Electrojet estimates from mesospheric magnetic field measurements

Karl Laundal\textsuperscript{1,1}, Jeng-Hwa Yee\textsuperscript{2,2}, Viacheslav G. Merkin\textsuperscript{3,3}, Jesper W. Gjerloev\textsuperscript{4,4}, Heikki Vanhamäki\textsuperscript{5,5}, Jone Peter Reistad\textsuperscript{6,6}, Michael Madelaire\textsuperscript{6,6}, and Kareem Sorathia\textsuperscript{3,3}

\textsuperscript{1}University in Bergen
\textsuperscript{2}Johns Hopkins Applied Physics Laboratory
\textsuperscript{3}The Johns Hopkins University Applied Physics Laboratory
\textsuperscript{4}APL-JHU
\textsuperscript{5}University of Oulu
\textsuperscript{6}Birkeland Centre for Space Science, University of Bergen

November 30, 2022

Abstract

The auroral electrojet is traditionally measured remotely with magnetometers on ground or in low Earth orbit (LEO). The sparse distribution of measurements, combined with a vertical distance of some 100 km to ground and typically $>300$ km to LEO satellites, means that smaller scale sizes can not be detected. Because of this, our understanding of the spatiotemporal characteristics of the electrojet is incomplete. Recent advances in measurement technology give hope of overcoming these limitations by multi-point remote detections of the magnetic field in the mesosphere, very close to the electrojet. We present a prediction of the magnitude of these disturbances, inferred from the spatiotemporal characteristics of magnetic field-aligned currents. We also discuss how Zeeman magnetic field sensors (Yee et al., 2020) onboard the Electrojet Zeeman Imaging Explorer (EZIE) satellites will be used to essentially image the equivalent current at unprecedented spatial resolution. The electrojet imaging is demonstrated by combining carefully simulated measurements with a spherical elementary current representation using a novel inversion scheme.
Electrojet estimates from mesospheric magnetic field measurements

K. M. Laundal\textsuperscript{1}, J. H. Yee\textsuperscript{2}, V. G. Merkin\textsuperscript{2}, J. W. Gjerloev\textsuperscript{1,2}, H. Vanhamäki\textsuperscript{3}, J. P. Reistad\textsuperscript{1}, M. Madelaire\textsuperscript{1}, K. Sorathia\textsuperscript{2}, P. J. Espy\textsuperscript{4}

\textsuperscript{1}Birkeland Centre for Space Science, University in Bergen, Norway
\textsuperscript{2}Applied Physics Laboratory, Johns Hopkins University, Laurel, MD 20723, USA
\textsuperscript{3}Space Physics and Astronomy Research Unit, University of Oulu, Finland
\textsuperscript{4}Birkeland Centre for Space Science, Norwegian University of Science and Technology, Norway

Key Points:

\begin{itemize}
  \item We describe a technique to image the electrojet from low Earth orbit using the Zeeman effect
  \item Simulation results show that the technique can resolve meso-scale structures in the electrojet
  \item A novel inversion scheme for spherical elementary current representation is presented
\end{itemize}

Corresponding author: Karl Laundal, karl.laundal@uib.no
Abstract

The auroral electrojet is traditionally measured remotely with magnetometers on ground or in low Earth orbit (LEO). The sparse distribution of measurements, combined with a vertical distance of some 100 km to ground and typically >300 km to LEO satellites, means that smaller scale sizes can not be detected. Because of this, our understanding of the spatiotemporal characteristics of the electrojet is incomplete. Recent advances in measurement technology give hope of overcoming these limitations by multi-point remote detections of the magnetic field in the mesosphere, very close to the electrojet. We present a prediction of the magnitude of these disturbances, inferred from the spatiotemporal characteristics of magnetic field-aligned currents. We also discuss how Zeeman magnetic field sensors (Yee et al., 2021) onboard the Electrojet Zeeman Imaging Explorer (EZIE) satellites will be used to essentially image the equivalent current at unprecedented spatial resolution. The electrojet imaging is demonstrated by combining carefully simulated measurements with a spherical elementary current representation using a novel inversion scheme.

Plain Language Summary

The interaction between the solar wind and the Earth’s magnetic field produces electric currents in the ionosphere which are closely associated with auroral activity. The magnetic effects of these currents have so far been measured remotely, with ground magnetometers which are about 100 km below the currents, or with satellite magnetometers that are even further away, but above the currents. Since the currents have only been measured from a distance, we only know their large-scale structure. This limitation can be overcome by using new sensor technology that can be carried on small satellites in low Earth orbit. Such an instrument would measure oxygen emissions from the upper atmosphere, just below the currents. These emissions change in the presence of a magnetic field due to quantum effects, and can therefore be used to infer magnetic disturbances. We demonstrate a technique to create high-resolution 2D maps of the magnetic field disturbances, using simulated data from a proposed satellite mission.

1 Introduction

The first attempts to relate ground magnetic disturbances to electric currents in space were carried out more than a century ago. Birkeland (1908) presented a horizontal two-cell equivalent current system which is reminiscent of maps derived from modern magnetometer networks (Waters et al., 2015). Birkeland further proposed a 3D structure of the space currents that involved magnetic field-aligned currents. This idea remained controversial until it was confirmed by early magnetometer measurements in space (Zmuda et al., 1966). We now view the 3D ionospheric current system as composed of Birkeland currents, that flow along magnetic field lines, and a horizontal current that is confined to a thin conducting layer of the ionosphere, mainly around 100–120 km altitude. The relationship between ground magnetic field observations and this 3D current system is ambiguous, and we therefore often interpret ground magnetic field observations in terms of an equivalent 2D current system. At high latitudes, the equivalent current is nearly identical with the divergence-free part of the horizontal current (e.g., Untiedt & Baumjohann, 1993; Fukushima, 1976). In this paper, we use the term electrojet as synonymous with the equivalent current, although it is sometimes used to refer to specific parts of it.

