

Conditions Leading to Anomalously Early Skywave

Peter B. Morris
Northrop Grumman Information Technology
Defense Enterprise Solutions
55 Walkers Brook Drive
Reading, MA 01867

Presented to

**The International Loran Association
32nd Annual Convention and Technical Symposium
Boulder, Colorado
03 – 07 November 2003**

**The Work Presented in This Paper
Was Supported by**

**Volpe National Transportation Systems Center
55 Broadway
Kendall Square
Cambridge, MA 02142**

**Under Contract Number
DTRS57-98-D-00050
Task Order CA9301**

Conditions Leading to Anomalously Early Skywave

Abstract

The skywave component of the Loran signal can adversely affect identification of the correct cycle if it appears early enough and strong enough with respect to the groundwave pulse. The ideal conditions for early skywave are long paths (>~1000 km) and relatively low ionospheric effective reflection heights. Normally, the effective reflection height for the skywave is lowest (maximum ionization) for propagation paths with low average solar zenith angles, such as might be experienced on a summer day near local noon. However, anomalous external occurrences, such as solar proton events (SPEs), often result in much lower ionospheric effective reflection heights. Normally, these events only affect the skywave for transmitter-to-receiver paths whose mid-point lies within the polar/auroral zone (>~ 60° geomagnetic latitude). However, geomagnetic storms accompanying the SPE often cause the auroral zone boundary to move to much lower geomagnetic latitudes. This phenomenon exposes the lower latitudes to the excess proton-induced ionization and anomalously low ionospheric reflection heights. We have looked at SPE events that occurred in past solar maximum years and accompanying geomagnetic storm data that is well-correlated with the equatorward auroral boundary. With this data we estimate the frequency of occurrence of conditions leading to anomalously early skywave. The implications of these results for Loran integrity and availability are discussed. Warning systems are discussed to address the integrity issue and possible receiver skywave identification and exclusion techniques are considered to address availability issues.

I. Introduction

The Loran signal is an electromagnetic field whose time variation depends partly on the transmission pattern (signal format) and on the distance from the signal source. The spatial variation depends strongly on the signal *path*, i.e., the electromagnetic environment through which the signal ray, or ray *bundle*, travels from source to receiver antenna. The *groundwave* is that component of the radiated signal that is observed slightly above the earth's conducting surface and is polarized in approximately the same direction (radial) as the transmitted signal. Stated another way, the groundwave is that component of a signal radiated from a vertical dipole on the surface of a conducting sphere that would be present at/near the entire surface of the sphere in the absence of an ionosphere. Because the groundwave results from a scattering process, it is effective only at lower frequencies, e.g., 100 kHz.

As shown in Figure 1, other components of the Loran signal that may be received at/near or above the earth's surface include

- Direct wave – propagated directly from the source via a refractive path through the air

- Trapped wave (not shown) – depending on the earth's local conductivity profile with depth, a scattered wave traveling just below the surface
- Skywave - the source signal scattered through/reflected from the ionosphere on its transit to the receiver

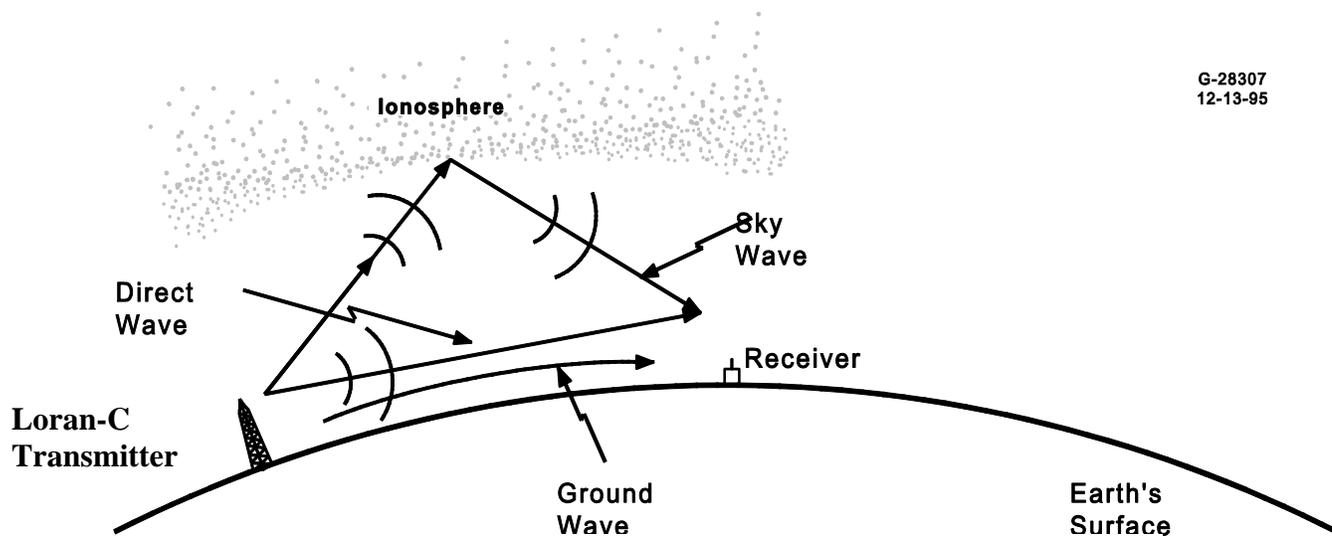


Figure 1-1 Components of the Transmitted Signal

The direct wave exists only at very close ranges to the transmitting antenna and the trapped wave is also usually quite range-limited. At ranges of conventional Loran use, only two components of the signal are normally available: groundwave and skywave. It is important to note that the polarization of the two waves will be the same (magnetic field normal to the plane of propagation) except that the skywave electric field vector is tipped by an angle approximately equal to its launch angle at the source.

At the user receiver/antenna, the skywave component of the Loran pulse is normally well-separated in time from the groundwave. However, for ionospheres with lower effective reflection heights and longer ranges, the delay of the longer-path skywave relative to the groundwave becomes smaller. In such cases, the groundwave is very difficult to resolve from the received superposition of the two signal components. As a result, the navigational receiver may select the wrong cycle (cycle other than the third cycle) to make its TOA or TD measurement. This possibility leads to a *failure of cycle integrity* and is one of the largest error sources in the use of Loran. Hence it is important to know when and where these anomalous conditions (combination of effective ionospheric reflection height and groundwave range that leads to cycle integrity failure) exist for Loran users.

In this paper we will show how these conditions can be identified by first characterizing ionospheric reflection heights (Section 2). In Section 3 we analyze both

the groundwave and skywave components and the parameters they depend on. With this information, we consider the behavior of skywave phase delay and skywave-to-groundwave amplitude ratios under normal (day/night) and anomalous (disturbed) ionospheres (Section 4). In this section we also consider the sources of the disturbances themselves to determine how frequently and where in the Loran coverage area these effects are likely to be experienced. The paper is summarized in Section 5, along with conclusions and recommendations.

II. Effective Ionospheric Height Profiles

2.1 Signal Component Geometry

Because the useful groundwave range of Loran signals is on the order of $0.1R_E$ and the ionospheric height is on the order of $0.01R_E$, we normally consider a spherical earth geometry when calculating distances and angles. Figure 2-1 illustrates the geometry for one-half of a Loran signal path, showing both the groundwave and skywave components. The ionosphere is shown as a concentric shell of radius $R_E + h$ from the center of the earth.

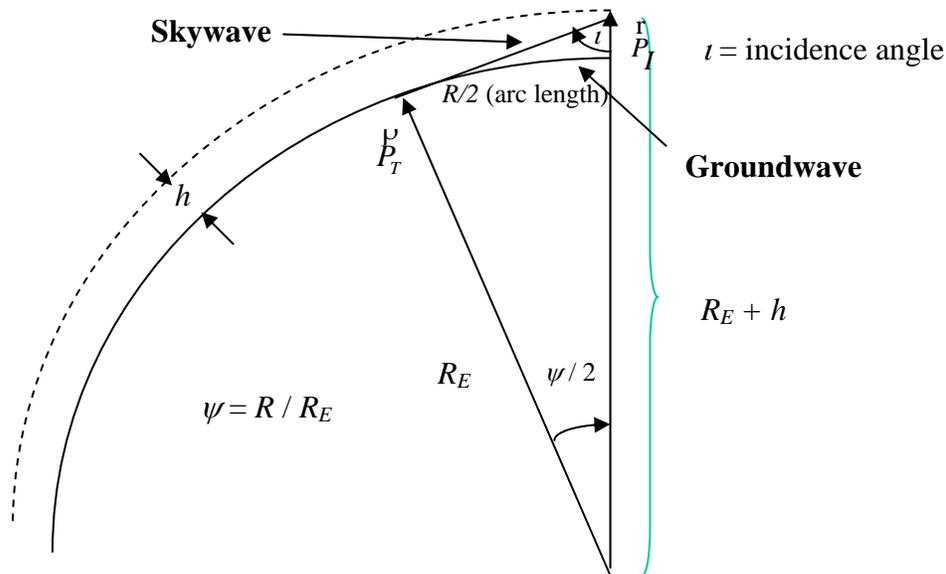


Figure 2-1 Groundwave and Skywave Geometry

The length of the groundwave path is just $R = R_E \psi$ and the skywave path vector from launch point to intersection with the ionosphere is $\frac{1}{P_T} - \frac{1}{P_T}$. Taking the magnitude of this vector and multiplying by 2 yields the skywave path length

$$\begin{aligned}
s &= 2 \left| \vec{P}_I - \vec{P}_T \right| = 2 \left[R_E^2 + (R_E + h)^2 - 2R_E(R_E + h) \cos(\psi / 2) \right]^{1/2} \\
&= 2 \left[h^2 + 4R_E(R_E + h) \sin^2(\psi / 4) \right]^{1/2}
\end{aligned} \tag{2-1}$$

where the last step is obtained with the use of the cosine double angle formula. The incidence angle at the ionosphere is expressed in terms of the earth's radius, ionospheric height, and range angle, ψ , using the Law of Sines as

$$\sin \iota = \frac{2R_E \sin(\psi / 2)}{s} \tag{2-2}$$

where s is given by Eq. (1). In terms of the previously computed quantities, the launch angle of the skywave relative to the local tangent plane at the transmitting station is just

$$\eta = \pi / 2 - \psi / 2 - \iota \tag{2-3}$$

The maximum range for a single-hop skywave occurs when the launch angle, η , is zero. At that range, the triangle shown in Fig. 2-1 is a right triangle, so that $\iota = \pi/2 - \psi/2$ and

$$\cos(\psi / 2) = \frac{R_E}{R_E + h} \tag{2-4}$$

From Eqs. (1) and (2), the maximum single-hop skywave distance is

$$s_{\max} = 2R_E \tan(\psi / 2) = 2\sqrt{h(h + 2R_E)} \tag{2-5}$$

2.2 Effective Ionospheric Reflection Height

The effective ionospheric height is not constant over the entire signal path as shown in Fig.2-1. By this we mean that, if a hypothetical receiver/antenna were placed at any position between the signal source and the point at groundwave range R , the effective ionospheric height at the corresponding path mid-point (with range between 0 and $R/2$) will vary. In this sense, the effective ionospheric height depends on

