is unchanged, this means that flux is accumulating in the northern lobe at a rate that is \( \phi_{MN} \) greater than it would have done if \( V_{NP} \) had been zero. Conversely, in the other 12-hr phase of the diurnal cycle, the transpolar voltage is enhanced because \( V_{SP} \) and \( \phi_{MS} \) are negative and the rate of flux storage in the tail lobe is reduced. Hence the transpolar voltage in the northern polar cap is

\[
\Phi'_{NC} = \Phi_{PC} - \phi_{MN} = \Phi_{PC} - V_{NP} B_i d_{PC}
\]

and in the southern hemisphere

\[
\Phi'_{SC} = \Phi_{PC} - \phi_{MS} = \Phi_{PC} - V_{SP} B_i d_{PC}
\]

where \( V_{NP} \) and \( V_{SP} \) are the sunward components of motion of the northern and southern polar caps.

If the variations of \( V_{NP} \) and \( V_{SP} \) were of the same amplitude and in antiphase then although there would be more power deposited in the northern polar cap (and less stored in the northern tail lobe) for half the day, and then more in the southern hemisphere (and less stored in the southern tail lobe) for the other half so that the sum of the two would be constant. However these motions are not of equal amplitude, nor are they exactly in antiphase and this leaves a net UT variation in power deposited (and that stored in that tail) globally (Lockwood et al., 2021).

Note that the voltage imposed by the solar wind flow, \( \Phi_{PC} \), across the polar caps of the two hemispheres can be different at any one instant of time because of different inductive changes in the magnetic fields of the two tail lobes. However, on average they must be the same: by Faraday's law, a voltage is synonymous with a magnetic flux transfer rate and maintaining \( \nabla \cdot \mathbf{B} = 0 \) means that the long-term average of flux transport over the south pole must equal that over the north pole. In other words \( \langle \Phi_{PC} \rangle \) must be the same in the two hemispheres on average. A number of statistical studies of satellite data report differences in the long-term averages of the transpolar voltage in the two hemispheres (e.g., Förster & Haaland, 2015) which almost certainly reflect aliasing of orbit paths with seasonal, UT and activity level variations and not a violation of \( \nabla \cdot \mathbf{B} = 0 \).

3. The Event of 3–4 February 2022

Figure 5 shows the variations of various geomagnetic activity indices during the events of early February 2022. The gray bands mark the times of the CME passages past the Earth, as defined from interplanetary data by Dang et al. (2022). The top panels shows the 3-hourly planetary range indices, \( am, an \), and \( as \) with peaks near the ends of each of the CME events. Both peaks reach 86 nT in the global \( am \) index. The variations of the hemispheric sub-indices, \( an \), and \( as \), are very similar to that for \( am \). Panel (b) shows large enhancements in the nightside auroral electrojet index -SML at the times of the peaks in \( am \). The bottom panel shows the SMR and SYM-H indices compiled from low-latitude stations and mainly responding to the ring current. These show intensifications that peak shortly after the peaks in -SML but which decay more slowly. Note the data in panels (b) and (c) have been smoothed with 20-min running (boxcar) means to reveal the variations on timescales that can be compared to the \( am \) index and below are compared to satellite observations of the polar cap which are available, on average, about every 40 min.

Figure 6a gives the predicted dayside reconnection voltage \( \Phi_D \) computed from the interplanetary measurements: the other panels give the normalized variations of the terms from which it is computed. It can be seen that the mass density in each CME was slightly depressed and the transverse component of the IMF was enhanced, particularly during the first CME. The solar wind speed factor varied by about ±10% in the interval: it was increased by the CME1 passage, giving higher \( \Phi_D \) at these times. The first CME gives a particularly large peak in \( \Phi_D \) near its trailing edge, reaching 150 kV and exceeding 85 kV for 5.7 hr (6.1–11.8 hr UT on d-o-y 34). The second CME gives a more sustained period of somewhat smaller magnetopause reconnection voltage, being between 85 and 103 kV for most of an interval of duration 11.3 hr (9.4–20.7 hr UT on d-o-y 35).