It can currently be argued that the spatiotemporal structure of Birkeland currents is better known than the electrojet. Since the Birkeland current magnetic fields are measured in-situ with high-frequency satellite magnetometers, spatial structures as small as ~ 1 km can be investigated (Neubert & Christiansen, 2003). On the other hand, the electrojet magnetic field is measured at least ~ 100 km below the currents using ground
magnetometers, or even further away above the currents using satellites (Olsen, 1996; Laundal et al., 2016). Due to the large distance between the current and measurements, the small-scale structures of the electrojet is unknown. Measurements of the magnetic field at high altitudes, close to the horizontal ionospheric currents, would therefore provide new insight into structure and evolution of the electrojet system. Magnetic field measurements from the upper atmosphere would also represent an electrojet measurement with less contribution from ground induced currents (Juusola et al., 2020).

There are ongoing efforts to develop measurement techniques that would allow for regular sampling of the magnetic field closer to the ionospheric currents. Kane et al. (2018) demonstrated a technique to measure the magnetic field at about 100 km using a high-power pulsed laser beam to optically pump the mesospheric sodium to a spin-polarized state, and a telescope to detect backscattered light. By changing the laser pulse frequency, a resonant frequency was detected which matches the Larmor frequency. The magnetic field was inferred from the Larmor frequency. The technique currently requires integration times that are longer than typical variations in the polar electrojet, and further development would therefore be needed to become truly useful for investigating small-scale variations in the current. Such efforts are underway by several groups.

Another approach was demonstrated by Yee et al. (2017). They used the Microwave Limb Sounder (MLS) on the Aura spacecraft to infer magnetic field disturbances based on the Zeeman effect. The MLS measures radiance spectra from the O$_2$ 118 GHz line in order to infer atmospheric properties. However, the emissions are strongly affected by the Zeeman effect, which creates a split in the emission line that depends on the ambient magnetic field. Yee et al. (2017) showed that magnetic fields could be retrieved from these microwave spectra, and that variations in the magnetic field are in agreement with well-known electrojet properties.

Yee et al. (2017) also discussed how future more compact instruments for Zeeman magnetic field sensing could give new insight into the spatiotemporal behavior of the electrojet. Yee et al. (2021) presented a new conceptual instrument design that could be miniaturized and placed on a CubeSat. Such an instrument will fly on NASA’s Electrojet Zeeman Imaging Explorer (EZIE). In this paper, we use simulated data from the EZIE mission, with a realistic ionospheric current system, main magnetic field, and instrument response, to show how it could be used to essentially image the electrojet and associated magnetic field.

In Section 1.1 we give a more quantitative description of the electrojet magnetic field on different heights. In Section 2 we describe the EZIE satellite mission. We give a brief review of the proposed measurement principles, which allow us to derive magnetic field disturbances and equivalent currents from observations of the Zeeman split of microwave emissions from mesospheric O$_2$. The main purpose of this paper is to present a novel technique, detailed in Section 2.3, to utilize such magnetic field measurements to image the electrojet. We demonstrate the technique’s feasibility using simulated data. In Section 3 we discuss potential improvements, challenges, and implications of the technique. Section 4 concludes the paper.

### 1.1 Electrojet magnetic field radial dependence

Since the mesosphere is resistive and presumably free of electric current, it is expected that magnetic disturbances there are associated with the same part of the ionospheric electric current system as is observed from ground. The variation in magnetic field strength as function of distance from the electrojet depends on scale size; large-scale currents are seen at greater distances than small-scale currents (e.g., Pulkkinen et al., 2006). Mesospheric magnetic field measurements will therefore enable us to resolve smaller spatial scales than what can be achieved with ground measurements. The purpose of this
Figure 1. A) Average magnitude of the magnetic field disturbances associated with FACs as function of scale-size. The bold curve represents estimates by Gjerloev et al. (2011) based on data from the ST-5 mission. The gray curve is an extrapolation to larger scale-sizes based on a linear fit in log-log space. B) Bold gray: Altitude variation of the magnetic field associated with an electrojet whose spatial structure is given by the gray curve in panel A. Dashed: Altitude variation of magnetic field associated with equivalent current of certain scale sizes, as determined by spherical harmonic degree $n$. Each curve is normalized so that the magnetic field perturbation is 500 nT at 110 km. C) Contour plots of magnetic field of an electrojet whose spatial structure is given by the gray curve in panel A but is otherwise random. The magnetic field is shown at 85 km and on ground, and a map of North America is provided for scale.

The key assumption that we use is that the spatial power spectrum of the electrojet is proportional to the spatial power spectrum of the field-aligned electric currents (FACs). In contrast to the electrojet power spectrum, empirical estimates of the FAC spatial power spectrum are available. Figure 1A shows the spatial power spectrum of the magnetic field associated with FACs in bold black, from Gjerloev et al. (2011). The spectrum is based on magnetic field measurements from the three ST-5 satellites, which flew in a pearl-on-a-string configuration in polar low Earth orbit. Since this orbit intersects the FACs, the determination of small spatial scales is not restricted by distance, as for the electrojet. This particular spatial power spectrum is valid for the nightside during disturbed conditions (AL index $<-100$ nT). The spectrum is close to linear on a log-log scale, and we use this property to extrapolate to scale sizes which are longer than those considered by Gjerloev et al. (2011). The fitted and extrapolated curve is shown in gray.

