

# JGR Space Physics

## RESEARCH ARTICLE

10.1029/2021JA029272

### Key Points:

- Underlying modes of ionospheric plasma flow velocity discovered using novel Empirical Orthogonal Function infill technique tailored to non-orthogonal Super Dual Auroral Radar Network data
- Plasma velocity modes similar to those of surface magnetic field variation, with a different hierarchy of mode importance
- Interplanetary Magnetic Field (IMF)  $B_y$  has a stronger impact on plasma velocity variability than IMF  $B_z$ , which is opposite to the magnetic field variability behavior

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

R. M. Shore,  
[robore@bas.ac.uk](mailto:robore@bas.ac.uk)

### Citation:

Shore, R. M., Freeman, M. P., & Chisham, G. (2021). Data-driven basis functions for SuperDARN ionospheric plasma flow characterization and prediction. *Journal of Geophysical Research: Space Physics*, 126, e2021JA029272. <https://doi.org/10.1029/2021JA029272>

Received 3 MAR 2021

Accepted 17 JUN 2021

©2021. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

# Data-Driven Basis Functions for SuperDARN Ionospheric Plasma Flow Characterization and Prediction

R. M. Shore<sup>1</sup> , M. P. Freeman<sup>1</sup> , and G. Chisham<sup>1</sup> 

<sup>1</sup>British Antarctic Survey, Cambridge, UK

**Abstract** The archive of plasma velocity measurements from the Super Dual Auroral Radar Network (SuperDARN) provides a rich data set for the investigation of magnetosphere-ionosphere-thermosphere coupling. However, systematic gaps in this archive exist in space, time, and radar look-direction. These gaps are generally infilled using climatological averages, spatially smoothed models, or a priori relationships determined from solar wind drivers. We describe a new technique for infilling the data gaps in the SuperDARN archive which requires no external information and is based solely on the SuperDARN measurements. We also avoid the use of climatological averaging or spatial smoothing when computing the infill. In this regard, our approach captures the true variability in the SuperDARN measurements. Our technique is based on data-interpolating Empirical Orthogonal Function analysis. This method discovers from the SuperDARN data a series of dynamical modes of plasma velocity variation. We compute the modes of a sample month of northern hemisphere winter data, and investigate these in terms of solar wind driving. We find that the  $B_y$  component of the Interplanetary Magnetic Field (IMF) dominates the variability of the plasma velocity. The IMF  $B_z$  component is the dominant driver for the background mean field, and a series of non-leading modes, which describe the two-cell convection variability, and the substorm. We recommend our new technique for reanalysis investigations of polar-scale plasma drift phenomena, particularly those with rapid temporal fluctuations and an indirect relationship to the solar wind.

**Plain Language Summary** Naturally occurring electric fields surround us in near-Earth space, especially concentrated around the magnetic north and south poles. These electric fields are ultimately driven by the turbulent magnetic field of the Sun, and as such, they are constantly varying. From the Earth's surface, we can measure the electric field disturbances caused by Sun-Earth interaction, using arrays of radars distributed across the globe. However, not everywhere can be simultaneously measured in this way, and our archive of electric field measurements has many gaps. Because of this, we have until now had to guess at the content of the missing information using quasi-stationary patterns of electric field, assuming averaged solar conditions. This is known to be a simplification of the true case. In this study, we develop and apply a pattern-finding technique, which allows us the best-yet determination of the true nature of near-Earth space electric field variability.

## 1. Introduction

The interaction of the solar wind with the magnetosphere drives plasma circulation (convection) on a variety of spatial and temporal scales throughout near-Earth space (Abel et al., 2006, 2007, 2009). Within the strongly coupled magnetosphere-ionosphere system, communication of stresses via magnetic field lines leads to a mapping of plasma convection features between the collisional and collisionless environments. In this way, the ionosphere both modulates, and provides a projection surface for dynamics occurring throughout the magnetosphere-ionosphere system (Merkin et al., 2016). By observing the ionospheric plasma, we can infer the underlying physics of this large-scale natural plasma laboratory.

The Super-Dual Auroral Radar Network (SuperDARN, e.g., Chisham et al., 2007; Greenwald et al., 1995; Nishitani et al., 2019) is a globally distributed array of ground-based coherent-scatter radars which measure line-of-sight horizontal ionospheric plasma velocity, based on backscattered signals from ionospheric electron density irregularities. Owing to the scientific utility of SuperDARN, it has enjoyed widespread use in studies seeking to characterize the ionospheric plasma variations, and to describe these in terms of solar wind or magnetosphere-ionosphere drivers. SuperDARN offers a simultaneous sampling of ionospheric

plasma velocity over large areas of the polar regions, providing a combination of coverage and resolution which is yet unmatched by satellite constellations. Despite this, there are significant gaps in the SuperDARN data archive in both space and time, and much study has been devoted to imputing (infilling) the data over these gaps, as we will later see.

The study of the high-latitude ionosphere must overcome the limitation of data sets which give an incomplete picture of the system, while characterizing its electrodynamics such that we can understand and forecast its variations. Historically, there have been two main schools of thought for approaching this. Over decades, many studies (each using relatively few data) have combined to show that the surface magnetic field variations caused by ionospheric electrodynamics can be considered as a sum of orthogonal component patterns (e.g., Friis-Christensen & Wilhjelm, 1975; Iijima et al., 1984; Maezawa, 1976; Matsushita et al., 1973; Nishida, 1968b). An alternative approach (now in more common use) is to characterize the ensemble high-latitude convection in an empirical model, with solar wind measurements and geographic season as the organizing parameters. Again, decades of theoretical and empirical studies have evolved our understanding of the optimal parameters for modeling the convection and surface magnetic field (e.g., Cousins & Shepherd, 2010; Cowley, 1982; Cowley & Lockwood, 1992; Dungey, 1961; Laundal, Finlay, & Olsen, 2016; Laundal et al., 2018; Lester et al., 2006; Reiff & Burch, 1985; Reiff et al., 1981; Ruohoniemi et al., 2002; Shore, Freeman, Coxon, et al., 2019; Weimer, 1996, 2001, 2005; Weimer et al., 2010; Wygant et al., 1983). For ease of reference, we will call these two schools of thought the “component” and “parameterized” approaches.

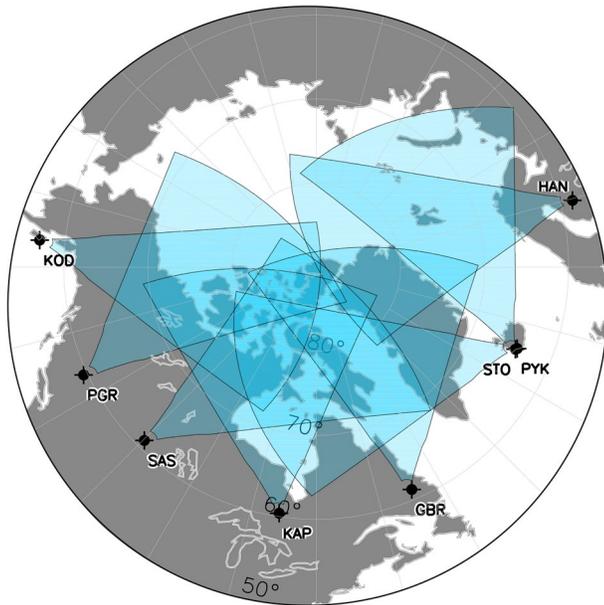
The parameterized approach is now ubiquitous in comparison to the component approach, and is driving progress in both characterization and forecasting, yet it has a shortcoming in that a relationship to the solar wind must be assumed. This has led to the potential for introducing biases. For example, climatological models which are developed assuming a relationship to the Interplanetary Magnetic Field (IMF) (e.g., Pettigrew et al., 2010; Ruohoniemi & Greenwald, 1996, 2005; Thomas & Shepherd, 2018) are used for both visual interpretation of quasi-stationary patterns of plasma velocity, and to populate the background field when the variability of the plasma velocity is unknown. Hence, the assumed relationship with the solar wind affects both the characterization and infill of the data. In comparison, the component approach need not make such assumptions when infilling the data.

We are now at a point whereby we can expand and automate the component approach, transitioning our understanding of ionospheric electrodynamics from a consensus of many studies, which each use few data, to new studies, which each use lots of data. Shore et al. (2017, 2018) have demonstrated this approach for magnetic field variations, using the Empirical Orthogonal Function (EOF) method. The EOF technique decomposes a large, highly variable data set into a series of independent spatio-temporal subgroups, called “modes” of variance. The modes are discovered from the data with no prior specification, thus they express the intrinsic dynamical properties of the system. Collectively, the modes describe the full variability of the data, and the majority of the variability is commonly given by a handful of mode patterns. These can be used to impute the missing data. This allows us to offer a technique to provide an infill solution without resort to external information (e.g., the solar wind). If desired, the dependence on the solar wind or substorms can later be determined separately for each mode (Shore, Freeman, & Gjerloev, 2019). This allows us to provide a semi-automated characterization of the electrodynamics in a manner similar to the historic component approach, but based on many more data. The goal of this study is to describe a variation of the EOF technique tailored for SuperDARN plasma velocity data, intended for self-consistent investigation of dynamic features in the ionospheric plasma resulting from both directly and indirectly driven sources.

In Section 2, we summarize the data used in the study, and in Section 3 we describe the application of the EOF method and the pre- and post-analysis processing. In Section 4, we describe our findings, which are placed in a wider context in Section 5. We summarize in Section 6.

## 2. Data

We use measurements from all SuperDARN northern hemisphere radars which contributed data in February 2001, chosen because it is one of the best months for SuperDARN data coverage in the archive (we are investigating the implications of poorer data coverage as part of a subsequent study). The SuperDARN radars



**Figure 1.** Nominal fields of view for each northern hemisphere radar, which contributed data during February 2001. Both data selection, and the natural variability of ionospheric backscatter, will act to reduce the effective sampling region for each radar. The acronyms are detailed further in Chisham et al. (2007).

transmit and receive high frequency radio signals (8–20 MHz) on oblique propagation paths that are refracted toward the horizontal in response to increases in electron density in the E- and F-regions of the ionosphere. When propagating perpendicular to the Earth's magnetic field, these signals can backscatter from magnetic field-aligned, decameter-scale ionospheric density irregularities that move with the background ionospheric plasma flow. This backscatter is measured along (approximately) discrete radar beams that are separated in horizontal azimuth, and sampled along each beam, at a series of distances (ranges) from the radar antenna array—the resulting measurement regions are illustrated in Figure 1.

The radars transmit a multi-pulse sequence to allow the determination of multi-lag complex autocorrelation functions for each sample (Greenwald et al., 1985; Ponomarenko & Waters, 2006). The autocorrelation functions are converted to line-of-sight Doppler velocities by applying linear fits to the autocorrelation function phase variations (Hanuise et al., 1985) using version 4.0 of the radar software toolkit (RST 4.0 and FITACF 2.5; SuperDARN Data Analysis Working Group et al., 2017). The resulting velocities are assigned an estimated geographic location (for each combination of beam azimuth and range) using a time-invariant “virtual height” model for the radio wave propagation through the ionosphere, described by Chisham et al. (2008). The uncertainties in the estimated location are summarized by Burrell et al. (2016). The data from a given radar now comprise a series of 3- or 7-s integrations of line-of-sight plasma velocity at a range of geographic locations, with a 1–2 min interval between repeat measurements at the same location (A variety of scanning modes are used by SuperDARN, but the aforementioned scan type is typical).

We apply data selection to these geo-located velocities in order to select good quality data from the F-region ionosphere. To ensure a high signal-to-noise ratio, we discard measurements with backscatter power less than 3 dB, data quality flag set to “true” (indicating low quality by RST 4.0), or which lie within 495 km of the radar array (corresponding to the first 11 discrete ranges) in which unwanted meteor echoes and E-region backscatter may be present. To reduce instances of near-zero velocity from the radar beam scattering off the Earth's surface instead of ionospheric irregularities, we discard measurements with an absolute velocity  $< 50 \text{ ms}^{-1}$ , or marked as ground scatter by RST 4.0. To avoid areas of low data coverage, which can affect the EOF solution, we remove all data from geomagnetic colatitudes greater than  $30^\circ$ .

