A Multiple-Mode Three-Dimensional Model of VLF Propagation in the Earth-Ionosphere Waveguide in the Presence of Localized D Region Disturbances

WILLIAM L. PoulSEN,1 UMRAN S. INAN, AND TIMOTHY F. BELL

Space, Telecommunications and Radioscience Laboratory, Stanford University, Stanford, California

Transient localized D region disturbances, such as those associated with lightning discharges, affect the characteristics of VLF waves propagating in the Earth-ionosphere waveguide. In particular, both phase and amplitude changes in the subionospheric signal can be observed at receiving sites as a result of the wave scattering that takes place in the disturbed region. In the present paper we present a multiple-mode three-dimensional model of VLF propagation in the Earth-ionosphere waveguide in the presence of localized D region disturbances. The model takes into account great circle (GC) propagation paths with realistic ground and ionospheric conductivity changes that result in mode conversion along the path. It is assumed that conductivity changes transverse to the GC paths are negligible except in the vicinity of the D region disturbance and that mode coupling is negligible within the disturbed region. This new model is applied to experimental observations and is found to be in general agreement. The diagnostics potential of the model for characterizing energetic particle precipitation events is discussed.

1. INTRODUCTION

Very low frequency remote sensing has recently been increasingly used to study ionospheric disturbances associated with lightning discharges [Inan et al., 1990 and references therein]. While extensive data have been acquired on the variations of amplitude and phase of subionospheric VLF signals on a variety of great circle paths, the interpretation of such data in terms of the physical characteristics of the disturbance (e.g., location, size, altitude profile) requires a quantitative model of VLF propagation in the Earth-ionosphere waveguide. Earlier theoretical models have treated the Earth-ionosphere waveguide, in which the VLF signal propagates, as infinite and homogeneous in the horizontal dimension transverse to the direction of propagation [Tolstoy, 1983; Inan et al., 1985; Tolstoy et al., 1986; Inan and Carpenter, 1987]. In this picture, the perturbation is assumed to lie on the great circle (GC) path of propagation and is assumed to extend to infinity in the transverse dimension. The problem is thus reduced to one having two dimensions: the height of the waveguide and the distance along the path of propagation as depicted in Figure 1. Analysis of experimental data has indicated the need for a more general treatment. Carpenter and LaBelle [1982] found that perturbations associated with ducted whistlers can at times be located at significant distances (up to ~200 km) transverse to the GC path between transmitter and receiver, and Inan and Carpenter [1987] recognized the need for more realistic three-dimensional models which include the effects of off-GC-path locations of ionospheric perturbations. Recently, Dowden and Adams [1988, 1989a] have put forward a heuristic three-dimensional model based on “echoes” from lightning-induced electron precipitation (LEP) ionization “patches” or ridges located off the GC path, and PoulSEN et al. [1990] have developed a three-dimensional model for scattering by D region perturbations based on waveguide mode theory.

PoulSEN et al. [1990] considered a single-mode method of modeling the effects of scattering caused by localized enhancement disturbances (also referred to as scatterers) in the lower ionosphere. The entire ambient path was assumed to be horizontally homogeneous, and the disturbances were assumed to involve gradual (i.e., slowly varying) changes in electron density over distances in the horizontal direction of the order of a wavelength, thus allowing the use of the WKB ( Wentzel, Kramers, and Brillouin) approximation. Also, the maximum change in electron density from the ambient values was assumed to be of sufficiently small magnitude that the Born approximation [Born and Wolf, 1965] could be used within the scattering region. In the present paper this single-mode scattering method is extended to the case of multiple modes in the WKB limit. We consider more realistic great circle propagation paths by taking into account the conversion between the different modes as the signal propagates along paths which have changing ionospheric as well as ground conductivities and permittivities (1) between the transmitter and receiver, (2) between the transmitter and the disturbance, and (3) between the disturbance and the receiver (Figure 1). We extend the single-mode method to take into account all non negligible modes that arrive at, and are scattered by, the disturbance. However, we continue to assume that the nature of the disturbance is such that the WKB approximation still holds within the disturbed region and that, therefore, mode conversion within the scattering disturbance region can be neglected.

For the case of a single dominant waveguide mode, it was found (as described by PoulSEN et al. [1990]) that the scattering effect of a localized electron density enhancement is sensitively dependent on the altitude profile of ionization in the disturbed region. In other words, it was found that the disturbance should not merely be treated as a reflective scatterer (simply leading to a change in the effective reflection height as for example assumed by Dowden and Adams [1989a]), but that rather, the actual altitude profile of electron density within the disturbance (as compared to the ambient density profile) must be considered in determining the effects
that a particular disturbance will have on the propagating VLF wave. Our results here show that the altitude profile of enhanced ionization within the disturbed region is an important determinant in the more general multiple-mode cases as well, especially since each waveguide mode is scattered a different amount relative to the other modes of the wave by the same disturbance.

The underlying methodology of the new multiple-mode three-dimensional model is depicted in Figure 1. Basically, a computer code, referred to as the Long Wave Propagation Capability (LWPC) and developed by the Naval Ocean Systems Center [Pappert and Snyder, 1972; Ferguson and Snyder, 1987], is used to calculate the complex field amplitude of the wave as it travels along the direct path between the transmitter and receiver ("direct path"), and along the paths from the transmitter to the disturbance ("leg 1") and from the disturbance (or scatterer) to the receiver ("leg 2"), taking into account the propagation and attenuation of multiple waveguide modes and the coupling between modes that occurs due to the changes in ground conductivity and ionospheric variations that may occur along each path. In addition, the single-mode three-dimensional method described by Poulsen et al. [1990] is used to calculate the magnitude and phase of the signal scattered by the disturbance toward the receiver on a mode-by-mode basis, assuming that no conversion among modes occurs within the disturbed region. (This assumes that the WKB approximation is valid, in other words, that any variations in the ionospheric or ground properties are gradual and relatively small over distances of ~1λ in the horizontal directions.)

In the following, we describe the basic formulation and the underlying assumptions for the new three-dimensional model and illustrate the application of the model to a particular path.

