LA-UR-14-27714

Approved for public release; distribution is unlimited.

Title:	The cause of outliers in electromagnetic pulse (EMP) locations
Author(s):	Fenimore, Edward E.
Intended for:	Documentation of work done
Issued:	2014-10-02

The Cause of Outliers in Electromagnetic Pulse (EMP) Locations

Ed Fenimore

National Security Education Center

Los Alamos National laboratory

LA-UR-14-xxxxx

Abstract

We present methods to calculate the location of EMP pulses when observed by 5 or more satellites. Simulations show that, even with a good initial guess and fitting a location to all of the data, there are sometime outlier results whose locations are much worse than most cases. By comparing simulations using different ionospheric transfer functions (ITFs), it appears that the outliers are caused by not including the additional path length due to refraction rather than being caused by not including higher order terms in the Appleton-Hartree equation. We suggest ways that the outliers can be corrected. These correction methods require one to use an electron density profile along the line of sight from the event to the satellite rather than using the total electron content (TEC) to characterize the ionosphere.

Introduction

Electromagnetic pulse (EMP) sources (for example lightning) can be located by observing them with a set of satellites (for example, the GPS satellites). One can use the time of arrivals to triangulate a location. If there are 4 satellites that observe the EMP, one can solve analytically for the location and time of the EMP. If there are only 3 satellites that observe the EMP, one can make an assumption such as that the EMP was on the surface of the Earth and again solve for the location and time analytically. When there are 5 or more satellites, one can do a best fit to all the observations to obtain a location and time. This has worked well in the past, giving accurate locations. However, sometimes the analysis results in a position that is much further than normal from known location, a so-called "outlier".

Here we present a method to obtain a location where there are 5 or more satellites observing the event. We start with a method to get an initial guess and then a fitting method that accurately finds the EMP location.

Initial Guess

Assume the event occurred at location X_{EMP}, Y_{EMP}, Z_{EMP} at time T_{EMP} and was seen by N satellites at times T_i and locations x_i, y_i, z_i. The goal is to find X_{EMP}, Y_{EMP}, Z_{EMP}, T_{EMP} from the satellite locations and the time of arrival at those satellites. The distance from the event to satellite "*i*" is

$$c(T_i - T_{EMP}) = c\Delta T_i = \sqrt{(x_i - X_{EMP})^2 + (y_i - Y_{EMP})^2 + (z_i - Z_{EMP})^2} \qquad \text{Eq 1}$$

Where c is the speed of light. There are N constraints like Equation 1. Direct solution of the N equations is difficult especially considering the observations contain noise so perhaps no event location and time satisfies all Equations 1. The usual solution is to use the time difference of arrival (DTOA) method where one works with ($T_i - T_j$) rather than (T_i-T_{EMP}). There are many references to this, here we follow the notation and method of Yang et al (2010) expanded to three dimensions.

One satellite is selected to use relative to the other satellites. We use satellite number 1 (i. e., j=1) which is usually the satellite with the smallest T_i, that is, the closest satellite to the event. The additional distance the signal travels to the ith satellite relative to the time it travels to the 1st satellite is

$$c(\Delta T_i - \Delta T_1) = c(T_i - T_1) = \sqrt{(x_i - X_{EMP})^2 + (y_i - Y_{EMP})^2 + (z_i - Z_{EMP})^2} - \sqrt{(x_1 - X_{EMP})^2 + (y_1 - Y_{EMP})^2 + (z_1 - Z_{EMP})^2} \qquad \text{Eq 2}$$

Rearranging

$$c(T_i - T_1) + \sqrt{(x_1 - X_{EMP})^2 + (y_1 - Y_{EMP})^2 + (z_1 - Z_{EMP})^2} = \sqrt{(x_i - X_{EMP})^2 + (y_i - Y_{EMP})^2 + (z_i - Z_{EMP})^2} \qquad \text{Eq 3}$$

Squaring both sides of equation 3 to linearize it and rearranging terms gives *N*-1 equations:

$$(x_i - x_1)(x_{EMP} - x_1) + (y_i - y_1)(y_{EMP} - y_1) + (z_i - z_1)(z_{EMP} - z_1) + c(T_i - T_1)r_1 = \frac{1}{2}\left((x_i - x_1)^2 + (y_i - y_1)^2 + (z_1 - z_1)^2 - c^2(T_i - T_1)^2\right) \qquad \text{Eq 4}$$