### 3. Method

#### 3.1. Pre-EOF-Analysis Processing

The SuperDARN data are well suited for EOF analysis, because the spatial and temporal behavior of the plasma velocity are strongly correlated, in a manner related to the underlying physics of the plasma flow, which we seek to discover. However, pre-processing of the data is essential in order to correctly obtain these patterns of covariance. When applying the EOF method to geophysical data, it is important that the measurement locations and epochs contribute uniformly to the total variance, and that there is no variance structure within the measurement basis itself. This means that the EOF method will then decompose the variance structure of the underlying field, instead of responding to correlations in the spatiotemporal measurement structure. To achieve this, we apply further processing to the line-of-sight velocity data described in Section 2 in order to create a regular, discrete basis in space, time, and for the velocity vector direction, as follows.

Prior to forming the regular spatial basis for our data, we transform the geo-located velocity measurement locations into an inertial (approximately Sun-synchronous) coordinate frame. This is done to reduce the number of EOF modes required to describe the variations in the plasma velocity, which is dominated by convective patterns that are to first order Sun-synchronous. To maintain consistency with our earlier reanalysis

of the surface-measured external and induced magnetic field (SEIMF; Shore et al., 2017, 2018), we choose the Quasi-Dipole (QD) system of inertial coordinates for this transformation (Emmert et al., 2010; Laundal & Gjerloev, 2014; Richmond, 1995), and we analyze a calendar month of data to allow multiple opportunities for measurements from a given longitude to contribute to a given local time. This time span is also longer than many magnetosphere-ionosphere time scales of interest such that representative modes may be captured. To provide a regular spatial basis in this coordinate frame, the velocity data locations are separated into equal-area spatial bins (Leopardi, 2007), which are defined in QD latitude and QD magnetic local time (MLT, given by the difference of the apex longitude and the geomagnetic dipole longitude of the subsolar point). Note that although we ultimately transform both the measurement locations and the measured vectors to the QD frame, we only transform the locations of the measurements into an inertial frame, while the velocity vectors remain defined in a frame co-rotating with the Earth, with no adjustment for the motion associated with Earth rotation.

We now describe the basis we use to represent the velocity vector direction. A discrete basis is needed for this quantity because at a given location and time, the relationship between the true plasma velocity vector, and the radar's line-of-sight sample of that vector, is unknown. This missing information can be recovered if we define a basis for the vector direction, using the horizontal angle between the line-of-sight direction for a given radar measurement, and the local QD north direction. This provides a fixed relationship between the measurement line-of-sight direction, and the QD frame. The velocity vector's direction and amplitude are then discoverable from the cumulative information given by many non-contemporaneous measurements in different look directions at the same location, as we describe later in Section 3.3. The non-orthogonal nature of the QD coordinate frame brings in additional complexities to defining this velocity direction basis in a self-consistent manner, as follows.

The horizontal QD basis vectors are termed  $\mathbf{f}_2$  and  $\mathbf{f}_1$  in the notation of Laundal and Gjerloev (2014), where  $\mathbf{f}_2$  points predominantly northward at most locations and  $\mathbf{f}_1$  predominantly eastward. In addition to being non-orthogonal, the QD basis vectors are not of unit length—both these factors vary with geographic location, and will affect the direction and magnitude of the line-of-sight plasma velocity measurements. Since we analyze the data in a non-geographic frame, we must rescale both the line-of-sight direction, and amplitude, of the velocity measurement as follows. The radar beam line-of-sight direction is defined by its horizontal clockwise angle ( $\theta_{beam}$ ) starting from the  $\mathbf{f}_2$  basis vector. This angle is normalized by the ratio between orthogonal axes ( $=90^\circ$ ) and the horizontal clockwise angle  $\gamma_{QD}$  from the  $\mathbf{f}_2$  basis vector to the  $\mathbf{f}_1$  basis vector:

$$\tilde{\theta}_{beam} = \frac{\theta_{beam}}{\gamma_{QD} / 90^\circ} \quad (1)$$

Following this normalization, a value of  $\tilde{\theta}_{beam} = 0^\circ$  describes a measurement in the direction of  $\mathbf{f}_2$ , and  $\tilde{\theta}_{beam} = 90^\circ$  points along  $\mathbf{f}_1$ . We use the rescaled north-offset angles  $\tilde{\theta}_{beam}$  to separate the data within each spatial bin into  $6^\circ$ -wide angular partitions defined clockwise from local QD north. This forms a regular basis for the different line-of-sight measurement directions, with minimal loss of precision.

Following Equations 2 and 3 in Laundal and Gjerloev (2014), we now rescale the magnitude of the velocity measurement  $v$  to take account of the non-unit-vector nature of the QD basis vectors:

$$\tilde{v} = v \left( \sqrt{\left( \cos(\tilde{\theta}_{beam}) (\|\mathbf{f}_2\| / F) \right)^2 + \left( \sin(\tilde{\theta}_{beam}) (\|\mathbf{f}_1\| / F) \right)^2} \right) \quad (2)$$

The operation scales the velocity measurement according to the “amount” that it is pointing in each of the  $\mathbf{f}_2$  and  $\mathbf{f}_1$  directions. The additional scaling factor  $F$  is given by

$$F = \mathbf{f}_1 \times \mathbf{f}_2 \cdot \mathbf{k} \quad (3)$$

where  $\mathbf{k}$  is a unit vector pointing upward from the Earth's surface.

Lastly, we construct a regular temporal basis for the velocity data. For all the data in a given spatial bin and in a given  $6^\circ$  line-of-sight bin, we compute a time series of the median line-of-sight velocity in successive

5-min bins. The first time bin in a given day is from 23:59:30 UT (on the previous day) to 00:04:30 UT (chosen to match the temporal basis of the Shore et al., 2018 model), so the center epoch is at 00:02:00 UT.

The above processing is applied to all available data in February 2001. Following this, the velocity data  $v$  for each  $n$ th 5-min epoch ( $n \in \{1, 2, \dots, N\}$ ),  $p$ th QD spatial location ( $p \in \{1, 2, \dots, P\}$ ), and  $i$ th line-of-sight partition ( $i \in \{1, 2, \dots, I\}$ ) for a given month are formed into a data matrix  $\tilde{\mathbf{X}}$  of dimension  $N$  by  $M$ , where  $M = P * I$

$$\tilde{\mathbf{X}} = \begin{bmatrix} \left( \begin{array}{cccc} v_{1,1,1} & v_{1,2,1} & \cdots & v_{1,p,1} \\ v_{2,1,1} & v_{2,2,1} & & v_{2,p,1} \\ \vdots & \vdots & & \vdots \\ v_{N,1,1} & v_{N,2,1} & \cdots & v_{N,p,1} \end{array} \right) & \cdots & \left( \begin{array}{cccc} v_{1,1,i} & v_{1,2,i} & \cdots & v_{1,p,i} \\ v_{2,1,i} & v_{2,2,i} & & v_{2,p,i} \\ \vdots & \vdots & & \vdots \\ v_{N,1,i} & v_{N,2,i} & \cdots & v_{N,p,i} \end{array} \right) \end{bmatrix} \quad (4)$$

Only the variance of the velocity data is analyzed, so prior to applying EOF, we remove the temporal mean  $\bar{\mathbf{X}}$  of the data in each column of  $\tilde{\mathbf{X}}$  (i.e., each  $m$ th combination of  $p$ th spatial location and  $i$ th partitioned line-of-sight direction now has zero mean).

Recall that our rationale for creating a regular basis for the velocity data was to ensure that all elements of a given dimension of the data matrix are uniform, in terms of their contribution to the total variance. However, the bases  $N$ ,  $P$ , and  $I$  of the data do not conform to this requirement, since regions of missing data are highly systematic in SuperDARN. For example, the data are relatively more numerous in the polar cap and cusp, and relatively sparser where radar beams pass through regions of high radio wave absorption due to strong auroral particle precipitation. To reduce the impact of this systematic bias on the EOF analysis, we weight the data in each  $n$ th row of  $\tilde{\mathbf{X}}$  by the number of measurements at that epoch (termed  $z$ ), divided by the number of possible data points  $M$

$$W_{jj} = \frac{z}{M} \quad (5)$$

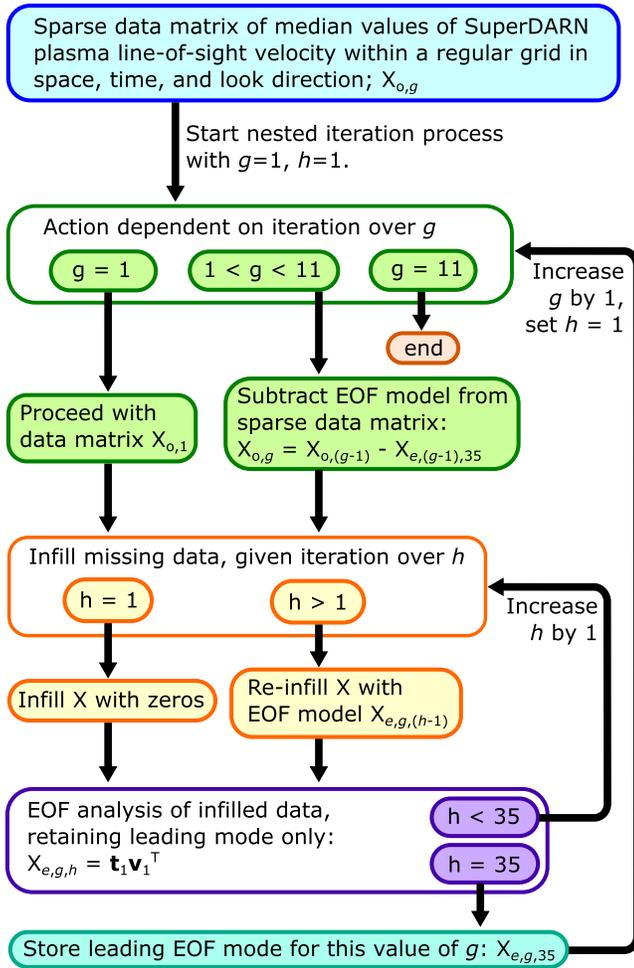
where  $\mathbf{W}$  is a square matrix with dimension  $N$ , and is zero on all off-diagonal elements. The weights range from a value of 1 where a given row of  $\tilde{\mathbf{X}}$  is fully filled, to 0 where that row is empty. They act to decrease the relative contribution of poorly sampled epochs to the solution for the variance of the data. We apply the weights via

$$\mathbf{X} = \mathbf{W}^{0.5} \tilde{\mathbf{X}} \quad (6)$$

where the exponent of 0.5 is used to account for squaring implicit in the EOF decomposition (described in the next section).

The EOF analysis requires that the input data set have no missing values. However,  $\mathbf{X}$  is highly sparse—the majority of the regular grid in space, time, and line-of-sight direction has no coverage. To infill the missing coverage we use a modification of the Data-Interpolating EOFs technique (DINEOFs, described by Beckers & Rixen, 2003), which was modified for application to magnetic field data by Shore et al. (2017, 2018). In the DINEOF technique, the missing values in the data matrix are initially infilled with zeros, and are then iteratively replaced with forecasts of the variability from successive EOF analyses, until the amplitude of the infill converges with that of the data coverage. For clarity, there are no measurements where we apply the infill, but since the EOF solution is computed over the full  $N$  by  $M$  extent of  $\mathbf{X}$ , there are regions where the infill and original data can be compared.