2. DESCRIPTION OF THE MODEL

The starting points for our development of the new multiple-mode three-dimensional model are (1) the expression for the total electric field (equation (1) below) as derived by Wait [1970], and (2) the scattering formulation of Poulsen et al. [1990] (equation (2) below). Wait [1970] showed that under undisturbed ionospheric conditions, with the space between the Earth and the ionosphere taken to constitute a spherically concentric waveguide, and with homogeneous conditions along the entire path, the vertical electric field \( E_{\text{total}} \) at a great circle path (GCP) distance \( d \) (much greater than a wavelength) from the transmitter is in the form

\[
E_{\text{total}}(d) \approx \frac{1}{\sqrt{\sin(d/R_E)}} \sum_{n} A_n^{T,R} e^{-ik_nS_nd} \tag{1}
\]

where \( n \) is the mode number, \( i = \sqrt{-1} \), and \( S_n = \sin \theta_n \) (\( \theta_n \) is the complex-valued "angle of incidence" of each mode on the ionosphere). \( S_n \) is equivalent to the complex index of refraction for mode \( n \) and thus determines the attenuation and phase velocity of that mode. In the general case, \( S_n \) is a function of the waveguide properties at each point along the propagation path. \( R_E \) is the radius of the Earth, \( k_n = 2\pi/\lambda_n \), where \( \lambda_n \) is the free space wavelength of the signal, and \( A_n^{T,R} \) is a complex-valued function which includes the excitation and the height gain factors at the transmitter and receiver and is dependent only on conditions at the transmitter and receiver locations. The factor

\[
\frac{1}{\sqrt{\sin(d/R_E)}}
\]

represents a cylindrical spreading factor on a spherical surface and is a first-order approximation resulting from the asymptotic expansions of Legendre functions.

All of the variables in equation (1) are functions only of the properties of the waveguide at the transmitter or receiver location, or of the transmitter and receiver antenna, except for \( S_n \). The factor \( A_n^{T,R} \) can be written out more completely as

\[
A_n^{T,R} = KT_n R_n
\]

where \( K = a\sqrt{P_f} \); \( a \) is a constant coefficient; \( P \) is the radiated power; \( f \) is the signal frequency;

\[
T_n = \sum_{j=1}^{3} t_j A_j^n T_j^n(z)
\]

and \( R_n = \sum_{j=1}^{3} t_j A_j^n R_j^n(x) \)

and \( j = 1, 2, 3 \) represent the three-dimensional coordinates \( x, y, \) and \( z \), where \( x \) is the distance along the GCP; \( t_j \) and \( r_j \) are factors which account for the orientation of the transmitter and the receiver antenna, respectively; and \( A_j^n T_j^n(z) \) and \( A_j^n R_j^n(x) \) are the "initial excitation" factors for each mode for the transmitter or receiver antenna, respectively [Budden, 1961]. These factors represent the relative amount of each mode that each component of the antenna excites (or, by reciprocity, receives). (The initial excitation factor \( \Lambda \) as written here is equivalent to \( \sqrt{\Lambda} \) as defined by Ferguson and Snyder [1980].) \( A_j^n T_j^n(z) \) are known as the "height gain" factors for each mode for the transmitter or receiver, respectively. They represent the relative amplitude and phase of each of the three components of each waveguide mode as a function of altitude, at the transmitter or receiver.

Equation (1) is a three-dimensional formulation which applies only to a homogeneous waveguide. In the case of the Earth-ionosphere waveguide we can expect that the waveguide properties may change both along the propagation path and transverse to the propagation path. Thus the \( S_n \) in (1) will be functions of position in the waveguide. Wait [1964a] has shown how to modify (1) appropriately for the case in which the waveguide properties vary slowly (WKB approximation) both along and transverse to the propagation path. This type of solution is three-dimensional and includes refraction effects due to changes in the \( S_n \) transverse to the GCP between transmitter and receiver. In the present work we are interested in localized \( D \) region disturbances which may be located many wavelengths off the GCP between transmitter and receiver. In order to focus attention on the three-dimensional scattering properties of these disturbances, we simplify our model by assuming that the \( S_n \) in (1) do not vary significantly over a distance of one wavelength (~10 km) transverse to the three great circle paths shown in Figure 1. This assumption is equivalent to
neglecting refraction effects which may occur on these paths. This assumption is discussed further in section 4.

The single-mode scattering formulation used by Poulken et al. [1990] was based on Equation (6) of that paper, which is reproduced below. The relevant coordinate system and vector relationships are given in Figure 2.

\[
\begin{align*}
\epsilon_n(x, y) &= \epsilon_n^d(x, y) + \epsilon_n^s(x, y) \\
&= \left( \frac{ik_0^2}{4} \right) \int_{\mathcal{P}} [S_n^2(x', y') - (S_n^s)^2] \frac{\epsilon_n(x', y')}{\Delta S^2} \left. \right|_{\text{field inside disturbance}} \\
&= \frac{H_0^2(k_0, S_n^R, R')}{\Delta S^2} \int_{\mathcal{P}} \epsilon_n(x', y') dz' \, dy'
\end{align*}
\]

(2)

where \( \epsilon_n \) is the total modal field for the mode of order \( n \) (i.e., the total field seen at \( (x, y) \) in the presence of some disturbance in the waveguide); \( \epsilon_n^d \) is the unperturbed modal field, also called the “direct” field (i.e., the field seen at \( (x, y) \) in the absence of any disturbance); \( \epsilon_n^s \) is the secondary or “scattered” field (i.e., the field seen at \( (x, y) \) due to the disturbance); and \( S_n^R \) is the ambient value of \( S_n \) in the absence of any disturbance. (In the single-mode analysis, it is a constant independent of \( x \) and \( y \)). As seen from the above the scattering is proportional to \( \Delta S^2 \), i.e., the change in the mode refractive index with respect to the ambient. \( \mathcal{P} \) is the region of integration, or “patch”, which extends over that portion of the \( x-y \) plane that encompasses the disturbance (i.e., where \( S_n \neq S_n^R \)); \( H_0^2(k_0, \cdot) \) is a Hankel function of the second kind of order zero; and

\[
R' = \sqrt{\left( x - x' \right)^2 + \left( y - y' \right)^2}
\]

is the distance from each integration point \( (x', y') \) within the disturbance to the observation point \( (x, y) \).