Where

$$r_1 = \sqrt{(x_1 - X_{EMP})^2 + (y_1 - Y_{EMP})^2 + (z_1 - Z_{EMP})^2} \qquad \text{Eq 5}$$

Equation 4 can be arranged into matrix form as

$$A\theta = B \qquad \text{Eq 6}$$

where

$$A = \begin{bmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 & c(T_2 - T_1) \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 & c(T_3 - T_1) \\ . & . & . & . \\ . & . & . & . \\ x_N - x_1 & y_N - y_1 & z_N - z_1 & c(T_N - T_1) \end{bmatrix} \qquad \text{Eq 7}$$

$$\theta = \begin{bmatrix} X_{EMP} - x_1 \\ Y_{EMP} - y_1 \\ Z_{EMP} - z_1 \\ c(T_{EMP} - T_1) \end{bmatrix} \qquad \text{Eq 8}$$

$$B = \frac{1}{2}\begin{bmatrix} (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 - c^2(T_2 - T_1)^2 \\ (x_3 - x_1)^2 + (y_3 - y_1)^2 + (z_3 - z_1)^2 - c^2(T_3 - T_1)^2 \\ . \\ . \\ (x_N - x_1)^2 + (y_N - y_1)^2 + (z_N - z_1)^2 - c^2(T_N - T_1)^2 \end{bmatrix} \qquad \text{Eq 9}$$

Equation 6 is N-1 equations with 4 unknowns. The unknowns are in $\boldsymbol{\theta}$ and are effectively the position and time of the EMP relative to satellite number 1. If there are 5 satellites, ***A*** is a square matrix. For 5 or more satellites, $\boldsymbol{\theta}$ can be found as

$$\theta = (A^T A)^{-1} A^T B \qquad \text{Eq 10}$$

Since x_1, y_1, z_1, and T_1 are known; θ gives X_{EMP}, Y_{EMP}, Z_{EMP} and T_{EMP}.

Equation 10 effectively gives a best fit to the N satellites that saw the event. This is likely better way to get an initial guess for the location than a strategy that takes 4 satellites at a time. One could, for example, take all permutations of 4 satellites and form an average of the resulting positions. This has the advantage that the 4-satellite solution is analytic (as is equation 10). However, four satellites often give two solutions due to the square root in equation 1. Equation 10 gives a unique solution and we have found that it is always close to the true location and close to the best fit answer we find below.

Best Fit Solution

The DTOA method is biased because satellite 1 is treated differently than the other satellites. We next find the best fit solution. We define a goodness of fit parameter, χ^2, which is the square of the distance from the EMP event to each satellite minus the light travel distance:

$$\chi^2 = \sum_{i=1}^{N} \{(x_i - X_{EMP})^2 + (y_i - Y_{EMP})^2 + (z_i - Z_{EMP})^2 - c^2 (T_i - T_{EMP})^2\}^2 \qquad \text{Eq 11}$$

Although symbolized with χ^2, it does not follow the chi-square statistic because we have not made it relative to the noise. If the x_i, y_i, and z_i are for GPS satellites, they are known to extremely high accuracy and would not contribute uncertainty compared to the uncertainty associated with the time. Not including noise in equation 10 is equivalent to assuming that the noise uncertainty on the timing is the same for all satellites.

There are several techniques for the best-fit solution that have been used for EMP in the past. It is a difficult problem because it requires a 4$^{\text{th}}$ order search of parameters that are nonlinearly coupled. Some methods are based on using 4 dimensional gradients, for example the Brent and Powell method from Numerical Recipes. The convergence criterion is that χ^2 falls below some level or the gradients become very small. Another method that has been used, the Nelder and Mead Simplex method, establishes a geometric figure in a 5 dimensional space and has rules to move the geometric figure towards the χ^2 minimum. When χ^2 changes slowly, those rules shrink the geometric figure to a small volume that should contain the χ^2 minimum. These methods depend on the value of χ^2 and can converge prematurely, resulting in not finding the best fit.