- local illumination condition as measured by solar zenith angle (which accounts for diurnal and seasonal effects)
- incidence angle of the propagation vector at the ionosphere (at lower angles the wave penetrates farther and at grazing angles, the penetration is less)
- magnetospheric/ionospheric activity condition, which may depend on both space and time (see Section 4)

2.2.1 Dependence on Incidence Angle

Let us first focus on the second bullet above and consider a height profile (over path length) that yields greater heights for low incidence angles (short path lengths/ranges) and smaller heights for higher incidence angles (greater ranges), viz.

$$h = \alpha \cos \iota + \beta \sin \iota \quad (2-6)$$

where ι is the incidence angle and α and β are constants to be determined. For a given $R = \psi/R_E$, incidence angle ι is a function of h through s (see Eqs. (1) and (2)). Thus, with given α , β and a series of R -values, Eq. 6 is an implicit function for h . The first term is a measure of the penetration of the electric field component parallel to the ionospheric interface into the plasma of the ionosphere. Boundary conditions dictate that the tangential component of the electric field (given by $|E| \cos \iota$) is preserved across the ionospheric interface; however, it exponentially decays as one proceeds vertically upward through the ionosphere. The wave is refracted as it passes through the ionosphere so that the propagation vector (\hat{k}) turns away from the normal and eventually $\hat{k} \cdot \hat{g} = 0$. At that point the wave is re-radiated (with a phase shift because the plasma is partially conducting) and returns to earth. The larger the initial value of $|E| \cos \iota$, the farther the vertically propagated wave component penetrates into the ionosphere before being scattered downward (reflected). The $\sin \iota$ term addresses the normal component of the signal electric field through the boundary condition that governs the normal component of the electric displacement.

Although the greatest ionospheric height would be expected at range $R = 0$ (where $i = 0$), Loran transmitting antennas have an effective dipole/monopole pattern, with a null vertically upward. There is actually a small field directly above the antenna, due to the departure from an ideal monopole pattern of a real Loran antenna, but the essential idea is that as the launch angle from the antenna decreases from 90° , the beam pattern implies greater energy in the beam, while the ionosphere becomes more distant and the incidence angle becomes larger. The net result is that the effective ionosphere height profile with groundwave range has a peak and gradually decreases with range for larger values of range as will be shown later.

2.2.2 Dependence on Solar Zenith Angle

Returning to the first bullet above, the effective ionospheric reflection height depends on the angle of illumination by the sun, since the altitude of the D-region is generally controlled by photoionization. The relationship is given by¹

$$h(\chi) = h(0) + H \ln(Ch(h, \chi)) \quad (2-7)$$

where $h(0)$ is the effective ionospheric reflection height when the sun is directly overhead, H is the scale height (~ 5 km with a pronounced semiannual variation), χ is the

solar zenith angle, and Ch is the Chapman Function^{*}. The scale height used in the calculation of the Chapman Function is much larger (~125 km) than that for calculating ionospheric height (Eq. (7)). Note also that, since h appears in the argument of the Chapman Function, Eq. (7) is an implicit function of h . Thus, given the solar zenith angle, χ , $h(0)$, and the two scale heights referred to above, one can iterate Eq. (7) until the solution converges.

There are several means for obtaining α and β (see Eq. 6). One is a series of measurements parameterized by groundwave range that involve the effective ionospheric reflection height. For example, measurements of the skywave delay relative to the groundwave depend (to first order) on skywave path length (Eq. (1)), ionospheric height, and groundwave range. For multiple measurements, one can use a fit procedure (see Section 4). One can also simply use vertical incidence sounding measurements from the literature³ to obtain a maximum height and assume that the ionospheric height at maximum 1-hop skywave range is a specific amount lower (10 km – night and 12 km – day)⁴. At least two measurements or assumed values are required for a given ionospheric condition, e.g., day or night. Once α and β are known, the implicit expression for h (Eq. (6)) is iteratively solved for various groundwave range values.

Fits to skywave delay data (see Section 4) and other measurements yield the following effective ionospheric heights associated with a given diurnal/seasonal condition:

Night:	Maximum – 99 km ; Minimum – 89 km
Winter Day:	Maximum – 81 km ; Minimum – 69 km
Summer Day:	Maximum – 72 km ; Minimum – 60 km

It should be recognized that “Winter Day” and “Summer Day” in the list above are labels only and do not necessarily apply to the indicated season. Winter Day actually refers to those daytime periods having a large average solar zenith angle. This would normally apply to daytime periods in the winter, but could also refer to times near path sunrise/sunset during the other three seasons. Similarly, Summer Day refers to path/times with small average solar zenith angle. Because a one-hop skywave intercepts the ionosphere at the path mid-point (between the source and receiver – see Fig 2-1), the solar zenith angle and/or diurnal condition for a given path is always referenced to the path mid-point.

The above discussion refers to effective ionospheric heights under “normal”, or solar-quiet conditions. So-called “anomalous” events generally occur during solar-active conditions. Although it will be discussed more extensively in Section 4, we will mention here that these anomalous events arise from active sources on the sun with enhanced x-ray/EUV (extreme ultra-violet) or high-energy particle (mostly proton) emissions. On the sunlit side of the earth, x-ray events cause what is known as sudden ionospheric disturbances (SIDs) that result in a depression of the effective ionospheric height. The magnitude of the SID depends on solar zenith angle and usually persists for 30-45 minutes. The high-energy particle events depress the ionosphere both day and night although usually more so during the day because of the more favorable positioning of the

^{*} Approximately $\sec \chi$ when $\chi \ll 90^\circ$. See Ref. 2 for the complete definition of the Chapman Function

geomagnetic field lines. Poleward of the auroral zone, these events persist much longer than x-ray events because the protons continue to pour in from the sun (though perhaps below designated event thresholds) for an extended period of time. A strong event (known as a polar cap disturbance or *PCD*) can depress the *D*-region of the ionosphere to as low as 50 km⁴. Since the ionospheric profile developed here depends on incidence angle and thus varies as a hypothetical receiver is moved from the transmitter to the path endpoint, we assume here that 50 km is the *average* ionospheric height along the path. Thus, we append to the above list of effective ionospheric heights associated with diurnal/season conditions the following entry for severe PCDs:

PCD: Maximum – 56 km ; Minimum – 44 km.

III. Comparison of Received Groundwave and Skywave Signals

3.1 Groundwave

For ranges out to a few hundred kilometers, the complex groundwave signal is given (in V/m) as⁵

$$E_z(d, t) \cong 9.5 \times 10^3 \frac{W_N \sqrt{P_r}}{d} E_0(t) ; t \geq \frac{d}{v_0} + \tau \quad (3-1)$$

where P_r is the radiated power (Watts), d is the groundwave distance in meters, $E_0(t)$ is the transmitted Loran pulse (observed after a time equal to the propagation delay)

$$E_0(t) = Ai(t - \tau)^2 e^{-\alpha(t-\tau)} e^{-2\pi i \beta t} ; t \geq \frac{d}{v_0} + \tau$$

where A is a measure of the pulse magnitude, $\alpha = 1/32.5 \mu\text{sec}^{-1}$, $\beta = 1/10 \mu\text{sec}^{-1}$, τ is the envelope-to-cycle delay (ECD), and

$$W_N ; 1 - i(\pi p)^{1/2} - 2p + \frac{1}{2}i\pi^{1/2} - \frac{1}{2}g^3 p^{3/2} + \frac{4}{3}(1 - g^3)p^2 - \frac{3}{2}g^3 p^{5/2} - \frac{1}{15}(8 - 16g^3 + 7g^6)$$

where

$$p = \frac{\epsilon_0 d \omega^2}{2v_0 \sigma}$$

is Sommerfeld's numerical distance, ϵ_0 is the permittivity of free space, ω is the radian frequency, $v_0 = c/n_a$ (c = speed of light *in vacuo*, n_a = index of refraction of air = 1.000338), and σ is the ground conductivity in mho/m,

$$g = e^{-\frac{3\pi}{4}} \sqrt{\frac{\sigma}{\epsilon_0 \omega}} (ka')^{-1/2}$$

where k is the wave number in km⁻¹ and a' is the effective earth radius = $(4/3)R_E$, where R_E is the earth radius in km.

3.1.1 More General Groundwave Amplitude Calculation

Since Eq. (3-1) is complex, it provides the phase (relative to the harmonic variation, $e^{i\omega t}$ not shown in Eq. (3-1)) as well as the amplitude. For distances greater than a few hundred kilometers, however, other forms must be used⁵. For calculations used in this paper, digitized forms of published curves (amplitude vs. distance for various ground conductivities) were used (see Reference 2). These results are applicable to a much wider spectrum of range values (see Fig. 3-1). The range domain of validity for Eq. (3-1) is shown in Fig. 3-2 for a homogeneous path having a ground conductivity value of 10^{-3} mho/m.

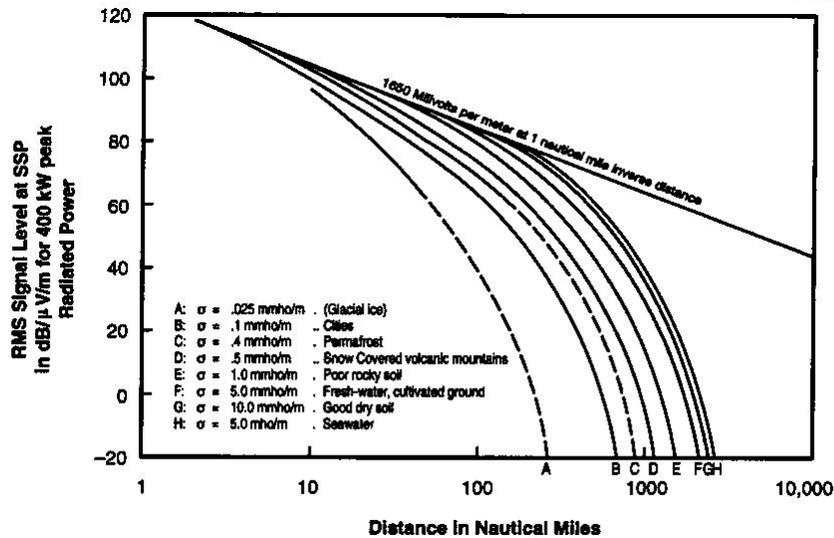


Figure 3-1 Groundwave Amplitude as a Function of Range for 8 Conductivity Levels

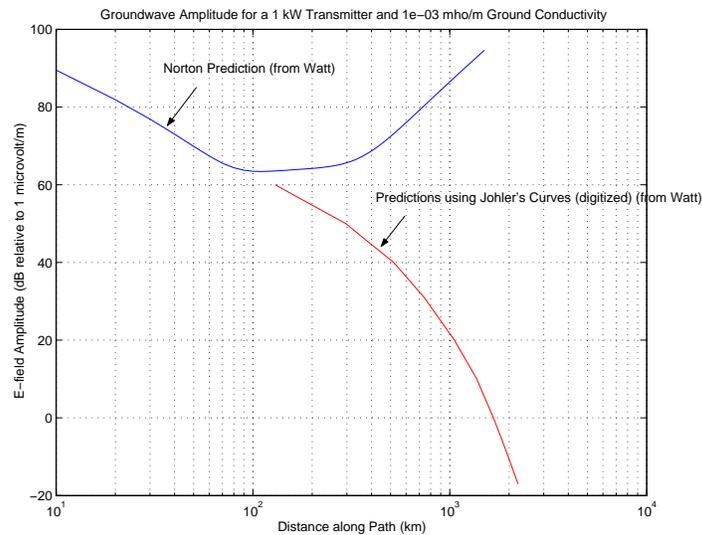


Figure 3-2 Groundwave Amplitude for a 1 kW Transmitter and 10^{-3} mho/m Ground Conductivity Path

