Chapter 1 – MORPHOLOGY AND DYNAMICS

1.1 GENERAL INTRODUCTION

1.1.1 The Ionosphere

(Stefan Hawlitschka, FGAN Germany)

The Ionosphere is a constituent of the atmosphere where at altitudes in excess of 50 km the earth's atmosphere becomes increasingly rarefied and can support free electrons for short periods. This region has the properties of a gas and a plasma which exhibit some interesting effects. Neutral atoms and molecules, such as N_2, N, NO, O_2, O, H and He are ionised by impact of photons of the *extreme ultraviolet* (EUV) and *X-ray* radiation, emanating from the sun. Different mechanisms are predominant at specific altitudes; at the outer reaches of the atmosphere the solar radiation amounts to some 1370 Watts per square metre but the density of atoms is low resulting in a small amount of ionisation. At lower altitudes atoms are more numerous increasing the rate of ionisation but at the same time allowing increased rate of recombination. A further mechanism is the absorption of the incident energy which is a function of the gas density and composition; causing attenuation to increase at lower altitudes. The resultant effect is horizontal stratification denoted as the *D*-, *E*-, and *F-layer* of the ionosphere (Figure 1-1). The *C-layer* also exists generated by cosmic rays but plays a minor role for our purposes.

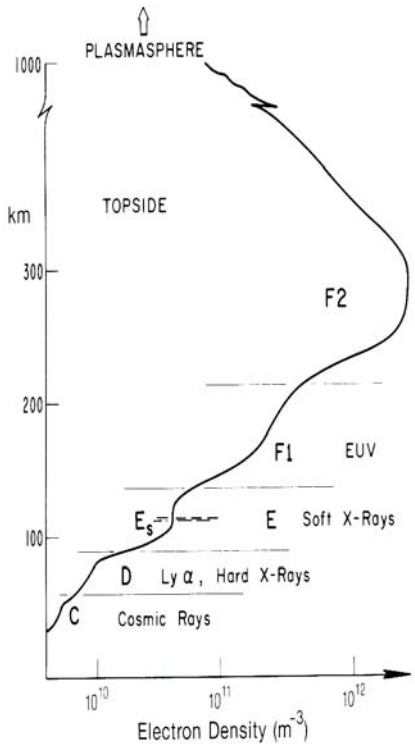

Figure 1-1: Ionosphere Structure on a Summer Day in a Middle Latitude, and the Main Bands of Solar and Cosmic Ionizing Radiation – from Davies [4].

Electromagnetic *high frequency* (HF) waves propagating through the ionosphere may experience reflection, refraction, attenuation and modification of the plane of polarisation. Reflection of HF waves occurs depending on their frequency, the degree of ionisation of the layers of the ionosphere, and on the angle of propagation. The smaller the elevation of the HF wave the higher is the maximum frequency that

can be reflected. As characteristic parameter the *critical frequency* is specified which denotes the highest frequency that is being reflected at vertical incidence. It is subject to high variability and depends on the production rate of free electrons, the loss rate which is mainly due to recombination and the transport of the charged particles.

The ionosphere is not only controlled by the radiation of the sun and the coupling with the neutral atmosphere, but also affected by the solar wind, the earth's magnetic field as well as the interplanetary magnetic field and other phenomena. The resulting high variability of the density of charged particles impedes a detailed forecast of the actual ionosphere. Modelling is therefore mostly constrained to predict in general terms only, average states of the ionosphere. The *International Reference Ionosphere* (IRI) is a climatological model describing the ionospheric parameters for quiet conditions, which is subject to constant review and has recently been modified to cover disturbed conditions. It is recognised that describing the necessary parameters to characterise a storm is not a trivial exercise and knowledge of the heterogeneities of the ionosphere an important and challenging task.

1.1.1.1 The Layers of the Ionosphere

The primary process to contribute to the production of free charge carriers is *photoionization* of gases by the solar extreme ultraviolet (EUV) and X-ray radiation. At polar latitudes, precipitation of energetic particles from the solar wind also plays an essential role in ionization production. Simultaneously a *recombination process* counters the ionisation Therefore at night, the lower layers D and E, which lie in a comparatively dense atmosphere, disappear while the F-layer persists because of its greater height where the atmosphere exhibits a smaller density and thus the collision frequency of the particles is low. At dawn the production of charged particles is restarted and the free electron density N_e increases until the production rate equals the recombination rate.

The lowest noticeable ionospheric region (or layer) is the D region between 50 km and 90 km above ground. The recombination rate is high in this region and as a result the free electron and ion densities are small. Electron density is not high enough for HF waves to be reflected, but electromagnetic waves may be attenuated. Typical day-time densities of free electrons lie around 10^9 to 10^{10} electrons/m^3. The D region appears sun-synchronously and nearly disappears at night. During a radiative eruption from the sun's surface, known as a solar flare, a sudden ionospheric disturbance (SID) can occur. A great enhancement of solar X-rays which penetrate deeply into the atmosphere causes much higher ionization of the D region. It may occur within a few seconds and result in strong absorption of HF waves and interrupt of communication circuits (Mögel-Dellinger effect). The ionization decreases again and returns to its normal density in about an hour.

The E-layer is located between 90 km and 120 km above ground. The electron densities may exceed 10^{11} electrons/m^3 during the day while at night densities are around 3 x 10^9 electrons/m^3. The daytime electron and ion concentrations are only about 10^{-11} of the neutral gas. Thus the structure of the plasma is strongly influenced by processes in the neutral atmosphere. Slant HF waves up to 10 MHz are reflected by this layer, the daytime critical frequency foE (vertical incidence) is about 4 MHz while during the night foE reaches a minimum of about 0.5 MHz. The E-region exists predominantly in daytime and begins to disappear at dusk due to recombination. The seasonal variations are in phase with solar elevation, but the critical frequency foE varies according to season but is less than 1 MHz. The solar cycle dependent variat$_i$on of foE is about 0.5 MHz.

A sporadic E layer can be observed, which as the name implies is not always present The mechanisms of generation are attributed to thunderstorms and/or wind shears but are not yet fully understood and is not predictable. Electron densities can reach 10^{14} electrons/cubic metre and is significant in that it will reflect waves up to a frequency of approximately 200MHz.

The F-layer is the most dense regular and the most important ionospheric region regarding the influence onto the propagation of electromagnetic waves. Electron densities are of the order of 10^{12} electrons/m^3. The daytime electron and ion concentrations are about 10^{-5} of the neutral gas which again strongly influences the structure of the plasma. But due to the higher electron and ion densities the impact of the magnetic field is stronger than in the E-region. At heights of about 300-400 km the gravitational and the magnetic forces are equal. The F-region extends between 120 km and 1000 km above ground with a maximum of the ionization in the range of 200 km to 400 km. It can split up into the F_1 and F_2 layer. The F_1 layer is observed only during daytime. It is more pronounced during summer than winter, at low sunspot numbers, and during ionospheric storms. It's observed best, when f_oF_2 is low. The seasonal variation of the critical frequency f_oF_1 is less than 1 MHz. The solar cycle dependent variation of f_oF_1 can reach 2 MHz.

The most variable of the regular ionospheric layers is the F_2 layer. Whereas variations of f_oE and f_oF_1 are comparatively regular, i.e., diurnal and seasonal, and in phase with the elevation of the sun, the maximum f_oF_2 at noon maybe twice in winter than in summer. The peak electron density has thus to be four times higher in winter than in summer what is called the *winter anomaly* and occurs in daytime only. The diurnal variation of f_oF_2 is much bigger in winter than in summer. During the maximum of the solar cycle the maximum noon f_oF_2 can reach twice the value during the solar minimum. The F_2 layer is influenced by the geomagnetic field. This can be seen in the fact that the worldwide distribution of the critical frequency is aligned with the magnetic equator.

At heights above about 1000 km the atmosphere is fully ionised and therefore the earth's magnetic field controls the structure of the plasma.

The critical frequencies of all layers grow with increasing sunspot numbers. The variability of the F_2 layer with the 11-year sunspot cycle is the highest among the regular layers. In Figure 1-2, the seasonal as well as the solar cycle variations are seen. The seasonal variation of f_oE and f_oF_1 are in phase with solar elevation whereas f_oF_2 is in antiphase (winter anomaly). The F_1 lar disappears in winter. There is a marked increase of the critical frequencies with the 11-year sunspot cycle [4].

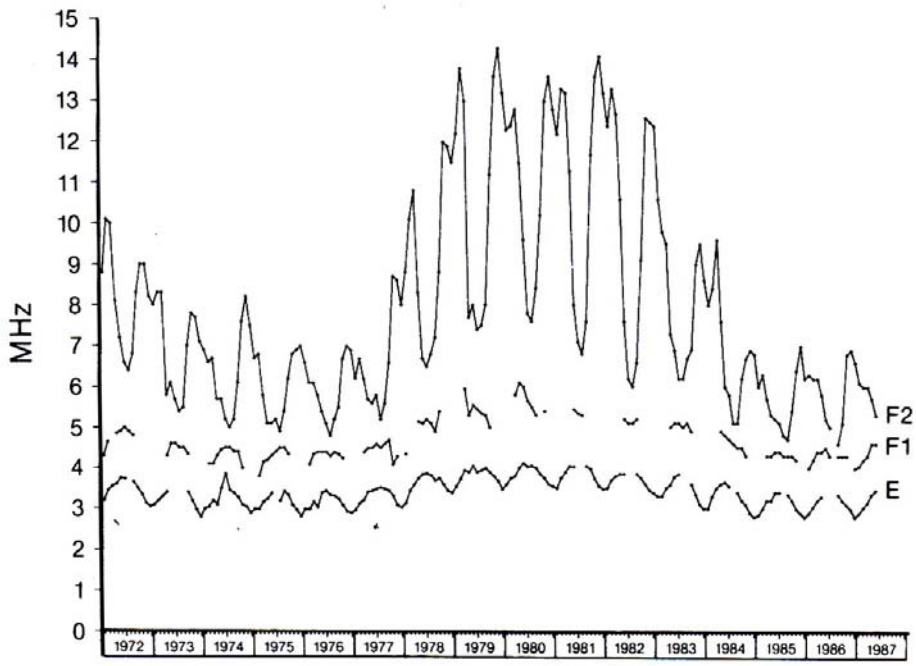

Figure 1-2: Monthly Median Noon Critical Frequencies at Washington DC, from 1972 to 1987 – from Davies [4].

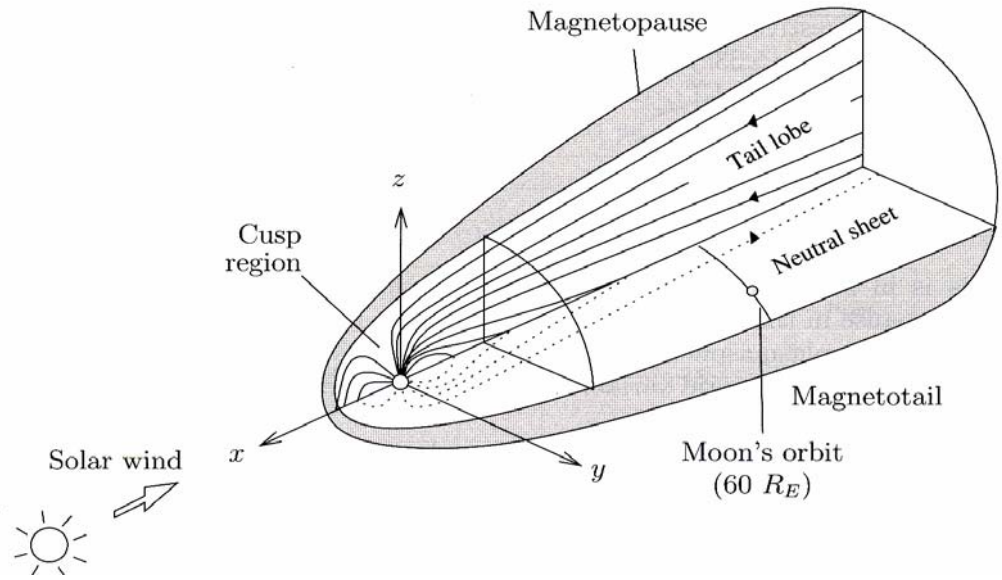

Figure 1-3: General Configuration of the Magnetosphere – from Proelss [12].

The dipolar terrestrial magnetic field is distorted by the interaction with the solar wind and the *frozen-in* interplanetary magnetic field (IMF) that significantly modify the structure of the terrestrial magnetic field. As a result, the earth magnetic field is confined to a finite volume of asymmetric shape. The outer boundary is called magnetopause. On the sunward dayside it assumes an ellipsoidal shape whereas on the opposite side it is characterized by a long cylindrical tail reaching far into space. The cusp region denotes a region of weak magnetic field where the magnetosphere is "open" for the input of the solar particle flux.

1.1.1.2 Disturbances

The ionosphere interacts with solar events, the terrestrial and interplanetary magnetic field as well as with the neutral atmosphere. Solar flares are characterized by short-duration (? 1h) strong enhancement of solar radiation. The radiation flux enhancement in the EUV ranges up to a factor of two and in the X-rays ranges up to a factor of 10^4 and has a significant impact upon the ionosphere. The occurrence frequency of solar flares generally follows the rhythm of the solar cycle, whereby about 5 major events occur every year when the solar activity is low and 70 such events are recorded yearly around the solar maximum.

Ionospheric storms result from very large energy inputs from the sun to the polar atmosphere and the interaction with *geomagnetic storms*. During *coronal* or *solar mass ejections* (CME/SME) a huge amount of solar material is ejected from the sun which propagates through interplanetary space with *frozen-in* magnetic fields forming *magnetic clouds*. When they impinge on the terrestrial magnetic field, initial conditions for the rise of a geomagnetic storm are given, especially when the interplanetary magnetic field is pointing southward (negative B_z). During a solar mass ejection magnetic clouds consisting of plasma are being ejected into space, and strong negative B_z values can occur. The resulting changes of the terrestrial magnetic field are described as *magnetic activity* whose strength is given by the *Dst* (disturbance storm) index or the three hourly local *K*- or the planetary K_p-indices. The terrestrial magnetic field deviates the charged particle flux around the earth. Over the poles the particles may travel along magnetic field lines and penetrate into the *magnetosphere* including the *magnetotail*. An *ionospheric storm* is triggered. One would expect that the input of particles leads to an increase of the ionization of the ionosphere. However at mid-latitudes both an increase as well as a decrease is observed. The former is referred to as a *positive ionospheric storm* while the latter as a *negative ionospheric storm*. A negative storm occurs when the atmospheric (thermospheric) gas composition changes during the storm. The density of atomic oxygen

decreases at F_2 layer heights and thus the oxygen ion production rate decreases too, whereas the N_2 density increases and, as a result, the recombination rate of the respective ions increases. A positive storm is triggered by traveling *atmospheric* disturbances (TAD) at high latitudes. The ionized component of the traveling atmospheric disturbance, the *traveling ionospheric disturbance* (TID) is deviated by the magnetic field to higher heights and with it the F-region is lifted. The production rate of ions and free electrons decreases more slowly with height than the loss coefficient. Thus the densities of charged particles are augmented. Many questions concerning the generation and the propagation of TIDs are not yet fully understood. But it is generally accepted that during high magnetic activity at high latitudes the atmosphere is being heated during particle precipitation (Joule heating). The friction between the drifting charge carriers and the neutral gas particles heats up the thermospheric gases. TADs with an average propagation velocity of 600 m/s (2160 km/h) are generated during *substorms*, the ionized component can be detected at mid-latitudes as large-scale TIDs which arrive at equatorial latitudes after about 4 h. The rise of a substorm during a storm phase is complicated and is connected to currents in the tail of the magnetosphere. The density disturbance propagation towards the equator is called convective disturbance transport. It is usually stronger on the night-side part of the earth because on the day-side winds are predominantly poleward directed.

As stated above a common feature of *large-scale* TIDs that are generated in the auroral zone is that they are the ionospheric signature of internal gravity waves which have been triggered by sudden expansions of the neutral gas density. Large-scale TIDs are characterised by horizontal wavelengths of up to a few thousands kilometer and periods of about half an hour up to about 4 hours. They are connected with winds propagating equatorwards at velocities of 50-200 m/s (180-720 km/h).

Medium-scale atmospheric gravity waves are oscillations of the neutral atmosphere which can be observed by their traces in the ionospheric plasma. MS Gravity waves cause medium-scale traveling ionospheric disturbances (MSTIDs), which are manifested through changes of the local plasma density, plasma temperature and field-aligned plasma velocity. Changes of the plasma density have been observed up to 40% (Bristow and Greenwald 1995). They estimated gravity wave parameters by estimation of oblique backscatter and its variations due to TIDs using a HF Radar. The wave amplitudes are inferred by comparing the HF radar observations with simulations.

1.1.1.3 Spread-F and Sporadic-E

Mid-latitude spread-F is identified in ionograms by multiple and spread F layer traces. These are caused by strong undulations and irregular plasma structuring in the F region electron density (Haldoupis et al. 2003). In the literature several generation mechanisms are described, ranging from the passage of large scale atmospheric gravity waves, electrodynamic forces and large-scale plasma instabilities. Haldoupis et al. (2003) postulated that spread-F can be generated from eastward electric fields and that there is a link between meso-scale spread-F and unstable sporadic-E layers (E_s).They present a mechanism spread-F generation in which the key role is played by polarization electric fields set up within westward-drifting E_s plasma patches and mapping up to the F region.

1.1.2 References

[1] Afraimovich, E. L., Kosogorov, E. A., Lesyuta, O. S., Ushakov, I. I., Yakovets, A. F. (2001): *Geomagnetic Control of the spectrum of travelling ionospheric disturbances based on data from a global GPS Network*. Ann. Geophys. 19, 723-731.

[2] Bristow, W. A. and R. A. Greenwald (1995): *Estimating gravity wave parameters from oblique high-frequency backscatter: Modeling and analysis*. J. Geophys. Res., 100(A3), 3639-3648.

[3] Budden, K. G. (1985): *The propagation of radio waves*. Cambridge University Press, ISBN 0 521 36952 2.

[4] Davies, K. (1990): *Ionospheric Radio*. Peter Peregrinus Ltd, ISBN 0 86341 186 X.

[5] Francis, S. H. (1975): *Global propagation of atmospheric gravity waves : A review*. JASTP 37, 1011-1054.

[6] Haldoupis, C., M. C. Kelley, G. C. Hussey and S. Shalimov (2003): *Role of unstable sporadic-E layers in the generation of mid-latitude spread F*. J. Geophys. Res., 108(A12), 1446, doi:10.1029/ 2003JA009956.

[7] Hawlitschka, S. (2006a): *Travelling ionospheric disturbances (TIDs) and tides observed by a super-resolution HF direction finding system*. JASTP 68 (2006) 568-577.

[8] Hawlitschka, S. (2006b): *Direction finding errors induced by plasmawaves of the ionosphere. Specialists Meeting/Symposium on Characterising the Ionosphere*, 12-16 June 2006, Fairbanks, USA.

[9] Hocke, K., Schlegel, K. (1996): *A review of atmospheric gravity waves and travelling ionospheric disturbances: 1982-1995*. Ann. Geophys. 14, 917-940.

[10] Hunsucker, R. D. (1982): *Atmospheric Gravity Waves Generated in the High-Latitude Ionosphere: A Review*. Reviews of Geophysics and Space Physics, Vol. 20, No. 2, 293-315.

[11] Kalikhman, A. D., (1980): Medium-scale travelling ionospheric disturbances and thermospheric winds in the F-region. JASTP 42, 697-703.

[12] Proelss, G. W. (2004): *Physics of the Earth's Space Environment*. Springer, ISBN 3-540-21426-7.

[13] Samson, J. C., R. A. Greenwald, J. M. Greenwald, J. M. Ruohoniemi, A. Frey and K. B. Baker (1990): *Goose Bay Radar Observations of Earth reflected, Atmospheric Gravity Waves in the high-Latitude Ionosphere*. JGR, Vol. 95, NO. A6, 7693-7709.

[14] Tsybulya, K. and N. Jakowski (2005): *Medium- and small-scale irregularities detected by GPS radio occultation method*. Geophysical Research Letters, 32 (9).

[15] Waldock, J. A., Jones, T. B. (1987): *Source regions of medium-scale travelling disturbances observed at mid-latitudes*. JASTP 61, 629-639.

[16] Yakovets, A. F., Kaliev, M. Z., Vodiannikov, V. V., (1999): *An experimental study of wave packets in travelling ionospheric disturbances*. JASTP 61, 629-639.

1.2 MID-LATITUDE

1.2.1 Neutral Atmosphere

(Ulf-Peter Hoppe, Norwegian Defence Research Establishment (FFI))

The ionosphere is embedded in the neutral atmosphere. Even at the maximum of the F-region, with a typical electron number density of 10^{12} m^{-3}, the neutral number density is about 3×10^{14} m^{-3}. At lower altitudes, in the E- and D-regions, the neutral number density overwhelms the electron number density by as much as 10 orders of magnitude. For this reason, neutral atmosphere processes such as the general circulation, tides, planetary and gravity waves as well as the chemical composition down to the number densities of source constituents like O_3 and NO sometimes play important roles for the characteristics of the ionosphere (e.g., Fuller-Rowell et al., 1996).

The neutral atmosphere is described in a statistical sense by model atmospheres such as the COSPAR International Reference Atmosphere (CIRA-86; Fleming et al., 1988; Fleming et al., 1990). For some altitudes, latitudes, and for some seasons, better statistical datasets are available (e.g., Lübken, 1999; Lübken and von Zahn, 1991). Natural fluctuations in all parameters of the atmosphere make individual observations differ from any of these average descriptions.

A number of recent studies address long-term trends in the upper atmosphere and ionosphere. Rishbeth (1997) pointed out how essential it was to maintain a basic world-wide programme of upper atmospheric and geomagnetic monitoring in order to detect any such changes. Lean (1997) came to a similar conclusion for the sun-earth system. A nearly linear decrease in the altitude of the F_2-layer peak over 39 years at Sodankylä, Finland, is indicative of a greenhouse cooling of the upper atmosphere and ionosphere, as is to be expected together with a greenhouse warming in the troposphere (Ulich and Turunen, 1997; Ulich et al., 2003). Cannon et al. (2004) have shown that a 16 km drop in $h(m)F_2$ will introduce ground errors of similar to 100 km. They point out an urgent need to reach a consensus understanding of long-term ionospheric trends. A reassessment of the CCIR and URSI worldwide maps of ionospheric characteristics is likely necessary. Hall and Cannon (2002) suggest that the greenhouse cooling at high altitudes is at least partly causal for the observed negative trend in the critical frequency of the F_2-layer over Tromsø. Bremer et al. (2004) argue that these findings are not global in nature, but limited to individual stations. These authors point out that in the F_1-layer the derived mean global trend in f_oF_1 is in good agreement with model predictions of an increasing greenhouse effect. The trend results in the D-region derived from ionospheric reflection height and absorption measurements in the LF, MF and HF ranges can at least partly be explained by an increasing atmospheric greenhouse effect.

Sprites and elves are recently discovered manifestations of neutral atmosphere-ionosphere interactions (e.g., Pasko et al., 2002; Moudry et al., 2003; Valdivia, 2003; Tong et al., 2004).

Tidal motions and planetary waves in the thermosphere have significant influence on ionospheric properties (Meyer and Forbes, 1997). For instance, the semi-diurnal and quarterdiurnal tides cause ionisation convergence into layers in the equatorial anomaly region of the lower ionosphere (Lee et al., 2003). Planetary waves propagating upward from the stratosphere and mesosphere influence the E- and F-regions (Lawrence and Jarvis, 2003) by transport and mixing.

The neutral atmosphere being 100 times to 10^{10} times more massive than the ionosphere, the former influences the latter much more than vice versa (e.g., Titheridge, 1995). However, the magnetospherically imposed ion convection pattern actually sets up a neutral convection pattern in the thermosphere which in turn modifies the distribution of neutral tracer constituents (Burns et al., 2004). Aruliah et al. (2004) report that the magnitude of the upper thermosphere neutral wind dynamo field is on average 50% of the magnetospheric electric field.

Some of the well-known F_2-layer anomalies might have their origins in the lower atmosphere. It has been suggested that phenomena such as storms, earthquakes and volcanic explosions may produce F_2-layer signatures (Rishbeth, 2006). Kazimirovsky et al. (2003) have reviewed such possible effects. Pulinets et al. (2006) have published a case study of anomalous variations of the total electron content (TEC) registered over the epicenter of an impending earthquake three days prior to the main earthquake event. Calais and Minster (1998) have demonstrated the GPS capabilities to detect the ionosphere perturbations immediately after an earthquake. Underground nuclear explosions emit infrasonic acoustic waves that propagate even to the ionosphere and may perturb the electron number density by as much as 10% (Krasnov and Drobzheva, 2005). Pokhotelov et al. (1995) have reviewed acoustic-ionosphere coupling in general. Aburjania and Machabeli (1998) have examined the coupling of infrasound waves into the ionosphere and mention earthquake precursors in this context.

Ice particles in the mesosphere, such as in noctilucent clouds (NLC) and Polar Mesosphere Summer Echoes (PMSE) modify the D-region at high latitudes in summer quite dramatically (e.g., Gumbel et al., 2003) – see also the section on PMSE below.

1.2.2 References

Aburjania, G.D., Machabeli, G.Z., Generation of electromagnetic perturbations by acoustic waves in the ionosphere, J. Geophys. Res. 103 (A5), 9441-9447, 1998.

Aruliah, A.L., Griffin, E.M., McWhirter, I., Aylward, A.D., Ford, E.A.K., Charalambous, A., Kosch, M.J., Davis, C.J., Howells, V.S.C., First tristatic studies of meso-scale ion-neutral dynamics and energetics in the high-latitude upper atmosphere using collocated FPIs and EISCAT radar, Geophys. Res. Lett. 31 (3), Art. No. L03802, 2004.

Bremer J., Alfonsi, L., Bencze, P., Lastovicka, J., Mikhailov, A.V., Rogers, N., Long-term trends in the ionosphere and upper atmosphere parameters, Ann. Geophys. 47 (2-3), 1009-1029, 2004.

Cannon, P.S., Rogers, N.C., and Hall, C.M., Trends in critical frequencies and layer heights over Tromsø and their consequential impact for radio system modeling, IRI: Quantifying Ionospheric Variability, Adv. Space Res. 34 (9), 2085-2091, 2004.

Burns, A.G., Wang, W., Killeen, T.L., and Solomon, S.C., A “tongue” of neutral composition, J. Atmos. Solar-Terr. Phys. 66 (15-16), 1457-1468, 2004.

Calais, E., and Minster, J.B., GPS, earthquakes, the ionosphere, and the Space Shuttle, Phys. Earth Planet. Interiors 105 (3-4), 167-181, 1998.

CIRA-86, http://badc.nerc.ac.uk/data/cira/.

Fleming, E.L., Chandra, S., Shoeberl, M.R., and Barnett, J.J., Monthly mean global climatology of temperature, wind, geopotential height and pressure for 0-120 km, National Aeronautics and Space Administration, Technical Memorandum 100697, Washington, D.C., 1988.

Fleming, E.L., Chandra, S., Barnett, J.J., and Corney, M., Zonal mean temperature, pressure, zonal wind, and geopotential height as a function of latitude, Adv. Space Res., 10(12), 11-59, 1990.

FullerRowell, T.J., Codrescu, M.V., Rishbeth, H., Moffett, R.J., Quegan, S., On the seasonal response of the thermosphere and ionosphere to geomagnetic storms, J. Geophys. Res. 101 (A2), 2343-2353, 1996.

Gumbel, J., Siskind, D.E., Witt, G., Torkar, K.M., Friedrich M., Influences of ice particles on the ion chemistry of the polar summer mesosphere, J. Geophys. Res. 108 (D8), Art. No. 8436, 2003.

Hall, C.M. and Cannon P.S., Trends in f_oF_2 above Tromso (69 degrees N 19 degrees E), Geophys. Res. Lett. 29 (23), Art. No. 2128, 2002.

Kazimirovsky, E., Herraiz, M., De la Morena, B.A., Effects on the ionosphere due to phenomena occurring below it, Surveys Geophys. 24 (2), 139-184, 2003.

Krasnov, V.M. and Drobzheva, Y.V., The acoustic field in the ionosphere caused by an underground nuclear explosion, J. Atmos. Solar-Terr. Phys. 67 (10), 913-920, 2005.

Lawrence, A.R. and Jarvis, M.J., Simultaneous observations of planetary waves from 30 to 220 km, J. Atmos. Solar-Terr. Phys 65 (6), 765-777, 2003.

Lean, J., The Sun's variable radiation and its relevance for Earth, Ann. Rev. Astron. Astrophys. 35, 33-67, 1997.

Lee, C.C., Liu, J.Y., Pan, C.J., Hsu, H.H., The intermediate layers and associated tidal motions observed by a digisonde in the equatorial anomaly region, Ann. Geophys. 21 (4), 1039-1045, 2003.

Lübken, F.-J., Thermal structure of the Arctic summer mesosphere, J. Geophys. Res., 104, 9135-9149, 1999.

Lübken, F.-J. and U. von Zahn, Thermal structure of the mesopause region at polar latitudes, J. Geophys. Res. 96, 20, 841-20, 857, 1991.

Pasko, V.P., Stanley, M.A., Mathews, J.D., Inan, U.S., Wood, T.G., Electrical discharge from a thundercloud top to the lower ionosphere, Nature 416 (6877), 152-154, 2002.

Meyer, C.K., Forbes, J.M., Natural oscillations of the ionosphere-thermosphere-mesosphere (ITM) system, J. Atmos. Solar-Terr. Phys. 59 (17), 2185-2202, 1997.

Moudry, D., Stenbaek-Nielsen, H., Sentman, D., Wescott, E., Imaging of elves, halos and sprite initiation at 1 ms time resolution, J. Atmos. Solar-Terr. Phys. 65 (5), 509-518, 2003.

Pokhotelov, O.A., Parrot, M., Fedorov, E.N., Pilipenko, V.A., Surkov, V.V., Gladychev, V.A., Response of the Ionosphere to Natural and Man-Made Acoustic Sources, Ann. Geophys. 13 (11), 1197-1210, 1995.

Pulinets, S.A., Ouzounov, D., Ciraolo, L., Singh, R., Cervone, G., Leyva, A., Dunajecka, M., Karelin, A.V., Boyarchuk, K.A., Kotsarenko, A., Thermal, atmospheric and ionospheric anomalies around the time of the Colima M7.8 earthquake of 21 January 2003, Ann. Geophys. 24 (3), 835-849, 2006.

Rishbeth, H., Long-term changes in the ionosphere, Middle Atmosphere: Changes and Electrodynamics, Advances in Space Research 20 (11), 2149-2155, 1997.

Rishbeth, H., F-region links with the lower atmosphere? J. Atmos.Solar-Terr.Phys. 68 (3-5), 469-478, 2006.

Titheridge, J.E., Winds in the Ionosphere - A Review, J. Atmos. Terr. Phys. 57 (14), 1681-1714, 1995.

Tong, L.Z., Nanbu, K., Hiraki, Y., Fukunishi, H., Particle modeling of the electrical discharge in the upper atmosphere above thundercloud, J. Phys. Soc. Japan 73 (9), 2438-2443, 2004.

Ulich, T. and E. Turunen, Evidence for long-term cooling of the upper atmosphere in ionosonde data, Geophys. Res. Lett. 24 (9), 1103-1106, 1997.

Ulich, T., Cliverd, M.A., Rishbeth, H. Determining the long-term change in the ionosphere. EOS Trans. 84(52), 581-585, 2003.

Valdivia, J.A., Lightning induced optical emissions in the ionosphere, Space Sci. Rev. 107 (1-2), 273-291, 2003.

1.3 HIGH-LATITUDE IONOSPHERE

1.3.1 Overview of Important Effects

(Paul Prikryl, Communications Research Centre Canada)

The high-latitude ionosphere differs significantly from its mid-latitude counterpart. The basic processes of ionization production, loss and transport that shape the ionosphere at mid latitudes also operate at high latitudes thus producing main ionospheric layers as described above. However, given the geometry of the geomagnetic field which is to the first approximation dipolar, the high-latitude ionosphere is modified by dynamic magnetospheric processes that are largely controlled by the interaction between the solar wind and the earth's magnetosphere. In fact, many ionospheric phenomena that occur at high latitudes are "footprint signatures" of this interaction and the ionosphere is sometime likened to a nature's TV screen for viewing the images of auroral displays. Through the coupling to the interplanetary magnetic field (IMF) the polar ionosphere gains direct access to solar wind plasma and particles that have been energized by magnetospheric processes and are sources of ionization in the auroral region in particular. By the same token, the ionospheric plasma can escape to space (polar wind and auroral bulk upflows of ions with energy of a few eV) or be trapped in the magnetosphere. A fraction of up-flowing ions can be accelerated to high energies through various processes that include ion beams, ion conics, upwelling ions and transversely accelerated ions.

1.3.1.1 Ionization Sources

At high latitudes, the main source of ionization, which is solar EUV and X rays, is supplemented by energetic particle precipitation. The electrons that are accelerated by magnetospheric processes (e.g. substorms) are the cause of aurora and a source of ionization. The rate of electron energy loss and thus of ion production is a function of the particle energy and atmospheric density. More energetic particles penetrate deeper into the atmosphere producing the maximum ionization at lower altitudes. Ion production rates due to different electron energy are calculated in a model atmosphere (Rees, 1989) and the production rate must be integrated over energy of all precipitating particles. Although protons also produce aurora (proton aurora) a much more significant source of ionization are high energy protons emitted by solar flares. Enhanced fluxes of energetic protons are the cause of polar-cap absorption (PCA) events (Reid, 1974) that result in HF radio signal blanketing.

1.3.1.2 The Influence of Magnetic and Electric Fields on Plasma Convection

The geomagnetic poles are located at high geographic latitudes where the geomagnetic field axis is oriented vertically. The magnetospheric electric field **E** maps along the magnetic field lines down to the high-latitude ionosphere where **E** becomes basically horizontal. The mobility of charged particles is quite different in the directions perpendicular and parallel to the magnetic field **B**. Barring collisions, particles can move freely along the magnetic field, but can only move perpendicular to the magnetic field in the presence of electric fields. At high latitudes, the electric fields that are imposed on the ionosphere through interaction between the solar wind and magnetosphere drive horizontal motions of the ionospheric plasma due to $\mathbf{E} \times \mathbf{B}$ drift, so called ionospheric convection. The convection pattern mainly depends on the orientation of the IMF since the latter controls the solar wind coupling to the magnetosphere. In the simplest case, for purely southward IMF($B_z < 0$ and $B_y = 0$), the F-region plasma is convected antisunward across the polar cap and returns sunward along the dawn and dusk portions of the auroral oval. This two-cell convection pattern is tilted with respect to sun-earth line for IMF $B_y \neq 0$. The pattern is more complex for the northward IMF ($B_z > 0$) when the ionospheric convection is less intense and confined to higher latitudes (smaller polar cap) but more structured. The convection pattern may consist of three or four cells (Cowley and Lockwood, 1992). For the northward IMF, the magnetic reconnection between IMF and magnetosphere, which is thought to be the principal coupling process acting at the magnetopause, is significantly reduced and limited to tail lobe reconnection producing a circulatory motion of plasma

(lobe cell) in the central polar cap. In addition, viscous-like coupling process acting simultaneously at the magnetopause boundary may further complicate the convection morphologies. The ionospheric convection plays an important role not only in the plasma transport but also in structuring of the high-latitude ionospheric plasma (see, e.g., Crowley, 1996).

The motion of the plasma convecting at a velocity $\mathbf{v} = \mathbf{E} \times \mathbf{B} / B^2$ can be measured by a variety of techniques, including coherent- and incoherent-scatter radars, digisondes, satellites and rockets. The convection patterns can also be derived from equivalent currents based on ground magnetometer data. All of the techniques have their limitations but a network of HF radars is most suitable for large-scale mapping of the convection at high spatial and temporal resolution. Such capabilities have been successfully employed by Super Dual Auroral Radar Network (Greenwald et al., 1995). Statistical convection/potential patterns were first derived using the single radar data (Ruohoniemi and Greenwald, 1996) but many more details of convection dynamics are revealed from global imaging of the instantaneous convection using the full network of radars (Ruohoniemi and Baker, 1998). These observations are in general agreement with the conceptual model proposed by Cowley and Lockwood (1992).

1.3.1.3 Ionospheric and Field-Aligned Currents

Unlike the F region where the ionospheric plasma is to a first approximation collisionless (ion collision frequency ν_i is much smaller than the ion gyrofrequency Ω_i) the ionosphere at E-region altitudes is a resistive medium. The ions are unmagnetized ($\nu_i >> \Omega_i$) while the electrons are still "collisionless" ($\nu_e << \Omega_e$). At the E-region altitudes, the ion-neutral collisions are also important but the electron-neutral collisions are not. Thus, the electrons can move perpendicularly to magnetic field while the $\mathbf{E} \times \mathbf{B}$ drift of ions is impeded by collisions. This leads to an ionospheric (Hall) current. Generally, the dominant electric field is northward in the dusk sector and southward in the dawn sector of the auroral oval and the Hall conductivity (perpendicular to $\mathbf{E}$) dominates over the Pedersen conductivity (parallel to $\mathbf{E}$). Thus, the dominant electric current is an eastward electrojet in the afternoon and a westward electrojet in the morning. The ionospheric currents can be observed by ground magnetometers measuring the magnetic perturbations due to the horizontal electric currents, including the auroral electrojets.

The large value of the direct or longitudinal (along the magnetic field) conductivity suggests that significant currents should be able to flow along the magnetic field. In fact, field-aligned currents (FACs), sometimes called Birkeland currents after K. Birkeland who postulated their existence in 1908, were disputed by Chapman (1935) and promoted by Alfvén (1939) before the concept was finally confirmed in 1960s. The currents were first detected in the early 1970s by satellite-borne magnetometers that measured magnetic perturbations perpendicular to the field. Field-aligned currents provide an important electric circuit between the plasmas in the solar wind, magnetosphere and the high-latitude ionosphere. The system of large-scale field-aligned currents determined from low-altitude satellite measurements is actually a superposition of several current circuits connecting the ionosphere with different regions of the magnetosphere. The main FAC systems are composed of so-called Region 1 and Region 2, east-west aligned sheet currents that are connected in the ionosphere by the north-south Pedersen current.

1.3.1.4 Ionospheric Density Irregularities

In the E-region, strong auroral electrojets are sources of field-aligned ionospheric irregularities on perpendicular spatial scales from a fraction of 1 m to more than 100 m. If a differential electron-ion drift velocity $\mathbf{V}_d = \mathbf{V}_i - \mathbf{V}_e$ which is often taken approximately equal to the electron Hall drift $\mathbf{E} \times \mathbf{B} / B^2$, exceeds the ion-acoustic speed c_s (in the absence of horizontal electron density gradients) the plasma becomes unstable to Farley-Buneman (two-stream) instability. Electrostatic plasma density waves (fluctuations) are excited propagating nearly perpendicular to the magnetic field. When electron density gradients are present, two-stream and gradient drift theory is required to describe the instabilities (Farley and Fejer, 1975).

In the F-region, in the presence of horizontal electron density gradients, irregularities are produced by convective plasma processes and in particular, by the fluid **E** × **B** (gradient drift) interchange instability (Tsunoda, 1988). The irregularity scale-size spectrum extends from ~10 km down to the ion gyroradius. Because plasma density gradients are necessary for generating the F-region irregularities, a large-scale ionospheric structure is often of primary importance in the generation of small-scale structure. Much recent work has focussed on the modelling of plasma structuring in the F region using ever more sophisticated numerical codes (e.g., Gondarenko and Guzdar, 2004).

1.3.1.5 Large-Scale Ionospheric Structure

Large-scale plasma structure in the high-latitude ionosphere includes the auroral oval, the location of diffuse and discrete auroras, auroral and sun-aligned (transpolar) arcs, polar cap patches, boundary blobs, ionospheric troughs or plasma density depletions, and traveling ionospheric disturbances. The magnetospheric, ionospheric and atmospheric sources of this rather complex structure of the high-latitude ionosphere include solar induced ionization, soft and energetic particle precipitation, field-aligned and ionospheric currents, electric field enhancements, plasma convection, magneto-hydrodynamic waves (MHD) and atmospheric gravity waves (AGWs). However, the overarching process that controls, influences or modulates all of these sources is the solar wind coupling to the magnetosphere.

The most dynamic spatial and temporal changes of the ionospheric structure occur in the auroral zone, particularly during substorms. The optical manifestation of this structure is aurora borealis and aurora australis. The statistical distribution of aurora occurrence led to a concept of the auroral oval (Feldstein, 1963). The all-sky observations showed that at any given time the locus of aurora is not circular but oval, extending further equatorward in the midnight sector than on the dayside. Also, the auroral oval is widest near local midnight. Although originally proposed as a statistical concept, the auroral oval is present at all times, although its intensity varies with changing activity levels as can be seen from satellite images. The position and width of the auroral oval are both functions of the level of the solar wind-magnetosphere interaction. The aurora is a consequence of a complex interplay among energetic particle fluxes, electromagnetic waves that may alter the particle pitch angles, and the ionization potentials of the atmospheric atoms and molecules in the altitude range where the emissions originate. The conductivity changes that are caused by precipitating particles (mainly keV electrons) together with the convection electric field that maps from deep space into the auroral ionosphere, act together to regulate field-aligned currents and the ionospheric currents to which they connect. While most of the energy dissipated though Joule heating is associated with the large scale slowly varying Birkeland currents, Alfven waves with a Poynting flux into the ionosphere can also contribute significantly to ionospheric heating (Keilling et al., 2003). Non-visual techniques are now in place for routinely monitoring the location of auroral boundaries, e.g., Newell et al. (2002) and link to http://sd-www.jhuapl.edu/Aurora/ ovation_live/.

In the magnetospheric context, the auroral oval lies predominantly on closed field lines, with its poleward edge approximately marking the boundary between open and closed field lines. Energy is transported into the closed field line region from the magnetospheric boundary layers and through reconnection of magnetic field lines at a neutral line located beyond ~ 20 R_E in the magnetotail. Although the transport processes occur over huge volumes of space, regions where spatial gradients in pressure and electric field are present can feature parallel electric fields that accelerate electrons into the ionosphere in localized regions causing auroral arcs.

The nightside of the auroral oval can exhibit explosive auroral intensifications and rapid changes in arc structure, which are the mark of the auroral substorm. The auroral substorm is part of a global disturbance called a magnetospheric substorm, which features disturbances in all portions of the electromagnetic spectrum. A magnetospheric substorm features two distinctly different processes, namely directly driven and loading-unloading. Directly driven activity follows closely the increase and subsequent decrease in energy provided to the magnetosphere from the solar wind. Some of the energy entering the magnetosphere is stored

in the magnetotail and then is released at some later time. The auroral substorm is the visible signature of the release of that stored energy. Substorms last for 1-2 hours and may recur as often as every few hours; they are common when the IMF is southward and infrequent when the IMF is northward.

Ionospheric troughs are regions of depleted F-layer plasma density limited in latitude but extended in longitude that are observed in the sub-auroral and high-latitude ionosphere including the polar cap. Traditionally, a distinction is made between the mid-latitude (main) trough that is found adjacent and equatorward of the auroral oval and the high-latitude troughs that are observed inside the auroral oval or in the polar cap. Ionospheric convection plays a crucial role in trough formation through chemical and dynamical processes that include ion recombination, Joule heating and frictional heating. A subset of mid-latitude troughs are caused by rapid sub-auroral ion drift (SAID) due to large, short-lived electric fields occurring just equatorward of the auroral oval (Rodger et al., 1992).

The morphology of the polar ionosphere is largely determined by the orientation of the IMF which causes it to have one of two idealized characteristic states (Carlson, 1994). For southward IMF ($B_z < 0$), mesoscale (100-1000 km) plasma density enhancements known as polar patches are produced in the dayside ionosphere near the magnetospheric cusp footprint. The patches then convect antisunward in the polar cap where they can be observed by various techniques, principally radio and optical. Patches are associated with steep density gradients generating small-scale irregularities causing HF radar backscatter and severe scintillations of VHF and UHF radio signals. Although processes responsible for the formation of patches have been proposed a controversy still remains and several mechanisms are probably active in causing patches (Crowley, 1996). For northward IMF ($B_z > 0$), the polar ionosphere is characterized by considerable plasma convection structure and stable sun-aligned arcs. The sun-aligned or transpolar arcs delineate the polar cap plasma flow gradients/shears and provide valuable insight into magnetospheric topology and solar wind coupling processes (Carlson, 1994). It can be said that the polar ionosphere alternates between the two states depending on the IMF orientation but when IMF B_z switches sign often, both patches and sun-aligned arcs may coexist in the polar cap at any given time.

As already discussed in Section 1.2 for the mid-latitude ionosphere, the ionospheric F-region is often structured into wavelike fluctuations of electron density – travelling ionospheric disturbances (TIDs) – caused by atmospheric gravity waves (AGWs). While AGWs that originate in the lower atmosphere and propagate their energy upward may cause TIDs the bulk of large- and medium-scale TIDs are generated in the auroral oval by surges of auroral electrojet currents or convection electric field (Hocke and Schlegel, 1996). TIDs can be detected at all latitudes using various radio techniques including HF Doppler sounders, ionosondes, HF radars, incoherent scatter radars, or optically as airglow emissions modulated by AGWs. More recently TID-induced electron densities have been imaged using tomographic techniques (e.g., Pryse et al., 1995) and TEC measurements by GPS networks (e.g., Tsugawa et al., 2004). Modelling of gravity waves generated by enhancements in the ionospheric electric field shows that each electric field enhancement causes a Joule heating pulse, which in turn launches a gravity wave propagating equatorward and poleward of the source region (Millward et al., 1993a, b). To elucidate the physics of the AGW-TID relationship, theoretical simulations were conducted and compared with incoherent scatter radar data (Kirchengast 1996; Kirchengast et al., 1996). This comparison allowed comprehensive information on AGWs and the AGW-induced TID ionospheric parameters (electron density, field-aligned ion drift, and ion and electron temperatures), or rather their variations with respect to background values based on a model of ionosphere, to be deduced by means of careful physical modelling.

The following sections focus on specific phenomena/processes and their effects on ionosphere and its structure in more detail. Considerable efforts that have been expended to understand the morphology of the ionosphere as well as magnetospheric processes have produced a vast scientific literature. Appropriate references for further details are made wherever deemed necessary but this chapter does not claim to be an exhaustive review. Rather it should be considered as a concise summary to provide an interested reader with up-to-date information (or a source of it) on the morphology of high-latitude ionosphere.

1.3.1.6 References

Alfvén, H., A theory of magnetic storms and of the aurorae, *K. Sven. Vetenskapakad.*, Handl., Ser. 3, 18(3), 1939. (Reprinted in part with comments by Alex Dessler and John Wilcox in *Eos*, 51, 180-194, 1970).

Carlson, H. C., Jr., Dark polar ionosphere: Progress and future challenges, *Radio Sci.*, 29(1), 157-165, 1994.

Chapman, S., The electric currents system of magnetic storms, *Terr. Mag. Atmos.*, 40, 349-370, 1935.

Cowley, S. W. H., and M. Lockwood, Excitation and decay of solar wind-driven flows in the magnetosphere-ionosphere system, *Ann. Geophys.*, 10, 103-115, 1992.

Crowley, G., Critical review of ionospheric patches and blobs, *Review of Radio Science 1993-1996*, 619-648, Oxford Univ. Press, New York, 1996.