The curves in Figure 1A, which represents the average magnitude of FAC magnetic field disturbances as function of scale size, is now assumed to also describe the spatial scale of the electrojet at zero distance from the current sheet. With this assumption, we can derive the radial variation of the magnetic field using results from spherical harmonic analysis. Equation 118 of Sabaka et al. (2010) describes the squared magnetic field, averaged over a sphere at radius $r \leq R$, where $R$ is the current sheet radius, as

$$\langle B(r)^2 \rangle = \sum_{n=1}^{\infty} n \left( \frac{r}{a} \right)^{2n-2} \sum_{m=0}^{n} \left[ (q_n^m)^2 + (s_n^m)^2 \right] = \sum_{n=1}^{\infty} n \left( \frac{r}{a} \right)^{2n-2} A_n. \quad (1)$$

$n$ and $m$ are spherical harmonic degree and order, respectively; $a = 6371.2$ km is a reference radius; and $q_n^m$ and $s_n^m$ are spherical harmonic coefficients. On the right hand side, the sum over spherical harmonic order $m$ is written as $A_n$. The terms in the sum rep-
resent the extent to which each degree $n$ contributes to the squared magnitude of the magnetic field.

The two $x$ axes of Figure 1A represent scale size $L$ (bottom) and spherical harmonic degree $n$ (top). They are related by assuming that the scale size is equal to the meridional wavelength $\lambda$ of the spherical harmonics, which is related to wavenumber $n$ by Jean’s formula $n = \frac{2\pi}{\lambda} = \frac{1}{2}$. $R_{FAC}$ is the radius at which the FAC power spectrum was evaluated by Gjerloev et al. (2011), 200 km altitude. This relationship allows us to find an estimate for $A_n$ in equation (1), and thus calculate the average magnetic field magnitude at other radii using the same equation. The result is shown as a gray line in Figure 1B. This curve can be interpreted as a prediction of how the electrojet magnetic field decreases with distance, assuming that the electrojet has a similar spatial structure as FACs, and that ground induced current contributions are negligible. The dashed lines show the altitude variation of the magnetic field associated with an electrojet of specific scale sizes, calculated by evaluating the terms in equation (1) for $n = 40, 80, \text{ and } 200$. These wavenumbers correspond to scale sizes of approximately 1000 km, 500 km, and 200 km, respectively. All curves in Figure 1B are normalized so that their magnitudes are 500 nT at 110 km. The figure shows that small scale magnetic field structures are reduced much more quickly than large scale features, and suggests that more detail could be resolved with magnetic field measurements from mesospheric altitudes than with ground magnetometers.

To further visualize the difference in the magnetic field structure at 85 km and on ground, we show in Figure 1C contour plots for a random electrojet whose spatial structure is given by the gray line in Figure 1A. The random magnetic field is constructed by assigning random spherical harmonic coefficients $q_0^n$ and $s_0^n$ which obey $(q_0^n)^2 + (s_0^n)^2 = A_n$. Longitudinal variations ($m > 0$) are ignored. Although the magnetic field in this figure is random, it has a realistic spatial power spectrum. The figure thus visualizes the difference in spatial structure in the magnetic field disturbances at ground and in the mesosphere.

The EZIE magnetic field measurements are conceptually very different from traditional in-situ magnetometer measurements. As described above, traditional magnetometers’ capacity to resolve spatial scales in the corresponding current is limited by the distance between the current and the magnetometer. With EZIE, the ability to resolve spatial scales in the corresponding current is limited by the distance between the current and the molecules that emit the observed radiation ($\approx 30$ km in the case of EZIE), and by the resolution with which the instrument can resolve the radiation. This resolution will deteriorate with increasing distance between the emissions and detector, but since the distance from O$_2$ emission to current remains the same, this will be at a much slower rate than the deterioration of spatial information in the magnetic field with distance from the current. A good example of this is the magnetograms produced by an optical instrument on the Solar Dynamics Observatory (SDO) spacecraft. Such magnetograms reveal spatial structures in the Sun’s magnetic field which are several orders of magnitude smaller than the distance between the Sun and SDO.

2 Estimating the electrojet from simulated mesospheric magnetic field measurements

EZIE is a NASA Heliophysics Mission of Opportunity scheduled to launch in 2024. It will consist of three satellites in low Earth orbit, equipped with four Zeeman magnetic field sensors each. These sensors will point towards the mesosphere separated in the cross-track direction (perpendicular to the satellite ground track), and thus observe emissions along four tracks as the satellite passes.
Figure 2. An illustrated outline of this paper: A high-resolution MHD simulation snapshot (left, described by Sorathia et al. (2020) and in Section 2.1) of the radial magnetic field at 80 km (color contours) and horizontal ionospheric current densities (green vectors). The MHD simulation is used to calculate realistic magnetic field perturbations in the field-of-view covered by an EZIE satellite during a 4 min time window (second panel). These perturbations are used to simulate the Zeeman split of the 118 GHz O$_2$, originating from the mesosphere at about 80 km (third panel, described by Yee et al. (2021)). The emissions are then used to produce realistic estimates of the magnetic field disturbance measurements, including noise (fourth panel, described in Section 2.2). This is described in Yee et al. (2021) and in Section 2.2. Finally, the simulated measurements shown in the fourth panel are used to estimate 2D maps of the electrojet and associated magnetic field (right panel, described in Section 2.3). The last step is the main focus of this paper.