To extend to the case of multiple-mode propagation and scattering we combine equations (1) and (2) to write

\[
E_{\text{total}} \simeq \sum_{n} A_{n}^T R \left[ e_n^d + e_n^s \right] = \sum_{n} A_{n}^T R \epsilon_n^d + \sum_{n} A_{n}^T R \frac{H_0^2(k_0, S_n^R)}{\sqrt{\sin d / R_E}} \left( \epsilon_n^s \right)
\]

(3)

where \( \epsilon_n^s \) is as in (2). Making the same substitutions as were made by Poulken et al. [1990], in particular, using the Born approximation (substituting \( \epsilon_n^s \simeq K e^{-ik_0 S_n^R d/\sqrt{d}} \) for \( \epsilon_n^s \) in the integrand and substituting the far-field approximations for \( \epsilon_n^d \) and \( H_0^2 \) (i.e., \( H_0^2 \simeq \frac{2i}{\pi k_0 S_n^R} R' \sqrt{\sin d / R_E} \)), respectively, we can write

\[
E_{\text{total}} \simeq \sum_{n} A_{n}^T R \left( \frac{e^{-ik_0 S_n^R d}}{\sqrt{\sin d / R_E}} \right) + \sum_{n} A_{n}^T R \left( \frac{-ik_0^2}{4} \right) \int_{\mathcal{P}} [S_n^2(x', y') - (S_n^s)^2] \frac{dz' \, dy'}{\Delta S^2}
\]

(4)

where all the variables have been previously defined.

2.1. Cylindrical Scattering on a Spherical Surface

and the LWPC Code

The far-field approximation for a cylindrical spreading wave as utilized by Poulken et al. [1990] (i.e., \( \epsilon_n^s \simeq K e^{-ik_0 S_n^R d/\sqrt{d}} \)) and in (5) above does not account for the fact that, in the earth-ionosphere waveguide, the wave cylindrically spreads over a spherically curved surface. Equation (4) is a more accurate representation of the wave’s spreading with propagation distance on the surface of the Earth [Watt, 1970; Ferguson and Snyder, 1980]. If we equate the far-field approximation for \( \epsilon_n^s \) (i.e., \( \epsilon_n^s \simeq K e^{-ik_0 S_n^R d/\sqrt{d}} \)) and (4) above and use the small angle approximation \( \sin x \approx x \) for \( x \ll 1 \), then we find that \( K = R_E \) for \( d \ll R_E \).

The LWPC code calculates the electric field versus distance using equation (1) for the total field and equation (4) for each waveguide mode. Since we use the LWPC code to calculate the field components along the three propagation paths illustrated in Figure 1, we need to match the two representations for \( \epsilon_n^s \) in the final formulation of the scattered field integral. Thus, we define

\[
e_n^{\text{LWPC}}(x) \equiv A_{n}^T R \frac{K e^{-ik_0 S_n^R x}}{\sqrt{\sin x / R_E}} \implies e_n(x)
\]

(6)

and then using \( K = \sqrt{R_E} \), we make the following substitution:

\[
A_{n}^T R \frac{K e^{-ik_0 S_n^R x}}{\sqrt{x}} \implies e_n^{\text{LWPC}}(x).
\]

(7)

Using this substitution in equation (5) (and rearranging), we find
\[ E_{\text{total}} \simeq \sum_{n} c_{n} \text{LWPC}(d) + \sum_{n} \left( \frac{-i k_{a} S_{n}^{2}}{4} \right) \left( \frac{2i}{\pi k_{a} S_{n}^{2}} \right)^{1/2} \times \int_{p} \left[ S_{n}^{2}(x', y') - (S_{n}^{2})^{2} \right] \{ e_{n} \text{LWPC}(R_{o}) \} \frac{e^{-ik_{a} S_{n}^{2} R_{o}'}}{\sqrt{R_{o}'}} dx' dy'. \]  

\begin{equation}
(8)
\end{equation}

The Hankel function term in the integrand of equation (2) essentially represents the retardation of each modal field toward the receiver over a flat Earth by each infinitesimal area element of the disturbance under the double integral. The double integral essentially sums the contributions to the scattered signal of all the infinitesimal radiators making up the disturbed region. We can extend the \( e^{-ik_{a} S_{n}^{2} R_{o}'}/\sqrt{R_{o}'} \) portion of the asymptotic expansion of the Hankel function term to account for the sphericity of the Earth's surface again by using equation (6). This means that we can take advantage of the LWPC code to calculate the propagation of the scattered signal from the disturbance to the receiver location and account for the spherical surface over which it propagates. Equation (8) can thus be rewritten as

\[ E_{\text{total}} \simeq \sum_{n} c_{n} \text{LWPC}(d) + \sum_{n} c_{n} \frac{\text{LWPC}(R')}{A_{n}^{2} K} dx' dy' \]

\begin{equation}
(9)
\end{equation}

where \( K = \sqrt{R_{o}} \) and

\[ c_{n} = \left( \frac{-i k_{a} S_{n}^{2}}{4} \right) \left( \frac{2i}{\pi k_{a} S_{n}^{2}} \right)^{1/2}. \]

\begin{equation}
(10)
\end{equation}

2.2. Simplifying Computational Approximations

Given enough computer time and patience, equation (9) can be used as written to calculate the total perturbed electric field caused by a lower ionospheric disturbance taking into account multiple-mode propagation and mode coupling over a real Earth but neglecting mode conversion within the disturbed region. However, the LWPC code must then be run to and from each infinitesimal area element of the double integration region \( P \). In order to reduce the very large amount of computer time required to do so, we assume that the spatial variation of the underlying ground conductivity along all paths going from the transmitter to all points within the disturbance region is the same. We further assume the same to be true for all paths leading from all the points within the disturbance region to the receiver, and choose the paths leading to (i.e., \( R_{oc} \)) Figure 3, and from (i.e., \( R_{c} \)), the center of the disturbed region as the representative paths. Thus, the ground conductivity along any radius within one of the two sectors shown in Figure 3 is assumed to be identical to that along the center radius of that sector. This is a reasonable approximation for disturbance regions of small radius \( a \) since it only requires that the ground conductivity vary slowly over distances \( r \) perpendicular to \( R_{oc} \) and \( R_{c} \). In the rare instances that the boundary between two regions of significantly different conductivity is nearly parallel to the centerline of the sector and lies under the sector area in ques-