Farley, D. T., and B. G. Fejer, The effect of the gradient drift term on Type 1 irregularities, *J. Geophys. Res.,* 80, 3087-3090, 1975.

Feldstein, Y. I., On Morphology and Auroral and Magnetic Disturbances at High Latitudes, *Geomagn. Aeron.* 3, 138, 1963.

Gondarenko, N.A., and P.N. Guzdar, Plasma patch structuring by the nonlinear evolution of the gradient drift instability in the high-latitude ionosphere, *J. Geophys. Res.*, 109, A09301, doi:10.1029/ 2004JA010504, 2004.

Greenwald, R.A., K.B. Baker, J.R. Dudeney, M. Pinnock, T.B. Jones, E.C. Thomas, J.-C. Villain, J.-C. Cerrisier, C. Senior, C. Hanuise, R.D. Hunsucker, G. Sofko, J. Koehler, E. Nielsen, R. Pellinen, A.D.M. Walker, N. Sato, and H. Yamagishi, DARN/SUPERDARN A global view of the dynamics of high-latitude convection, *Space Sci. Rev.*, 71, 761 796, 1995.

Hocke, K., and K. Schlegel, A review of atmospheric gravity waves and traveling ionospheric disturbances: 1982-1995, *Ann. Geophysicae*, 14, 917-940, 1996.

Keiling, A., J. R. Wygant, C. A. Cattell, F. S. Mozer, and C. T. Russell, The Global Morphology of Wave Poynting Flux: Powering the Aurora, *Science*, 299, 383-386, 2003.

Kirchengast, G., Elucidation of the physics of the gravity wave-TID relationship with the aid of theoretical simulations, *J. Geophys. Res.*, 101, 13353-13368, 1996.

Kirchengast, G., K. Hocke, and K. Schlegel, The gravity wave-TID relationship: insight via theoretical model - EISCAT data comparison, *J. Atmos. Terr. Phys.*, 58, 233-243, 1996.

Millward, G.H., S. Quegan, R.J. Moffett, T.J. Fuller-Rowell, and D. Rees, A modeling study of the coupled ionospheric and thermospheric response to an enhanced high-latitude electric field event, *Planet. Space Sci.*, 41, 45-56, 1993a.

Millward, G.H., R.J. Moffett, S. Quegan, and T.J. Fuller-Rowell, Effect of an atmospheric gravity wave on the mid-latitude ionospheric F layer, *J. Geophys. Res.*, 98, 19173-19179, 1993b.

Newell, P.T., T. Sotirelis, J.M. Ruohoniemi, J.F. Carbary, K. Liou, J.P. Skura, C.-I. Meng, C. Deehr, D. Wilkinson, and F.J. Rich, OVATION: Oval variation, assessment, tracking, intensity, and online nowcasting, *Annales Geophysicae*, 20, 1039-1047, 2002.

Pryse, S. E., C. N. Mitchell, J. A. T. Heaton and L. Kersley, Travelling ionospheric disturbances imaged by tomographic techniques, *Ann. Geophysicae*, **13**, 1325-1330, 1995.

Rees, M. H., *Physics and Chemistry of the Upper Atmosphere*, Cambridge Univ. Press, New York, 1989.

Reid, G.C. 1974. Polar cap absorption - Observations and theory. *Fundamentals of Cosmic Physics* 1:167-202.

Rodger, A.S., R.J. Moffett, and S. Quegan, The role of ion drift in the formation of ionisation troughs in the mid- and high-latitude ionosphere – a review, *J. Atmos. Terr. Phys*., 54, 1-30, 1992.

Ruohoniemi, J. M. and R. A. Greenwald, Statistical patterns of high-latitude convection obtained from Goose Bay HF radar observations, *J. Geophys. Res.*, 101, 21743, 1996.

Ruohoniemi, J. M. and K. B., Baker, Large-scale imaging of high-latitude convection with Super Dual Auroral Radar Network HF radar observations, *J. Geophys. Res.*, 103, 20797-20811, 1998.

Tsugawa, T., A. Saito, and Y. Otsuka, A statistical study of large-scale traveling ionospheric disturbances using the GPS network in Japan, *J. Geophys. Res.*, 109(A6), A06302, doi:10.1029/2003JA010302, 2004.

Tsunoda, R. T., High latitude F-region irregularities: A review and synthesis, *Rev. Geophys., 26,* 719, 1988.

1.3.2 Particle Precipitation

(Dirk Lummerzheim, Geophysical Institute, University of Alaska)

1.3.2.1 Introduction

The most spectacular manifestation of particle precipitation into the thermosphere is the aurora. Energetic electrons in the energy range of a fraction of a keV to several keV are the cause of the visible aurora, but energetic proton precipitation also causes diffuse auroral emissions. The electron precipitation averaged over the entire auroral oval dominates the auroral precipitation. The spatial distribution of protons and electrons is, on average, different, and the equatorward edge of the evening aurora tends to be dominated by proton precipitation. During large substorms, the hemispheric power in the auroral electron precipitation can reach several hundred GW. Locally, energy fluxes in small scale auroral filaments are observed to reach up to several hundred mW/m^2. These precipitating particles deposit their energy and cause ionization at altitudes between 80 and 500 km. The source of the auroral particles is in the magnetosphere. In the inner magnetosphere pitch angle scattering produces electrons with high mean energy (tens of keV) and low energy fluxes, which in turn give rise to the diffuse aurora at the equatorward boundary of the auroral oval. The plasma sheet and tail of the magnetosphere is the region where substorms originate. On the dayside in the cusp, the auroral precipitation originates at the magnetopause. While nightside (substorm) precipitation typically has mean energies of several keV to more than 10 keV for electrons and several tens of keV for protons, the precipitation in the cusp is softer and causes higher altitude aurora with dominant OI(630 nm) emissions from the F-region.

Non-auroral precipitation occurs inside the polar cap and in regions that are magnetically connected to the sunlit part of the other hemisphere. Photoelectrons that are generated inside the thermosphere escape upwards and travel along the magnetic field to the other hemisphere. Outside of the auroral region and on the nightside, these fluxes can be the dominant precipitation, although the energy fluxes are very low (below 1 mW/m^2) with mean energies of a fraction of a keV.

The polar caps are magnetically connected to the magnetosheath and the solar wind has direct access to the polar cap ionosphere. Electrons of a few hundred eV that originate in the solar corona precipitate here. Figure 1-4 shows an example of the auroral electron and ion precipitation as well the soft electrons in the polar cap. During large solar events, high energy protons (10 MeV) also penetrate into the lower ionosphere at high latitudes.

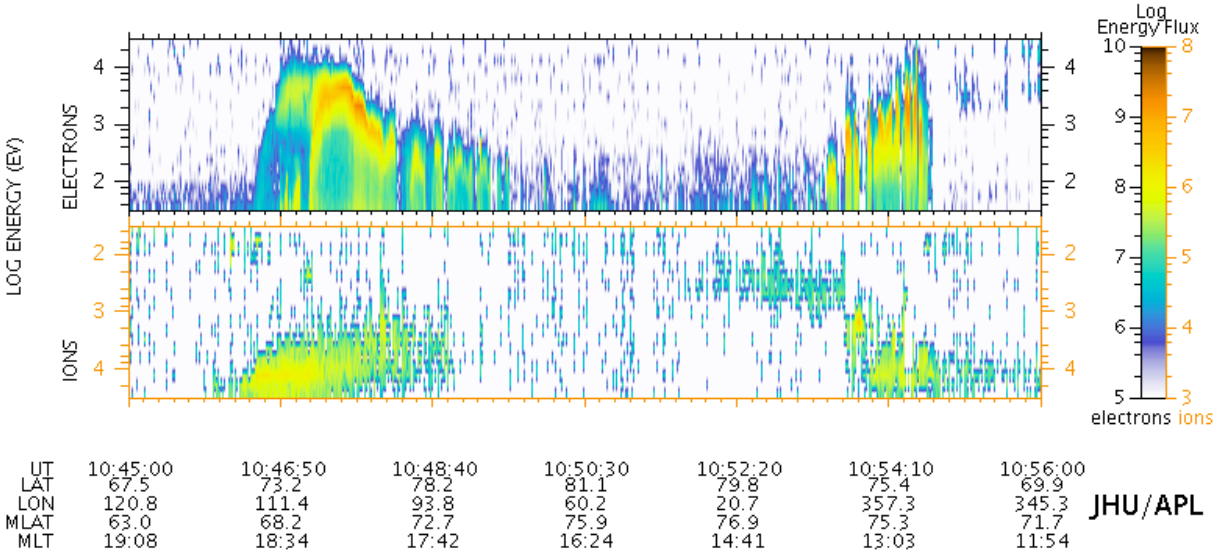

Figure 1-4: Example of the Downward Electron and Ion Flux from a DMSP Pass Over the Northern Auroral Oval.

The satellite enters the aurora on the evening side and observes electrons of several keV energy. As it enters the polar cap, low fluxes in the range of about 100 eV are seen, until eventually the satellite exits the auroral oval near local noon (13:00 MLT). Close inspection of the ion fluxes show that in the evening sector the ion precipitation extends to latitudes below the electron precipitation.

1.3.2.2 Auroral Electron Precipitation

1.3.2.2.1 Modeling

Auroral electrons are guided by the Earth's magnetic field into the thermosphere. For the field strength and average mean free path at auroral altitudes the gyrofrequency is much larger than the collision frequency. This allows specifying the motion of an individual electron by the motion of its guiding center. Consequently, the motion is symmetric with respect to azimuth and one-dimensional along the direction of the magnetic field. The acceleration of the electrons takes place in the magnetosphere, while electrons only lose energy as they penetrate the collision dominated thermosphere. The time that it takes the electron to penetrate the thermosphere and deposit its energy is small compared to the time it takes for the host medium to change in response to the precipitating electrons. The collision processes are dominated by the electron neutral collisions, and the non-linearity that would arise from collisions with aurorally produced ions are negligible. All collisions are assumed to be binary collisions, i.e., no three-body collisions need to be considered. The density of streaming particles is low compared to the ambient density. The mirror force has a small effect on the precipitating electrons and is also usually neglected. These assumptions allow to simplify the non-linear Boltzmann equation for auroral electrons and to derive a linear, steady state transport equation:

$$\mu \frac{\partial I(E,\mu,\tau)}{\partial \tau} = I(E,\mu,\tau) - \frac{\omega}{2}\int_{-1}^{1} d\mu' p(\mu,\mu') I(E,\mu',\tau) - Q\left(I(E'>E),\mu,\tau\right)$$

In this transport equation, the change of the electron intensity (or differential number flux) I as a function of energy E, cosine of the pitch angle μ, and scattering depth τ, as it penetrates deeper into the atmosphere is due to loss from collisions (first term on the right), angular redistribution from elastic scattering (second term) and energy degradation and secondary electron production Q (third term). The angular redistribution in the elastic scattering term is given by the phase function p, while the scattering albedo ω is given by the ratio of the total elastic cross-section σ^{el} to the total cross section σ^{tot}:

$$\omega = \frac{\sum n(z)\sigma^{el}(E)}{\sum n(z)\sigma^{tot}(E)}$$

where *n(z)* is the neutral density as function of altitude and the sums extend over all neutral species. The scattering depth is defined as:

$$\tau(z,E) = \sum \sigma^{tot}(E) \int_{z}^{\infty} n(z')\, dz'$$

Because of the assumption that electrons do not gain energy once they enter the thermosphere, the energy redistribution term can be expressed in terms of the electron intensity at higher energies, and the transport equation can be solved starting at some high energy and progressing to lower energies. The numerical solution requires that neutral densities, cross sections for various collision processes, phase functions, energy redistribution functions, and secondary electron production from ionization processes are specified. Lab measurements of many of these processes are available but often have large uncertainties. Neutral density models that are based on global observations such as MSIS [*Hedin*, 1991, 1992; *Picone et al.*, 2002] also have a poor representation of the thermosphere in the auroral oval and introduce uncertainties into an auroral model.

Two complications in a numerical treatment of the transport equation arise from the range of densities, and thus the scattering depth for the altitude region in which the equation must be solved, and from the large difference between the mean energy of the incident electrons and the energy loss in individual collisions. Either the numerical grid has to resolve these large ranges of the required variable space or numerical methods have to be found to accurately treat these mismatched variables.

A number of different approaches to the solution of the transport equation are presented in the literature. Some models relax the assumptions listed above, usually with a penalty of introducing other assumptions. Electron transport models for Earth's aurora have been developed by *Banks and Nagy*, [1970; see also *Nagy and Banks*, 1970] using a two-stream approach, *Banks et al.* [1974] using Monte-Carlo techniques, *Jasperse* [1976] using a theoretical method of solution, *Strickland et al.* [1976] using a multi-stream approach (i.e., with pitch angle resolution), *Stamnes* [1980, 1981] using a two-stream discrete ordinate method, *Link* [1992] applying a Feautrier solution, *Lummerzheim and Lilensten* [1994] using a discrete ordinate technique, and *Solomon* [1993] using a Monte-Carlo technique. *Min et al.* [1993] took the discrete ordinate method of *Lummerzheim and Lilensten* [1994] to include small field aligned ionospheric electric fields in order to study the influence of the ambipolar diffusion field on electron precipitation. *Peticolas and Lummerzheim* [2000] have developed a time-dependent electron transport model which can simulate flickering aurora or fast moving auroral filaments.

The solution of the transport equation provides the electron intensity as a function of altitude, given an electron source. Figure 1-5 shows an example of the downward electron flux from a model calculation.

For photoelectrons this source would be distributed in altitude and must be calculated from a radiative transfer calculation of solar EUV radiation, while for aurora the incident electron precipitation can be specified as a downward directed electron flux at the upper boundary of the model. From the altitude profile of the electron intensities, one can calculate the excitation and ionization rates, energy deposition and heating rates, electron particle and energy fluxes, and a number of auroral emission rates.

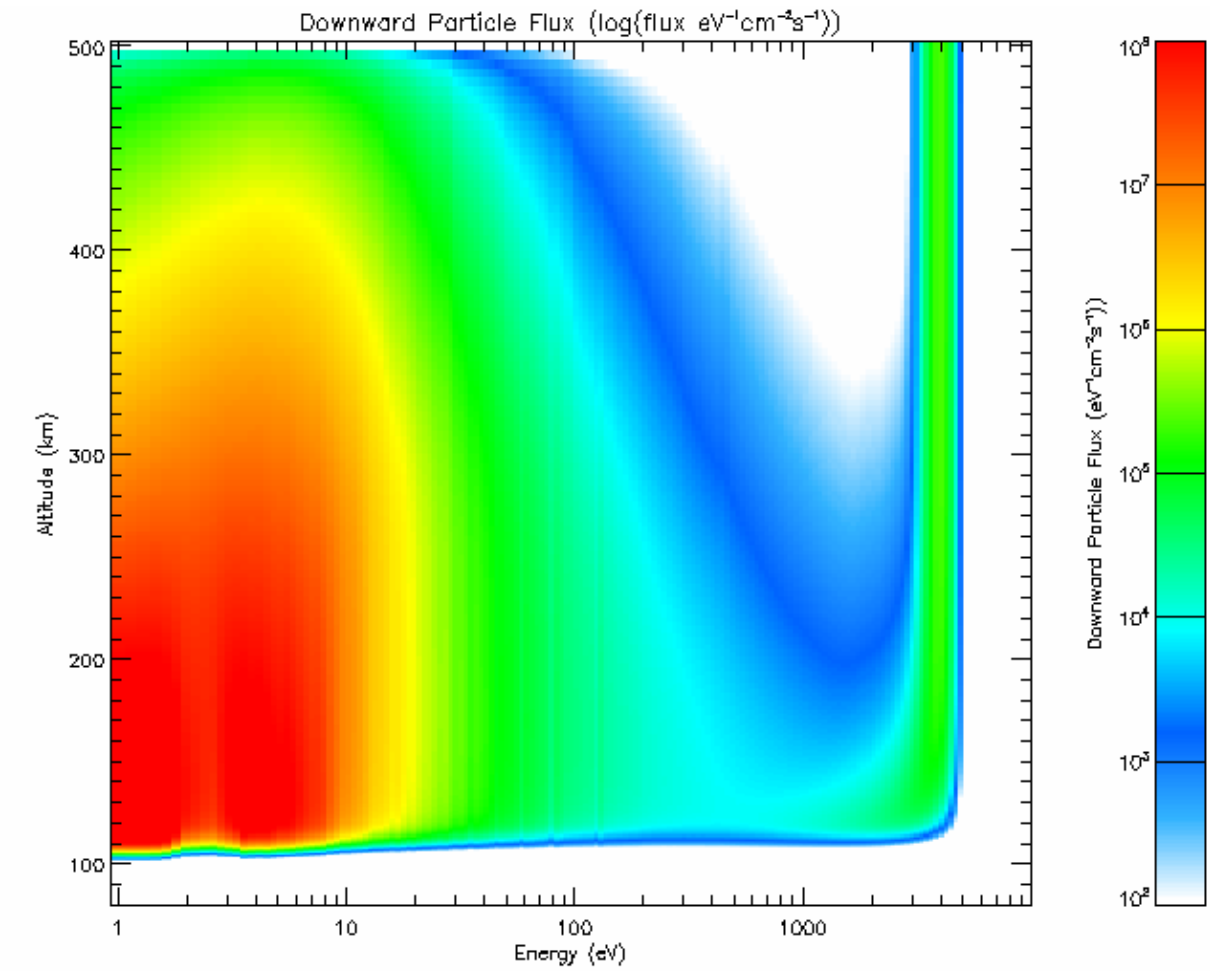

Figure 1-5: Example of the Downward Electron Flux as a Function of Altitude and Energy from a Model Calculation.

The incident flux was specified at 500 km as a Gaussian energy distribution at 4 keV with an energy flux of 1 mWm^{-2}. As the electron penetrate the atmosphere, low energy electrons from ionization collisions are generated. The high energy flux diminishes through excitation and ionization collisions as can be seen at about 120-150 km altitude. Auroral emissions from excited states are mostly generated by collisions of the secondary electrons in the energy range of a few eV to several tens of eV, where cross sections maximize and the secondary electron fluxes are large.

1.3.2.2.2 Observations

The spectrum of the precipitating electrons can be obtained either by direct in situ observations from sounding rockets or satellites, or by inferring from optical auroral or ionospheric radar observations. In situ observations are the least ambiguous, while remote sensing using optical observations and modeling can provide simultaneous large scale maps of auroral particle precipitation characteristics.

Low earth orbit satellites typically move at velocities of 7-10 km/s, and the in situ observations have to be made at a high frequency in order to resolve the spatial structure of auroral precipitation. An example of the extreme fine structure is presented by *Boehm et al.* [1995]. The Freja satellite had a high resolution electron particle spectrometer which obtained the energy spectrum by sweeping the high voltage in the detector. During a transition through an auroral precipitation region, the instrument encountered a rapid variation in the observed electron flux as a function of energy which *Boehm et al*. interpreted as a sharp spatial rather than temporal gradient that the satellite penetrated while sweeping though the energy channels of the detector. The 2 ms rise time translates into a spatial length scale of 10 m for the auroral electron precipitation.

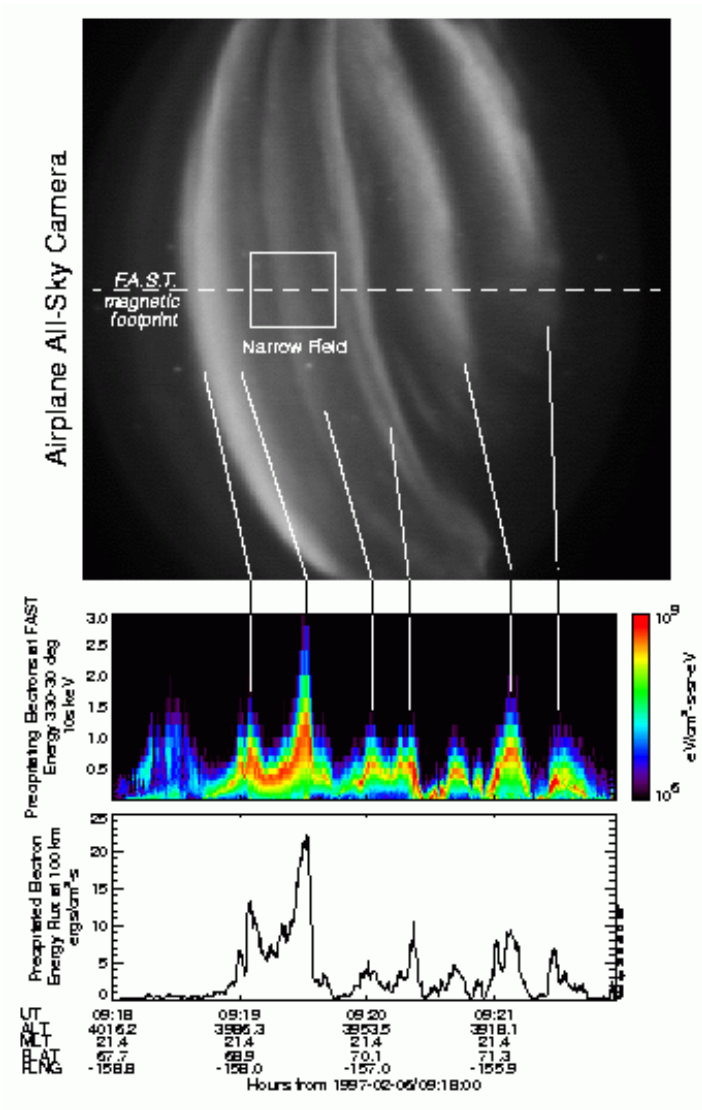

Figure 1-6: Electron Precipitation from FAST Observations with Coincident Auroral Imaging from an Airplane – from Stenbaek-Nielsen et al., 1998.

Such high resolution measurements are not made routinely, and most in situ observations are therefore averages over a limited spatial region. The DMSP satellites carry particle detectors that observe the incident electron and proton spectra in the vertical direction and collect one spectrum over an integration period of 1 s [*Hardy et al.,* 1985]. These electron precipitation spectra have been compiled into statistical

descriptions of the auroral oval [*Hardy et al.,* 1985], have been used to identify precipitation boundaries [*Newell and Meng,* 1988; *Newell at al.,* 1991], and have provided the electron precipitation in numerous case studies of aurora. A statistical analysis of the DMSP obtained particle spectra lead *Newell et al.* [1996] to the conclusion that the mean energy of auroral electrons in sunlit aurora is on average less than in aurora in the dark thermosphere. These findings are supported by subsequent publications of the analysis of UVI auroral images [*Liou et al*., 1997], and the re-analysis of an old data set of ground based observations [*Deehr et al.,* 2005].

Most particle detectors for auroral precipitation measurements cover the energy range from a few eV to about 32 keV. This is the energy range where most of the energy flux is and which contains the mean energy of the precipitation. The UARS PEM (particle environment monitor) and the NOAA/TIROS satellites also carry integral detectors to extend the energy sensitivity into the MeV range. Although the electrons in the hundreds of keV and MeV range contribute little to the auroral emissions, these electrons penetrate deeper into the middle atmosphere where they may be the dominant source for excitation, ionization, and energy deposition [*Winningham et al.,* 1993, *Codrescu et al.,* 1997]. Ground based cosmic noise absorption measurements are suitable to determine particle fluxes that penetrate into the D-region [*Tanaka et al*., 2005].

The FAST satellite was designed to obtain high resolution observations of auroral particle precipitation and other auroral parameters [*Carlson et al.,* 1998]. The particle detector obtains a full spectrum (pitch angle and energy distribution) every 70 ms. A high time resolution mode allows a limited number of energies and pitch angle measurements with integration times of 1.7 ms. A review of the large number of publications and science results from the FAST satellite exceeds this article by far. There are numerous other satellites and sounding rocket observations of precipitating particles and their characteristics.

Indirect methods to obtain the auroral electron precipitation are based on spectroscopic optical observations, or incoherent scatter radar observations. Incoherent scatter radar measures the altitude profile of the plasma density in the ionosphere. In aurora, this can be related to the ionization rate profile with some assumptions and using an auroral model one can deduce some parameters describing the incident electron flux [*Robinson and Vondrak,* 1985, *Doe at al.,* 1997]. *Lanchester et al.* [1994, 1997] used high resolution (200 ms temporal resolution) EISCAT measurements of the plasma density profile and an ionospheric model to conclude that small auroral filaments are characterized by near mono-energetic electron precipitation of several keV mean energy and energy fluxes in excess of 100 mWm^{-2}.

Optical methods to obtain the mean energy of the auroral electrons are based on the fact that the altitude of maximum emission is a function of the mean energy. The neutral composition changes with altitude, and ratios of emissions can therefore be used to measure this altitude and thus infer the mean energy. Because of the lighter mass and greater scale height of the atomic oxygen, the ratios of molecular nitrogen (or molecular nitrogen ion) emissions and atomic oxygen emissions are most useful. The disadvantage of this method is that a good model of the neutral atmosphere is required. The aurora modifies the ionosphere and subsequently the neutral composition sufficiently, that the atomic oxygen density in aurora is not very well represented in statistical neutral atmosphere models like MSIS. *Hecht et al.* [1989] has developed techniques where the comparison of three different auroral spectral brightnesses can be used to obtain the auroral electron energy and a modification of the atomic oxygen density profile at the same time.

Remote sensing from satellite in the UV wavelength range has the advantage that the auroral brightness can be analyses on the entire auroral oval, including the sunlit portion [*Lummerzheim et al.,* 1997]. The altitude of the auroral emission maximum can then be inferred from the N_2LBH spectral brightness. The LBH bands cover a sufficient wavelength range that the short wavelength end of the bands are strongly absorbed by O_2 Schuman-Runge extinction, while the long wavelength bands propagate nearly without extinction. This is used in the analysis of Dynamics Explorer images [*Rees et al,* 1988], Polar UVI images [*Germany et al,* 1998], and TIMED GUVI [*Strickland et al.,* 1992].

1.3.2.3 Auroral Ion Precipitation

1.3.2.3.1 Modeling

In order to model proton aurora, the assumptions made to derive an electron transport equation lead to a similar equation for protons. There is, however, one significant complication. Protons easily charge-exchange in collisions with atomic oxygen, thus turning into hydrogen. The newly created hydrogen atom keeps nearly the same kinetic energy as the original proton had, but does not gyrate around the magnetic field. The precipitation therefore spreads horizontally as soon as collisions with the neutrals in the thermosphere occur. The energetic hydrogen can be stripped of its electron in subsequent collisions. This process can repeat several times before the particle has deposited its energy. The energetic hydrogen is described by the same transport equation as the proton, except that the charge-exchange collision term is replaced by the stripping collision term. Ionization collisions as well as stripping collisions create electrons which have energies up to several hundred eV. These electrons must be included in a proton aurora model by solving an electron transport equation with a source term to represent this secondary electron production.

Because the proton and hydrogen transport equations are strongly coupled, the numerical solution of these equations offers additional challenges over the modeling of electron aurora. The spreading across the magnetic field direction makes this an inherently two-dimensional problem. However, there are few two or three dimensional numerical simulations available. Most authors make the additional assumption that the precipitation region is large and can be treated as a one-dimensional problem. To account for the finite extend of incident proton beams, a spreading factor which is either derived from observations of proton aurora or from parameterizations of three-dimensional calculations is used.

One of the first numerical techniques to solve these transport equations was the Monte-Carlo method (*Davidson, 1965)*. Davidson's model included the coupled proton-hydrogen transport in two dimensions to assess the spreading of the incident proton precipitation. More recent Monte-Carlo simulations with different emphasis were performed by *Kozelov* [1993] who developed a proton-hydrogen auroral model which includes a converging magnetic field, *Decker et al.* [1996] who compared a number of different approaches to solve the coupled transport equations and used the Monte-Carlo simulation as a validation tool, *Lorentzen et al.* [1998] looked in particular at hydrogen emissions in the cusp, *Synnes et al.* [1998] developed a Monte-Carlo simulation including the angular redistribution to study the proton aurora as a source for ENA (energetic neutral atoms) observed from space, and *Gérard et al.* [2000] who developed the proton auroral model for the interpretation of space based auroral images. *Jasperse and Basu* [1982] and *Basu et al.* [1987] constructed quasi-analytical solutions for the coupled one-dimensional transport equations. Explicit numerical solutions to the coupled transport equations for proton and hydrogen precipitation, and the additional coupled equation for the secondary electrons were published by *Basu et al.* [1993], *Strickland et al.* [1993] and *Galand et al.* [1997, 1998].

Hydrogen emissions are negligible in electron aurora because the ambient hydrogen density is too low to lead to any significant emission. Proton aurora produces its own hydrogen, and these emissions are characteristic for proton precipitation. The hydrogen that is produced by charge exchange collisions has the kinetic energy of the proton precipitation. The hydrogen emissions have therefore a distinctive Doppler shifted emission line profile. As the proton-hydrogen beam forms in the atmosphere, energy degradation and scattering will produce a smooth energy distribution of particles even from a mono-energetic incident beam. This energy distribution is reflected in the hydrogen emission line profile. Scattering processes and the convergent magnetic field also produce angular redistribution of the incident beam. Part of the energetic hydrogen is then moving upwards, giving rise to a Doppler shift in the opposite direction. For a ground based observer, the bulk of the emissions come from downward moving particles, thus creating a blue shifted line profile. Angular distribution generates upward moving particles which cause a red-shifted emission. *Gérard et al.* [2000] make use of this Doppler shift to separate the auroral Lyman-α emissions from the geocorona emissions which are not Doppler shifted. The IMAGE satellite makes use of this

method to obtain global images of the proton aurora [*Mende et al*., 2000]. *Galand et al.* [1997, 1998] have studied the angular distribution in detail with their multi-stream proton aurora model. High spectral resolution ground based observations of the Doppler shifted H_β line show the expected blue shifted line profile and the small red-shifted component [*Lummerzheim and Galand,* 2001]. Figure 1-7 demonstrates this by comparing an auroral H_β line profile with a stationary calibration lamp.

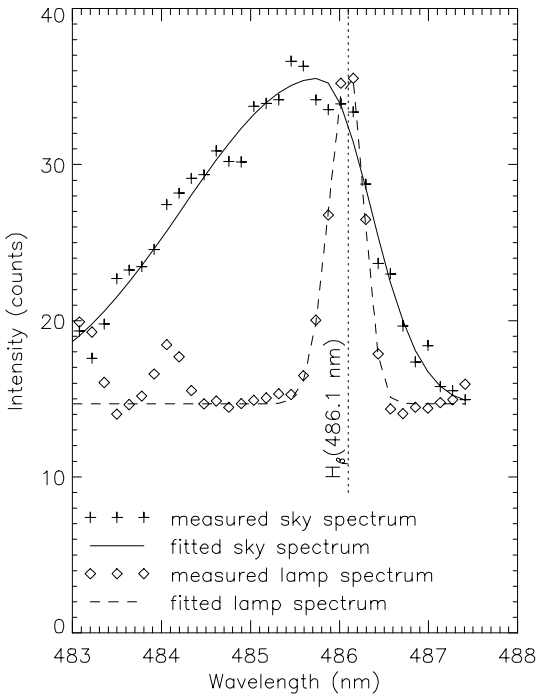

Figure 1-7: High Spectral Resolution Observations of the Doppler H_β Line Profile from Proton Aurora. Superimposed is the unshifted line shape from a stationary hydrogen lamp, taken with the same instrument and resolution.

The proton-hydrogen stream and the secondary electrons also cause excitation of the ambient neutral constituents of the thermosphere in aurora and give rise to all the emissions that are seen in electron aurora. Protons are unlikely to produce excitation without ionization, but the energetic hydrogen can cause excitation of all neutral species. The secondary electrons have a lower mean energy than secondary electrons in electron aurora and can thus cause slightly different brightness ratios between auroral emissions [Lummerzheim et al., 2001]. This can cause erroneous interpretation of auroral brightness ratios in terms of mean energy, if the brightness is assumed to be due to electron precipitation [Galand and *Lummerzheim,* 2004].

1.3.2.3.2 Observations

Most of the satellites and sounding rockets that carry electron detectors also have detectors for energetic ions. These may include detectors for measuring the ion mass and pitch angle distribution in addition to the energy spectrum as on FAST [*Carlson et al.,* 1998], or they may only determine the energy flux and mean energy from onboard processing of a measured energy spectrum (e.g. NOAA-TIROS Total Energy Detector (TED) [*Evans*, 1987]). Like the electron spectra, statistical analysis of a large number of satellite passes have been conducted [e.g. *Hardy et al.,* 1985; *Newell and Meng,* 1988; *Newell at al.,* 1991]. Case studies where in situ proton precipitation, satellite Lyman-α images, and ground based high resolution spectroscopy of the Doppler shifted H_β line have been use to validate model calculations [*Lanchester et al*., 2003] rely on measured proton precipitation energy spectra as model input.

Although the majority of the auroral ion precipitation is in the form of protons, great geomagnetic storms have also caused energetic oxygen ion and neutral precipitation [*Stephan et al*., 2004], usually at subauroral latitudes. The source of these oxygen atoms is in the ring current. This oxygen is observed by enhancements of atmospheric oxygen emissions and by ENA (energetic neutral atom) imaging from space [*Stephan et al*., 2000].

1.3.2.4 Polar Cap

Particle precipitation also occurs outside the auroral oval. The polar cap is magnetically connected to the solar wind, and solar wind electrons and ions have direct access to the ionosphere. This precipitation is called polar rain [Winningham and Heikkila, 1974]. It is characterized by very low energy fluxes (compared to auroral precipitation) of typically 10^{-3} mWm^{-2} and a mean energy typically in the range of a few hundred eV. On rare occasions, the electron flux can be more energetic. Newell and Meng [1990] conducted a study over three years and found 17 days where electron flux exceeded a threshold of 10^6 cm^{-2}s^{-1} at an energy above 1 keV. Intense polar rain events are not associated with geomagnetic storms, but rather occur during extreme quiet times [Hardy, 1984]. One the "day when the solar wind vanished" (11 May, 1999) intense polar rain was observed with energies up to 10 keV, causing X-ray emissions at extremely high latitude [Anderson et al., 2000]. Because of the usually low energy, the ionospheric effects of polar rain are mostly in the F-region. The polar rain can cause high airglow emissions of OI(630 nm) inside the polar cap [McEwen and Harrington, 1991].

On the other extreme of particle precipitation are the very high energy protons during Solar Proton events (SPE). These protons have energies in excess of 10 MeV, and have fluxes of more than 10 particles cm^{-2} s^{-1} ster^{-1}. SPE are rare, however, the Space Environment Center (SEC) has recorded 134 such events in the 31-year period from 1976-1997 [Kunches and Zwickl, 1999]. Since these protons penetrate into the lower ionosphere they have the potential to disrupt HF radio communication. Because of the large gyroradius of the high energy protons, the geomagnetic field limits access of this precipitation in latitude as a function of the rigidity. Ground based global riometer observations can thus be used to determine some energy characteristics based on the latitudinal extend of the D-region ionization [Rodger et al., 2006]. Depending on their energy, these energetic protons also penetrate into the middle atmosphere where they can have a large influence on the local chemistry, including the ozone layer [*Jackman et al*., 2004]. These very high energy protons are also a health risk for astronauts [e.g. *Hoff et al*., 2004].

1.3.2.5 References

Anderson, P. C., D. L. McKenzie, D. W. Datlowe, J. D. Hawley, S. M. Petrinec, M. Schulz, and D. E. Larson, 2000, Polar cap x-ray and electrons under low density solar wind conditions: Coordinated PIXIE and DMSP observations on 11 May 1999, *Geophys. Res. Lett., 27,* 4021.

Banks, P. M., and A. F. Nagy, 1970, Concerning the influence of elastic scattering upon photoelectron transport and escape, *J. Geophys. Res.*, *75*, 1902.

Banks, P. M., C. R. Chappell, and A. F. Nagy, 1974, A new model for the interaction of auroral electrons with the atmosphere: spectral degradation, backscatter, optical emission, and ionization, *J. Geophys. Res.*, *79*, 1459.

Basu, B., J. R. Jasperse, R. M. Robinson, R. R. Vondrak, and D. S. Evans, 1987, Linear transport theory of auroral proton precipitation: A comparison with observations, *J. Geophys. Res.*, *92*, 5920.

Basu, B., J. R. Jasperse, D. J. Strickland, and R. E. Daniell Jr., 1993, Transport-theoretic model for the electron-proton-hydrogenatom aurora, 1, Theory, *J. Geophys. Res.*, *98*, 21,517.

Carlson, C. W., R. F. Pfaff, and J. G. Watzin, 1998, The Fast Auroral SnapshoT (FAST) mission, *Geophys. Res. Lett., 25*, 2013.

Codrescu, M. V., T. J. Fuller-Rowell, R. G. Roble, and D. S. Evans, 1997, Medium energy particle precipitation influences on the mesosphere and lower thermosphere, *J. Geophys. Res., 102*, 19, 977.

Davidson, G. T., 1964, Expected spatial distribution of low energy protons precipitated in the auroral zones, *J. Geophys. Res*., *70*, 1061.

Deehr, C. S., M. H. Rees, A. E. H. Belon, G. J. Romick, and D. Lummerzheim, 2005, Influence of the ionosphere on the altitude of discrete auroral arcs, *Ann. Geophys., 23,* 759.

Doe, R. A., J. D. Kelly, D. Lummerzheim, M. Brittnacher, G. A. Germany, G. K. Parks, and J. F. Spann, 1997, Initial comparison of POLAR UVI and Sondrestrøm IS radar estimates for auroral electron energy flux, *Geophys. Res, Lett.*, *24*, 999.

Evans, D. S., 1987, Global Statistical Patterns of Auroral Phenomena, in *Proceedings of the Symposium on Quantitative Modeling of Magnetospheric-Ionospheric Coupling Processes*, 325, Kyoto.

Galand, M., J. Lilensten, W. Kofman, and R. B. Sidje, 1997, Proton transport model in the ionosphere, 1, Multistream approach of the transport equations, *J. Geophys. Res.*, *102*, 22, 261.

Galand, M., J. Lilensten, W. Kofman, and D. Lummerzheim, 1998, Proton transport model in the ionosphere, 2, Influence of magnetic mirroring and collisions on the angular redistribution in a proton beam, *Ann. Geophys.*, *16*, 1308.

Galand, M., and D. Lummerzheim, 2004, Contribution of proton precipitation to space-based auroral FUV observations, *J. Geophys. Res., 109*, A03307 10.1029/2003JA010321.

Gérard, J.-C., B. Hubert, D. V. Bisikalo, and V. I. Shematovich, 2000, A model of the Lyman-α line profile in the proton aurora, *J. Geophys. Res.*, *105*, 15, 795.

Germany, G. A., G. K. Parks, M. J. Brittnacher, J. F. Spann, J. Cumnock, D. Lummerzheim, F. Rich, and P. G. Richards, 1998, Energy characterization of a dynamic auroral event using GGS UVI images, AGU Monograph 104 "Encounter Between Global Observations and Models in the ISTP Era", Jim Horwitz, Dennis Gallagher, and Bill Peterson, editors, *143*.

Hardy, D. A, 1984, Intense fluxes of low-energy electrons at geomagnetic latitudes above 85 deg, *J. Geophys. Res., 89,* 3883.

Hardy, D. A., M. S. Gussenhofen, and E. Holman, 1985, A statistical model of auroral electron precipitation, *J. Geophys. Res., 90*, 4229.

Hecht, J. H., A. B. Christensen, D. J. Strickland, and R. R. Meier, 1989, Deducing composition and incident electron spectra from ground-based auroral optical measurements: Variations in oxygen density, *J. Geophys. Res., 94*, 13,553.

Hedin, A. E., 1991, Extension of the MSIS thermosphere model into the middle and lower atmosphere, *J. Geophys. Res.*, *96*, 1159.

Hedin, A. E., 1992, MSIS model (1986), *Planet. Space Sci*., *40*, 555.

Hoff, J. L., L. W. Townsend, and E. N. Zapp, 2004, Interplanetary crew doses and dose equivalents: variations among different bone marrow and skin sites, *Adv. Space Res., 34,* 1347.

Jackman, C. H., M. T. Deland, G. J. Gordon, E. L. Fleming, D. K. Weisenstein, M. K. W. Ko, M. Sinnhuber, J. Anderson, and J. M. Russell, The influence of the several very large solar proton events in years 2000 2003 on the neutral middle atmosphere, *Adv. Space Res., 35,* 445.

Jasperse, J. R., 1976, Boltzmann-Fokker-Planck model for the electron distribution function in the earth's atmosphere, *Planet. Space Sci*., *24*, 33.

Jasperse, J. R., and B. Basu, 1982, Transport theoretic solutions for auroral proton and H atom fluxes and related quantities, *J. Geophys. Res*., *87*, 811.

Kozelov, B. V., 1993, Influence of the dipolar magnetic field on transport of proton-H atom fluxes in the atmosphere, *Ann. Geophys*., *11*, 697.

Kunches, J. M, and R. D. Zwickl, 1999, The effects of coronal holes on the propagation of solar energetic protons, *Radiat. Meas*., *30*, 281.

Lanchester, B. S., J. R. Palmer, M. H. Rees, D. Lummerzheim, K. Kaila, and T. Turunen, 1994, Energy flux and characteristic energy of an elemental auroral structure, *Geophys. Res. Lett.*, *21*, 2789.

Lanchester, B. S., M. H. Rees, D. Lummerzheim, A. Otto, H. U. Frey, and K. U. Kaila, 1997, Large fluxes of auroral electrons in filaments of 100m width, *J. Geophys. Res.*, *102*, 9741.

Lanchester, B. S., M. Galand, S. C. Robertson, M. H. Rees, D. Lummerzheim, I. Furniss, L. M. Peticolas, H. U. Frey , J. Baumgardner, and M. Mendillo, 2003, High resolution measurements and modeling of auroral hydrogen emission line, *Ann. Geophys.*, *21*, 1.

Link, R., 1992, Feautrier solution of the electron transport equation, *J. Geophys. Res.*, *97*, 159.

Liou, K., P. T. Newell, C.-I. Meng, M. Brittnacher, and G. Parks, 1997, Synoptic auroral distribution: A survey using Polar ultraviolet imagery, *J. Geophys. Res., 102*, 21, 197.

Lorentzen, D. A., F. Sigernes, and C. S. Deehr, 1998, Modeling and observations of dayside auroral hydrogen emission Doppler profiles, *J. Geophys. Res.*, *103*, 17, 479.

Lummerzheim, D., and J. Lilensten, 1994, Electron transport and energy degradation in the ionosphere: Evaluation of the numerical solution, comparison with laboratory experiments and auroral observations, *Ann. Geophys.*, *12*, 1039.

Lummerzheim, D., M. Brittnacher, D. Evans, G. A. Germany, G. K. Parks, M. H. Rees, and J. F. Spann, 1997, High time resolution study of the hemispheric power carried by energetic electrons into the ionosphere during the May 19/20, 1996 auroral activity, *Geophys. Res. Lett.*, *24*, 987.

Lummerzheim, D., and M. Galand, 2001, The profile of the hydrogen Hβ emission line in proton aurora, *J. Geophys. Res.*, *106*, 23.

Lummerzheim, D., M. Galand, J. Semeter, M. J. Mendillo, M. H. Rees, F. J. Rich, 2001, Emission of OI(630 nm) in proton aurora, *J. Geophys. Res.*, 106, 141.

McEwen, D. J., and D. A. Harrington, 1991, Polar airglow and aurora, *Can. J. Phys., 69,* 1055.

Mende, S.B., H. Heetderks, H.U. Frey, M. Lampton, S.P. Geller, S. Habraken, E. Renotte, C. Jamar, P. Rochus, J. Spann, S.A. Fuselier, J.-C. Gerard, R. Gladstone, S. Murphree, L. Cogger, 2000, Far Ultraviolet Imaging from the IMAGE Spacecraft: 1. System Design, Space Sci. Rev., IMAGE Special Issue, Vol. *91*, 243.

Min, Q. L., D. Lummerzheim, M. H. Rees, and K. Stamnes, 1993, The Effects of a Parallel Electric Field and the Geomagnetic Field in the Topside Ionosphere on Auroral and Photoelectron Energy Distributions, *J. Geophys. Res.*, *98*, 19,223.

Nagy, A. F., and P. M. Banks, 1970, Photoelectron fluxes in the ionosphere, *J. Geophys. Res*., *75*, 6260.

Newell, P. T. and C.-I. Meng, 1988, The cusp and the cleft/boundary layer: Low-altitude identifications and statistical local time variation, *J. Geophys. Res*., *93*, 14549-14556.

Newell, P. T. and C.-I. Meng, 1990, Intense keV energy polar rain, *J. Geophys. Res.*, 95, 7869.

Newell, P. T., S. Wing, C.-I. Meng, and V. Sigillito, 1991, The auroral oval position, structure, and intensity of precipitation from 1984 onwards: An automated online data base, *J. Geophys. Res*., *96,* 5877-5882.

Newell, P. T., C.-I. Meng, and K. M. Lyons, 1996, Suppression of discrete aurora by sunlight, *Nature,* 381, 766.

Peticolas, L. M., and D. Lummerzheim, 2000, Time-dependent Transport of Field-aligned Bursts of Electrons in Flickering Aurora, *J. Geophys. Res.*, *105*, 12, 895.

Picone, J. M., A. E. Hedin, D. P. Drob, and A. C. Aikin, 2002, NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, *J. Geophys. Res*., DOI 10.1029/2002JA009430.

Rees, M. H., D. Lummerzheim, R. G. Roble, J. D. Winningham, J. D. Craven, and L. A. Frank, 1988, Auroral Energy Deposition Rate, Characteristic Electron Energy, and Ionospheric Parameters Derived from Dynamics Explorer 1 Images, *J. Geophys. Res.*, *93*, 12841.

Rodger, C. J., M. A. Clilverd, P. T. Verronen, T. Ulrich, M. J. Jarvis, and E. Turunen, 2006, Dynamic geomagnetic rigidity cutoff variations during a solar proton event, *J. Geophys. Res*., DOI 10.1029/ 2005JA011395.

Robinson, R. M. and R. R. Vondrak, 1985, Characteristics and sources of ionization in the continuous aurora, *Radio Sci., 20*, 447.

Solomon, S. C., 1993, Auroral electron transport using the Monte Carlo method, *Geophys. Res. Lett.*, *20*, 185.

Stamnes, K., 1980, Analytic approach to auroral electron transport and energy degradation, *Planet. Space Sci.*, *28*, 427.

Stamnes, K., 1981, On the two-stream approach to electron transport and thermalization, *J. Geophys. Res.*, *86*, 2405.

Stenbaek-Nielsen, H. C., T. J. Hallinan, D. L. Osborne, J. Kimball, C. Chaston, J. McFadden, G. Delory, M. Temerin, and C. W. Carlson, 1998, Aircraft observations conjugate to FAST of auroral arc thicknesses, *Geophys. Res. Lett., 25*, 2073.

Stephan, A.W., S. Chakrabarti, and D. M. Cotton, 2000, Evidence of ENA precipitation in the EUV dayglow, *Geophys. Res. Lett., 27*, 2865.

Stephan, A.W., K.F. Dymond, S.A. Buszien, S. E. Thonnard, and R.P. McCoy, 2004, Oxygen aurora during the recovery phase of a major geomagnetic storm, *J. Geophys. Res., 109*, A09208.

Strickland, D. J., D. L. Book, T. P. Coffey, and J. A. Fedder, 1976, Transport equation techniques for the deposition of auroral electrons, *J. Geophys. Res.*, *81*, 2755.