2.1 Ionospheric current and magnetic field simulation

The electric currents and magnetic field disturbances that we use to simulate EZIE measurements are taken from a snapshot of a global magnetosphere simulation published recently by Sorathia et al. (2020). This simulation of a synthetic substorm used the Grid Agnostic Magnetohydrodynamics with Extended Research Applications (GAMERA) code (Zhang et al., 2019) at an unprecedentedly high spatial resolution approaching the ion kinetic scales in the central plasma sheet and ~30 km azimuthally in the auroral ionosphere. We take advantage of this unprecedented high spatial resolution by selecting a region with prominent meso-scale (~100~500 km) electrojet structures for simulated EZIE overflight discussed further in the following sections.
While the original simulation by Sorathia et al. (2020) used a uniform Pedersen ionospheric conductance, for the numerical experiment presented here, we used the same field-aligned currents but replaced the conductance with the full auroral model (Fedder et al., 1995) to produce a realistic distribution of both Pedersen and Hall conductances. Using this conductance model, the standard ionospheric potential solution was obtained using a version of the Magnetosphere-Ionosphere Coupler/Solver (MIX) code (Merkin & Lyon, 2010) rewritten for GAMERA (dubbed REMIX). Thus, the distribution of the horizontal ionospheric currents was derived. In combination with the field-aligned currents, it was then used to derive the magnetic perturbation vectors at the EZIE measurement altitude using the Biot-Savart integration (Rastätter et al., 2014).

The magnetic field perturbations from the MHD simulation are used as input to obtain a set of simulated measurements, described in Section 2.2. In Section 2.3 we use these simulated measurements to estimate the corresponding electrojet, and compare the estimated electrojet to the original MHD simulation. The MHD simulation electrojet is calculated by extracting the divergence-free part the ionospheric currents $j_{df}$. This is achieved by use of Helmholtz’ theorem which implies that $j_{df} = j - j_{cf}$. The ionospheric horizontal current $j$ and the field-aligned current $j_{\parallel}$ of the MHD simulation are well defined everywhere, and the curl-free part $j_{cf}$ can be calculated from $j_{cf} = \nabla \Psi$ where $\Psi$ is the solution to $\nabla^2 \Psi = -j_{\parallel}$ (e.g., Laundal et al., 2015).

2.2 Radiation simulation and magnetic field inversion

The magnetic field perturbations described in the previous section are used here to calculate synthetic EZIE magnetic field measurements. The measurement concept is based on the observation of spectra and polarization of microwave emissions near the 118 GHz O$_2$ line which, because of the Zeeman effect, depend on the magnetic field in the mesosphere where the emissions are produced.

The Zeeman effect, discovered in 1897 (Zeeman, 1897), is a splitting of the spectral lines of atomic or molecular emissions that depends on the ambient magnetic field. The 118 GHz O$_2$ emission line is split in three in the presence of a magnetic field: One unshifted line ($\pi$), and two lines that are shifted to higher or lower frequencies, $\sigma$. The magnitude of the shift depends on the magnetic field strength $B$: $\sigma = \pi \pm 14.012 B$ Hz, with $B$ given in nT (Yee et al., 2021). Furthermore, the relative intensities of the three lines change with polarization and viewing angle relative to the orientation of the magnetic field where the emissions occur. By measuring the spectrum of polarized electromagnetic radiation near 118 GHz, the magnetic field intensity and orientation can be inferred.

To simulate the radiation that will be observed by EZIE we use the Atmospheric Radiative Transfer Simulator (based on formulations by Larsson et al. (2014)) with an MSIS atmosphere. The modifications of the radiation introduced by the Zeeman effect are modeled by assuming that the magnetic field is a sum of the International Geomagnetic Reference Field model (Thébault et al., 2015) and the MHD simulation magnetic field described in the previous section. The simulated radiation is finally passed through a simulation of a realistic instrument response to yield synthetic measurements of microwave spectra and polarization. The result is a set of spectra that mimic what EZIE will observe, including noise.

These simulated microwave measurements are then used to retrieve magnetic field disturbances that correspond to the observed Zeeman shift and polarization. To do this, we handle the synthetic microwave measurements in the same way as we plan to handle real measurements: The forward model described above, except for the MHD part, is constrained in an inversion such that the disturbance magnetic field fits the observed spectra. The perturbation magnetic field components are thus simultaneously and iteratively retrieved along with their error covariances from the microwave measurements.
and their estimated uncertainties. This concept is described in more detail by Yee et al. (2021).

The precision of the resulting magnetic field estimates depends on how well the frequencies and relative intensities in different polarizations can be resolved. The 118 GHz O$_2$ line is an ideal choice for this purpose (Yee et al., 2021), since it is bright relative to the background and since the Zeeman split is relatively strong. The measurements will be most accurate during polar summer when low mesopause temperatures ensure that the split emission lines are easily distinguishable. These emissions are also uniformly distributed geographically and present at all local times, a requirement for the interpretation of magnetic field observations.

Each of the three EZIE satellites will carry four sensors each allowing the determination of the mesospheric magnetic field. The measurement concept builds on the Microwave Limb Sounder on the Aura satellite (Waters et al., 2006), which measures radiance spectra from the atmospheric limb. Yee et al. (2017) demonstrated that MLS spectra near the 118 GHz O$_2$ line could be used to retrieve the magnetic field in the mesosphere where the emissions originate. The EZIE sensors differ from the MLS in that they focus only on the O$_2$ emissions, and therefore are far more compact, weigh less, and require less power. Instead of pointing at the limb they will observe in a near nadir direction, providing vastly improved geo-location of the emissions and thus the magnetic field measurement.

