Role of gravity wavelike seed perturbations on the triggering of ESF – a case study from unique dayglow observations

V. Sreeja, C. Vineeth, Tarun Kumar Pant, Sudha Ravindran, and R. Sridharan

Space Physics Laboratory, Vikram Sarabhai Space Centre, Trivandrum, 695022, India

Received: 12 November 2008 – Accepted: 4 December 2008 – Published: 19 January 2009

Abstract. First observational evidence, from the Indian longitudes, for the presence of gravity wavelike perturbations with periods of 20–30 min, acting as probable seeds for Equatorial Spread F (ESF) irregularities is described. The study is based on the daytime optical measurements of the mesopause temperature and the intensity of the thermospheric O(1D) 630.0 nm dayglow emissions using the unique MultiWavelength Dayglow PhotoMeter from Trivandrum (8.5°N; 77°E; dip lat ~0.5°N), a dip equatorial station. Measurements during the equinoctial months of a solar maximum (2001) and a solar minimum year (2006) have been used in this study. It is shown that under identical background ionospheric conditions within a solar epoch, the power of the gravity waves have a deterministic role in the generation of ESF. The mesopause temperature simultaneously observed, indicate that possible source regions for these perturbations lie in the lower atmosphere.

Keywords. Ionosphere (Equatorial ionosphere; Ionospheric irregularities; Wave propagation)

1 Introduction

The nighttime phenomenon of Equatorial Spread F (ESF) is characterized by plasma density irregularities spanning a wide range of scale sizes. ESF manifests itself as bottomside spread in the ionograms, bite outs in the thermospheric airglow, plumes in the UHF/VHF and HF radars and scintillations of VHF and UHF frequencies. The presence of ESF irregularities has significant influence even at microwave frequencies, which has implications in satellite communication, navigation and geodetic systems. The occurrence of ESF varies from day to day, with season, geographical location and solar and geomagnetic activity. These aspects of ESF have been studied extensively over the past several years (Abdu, 2001; Huang et al., 2002, and the references therein).

The primary generation mechanism for the ESF irregularities is the Collisional Rayleigh-Taylor (CRT) instability operating in the bottomside of the F-region (Haerendel, 1973; Hysell and Kudeki, 2004), followed by a hierarchy of other instabilities. The instability is excited where the electron density gradient steepens after local sunset due to the rapid ion-electron recombination effects, and the ambient $E \times B$ uplifting of the base of the F-region due to the F-region dynamo generated electric fields (Rishbeth, 1981; Farley et al., 1986). The irregularities thus generated rise non-linearly through the F-region peak and manifest as ESF. The local growth rate of the R-T instability is given by the equation (Sekar and Raghavarao, 1987):

$$\gamma = \frac{1}{L} \left[ \frac{g}{\nu_{in}} + \left[ \frac{E_X}{B} \right] + \left[ \frac{W_X}{W} \left( \frac{\nu_{in}}{\Omega_i} \right) \right] - W_Z \right]$$

where $W_X$ and $W_Z$ are the zonal and vertical winds respectively, $E_X$ – zonal electric field in the F-region, $L$-plasma scale length, $\nu_{in}$ – ion-neutral collision frequency, $B$ – geomagnetic field and $\Omega_i$ – ion gyro- frequency.

The primary requisites for the generation of the irregularities is the lifting of the F-region to a region of lower $\nu_{in}$ and the presence of a seed perturbation at the altitude of the steep electron density gradient. It had been shown that for a realistic estimation of the growth rate one should also take into account the flux tube integrated ionospheric conductivity extending from the E-region to the F-region (Sultan, 1996). The growth rate was also seen to be controlled by the neutral winds (zonal and vertical), with an opposing gradient enhancing the growth rate (Raghavarao and Gupta, 1987). Maruyama (1988) and Mendillo et al. (1992) studied the effects of the meridional winds on the growth rate and showed that a poleward wind would inhibit ESF by pushing the ionization along the field lines to the E-region and thereby load the F-region dynamo. However, the polarity (equatorward/poleward) and magnitude of the meridional winds have been shown to play a significant role only when...
the maximum height of the post sunset F-layer (h′F) is below a critical height. Above this height, the polarity of the winds does not have a direct control on the generation of ESF (Devasia et al., 2002). Thus, the occurrence/non occurrence of ESF is essentially controlled by a large number of factors, thereby making its prediction an enigmatic problem till this date.