We use a 4 dimensional Golden Search technique. It establishes a minimum and maximum range that brackets each parameter. Our initial guess is used with a generous possible range. We evaluate χ^2 at the points that split the range by the Golden ratio (0.618, 0.382) and use them to reduce the range that is bracketed. This is repeated until the range that is bracketed for each parameter is arbitrarily small. This method does not dependent on the value of χ^2, only that it has a minimum. At the end of the process we know the location of the minimum without depending on gradients, the magnitude of χ^2, or the shape of the χ^2surface. If the location of the minimum is at the edge of the range (i. e., the actual minimum is outside the range), the range is increased. Unless there is a secondary minimum, we are guaranteed to find the minimum to arbitrary accuracy. We have never had a case of a secondary minimum using this technique.

Simulations

To evaluate how well the EMP location algorithm works, we ran Monte Carlo simulations using the CONSIM code. Random locations on the Earth between -60 degree latitude and +60 degree latitude were selected for 1000 EMP events. The altitude was set to 2 km and the times were separated by at least 1 sec with some randomness. A GPS constellation of 28 satellites was established. On average, each event was seen by about 9 satellites so there were ~ 9000 simulations of instrumental responses.

The source was a double exponential in the time domain and was propagated through the ionosphere using the CONSIM ray tracing algorithm and the full Appleton-Hartree equation. The International Reference Ionosphere (IRI) was used to define the electron density and magnetic field as a function of time, latitude, longitude, and altitude.

The instrument had 5 receivers with frequencies similar to the BDW instrument on GPS. Each receiver measures the power within a narrow bandpass (typically 2 or 3 MHz). The ionosphere disperses the frequencies such that the signal is in the bandpass for a short period of time. Due to the dispersion, high frequencies show up first and then low frequencies. Let t_k be the time in the instrument when the k^{th} frequency peaks. The T_i needed by equations 10 and 11 is effectively when an infinite frequency would arrive at the satellite. T_i can be fit to the t_k's by minimalizing a goodness of fit parameter:

$$\chi^2 = \sum_k \left(T_i + \frac{8450 TEC}{2\pi f_k^2} - t_k \right)^2 \qquad \text{Eq 12}$$

Here, TEC is the total electron content which is the integrated column density of electrons on the line of sight path from the event to the satellite. Both TEC and T_i are free parameters easily found by solving two equations with two unknowns ($\frac{\partial \chi^2}{\partial T_i} = 0, \frac{\partial \chi^2}{\partial TEC} = 0$). The TEC/$f^2$ dependency comes from a series expansion of the Appleton-Hartree equation which gives the phase delay as a function of frequency due to the ionosphere.

For each event we had N measurements of x_i, y_i, z_i, and T_i. Equation 10 was used to get an initial guess where the event was located and equation 11 was used in CONSIM's 4-dimensional Golden Search package to get a best fit X_{EMP}, Y_{EMP}, Z_{EMP} and T_{EMP} for each of the 1000 events.

From X_{EMP}, Y_{EMP}, and Z_{EMP} and the known random location, the error in X and Y and the total error ($(X^2+Y^2)^{1/2}$) was found for each of the 1000 events. Figure 1 shows a typical distribution of errors. Most of the events have errors that are small. The green circle is the circular error probable (CEP) and contains half the events. However, there are outliers with much larger errors.