The curves in Figs. 3-1 and 3-2, of course, are applicable only to a homogeneous path, i.e., one in which the conductivity is the same throughout the entire path. For inhomogeneous paths, i.e., those with mixed conductivity (including nearly all real paths), the problem is more complex because at each conductivity boundary, there is a change in wave tilt, energy lost to the ground, and phase and group velocities at the ground interface. An approximate technique known as *Millington's Method*⁶ has been developed that is reasonably accurate and affords relatively rapid computation times. The method ensures that the reciprocity condition is satisfied by calculating both forward and backward along the path and taking the average of the two calculations. For one direction (say the forward direction), the basic idea is that the field at the receiver is first calculated assuming that the conductivity at the receiver is the same along the entire path (Assumption 1). At the first conductivity interface (receiver conductivity meeting a new conductivity value) encountered in moving backward along the path towards the transmitter, the difference between Assumption 1 and Assumption 2 is calculated. Assumption 2 is that the path up to the first conductivity interface is homogeneous with conductivity equal to that of the new value at the interface. The initial calculation of the field at the receiver is corrected by this difference and the same procedure is repeated at each conductivity interface. Figure 3-3 is a sample calculation using Millington's Method.

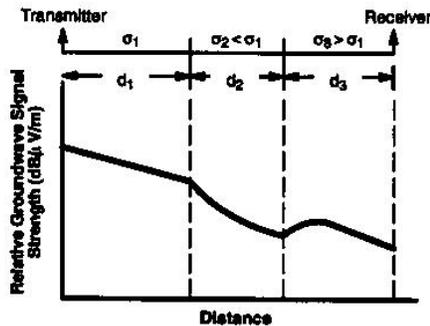


Figure 3-3 Relative Groundwave Signal Strength over a Path with 3 Conductivity Levels using Millington's Method

3.1.2 Groundwave Phase Calculation

For groundwave ranges greater than about 100 km, the phase calculation is also complex and curves are available from many of the sources cited above. Figure 3-4 is a plot of 100 kHz signal phase vs. distance. Since the signal frequency is fixed, the phase is proportional to time, e.g., a phase shift of $\pi/2$ corresponds to 2.5 μsec . Note that the range axis in Fig. 3-4 is in units of statute miles.

Millington's Method can also be used with signal phase in the same way as amplitude. Figures 3-5 and 3-6 are examples of groundwave signal amplitude and phase calculations along actual signal paths. In these two figures, conductivities of successive path segments (about 30 km in length) are indicated by circles (values in mhos/m are

given by the ordinate scaled by 10^4 for amplitude and 10^3 for phase). The forward path calculations are shown in red, the reversed path calculations in blue, and the (logarithmically) averaged calculations, corresponding to Millington's Method are shown in green. Note that the differences between the forward and backward calculations are

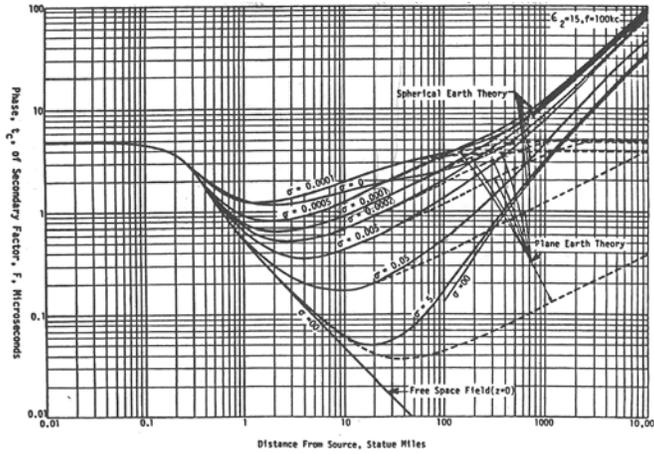
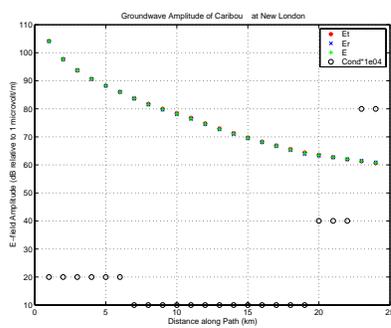
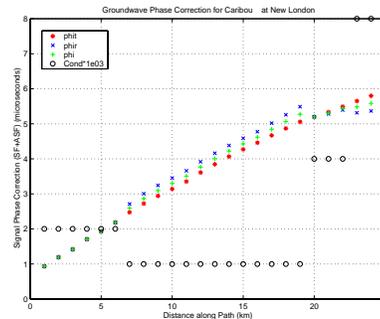


Figure 3-4 Groundwave Signal Phase (microseconds) vs. Groundwave Range (statute miles) for Various Conductivity Levels (mhos/m)

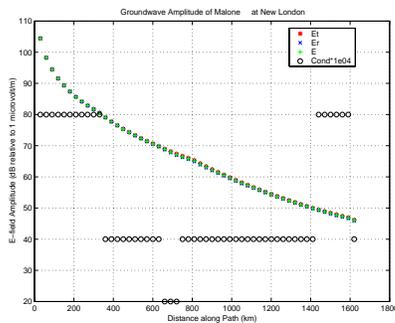


a) Amplitude

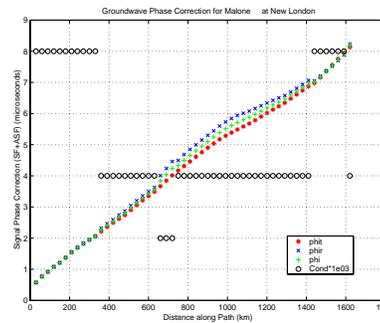


b) Phase

Figure 3-5 Predicted Signal Behavior along the Caribou, ME to New London, CT Path



a) Amplitude



b) Phase

Figure 3-6 Predicted Signal Behavior along the Malone, FL to New London, CT Path

greater for phase than for amplitude. This is because the phase plots shown in Figs. 3-5 (b) and 3.6 (b) are actually corrections to the cumulative phase kr , where k is the wave number for the 100 kHz signal and r is the range, whereas Figs. 3.5 (a) and 3.6 (a) show the full signal amplitude.

3.2 Skywave

The single-hop skywave signal sensed by a vertical E-field antenna (in mV/m) in the radiation field is given by⁷

$$E_z(s,t) = 300\sqrt{P_T} \cos^2 \alpha \, {}_{\parallel}R_p D F_t F_r E_0(t) / s \quad ; \quad t \geq \frac{s}{c} + \tau \quad (3-2)$$

where, again, P_T is the transmitting station power in kW, α is the signal launch angle, ${}_{\parallel}R_{\parallel}$ is the parallel-to-parallel component of the reflection coefficient matrix, D is the ionospheric focusing factor, F_t is the transmitting antenna factor, F_r is the receiving antenna factor, $E_0(t)$ is the Loran pulse defined above, and s is the skywave path length. This form presumes that both the transmitting and receiving antennas have a beam pattern that varies as $\cos \alpha$. The parallel-to-parallel component of R is given by

$${}_{\parallel}R_{\parallel} = \frac{\mu^2 C - \sqrt{\mu^2 - S^2}}{\mu^2 C + \sqrt{\mu^2 - S^2}} \quad (3-3)$$

where $C = \cos \iota$, $S = \sin \iota$, and

$$\mu^2 = 1 - i \frac{\omega_r}{\omega} \quad ; \quad \omega_r = \frac{\omega_0^2}{\nu}$$

where ω_0 is the plasma frequency, ν is the collision frequency, ι is the incidence angle as defined earlier, and ω_L is the electron gyrofrequency. The parameter ω_r is sometimes called the ionospheric conductivity (scaled by the free space permittivity, ϵ_0) and μ is the complex index of refraction.

3.2.1 Ionospheric Reflection Coefficients

The reflection coefficient given by Eq. (3-3) applies to the case where both the incident and scattered (reflected) signal is linearly polarized with their electric field vectors, \vec{E} , lying in the plane of incidence, i.e., the plane containing both the propagation vector, \vec{k} and the normal to the ionosphere. Although the transition represented in Eq. (3-3) is dominant, there are other reflection matrix components that govern the transition between incident signal fields perpendicular and parallel to the plane of incidence and the two possible kinds of reflected fields. The expression, Eq. (3-3), also does not assume that the geomagnetic field permeates the ionosphere. Reflection coefficients for the more general case of a non-zero geomagnetic field are given in Appendix B.

Because the reflection coefficients depend on ionospheric parameters and processes that are not well understood (e.g., multiple species are involved in photodissociation), reflection coefficient magnitudes have been measured for several

diurnal/seasonal conditions. Unfortunately, these measurements are not available under PCD conditions. As a consequence, we extrapolate the reflection coefficients assuming that the logarithm of the reflection coefficient is roughly proportional to the effective ionospheric reflection height. This means that if the height difference for two different diurnal/seasonal conditions is the same, the ratio of reflection heights is also the same. Fig.3-7 shows reflection coefficient values taken from curves that approximate the measurements of reflection coefficients. Note that the reflection coefficients are plotted

