Faith in a seed: on the origins of equatorial plasma bubbles

J. M. Retterer¹ and P. Roddy²

¹Institute for Scientific Research, Boston College, 140 Commonwealth Avenue, Chestnut Hill, Massachusetts, 02467, USA
²Space Vehicles Directorate, Air Force Research Laboratory, Kirtland AFB, New Mexico, USA

Correspondence to: J. M. Retterer (john.retterer@bc.edu)

Received: 15 November 2013 – Revised: 19 February 2014 – Accepted: 18 March 2014 – Published: 15 May 2014

Abstract. Our faith in the seeds of equatorial plasma irregularities holds that there will generally always be density perturbations sufficient to provide the seeds for irregularity development whenever the Rayleigh–Taylor instability is active. When the duration of the time of the Rayleigh–Taylor instability is short, however, the magnitude of the seed perturbations can make a difference in whether the irregularities have a chance to grow to a strength at which the nonlinear development of plumes occurs. In addition, the character of the resulting irregularities reflects the characteristics of the initial seed density perturbation, e.g., their strength, spacing, and, to some extent, their spatial scales, and it is important to know the seeds to help determine the structure of the developed irregularities. To this end, we describe the climatology of daytime and early-evening density irregularities that can serve as seeds for later development of plumes, as determined from the Planar Langmuir Probe (PLP) plasma density measurements on the C/NOFS (Communication and Navigation Outage Forecast System) satellite mission, presenting their magnitude as a function of altitude, latitude, longitude, local time, season, and phase in the solar cycle (within the C/NOFS observation era). To examine some of the consequences of these density perturbations, they are used as initial conditions for the PBMOD (Retterer, 2010a) 3-D irregularity model to follow their potential development into larger-amplitude irregularities, plumes, and radio scintillation.

Keywords. Ionosphere (equatorial ionosphere; ionospheric irregularities; modeling and forecasting).

"Though I do not believe that a pla[sm]a bubble will spring up where no seed has been, I have great faith in a seed. Convince me that you have a seed there, and I am prepared to expect wonders." – Henry David Thoreau

1 Introduction

There is much that remains poorly understood about the origins of plasma density irregularities in the low-latitude region of Earth’s ionosphere (Kelley et al., 2011). These irregularities, first observed in ionosonde traces, have subsequently been observed by incoherent-scatter radars and in airglow images of the night sky, in situ measurements of plasma density and velocity, and by numerous other means. They are believed to begin as small-amplitude density variations which are then amplified by the exponential growth in time that is characteristic of plasma instabilities like the Rayleigh–Taylor (RT) instability of the bottomside of the ionospheric F layer (Kelley, 1989) until their amplitudes are large enough that nonlinear interactions significantly modify their structure, leading to rising plumes of low-density plasma filled with turbulence.

Bubble formation can occur anytime that the Rayleigh–Taylor growth rate is large enough for a long enough time for significant irregularities to develop. Generally they form at night, when the E region conductivity is small (otherwise, the E region conductivity will short out the growing polarization electric fields and prevent the feedback that leads to the instability). Uplift of the F layer aids the instability by moving the plasma layer into altitudes where the neutral collision frequencies are weaker and the RT growth rate is consequently larger. Thus, bubbles are often seen shortly after sunset, when the prereversal enhancement of the upward drift leads to stronger growth rates, but they can be seen at other times of night when other mechanisms produce electric fields that cause the uplift of the F layer (Yizengaw et al., 2013).