Strickland, D.J., R. Link, and L.J. Paxton, 1992, Far UV remote sensing of thermospheric composition, *Ultraviolet Technology IV, SPIE vol. 1764,* 117-131.

Strickland, D. J., R. E. Daniell Jr., J. R. Jasperse, and B. Basu, 1993, Transport-theoretic model for the electron-proton-hydrogen atom aurora, 2, model results, 1993, *J. Geophys. Res.*, *98*, 21,533.

Synnes, S. A., F. Søraas, and J. P. Hansen, 1998, Monte-Carlo simulation of proton aurora, *J. Atmos. Sol.-Terr. Phys.*, *60*, 1695.

Tanaka, Y.-M., M. Ishii, M. Kubota, Y. Monzen, Y. Murayama, H. Mori, and D. Lummerzheim, 2005, Two-dimensional distribution of medium-energy precipitating electrons estimated from optical and CNA measurements, *Ann. Geophys., 23*, 1555.

Winningham, J. D., and W. J. Heikkila, 1974, Polar cap auroral electron fluxes observed with ISIS 1, *J. Geophys. Res.*, *79*, 949.

Winningham, J. D., J. R. Sharber, R. A. Frahm, J. L. Burch, N. Eaker, R. K. Black, V. A. Belvins, J. P. Andrews, J. Rudzki, D. L. Chenette, D. W. Datlowe, E. E. Gaines, W. I. Imhof, R. W. Nightingale, S. S. Prasad, R. M. Robinson, T. L. Schumaker, R. R. Vondrak, P. F. Bythrow, B. J. Anderson, T. A. Potemra, L. J. Zanetti, D. B. Holland, M. H. Rees, D. Lummerzheim, G. C. Reid, R. G. Roble, C. R. Clauer, and P. M. Banks, 1993, The UARS Particle Environment Monitor, *J. Geophys. Res.*, *98*, 10, 649.

1.3.3 Ionospheric Plasma Convection

(Mike Ruohoniemi, The John Hopkins University Applied Physics Laboratory)

1.3.3.1 Introduction

The ionosphere is the lowest level of the atmosphere that displays significant electrical properties. Even so, the ionized component constitutes only a very small percentage of the total mass. In the absence of electrical and magnetic fields the motion of the charged particles would be dictated by the dynamics of the neutral atmosphere. With the imposition of a magnetic field alone, the motion of the charged particles becomes subject to the $q\mathbf{v}\times\mathbf{B}$, or Lorentz, force, which compels them to gyrate about field lines. When an electric field is added, the particles acquire a drift in the $\mathbf{E}\times\mathbf{B}$ direction irrespective of particle charge. The motion of ionospheric plasma under the influence of electric field perpendicular to **B** constitutes convection. At high latitudes the geomagnetic field lines are almost vertical and so the plasma convection is effectively confined to the horizontal plane. At their upper ends the field lines are in contact with large-scale electric fields that are generated by the flow of the solar wind past the Earth. These fields map with

little attenuation along the magnetic field lines into the ionosphere. The **E**×**B** motion of plasma in the outer magnetosphere is thus projected onto a high-latitude ionospheric 'screen'. The same connection provides a path for energetic particles to precipitate from the magnetosphere and cause the aurora. Both the large-scale ionospheric precipitation and convection can be viewed advantageously within a reference frame that is centered on the geomagnetic pole and extends down in latitude to an auroral boundary that is typically at 60 deg Λ.

Figure 1-8 shows a model of convection as envisioned originally by *Dungey* [1961]. The solid lines represent contours of constant electrostatic potential, Φ, to which the electric field is related via **E** = -gradΦ. The plasma circulates along the contours in a manner analogous to the movement of air along isobars of constant pressure in conventional weather maps. The large-scale pattern of convection consists of two cells with antisunward flow over the polar cap closed by return flows along the dawn and dusk flanks. *Dungey* [1961] hypothesized that such a circulation would result from a process on the dayside that merges solar and geomagnetic magnetic field lines with an accompanying transfer of energy and momentum into the magnetosphere. *Axford and Hines* [1961] suggested that a viscous interaction takes place between the solar wind and magnetospheric plasma along the magnetospheric flanks; a similar pattern of large-scale circulation is obtained.

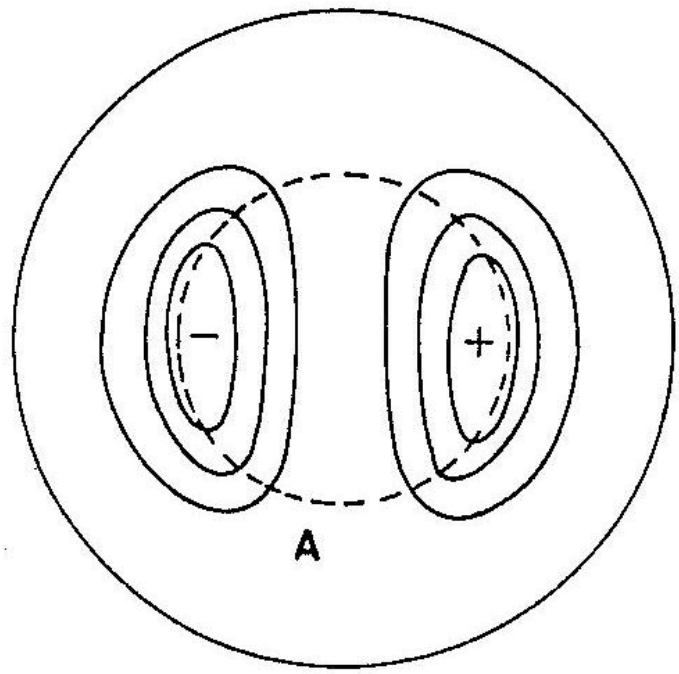

Figure 1-8: Theoretical Pattern of Ionospheric Plasma Convection as Proposed by *Dungey* [1961].

At altitudes within the *E* region (90-150 km), the electrons partake fully in **E**×**B** convection but the ions are partially or fully immobilized by collision with neutrals. The velocity difference gives rise to an electrical current. The ionosphere functions as a resistive load on the magnetospheric generators of convection because collisions between charged particles and neutrals dissipate energy through heating. Convection also redistributes ionization and plasma structure throughout the convection zone. The velocities associated with high-latitude convection are typically hundreds of meters per second, often exceed 1500 m/s, and can reach 3-4 km/s. The redistribution of plasma under the control of convection electric field takes place on time scales of hours to days.

1.3.3.2 Methods of Measuring Ionospheric Plasma Convection

A variety of methods have been applied to study the **E**×**B** drift of ionospheric plasma. Some of the first direct observations were made with rockets, e.g., *Haerendel et al.* [1967]. These released clouds of barium gas that rapidly ionized and then drifted at the convection velocity. In situ measurements are now more commonly obtained from satellites that carry instruments to detect ion drift. These have included Ogo 6 [*Heppner*, 1972], Atmosphere Explorer C [*Heelis*, 1984], and Dynamics Explorer 2 [*Heelis et al*., 1986]. The DMSP series of low-altitude satellites [*Hairston and Heelis*, 1995] in particular have provided nearly continuous observations for many years. Ionospheric convection can also be monitored from the ground using remote sensing techniques. Powerful incoherent-scatter radars generate backscatter from thermal fluctuations in the ambient plasma and analysis of the spectra yields information on the **E**×**B** drift as well as on many other ionospheric parameters [*Banks et al.*, 1973]. A less direct ground-based method utilizes magnetometers. These instruments detect magnetic disturbances at the ground that can be related to the currents flowing in the *E* region ionosphere. The convection is obtained by deconstructing the current signatures into electric field and conductivity components [*Richmond et al.*, 1988].

Another radar method involves coherent backscatter from irregularities in the ionospheric ionization that have been amplified far beyond thermal levels by plasma instability processes. Much lower levels of power are required than is the case for incoherent scatter but the technique does require that irregularities exist and that the radar signal encounters them while propagating nearly orthogonal to the magnetic field lines. The Doppler shift in the backscattered signal characterizes one component of the **E**×**B** drift. One early system operated on VHF frequencies in Scandinavia and used pairs of receivers to resolve stereoscopically the two-dimensional convection velocity vector [Greenwald et al., 1978]. The geometry of the geomagnetic field restricts VHF operations to the E region ionosphere. A system based on operations at HF frequencies was subsequently developed [Greenwald et al., 1985]. The advantage of HF is that the vertical density gradients in the ionosphere can refract the transmissions to propagate orthogonally to geomagnetic field lines at F region altitudes (150-1000 km); this vastly increases the area that is accessible to coherent scatter. An international collaboration known as SuperDARN was assembled to pursue development of chains of HF radars in both the northern and southern hemispheres [Greenwald et al., 1995a].

Figure 1-9 shows the sites and fields of view of the radars as of early 2005. For the most part the radars have been organized into pairs with overlapping fields of view to facilitate stereoscopic observation of the convection velocity. Data from the radar chains are routinely merged with statistical model information and solved for a best-fit image of the large-scale convection pattern in the high-latitude ionosphere [Ruohoniemi and Baker, 1998; Shepherd and Ruohoniemi, 2000]. A 'nowcast', or near real-time specification, of the convection pattern in the northern hemisphere is generated using internet links to the northern radars [Ruohoniemi et al., 2001] and posted to a public access web site (http://superdarn.jhuapl.edu/). The images are generated continuously at a cadence of 1 or 2 minutes. Figure 1-10 shows an example of a convection map that was derived for January 21, 2001, 1540-1542 UT. Velocity vectors are drawn where direct measurements were available from at least one of the radars. As indicated, the prevailing IMF was predominantly southward and a two-cell pattern was obtained. The total variation in the electrostatic potential across the polar cap, ΦPC, was 67 kV. The value of ΦPC typically ranges from 15 to 300 kV, with the lower end encountered for prolonged periods of northward IMF that generally lead to geomagnetic quiescence while the higher end is associated with very large geomagnetic disturbances.

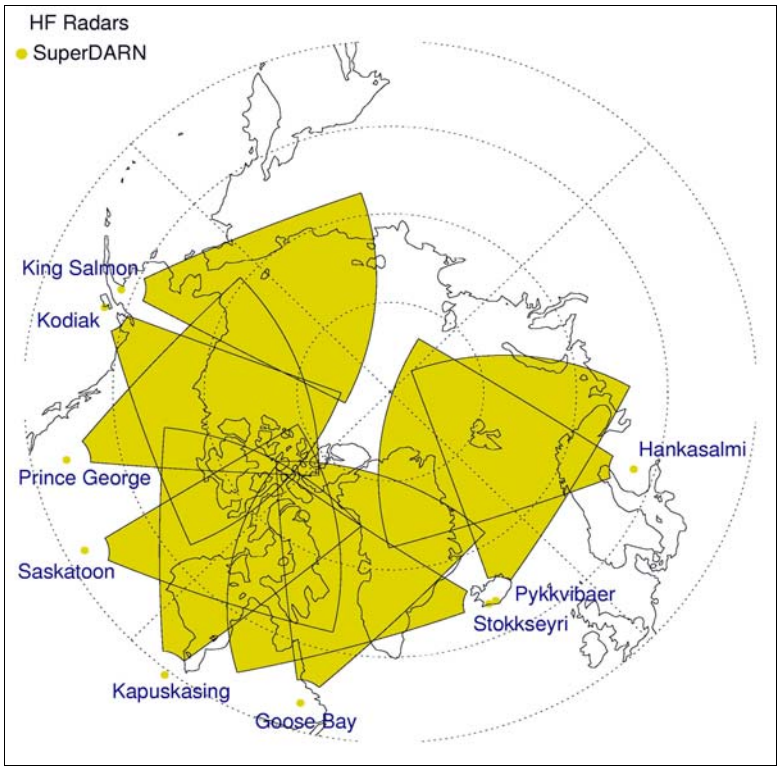

Figure 1-9: Fields of View of the Northern Component of the Super Dual Auroral Radar Network as of Early 2005.

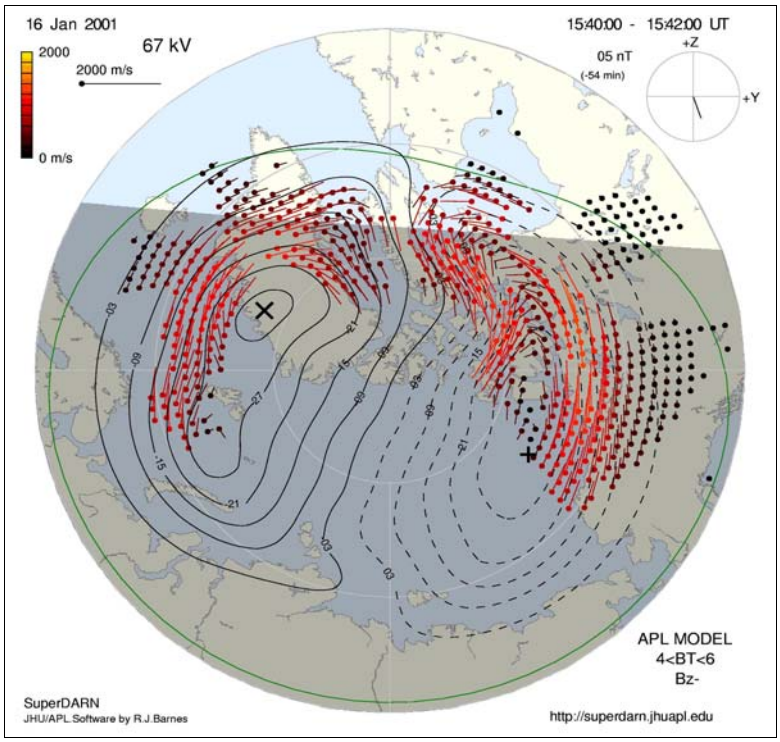

Figure 1-10: An Example of a Map of Ionospheric Plasma Convection Derived from Observations with the SuperDARN HF Radars as Described in the Text. The circular inset map in the upper right hand corner indicates the orientation of the IMF in the GSM *y-z* plane.

1.3.3.3 Dynamics of Plasma Convection

The convection of ionospheric plasma at high latitudes most often conforms to a two-cell pattern of circulation with antisunward flow over the polar cap. It is thought that the merging process envisioned by *Dungey* [1961] usually accounts for more of the convection than the viscous interaction mechanism advanced by *Axford and Hines* [1961]. The essential factors in setting the convection morphology is that geomagnetic field lines merge on the dayside magnetopause and are then subject to an antisunward tug due to their projections into the solar wind. As shown in Figure 1-11, the process of merging, also called reconnection, transforms a geomagnetic field line that was 'closed' between the hemispheres into a field line that 'opens' into interplanetary space. The field line is dragged across the pole and the plasma attached to the line shares in this motion. Open magnetic flux accumulates in the polar cap or is returned to a closed condition by nightside reconnection processes that can include the explosive auroral substorm. The cycle is completed by flows along the dawn and dusk flanks that return closed flux and plasma to the dayside.

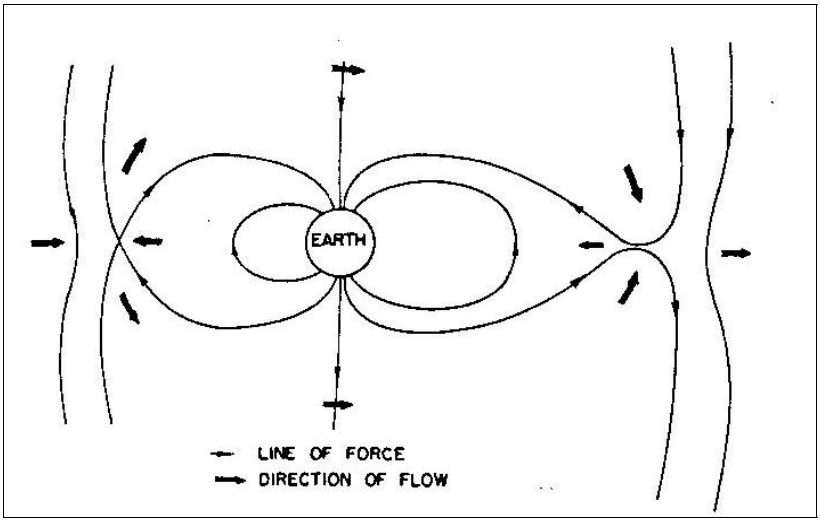

Figure 1-11: Schematic of the Process whereby Geomagnetic Lines on the Dayside Merge with Southward Interplanetary Magnetic Field. The open field lines are carried tailward by the solar wind plasma and are eventually returned to closed dipolar form by merging across the magnetotail – from *Dungey* [1961].

A vital aspect of the merging process is the orientation of the solar wind field line at the magnetopause; merging is favored when the IMF and geomagnetic field lines are anti-parallel. As the geomagnetic field line connects between the hemispheres in a south-to-north sense, *Dungey* [1961] proposed his merging scenario for a southward orientation of the IMF. Later, he considered the possibility of merging for northward IMF [*Dungey*, 1963]. The geomagnetic field at the magnetopause converges at a point and forms a funnel-like feature called the cusp that descends into the high-latitude ionosphere. The convoluted geometry of the cusp field lines combined with tilting of the Earth leads to anti-parallel merging sites even for northward IMF. However, in this case the initial tug felt by the merged field lines is in the sunward direction owing to the stresses on the highly kinked field lines. Eventually the motion of the solar wind dominates and the flow turns antisunward. In the vicinity of the cusp, two small convection cells form that cycle plasma in the reverse sense, i.e., sunward at noon and antisunward along the dawn and dusk flanks, with sunward return flow from midnight to noon. When coupled with weak viscous cell convection of the usual sense, a four-cell pattern of circulation is obtained. If a cell is very weak it may practically disappear, leaving a three-cell pattern.

Other merging scenarios for driving high-latitude convection are possible (see *Crooker* [1992] for a review). For example, within the polar cap, the open field lines may themselves present merging sites to

the incoming IMF, leading to open-to-open conversions where a geomagnetic field line exchanges its extension into the solar wind in its own hemisphere for a line that overdrapes the magnetopause on the dayside and extends into the solar wind in the opposite hemisphere. The resulting plasma motion, which is reversed from the usual two-cell sense, amounts to a continual stirring of open field lines within the polar cap and is referred to as lobe cell convection. The character of the global convection (two-cell, distorted two-cell, multi-cell, reverse, etc.) is observed to hinge on the relative magnitudes of the north-south and azimuthal components of the IMF, e.g., *Knipp et al.* [1993].

These considerations indicate the paramount importance that attaches to the solar wind, especially the IMF, in determining plasma convection. Many measurement techniques have been applied to study the nature of the IMF effect. These have lead to numerous statistical studies that describe the average pattern of ionospheric plasma convection as a function of IMF, e.g., *Heppner and Maynard* [1987], *Holt et al.* [1987], *Papitashvili et al.* [1994], *Weimer* [1995], *Ruohoniemi and Greenwald* [1996]. Figure 1-12 shows an example from *Ruohoniemi and Greenwald* [1996] that was derived on the basis of measurements from the prototype SuperDARN HF radar located at Goose Bay, Labrador. Note the intensification of flows and the enlargement of the two cell pattern for increasingly southward IMF and the emergence of more convoluted dayside flows for increasingly northward IMF. The azimuthal (east-west) component of the IMF is seen to play a major role in orienting the dayside flows and defining the shapes of the cells. Figure 1-13 shows the average pattern derived for strongly northward IMF conditions; two small but distinct reverse dayside convection cells emerge bounded at lower latitude by cells with the usual sense of circulation. It turns out that the various studies are remarkably consistent in their depictions of the average IMF dependencies. The statistical characterization of the convection pattern in terms of the IMF can be considered solved and interest has moved on to the impact of secondary factors, such as season, e.g., *Ruohoniemi and Greenwald* [2005].

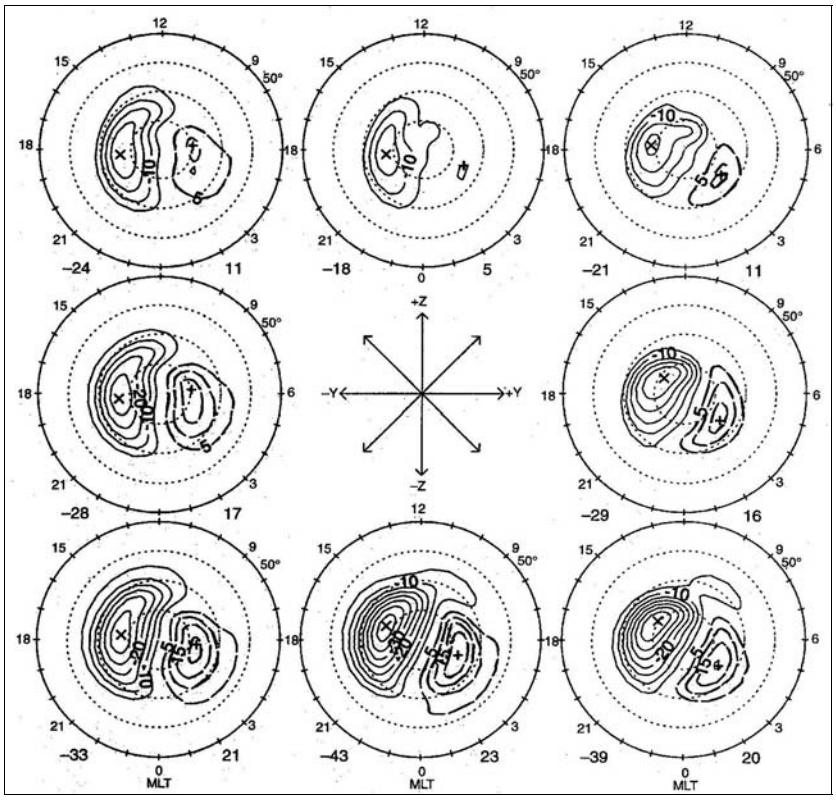

Figure 1-12: Average Patterns of Ionospheric Plasma Convection Sorted by the Orientation of the IMF in the GSM y-z Plane for a Moderate Level of Geomagnetic Disturbance. The values of the potential extrema are indicated (in kV) – from *Ruohoniemi and Greenwald* [1996].

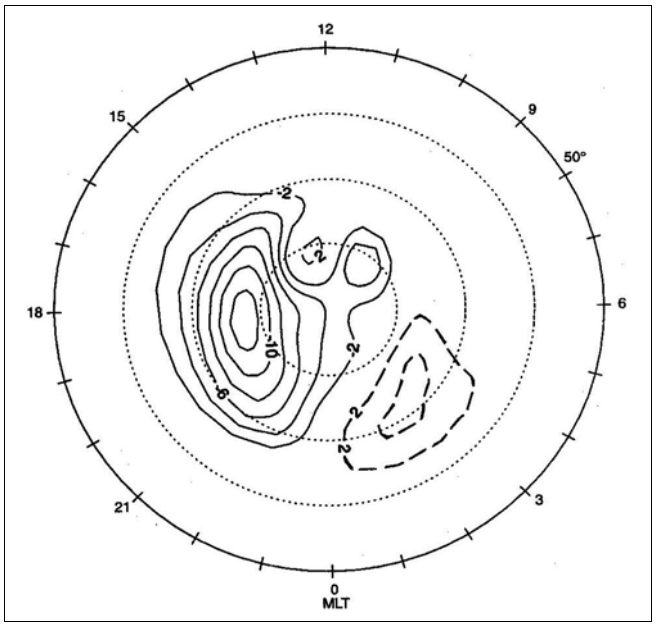

Figure 1-13: Average Pattern of Ionospheric Plasma Convection Derived for Strongly Northward IMF – from *Ruohoniemi and Greenwald* [1996].

It has been more challenging to study and understand the instantaneous convection pattern, i.e., that which prevails in the ionosphere at an instant in time. It has not been possible to make observations everywhere all at once although the incorporation of individual instruments into larger networks such as SuperDARN is greatly increasing our capabilities. The first large-scale depictions of instantaneous convection patterns were generated with the AMIE technique which relies on the inversion of magnetometer data [*Richmond and Kamide*, 1988]. These illustrated that the instantaneous convection was quite variable and sensitive to IMF variations, e.g., *Knipp et al.* [1991]. An early triumph of the SuperDARN concept was the first direct mapping of a reverse dayside convection cell for northward IMF [*Greenwald et al.*, 1995*b*]. With the continuing improvement in coverage, the nature of the convection response to changed IMF came under intense scrutiny. It had been argued that the initial response in ionospheric convection would be confined to the dayside cusp region where the plasma has the most direct connection to conditions at the magnetopause and then progress towards the nightside with delays measured in the tens of minutes to hours, reflecting the relatively slow progress of convecting flux tubes. However, observations with distributed sets of magnetometers and HF radars indicated that the convection begins to respond almost simultaneously (within minutes) at all longitudes [*Ridley et al.*, 1998; *Ruohoniemi et al.* 1998]. A lively controversy ensued [*Lockwood*, 1999; *Ridley et al.*, 1999]. Continued observations with the expanding instrument networks decided the issue in favor of simultaneous onset [*Murr and Hughes*, 2001; *Ruohoniemi et al.*, 2002]. The manner in which ionospheric convection evolves following the onset of a response remains an active study area, e.g., *Lu et al.* [2002].

Plasma convection is in general a highly variable phenomenon. Part of the variability is attributable to variability in the solar wind driver. The magnitude and direction of the IMF vary typically on time scales of hours but can occur on steps as short as one minute. Advance knowledge of the IMF gives a basis for predicting the convection using one of the IMF-keyed statistical models discussed earlier. Unfortunately, the most advance warning that can be managed currently is about 1 hour. This is set by the propagation time of the solar wind from the position of the ACE satellite at the Lagrangian L1 point to effects at the magnetopause. Furthermore - and very significantly for current directions in research - the convection is not completely determined by the IMF or solar wind factors. Marked variability in convection for stable IMF has been noted in both fine-scale [*Codrescu et al.*, 2000; *Shepherd et al.*, 2003] and large-scale

[*Matsuo et al.*, 2003; *Bristow et al.*, 2004] observations. Indeed, the ability of the standard statistical convection models to predict the instantaneous local convection is not very high, e.g., *Eriksson et al.*, [2002]. This likely points to the influence of factors internal to the magnetosphere-ionosphere (M-I) system. It is perhaps not surprising that, as an auroral phenomenon, ionospheric plasma convection at high latitudes exhibits pronounced variability in time and space. There are notable impacts on convection during auroral substorm activity, e.g., Provan et al. [2004] that appear to be a consequence of balancing electric fields and conductivities in the ionosphere with currents supplied by the magnetosphere. Understanding the origins and effects of the variability in convection is a very active research area and presents a particular challenge for efforts to model the coupling of the solar wind to the M-I system using MHD codes.

1.3.3.4 Conjugacy

The northern and southern hemispheres are linked by the lines of force of the geomagnetic field. This coupling is significant as charged particles move relatively freely along the field lines and so auroral processes tend to mirror between the hemispheres in regions that are magnetically conjugate. Ionospheric convection also exhibits conjugacy in that the electrostatic potential, Φ, associated with the large-scale convection electric fields maps between the hemispheres along the geomagnetic field lines. Thus convection in the northern hemisphere should mirror in the southern hemisphere within the region of closed field lines. At the higher latitudes associated with the polar cap the field lines are open to the IMF in the solar wind and interhemispheric conjugacy cannot be expected.

The mapping of geomagnetic flux between the hemispheres varies with season and UT and reflects the secular variation in the geomagnetic field, e.g., *Stassinopoulos et al.*, [1984]. Surprisingly perhaps, given the long history of observations in Antarctica, the extent of conjugate behaviour in auroral phenomenon has still not been fully worked out. This is especially true of ionospheric plasma convection. Of course, proper conjugacy is limited to field lines that close between the hemispheres and convection dynamics are largely controlled by processes that involve the transformation of field lines between closed and open. There is also the difficulty of determining which ionospheric plasma parcels are truly conjugate as the geomagnetic field lines can depart dramatically from their nominal dipolar geometry especially in the vicinity of the open/closed field line boundary.

An overall view of the relation between convection in the northern and southern hemispheres can be obtained by comparing convection patterns. It was noted by *Heppner and Maynard* [1987] that the dayside ionospheric flows conform to signatures ordered by the sign of the azimuthal component of the IMF, i.e., B_y, and that the convection in the northern hemisphere for B_y+ resembles the convection in the southern hemisphere for B_y- and vice versa. The first statistical convection model derived exclusively for the southern hemisphere was based on observations with an HF radar located at Halley Station, Antarctica [*Leonard et al.*, 1995]. Comparison with the northern observations confirms that the large-scale convection pattern is asymmetric between the hemispheres with respect to the sign of B_y. This is related to the anti-parallel condition for merging between the IMF and the geomagnetic field at the magnetopause and the reversal in the direction of the azimuthal component of the geomagnetic field lines between the hemispheres. The B_y-sign factor was demonstrated in instantaneous convection by *Greenwald et al.* [1990], who observed the emergence of distinctive cusp flows in the two hemispheres following changes in the sign of B_y using observations from the Goose Bay and Halley Station HF radars. Figure 1-14 shows an example. Flows on the dayside were oriented more westward (dawnward) in the northern/southern hemisphere for B_y+/B_y- conditions but acquired a pronounced shear in latitude when the sign of B_y subsequently reversed. An analogous interhemispheric asymmetry in nightside convection was recently reported by *Grocott et al.* [2005].

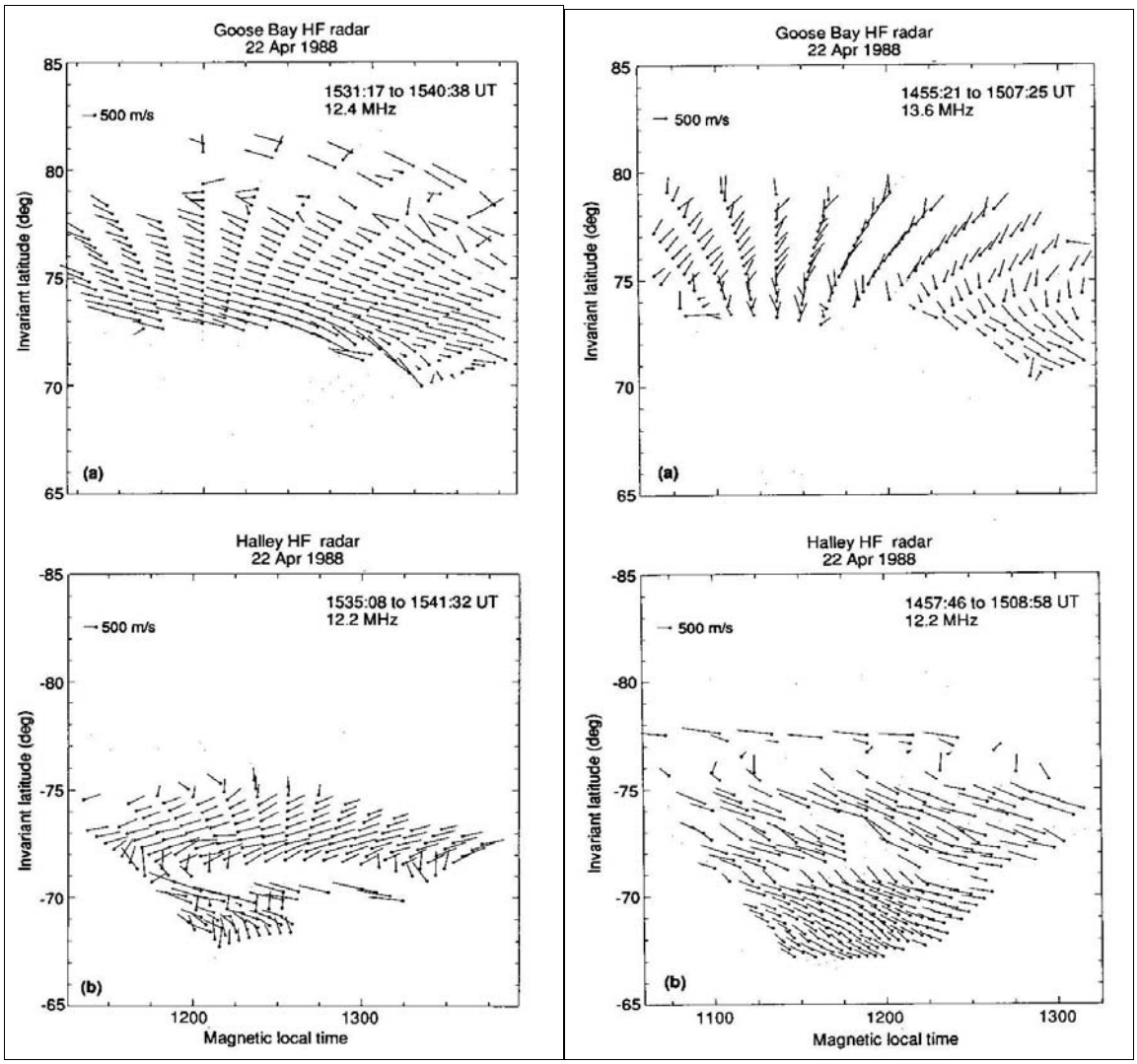

Figure 1-14: Instantaneous Patterns of Dayside Plasma Convection Obtained in Conjugate Hemispheres for Distinct B_y Conditions.

For B_y+, the northern radar (Goose Bay) observes westward flow while the southern radar (Halley) observes a shear reversal in latitude. For subsequent B_y- conditions, the flows in the northern hemisphere are sheared in latitude while those in the southern are westward and poleward – after Greenwald et al. [1991].

1.3.3.5 Impacts of Convection

Convection has a range of significant effects in the high-latitude ionosphere. At *F* region altitudes (> 200 km) the lifetime of ionization is fairly long (~hours) and plasma can be transported large distances with little change. Consequently, the distribution of high-latitude plasma density at any instant represents the cumulative effect of convection acting on the sources of ionization over many hours. Since convection pattern varies on time scales of minutes to hours, redistribution of plasma density can be expected over a range of spatial scales and there are many possibilities for mixing dissimilar plasma populations to create structure. Figure 1-15 shows a particularly dramatic example of structuring due to *Foster et al.* [2005] who showed in combined GPS TEC/convection observations the propagation of a storm-enhanced density feature (SED) from mid to high latitudes during a major geomagnetic storm. Because of convection the ionization overhead at a location deep on the nightside can in fact be due to processes that took place hours earlier on the dayside.

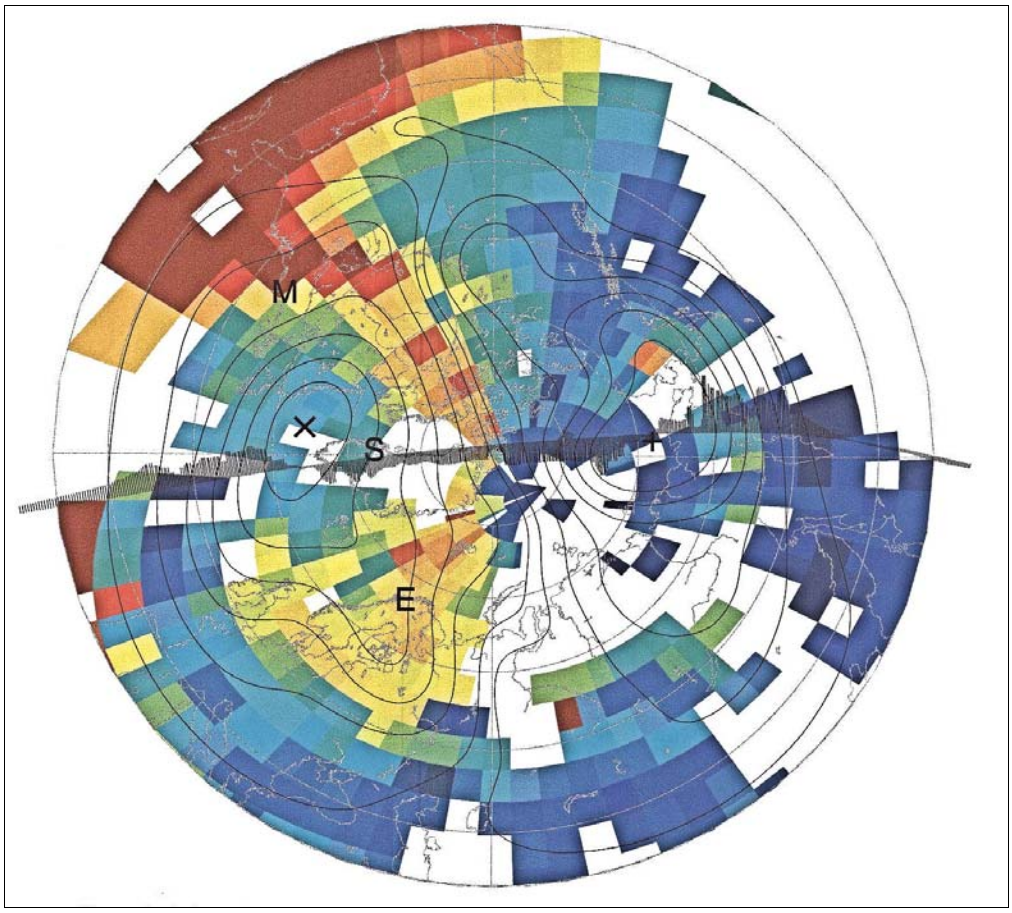

Figure 1-15: Map of GPS Total Electron Content (TEC) Data in Polar Projection Showing a Storm-Enhanced Density Plume Extending Continuously from a Low-Latitude Source Region in the Postnoon Sector through the Polar Cap into the Midnight Sector.

The locations of large incoherent scatter radar facilities that observed related effects are shown by the letters (M: Millstone Hill, S: Sondrestrom, E: EISCAT). Overplotted are convection contours deduced from combined SuperDARN radar and DMSP satellite observations – after *Foster et al.* [2005].

Large-scale models of the ionosphere attempt to account comprehensively for the condition of the high-latitude ionosphere by incorporating the sources of ionization, convection, and chemical effects in numerical constructs with long timelines, e.g., the Utah State University (USU) model described by *Schunk and Sojka* [1987]. The evolution of enhancements in ionospheric ionization known as patches has been modeled. These features form on the dayside by processes that may involve convection dynamics (e.g., *Valladares et al.* [1994]) and are transported by convection across the polar cap to the nightside where they become elongated in longitude as determined by the convection trajectories and form what are known as boundary blobs, e.g., *Anderson et al.*, [1996], *Crowley et al.* [2000]. There are currently major efforts to model the global ionosphere in near real-time using, among other inputs, as much information on convection as can be obtained, e.g., *Schunk et al.* [2004].

In addition to transporting bulk ionospheric plasma, the electric fields associated with convection contribute to plasma instabilities that generate irregularities in the ionization. Most often the irregularities are thought of as waves in the plasma that range in scale from tens of kilometers to centimetres. Other factors that contribute to the formation of irregularities include gradients in plasma density and particle temperatures (see reviews by *Fejer and Kelley* [1980] and *Keskinen* [1983]). The irregularities can be observed with

radars that operate on wavelengths that generate Bragg scattering from the plasma waves, generally on frequencies from HF to UHF. For conventional surveillance and communications radars the irregularities cause clutter that can severely impair operations. GPS systems experience phase fluctuations when transiting affected volumes and their performance is degraded. Efforts to understand the interplay of convection and plasma structure in generating irregularities are becoming more sophisticated with the increasing availability of high powered computer systems for numerical modeling, e.g., Guzdar et al. [2001].

1.3.3.6 References

Anderson, D.N., D.T. Decker, and C.E. Valladares, Modeling boundary blobs using time varying convection, *Geophys. Res. Lett.*, *23*, 579-582, 1996.

Axford, W.I., and C.O. Hines, A unifying theory of high-latitude geophysical phenomena, *Canadian Journal of Physics*, *39*, 1433-1464, 1961.

Baker, J.B.H., R.A. Greenwald, J.M. Ruohoniemi, M. Forster, G. Paschmann, E.F. Donovan, N.A. Tsyganenko, J.M. Quinn, and A. Balogh, Conjugate comparison of Super Dual Auroral Radar Network and Cluster electron drift instrument measurements of ExB plasma drift, *Journal of Geophysical Research*, *109*, doi: 10.1029/2003JA009912, 2004.

Banks, P.M., J.R. Doupnik, and S.I. Akasofu, Electric field observations by incoherent scatter in the auroral zone, *J. Geophys. Res.*, *78*, 6607, 1973.

Bristow, W.A., R.A. Greenwald, S.G. Shepherd, and J.M. Hughes, On the observed variability of the cross-polar cap potential, *J. Geophys. Res.*, *109*, A02203, doi: 10.1029/2003JA010206, 2004.

Codrescu, M.V., T.J. Fuller-Rowell, J.C. Foster, J.M. Holt, and S.J. Cariglia, Electric field variability associated with the Millstone Hill electric field model, *Journal of Geophysical Research*, *105*, 5265-5273, 2000.

Crowley, G., A.J. Ridley, D. Deist, S. Wing, D.J. Knipp, B.A. Emery, J. Foster, R. Heelis, M. Hairston, and B.W. Reinisch, Transformation of high-latitude ionospheric F region patches into blobs during the March 21, 1990, storm, *J. Geophys.l Res.*, *105*, 5125-5230, 2000.

Dungey, J.W., Interplanetary magnetic fields and the auroral zones, *Physical Review Letters*, *6*, 47-48, 1961.

Dungey, J. W., The structure of the exosphere or adventures in velocity space, in *Geophysics, The Earth's Environment*, edited by E.C. Dewitt, J. Hieblot, and A. Lebeau, pp. 505-550, Gordon and Breach, New York, 1963.

Eriksson, S., L.G. Blomberg, and D.R. Weimer, Comparing a spherical harmonic model of the global electric field distribution with Astrid-2 observations, *Journal of Geophysical Research*, *107*, 1391, doi:10.1029/2002JA009313, 2002.

Fejer, B.G., and M.C. Kelley, Ionospheric irregularities, *Rev. Geophysics Space Res.*, *18*, 401, 1980.

Foster, J.C., A. J. Coster, P. J. Erickson, J. H. Holt, F. D. Lind, W. Rideout, M. McCready, A. van Eyken, R. J. Barnes, R. A. Greenwald, and F. J. Rich, Multiradar observations of the polar tongue of ionization, *Journal of Geophysical Research*, *110*, doi: 10.1029/2004JA010928, 2005.

Greenwald, R.A., W. Weiss, E. Nielsen, and N.R. Thomas, STARE: a new radar auroral backscatter experiment in northern Scandinavia, *Radio Science*, *12*, 1021-1039, 1978.

Greenwald, R.A., K.B. Baker, R.A. Hutchins, and C. Hanuise, An HF phased-array radar for studying small-scale structure in the high-latitude ionosphere, *Radio Science*, *20*, 63-79, 1985.

Greenwald, R.A., K.B. Baker, J.M. Ruohoniemi, J.R. Dudeney, M.J. Pinnock, N. Mattin, J.M. Leonard, and R.P. Lepping, Simultaneous conjugate observations of dynamic variations in high latitude dayside convection due to changes in IMF B_y, *Journal of Geophysical Research*, *95*, 805708072, 1990.

Greenwald, R.A., K.B. Baker, J.R. Dudeney, M. Pinnock, T.B. Jones, E.C. Thomas, J.-P. Villain, J.-C. Cerisier, C. Senior, C. Hanuise, R.D. Hunsucker, G. Sofko, J. Koehler, E. Nielsen, R. Pellinen, A.D.M. Walker, N. Sato, and H. Yamagishi, DARN/SuperDARN: A global view of high-latitude convection, *Space Science Reviews*, *71*, 763-796, 1995.

Greenwald, R.A., W.A. Bristow, G.J. Sofko, C. Senior, J.-C. Ceriser, and A. Szabo, SuperDual Auroral Radar Network radar imaging of dayside high-latitude convection under northward interplanetary magnetic field: Toward resolving the distorted two-cell versus multicell controversy, *Journal of Geophysical Research*, *100*, 19661-19674, 1995.

Grocott, A., T.K. Yeoman, S.E. Milan, and S.W.H. Cowley, Interhemispheric observations of the ionospheric signature of tail reconnection during IMF-northward non-substorm intervals, *Ann. Geophys.*, *23*, 1763-1770, 2005.

Guzdar, P.N., X. Shao, C.C. Goodrich, K. Papadopoulos, M.J. Wiltberger, and J.G. Lyon, Three-dimensional MHD simulations of the steady state magnetosphere with northward interplanetary magnetic field, *Journal of Geophysical Research*, *106*, 275-287, 2001.

Hairston, M.R., and R.A. Heelis, Response time of the polar ionospheric convection pattern to changes in the north-south direction of the IMF, *Geophysical Research Letters*, *22*, 631-634, 1995.

Haerendel, G.R., R. Lust, and E. Reiser, Motion of artificial ion clouds in the upper atmosphere, *Planetary Space Science*, *15*, 1, 1967.

Heelis, R. A., The effects of interplanetary magnetic field orientation on day side high-latitude convection, *J. Geophys. Res.*, *89*, 2873-2880, 1984.

Heelis, R.A., P.H. Reiff, J.D. Winningham, and W.B. Hanson, Ionospheric convection signatures observed by DE 2 during northward interplanetary magnetic field, *J. Geophys. Res.*, *91*, 5817-5830, 1986.

Heppner, J.P., Polar cap electric field distributions related to the interplanetary magnetic field directions, *J. Geophys. Res.*, *77*, 4877, 1972.

Heppner, J.P., and N.C. Maynard, Empirical high-latitude electric field models, *Journal of Geophysical Research*, *92*, 4467-4489, 1987.

Holt, J.M., R.H. Wand, J.V. Evans, and W.L. Oliver, Empirical models for the plasma convection at high latitudes from Millstone Hill observations, *Journal of Geophysical Research*, *92*, 203-212, 1987.

Keskinen, M.J., and S.L. Ossakow, Theories of high-latitude ionospheric irregularities: A review, *Radio Sci.*, *18*, 1077-1091, 1983.

Knipp, D.J., A.D. Richmond, B. Emery, N.U. Crooker, O.D.L. Beaujardiere, D. Evans, and H. Kroehl, Ionospheric convection response to changing IMF direction, *Geophysical Research Letters*, *18*, 721-724, 1991.

Knipp, D.J., B.A. Emery, A.D. Richmond, N.U. Crooker, M.R. Hairston, J.A. Cumnock, W.F. Denig, F.J. Rich, O.d.l. Beaujardiere, J.M. Ruohoniemi, A.S. Rodger, G. Crowley, B.-H. Ahn, D.S. Evans, T.J. Fuller-Rowell, E. Friis-Christensen, M. Lockwood, H.W. Kroehl, C.G. MacLennan, A. McEwin, R.J. Pellinen, R.J. Morris, G.B. Burns, V. Papitashvili, A. Zaitzev, O. Troshichev, N. Sato, P. Sutcliffe, and L. Tomlinson, Ionospheric convection response to slow, strong variations in a northward interplanetary magnetic field: A case study for January 14, 1988, *Journal of Geophysical Research*, *98*, 19273-19292, 1993.

Leonard, J.M., M. Pinnock, A.S. Rodger, J.R. Dudeney, R.A. Greenwald, and K.B. Baker, Ionospheric plasma convection in the southern hemisphere, *J. Atmos. Terr. Phys.*, 57, 889-897, 1995.

Lockwood, M.A.S.W.H.C., Comment on "A statistical study of the ionospheric convection response to changing interplanetary magnetic field conditions using the assimilative mapping of ionospheric electrodynamics technique" by A. J. Ridley et al., *Journal of Geophysical Research*, *104*, 4387-4391, 1999.