In this paper, we use simulated measurements from the four sensors on-board a single satellite to retrieve the auroral electrojet in the region scanned by the satellite. The viewing geometries correspond to sensors mounted on a sun-synchronous satellite at 500 km, and pointing at fixed angles in the cross-track direction. The EZIE measurement concept would give one magnetic field measurement per sensor every 2 seconds. With four simultaneous measurements for each satellite, we get measurements along four tracks in a push-broom configuration. The colored lines in the top panel of Figure 3 shows the paths formed by the measurement points in a four minute interval as the simulated satellite crosses the auroral zone. The lower panels in Figure 3 show the magnetic field components along each track, estimated from simulated microwave emissions. The colors correspond to the trajectory of the same color in the top plot. The solid lines show the magnetic field according to the MHD simulation, and the dots show the realistic simulated measurements including noise and other complicating effects. Notice that the vertical (nearly magnetic field-aligned) component is much more precisely determined than the two horizontal components. This is due to the viewing geometry with respect to the orientation of the magnetic field (Yee et al., 2021).

### 2.3 Electrojet inversion

Here we describe a procedure to use the measurements of the previous section to estimate continuous functions to represent the disturbance magnetic field and associated equivalent current in the region spanned by the Zeeman magnetic field sensors as the satellite traverses. To accomplish this, we use a divergence-free spherical elementary current representation. The spherical elementary current system (SECS) technique was developed by Amm (1997), and has since been widely used to estimate ionospheric currents from magnetic field measurements on ground and in space (Amm et al., 2015). The key idea is to model the ionospheric current as the sum of contributions from a set of basis functions that are centered at nodes strategically placed on a spherical shell at an ionospheric radius $R_I$. In our case, the nodes are placed at the center of the grid cells shown in Figure 3. Each basis function describes a horizontal surface current that circulates the node.

The magnetic field at location $r$ can then be modeled as the combined effect of a set of $K$ divergence-free elementary currents. A Biot-Savart integral over these surface
Figure 3. Top: SECS grid (black mesh) and measurement locations (colors) for a 4-min segment of simulated EZIE measurements. The spherical coordinate grid represents geomagnetic latitude and longitude. Bottom: The three components of the magnetic field retrieved from the simulated observed microwave spectra (dots) and the magnetic field according to the MHD simulation (solid lines). The colors correspond to the trajectory of the same color in the top plot. The x axis is common among all plots, and the space between tick marks is 200 km. Notice the different scales of the y axes in the horizontal and radial components. The lines represent the magnetic field disturbances from the MHD simulation along the measurement tracks.
currents is expressed as

\[ \mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_S \left[ \sum_{j=1}^{K} S_j \cot \left( \frac{\phi_j}{2} \right) \hat{\phi}_j \right] \times \hat{\mathbf{r}}' \quad \frac{dS}{\| \mathbf{r} - \mathbf{r}' \|^2} \]  \tag{2} \]

where primes denote the variable of integration \((\mathbf{r}')\), and the integral is over the entire spherical shell at \(r = R_I\). The expression in square brackets is the divergence-free surface current density, and the summation index refers to the node at the center of each grid cell in Fig. 3. The nodes have amplitudes \(S_j\) and are located at \(\mathbf{r}_j\). \(\theta_{\mathbf{r}\mathbf{r}'_j}\) is the angle between \(\mathbf{r}_j\) and \(\mathbf{r}'\), and \(\hat{\phi}_j\) is a unit vector in the eastward direction in a coordinate system whose north pole is at \(\mathbf{r}_j\). Amm & Viljanen (1999) presented a closed-form solution to the integral which does not depend on primed variables, which is what we use here. See also the review paper by Vanhamaki & Juusola (2020) for a more detailed overview of the technique and for the full set of relevant equations.

Given a set of \(K\) node locations \((\mathbf{r}_j\) in equation 2), the amplitudes \(S_j\) of each divergence-free current basis function can be estimated from a set of \(N\) magnetic field component measurements by solving an \(N \times K\) set of linear equations. However, our measurements are non-uniformly distributed and fewer than the number of nodes. The solution to the under-determined set of equations is therefore highly dependent on the choice of grid (node locations), and on the way that the inverse problem is regularized (Vanhamaki & Juusola, 2020). When applying the SECS technique to EZIE data, it is critical that the grid and regularization technique are selected such that variations reflect geophysical changes and not changes in geometry, for example as the satellite moves.

We solve this problem by choosing a grid of nodes that changes minimally relative to the measurements as the satellite moves. We use a grid which is regular in a cubed sphere projection (Ronchi et al., 1996). This projection maps points on the sphere to a circumscribed cube. We only need to project to one of the sides of the cube since we focus on a relatively small region. We center this side on the satellite location at the time in the middle of the measurement segment, and align it with the satellite velocity vector. The grid, satellite track, measurement tracks, and geomagnetic coordinate contours are all shown in this projection at the top of Figure 3. Notice that the grid extends beyond the measurement tracks. The purpose of this is not to extrapolate, but to allow the exterior nodes to represent a uniform background current density (Vanhamaki & Juusola, 2020).