4. Analysis of Polar Cap Behavior During the Event

Part (a) of Figure 7 shows the variation of \( \Phi_D \) computed using Equation 5 in red. In blue is the estimated variation of \( \Phi_N \), computed from the observed SML index variation, shown in part (b), using Equations 6–8. Part (d)
shows the variation in open flux $F_{PC}$ derived using Equation 9. These variations use the optimum combination of $\tau_N = 6.81$ min and $k(F) = 0.9972$. It was found that $\tau_N \leq 5.33$ min gave times when $F_{PC}$ became negative during the calibration interval and that $\tau_N \geq 10.52$ min gave times when $F_{PC}$ exceeded 1.2 GWb, which is larger than the values reported in the literature and then a proposed upper limit for a saturated polar cap. Substorm onsets are typically initiated when $F_{PC}$ reaches about 0.9 GWb (Milan et al., 2008). Indeed, Boakes et al. (2009) find that the probability of a substorm onset occurring is zero for $F_{PC}$ below about 0.3 GWb and increases linearly with $F_{PC}$ at higher values to near unity at 0.9 GWb, but the numbers of available events in their study mean the probability at and above this open flux are not well defined. However, larger values of $F_{PC}$, up to about 1.1 GWb, have been deduced in sawtooth events and steady convection events (Brambles et al., 2013; DeJong et al., 2007; Lockwood et al., 2009). It has been estimated that in large superstorms, $F_{PC}$ effectively saturates at 1.2 GWb (Mishin & Karavaev, 2017). In order to set the value of $\tau_N$ within this allowed range, we look at the modeled diameter of the polar cap in the calibration interval, as this is a strong function of $\tau_N$. The black line in Part (e) shows the variation in $d_{PC}$ for the variation of $F_{PC}$ shown in Part (d). This is obtained by PCHIP (Piecewise Cubic Hermite Interpolation Polynomial) interpolation of the variation of $F_{PC}$ with $d_{PC}$ given by Equation 3. PCHIP was used because it gives a monotonic variation without the erroneous overshoot that many splines can generate (it is continuous in both value and first derivative). Like linear interpolation, PCHIP ensures that the fitted polynomial passes through the data points and to check values between points are sensible, the analysis was repeated using linear interpolation and the results were not radically different. The green line in Figure 7d is from an empirical fit to DMSP ionospheric convection data given in Figure 5 of Hairston and Heelis (1990) (H&H90) which yields $d_{PC}$ as a function of transpolar voltage $\Phi_{PC}$ and which is here evaluated from $\Phi_{PC}$, $\Phi_N$, and $\Phi_V$, using Equation 2. It can be seen that the variation of the two estimates of $d_{PC}$ have similarities, but that the empirical model shows less variation than the one derived here. Figure 5 of H&H90 shows considerable scatter about the fitted line and so this difference is not unexpected. It is also worth noting that the average variation deduced from that plot is quite similar to the variation derived theoretically by Siscoe (1982). The value of $\tau_N = 6.81$ min adopted here makes
the averages of these two variations the same over the calibration interval. It yields an open flux $F_{PC} = 0.26$ GWb at the start and end of the calibration interval, which is lower than the average value near 0.4 GWb in the surveys by Boakes et al. (2009) and Milan et al. (2021). It also yields peak values of 1.01 GWb at the end of CME1 and 1.03 GWb at the end of CME2. Valuable confirmation of the value of $\tau_N$ comes from images of the northern and southern auroral ovals by the Special Sensor Ultraviolet Spectrographic Imager (SSUSI) on board the Defense Meteorological Satellite Program (DMSP) F-17 and F-18 satellites (Paxton et al., 2002, 2021). These images are in the Lyman-Birge Hopfield (LBH) short band (LBHS, 140–160 nm) of molecular nitrogen and the observed swathes usually show the full extent dawn-dusk diameter of the dark region poleward of the auroral oval. The poleward edge of the aurora seen in Far Ultraviolet (FUV) is often used as a proxy indicator of the OCB (Longden et al., 2010) and the six DMSP/SSUSU FUV images presented in Figure 2 of Y. Zhang et al. (2022) for the period studied here indicate considerable variability of the polar cap diameter.

In general, there is a difference between the latitude of the OCB, as identified in particle precipitation data and the poleward edge of the aurora, giving a dark ring of closed field lines poleward of the poleward edge of the aurora. The latitudinal width of this offset, $\delta\Lambda$ will, to some extent, depend on the imager, the magnetic local time (MLT) and the intensity of the auroral precipitation. Carbary et al. (2003) found that $\delta\Lambda$ did vary systematically with MLT but argued a constant value of 3.5° could be used for the purposes of computing $F_{PC}$. Boakes et al. (2008) found $\delta\Lambda \leq 1^\circ$, except in the predawn and evening sectors, where values up to 2° may apply. Longden et al. (2010) find that near dawn and dusk, the sectors of interest here, $\delta\Lambda \approx 1^\circ$ applies and that is what we adopt here.

There is an interesting minimum in the modeled $F_{PC}$ at $UT$ of 9.5 hr on d-o-y 33 which is as low as 0.05 GWb. This followed an interval of duration 4 hr of strongly northward IMF when $\Phi_D$ was essentially zero. This is