\begin{equation}
(11)
\end{equation}

\[ e_{n} \text{LWPC}(R_{o}) e_{n} \text{LWPC}(R') = F_{n}(x', y') e_{n} \text{LWPC}(R_{o}) e_{n} \text{LWPC}(R') \]

where \( R_{o} \) and \( R' \), as shown in Figure 1, are functions of \( x' \) and \( y' \), but \( R_{oc} \) and \( R_{c} \), as defined above, are constants and can be moved outside the double integral. If we substitute equation (6) into equation (11) and solve for \( F_{n}(x', y') \), we find that

\[ F_{n}(x', y') = \left[ \begin{array}{c}
\sin \left( \frac{R_{oc}}{R_{o}} \right) \sin \left( \frac{R_{c}}{R_{o}} \right) \\
\sin \left( \frac{R_{oc}}{R_{o}} \right) \sin \left( \frac{R_{c}}{R_{o}} \right)
\end{array} \right]^{1/2} e^{-ik_{a} S_{n}^{2} [R_{o} - R_{oc}] - [R_{oc} + R_{c}]} \]

\begin{equation}
(12)
\end{equation}

where the variables are as previously defined.

Substituting equation (12) in equation (9) we obtain the final formulation that we use to describe the total electric field at the receiver in the general case of multiple-mode three-dimensional scattering from a disturbance region \( P \):

\[ E_{\text{total}} \simeq E_{\text{LWPC}}(d) + \sum_{n} c_{n} \frac{\text{LWPC}(R_{o})}{K A_{n}^{2} R} e_{n} \frac{\text{LWPC}(R_{c})}{K A_{n}^{2} R} \]

\begin{equation}
(13)
\end{equation}

\[ \int_{P} \left[ S_{n}^{2}(x', y') - (S_{n}^{2})^{2} \right] F_{n}(x', y') dx' dy'. \]
where $E_{\text{LWPC}}(d) = \sum_{n} \epsilon_n \epsilon_n^{\text{LWPC}}(R_{\text{sc}})$, which is the electric field that is obtained using the standard LWPC code along the direct transmitter-receiver path in the absence of any disturbances; $\epsilon_n^{\text{LWPC}}(R_{\text{sc}})$ is the electric field for mode $n$ calculated with the LWPC code along "leg 1" (Figure 1) from the transmitter to the center point of the disturbance; and $\epsilon_n^{\text{LWPC}}(R_{\text{sc}})/K_{A_n}$ is the electric field, also for mode $n$, calculated using the LWPC code along "leg 2" from the center point of the disturbance to the receiver, but with the transmitter and receiver antenna height gain and initial excitation factors removed from the LWPC calculations. ($K_{A_n}$ is calculated by LWPC when evaluating $\epsilon_n^{\text{LWPC}}(R_{\text{sc}})$ and only needs to be included once in the transmitter-scatterer-receiver propagation path calculation.) The factor $K$ is a result of the geometrical spreading factor substitution explained in section 2.1. The other variables have been previously defined.

The summation in equation (13) is in principle over an infinite number of modes, but beyond a finite number $N$ the mode field components become negligible in magnitude.

2.3. The Methodology

The methodology of our multiple-mode scattering formulation, based on equation (13), can now be summarized as follows, where steps 1–6 correspond to the label numbers in Figure 1:

1. LWPC is used along the GCP from the transmitter to the receiver to find the total electric field arriving along the "direct path" ($E_{\text{direct}}$).

2. LWPC is used along "leg 1" from the transmitter to the disturbance to find the value of the electric field for each mode $n$ arriving at the center of the disturbance region ($\epsilon_n$). This calculation includes the effects of both the transmitter and the receiver at which the scattered signal will eventually arrive.

3. A factor representing the signal strength scattered by the entire disturbance toward the receiver is calculated for each mode $n$, one mode at a time.

4. LWPC is used along "leg 2" from the center of the disturbed region to the receiver, but renormalized to account for only the propagation effects on each mode in traveling from the disturbance to the receiver. (The usual antenna excitation values for both the transmitter and the receiver have already been calculated in step 2).

5. The results of steps 2, 3, and 4 are combined to find the total electric field scattered by the disturbance which arrives at the receiver ($E_{\text{scattered}}^s$).

6. The direct and scattered electric fields are summed to obtain the total perturbed value of the electric field ($E_{\text{total}} = E_{\text{direct}} + E_{\text{scattered}}^s$).

7. The total perturbed value of the electric field at the receiver, $E_{\text{total}}$, is compared to the total ambient electric field at the receiver, $E_{\text{direct}}$, to calculate the change in amplitude $\Delta A$ and change in phase $\Delta \phi$ caused by the disturbance. The latter quantities are directly measured in experimental data when a localized disturbance suddenly appears, for example due to a burst of precipitating energetic electrons.