As on date, the statistical occurrence pattern of ESF is well understood (Aarons, 1993). The theoretical foundation based on the R-T instability mechanism is both widely accepted and used successfully in nonlinear numerical simulations (Ossakow and Chaturvedi, 1978; Sekar and Raghava Rao, 1997). One of the unaddressed problems is the day-to-day variability of ESF (Kelley, 1989). To address this problem, the approach all along had been to examine the various physical processes described by the various terms in the R-T growth rate equation. In the earlier studies, it has inherently been assumed that the seed perturbation is omnipresent.

In the simulations of the nonlinear evolution of the R-T instability under different conditions by Huang and Kelley (1996, and the references therein), gravity waves were considered as seed perturbations for initiating the ESF. They studied the evolution of ESF for different parameters of the seeding gravity waves like amplitude, wavelength, timing and also its propagation direction. The first in situ evidence for the presence of wavelike ion density structures that could serve as a seed for the development of plasma bubbles, was shown by Singh et al. (1997) using the data from the Atmosphere Explorer E satellite. McClure et al. (1998) used a comprehensive analysis to propose that, given favorable background conditions, the presence/absence of gravity waves would determine if ESF would occur or not on a given night. In this paper, we report the first observations on the role of gravity wave like perturbation of 20–30 min periodicity in the generation of ESF. The database used is the daytime thermospheric intensity and the mesopause temperature.

2 Experimental database and method of analysis

The daytime airglow intensity measurements were made using the Multi Wavelength Dayglow PhotoMeter (MWDPM) on two rotational lines at 731.6 and 740.2 nm in the OH Meinel (8-3) band and at O(3D) 630.0 nm from the magnetic equatorial location of Trivandrum (8.5° N; 77° E; dip lat ~0.5° N) in India. The main objectives had been to detect the presence of wave like perturbations well before the occurrence of ESF and to identify the possible source regions through multiple airglow emissions originating at different altitude regions. The measurements for the zenithal sky were made between 08:00 h and local sunset (18:00 h). The thermospheric O(3D) 630 nm emissions are typically centered at ~220 km (Hays et al., 1988). The daytime OH emission intensity measurements were used to estimate the mesopause temperature. These mesopause temperatures have already been compared and validated (Vineeth et al., 2005). Trivandrum, being a tropical station, the dayglow measurements are largely restricted to the clear sky conditions. With the onset of the Indian monsoon, the airglow data collection becomes rather difficult. In this context, uninterrupted data were available only in the northern hemispheric winter/equinoctial months (December–March). The post sunset ionospheric F-layer bottom height (h′F) at every 15 min interval was obtained using a collocated Ionosonde.

The equinoctial months of the solar maximum year are characterized by the occurrence of ESF on almost all the days in the Indian longitudes, whereas during the solar minimum year, its occurrence is significantly less. ESF is a multi-dimensional problem and therefore the effect of seed perturbations can be studied only if we have days of ESF and non-ESF under nearly identical ionospheric conditions with nearly the same base height. It’s very difficult to get a large number of days satisfying this condition, which limits the database used in this study. Nevertheless, the relevance and usefulness of the results obtained from this case study, with this limited database, makes them important and unique.