In the modified version of DINEOFs applied here (described by Shore et al., 2017, 2018), the infill is a nested two-stage iteration process. For clarity, we provide a schematic illustration of the iterative data infill procedure in Figure 2. In the first stage of the iteration process, the infill values start at zero and are iteratively replaced with EOF predictions of increasing amplitude—we signify the iteration number with the subscript  $h$ . When the iteration over  $h$  is complete (see below), the infilled model prediction is subtracted from the original data, thus iteratively modifying the values of the original (non-infilled) data—we signify this with the subscript  $g$ . This modification is performed because it means that the leading EOF modes obtained at successive iterations of  $g$  are no longer orthogonal, which allows for the variance of the underlying system



**Figure 2.** Schematic of the data infill procedure, described throughout Section 3.

to be described using fewer iterations of  $g$ . To distinguish the original and altered data matrices, we use the subscripts  $o$  (for “original”) and  $f$  (for “filled”), and we use the subscripts  $g$  and  $h$  to describe the different infill stages. So the sparse data matrix from Equation 4 is referred to as  $\mathbf{X}_{o,1}$ , where 1 is the value of  $g$ , signifying that these data have not yet had an EOF prediction subtracted from them. We refer to data matrices which have undergone infill as a variant of  $\mathbf{X}_{f,g,h}$ . As an initial infill choice we define a matrix  $\mathbf{X}_{f,1,1}$ , for which the null values of  $\mathbf{X}_{o,1}$  are replaced by zeros (i.e., the mean value of the processed data)

$$\begin{aligned} \mathbf{X}_{f,1,1} \Big|_{\mathbf{x}_{o,g} \text{ exists}} &= \mathbf{X}_{o,1} \\ \mathbf{X}_{f,1,1} \Big|_{\mathbf{x}_{o,g} \text{ null}} &= 0 \end{aligned} \quad (7)$$

The processed data are now ready for EOF analysis.

### 3.2. Eigen-Decomposition Method

The following summary of EOF analysis is described in full by Bjornsson and Venegas (1997, page 12), von Storch and Zwiers (2002, pages 294–295) and Jolliffe (2002, page 5). Given a mean-removed vector field  $\mathbf{X}_{f,g,h}$ , measured at  $N$  times and  $P$  locations in  $I$  different directions, EOF decomposes this into  $N$  spatial basis vectors of length  $P * I$ , and  $N$  temporal oscillations of length  $N$ . Each pair of a spatial pattern  $\mathbf{v}$  and its temporal oscillation  $\mathbf{t}$  is termed a “mode” of the analysis. These modes are standing wave components of the vector field  $\mathbf{X}$  which together reconstruct the original data:

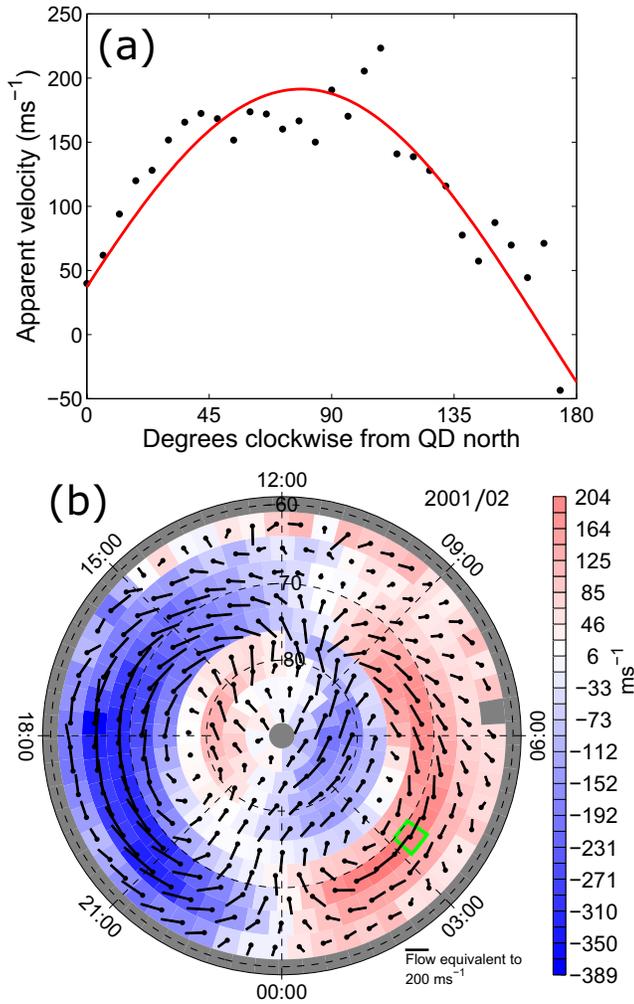
$$\mathbf{X}_{f,g,h} = \sum_j^N \mathbf{t}_j \mathbf{v}_j^T = \mathbf{T} \mathbf{V}^T \quad (8)$$

where  $\mathbf{V}$  are eigenvectors of the covariance matrix  $\mathbf{X}_{f,g,h}^T \mathbf{X}_{f,g,h}$ , and  $\mathbf{T}$  are eigenvectors of  $\mathbf{X}_{f,g,h} \mathbf{X}_{f,g,h}^T$ . Each column of  $\mathbf{V}$  and  $\mathbf{T}$  describe a different mode of the EOF analysis:  $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N)$  and is of size  $(P * I \text{ by } N)$ , and  $\mathbf{T} = (\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_N)$ , which is of size  $(N \text{ by } N)$ . The modes are ranked by eigenvalue, so mode 1 has the largest eigenvalue, and it describes the largest proportion of the variance of  $\mathbf{X}$ .

To iteratively infill the missing data with an EOF model prediction, we use Equation 8 with  $j = 1$  to reconstruct the infilled data matrix using mode 1 only:  $\mathbf{X}_{e,1,1} = \mathbf{t}_1 \mathbf{v}_1^T$ . Here the subscript  $e$  means that the data matrix is based on the EOF model only, with no direct inclusion of the original data values. The values of  $\mathbf{X}_{e,1,1}$  are used to re-infill the data matrix (afterward termed  $\mathbf{X}_{f,1,2}$ ) via

$$\begin{aligned} \mathbf{X}_{f,1,2} \Big|_{\mathbf{x}_{o,g} \text{ exists}} &= \mathbf{X}_{o,1} \\ \mathbf{X}_{f,1,2} \Big|_{\mathbf{x}_{o,g} \text{ null}} &= \mathbf{X}_{e,1,1} \end{aligned} \quad (9)$$

The process of infilling  $\mathbf{X}_{f,1,2}$  as per Equation 9 and performing the EOF decomposition of Equation 8 is repeated until the amplitude of  $\mathbf{X}_{e,1,h}$  converges with that of  $\mathbf{X}_{o,1}$ , where both exist. Similar to Shore et al. (2017), we find that convergence occurs well within  $h = 35$ , with diminishing returns for greater  $h$ . After convergence, the information represented by the leading mode is complete across all parameters  $N$ ,  $P$ , and  $I$ . We do not want this added information to control the infill for subsequent modes, so when incrementing the iteration over  $g$  we remove the iterated EOF solution from the original data via



**Figure 3.** (a) The cosine-fitting process demonstrated for the mean field data ( $\bar{\mathbf{X}}$ ) at  $68.5^\circ$  latitude and 03:27 magnetic local time (MLT) from February 2001. The black dots are the values of  $\bar{\mathbf{X}}$  at this location, as a function of line-of-sight angle (clockwise from Quasi-Dipole north). The red curve is the fitted sinusoid—the evaluation of this curve at  $0^\circ$  and  $90^\circ$  gives the  $N_{\text{QD}}$  and  $E_{\text{QD}}$  components of the horizontal mean plasma velocity vector. (b) The spatial pattern of the horizontal mean plasma velocity from February 2001, after sinusoid fitting. Background colors are the east component of the velocity, chosen to highlight regions of shear. Bins without data are shown in gray (the polar bin data are discarded due to their poor sinusoid fit). The bin at  $68.5^\circ$  latitude and 03:27 MLT has a green border.

$I$  line-of-sight directions at each  $p$ th location. We compute the relationship of this multi-direction spatial basis to the cardinal QD north and east vector components ( $N_{\text{QD}}$ ,  $E_{\text{QD}}$ ) as follows, using the mean field from February 2001 as an example. At a single location in QD latitude and QD MLT, the mean field data give the proportional amplitude of the background plasma velocity in all  $6^\circ$ -wide line-of-sight partitions  $i$ . These values are shown for a single location in Figure 3a, as black dots. We fit a sinusoid to these data (illustrated with the red curve in Figure 3a), using an unconstrained nonlinear minimization to give the best-fitting amplitude  $a$  and phase  $\theta_{\text{phase}}$ . Given the scaling and normalization applied to the pre-EOF data described in Section 3.1, the QD north and east vector components at the spatial location are then given by

$$\mathbf{X}_{o,2} = \mathbf{X}_{o,1} - (\mathbf{M}\mathbf{X}_{e,1,35}) \quad (10)$$

where  $\mathbf{X}_{o,2}$  represents the original data minus the information contained in the first iterated mode, and  $\mathbf{M}$  is a “mask” matrix used to assign null values to a given matrix, such that  $\mathbf{X}_{o,g} = \mathbf{M}\mathbf{X}_{f,g,h}$ . We then repeat the iterative infill process, beginning with zeros as before

$$\begin{aligned} \mathbf{X}_{f,2,1} & \Big|_{\mathbf{X}_{o,g} \text{ exists}} = \mathbf{X}_{o,2} \\ \mathbf{X}_{f,2,1} & \Big|_{\mathbf{X}_{o,g} \text{ null}} = 0 \end{aligned} \quad (11)$$

$\mathbf{X}_{f,2,1}$  is then decomposed according to Equation 8, and the mode with the largest eigenvalue is used to refresh the infill, repeating until  $h = 35$ .

In summary, during this nested iteration over  $g$  and  $h$ , we iteratively infill empty bins using the amplitudes of the first mode, then we remove this mode from the original data, and repeat the process. The results discussed in this manuscript are the leading modes from each stage of the removal process, once convergence has been attained. We refer to the set of all fully iterated  $\mathbf{v}$  as  $\mathbf{Z}$ , and all fully iterated  $\mathbf{t}$  as  $\mathbf{U}$ . Although all columns of  $\mathbf{Z}$  are technically “mode 1” of a given removal stage, we refer to them as modes 1, 2, and so on, according to the stage (i.e., value of  $g$ ) they originated from.

In the remainder of this report, we describe the EOF model with the weighting removed, via

$$\mathbf{Y} = \mathbf{W}^{-0.5}\mathbf{Z} \quad (12)$$

Following this de-weighting, the reconstruction  $\mathbf{U}\mathbf{Y}^T$  will produce a model of the data matrix with the same units as  $\bar{\mathbf{X}}$ .

For display, we compute a version of the EOF basis vectors in which the spatial patterns have normalized amplitudes, and the temporal eigenvectors have units of  $\text{ms}^{-1}$ , as follows

$$\begin{aligned} \tilde{\mathbf{u}}_j & = \mathbf{u}_j |Y_j| \\ \tilde{\mathbf{y}}_j & = \frac{\mathbf{y}_j}{|Y_j|} \end{aligned} \quad (13)$$

where  $\mathbf{u}_j$  is the  $j$ th column vector of  $\mathbf{U}$ , and the scalar  $Y_j$  is the maximum-amplitude element of  $\mathbf{y}_j$ , itself the  $j$ th column vector of  $\mathbf{Y}$ .

### 3.3. Post-EOF-Analysis Processing

At this point, both the mean field  $\bar{\mathbf{X}}$  which we removed prior to EOF analysis, and the individual EOF spatial basis vectors  $\tilde{\mathbf{y}}_j$ , are defined over

$$\begin{aligned} N_{QD} &= a \cos(\theta_{phase}) \\ E_{QD} &= a \sin(\theta_{phase}) \end{aligned} \quad (14)$$

The sinusoid fitting procedure is repeated for all  $p$  spatial locations of the mean field. The horizontal vector map of fitted mean velocities (encompassing all locations) is shown in Figure 3b.