---

**Figure 6.** Variations of 1-min integrations of solar wind parameters during the interval shown in Figure 5. Panel (a) shows the derived magnetopause reconnection voltage $\Phi_D$, given by Equation 5, and the panels beneath show the component terms, normalized to their mean value in the interval. (b) $t_1 = B^a_t$ where $B_t = (B^a_t + B^b_t)^{1/2}$ is the IMF transverse to the X direction of the Geocentric Solar Magnetospheric (GSM) frame in which the $Y$ and $Z$ components are $B_p$ and $B_{sw}$, (c) $t_2 = \rho_{sw}$ where $\rho_{sw}$ is the solar wind mass density, (d) $t_3 = V_{\perp sw}$ where $V_m$ is the X-component of the solar wind velocity, (d) $t_4 = \sin^d(\theta/2)$ where $\theta = \tan^{-1}(|B_y/B_z|)$ is the IMF clock angle in the GSM frame. The best-fit exponents are $a = 0.6554$, $b = 0.0522$, $c = 0.6676$, $d = 2.5$. The vertical pink bands show the times of the CME passages. All parameters are lagged by the inferred propagation time to the dayside ionosphere of $\delta t = 5$ min.
slightly lower than the lowest reported value for their value for $\tau_N$ is slightly too low. Nevertheless, the $\tau_N$ used in Figure 7 does give a range of $F_{PC}$ values that matches the distribution previously reported in the literature. It is worth noting that panel (c) of Figure 7 shows that the solar wind dynamic pressure, $p_{SW}$ is raised above previous levels at this time and, as mentioned above, this could have enhanced $\Phi_N^2$, reducing $F_{PC}$ values.

Figure 7e gives the modeled dawn-dusk polar cap diameter $d_{PC}$ predicted using the assumption that the polar cap remains circular. We here used the dawn-dusk diameter deduced from DMSP/SSUSI images only as a rough check on that modeling. A major reason it can be of no greater significance than this is that, in reality, the polar cap is far from circular at some times. This is illustrated by Part (a) of Figure 8. This shows the aurora as seen by the F-17 satellite during a pass over the quiet, contracted polar cap in the northern hemisphere at 07:18 UT on d-o-y 33: this is during the descent toward the deep minimum in $F_{PC}$ at UT = 9.5 hr noted above. This pass reveals a horse-collar aurora with the putative OCB marked by narrow arcs that are almost parallel to the noon-midnight meridian (Wang et al., 2022). The polar cap is very far from circular at this time. Figure 8d shows a later pass at 01:56 UT on d-o-y 33, close to the time of the arrival of the first CME. The horse-collar form is still present and the dawn-dusk dimension of the dark polar cap gives a value of $d_{PC}$ that is shown by the orange point in Figure 8e: this is lower than the value predicted for a circular polar cap because of the horse-collar form is still present to some degree. In general, polar caps becomes more circular as the open flux increases. The other panels of Figure 8 give examples. Parts (c), (b) and (f) are northern hemisphere passes for, respectively, small, moderate and large open flux, in which the observed part of the polar cap is more circular in form. Part (e) is an
example of a southern hemisphere pass of F-17 over a relatively large polar cap. Southern hemisphere passes of both F-17 and F-18 tend to be toward the nightside. We use passes where the poleward edge of the aurora is detectable within 1 hr of the dawn-dusk MLT meridian and extrapolate the poleward boundary over up to 1 hr of MLT if it does not actually cross the meridian in the observed swathe. We apply $\delta \Lambda$ of 1° at both 18 and 06 MLT. The results are shown for pases of F-17 and F-18 over the northern and southern hemisphere polar caps by, respectively, red and blue points in Figure 8e. Cases where the polar cap is far from circular, such as in parts (a) and (d) are not used. The results show considerable scatter which is readily explained by the changes in shape of the polar cap, but do reveal the polar cap expansion and contraction during and after the CME impacts. They also provide confirmation that the value for $\tau_N$ used is appropriate.

The variations in the nightside voltage $\Phi_N$ and the polar cap flux $F_{PC}$ in Figure 7 follow that in the magnetopause voltage $\Phi_D$ in ways that we would expect. Toward the end of d-o-y 30 there is a sharp rise in $\Phi_D$, The high $\Phi_D$ persists for almost a day, declining only slightly. The mismatches in these voltages causes the polar cap flux $F_{PC}$ to rise to a peak of 0.68 GWb at 8.4 UT of d-o-y 31 and then fall back to its initial value near 0.26 GWb. This day-long event appears to be a period of driven convection and there is a slight rise in the -SML-related nightside reconnection voltage $\Phi_{N1}$ at its end. The next interesting feature is a sharp spike in $\Phi_{D}$ just before the end of d-o-y 32. This generates a response in -SML and hence $\Phi_{N1}$ and the rise in $F_{PC}$ is small and short-lived. This appears to be a small isolated substorm cycle in which -SML does not rise above 478 nT. The decay in $F_{PC}$ to the deep minimum discussed above occurs after this event: we see $\Phi_D$ and SML fall to essentially zero for almost a day and the exponential decay of open flux due to $\Phi_{N2}$ can clearly be seen. After this, a second, stronger isolated substorm cycle occurs in which -SML rises to 763 nT but, again, the rise in $F_{PC}$ is small and short-lived.