Lu, G., T.E. Holzer, D. Lummerzheim, J.M. Ruohoniemi, P. Stauning, O. Troshichev, P.T. Newell, M. Brittnacher, and G. Parks, Ionospheric response to the interplanetary magnetic field southward turning: Fast onset and slow reconfiguration, *Journal of Geophysical Research*, *107*, 10.1029/2001JA000324, 2002.

Matsuo, T., A.D. Richmond, and K. Hensel, High-latitude ionospheric electric field variability and electric potential derived from DE-2 plasma drift measurements: Dependence on IMF and dipole tilt, *Journal of Geophysical Research*, *108*, 1005, doi:10.1029/2002JA009429, 2003.

Murr, D.L., and W.J. Hughes, Reconfiguration timescales of ionospheric convection, *Geophysical Research Letters*, *28*, 2145-2148, 2001.

Papitashvili, V.O., B.A. Belov, D.S. Faermark, Y.I. Feldstein, S.A. Golyshev, L.I. Gromova, and A.E. Levitin, Electric potential patterns in the northern and southern polar regions parameterized by the interplanetary magnetic field, *Journal of Geophysical Research*, *9*, 13251-13262, 1994.

Provan, G., M. Lester, S.B. Mende, and S.E. Milan, Statistical study of high-latitude plasma flow during magnetospheric substorms, *Annales Geophysicae*, *22*, 3607-3624, 2004.

Richmond, A.D., and Y. Kamide, Mapping electrodynamic features of the high-latitude ionosphere from localized observations: techniques, *J. Geophys. Res.*, *93*, 5741-5759, 1988.

Ridley, A.J., G. Lu, C.R. Clauer, and V.O. Papitashvili, A statistical study of the ionospheric convection response to changing interplanetary magnetic field conditions using the assimilative mapping of ionospheric electrodynamics technique, *Journal of Geophysical Research*, *103*, 4023-4039, 1998.

Ridley, A.J., G. Lu, C.R. Clauer, and V.O. Papitashvili, Reply, *Journal of Geophysical Research*, *104*, 4393-4396, 1999.

Ruohoniemi, J.M., and R.A. Greenwald, Statistical patterns of high-latitude convection obtained from Goose Bay HF radar observations, *Journal of Geophysical Research*, *101*, 21743-21763, 1996.

Ruohoniemi, J.M., and K.B. Baker, Large-scale imaging of high-latitude convection with Super Dual Auroral Radar Network HF radar observations, *J. Geophys. Res.*, *103*, 20797-20811, 1998.

Ruohoniemi, J.M., and R.A. Greenwald, The response of high-latitude convection to a sudden southward IMF turning, *Geophysical Research Letters*, *25*, 2913-2916, 1998.

Ruohoniemi, J.M., R.J. Barnes, R.A. Greenwald, and S.G. Shepherd, The response of the high-latitude ionosphere to the coronal mass ejection event of April 6, 2000: A practical demonstration of space weather nowcasting with the Super Dual Auroral Radar Network HF radars, *Journal of Geophysical Research*, *106*, 30085-30097, 2001.

Ruohoniemi, J.M., S.G. Shepherd, and R.A. Greenwald, The response of the high-latitude ionosphere to IMF variations, *Journal of Atmospheric and Solar-Terrestrial Physics*, *64*, 159-171, 2002.

Ruohoniemi, J. M., and R. A. Greenwald, Dependencies of high-latitude plasma convection: Consideration of interplanetary magnetic field, seasonal, and universal time factors in statistical patterns, *J. Geophys. Res.*, *110*, doi: 10.1029/2004JA010815, 2005.

Schunk, R.W., L. Scherliess, J.J. Sojka, D.C. Thompson, D.N. Anderson, M. Codrescu, C. Minter, T.J. Fuller-Rowell, R.A. Heelis, M. Hairston, and B.M. Howe, Global assimilation of ionospheric measurements (GAIM), *Radio Sci.*, *39*, doi:10.1029/2002RS002794, 2004.

Schunk, R.W., and J.J. Sojka, A theoretical study of the lifetime and transport of large ionospheric density structures, *Journal of Geophysical Research*, *92*, 12343-12351, 1987.

Shepherd, S.G., J.M. Ruohoniemi, and R.A. Greenwald, Direct measurements of the ionospheric convection variability near the cusp/throat, *Geophysical Research Letters*, *30*, 2109, doi:10.1029/2003GL017668, 2003.

Shepherd, S.G., and J.M. Ruohoniemi, Electrostatic potential patterns in the high-latitude ionosphere constrained by SuperDARN measurements, *J. Geophys. Res.*, *105*, 23005-23014, 2000.

Valladares, C.E., S. Basu, J. Buchau, and E. Friis-Christensen, Experimental evidence for the formation and entry of patches into the polar cap, *Radio Science*, *29*, 167-194, 1994.

Weimer, D.R., Models of high-latitude electric potentials derived with a least error fit of spherical harmonic coefficients, *Journal of Geophysical Research*, *100*, 19595-19607, 1995.

1.3.4 Substorms

(Gordon Rostoker, Department of Physics, University of Alberta, Edmonton, Alberta, Canada T6G 2J1 rostoker@space.ualberta.ca)

Abstract

In the science of space weather, one of the most important disturbances under study is the magnetospheric substorm. In this review, the historical development of the study of substorms will be traced, the terminology used to describe the phenomenon will be introduced, and methods of modeling substorms will be presented. While the various frameworks of describing the substorm will not be reviewed here, the underlying physics of the process will be discussed and constraints limiting one's ability to interpret the existing data will be stressed.

1.3.4.1 Early Views of the Substorm

The term substorm was introduced by Akasofu [1] to describe a series of auroral activations observed using allsky cameras located at and around Fairbanks, Alaska. The activations followed a repeatable pattern and the terms used to describe this pattern are in use to this day. Figure 1-16 shows the sequence of events in what Akasofu termed an *auroral substorm*. The development of the substorm was initiated by a brightening of an auroral arc somewhere near the equatorward region of the region of auroras as detected by the allsky

cameras of the day. The initial brightening was followed by a rapid poleward evolution of the disturbed region, with the maximum poleward displacement being reached over a period of time averaging ~ 30 minutes. The period of time during which this poleward expansion took place was termed the *expansion phase* of the substorm. During the expansion *phase*, the western edge of the disturbed region developed a characteristic S-shaped form and was said to expand westward at an average speed of ~ 1km/s: this expanding auroral form was termed the *westward traveling surge*, and was deemed to move westward as far as the dusk meridian and beyond. After some minutes, the auroral arcs in the disturbed region began to drift equatorward, doing so until the original pre-substorm auroral configuration was reached after some tens of minutes. This interval of equatorward drift of auroral arcs was termed the *recovery phase*.

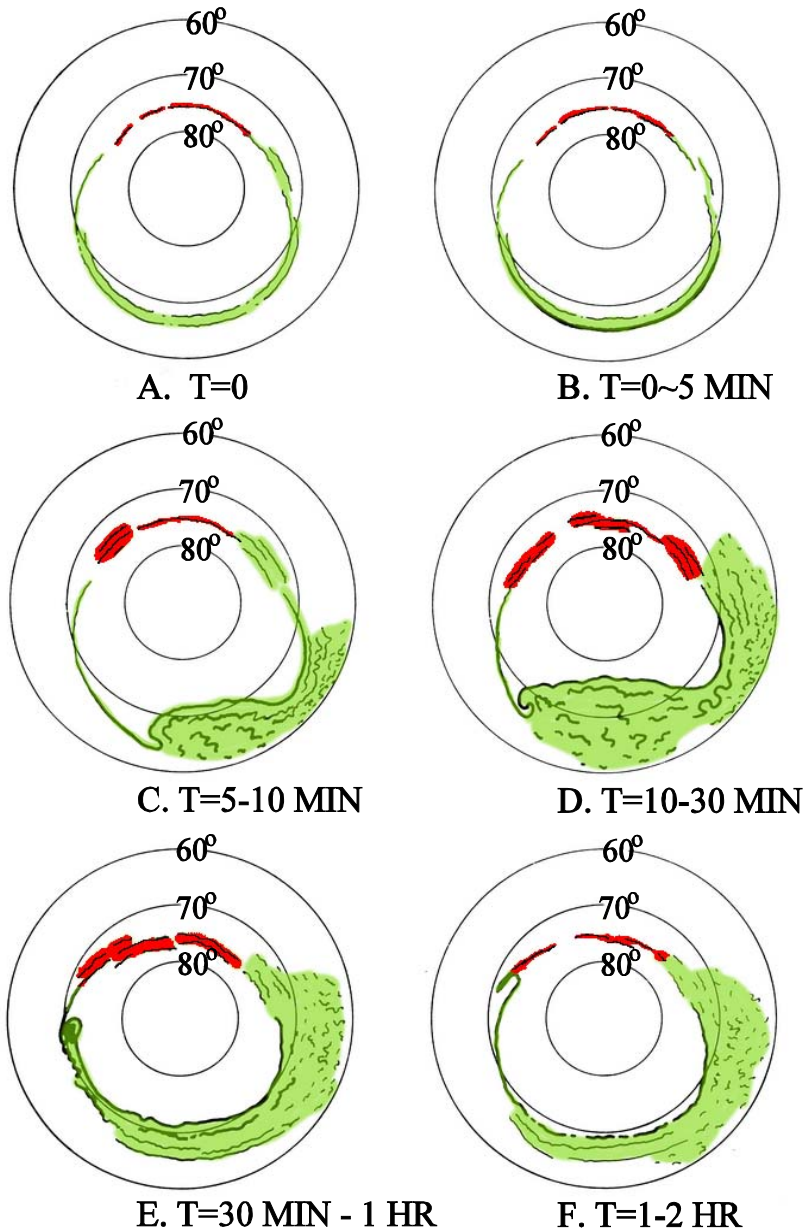

Figure 1-16: Sequence of Development of Discrete Auroral Arcs that Forms the Basis for the Substorm Expansion Phase Phenomenon (modified after Akasofu [1]). Since 1964, information has come to light about the diffuse aurora (green shading) in which the discrete auroral arcs are immersed and the dayside auroras (red shading) which tend to be enriched in red emissions.

It was well recognized that the substorm was accompanied by a marked intensification of electric current in the region of the ionosphere where the substorm arcs were observed. This increase in electric current produced a clear disturbance on a magnetogram (cf. Figure 1-17) which was called a magnetic bay (because its appearance on a magnetogram was similar to a bay on the coastline of a continent). Such disturbances had also been noted by researchers as far back as the late 19th century, and had been given the name *polar elementary storm* by Birkeland [2]. In fact, the auroral substorm was accompanied by many disturbances across the electromagnetic spectrum and Akasofu [3] ultimately brought together all these disturbances under the umbrella of the *magnetospheric substorm*. Within that umbrella, the low frequency magnetic disturbance was re-labeled the *polar magnetic substorm*, and most studies of substorms to this day have relied on the auroral and magnetic signatures of the substorm to define its presence or absence.

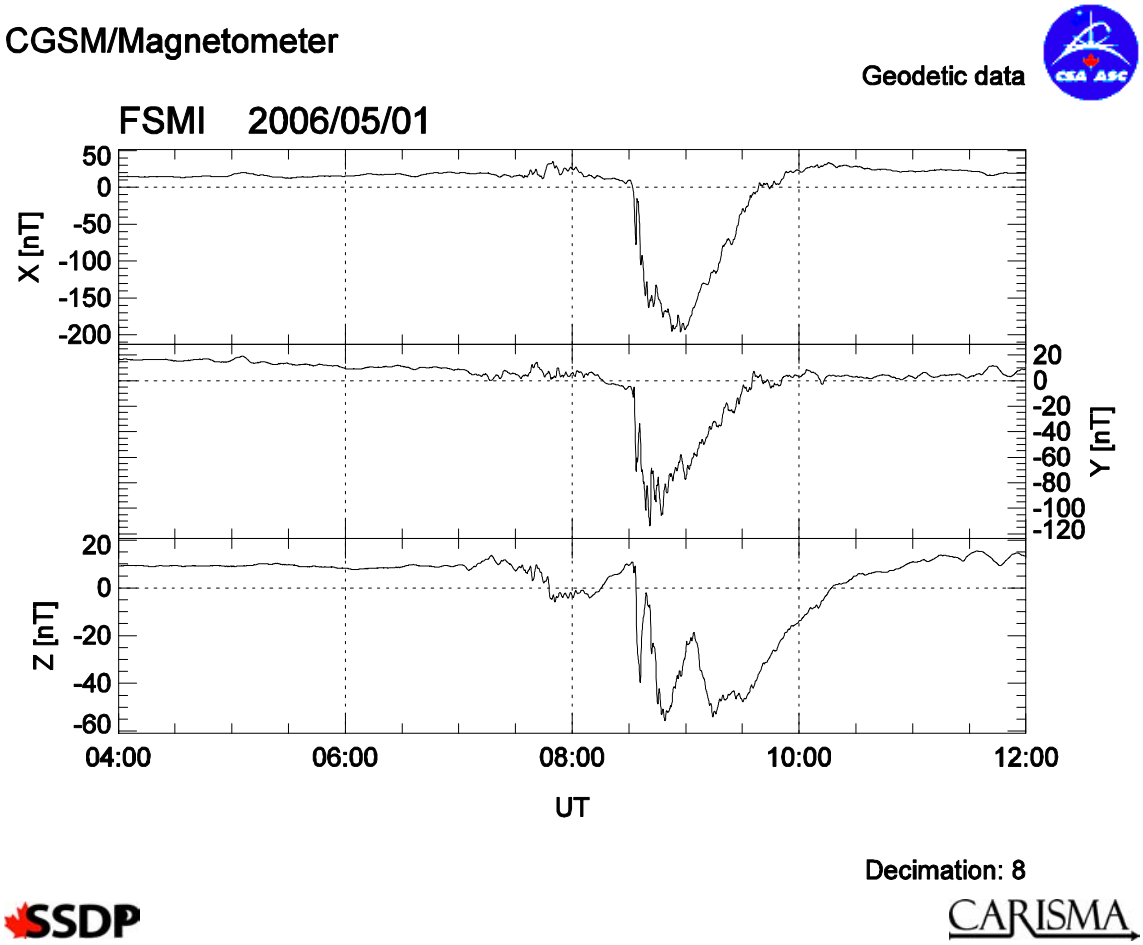

Figure 1-17: Typical Auroral Zone Magnetogram from Fort Smith, NWT Canada Showing the Development and Subsequent Decay of a Polar Magnetic Substorm. The shape of the time series of the negative X-component disturbance between ~083-0930 UT resembling a bay on a coastline gave rise to the term "geomagnetic bay" (data courtesy the Canadian Space Agency).

Based on the observations described above, researchers developed a paradigm for substorm physics at the beginning of the 1960's that involved energy transfer from the solar wind to the magnetosphere through the process of magnetic field line reconnection (cf. Dungey [4]). In this process, solar wind magnetic field lines merged at the front side magnetopause with terrestrial magnetic field lines. The merged field lines were swept over the poles to form a long stretched region of magnetic field behind the earth (i.e. the magnetotail) where they ultimately reconnected , sending the plasma on those field lines toward and away

from the earth on either side of the region of reconnection (i.e. the X-line or neutral line). The plasma that moved earthward became part of a large scale circulation pattern (cf. Axford and Hines [5]) in which plasma was transported into the near-Earth region while simultaneously being energized. In this process, the energy provided to the magnetosphere from the solar wind was often out of balance with the energy dissipation in the magnetosphere-ionosphere system. When the rate of energy input exceeded the rate at which energy could be dissipated, the excess energy was stored in the magnetotail magnetic field (i.e. the tail lobes). McPherron [6] ultimately proposed that there interval of time during which energy was stored in the tail lobes should be termed the *growth phase* of the substorm. From time to time, the stored energy would be suddenly deposited into the high latitude ionosphere, and these episodes of energy dissipation were thought to be associated with the substorm expansion phase through the engagement of the tail lobes and the reduction of the magnetic field energy stored in that region (cf. Atkinson [7]). This would be seen, in the ionosphere, as a decrease in the area of the polar cap working under the assumption that the (open) tail lobe field lines mapped to the polar cap. The reader is referred to Baker et al. [8] for a full review of the reconnection paradigm and the observations behind it, and to Kennel [9] for a review of the theoretical concepts behind the paradigm.

The relationship between the auroral and magnetic activity during intervals of substorm activity was the subject of strong controversy over the early years after the term *substorm* was introduced. Prior to the introduction of the term, researchers had studied the phenomenon though its magnetic signatures (called geomagnetic bays) that accompanied the intervals of enhanced auroral activity. Two characteristic types of magnetic disturbance were noted, as characterized by the equivalent current systems that were determined from available ground magnetometers distributed around the world. [An equivalent current system is drawn under the assumption that all magnetic perturbations are due to sheet overhead current flow. The direction of the equivalent current is at 90° to the direction of the horizontal magnetic field perturbation at each measurement site and the separation of the equivalent current lines is inversely proportional to the strength of the magnetic disturbance.] The two systems featured two cells and one cell of equivalent current (cf. Figure 1-18) and have been given different names by different researchers over the years. The two cell system has been called SD (Chapman and Bartels [10]), DP2 (Obayashi [11]) and the directly driven current system (Akasofu [12]). The one cell system has been called DP1 (Obayashi [11]), the unloading component of the substorm (Akasofu [12]) and the substorm current wedge (Baumjohann [13]). Fukushima [1953] had noted that sometimes geomagnetic activity was best represented by a two cell system and sometimes by a one cell system. After the introduction of the term substorm by Akasofu [1], Sugiura and Heppner [14] suggested that this disturbance was best characterized by the two cell system while Akasofu et al. [15] suggested that the substorm should be associated with the one cell system. Ultimately, Rostoker [16] demonstrated that both systems coexisted during substorm activity, setting the stage for the generally held view at present that a substorm involves both directly driven activity an a process for storage and release of energy from the magnetotail, the release being responsible for the substorm expansion phase. Figure 1-19 shows a cartoon of the two components of substorm activity as reflected in the auroral electrojets and the field-aligned currents that connect those electrojets to the source regions of the outer magnetosphere.

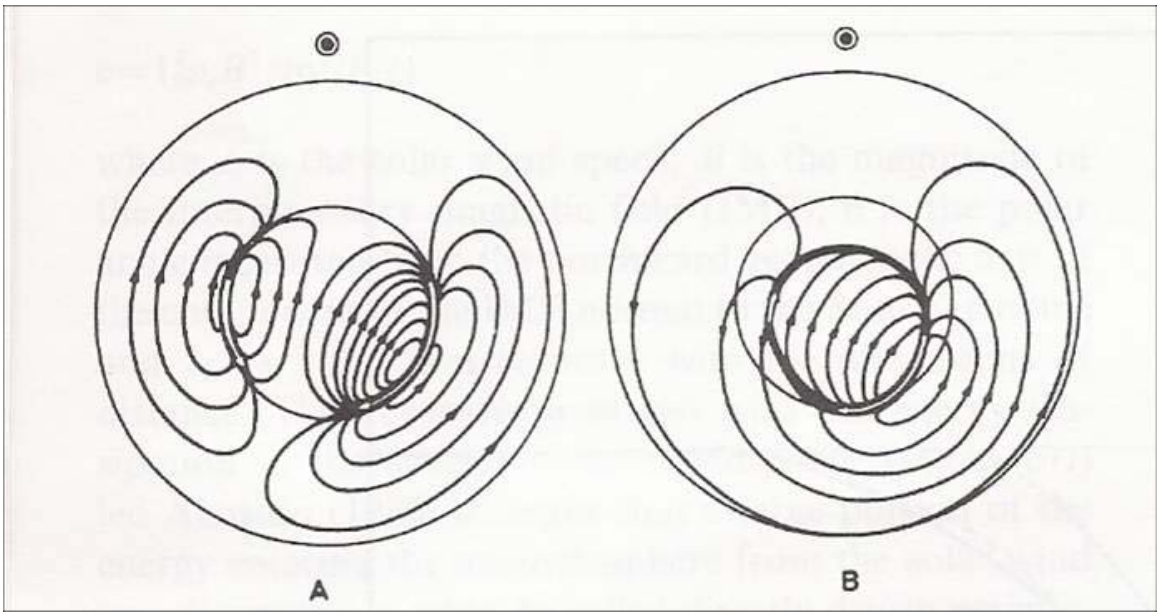

Figure 1-18: The Two Cell and One Cell Equivalent Current Systems which Both Had, at One Time or Another, Been Considered to be Representative of the Substorm Disturbance (after Rostoker [29]).

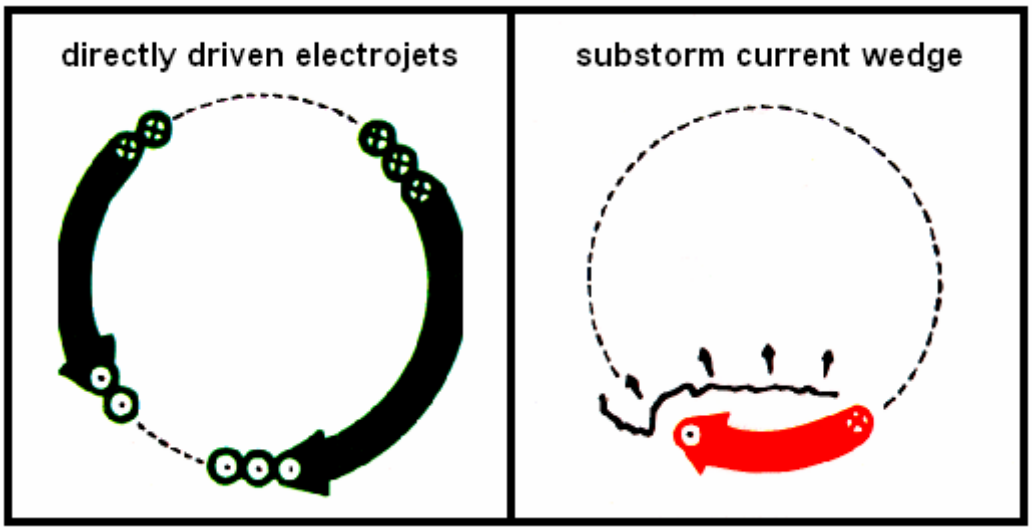

Figure 1-19: Ionospheric and Field-Aligned Currents Associated with the Auroral Electrojets for the Two Components of the Substorm Disturbance (modified after Baumjohann [13]).

1.3.4.2 Present View of the Substorm

It should come as no surprise that the original template for a substorm constructed by Akasofu [1] has undergone considerable evolution. The measurement tools of the time (i.e. allsky cameras) suffered from low film sensitivity and a very sparse distribution of the earth's surface. They were unable to distinguish the presence of diffuse aurora and hence fully define the latitudinal range of energetic particle precipitation into the auroral ionosphere. Vastly improved instrumentation including global imaging from orbiting satellites has afforded researchers a much more detailed picture of the evolution of a substorm, and theoretical studies have provided a framework in which many substorm. observations can be understood in the context of solar wind-magnetosphere coupling and magnetosphere-ionosphere coupling. In particular, knowledge of how

substorm activity develops in response to changes in solar wind energy input has provided the opportunity to better place the Akasofu [1] substorm development in a larger global context.

The understanding of the physical processes which lead to substorm activity requires that the evolution of the disturbance be placed in the framework of changes in the source of energy, viz. the solar wind. As mentioned in the previous section, the development of substorm activity is associated with an increase in the input to the magnetosphere of solar wind energy. This increase is most often affected by the IMF becoming more southward, increasing the rate of reconnection at the front side magnetopause. The increased energy is partially dissipated in the ionosphere through Joule heating and particle precipitation associated with the directly driven currents (c.f. left panel of Figure 1-19) with the balance of the energy being stored in the magnetotail magnetic field. Figure 1-20 shows the development of activity in the auroral oval in response to the increase in energy input from the solar wind. After the southward turning of the IMF, the size of the polar cap grows and the auroral region is forced to lower latitudes. During this growth phase of the substorm, the inner edge of the plasma sheet moves earthward and the plasma sheet, itself, is observed to become thinner. Ultimately, at some later time, the IMF returns to its previous configuration, often by a sudden northward turning. This change in external conditions often triggers the substorm expansion phase, which initially involves the brightening of auroral arcs near the equatorward edge of the midnight sector auroral oval.

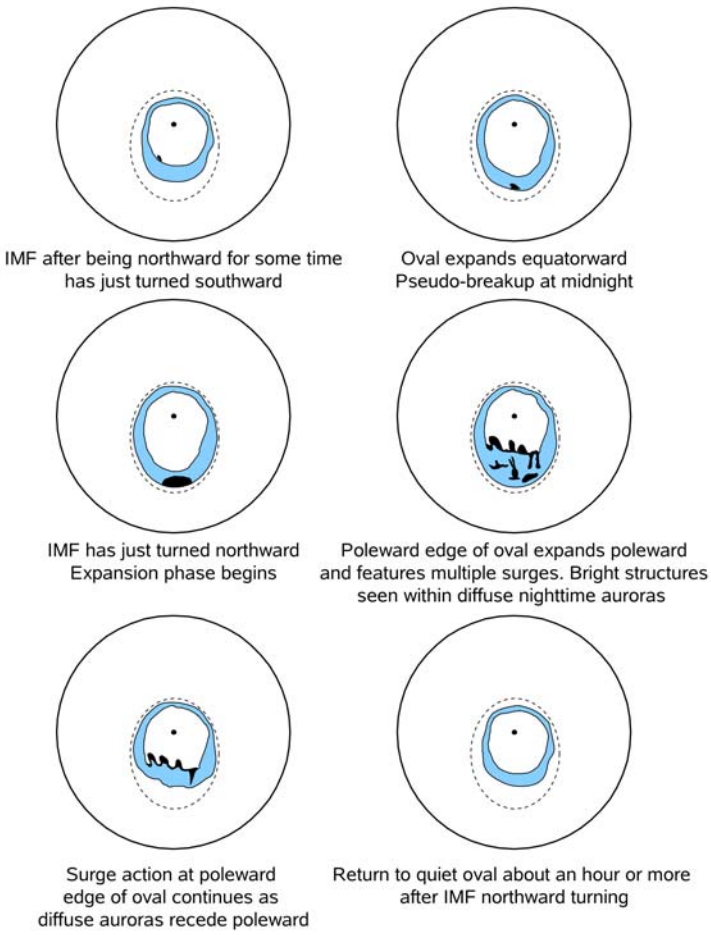

Figure 1-20: A Global Perspective of the Behaviour of the Auroral Oval During the Growth and Expansion Phases of a Substorm. The Akasofu [1] auroral substorm development characterizes the behaviour of the auroras around midnight after a reduction in energy flow into the magnetosphere (often triggered by a northward turning of the IMF), featuring an intensification near the equatorward edge of the oval and subsequent explosive poleward motion of the high latitude edge of the disturbed region.

The explosive poleward motion of the breakup region is well described by the original Akasofu framework shown in Figure 1-16. In a matter of less than 5 minutes, the disturbed region has spread to the poleward edge of the oval, and from then on the polar cap collapses back to its pre-substorm configuration. During the minutes to tens of minutes over which time the polar cap collapse takes place, bright auroral loops and surges appear along the poleward edge of the evening sector oval. These disturbances have been termed *poleward boundary intensifications* (PBIs) and their auroral luminosities and associated magnetic perturbations are commensurate with those detected during the explosive development of the substorm disturbed region at onset.

One of the most important realizations coming from satellite borne imagers is that there are two distinct regions of auroral activations in the night side auroral oval. One region is located near the equatorward edge of the midnight sector auroral oval, and is the site of substorm expansive phase onset and pseudo-breakups (which have the appearance of an expansive phase onset but engage only the closed field line portion of the magnetotail and hence never develop to the extent of causing a reduction in the area of the polar cap). Figure 1-21 shows the two zones of the double oval, from which it is seen that the regions of auroral luminosity in the midnight/morning and evening sectors are the sites of unbalanced upward field-aligned current (FAC). These two zones can both be activated and may feature auroral loops and surges whose magnetic and auroral signatures are quite similar. Figure 1-22 shows a moment during an interval of strong magnetospheric activity. In Panel a of that Figure, the two branches of the evening sector are clearly evident and an auroral surge form is seen on the most poleward of the two branches. This surge form is a PBI, and may occur either with or without the presence of substorm expansive phase activity. In Panel b of the Figure, a substorm expansion phase is seen to be initiated in the late evening sector of the equatorward branch of the double oval.

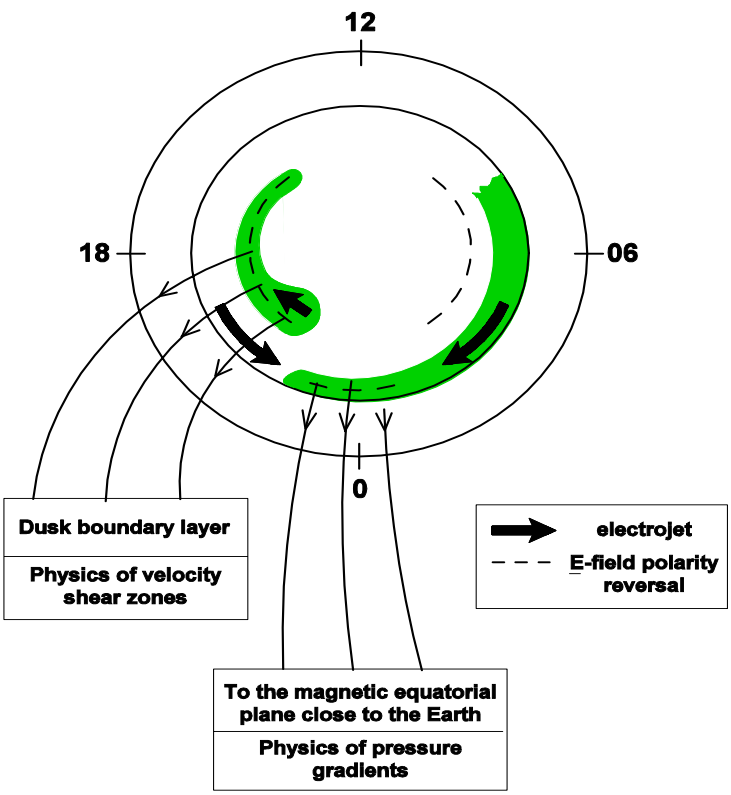

Figure 1-21: The Two Zones of Auroral Luminosity (shown in green) Associated with the Double Oval. These regions of discrete auroral arcs mark the presence of upward field-aligned current that connects the ionosphere to the source regions in the magnetosphere where energy is provided. The eastward and westward electrojets are bounded by anti-parallel Birkeland current sheets (cf. Zmuda and Armstrong [41]) not shown in this figure.

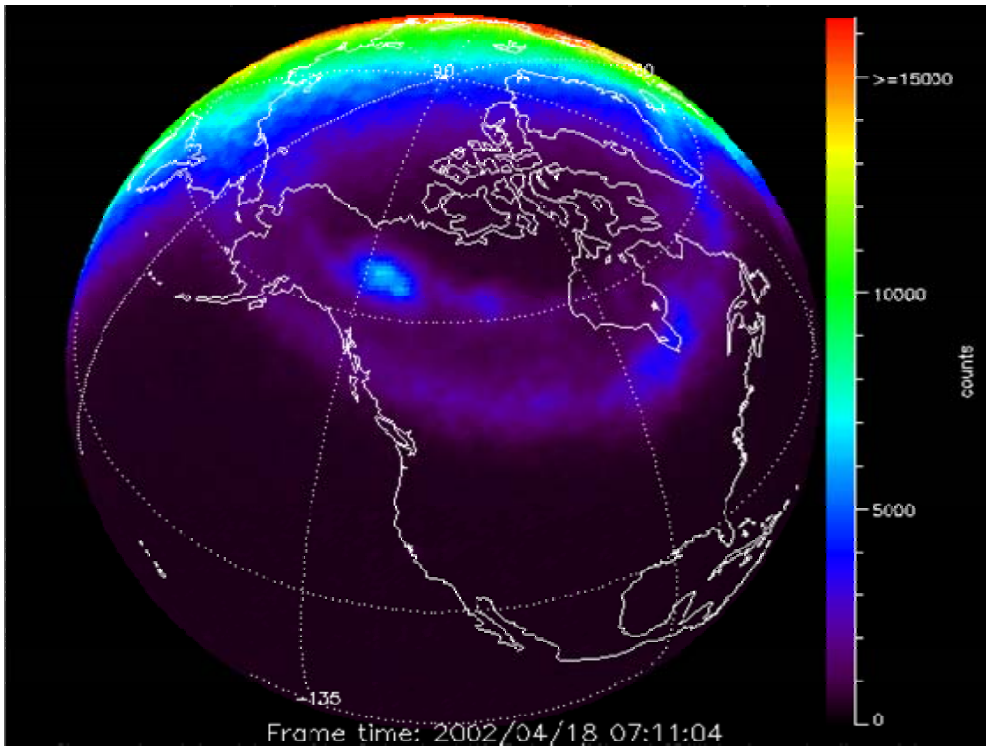

Figure 1-22a: A Poleward Boundary Intensification (PBI) on the Poleward Branch of the Double Oval as Observed by the UV Camera Aboard the IMAGE Satellite (courtesy S. Mende [University of California, Berkeley] and CDAWeb).

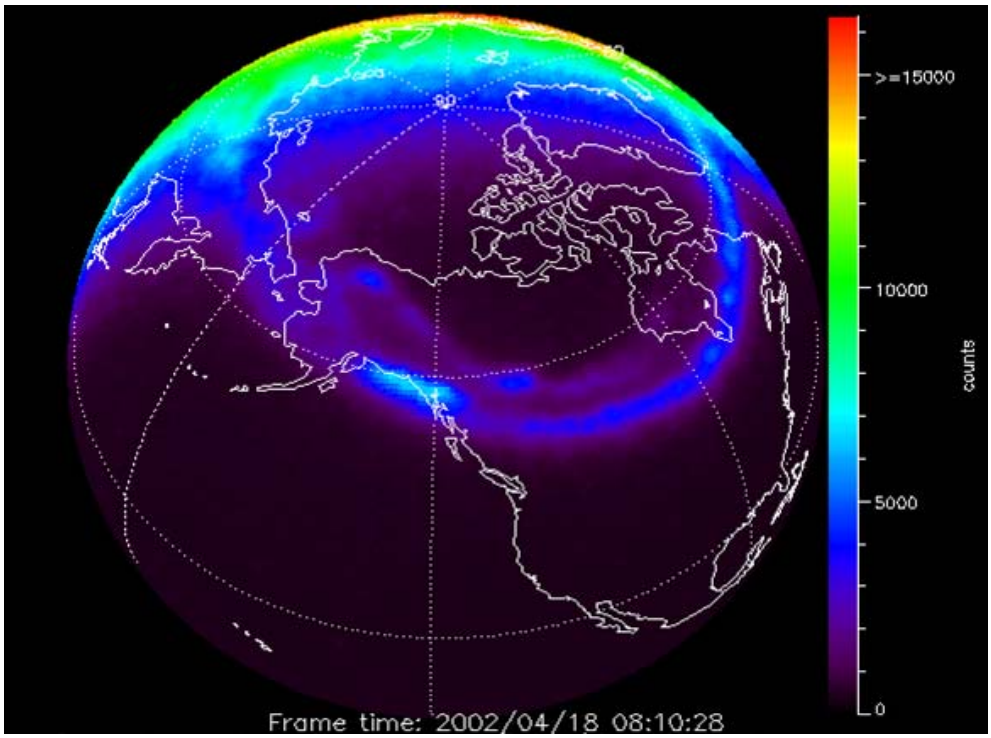

Figure 1-22b: A Substorm Expansion Phase Onset on the Equatorward Branch of the Late Evening Sector Auroral Oval as Observed by the UV Camera Aboard the IMAGE Satellite (courtesy S. Mende [University of California, Berkeley] and CDAWeb).

Finally, it should be noted that one of the most spectacular auroral features, the so-called westward traveling surge (WTS), may not travel much at all in the sense of the moment of the auroral luminosity shifting westward over significant distances. The concept of the WTS originated from observations by allsky cameras of the local field of view. The surge appeared as a bright luminous form at the eastern edge of the field of view and its western edge would shift westward across the field of view at speeds of ~ 1 km/s on the average. As early as the 1970's, Kisabeth and Rostoker [17] noted that, while the western edge of the surge form expanded westward, the eastern edge might expand eastward. This behavior is characteristic of a spatial confined region of luminosity that grows by expanding in all directions from its initial site of formation. Subsequently, Tighe and Rostoker [18] showed that the initial westward expansion of the western edge of a surge form could occur at speeds of several km/s in the early stages of development, but that the expansion would cease after a few minutes leaving behind a quasi-stationary bright auroral form. There could, in fact, be several such forms arrayed along the high latitude boundary of the evening sector auroral oval during the course of a substorm event. The results of Craven et al. [19] support this revised view of the WTS, and it would appear that the concept of a monolithic surge form moving over large longitudinal distances during the course of a substorm is no longer tenable. It does appear that the surge forms are actually poleward boundary intensifications (PBIs) in the light of their location at the poleward edge of the disturbed auroral oval after the poleward expansion of the auroral bulge during the substorm expansion phase.

1.3.4.3 Storm-Substorm Relationship

Much of the early work on the solar-terrestrial interaction involved the study of the response of the ionosphere to geomagnetic storms. These storms featured a significant decrease in the horizontal component of the low latitude magnetic field that grew in a matter of hours and decayed over a period of tens of hours to a few days. These disturbances, ranging magnitude from a few tens of nT to hundreds of nT were understood as the consequence of the formation of a symmetric ring current flowing around the earth at distances of ~ 3-8 Re. The strength of this ring current was quantified through the Dst index (cf Sugiura [20]), which was computed from the horizontal (H) component of the magnetic field from ~ 4 – 5 low latitude stations distributed with approximately equal longitudinal separation around the world.

While it was clear from the beginning (cf. Rostoker and Fälthammar [21]) that southward interplanetary field was required for the development of a geomagnetic storm, the role of substorms in that development has been somewhat controversial. Akasofu et al. [22] argued that substorms played an integral role in the formation of the storm time ring current by injecting energetic particles from the plasma sheet to regions much closer to the earth where they ultimately became the ring current carriers. However, shortly thereafter, Burton et al. [23] provided convincing evidence that simply the presence of a sufficiently strong dawn-to-dusk interplanetary electric field for a sufficiently long period of time was sufficient to produce the observed ring current growth. In the scheme of Burton et al., there was no role for substorms in the ring current development, although it was clear that substorms and the ring current derived their energy from the same source, i.e. the solar wind.

Since that time, the role of substorms in the creation of the storm time ring current has been a subject of intense research activity. This question has been highlighted in reviews of the putative storm-substorm relationship that have appeared in recent years (cf. Gonzalez et al. [24]; Kamide et al.[25]; Sharma et al.[26]) but no definitive assessment had emerged. However, Tsurutani et al. [27] took advantage of the topside imaging capability of the Polar spacecraft to provide evidence that, during the main phase of a magnetic storm, there could be no substorm activity although many spatially confined auroral structures (e.g. pseudo-breakups, torches, poleward boundary intensifications) were in evidence. Zhou et al. [28] explored this question further and reached the same conclusion, namely substorms were not an essential element of storm time ring current growth.

On the other hand, Rostoker [29] has argued that the process whereby substorm expansion phase is initiated near the inner edge of the plasma sheet plays an integral role in the development of the storm time

ring current. He contends that pseudo-breakups and substorm expansion phase onsets both reflect the results of an instability near the inner edge of the crosstail current sheet which initiates the breakdown of the shielding electric field (cf. Senior and Blanc [30]). This breakdown permits the penetration of the convection electric field closer to the earth, leading to additional acceleration of plasma sheet particles and increasing the possibility that they can be injected into stable trapped orbits as part of the ring current population. For a classical substorm, triggered by a reduction of energy input into the magnetosphere from the solar wind (cf. Rostoker [31]), the effects of the earthward penetration of the convection electric field is mitigated by the fact that that electric field begins to decline in strength. For storm conditions, when the driving electric field maintains its strength for lengthy periods of time, there can be a series of breakdowns of shielding, each episode occurring closer and closer to the earth. In this way, the inner edge of the plasma sheet can get relatively close to the earth and the probability increases that injected energetic particles can become part of the ring current population.

1.3.4.4 Conductivity of the Auroral Ionosphere

One of the primary mechanisms for dissipation of energy entering the magnetosphere from the solar wind is Joule heating associated with the flow of electric current through the resistive ionosphere. The entire process starts with the penetration of solar wind plasma into the outer edges of the magnetosphere. The kinetic energy of flow of this plasma is converted to electromagnetic energy, with a major portion of this being stored in the magnetic field of the tail lobes and in the magnetic field associated with the large scale current systems that couple the source region of the outer magnetosphere to the auroral ionosphere.

The Joule heating takes place in a relatively limited region of altitude extending from ~ 100 – 300 km above the earth's surface. Here, there is a significant electric conductivity as a consequence of collisions between the neutral particles of the atmosphere and the ions and electrons that are there as a consequence of collisions with energetic particles precipitating from the magnetosphere and with the interaction of solar visible and UV radiation with the atmospheric neutrals. The precise number densities of ions and electrons represent a balance between the production rates due to particle precipitation and solar radiation and recombination rates due to collisions among the atmospheric constituent particles.

The ionospheric conductivity at a point is space is anisotropic, and is typically expressed in terms of current flow parallel to the component of the electric field normal to the ambient magnetic field and perpendicular to that component of the electric field. These are termed the Pedersen and Hall conductivities respectively, and are defined by the expressions:

$$\sigma_P = \frac{n_e}{B}\left[\frac{\nu_e\omega_e}{\omega_e^{\,2}+\nu_e^{\,2}} + \frac{\nu_i\omega_i}{\omega_i^{\,2}+\nu_i^{\,2}}\right]$$

$$\sigma_H = \frac{n_e}{B}\left[\frac{\omega_e^{\,2}}{\omega_e^{\,2}+\nu_e^{\,2}} - \frac{\omega_i^{\,2}}{\omega_i^{\,2}+\nu_i^{\,2}}\right]$$

where n_e is the number density of electrons, B is the magnetic field, ν_e and ν_i are the electron-neutral and ion-neutral collision frequencies, and ω_e and ω_i are the electron and ion gyrofrequencies. There is, of course, a conductivity parallel to the magnetic field, but this is generally assumed to be extremely high consistent with the assumption in magnetohydrodynamics of frozen field conditions over most of the length of the magnetic field lines linking the ionosphere to the outer magnetosphere.

Often, one sees, rather than the conductivity at a point, the height integrated conductivities:

$$\Sigma_P = \int \sigma_P dz$$

$$\Sigma_H = \int \sigma_H dz$$

where the integration typically extends from the bottom of the ionosphere (z~ 90 km) to the top of the ionosphere (z ~ 300 km). All the conductivities measured observationally or used in theoretical expressions describing magnetosphere-ionosphere coupling are height-integrated and it is generally assumed that the current in the ionosphere flows in an infinitesimally thin sheet for which the current flow is described by:

$$\mathbf{J} = \Sigma_P \mathbf{E} + \Sigma_H \mathbf{n}_r \times \mathbf{E}$$

where $\mathbf{n}_r$ is a unit radial vector, **E** is the component of the electric field perpendicular to **B** and frozen-field conditions apply, viz. $\mathbf{V} = (\mathbf{E} \times \mathbf{B})/B^2$. From the assumption that div **J** = 0, one can obtain an expression for the field-aligned current density:

$$j_{\parallel} = \Sigma_P \text{ div } \mathbf{E} + \mathbf{E} \cdot \text{ grad } \Sigma_P + \mathbf{n}_r \times \mathbf{E} \cdot \text{ grad } \Sigma_H$$

where div **V** = 0. From this expression, it can be seen that knowledge of gradients in the Hall and Pedersen conductivities are crucial in determining where field-aligned currents flow. Under conditions of steady magnetospheric convection, the location of the discrete aurora in Figure 1-21 follows automatically from consideration of the Pedersen conductivity terms in Eq (4) when one understands that the discrete aurora mark the location of upward field-aligned current.

A number of researchers have produced semi-empirical models of ionospheric conductivity that are available for use in simulations of electric current flow in the magnetosphere-ionosphere system. One of the earliest of these models was produced by Wallis and Budzinski [32], who used measurements of precipitating energetic electron fluxes from the Isis 2 satellite together with the theoretical estimates of ionization rate per unit volume formulated by Rees [33, 34] to obtain estimates of the Hall and Pedersen conductivities every two hours of magnetic local time at one degree intervals of geomagnetic latitude from 60-84°. They included in their calculations the conductivity effected by solar radiation and considered conditions of relatively low activity (0 < Kp < 3) and of high activity (3 < Kp < 9). Since the work of Wallis and Budzinski, there have been several other efforts to estimate ionospheric conductivity, most of them calling on measurements of precipitating electron flux (e.g. Spiro et al.[35]; Robinson et al. [36]; Hardy et al. [37]).

For example, Robinson et al.[36] developed semi-empirical expressions relating the height-integrated Hall and Pedersen conductivities to the average energy and energy flux of the precipitating electrons responsible for the conductivity, i.e.

$$\Sigma_P = \frac{40\overline{E}}{16 + \overline{E}^2} \Phi_E^{1/2}$$

$$\frac{\Sigma_H}{\Sigma_P} = 0.45 \overline{E}^{0.85}$$

A different approach was developed by Ahn et al. [38] who developed semi-empirical relationships between the height-integrated Hall and Pedersen conductivities and the magnetic perturbation |ΔH| in the north-south component of the magnetic field, viz.

$$\Sigma_H, \Sigma_P = a \mid \Delta H \mid^b$$

that were parameterized to apply to the poleward and equatorward portions of the auroral oval for the evening (eastward electrojet) and morning (westward electrojet) sectors.

The question now arises as to how useful these statistical procedures are in addressing how substorms influence ionospheric conductivity and hence the flow of electric current between the generator regions in the outer magnetosphere and the sink of energy in the upper atmosphere. A major source of difficulty is that all relationships were developed based on the statistical properties of the individual parameters in terms of their variation with activity level as quantified by an index (e.g. Kp) or by individual site measurements (e.g. perturbations in the north-south component of the magnetic field |ΔH| measured on the ground). Because the currents in the auroral oval can produce similar magnetic perturbations for different latitudinal positioning of the auroral oval, there is certain to be a smearing for any parameter inferred from such data. The same is true for the flux and average energy of precipitating electrons for a given activity level. Figure 1-23 shows Hall conductivities inferred from precipitating energetic electron fluxes by Hardy et al. [39] for varying levels of magnetic activity as quantified by the Kp index. One would expect to see in the evening sector two zones of enhanced conductivity, consistent with the distribution of upward field-aligned current shown in Figure 1-21. For the Kp = 2 pattern, this is in fact observed, but for the Kp = 3 pattern smearing of the data for both branches of the double oval makes the distribution appear more continuous. In reality, as one moves to higher activity levels, the probability of the data being taken during an substorm expansion phase is enhanced, and for such circumstances one would expect the midnight and late evening auroral oval to feature many spatially localized regions of high luminosity, varying rapidly in space and time. Any statistical study of energetic electron precipitation binned according to latitude is very likely to produce a smeared out latitudinally confined region of enhanced fluxes, even though individual events may actually feature discrete structures such as the double oval or longitudinally confined north-south streamers (cf. Rostoker et al. [40]).