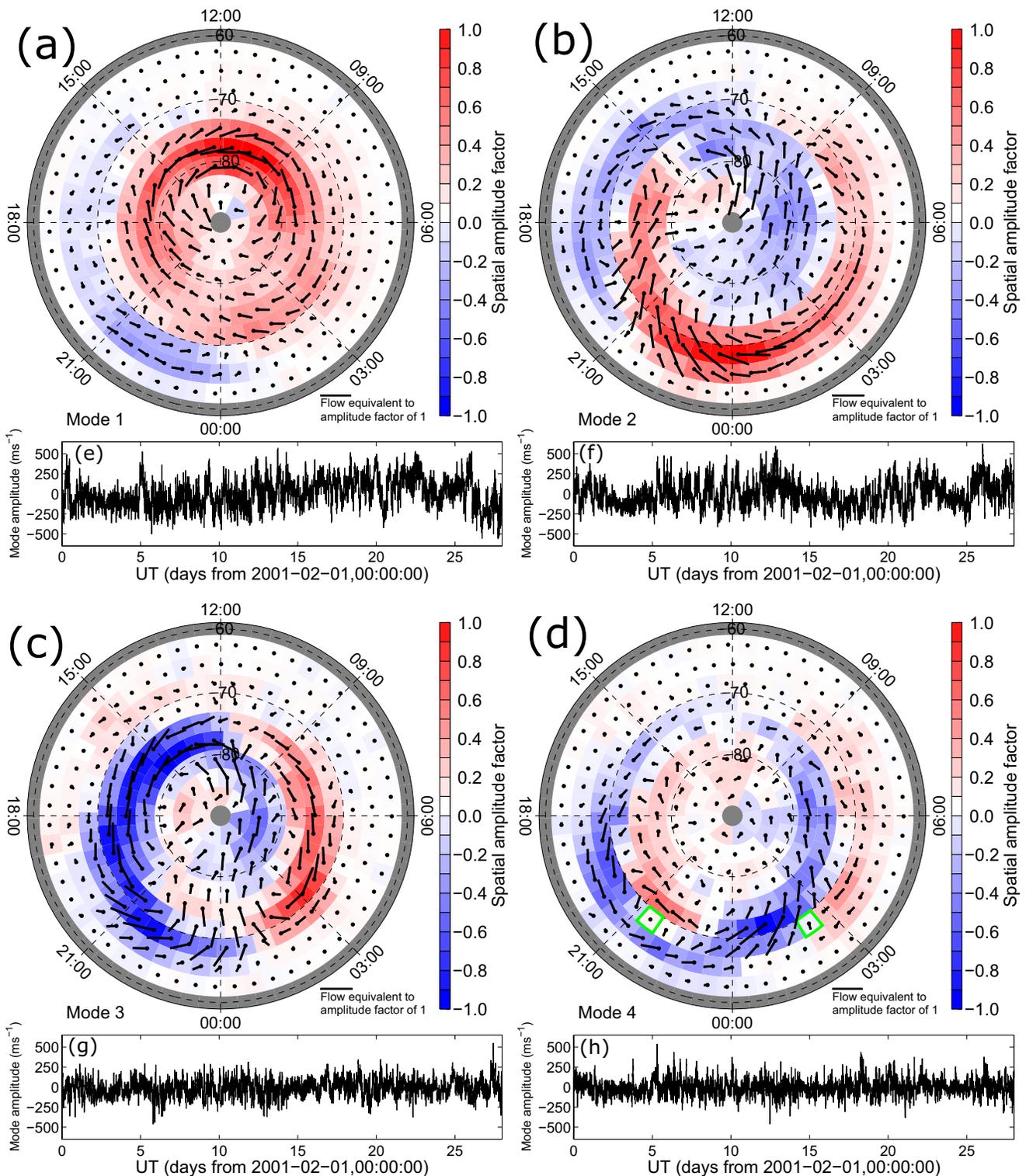
The sinusoid fitting procedure is likewise applied to each location in each of the  $g$  mode spatial patterns (our model is based on 10 fully iterated modes). We note that the fitting of a smooth (sinusoid) model to the plasma velocity happens only after the EOF-driven infill of the variability in un-measured locations, and does not affect the mode time series (which retain their 5 min cadence). In other words, the EOF modes recover the plasma velocity vector based on an entire month's worth of radar measurements at all locations, without necessitating any temporal smoothing. In this way, the sinusoid fitting provides a natural complement to the decomposition of the data variability into independent standing waves of spatial and temporal information.

Lastly, we note that a given  $p$ th location may not include coverage from all  $l$  line-of-sight directions. The combinations of  $p$  and  $l$  which lack all measurements cannot be infilled using DINEOFs, and were not included in the EOF analysis of  $\mathbf{X}$ . The extent of this missing data varies with data selection choices, and hence this is not reflected in the generalized mathematics of Section 3.2. In extreme cases, the sinusoid fitting is affected by systematic gaps in particular line-of-sight directions, particularly where data coverage is very low. As part of the sinusoid fitting process for each bin, we record the sum of squared differences between the sinusoid model fit and the mode spatial pattern data. If this misfit value exceeds 6 times the standard deviation of the misfits for all bins, we reject the sinusoid fit for that bin—this is the reason for the grayed-out bin near dawn in Figure 3b. The polar bin (where line-of-sight direction is undefined), and bins at colatitudes over  $30^\circ$ , are gray in Figure 3b since they were removed as part of the data selection process. For reference, the sum of squares misfit value of the example shown in Figure 3a is  $19,662 \text{ (ms}^{-1}\text{)}^2$  (with a root mean square misfit of  $26 \text{ ms}^{-1}$ ), and the bin in Figure 3b with the poorest permitted fit has a sum of squares misfit of  $1,670,533 \text{ (ms}^{-1}\text{)}^2$  (with a root mean square misfit of  $240 \text{ ms}^{-1}$ ). This poorest permitted fit is shown in Figure S1 in the Supporting Information, to illustrate the range of quality in the fits which are permitted.

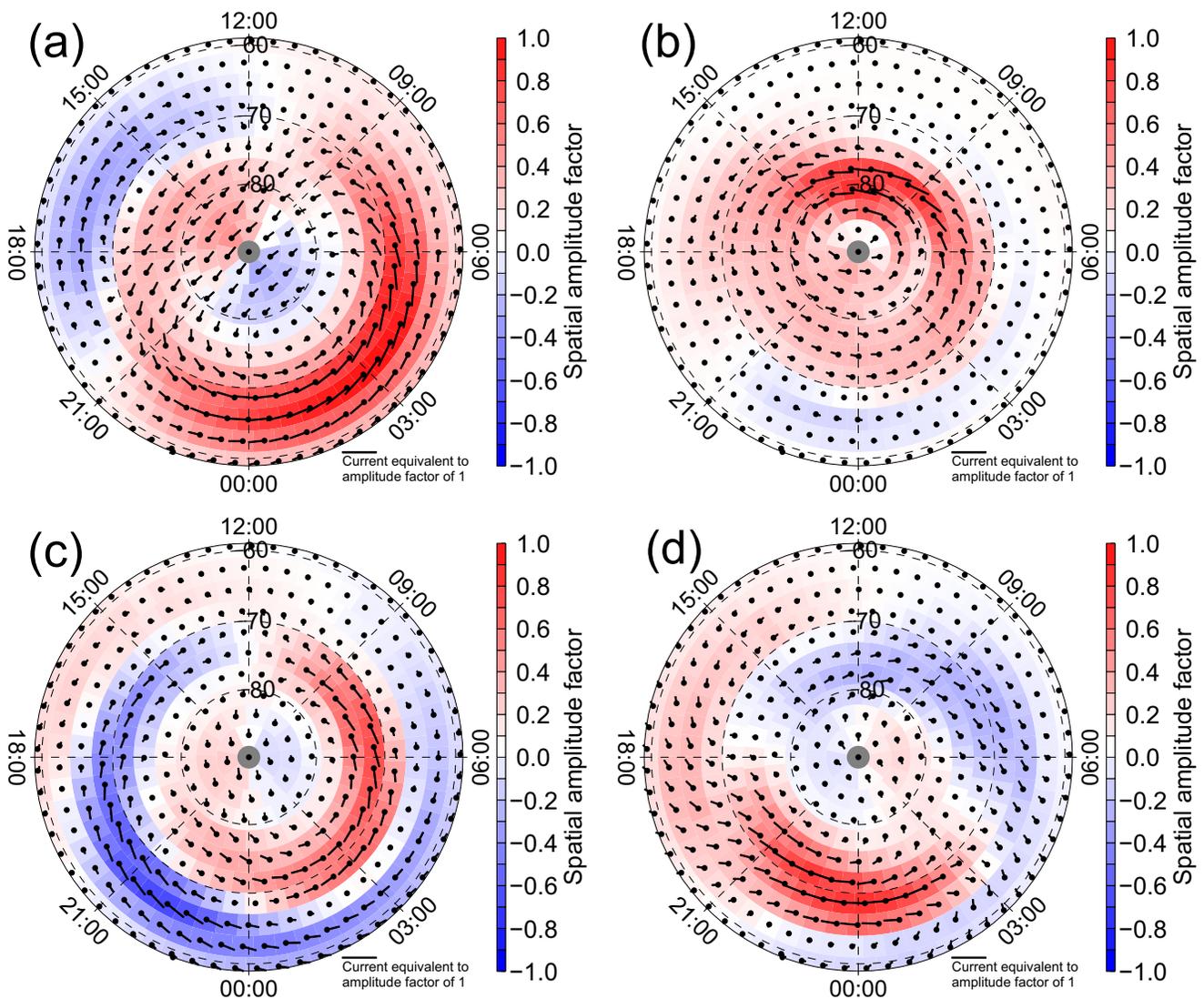
#### 4. Results

In Figure 4 we show the first four modes (those with the largest four eigenvalues) from the EOF analysis of SuperDARN data spanning February 2001. Figures 4a–4d show the spatial patterns of normalized horizontal plasma velocity for each mode ( $\tilde{\mathbf{y}}$  from Equation 13), following the evaluation of Equation 14 to give the QD horizontal components of the localized velocity vector. Figures 4e–4h show the time series of amplitude oscillations for each mode ( $\tilde{\mathbf{u}}$  from Equation 13). These time series describe the amplitude of the associated spatial pattern at each 5-min epoch of the data, that is, Figures 4a and 4e together comprise mode 1. Hence, the mode time series define the epoch-by-epoch contribution of that specific spatial pattern of velocities to the total variability of the horizontal plasma velocity.

Following the approach of Shore et al. (2017), we can physically interpret the modes in Figure 4 to determine which solar wind or magnetospheric driver best explains each of them. For this subjective interpretation, we rely on the expert knowledge garnered from decades of study of the magnetosphere-ionosphere system behavior (refer to Section 1). We also refer to the discovered modes from a series of EOF analyses of magnetic field data spanning solar cycle 23, performed by Shore et al. (2018), which are shown in Figure 5. These vector patterns are shown in the form of the equivalent ionospheric current which would lead to the observed magnetic perturbation. As discussed by Shore, Freeman, Coxon, et al. (2019), there are a number of definitions for this equivalent current. We choose the simplest, which is to rotate the horizontal magnetic vector by  $90^\circ$  clockwise (this transformation does not affect the time variation of these modes). These equivalent currents may not reflect the true current distribution equally well in all seasons and locations (e.g., Laundal et al., 2015; Laundal, Finlay, & Olsen, 2016; Laundal, Gjerloev, et al., 2016; Laundal et al., 2018). However, the transformation allows a simple visual comparison between the equivalent current and the plasma velocity, which are approximately anti-parallel for uniform ionospheric conductance.