The SECS current amplitudes, \(m = [S_1, S_2, \ldots, S_K]^\top\), are solutions to the set of equations

\[ \mathbf{d} = \mathbf{G}m \]  \tag{3} \]

where \(\mathbf{d}\) is a column vector of measured magnetic field components and \(\mathbf{G}\) is the design matrix relating the magnetic field components and the DF SECS amplitudes according to equation 2. In this case, since all three magnetic field components are the result of an inversion from the same spectrum, the errors are correlated. That means that the effective number of equations in (3) is less than the number of elements in \(\mathbf{d}\). The system of equations can in principle be solved for \(m\) by generalized least squares (e.g., Riley et al., 2006, Chapter 31) by minimizing

\[ f_0 = (\mathbf{d} - \mathbf{G}m)^\top \mathbf{V}^{-1} (\mathbf{d} - \mathbf{G}m), \]  \tag{4} \]

where \(\mathbf{V}\) is the data covariance matrix, which contains off-diagonal terms due to the correlated errors. \(\mathbf{V}\) is known from the magnetic field inversion.

However, the inverse problem is under-determined since the density of SECS poles is higher than the density of measurements almost everywhere. Therefore, additional information must be provided to yield physically meaningful solutions. We choose to add two terms to the cost function (4), to minimize the norm of the amplitude vector \(L_2\).
regularization), and the gradient of SECS poles along magnetic circles of latitude. The total cost function is
\[ f = f_0 + \lambda_1 \|Im\|_2 + \lambda_2 \|Lem\|_2, \]  
(5)
where \(\lambda_1\) and \(\lambda_2\) are damping parameters, \(I\) is the \(K \times K\) identity matrix, and \(\| \cdot \|_2\) indicates the Euclidean norm. \(L_e\) is a matrix that, when multiplied by \(m\) gives an estimate of the gradient of the SECS amplitudes in the magnetic eastward direction. We choose to penalize solutions that show variations in the magnetic longitudinal direction knowing that the electrojet tends to be extended in the longitudinal direction or in other words gradients in magnetic latitude typically exceed those in magnetic longitude.

In our implementation, the longitudinal gradient estimates are based on a first-order central difference scheme. Our choice of a regular grid of SECS poles in cubed sphere projected coordinates, together with the equations provided by Ronchi et al. (1996), greatly simplifies the calculation of \(L_e\). The grid resolution is to some extent linked to the gradient evaluation, since the difference scheme accuracy increases with grid density. The number of nodes may thus potentially be reduced by increasing the order of the difference scheme. The need for a regular grid ostensibly removes one of the advantages of using the SECS representation: That pole density can be adjusted according to data density. However, variations in data density could be taken into account via regularization, by damping variations more strongly in regions with sparse data. We forgo this option here for simplicity.

The solution \(m\) that minimizes the cost function (5) can be written as
\[ m = (G^TV^{-1}G + \lambda_1 I + \lambda_2 L_e^T L_e)^{-1}(G^TV^{-1}d). \]  
(6)
This way of solving the under-determined problem is different from the approach traditionally taken in SECS analysis, which is to ensure a smooth solution by truncated singular value decomposition. In our scheme, the damping parameter \(\lambda_1\) plays the role of the singular value truncation level in traditional SECS analysis.

The flowchart in Figure 4 gives an overview of the inputs and output of the electrojet inversion algorithm described in this section.

2.3.1 Results

Figure 5 shows six views of the region covered by the interior part of the grid in Figure 3. The grid is rotated 90°, and the satellite tracks are shown in the right column with similar color as in Fig. 3. The left column shows maps of the magnetic field perturbations of the MHD simulation at 80 km, with the component indicated in the top left corners. The gray vector field is the same in all six plots, and indicates the divergence-free part of the horizontal ionospheric current of the MHD simulation. The right column shows inversion results based on the data points shown in Fig 3, and the inversion scheme described above. The black vector field is the associated equivalent current.

Comparisons between inversion results and MHD simulation output show that the meso-scale features of the disturbance magnetic field are retrieved by the inversion. This is true for all three components despite the significant differences in noise between the components demonstrated in Figure 3. This is possible because the magnetic field components are not independent, but manifestations of the same electrojet. In our case, the radial component is most precisely measured, and therefore most important in the SECS inversion. Thus in principle, the relatively accurate measurement of the radial component can help increase the precision of the horizontal components via their relationship to an equivalent current.

Figure 5 shows that the match between MHD simulation and inversion result is better where the spacing between measurement tracks is small. Since the cost function (equa-
Figure 4. Flowchart of the electrojet inversion described in Section 2.3.

The comparison in Figure 5 shows that the EZIE measurement concept will give 2D maps which reflect prominent meso-scale features in the electrojet and magnetic field disturbances in the region enclosed by measurement tracks. This will allow us to address outstanding science questions about the structure and evolution of the substorm current wedge (Kepko et al., 2015). The quantitative agreement between the inversion output and the true disturbance field will depend on the proximity to the measurement tracks.

3 Discussion

We have shown how a low Earth orbiting satellite equipped with four sensors that observe the Zeeman split of mesospheric 118 GHz O$_2$ emissions can be used to produce 2D maps of the auroral electrojet. The background for this study is the demonstration of the measurement concept by Yee et al. (2017), the recent development of instrument technology (Yee et al., 2021), and NASA’s decision to implement the Electrojet Zeeman Imaging Explorer (EZIE) satellite mission.