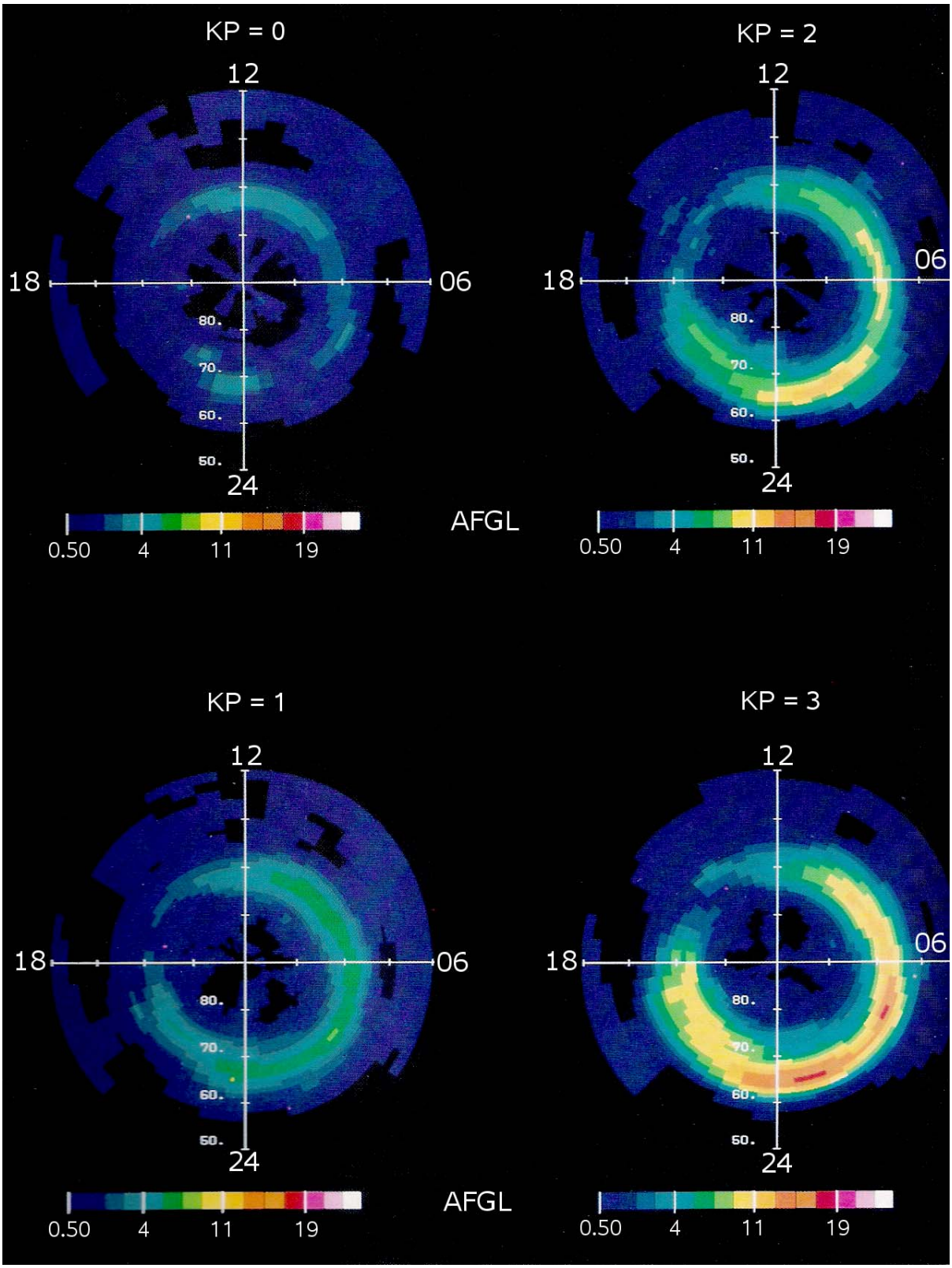

Figure 1-23: The High Latitude Hall Conductivity Distribution Inferred from Precipitating Energetic Electron Fluxes at Different Magnetic Activity Levels as Quantified by the Kp Index (Top left Kp=0, Bottom left Kp=1, Top right Kp=2 , Bottom right Kp=3). The colors are assigned on a logarithmic scale (after Hardy et al. [37]).

1.3.4.5 Conductivity in the Ionosphere During Substorm Activity

In section 1.2 above, the general conductivity distribution during periods of magnetospheric activity has been presented. Technically, one should try to decouple the conductivity distribution during times of purely directly driven activity (i.e. steady magnetospheric convection) from that in evidence during substorm expansion phase activity. However, this has never been done because of the difficulty in identifying periods of purely directly driven activity. There are, however, two aspects of substorm activity in which specific conductivity situations have been identified and discussed in the context of coincident electric field changes.

The first of these relates to the onset of the substorm expansion phase, at which time breakup arcs appear near the equatorward edge of the midnight sector auroral oval. At this time, it is suggested that the westward convection electric field penetrates closer to earth as a consequence of a reduction in the shielding electric field (cf. Rostoker [29]). A consequence of this is the development of a southward polarization electric field across the region of auroral breakup, so that the total horizontal electric field in the midnight sector auroral ionosphere points in the south-west direction. When the orientation of this electric field is such that the southward Pedersen current exactly cancels the northward Hall current, the intensity of the westward electrojet is a maximum. Under these circumstances, the westward electrojet is called a Cowling current and its current density is written

$$J_C = \Sigma_C E_w$$

where $\Sigma_C = (\Sigma_H^2 / \Sigma_P) + \Sigma_P$ is the Cowling conductivity.

The second issue, somewhat related to the first, is the electric current flow in the vicinity of an auroral surge form. As mentioned earlier, auroral surges often turn out to be PBIs located at the poleward boundary of the evening sector auroral oval. An auroral surge form represents a concentrated region of precipitating energetic electrons, and hence a region of upward field-aligned current. The scale size of surges is ~ 500 km when fully developed, and the western edge of a surge can expand westward initially at speeds of ~ 5 km/s although, after a few minutes, the western edge comes to a halt (cf. Tighe and Rostoker [18]). A surge can be viewed as a core region of upward field-aligned current, encircled by an annular region of downward field-aligned current. The downward current connects to the upward current in the ionosphere through radially flowing Pedersen current, and in the magnetosphere with a generator current in the source region where the energy is available to accelerate particles into the ionosphere and to permit the degrading of stored energy in the magnetosphere to heat in the ionosphere though the process of Joule heating. There is also a Hall current that circulates around the luminous region of the surge, flowing equatorward at the western head of the surge and poleward at its eastern trailing edge. The magnetic effects of this vortical Hall current are what produce the characteristic signatures observed by magnetometers on the ground in the vicinity of the surge form. Figure 1-24 shows ground magnetometer data from two sites in northern Alaska which are fortuitously located at the western edge (Barrow) and the eastern edge (Katovik) of a large auroral surge form detected at ~0600 UT (aurora not shown here but observed by the UV imager aboard the Polar satellite). The positive D-component at Barrow is caused by equatorward ward flowing Hall current at the western edge of the surge, while the negative D-component at Katovik is caused by poleward flowing Hall current at its eastern edge. The positive H-component is consistent with the two stations being at the equatorward edge of the bright auroras associated with the surge.

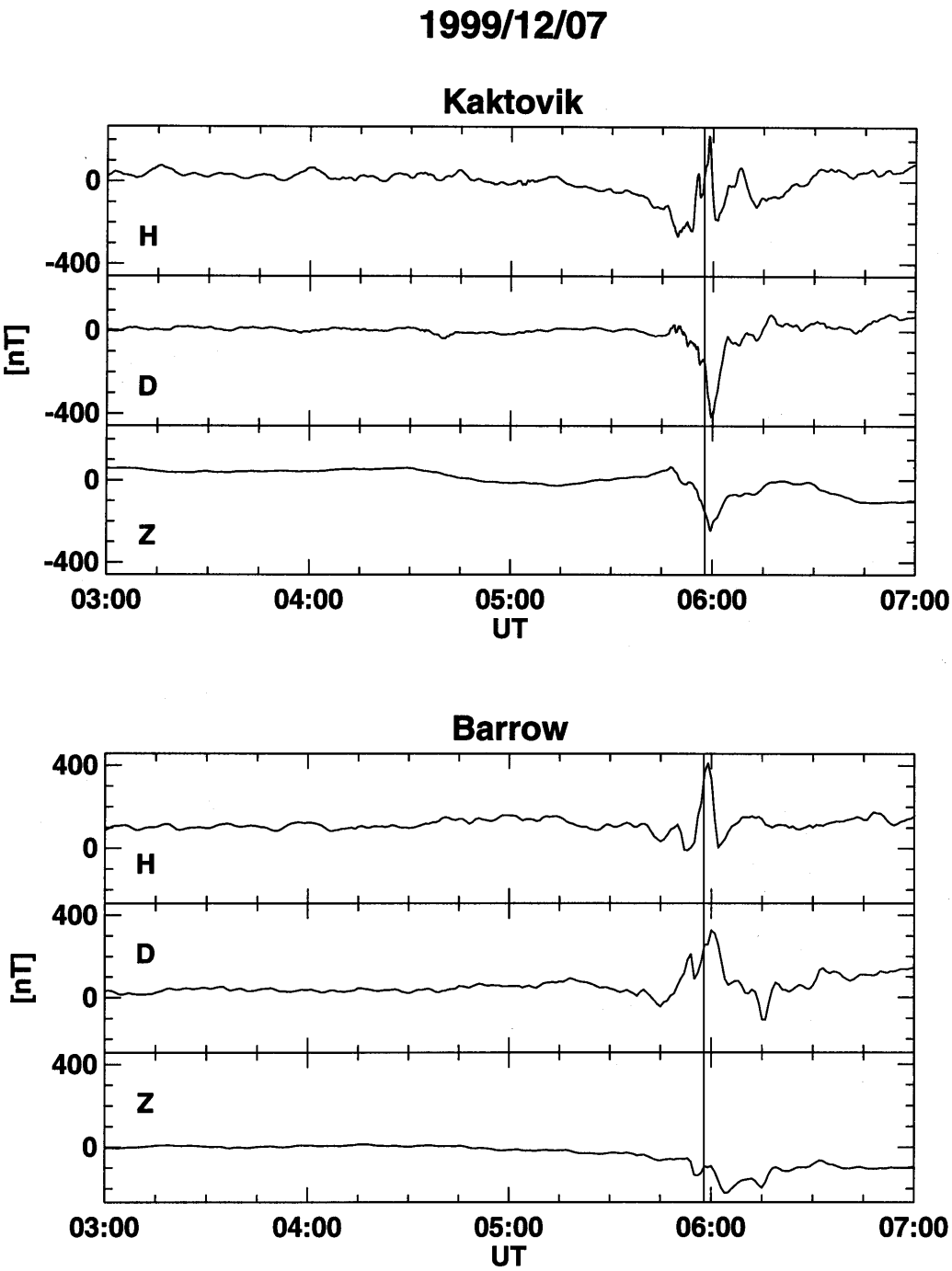

Figure 1-24: Ground Magnetometer Data from Two Stations in Northern Alaska (Barrow and Katovik), Located at Approximately the Same Magnetic Latitude but Separated in Longitude. Fortuitously, Katovik is situated on the eastern edge of an auroral surge while Barrow is situated on the western edge of the same surge.

Finally, one should note that the aurora during substorm expansive phase activity involve discrete arcs, often separated by regions of darkness. Therefore, the conductivity distribution can be expected to be very complex and images obtained from high altitude satellites some thousands of km above the aurora cannot be expected to provide information on gradients with a resolution of any better than ~100 km.

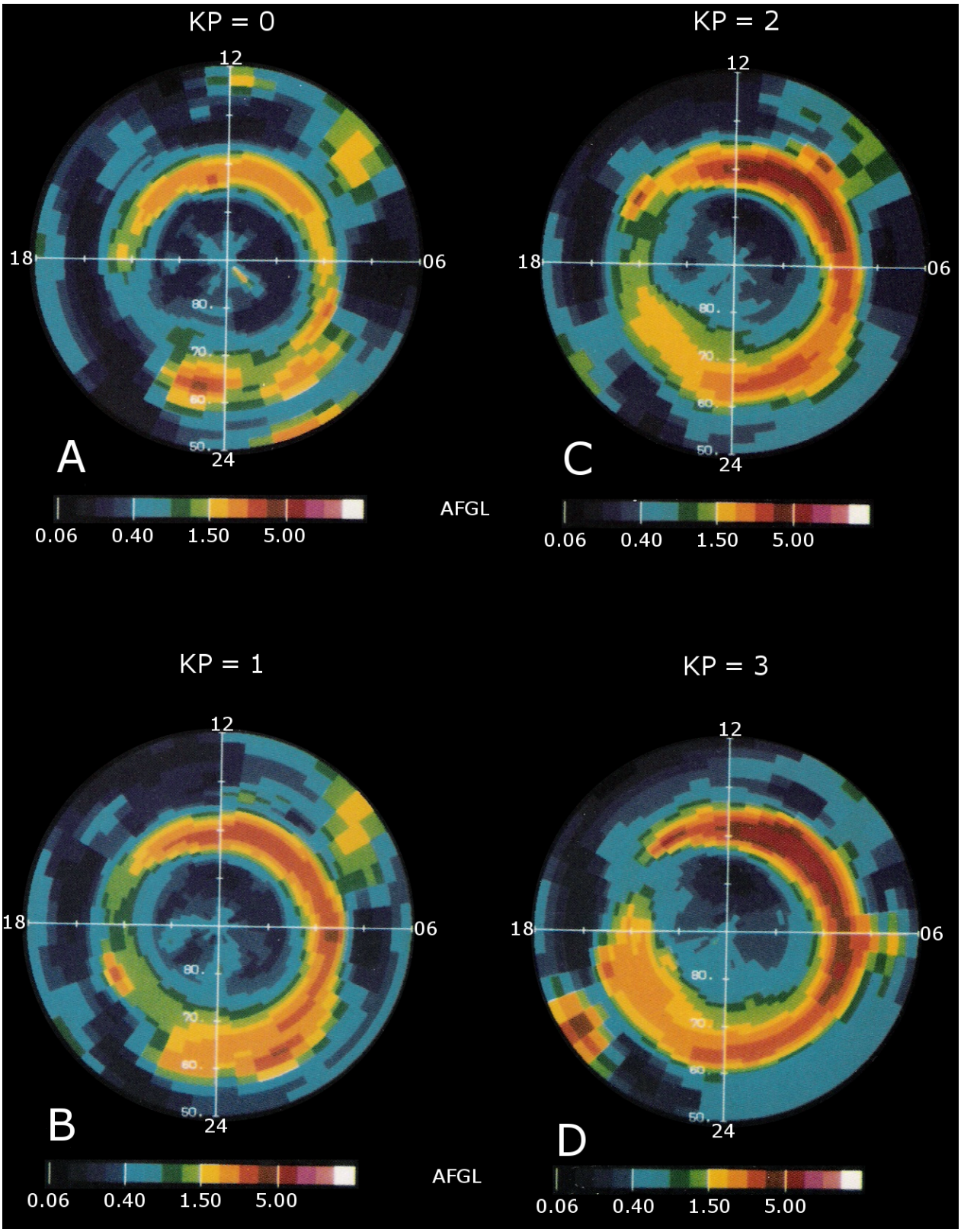

Figure 1-25: Polar Plots of the Average Energy of Precipitating Electrons, in Units of keV, for Different Activity Levels as Quantified by Kp (Panel A Kp=0, Panel B Kp=1, Panel C Kp=2, an Panel D Kp=3). The colors are assigned on a logarithmic scale (after Hardy et al. [39]).

1.3.4.6 Precipitation of Energetic Particles into the Ionosphere

For practical purposes, the primary way in which substorm activity influences the high latitude ionosphere is to alter the electrical conductivity in the auroral oval in a way that changes the ionospheric current flow and the distribution of field-aligned upward and downward currents that link the source regions of the outer magnetosphere to the dissipative upper atmosphere. To quantify that behaviour, efforts have been made in the past to establish the local time distribution of energetic electron precipitation for varying levels of magnetospheric activity. The most detailed data base analysed thus far was acquired using the DMSP satellites and published by Hardy et al. [39] who binned the data by activity level as defined by the Kp index. Examples of the global distribution of the average energy of precipitating electrons are shown in Figure 1-25 for four different levels of Kp ranging from quiet to disturbed conditions. For quieter conditions (panels A and B), it is apparent that the energetic electron precipitation maximizes at lower latitude in the midnight and post midnight sectors and at higher latitudes in the pre-midnight sector. This is consistent with the concept of the double oval mentioned earlier in this review, and with the observations of upward field-aligned current made by polar orbiting satellites carrying magnetometers (cf. Zmuda and Armstrong [41]) since strong upward field-aligned currents are carried by energetic electrons precipitating into the ionosphere.

It should be pointed out that the simple patterns shown in Figure 1-25 are what one would expect in the absence of substorm expansion phase activity. When expansion phase activity is in progress, it is often difficult to separate the poleward and equatorward branches in the double oval and statistical representations of precipitating energetic electron fluxes seem to indicate a smooth connection across midnight of the low latitude region of energetic electron precipitation in the morning sector with the high latitude region of precipitation in the evening sector. Great care has to be exercised in the study of particles and fields in the hours around midnight when substorm expansion phase activity is in progress.

1.3.4.7 Modeling of Substorm Current Systems

The product of several decades of experimental and theoretical studies of the solar-terrestrial interaction has been the development of models that synthesize the data from observation points in space and on the ground and produce global pictures of the disturbances associated with the transport and dissipation of energy derived from the solar wind. Since there is no long term stable configuration of satellites available for modeling studies, it has fallen on ground based data to provide of the core of information required to produce models of real electric currents and electric potentials which are required to quantify the character of the coupling between the magnetosphere and ionosphere which determines the dispensation of the solar wind energy that penetrates into the magnetosphere.

Until the late 1970's, the only way in which global patterns of energy dissipation were described in a quasi-quantitative fashion was through the use of equivalent current systems of the type shown in Figure 1-18. While these were useful, they did not take into account the fact that a significant portion of the horizontal component of the magnetic field disturbance could be due to field-aligned rather than ionospheric current flow. Techniques developed independently by Mishin et al. [42] and Kamide et al. [43] provided the ability to obtain, simply from ground based measurements, estimates of the real current flow (both ionosphere and field-aligned) and the cross-polar cap potential drop. Following the approach of Kamide et al. [43], the flow lines of the equivalent current system were characterized by a scalar potential function Ψ so that the equivalent current density (in A/m) is given by

$$\mathbf{J}_T = \mathbf{n}_r \times \text{grad}\, \Psi$$

The total height-integrated current is represented as the sum of the equivalent current and a potential current $\mathbf{J}_p$, i.e.

$$\mathbf{J} = \mathbf{J}_p + \mathbf{J}_T$$

where div $\mathbf{J}_T = 0$ and curl $\mathbf{J}_p = 0$. The field-aligned current density is then given by

$$j_{\parallel} = \text{div}\, \mathbf{J} = \text{div}\, \mathbf{J}_p$$

and $j_{\parallel}$ is positive downwards.

The real ionospheric current, expressed in the frame of reference co-rotating with the earth, is written as

$$\mathbf{J} = \Sigma_P \mathbf{E} + \Sigma_H \, \mathbf{E} \times \mathbf{n}_r$$

and the electric field is assumed to be electrostatic, i.e. $\mathbf{E}$ = -grad Ø.

Using the above equations, one can write a partial differential equation for Ø in terms of Ψ

$$A\frac{\partial^2 Ø}{\partial \theta^2} + B\frac{\partial Ø}{\partial \theta} + C\frac{\partial^2 Ø}{\partial \lambda^2} + D\frac{\partial Ø}{\partial \lambda} = F$$

where θ is the colatitude and λ is the east longitude, while (A,B,C,D) are complicated functions of $(\theta,\lambda,\Sigma_P,\Sigma_H,\Psi)$. This partial differential equation can be solved for Ø using reasonable boundary conditions at the pole and at the equator, at which point it is possible to calculate $\mathbf{E}$, $\mathbf{J}$, and $j_{\parallel}$ at all points in the ionosphere assuming some ionospheric conductivity distribution. Figure 1-26 shows calculated electric potential and current distributions in the high latitude ionosphere for averaged activity levels for disturbed days in May and June 1965.

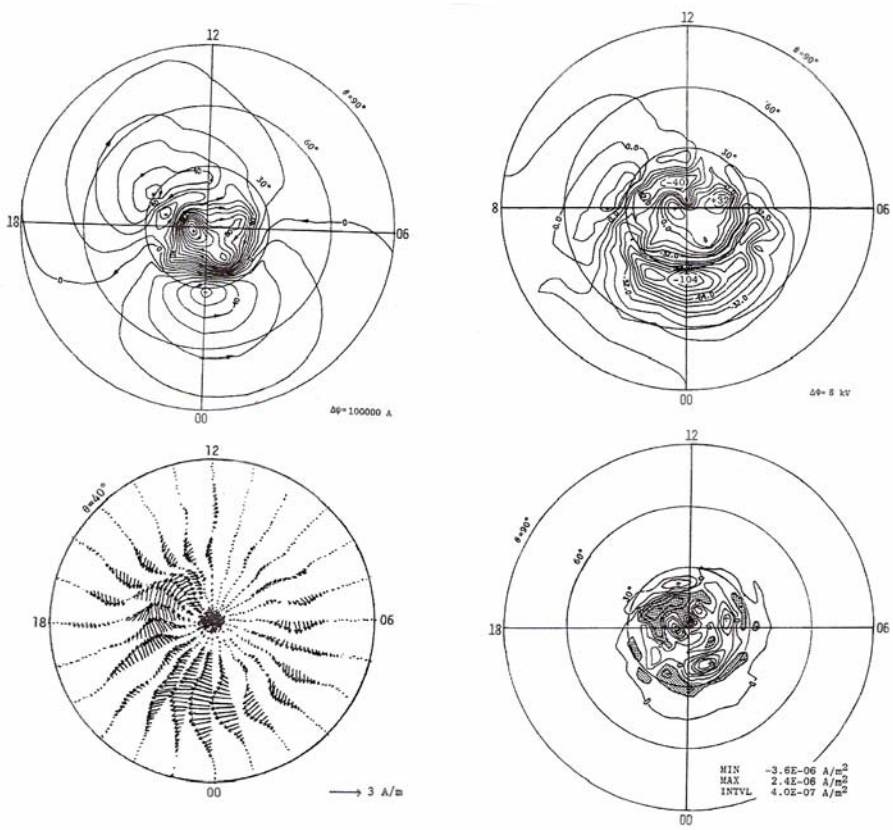

Figure 1-26: Input and Output from the KRM Routine.

Top left panel is the equivalent current system (10^5 A contour interval) obtained from ground magnetometer data for the instant 1252 UT during a substorm on January 12, 1973. The top right panel shows the pattern of electric potential (8 kV between potential contours), the bottom left panel shows the ionospheric current vectors (the length of the current vector in A/m as shown in the panel), and the bottom right panel shows the field-aligned current distribution (4 x 10^{-7} A/m^2 contour interval) with downward current indicated by stippled shading (after Kamide et al. [43]).

The technique described above is called the KRM method, and demands the assumption of a conductivity distribution. Since this is rarely known at any instant, it makes it difficult to apply this technique to given moments in time. Some years after the KRM technique was implemented, Richmond and Kamide [44] enhanced its capability by allowing additional information to be incorporated into the basic methodology (e.g. electric field data, conductivity data, etc.) creating the method of Assimilative Mapping of Ionospheric Electrodynamics (AMIE). The AMIE technique has been able to incorporate UV luminosity data from high altitude orbiting satellites (e.g. Polar, IMAGE) to infer more detailed information about the ionospheric conductivity distribution using the techniques described in Lummerzheim et al. [45].

1.3.4.8 Separation of Substorm Expansion Phase Effects from Directly Driven Activity

Despite the major advances in providing a quantitative description of the behaviour of key electrodynamic parameters in the magnetosphere-ionosphere system during intervals of activity, the problem still persists in being able to decouple the effects of substorm expansion phase activity from the background directly driven activity. This kind of problem is evident in examining the use of (for example) the AL index in tracking substorm activity. The AL index is a time series of points made from a superposition of the H-components of approximately 12 stations distributed as evenly as possible in longitude across all local times at average auroral zone latitudes. Each point represents the maximum of all the negative perturbations detected at the 12 stations, regardless of what local time the station that produced that maximum value is located. Negative perturbations arising from directly driven activity tend to peak near dawn, while those due to substorm expansion phase development tend to peak near midnight. Circumstances often arise in which the maximum perturbation due to directly driven activity is larger than the maximum perturbation due to directly driven activity. Then the expansion phase onset is not visible in the AL time series, which can lead to significant problems in associating disturbances measured by satellites with those observed on the ground.

An obvious solution to this problem is to examine individual magnetograms when studying any given event; however, this poses a formidable logistical problem because of the number of ground magnetograms required to optimize the likelihood of identifying the actual onset time of a typical substorm expansion phase. One way to circumvent that problem is to develop a technique that separates the effects of directly driven activity from expansion phase disturbances. This problem has been addressed by Sun et al. [46] who have developed a technique called the method of natural orthogonal components (MNOC).

The technique involves the representation of the horizontal component of the magnetic field perturbations at all available magnetometer stations over the earth's surface by an equivalent current system which is described by a function E (t_i,r_j) which is a function of time t_i (I = 1,2,……m) and position r_j (j = 1,2,…n). The equivalent current is computed for latitudes from the pole to 50°N at 2° intervals and for 15° intervals of longitude, so that spatially j = 1,……480. The equivalent current function is expressed as an expansion in terms of functions X_j^k (k = 1,2,……h≤ n), viz.

$$E_{ij} = \sum_{k=1}^{h} T_i^* X^j_k$$

where T_i* is a function of time and the X_j^k values are orthogonal over the observation points, viz.

$$\sum_{k=1}^{n} X^l_j \, X^p_j = 0 \quad l \neq p$$

Using a minimization procedure, and defining

$$A_{lp} = (1/m) \sum_{i=1}^{h} E_{il} \, E_{ip}$$

$$\lambda_k = (1/m) \sum_{i=1}^{m} (T_i^*)^2 \sum_{p=1}^{n} (X^k_p)^2$$

$$Xk = (X^k_1, X^k_2, \ldots X^k_n)^T$$

where (1/m) is a normalization factor, one ends up solving the equation

$$AX^k = \lambda_k X^k$$

where λ_k is a function of the order number k (=1,2,….20). It turns out λ_1 and λ_2 dominate over all the higher order values of λ_k , with λ_1 reflecting the two cell equivalent current of directly driven activity and λ_2 reflecting the one cell equivalent current of the substorm current wedge. Figure 1-27 shows a typical equivalent current system at an instant of time during substorm expansion phase activity, together with the two equivalent current patterns associated with directly driven and substorm current wedge activity as separated by the MNOC technique.

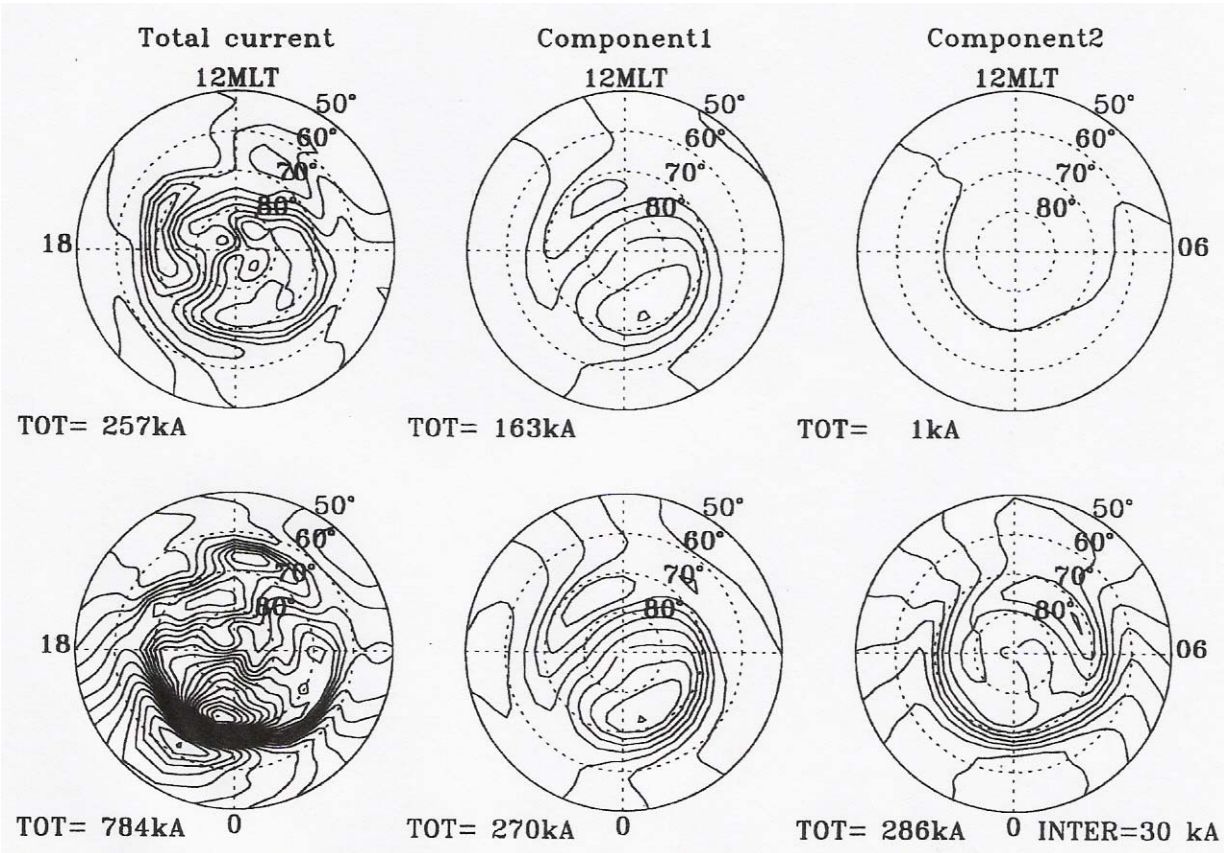

Figure 1-27: The Top Row of Polar Plots Show the Equivalent Current System for a Growth Phase Interval Followed by the Two Lowest Order Functions λ1 and λ2 Representing the Two Cell and One Cell Patterns Respectively.

The bottom row of polar plots represents the same information as the top row except for expansion phase conditions. Note, during growth phase conditions, the one cell system is virtually absent, while both one cell and two cell systems are present during expansion phase conditions (after Sun et al. [46]).

In closing, it should be said that the success of this technique depends greatly on the ground based magnetometer array coverage. If there are large gaps in the global array, this will reduce the effectiveness of the technique, since the minimization technique will tend to place the largest errors in regions where there are no data to constrain the result. However, it represents the only effective technique to date that has the potential of separating the effects of steady magnetospheric convection from effects associated with substorm expansion phase activity.

1.3.4.9 Final Comments on Substorms

In the sections above, I have outlined the concept of the substorm from the time of early observations at the beginning of the 20th century until modern times, trying to tie the terminology of the early days to the terminology employed in modern day substorm research. The original substorm definition in the context of auroral observations by Akasofu [1] has changed considerably due to the recognition that substorm activity features two characteristically different disturbances that co-exist (viz. directly driven activity and the loading-unloading behavior of the magnetotail the unloading component of which is associated with the substorm current wedge). At the same time, the recognition of the double oval structure especially evident in the evening hours has made it necessary to be specific about the nature of any auroral intensifications that are observed in the midnight to evening sector of the auroral oval. The researcher now must ask whether any observed auroral activation is a legitimate expansion phase onset, a pseudo-breakup or a poleward boundary intensification (PBI).

Another important point to recognize is that substorm research is not sufficiently advanced to permit the decoupling of effects due to directly driven activity from those due to expansion phase activity. In principle, the mathematical tools are available to attempt such a decoupling (cf. the MNOC technique of Sun et al. [1998]); however, the distribution of ground observatories whose data are needed to establish the equivalent current system for any given event at any moment in time is inadequate to provide the input to the mathematical routines that would produce believable results in detail. The same can be said of the advanced modeling techniques such as KRM (cf. Kamide et al. [43]) or AMIE (cf. Richmond and Kamide [44]) which would be extremely effective if the input data were adequate. Further progress in substorm research using these modeling tools awaits the deployment of adequate global arrays of instrumentation that will provide the input data required to produce detailed estimates of key parameters such as the cross polar cap potential drop and the strength and location of the field-aligned currents that connect the source regions of the outer magnetosphere to the regions of energy dissipation in the upper atmosphere.

Finally, it should be noted that the past decade has provided the best data set that space scientists have had to work with in attempting to develop a clear understanding of the substorm phenomenon. With continuous solar wind coverage from the Wind and ACE satellites and superb topside imaging from the Polar and IMAGE spacecraft, researchers have had the best opportunity ever to examine the details of substorm development and the response of substorm activity to the input of energy from the solar wind. This type of information, together with better ground based arrays of instrumentation, are ultimately going to be required to provide a final understanding of the role that substorms play in the solar-terrestrial interaction.

1.3.4.10 Acknowledgements

This research was supported by the Natural Sciences and Engineering Research Council of Canada.

1.3.4.11 References

[1] Akasofu, S.-I., (1964) The development of the auroral substorm, Planet. Space Sci., 12, 273.

[2] Birkeland, K, (1908) *The Norwegian Aurora Polaris Expedition 1902-1903*, 1, section 1, Aschhoug and Co., Christiania.

[3] Akasofu, S.-I., (1968) *Polar and Magnetospheric Substorms*, D. Reidel, Norwell, Mass.

[4] Dungey, J.W., (1961) Interplanetary magnetic field and the auroral zones, Phys. Rev. Lett., 6, 47.

[5] Axford, W.I. and C.O. Hines, (1961) A unifying theory of high latitude phenomena and geomagnetic storms, Can. J. Phys., 39, 1443.

[6] McPherron, R.L., (1970) Growth phase of magnetospheric substorms, J. Geophys. Res., 75, 5592.

[7] Atkinson, G., (1966) A theory of polar substorms, J. Geophys. Res., 71, 5157.

[8] Baker, D.N., et al. (1996) Neutral line model of substorms: Past results and present view, J. Geophys. Res., 101, 12,975.

[9] Kennel, C.F., (1995) *Convection and Substorms: Paradigms of magnetospheric phenomenology*, Oxford Univ. Press, New York.

[10] Chapman, S., and Bartels, J. (1940) *Geomagnetism*, vol. 1, Clarendon, Oxford.

[11] Obayashi, T. (1967) The interaction of solar plasma with geomagnetic field, disturbed condition, in *Solar Terrestrial Physics*, p.107, ed. by J.W. King and W.S. Newman, Academic Press, London.

[12] Akasofu, S.-I. (1980) What is a magnetospheric substorm?, in *Dynamics of the Magnetosphere*, ed. by S.-I. Akasofu, p. 447, D. Reidel Publ. Co., Dordrecht, Holland.

[13] Baumjohann, W. (1983) Ionospheric and field-aligned current systems in the auroral zone: A concise review, Adv. Space Res., 2, 55.

[14] Sugiura, M., and Heppner, J.P. (1965) The Earth's magnetic field, in *Introduction to Space Science*, ed. by W.N. Hess, p. 5, Gordon and Breach, New York.

[15] Akasofu, S.-I., Chapman, S. and Meng, C.-I. (1965) The polar electrojet, J. Atmos. Terr. Phys., 30, 227.

[16] Rostoker, G. (1969) Classification of geomagnetic disturbances, J. Geophys. Res., 74, 5161.

[17] Kisabeth, J.L. and Rostoker, G. (1973) Current flow in auroral loops and surges inferred from ground-based magnetic observations, J. Geophys. Res., 78, 5573.

[18] Tighe, W.G., and Rostoker, G. (1981) Characteristics of westward traveling surges during magnetospheric substorms, J. Geophys., 50, 51.

[19] Craven, J.D., Frank, L.A. and Akasofu, S.-I. (1989) Propagation of a westward traveling surge and the development of persistent auroral features, Geophys. Res. Lett., 94, 6961.

[20] Sugiura, M. (1964) Hourly values of equatorial Dst for the IGY, *Ann. Int. Geophys Year*, 35, 9.

[21] Rostoker, G., and Fälthammar, C.-G. (1967) Relationship between changes in the interplanetary magnetic field and variations in the magnetic field at the Earth's surface, J. Geophys. Res., 72, 5853.

[22] Akasofu, S.-I., DeForest, S., and McIlwain, C.E. (1974) Auroral displays near the "foot" of the field line of the ATS-5 satellite, Planet. Space Sci., 22, 25.

[23] Burton, R.K., McPherron, R.L., and Russell, C.T. (1975) An empirical relationship between interplanetary conditions and Dst, J. Geophys. Res., 80, 4204.

[24] Gonzalez, W.D., et al. (1994) What is a geomagnetic storm?, J. Geophys. Res., 99, 5771.

[25] Kamide, Y., et al. (1998) Current understanding of magnetic storms: Storm-substorm relationship, J. Geophys. Res., 103, 17,705.

[26] Sharma, A.S., et al. (2003) The storm-substorm relationship: Current understanding and outlook, in *Disturbances in Geospace: The Storm-Substorm Relationship*, ed. by Y. Kamide et al., Geophysical Monograph 142, p. 1, American Geophysical Union, Washington, DC.

[27] Tsurutani. B.T., Zhou, X.-Y., and Gonzalez, W.D. (2003) A lack of substorm expansion phases during magnetic storms induced by magnetic clouds, in *Disturbances in Geospace: The Storm-Substorm Relationship*, ed. by Y. Kamide et al., Geophysical Monograph 142, p.23, American Geophysical Union, Washington, DC 2003.

[28] Zhou, X.-Y., et al. (2003) Ring current intensification and convection-driven negative bays: Multisatellite studies, *J. Geophys. Res.*, *108*(A11), 1407, doi: 10.1029/2003JA009881.

[29] Rostoker, G., (1996) Phenomenology and physics of magnetospheric substorms, J. Geophys. Res., 101, 12,955.

[30] Senior, C., and Blanc, M. (1984) On the control of magnetospheric convection by the spatial distribution of ionospheric conductivities, J. Geophys. Res., 89, 261.

[31] Rostoker, G. (1983) Triggering of expansive phase intensifications of magnetospheric substorms by northward turnings of the interplanetary magnetic field, J. Geophys. Res., 88, 6981.

[32] Wallis, D.D. and Budzinski, E.E. (1981) Empirical models of height integrated conductivities, J. Geophys. Res., 86, 125.

[33] Rees, M.H. (1963) Auroral ionization and excitation by incident energetic electrons, Planet. Space Sci., 11, 1209.

[34] Rees, M.H. (1969) Auroral electrons, Space Sci. Rev., 10, 413.

[35] Spiro, R.W., Reiff, P.H., and Maher, Jr., L.J. (1982) Precipitating electron energy flux and auroral zone conductances – an empirical model, J. Geophys. Res., 87, 8215.

[36] Robinson, et al. (1987) On calculating ionospheric conductances from the flux and energy of precipitating electrons, J. Geophys. Res., 98, 2565.

[37] Hardy, D.A., Gussenhoven, M.S., Raistrick, R., and McNeil, W.J. (1987) Statistical and functional representations of the pattern of auroral energy flux, number flux and conductivity, J. Geophys. Res., 92, 12,275.

[38] Ahn, B.-H., et al. (1998) An ionospheric conductance model based on magnetic disturbance data, J. Geophys. Res., 103, 14,769.

[39] Hardy, D.A., Gussenhoven, M.S., and Holeman, E. (1985) A statistical model of auroral electron precipitation, J. Geophys. Res., 90, 4229.

[40] Rostoker, G., Lui, A.T.Y., Murphree, J.S., and Anger, C.D. (1987) North-south structures in the midnight sector auroras as viewed by the VIKING imager, Geophys. Res. Lett., 14,407.

[41] Zmuda, A.J., and Armstrong, J.C. (1974) The diurnal flow pattern of field-aligned currents, J. Geophys. Res., 79, 4611.

[42] Mishin, V.M., Bazarzhapov, A.D. and Shpynev, G.P. (1980) Electric fields and currents in the earth's magnetosphere, in *Dynamics of the Magnetosphere*, ed. by S.-I. Akasofu, p. 249, D. Reidel Publ. Co., Hingham, MA.

[43] Kamide, Y., Richmond, A.D., and Matsushita, S. ([1981) Estimation of ionospheric electric fields, ionospheric currents, and field-aligned currents from ground magnetic records, J. Geophys. Res., 86, 801.

[44] Richmond, A.D., and Kamide, Y. (1988) Mapping electrodynamic features of the high-latitude ionosphere from localized observations, J. Geophys. Res., 93, 5741.

[45] Lummerzheim, D., et al. (1996) High time resolution study of the hemispheric power carried by energetic electrons into the ionosphere during the May 19/20, 1996 auroral activity, Geophys. Res. Lett., 24, 987.

[46] Sun, W., Xu, W.-Y., and Akasofu, S.-I. (1998) Mathematical separation of directly driven and unloading components in the ionospheric equivalent current during substorms, J. Geophys. Res., 103, 11695.

1.3.5 Ionospheric Troughs at F-Region Altitudes

(Alan S. Rodger, British Antarctic Survey, Cambridge, UK)

1.3.5.1 Introduction

The term ionisation trough is applied to a region of the mid- to high-latitude F layer, typically a few degrees wide in latitude wide, where the plasma concentration is low, with the plasma concentrations both poleward and equatorward of this region being higher. Troughs in the polar ionosphere are normally considered in two forms, the mid-latitude or main ionospheric trough, usually occurring just equatorward of the auroral precipitation boundary, and the high latitude trough observed within the auroral oval or polar cap. Troughs have major effects on the propagation of radio waves both because of the strong gradients in plasma concentration and also because the edges are often regions of high plasma irregularity occurrence. Troughs comprise three parts, the equatorward, and poleward edges and the trough minimum (Figure 1-28). To date there is no agreed quantitative definition of a trough making intercomparisons of results from different analyses subjective. Despite these limitations, the overall morphology of both high- and mid-latitude troughs is reasonably well known, and summarised in Rodger et al. (1992). The key features of trough morphology, and an overview of the physics of trough formation are summarised here, and some areas for further studies in order to improve the predictions of troughs characteristics are identified.

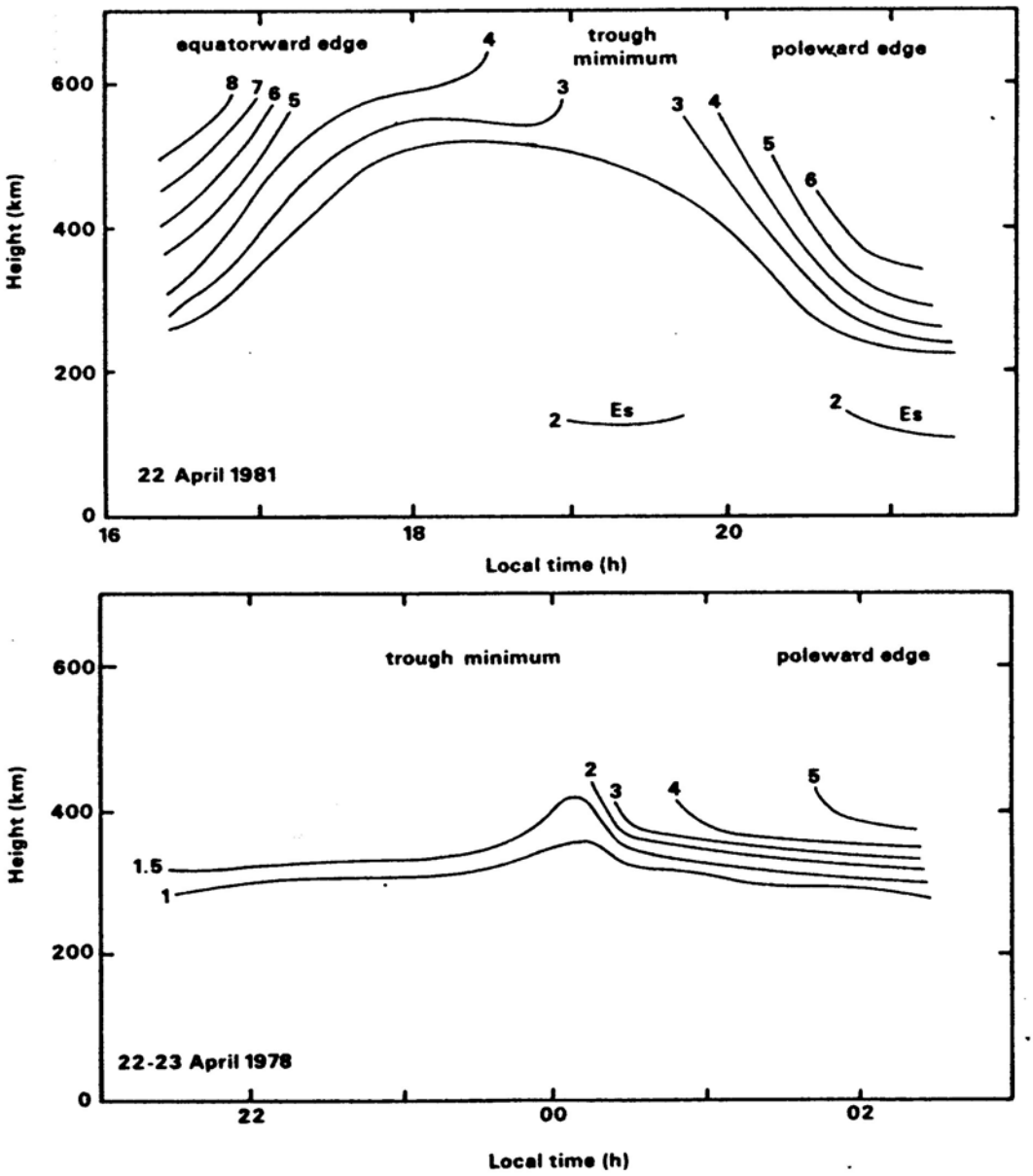

Figure 1-28: Schematic Diagram of the Mid-Latitude Trough for Two Occasions; the upper panel is representative of an evening sector trough with a well-formed equatorward edge, whereas the lower panel is more typical of a morning sector trough with an ill-defined equatorward edge. The numbers associated with the contours indicate plasma frequency in MHz.

1.3.5.2 Mid-Latitude Trough Morphology

The mid-latitude trough is primarily a phenomenon that occurs in darkness thus is seen mainly in winter and equinox and is less common in summer. The local time extent of the trough is small in summer, centred around midnight and extends further towards dawn and dusk with progression towards winter. In some longitudes, and under the quietest magnetic conditions, the trough can be observed at all local times around mid-winter. An excellent example of this dayside main trough is presented by Pryse et al. (2005) who show a deep trough at the equatorward edge of the auroral oval. These authors use the term ‘high-latitude trough’ owing to the geographic latitude, rather than the physical definition suggested above.

During steady geomagnetic conditions, the trough moves progressively to lower latitudes with later magnetic local times and is observed at lower latitudes at the same magnetic local time as geomagnetic activity increases. There are several empirical equations that describe this motion reasonably well (Dudeney et al., 1983). Most formulae use the Kp magnetic index, which has a number of limitations, an index related to the solar wind conditions might lead to improved representations. As solar activity increases the electron concentration in trough minimum and its edges also increases.