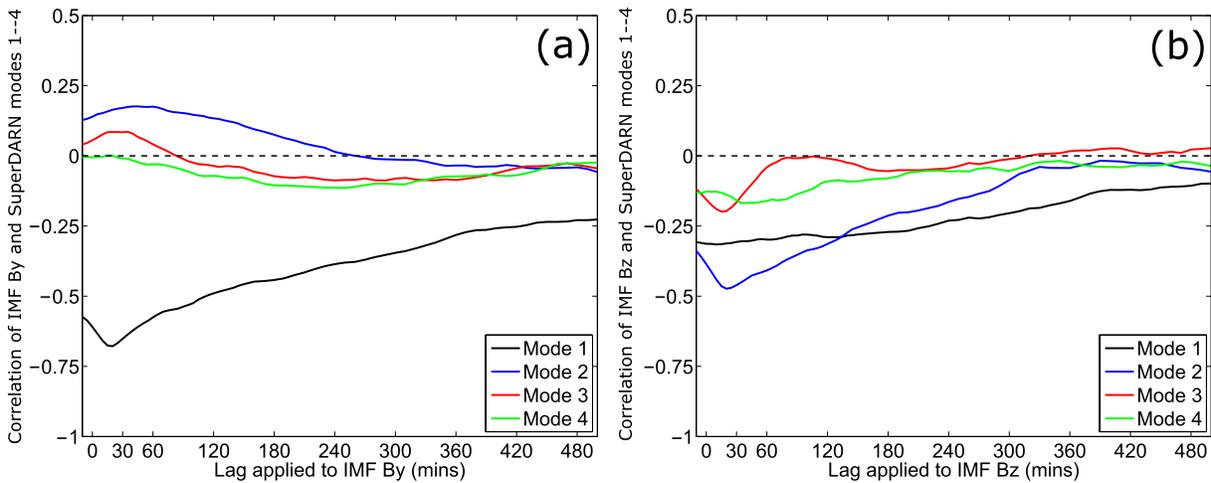
**Figure 4.** (a–d) spatial patterns of modes 1–4 of the Empirical Orthogonal Function analysis of Super Dual Auroral Radar Network data from February 2001. The patterns are shown as vectors of perturbations to the horizontal plasma flow (background colors are the east component of the vectors). (e–h) temporal patterns of modes 1–4, showing the time series of amplitudes for each associated spatial pattern (i.e., (a) and (e) together comprise mode 1). We define the modes as (a) Velocity Perturbation Type Y (VPY), (b) Velocity Perturbation Type 2 expanded (VP2e), (c) Velocity Perturbation Type 2 contracted (VP2c), and (d) Velocity Perturbation Type 1 (VP1). The two green squares in panel (d) at 21:30 MLT and 02:30 MLT, and both at 68° latitude, highlight the centers of velocity rotation associated with substorm field-aligned currents, as discussed in the main text.



**Figure 5.** (a–d) spatial patterns of the four dominant modes of magnetic field variation, discovered by Empirical Orthogonal Function (EOF) analysis of SuperMAG data. These patterns are each the mean of up to 144 individual EOF patterns. Those constituent data were obtained from 12 years of monthly EOF analyses (spanning solar cycle 23), and subsequently grouped according to the similarity of their spatial patterns by applying network analysis—refer to Shore et al. (2018) for a full description of the technique. These patterns are shown as vectors of the equivalent current, which would cause the observed perturbation in the geomagnetic field. Background colors are the south component of the magnetic perturbation, associated with westward ionospheric equivalent current. The modes have been previously identified or defined as (a) Disturbance Polar Type 2 (DP2), (b) Disturbance Polar Type Y (DPY), (c) Disturbance Polar Type 2 expansion contraction (DP2EC), and (d) Disturbance Polar Type 1 (DP1).

We will now contrast the magnetic field modes from Figure 5 with the SuperDARN plasma velocity mean field and modes shown in Figures 3 and 4, to draw some conclusions about the different structures of variance which comprise the plasma drift and the magnetic field, and the underlying physics.

Based on the spatial pattern of mode 1 in Figure 4a, we consider this mode to represent the horizontal motion of ionospheric plasma driven by the IMF  $B_y$  (dawn-dusk) component, following dayside magnetic reconnection at the magnetopause. For ease of reference, we term this mode the “Velocity Perturbation Type Y” (VPY), to reflect its driving by IMF  $B_y$ . The VPY feature is spatially very similar to the “Disturbance-Polar Type Y” (DPY; Friis-Christensen et al., 1985; Friis-Christensen & Wilhelm, 1975; Greenwald et al., 1990) mode shown in Figure 5b. We can contrast these patterns to make inferences about the different behavior of the electric field (which drives the plasma convection) and the magnetic field, as follows.



**Figure 6.** The cross correlation of the leading four Super Dual Auroral Radar Network (SuperDARN) mode time series with (a) Interplanetary Magnetic Field (IMF)  $B_y$  and (b)  $B_z$ , as a function of lag. Positive lags  $\tau$  indicate that the ionosphere responds  $\tau$  minutes following a perturbation at the bow shock nose. The SuperDARN modes time series used are those from Figures 4e–4h. Our interpretation of the modes is: 1: Velocity Perturbation Type Y (VPY) directly driven primarily by IMF  $B_y$ , at 20 min lag, 2: Velocity Perturbation Type 2 expanded (VP2e) directly driven primarily by IMF  $B_z$  at 20 min lag, 3: Velocity Perturbation Type 2 contracted (VP2c) directly driven primarily by IMF  $B_z$  at 15 min lag, 4: substorm mode Velocity Perturbation Type 1 (VP1) indirectly driven by IMF  $B_z$  at 40–80 min lag.

It is worth noting that the dominant (VPY) pattern of plasma velocity variability (Figure 4a, which is controlled by IMF  $B_y$ ) is spatially dissimilar to the monthly mean velocity (Figure 3b, which is controlled by IMF  $B_z$ ). Conversely, for the magnetic field, the dominant variability pattern is spatially alike to the mean field (e.g., compare Figures 2 and 4a of Shore et al., 2017, both of which show a two-cell convection). This implies that for the magnetic field, the (monthly) mean better represents the value at a given epoch, whilst the plasma drift (and electric field) is less well suited to representation by climatological averages. This underscores the importance of accounting for the variability of SuperDARN data when performing imputation to infill data gaps.

In terms of contributions to the total variability, Shore et al. (2018) found that the DPY pattern was only identifiable in the leading modes of geomagnetic perturbation during local summer, when it was always a secondary contributor compared to the two-cell convection pattern (driven by IMF  $B_z$ ). In contrast, the VPY feature is resolved here even in local winter, where it dominates the SuperDARN variation, while variance features related to the two-cell convection (detailed below) are secondary. This implies a greater importance for IMF  $B_y$  in describing the plasma drift (or electric field) variations than in describing the magnetic field variations. The different positions in the variance hierarchy of the VPY and DPY modes are evidence that the IMF  $B_y$  input acts as a voltage source, whose contribution to the magnetic field variance is modulated by ionospheric conductivity.

Further evidence for our physical interpretation of the plasma velocity variance structure is provided by comparing the modes with independent measures of solar-terrestrial coupling. In Figure 6, we cross-correlate the four modes shown in Figures 4e–4h with the month-long time series of the IMF  $B_y$  and  $B_z$  components, as previously done for the magnetic field by Shore, Freeman, and Gjerloev (2019). These correlations are performed at a range of lags (from 0 to 500 min), where a positive lag  $\tau$  indicates that the ionosphere responds  $\tau$  minutes after a perturbation in the IMF at the bow shock nose. The correlations between the modes and the IMF are expected to peak at lags which reflect the time taken to reconfigure the polar ionospheric electrodynamics, following a change in the IMF.

From the SuperDARN mode correlations with IMF  $B_y$ , shown in Figure 6a, we see that only mode 1 (the VPY pattern, corresponding to the black line) has a strong response to this IMF component. The VPY response to IMF  $B_y$  driving is strongest at 20 min' lag—similar to that found for the DPY pattern (Shore, Freeman, & Gjerloev, 2019). In other words, we find that both the plasma drift (i.e., the electric field) and the magnetic field respond to direct driving with the same latency, when ionospheric conductance is high

enough to produce a detectable magnetic perturbation. The long lag time taken for VPY and IMF  $B_y$  to become decorrelated is due to the long autocorrelation of IMF  $B_y$  with increasing lag (see Shore, Freeman, & Gjerloev, 2019).

In their analysis of the geomagnetic variability, Shore et al. (2018) showed that two modes (shown here in Figures 5a and 5c) each represent a separate aspect of the ionospheric two-cell convection, which is driven when IMF  $B_z$  is directed southward (Friis-Christensen & Wilhjelm, 1975; Hairston et al., 2005; Nishida, 1968a). The pattern of geomagnetic variation shown in Figure 5a represents the strength of the two-cell convection. Shore et al. (2017); Shore et al. (2018) termed this mode “Disturbance Polar Type 2” (DP2), to reflect its relationship to the associated equivalent current system (e.g., Nishida, 1968a). According to the Expanding-Contracting Polar Cap (ECPC) paradigm (Cowley & Lockwood, 1992; Lockwood & Cowley, 1992; Siscoe & Huang, 1985), the open-closed field line boundary (and with it, the two-cell convection system) expands and contracts on a variety of timescales, owing to the balance of dayside and nightside magnetic reconnection (Milan et al., 2017). Here, this expansion and contraction is represented by the geomagnetic variation pattern shown in Figure 5c. This mode describes the spatial rate of change of DP2 (when the two patterns are summed), and was termed “DP2 Expansion-Contraction” (DP2EC) by Shore et al. (2017). The DP2EC pattern is essentially the first derivative in latitude of the DP2 pattern. We find that both the amplitude and the spatial extent of the two-cell convection are represented in the SuperDARN modes, albeit in a slightly different manner, as we outline below.

Mode 2 of the plasma velocity (shown in Figure 4b), and mode 3 (Figure 4c), represent an expanded and contracted form of the two-cell convection pattern respectively. For instance, in Figure 4b the plasma velocity direction in the dawn-dusk meridian changes from east to west at latitudes of 75° (dawn) and 72° (dusk). Conversely, in Figure 4c these reversals occur at 79° (dawn) and 81° (dusk). So in the SuperDARN modes, the two-cell convection system is represented in the form of two 2-cell convection patterns, with mode 2 representing the expanded system, and mode 3 describing the contracted system. We term mode 2 “Velocity Perturbation Type 2 expanded” (VP2e), and we term mode 3 “Velocity Perturbation Type 2 contracted” (VP2c). Like the geomagnetic modes DP2 and DP2EC, the VP2e and VP2c modes (when summed with different relative amplitudes) describe the range of expanded or contracted states of the convection system, and its differing intensities. We provide proof of our interpretation of VP2e and VP2c in Figure 7, where the mean field is added to modes 2 and 3, whilst the modes have had different amplitudes (and signs) assigned. The superposition of these three fixed patterns is sufficient to recover the range of the two-cell convection system’s spatial variations: either strong (S) or weak (W), and either expanded (E) or contracted (C) in latitude.

We now compare the response of the different convection modes to IMF  $B_z$  driving. VP2e and VP2c each represent the strength of the two-cell convection pattern, so we expect them to have a direct relationship to IMF  $B_z$ . The correlation between VP2e (i.e., mode 2, in Figure 4f) and IMF  $B_z$  is shown in Figure 6b (blue line). We see that the plasma drift reconfiguration response for this mode peaks with a relatively high correlation coefficient at a lag of 20 min and has a long decorrelation timescale, similar to the driving of VPY by IMF  $B_z$ .

The correlation between VP2c (i.e., mode 3 in Figure 4g) and IMF  $B_z$  is shown via the red line in Figure 6b. This correlation peaks at 15 min lag, which is slightly sooner than the response seen for VP2e. The longer response of VP2e is likely because the development of an expanded convection system requires more driving from IMF  $B_z$ . We also see that VP2c decorrelates (with IMF  $B_z$ ) faster than VP2e does (i.e., there is zero correlation in the red line in Figure 6b between lags of 60–120 min). A possible explanation is that following sustained IMF  $B_z$  driving, a substorm will likely occur, and the VP2c mode would then describe the associated contraction of the convection system. From the low correlation between VP2c and IMF  $B_z$  during lags of 60–120 min, it appears that the nature of the polar cap contraction is not directly related to the IMF  $B_z$  driving in the preceding 1–2 h.