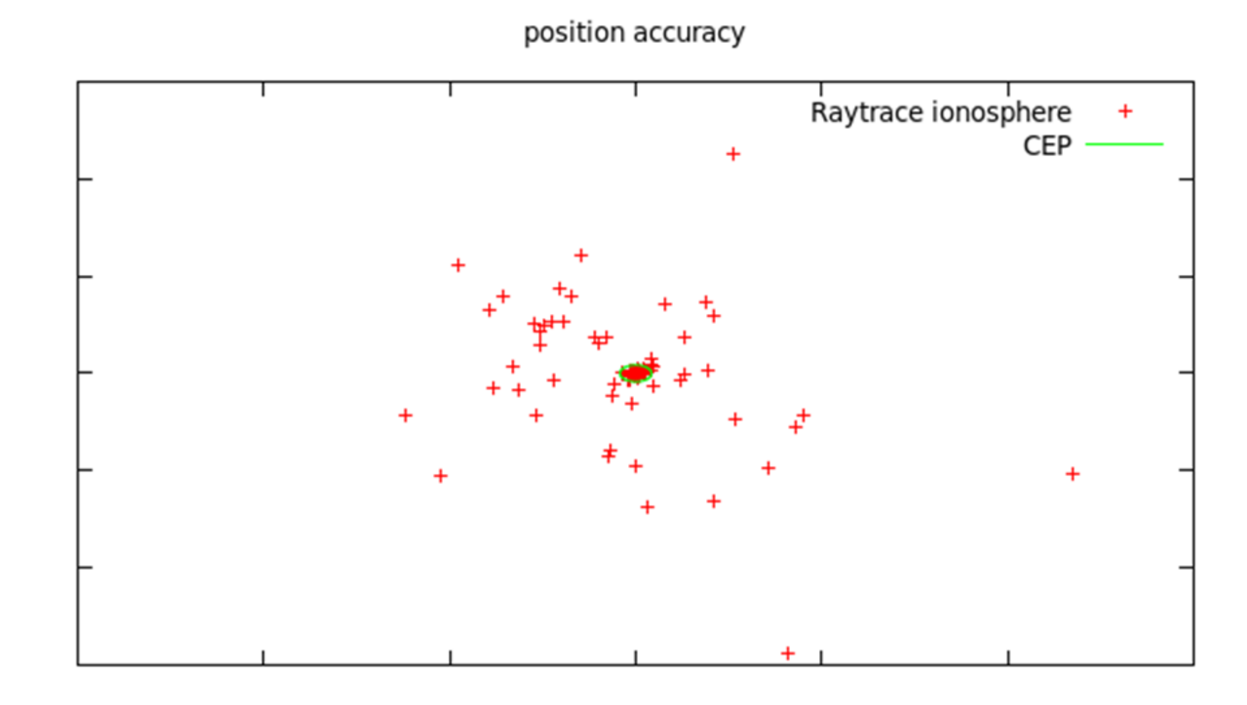

Figure 1: A typical result of simulating 1000 random events into a constellation of 28 GPS satellites. The red crosses show the analyzed position relative to the true location which is at the center. This uses ray tracing to calculate the phase delay for each frequency through the ionosphere. The green circle contains one half of the events. Note most are found accurately, but there are outliers.

What causes the Outliers?

The outliers as shown in Figure 1 have been a long standing problem. It is usually assumed that they occur because the TEC/f^2 term does not adequately represent the phase changes caused by the ionosphere. There have been attempts to correct the outliers by adding higher order terms from the Appleton-Hartree equation into equation 12.

CONSIM has a unique ability to model the ionosphere many different ways. Figure 2 shows four of those ways.

The “simple” ionosphere transfer function (ITF) uses a first order series expansion of the Appleton-Hartree equation including mode splitting due to the Earth’s magnetic field. The ITF as a function of frequency is