1.3.5.3 Formation Processes of the Mid-Latitude Trough – Steady State

During steady geomagnetic conditions, the trough in the evening sector forms immediately equatorward of the low latitude boundary of auroral precipitation. This is a region where the corotation and convection electric fields are oppositely directed (i.e. quasi stagnation) and hence the plasma residence time at a particular magnetic local time is greater than at lower latitudes. Provided the plasma is in darkness, then normal F-region plasma recombination processes cause the plasma to decay further compared with the slightly lower latitudes. Poleward of this region in the evening sector there are two major causes of the enhancement of plasma. First energetic particle precipitation causes additional ionisation both at E- and F-region altitudes. Sporadic E layers can be formed in addition to enhancements of the F-region concentration (see Figure 1-28). Secondly, some plasma forming the poleward edge has been transported from the dayside ionosphere, through the cusp, polar cap and nightside oval contributing to the elevated plasma concentration. Hence a trough is formed i.e. a region of low plasma concentration between two regions of higher plasma concentration.

In the morning sector, the trough minimum is the extension of the evening sector trough that is slowly convecting and/or corotating towards dawn. This convection can cause the trough to be observed slightly (~1 h) beyond the day-night terminator. The equatorward edge is formed by corotating plasma which continues to decay with time becoming progressively less distinct from the trough minimum with time (Figure 1-28 lower panel). The poleward edge is formed by plasma that has drifted across the polar cap from the dayside ionosphere with some enhancement in the plasma concentration as the flux tubes drift through the nightside oval and then to lower latitudes. The poleward edge occurs equatorward of the auroral oval and in this respect is different from the evening sector trough.

Thermosphere-ionosphere general circulation models reproduce the steady-state conditions reasonably well. This includes the variations that occur as a function of season, Universal Time (UT) (including the effects of the offset of the geographic and geographic poles), and the variations of the B_z and B_y components of the interplanetary magnetic field. The UT effects lead to particularly strong seasonal variations in trough shape, longitude extent, and gradients of the equatorward and poleward walls of the trough; these features vary as a function of solar activity.

In summary, the mid-latitude trough is a delicate balance between plasma production in the vicinity of the poleward edge of the trough, horizontal transport arising from the convection electric field around the trough minimum, and recombination processes particularly in the trough minimum and the equatorward edge. The incorporation of ionospheric outflow during the daytime and its subsequent return to the F region at night affects markedly the plasma concentration in the trough minimum, and is probably the most difficult parameters to model realistically.

1.3.5.4 Formation Processes of the Mid-Latitude Trough – Geomagnetically-Active Periods

During geomagnetically-active periods, the detailed morphology of troughs becomes much more difficult to predict although the physics is essentially identical. The fundamental problem is that the spatial and temporal variations of critical processes are not sufficiently well known to allow accurate modelling to be undertaken. The most important unknown is the spatial and temporal variation of electric fields. The precise nature of the trough morphology depends critically on the time history of the convection

pattern and hence the trajectory of plasma in the sub-auroral and high latitude ionospheres. As the lifetime of F-region plasma is many hours, especially in winter, and plasma only takes 1-2 hours to convect from the cusp into the nightside auroral oval, the electric field over the entire high latitude region is required, probably with a temporal resolution of a few minutes, to be able to predict the electron concentrations and gradients accurately.

There is a wide array of space and ground-based measurements now available in near-real time to assimilate into prediction models. Particularly valuable are the ever-increasing number of SuperDARN radars, TEC receivers, ionospheric tomography (Kersley et al., 1997) and magnetometers, but still this is insufficient information to give reliable global forecasts. Particular challenges occur where plasma velocities (**V**) exceed about 500 ms^{-1} in the rest frame of the neutral wind (**U**). On these occasions, enhanced plasma loss occurs (Schunk et al., 1975). It is a highly non-linear process: an increase of plasma velocity in the rest frame of the neutrals (**V**-**U**), from 1 to 2 kms^{-1} leads to an order of magnitude increase in the plasma loss rate. The neutral wind is often assumed to be zero but this too can lead to serious errors both in the loss rate, and in the field-aligned plasma motion. Some plasma velocity enhancements occur in narrow regions (a few km) associated with auroral arcs. Thus deep narrow troughs can be formed in minutes which are not detectable by many types of instruments, but they affect plasma concentrations and irregularity occurrence and hence radio wave propagation.

Penetration of the electric field from high to mid and low latitudes is one of the areas of most active research at present and a topic that has great relevance for the formation and morphology of the mid-latitude trough. A wide variety of terminology for the phenomenon has been used over the last three decades, including polarisation jet (Galperin et al. 1973), sub-auroral ion drift (SAID) (Spiro et al., 1979), substorm-associated radar auroral surges (SARAHs) (Freeman et al., 1992), and sub-auroral polarisation streams (SAPS) (Foster and Burke, 2002). The plasma jets not only are regions of enhanced loss, but they also transport low concentration, night-side plasma into the dayside ionosphere. As geomagnetic increases and the auroral oval and convection patterns expand, high plasma concentration mid-latitude plasma is entrained into the high-latitude convection pattern and lead to storm enhanced densities (SEDs), particularly in the afternoon sector (Foster et al., 2002). GPS measurements of total electron content have helped very significantly in determining the spatial and temporal variations of plasma concentrations during storm time but accurate prediction of the penetration fields is still some way off.

Energetic particle precipitation is particularly important in forming the poleward edge of the mid-latitude trough. As with electric fields, the general morphology is well known but the detail of where, when and the spectrum of the precipitating particle flux is less predictable. As a simple example, the first substorm normally occurs within about 1 hour after a southward turning of the interplanetary magnetic field but the longitude of initiation and the extent of the substorm cannot yet be predicted, so the pattern energy deposition from the magnetosphere into the ionosphere cannot be predicted.

Energetic particle precipitation also leads to increased conductivity, which in turn leads to enhanced Joule heating in the presence of an electric field. This is important for trough formation in two ways. First it can increase or decrease the neutral wind locally causing the altitude of the plasma to change, and hence affect the lifetime of plasma. Also regions of high Joule heating is where there is rapid ion outflow which contributes to plasma depletion at F-region altitudes i.e. a trough. These factors contribute significantly to the difficulty in predicting accurately trough characteristics.

1.3.5.5 High Latitude Trough Morphology

The morphology of high latitude troughs is less well defined mainly because of the less comprehensive sets of ground-based observations in very high latitudes due to the logistic difficulties of establishing instrumentation near both poles. Satellite data have played a key role in trough identification. Figure 1-29 gives an example of a trans-polar crossing showing trough structures both within the polar cap and the

auroral oval. This figure also illustrates the problem of having no agreed definition of a trough, as many trough-like features can be identified. The physical processes for high-latitude trough formation are essentially the same as those of the mid-latitude trough, i.e. a very delicate balance between various plasma transport processes, production and loss of plasma and again the combination of them is unique on each occasion.

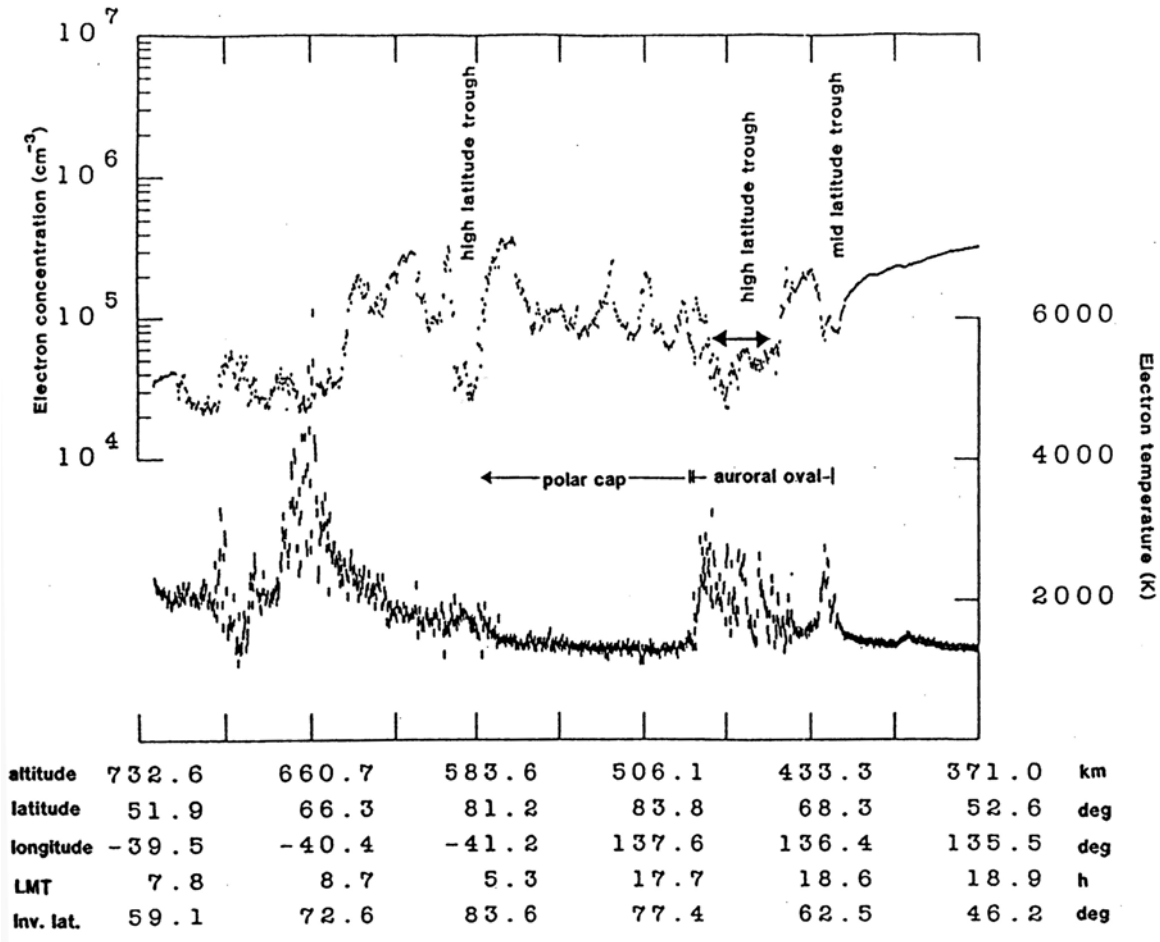

Figure 1-29: Latitude Profile of Electron Concentration (left-hand scale) and Electron Temperature (right-hand scale) from Dynamics Explorer 2 Spacecraft, 22 November 1981 Showing the Relationship of High-Latitude Troughs with Respect to the Mid-Latitude Trough, Auroral Oval and Polar Cap.

The most familiar high-latitude troughs are those that form deep in the polar cap when plasma convection trajectories do not experience sunlight for very extended intervals (many hours) and plasma concentration in the F region decays with time – these are called polar holes.

High-latitude troughs have also been found to be associated with the boundary between a lobe cell and a viscous convection cell during periods of IMF northward conditions (Pryse et al., 2006). This is a region where there is a large horizontal velocity shear, and hence field-aligned currents and intense energetic particle precipitation. Adjacent convecting flux tubes have a very time history, which can result in trough-like features at F-region altitudes. These effects can be sufficiently large that troughs can be observed even when the in the high-latitude ionosphere is in total sunlight in summer (see Figure 1-30 for example).

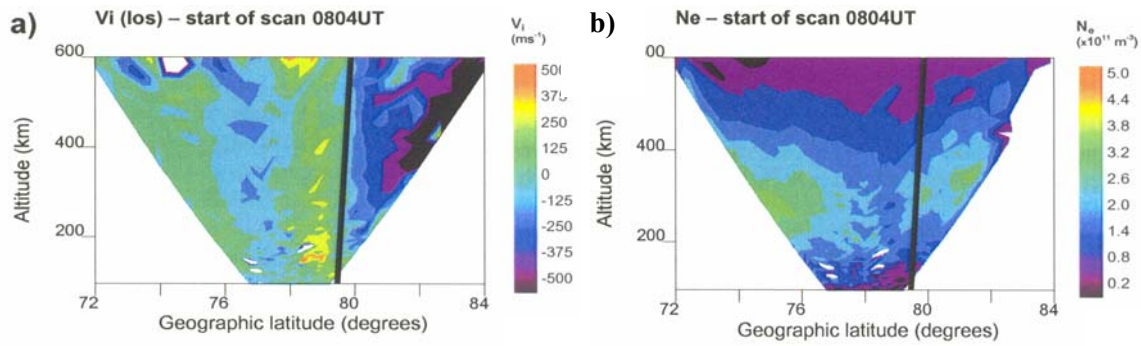

Figure 1-30: The Line-of-Sight Velocities Determined from the EISCAT Svalbard Radar at 0804 UT on 24 August 2000 with the Corresponding Electron Concentrations. The black line marks the centre of a significant velocity shear with a high-latitude trough lying close to the shear (after Pryse et al., 2005).

Considerable structuring of F-region plasma occurs in the vicinity of the polar cusp. High-concentration plasma enters the polar cap on the dayside, but can become highly structured by the effects of variable, and large magnitude (>1 km s^{-1}) electric fields, resulting in the formation of polar patches (Rodger et al., 1994a). These structures persist right across the polar cap and enter the nightside oval (Rodger et al., 1994b).

A further way of forming troughs in the polar cap is the effects of high plasma velocities (500 ms^{-1}) associated with trans-polar arcs that cause rapid recombination and hence reduce the plasma concentration at F-region altitudes (Schunk et al., 1975).

The relative importance of these various mechanisms, or the precise geophysical conditions under which they occur has yet to be fully established.

1.3.5.6 Focus for the Future

Scientific progress is often made by the inter-linking of four approaches. These are observations (the crucial link to reality), data assimilation techniques, sophisticated modelling developments reflecting an ever-increasing understanding of the underlying science, and new scientific theories. Improved trough predictions require some contribution from each of these elements.

There are ever-increasing quantity of high-quality ionospheric, magnetospheric and solar wind observations available in near-real time that can be assimilated into mature general circulation models. However observations of key parameters, such as electric fields, are not likely to be available at a sufficiently small scale to describe very detailed plasma motions. An alternative approach would to parameterise the electric field using the knowledge that the amplitude of electric field fluctuations in the polar cap can be described by a power law relationship that is very similar to that of the Poynting flux at 1 AU in the solar wind (Freeman et al., 2000). Variations of the electric field within the auroral oval also can be described by a power law but with a different exponent (Abel, private communication). However there are significant challenges to know how to relate the variations in electric field to those of ionospheric conductivity, recognising, for example, that the peak in the electric field associated with an auroral arc is not collocated with the peak conductivity. This makes the accurate calculation of Joule heating particularly challenging.

The second area that is in most need of scientific progress is a quantitative understanding of a few key processes, the most important of which are probably energy deposition as a function of space and time in a substorm, and the geophysical factors that control the penetration of the electric field from high to lower latitudes and flow into and from the topside ionosphere. Substorms have been studied for half a century

but the forthcoming THEMIS spacecraft mission with very comprehensive, complementary ground-based measurements offer some new possibilities to understand the instability processes that initiate substorms in the tail of the magnetosphere and hence predict energy deposition. The spatial/temporal variations of the region 2 and ring current regions must hold the key to the penetration of the electric field to lower latitudes. Modelling the ionospheric outflows needs further work to understand the acceleration processes at low altitude, and more realistic treatment of the upper boundary condition of thermosphere-ionosphere general circulation models is also necessary.

1.3.5.7 References

Dudeney, J. R., A. S. Rodger and M. J. Jarvis, Radio studies of the main F-region trough in Antarctica. Radio Science, 18, 927-936, 1983.

Foster, J. C., P. J. Erickson, A. J. Coster, J. Goldstein and F. J. Rich, Geophys. Res. Lett., 29, 10.1029/ 2002GL015067, 2002.

Foster J. C. and W. J. Burke, SAPS: a new categorization of sub-auroral electric fields, EOS. Transactions, American Geophysical Union 83, 393-394, 2002.

Freeman, M. P., D. J. Southwood, M. Lester, T. K. Yeoman and G. D. Reeves, Substorm-associated radar auroral surges, J. Geophys. Res. 97, 12173-12185, 1992.

Freeman M., N. W. Watkins and D. J. Riley Power law distributions of burst duration and interburst interval in the solar wind: Turbulence or dissipative self-organized criticality? Physical Review E 62 (6): 8794-8797 Part B, December 2000.

Galperin, Y. I., V. L. Khalipov and A. G. Zosimova, Direct measurement of ion drift velocity in the supper ionosphere during a magnetic storm, 2. Results of measurements during the November 3, 1967 magnetic storm, (In Russian), Cosmicheskie Issledovanija, 11, 273, 1973.

Kersley, L., S. E. Pryse, I. K. Walker, J. A. T. Heaton, C. N. Mitchell, M. J. Williams and C. A. Wilson, Imaging of electron density troughs by tomographic techniques, Radio. Sci. 32, 1607-1621, 1997.

Pryse S. E., K. L. Dewis, R. L. Balthazor, H. R. Middleton and M. H. Denton, The dayside high –latitude trough under quiet geomagnetic conditions: radio tomography and the CTIP model, Ann. Geophys. 23, 1199-1206, 2005.

Pryse S. E., R. W. Sims, J. Moen and K. Osavik, Ionspheric signatures of low-latitude the boundary layer under conditions of northward IMF and small clock angle, Ann. Geophys. 24, 2169-2178, 2006.

Rodger, A. S., R. J. Moffet, and S. Quegan, The role of ion drift in the formation of ionisation troughs in the mid- and high-latitude ionosphere – a review, J. Atmos. Terr. Phys., 54, 1-30, 1992.

Rodger, A. S., M. Pinnock, J. R. Dudeney, K. B. Baker, R. A. Greenwald, A new mechanism for polar patch formation. J. Geophys. Res., 99, 6425-6436, 1994a.

Rodger, A. S., M. Pinnock, J. R. Dudeney, J. Watermann, O. de la Beaujardiere and K. B. Baker, Simultaneous two-hemisphere observations of the presence of polar patches in the night-side ionosphere, *Annales Geophysicae*, 12, 642-648, 1994.

Schunk, R. W., P. M. Banks and W. J. Raitt, Effects of electric fields and other processes upon the night-time high latitude F layer, J Geophys, Res., 80, 3121-3130, 1975.

Spiro, R. W., R. A. Heelis and W. B. Hanson, Rapid sub-auroral ion drift events observed by Atmospheric Explorer C, Geophys. Res Lett. 6, 660-663, 1979.

1.3.6 Small Scale Irregularities at High Latitudes

(J.-P. St.-Maurice, Institute of Space and Atmospheric Studies, University of Saskatchewan; A.M. Hamza, Department of Physics, University of New Brunswick)

The text for this section was been provided as a separate PDF document – *click here to view Section 1.3.6.*

1.3.7 Characterizing the Polar-Cap Ionosphere

(Herbert C. Carlson – CASS, Utah State University, 4405 Old Main Hill, Logan UT, 84322-4405 and AFOSR Chief Scientist, 875 N Randolph St, Arlington VA, 22203-1954)

1.3.7.1 Introduction and Motivation

Here we address polar-cap phenomena, notably polar-cap patches and sun-aligned arcs, as they affect the ionosphere. The intent is to summarize current knowledge without the extensiveness of a full review. Several full reviews of knowledge of polar-cap patches and polar-cap sun-aligned arcs have appeared in the last decade. For patches and blobs the reader is referred to Crowley (1996). For polar-cap patches and arcs, Radio Science special section on CEDAR HLPS/STEP GAPS, 31, 573-678, (1996); Basu and Valladares, (1999). For sun-aligned arcs, Smith and Lockwood, (1996); Zhu et al., (1997); Sandholt et al., (2002). Also, Space Weather issues in general are discussed by Schunk and Sojka, (1996).

Because understanding of the polar-cap ionosphere is of practical importance to communication, GPS navigation, and satellite radar imagery, description of its phenomena and physics here includes reference to how research advances improve our ability to utilize or circumvent its effects on radio frequency propagation.

It was in fact the launch of a polar-orbiting satellite, experiencing communication outages, which led to the widespread recognition that the polar-cap ionosphere could be of any consequence to radio-frequency signals from satellites. Such outages had been predicted five years earlier by a small group of researchers, based on theoretical extrapolation of sunspot minimum data. However, without actual sunspot maximum data, the prediction could not be definitively proven; that is, not until the other half of the 11-year solar cycle had passed. A robust research program based on an experiment-theory-model has since led to actionable predictions, and operations in the polar cap have benefited. Improved definition of morphology, correlations and patterns, and understanding leading to physical models, now-casts, and forecasts, have aided design, and enabled mitigation and circumvention of many operational issues.

To understand the polar ionosphere, it must be viewed as alternating between two fundamentally different states. It changes its basic character in response to reversals of the sign of the north-south component (Bz) of the interplanetary magnetic field (IMF). The other components (By, Bx) of the IMF have important but lower-order impact on its properties. Changes can occur over scales of minutes or less through the IMF control of phenomena within the polar cap. These rapid changes are perhaps what most distinguish the polar-cap ionosphere from the rest of the globe, with which it shares other strong dependencies on the solar cycle, and on seasonal and diurnal variations. Distinction blurs, though, between the latter two when deep in the polar cap, where there are months of continuous darkness or daylight at midwinter and midsummer.

To quantify the degree of degradation of radio frequency signals from satellites, see Figure 1-31. It shows eight years of polar-cap scintillation data from Thule, Greenland, monitoring 244 MH signals from a

USAF polar beacon satellite, at high elevation angles which do not accentuate ionospheric scintillation effects. Fifteen-minute fade depths exceed 10 db over 50% of the time for each of the five winters sampled during the sunspot maximum years 1979-1982, but are present less than 5% of the time for sunspot minimum years 1985-1986. The fades reach 20 db at least 10% of the time in mid-winter for these same five solar max years.

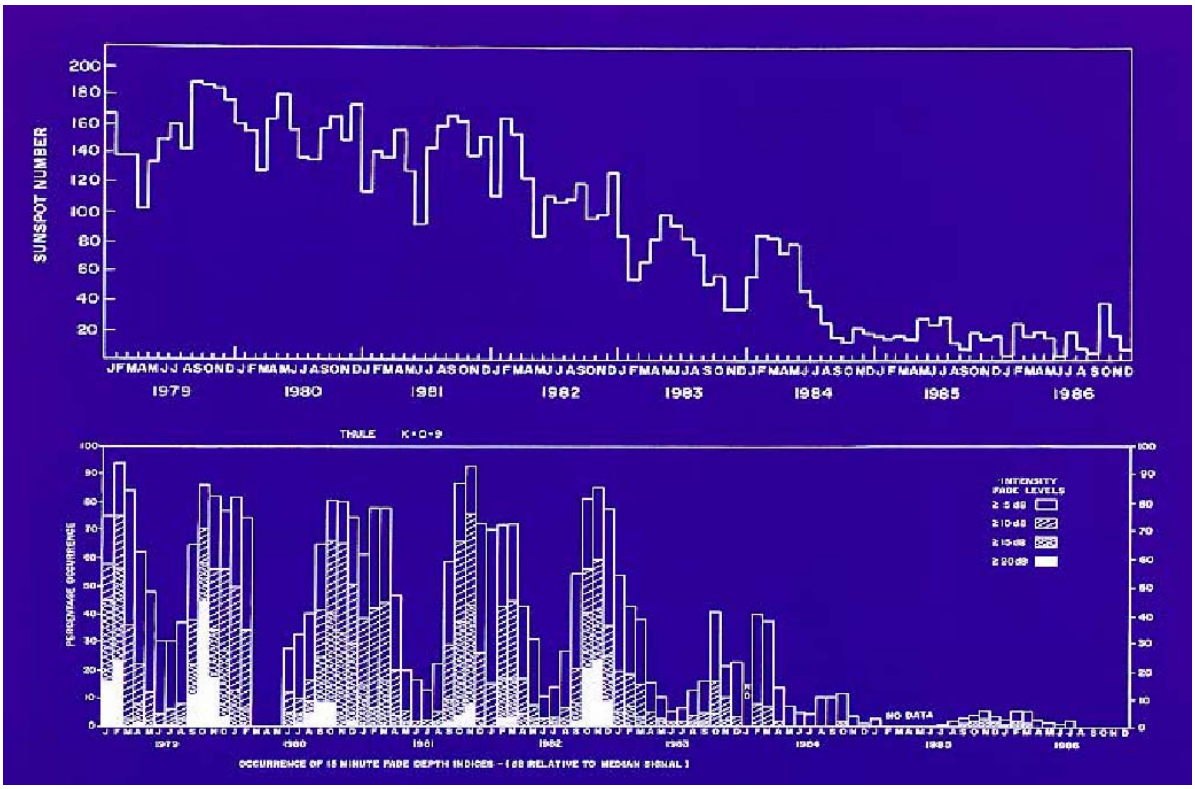

Figure 1-31: Polar-Cap Scintillation Variation with Solar Cycle, for 244 MHz Signals from a High-Elevation USAF Polar Beacon Satellite, Received at Thule, Greenland During 1979-1986. Sunspot number is in the upper plot, percent occurrence of 15-minute fade depth indices are in the lower plot for 5, 10, 15, and 20 db fade relative to median signal. (Basu et al., 1988)

1.3.7.2 The Two States of the Polar Cap

The polar ionosphere and upper atmosphere alternate between two states, depending on whether the IMF is southward or northward. This alternation is due to the very different degree of coupling between the solar wind and the magnetosphere, for opposite signs of the Z component of the IMF B_z. For southward IMF, called negative B_z, the coupling is very strong, and such strong coupling can have severe effects on radio signals. Ground-based optical and radio measurements in the polar cap respond in a way so reliably and so strongly coupled that their measurement can show which of these two states prevails. This correlation was established through the use of satellites measuring the full IMF vector (and solar wind speed, density and temperature). Satellites "parked" at the location where earth's and sun's gravitational fields balance are upstream of earth in the solar wind, and can therefore give us early warning of IMF changes about an hour before corresponding conditions switch at earth.

Below we summarize the salient optical and radio signatures that characterize the southward and northward states.

1.3.7.2.1 Southward IMF State

The state for southward IMF is characterized in midwinter by ~100-1000 km islands, or patches, of enhanced F-region plasma, surrounded by lower-density plasma. As these patches pass overhead, the overhead ionosphere can switch sharply between peak plasma densities ~10^{11} m^{-3} to ~10^{12} m^{-3} (~10^5 cm^{-3} to ~10^6 cm^{-3}) (Weber et al., 1984), values which correspond to an ionospheric f_oF_2 of 3 to 9 MHz in the HF band. This excursion is literally from night to day (from typical electron densities found near midnight to near midday at mid-latitudes). The term "patches" derives from their patchy appearance in 630.0 nm imagers, which discovered them (Weber and Buchau, 1981). By convention (Crowley, 1996) high electron density islands, double or more their surrounding densities, are termed patches. Patches move antisunward across the polar cap, at ~1 km/s from near noon toward midnight. A patch formed over Svalbard, Norway, passes over Alaska about an hour or two later. Its transpolar trajectory is taken to be consistent with a normal two-cell convection pattern (Heelis, 1984) defined by averaged flow patterns of ionospheric plasma in the polar cap for southward IMF.

Figure 1-32 shows a representative patch (Weber et al., 1984). The upper sequence of frames shows 630.0 nm intensity images overhead. The field of view is about 155° of an overhead hemisphere in this fish-eye view of the All-Sky Imaging Photometer (ASIP). The bright form moving from noon (bottom of first frame) toward midnight (top of fifth frame) is the signature of the high-density plasma patch. That it is associated with high-density ionospheric plasma is established by its coincidence with approaching, then passing overhead, then receding, Doppler-shifted HF echoes as shown in the lower frames of the figure. Approaching plasma, as indicated by upward Doppler shift, is in the lower set of ionogram panels; receding plasma is in the upper set of ionogram frames. The enhanced plasma is of an ionospheric f_oF_2 of ~9 MHz, the background ~3 MHz. The 630.0 nm emission is due to the chemically-stored energy released by dissociative recombination of the ionospheric plasma, which first excites the lowest excitation state of atomic oxygen $(OI)^1D$ upon recombination, then within about a minute emits a 630.0 photon as the excited state relaxes back to its ground state. The 9 MHz ionosphere has ~ten times the number of recombining electron-ion pairs per unit time, hence about ten times the atomic oxygen $(OI)^1D$ excitation rate as the 3 MHz, and thus is ~ten times as bright in the 630.0 nm emission. Hence the bright patch. It is not aurora, it is just standard atomic oxygen recombination red-line airglow from nighttime F-region ionospheric recombination. For the nominal 250 km altitude of this airglow, the ASIP field of view is ~1000 km.

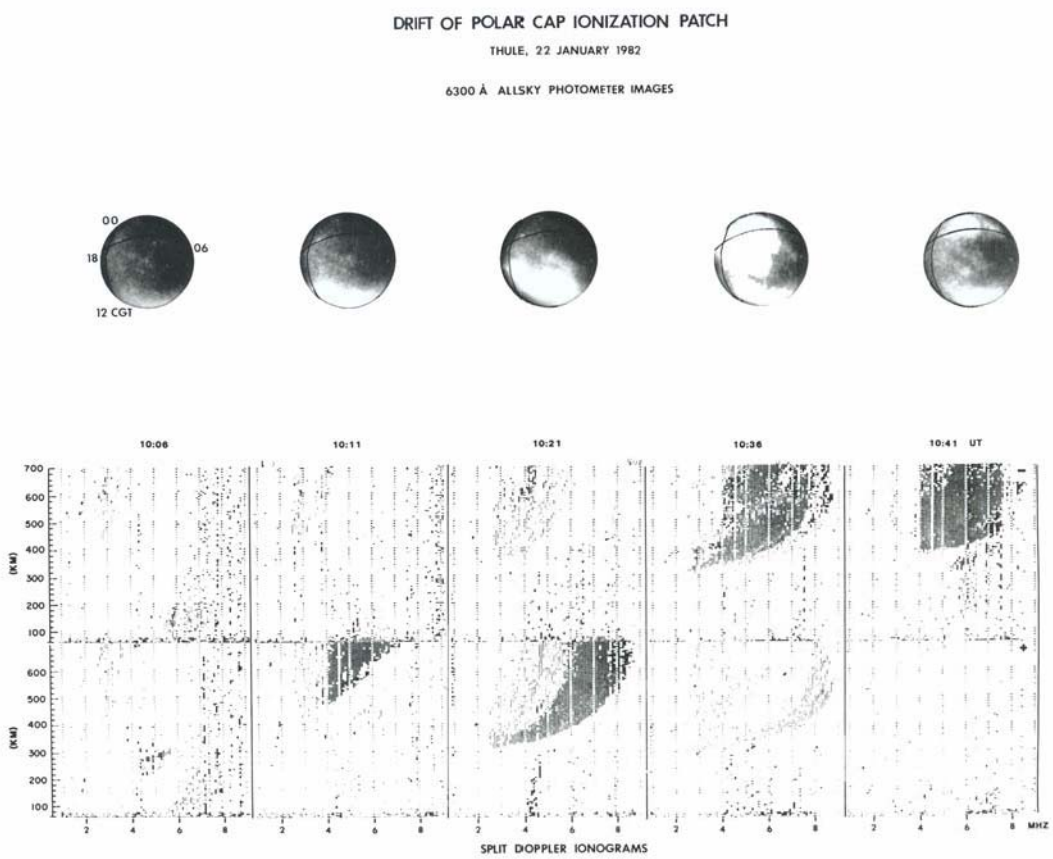

Figure 1-32: The Overhead Passage of the 630.0 nm Bright Patch Seen on the ASIP Coincides with the Overhead Passage of the High-Density Plasma Enhancement (from 3 to 9 MHz f_oF_2) Seen on a Digisonde Sensitive to Doppler Intensity and Range Discrimination (Weber et al., 1984).

The plasma within these enhanced plasma patches is highly structured over scale sizes spanning more than 0.l to tens of km, producing severe scintillation (Basu et al., 1988). These electron-density irregularities are produced by a plasma instability, the gradient drift instability, with the horizontal electrical force acting across the density gradients. The patches are less obvious in summer when the patches, still present, are of lower contrast against the higher background sunlit F-region polar plasma (Buchau et al., 1985). They are also far less structured, because the conducting E-region plasma largely shorts out the instability process.

It is thought that these patches exit the polar cap near midnight, to become the “blobs” (Vickrey et al., 1988) often seen on winter nights, in the post-sunset subauroral return-flow region. These cigar-shaped high plasma density features are also highly structured in plasma-density irregularities. Thus they represent a source of transient rf disruption, just equatorward of the polar-cap boundary.

1.3.7.2.2 Northward IMF State

The other state of the polar-cap ionosphere is characterized in optical emission by sun-aligned arcs. Sample observations, to contrast arcs with patches, are illustrated in the top pair of panels of Figure 1-33. The northward IMF panel shows a sun-aligned arc emerging from the dawn edge of the auroral oval and moving duskward toward the central polar cap. The motion is tracked at equal five-minute intervals per frame over a half-hour period. The arc moved steadily, persisting nearly an hour before finally starting to

fade. The bright edge at the left of the frame is the poleward edge of the polar cap (highly contracted for these northward IMF conditions) near dawn. The field of view is as in Figure 1-32. Observations from selected pairs of stations within the nighttime polar cap show that single discrete sun-aligned emissions can extend very large distances, > 2000 km for the nominal 250 km of typical F-region 630.0 nm emission, in the noon-midnight direction.

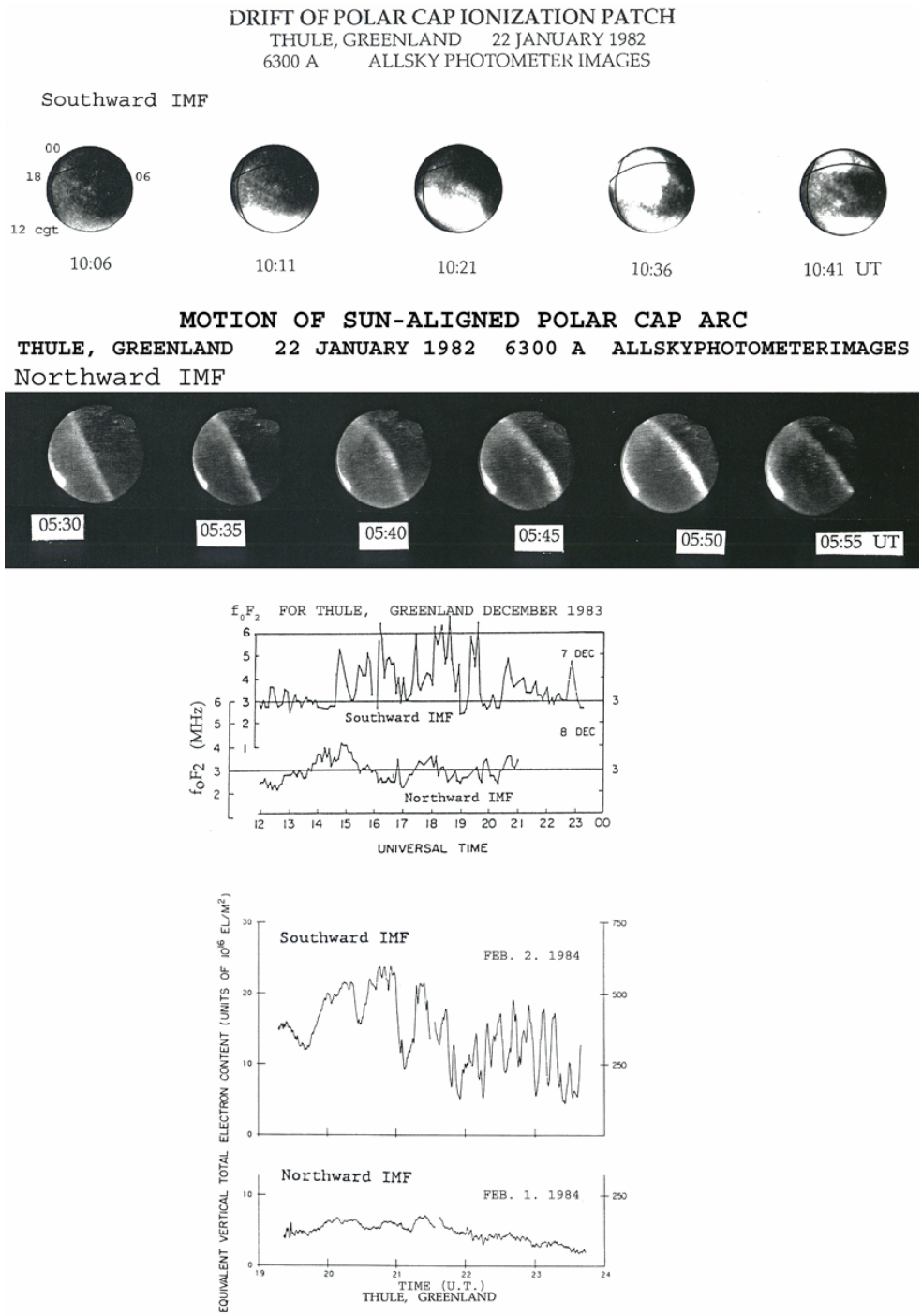

Figure 1-33: This Figure Contrasts the Signatures of the Southward vs. Northward IMF Conditions of the Polar-Cap Ionosphere – in, respectively, the upper vs. lower portions of each panel pair from: ASIP images, a digital ionosonde, and a TEC polarimeter.

The data are all overhead Thule, Greenland, in the central polar cap, during local winter. The ASIP shows an antisunward drifting patch in top panel vs. dawn-dusk drifting sun-aligned arc. For IMF south the digisonde shows sporadic strong enhancements of f_oF_2 (up to ~9MHz) above the low background nighttime level of ~3 MHz, and the TEC data shows a series of TEC enhancements about an average near 5 TEC units, both signatures of patches. For IMF north, the levels are relatively flat near the baseline f_oF_2 and TEC values on which patch enhancements sit for southward IMF. Strongly enhanced scintillation due to plasma irregularities of 0.1 to tens of km scale size coincide in time and space with overhead patches seen by any of the three diagnostics for southward IMF, and with sun-aligned arcs for sunspot maximum northward IMF conditions.

Complementary measurements by other ground-based and satellite sensors have allowed us to determine that the sun-aligned arcs are excited by sun-aligned sheets of precipitating electrons, which mark shear lines of transpolar ionospheric plasma flow. The fact that the arcs are the optical signature of electron-impact excitation of the upper atmosphere by suprathermal electron current-carriers, means that the sun-aligned arcs are true aurora. Their low intensity (100s of R) is due to weak fluxes of soft, low-energy electrons, order hundreds of eV. The state of the polar-cap ionosphere for positive B_z is characterized by considerable structure along the dawn/dusk direction in its plasma velocities, and weak but stable sun-aligned arcs seen in optical emissions.

Sun-aligned arcs were initially thought to be present only a few percent of the time (Ismail et al., 1977; Gussenhoven, 1982; Frank et al., 1986) based on observations with a sensitivity threshold of order I kR. Now we know they are prevalent for northward IMF, present in that half of the time when B_z is positive. Studies have led them to be known now as markers of highly anisotropic transpolar velocity gradients (Carlson et al., 1988). Antisunward plasma flow, decreasing (or even reversing) along a dawn-to-dusk direction, requires a vertical upward current to maintain divergence-free current across the arc. The arc emissions are impact-excited by the downgoing current-carrying electrons. Near the central part of the polar cap, the arcs drift duskward (dawnward) for positive (negative) IMF B_y at a non-uniform speed ~0.1 - 0.2 km/s) (Weber et al., 1984; Frank et al., 1986; Valladares et al., 1994). The F-region plasma above these arcs is also structured, in this case driven by a different class of plasma instability, a shear-driven instability (Basu et al., 1998). Scintillation is strong over these arcs only during sunspot maximum. During sunspot minimum, the F-region densities are much lower and the scintillation much weaker. These arcs are presumed to connect to the dayside auroral oval and are known to connect to the midnight auroral oval, although most observations are confined to the central-polar regions, with very little data available for the area of the dayside auroral region. Above the optical signature of the arcs, current sheets must then extend along the earth-sun direction, over the quasi-transpolar length of the arc, reaching upwards into the magnetosphere to as far as the electron current source.

1.3.7.2.3 Framework for Character and Morphology

A framework for considering the polar-cap ionosphere in terms of northward vs. southward IMF is summarized in Figure 1-33 and Figure 1-34. Figure 1-33 shows what one sees from a ground-based or airborne polar-cap location on earth, i.e., the near-earth frame of reference. Figure 1-34 shows how the field of view of Figure 1-33 fits into the context of temporal dynamics across the full polar cap, as viewed from above the polar cap in the earth-sun frame of reference.

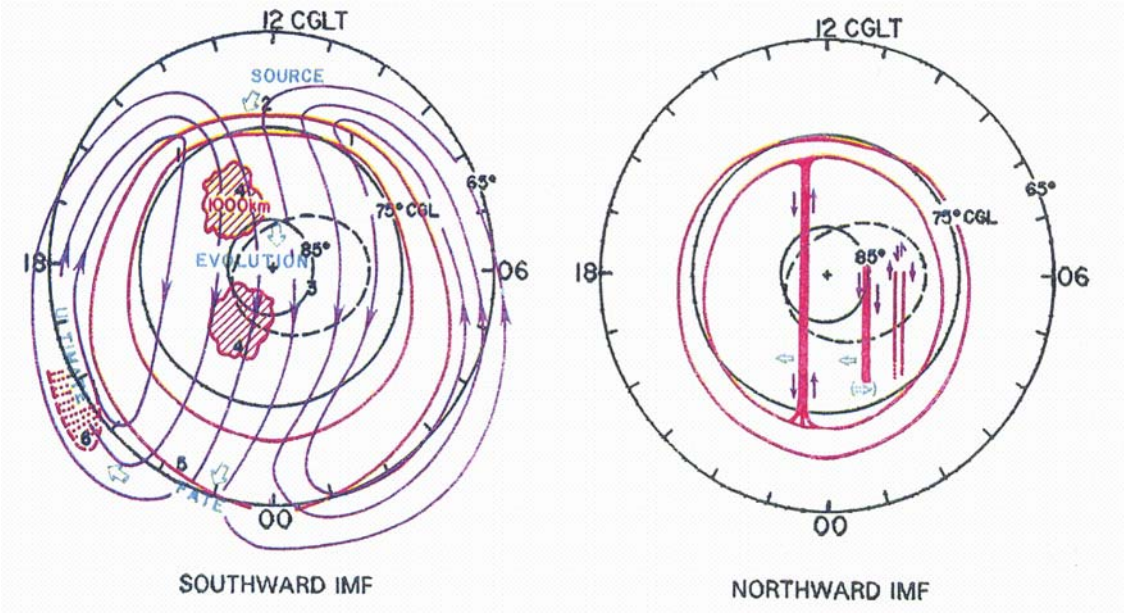

Figure 1-34: View in Earth-Sun Frame of Reference of Conditions Characterizing the State of the Polar-Cap Ionosphere for Southward and Northward IMF and Positive B_y.

For negative B_y the patterns are mirror-imaged about a noon-midnight axis of symmetry. A nominal auroral oval location is shown for each condition, on CGLT coordinates from 65o magnetic latitude, looking down on the north magnetic pole. Convection is shown by closed flow lines with arrows overlaid (southward IMF) or thin adjoining arrows (northward IMF). Cross-hatched areas show recombination 630.0 nm-enhancements from high plasma density patches (southward IMF); sun-aligned bars show electron impact 630.0 nm-enhancements (northward IMF). Open arrows show the direction of motion of these polar 630.0 nm features. The dashed-line ellipse near the central polar cap is the field of view of an ASIP at the latitude of Thule, for 250 km emissions. Note that for this CGLT coordinate system, fixed magnetic latitudes are circles, so circular fields of view appear as ellipses. (Carlson, 1994)

All these descriptions are for dark northern hemisphere polar-cap winter conditions. The same physical phenomena must occur in the summer sunlit pole, but the degree of change in ionospheric properties and signatures will be largely masked and/or subdued by the ionospheric plasma locally produced by the sunlight (EUV for the F-region ionosphere). Plasma transported from large distances dominate the character of the winter conditions, whereas locally-produced plasma dominate the summer conditions.

The polar ionosphere for southward, vs. northward IMF, is characterized by quite different properties of the overhead sky seen by: an ASIP, a digital ionosonde (digisonde), or a total electron content (TEC) polarimeter receiver of signals from a satellite radio beacon. Thus, such ground-based measurements can serve as surrogate indicators of the north-south sense of the IMF.

1.3.7.3 Outline of Effects on rf Signals

The primary concerns at HF and VHF are clutter, scatter, and multi-path. Whereas, at and above VHF through GHz, they are amplitude and phase scintillation, leading to signal drop out and loss of phase lock.

Patches (and sun-aligned arcs at sunspot maximum) affect radio frequency signals at all wavelengths, so frequency diversity is of little value. The higher the frequency, the weaker the scintillation. Both scatter and scintillation stem from gradients in refractive index at a wide range of scale sizes, the relevant refractive index gradients here being due to electron-density irregularities over scales of meters (and less) to ~10s km. There exist good databases on irregularity scale size spectra, and scintillation index vs.

frequency, for a wide range of locations, time of day and season, and phase of solar cycle, e.g., Basu et al. (1988). Fades in amplitude and phase rotation rates are strong enough to lose lock on GPS signals.

Scintillation produced by both patches and arcs, which have relatively steep spatial boundaries, tends to switch on/off as a patch or arc moves across the line of sight from a receiver to a satellite rf beacon. Spatial diversity is the best mitigation (other than very large fade margins), and progress on patch characterization, location tracking, now-cast, and prediction can considerably improve system reliability. The concept, in a real-time patch-tracking scenario, is to anticipate and switch to another satellite with a free line-of-sight. Switching from one satellite to another, as different patches drift through different lines-of-sight, can significantly improve continuity of signal. Likewise for arcs. Even for "blind" operations, knowledge of the statistics on typical scale sizes and velocities of patch (and arc) envelopes, frequency of occurrence, and spatial separations, can guide message/backup strategies. Readily available real time IMF monitoring can refine this.

At HF, even at times of season when clutter seems impossible to deal with, novel ways to circumvent the issue can emerge from an understanding of patch/arc morphology and size/location of the auroral oval with time. If uninterrupted HF communication is a priority, a third station, adequately far equatorward of the auroral oval, can circumvent polar patch HF scatter problems, even while direct point-to-point communication may be impossible. The solution could be propagation via two legs of a triangle. Polar transmission to a transceiver/repeater, farther equatorward than half the distance to the far edge of the auroral oval, could give reflection off a smooth ionosphere to a third station which retransmits to the intended polar receiver. Patches and arcs can also add an HF communication path, provided the patch or arc location is adequately known or forecast. Patches and arcs move in well-defined directions with nominal steady velocities. Sun-aligned arcs can provide a sun-aligned slowly-moving reflecting wall, with persistence up to order an hour.

1.3.7.4 Patches: Theory and Modeling Context

To model patch behavior, one needs to be concerned with the source reservoir of plasma for patches; a mechanism for patch creation; the transpolar trajectory of patch path after entry into polar cap; the structuring of patch from creation to exit; the exit process for the patch; and its return sunward flow, with the possibility of recycling.