Because the RT instability initially merely amplifies the structure of the perturbations, a number of the characteristics of the original irregularities (for example, the zonal spacing of plumes) are carried over into the developed turbulent
plumes. Moreover, when the time interval in the post-sunset sector in which the RT growth rate is significant is short, the strength of evening irregularities can be as much dependent on the initial amplitudes of the irregularities as it is on the strength of the RT instability because the stronger the initial perturbation is, the less Rayleigh–Taylor growth is required to boost the irregularities into the nonlinear regime. Figure 1 illustrates this situation. Several (two-dimensional, for simplicity) simulations (Retterer, 2010a) of plume formation are performed, with both different strengths of the RT instability (set by the magnitude of the preretal vertical drift of the ambient plasma – higher drift means stronger RT growth rate) and initial density perturbations of different magnitudes, with the peak density perturbation 0.003, 0.01, or 0.03 times the F peak ambient density. The quantity plotted is the peak vertical plasma drift within the simulation volume as a function of time. A rapid growth of this drift indicates that a plume of low-density, rapidly upward-rising plasma has formed. The top panel shows that when the RT instability is weak, none of these perturbations is sufficient to trigger plume formation. With the next largest RT growth rate, only the biggest initial density perturbation triggers plume development. With stronger RT growth, smaller initial perturbations can trigger plume development, but it is always the biggest initial perturbation (the red curves) that initiates plume formation sooner and produces the strongest plumes (as marked by the peak magnitude of the vertical drift).

But what are the origins of the initial density fluctuations? What are their expected spatial and temporal scales? What are their seasonal, latitudinal, and longitudinal occurrence climatologies, and can knowing this help us forecast the probability of plume formation and radio scintillation? One of the processes that lead to ion density variations is momentum exchange with neutral constituents of the thermosphere, whereby the friction with the neutral component can induce the ions to drift in the same way as the neutral wind. When the neutral wind varies periodically as in a wave, the ion drift can be found to vary periodically in the same way. As the ion drift responds to the varying neutral wind, the plasma density responds in the way dictated by the continuity equation, producing density perturbations that can be amplified in the RT instability.

Wave perturbations in the winds and density are pervasive features of the neutral atmosphere that exist over a wide range of spatial and temporal scales. Winds are responsible for many of the major electrodynamic features of the ionosphere through the production of dynamo current systems (Fejer, 2011). In addition to the larger-scale tidal oscillations and planetary waves (Maute et al., 2012), there also exist smaller-scale gravity waves (Hines, 1974), traveling neutral disturbances that induce corresponding wave variations in the ionospheric plasma (Hooke, 1968). There are a number of sources for these waves, ranging from storm-time energy inputs at high latitudes to tropospheric storm systems at low and middle latitudes (Hagan and Forbes, 2003).

Note that as the wave-like plasma density perturbations are supported and maintained by polarization electric fields that develop within their structure, the notion that, for the Rayleigh–Taylor instability, the strength of the growth rate and the initial density perturbation are independent quantities is illusory. Large-scale wave structures (Tsunoda, 2010) can provide both the initial density perturbation and enhancements in the upward plasma drift that enhance the Rayleigh–Taylor growth rate. Typically, what is seen in the simulations (Retterer, 2010a) is that plumes initially develop from the points where the bottomside of the F layer is the highest in

---

**Fig. 1.** The sensitivity of bubble development to the magnitude of the initial perturbation as well as the magnitude of the Rayleigh–Taylor growth rate. Vz peak is the maximum vertical velocity (in m s⁻¹) over the grid of the 2-D simulation of plume development; the upward jump in Vz peak indicates the formation of plumes. The panels of the plot show the results for the different preretal ambient drifts, which control the strength of the Rayleigh–Taylor growth rate. The different color curves represent the results of simulations with different initial magnitude perturbations: black, blue, and red represent a maximum perturbation 0.003, 0.01, and 0.03 times the peak ambient density, respectively.
Empirical knowledge of the phenomenon is the first key to understanding it. Here we take advantage of the survey of low-latitude plasma densities provided by the Communication and Navigation Outage Forecast System (C/NOFS) satellite from the solar minimum of 2008 through the rising phase of the solar cycle into 2012. We employ a high-pass filtering technique to isolate the shorter-term density irregularities in which we are interested from the smoother variations of the ambient plasma as the satellite moves around its low-inclination orbit, sampling the plasma density through longitudes and local times. We accumulate statistics of the amplitude of the density fluctuations and plot it as a function of altitude, latitude, local time, longitude, season, and phase of the solar cycle. Lastly, we use snapshot examples of the density perturbations as an initial condition for a 3-D plume model to demonstrate how the fluctuations will evolve and how irregularity-filled plumes and radio scintillation can develop from them.