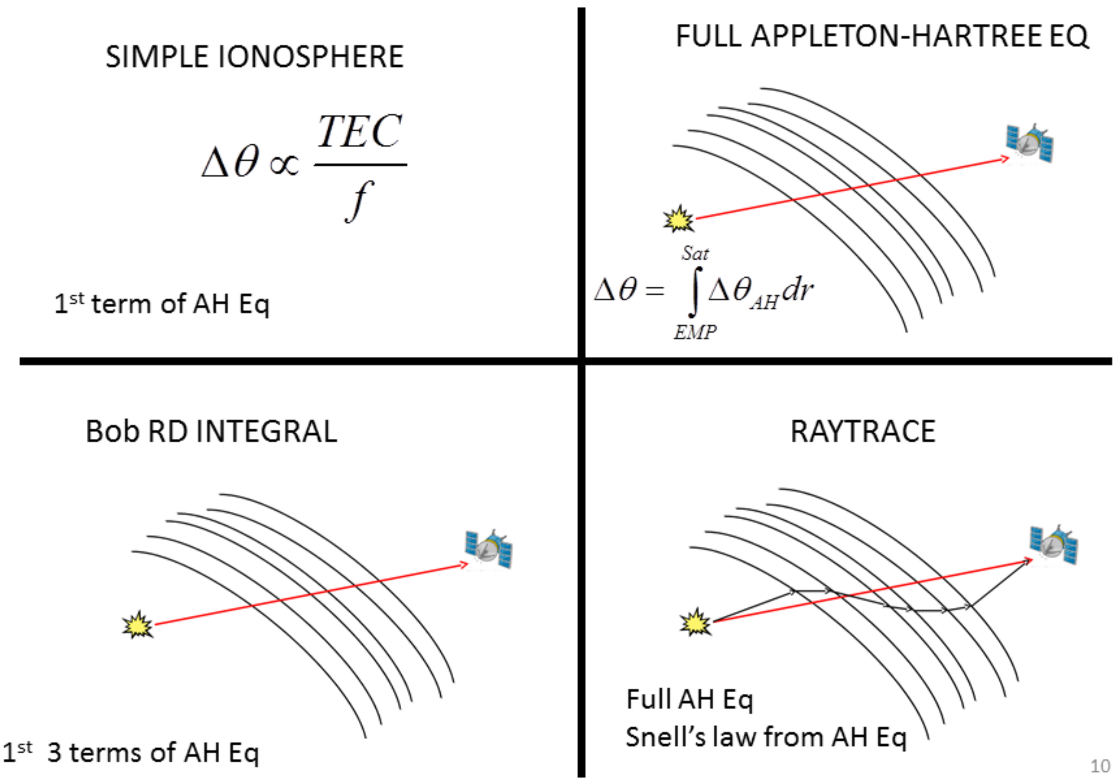

Figure 2: The four different methods in CONSIM for calculating the phase delays due to the ionospheres. The simple ionosphere uses the first term from the Appleton-Hartee equation. The Bob Roussell-Dupre (Bob RD) method uses the first 3 terms integrated on the line of sight to the satellite and the full Appleton-Hartee method uses effectively an infinite number of terms in the Appleton-Hartree equation integrated along the line of sight. The Ray trace method uses the full Appleton-Hartree equation and solves for what path one must start out on in order to hit the satellite when the ray bends due to changes in the electron density.

$$ITF_{\pm}(f) = \frac{8450TEC}{f(1 \pm \frac{17.6B\cos(\phi)}{4\pi f})} \qquad \text{Eq 13}$$

The TEC is found by integrating the electron density (from IRI) along the line of sight from the EMP event to the satellite. The magnetic field (B) and angle of the line of sight with respect to the magnetic field (ϕ) are found from weighted averages on the line of sight.

Roussel-Dupre, Jacobson, and Triplett (2001) developed an ITF (called Bob RD in Figure 2) from the first three terms of a series expansion of the Appleton-Hartree equation. That method integrates along the line of sight from the EMP event to the satellite to determine the phase change for each frequency.

The “full Appleton-Hartree” ITF in Figure 2 integrates the full Appleton-Hartree equation along the line of sight from the EMP event to the satellite. It is equivalent to an infinite series expansion of the Appleton-Hartree.

The “ray trace” ITF in Figure 2 is the only one that does not integrate along the line of sight from the EMP event to the satellite. Rather, it iteratively searches for a ray with a “shooting” angle relative to the line of sight that is propagated from shell to shell in an onion shell model until it hits the satellite. At each shell, Snell’s law is used to determine how the ray refracts. The phase delay is integrated along each path segment.

These ITF’s have been extensively validated in a two year effort involving Los Alamos and Sandia (see Fenimore, 2014). The “simple” ITF code in CONSIM was analyzed and compared to analytic calculations (see Appendix D, Fenimore 2014). The Bob RD ITF was integrated into CONSIM from the “TIPC” code that was written by Bob Roussel-Dupre. Its close agreement with the ray trace code (see below) gives credence to its accuracy. The ray trace code was compared directly to the ISP ray trace code of Sandia and it agreed to within one part in 10^4 for all frequencies over a range of conditions (See Fenimore, 2014). The “full Appleton-Hartree” ITF uses the same infrastructure (IRI, onion shells, Appleton-Hartree routines) that the “simple”, ray trace, and “Bob RD” ITFs use. It agrees closely with the simple ITF when the frequency is well above the plasma frequency.

Figure 3 shows the simulation of 1000 random events into a GPS constellation using the four ITFs. The simulation using the “simple” ITF has no outliers. This is not surprising. The analysis that gives the T_i (i. e., equation 13) assumes a simple ITF. Since the simulation assumes a simple ITF and the analysis assumes a simple ITF, it always gets an accurate EMP location. The ray trace results are also not surprising. It has been known for a long time that outliers occur in reality and the ray trace code is the highest fidelity simulation code that we have.