We have presented a novel technique to use spherical elementary current systems (SECS, Amm (1997)) to represent the electrojet which corresponds to simulated EZIE magnetic field measurements from 80 km altitude. The simulations involve high-resolution
Figure 5. Comparisons between the MHD model output (left) and the SECS inversion results based on simulated EZIE measurements (right). Each column correspond to different magnetic field components. The divergence-free part of the MHD simulation ionospheric horizontal current is shown as gray arrows, repeated in all panels. Compare this to the divergence-free current of the inversion shown as black arrows in the right column. A $200 \times 200$ km grid is shown in the top left panel to indicate the scale sizes of the structures.
realistic background (Thébault et al., 2015) and perturbation (Sorathia et al., 2020) magnetic fields, and a realistic instrument response. The electrojet has so far been measured almost exclusively from distances of ∼100 km (ground based magnetometers) and with a network of sparse non-uniformly distributed stations that are fixed in the rotating Earth frame. These observational limitations can be overcome by a pearls-on-a-string mission such as EZIE. The good match between our SECS representation of the electrojet and the simulation shows that the measurement technique has the potential to fill this important knowledge gap.

3.1 Physical interpretation of the electrojet

As mentioned in the introduction, the equivalent current / electrojet is a theoretical horizontal sheet current whose magnetic perturbations are equivalent with the observed magnetic field perturbations under the ionosphere. Although the relationship to real 3D current systems is ambiguous, certain properties are helpful in the physical interpretation of equivalent currents: First of all, at high latitudes, where magnetic field lines are almost vertical, the equivalent current is nearly identical to the divergence-free part of the horizontal current. However, the divergence-free part of the horizontal current is also a rather abstract quantity, and it can be non-zero in regions which are current-free (e.g., Laundal et al., 2015). If we can assume that the electric field in the frame of the neutral wind is a potential field and parallel with conductance gradients, the equivalent current at high latitudes is equal to the Hall current. Furthermore, if the Hall/Pedersen conductance ratio is constant, field-aligned currents are directly proportional to the curl of the equivalent current/electrojet (e.g., Amm et al., 2002). If these assumptions are violated, the relationship between the electrojet and the true 3D current system may be determined by combining with other measurements (auroral precipitation, ionospheric convection, field-aligned currents) (Richmond & Kamide, 1988).

Dependent on the science issue at hand, these subtleties may not be relevant. For example, present theories concerning the composition of the horizontal segment of the substorm current wedge are distinguishable by their predictions of a continuous versus structured horizontal current channel. Such differences would be directly reflected in the 2D equivalent current. In this case the main difference between the 2D equivalent current and the true 3D current would be their closure. Since the equivalent current by definition is horizontal, current channels that in reality connect to field-aligned currents will appear to close via large-scale horizontal return currents that enclose the channels (e.g., Laundal et al., 2018).

3.2 Effects of temporal variations

The electrojet retrieval presented here implicitly assumes that the current system remains static in the 4 min interval of the analysis. This assumption was automatically fulfilled, since the simulated magnetic field measurements are based on a single snapshot of magnetic field disturbances from an MHD simulation. However, as a mission such as EZIE traverses the region of interest, the 2D inversion will include measurements that are separated in space as well as time. It will take the satellite around 2 min to traverse the auroral region. To determine to what extent the electrojet is static during this time we refer to the analysis presented by Gjerloev et al. (2011). They found that on the nightside features with scale sizes less than 250 km could on average be considered static over a 2 min period. They had no way of determining if a particular process took place (e.g., north-south streamers, polar boundary intensifications (PBIs) or other meso-scale features) but merely determined this scale size-variability relationship as an average over all conditions. We do, however, know both the scale size at ionospheric altitudes and the lifetime of several meso-scale features: PBIs (∼500 km / 8 min (e.g., Zou et al., 2014)); Streamers (∼350 km / 20 min (e.g., Sergeev et al., 2004)); and, Omega Bands (∼500 km/20 min (e.g., Partamies et al., 2017)). For a LEO satellite (∼8 km/s) it is thus question-
able if PBIs can be considered static while streamers and omega bands may at first glance appear to fall into the static category. The concern, however, may be that for example streamers move which complicates the static assumption. This concern could potentially be checked if auroral imaging with sufficient spatial and temporal resolution as well as sufficiently large field-of-view was available. For a mission like EZIE these concerns imply that the science focus should be on processes with characteristics that suit the observational capabilities. A problem suitable for EZIE could be the structure and evolution of the substorm current wedge which is still being debated despite decades of studies. As a final note it should be mentioned that smaller scale sizes are typically more variable and should not be considered as static over the 2 min traversal time. However, as measurements are made some 30 km separated in altitude from the actual ionospheric currents these smaller scale sizes may not be captured anyway (see Figure 1).

3.3 Effects of volume emissions

The microwave spectrum observed with a Zeeman magnetic field sensor in low Earth orbit would represent a weighted average of the emissions in the sensor’s field of view. Nevertheless, in the inversion presented above, each measurement was assigned the precise position of the center of the field of view. In this section we replace this precise value with a distribution, and investigate the corresponding distribution of solutions. The purpose of this is to give a rough estimate of how the fuzziness in measurement locations maps to a spread in the SECS magnetic field.

To do this, we apply the bootstrap method to the inversion described in Section 2.3: Instead of using the precise measurement locations, we draw random values according to probability distributions that mimic the EZIE field of view, and repeat the inversion many times. The EZIE sensors’ field of view in the mesosphere will be around 40 km, and we define a corresponding probability distribution by sliding a 2D Gaussian with $\sigma = 10$ km along the measurement track by a distance of $\approx 15$ km, corresponding to the 2 second integration time used in this study. To represent the vertical variation of emissions we use a Gaussian with $\sigma = 2$ km. A set of random samples from these 3D distributions give a set of measurement locations, and a corresponding solution vector from equation (6). An example of the magnetic field of such a random dataset is shown in the second column of Figure 6, next to the inversion result based on an exact measurement location in the first column (a copy of the right column from Figure 5). The magnetic field structures are clearly very similar in the two columns.