2.4. Application to the NSS-Stanford Path

An example of the application of the three-dimensional methodology (including the effects of ground variation along the 'direct path' and 'leg 1') is presented in Figures 5 and 6 for the propagation path between the NSS transmitter (Annapolis, Maryland, $f = 21.4$ kHz) and Stanford University and a particular example of an ionospheric disturbance. The altitude profiles of the ambient electron-neutral particle $\nu_+(\epsilon)$, positive ion-neutral particle $\nu_-(\epsilon)$, and negative ion-neutral particle $\nu_-(\epsilon)$ collision frequencies were assumed to be as described by Wait and Spier [1964] and Morfit and Shellman [1976], namely, $\nu_+ = 1.816 \times 10^{-5} \epsilon^{-0.15}$; $\nu_- = 4.540 \times 10^{-5} \epsilon^{-0.15}$; and $\nu_-(\epsilon) = 4.540 \times 10^{-5} \epsilon^{-0.15}$, where $\nu$ is in s$^{-1}$ and $\epsilon$ is in kilometers. The center of the disturbed region for this case was assumed to have the disturbed density profile shown in Figure 4, and the density enhancement was assumed to decrease with radial distance from the center and merge back to the ambient density proportional to a Gaussian function to produce a cylindrically symmetric transverse profile with $\sim 100$ km effective radius (i.e., $\Delta N_e(r', h') = \Delta N_e(r, h)e^{-r'^2/h'^2}$) with $a = 100$ km. The center of the disturbance is taken to be 100 km in the transverse direction away from a point 3000 km along the GCP from the NSS transmitter and on the northerly side of the NSS-Stanford path. Figure 5 shows plots of the amplitude and phase of the total vertical electric field (as given by equation (11)) along the three propagation paths ("legs") of Figure 2. The density profile of Figure 4 represents the electron density as a function of altitude that would result from electron precipitation bursts induced by lightning-generated whistlers propagating at $f \approx 2.0$ as calculated in the manner described by Inan et al. [1988a] for a total precipitated energy flux density for each case of $1.5 \times 10^{-2}$ ergs cm$^{-2}$ s$^{-1}$. The signal propagation along "leg 1" and the direct path are obtained using the LWPC software which takes account of the realistic variation of the ground parameters (using a conductivity map of the Earth’s surface) [Ferguson and Snyder, 1987]. The amplitude variation along "leg 1" is the same as that along the first ~1000 km of the direct path. For simplicity, we assume uniform ground conductivity along the scattered path (leg 2), where a larger number of modes of comparable amplitude (beginning at the scatterer) result in a more complicated amplitude variation. It should also be noted that the amplitude scale for leg 2 is different from that for the other two legs since the field scattered toward the receiver is generally much smaller than the field incident on the disturbance.

We note here that our assumption of uniform ground conductivity along leg 2 is merely to facilitate relatively simple interpretation of the results of Figures 5 and 6. The multiple-mode three-dimensional model is fully capable of accounting for mode coupling and ground conductivity variations along all three paths shown in Figure 1.

Fig. 4. Plot of the electron density distribution versus altitude at the center of a disturbance produced by an LEP event induced by whistlers propagating at $L \sim 2.0$. A typical value of 200 ma for the duration of the lightning discharge and subsequent LEP burst and a precipitated electron flux of $\sim 1.5 \times 10^{-2}$ erg cm$^{-2}$ s$^{-1}$ have been assumed in generating the profile, as discussed by Inan et al. [1988a]. The amont density profile used in the study is also shown.
Further insight into the propagation along the different path segments and the scattering can be gained upon examination of Figure 6. Here we show the relative signal strength and phase for each of the strongest propagating waveguide modes at various points along the three paths of propagation for the specific example described in the previous paragraph. The QTM and QTE designations refer to quasi-transverse magnetic and quasi-transverse electric modes, respectively, and the mode index n is defined such that, for example, the electric field for QTM1 exhibits two maxima (one at ground and one at ~60 km) in the waveguide. The units of vector length for each phasor diagram are in microvolts per meter, assuming a total radiated transmitter power of 1 kW. (We note here that this is simply a normalization for the purpose of discussing relative magnitudes. The typical radiated power levels for operational VLF transmitters are 10 kW to 1 MW, and the NSS transmitter nominally radiates ~250 kW.) As was noted in the previous paragraph, the phasor diagrams show generally larger amplitudes of higher-order modes being scattered by the disturbance that subsequently propagate along leg 2. Even at the receiver, there are substantial amounts of these higher-order modes still present. It is interesting, for example, that mode QTE2 has a larger magnitude than modes QTE3 or QTE4 incident on the disturbance (at the end of leg 1), but that upon scattering, mode QTE4 has a larger magnitude than either of modes QTE3 or QTE2, with mode QTE3 now having the smallest magnitude of the three and mode QTE4 having a magnitude nearly comparable to that of mode QTM2 (one of the two “dominant” modes). Mode QTE4 is still the third largest mode of the scattered signal that arrives at the receiver (end of “leg 2”), but plays a minor role in the signal that arrives at the receiver along the “direct path”. For this example the calculated change in total received amplitude ΔA of the signal is ~0.05 dB, accompanied by a phase change Δφ of ~0.06°. These values are in the range of amplitude and phase changes that have been observed on the NSS signal at Stanford during LEP events [Wolf, 1990; Wolf and Iman, 1990].

The procedure outlined above can be employed repeatedly to obtain values for ΔA and Δφ as the location of the disturbance is moved over the region between NSS and Stanford along the GCP, and on both sides of the GCP out to transverse distances where the effect of the disturbance becomes negligible (as discussed by Poulsen et al. [1990]). Thus, one could obtain a contour map of ΔA and Δφ due to this particular disturbance similar to that shown for the single-mode case (Figure 4 of Poulsen et al. [1990]), except that it would show predicted values for an actual path, accounting for the effects of the changing terrain that occurs across mid-North America. Given sufficient computer time, one could produce a “data base” of calculated values of ΔA and Δφ for a variety of disturbance sizes and disturbed electron density profiles for every transmitter-receiver path of interest. The data base of theoretical values could then be used in conjunction with actual experimental measurements of signal perturbations on those paths to make a first-order prediction of the possible locations, sizes, and disturbed electron density profiles of the disturbances believed to have caused those signal perturbations.

3. SOME RESULTS OBTAINED WITH THE MULTIPLE-MODE MODEL

In this section, the new three-dimensional multiple-mode model is used to investigate the physics of VLF wave propagation in the Earth–ionosphere waveguide in the presence of the type of lower ionosphere electron density enhancement disturbances described by Poulsen et al. [1990]. In order to concentrate on the effects on propagation of ionospheric variations and spatial parameters (i.e., distance between transmitter and receiver, location, and site of disturbance), we limit the Earth surface parameters (con-
ductivity and relative dielectric constant) to being homogeneous over the entire region underlying the propagation paths. Note that the assumption of an electrically homogeneous Earth is made simply for the purpose of clarity and is not necessary for the use of the three-dimensional model. Indeed, the NSS–Stanford example discussed in the previous section was based on a realistic Earth conductivity map for the direct path and leg 1.