Comparing the plasma drift and magnetic field responses to IMF  $B_z$  at long lag allows us to contrast the differing impact of ionospheric conductivity on these fields. In contrast to the rapid responses to IMF  $B_z$  driving seen for VP2e and VP2c, the response of DP2 to IMF  $B_z$  reported by Shore, Freeman, and Gjerloev (2019, their Figure 2g) exhibits a much broader correlation peak, centered on a lag of 35 min. This long-lag

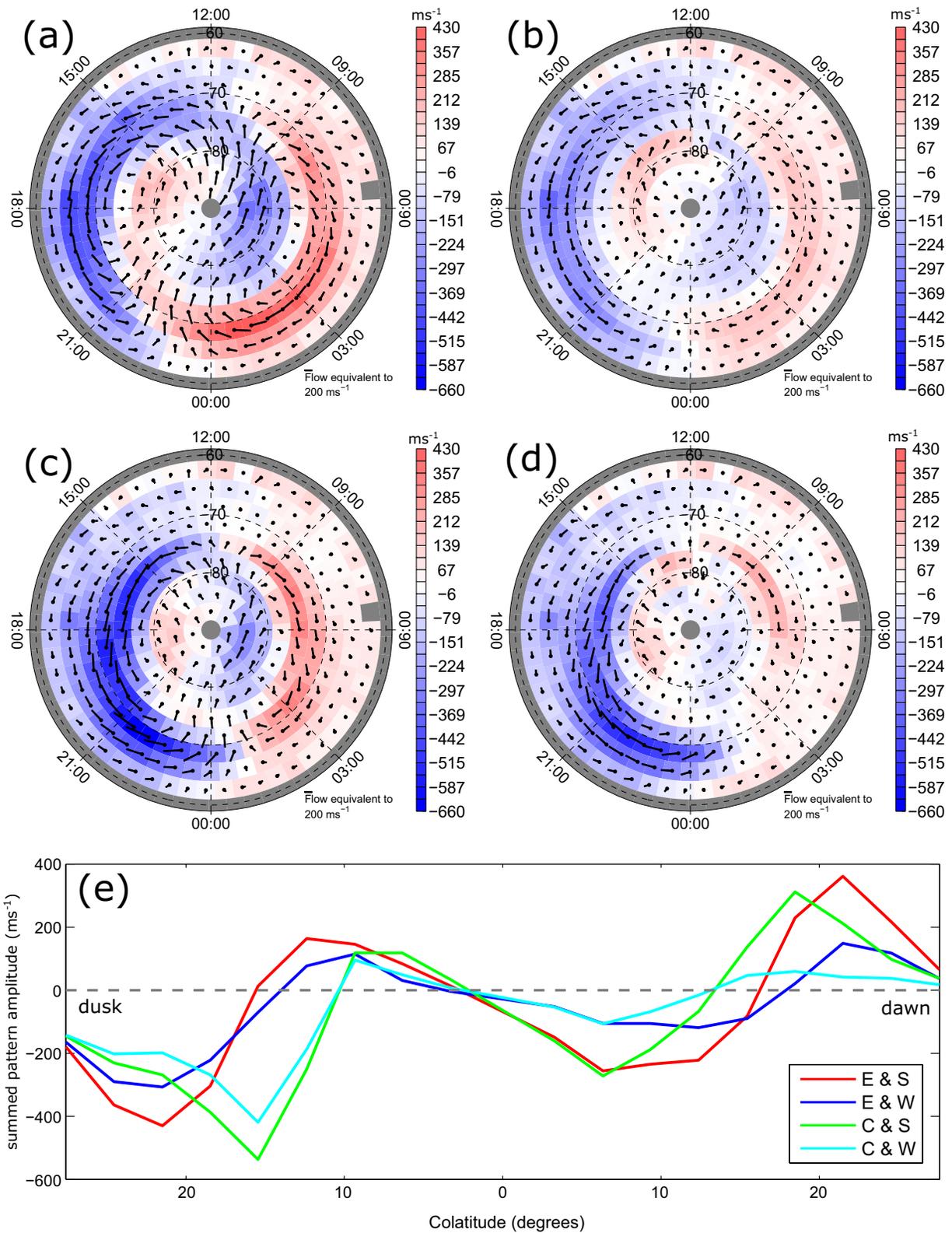


Figure 7

response of the geomagnetic perturbation to IMF driving was also identified by Weimer et al. (2010), with two key findings of relevance to the discussion here. First, Weimer et al. (2010) found a consistent difference between dayside and nightside IMF response timescales, with nightside longer on average. This is expected from Dungey cycle theory (Dungey, 1961), and has since been investigated in more detail by Shore, Freeman, Coxon, et al. (2019). Second, Weimer et al. (2010) found higher correlations between the IMF and nightside geomagnetic variations, compared to dayside variations. The cause of this correlation difference was identified by Shore, Freeman, and Gjerloev (2019). The authors proposed that because sustained IMF  $B_z$  driving will likely lead to a substorm event, then on average we expect increased particle precipitation and conductivity in the auroral zone at longer lags. Following a precipitation event, a unit change in the solar wind driver will cause a larger geomagnetic response. This will increase the correlation between DP2 and IMF  $B_z$  at around 60 min' lag, which is the most common time taken for IMF  $B_z$  driving to create substorm conditions. Since both VP2e and VP2c have narrow peaks in their correlation with lagged IMF  $B_z$ , we infer that the plasma drift (i.e., electric field) variation does not have this feedback from ionospheric conductivity.

Lastly, we highlight mode 4 of the plasma velocity variation, shown in Figure 4d. This pattern is dominated by a nightside signal, which contains two opposing-direction vortices, both at 68° latitude, with one centered near 02:30 MLT, and the other centered near 21:30 MLT. These may be the plasma velocity perturbations associated with the incoming and outgoing field-aligned currents which characterize the substorm current wedge (Frey et al., 2004; McPherron et al., 1973). This interpretation is upheld by the correlation of this pattern's time series (from Figure 4h) with lagged IMF  $B_z$ , which is shown in Figure 6b (green line). The correlation is weakly negative, implying that the mode 4 pattern is driven indirectly by southward IMF. Also, the correlation peaks around 60 min' lag, which is the typical time taken for the IMF to drive a substorm. Further evidence that mode 4 is a substorm mode is given by a superposed epoch analysis and correlation with SMU and SML in the Supporting Information.

## 5. Discussion

In this study, we have described a technique to solve the following issues, which are intrinsic to the SuperDARN archive: incomplete sampling of the plasma velocity vector (i.e., only in the line-of-sight direction), incomplete temporal coverage, and incomplete spatial coverage. To achieve this goal, we have developed a version of data interpolating EOFs (DINEOFs) that is tailored to SuperDARN data. DINEOFs were originally proposed by Beckers and Rixen (2003), but here they are based on the slightly altered formulation of Shore et al. (2017). The primary advance is that the EOF analysis imputes the spatial, temporal and velocity-direction information at any given point based solely on the covariances of the data over a month. In other words, the missing information in time, space, and velocity perpendicular to the radar line-of-sight is infilled at a given location based on the behavior of both the co-located line-of-sight plasma velocity measurements from different line-of-sight directions throughout the month, and also the contemporaneous line-of-sight velocity at other locations. This allows us to avoid the use of external information (e.g., the solar wind), climatological averages, or smooth global models (e.g., spherical harmonic analysis) when constructing a complete reconstruction from the available measurements.

Once the spatial and temporal basis of the data has been completed at every location in the gridded data set by EOF analysis, a sinusoid is fitted to each spatiotemporal location to gain the full velocity vector information from the incomplete directional measurement basis given by the radar network. This is required because the radars can only return line-of-sight velocity, and do not measure the full velocity vector. Since the sinusoid fitting extracts for us a cardinal (north, east) vector basis from the n-dimensional mode spatial patterns, we have subsequently interpreted these patterns physically, to infer information about the behavior

**Figure 7.** The sum of the mean field with modes 2 and 3, where the modes have been assigned differing amplitudes and signs. (a–d) show the two-cell convection system in a range of states: expanded or contracted, and strong or weak. (a) Expanded and strong, (b) expanded and weak, (c) contracted and strong, (d) contracted and weak. (e) Shows a latitudinal cross-section of the plasma velocity vector's east-west component in (a–d), along meridians at 17:20 MLT (dusk) and 03:20 MLT (dawn), to highlight the position of the centers of the two-cell convection system vortices, via the zero-crossing latitude of the east-west component. For reproducibility: if we adopt a normalized maximum amplitude for the mean field pattern (i.e., a value of 1), then the relative maximum amplitudes of the VP2e (mode 2) and VP2c (mode 3) patterns is as follows for each of the cases shown in (a–d), using the format (mode 2 value, mode 3 value) in each case. E&S; (1.03, 0), E&W; (0, -0.51), C&S; (-0.51, 1.29), C&W; (-1.03, 0.26).

of the plasma drift (and electric field) in the polar ionosphere. The temporal behavior of the modes is unaffected by the sinusoid fitting process, and has been used to investigate the solar wind driving of the modes.

EOF analysis has been used in multiple prior studies of high-latitude electric field variability (i.e., plasma velocity variability), for instance by Kim et al. (2012), Cousins et al. (2013), and Walach et al. (2021) using SuperDARN data, and by Matsuo et al. (2002) using Dynamics Explorer-2 satellite data. We can compare their approaches and results with ours. In agreement with our finding that IMF  $B_y$  dominates the plasma velocity variability, Matsuo et al. (2002) found that their primary mode was also best associated with IMF  $B_y$ . In contrast, Kim et al. (2012), Cousins et al. (2013), and Walach et al. (2021) found that their leading mode was best described by IMF  $B_z$ . We suggest two possible reasons for this discrepancy. First, the data set used by Cousins et al. (2013) is considered to be representative of geomagnetically quiet conditions with a bias toward summer, whereas we use data from an entire winter month at solar maximum. In this case, different dynamical modes should be expected from EOF analysis of these different datasets. It may be that the lower variability in the Cousins et al. (2013) data set makes the dynamic behavior closer to the mean behavior, which is dominated by IMF  $B_z$  (refer to Figure 3b). A second possible reason for the discrepancy is that the studies of Kim et al. (2012), Cousins et al. (2013), and Walach et al. (2021) apply smoothing to the SuperDARN data prior to the computation of the EOF modes, either via spatial and temporal regridding to reduce the variability of the data, or via smooth model fits. Kim et al. (2012) additionally used an IMF-dependent model to infill their data gaps. In contrast to these, Matsuo et al. (2002) compute their EOF modes directly on the along-track satellite data, prior to computing the (smoothed) global electric potential from these EOF modes. This matches our approach, in that we seek to allow the EOF analysis maximum access to the full variability of the data, before we apply any smooth models. It may be that this approach better represents the true plasma flow variability. Indeed, we resolve higher peak correlations between our modes and IMF  $B_y$  ( $-0.68$ ) and  $B_z$  ( $-0.47$ ) than either Kim et al. (2012) ( $-0.48$  and  $0.42$  respectively) or Cousins et al. (2013) ( $-0.42$  and  $-0.38$  respectively). Lastly, we note that we resolve a mode which describes the substorm, whereas each of the studies mentioned above do not. This further implies that our technique is more sensitive to the electrodynamic variability.

Regarding the use of our EOF reanalysis as a solution for infilling gaps in the SuperDARN archive, again there is a range of possible options for the user to choose between. Our primary motivation in describing our technique is to create an alternative to the “Map Potential” method for obtaining the global plasma flow pattern, introduced by Ruohoniemi and Baker (1998), and further improved by Thomas and Shepherd (2018). Recently, several additional alternative infill solutions have been proposed, for example, by Gjerloev et al. (2018) and Nakano et al. (2020). These techniques each contrast with ours in a similar manner, in that they rely on additional measurement information to complete the SuperDARN data coverage. For Thomas and Shepherd (2018) and Nakano et al. (2020), this information is climatological averages of the flow for different solar wind states, while Gjerloev et al. (2018) relied on a principal component regression of the flow history onto geomagnetic indices. Conversely, our infill solution is entirely self-consistent, and to our knowledge is the only SuperDARN imputation which requires only SuperDARN data. However, we stress that EOF meets a different need for the SuperDARN community to these other models. We suggest that Map Potential (or the models of Gjerloev et al., 2018 or Nakano et al., 2020) remains the go-to model for nowcasts, i.e., maps of the instantaneous representation of the electric field. EOF offers no intrinsic nowcast or forecast potential (although the dependence of these modes on solar wind variables could be used to develop a forecast model, which we will investigate in future studies). Hence, we suggest that our model technique be used to provide a reanalysis of SuperDARN, for an accurate reconstruction of past variability.