It is after this that CME1 arrives. $\Phi_D$ ramps up considerably, as does -SML and $\Phi_{N1}$ but with a lag and by a smaller degree. Hence the open flux grows to a peak of 1.01 GWb. The SML index shows a series of spikes that suggest substorms, but $\Phi_D$ is large and keeps increasing, causing it to always exceed $\Phi_N$ and so $F_{PC}$ keeps rising. Only

Figure 8. Selected sample auroral images recorded by the Special Sensor Ultraviolet Spectrographic Imagers (SSUSI) instruments on board the Defense Meteorological Satellite Program (DMSP) F-17 and F-18 satellites in the Lyman-Birge Hopfield (LBH) short wavelength band (LBHS, 140–160 nm). See text for details.
after the CME has passed and $\Phi_D$ declines sharply does $\Phi_N$ dominate and reduce the open flux. This rapid decline is caused by the loss rate $\Phi_{N1}$ but between enhancements in $-SML$ we see that $\Phi_{N2}$ also contributes. Again we note that in this interval between CME1 and CME2 Figure 7c shows considerably enhanced $p_{SW}$ and this may have contributed to the decline in $F_{PC}$ by enhancing $\Phi_{N2}$.

CME two is somewhat different. The rises in $\Phi_D$, $\Phi_N$ and $-SML$ are more modest than in CME1. Again spikes in $SML$ suggest a series of substorms. For this second event, the response in $\Phi_{N2}$, is slower and so $F_{PC}$ grows to levels that slightly exceed those attained in CME1 (1.04 GWb), even though the driving voltage $\Phi_D$ is not as large in this second event. Toward the end of the passage of CME2, $\Phi_D$ and $\Phi_N$ are approximately balanced and the peak open flux is maintained. After the passage of CME2, intermittent $\Phi_D$ means that the decline in $F_{PC}$ takes considerably longer and $-SML$ remains low so the slow decline in $F_{PC}$ is associated with $\Phi_{N2}$ more than $\Phi_{N1}$.

It is useful to look at the relative contributions of the two open flux loss rates $\Phi_{N1}$ and $\Phi_{N2}$. Figure 9 looks at the relationship of the two. Comparison of Parts (a) and (d) of Figure 9 shows that, in general, $\Phi_{N1}$ is greater than $\Phi_{N2}$. Both increase with increasing transpolar voltage $\Phi_{PC}$ and hence the sum of the two does the same (Figure 9b). Part (e) shows the ratio of the two, $\Phi_{N1}/\Phi_{N2}$, as a function of $\Phi_{PC}$, the mauve line is unity for this ratio and so $\Phi_{N1}$ always dominates for $\Phi_{PC} \geq 85$ kV. Figure 9f is the same ratio as a function of $-SML$ and shows that $\Phi_{N1}$ always dominates for $-SML \geq 27$ nT; however, the plot also shows great variability in this ratio which is the effect of the amount of open flux $F_{PC}$ on $\Phi_{N2}$. This is also seen in Figure 9c which plots the total loss rate as a function of $-SML$. This reveals the form of $\Phi_{N1}$ given by Equation 6, but that $\Phi_{N2}$ has indeed added scatter; however, the scatter in Figure 9c is not as great as in Figure 3 because it is for 7 days of 1-min data whereas Figure 3 is for 25 years of hourly data.

We conclude that the loss rate of open flux is largely dependent on the $SML$ auroral electrojet index and although the loss $\Phi_{N2}$ that is not captured by $SML$ is relatively small, it is still important for the application of the ECPC model because otherwise estimated open fluxes rise to levels that are not seen in the real magnetosphere.

It is interesting to note that Figure 7d shows that the open flux $F_{PC}$ between the CMEs fell to 0.20 GWb, just below the value at the start and end of the calibration interval. This eliminates preconditioning effects of CME1 on the response to CME2, at least in terms of residual open flux. In this context, we also note that Figure 5c shows...
that the ring current indices -SMR and -SYM-H fell back down after CME1 to only just above their values before the arrival of CME1. (Their decay after CME2 was noticeably slower.) This argues against major preconditioning in terms of ring current energy as well. There are other possible pre-conditioning effects such as the thickness of the plasma sheet and the speeds of polar thermospheric winds. The latter, however, would tend to reduce the heating effect of CME2 not increase it.

5. The Effect of Earth's Eccentric Magnetic Field

The previous section shows that the polar cap expanded and contracted during the events in early February 2022. In this section we add to the effects of this the UT variations caused by the diurnal rotation of Earth's eccentric magnetic field. The motions are computed using the eccentric dipole model of the geomagnetic field by Koochak and Fraser-Smith (2017), as discussed by Lockwood et al. (2021).

Figure 10 demonstrates the effect of polar cap motions. Panel (a) shows the transpolar voltage $\Phi_{PC}$ computed using Equation 2 from the variations of $\Phi_p$ and $\Phi_s$ (with constant $\Phi_i$) derived in the previous section. Part (b) shows the dawn-dusk voltages induced in the north and south polar cap, in a geocentric frame, by the diurnal cycle of sunward/antisunward polar cap motion, respectively $\Phi_{MN}$ and $\Phi_{MS}$. Note that for constant dawn-dusk polar cap diameter, these would be sine waves and the dashed lines give the variation for the mean value of the polar cap diameter over the calibration interval $\tau$, $\langle d_{PC} \rangle = 3.08$ Mm. (c) The variations in voltages in the GSE/GSM frames allowing for pole motions in the northern hemisphere, $\Phi'_{NC}$ (in red) and in the southern hemisphere, $\Phi'_{SC}$ (in blue). The black line is the average of the two ($\Phi'_{PC}$) and the gray area gives the uncorrected voltage, $\Phi_{PC}$, repeated from panel (a) to aid comparison.
because for the eccentric dipole field model used the magnetic poles are not 180° apart in longitude (as they would be for a geocentric dipole).