3.1. The Role of the Disturbed Electron Density Profile

Figure 7 shows one set of examples of the effect that differences in the altitude profile of electron density within a disturbance can have on the received signal perturbations caused by that disturbance. Each of the four disturbed density profiles shown in Figure 7a was used as the density profile at the center of a disturbance of effective radius \( a = 150 \text{ km} \) (i.e., the density enhancement in the transverse direction varying as \( \Delta N_e(r', h) = \Delta N_0 \frac{r'}{a} e^{-r'/a^2} \)) centered at the midpoint of a 6000-km path with ground conductivity \( \sigma = 4 \text{ S/m} \) and \( \epsilon_r = 15 \) along the entire path. The profiles I, II, III, and IV represent the electron density as a function of altitude that would result from electron precipitation bursts induced by lightning-generated whistlers propagating at \( L = 3, 2.5, 2, \) and 1.6, respectively, calculated in the manner described by Inan et al. [1988a] for a total precipitated energy flux density for each case of \( \sim 1.5 \times 10^{-2} \text{ ergs cm}^{-2} \text{ s}^{-1} \). The electron precipitation bursts have a duration of \( \sim 200 \text{ ms} \). During the burst the local electron temperature increases due to the kinetic energy deposition. However, shortly (<1 ms) after the burst the electron temperature relaxes back to the ambient level, and it is to this immediate postburst period that our results apply.

The change in amplitude \( \Delta A \) and change in phase \( \Delta \phi \) of a 25-kHz signal caused by this disturbance were calculated for each of the four disturbed density profiles. The results of these calculations are displayed in Figure 7b. We note that some disturbed profiles cause relatively larger changes in amplitude \( \Delta A \) than other profiles accompanied by relatively smaller changes in phase \( \Delta \phi \).
ground parameters which were assumed to be $\sigma = 10^{-3}$ S/m and $c_r = 15$. While the $\Delta A$ for profile IV is larger in magnitude than for profile III, the $\Delta \phi$ for profile III is larger than that for profile IV and calculations indicate that the magnitude of the total signal scattered toward the receiver by a disturbance having disturbed profile III is only 4.3 dB lower than the direct signal, while the magnitude for that having profile IV is 5.9 dB lower than the direct signal. Thus, a disturbance with profile III again scatters a stronger signal than do disturbances having profiles I, II, or IV.

The physical reason for the results of Figures 7b and 8 becomes clear upon examination of the eigenangle solutions (i.e., $\theta_n$ in $S_n = \sin \theta_n$) as is shown in Figure 9 for a 25-kHz signal and $\sigma = 4$ S/m for two of the profiles of Figure 7. These solutions were obtained using the MODEFDR software (which solves the mode equation) [Morfats and Shellman, 1976] as part of the LWPC software [Ferguson and Snyder, 1987]. We see that most of the mode solutions for profile III are much further away from the corresponding ambient mode solutions than are those for profile IV, particularly for the QTM "branch" of solution angles. This configuration of $\theta_n$ indicates that most of the modes are scattered more strongly by a disturbance having profile III than by having profile IV. The separation between the disturbed and ambient mode solutions for profile III is also larger than for profiles I and II (not shown). The same result is found when the corresponding calculations and comparisons are made for frequencies of 15kHz and 50 kHz.

Further insight can be gained by examining the altitude pattern of the electric field as shown in Figures 10 and 11 for two representative waveguide modes (QTE$_7$ and QTM$_7$, respectively) for the 25-kHz case and for disturbances represented by profiles III and IV. When compared with the ambient electric field patterns for the same modes shown in the top panels, the patterns of Figures 10 and 11 illustrate the modification of the modal field component structures as a result of the disturbances. In particular, the QTM$_7$ field components are significantly different from the ambient for both the profile III and IV cases. For example, the $e_y$ component for QTM$_7$ is larger at all altitudes than the $e_y$ component as compared to the ambient component fields. And in the case of profile III, the QTM$_7$ field structure undergoes dramatic change from the ambient field structure. We note in particular that the QTM$_7$ field structure appears to undergo the greatest change from the ambient structure in the 50-80 km altitude region for profile IV, while the maximum modification with respect to the ambient for the case of profile III is near 70 km in altitude.

This illustrates the sensitive dependence of scattering caused by a disturbance on differences in the vertical electron density profile within the disturbed region. For example, profiles I and II cause a larger $\Delta A$ than profile IV, while profile IV causes a larger $\Delta \phi$ than either profile I or II. (The use of the terminology "profile X 'caused' a $\Delta A$" is a shorthand notation to refer to the change in the received signal amplitude $\Delta A$ caused by a disturbance having a vertical electron density profile at its center given by profile X.) However, we note that profile III causes both the largest $\Delta A$ and $\Delta \phi$, even though it represents an electron density enhancement which does not penetrate as deeply into the atmosphere as does that represented by profile IV.

An investigation of the amplitude and phase changes produced by these same profiles for different ground conductivities $\sigma$ and effective disturbance radii $a$ showed that the scattered signal represented by $\Delta A$ and $\Delta \phi$ was always largest for a disturbance having the disturbed density profile III. Figure 8 shows the resulting $\Delta A$ and $\Delta \phi$ for the same conditions as in Figure 7 except for the

![Figure 8](image-url)  
**Fig. 8.** Change in the received amplitude and phase caused by a lower ionospheric disturbance versus the four disturbed density profiles shown in Figure 7a for the same conditions as in Figure 7 except that $\sigma = 10^{-3}$ S/m.

![Figure 9](image-url)  
**Fig. 9.** Eigenangle solutions for $f = 25$ kHz and $\sigma = 4$ S/m for vertical electron density profiles III and IV of Figure 7a.
Fig. 10. Variation of electric field components with altitude for mode QTM of a 25-kHz signal for $\sigma = 4$ S/m and ionospheric electron density profiles III and IV of Figure 7a. The field structures for undisturbed ambient ionospheric conditions are shown in the top panel.

Fig. 11. Variation of electric field components with altitude for mode QTE of a 25-kHz signal for $\sigma = 4$ S/m and ionospheric electron density profiles III and IV of Figure 7a. The field structures for undisturbed ambient ionospheric conditions are shown in the top panel.

Fig. 12. Plots of conductivity $\sigma_o$ versus altitude corresponding to the density profiles shown in Figure 7a for a 25-kHz wave.
density profiles shown in Figure 7 for a wave frequency of 25 kHz and for a collision frequency profile as defined above in section 2.4.