## 6. Conclusions

We present a novel solution for a complete specification of the 2-D plasma velocity field over the entire mid- and high-latitude polar regions from sparse SuperDARN radar data. This is the first SuperDARN infill solution which is entirely self-consistent, in that it relies only on the SuperDARN data, in contrast to the more popular climatological averaging technique of computing the mean velocity field for a given exogenous (e.g., IMF) state. By avoiding climatological averaging, our method also provides superior estimates of the ionospheric electrodynamic variability.

Our technique decomposes the velocity field variability into mathematically orthogonal modes, analogous to Fourier analysis. By investigating the dependence of these discovered modes on the IMF, we find that the dominant mode of variability of the plasma velocity is driven by the IMF  $B_y$  component, while IMF  $B_z$  is the dominant driver of the next three most important modes of variability. Following a perturbation in IMF  $B_y$  at the bow shock nose, the reconfiguration timescale for the plasma velocity is 20 min. For a perturbation in IMF  $B_z$ , the reconfiguration occurs between characteristic timescales of 15 and 60 min.

Based on the spatial morphology of the velocity modes, their dependence on the IMF, and their relationships to previously reported surface magnetic field modes; we define and interpret the mean and the leading four velocity modes as follows:

1. Mode 0 (Mean) = Velocity Perturbation Type 2 (VP2). This is the (monthly) average two-cell convection pattern.
2. Mode 1 = Velocity Perturbation Type Y (VPY) that is directly driven primarily by IMF  $B_y$  with a lag of 20 min from the bow shock. The addition of this mode to the mean two-cell convection mode VP2 creates a dawn-dusk asymmetry and the familiar “banana” and “orange” convection cells in the polar cap. It represents the velocity perturbation caused by dawn-dusk magnetic tension on newly reconnected magnetic field lines at the magnetopause. VPY is analogous to the Disturbance Polar Y mode of surface magnetic field variation.
3. Mode 2 = Velocity Perturbation Type 2 expanded (VP2e) that is directly driven primarily by IMF  $B_z$  with a lag of 20 min. The addition of this mode to the mean two-cell convection mode VP2 expands and enhances the global convection. It represents the velocity perturbation caused by variable magnetic reconnection at the magnetopause and magnetotail.
4. Mode 3 = Velocity Perturbation Type 2 contracted (VP2c) that is directly driven primarily by IMF  $B_z$  with a lag of 15 min. The addition of this mode to the mean two-cell convection mode VP2 contracts and strengthens the global convection. The sum of VP2, VP2e, and VP2c is analogous to the Disturbance Polar 2 and Disturbance Polar Expansion-Contraction modes of surface magnetic field variation.
5. Mode 4 = Velocity Perturbation Type 1 (VP1) that is indirectly driven by IMF  $B_z$  with a lag of 40–80 min. This mode represents the velocity perturbation caused by the substorm current wedge. VP1 is analogous to the Disturbance Polar 1 mode of surface magnetic field variation.

Our technique appears to capture more of the underlying electrodynamic behavior in fewer dynamic modes than the analyses of previous studies. This is probably because we infill the data before applying any smooth model and avoid specifying prior relationships to the solar wind. We envisage that our technique will be used either as a standalone data set for investigating polar-scale phenomena, or as an assumptions-free data set for infilling gaps in the (higher-resolution) SuperDARN measurement set.

### Data Availability Statement

The SuperDARN EOF reanalysis model produced in this study is available at <https://doi.org/10.5285/f4245a21-dee9-46cf-85b2-114798cb7ebc>. OMNI IMF data were downloaded from [ftp://spdf.gsfc.nasa.gov/pub/data/omni/high\\_res\\_omni/monthly\\_1min/](ftp://spdf.gsfc.nasa.gov/pub/data/omni/high_res_omni/monthly_1min/) on 2014-03-06. The geomagnetic EOF model based on SuperMAG data is available in the supporting information of Shore et al. (2018). The SuperDARN data are available from the BAS SuperDARN data mirror (<https://www.bas.ac.uk/project/superdarn>). To gain access to the SuperDARN data, email [bassuperdarn@bas.ac.uk](mailto:bassuperdarn@bas.ac.uk) to set up a user account for the BAS SuperDARN data mirror.

### References

Abel, G. A., Freeman, M. P., & Chisham, G. (2006). Spatial structure of ionospheric convection velocities in regions of open and closed magnetic field topology. *Geophysical Research Letters*, 33(24). <https://doi.org/10.1029/2006GL027919>

Abel, G. A., Freeman, M. P., & Chisham, G. (2009). IMF clock angle control of multifractality in ionospheric velocity fluctuations. *Geophysical Research Letters*, 36(19). <https://doi.org/10.1029/2009GL040336>

Abel, G. A., Freeman, M. P., Chisham, G., & Watkins, N. W. (2007). Investigating turbulent structure of ionospheric plasma velocity using the Halley superdarn radar. *Nonlinear Processes in Geophysics*, 14(6), 799–809. <https://doi.org/10.5194/npg-14-799-2007> Retrieved from <https://npg.copernicus.org/articles/14/799/2007/>

Beckers, J. M., & Rixen, M. (2003). EOF calculations and data filling from incomplete oceanographic datasets. *Journal of Atmospheric and Oceanic Technology*, 20(12), 1839–1856.

### Acknowledgments

This work was funded by the Natural Environment Research Council under grants NE/N01099X/1 (THEMES) and NE/V002732/1 (SWIMMR-T). We would like to thank Angeline Burrell, Steve Browett, Maria-Theresia Walach, and Andrew Kavanagh for discussions, which improved the manuscript, and Evan Thomas and Paul Breen for RST code expertise and assistance. For the use of the SuperMAG ground magnetometer data we gratefully acknowledge: INTERMAGNET (we thank the national institutes that support its contributing magnetic observatories, and INTERMAGNET for promoting high standards of magnetic observatory practice ([www.intermagnet.org](http://www.intermagnet.org))); USGS, Jeffrey J. Love; CARISMA, PI Ian Mann; CANMOS; The S-RAMP Database, PI K. Yumoto and Dr. K. Shiokawa; The SPIDR database; AARI, PI Oleg Troshichev; The MACCS program, PI M. Engebretson, Geomagnetism Unit of the Geological Survey of Canada; GIMA; MEASURE, UCLA IGPP and Florida Institute of Technology; SAMBA, PI Eftyhia Zesta; 210 Chain, PI K. Yumoto; SAMNET, PI Farideh Honary; The institutes who maintain the IMAGE magnetometer array, PI Eija Tanskanen; PENGUIN; AUTUMN, PI Martin Connors; DTU Space, PI Dr. Juergen Matzka; South Pole and McMurdo Magnetometer, PI's Louis J. Lanzarotti and Alan T. Weatherwax; ICESTAR; RAPIDMAG; PENGUIN; British Antarctic Survey; MacMac, PI Dr. Peter Chi; BGS, PI Dr. Susan Macmillan; Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation (IZMIRAN); GFZ, PI Dr. Juergen Matzka; MFGI, PI B. Heilig; IGF PAS, PI J. Reda; University of L'Aquila, PI M. Vellante; SuperMAG, PI Jesper W. Gjerloev. The authors acknowledge the use of SuperDARN data. SuperDARN is a collection of radars funded by the national scientific funding agencies of Australia, Canada, China, France, Italy, Japan, Norway, South Africa, UK, and United States.

- Bjornsson, H., & Venegas, S. (1997). A manual for EOF and SVD analyses of climatic data. *CCGCR (Centre for Climate and Global Change Research) Report*, 97(1).
- Burrell, A. G., Yeoman, T. K., Milan, S. E., & Lester, M. (2016). Phase calibration of interferometer arrays at high-frequency radars. *Radio Science*, 51(9), 1445–1456. <https://doi.org/10.1002/2016RS006089>
- Chisham, G., Lester, M., Milan, S., Freeman, M., Bristow, W., Grocott, A., & Walker, A. (2007). A decade of the super dual auroral radar network (superdarn): Scientific achievements, new techniques and future directions. *Surveys in Geophysics*, 28(1), 33–109. <https://doi.org/10.1007/s10712-007-9017-8>
- Chisham, G., Yeoman, T. K., & Sofko, G. J. (2008). Mapping ionospheric backscatter measured by the SuperDARN HF radars—Part 1: A new empirical virtual height model. *Annales Geophysicae*, 26(4), 823–841. <https://doi.org/10.5194/angeo-26-823-2008>
- Cousins, E. D. P., Matsuo, T., & Richmond, A. D. (2013). Mesoscale and large-scale variability in high-latitude ionospheric convection: Dominant modes and spatial/temporal coherence. *Journal of Geophysical Research: Space Physics*, 118(12), 7895–7904. <https://doi.org/10.1002/2013JA019319>
- Cousins, E. D. P., & Shepherd, S. G. (2010). A dynamical model of high-latitude convection derived from superdarn plasma drift measurements. *Journal of Geophysical Research*, 115(A12). <https://doi.org/10.1029/2010JA016017>
- Cowley, S. W. H. (1982). The causes of convection in the earth's magnetosphere: A review of developments during the ims. *Reviews of Geophysics*, 20(3), 531–565. <https://doi.org/10.1029/RG020i003p00531>
- Cowley, S. W. H., & Lockwood, M. (1992). Excitation and decay of solar wind-driven flows in the magnetosphere-ionosphere system. *Annales Geophysicae-Atmospheres Hydrospheres and Space Sciences*, 10(1–2), 103–115.
- Dungey, J. W. (1961). Interplanetary magnetic field and the auroral zones. *Physical Review Letters*, 6, 47–48.
- Emmert, J. T., Richmond, A. D., & Drob, D. P. (2010). A computationally compact representation of Magnetic-Apex and Quasi-Dipole coordinates with smooth base vectors. *Journal of Geophysical Research*, 115(A8). <https://doi.org/10.1029/2010JA015326>
- Frey, H. U., Mende, S. B., Angelopoulos, V., & Donovan, E. F. (2004). Substorm onset observations by image-fuv. *Journal of Geophysical Research*, 109(A10). <https://doi.org/10.1029/2004JA010607>
- Friis-Christensen, E., Kamide, Y., Richmond, A. D., & Matsushita, S. (1985). Interplanetary magnetic field control of high-latitude electric fields and currents determined from Greenland Magnetometer Data. *Journal of Geophysical Research*, 90(A2), 1325–1338. <https://doi.org/10.1029/JA090iA02p01325>
- Friis-Christensen, E., & Wilhjelm, J. (1975). Polar cap currents for different directions of the interplanetary magnetic field in the Y-Z plane. *Journal of Geophysical Research*, 80(10), 1248–1260. <https://doi.org/10.1029/JA080i010p01248>
- Gjerloev, J. W., Waters, C. L., & Barnes, R. J. (2018). Deriving global convection maps from superdarn measurements. *Journal of Geophysical Research: Space Physics*, 123(4), 2902–2915. <https://doi.org/10.1002/2017JA024543>
- Greenwald, R. A., Baker, K. B., Dudeney, J. R., Pinnock, M., Jones, T. B., Thomas, E. C., & Yamagishi, H. (1995). Darn/superdarn. *Space Science Reviews*, 71(1), 761–796. <https://doi.org/10.1007/BF00751350>
- Greenwald, R. A., Baker, K. B., Hutchins, R. A., & Hanuise, C. (1985). An hf phased-array radar for studying small-scale structure in the high-latitude ionosphere. *Radio Science*, 20(1), 63–79. <https://doi.org/10.1029/RS020i001p00063>
- Greenwald, R. A., Baker, K. B., Ruohoniemi, J. M., Dudeney, J. R., Pinnock, M., Mattin, N., & Lepping, R. P. (1990). Simultaneous conjugate observations of dynamic variations in high-latitude dayside convection due to changes in IMF By. *Journal of Geophysical Research*, 95(A6), 8057–8072.
- Hairston, M. R., Drake, K. A., & Skoug, R. (2005). Saturation of the ionospheric polar cap potential during the October–November 2003 superstorms. *Journal of Geophysical Research*, 110(A9). <https://doi.org/10.1029/2004JA010864>
- Hanuise, C., Greenwald, R. A., & Baker, K. B. (1985). Drift motions of very high latitude F region irregularities: Azimuthal Doppler analysis. *Journal of Geophysical Research*, 90(A10), 9717–9725. <https://doi.org/10.1029/JA090iA10p09717>
- Iijima, T., Potemra, T. A., Zanetti, L. J., & Bythrow, P. F. (1984). Large-scale Birkeland currents in the dayside polar region during strongly northward IMF: A new Birkeland current system. *Journal of Geophysical Research*, 89(A9), 7441–7452. <https://doi.org/10.1029/JA089iA09p07441>
- Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed.). Springer.
- Kim, H.-J., Lyons, L. R., Ruohoniemi, J. M., Frissell, N. A., & Baker, J. B. (2012). Principal component analysis of polar cap convection. *Geophysical Research Letters*, 39(11). <https://doi.org/10.1029/2012GL052083>
- Laundal, K. M., Finlay, C. C., & Olsen, N. (2016). Sunlight effects on the 3D polar current system determined from low Earth orbit measurements. *Earth, Planets and Space*, 68(1), 142.
- Laundal, K. M., Finlay, C. C., Olsen, N., & Reistad, J. P. (2018). Solar wind and seasonal influence on ionospheric currents from swarm and CHAMP measurements. *Journal of Geophysical Research: Space Physics*. <https://doi.org/10.1029/2018JA025387>
- Laundal, K. M., & Gjerloev, J. W. (2014). What is the appropriate coordinate system for magnetometer data when analyzing ionospheric currents? *Journal of Geophysical Research: Space Physics*, 119(10), 8637–8647. <https://doi.org/10.1002/2014JA020484>
- Laundal, K. M., Gjerloev, J. W., Ostgaard, N., Reistad, J. P., Haaland, S., Snekvik, K., & Milan, S. E. (2016). The impact of sunlight on high-latitude equivalent currents. *Journal of Geophysical Research: Space Physics*, 121(3), 2715–2726. <https://doi.org/10.1002/2015JA022236>
- Laundal, K. M., Haaland, S. E., Lehtinen, N., Gjerloev, J. W., Ostgaard, N., Tenfjord, P., & Anderson, B. J. (2015). Birkeland current effects on high-latitude ground magnetic field perturbations. *Geophysical Research Letters*, 42(18), 7248–7254. <https://doi.org/10.1002/2015GL065776>
- Leopardi, P. (2007). *Distributing points on the sphere: Partitions, separation, quadrature and energy* (Unpublished doctoral dissertation). University of New South Wales.
- Lester, M., Milan, S., Provan, G., & Wild, J. (2006). Review of ionospheric effects of solar wind magnetosphere coupling in the context of the expanding contracting polar cap boundary model. In *Solar dynamics and its effects on the heliosphere and earth* (pp. 117–130). Springer.
- Lockwood, M., & Cowley, S. W. H. (1992). Ionospheric convection and the substorm cycle. In *Proceedings of the international conference on substorms (ics-1)* (Vol. 99, p. 109).
- Maezawa, K. (1976). Magnetospheric convection induced by the positive and negative Z components of the interplanetary magnetic field: Quantitative analysis using polar cap magnetic records. *Journal of Geophysical Research*, 81(13), 2289–2303. <https://doi.org/10.1029/JA081i013p02289>
- Matsuo, T., Richmond, A., & Nychka, D. (2002). Modes of high-latitude electric field variability derived from DE-2 measurements: Empirical Orthogonal Function (EOF) analysis. *Geophysical Research Letters*, 29(7). <https://doi.org/10.1029/2001GL014077>
- Matsushita, S., Tarpley, J. D., & Campbell, W. H. (1973). IMF sector structure effects on the quiet geomagnetic field. *Radio Science*, 8(11), 963–972. <https://doi.org/10.1029/RS008i011p00963>