First though, geomagnetic vs. geographic coordinates are essential to patch morphology. The frequency of patch occurrence in the polar cap has been studied from ground stations measuring F-region peak electron density and total electron content, and by satellites passing through the topside F region. Rodger et al. (1994) give a good summary of the characteristics of patches. Many ground-based statistics have been published (Buchau and Reinish, 1991; McEwen and Harris, 1996; Dandekar and Bullett, 1999). Figure 1-35 shows results of the comprehensive study by Coley and Heelis (1998), using DE in-situ satellite at about 840 km. The occurrence frequency peaks during winter, around UT ~ 12 to 24. Patch occurrence is suppressed in summer because the daytime sunlit polar-cap electron densities are already large, and difficult to double by transporting plasma from elsewhere. The UT dependence is due to the magnetic pole displacement (about 10 degrees in the northern hemisphere) from the earth's axis of rotation. The plasma flow lines in Figure 1-34 are in the geomagnetic latitude frame of reference, reflecting transpolar plasma transport bringing plasma into the polar cap through the cusp region, near 75 degrees magnetic latitude near noon. The reservoir of high-density sunlit plasma is in geographic coordinates, reflecting solar EUV plasma production. The magnetic pole (and cusp entry region) then dip 10 degrees more deeply into lower-latitude sunlit plasma about Canada, and 10 degrees farther from sunlit plasma about Siberia. For favored UTs, the magnetic entry region tips more deeply into lower-latitude sunlit high density plasma; for unfavored UTs, it draws from dark low-density plasma. This difference largely explains the northern hemisphere UT dependence of patch occurrence. The magnetic pole offset is not symmetric north to south pole, nor is the patch frequency of occurrence. As the auroral oval expands (or contracts) sunlit plasma will be drawn upon for a longer (or shorter) fraction of the day.

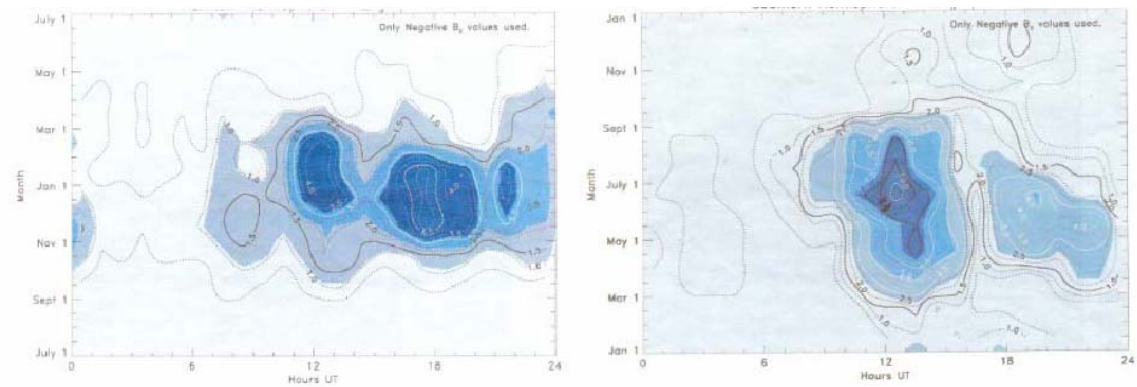

Figure 1-35: LH Side, Contour Plot of Occurrence Frequency of Polar Patches as Measured by DE 2 in the Northern Hemisphere vs. UT and Month of the Year, when the z Component of the IMF is Negative. The units on the contour lines are the number of patches observed per hour of spacecraft observation above 75 ILAT (from Coley and Heelis, 1998; Basu and Valladares, 1999.) RH side same but for southern hemisphere, showing strong hemispheric asymmetry.

Source Reservoir: Solar EUV-produced plasma density in this subauroral reservoir has well-documented seasonal and solar cycle variations, which reflect in variations in peak densities drawn into the polar cap. At times, events significantly enhance the plasma densities in the reservoir, and in its path between its region of solar EUV production and the cusp. These events can have very dramatic temporal/longitudinal consequences for the reservoir (Foster et al., 2005) that trace all the way across the polar cap. In the last decade or so, major progress has been made in understanding such processes, in the observing networks to track them, and in modeling them. Afternoon to pre-midnight electron densities at the equatorward edge of the trough can become significantly enhanced early in magnetic storms, a phenomena referred to as storm-enhanced density (SED) (Foster, 1993). This enhanced plasma is then transported towards noon by electric fields referred to as subauroral polarization streams (SAPS). This westward-flowing plasma, which can have major density increase through downward drainage of plasmaspheric plasma into the ionosphere, is not only of major interest to global ionospheric behavior, but can dramatically increase patch densities (Foster et al., 2005). We distinguish here between the plasma reservoir, and the mechanism that draws upon this reservoir to form patches, by pulling them into the polar cap in segmented islands. The physics of the subauroral plasma reservoir is treated in more detail elsewhere in this document.

When the magnetic pole tips towards the sun, so the magnetic cusp entry region continuously taps high-density plasma, one does not get a continuous smooth tongue of high-density plasma crossing the polar cap. To the contrary, the plasma is segmented, and the segments are severely structured, where the segmentation process is likely critical to the structuring. What leads to this segmentation?

Mechanisms for Patch Production: While there are no doubt many ways in which polar-cap patches can be formed, and at least six mechanisms have been published, there is no agreement as to which one(s) are most important or dominate (Lockwood and Carlson, 1992; Sojka et al., 1993; Rodger et al., 1994; Valladares et al., 1994, 1998).

Mechanisms proposed include, discrete changes in IMF B_y, B_z, and solar wind density/speed/pressure; time variation of average-flow-pattern models (e.g., Anderson et al., 1988); plasma flow jets channels in which relative-velocity-dependent recombination rates cut continuous tongues of plasma into segments and IMF reversals (Valladares et al., 1998); plasma production by cusp particle precipitation (Rodger et al., 1994, Millward et al., 1999); and transient magnetopause reconnection (Lockwood and Carlson, 1992), followed by Alfven wave coupling (Prikryl et al, 1999) of the solar wind to the ionosphere. The problem is not how patches might be produced, because a discrete change in many parameters can do that. The real question needs to be, how are most observed patches produced?

Cusp particle precipitation can produce patches up to ~ 5MHz f_oF_2, but models show it cannot explain the common 9 MHz patches, which produce severe scintillation of most concern. The increased recombination in high-speed flow channels can, in principle, segment tongues of high-density plasma to produce patches. However, this mechanism does not work, at least in the European sector, where high-speed flow channels produce higher plasma densities, not lower. The chemistry applies only where neutral densities are high enough to give significant recombination; it applies to plasma near and below 250 km, not much above. Anderson et al. (1988) have modeled patches by switching between different average large-scale flow patterns, but most patches appear to derive from transient events. Finding the dominant patch production mechanism, though challenging, is of more than academic interest. The production mechanism defines signatures for early identification, early tracking of direction, and even where in the polar cap scintillation-producing irregularities set in (Moen et al., 2001).

Recently one mechanism, the transient magnetopause reconnection mechanism, has been proven to work, to create patches, and furthermore to be a dominant patch-generation mechanism. Carlson et al. (2006), using the Svalbard EISCAT Radar (ESR) with supporting ASIP data, showed detailed agreement between model prediction and observations during an extended series of patches within the polar cap, observed coming directly from a subauroral plasma reservoir. The mechanism is a transient poleward excursion of the boundary between co-rotating and poleward moving plasma. The boundary then relaxes poleward. Observations confirmed the theoretical prediction of detailed coincidence in time and space of six plasma parameter boundaries, including optical arc boundaries that established motion of the patch boundary at formation speeds of ~2km/s. These observations establish clear patch onset signatures in parameters ideal for remote detection and early warning of patch onset.

Plasma Structuring and Scintillation: Keskinen and Ossakow (1983) have developed a well-tested theory for high-latitude ionospheric F-region irregularities (Tsunoda, 1988). This theory has led to the common acceptance of gradient drift instability as the mechanism that produces scintillation in mature patches. A further significant advance of experiment and theory has been the differentiation of this gradient drift instability and the velocity shear instability, which produces scintillation in sun-aligned arcs (Basu et al., 1998). These two plasma instabilities, which are triggered under very different conditions, were separated by in situ satellite measurements showing that the fractional fluctuation in electron density varied with fractional fluctuation in associated electric field, with tenfold more sensitivity in the velocity shear than in the gradient instability. Further work remains in theory and modeling of capacitive coupling to the magnetosphere above, coupling to variable conductivity below, and onset times. Also, irregularities in patches, attributed to the gradient drift instability, are observed earlier in the polar cap than is easily explained. While many patch-generating mechanisms leave this dilemma unresolved, the transient magnetopause mechanism for patch production would explain production of a patch filled with scintillation-producing irregularities by the process of patch creation.

Transpolar Trajectory, Exit, and Return Flow: Entry to the polar cap will differ significantly in initial velocity and plasma stress from one production mechanism to another. For example, magnetopause reconnection has initial >~2 km/s strong east or west B_y -dependent motion, compared to slower entry and weaker east-west motion for some other mechanisms. Once inside the polar cap, though, patches are assumed to follow well-established IMF-dependent transpolar flow paths (Heelis, 1984). Real time data can significantly improve accuracy and detail of large scale plasma flow (Ruohoniemi, 1993). Patch entry and exit have been modeled using time-varying patterns of flow, for flow patterns derived from averaged flow. Whether the exit of plasma through the nightside aurora is smooth or transient has been questioned.

“Boundary blobs” is the name given to enhanced ionospheric F-region plasma densities that are observed moving sunward near the equatorward edge of the auroral oval, primarily in the late afternoon and evening hours. These so-called blobs also represent walls of irregularities that produce scatter and scintillation. They were presumed to be patches that had exited from the midnight sector of the auroral oval. This assumption has not been confirmed, leaving their origin unproven. Their assumed exit from the polar cap, distortion of

shape, and return sunward flow has been tested once by Anderson et al. (1996). That work successfully models one representative patch example and identifies remaining issues for other scenarios.

1.3.7.5 Sun-Aligned Arcs and Theta Aurora: Theory and Modeling Context

The morphology of sun-aligned arcs has been well described by Valladares et al. (1994). That work describes frequency and location of IMF-dependent occurrence, orientation, motion, and other morphological properties. Much has also been learned about the near-earth electrodynamics, thermal structure, and energetics of these arcs (Valladares and Carlson, 1991). Arcs and patches have been observed in the same field of view after an IMF reversal, but only one from south to north, never from north to south, consistent with theoretical expectation that patches had been formed within the hour before the IMF reversal and had not had time to exit before the rapidly-forming arcs entered the field of view. Morphology of Theta aurora has been well described by Frank et al. (1986), within the context of associated particle energy distributions, currents, and the full suite of in situ measurements on board the DE satellite. Many characteristics mimicked those of sun-aligned arcs, but all were less dramatic. The number of sun-aligned arc events decreases with increasing intensity, and an intensity- vs. frequency-of-occurrence plot would put the intensity of Theta aurora at about their observed frequency of occurrence. Therefore, morphology alone leads us to ask, are they fundamentally the same, or fundamentally different? Distinctions will be discussed later below, between common, weak sun-aligned arcs and rare, intense Theta aurora, within the context of modeling and theory. Figure 1-36 shows representative sun-aligned arcs and Theta aurora.

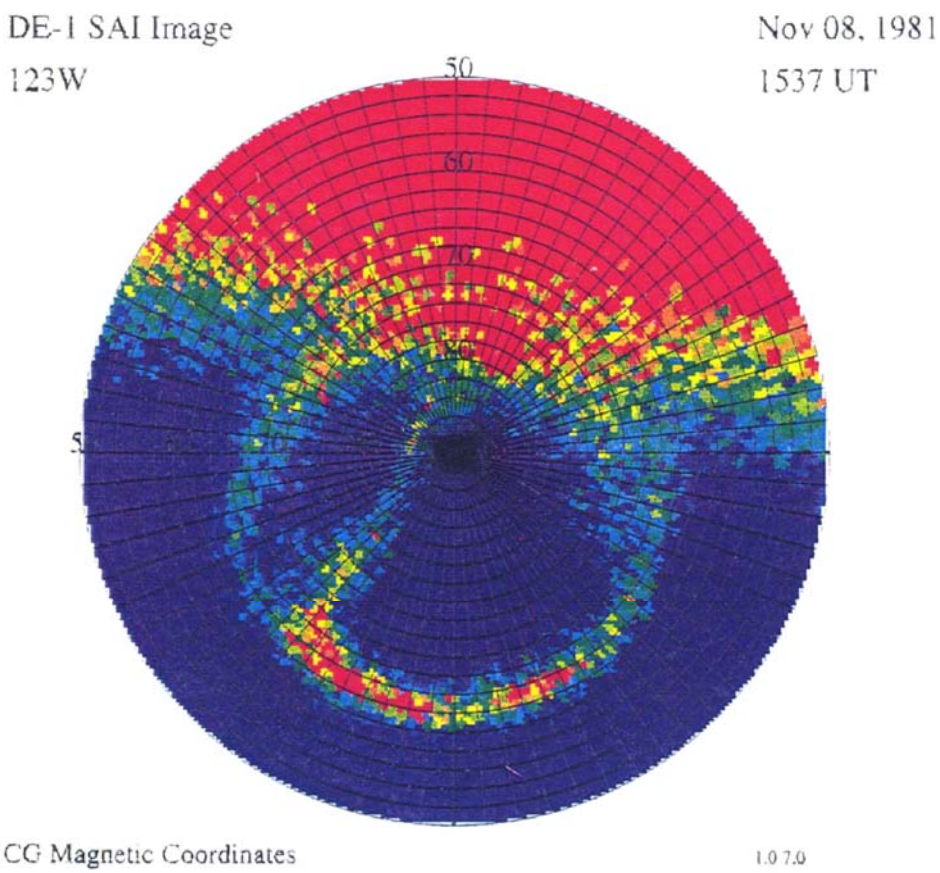

Figure 1-36a: A Theta Aurora Observed by the DE Satellite, Spanning the Polar Cap from the Midnight to the Dayside Auroral Oval, or Cusp, Near the Earth Limb (Frank et al., 1986). Theta aurora are seen a few percent of the time, and only after extended northward IMF.

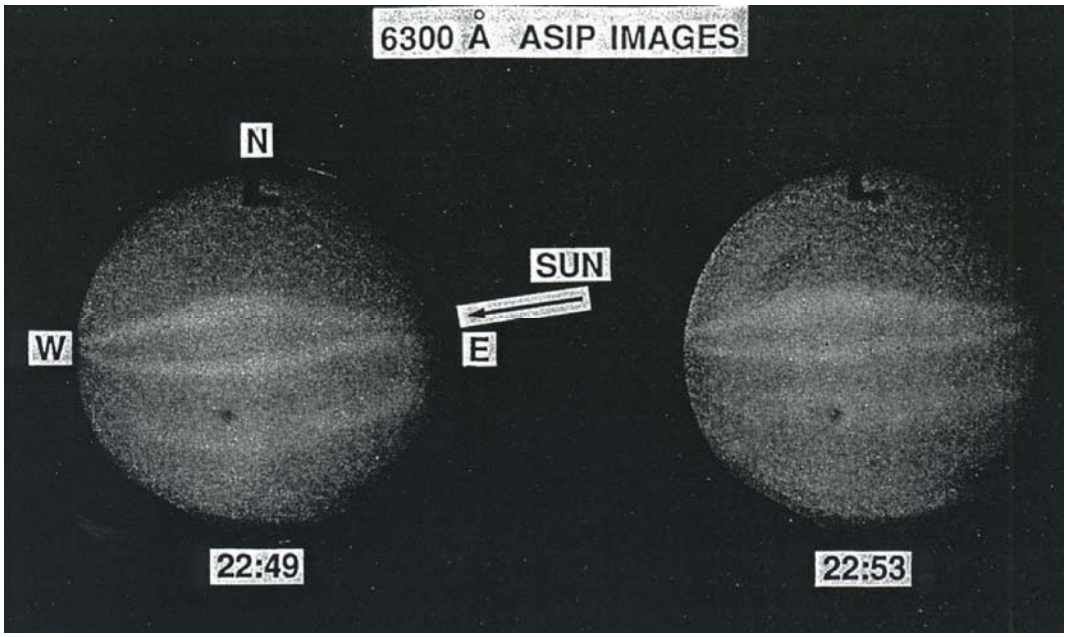

Figure 1-36b: Sun-Aligned Arcs in 1000 km Field of View are Present Nearly Half the Time, that Half when the IMF is Northward. The black dot is a fixed point on the ground, with scintillation-free field of view overhead at 22:40 UT and strong scintillation at 22:53 UT. Note the difference in both field of view, and spatial resolution, between this ground-based SIP and satellite-borne platforms (Valladares et al. 1994).

Outstanding theoretical and modeling issues have been the mapping of boundaries between the ionosphere and magnetosphere, the direction of drift and symmetry/asymmetry between northern and southern hemisphere (Obara et al. 1988), particle sources for current sheets, the fundamental mechanism for formation, and what, if any, is the fundamental difference between sun-aligned arcs and Theta aurora. Resolution of these questions is key to moving on to a predictive modeling capability and quantitative understanding.

1.3.7.5.1 Mapping to the Magnetosphere

Studies of data from satellite-borne particle detectors, sometimes including coincident optical imagery from satellites or the ground, have led to seemingly contradictory interpretations or conclusions about where polar-cap arcs map to in the magnetosphere. Examples are the 12 "cartoon" models shown in the figures described in section 6.2.3 of the book by Sandholt et al., 2002. The sun-aligned arcs have been either directly observed with coincident images, or inferred to have been present on the basis of low-earth-orbit satellite particle measurements. The different magnetospheric regions to which sun-aligned arcs have thus been proposed to connect include, closed field lines threading an expanded (e.g., Meng, 1981) or bifurcated (e.g., Frank et al., 1986) plasma sheet and associated with sunward flow; open field lines linking the mantle and magnetosheath (e.g., Hardy et al., 1982), or closed field lines mapping to the lower-latitude boundary layer (LLBL) and associated with antisunward flow (e.g., Lundin et al., 1991). Weiss et al., (1993), have schematically mapped from the high-latitude ionosphere to the distant magnetotail for five models of the polar-cap morphology, namely open, bifurcated, expanded plasma sheet, expanded LLBL, and rotated X line. There is no consensus on the mapping to the magnetosphere.

1.3.7.5.2 Drift

The sun-aligned arcs are observed to drift sideways, at ~0.1 km s^{-1}. The direction of dawn-dusk drift of the sun-aligned arcs, duskward (dawnward) for $B_y > 0$ ($B_y < 0$), has been suggested as being due to dayside

merging (Valladares et al., 1994; Newell and Meng, 1995). Many models are based on Crooker's antiparallel merging hypothesis (Crooker, 1979). Reconnection topology presented by Reiff and Burch (1985) can drive convection cells by merging in the cusps, where, if the magnetotail field lines are open, they recover the single-lobe cell of Crooker (1979) with B_y-dependent sense of rotation.

1.3.7.5.3 Currents

The near-earth ionospheric behavior is much better understood. The generation of large-scale regions of auroral currents and particle precipitation by the divergence of convection electric fields was treated in detail by, and explicitly applied to the mesoscale electrodynamics of observed transpolar and sun-aligned arc segments by Carlson et al. (1988). The essential electrodynamic, thermal, and energetic properties of sun-aligned arcs at ionospheric altitudes were calculated by Valladares and Carlson (1991). Sun-aligned arcs exist within plasma velocity gradients of a sense which produces converging electric fields. In the northern polar cap, antisunward plasma flow velocity decreases across the arc in the dawn to dusk direction, with this gradient sometimes observed to extend into sunward flow on the dusk edge. The particle precipitation and consequent optical emission ends where the gradient reverses or ends, as for an Ohm's law arc. The arcs are called sun-aligned arcs because the flow lines are strongly anisotropic, with scale sizes >1000 km along the sun-earth line, while only approximately tens to hundreds of kilometers in the dawn-dusk direction.

1.3.7.5.4 Formation

Theta auroras are sun-aligned arcs which have intense emissions, well exceeding a kR in 630.0 nm emission, the electrons exciting them having relatively high energy, exceeding a keV. They connect fully from the day to nightside auroral oval, and are seen to drift dawn to dusk as do weaker, common, quickly-forming sun-aligned arcs. Theta aurora are seen only after many hours of uninterrupted northward IMF, and are best, if not only, seen from satellites. They are rare, being seen only a few percent of the time, but they are dramatic and have attracted considerable attention in the literature since first noted (Frank et al., 1986). These authors, on the basis of the event they reported on, argued that the theta aurora was located on closed field lines. They argued for a polar cap consisting of two open cells bifurcated by a closed region mapping to the plasma sheet boundary layer or the distant plasma sheet.

For a recent sketch of the large number (nine) of models devoted to explaining the formation of Theta aurora, their conflicting predicted behavior, and consistency with data (e.g. Carlson and Cowley, 2005).

For common sun-aligned arcs, Carlson and Cowley (2005) applied a kinetic theory perspective of electrons precipitating into the polar cap from the electron reservoir in the distant magnetospheric tail, for currents sustainable without electron acceleration. They concluded that for typical sun-aligned arc conditions, there is no need to look further than the magnetospheric tail for the source reservoir of electrons needed to provide observed fluxes in either polar rain or sun-aligned arcs. Applying this theory to this ready reservoir, they suggested that the sun-aligned arcs are indeed produced by field-aligned acceleration of polar rain electrons. The implied field-aligned voltages are ~200–600 V, with inferred minimum heights of the acceleration region of ~0.7–0.9 R_E above the polar ionosphere. Finding the likely source of the electrons drawn upon to produce sun-aligned arcs should be seen as a significant step toward the resolution of the long- question standing about the topology of the magnetosphere under northward IMF conditions.

Recent work (Milan et al., 2005 and references therein) has turned more serious effort towards sun-aligned arc formation, using concepts such as those of Cowley and Lockwood (1992) and Reiff and Burch (1985), at a high-latitude or lobe-reconnection site, as expected when the IMF is northward. It is expected that this approach will lead to new insights in the near future, regarding the ubiquitous sun-aligned arcs.

1.3.7.5.5 Contrasting Formation of Common Sun-Aligned vs. Theta Aurora

Interpretation of many observations of Theta aurora particle properties indicate that plasma sheet precipitation on closed field lines is the cause of Theta aurora. Theory and modeling support this conclusion. Theta aurora are the class of polar-cap sun-aligned arcs that are more than a kR in intensity (and thus bright enough to be visible to the naked eye), and which result after hours of northward IMF conditions, affording enough time for significant magnetospheric topological reconfiguration. The location of theta aurora on closed field lines suggests that the magnetospheric dynamics has resulted in a bifurcation (Frank et al., 1986) of the tail lobe, due, for example, to a reversal in B_y during the interval which takes time to develop. Overall, theta auroras are found in the polar cap a few percent of the time.

In sharp contrast, sun-aligned arcs are present in the polar cap nearly half the time, that half of the time that the IMF is northward. These arcs are sub-visual intensity, typically hundreds of R. They have received less attention in the literature, perhaps because they are difficult to observe from satellites, and are less dramatic in commonly available images. However, they define the dominant character of the northward IMF polar cap, including its highly anisotropic shear flow character (Carlson et al., 1988). The ubiquitous sub-visual sun-aligned arcs are consistent with accelerated polar rain on open field lines. These arcs can be generated by any modest mechanism that drives a flow shear across the open field lines. Such mechanisms include, for example, the boundaries between lobe cells and merging cells (open flux production does not switch off completely until the IMF is quite close to pure north, perhaps within ~30–40 degrees), or flows from differing magnetopause reconnection sites. Such conditions can easily occur. Any such process giving rise to flow shears on open field lines, leading to upward field-aligned currents bigger than 0.01 mA m^{-2}, will require acceleration, and will produce polar shower precipitation and a corresponding aurora.

In short, present theory indicates that the rare Theta aurora are on closed field lines, and the ubiquitous sun-aligned arcs are on open field lines.

1.3.7.6 Future Work

Our understanding today of polar-cap phenomena includes a detailed characterization, much morphology, and definition of many of the processes determining behavior. Consequently, we have now developed valuable techniques to mitigate the effects of disruptive polar phenomena on high-latitude communication and GPS navigation systems. However, only a physics-based model, tested and validated, can deliver a predictive capability that will allow uninterrupted connectivity where needed. Much work remains to achieve this.

We do not have agreed-upon physical models of the entry, transit, exit, or return flow of the patches characteristic of southward IMF conditions. Neither do we have agreed-upon physical models for rare Theta aurora or the sun-aligned arcs characteristic of northward IMF conditions. We have much to discover about the processes that couple the polar ionosphere to the magnetosphere, and to the solar wind looking outward towards far space, and about how polar-cap coupling impacts the dynamics of the global ionosphere, our near-space earth environment. The mechanism for the creation of polar-cap patches has been debated for 20 years now, without a consensus. The one validated mechanism has limited early-warning time, and we are unclear on the significance of other candidate mechanisms. Accurate prediction of patch trajectory across the polar cap is key to the timing and location of downstream outages of 10 minutes or more. Yet transpolar models are statistical, and their application to patches exiting through the nightside aurora and returning with a westward flow is questionable. Conflicting models exist for Theta aurora and sun-aligned arc formation, with conflicting predicted behavior. We must even ask ourselves, to what degree is polar plasma large-scale flow dominated by large-scale fields rather than being the sum of many small-scale transient reconnection events.

Recent advances in experimental and instrumental capability, and computer modeling, make this a time ripe for a surge of progress. Recent ground-based measurements finally have the temporal and spatial

resolution to diagnose the evolution of key transient events, rather than only seeing before-and-after states. Satellite constellations can now sample fields and particles at sites of critical phenomena, with the multipoint ability to separate temporal from spatial structure. Computer power now dissects complex interactive processes, and focuses instrument development and experiment deployment on the most critical challenges. It is not unreasonable to suggest that the next decade might even reveal the nature of magnetic reconnection itself, a discovery of fundamental importance to plasmas from laboratory to galactic scales. Strong dialogue between the radio and space science communities will compound the gains.

1.3.7.7 Caveat

There is a great amount of on-going progress, from the work of a large and vigorous research community, in the area of observing, modeling, and understanding the character and space physics of the polar-cap ionosphere. In a sketch as short as this, it is inevitable that much excellent and important work could not be covered, in some cases not even noted. The intent is to present to radio professionals concerned with phenomena and effects of the polar-cap ionosphere, a framework, and enough information to pursue those areas of most concern to them.

1.3.7.8 Acknowledgements

The author wishes to acknowledge support from AFOSR under contract number FA9550-06-1-0217.

1.3.7.9 References

Anderson, D. N., J. Buchau, and R. A. Heelis, Origin of density enhancements in the winter polar cap ionosphere, *Radio Sci., 23*, 513-519, 1988.

Anderson, D. N., D. T. Decker and C. E. Valladares, Modeling Boundary Blobs using time varying convection, *Geophys. Res. Lett., 23*, 579-582, 1996.

Basu, S., E. MacKenzie, and Sunanda Basu, Ionospheric constraints on VHF/UHF communications links during solar maximum and minimum periods, *Radio Sci., 23*, 363-378, 1988.

Basu, S. and C. E. Valladares, Global aspects of plasma structures, *J. Atmos. Solar-Terr. Phys., 61*, 127, 1999.

Basu, S., S. Basu,, E. MacKenzie, P. F. Fougere, W. R. Coley, N. C. Maynard, J. D. Winningham, M. Sugiura, W. B. Hanson and W. R. Hoegy, Simultaneous density and electric field fluctuation spectra associated with velocity shears in the auroral oval, *J. Geophys. Res*. *93*, 115-136, 1998.

Buchau J., et al., Ionospheric structures in the polar cap: Their origin and relation to 250-MHz scintillation, *Radio Sci*., *20*, 325-338, 1985.

Buchau, J., and B. W. Reinish, Electron density structures in the polar F-region, *Adv. Space Res*., *11* (10), 29, 1991.

Carlson, H. C., R. A. Heelis, E. J. Weber, and J. R. Sharber, coherent mesoscale convection patterns during northward IMF, *J. Geophys. Res*., *93*, 14,501-14,514, 1988.

Carlson, H.C., The dark polar ionosphere: Progress and future challenges, *Radio Sci*. *29*, 157-165, 1994.

Carlson, H. C., J. Moen, K. Oksavik, C. Nielsen, I. McCrea, T. Pedersen, P. Gallop, Direct observations of injection events of subauroral plasma into the polar cap, *Geophys. Res. Lett, 33*. L05103, 2006.

Carlson, H. C. and S. W. H. Cowley, Accelerated polar rain electrons as the source of sun-aligned arcs in the polar cap during northward IMF conditions, *J. Geophys. Res., 110*, A05302 (doi: 10.1029/ 2004JA010669), 2005.

Coley, W. R. and R. A. Heelis, Structure and occurrence of polar ionization patches, *J. Geophys. Res., 103*, 2201, 1998.

Cowley, S. W. H., and M. Lockwood, Excitation and decay of solar wind-driven flows in the magnetosphere-ionosphere system, *Ann. Geophys*, *10*, 103, 1992.

Crooker, N.R., Dayside merging and cusp geometry, *J. Geophys. Res., 84*, 951, 1979.

Crowley, G., A critical review of ionospheric patches and blobs, *Rev of Radio Science 1993-96*, Oxford Univ. Press, 619-648, 1996.

Dandekar, B. S., and T. W. Bullett, Morphology of polar-cap patch activity, *Radio Sci., 34*, 1187, 1999.

Foster, J. C., Storm time plasma transport at middle and high latitudes, *J. Geophys. Res., 98*, 1675, 1993.

Foster, J. C., et al, Multiradar observations of the polar tongue of ionization, *J. Geophys. Res., 110*, A09S31, 2005.

Frank, L., et al., The theta aurora, *J. Geophys. Res*., *91*, 3177, 1986.

Gussenhoven, M. S., Extremely high latitude auroras, *J. Geophys. Res*., *87*, 2401, 1982.

Hardy, D. A., W. J. Burke, and M.S. Gussenhoven, DMSP optical and electron measurements in the vicinity of polar-cap arcs, *J. Geophys. Res., 87*, 2413, 1982.

Heelis, R.A., The effects of interplanetary magnetic field orientation on dayside high-latitude convection, *J. Geophys. Res*. *89*, 2873, 1984.

Ismail, S., D. D. Wallis, and L. L. Cogger, Characteristics of polar-cap Sun-aligned arcs, *J. Geophys. Res*. *82*, 4741, 1977.

Keskinen, M. J., and S. I.. Ossakow, Theories of high latitude ionospheric irregularities, *Radio Sci*, *18*, 1077, 1983.

Lockwood, M. and H. C. Carlson, Production of polar cap electron density patches by transient magnetopause reconnection, *Geophys. Res. Lett*., *19*, 1731, 1992.

Lundin, R., L. Eliasson, and J. S. Murphree, The quiet-time aurora and the magnetosphere configuration, in Auroral Physics, edited by C.-I. Meng, M. J. Rycroft, and L.A. Frank, p. 177, Cambridge Univ. Press, New York, 1991.

McEwen, D. J., and D. P. Harris, Occurrence patterns of F region layer patches over the north magnetic pole, *Radio Sci., 31*, 619-628, 1996.

Meng, C.-I., Polar-cap arcs and the plasma sheet, *Geophys. Res. Lett.*, *8*, 273, 1981.

Milan, S. E., B. Hubert and A. Grocott, Formation and motion of a transpolar arc in response to dayside and nightside reconnection, *J. Geophys. Res., 110*, A01212, 2005.

Millward, C. H., et al., Modeling the ionospheric effects of ion and electron precipitation in the cusp, *J. Geophys. Res., 104*, 24603. 1999.

Moen et al, On the collocation between dayside auroral activity and HF Radar backscatter, *Ann. Geophys., 18*, 1531-1549, 2001.

Newell, P.T., and C.-I Meng, Creation of theta-auroras: The isolation of plasma sheet fragments in the polar cap, *Science, 270*, 1338, 1995.

Obara, T., M. Kitayama, T. Mukai, N. Kaya, J. S. Murphree and L. L. Cogger, Simultaneous observations of Sun-aligned polar-cap arcs in both hemispheres by EXOS-C and Viking, *Geophys Res. Lett., 15*, 713, 1988.

Prikryl, P., J. W. MacDougal, I. F. Grant, D. P. Steel, G. J. Sofko, and R. A. Greenwald, Observations of polar patches generated by solar wind Alfven wave coupling to the Datside magnetosphere, *Ann. Geophysicae, 17*, 463-489, 1999.

Reiff, P.H. and J. L. Burch, IMF B_y-dependent on dayside plasma flow and Birkeland currents in the dayside magnetosphere 2: A global model for northward and southward IMF, *J. Geophys. Res., 90.* 1595, 1985.

Rodger, A. S., et al., A new mechanism for polar patch formation, *J. Geophys. Res., 99*, 6425, 1994.

Ruohoniemi, J. M. and K. B. Baker, Large-scale imaging of high latitude convection with Super Dual Auroral Radar Network HF radar observations, *J. Geophys. Res.*, *103*, 20797, 1993.

Sandholt. P. E., H. C. Carlson, and E. Egeland, *Dayside and Polar-cap Aurora*, Kluwer Academic Publishers, Dordrecht Netherlands, 2002.

Schunk, R. W., and J. J. Sojka, Ionosphere-thermosphere space weather issues, *J. Atmos. Terr. Phys*., *58*, 1527, 1996.

Smith, M. F., and M. Lockwood, Earth's Magnetospheric Cusps, *Rev. Geophys*., *34*, 233-260, 1996.

Sojka, J. J., et al., Modeling polar cap F-region patches using time varying convection, *Geophys. Res. Lett*., *20*, 1783, 1993.

Sojka, J. J., M. D. Bowline and R. W. Schunk, Patches in the polar-cap ionosphere: UT and seasonal dependence, *J. Geophys. Res., 99*, 959, 1994.

Tsunoda, R. T., High-latitude F region irregularities: a review and synthesis, *Rev. Geophys 26*, 719, 1988.

Valladares, C.E. and H, C, Carlson, The electrodynamic thermal and energetic character of intense sun-aligned arcs in the polar cap, *J. Geophys. Res.* 96, 1379-1400, 1991.

Valladares, C. E., H. C. Carlson and K. Fukui, Interplanetary magnetic field dependency of stable sun-aligned polar-cap arcs, *J. Geophys. Res., 99*, 6247-6272, 1994.

Valladares, C. E., et al., Formation of polar cap patches associated with north-to-south transitions of the IMF, *J. Geophys. Res.*, *103*, 14,657, 1998.

Vickrey, J. F., C. L. Rino, and T. A. Potemra, Chatanika, TRIAD observations of unstable ionization enhancements in the auroral F-region, *Geophys. Res. Lett. 7*, 789-792, 1988.

Weber, E.J. and J. Buchau, Polar-cap F-layer auroras, *Geophys. Res Lett. 8*, 125, 1981.

Weber et al., F-layer ionization patches in the polar cap, *J. Geophys. Res., 89*, 1683, 1984.

Weber, E.J., J. A. Klobuchar, J. Buchau, H. C. Carlson, R. C. Livingston, O. de la Beaujardiere, M. McCready, J. G. Moore and G. J. Bishop, Polar cap it F layer patches: structure and dynamics, *J. Geophys. Res. 91*, 121, 1986.

Weiss, L.A., E. J. Weber, P.H. Reiff, J. L. Sharber, J. D. Winningham, F. Primdahl, I. S. Mikkelsen, C. Seifring, and E. M. Wescott, Convection and electrodynamic signatures in the vicinity of a Sun-aligned arc: Results from the polar acceleration regions are convection study (polar arcs), in *Auroral Plasma Dynamics, Geophys. Mongr. Ser.*, *80* edited by R. L. Lysak, p. 69, AGU, Washington, DC, 1993.

Zhu, L., R. W. Schunk, and J. J. Sojka (1997), Polar-cap Arcs: A review, J. Atmos. Solar-Terr Phys., 59, 1087, 1997.

1.3.8 Traveling Ionospheric Disturbances (TIDs)

(Paul Prikryl, Communications Research Centre Canada)

1.3.8.1 Relationship between Atmospheric Gravity Waves and TIDs

Atmospheric gravity (buoyancy) waves (AGWs) with sources in the lower and upper atmosphere, particularly those originating in the auroral electrojets, produce traveling ionospheric disturbances (TIDs). As the gravity wave propagates in the thermosphere the ionosphere is structured into quasi-periodic plasma density enhancements and depletions moving with the horizontal phase velocity of the wave. The neutral motions due to the gravity wave are transferred to ions and electrons through collisions (ion drag). Because the collision rate depends on the altitude, the ionospheric response to gravity waves is altitude dependent. In the F-region, where the collision frequency is small compared to the gyrofrequency, the ion motion perpendicular to the magnetic field is inhibited. The ions move parallel to the geomagnetic field, which has a tilt from the vertical that increases with decreasing latitude. Although the effect of a passing gravity wave is manifested in the ionosphere as a TID, the observed F-region motions and densities of ions, which are coupled to electrons, by no means precisely represent the gravity-wave-induced neutral motions and densities. Nevertheless, some of the wave parameters, namely wavelength and period, can be inferred from observations of TIDs (Kirchengast, 1996).

In general, AGW/TIDs are divided into two broad classes according to their phase velocity and period, namely large- and medium-scale waves (see, e.g., Hocke and Schlegel, 1996). The large-scale waves propagate in the thermosphere with horizontal phase velocities between 400 and 1000 m/s (comparable with the sound velocity in the thermosphere), have horizontal wavelengths greater than 1000 km and periods between about 1 and 3 h. The medium-scale waves, which may also propagate in the lower atmosphere before they are detected in the ionosphere, have horizontal phase velocities between 100 and 300 m/s (less than the sound velocity in the lower atmosphere), wavelengths of several hundred kilometers and periods between 15 and 60 min (Hunsucker, 1982).

1.3.8.2 Dispersion Relation and Ray Tracing of Gravity Waves

The theoretical basis for the understanding of AGWs and their role in the ionosphere was developed by Hines (1960). The gravity wave dispersion relation is consistent with both upward group (downward phase) and downward group (upward phase) propagation. Yeh and Liu (1974) reviewed propagation of AGWs in a realistic atmosphere using the ray theory approach and the WKB approximation. A simplified ray tracing procedure based on Snell's law is applicable for a horizontally stratified atmosphere. Assuming

a model atmosphere and a source of gravity waves in the auroral ionosphere, a simple two-dimensional ray tracing of the gravity waves can be conducted (Prikryl et al., 2005). Viscosity, thermal conductivity and neutral winds are neglected. Snell's law for a spherically stratified atmosphere with no horizontal gradients (Muldrew, 1959) is given by $nr \sin \alpha =$ constant where n is the refractive index, r is the distance from the centre of the earth to a point on the ray and α is the angle of incidence. The dispersion relation between the gravity wave frequency ω and wave vector **k** given by Hines (1960) is used:

$$(\omega^2 - \omega_a^2)\, \omega^2 / C^2 - \omega^2(k_x^2 + k_z^2) + \omega_b^2 k_x^2 = 0, \tag{1}$$

where $\omega_a = \gamma g/2C$ is the acoustic cutoff frequency, C denotes the speed of sound, and k_x and k_z are the horizontal and vertical components of the wave vector **k**. For a nonisothermal atmosphere (one in which the atmospheric scale height varies with altitude) the Brunt-Väisälä (buoyancy) frequency ω_b is defined as

$$\omega_b^2 = (\gamma\text{-}1)g^2/C^2 + (g/C^2)(dC^2/dz) \tag{2}$$

Figure 1-37 shows an example of group paths for gravity waves having a period of 30 min and originating at an altitude of 110 km, using the MSIS90 atmosphere model. The ray tracing is initiated at an altitude of 110 km using selected initial values α_0 of the **k** directions α, where α is taken to be positive upward from the southward horizontal direction in the northern hemisphere. The group velocity $\mathbf{V}_g \equiv (\partial\omega/\partial k_x, 0, \partial\omega/\partial k_z)$ or path of energy flow of the gravity wave is obtained from (1) numerically. The rays are the same in all panels but are colour-coded for (a) time of travel, (b) phase velocity V_x^{ph}, (c) horizontal wavelength λ_x and (d) vertical wavelength λ_z (Figure 1-37). For each ray, the wave vector directions are shown at the start of the ray tracing (α_0) and after one wave period. The ray tracing shows three distinct group paths that reach the F region, namely the rays that travel directly upward, rays that are reflected in the lower thermosphere, and rays that are reflected in the troposphere or from the earth's surface.

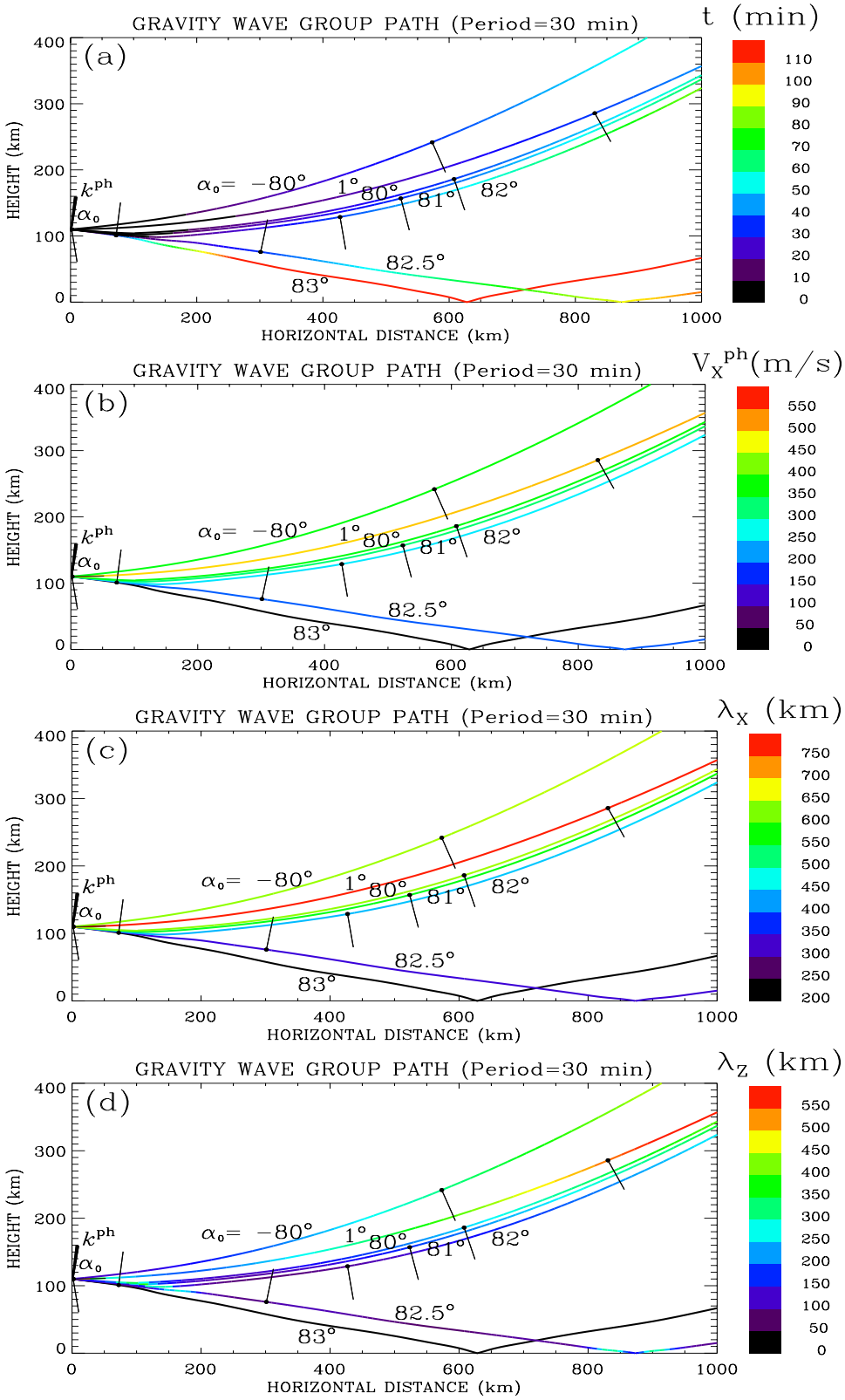

Figure 1-37: Ray Tracing of Gravity Waves (Group Path). When the phase propagation direction k^{ph} is upward ($\alpha_0 > 0$), the group velocity is downward and vice versa. The rays are the same in all panels, but are colour-coded for time of travel, horizontal wavelength and horizontal phase velocity (Prikryl et al., 2005).

1.3.8.3 Radio and Optical Techniques of AGW/TID Detection

AGW/TIDs can be observed using various techniques (see, e.g., Hunsucker, 1982; Davies, 1990; Hargreaves, 1992). In the ionosphere, the principal methods are based on radio techniques, namely ionosode, HF Doppler, HF radar (ground/sea scatter), incoherent scatter radar, total electron content (TEC) now routinely measured with GPS receivers and ionospheric tomography. Also, because the gravity waves modulate the night sky airglow (Mendillo et al., 1997), optical techniques provide a wealth of information on gravity waves.

a) HF Doppler

A very sensitive method for detecting transient variations in the vertical motion of ionospheric layers is the HF Doppler technique operating in the 3-10 MHz band (Georges, 1968; Davies, 1990). Networks of HF Doppler systems were used to observe bottom-side F-region undulations due to TID/AGW and derive periods and horizontal phase velocities (see references in Lewis et al., 1996). The advantages of this technique over other techniques of TID measurement are relatively high temporal resolution (30 sec) and lower power consumption than traditional ionosonde chains.

b) Ionosonde

Traditional ionosonde measurement of the virtual height displacement of the F region was traditionally used as indicator for large-scale TIDs. Changes in N_mF_2 and h_mF_2 can be detected. Extensive networks or chains of ionosondes have been deployed to study propagation of TIDs and to identify the sources of AGWs, particularly in the auroral electrojet, by measuring delays, phase velocities, wavelengths and amplitudes. Worldwide atmospheric gravity wave studies (WAGS) have been conducted in coordination with other instruments. Results from these studies that were summarized by Hocke and Schlegel (1996) provided a major improvement in understanding of TIDs. A modern digital ionosonde (see Chapter 2, and http://cadiweb.physics.uwo.ca) combines the capabilities of the HF Doppler and traditional ionosonde technique (Grant et al., 1995), thus allowing improved observations of TIDs (MacDougall et al., 1997).

c) HF Radar

Radar backscatter sounding techniques for observation of motions of large-scale F-region irregularities date back to 1950s (Tveten, 1961; and references therein). Hunsucker and Tveten (1967) used the HF backscatter technique to observe large scale TIDs on ground-scatter records. At the ionospheric reflection point the signals are modulated (focused/defocused) by TIDs which are attributed to imperfectly-ducted atmospheric gravity waves generated by auroral sources. More recently, HF radars are used to observe ionospherically generated medium-scale gravity waves, which may be earth-reflected and then produce trains of TIDs many hundreds of kilometers equatorward of the original source region (Samson et al. 1989; 1990). These radars are now a part of SuperDARN (Super Dual Auroral Radar Network) (Greenwald et al., 1995; Chisham et al., 2007). SuperDARN has become a useful instrument for the study of TIDs (Bristow et al., 1994; Hall et al., 1999) because of the large spatial coverage. The combined radar field-of-view covers a large fraction of the auroral and polar region, and most recently, new radars have been installed in mid latitudes. Because SuperDARN was primarily designed to monitor ionospheric convection (Greenwald et al., 1995), the SuperDARN radars can observe the sources of AGW/TIDs, namely convection intensifications at high latitudes (Prikryl et at., 2005).