2 Data analysis

The C/NOFS satellite was launched in April 2008 as a component of a system to specify and forecast the occurrence of radio scintillations that interfere with the systems that rely on transionospheric radio propagation (de La Beaujardiere et al., 2004). Its orbit was elliptical (400 × 850 km) with low inclination (13°) so as to sample the ionospheric plasma in the region near the geomagnetic equator, one of the regions where
the density irregularities responsible for the scintillation are found. Among the instruments onboard the satellite, we employ here the plasma density measurements, \( N_e(t) \), from the planar Langmuir probe (PLP) (Roddy et al., 2010), at a one sample per second resolution.

To isolate the faster density variations of interest, \( \Delta N_e \), from the longer spatial scale or slower variations of the ambient density (considered to be at along-track scales of 300 km or more), a smoothed version of the PLP density data series is formed, \( N_{e0} \), which is then subtracted from the original data series, \( N_e \), to leave the variations at the shorter scales, \( \Delta N_e = N_e - N_{e0} \). The smoothed density can be estimated either by a running mean or, if the irregularities expected are overwhelmingly depletions (reductions in density), estimated by the envelope of density maxima spaced on an ambient-scale grid throughout the time series. Dao et al. (2011) employed similar techniques, but looked solely at \( \Delta N_e / N_{e0} \), which is always small in the daytime, to focus on the nighttime irregularities. A sample of this density-irregularity calculation using data from day 281, year 2010, around 18 UT, is shown in Fig. 2. The top panel gives \( \Delta N_e \), the density irregularity in number per cm\(^3\) as calculated by subtracting the smoothed density from the original signal, and the second panel gives this density irregularity normalized by the ambient density, \( \Delta N_e / N_{e0} \). The ambient density in this example was determined using a set of density points sampled every 40 s, providing a spatial resolution of about 300 km; thus, \( \Delta N_e \) includes all spatial scales smaller than that (down to the 8 km resolution due to the sampling rate). The other panels give information about the time and location of the observations. The altitude of the satellite is shown in black in panel three, along with the apex altitude, shown in red (the altitude at the geomagnetic equator to which the location of the satellite maps; where the two curves touch, the satellite is at the geomagnetic equator). The fourth panel shows the latitude of the satellite, and the fifth panel gives the longitude (black curve, with axis on the left) and local time (blue curve, with axis on the right) of the satellite. The gray lines mark the points where the local time is 18:00 LT, while the yellow lines mark the points where the local time is 06:00 LT.

We first see in this segment of data, at 21.8 UT (around 10:00 LT), a burst of irregularities with an amplitude of a few 1000 cm\(^{-3}\), followed by a few more bursts of smaller amplitude throughout the afternoon. Because the ambient density is high during the day, these irregularities have small fractional strength, \( \Delta N_e / N_{e0} \), and may not have much radio-propagation effect themselves, but they persist beyond 18:00 LT into the early evening, where they can act as seeds for the development of stronger irregularities amplified by the RT instability, such as can be seen later in the orbit, around 22.5 UT (around 20:00 LT).

Another case, from day 282, 2010, is shown in Fig. 3, in exactly the same format as Fig. 2. Here the sampled daytime irregularities are smaller, perhaps because the satellite is farther away from the geomagnetic equator (note the distance between the satellite altitude and apex altitude in the third panel), but the early evening irregularities might be a little larger and have a slower temporal variation (corresponding to larger spatial scales) than the case in Fig. 2. The spectral density of the power in these irregularities near 18:00 LT as a function of wave number (mapping from sampled time to space assuming static structures sampled by the satellite moving at 8 km s\(^{-1}\)) is shown in Fig. 4. We see that the spectrum is close to a power law, as comparison with the straight line in this log-log plot suggests, with a spectral index close to \(-2\). Unfortunately, this seemingly ubiquitous spectral shape in atmospheric fluid turbulence leaves few clues (like a preferred wavelength) as to the origins of the waves.