- McPherron, R. L., Russell, C. T., & Aubry, M. P. (1973). Satellite studies of magnetospheric substorms on August 15, 1968: 9. Phenomenological model for substorms. *Journal of Geophysical Research*, 78(16), 3131–3149. <https://doi.org/10.1029/JA078i016p03131>
- Merkin, V. G., Kondrashov, D., Ghil, M., & Anderson, B. J. (2016). Data assimilation of low-altitude magnetic perturbations into a global magnetosphere model. *Space Weather*, 14(2), 165–184. <https://doi.org/10.1002/2015SW001330>
- Milan, S. E., Clausen, L. B. N., Coxon, J. C., Carter, J. A., Walach, M.-T., Laundal, K., & Anderson, B. J. (2017). Overview of solar wind–magnetosphere–ionosphere–atmosphere coupling and the generation of magnetospheric currents. *Space Science Reviews*, 206(1), 547–573. <https://doi.org/10.1007/s11214-017-0333-0>
- Nakano, S., Hori, T., Seki, K., & Nishitani, N. (2020). A framework for estimating spherical vector fields using localized basis functions and its application to SuperDARN data processing. *Earth, Planets and Space*, 72(1), 1–13.
- Nishida, A. (1968a). Coherence of geomagnetic DP 2 fluctuations with interplanetary magnetic variations. *Journal of Geophysical Research*, 73(17), 5549–5559. <https://doi.org/10.1029/JA073i017p05549>
- Nishida, A. (1968b). Geomagnetic DP 2 fluctuations and associated magnetospheric phenomena. *Journal of Geophysical Research*, 73(5), 1795–1803. <https://doi.org/10.1029/JA073i005p01795>
- Nishitani, N., Ruohoniemi, J. M., Lester, M., Baker, J. B. H., Koustov, A. V., et al. (2019). Review of the accomplishments of mid-latitude Super Dual Auroral Radar Network (SuperDARN) HF radars. *Progress in Earth and Planetary Science*, 6(1), 1–57.
- Pettigrew, E. D., Shepherd, S. G., & Ruohoniemi, J. M. (2010). Climatological patterns of high-latitude convection in the northern and southern hemispheres: Dipole tilt dependencies and interhemispheric comparisons. *Journal of Geophysical Research*, 115(A7). <https://doi.org/10.1029/2009JA014956>
- Ponomarenko, P. V., & Waters, C. L. (2006). Spectral width of superdarn echoes: Measurement, use and physical interpretation. *Annales Geophysicae*, 24(1), 115–128. <https://doi.org/10.5194/angeo-24-115-2006>
- Reiff, P. H., & Burch, J. L. (1985). IMF By-dependent plasma flow and birkeland currents in the dayside magnetosphere: 2. A global model for northward and southward IMF. *Journal of Geophysical Research*, 90(A2), 1595–1609. <https://doi.org/10.1029/JA090iA02p01595>
- Reiff, P. H., Spiro, R. W., & Hill, T. W. (1981). Dependence of polar cap potential drop on interplanetary parameters. *Journal of Geophysical Research*, 86(A9), 7639–7648. <https://doi.org/10.1029/JA086iA09p07639>
- Richmond, A. D. (1995). Ionospheric electrodynamics using magnetic apex coordinates. *Journal of Geomagnetism and Geoelectricity*, 47(2), 191–212.
- Ruohoniemi, J. M., & Baker, K. B. (1998). Large-scale imaging of high-latitude convection with super dual auroral radar network HF radar observations. *Journal of Geophysical Research*, 103(A9), 20797–20811. <https://doi.org/10.1029/98JA01288>
- Ruohoniemi, J. M., & Greenwald, R. A. (1996). Statistical patterns of high-latitude convection obtained from goose bay HF radar observations. *Journal of Geophysical Research*, 101(A10), 21743–21763.
- Ruohoniemi, J. M., & Greenwald, R. A. (2005). Dependencies of high-latitude plasma convection: Consideration of interplanetary magnetic field, seasonal, and universal time factors in statistical patterns. *Journal of Geophysical Research*, 110(A9). <https://doi.org/10.1029/2004JA010815>
- Ruohoniemi, J. M., Shepherd, S. G., & Greenwald, R. A. (2002). The response of the high-latitude ionosphere to IMF variations. *Journal of Atmospheric and Solar-Terrestrial Physics*, 64(2), 159–171.
- Shore, R. M., Freeman, M. P., Coxon, J. C., Thomas, E. G., Gjerloev, J. W., & Olsen, N. (2019). Spatial variation in the responses of the surface external and induced magnetic field to the solar wind. *Journal of Geophysical Research: Space Physics*, 124(7), 6195–6211. <https://doi.org/10.1029/2019JA026543>
- Shore, R. M., Freeman, M. P., & Gjerloev, J. W. (2018). An empirical orthogonal function reanalysis of the northern polar external and induced magnetic field during solar cycle 23. *Journal of Geophysical Research: Space Physics*. <https://doi.org/10.1002/2017JA024420>
- Shore, R. M., Freeman, M. P., & Gjerloev, J. W. (2019). Interplanetary magnetic field control of polar ionospheric equivalent current system modes. *Space Weather*, 17(7), 976–988. <https://doi.org/10.1029/2019SW002161>
- Shore, R. M., Freeman, M. P., Wild, J. A., & Gjerloev, J. W. (2017). A high-resolution model of the external and induced magnetic field at the Earth's surface in the northern hemisphere. *Journal of Geophysical Research: Space Physics*. <https://doi.org/10.1002/2016JA023682>
- Siscoe, G. L., & Huang, T. S. (1985). Polar cap inflation and deflation. *Journal of Geophysical Research*, 90(A1), 543–547. <https://doi.org/10.1029/JA090iA01p00543>
- SuperDARN Data Analysis Working Group., Thomas, E. G., Ponomarenko, P. V., Bland, E. C., Burrell, A. G., Kotyk, K., et al. (2017). *SuperDARN radar software toolkit (RST) 4.0*. GitHub repository, Zenodo. <https://doi.org/10.5281/zenodo.801459>
- Thomas, E. G., & Shepherd, S. G. (2018). Statistical patterns of ionospheric convection derived from mid-latitude, high-latitude, and polar SuperDARN HF Radar Observations. *Journal of Geophysical Research: Space Physics*, 123(4), 3196–3216. <https://doi.org/10.1002/2018JA025280>
- von Storch, H., & Zwiers, F. (2002). *Statistical analysis in climate research*. Cambridge University Press. Retrieved from [http://books.google.co.uk/books?id=\\_VHxE26QvXgC](http://books.google.co.uk/books?id=_VHxE26QvXgC)
- Walach, M.-T., Grocott, A., & Milan, S. E. (2021). Average ionospheric electric field morphologies during geomagnetic storm phases. *Journal of Geophysical Research: Space Physics*, 126(4), e2020JA028512. <https://doi.org/10.1029/2020JA028512>
- Weimer, D. R. (1996). A flexible, IMF dependent model of high-latitude electric potentials having “Space Weather” applications. *Geophysical Research Letters*, 23(18), 2549–2552. <https://doi.org/10.1029/96GL02255>
- Weimer, D. R. (2001). An improved model of ionospheric electric potentials including substorm perturbations and application to the geospace environment modeling November 24, 1996, event. *Journal of Geophysical Research*, 106(A1), 407–416. <https://doi.org/10.1029/2000JA000604>
- Weimer, D. R. (2005). Improved ionospheric electrodynamic models and application to calculating Joule heating rates. *Journal of Geophysical Research*, 110(A5). <https://doi.org/10.1029/2004JA010884>
- Weimer, D. R., Clauer, C. R., Engebretson, M. J., Hansen, T. L., Gleisner, H., Mann, I., & Yumoto, K. (2010). Statistical maps of geomagnetic perturbations as a function of the interplanetary magnetic field. *Journal of Geophysical Research*, 115(A10). <https://doi.org/10.1029/2010JA015540>
- Wygant, J. R., Torbert, R. B., & Mozer, F. S. (1983). Comparison of s3-3 polar cap potential drops with the interplanetary magnetic field and models of magnetopause reconnection. *Journal of Geophysical Research*, 88(A7), 5727–5735. <https://doi.org/10.1029/JA088iA07p05727>