What is surprising is that the “full Appleton-Hartree” ITF does not show outliers and the Bob RD ITF does show outliers. The full Appleton-Hartree ITF has the first three terms in a series expansion of the Appleton-Hartree equation (as does the Bob RD ITF) as well as all higher order terms. We conclude that outliers are not caused by omitting higher order terms in equation 13 and corrections based on the higher order terms are unlikely to correct the outliers. It is an underappreciated fact that the Roussel-Dupre, Jacobson, and Triplett (2001) integrals also account for the additional path length due to refraction. The ray trace ITF accounts for the refraction by finding the actual path the rays take to the satellite. The Bob RD integrals includes the additional path length through the α terms in equation 10 of Roussel-Dupre, Jacobson, and

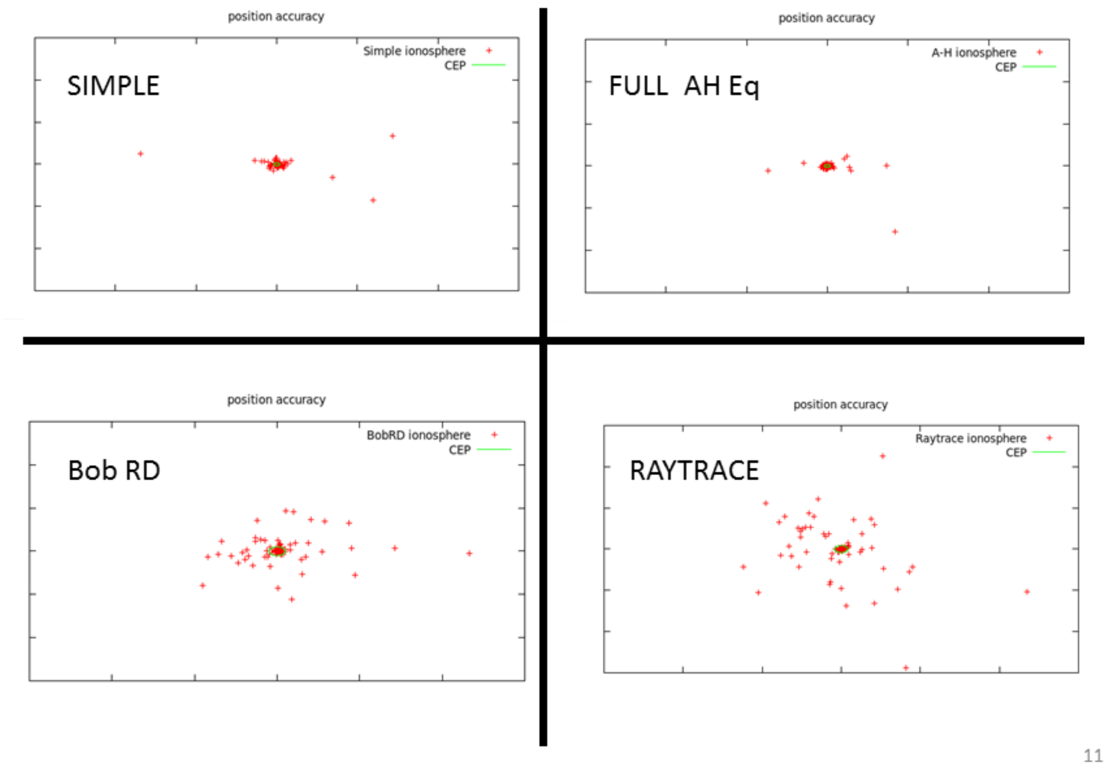

Figure 3: Simulations of 1000 random events using four different ITFs. It is no surprise that the simple ITF gives few outliers since the processing of the receivers effectively assumes the simple ITF. The ray trace results are also not surprising since outliers occur in reality. The “full Appleton-Hartree” and “Bob RD” results are quite surprising. Previous work has assumed that the cause of the outliers was insufficient high order terms from the Appleton-Hartree equation in the processing of the receivers. The lack of outliers with the “full Appleton-Hartree” simulations says that high order terms are *not* the cause of outliers. The presence of outliers with the “Bob RD” ITF and the fact that the “Bob RD” ITF includes the phase delay due to refraction strongly suggests that the cause of the outliers is not accounting for the additional path length due to refraction.

