

RESEARCH ARTICLE

10.1002/2017JA024907

Key Points:

- Polarization ellipse analysis combines two horizontal magnetic field phasors and is used to remote sense the lower ionosphere
- Synchronized MSK demodulation enables polarization ellipse analysis and improves performance in low-SNR conditions
- We observe the day/night terminator, early/fast event, LEP event, and a solar flare, and find variable polarization signatures

Correspondence to:

N. C. Gross,
nicholas.gross@gatech.edu

Citation:

Gross, N. C., Cohen, M. B., Said, R. K., & Golkowski, M. (2018). Polarization of narrowband VLF transmitter signals as an ionospheric diagnostic. *Journal of Geophysical Research: Space Physics*, 123, 901–917. <https://doi.org/10.1002/2017JA024907>

Received 20 OCT 2017

Accepted 4 JAN 2018

Accepted article online 10 JAN 2018

Published online 26 JAN 2018

Polarization of Narrowband VLF Transmitter Signals as an Ionospheric Diagnostic

N. C. Gross¹ , M. B. Cohen¹ , R. K. Said² , and M. Golkowski³ 
¹School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA, ²Vaisala Inc, Louisville, CO, USA, ³Department of Electrical Engineering, University of Colorado Denver, Denver, CO, USA

Abstract Very low frequency (VLF, 3–30 kHz) transmitter remote sensing has long been used as a simple yet useful diagnostic for the *D* region ionosphere (60–90 km). All it requires is a VLF radio receiver that records the amplitude and/or phase of a beacon signal as a function of time. During both ambient and disturbed conditions, the received signal can be compared to predictions from a theoretical model to infer ionospheric waveguide properties like electron density. Amplitude and phase have in most cases been analyzed each as individual data streams, often only the amplitude is used. Scattered field formulation combines amplitude and phase effectively, but does not address how to combine two magnetic field components. We present polarization ellipse analysis of VLF transmitter signals using two horizontal components of the magnetic field. The shape of the polarization ellipse is unchanged as the source phase varies, which circumvents a significant problem where VLF transmitters have an unknown source phase. A synchronized two-channel MSK demodulation algorithm is introduced to mitigate 90° ambiguity in the phase difference between the horizontal magnetic field components. Additionally, the synchronized demodulation improves phase measurements during low-SNR conditions. Using the polarization ellipse formulation, we take a new look at diurnal VLF transmitter variations, ambient conditions, and ionospheric disturbances from solar flares, lightning-ionospheric heating, and lightning-induced electron precipitation, and find differing signatures in the polarization ellipse.

1. Introduction

The *D* region ionosphere (60–90 km) is known to be highly variable. During the daytime, the Sun's extreme ultraviolet and X-ray radiation dominates the formation of electrons and ions in the *D* region. While the ionosphere varies during the daytime due to, for instance, changes in the solar zenith angle, the *D*-region is relatively stable with the exception of solar flares (Thomson & Clilverd, 2001), which heavily change the ionization levels. During the nighttime, however, the ionosphere is both variable and highly unpredictable. The ambient electron densities are set by background cosmic ray radiation, but are also modified by a number of geophysical processes, like lightning-generated heating (Inan et al., 1991), lightning-induced electron precipitation (LEP) from the radiation belts (Voss et al., 1984), atmospheric gravity waves (Hunsucker, 1982), earthquake onset (Nemec et al., 2009), and cosmic gamma ray bursts (Fishman & Inan, 1988).

Electrical properties of the *D* region are generally difficult to directly measure. The altitude is too high for balloons and aircraft but too low for spacecraft. One of the more promising techniques is to use very low frequency (VLF, 3–30 kHz) radio transmitters for ionospheric remote sensing. VLF waves efficiently reflect from both the Earth's surface and the *D* region ionosphere. These two reflecting boundaries form a concentric-sphere waveguide known as the Earth-ionosphere waveguide and allow VLF waves to propagate to global distances. Most high-power VLF radio transmitters are used for submarine communications (Watt, 1967). VLF receivers are also used for terrestrial navigation, such as the Russian *Alpha* VLF system. The U.S. *Omega* VLF system was used for terrestrial navigation (Swanson, 1983), but advances in satellite GPS technology led to the decommissioning of the system, and a few transmitters were repurposed for submarine communications.

When VLF waves propagate in the Earth-ionosphere waveguide, their properties are altered through interactions with the *D* region. Ionospheric remote sensing uses these alterations to infer characteristics about

ionosphere dynamics. This has been done with lightning-generated VLF radio atmospherics (Cummer et al., 1998; Han & Cummer, 2010a, 2010b; Han et al., 2011; Lay & Shao, 2011) but more commonly, with VLF transmitters. For instance, Thomson (1993) and Thomson et al. (2007) measured VLF transmitter signals at a long distance and Thomson et al. (2014) used a mix of short and long distances, and after comparing measurements with a theoretical VLF propagation model, estimated the average ionospheric electron density profile along the path during typical/ambient conditions. Likewise, low frequency (LF, 30–300 kHz) transmitters are used to remotely sense the ionosphere on smaller spatial scales (Higginson-Rollins & Cohen, 2017).

VLF transmitters are also used to characterize *D* region disturbances such as solar flares and solar eclipses. Ambient daytime *D* region ionospheric conditions are maintained primarily by Lyman- α radiation. However, solar flares produce strong X-ray emissions, which are absorbed at *D* region altitudes and greatly enhance ionization of neutral species, including oxygen and nitrogen (Mitra, 1974). Estimating change in the *D* region electron density profile from solar flares has been done by remote sensing the ionosphere with VLF transmitters (Burgess & Jones, 1967; McRae & Thomson, 2004; Thomson & Clilverd, 2001). Conversely, a solar eclipse blocks Lyman- α radiation over a small region of the ionosphere and affects electron recombination and attachment rates. Bracewell (1952) estimated changes in the ionospheric reflection height of a VLF signal during a partial eclipse, and Clilverd et al. (2001) measured multiple VLF transmitters during a total eclipse and estimated changes in the *D* region's electron density profile.

Two types of lightning-induced *D* region disturbances have received particular focus and have been analyzed using VLF remote sensing. The first is Early VLF events, which occur from heating and ionization directly over the region of a lightning stroke. Johnson et al. (1999) used an array of VLF receivers in the midwestern U.S. to characterize the size and location of Early VLF events. These events have been closely associated with sprites and quasi-static field changes with certain lightning strokes (Haldoupis et al., 2004; Moore et al., 2003) but have also been linked to electromagnetic pulses from cloud-to-ground and in-cloud lightning activity (Marshall et al., 2010). The second type of event is known as lightning-induced electron precipitation (LEP). These can be distinguished from Early VLF events because the LEP is often displaced from the lightning location and has a characteristic onset delay and duration (Peter & Inan, 2007). A review of lightning-generated ionospheric disturbances can be found in Inan et al. (2010).

Analysis of *D* region disturbances are greatly complicated by the multimode nature of VLF propagation in the Earth-ionosphere waveguide (NaitAmor et al., 2013). For a given frequency, the multimode signal can be composed of many individual transverse magnetic and transverse electric modes, each with its own amplitude and phase. When conditions in the waveguide change, the phase velocity, group velocity, and attenuation rate, for each mode, are individually affected. The resultant signal, or the sum of all modes, then either increases or decreases in amplitude and advances or regresses in phase, depending on how the perturbed set of modes constructively or destructively interfere with each other. Directly diagnosing this change is therefore extremely difficult, and requires either multiple independent measurements (Bainbridge & Inan, 2003) or a set of assumptions and simplifications that may not be warranted. Measurements of both amplitude and phase of a narrowband signal are required to properly characterize the phasor sum of all modes within the waveguide, but acquisition of phase data is not always possible, often due to receiver limitations or phase instabilities of the transmitter.

One technique that helps isolate the impact of the ionospheric disturbance was introduced by Dowden and Brundell (1996) and Dowden et al. (1997) and is known as scattered field formulation. The concept is to wait for a change in the signal, and then subtract the phasor signal before the disturbance from the postdisturbance phasor. The residual phasor gives a measure of what additional electromagnetic energy appeared as a result of the disturbance, known as the scattered field. In order to apply the scattered field technique, amplitude and phase must be recorded together and combined, whereas most published studies of VLF signals use only amplitude, or perhaps amplitude and phase separately. The scattered field formulation was recently reinvigorated by its use in Golkowski et al. (2014) and Kotovsky and Moore (2015).

We introduce an approach of narrowband signal tracking which uses the complete horizontal polarization of the VLF signal. Scattered field formulation is effective at combining both amplitude and phase from a single antenna, but polarization has the additional advantage of combining multiple colocated antenna measurements and provides the decomposition of the magnetic flux density in both the radial and azimuthal directions. In addition, we show that this approach isolates VLF transmitter phase instabilities and removes phase ambiguities between receiver sites.

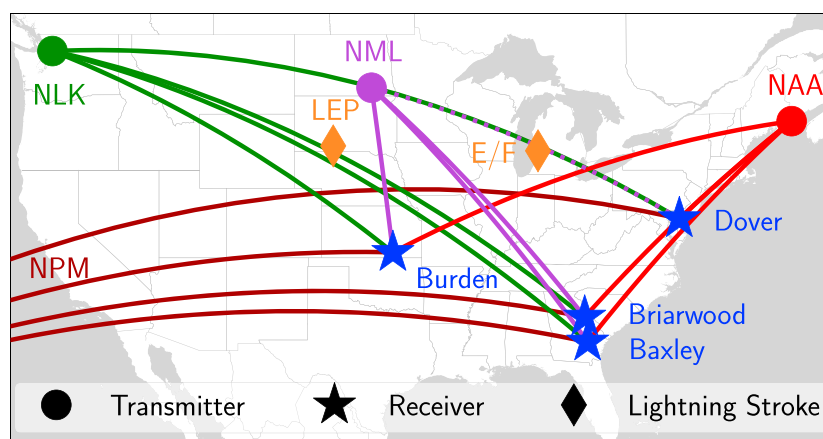


Figure 1. Map of transmitters, receivers, great-circle paths, and lightning strokes used in this paper. The circle markers show the locations of three VLF transmitters. A fourth VLF transmitter located in Hawaii is not shown, but the great-circle paths to the receivers are drawn. The four star markers show the location of the Georgia Tech VLF receivers used in this work. NLDN geolocated two lightning strokes that are associated with an early/fast event and an LEP event, shown as diamond markers.