3 Observations

In order to find out the dominant periods present, the “Morlet” wavelet analysis (Torrence and Compo, 1998) has been performed on each time series of the mesopause temperature and the ratio I′/I, where I’ are the perturbations and I is the background 630 nm emission intensities. Emphasis has been given to periodicities less than 2 h. Figure 1a and b represents the periodogram of the ratio I′/I during three days of the solar maximum year of 2001 and two days of the solar minimum year of 2006 respectively. The white solid contours denote the regions of the wavelet spectrum above the 95% confidence level. The top and the middle panels in Fig. 1a are for the two ESF days (13 and 14 March) and the bottom panel (4 March) is for a non-ESF day. In Fig. 1b, the top panel represents an ESF day (3 March) and the bottom panel (9 March) a non-ESF day. The colorbar indicates the amplitude of the various periodicities. The maximum post sunset height of the base of the F-layer (h′F) and the onset time of ESF is shown in parenthesis for all the days. The background conditions on the three days during the solar maximum year and on the two days during the solar minimum year are nearly the same, as indicated by the h′F. Having ascertained this aspect, the other two parameters, which could play a role in the day-to-day variability of ESF occurrence, are the polarity of the meridional winds and the presence/absence of the seed perturbation. Manju et al. (2007) has shown that, during the equinoctial months, the threshold height below which the polarity of the meridional winds is important for the generation of ESF irregularities is ~314 km for the solar maximum year and ~225 km for the solar minimum year. In our case, during both the solar maximum as well as the minimum year, h′F...
had been much above the threshold height, which means that
polarity of the meridional winds would not have any role. So
the occurrence/non-occurrence of ESF can now be construed
to be solely determined by the presence/absence of the seed
perturbation.

Figure 1a, representing three days of the solar maximum
year, clearly shows a dominant periodicity of 20–30 min in
the late evening hours. The amplitude of this periodicity was
∼2 on the ESF days, but on the non-ESF day, the amplitude
was ∼0.1 (20 times smaller as compared to that on the ESF
days). In Fig. 1b, the periodicity of 20–30 min is seen on the
ESF day, with amplitude of ∼0.1. On the non-ESF day, the
amplitude was only ∼0.05 (a factor of 2 smaller compared
to the ESF day). A comparison between the amplitude of the
20–30 min periodicity on the two ESF days during the sol-
lar maximum and one day of the solar minimum year shows
that it is higher by a factor of ∼20 during the solar maximum
year. It is to be noted that during two days of the solar max-
imum year presented here, even an amplitude of ∼0.1 is not
sufficient to trigger ESF. Whereas, during one day of the sol-
lar minimum year, an amplitude of ∼0.1 was found to be suf-
cient to trigger ESF, implying that the power of the waves
required for the irregularity generation shows considerable
variation with the solar activity. This point is discussed later.

When it comes to the source region of the perturbation,
especially while ascertaining whether they are insitu gen-
erated or propagated from below, simultaneous monitoring
of parameters at different altitudes becomes handy. The

mesopause temperature data has also been subjected to sim-
ilar analysis in order to check whether waves/perturbations
similar to the 630 nm dayglow intensity were present in it or
not.

The periodogram of the mesopause temperature data for
the same days in the solar maximum year of 2001 and the
minimum year of 2006 is shown in Fig. 2a and b, respec-
tively. The white solid contours denote the regions of the
wavelet spectrum above the 95% confidence level. The top
panels in Fig. 2a represents two ESF days and the bottom
panel a non-ESF day. The colorbar indicates the amplitudes
of the various periodicities in degree Kelvin (K). A domi-
nant periodicity of 20–30 min is seen to get enhanced dur-
ing the evening hours on the two ESF days with an ampli-
tude of ∼250 K, whereas such a periodicity is conspicuously
absent on the non-ESF day. In Fig. 2b, the top panel rep-
resents an ESF day and the bottom panel a non-ESF day.

The mesopause temperature in 2006 showed the presence of
higher periodicities as well. The 20–30 min periodicity al-
lant periodicity of 20–30 min is seen to get enhanced dur-
ing the evening hours on the two ESF days with an ampli-
tude of ∼250 K, whereas such a periodicity is conspicuously
absent on the non-ESF day. In Fig. 2b, the top panel rep-
resents an ESF day and the bottom panel a non-ESF day.