Triplett (2001). This strongly suggests that the cause of the outliers is not accounting for the additional path length due to refraction rather than not accounting for the higher order terms in the Appleton-Hartree equation.

Figure 4 takes each of the four sets of simulations in Figure 3 and finds the sum (normalized to unity) of the number of events as a function of the distance from the origin. Note how similar “simple” and “full Appleton-Hartree” are and how similar “Bob RD” and “ray trace” are.

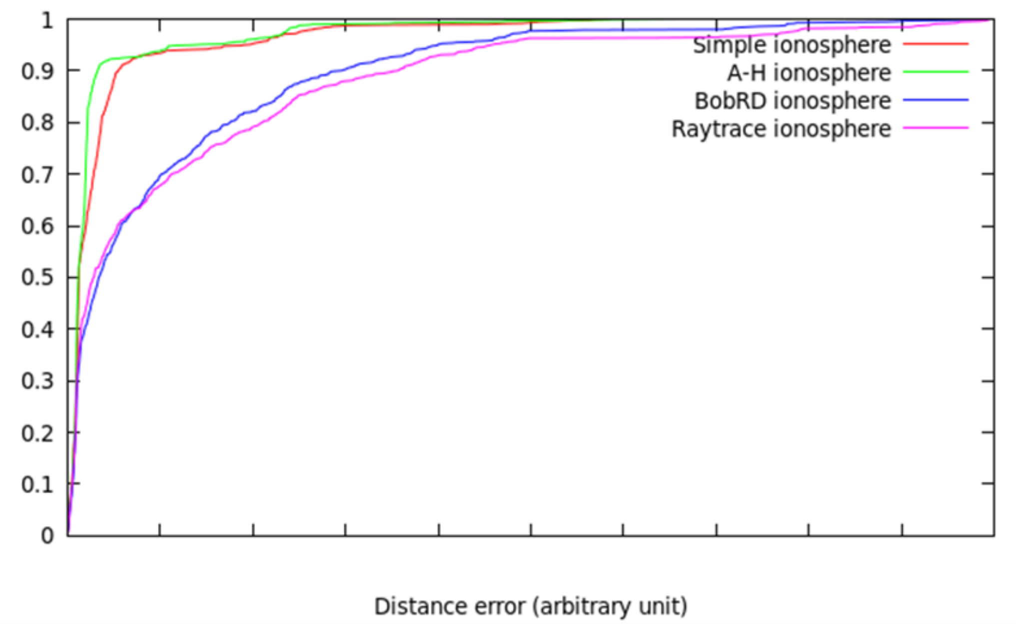

Figure 4: The integral distribution functions for the four simulations in Figure 3.

The initial guess (equation 10) gives distributions that are about a factor of 3 wider (see Figure 5). In all the simulaitons done, there was never a problem with convergence or secondary minimum. The initial guess is actually quite accurate.

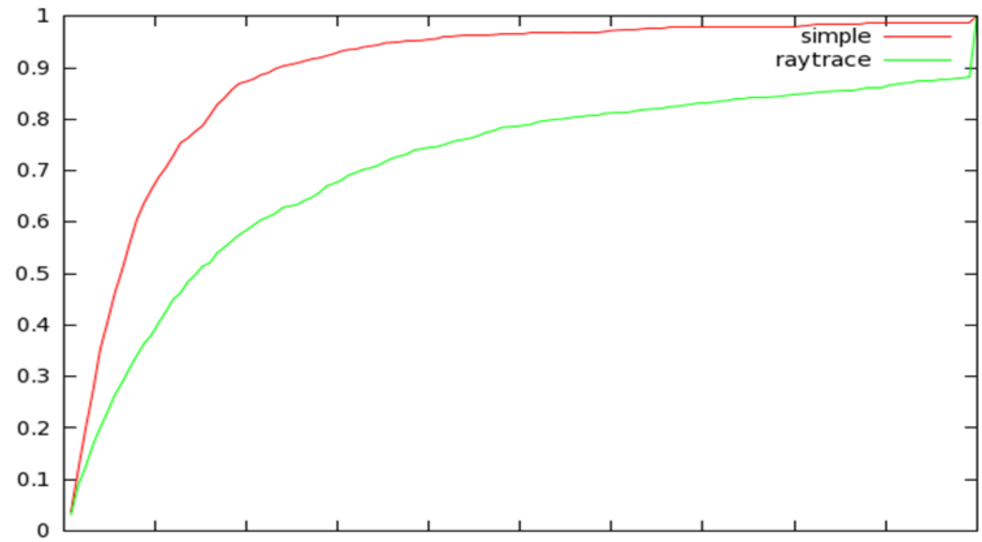

Figure 5: The integral probability distribution function for the initial guess with the "simple" ITF and the "ray trace" ITF.