Figure 10c shows the average transpolar voltage allowing for the motions of both polar caps, computed using Equations 10 and 11 for the northern and southern hemisphere, $\Phi_{\text{NC}}$ and $\Phi_{\text{SC}}$, shown in red and blue, respectively. The black line is the average of the two, $\Phi_{\text{PC}}$ which is different from the transpolar voltage in the absence pole motions $\Phi_{\text{PC}}$ which is shown in Figure 10a and by the gray area in Figure 10c. If $\Phi$ in a hemisphere is smaller than $\Phi_{\text{PC}}$ it means that polar cap is moving toward the Sun. Let us apply Faraday’s law to a loop PCGS, fixed in a geocentric-solar frame, where P and C are the dawn and dusk flanks of the ionospheric polar cap which map along open geomagnetic field lines to the points G and S, respectively, just outside the bow shock in interplanetary space, often referred to as the “Stern Gap”, SG (see Lockwood & Cowley, 2022). The antisunward flow of the solar wind, with frozen-in open magnetic field, between S and G adds to the magnetic flux threading the loop PCGS and the sunward convection of frozen-in field in the F-region ionosphere between P and C removes flux from that loop. Hence if sunward polar cap bulk motion is slowing the rate that magnetic flux is transferred antisunward across PC in the ionosphere, it is reducing the rate at which flux removed from the loop. Hence this situation means that less energy is being directly deposited in the ionosphere but more magnetic energy is being stored in that hemisphere of the tail lobe. Conversely, if $\Phi'$ exceeds $\Phi_{\text{PC}}$, the polar cap is moving away from the Sun (i.e., with the solar wind) so that more energy is being directly deposited in ionosphere but less energy is being stored in that hemisphere of the tail lobe.

It is worth noting that some of these diurnal cycles may have been missed in some magnetometer observations of geomagnetic activity as they were attributed to the Sq variation. This is because to reveal geomagnetic activity, magnetometer data usually has subtracted from it a quiet diurnal variation to remove the effects of dynamo action, of particular, solar thermal tides and the equatorial electrojet that give the Sq variation (Yamazaki & Maute, 2017). The polar cap diameter $d_{\text{pc}}$ will be small in quiet times but not zero and so the pole motions may have added to the quiet day diurnal variation that is subtracted. This is most likely to be the case inside the polar cap and so a factor for the Polar Cap Indices (PCI). We searched for an effect of the different polar motions during this event in the published provisional Northern and Southern hemisphere PCI. Their variations are both very similar to the -SML index and so appear to be dominated by the auroral electrojet in this event. We note that some other studies have almost certainly detected signatures of the pole-motion effect but have generally attributed it to ionospheric conductivity effects.

### 6. The Importance of the Universal Time of CME Arrival

Owens et al. (2020) have discussed the value of accurate prediction of the arrival time of CME impacts on Earth’s magnetosphere. They make the point that if false-alarms are a serious problem, accurate arrival time information is only valuable if the geoeffectiveness of the CME can also be forecast. The analysis presented in this paper adds a further complication to that discussion in that the geoeffectiveness of a CME is shown to depend upon the Universal Time of the CME arrival.

To investigate the effect of pole motions in isolation, we here consider that the voltages $\Phi_D$ and $\Phi_N$ would not depend on the Universal Time of the CME impact. The phase of the diurnal cycles of pole motions have an influence on geoeffectiveness through the modification of $\Phi_{\text{PC}}$ to $\Phi_{\text{NC}}$, $\Phi_{\text{SC}}$ and $\Phi_{\text{PC}}$.

However, in reality, the dipole tilt (and hence Universal Time) will influence $\Phi_N$, through its effect on the magnetic shear at, and length of, the magnetopause reconnection X-line via the Russell-McPherron (R-M) effect (Lockwood et al., 2020; Russell & McPherron, 1973). In our synthesis of the effects of a delayed arrival of a CME, we allow for this by lagging (by a delay $\delta_t$) the variations in the factors $B_{\perp,t}$, $P_{\text{sw}}$, and $V_{\text{sw}}$ in Equation 5; however, the term $\sin^2(\theta/2)$ cannot be simply lagged in the same way. The reason is that the clock angle $\theta$ is computed in the GSM frame and because of the $UT$ variation in the rotation angle $\gamma$ between the GSE and GSM frames (caused by the dipole tilt variation with UT), the lagged values of IMF in the GSE frame, $|B_{Z,GSE}|$ and $|B_{Y,GSM}|$, have to be transformed for the delayed arrival at Earth.