2. VLF Transmitter Measurements

2.1. Transmitters and Receivers

Measurements are taken in the VLF/LF radio range with an upgraded version of the AWESOME receiver originally described by (Cohen et al., 2010), which uses orthogonal air core loop antennas sampled with 16 bits at 1 MHz. The receiver is sensitive to signals well below 1 fT/rt-Hz over the frequency band 1–470 kHz, depending somewhat on antenna size, and provides GPS-conditioned timing accuracy of 15–20 ns. The receiver consists of two orthogonally oriented antennas, each of which is sensitive to one horizontal component of the magnetic flux density, which we referred to as the north-south and east-west channels. The two channels are sampled synchronously within ~ 2 ns.

Georgia Tech currently operates 10 of these VLF/LF receivers continuously. Figure 1 shows a subset of this network whose data are shown in this paper, along with great-circle paths to VLF transmitters. Three of the transmitters reside in the continental United States: NAA, at 24.0 kHz, is located near Cutler, Main (44.64°N, 67.28°W), NML, at 25.2 kHz, is in LaMoure, North Dakota (46.37°N, 98.34°W), and NLK, at 24.8 kHz, is in Jim Creek, Washington (48.20°N, 121.92°W). The final transmitter, NPM, is located in Lualualei, Hawaii at 21.4 kHz (21.42°N, 158.15°W), and is not shown in the figure, although the great-circle paths from NPM to the receivers are shown. Two of the receivers, Briarwood (33.43°N, 82.58°W) and Baxley (31.88°N, 82.36°W), are located in the southeast United States, while the final two receivers are located in Burden, Kansas (37.32°N, 96.75°W) and Dover, Delaware (39.28°N, 75.58°W). The two diamond-shaped markers represent the location of the causative lightning stroke for an Early event and an LEP event that are discussed in section 4.

2.2. Synchronized MSK Demodulation

All four of the VLF transmitters in Figure 1 use a Minimum-Shift Keying (MSK) modulation scheme at 200 baud. Appendix A gives an overview of the well-established mathematical description for demodulating an MSK signal and estimating the clock phase and carrier phase. Clock phase is proportional to the difference between the phase of the upper and lower tones of the MSK signal. If a single propagating mode dominates, then clock phase is proportional to the group delay at the center frequency of the signal (at 200 baud the group delay proportionality constant is $-0.01/\pi$ s/rad), modulo 360°. Carrier phase is the phase of the center frequency after removal of the MSK modulation, effectively converting the MSK signal into a quasi-CW signal. If a single propagating mode dominates, this parameter is proportional to the phase delay at the center frequency plus transmitter phase changes, modulo 90°. Past VLF remote sensing work has focused heavily on carrier phase, so we use the term *phase* to represent carrier phase unless specifically stated otherwise.

Figure 2 shows an example of the NML VLF transmitter being received at Briarwood. The north/south antenna (N/S channel), which physically lies in the N/S plane, is sensitive to transverse magnetic (TM) waves arriving from those directions. The east/west antenna (E/W channel), which physically lies in the E/W plane,

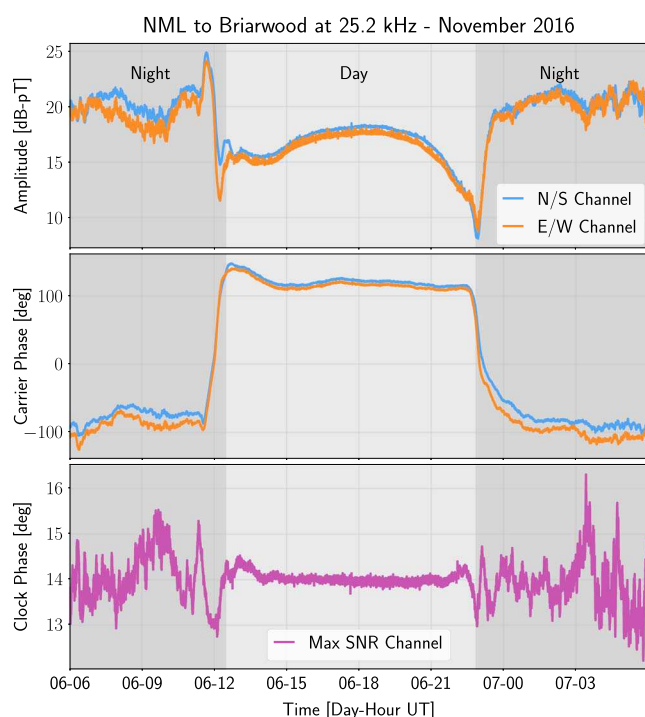


Figure 2. Synchronized MSK demodulation of the broadband data measured at Briarwood, tuned to NML's center frequency of 25.2 kHz. The (top) amplitude, (middle) carrier phase, and (bottom) clock phase are shown over a full day for both the N/S and E/W channels. These diurnal curves are typical during quiet ionospheric conditions.

is sensitive to TM signals from the east or west directions. Figure 2 (top) shows the amplitude data for the N/S and E/W channels. The data shows a classic and well-understood curve due to diurnal changes in the ionosphere. The night-to-day and day-to-night transition periods, known as the day/night terminators, form distinctive nulls and peaks in the diurnal curves and can easily be seen. The day/night terminator lasts about as long as it takes for the sunrise or sunset terminator to traverse the entire transmitter-receiver path. Since the change in longitude along this path is relatively small, the terminator effects are only seen for a short amount of time. Figure 2 (middle) shows the carrier phase, for both channels, advance and regress from the diurnal variations of the D region ionosphere. During the daytime, the carrier phase is smooth and a small offset is maintained between the two channels. The nighttime carrier phase is less smooth and the offset between the two channels varies more, due to the turbulent nature of the nighttime ionosphere. Figure 2 (bottom) shows the clock phase on a single channel where the data have been rotated to maximize the SNR (an in-depth discussion about maximizing the SNR is presented later in this section). Like carrier phase, the clock phase is smooth during the daytime and more chaotic during the nighttime.

An MSK constellation diagram (the in-phase and quadrature mapping of the symbols) contains four symbols, which are evenly spaced around the origin by 90° . These symbol states do not represent the bits that are being transmitted, rather a $\pm 90^\circ$ transition to one of the two adjacent symbols (i.e., clockwise or counterclockwise), relative to the current symbol, determines whether a 1 or 0 is being transmitted. Often, the time series transitions from one symbol state to the next are mapped onto a phase trellis.

To obtain the carrier phase, the estimated phase trellis is subtracted from the baseband MSK signal. However, it is possible, especially in low-SNR environments, to improperly estimate the phase trellis (meaning that one or many bit estimation errors occurred). When this happens, a phase jump of $\pm 180^\circ$ occurs. Furthermore, at the start of the MSK demodulation, an initial phase must be assumed, so a value of 0° on the phase trellis is chosen. This assumption inherently implies a 90° ambiguity when comparing phase measurements between receivers, channels, or time series data sets that are not synchronized during demodulation. In principle, this ambiguity is removed if the initial trellis position when the transmitter last began broadcasting is known, as well as the entire sequence of bits transmitted since then, but in practice this is not often known.

An MSK signal is composed of two tones, an upper tone ϕ_+ and a lower tone ϕ_- (at 200 baud these tones are 50 Hz above and below the center frequency), see (A8) for more details. The initial estimate of the carrier phase $\hat{\phi}_0$ is made using these tones, $\hat{\phi}_0 = (\phi_+ + \phi_-)/4$. The phase of these two tones can have a 2π wrapping associated with them, so they must first be precisely unwrapped before the clock phase and carrier phase are calculated. If the phase of the tones are not correctly unwrapped, then an error that is a multiple of 90° is induced in the carrier phase. Correct unwrapping is particularly difficult during periods of low SNR or when the transmitter is toggled between off and on. The example shown in Figure 2 is a case where the transmitter was on and a relatively high SNR was recorded for the entire 24 h period, but this is often not the case.

For the purposes of calculating the polarization ellipse, 90° phase jumps and 180° phase jumps are troublesome if they occur on one antenna channel, but not the other. If one channel has a 90° jump, but the other does not, then a linearly polarized signal will appear to be a circularly polarized signal (or vice versa). If one channel has a 180° jump, but the other does not, then the apparent arrival direction (or the major axis direction) will be flipped about either the x axis or y axis.

These 90° and 180° phase jumps between channels result from the MSK demodulation being applied independently on each antenna channel. Our proposed solution is therefore to synchronize the bit and clock phase estimation between the two channels and ensure that the carrier phase between the two channels is devoid of the 90° ambiguity. To achieve this, we digitally rotate the data to a direction that maximizes the SNR of the MSK signal, estimate the phase trellis on the high-SNR channel, and then apply the high-SNR phase trellis

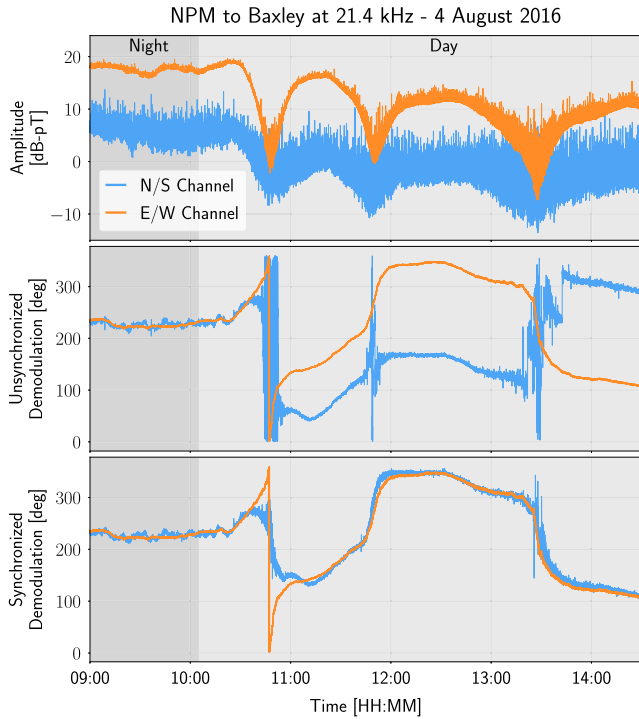


Figure 3. Comparison of the unsynchronized and synchronized MSK demodulation algorithm over a long west-to-east path (Hawaii to Georgia). (top) Amplitude is shown and has multiple peaks and nulls due to the night-to-day terminator. Also shown are the (middle) unsynchronized carrier phase and (bottom) synchronized carrier phase. The nighttime to daytime transition is defined as the point in time where the terminator is directly above the receiver.

to two orthogonal channels that are $\pm 45^\circ$ off of the maximum SNR direction. The 90° and 180° ambiguities still exist but affect both magnetic field components synchronously. An added benefit to the synchronized approach is that the initial bit and clock phase estimation is done at the highest possible SNR, thanks to the digital rotation of the MSK signal. A high SNR reduces the chance of estimating an incorrect bit and gives a more precise unwrapping for the phase of the upper and lower MSK tones.