We note that the largest increase in conductivity (with respect to the ambient) of any of the profiles occurs for profile III and lies in the ~68–75 km altitude region. Proportional increases in $\sigma_T$ and $\sigma_p$ also occur at these same altitudes, and this is the same region where we see the greatest change in the field structure of the QT$M_7$ mode for profile III shown in Figure 11. The largest increase in conductivity for profile IV occurs in the ~50–75 km altitude region, which is the same altitude range where the greatest change in the field structure of the QT$M_7$ mode for profile IV occurs. Thus, as we would expect on physical grounds, the scattering caused by a given ionospheric disturbance is sensitive to modifications in the altitude profile of the ionospheric "conductivity" within the disturbed region. Furthermore, for a given disturbance and for each waveguide mode, larger electrical conductivity changes with respect to the ambient lead to larger scattered field magnitudes.

3.2. Dependence on the Receiver Location

While the results shown in Figures 5 and 6 are for one particular size and location of the disturbance, the amplitude and phase changes seen at the receiver in the multiple-mode case also depend strongly on the location of the receiver, as well as on the location, size, and makeup of the disturbed region [Tolstoy, 1983; Tolstoy et al., 1986; Poulsen et al., 1990]. Thus, if the receiver were moved along the GCP beyond the disturbance, the measured change in amplitude and/or phase, $\Delta A$ and $\Delta \phi$, may typically vary from $\Delta A = 0$ dB or $\Delta \phi = 0^\circ$ to $|\Delta A| > 10$ dB or $|\Delta \phi| > 30^\circ$ [Poulsen, 1991]. We illustrate below this sensitive dependence on receiver location for an example path.

Figure 13a shows a plot of the calculated amplitude of a 25-kHz signal propagating over a homogeneous, $\sigma = 10^{-3}$ S/m, $\epsilon_r = 15$, ground surface. A cylindrically symmetric (three-dimensional) disturbance with effective radius $\alpha = 50$ km is centered directly over the GCP at a point 3000 km away from the transmitter. The electron density profile at the center of the disturbance was taken to be that of profile III of Figure 7a. The receiver location is then varied along the GCP from a point just beyond the disturbance out to ~6000 km away from the transmitter (or 3000 km away from the center of the disturbance). Figures 14b and 14c show the changes in amplitude ($\Delta A$) and in phase ($\Delta \phi$) that would be measured at the receiver due to the disturbance as a function of the receiver location. We note that $\Delta A$ and $\Delta \phi$ vary in both magnitude and sign as a function of the receiver location along the GCP away from the disturbance. An interesting difference from the single-mode behavior that is evident from Figures 13b and 13c is that, although each waveguide mode component suffers a reduction in amplitude and an advance in phase due to disturbances centered on the GCP (i.e., $\Delta A_h$ is negative and $\Delta \phi_h$ is positive for each mode $n$), the total vector sum of the modes may produce overall changes in the received signal amplitude and phase, $\Delta A$ and $\Delta \phi$, of both polarities. The largest magnitudes of $\Delta A$ and $\Delta \phi$ occur when the receiver is located in the deep null in the ambient signal strength located just before the 4000-km point along the GCP (see Figure 13a). The reason for this result is the fact that small changes in the mode structure of the wave at this location will produce large changes in the amplitude and phase relative to the small signal level at this point [Tolstoy, 1983; Tolstoy et al., 1986; Barr et al., 1985]. In terms of interpretation of experimental data and the design of new experiments, this result is important since, for a given signal-to-noise environment, the sensitivity to small ionospheric changes is greatly enhanced by locating the receiver near a null in the signal amplitude.

3.3. Dependence on the "Ambient" Ionospheric Density Profile

The potential importance of the result described in the previous section raises the question of the dependence of the ambient null locations on the ionospheric conditions. For the wave frequencies of interest here, ions do not play a significant role in determining the ionospheric conductivity, and only the electron density and electron–neutral collision frequency are important. Among these, the nighttime electron density is the one that is highly variable and difficult to determine without independent measurements. Figure 14 shows that differences in the ambient (undisturbed) ionospheric electron density profile significantly affect or alter the undisturbed signal strength versus distance "pattern" and hence the location of the nulls. Figure 14a shows two different "ambient" lower ionospheric electron density profiles used for purposes of comparison. Profile A is the simple exponential profile as used by Inan and Carpenter [1987] for a nighttime $D$ region ionosphere where $\beta = 0.5$ km$^{-1}$ and $h_0 = 85$ km. Profile B is the ambient nighttime profile used for the profiles of Figures 4 and 7 as well as in the work by Poulsen et al. [1990]. Figure 14b shows a plot of the eigenvalue solutions (i.e., $\theta_n$ in $S_n = \sin \theta_n$) of a 25-kHz wave for both profiles of Figure 14a and a seawater ($\sigma = 4$ S/m, $\epsilon_r = 81$) Earth surface. We note that almost none of the eigenvalue solutions of corresponding modes are identical or even close to one another in the complex plane. Figure 14c shows the signal strength versus distance corresponding to both sets of mode...
solutions shown in Figure 14b for homogeneous waveguide conditions. We note that the location, multiplicity, and depth of the nulls in the field strength are quite different for the two ambient ionospheric density profiles. For example, there is a null at a distance of approximately 2200 km from the transmitter for profile A whereas there is a relative maximum in the field strength at the same location for profile B.

The sensitive dependence of the $\Delta A$ and $\Delta \phi$ to be received at a fixed location on the ambient ionospheric conditions underscores the need to use a quantitative model, such as that presented here, in interpreting experimental data. In doing so, and in view of the general lack of knowledge of the lower ionospheric ambient density [Forbes, 1989; Ferguson et al., 1989], it is important that careful measurements of the signal under ambient conditions be made and compared with the model predictions to assess the ambient ionospheric profile. Once this is done, amplitude and phase changes due to transient disturbances can be properly interpreted.