The third column in Figure 6 shows the mean magnetic field of 30,000 randomly re-sampled datasets, and the fourth (rightmost) column shows the standard deviation. An analysis of the distributions of the model parameters $S_j$ show that they converge after about 7,000 repetitions, which means that the statistics presented in these columns are reliable. We see that the difference between the mean magnetic field and the magnetic field in the left column is very small. The standard deviation is mostly in the order of $\approx 10$ nT, but reaches $\approx 50$ nT near the middle measurement track. This is approximately the magnitude of the error that we can expect in the predicted magnetic field due to uncertainties in measurement locations. The black arrows in Figure 6 represent the SECS equivalent current, and the gray arrows the MHD divergence-free current. The mean equivalent current vectors from the 30,000 re-sampled datasets are visually indistinguishable from the SECS currents based on precise measurement locations.

3.4 Possible improvements of the electrojet estimation

The electrojet estimation technique presented in Section 2.3 involves a novel approach for regularization of the SECS amplitude inversion. Our approach is arguably more flexible than the traditional method (Vanhamaäki & Juusola, 2020), since it allows us to impose constraints other than a uniform spatial smoothing. In this paper we included
Figure 6. Results of a bootstrap experiment to investigate the effect of imprecise measurement locations. Each row corresponds to eastward, northward, and upward components of the magnetic field, respectively, while the columns correspond to (from left to right): 1) The magnetic field according to the method described in Section 2.3. This is identical to the right column of Figure 5, repeated here to help comparison with the other columns. 2) The magnetic field from one of the re-sampled datasets (notice scattered measurement locations). 3) The average magnetic field of all 30,000 re-sampled datasets. 4) The standard deviation of the magnetic field in all the datasets. The black vectors show the corresponding equivalent currents, while the gray vectors show the MHD simulation divergence-free horizontal current.
a penalty for solutions that vary in the magnetic east-west direction, but we expect that there are additional ways in which knowledge about the physics of the electrojet could help inform the inversion.

One strategy could be to use knowledge about the ionospheric conductivity. We know that in the winter, magnetic field perturbations on ground (and hence in the mesosphere) are largely related to currents in the auroral zone (Laundal et al., 2015) where the conductivity is enhanced by ionizing particle precipitation. Knowledge about the location of the auroral oval would enable us to confine SECS amplitudes to this region. Simultaneous observations of the aurora could also be used to make a more precise determination of the preferred direction of variation; instead of penalizing variation in the magnetic east-west direction we could add a penalty for variations along the observed auroral arcs. A similar idea but different application and implementation was used in a recent study by Clayton et al. (2019).

It could also be beneficial to use knowledge about the magnetic field-aligned current (FAC) system in the inversion. As mentioned above, the divergence-free SECS amplitudes are proportional to field-aligned currents under certain conditions. Thus we expect that the SECS amplitudes and FACs are spatially correlated. Global FAC estimates are available from the AMPERE (Waters et al., 2020) project at 2 min cadence, based on 10 min of data from the fleet of Iridium satellites. EZIE will be able to provide spatial resolution far better than AMPERE but they could nevertheless help provide a baseline for the map of SECS amplitudes. It would also be straightforward to include ground magnetic field measurements in the electrojet estimates. This could improve the estimates of large-scale structures and mitigate boundary effects related to uniform electrojets that flow through the analysis area.

4 Conclusions

While the Zeeman magnetic field measurement technique is well established for sensing cosmic magnetic fields, it is new in the context of geospace. The primary benefits of such measurements are the close proximity between the detected magnetic field and the electric current, and the ability to remotely measure the magnetic field at multiple points simultaneously. The EZIE mission concept involves three satellites that scan the electrojet magnetic field as they pass over the auroral zone. In comparison to traditional techniques used for electrojet analyses, the measurement precision is poor, and even the source location is inexact. The electrojet inversion technique presented here uses statistics and knowledge about the nature of the electrojet to overcome these challenges.

Although the inversion scheme in Section 2.3 was developed with the EZIE satellite concept in mind, it would be straightforward to combine with data from ground magnetometers in the vicinity of the satellite. We also believe that the ideas behind the grid and inversion, including possible improvements described in Section 3.4, will be useful in other analyses of regional ionospheric electrodynamics. The technique could be applied with both ground and space magnetometers, or for estimating ionospheric convection using the The Super Dual Auroral Radar Network (e.g., Reistad et al., 2019).

Acknowledgments

K.M. Laundal, J.P. Reistad, and M. Madelaire are supported by the Research Council of Norway/CoE under contracts 223252/F50 and 300844/F50, and by the Trond Mohn Foundation. H. Vanhamäki was supported by Academy of Finland project 314664. V. G. Merkin and K. A. Sorathia acknowledge support from the NASA DRIVE Science Center for Geospace Storms (CGS) under Grant 80NSSC20K0601. The authors gratefully acknowledge the work done by the EZIE mission concept team in support of this paper.
The simulation dataset used in this study is available at https://doi.org/10.18710/8POYRY. Python code for working with spherical elementary current systems and cubed sphere coordinates is available at https://github.com/klaundal/secsy

References