The carrier phase ϕ_0 of an MSK signal is estimated by subtracting the phase trellis $\hat{\phi}_{\text{MSK}}(t - \hat{\tau})$ from the phase of the baseband MSK signal $\angle z(t)$, shown in (A13). We extend the single-channel mathematics from Appendix A to a two-channel synchronized demodulation method by estimating the bit sequence, $\hat{b}_M(t)$ and $\hat{\theta}_M(t)$, and clock phase $\hat{\phi}_{1,M}$ on a single channel with maximum SNR, and use these estimates to demodulate the MSK signal on two orthogonal channels. Subscript M denotes the digitally rotated channel that maximizes the SNR of the MSK signal.

The phase trellis on the maximum SNR channel

$$\hat{\phi}_{\text{MSK},M}(t - \hat{\tau}) = \hat{b}_M(t - \hat{\tau}) \left(\frac{\pi t}{2T} + \hat{\phi}_{1,M} \right) + \hat{\theta}_M(t - \hat{\tau}) \quad (1)$$

is calculated by estimating the bit sequence and clock phase (see (A3) for details) on the maximum SNR channel. The carrier phase is found by subtracting the maximum SNR phase trellis from the baseband MSK signal for an arbitrary channel $z(t)$,

$$\phi_0(t) \approx \angle[z(t)] - \hat{\phi}_{\text{MSK},M}(t - \hat{\tau}). \quad (2)$$

Typically, the carrier phase is averaged over a period a ,

$$\phi_{0,\text{ave}} \approx \angle \left\{ \frac{1}{a} \int_t^{t+a} \frac{z(\xi)}{|z(\xi)|} e^{-j\hat{\phi}_{\text{MSK},M}(\xi - \hat{\tau})} d\xi \right\}, \quad (3)$$

or resampled to a larger sampling period. The baseband MSK signal $z(t)$ can, in general, be chosen from any direction. However, to equally distribute the noise across both orthogonal channels, $z(t)$ should be $\pm 45^\circ$ away from the maximum SNR direction, $z_{M+45^\circ}(t)$ and $z_{M-45^\circ}(t)$. Applying $z_{M+45^\circ}(t)$ and $z_{M-45^\circ}(t)$ to (2) results in the synchronization of the two orthogonal channels. For this work, we then rotate $z_{M+45^\circ}(t)$ and $z_{M-45^\circ}(t)$ to the N/S and E/W directions, but this rotation is not generally necessary. With the N/S and E/W channels synchronized, the 90° and the 180° phase jumps between the two channels no longer exist and a more confident estimate of the carrier phase is found, even in low-SNR conditions.

We demonstrate the modified algorithm's utility in Figure 3, which shows the NPM VLF transmitter, in Hawaii, recorded at Baxley. The amplitude on both channels is displayed in Figure 3 (top). The path from NPM to Baxley has a large longitudinal span, meaning the night-to-day terminator takes several hours to traverse westward along the transmitter-to-receiver path, and this transition is apparent in the received signal between $\sim 10:30$ and $13:45$ UT. Three nulls are seen in the amplitude data on both channels and are the result of modal interference that is commonly observed in long east-to-west or west-to-east paths (Pappert & Morfitt, 1975). During these nulls, the SNR of the VLF signal drops which makes carrier phase estimation more difficult. In Figure 3 (middle), the carrier phase has been extracted from each channel individually, using the original single-channel algorithm. Before the first null occurs, the phase difference between channels is typically within $\sim 5^\circ$ of each other. During the first null, the N/S channel carrier phase begins to sporadically move, showing that the carrier phase is being improperly estimated due to the low-SNR condition. Once the first null has passed, a 90° offset between the two channels appears. This offset is the result of the 90° phase ambiguity from the two channels not being synchronized. During the second and third nulls, the carrier phase is again uncertain and more 90° phase shifts occur, compounding the 90° phase ambiguity problem. Carrier phase in Figure 3 (bottom) is calculated using the modified algorithm with the bits and clock phase synchronized. During the first two nulls, the carrier phase is estimated well, and during the third null, nearly all points in time

northward direction ($\hat{y}$) to the direction of the transmitter ($-\hat{r}$). A clockwise rotation matrix is used to rotate the magnetic flux density from the $\hat{x}$ and $\hat{y}$ coordinate system to the $\hat{r}$ and $\hat{\phi}$ coordinate system,

$$\begin{bmatrix} B_r \\ B_\phi \end{bmatrix} = \begin{bmatrix} \cos(\theta_{az} + \frac{\pi}{2}) & \sin(\theta_{az} + \frac{\pi}{2}) \\ -\sin(\theta_{az} + \frac{\pi}{2}) & \cos(\theta_{az} + \frac{\pi}{2}) \end{bmatrix} \begin{bmatrix} A_{NS} e^{j\psi_{NS}} \\ -A_{EW} e^{j\psi_{EW}} \end{bmatrix}, \quad (5)$$

where B_r and B_ϕ are the phasor form of the magnetic flux density.

The blue trace in Figure 4 shows an example polarization ellipse, it is equivalent to the time domain magnetic flux density, equation (4), being drawn on the xy -plane over a full period. There are four properties that together specify the ellipse: the major axis (which is double the semimajor), the tilt angle (τ), the ellipticity (χ), and the start phase. The polarization ellipse, in general, is not aligned perfectly with the azimuthal direction, but for TM-dominated signals it is often reasonably close. Tilt angle

$$\tau = \frac{1}{2} \tan^{-1} \left[\frac{2\gamma}{1 - \gamma^2} \cos(\psi_0) \right], \quad (6)$$

where

$$\gamma = \frac{|-B_r|}{|B_\phi|} \quad (7)$$

and

$$\psi_0 = \angle \left[\frac{-B_r}{B_\phi} \right], \quad (8)$$

is defined as the counterclockwise angle from $\hat{\phi}$ to the major axis direction, with a range of $-90^\circ \leq \tau \leq 90^\circ$. Tilt angle can also be thought of as the error in angle of arrival when using classic magnetic direction finding.

The ellipticity angle

$$\chi = \frac{1}{2} \sin^{-1} \left[\frac{2\gamma}{1 + \gamma^2} \sin(\psi_0) \right], \quad (9)$$

of an ellipse encapsulates two parameters: how linear (or circular) the ellipse is and the rotational sense of the ellipse. Since the minor axis is by definition smaller than the major axis, the ellipticity angle exists over the interval $-45^\circ \leq \chi \leq 45^\circ$. The rotational sense of the ellipse is defined by the sign of χ , a positive value ($\chi > 0$) indicates that the ellipse is right-hand polarized (or rotating counterclockwise) and a negative value ($\chi < 0$) means that the ellipse is left-hand polarized (or rotating clockwise). For the example in Figure 4 the *Rotation Direction* vector is showing a right-hand polarization, so χ is positive. The ellipticity angle also describes how linear the ellipse is. When χ is close to zero, the ellipse is highly linear. Conversely, the ellipse is highly circular when the absolute value of the ellipticity ($|\chi|$) is near 45° .

Start phase, ρ , is the phase difference between the vector parallel to the semimajor axis closest to $\hat{\phi}$, written as $\underline{B}(t=t_{\text{Maj}})$, and the initial point of the magnetic flux density vector $\underline{B}(t=0)$, which gives

$$\rho = \omega t_{\text{Maj}}. \quad (10)$$

Using a counterclockwise rotation matrix, the magnetic flux density phasors along the semiminor axis and semimajor axis, B_{Min} and B_{Maj} , respectively, can be solved for

$$\begin{bmatrix} B_{\text{Min}} \\ B_{\text{Maj}} \end{bmatrix} = \begin{bmatrix} \cos(\tau) & -\sin(\tau) \\ \sin(\tau) & \cos(\tau) \end{bmatrix} \begin{bmatrix} B_r \\ B_\phi \end{bmatrix}. \quad (11)$$

By definition, $B_{\text{Maj}} = |B_{\text{Maj}}| e^{-j\omega t_{\text{Maj}}}$, which yields $\omega t_{\text{Maj}} = -\angle B_{\text{Maj}}$, and allows the start angle to be written as

$$\rho = -\angle B_{\text{Maj}}. \quad (12)$$

Start phase is an important metric, because as we will show, it captures transmitter phase changes but is independent of the geometric shape of the ellipse. If the carrier phase changes equally on both the azimuthal and radial channels, then B_r and B_ϕ experience equal phase changes. On the other hand, the shape of the ellipse, whether it is highly linear or circular, is closely connected to the phase difference between the two components. This is useful when dealing with transmitters that have drifting or unstable phase at the source. The phase instabilities are captured in the start phase, but they do not contaminate the geometric properties of the ellipse. Note that since $\underline{B}(t)$ forms an ellipse, the start phase is not the same as the geometric angle between $\underline{B}(t=t_{\text{Maj}})$ and $\underline{B}(t=0)$, unless $\underline{B}(t)$ is circularly polarized.