$\theta = \tan^{-1}(|B_{Y,GSM}|/|B_{Z,GSM}|)$ and, hence, $\sin^2(\theta/2)$ and $\Phi_D$ for the delayed arrival at Earth.

In general, there will probably also be $UT$ effects on $\Phi_N$. A number of papers have discussed mechanisms by which the dipole tilt can influence tail structure and dynamics and so introduce $UT$ effects into reconnection responses in the tail (Danilov et al., 2013; Kivelson & Hughes, 1990; Kubyshkina et al., 2015; Lockwood et al., 2021;
However as these effects are less well established than the R-M effect, we here simplify by investigating the effect of dipole tilts on $\Phi_D$ only. We do this by assuming the form of the temporal variation in $\Phi_N$ response is the same as for the unlagged $\delta t = 0$ case, as shown in Figure 7. This allows us to simply lag the $\Phi_N$ variation by $\delta t$. We then scale these lagged values of $\Phi_N$ so that the integral over the calibration interval (which is also lagged by $\delta t$) is equal to that of the recalculated, lagged $\Phi_D$. This gives us the revised variation of $\Phi_{PC}$ for the $\delta t$ considered, from Equation 2. We can study the R-M effect in isolation from this $\Phi_{PC}$ variation but also transform it into $A\Phi'_{PC}$ (by evaluating $A\Phi'_{NC}$ and $A\Phi'_{SC}$ and averaging them) to study the combined effect of the R-M effect and the pole motions.

Figure 11a shows four variations associated with the CME events. These are: (a) the observed SYM-H geomagnetic index (in mauve); (b) the power deposited globally in the ionosphere and thermosphere by Joule heating, $P_i$, as modeled for these events by (Dang et al., 2022) using the empirical transpolar voltage and Joule heating model of (Weimer, 2005); (c) the square of the voltage derived here from the ECPC model but not allowing for polar cap motions $A\Phi'^2_{PC}$; and (d) the square of the transpolar voltage derived here when polar cap motions are included $A[\Phi'_{PC}]^2$ (in black). All four are normalized by their peak value toward the end of the passage of CME1 and all three are shown in the panel without any synthetic lag $\delta t$ being introduced other than for $P_i$ which has been lagged by 0.6 hr. (b) shows the variations of the square of the transpolar voltages in the northern and southern polar caps allowing for pole motions, $A[\Phi'_{NC}]^2$ and $A[\Phi'_{SC}]^2$, respectively. In (c) the interplanetary data sequences are lagged by $\delta t$ but the IMF orientation factor in GSM recomputed for the different UT of arrival in evaluating $\Phi_D$ (the Russell-McPherron effect, R-M) as well as the different phase of the diurnal cycles of pole motions (the pole motion effect). The response of the nightside reconnection voltage $\Phi_N$ is assumed to be the same in waveform (but lagged by $\delta t$) as that for $\delta t = 0$ but is scaled such that the integral over the calibration interval matches the integrated revised $\Phi_D$. For each lag $A[\Phi'_{NC}]$ and $A[\Phi'_{SC}]$ are computed. The plot shows the square of the average of the two, $A[\Phi'_{PC}]^2$ as a function of time for $\delta t$ which is varied between 0 (black line, as also shown in part a) and 12 hr (blue line).
accounts for propagation and response lags. We can see that, with this lag, $P_i$ agrees very well with the variations in $\Phi'^2_{PC}$ and $\left[\Phi'_{SC}\right]^2$ predicted here. $P_i$ was derived from the observed solar wind and IMF parameters using the statistical and empirical convection model of Weimer (2005). If we use a simple resistor analogy the power deposited is proportional to the square of the voltage and so the square of the transpolar voltage is a simple proxy metric of Joule heating rate and indeed Robinson and Zanetti (2021) do find a good statistical square-law relationship between the two for the 27 geomagnetic storm events that they studied (see their Figure 2d). Hence the use of a steady-state convection model, driven by the upstream solar wind and IMF, yields a very similar global heating variation to that inferred here using the ECPC model. Note that SYM-H is of relevance to magnetospheric energetics being related to the energy content of the ring current; however, that relationship is not a straightforward one (Robinson & Zanetti, 2021; Sandhu et al., 2021). Of interest in the time series studied here is the fact that peaks in Joule heating of the ionosphere after each CME are associated with the time of enhanced auroral activity as identified by the bursts of enhanced -SML and enhanced $\Phi_n$. However, a lower level of Joule heating does continue after -SML declines when SYM-H remains enhanced and the polar cap flux $F_{PC}$ is decaying slowly. This appears to be due to the “quiet” open flux loss that is not associated with SML and energy stored in the tail lobes is quietly (without auroral electrojet activity) deposited in both the ionosphere and ring current.