The most important outcome of the present case study,
even with the limited database, is that, given identical iono-
spheric conditions within a solar epoch, the power of the
gravity wave like perturbations seems to control the trigger-
ing of ESF. It also provides the first observational evidence
for the presence of gravity wave like perturbations, serving
as seeds for ESF. The presence of such perturbations simul-
taneously in the mesopause temperature and also in the in-
tensity of the thermospheric 630 nm emissions suggest that
they could have originated lower below and propagated to
the thermosphere. It is also seen that the power of the seeder
perturbation has a control on the ESF irregularity generation,
while bringing out distinctly different requirements in terms
of amplitude and power for solar maximum and minimum

Figure 1a. Wavelet periodogram of the intensity of the thermospheric
630 nm airglow emissions on ESF days (top and middle panels) and
a non-ESF day (bottom panel) during the solar maximum year of
2001. The white solid contours denote the regions of the wavelet
spectrum above the 95% confidence level. The numbers in the
parenthesis indicates the maximum post sunset $h^{' } F$ and the onset
time of ESF.

Figure 1b. Same as in Fig. 1a, but for the solar minimum year of 2006.
4 Discussion

The nature of the seed perturbation for ESF is widely believed to be gravity waves (Kelley et al., 1981), but their precise sources, variabilities therein and their deterministic role is still unknown and least explored. Hence, the unresolved questions related to the seed perturbation could be stated as: (1) possibility of insitu generation of the gravity waves through (a) motion of the solar terminator across the lower thermosphere, (b) evening F-layer rise permitting zonal wind acceleration and associated strong vertical shear (Anderson et al., 1982); and (c) deposition of momentum at z∼180 km from the dissipation of gravity waves from the lower atmosphere (Vadas and Fritts, 2006), (2) remote sources such as tropospheric convection activity that could produce upward propagating gravity waves (Rottger, 1981; Larsen and Swartz, 1982; Lane et al., 2003).

Gravity waves produce turbulence by depositing their energy and momentum through breaking, thereby influencing the general circulation, thermal regime and composition of the middle atmosphere (Fritts and Alexander, 2003). Though the background atmosphere would significantly modulate the gravity wave spectrum at any level, short period gravity waves generated by convective storms have the potential to influence regions extending even up to mesospheric altitudes as had been demonstrated through numerical simulations (Alexander et al., 1995). Gravity waves with periods of 30 to 90 min have been detected in the mid-latitude thermosphere using the MU radar (Oliver et al., 1994). Vadas (2007) showed that dissipative filtering is the reason that only gravity waves with periods less than an hour can propagate well into the thermosphere, consistent with the MU radar measurements. Another possibility could be the interaction of gravity waves and the tides thus influencing the F-region through electrodynamical processes. Upward propagating tides over the dip equator are known to significantly influence the background wind in the lower ionosphere, which in turn through the E-region dynamo leads to the generation of the global electric field. The interaction of the tides with the gravity waves are known to modify the amplitude and phase of the tides (Hagan, 2000), which would cause global electric field modulations and vice versa. Also, any perturbation in the dynamo region away from the equator would get promptly imprinted in the F-region above along the magnetic field lines, thus providing one more means of imprinting the perturbations as seen in the 630 nm dayglow.

When it comes to the significant differences in the power of the seed perturbations, especially during different solar activity levels, it is as per the expectations based on the background ionosphere-thermosphere conditions which in turn have large solar cycle dependence. Both the parameters viz, the 630 nm dayglow and the mesopause temperatures are confined to their respective altitude regions of ∼220 km and ∼87 km. As a consequence of significant increase in both the neutral and plasma densities in a given altitude region with solar activity, the stabilizing factors mainly the ion-neutral collision frequency dominates the instability process calling for larger amplitude perturbations so as to trigger the instability. Our present results clearly bring out that perturbation amplitudes that could trigger the instability during low sunspot years are grossly inadequate during high sunspot years and even amplitudes larger by a factor of two would not suffice. Typically, one encounters amplitudes larger by a factor of 20 during solar maximum conditions. Such large amplitude requirements in the initial perturbation are suggested to be mainly due to the large increase one encounters in the background ionospheric/thermospheric densities.