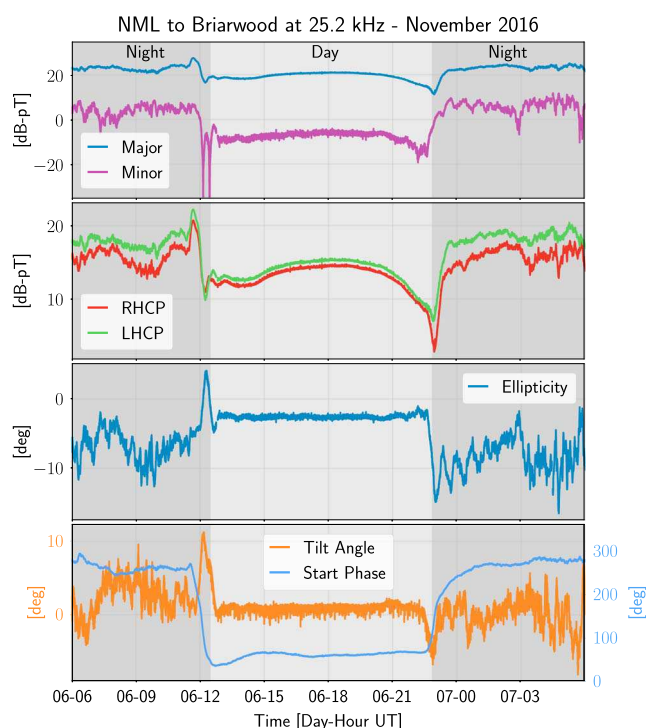


Figure 5. Data from Figure 2 transformed into ellipse parameters. From top to bottom, (first panel) the semimajor axis and semiminor axis lengths of the time varying ellipse; (second panel) the decomposed LHCP and RHCP ellipse components. (third panel) Ellipticity is the rotation sense and how linear (or circular) the ellipse is. (fourth panel) The ellipse tilt angle and start phase.

3.2. Example Data

Figure 5 shows the same 24 h VLF transmitter data set from Figure 2 reformulated into the polarization ellipse parameters. Figure 5 (first panel) shows the semimajor axis and semiminor axis lengths. The major axis is much stronger than the minor axis because the TM mode of VLF wave propagation dominates at longer distances, so the ellipse is close to linearly polarized. Also, the major axis is relatively steady with time, but the minor axis fluctuates greatly, and at times disappears almost entirely as the signal becomes linearly polarized.

An ellipse can also be decomposed into two counter-rotating circles, commonly referred to as right-hand circular polarization (RHCP) and left circular polarization (LHCP), and these are shown in Figure 5 (second panel). The LHCP signal is slightly stronger for the entire day, except for a short amount of time during the night-to-day terminator. The change in rotation direction can also be seen in the ellipticity data. The ellipticity angle is greater than zero for a short time during the night-to-day terminator, indicating that the rotation direction briefly changed to a counterclockwise rotation. Also, the ellipticity angle shows that the ellipse is quite linear during the day but slightly more circular during the night. This result is not surprising since nighttime conditions likely support a higher number of propagating modes.

Figure 5 (fourth panel) shows the last two parameters, the tilt angle and start phase. The tilt angle is roughly $+1.5^\circ$ during the daytime, which is reasonable for angle of arrival measurements. It should be noted that we use the sferic calibration method in Wood and Inan (2004) to remove the effects of terrain and antenna misalignment, which has also been done in Said (2009) and Zoghzoghy (2015). Interestingly, the tilt angle exhibits a brief change at each of the day-night transitions, but in opposite direc-

tions. It is possible that these quick changes in tilt angle are caused by scattering from the oblique angle between the terminator and the great-circle path. The final parameter, start phase, is related to the phase velocity of the signal (along with the transmitter source phase variations), and therefore provides some information on both. If we assume the magnetic flux density is dominated by a single mode and the phase of the transmitter is stable, then the start phase in Figure 5 implies that the phase velocity is $\sim 0.003c$ m/s faster during the day. However, these assumptions are not generally true. The gradual change in tilt angle and start phase may also be the result of mode conversion during the day/night terminators (Kaiser, 1968).

Every parameter in Figure 5 shows steady and smooth trends during the daytime. Conversely, the nighttime ellipse parameters are variable and reflect the more chaotic nature of the nighttime *D* region ionosphere. As expected, the ellipse parameters are highly variable during the day/night terminator, showing that the sharp transition region between the daytime ionosphere and nighttime ionosphere greatly affects the propagation of VLF waves traversing the terminator and can significantly increase or decrease the power received from a transmitter.

3.3. Phase Stability of VLF Transmitters

Even though VLF transmitters have been used in decades of remote sensing studies, Figure 6 shows clear evidence that the phase of at least some VLF transmitters is not always stable. The figure shows the carrier phase and the polarization parameters for the transmitter NAA recorded at Baxley, Georgia. Note that the time period is during the day and no significant solar activity is present. The first column presents the carrier phase measured on the N/S (top) and E/W (bottom) channels, both of which exhibit sharp phase transitions. Though not shown, these phase changes are also seen at other receivers at the exact same time, but only on the NAA transmitter. These sharp phase transitions are certainly a result of an unstable carrier phase at the transmitter and not from transient ionospheric effects. Although these phase transitions are sharp and plainly obvious when looking at phase measurements from multiple sites, it may be easy to mistake one of these transitions as an actual ionospheric disturbance, particularly if only one such sharp transition occurs and if data from only one receiver is considered. It is worth noting that at least some ionospheric disturbances

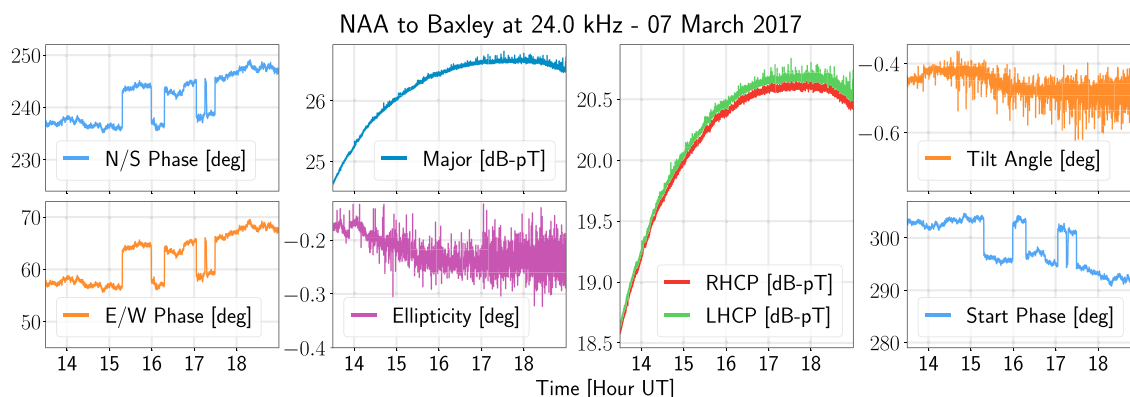


Figure 6. Evidence of unstable carrier phase in a VLF transmitter. (first column) The carrier phase on both the N/S and E/W channels has strong step-like changes that are caused by phase instabilities at the transmitter. (second to fourth columns) The ellipse analysis isolates the transmitter phase instability to the start phase, and the other ellipse parameters are uncontaminated by the instability.

have been known to affect only the phase, and not the amplitude (Cotts & Inan, 2007). We are not suggesting that these earlier observations may be transmitter artifacts, but rather we are underscoring the importance of considering transmitter source phase transitions in identifying ionospheric events. More importantly, the sharp transitions are likely accompanied by much more gradual drifts in the transmitter source phase that cannot be inferred at the receiver.

As a consequence, carrier phase measurements are not invariant to transmitter source phase wanderings, whereas amplitude measurements can be more carefully calibrated (Cohen et al., 2010). Instead, carrier phase measurements are useful in a relative sense and on short time scales. That is, if we assume the transmitter source phase variations are relatively slow compared to the time scale of ionospheric changes under investigation, then *changes* in carrier phase are a reliable measurement to use. However, if we study hour-long trends or day-to-day variations in ionospheric conditions, measured phase changes may be unreliable. Thomson et al. (2014) use a direct approach to mitigate the source phase problem by placing a receiver near the transmitter and measuring the phase; they then use the phase from the short path to calibrate the received transmitter phase on the longer paths. Unfortunately, this requires significant effort to maintain sites next to every VLF transmitter. Instead, the polarization ellipse technique can be applied to the narrowband data and the ellipses from each site can be directly compared, even though each site still has an ambiguous phase offset.

Figures 6 (second column)–6 fourth column) contain the polarization ellipse parameters. As seen, only the start phase measurement contains sharp phase transitions that are correlated with the carrier phase measurements on the N/S and E/W channels. This shows that the polarization ellipse technique is a useful way to characterize magnetic flux density measurements on two channels even when the source phase is unstable.

4. Ionospheric Remote Sensing

We now apply the polarization ellipse technique to narrowband data associated with *D* region ionospheric disturbances in order to demonstrate features of the received VLF signal that may not be apparent in the single-channel amplitude and phase data. The first example is modal changes of VLF propagation due to the day/night terminator, a somewhat predictable and repeatable phenomenon. The other three observation examples show different types of ionospheric perturbations with different ionospheric modification mechanisms: Early/fast, LEP, and solar flare. Each type of perturbation produces a distinctly different set of polarization parameters. Note that these results should not be generalized since we are presenting only an example case study of each type of disturbance. Since these phenomena are all dependent on, for example, path geometry of the transmitter-disturbance-receiver, an in-depth statistical study is required to understand how the changes in polarization ellipses link to causative physical mechanisms. At the same time, the markedly different responses of the polarization ellipses for these examples are representative of the value of utilizing the polarization ellipse method.

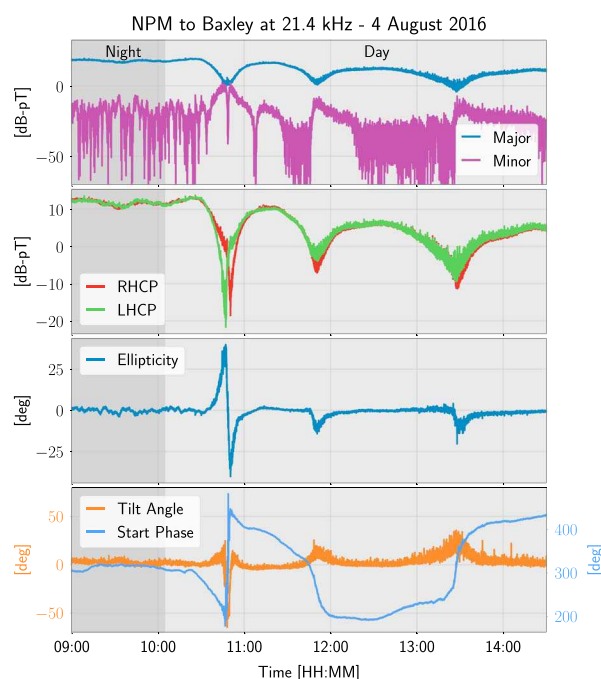


Figure 7. Data from Figure 3 transformed into ellipse parameters. From top to bottom, (first panel) the phase-synchronized E/W channel in Figure 3 shows an interesting artifact during the first null, which is also seen in the semimajor curve. (second panel) Using ellipse analysis, the artifact is determined to be caused by a separation between the LHCP and RHCP nulls (about 2 min before and 2 min after the first null, respectively). A strong change in (third panel) ellipticity and (fourth panel) tilt angle is also seen during the first null, but the severity in change is not seen during the other two semimajor nulls. Changes in start phase (Figure 7, fourth panel) are strongest around each semimajor null.