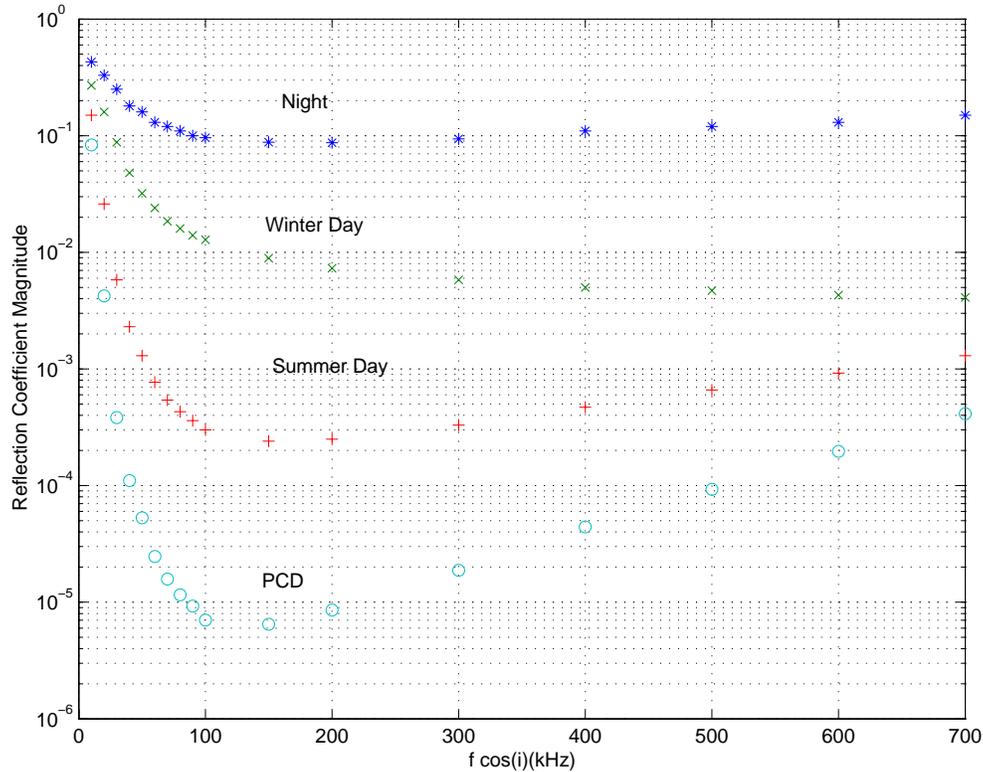


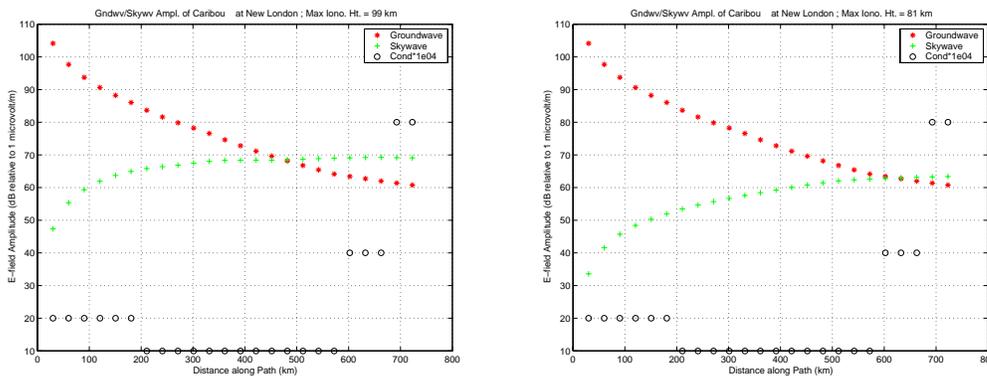
Figure 3-7 Digitized Curves Fit to Measurements of Reflection Coefficients (PCD data is extrapolated)

as a function of $f \cos \iota$, where f is the signal frequency and ι is the incident angle defined earlier. Since we are concerned with $f = 100$ kHz and $\cos \iota \leq 1$, the region of interest in Fig. 3-7 are values of $f \cos \iota$ between 0 and 100. In this region, the reflection coefficient monotonically decreases with increasing $\cos \iota$ (decreasing ι). Since incidence angle increases with increasing groundwave range, the reflection coefficients increase monotonically with increasing range.

3.2.2 Skywave Phase Calculation

As in the case of the groundwave, the skywave phase is generally given as a correction to the cumulative phase ks , where k is the wave number for the 100 kHz signal and s is the skywave path length. The phase corrections come from the complex form of the reflection coefficient, $||R_{||}$. A perfectly conducting reflector would have $||R_{||}|| = 1$ and the phase shift would be π since the incident and reflected waves must cancel at a perfect conductor where no electric field can exist. Skywave phase corrections are sometimes given as corrections to $ks + \pi$. Here k is only approximately the free-space wave number (k_0) because the skywave actually travels through a portion of the atmosphere, both after launch and before reception (especially true for long ranges where the launch angles are nearly grazing). Thus, at the path ends, the wave number is $k = k_0 n$ where n is the index of refraction of air. In most cases, this effect is small and is not usually included in the calculations of skywave phase.

Figure 3-8 illustrates the skywave (green points) amplitude overlaid with the groundwave amplitude (red points). Note that the groundwave amplitude does not change

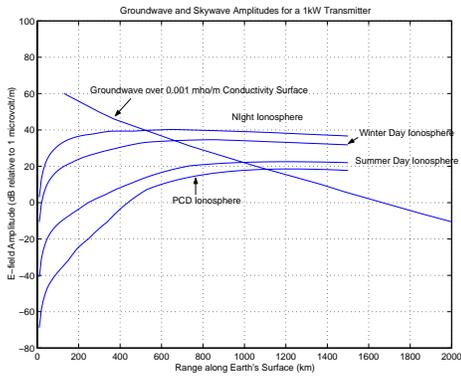


a) Night (Max. Iono. Ht. = 99 km) ; b) Winter Day (Max. Iono. Ht. = 81 km)

Figure 3-8 Skywave and Groundwave Amplitude for the Caribou to New London Path

with diurnal/seasonal condition whereas skywave amplitude is larger for night than day. Although somewhat difficult to see on this figure, the skywave amplitude frequently has a maximum value for longer paths. Initially, the amplitude grows as the fast-growing reflection coefficient offsets the inverse-distance decrease. Eventually, the reflection coefficient increase slows (see Fig. 3-7) and the inverse-distance factor causes the amplitude to decrease. The groundwave range at the skywave maximum value increases with decreasing ionospheric height as a result of the steeper growth (with increasing groundwave range) in the reflection coefficients at the lower ionospheric heights (see Fig. 3-7).

Fig. 3-9 shows the skywave and groundwave amplitudes for four diurnal/seasonal conditions for a ground path having a conductivity of 0.001 mho/m along the entire path and a 1-kW transmitting station. As in the case of the groundwave range at maximum skywave amplitude, the point where the groundwave and skywave amplitudes are equal has a groundwave range that increases with decreasing ionospheric height.



0.01 mho/m ground path
1-kW transmitting station

Figure 3-9 Skywave and Groundwave Amplitudes for Various Diurnal/Seasonal Conditions

As noted above, the phase of the skywave signal is ks (or $ks + \pi$, depending on convention) with corrections given by the phase introduced by the complex reflection coefficient. The phase of the parallel-parallel component of the complex reflection coefficient (modified form of Eq. (3-3)) is plotted in Figure 3-10.

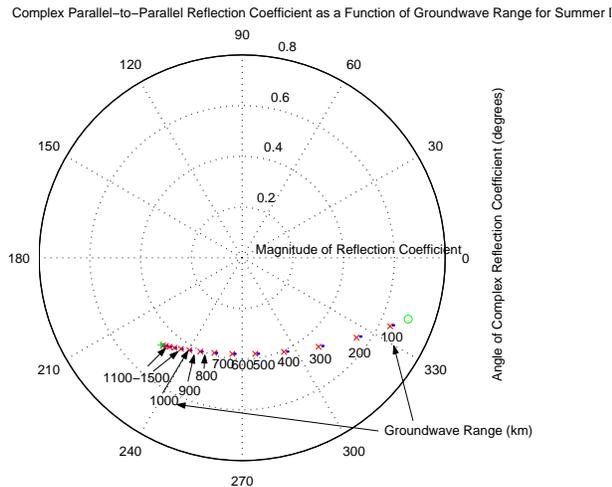


Figure 3-10 Complex Parallel-Parallel Reflection Coefficient for Summer Day Paths of Various Groundwave Ranges

The figure shows that the phase decreases from about 340° to 225° as the range increases from 100 km to 1500 km. The magnitude of the reflection coefficient has a minimum at about 500 km, where the phase is about 280° .

3.2.3 Group Delay

As a final topic in this section, we consider the skywave *group delay*. The group delay is defined as time of arrival of an electromagnetic pulse relative to s/c , where s is the skywave path length and c is the free-space speed of light. Since c is the maximum speed that information can be transmitted, s/c is the minimum transmission time of a pulse and the group delay is the additional time required for a real pulse to propagate through path length s . The group velocity, defined as the rate of change of skywave path length covered by a pulse with time is given by $d\omega/dk$, where ω is the signal frequency and k is the wave number. From this definition, it is shown in Appendix A that, for a signal with bandwidth, $\Delta\omega$, the group delay is given approximately as

$$\Delta t \cong \frac{\Delta(\text{phase})}{\Delta\omega} \quad (3-4)$$

where Δ (phase) is the phase change in the signal with frequency ω compared to a signal with frequency $\omega - \Delta\omega$. In practice, one decomposes a pulse, such as the Loran pulse, into its Fourier components and computes the phase associated with each component. For a series of bandwidths, one then calculates the phase of the upper and lower bound component to determine the group delay. Since $d\omega/dk$ is the second term in a Taylor series expansion of ω about k_0 , the form, Eq. (3-4) is valid only for weakly dispersive media. For the Loran skywave interacting with the ionosphere, the group delay has a magnitude between about 0.1 and 0.3 microsecond.

IV. Characteristics and Sources of Early Skywave

As mentioned in Section I, the skywave is problematic for Loran groundwave navigation and tracking only when its strength is comparable to the groundwave and the skywave delay is short relative to the groundwave. We will first consider the skywave delay relative to the groundwave to determine under what conditions this quantity is minimized.

4.1 Skywave Delay relative to Groundwave

From Eq. (2-1), we may express the difference between the groundwave and skywave signal arrival times as

$$\Delta t_{D,N} = 2[h_{D,N}^2 + 4R_E(R_E + h_{D,N}) \sin^2(\psi/4)]^{1/2} / v_{D,N} - n_a R / v_g \quad (4-1)$$

where the D, N subscripts refer to day and night, respectively (or any other solar zenith angle condition at path midpoint). The ionospheric height at the path mid-point is h and R_E is the earth's radius. The groundwave range, R , is expressed in terms of the subtended path angle, ψ , as $R = R_E \psi$. The ionospheric group velocity is given by v (with the D, N subscripts) and the signal velocity over the ground is given by v_g . The index of refraction of the lower stratosphere is expressed as n_a . The skywave day and night velocities are taken to be equal to the speed of light *in vacuo* although a portion of the path (especially

at grazing launch angles) passes through the lower stratosphere with index of refraction n_a (see Section 3.2.2 above).