3 Climatology

To produce a more general picture of the occurrence of these daytime density irregularities, we calculate their climatology, the statistical description of how their magnitudes vary with parameters such as altitude, local time, latitude, longitude, season, and phase of solar cycle, by binning large numbers of observations of them over a significant period of time. We choose periods of accumulation that are multiples of the orbital precession period of the satellite (67 days) to insure better uniformity of sampling.
Figure 5 shows the climatology of irregularity strength as a function of longitude and local time through the first four years of the C/NOFS era. Each panel gives the root-mean-square magnitude of the irregularities as a function of longitude and local time, integrated over all altitudes (although we will see later that most of the contribution will come from the lowest altitudes that the C/NOFS samples). Each row gives the results for a year, while each column shows the results for a given season: summer (Northern Hemisphere), fall, winter, and spring, in order. The sampling for the summer plot was over 134 days ending on day 249, the sampling for the fall seasons was over 67 days ending on day 307, that for winter 134 days ending on day 67, and that for spring 67 days ending on day 123. Through the course of the C/NOFS era, we see the increases in density due to the stronger plasma photo production in the rising phase of the solar cycle, as well as seasonal variations, e.g., the daytime irregularities are weakest in the summer season. Throughout the data set we see the strong evening and early morning irregularities produced by the RT amplification, which appear to follow the well-known climatology for scintillation and equatorial plasma bubbles (Gentile et al., 2006): strong through many longitudes at the spring and fall equinoxes, strong in the American sector in winter, strong in the Pacific sector in summer, plus the general tendency to increase in intensity as the solar activity increases. Note how the daytime irregularities persist into the evening, where they can serve as seeds for the development of the strong irregularities and plumes under the influence of the RT instability. There is some suggestion of the four-node pattern in longitude that is a sign of coupling with the lower atmosphere (Immel et al., 2006), which is even more strongly visible in the longitudinal modulations.
Fig. 6. Climatology of the ambient density as a function of longitude and local time, integrated over altitudes and averaged over seasons from (Northern Hemisphere) summer 2008 (upper left) to spring equinox 2012 (lower right). The density scale is logarithmic, ranging from $3 \times 10^3$ to $3 \times 10^6$ cm$^{-3}$.

Figure 8 shows the climatology of the rms density-irregularity magnitudes as a function of altitude and local time, showing how the irregularities in the midday and afternoon taper off in magnitude with altitude. The amplified irregularities due to RT growth in the evening and post-midnight times are apparent here as well. As the solar cycle advances toward its peak, the amplified irregularities extend to larger altitudes as freely convecting plumes develop and rise to greater heights.

Finally, Fig. 9 gives the climatology of the rms density-irregularity magnitudes as a function of latitude and local time. To be able to combine observations from different longitudes, the latitude parameter that is used is the latitudinal distance from the geomagnetic equator at the satellite longitude. We expect that the geomagnetic equator is the center point of organization of much of the plasma structure, so by...
Fig. 7. Climatology of the ratio of rms density irregularity to ambient plasma density as a function of longitude and local time (LT), integrated over altitudes and averaged over seasons from (Northern Hemisphere) summer 2008 (upper left) to spring equinox 2012 (lower right). The scale of this dimensionless quantity (0 to 0.05) was chosen to show the daytime irregularities.

using the distance from that point, we may put equivalent points at different longitudes on an equal footing to be combined in one picture. Note, however, that offset between the geographic and geomagnetic equators and the low inclination of the C/NOFS orbit means that the north side of the geomagnetic equator is not sampled to distances as great as the south side is. There is no consistent seasonal pattern here. There are lobes in the daytime irregularities that appear on either side of the equator at different times, with a reversed pattern the next year. One consistent feature is the lozenge shape of the amplified irregularities in the evening, due to the rising of the plumes, and thus the spreading out in latitude of their motion.

4 Seeding

We next use the low-level density irregularities observed just before sunset as the initial perturbation in a calculation of the development of the RT instability in the evening time sector. To follow the development we use the three-dimensional, fully nonlinear model of Retterer (2010a). This model assumes that the ions are inertia-less to solve the momentum equation (an assumption which is appropriate for F layer heights). Under this approximation, the ion drifts at an instant are specified by the instantaneous accelerations due to the electric field and other forces affecting the ions and electrons (modified by the mobilities that depend on the collision and gyrofrequencies). The electric field, being the solution of an elliptic equation with coefficients dependent on the conductivities and external driving terms (like gravity and
the neutral winds), in turn is dependent on the variations of the plasma density, so the initial condition can be specified solely in terms of the plasma density. For this we use the C/NOFS observations to describe the longitudinal (east–west along the geomagnetic equator) variation of the density, using specific cases from particular times and dates since each is unique. Because of the satellite’s low-inclination orbit, its motion allows it to sample the longitudinal variation of the plasma density; because it moves quickly compared to the evolution timescales of the plasma, its sequential sampling through a time interval can be taken as a snapshot of the variation through the range of longitudes it visits through that time. With the orbital speeds of low earth orbit (8 km s\(^{-1}\)), the 1 Hz sampling of the PLP instrument is just about right for plume models of this kind, which typically have grid spacings of 4–8 km.