4.1. Modal Changes Due To Terminator Effects

Understanding the day/night terminator and its effect on VLF wave propagation is, in general, a difficult problem. The strong inhomogeneity of the *D* region along with the sharp change in electron density has made modeling propagation across the terminator challenging. Figure 3, which is the signal from the NPM transmitter in Hawaii recorded at Baxley, Georgia, is an example of the complex changes in VLF wave propagation as the terminator traverses the ionosphere. In particular, the first null in Figure 3, which occurs at about 10:48 UT and is referred to as *the first null* for the rest of this section, shows a peculiar deviation between the synchronized N/S and E/W phases. The first null occurs about 43 min after the terminator passes over the receiver and is well over the great-circle path. It is difficult to tell what this deviation means when analyzing the amplitude and phase. However, when looking at the first null in terms of ellipse parameters, shown in Figure 7, it is obvious that the first null is undergoing a change in elliptical polarization.

The LHCP and RHCP components, in Figure 7 (second panel), have nulls at different times. The LHCP component of the ellipse reaches its minimum about 2 min before the major axis null, but the RHCP portion reaches a null about 2 min after the major axis null. Hence, the RHCP component dominates over the LHCP component for a brief period just before the major axis null, whereas the LHCP component dominates for a brief period after the major axis null. The signal then returns to being almost linearly polarized. This transition can also be seen in the ellipticity curve. Before the first null, the ellipticity is much greater than zero, meaning that it is highly circular and right-hand polarized. Then, as the first null passes, the ellipticity quickly becomes much less than zero, which indicates the ellipse is again highly circular, but this time left-hand polarized. Note that the quick change, during the first null, in ellipticity angle and start phase is not instantaneous and is not caused by phase wrapping; instead, the change from one extreme to the other is quick, smooth, and continuous.

The cause for the VLF wave to transition from LHCP to RHCP during the first null is not fully understood. Mode conversion from the night-to-day transition does affect the RHCP and LHCP components differently, due to ionospheric anisotropy from Earth's background magnetic field, which may also explain why the LHCP component is significantly stronger in the second and third nulls. Also, given that the terminator forms a sharp gradient in the *D* region electron density profile, the receiver may be measuring signals that are obliquely forward scattered from the terminator (i.e., reflecting off the terminator at a region which is not on the great-circle path, causing multipath propagation). It is worth noting that this flip in the rotational sense only occurs for the first null, when the day/night terminator has drifted a few hundred kilometers past the receiver toward the transmitter. We will show in the next subsection an Early/fast event, which is generally caused by ionospheric scattering, that also produces an increase in circular polarization and a strong change in tilt angle. The changes for the case study Early/fast event are similar, but smaller in scale to what is seen here during the first null.

4.2. Early/Fast Event

Early events are VLF perturbations caused by powerful or intense lightning near the transmitter-receiver path. Early events result from the coupling of a lightning stroke to the bottom *D* region ionosphere, which in turn creates localized heating and ionization of the lower ionosphere (Inan et al., 1996). Proposed physical mechanisms include scattering from a sprite halo (Moore et al., 2003), scattering off a sprite body (Dowden et al., 1997), ionization from the electromagnetic pulse or "elve" (Haldoupis et al., 2013; Mika et al., 2006), quasi-electrostatic quiescent heating (Pasko et al., 1996; Kabirzadeh et al., 2017), and heating from intracloud pulses (Marshall et al., 2008). Recovery for Early events is typically 10–100 s but sometimes much longer (Cotts & Inan, 2007), and may occur as far away as 300 km from the path (Salut et al., 2013). Early events are seen in narrowband data, because the perturbed region of the ionosphere scatters a signal that is not on the great-circle path toward the receiver (Dowden & Brundell, 1996). The receiver measures the superposition of

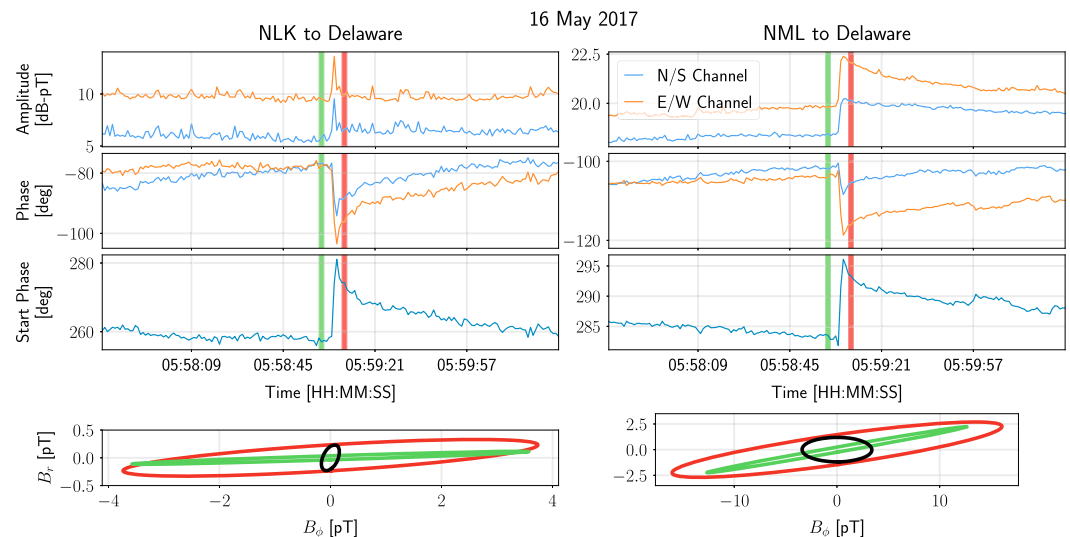


Figure 8. Polarization ellipse technique used on an early/fast event, recorded at a single receiver from two different transmitters. Shown are (first row) amplitude, (second row) carrier phase, and (third row) start phase. Figure 8 (fourth row) shows the ellipses before (green) and after (red) the event occurred, along with the scatter field ellipses (black). Vertical green and red bars in the top three rows show the time periods that the preonset ellipses and postonset ellipses are windowed over.

the ambient signal that propagates directly along the great-circle path and a scattered signal from the perturbed region. The recorded amplitude and phase either increases or decreases depending on how the two signals interfere with each other (Moore et al., 2003).

Figure 8 shows a typical Early/fast event, with a rapid (<20 ms) onset duration. The location of the causative lightning stroke is shown in Figure 1 and labeled as *E/F*. The National Lightning Detection Network (NLDN) (Cummins & Murphy, 2009) measured the peak current of the stroke to be +156 kA and located the stroke at 43.05°N and 86.03°W. Each column shows the measured narrowband data at Delaware from the NLK and NML transmitters, respectively. This particular event occurred along two transmitter-to-receiver paths which almost exactly overlap.

Figure 8 (first row) and 8 (second row) show the amplitude and phase on each channel, respectively. The data are recorded at both 1 s and 20 ms time resolutions, but our plots here show only the low resolution since they are less noisy. We observe the typical response for an Early/fast event: a quick (<20 ms) onset duration followed by a slower recovery. Figure 8 (third row) shows the start phase at both receivers. The curve of the start phase closely follows the phase on the N/S and E/W channels, but with opposite polarity. This strong change in start phase is the result of carrier phase changes with respect to the sum of all modes from the direct path and the scattered path.

Narrowband data are inherently susceptible to impulsive noise from, in particular, lightning-generated sferics. We reduced this impulsive noise by applying a three-point median filter to the narrowband data before using the ellipse analysis. The polarization ellipse results are shown in Figure 8 (fourth row). The green ellipse represents the narrowband signal before the causative stroke, and the red ellipse is after the onset of the Early/fast event. The panels in the top three rows show red and green vertical bars to indicate the time period over which the ellipse parameters were determined. By subtracting the ambient signal (green) from the disturbed signal (red), we can obtain a scattered field polarization ellipse, which is shown in black on the bottom two panels. The scattered field ellipse captures the change in major axis, minor axis, and tilt angle, as a result of the ionospheric disturbance.

For the NLK to Delaware path, the Early/fast event caused a rotation of the ellipse and a significant increase of the semimajor axis. Likewise, the NML to Delaware path also has a change in tilt angle but opposite in direction, and an increase on both the semimajor and semiminor components. These changes can also be seen in the scattered ellipses. The increase in the semimajor component means that an increase of VLF power is received at the antenna close to the semimajor direction. This power increase may be due to the scattering

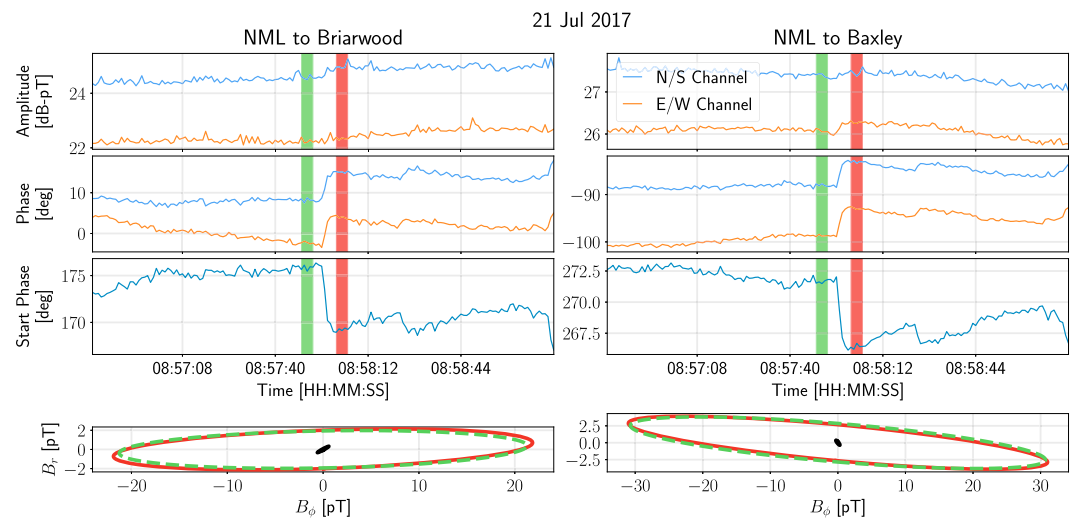


Figure 9. Polarization ellipse technique used on an LEP event, measured at two different receivers from a single transmitter. See the caption in Figure 8 for details on this figure layout.

effect (i.e., multipath) that Early/fast perturbations produce, since this particular event was relatively close (25.7 km) to the great-circle path.