3.4. Note on the Dominance of a "Dominant" Mode

Again referring to Figure 14c, we note that a deep null (an approximately 20-dB drop in signal level) occurs near the 4000-km point (for either of the two cases) even though the signal has propagated a relatively long distance over an all-sea path and in such cases we expect [Inan and Carpenter, 1987] that one, or maybe two, modes are "dominant." Indeed, in the case shown in Figure 14c, the strongest mode at the 4000-km point is the QTM$_2$ mode while the QTE$_2$ mode is just ~5-dB smaller in magnitude. Even if these two modes are exactly opposite in phase, their destructive interference would produce a drop in total signal amplitude of only ~7-dB. However, at the 4000-km point there are three other modes, QTM$_1$, QTM$_3$, and QTE$_3$, whose amplitudes are 15, 12, and 12 dB smaller, respectively, than the amplitude of the "dominant" QTM$_2$ mode. These three modes could, in tandem, combine to destructively interfere with the QTM$_2$ mode and cause an ~10-dB drop in the total signal amplitude. Thus, these "minor" modes, together with the QTE$_2$ mode, can (and do) sometimes combine together in phase so as to produce a 20-dB drop in total signal level at large propagation distances. Even if there were one "dominant" mode and the next three smaller modes were all >10 dB lower in amplitude, if they were in phase with each other and in opposition phase with the dominant mode, they could theoretically cause a 26-dB drop in the total signal level. Thus, care must be taken in deciding if a "dominant" mode is truly dominant at every point by taking into account the relative levels (amplitude and phase) of all the other nonnegligible modes.

4. DISCUSSION AND SUMMARY

In our model, in order to focus attention upon the three-dimensional wave scattering characteristics of localized $D$ region perturbations it is assumed that the $S_n$ functions in (1)-(13) do not vary significantly over the distance of one wavelength (~10 km) transverse to the great circle paths shown in Figure 1. In this case we are able to use realistic two-dimensional solutions to the waveguide propagation problem in order to determine the wave field that reaches the disturbed region and the scattered wave field that reaches the receiver.

Since the disturbed region can lie many wavelengths away from the GCP between transmitter and receiver, our model gives a three-dimensional representation of the effects of localized $D$ region perturbations on VLF wave propagation in the Earth-ionosphere waveguide. The assumption of no significant change in the $S_n$ functions transverse to the GCP would appear to be reasonable in many cases, for example, on a GCP that lies entirely over the ocean. On the other hand this assumption would not be reasonable on a GCP that is tangent at some point to a coastline. In the general case, if the $S_n$ vary rapidly with distance transverse to the GCP, then a completely numerical solution to the propagation problem must be carried out.

Some improvement to the present work could be obtained by allowing for slow variations in the $S_n$ transverse to the GCP. In this case, WKB wave solutions [Wait, 1964c] could be used in (2) to provide a somewhat more accurate representation of the zero-order (ambient) fields for those cases where refraction may be important. However, in general we do not expect that this small change would have a major effect upon the results reported here.

The usefulness of our three-dimensional multiple-mode model as a diagnostic tool to study localized $D$ region disturbances will depend upon our knowledge of four factors: (1) the ambient ionospheric electron density profile, (2) the ambient ionospheric electron-neutral collision frequency, (3) the transverse distance of the perturbed region from the GCP between transmitter and receiver, and (4) the horizontal scale of the perturbed region. For VLF frequencies we can safely neglect the effects of ions. As
mentioned in section 3.3, information on the ambient electron density and collision frequency profiles can be obtained from careful measurements of the unperturbed waveguide signal under ambient conditions. Furthermore, additional information concerning the four factors can be obtained through multiple frequency and spaced-station measurements.

Multiple frequency measurements entail the simultaneous monitoring at one site of the phase and amplitude of Earth-ionosphere waveguide signals from a number of VLF/LF transmitters of differing frequency. The transmitters can often be chosen so that their propagation paths to the observing site are roughly collinear [e.g., Inan and Oh, 1990; Tip et al., 1991]. In this case, localized D region disturbances near these propagation paths may significantly perturb each of the separate waveguide signals. If perturbations appear on N separate VLF/LF waveguide signals, there would be 2N independent measurements (phase and amplitude) available to characterize the unknowns in the system. Another strategy to obtain additional information is to monitor a particular waveguide signal at two closely spaced sites in order to obtain the direction of arrival of the signal scattered by the D region disturbance. Determination of the group delay of the scattered signal then allows the position of the disturbance to be determined [Dowden and Adnart, 1990]. Thus, for example, if three waveguide signals are monitored at two closely spaced (~10 km) sites and the ambient profiles are known through measurement of the ambient waveguide signals, then six independent pieces of data (three amplitudes and three phases) would be available to characterize the disturbed D region electron density profile and horizontal scale.

In summary, a three-dimensional methodology to account for the effect of localized, lower ionosphere electron density disturbances on propagating subionospheric VLF waves has been developed. It utilizes the LWPC software developed by the Naval Ocean Systems Center and the three-dimensional VLF scattering formulation of Poulsen et al. [1990]. In the few cases considered so far, model results were found to be in general agreement with observations of lightning-induced electron precipitation (LEEP) events affecting typical paths. The new model is applicable to a wide range of disturbance types in the lower ionosphere that can be modeled by a vertical density profile and that satisfy the WKB and Born approximations. The use of this new model allows the investigation of the physics underlying the connections between the Earth-ionosphere waveguide characteristics and the individual waveguide modes that comprise the propagating wave. Further application of the new propagation model to the interpretation of experimental data can potentially lead to a better understanding of the physics behind the type of disturbances thought to be created by LEEP events.

Acknowledgments. This research was sponsored by the Office of Naval Research under grant N00014-87-K-0259 to Stanford University. We greatly appreciate the interest in and support of this program by Gra- cen R. Joiner of ONR. The MODEP/LDR program used to determine the mode refractive indices and the LWPC software were provided to us by J. A. Ferguson and F. P. Snyder of the Naval Ocean Systems Center (NOSC). We appreciate their support of this effort and our consultations with them on the use of the NOSC programs. We also appreciate discussions with our colleagues in the STAR Laboratory. The manuscript was prepared by Zheng Xu. The Editor thanks K. Papadopoulos and D. D. Seatman for their assistance in evaluating this paper.

References


Forbes, J. M., The lower thermosphere coupling study of CEDAR and WITS programs, an attempt to better understand the “ignorosphere”, EOS Trans. AGU, 70, 905, 1989.


U. S. Inan and T. F. Bell, STAR Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305.

W. L. Poulson, Jet Propulsion Laboratory, Mail Stop 300-319, 4800 Oak Grove Drive, Pasadena, CA 91109.

(Received February 14, 1992; revised June 22, 1992; accepted June 24, 1992.)