To fully specify the plasma perturbation for the simulation’s initial condition, however, we need to specify the variation of the perturbation with altitude and latitude as well. A linearization of the continuity equation shows that when the perturbations are small, the altitude variation is proportional to the altitude gradient of the ambient density. (This means, for example, that the perturbations are small near the F layer peak because any displacement of the layer brings plasma of nearly the same density to that height.) For the latitudinal variation, we appeal to our climatology and employ a Gaussian distribution over latitudes about the geomagnetic equator, with a half-width of 5°. Thus, we first map out the perturbations in the equatorial plane, assuming that the dependence on altitude and longitude can be regarded as separable. The longitudinal variation is provided by the satellite observations while the variation in altitude is specified by the
Fig. 9. Climatology of plasma density irregularities as a function of latitudinal distance from the geomagnetic equator and local time (LT), integrated over all longitudes and averaged over seasons from (Northern Hemisphere) summer 2008 (upper left) to spring equinox 2012 (lower right). The density scale is logarithmic, ranging from $10^2$ to $10^5$ cm$^{-3}$ for the density irregularities.

The variation of the absolute value of the derivative of the density profile with respect to altitude in the equatorial plane, normalized by the magnitude of the derivative at the satellite observation point. Then, the variation along the length of each flux tube crossing the equatorial plane is described by the Gaussian function of latitude, because we expect the fast flux-tube-aligned parallel transport to maintain correlation of densities at points along each flux tube.

We simulate the ionospheric plasma on days 281 and 282 of 2010, near the longitude 280°, using the ambient model of PB-MOD (Retterer, 2005) as a driver. The simulation was performed with a $256 \times 256$ grid in the equatorial plane, thus with grid cells about 4 km by 4 km, with 201 points describing the variation of the density along the magnetic field line in each flux tube intersecting that plane. The density variations would be too small to discern if the total density were plotted, we plot here instead the deviation from the mean density at each altitude, which allows the longitudinal structure of the small irregularities and their evolution to be seen. The top
The density irregularities remains small, filamentation of the original density structures into shorter-wavelength features is apparent at later times in the simulation. The formation of these roughly 30 km-long irregularities may be caused by collisional shear instabilities driven by the shear in the neutral wind and the difference between the wind and the ion drifts (Hysell and Kudeki, 2004). (Note that although the observations in Fig. 2 show amplified irregularities later in the evening, the satellite has moved into a completely different longitude sector by then, with different stability conditions.)

The results of the simulation of the structure on the unstable day 281, 2010, are shown in Figs. 12 and 13. Figure 12 shows snapshots of the density perturbations calculated as done for Fig. 11. Note the rapid increase in the range of the density variations (about a factor of ten from the initial range) as the plume develops around 25 UT. These density variations become strong enough to view in a plot of the total density (Fig. 13).

The spectrum of density irregularities in the simulations can be extrapolated into the scale range relevant to interaction with propagating radio waves to estimate the strength of scintillation that they can cause (Retterer, 2010b). We do not expect strong scintillation from these daytime irregularities because daytime conductivities prevent the occurrence of the Rayleigh–Taylor instability, whose nonlinear development leads to the cascade that strengthens the turbulence at the shorter wavelengths ($\lambda \approx 1$ km) that interact with radio propagation at relevant frequencies. Weak scintillation ($S_1 < 0.4$) during the daytime is in fact observed at the Ancon, Peru site of the US Air Force SCINDA (Scintillation Decision Aid) network (Groves et al., 1997) on the days presented here. Daytime F region irregularities that produce spread-F-like effects have also been observed by radar and reported by Woodman et al. (1985), Chau and Woodman (2001), and Shume et al. (2013).