The observed rotation and ellipticity change of the polarization ellipse is worth highlighting. This effect implies that the Early/fast event has different characteristics if observed independently on each of the two channels, both in terms of the amplitude and phase changes, as well as the recovery event. Past works have observed, for instance, that some Early/fast events are only observed on one channel or that the recovery time differs on the two antennas. Polarization observations indicate at least observationally why this occurs, and observation of a rotating ellipse may be useful in identifying properties of the scattering region, as well as new scattered modes that are created at the disturbed ionosphere.

4.3. LEP Event

Lightning-induced electron precipitation (LEP) events are also driven by lightning strokes, but the ionospheric disturbance is triggered in a different way: lightning produces VLF energy that leaks into the magnetosphere and interacts with geomagnetically trapped particles, causing some electrons to precipitate into the lower ionosphere, ultimately increasing the electron density of the *D* region (Voss et al., 1984). Compared to Early events, LEP events typically impact much larger regions of the ionosphere (sometimes >1,000 km (Peter & Inan, 2004)). Since the increased ionization occurs over a larger region, two types of VLF perturbations, or a combination of both, is seen at the receiver (Cotts, 2011). First, the disturbed ionospheric region may scatter some VLF energy toward the receiver, similar to Early event scattering. Second, if the ionosphere is significantly disturbed over a large (much greater than a wavelength) portion of the great-circle path, then the modal properties of the VLF wave change.

Figure 9 shows an LEP event measured by two separate receivers, Briarwood and Baxley, from a single transmitter, NLK. The causative stroke, labeled as *LEP* in Figure 1, had a peak current of +204 kA and was located at 43.32°N and 101.17°W. The onset delay (the time difference from when the stroke occurred to when the perturbation is seen in the receiver data) is ~0.7 s, and the onset duration (period from when the perturbation occurs until the recovery begins) persists for over 1 s. Noise from other lightning strokes was particularly high during this event, so we applied a nine-point median filter to the low-resolution (1 Hz) data for visual clarity. The before and after ellipses from the NLK to Burden path show almost no change, and yet a strong jump in the start phase occurred. This produced a very small scatter field ellipse shown at the origin as a small dot. For the other path, NLK to Baxley, a slight change in tilt angle and modification along the semiminor axis occurred. Compared to the ellipse changes in the Early/fast event, the LEP event produced almost no change to the shape of the polarization ellipse, just a change in start phase. The difference between this observations and the Early/fast event shown earlier may be due to the fact that Early/fast events result from scattering, whereas the LEP was caused by modal changes. In particular, the fact that the start phase changed whereas the shape did not may indicate that the phase velocity of the dominant propagating modes were

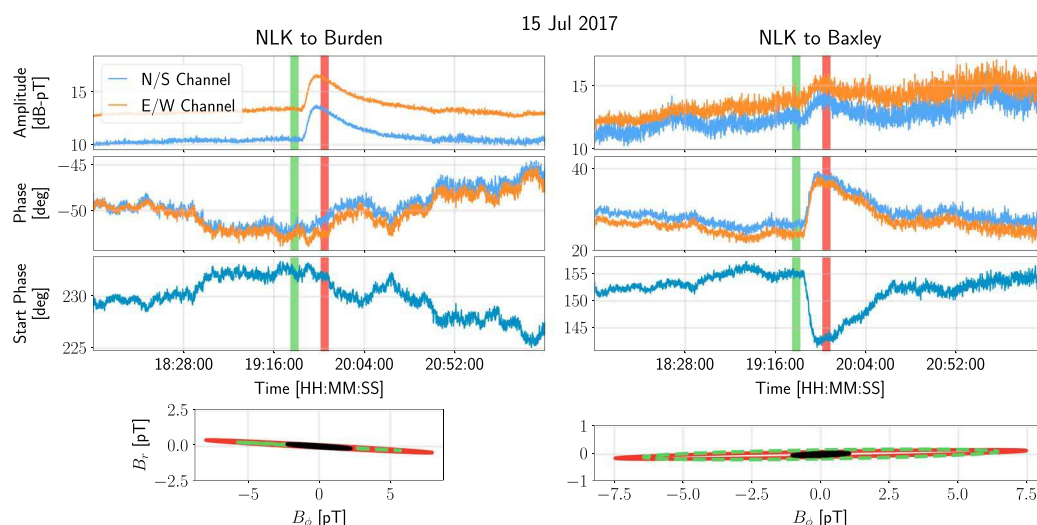


Figure 10. Polarization ellipse technique used on an ionospheric disturbance caused by a C-class solar flare, measured at two different receivers from a single transmitter. See the caption in Figure 8 for details on this figure layout.

impacted by the LEP event in a roughly equal way but that the attenuation rate was not much affected. However, the effect of geometry of the transmitter, disturbance, and receiver, is also important and should not be discounted (NaitAmor et al., 2010, 2013).

4.4. Solar Flare Event

The Sun is the chief driving force for the diurnal variations of the ionosphere. During the daytime, the Sun increases the electron density of the *D* region by orders of magnitude and creates a relatively stable and predictable ionosphere as the Sun traverses across the sky. LEP and Early/fast events are rarely, if ever, observed during the daytime. However, abnormally strong X-ray emissions from the Sun, known as solar flares, can perturb the daytime ionosphere an appreciable amount (Thomson & Clilverd, 2001). Solar flares affect at least the entire daytime side of the ionosphere and enhance the electron density through the entire *D* region.

Figure 10 shows VLF perturbations from a C-class solar flare event over the continental United States, which is measured along the paths NLK to Burden and NLK to Baxley. The ellipses are median filtered over 4 min windows. This example is interesting because the NLK to Burden narrowband clearly shows the solar flare in the amplitude data, but no change is visible in the phase data. In contrast, the NLK to Baxley path shows a disturbance plainly obvious in the phase, but barely detectable in the amplitude. This fact reinforces the idea that analyzing amplitude and narrowband data individually may not give a sufficiently comprehensive picture of the underlying phenomena and could either miss or mischaracterize ionospheric disturbances. Techniques such as scattered field method and polarization ellipse method may provide crucial information in the analysis of these types of events.

The polarization ellipse shape before and after the solar flare ellipses are quite similar for both paths. Both experience minimal minor-axis change and no tilt angle change, but both present strong increases in the major-axis component. Likewise, the scattered field ellipses are almost linear and are close to parallel to the major axis of the before and after ellipses. The start phase on the NLK to Burden path shows almost no change from the solar flare, but the NLK to Baxley start phase has a strong change that is similar to the phase data but with opposite polarity. Since this event is during the daytime and over a long path, the transverse magnetic field dominates over the strongly attenuated radial magnetic field (Bainbridge & Inan, 2003), which is why the ellipses are highly linear and almost no change occurs on the semiminor component.

5. Discussion and Conclusion

We have shown that inferring ionospheric conditions using VLF narrowband amplitude and phase data independently can lead to ambiguous results or even missed events. When using VLF for remote sensing, the amplitude and phase data are best considered concurrently, such as in the scattered field analysis. However,

scattered field analysis is best used with a single isotropic antenna. If a pair of colocated orthogonal directional antennas are being used, such as two-loop antennas, then the polarization ellipse technique is well suited for the data analysis.

We also showed evidence of unstable carrier phase from VLF transmitters and discussed that care must be taken to avoid having the transmitter phase instabilities contaminate the remote sensing results. Furthermore, 90° and 180° phase jumps are present in some narrowband data due to the inherent nature of MSK modulation. To circumvent the 90° ambiguity, we developed a technique for synchronizing the MSK demodulation on a multichannel receiver, which removes the 90° and 180° jumps between channels and can significantly improve the phase estimation on channels with low SNR. With the channels of the colocated antennas synchronized, the polarization ellipse technique can be applied and any VLF phase instabilities are shifted into the start phase parameter, leaving the other ellipse parameters uncontaminated.

We demonstrated the properties of the polarization ellipse method for one particular day/night transition, and then observed three separate examples of ionospheric disturbances: Early/fast, LEP, and solar flare. These examples showed that polarization ellipses can be directly compared between different transmitter and receiver combinations, along with measurements from different times without phase instabilities corrupting the ellipses or concerns about a 90° ambiguity. Each event showed unique characteristics when comparing the polarization ellipse before and after the initial onset of the event. We emphasize that these case studies are not necessarily representative of all ionospheric disturbances. Future work should consider a more in-depth statistical study to see if specific ellipse characteristics are unique to certain ionospheric disturbances, and a comparison between measured polarization ellipses and current VLF waveguide propagation models should be performed.

Appendix A: Single-Channel MSK Demodulation

The mathematics for MSK demodulation is well established, but is stated here for reference. The MSK signal can be written in the form

$$s(t) = \cos \left[2\pi f_c t + b(t) \left(\frac{\pi t}{2T} \right) + \theta(t) \right] \quad (\text{A1})$$

where f_c is the carrier frequency, T is the bit period, and

$$b(t) = \begin{cases} +1, & a_I(t) \neq a_Q(t) \\ -1, & a_I(t) = a_Q(t) \end{cases}, \quad \theta(t) = \begin{cases} 0, & a_I(t) = 1 \\ \pi, & a_I(t) = -1 \end{cases} \quad (\text{A2})$$

where $a_I(t)$ and $a_Q(t)$ represent the in-phase and quadrature bit streams (encoded with ± 1), respectively, and the signals $b(t)$ and $\theta(t)$ have pulse widths of T . In this view, the MSK signal may be thought of as four separately transmitted sine waves: two are transmitted at $f_c + 1/(4T)$ and are π out of phase, while the other two are at $f_c - 1/(4T)$ and are also π out of phase. At each interval boundary Tk , $k = 0, \pm 1, \pm 2, \dots$, the signal $s(t)$ transitions to one of these four sine waves in a phase-continuous fashion.