Figure 11a shows that the proxy for global heating rate from our analysis $\left[\Phi'_{PC}\right]^2$ is similar to the results of Dang et al. (2022) who employed the steady-state empirical model of transpolar voltage and Joule heating by Weimer (2005). However, Figure 11b shows there is a major difference. The Weimer model predicts the same transpolar voltage in the two polar caps and only some dipole tilt effects on conductivity would give any difference in Joule heating between the two hemispheres. We here predict that both the transpolar voltage and the Joule heating have strong Universal Time variations that are close to being in antiphase in the two polar caps. The precise behavior depends on the Universal Time of the storm-time rise in $\Phi_n$ and that is one reason why CME1 and CME2 have such different effects in these events. Figure 11b shows that during CME1 the two hemispheres show similar temporal variations in $\left[\Phi'_{NC}\right]^2$ and $\left[\Phi'_{SC}\right]^2$ but the southern hemisphere value is considerably larger. During CME2, $\left[\Phi'_{SC}\right]^2$ again dominates initially but for the peak at the end of the passage of CME2 the Joule heating is dominated by $\left[\Phi'_{SC}\right]^2$. Therefore, although the two methods predict very similar global Joule heating power, the temporal variations of the deposition into the two polar caps is more complex in our analysis. The behavior during CME2 is more complex than for CME1 because, in addition to the pole motions, the $\Phi_n$ rise is delayed after the causal rise in $\Phi_n$ (see Figure 7a).

The bottom panel of Figure 11 shows the combined effect of introducing synthetic lags $\delta t$ in arrival time at Earth in the range [0:2:12] hrs, via both the R-M and pole-motion effects. It shows the $\left[\Phi'_{PC}\right]^2$ variations for each value of $\delta t$, derived as described above.

Figure 11c demonstrates that the peak $\left[\Phi'_{PC}\right]^2$ at the end of CME1 would have been considerably smaller if the event had arrived some hours later. Figure 12 analyses the relative contribution of the R-M effects and the pole motions to the combined variation with arrival time shifts $\delta t$ between $-12$ hr and $+12$ hr ($\delta t = [-12:0.1:12]$ hrs). The event is here taken to be between the start of d-o-y 34 and d-o-y 36.5 (in unshifted time) which runs from the start of CME1, to near the end of the recovery from CME2. Figure 11a shows that $\left[\Phi'_{PC}\right]^2$ is a plausible approximate proxy for the global Joule heating rate and the maximum and event-integrated values are computed for each value of $\delta t$. Part (a) of Figure 12 shows the variations for the peak $\left[\Phi'_{PC}\right]^2$ which occurs at the end of CME1. Part (b) shows the integral of $\left[\Phi'_{SC}\right]^2$ over the event interval, as defined above. It can be seen in Part (a) that the R-M effect (in mauve) is relatively minor for the induced UT changes in the peak power deposited, the total variation (in black) of the peak being dominated the pole-motion effect (in green). In this case, the R-M effect causes the opposite effect on the peak to the pole motion: this is not a general result, for example, the R-M effect would be reversed if the IMF $[B_t]_{GSE}$ had the opposite polarity. Hence the two effects tend to cancel in this case, but because the pole-motion effect is larger the net effect is still considerable. However for the integrated power deposited by the two CMEs, shown in Part (b), the variations caused by the two effects are more similar in amplitude and not so close to being in antiphase and the net effect is smaller but still present.

The combination of the two effects causes a variation in the peak $\left[\Phi'_{PC}\right]^2$ between 82.7% at $\delta t = 10.8$ hr (at $\delta t = -13.2$ hr) and 100.4% at $\delta t = 0.8$ hr showing that the CME events arrived at almost the optimum UT for generating maximum heating effect in the thermosphere and that the peak would have been 17.3% smaller if the
event had arrived 13.2 hr earlier or 10.8 hr later. It should be noted that the R-M effect depends on the temporal variation of the IMF Y and Z components in GSE just before or after the peak (for \( \delta t > 0 \) and \( \delta t < 0 \), respectively, and so the R-M effect on the peak in other events could be considerably different. The integrated power over the event varies between 92.5% at \( \delta t = -11.6 \) hr (and 12.4 hr) and 100.05% at \( \delta t = 0.2 \) hr. This is a smaller effect than that on the peak value but still considerable.

It is interesting that the maximum in both curves is close to \( \delta t = 0 \), which shows that the thermospheric heating would have been smaller had the CMEs arrived at almost any other UT. To understand the full implications for the space weather event would require numerical modeling of the coupled ionosphere-thermosphere system for the various UT of event arrival with allowance for the UT-dependent total power deposition, and for the UT-dependent division of that power between the two hemispheres. Without such modeling it is not possible to say how critical to the outcome the UT of event arrival was. Nevertheless, it is interesting to note that 11 of the 40 satellites launched did reach their intended orbit, almost certainly because by the time of the event they were the ones had been sufficiently elevated in altitude (Tsurutani et al., 2022). This strongly implies that fewer of the satellites would have been lost had the event occurred at a different UT.