Correcting the Outliers

If the outliers are caused by the additional path length due to refraction, there are likely ways to correct those outliers. Since the refraction is caused by the electron density profile, it is inevitable that we will need the electron density as a function of distance along the line of sight rather than just the TEC. This has not been previously tired. Previously, polynomials (f^2, f^3, f^4) have been fit to simulations or t_k's in order to generate correction functions. Here, we will estimate correction terms from an estimated electron density profile for the geometry of the event.

The electron density profile for the geometry of the event is not difficult to obtain even in real time. One can go to the NeQuick2 web site (http://t-ict4d.ictp/nequick2/nequick2-2-web-model) and use the initial guess as the event location and the known satellite position and time. For real time electron density profiles, one can select "Daily Solar Radio Flux" and get the real time F10.7 cm flux from www.spaceweather.com. NeQuick2 quickly runs a model and the tabular results can be read directly into CONSIM and used in any of the ITFs. Alternatively, the NeQucik2 source code is likely available and it only needs the time, initial guess, satellite position, and F10.7 cm flux to give the electron density along the path from the event to the satellite.

Let $\Delta\varphi(\overline{r_i}, T_i, f_k, \lambda)$ be the phase delay from the Roussel-Dupre (2001) integrals as a function of the path from the initial guess to the satellite ($\overline{r_i}$), the time (T_i), the receiver frequency (f_k), and a unitless scale factor (λ). The unitless scale factor λ is used to scale the NeQuick2 electron density profile, if necessary. The time needs only be accurate to a minute or so, it is effectively known.

There are two cases: the first case is if the individual t_k's are available on the ground and the second case is if one only has a T_i from the satellite. Here we suggest ways that T_i could be found in these cases. Future work should test these suggestions. If one has the individual t_k's one can find T_i directly by minimizing

$$\chi^2 = \sum_k \left(T_i + \frac{\Delta\varphi(\overline{r_i}, T_i, f_k, \lambda)}{2\pi f_k} - t_k \right)^2 \qquad \text{Eq 14}$$

NeQuick2 will not perfectly set the electron density profile, so likely one should treat λ as a free parameter. In the first term of the series expansion of the Appleton-Hartree equation, λ acts like TEC. The two free parameters (T_i and λ) could be found by a 2-dimensional Golden Search. To correct the outliers, one would use the T_i from equation 14 in equation 11.

In the second case we do not have the t_k values. We could solve for the effect of assuming the simple ITF in equation 12 versa using the Roussel-Dupre (2001) integrals by minimizing

$$\chi^2 = \sum_k \left(\Delta T_i + \frac{8450 TEC}{2\pi f_k^2} - \frac{\Delta\varphi(\overline{r_i}, T_i, f_k, \lambda)}{2\pi f_k} \right)^2 \qquad \text{Eq 15}$$

Here there would be two free parameters: ΔT_i and λ (λ and the electron density profile sets TEC). Then, to correct the outliers, one revises equation 11 to be

$$\chi^2 = \sum_{i=1}^{N} \{(x_i - X_{EMP})^2 + (y_i - Y_{EMP})^2 + (z_i - Z_{EMP})^2 - c^2(T_i + \Delta T_i - T_{EMP})^2\}^2 \qquad \text{Eq 16}$$

More studies should be done to verify if these suggestions correct the outliers.

References

Fenimore, "The Verification and Validation of the EMP Simulations in CONSIM during Red Team 4 and 5", LA-UR 14-20104, 2014

Roussel-Dupre, Jacobson, and Triplett, “Analysis of FORTE data to extract ionospheric parameters”, Radio Science **36**, pg 1615, 2001.

Yang et al., “Constrained Total Least-Squares Location Algorithm Using Time-Difference-of-Arrival Measurements”, IEEE Transactions on Vehicular Technology **59**, pg 1558, 2010.