5 Conclusions

In this paper we have shown how important it is to know the initial perturbation of the ionospheric plasma as well as the RT growth rate to model the development of equatorial plasma plumes and the turbulence embedded within them. There are many possibilities for the source of these initial perturbations, traveling ionospheric disturbances driven by atmospheric gravity waves being one of them. But systematic knowledge of the daytime irregularities has been limited, so we appealed to observations of densities by the PLP instrument on the C/NOFS satellite to provide a climatological description of them.

Using the data, we calculated the mean values of the density irregularities as a function of altitude, latitude, longitude, local time, season, and phase of solar cycle through the C/NOFS era: so far, 2008–2012. The next step in the analysis will be to calculate occurrence probabilities because...
the case studies we presented showed that the daytime irregularities can be quite variable, being stronger one day than the next. (We have hesitated to do this so far because unevenness in sampling is a more difficult problem for occurrence-percentage calculations than it is for calculations of mean values.)

Our analysis of the data showed a nearly ubiquitous presence of density fluctuations in the low-latitude ionosphere with sufficient amplitude to form reasonable seeds for the formation of equatorial plumes as these irregularities get amplified by the Rayleigh–Taylor instability. We used the irregularities observed on two days to seed a plume model to demonstrate the nature and properties of the mature irregularities that can form from them, on both stable and RT-unstable days – the spacings of the plumes, etc. The next step here would be to perform these calculations for a large number of days to get a statistical characterization of such parameters to compare to the observational database. The near ubiquity of the density fluctuations may explain the success of calculations of the probability of equatorial plasma bubbles within the DMSP (Defense Meteorological Satellite Program) satellite data set using bubble simulations employing uniform seeding with longitude and season (Retterer and Gentile, 2009).

We isolated relatively small (shorter than 300 km) wavelengths in our separation of finer scales from ambient scales (to better characterize the local structure of the irregularities).

It has been suggested (Tsunoda, 2005) that very long wavelength waves are responsible for the seeding, and indeed it would be possible to extend the present study to look for these. In fact, the hints of coupling with the lower atmosphere within the data suggest that it would be interesting to look for planetary-wave scale variations in the ionospheric density measurements, which might help explain some of the long-term variations in scintillation occurrence.

It has recently come to be appreciated that because the spatial filtering of the upward propagating gravity waves (Fritts and Vadas, 2008) limits the spectrum of gravity waves that reach ionospheric heights to those of 100 km wavelengths or longer, it is practical to conduct a simulation of the whole atmosphere (Fuller-Rowell et al., 2008) with reasonably sized grids that permits these waves to affect the ionosphere. Using these gravity waves to produce the ionospheric disturbances that can be amplified by plasma instabilities in a plume simulation (Retterer et al., 2014) removes the need to specify an artificially imposed initial perturbation altogether. It remains to be seen whether the variability in irregularity occurrence produced by this lower-atmospheric driving is sufficient to explain the observed day-to-day variability of scintillation occurrence.
Fig. 12. Snapshots of the density perturbations taken at a sequence of times in a simulation in the plane of the geomagnetic equator as a function of longitude and apex altitude, beginning with the density perturbations observed by the C/NOFS satellite around 23.8 UT on day 281 of 2010 near magnetic longitude 280°. Note the changing scale from snapshot to snapshot for the density perturbations (permitted to vary to capture all the variations of the density); the magnitude of the limits of the range of density perturbations (in number per cm³) spanned by the color bar are indicated in each panel. In this case, plumes do develop.

Fig. 13. Snapshots of the density taken at a sequence of times in a simulation in the plane of the geomagnetic equator as a function of longitude and apex altitude, beginning at the time of the density perturbations observed by the C/NOFS satellite around 23.8 UT on day 281 of 2010 near magnetic longitude 280°. (These density snapshots correspond to the density perturbations in Fig. 12.)
Acknowledgements. We thank the C/NOFS instrument teams for making their data available for this study. Support for this work was provided by AFRL and NASA.

Topical Editor F. Rodrigues thanks E. Shume and one anonymous referee for their help in evaluating this paper.

References