### 7. Discussion and Conclusions

The introduction to this paper gave the occurrence probability of events of the magnitude of the 3–4 February 2022 events in the \( kp \) and \( am \) geomagnetic index datasets that extend back to 1932 and 1959, respectively. These probabilities were 3.5% for \( kp \) and 2% for \( am \). We have near-continuous IMF (with continuous SML data) since 1995 and to compare with the geomagnetic data occurrence statistics it is important to use the subset of the index datasets over the same interval. The reason is that the decline in the open solar flux since 1985 has caused a similar long-term decline in geomagnetic activity (Lockwood et al., 2022). The \( kp \) and \( am \) indices after both the CME impacts in the events studied here peaked at 5+ and 84 nT, levels that have been detected or exceeded.
for, respectively, 1.22% and 1.13% of the years 1995–2021, inclusive. The figure for the $am$ index is the most significant because it has the most uniform response to solar wind forcing with time-of-year and Universal Time of all the geomagnetic indices because it employs the most uniform geographical network of stations, using two hemispheric rings of near-equispaced mid-latitude stations (Lockwood, Chambodut, et al., 2019).

It is interesting to compare with the $-SML$ index which, in 10-min running means, peaked at 1348 and 1059 nT after CME1 and CME2, respectively. These values have been detected or exceeded 0.06% and 0.21% of the interval 1995–2021. Notice that the choice of 10 min is somewhat arbitrary and the corresponding values for the raw 1-min $-SML$ data are 0.03% and 0.09%; however, the maximum $-SML$ in 1-min data is not a good measure of the integrated thermospheric heating effect in the event. For 1-hr running means the values are 0.06% and 0.23% which are very similar to the 10-min values. Similarly the 10-min running means of the inferred dayside magnetopause reconnection voltage $\Phi_p$ peaked at 148.2 and 103.7 kV after CME1 and CME2, values that were found for 0.13% and 0.63% of the same interval. Hence the events were significantly more unusual and extreme in the 1-min $-SML$ and $\Phi_p$ values than in the 3-hourly planetary indices.

We have here investigated the magnetospheric response to the two CMEs that impacted Earth on 3 and 4 February 2022 and famously led to the loss of many recently-launched Starlink satellites. Using statistical relationships to derive the variations in reconnection rate in the dayside magnetopause and in the cross-tail current sheet from the observed variations of interplanetary space and of the $SML$ auroral electrojet index. We find that, in addition to a loss rate of magnetospheric open flux related to $SML$, that a smaller loss rate, not detected by $SML$ but proportional to the open flux is needed to prevent unfeasibly large polar cap fluxes (above about 1 GWb). Using a loss rate time constant of $\tau_N \approx 6.8$ min we match the polar cap diameters inferred from auroral images, making the simplifying assumption of a circular polar cap.

The dawn-dusk diameter of the open polar cap inferred in this way was than used to compute the voltage contributions made by the diurnal pattern of polar cap motions in a geocentric-solar frame. This voltage was then combined with that due to the solar wind-magnetosphere interaction, which also depends on UT because of the well-understood Russell-McPherron effect. This has enabled us to compute what the transpolar voltage and Joule heating responses would have been, had the CMEs arrived at different Universal Times.

We have shown that the CMEs in the events of early February 2022 arrived at close to the Universal Time which gave optimum heating of the thermosphere. This is particularly true for the peak of the heating burst at the end of the first CME, but is also true for the integrated heating over the duration of both CME events and their immediate aftermath. We show that the heating peak could have been lower by up to 17.3%, and the integrated heating lower by up to 7.5%, had the events arrived roughly 12 hr earlier or 12 hr later. It is not possible from the study presented here to make definitive statements about what difference this would have made to the fate of the Starlink satellites. However, our results could be used with numerical modeling of the coupled ionosphere-thermosphere system to evaluate how the UT of event arrival could have influenced the subsequent orbit changes. This modeling would need to take account of the UT-dependent total power deposition, and for the UT-dependent division of that power between the two hemispheres. Hence, at this stage, it is not possible to say how critical to the outcome for the satellites the UT of event arrival was. Nevertheless, noting that 11 of the 40 satellites launched did reach their intended orbit does strongly suggest that more would have survived had the events occurred at a different UT. We note that the modeling of the heating during this event by (Dang et al., 2022) shows that there is structure in MLT, latitude (including the hemisphere). The phenomena we describe has implications for this structure which will only be resolved by further modeling, allowing for the UT dependence of the Joule heating and the differences between the two hemispheres.

This adds another dimension to the prediction and understanding of terrestrial space weather events: we have known for many years that we need to develop techniques to better predict the IMF field strength and orientation embedded in events, but we need to also predict the time-of arrival with some accuracy to make best use of that information and predict the geoeffectiveness of the event.

**Data Availability Statement**

The data used in this study are all openly available. The interplanetary data are available from the Physics Data Facility (SPDF) at NASA’s Goddard Space Flight Center as the Omni composite from https://omniweb.gsfc.nasa.gov/ow_min.html. The am indices are available from the International Service of Geomagnetic Indices (ISGI) at http://isgi.unistra.fr/data_download.php. The SuperMAG SML and SMR indices are available from the SuperMAG project website at the Johns Hopkins University Applied Physics Laboratory at
References