Historical measurements of the skywave delay (relative to groundwave) are relatively scarce, especially those that are sampled at different groundwave ranges along the same, or similar, paths. One important source is the so-called “DOTDMA” data⁸ in which groundwave and skywave arrival time differences were extracted from DMA groundwave/skywave correction tables. These tables were purportedly based on measurement data so that the differences in the corrections at a fixed range are equal to the differences in total path arrival time. From Eq. (4-1), we see that measurements of skywave delay at one or more range values can determine ionospheric height (if assumed fixed). Of course, the ionospheric height would be given at the path mid-point where the skywave interacts with the ionosphere. However, if the form of the ionospheric height profile is given as in Eq. (2-6) and one assumes the difference between the maximum and minimum ionospheric height is known for a given illumination condition, then the entire ionospheric height profile along the path can be estimated. This is shown in Figure 4-1

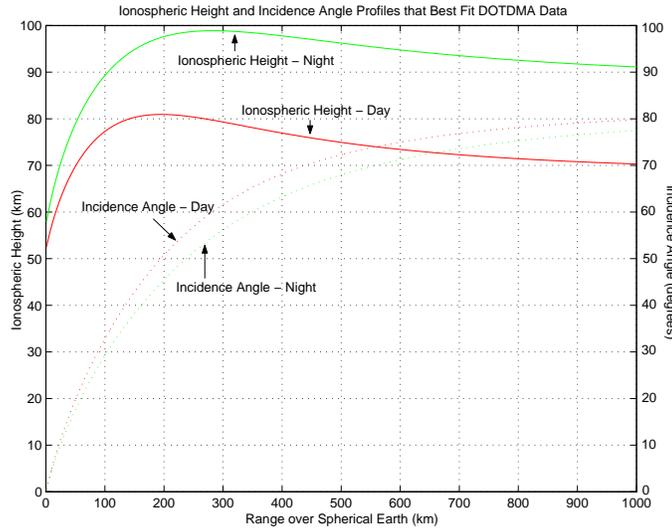


Figure 4-1 Ionospheric Height and Incidence Angle Profiles Based on DOTDMA Data

where the ionosphere height profile has the 2-parameter form of Eq. (2-6), the maximum and minimum height differences are assumed to be those given in Section 2.2.2 (10 km – night and 12 km – day), and the remaining parameter is best-fit to the DOTDMA data. Note carefully that the groundwave “range” shown on the abscissa in Figure 4-1 is the length of a ground path (over a spherical earth) corresponding to a given received skywave. In all cases, the 1-hop skywave interacts with the ionosphere at the path mid-point, so that the actual groundwave ranges of the ionosphere interaction points in Figure 4-1 are one-half the indicated abscissa values (same is true for the incidence angles). Note that the daytime peak ionospheric height occurs when a receiver has a groundwave range of about 200 km (surface range of ionospheric interaction point is 100 km) whereas the nighttime peak is at a groundwave range of about 300 km (150 km for the surface

range of the ionospheric interaction point). Also, the “Day” profile shown in the figure actually corresponds to the “Winter Day” profile* described in Section 2.2.2.

Figure 4-2 illustrates the results of the fit of the ionospheric profile to the skywave delay data by showing the skywave delay calculated from a profile that fit the day and night DOTDMA data at long ranges. Note that the data itself and the predicted skywave

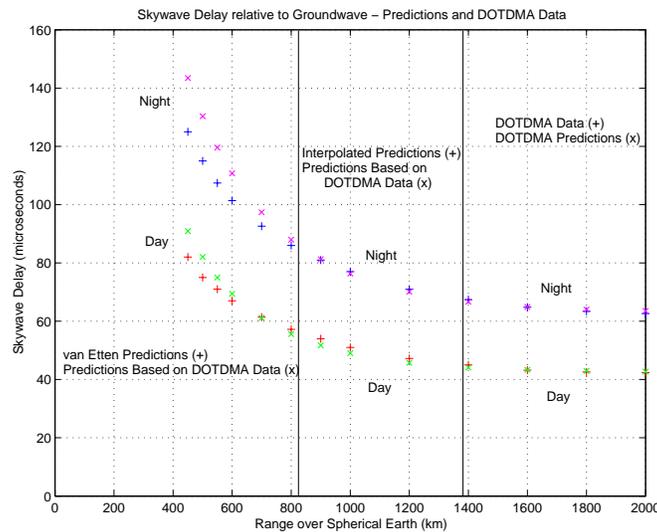


Figure 4-2 Skywave Delay relative to the Groundwave: Predictions and Observations

delay based on the data is given on the right-hand panel of the plot. The left panel of the figure shows data obtained from (1) a model with a fixed ionospheric height and (2) predictions based on the DOTDMA data and the variable ionospheric height profiles described in Section 2.2.1. The wide discrepancy between the two prediction models at relatively short ranges is likely due to errors in both models: Model (1) because a fixed ionospheric height ignores ionospheric penetration dependent on incidence angle and Model (2) because its profile is based on data taken only on long-range paths. The results for the maximum and minimum ionospheric height are given in Section 2.2.2.

Figure 4-2 does show, however, four main features of skywave delay:

- skywave delay becomes smaller at longer groundwave ranges
- the rate of skywave delay decrease slows considerably for longer ranges
- nighttime delay is greater than daytime delay
- daytime delay is never less than 40 μ secs

* The DOTDMA data applies to paths from Loran Stations in the Norwegian Sea chain and thus “day” at these high latitudes equates to “Winter Day” where the solar zenith angles are generally larger than day paths at lower latitudes (see Section 2.2.2)

These features imply that, though small, the skywave delay changes little for groundwave ranges greater than 1000 – 1200 km. The features also indicate that early skywave is most prevalent in the daytime and at long ranges. The fact that daytime skywave delay is smaller than nighttime skywave delay suggests that greater ionization (at a fixed altitude) might lead to earlier skywave. This is true, since greater ionization means a lower effective reflection height and lower reflection height implies less difference between groundwave and skywave path lengths for a fixed groundwave range (see Eq. (4-1)).

Thus, we would expect an ionosphere subject to a polar cap disturbance (PCD) to have a skywave delay less than 40 μsecs. Calculations using the empirical reflection coefficients given in Figure 3-7 yield the values shown in Figure 4-3. This figure shows

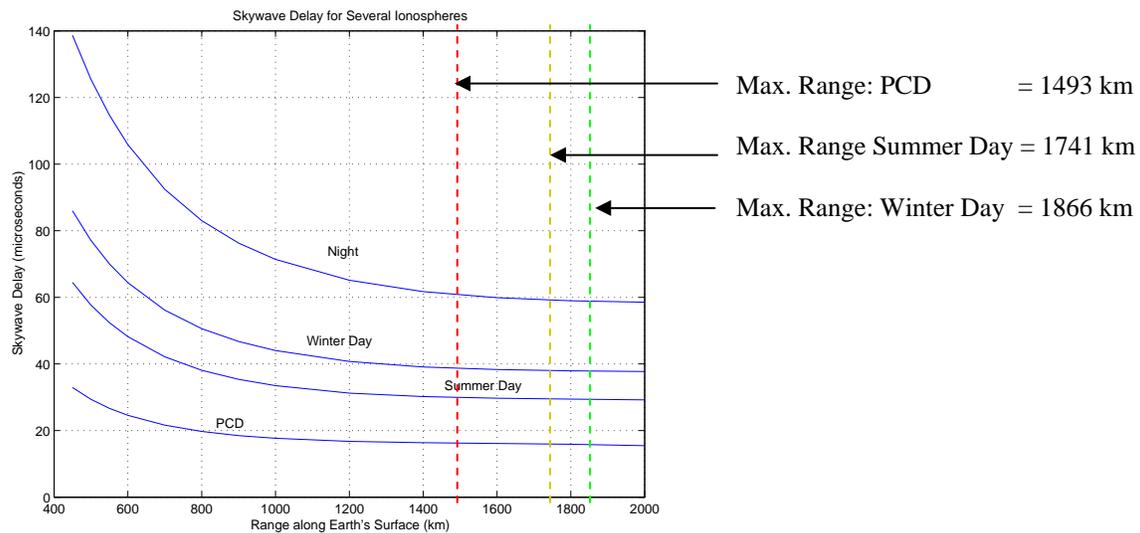


Figure 4-3 Skywave Delay for Night, Winter Day, Summer Day, and PCD Ionospheres

that the predicted skywave delay for the PCD is less than 20 microseconds. A delay this short can lead to an incorrect cycle selection, depending on a number of other factors. In fact, the general rule is that any skywave delay less than 30 μsecs has a significant risk of incorrect cycle selection⁹. Also shown in Figure 4-3 are the maximum possible groundwave ranges for a one-hop skywave for each of the diurnal/seasonal ionosphere conditions (Night is not shown as its max. range is greater than 1000 km)

4.2 Skywave Amplitude relative to Groundwave Amplitude

Another important factor in determining the risk of incorrect cycle selection due to early skywave is the relative amplitude of the groundwave and skywave components of the signal. Clearly if the skywave signal is sufficiently low in amplitude relative to the groundwave, a short skywave delay does not affect the probability of false cycle selection.

A plot of skywave-to-groundwave amplitude ratio for a typical long path (Shoal Cove to New London) is shown in Figure 4-4 for both a summer day and a PCD

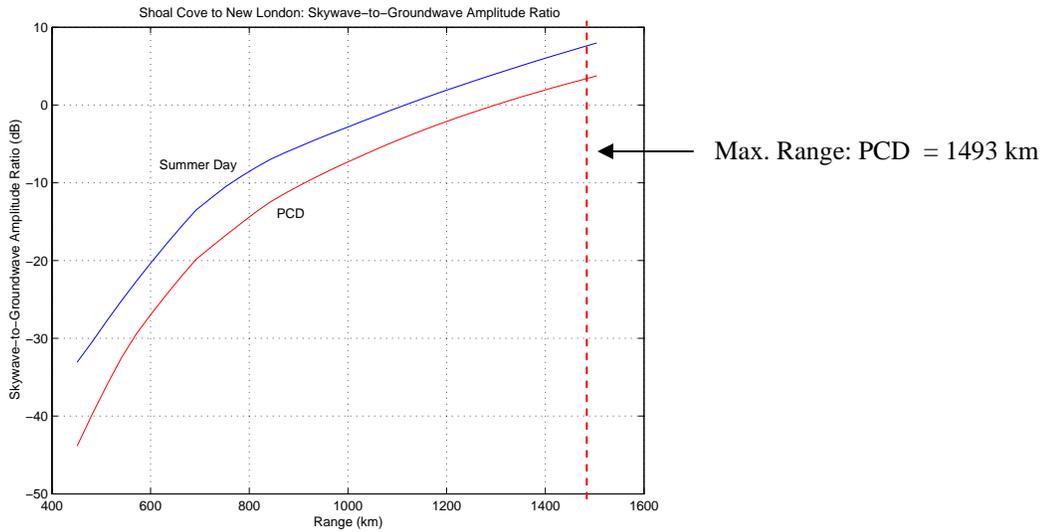
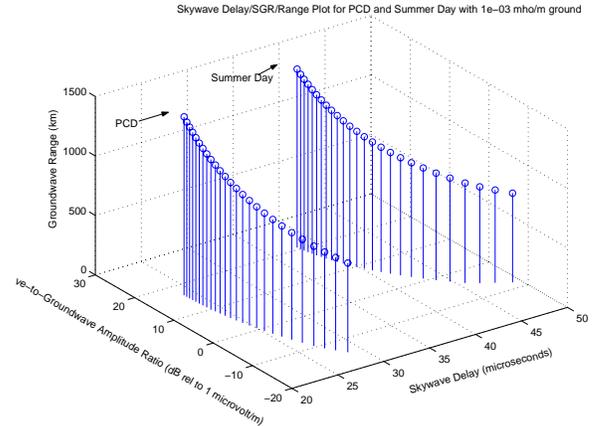
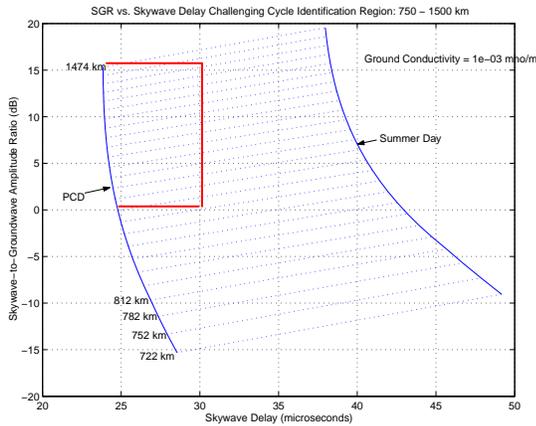


Figure 4-4 Skywave-to-Groundwave Amplitude Ratio: Shoal Cove to New London for PCD and Summer Day Ionospheres

ionosphere. A groundwave range of 400 to 1500 km appears in the abscissa of the plot to be consistent with the “danger” region where the skywave delay under a PCD ionosphere is less than 30 μ secs for groundwave ranges greater than 450 km (see Fig. 4-3).