It can be noted from Figs. 1 and 2 that there is an increased gravity wave activity at the mesospheric altitudes around 12:00–13:00 IST, but simultaneously, there is no gravity wave activity at the thermospheric region. Mesopause is the region which exhibits significant variability not just
temporally but spatially also. Theoretically a range of periodicities can be present in the mesosphere and below, however only certain wave modes can reach the mesospheric and thermospheric heights as a result of wave mean-flow, wave-wave and wave-tidal interactions during the upward propagation of these waves. In this context, the appearance of increased wave-activity in the mesopause around 12:00–13:00 IST can be construed as the result of wave-tidal interactions, as some of the tidal modes maximize in the lower MLTI region around this time.

In a study using rocket based chemical and vapour release experiments, Larsen et al. (2005) clearly show that very large wind shears exist in the lower thermosphere which may show significant temporal variability. As is known, these strong wind shears can lead to the dissipation of the upward propagating waves. Therefore, there can be times when waves seen in the mesopause region may not be seen in the thermosphere. Though the database used in this study is limited in order to unambiguously discuss these aspects, the simultaneous presence of the waves in the mesopause and the thermosphere is a strong indicator of the wave being generated in the lower atmosphere.

The most critical piece of our observation is that the gravity wavelike perturbations in the OH layer occur at almost the same time as in the 630 nm dayglow intensity. This observation matches well with the theoretical studies (Vadas and Fritts, 2006, and the references therein), indicating that the waves could be the ones propagating from the lower atmosphere. The simultaneous observation of the gravity waves in the OH layer as well as in the 630 nm dayglow is as expected from the theory, since the gravity waves that are observed in the OH layer have vertical group velocities of only 15 m/s, while those which are able to propagate to the bottomside of the F layer have vertical group velocities of 2 to 3 times that amount, 30–45 m/s. Therefore, these slower gravity waves propagate from 10 to 90 km in about the same time as the faster waves propagate from 10 to 200 km.

The modulation of the base of the F layer preceding the ESF plume generation as observed over ALTAIR radar by Tsunoda and White (1981) and gravity waves modulation of the large scale plumes by Kelley et al. (1981) substantiate the potential role of gravity wave like perturbations in the generation and/or modulation of ESF bubbles structures. Recent results from the Indian longitudes by Sekar et al. (2007) based on nighttime photometric and coherent VHF backscatter radar add credence to the above results. Huang et al. (1993) showed for the first time through numerical simulation that both atmospheric gravity wave seeding and explosive mode coupling of the R-T instability could be responsible for the generation of intense ESF. It was also shown that the instability initiated only by large-amplitude and long wavelength gravity waves can grow to substantial amplitudes. On the other hand, Thampi et al. (2006), in their attempt to predict the occurrence of ESF based on background ionospheric conditions were successful to a level of >95%, by defining a critical ionospheric parameter. They had indicated that 100% deterministic prediction would be possible if additional inputs on the seed perturbations were also available and this would be extremely crucial for operational forecasting purposes.

Our unique observations through dayglow measurements, well before the occurrence of ESF, reported herein substantiate that the amplitude of the seed perturbation has a significant role in deciding the day-to-day variability of ESF and the subsequent nonlinear evolution of the equatorial plasma bubbles.

5 Conclusions

The first observational evidence for the presence of gravity wave like perturbation acting as seed for ESF irregularities is presented through 630.0 nm dayglow variability and also the daytime mesopause temperature in the evening hours. The power of the waves shows considerable day-to-day variability and also significant solar activity dependence effectively having a control on ESF generation, which itself is determined by the combined effects of the prevailing neutral dynamical and electrodynamical conditions. Under nearly identical ionospheric and thermospheric conditions, the presence or absence of ESF has been conclusively shown to be due to the variability in the seed perturbations presumably originating lower below in the pre dusk hours, well before the actual event. The present study takes us one step ahead towards solving the enigmatic problem of the day-to-day variability of ESF.

Acknowledgements. This work was supported by Department of Space, Government of India. One of the authors, V. Sreeja, gratefully acknowledges the financial assistance provided by the Indian Space Research Organization through Research Fellowship.

Topical Editor M. Pinnock thanks one anonymous referee for her/his help in evaluating this paper.

References