We wish to estimate and remove the MSK phase modulation in order to track the phase variation of the carrier frequency. Propagation and receiver hardware effects introduce an apparent carrier offset $\phi_0(t)$ and a bit transition delay $\tau(t)$, where the latter is proportional to the clock phase $\phi_1(t)$ and is influenced by the group delay of the signal. Letting

$$\phi_{\text{MSK}} \equiv b(t)\pi t/(2T) + \theta(t), \quad (\text{A3})$$

we can write the received signal $x(t)$ as

$$x(t) = A(t) \cos \left[2\pi f_c t + \phi_0(t) + \phi_{\text{MSK}}(t - \tau(t)) \right]. \quad (\text{A4})$$

Since ϕ_0 is an offset to the carrier phase, it is a proxy for the phase velocity between the source and receiver. On the other hand, since τ tracks the delay of a narrowband signal that modulates the carrier phase, it serves as a proxy for the group delay at f_c .

The first step to estimate an MSK signal is to convert $x(t)$ into the baseband equivalent analytic signal. In practice, this conversion is easily accomplished by applying a low-pass filter to the frequency-shifted complex signal $2x(t) \exp(-j2\pi f_c t)$. Dropping the assumed time dependence of the carrier phase $\phi_0(t)$ and the bit transition time $\tau(t)$, we can write

$$z(t) = A(t)e^{j\phi(t)} = A(t) \exp[j\phi_0 + j\phi_{MSK}(t - \tau)] \quad (A5)$$

Using $z(t)$, we can immediately measure the amplitude $A(t)$ by taking the magnitude $|z(t)|$. The real signal $A(t)$ can be resampled to provide the desired resolution for the application in question. We also wish to estimate ϕ_0 from $z(t)$. With reference to (A2), squaring $z(t)$ removes the contribution of $\theta(t)$ since $2\theta(t)$ is a multiple of 2π . Hence, the angle of the squared signal reduces to

$$\angle[z(t)]^2 = 2\phi_0(t) \pm \frac{\pi t}{T} \pm 2\phi_1 \quad (A6)$$

where the sign of the last two terms are determined by the instantaneous sign of $b(t - \tau)$ and

$$\phi_1 \equiv -\pi\tau/(2T) \quad (A7)$$

That is, $[z(t)]^2$ is an FSK signal that alternates between two tones at $\pm 1/2T$ with respective phases

$$\phi_{\pm} = 2\phi_0 \pm 2\phi_1. \quad (A8)$$

If we estimate $\phi_{\pm}$ over a given interval, then the carrier and bit clock phases can be estimated as $\hat{\phi}_0 = (\phi_+ + \phi_-)/4$ and $\hat{\phi}_1 = (\phi_+ - \phi_-)/4$, respectively. A 2π jump in either ϕ_+ or ϕ_- in isolation adjusts each of $\hat{\phi}_0$ and $\hat{\phi}_1$ by $\pi/2$, which is equivalent to reversing the roles of the in-phase and quadrature components, plus a possible sign adjustment on the bit sequences. If both ϕ_+ and ϕ_- shift by 2π , then ϕ_0 and ϕ_1 change by a multiple of π . In applications that track the carrier or clock phase over time, such phase shifts can be mitigated by unwrapping 2π phase shifts on $\phi_{\pm}$ before the calculation of $\hat{\phi}_0$ and $\hat{\phi}_1$.

In general, $\phi_{\pm}$ varies with time as the channel impulse response changes due to fluctuations in the ionospheric profile along the propagation path. Assuming each phase is constant over a time duration a , integrating a shifted version of $[z(t)]^2$ by $\pm 1/2T$ over a gives an estimate of $\phi_{\pm}$:

$$\phi_{\pm}(t) \approx \angle \int_t^{t+a} [z(\xi)]^2 \exp(\mp j\pi\xi/T) d\xi \quad (A9)$$

With a lock on the carrier and bit transition phases over a time interval a , it is possible to coherently demodulate the MSK signal. Assuming Gaussian noise, the optimal demodulation uses matched filtering on the in-phase and quadrature components separately, where the integration period is $2T$ for each component. Writing the baseband signal $z(t)$ in terms of the in-phase and quadrature components with the bit transition delay represented with the phase offset ϕ_1 , we have

$$z(t) = A(t) \left[a_I(t) \cos\left(\frac{\pi t}{2T} + \phi_1\right) - ja_Q(t) \sin\left(\frac{\pi t}{2T} + \phi_1\right) \right] e^{j\phi_0} \quad (A10)$$

The in-phase signal $a_I(t)$ bit transition boundaries occur at $t = t_1 + T + 2Tk$, $k = 0, \pm 1, \dots$, where $t_1 = -4T\phi_1/2\pi$. Similarly, the quadrature signal $a_Q(t)$ bit transition boundaries occur at $t = t_1 + 2Tk$. For each index k , we have

$$\hat{a}_{Ik} = \frac{1}{T} \int_{t_1+2Tk}^{t_1+T+2T(k+1)} \Re \left\{ z(t) e^{-j\hat{\phi}_0} \right\} \cos\left(\frac{\pi t}{2T} + \hat{\phi}_1\right) dt \quad (A11)$$

$$\hat{a}_{Qk} = \frac{1}{T} \int_{t_1+2Tk}^{t_1+T+2T(k+1)} \Im \left\{ z(t) e^{-j\hat{\phi}_0} \right\} \sin\left(\frac{\pi t}{2T} + \hat{\phi}_1\right) dt \quad (A12)$$

Evaluating the sign of $\hat{a}_{Ik}$ and $\hat{a}_{Qk}$ at each index k gives the optimal estimate of the two digital sequences. Using (A2) and (A7) we can reconstruct a (shifted) version of the estimate $\hat{\phi}_{MSK}(t - \hat{\tau})$. Subtracting the resulting phase trellis from $z(t)$ gives an estimate of the phase variation due to propagation and receiver effects in the absence of MSK phase modulation,

$$A(t)e^{j\phi_0(t)} \approx z(t)e^{-j\hat{\phi}_{MSK}(t - \hat{\tau})}. \quad (A13)$$

Acknowledgments

This work was supported by the National Science Foundation under grants AGS 1451142 and AGS 1653114 (CAREER) to the Georgia Institute of Technology, and by grant AGS 1451210 to the University of Colorado Denver. We thank Vaisala, Inc. for providing NLDN data under a research agreement. Narrowband data may be downloaded from <https://doi.org/10.5281/zenodo.1134910>. GOES data, from <https://satdat.ngdc.noaa.gov/sem/goes/data/full>, were used to classify the solar flare. We also thank the referees for their assistance in evaluating this paper.