As indicated in Figure 4-3, the skywave delay curve for the PCD ionosphere is essentially flat for ranges greater than about 900 km. Similarly, the skywave delay curve for Summer Day changes little for ranges greater than about 1100 km. Thus, for these minimum ranges and ionospheric conditions, the effect of the skywave is governed by Figure 4-4. For ranges greater than 1100 km, the Summer Day SGR (skywave-to-groundwave amplitude ratio) is greater than 0 dB, which begins to be problematic from an interference viewpoint. However, from Figure 4-3, the skywave delay at this point is a little over 30 μ secs and decreases about 1 – 2 μ secs over the next 641 km. Thus, although the SGR increases about 7 – 8 dB over the same range interval, the net effect on cycle selection is marginal at best. For PCD ionospheres, however, the skywave delays are short enough to be quite dangerous, but the SGR doesn’t get high enough for the skywave to become a serious interferer until a range of about 1100 km (-5 dB; see Figure 4-4). From about 1250 km to the maximum range of 1493 km (where the SGR becomes positive), the probability of incorrect cycle selection becomes substantial.

Figure 4-5 shows the challenging cycle identification region in a more compact way. The left panel plots SGR vs. skywave delay for a “worst-case” path in which the ground conductivity is 10^{-3} mho/m everywhere on the path. The skywave results from a PCD or summer day ionosphere (or any ionosphere whose ionization profile is between the two specified). In Figure 4-5a, the groundwave range is indicated by points along



a) SGR vs. Skywave Delay (parametric in range)

b) SGR vs. Skywave Delay vs. Range (3-D)

Figure 4-5 Regions of Challenging Cycle Identification

the two curves (dotted lines connect points of constant groundwave range). The red “box” indicates the most challenging cycle identification region. Figure 4-5 b shows the same data with groundwave range as the third dimension.

4.3 Sources of Early Skywave

As noted several times in earlier discussion, the skywave component of the transmitted Loran pulse occurs earliest when the ionization at a fixed altitude is maximum resulting in a minimum effective reflection height. For the *D*-region, which reflects Loran signals, excess ionization is produced primarily as the result of two solar-induced anomalous ionospheric conditions:

- Sudden Ionospheric Disturbances (SIDs)
- Polar Cap Disturbances (PCDs)

SIDs are produced as a result of strong X-ray events on the sun. Clearly they affect only the dayside ionosphere and the magnitude of the effect (ionization produced) increases monotonically with decreasing solar zenith angle. These events are generally less than 30 minutes in duration and reduce the effective reflection height by about 10 – 12 km.

Although SIDs occur somewhat more frequently than PCDs, the latter events have a much greater overall impact. PCDs occur as the result of solar proton events (SPEs) in which excess protons are emitted from the surface of the sun and are conveyed to earth via the solar wind, a net outflow of charged particles that is parallel to the interplanetary magnetic field (IMF). The high-energy protons deposit most of their energy at the lower ionospheric layers, such as the *D*-region. Unlike the high-energy photons that cause SIDs, protons are charged particles and, as such, are usually deflected by the geomagnetic field. However, at high geomagnetic latitudes, the geomagnetic field lines are more parallel to the IMF so that the proton fluxes aimed toward the earth are not deflected. In particular,

the field lines of the earth's auroral and polar zones map outside the dayside plasmopause, and are especially susceptible to incident solar protons. Depending on the geomagnetic latitude, season, and B_z – component of the IMF*, the ionosphere at a given geomagnetic latitude and altitude under PCD conditions may experience excess ionization over a full diurnal period. As noted earlier, PCDs can result in decreases in D –region effective reflection heights down to 50 km (exact height depends on the local proton flux and other factors). Poleward of the lower latitude auroral zone boundary, PCD effects can persist for as long as 5 days¹⁰. These effects combine to make PCDs the primary source for conditions leading to anomalously early skywave for ionospheres poleward of the auroral zone (AZ).

At this point one may think that early skywave conditions caused by PCDs are strictly limited to high geomagnetic latitudes. Unfortunately, there are conditions under which the equatorward boundary of the AZ moves toward the equator. This occurs when magnetic structures on the solar surface change suddenly (e.g., disappearing filaments or coronal mass ejections) and the effect is propagated outward along the IMF field lines at about 1000 – 1500 km/sec. When this shock effect, or “kink” reaches earth's geospace, the entire magnetosphere responds, manifesting itself as a geomagnetic storm, or set of sub-storms. One of the effects of this storm is a compression of the dayside geomagnetic field bounded by the plasmopause. The plasmopause maps to the earth's surface at the equatorward edge of the AZ. Thus, a compression of the geomagnetic field leads to an equatorward displacement of the lower-latitude AZ boundary. If this event coincides with a PCD, the excess protons effectively depress the ionospheric D -region at much lower latitudes as compared to magnetic-quiet conditions.

It is known that the compression of the geomagnetic field is closely associated with a global-scale index known as Dst ¹¹. This index measures ring current energy and, based on a simple model¹², is actually proportional to the difference in magnetic energy of the displaced L - shell and its quiet-time value. A relationship has been derived (Ref.3) between the equatorward AZ boundary and Dst . Using the results of this paper (Ref.3), one can show that

$$\cos^6 \Lambda = 5.3409 \times 10^{-3} - 4.5455 \times 10^{-4} Dst \quad (4-2)$$

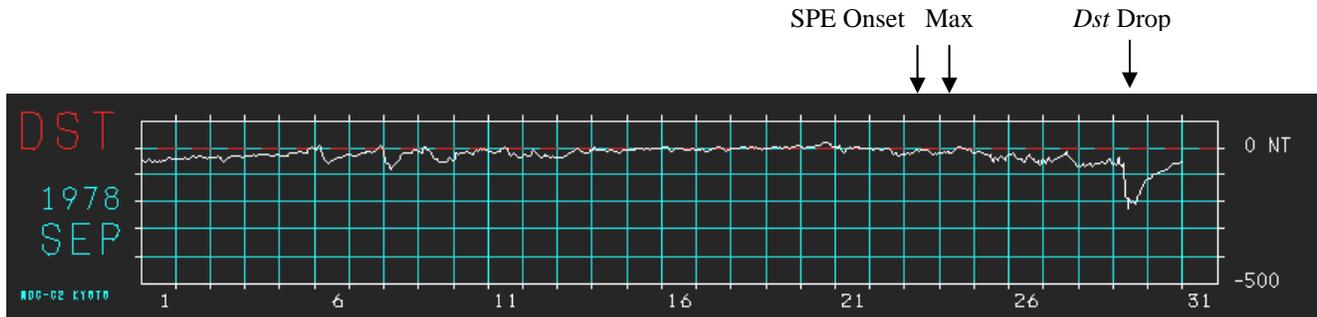
where Λ is the geomagnetic latitude of the equatorward AZ boundary and Dst is in nT. From Eq. (4-2), we see that for large negative values of Dst , the AZ boundary geomagnetic latitude, Λ , becomes smaller, corresponding to a southward movement of this boundary in the northern hemisphere.

To show that this equatorward movement of the AZ boundary occurs in conjunction with a PCD, SPE data was obtained from the Space Environment Laboratory in Boulder, CO for the years 1976 – 2002. Dst data is compiled by the World Data Center (C2) at Kyoto University (web-accessible). For each SPE event, Dst data was analyzed for the corresponding date and time. For nearly every event, a precipitous drop in Dst occurred 2 – 4 days after the SPE onset time. As nearly as one could discern from the hourly Dst data, the large negative Dst value persisted about 1 – 2 hours. Moreover, the

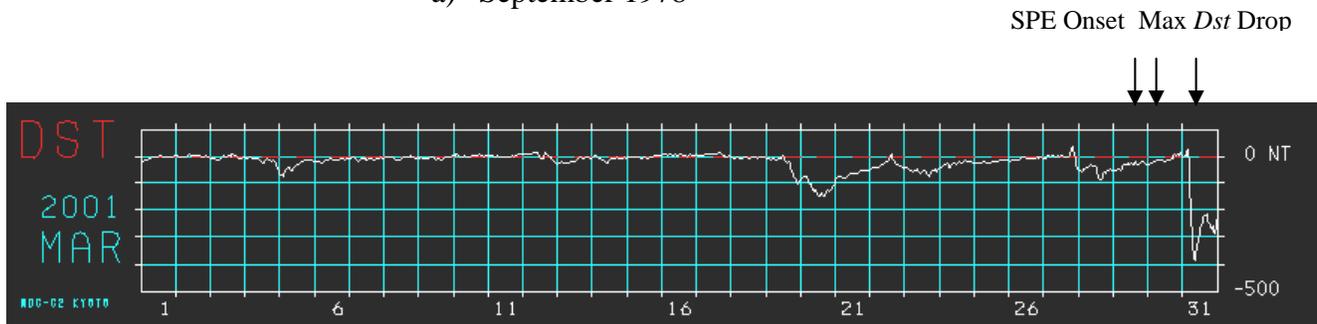
* The IMF is a dipole solar field stretched to an extreme in the vicinity of the solar equator by the solar wind. Since the solar equatorial plane is at a slight angle with respect to the ecliptic, the components of the IMF change with earth's position in its orbit about the sun.

response of the AZ boundary to the excursion in Dst is about -1 hour, i.e., the AZ boundary moves prior to the drop in Dst ¹¹.

Figure 4-6 shows a plot of Dst for September of 1978 and March of 2001. The



a) September 1978



b) March 2001

Figure 4-6 Monthly Plots of the Dst Index for SEP 1978 and MAR 2001

Dst decreases are fairly large: 200 – 400 nT and achieve maximum displacement within 1 – 2 hours. The corresponding SPE onset and maximum given in Table 4-1 are also shown above each panel. Note that the SPE maximum is generally only a few hours after onset

Table 4-1 SPE Data for Two Selected Events

Start (Day/UT)	Maximum	Proton Flux (pfu @ >10 MeV)	Importance (Xray/Opt.)	Min Dst	Date/Time of min Dst (DayMonthHour)	Geomagnetic Latitude of AZ Boundary (degrees)
1978						
Sep 23/1035	Sep 24/0400	2200	X1/3B	-224	29 SEP 1100	46.43
2001						
Mar 29/1635	Mar 30/0610	35	X1/1N	-387	31 MAR 0900	41.21

but the Dst drop can be from about one day to more than five days following SPE onset. The last column in Table 4-1 specifies the geomagnetic latitude of the AZ boundary

corresponding to the *Dst* drop. Since the nominal AZ boundary has a geomagnetic latitude of about 60° , these displacements of the AZ boundary are on the order of 1000 miles. This means that, for a period of 1 – 2 hours, a much larger area, including much of CONUS, is subject to a PCD ionosphere – and therefore the possibility of cycle mis-identification due to early skywave.

From a Loran user's perspective, the effects described above may be summarized as follows:

- For users poleward of the nominal AZ boundary ($\sim 60^\circ$ geomagnetic north), 2 – 22 significant SPE events occur annually (depending on the phase of the solar cycle) that result in PCD conditions for auroral ionosphere. Each event can have a duration of 3 – 5 days depending on the event and local geospace conditions. If the Loran station being used is such that the path mid-point is poleward of the quiet AZ boundary, then early skywave will always be present during PCD conditions. If the Loran station being used is such that the path mid-point is equatorward of the quiet AZ boundary, then early skywave will occur only when the geomagnetic storm occurs following the event (beginning 1-5 days after SPE onset and persisting for about 1 – 2 hours)
- For users equatorward of the nominal AZ boundary ($\sim 60^\circ$ geomagnetic north), an SPE results in PCD conditions for a path if the path mid-point is poleward of the AZ boundary. Near the end of an SPE event (1 – 5 days after onset), a geomagnetic storm may advance the AZ boundary equatorward substantially but only for a period of 1 – 2 hours.

The paths referred to in the above description must be at least 800 km in length and less than 1500 – 1600 km in order that early skywave be present on the received signals.

This is illustrated in Figure 4-7, which shows the affected region for a Loran

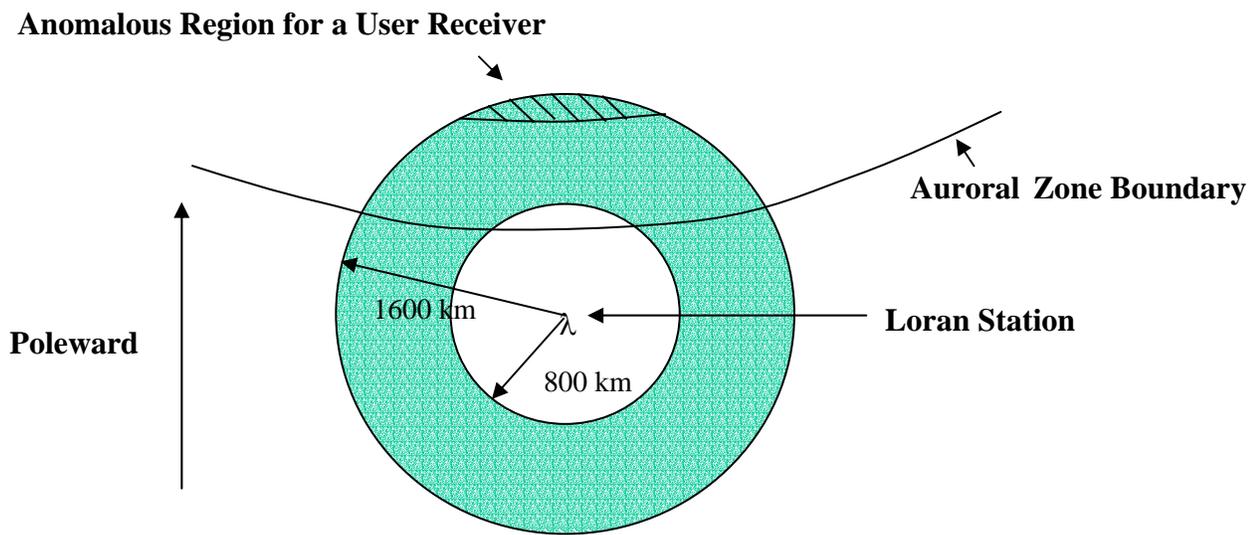


Figure 4-7 User Receiver Locations Subject to Early Skywave during a PCD Event

receiver poleward of the AZ boundary using a Loran station equatorward of the AZ boundary. This figure describes an annulus between 800 and 1600 km, which is conservative (1100 – 1500 would probably be more realistic). Note that the paths from the Loran station to user receivers in the anomalous region have mid-points at locations poleward of the AZ boundary. Recall that the user groundwave range cannot exceed about 1600 km (for a PCD ionosphere) - the limit for a one-hop skywave. Further note that, as the AZ boundary moves down so that it is coincident with the Loran station, the anomalous region would cover nearly half the annulus since the path mid-points would lie poleward of the AZ boundary and the user receiver would be located in the green annulus region.

These ideas are quantified and extended in Figure 4-8, which the fractional area of CONUS subject to the anomalous conditions for early skywave discussed above. Here we assume 24 Loran stations in CONUS, Canada, and Alaska. Given that an SPE has occurred and PCD conditions exist on the *D*-region ionosphere poleward of the AZ boundary, Figure 4-8 shows how the affected anomalous region in CONUS grows as the AZ boundary is displaced equatorward during a geomagnetic storm. The figure shows that, as the AZ boundary geomagnetic latitude reaches 50°, more than 90% of CONUS is affected. Note that somewhat under 10% of CONUS users are affected even when the AZ boundary has not moved at all. This because of paths from northern CONUS users to some Canada and Alaskan stations with mid-points above 60° geomagnetic latitude.

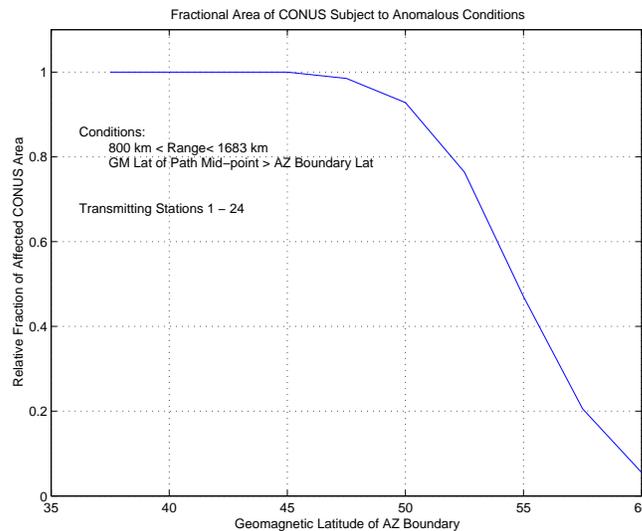


Figure 4-8 Fractional Area of CONUS Subject to Anomalous Conditions for Early Skywave

Although Figure 4-8 provides insight into the spatial dependence of the affected regions on AZ boundary movement, it is of interest to include the temporal dependence so as to gauge the total impact of SPEs on CONUS users. Toward this end, a calculation of the relative occurrence frequency, or probability that a given Loran user, at an arbitrary location in CONUS, and at an arbitrary time within a one-year period at the peak of the

solar cycle will experience anomalous conditions for early skywave. As above, a calculation was made for 24 Loran stations and SPEs for CY 2001. For each SPE of 2001, *Dst* data was compiled and the corresponding maximum AZ boundary excursion was computed as described above. Figure 4-9 is a histogram showing the number of SPEs that yielded AZ boundary excursions in the geomagnetic latitude ranges shown on the

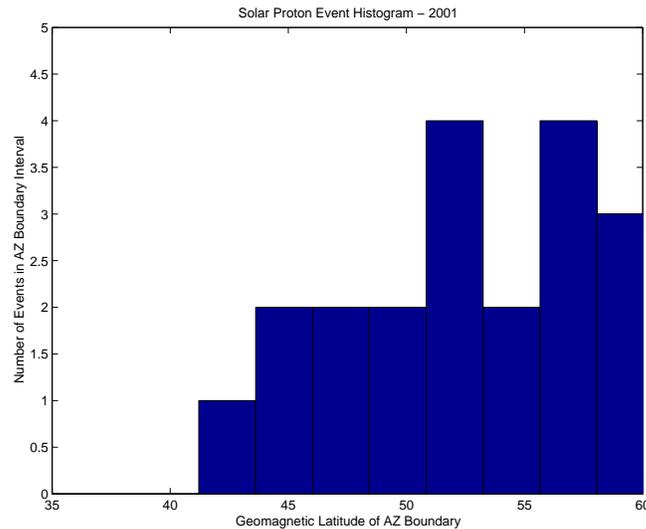


Figure 4-9 SPE/*Dst* Histogram for 2001 showing Minimum Geomagnetic Latitude of AZ Boundary

abscissa. For example, there were four SPEs with associated *Dst* drops that produced maximum AZ boundary excursions between 50° and 52.5° geomagnetic latitude.

In calculating the probability density, it is assumed that each SPE has a single AZ boundary excursion in which the AZ boundary remains at the minimum latitude for a duration of 2 hours. The calculation is made separately for each of the 24 Loran stations by looping over each user receiver location at 1° × 1° intervals. If the user location is within the annulus surrounding the station (see Fig. 4-7) and poleward of the AZ boundary, a one is assigned to that location (otherwise 0). The temporal component is incorporated by multiplying by the fraction of a year the AZ boundary is at its current position or equatorward of it. This is obtained from Fig. 4-9 by multiplying the total number of events with AZ boundary geomagnetic latitudes to the left of the geomagnetic latitude corresponding to the mid-point of the user location-Loran station path time two hours. These products are summed over all 1° × 1° user locations and normalized by the total area of CONUS (in 60 nm × 60 nm units) and hours in a year. The result then gives the probability density (per unit area per unit time) of a CONUS user being subject to an anomalous condition (for early skywave) in 2001 for the given Loran station.

The results of this calculation are shown in Table 4-2, which gives the station name, location, geomagnetic latitude, and probability density (per unit area per unit time).

Table 4-2 Probability that a CONUS Loran User is Subject to Anomalously Early Skywave Conditions in 2001

Station Name	Latitude (degrees)	Longitude (degrees)	Geomagnetic Latitude (degrees)	Probability of Anomalous Condition (2001)
Caribou	46.8075850	-67.9269890	57.9982	0.00090802
Nantucket	41.2533460	-69.9773710	52.4533	0.0008505
Cape Race	46.7756350	-53.1743350	57.3655	0.00035746
Fox Harbor	52.3764590	-55.7077390	63.1078	0.00033691
Williams L	51.9663540	-122.3671400	57.7492	0.00083818
Shoal Cove	55.4391500	-131.2553000	59.4098	0.00039444
George	47.0633600	-119.7441600	53.4814	0.0010272
Port Hardy	50.6082860	-127.3579100	55.5474	0.00073135
Malone	30.9941310	-85.1690980	41.7471	0.001208
Grangeville	30.7258750	-90.8286240	41.091	0.001171
RaymondMl	26.5319840	-97.8332050	36.2894	0.00076422
Jupiter	27.0329240	-80.1146880	38.0371	0.0008012
Carolina B	34.0628360	-77.9128060	45.1383	0.0012285
Havre	48.7440530	-109.9815600	56.6781	0.0015038
Baudette	48.6138740	-94.5549760	58.5157	0.0017092
Boise City	36.5057730	-102.8998600	45.6153	0.0016188
Gillette	44.0031400	-105.6233000	52.6586	0.0016517
Dana	39.8521270	-87.4865520	50.4282	0.0016517
Wildwood	38.9495380	-74.8669770	50.1009	0.00097787
Fallon	39.5518720	-118.8321700	46.3221	0.00089981
Middletown	38.7825310	-122.4955500	44.9606	0.00087105
Searchlight	35.3217510	-114.8046900	42.8124	0.0012203
Las Cruces	32.0717030	-106.8678900	40.7337	0.0013066
Seneca	42.7140880	-76.8259190	53.8158	0.00106

A scan of Table 4-2 shows that the stations with the lowest probabilities are the far northern/northwestern stations in Canada/Alaska, and, to some degree, the southernmost stations. For the far northern stations, the probabilities are lower because of range limitations (800 – 1600 km annulus about the station) to most locations in CONUS. For the southern stations, the lower probability stems more from the rare deep excursions of the AZ boundary into the lower geomagnetic latitudes. The scatter of probabilities for each station is shown in Figure 4-10.