References

- Bainbridge, G., & Inan, U. S. (2003). Ionospheric *D* region electron density profiles derived from the measured interference pattern of VLF waveguide modes. *Radio Science*, 38(4), 1077. <https://doi.org/10.1029/2002RS002686>
- Bracewell, R. N. (1952). Theory of formation of an ionospheric layer below *E* layer based on eclipse and solar flare effects at 16 kc/sec. *Journal of Atmospheric and Solar-Terrestrial Physics*, 2(4), 226–235.
- Burgess, B., & Jones, T. B. (1967). Solar flare effects and VLF radio wave observations of the lower ionosphere. *Radio Science*, 2(6), 619–626.
- Clilverd, M. A., Rodger, C. J., Thomson, N. R., Lichtenberger, J., Steinback, P., Cannon, P., & Angling, M. J. (2001). Total solar eclipse effects on VLF signals: Observations and modeling. *Radio Science*, 36(4), 773–788.
- Cohen, M. B., Inan, U. S., & Paschal, E. P. (2010). Sensitive broadband ELF/VLF radio reception with the AWESOME instrument. *IEEE Transactions on Geoscience and Remote Sensing*, 48(1), 3–17. <https://doi.org/10.1109/TGRS.2009.2028334>
- Cotts, B. R. T. (2011). Global quantification of lightning-induced electron precipitation using very low frequency remote sensing, (PhD thesis), Stanford, UK: Stanford University.
- Cotts, B. R. T., & Inan, U. S. (2007). VLF observation of long ionospheric recovery events. *Geophysical Research Letters*, 34, L14809. <https://doi.org/10.1029/2007GL030094>
- Cummer, S. A., Inan, U. S., & Bell, T. (1998). Ionospheric *D* region remote sensing using VLF radio atmospherics. *Radio Science*, 33(6), 1781–1792.
- Cummins, K. L., & Murphy, M. J. (2009). An overview of lightning locating systems: History, techniques, and data uses, with an in-depth look at the U.S. NLDN. *IEEE Transactions on Electromagnetic Compatibility*, 51(3), 499–518.
- Dowden, R. L., & Brundell, J. B. (1996). Are VLF rapid onset, rapid decay perturbations produced by scattering off sprite plasma? *Journal of Geophysical Research*, 101(D14), 19,175–19,183.
- Dowden, R. L., Brundell, J. B., & Rodger, C. J. (1997). Temporal evolution of very strong Trimpis observed at Darwin, Australia. *Geophysical Research Letters*, 24(19), 2419–2422.
- Fishman, G. J., & Inan, U. S. (1988). Observation of an ionospheric disturbance caused by a gamma-ray burst. *Nature*, 331, 418–420.
- Golkowski, M., & Inan, U. S. (2008). Multistation observations of ELF/VLF whistler mode chorus. *Journal of Geophysical Research*, 113, A08210. <https://doi.org/10.1029/2007JA012977>
- Golkowski, M., Gross, N. C., Moore, R. C., Cotts, B. R. T., & Mitchell, M. (2014). Observation of local and conjugate ionospheric perturbations from individual oceanic lightning flashes. *Geophysical Research Letters*, 41, 273–279. <https://doi.org/10.1002/2013GL058861>
- Haldoupis, C., Cohen, M. B., Arnone, R., Cotts, B. R. T., & Dietrich, S. (2013). The VLF fingerprint of elves: Step-like and long-recovery early VLF perturbations caused by powerful $\pm$ CG lightning EM pulses. *Journal of Geophysical Research: Space Physics*, 118, 5392–5402. <https://doi.org/10.1002/jgra.50489>
- Haldoupis, C., Neubert, T., Inan, U. S., Mika, A., Allin, T. H., & Marshall, R. A. (2004). Subionospheric early VLF signal perturbations observed in one-to-one association with sprites. *Journal of Geophysical Research*, 109, A10303. <https://doi.org/10.1029/2004JA010651>
- Han, F., & Cummer, S. A. (2010a). Midlatitude nighttime *D* region ionosphere variability on hourly to monthly timescales. *Journal of Geophysical Research*, 115, A09323. <https://doi.org/10.1029/2010JA015437>
- Han, F., & Cummer, S. A. (2010b). Midlatitude daytime *D* region ionosphere variability measured from radio atmospherics. *Journal of Geophysical Research*, 115, A10314. <https://doi.org/10.1029/2010JA015715>
- Han, F., Cummer, S. A., Li, J., & Lu, G. (2011). Daytime ionospheric *D* region sharpness derived from VLF radio atmospherics. *Journal of Geophysical Research*, 116, A05314. <https://doi.org/10.1029/2010JA016299>
- Higginson-Rollins, M. A., & Cohen, M. B. (2017). Exploiting LF/MF signals of opportunity for lower ionospheric remote sensing. *Geophysical Research Letters*, 44, 8665–8671. <https://doi.org/10.1002/2017GL074236>
- Hunsucker, R. D. (1982). Atmospheric gravity waves generated in the high-latitude ionosphere: A review. *Reviews of Geophysics*, 20(2), 293–315.
- Inan, U. S., Bell, T. F., & Rodriguez, J. V. (1991). Heating and ionization of the lower ionosphere by lightning. *Geophysical Research Letters*, 18(4), 705–708.
- Inan, U. S., Cummer, S. A., & Marshall, R. A. (2010). A survey of ELF/VLF research on lightning-ionosphere interactions and causative discharges. *Journal of Geophysical Research*, 115, A00E36. <https://doi.org/10.1029/2009JA014755>
- Inan, U. S., Pasko, V. P., & Bell, T. F. (1996). Sustained heating of the ionosphere above thunderstorms as evidenced in “early/fast” VLF events. *Journal of Geophysical Research*, 101, 1067–1070.
- Johnson, M. P., Inan, U. S., Lev-Tov, S. J., & Bell, T. F. (1999). Scattering pattern of lightning-induced ionospheric disturbances associated with early/fast VLF events. *Geophysical Research Letters*, 26(15), 2363–2366.
- Kabirzadeh, R., Marshall, R. A., & Inan, U. S. (2017). Early/fast VLF events produced by the quiescent heating of the lower ionosphere by thunderstorms. *Journal of Geophysical Research: Atmospheres*, 122, 6217–6230. <https://doi.org/10.1002/2017JD026528>
- Kaiser, A. B. (1968). Identification of a new type of mode conversion interference in VLF propagation. *Radio Science*, 3(6), 545–549.
- Kotovsky, D. A., & Moore, R. C. (2015). Classifying onset durations of early VLF events: Scattered field analysis and new insights. *Journal of Geophysical Research: Space Physics*, 120, 6661–6668. <https://doi.org/10.1002/2015JA021370>
- Lay, E. H., & Shao, X. M. (2011). High temporal and spatial-resolution detection of *D*-layer fluctuations by using time-domain lightning waveforms. *Journal of Geophysical Research*, 116, A01317. <https://doi.org/10.1029/2010JA016018>
- Marshall, R. A., Inan, U. S., & Chevalier, T. W. (2008). Early VLF perturbations caused by lightning EMP-driven dissociative attachment. *Geophysical Research Letters*, 35, L21807. <https://doi.org/10.1029/2008GL035358>
- Marshall, R. A., Inan, U. S., & Glukhov, V. S. (2010). Elves and associated electron density changes due to cloud-to-ground and in-cloud lightning discharges. *Journal of Geophysical Research*, 115, A00E17. <https://doi.org/10.1029/2009JA014469>
- Maxworth, A. S., Golkowski, M., Cohen, M. B., Chorsi, H. T., Gedney, S. D., & Jacobs, R. (2015). Multistation observations of the azimuth, polarization, and frequency dependence of ELF/VLF waves generated by electrojet modulation. *Radio Science*, 50, 270–278. <https://doi.org/10.1002/2015RS005683>
- McRae, W. M., & Thomson, N. R. (2004). Solar flare induced ionospheric *D*-region enhancements from VLF phase and amplitude observation. *Journal of Atmospheric and Solar-Terrestrial Physics*, 66(1), 77–87.
- Mika, A., Haldoupis, C., Neubert, T., Su, H. T., Hsu, R. R., Steiner, R. J., & Marshall, R. A. (2006). Early VLF perturbations observed in association with elves. *Annales Geophysicae*, 24(8), 2179–2189.
- Mitra, A. P. (1974). *Ionospheric effects of solar flares*. Dordrecht, Netherlands: D. Reidel.
- Moore, R. C., Barrington-Leigh, C. P., Inan, U. S., & Bell, T. F. (2003). Early/fast VLF events produced by electron density changes associated with sprite halos. *Journal of Geophysical Research*, 108(A10), 1363. <https://doi.org/10.1029/2002JA009816>

- NaitAmor, S., AlAbdoadain, M. A., Cohen, M. B., Cotts, B. R. T., Soula, S., Chanrion, O., ... Abdelatif, T. (2010). VLF observations of ionospheric disturbances in association with TLEs from the Eurosprite-2007 campaign. *Journal of Geophysical Research*, 115, A00E47. <https://doi.org/10.1029/2009JA015026>
- NaitAmor, S., Cohen, M. B., Cotts, B. R. T., Ghalila, H., Alabdoadain, M. A., & Graf, K. L. (2013). Characteristics of long recovery early VLF events observed by the North African AWESOME network. *Journal of Geophysical Research: Space Physics*, 118, 5215–5222. <https://doi.org/10.1002/jgra.50448>
- Nemec, F., Santolik, O., & Parrot, M. (2009). Decrease of intensity of ELF/VLF waves observed in the upper ionosphere close to earthquakes: A statistical study. *Journal of Geophysical Research*, 114, A04303. <https://doi.org/10.1029/2008JA013972>
- Pappert, R. A., & Morfitt, D. G. (1975). Theoretical and experimental sunrise mode conversion results at VLF. *Radio Science*, 10(5), 537–546.
- Pasko, V. P., Inan, U. S., & Bell, T. F. (1996). Blue jets produced by quasi-electrostatic pre-discharge thundercloud fields. *Geophysical Research Letters*, 23(3), 301–304.
- Peter, W. B., & Inan, U. S. (2004). On the occurrence and spatial extent of electron precipitation induced by oblique nonducted whistler waves. *Journal of Geophysical Research*, 109, A12215. <https://doi.org/10.1029/2005JA011346>
- Peter, W. B., & Inan, U. S. (2007). A quantitative comparison of lightning-induced electron precipitation and VLF signal perturbations. *Journal of Geophysical Research*, 112, A12212. <https://doi.org/10.1029/2006JA012165>
- Poorya, H., Golkowski, M., & Truner, D. L. (2017). Unique concurrent observations of whistler mode hiss, chorus, and triggered emissions. *Journal of Geophysical Research: Space Physics*, 122, 6271–6282. <https://doi.org/10.1002/2017JA024072>
- Said, R. K. (2009). Accurate and efficient long-range lightning geo-location using a VLF radio atmospheric waveform bank, (PhD thesis). Stanford, UK: Stanford University.
- Said, R. K., Inan, U. S., & Cummins, K. L. (2010). Long-range lightning geo-location using a VLF radio atmospheric waveform bank. *Journal of Geophysical Research*, 115, D23108. <https://doi.org/10.1029/2010JD013863>
- Salut, M. M., Cohen, M. B., Ali, M. A. M., Graf, K. L., Cotts, B. R. T., & Kumar, S. (2013). On the relationship between lightning peak current and early VLF perturbations. *Journal of Geophysical Research: Space Physics*, 118, 7272–7282. <https://doi.org/10.1002/2013JA019087>
- Silber, I., Price, C., Galanti, E., & Shuval, A. (2015). Anomalously strong vertical magnetic fields from distant ELF/VLF sources. *Journal of Geophysical Research: Space Physics*, 120, 6036–6044. <https://doi.org/10.1002/2015JA021141>
- Swanson, E. R. (1983). Omega. *Proceedings of the IEEE*, 71(10), 1140–1155.
- Thomson, N. R. (1993). Experimental daytime VLF ionospheric parameters. *Journal of Atmospheric and Solar-Terrestrial Physics*, 55, 831–834.
- Thomson, N. R., & Clilverd, M. A. (2001). Solar flare induced ionospheric D-region enhancements from VLF amplitude observations. *Journal of Atmospheric and Solar-Terrestrial Physics*, 63, 1729–1737.
- Thomson, N. R., Clilverd, M. A., & McRae, W. M. (2007). Nighttime ionospheric D region parameters from VLF phase and amplitude. *Journal of Geophysical Research*, 112, A07304. <https://doi.org/10.1029/2007JA012271>
- Thomson, N. R., Clilverd, M. A., & Rodger, C. J. (2014). Low-latitude ionospheric D region dependence on solar zenith angle. *Journal of Geophysical Research: Space Physics*, 119, 6865–6875. <https://doi.org/10.1002/2014JA020299>
- Voss, H. D., Imhof, W. L., Walt, M., Mobilia, J., Gaines, E. E., Reagan, J. B., ... Chang, H. C. (1984). Lightning-induced electron precipitation. *Nature*, 312, 740–742.
- Watt, A. D. (1967). *VLF radio engineering*. Oxford, New York: Pergamon Press.
- Wood, T. G., & Inan, U. S. (2004). Localization of individual lightning discharges via directional and temporal triangulation of sferic measurements and two distant sites. *Journal of Geophysical Research*, 109, D21109. <https://doi.org/10.1029/2004JD005204>
- Zoghzyghy, F. G. (2015). Statistical analysis and modeling lightning using radio remote sensing, (PhD thesis). Stanford, UK: Stanford University.