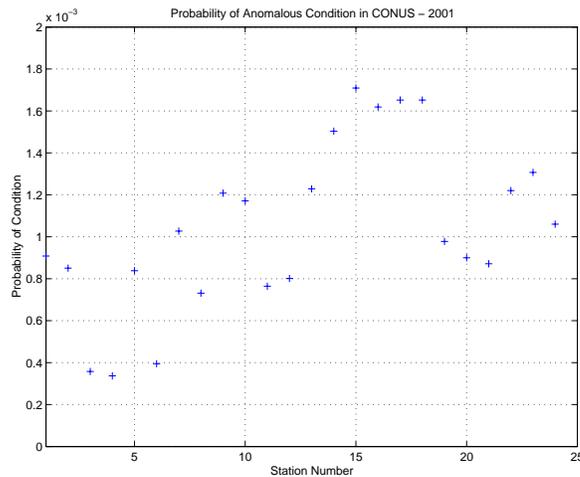


Figure 4-10 Probability of Anomalous Conditions in CONUS vs. Station Number - 2001

Table 4-2 and Figure 4-10 show that the probabilities of anomalous conditions for early skywave vary from about 3×10^{-4} to 1.7×10^{-3} . The average of the probabilities for these 24 Loran stations is about 0.001. One might ask about the probabilities of anomalous conditions on paths from the multiple stations normally available to a given user receiver. The probabilities are lower, as one would expect, but cycle mis-identification on only one path is necessary for a large integrity error.

V. Summary and Conclusions

In this paper we have considered in some detail both the skywave and groundwave components of the Loran signal and how they combine to produce a superposition that makes correct cycle identification difficult. Because the effective ionospheric reflection height is so critical for calculation of the skywave delay, the paper develops a variable ionospheric effective reflection height profile that depends on incidence angle at the ionosphere, solar zenith angle, and the station antenna beam pattern. The paper shows how this profile can be calibrated from skywave delay data and conventional assumptions regarding maximum and minimum ionospheric height differences. We described the received skywave signal's critical dependence on reflection coefficients in terms of amplitude reduction, phase shift, and group delay. Because the skywave delay depends also on the groundwave, important parameters of the groundwave and its calculation were also introduced.

In discussing the seasonal and diurnal dependence of the ionospheric reflection coefficients, we introduce the anomalous ionospheric conditions created by sudden ionospheric disturbances and, especially, polar cap disturbances (PCDs). Effective reflection height profiles were obtained for PCD ionospheres partly from data found in the literature. Because calculations predict very small skywave delays for these ionospheres, skywave-to-groundwave amplitude ratios (SGRs) were computed to see if the interference was substantial enough to result in incorrect cycle selection. A set of points in skywave delay/SGR space was identified that would most likely lead to cycle identification problems. These anomalous conditions could only occur in a certain interval of groundwave range (station-to-receiver) when the path mid-point (interaction point with the ionosphere) lies within a PCD ionospheric region.

Temporal conditions were identified for PCDs, which result from solar proton events (SPEs). PCDs affect the ionospheric *D*-region north of the auroral zone (AZ) boundary (in the northern hemisphere) and thus always affect Loran users in Canada and Alaska. SPEs are found to be associated with certain types of geomagnetic storms, whose effects can be measured by the *Dst* index. Based on a recently discovered relationship between the *Dst* index and AZ boundary, one can determine the southward excursion of the AZ boundary for each SPE. These southward excursions of the AZ boundary result in extension of PCD ionosphere to regions throughout CONUS. Based on these data, we can determine the spatial and temporal conditions regarding anomalously early skywave conditions for CONUS users of Loran stations in CONUS, Canada, and Alaska. The results indicate that the probability that a CONUS user in a $1^\circ \times 1^\circ$ cell experiences anomalously early skywave conditions is about 0.001 over a year near the peak of the solar cycle.

It was noted earlier that anomalously early skywave is likely to induce an incorrect cycle selection, an error which is placed in the category of an integrity error. In this paper, we have found that the probability that a CONUS Loran user is subject to anomalously early skywave conditions is fairly high (0.001) compared to typical integrity requirements. For example, an aviation integrity requirement for navigation during non-precision approach is 10^{-7} /hour, four orders of magnitude more restrictive. There are several possible solutions to this problem:

- Shape pulse for earlier rise time
- Receiver techniques
 - identify/remove through prediction and correlation
 - antenna polarization techniques
- Warning systems
 - internal monitors
 - external monitors

Of these possibilities, it appears that warning systems may be the appropriate response to the problem. If a PCD is detected by a monitor and the information rapidly disseminated to the user community, users will know which station paths to deselect. The deselection may result in a reduction of total signal availability, but not the more serious integrity errors. Typical availability requirements are much less restrictive than integrity requirements. Therefore, the recommendation of this paper is that, at least until receiver techniques to identify and remove early skywave become sufficiently robust, internal monitors be used to transmit warnings to users via modulation of the Loran signal itself.

References

- ¹ Rishbeth, H. and O. Garriott, Introduction to Ionospheric Physics, Academic Press, NY, 1969 (p.161)
- ² Morris, P., R. Gupta, P. Creamer, J. Pisano, *Loran-C Availability Models*, TASC TIM 07462-1, supported under Contract DTICG23-89-C-20008, DOT/USCG Omega Navigation System Center, December 1994
- ³ Ratcliffe, J., Physics of the Upper Atmosphere, Academic Press, NY, 1960
- ⁴ Davies, K., *Ionospheric Radio*, IEE Electromagnetic Waves Series 31, Peter Peregrinus, Ltd, 1989
- ⁵ Watt, A.D., VLF Radio Engineering, Pergamon Press, Oxford, 1967
- ⁶ Millington, G., *Ground-Wave Propagation over an Inhomogeneous Smooth Earth*, Proceedings of the Institution of Electrical Engineers, Part III, Radio and Communication Engineering, Vol. 96, 1949
- ⁷ CCIR/ITU, Recommendations and Reports of the CCIR, 1986, Vol. VI, Propagation in Ionized Media, XVIth Plenary Assembly, Dubrovnik, 1986
- ⁸ Last, J., R. Farnworth, M. Searle, *Effect of Skywave Interference on the Coverage of Loran-C*, IEE Proceedings-F, Vol. 139, No. 4, August 1992

- ⁹ Peterson, Benjamin, Early Skywave Effects on TOA and ECD, Presentation to the Lorán Integrity Performance Panel Working Group, 25 AUG 2003
- ¹⁰ Sauer, H., W. Spjeldvik, and K. Steele, *Relationship between Long-term Phase Advances in High-latitude VLF Wave Propagation and Solar Energetic Particle Fluxes*, Radio Science, Vol. 22, No. 3, May-June, 1987
- ¹¹ Yokoyama, N., Y. Kamide, and H. Miyaoka, *The Size of the Auroral Belt during Magnetic Storms*, Ann. Geophysicae 16 (566-573) 1998
- ¹² Dessler, A. and E. Parker, *Hydromagnetic Theory of Geomagnetic Storms*, J. Geophys. Res. **64**, 2239 – 4214, 1959
- ¹³ Wait, J., Electromagnetic Waves in Stratified Media, Int'l. Series of Monographs on Electromagnetic Waves, Vol. 3, Pergamon Press, Oxford, 2nd Edition, 1970

Appendix A

Relative Time Shifts of a Pulse in a Dispersive Medium

The group velocity for an electromagnetic pulse in a dispersive medium is given by

$$v_p = \frac{d\omega}{dk}$$

where ω is the signal (radian) frequency and k is the signal wave number. If only discrete values of ω and k are available, the above expression is approximated by

$$v_p \cong \frac{\Delta\omega}{\Delta k}$$

The pulse propagates a distance Δr in a time

$$\Delta t = \frac{\Delta r}{v_p} \cong \frac{\Delta r}{\Delta\omega/\Delta k} = \frac{\Delta k \Delta r}{\Delta\omega}$$

which is also known as the group delay. If we set $\Delta k = k_1 - k_2$, then

$$\Delta t \cong \frac{k_1 \Delta r - k_2 \Delta r}{\Delta\omega}$$

The signal phase accumulated over a distance Δr is given by $k\Delta r$, so that the above expression becomes

$$\Delta t \cong \frac{\Delta(\text{phase})}{\Delta\omega}$$

Here the frequency has units of radians/sec and the phase is in units of radians.

Appendix B

Ionospheric Reflection Coefficients in the Presence of a Geomagnetic Field

Consider the case of a sharply bounded homogeneous ionosphere with an impressed geomagnetic field. The general case is very complex, but under a well-known condition known as the *quasi-longitudinal (QL) approximation*, the reflection coefficients may be expressed in a reasonably compact form¹³. This condition holds when the geomagnetic field line at the local ionosphere is nearly vertical and the propagation path through the ionosphere is nearly parallel to the local geomagnetic field direction. With this condition, we have:

$${}_pR_p = \frac{A - B}{A + B}$$

where

$$A = (\mu_o + \mu_e)C^2 + \mu_o\mu_e(C_o + C_e)C;$$

$$B = (\mu_o + \mu_e)C_oC_e + (C_o + C_e)C$$

and

$$\mu_o = \sqrt{1 - i \frac{\omega_r}{\omega} e^{i\tau}} \quad ; \quad \mu_e = \sqrt{1 - i \frac{\omega_r}{\omega} e^{-i\tau}}$$

$$\omega_r = \frac{\omega_0^2}{\sqrt{\nu^2 + \omega_L^2}} \quad ; \quad \tan \tau = \frac{\omega_L}{\nu}$$

Here ω_0 is the plasma frequency, ω_L is the electron gyrofrequency, and ν is the electron-ion collision frequency. Finally,

$$C = \cos \theta \quad ; \quad C_o = \cos \theta_o \quad ; \quad C_e = \cos \theta_e$$

The subscripts *o* and *e* refer to the ordinary and extraordinary wave, respectively. Some of these quantities are related through Snell's Law at the atmosphere-ionosphere interface

$$\mu_o \sin \theta_o = \sin \theta = \mu_e \sin \theta_e$$

In the case of zero geomagnetic field ($B = 0$), then $\omega_L = eB/m = 0$ ($e =$ electric charge and $m =$ electron mass). In this case, $\tau = 0$ and $\mu_o = \mu_e$. With these results, it is straightforward to show that the expression for ${}_pR_p$ given above is equivalent to the expression derived in Eq. (3-3) of the main text.