Figure 1-38 shows a sequence of the Kapuskasing radar maps of the received power showing the ground scatter signatures of two successive medium-scale TIDs. The ground scatter band 2 spanning the radar field of view moved equatorward as the next band 3 appeared about 400 km poleward of it. A series of ionospheric convection bursts at high latitudes caused by pulsed magnetic reconnection at the dayside magnetopause was the source of TIDs. A band of ionospheric scatter at far ranges (near

80° magnetic latitude) is the TID source region where the ionospheric signature of the reconnection, namely a convection pulse that is sometimes called a flow channel, occurred.

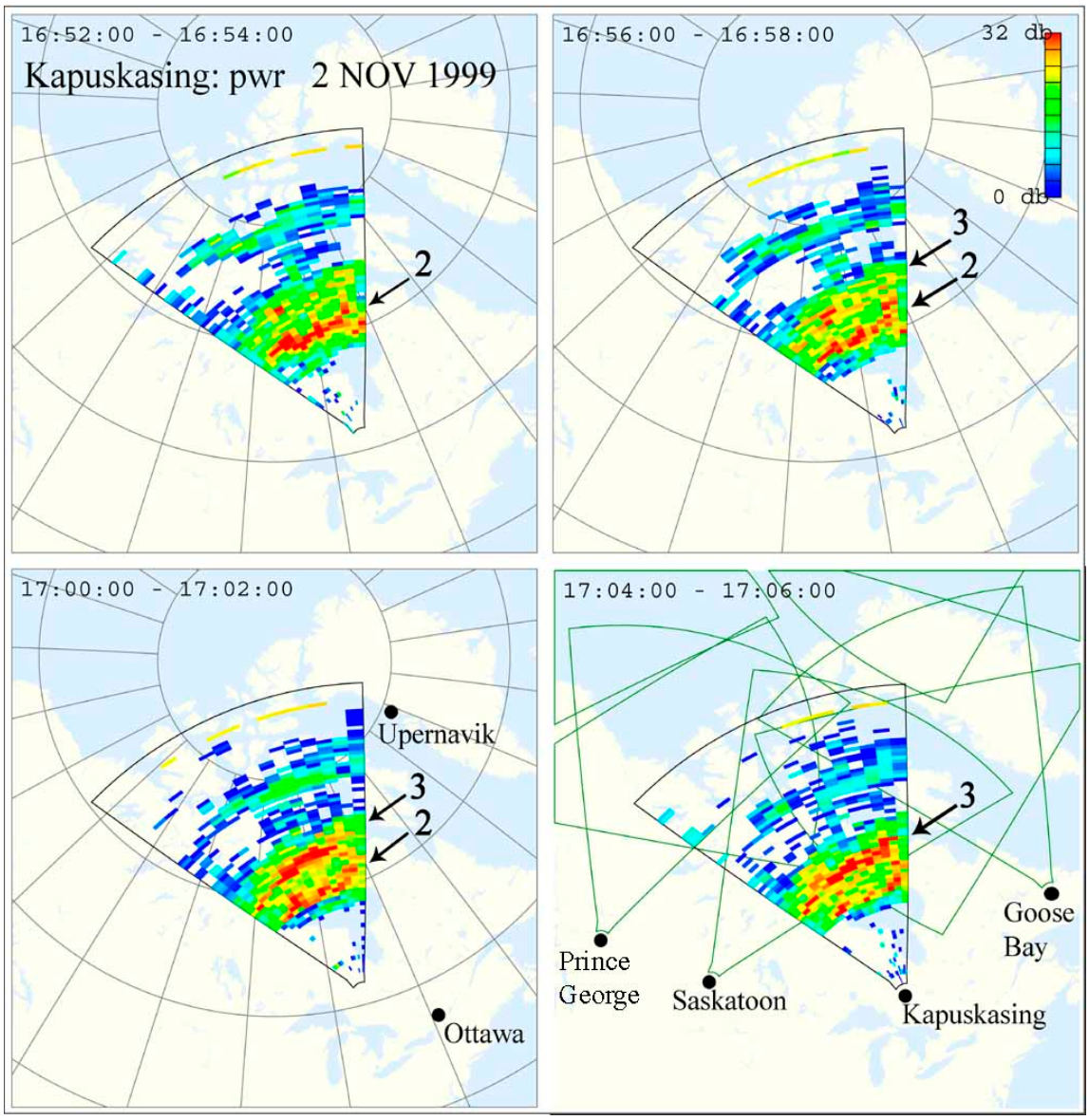

Figure 1-38: The Ground Scatter Signatures of Equatorward-Moving Medium-Scale TIDs in the Kapuskasing Radar Field-of-View.

Figure 1-39 shows the ionospheric convection/potential map constructed using the line-of-sight velocity data from all northern-hemisphere SuperDARN radars (Ruohoniemi and Baker, 1998). The fitted convection velocity vectors are shown colour-coded (see velocity scale bar top-left). The cross-polar-cap electric potential contour map is shown by solid and broken lines outlining the dusk and dawn convection cells, respectively. The value of the cross-polar-cap voltage is printed in the upper-left corner. Also, the IMF BZ versus BY clock-dial plot is shown (top-right). The convection map shows the ionospheric signature of a flux transfer event, namely an ionospheric flow channel extended over several hours of local time about 12 MLT.

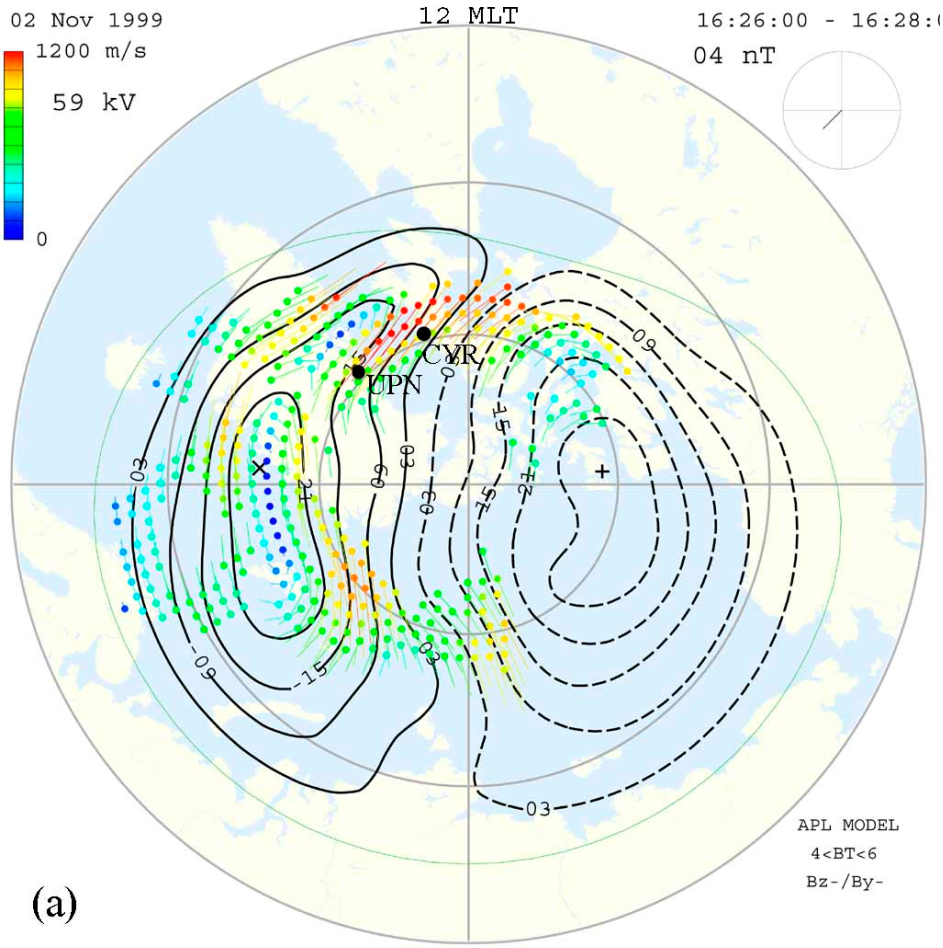

Figure 1-39: The Ionospheric Convection Intensification Centered at 12 MLT was the Source Region for the TID 2 shown in Figure 1-38.

d) Incoherent Scatter Radar

Simultaneous measurements by incoherent scatter radars of various ionospheric parameters, such as electron density, ion velocity, and ion/electron temperature, have been used to retrieve TID parameters by means of inversion techniques (Hocke and Schlegel, 1996). To elucidate the physics of the AGW-TID relationship, theoretical simulations were conducted and compared with incoherent scatter radar (ISR) data (Kirchengast 1996; Kirchengast et al., 1996). Based on such theoretical investigations AWG characteristics may be derived or inferred from TID modelling in conjunction with incoherent scatter radar measurements (Kirchengast et al., 1995).

Large-scale TIDs have been extensively imaged using ISR techniques and later compared with tomographic reconstruction (Nygrén et al., 1996) and GPS measurements of total electron content (e.g., Makela et al., 2000; Nicolls et al., 2004). Figure 1-40 shows an example of the Arecibo ISR observations of large-scale TIDs with a period of 2 to 3 hours (Nicolls et al., 2004) that were generated by a series of geomagnetic substorms and propagated equatorward at very high velocities. At least three oscillations of the F layer were observed over Arecibo (18.3°N) with the final one driving the F peak to over 450 km and leading to small-scale structures indicative of plasma instabilities at dawn. In the lower panel of Figure 1-40 (Nicolls et al., 2004; Figure 1), the F-region vertical ion velocities along with polynomial fits are plotted, showing evidence for AGW downward phase progression with a velocity of ~200–300 m/s (Nicolls et al., 2004).

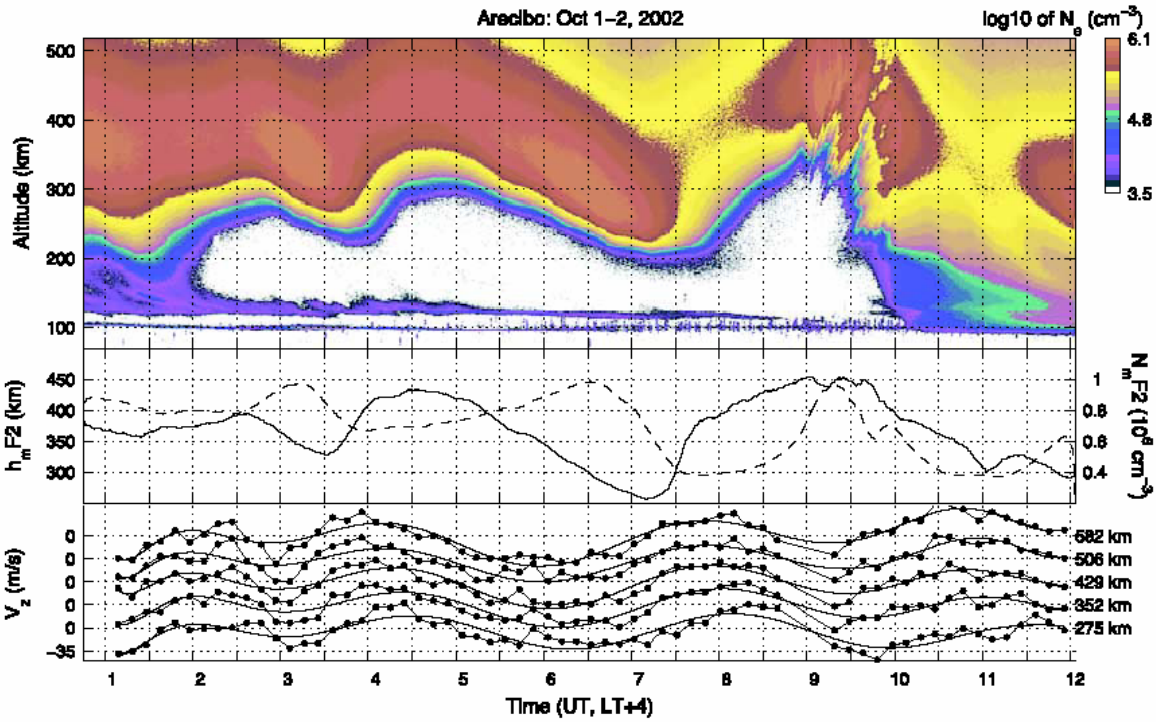

Figure 1-40: Summary of Arecibo ISR Data Showing Large-Scale TIDs for the Night of October 1–2, 2002. The top panel is $\log_{10}$ of electron density, the middle panel is h_mF_2 (solid) and N_mF_2 (dashed), smoothed over 12 minutes, and the bottom panel is vertical ion velocity (points) with polynomial fits (solid). The tick marks in the V_z panel are separated by 35 m/s (Nicolls et al., 2004).

e) Ionospheric Tomography

Radio tomography using satellite beacons was originally suggested by Austen et al. (1986). Several inversion techniques have been developed (e.g., Raymund et al., 1990, 1993 and 1994) and compared (Raymund, 1995). Further improvements of the computerized ionospheric tomography using GPS and low earth orbit (LEO) satellites have been achieved (Bernhardt et al., 1998). Satisfactory comparisons between the HF radar phase front observations and electron density measurements by EISCAT and tomographic techniques have been made (Pryse et al., 1995; Nygrén et al., 1996, Mitchell et al., 1998). Sequences of ionospheric images of TIDs can be analyzed to deduce the AGW spatial and temporal structure and the periods of the gravity waves.

f) Total Electron Content (TEC) from GPS Receiver Networks

While the ISR and tomographic techniques allow mapping the 2D structure of TIDs (Hunsucker, 1982; Hocke and Schlegel, 1996), additional techniques are needed to observe the 3D structure of TIDs. Total electron content (TEC) measurements in conjunction with other instrumentation like the MU radar (Shiokawa et al., 2002, 2003; Tsugawa et al., 2003) and ISRs (Nicolls et al., 2004) have been used to study large-scale TIDs. Total electron content (TEC) maps that respond to height-integrated TID effects are now provided by a dense GPS network (e.g., Coster et al., 2003). Individual receiver and satellite biases (Sardon et al., 1994; Wilson et al., 1999) are removed and slant TEC is converted to vertical TEC using a simple mapping function for an ionosphere stratified into thin layers. Difficulties of combining GPS TEC measurements with ISR observations of electron densities as a function of time and altitude are discussed by Nicolls et al. (2004), but useful techniques to study TEC perturbations due to TIDs have been developed (Saito et al., 1998; Ho et al. 1999; Afraimovich et al., 2000; Shiokawa et al., 2002a&b; 2003; Tsugawa et al. 2003; Nicolls et al., 2004). Recently,

TEC maps have been used to support observations of TIDs and tides by a super-resolution HF direction finding system (Hawlitschka, 2006).

As discussed by Nicolls et al. (2004), the TEC perturbation (TECP) at mid-latitudes caused by the ionospheric fluctuations associated with the neutral wind due to large-scale AGWs is not expected to be large. Using the GPS mapping technique, Nicolls et al. (2004; their Figure 3) showed nighttime TECPs of ~1-2 TECU (1 TECU = 10^{16} el/m^2) due to the large-scale TIDs on October 2, 2002 (Figure 1-41).

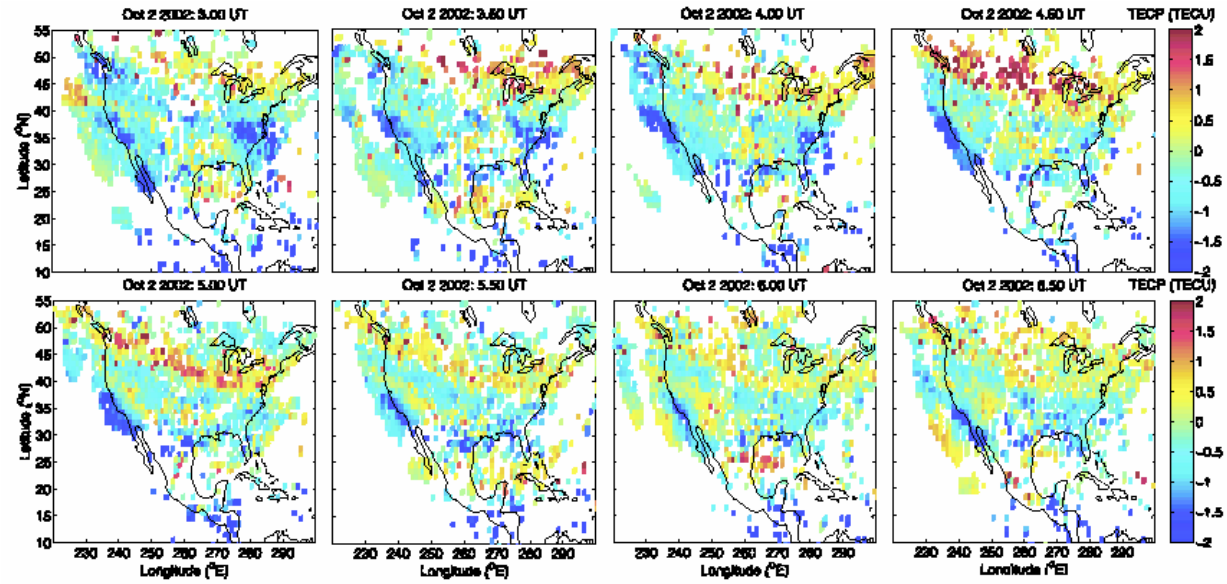

Figure 1-41: Examples of TECP Images Produced by the GPS Mapping Technique (Nicolls et al., 2004).

g) Airglow

Several naturally occurring airglow emissions originate in the mesosphere and lower thermosphere region 80-110 km (Hecht, 2004) as well as the bottom-side ionosphere region 100-300 km (Kubota et al., 2001). As the gravity waves propagate through the atmosphere they cause density and temperature variations, which are then detected by airglow imagers as quasi-periodic variations in the airglow emission intensity. Large- and medium-scale TIDs are commonly observed in the 630-nm airglow images as bands moving southwestward in the northern hemisphere (Mendillo et al., 1997; Shiokawa et al., 2000a) and have been detected at geomagnetically conjugate stations in the southern hemisphere (Otsuka et al., 2004; Shiokawa et al., 2005). Coordinated experiments using the Arecibo ISR, an airglow imager, and digital ionosondes addressed the electrodynamics of gravity waves in the thermosphere (Miller et al., 1997) and a new theory of gravity-wave-induced electric fields has been proposed (Miller, 1997; Kelley and Miller, 1997). However, at least one class of low-velocity TIDs is not well understood theoretically (Garcia et al., 2000).

1.3.8.4 Observations and Modeling of AGWs/TIDs and Their Sources

A causal relationship between the AGWs generated in the auroral zone (by Joule heating, Lorentz forcing or particle precipitation) and TIDs has been well established (Chimonas and Hines, 1970; Chimonas, 1970; Testud, 1970; Richmond, 1978; Hocke and Schlegel, 1996). The Worldwide Atmospheric Gravity-wave Studies (WAGS) program (see, e.g., Crowley and Williams, 1987; Williams et al., 1993) clearly demonstrated that large-scale TIDs observed at mid-latitudes originate in the auroral oval. The observed characteristic periods of TIDs (typically 30-70 min) are well correlated with auroral variations (electric

field and Joule heating) of similar periods. On the other hand, auroral fluctuations with time scales of 20 min or less are strongly attenuated in the observed TIDs. Thus the auroral plasma convection, which commonly exhibits quasi-periodic, recurrent bursts (e.g., Williams et al., 1992), is a likely source of gravity waves. Such a source then determines not only the magnitude but also the period of the gravity waves (Crowley and Williams, 1987).

Direct and earth-reflected gravity waves were theoretically described and distinguished by Francis (1974), who pointed out that the latter appear in the F region as wave packets (a number of nearly monochromatic wave cycles) while the former induce isolated (nonperiodic) TIDs, consistent with the modeling results by Millward et al. (1993a, b). Francis (1974) also suggested a dominance of upper-atmospheric (as opposed to tropospheric) sources of gravity waves. While the large-scale TIDs have been attributed to auroral electrojet surges, the identification of sources of medium-scale TIDs has remained difficult. Both auroral activity (e.g., Hunsucker, 1982, Lewis et al., 1996, Oyama et al., 2001) and tropospheric weather disturbances (e.g., Bertin et al., 1975; Waldock and Jones, 1987; Oliver et al., 1997) can generate medium-scale AGWs, particularly for TIDs that are observed at middle latitudes.

At high latitudes, Samson et al. (1990) suggested that the AGWs observed by SuperDARN originate near the ionospheric convection reversal boundary. Thus it has been hypothesized that the IMF variations that lead to auroral electrojet fluctuations generate AGWs/TIDs, but the actual identification of the distant sources of the observed TIDs has been posing difficulties in spite of the use of various observation techniques (Lewis et al., 1996; MacDougall et al., 1997; Hall et al., 1999, Oyama et al., 2001). One of the advantages of SuperDARN for the study of gravity waves is that it allows simultaneous observations of TIDs and of their high-latitude source activity. Various HF propagation modes frequently result in observations of the TID modulated ground scatter along with the ionospheric echoes at ranges immediately poleward of the ground scatter (Huang et al., 1998a; Sofko and Huang, 2000). A causality sequence of relationship among the IMF southward turnings, auroral electrojet currents, and gravity waves was reported for a case study (Huang et al., 1998b). The proximity of the observed TID signatures to the source region (auroral electrojet) indicated direct rather than ducted or earth-reflected gravity waves. Sofko and Huang (2000) associated each of the 100-min pulses observed in the ionospheric flow (electric field), the IMF and ground magnetic field with a pair of gravity waves (two wave pulses).

Modelling of gravity waves generated by enhancements in the ionospheric electric field shows that each electric field enhancement causes a Joule heating pulse, which in turn launches a gravity wave propagating equatorward and poleward of the source region (Millward et al., 1993a, b). This is consistent with some of the HF radar TID observations of TIDs which led Prikryl et al. (2005) to suggest that solar wind Alfvén wave coupling to the magnetosphere-ionosphere system is a source of auroral AGWs. Large scale traveling ionospheric/atmospheric disturbances generated by auroral surges have been observed and modeled (Hajkowicz, 1991; Balthazor and Moffett, 1997; Lee et al., 2004; Fujiwara and Miyoshi, 2006) to propagate as far as the equator. Medium-scale gravity waves can be ducted between the earth's surface and the lower thermosphere, propagate through the dissipationless lower atmosphere over long distances (Hocke and Schlegel, 1996) and leak from the lower atmosphere back to the thermosphere (Mayr et al., 1984a,b). Djuth et al. (2004) suggested that, near Arecibo, there was a continuum of gravity waves in the thermosphere, and that these showed "sets" of waves separated typically by 20-60 min. Livneh et al. (2007) used the incoherent scatter radar at the Arecibo Observatory to observe quasi-coherent continuous waves with periods of ~1 hour in the ionosphere. These are also typical periods of auroral AGWs (Bristow and Greenwald, 1997), which are likely generated by pulsed ionospheric convection (Prikryl et al., 2005) virtually at all times and sometimes observed as TIDs by SuperDARN (Bristow et al., 1996; He et al., 2004).

SuperDARN work on TIDs pursued studies of the relationship between auroral electrojet currents and TIDs (Sofko and Huang, 2000; Huang et al., 1998a,b) but also addressed the question of how one can deduce TID properties from SuperDARN ground scatter (Hall et al., 1999; Stocker et al., 2000;

MacDougall et al., 2001). In particular, the last named authors showed that the assumption that a TID must be located at approximately half the range of the radar ground scatter intensification that arises due to focusing by TIDs, may be valid only for low amplitude TIDs. Such an assumption, used in earlier SuperDARN studies of TIDs, implies that the TID horizontal phase velocity would be about half of the observed time rate of change of the distance to the ground backscatter intensification. This assumption of the factor 0.5 for making TID horizontal phase velocity estimations was challenged by Hall et al. (1999), who suggested that a more suitable value is 0.6. Further HF ray-tracing analysis and modeling of TIDs led MacDougall et al. (2001) to conclude that, depending on the TID amplitude, the appropriate proportionality factor may actually be close to 1.0, which almost doubles the horizontal TID phase velocities and wavelengths derived from the SuperDARN ground scatter signatures of TIDs.

Figure 1-42 shows an example of the SuperDARN Hankasalmi radar ground scatter power modulated by medium-scale gravity waves that were produced in the dayside auroral oval by ionospheric convection (Figure 1-39) that was pulsed by solar wind Alfvén wave coupling to dayside magnetopause (Prikryl et al., 2005). The slanted bands of strongly focused ground scatter power are due to equatorward-moving TID fronts in the F region. There is a one-to-one correspondence between the IMF-B_Y oscillations, the ground magnetic field perturbations and the TIDs.

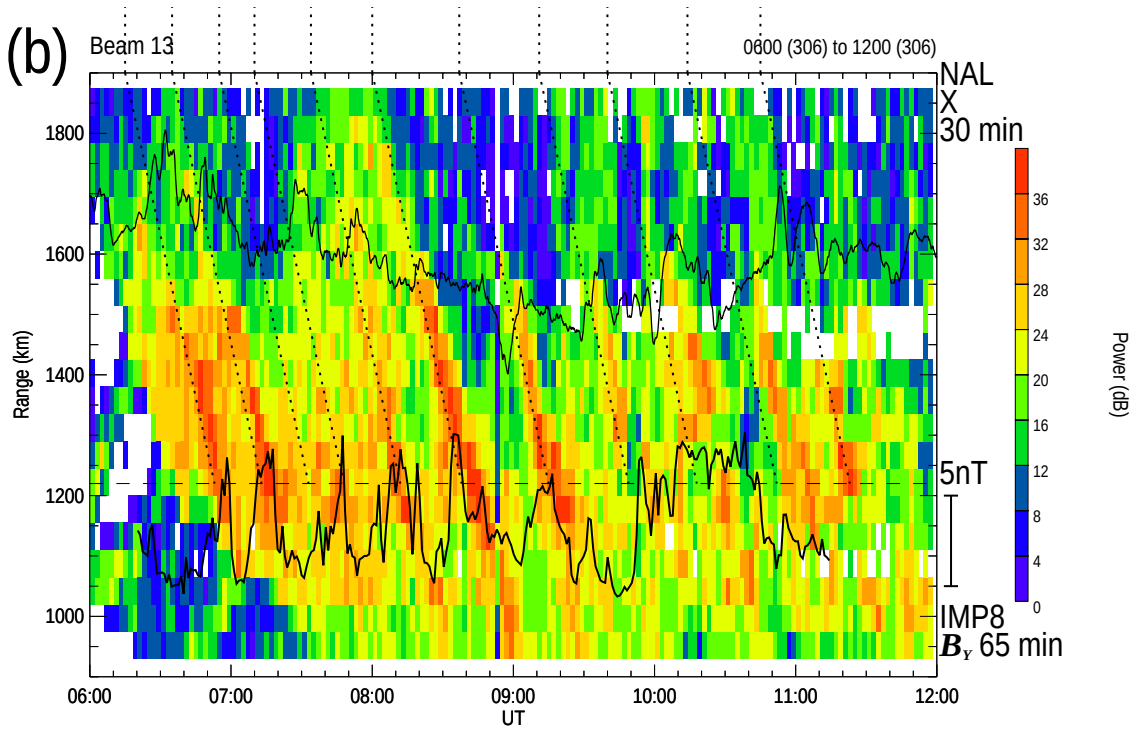

Figure 1-42: The SuperDARN Hankasalmi Radar Ground Scatter Power at Near Ranges.

Superposed are time series of the IMP-8 IMF BY (shifted by 65 min) and the Ny Åleslund (NAL) ground magnetic field X component (shifted by 30 min). The dotted lines along the medium-scale TID bands are superposed to show the correspondence between the ground magnetic field perturbations and the IMF-BY oscillations due to solar wind Alfvén waves (Prikryl et al., 2005).

While most of the above research has focused on identifying the wave source and wave properties, there have been studies, in addition to those mentioned in Sections 1.3.9.3 (d) and (e), that have attempted to determine the amplitude of the electron density fluctuation induced by the wave. By comparing the profiles obtained by ray-tracing calculations to model the received HF power with those measured by

SuperDARN, Bristow and Greenwald (1995) estimated that the peak density perturbations of TIDs were as great as 35%. This technique to estimate the electron density perturbations was extended by Stocker et al. (2000) who modeled the change in skip distance and obtained very similar amplitudes.

1.3.8.5 TID Effects on Transionospheric HF Propagation

While the reflection of HF radio waves from the bottom-side ionospheric layers that are modulated by TIDs are employed in TID detection techniques (Section 1.3.9.3), some of the HF power penetrates the ionosphere and can be detected by satellites. Transionospheric HF propagation can be significantly affected by TIDs. Transionospheric HF propagation experiments have been conducted in the past (James et al., 2006; James, 2006) and are planned for future satellite missions (Yau et al., 2006). James et al. (2006) investigated HF fades caused by multiple wave fronts detected by sounder receiver dipole antennas on board ISIS 1 and ISIS 2 spacecraft. Fades of the Faraday type, involving superposition of the ordinary and extraordinary wave components, as well as single-mode fades due to interference of multi-path signals, were observed.

Figure 1-43 shows rays at 9.303 MHz traced through a two-dimensional model ionospheric density based on tomographic data (Pryse et al., 1995) representing medium-scale TIDs with a horizontal wavelength of ~250 km. Figure 1-43 reveals the resulting shapes and orientations of mesoscale ionospheric irregularities, the most intense of which are elongated approximately parallel to the geomagnetic field. The low elevation rays are reflected and focused/defocused by the bottom-side ionospheric undulations caused by the TIDs. Because the density gradients of the elongated TID structure are approximately perpendicular to magnetic field, the poleward directed high-elevation rays (ray-gradient angles near 0º) suffer minimal refraction. In contrast, the ray-gradient angles are near 90º on the equatorward (left) side, so the structure is most effective in focusing or defocusing rays in that sector. Such rays were intercepted by the ISIS-2 spacecraft which observed interference fades as it was passing at altitude of 1400 km (James et al., 2006).

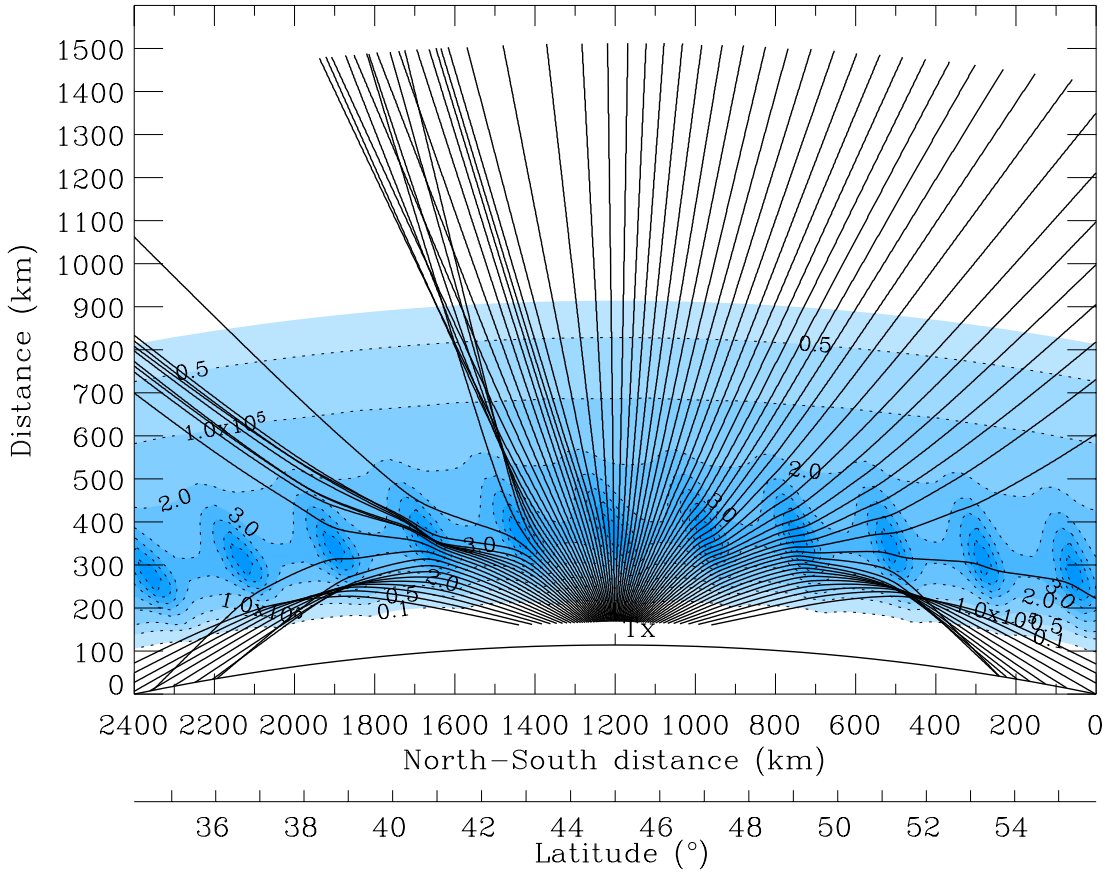

Figure 1-43: *O*-Mode Rays at 9.303MHz Traced in Two Dimensions Through a Model Ionosphere with Latitudinally Periodic Density Structure Resulting from Medium-Scale TIDs. The elevation of the lowest ray on the right side is 11° and increases in steps of 2°. The abscissa scale is distance along the chord of the Earth's surface, while the ordinate axis is at right angles to it. The latitude scale refers to the Earth's surface (James et al., 2006).

1.3.8.6 References

Afraimovich, E. L., Kosogorov, E. A., Leonovich, L. A., Palamartchouk, K. S., Perevalova, N. P., Pirog, O. M., Observation of large-scale traveling ionospheric disturbances of auroral origin by global GPS networks, *Earth, Planets and Space*, Volume 52, p. 669-674, 2000.

Austen, J. R., S. J. Franke, C.H. Liu, and K. C. Yeh, Application of computerized tomography techniques to ionospheric research, in *Radio beacon contribution to the study of ionization and dynamics of the ionosphere and corrections to geodesy*, Ed. A. Tauriainen, University of Oulu, Oulu Finland, Part 1, 25-35, 1986.

Balthazor, R. L. and R. J. Moffett, A study of atmospheric gravity waves and travelling ionospheric disturbances at equatorial latitudes, *Ann. Geophysicae*, Vol. 15(8), 1048-1056, 1997.

Bernhardt, P. A., McCoy, R. P., Dymond, K. F., Picone, J. M., Meier, R. R., Kamalabadi, F., Cotton, D. M., Chakrabarti, S., Cook, T. A., Vickers, J. S., Stephan, A. W., Kersely, L., Pryse, S. E., Walker, I. K., Mitchell, C. N., Straus, P. R., Na, H., Biswas, C., Bust, G. S., Kronschnabl, G. R., and Raymund, T. D.: Two-dimensional mapping of the plasma density in the upper atmosphere with computerized ionospheric tomography (CIT), *Phys. Plas.*, 5, 2010-2021, 1998.

Bertin, F., J. Testud, and L. Kersley, Medium-scale gravity waves in the ionospheric F-region and their possible origin in weather disturbances, *Planet. Space Sci.*, 23, 493-507, 1975.

Bristow, W. A., R. A. Greenwald, and J. C. Samson, Identification of high-latitude acoustic gravity wave sources using the Goose Bay HF radar, *J. Geophys. Res.*, 99, 319-331, 1994.

Bristow, W. A. and Greenwald, R. A.: Estimating gravity wave parameters from oblique high-frequency backscatter: modeling and analysis, *J. Geophys. Res.*, 100, 3639-3648, 1995.

Bristow, W. A. and Greenwald, R. A.: On the spectrum of thermospheric gravity waves observed by the Super Dual Auroral Radar Network, *J. Geophys. Res.,* 102(A6), 11585-11596, 1997.

Bristow, W. A., Greenwald, R. A., and Villain, J. P.: On the seasonal dependence of medium-scale atmospheric gravity waves in the upper atmosphere at high latitudes, *J. Geophys. Res.*, 101, 15685-15699, 1996.

Chimonas, G., and C. O. Hines, Atmospheric gravity waves launched by auroral currents, *Planet. Space Sci.*, 18, 565-582, 1970.

Chimonas, G., The equatorial electrojet as a source of long period travelling ionospheric disturbances, *Planet. Space Sci.*, 18, 583-589, 1970.

Chisham, G., Lester, M., Milan, S. E., Freeman, M. P., Bristow, W. A., Grocott, A., McWilliams, K. A., Ruohoniemi, J. M., Yeoman, T. K., Dyson, P. L., Greenwald, R. A., Kikuchi, T., Pinnock, M., Rash, J. P. S., Sato, N., Sofko, G. J., Villain J.-P., and Walker, A. D. M.: A decade of the Super Dual Auroral Radar Network (SuperDARN): scientific achievements, new techniques and future directions, *Surv. Geophys.*, 28, 33-109, DOI 10.1007/s10712-007-9017-8, 2007.

Coster, A. J., J. C. Foster, and P. J. Erickson, Space weather: Monitoring the ionosphere with GPS, *GPS World*, May, 40-45, 2003.

Crowley, G., and P.J.S. Williams, Observations of the source and propagation of atmospheric gravity waves, *Nature*, 328, 231-233, 1987.

Davies, K., Ionospheric Radio, Peter Peregrinus Ltd, London, 1990.

Djuth F. T., Sulzer, M. P., Gonzáles, S. A., Mathews, J. D., Elder, J. H., Walterscheid, R. L.: A continuum of gravity waves in the Arecibo thermosphere?, *Geophys. Res. Lett.*, 31, L16801, doi:10.1029/2003GL019376, 2004.

Francis, S. H., A theory of medium-scale traveling ionospheric disturbances, *J. Geophys. Res.*, 79, 5245-5260, 1974.

Fujiwara, H., and Y. Miyoshi, Characteristics of the large-scale traveling atmospheric disturbances during geomagnetically quiet and disturbed periods simulated by a whole atmosphere general circulation model, *Geophys. Res. Lett.*, 33, L20108, doi:10.1029/2006GL027103, 2006.

Garcia, F. J., M. C. Kelley, J. J. Makela, and C.-S. Huang, Airglow observations of mesoscale low-velocity traveling ionospheric disturbances at midlatitudes, *J. Geophys. Res.*, 105 (A8), 18,407-18,415, 2000.

Georges, T. M., HF Doppler studies of traveling ionospheric disturbances, *J. Atmos. Terr. Phys.*, 30, 735-746, 1968.

Grant, I. F., J. W. MacDougall, J. M. Ruohoniemi, W. A. Bristow, G. J. Sofko, J. A. Koehler, D. Danskin, and D. André, Comparison of plasma flow velocities determined by the ionosonde Doppler drift technique, SuperDARN radars, and patch motion, *Radio Sci.*, 30, 1537-1549, 1995.

Greenwald, R.A., K.B. Baker, J.R. Dudeney, M. Pinnock, T.B. Jones, E.C. Thomas, J.-C. Villain, J.-C. Cerrisier, C. Senior, C. Hanuise, R.D. Hunsucker, G. Sofko, J. Koehler, E. Nielsen, R. Pellinen, A.D.M. Walker, N. Sato, and H. Yamagishi, DARN/SUPERDARN A global view of the dynamics of high-latitude convection, *Space Sci. Rev.*, 71, 761-796, 1995.

Hajkowicz, L. A.: Auroral electrojet effect on the global occurrence pattern of large-scale traveling ionospheric disturbances, *Planet. Space Sci.*, 39, 1189-1196, 1991.

Hall, G.E., J.W. MacDougall, J.-F. Cecile, D.R. Moorcroft, and J.-P. St.-Maurice, Finding gravity wave source positions using the Super Dual Auroral Radar Network, *J. Geophys. Res.*, 104, 67-78, 1999.

Hargreaves, J. K., "The Solar Terrestrial Environment", Cambridge University Press, 1992.

Hawlitschka, S., Travelling ionospheric disturbances (TIDs) and tides observed by a super-resolution HF direction finding system, *Journal of Atmospheric and Solar-Terrestrial Physics*, 68(3-5), 568-577, 2006.

He, L.-S. , Dyson, P. L., Parkinson, M. L., and Wan, W.: Studies of medium scale travelling ionospheric disturbances using TIGER SuperDARN radar sea echo observations, *Ann. Geophysicae*, 22, 4077-4088, 2004.

Hecht, J. H., Instability layers and airglow imaging, *Rev. Geophys.*, 42, RG1001, doi:10.1029/ 2003RG000131, 2004.

Hines, C.O., Internal atmospheric gravity waves at ionospheric heights, *Can. J. Phys.*, 38, 1441-1481, 1960.

Ho, C. M., A. J. Mannucci, U. J. Lindqwister, X. Pi, and B. T. Tsurutani, Global ionosphere perturbations monitored by the worldwide GPS network, *Geophys. Res. Lett.*, 23(22), 3219-3222, 1996.

Hocke, K., and K. Schlegel, A review of atmospheric gravity waves and traveling ionospheric disturbances: 1982-1995, *Ann. Geophysicae*, 14, 917-940, 1996.

Huang, C. S., D. A. André, and G. J. Sofko, High-latitude ionospheric perturbations and gravity waves: 1. Observational results, *J. Geophys. Res.*, 103, 2131-2141, 1998a.

Huang, C. S., D. A. André, and G. J. Sofko, Observations of solar wind directly driven auroral electrojets and gravity waves, *J. Geophys. Res.*, 103, 23347-23356, 1998b.

Hunsucker, R. D., Atmospheric gravity waves generated in the high-latitude ionosphere: A review, *Reviews of Geophysics and Space Physics*, Vol. 20, 293-315, 1982.

Hunsucker, R.D., and L.H. Tveten, Large traveling-ionospheric-disturbances observed at midlatitudes utilizing the high resolution h.f. backscatter technique, *J. Atmos. Terr. Phys.*, 29, 909-916, 1967.

James, H. G., Effects on transionospheric HF propagation observed by ISIS at middle and auroral latitudes, *Advances in Space Research*, 38(11), 2303-2312, 2006.

James, H. G., R. G. Gillies, G. C. Hussey, and P. Prikryl, HF fades caused by multiple wave fronts detected by a dipole antenna in the ionosphere, *Radio Sci.*, 41, RS4018, doi:10.1029/ 2005RS003385, 2006.

Kelley, M. C., and C. A. Miller, Electrodynamics of midlatitude spread F 3. Electrohydrodynamic waves? A new look at the role of electric fields in thermospheric wave dynamics, *J. Geophys. Res.*, 102(A6), 11,539-11,548, 1997.

Kirchengast, G., K. Hocke, and K. Schlegel, Gravity waves determined by modeling of traveling ionospheric disturbances in incoherent-scatter radar measurements, *Radio Sci.*, 30(5), 1551-1567, 1995.

Kirchengast, G., Elucidation of the physics of the gravity wave-TID relationship with the aid of theoretical simulations, *J. Geophys. Res.*, 101, 13353-13368, 1996.

Kirchengast, G., K. Hocke, and K. Schlegel, The gravity wave-TID relationship: insight via theoretical model - EISCAT data comparison, *J. Atmos. Terr. Phys.*, 58, 233-243, 1996.

Kubota, M., H. Fukunishi, and S. Okano, Characteristics of medium- and large-scale TIDs over Japan derived from OI 630-nm nightglow observation, *Earth Planets Space*, 53, 741-751, 2001.

Lee, C.-C., J.-Y. Liu, M.-Q. Chen, S.-Y. Su, H.-C. Yeh, and K. Nozaki, Observation and model comparisons of the traveling atmospheric disturbances over the Western Pacific region during the 6-7 April 2000 magnetic storm, *J. Geophys. Res.*, 109, A09309, doi: 10.1029/2003JA010267, 2004.

Livneh, D. J., Seker, I., Djuth, F. T., Mathews, J. D.: Continuous quasiperiodic thermospheric waves over Arecibo, *Journal of Geophysical Research*, 112, A7, CiteID A07313, 2007.

Lewis, R.V., P.J.S. Williams, G.H. Millward and S. Quegan, The generation and propagation of atmospheric gravity waves from activity in the auroral electrojet, *J. Atmos. Terr. Phys.*, 58, 807-820, 1996.

MacDougall, J.W., G.E. Hall, and K. Hayashi, F region gravity waves in the central polar cap, *J. Geophys. Res.*, 102, 14513-14530, 1997.

MacDougall, J.W., D.A. Andre, G.J. Sofko, C.-S. Huang, and A.V. Koustov, Travelling ionospheric disturbance properties deduced from Super Dual Auroral Radar measurements, *Ann. Geophysicae*, 18, 1550-1559, 2001.

Makela, J. J., S.A. Gonzalez, B. MacPherson, X. Pi, M.C. Kelley, P.J. Sultan, Intercomparisons of total electron content measurements using the Arecibo incoherent scatter radar and GPS, *Geophys. Res. Lett.*, 27(18), 2841-2844, 2000.

Mayr, H. G., Harris, I., Varosi, F., and Herrero, F. A.: Global excitation of wave phenomena in a dissipative multiconstituent medium 1. Transfer function of the Earth's thermosphere, *J. Geophys. Res.*, 89, 10929-10959, 1984a.

Mayr, H. G., Harris, I., Varosi, F., and Herrero, F. A.: Global excitation of wave phenomena in a dissipative multiconstituent medium 2. Impulsive perturbations in the Earth's thermosphere, *J. Geophys. Res.*, 89, 10961-10986, 1984b.

Mendillo, M., J. Baumgardner, D. Nottingham, J. Aarons, B. Reinisch, J. Scali, and M. Kelley, Investigations of thermospheric-ionospheric dynamics with 6300 Å images from the Arecibo Observatory, *J. Geophys. Res., 102,* 7331, 1997.

Miller, C. A., W. E. Swartz, M. C. Kelley, M. Mendillo, D. Nottingham, J. Scali, and B. Reinisch, Electrodynamics of midlatitude spread F 1. Observations of unstable, gravity wave-induced ionospheric electric fields at tropical latitudes, *J. Geophys. Res.*, 102(A6), 11,521-11,532, 1997.

Miller, C. A., Electrodynamics of midlatitude spread F 2. A new theory of gravity wave electric fields, *J. Geophys. Res.*, 102(A6), 11,533-11,538, 1997.

Millward, G.H., S. Quegan, R.J. Moffett, T.J. Fuller-Rowell, and D. Rees, A modeling study of the coupled ionospheric and thermospheric response to an enhanced high-latitude electric field event, *Planet. Space Sci.*, 41, 45-56, 1993a.

Millward, G.H., R.J. Moffett, S. Quegan, and T.J. Fuller-Rowell, Effect of an atmospheric gravity wave on the midlatitude ionospheric F layer, *J. Geophys. Res.*, 98, 19173-19179, 1993b.

Millward, G.H., A resonance effect in AGWs created by periodic recurrent bursts in the auroral electric field, *Ann. Geophysicae*, 12, 94-96, 1994.

Mitchell, C. N., I. K. Walker, S. E. Pryse, I. Kersley, I. W. McCrea, T. B. Jones, First complementary observations by ionospheric tomography, the EISCAT Svalbard radar and the CUTLASS HF radar, *Ann. Geophysicae*, 16, 1519-1522, 1998.

Muldrew, D.B., An ionospheric ray-tracing technique and its application to a problem in long-distance radio propagation, *IRE Trans.*, Vol. AP-7, No. 4, 393-396, 1959.

Nicolls, M. J., M. C. Kelley, A. J. Coster, S. A. González, and J. J. Makela, Imaging the structure of a large-scale TID using ISR and TEC data, *Geophys. Res. Lett.*, 31, L09812, doi: 10.1029/ 2004GL019797, 2004.

Nygrén, T., Markkanen, M., Lehtinen, M., Tereshchenko, E. D., Khudukon, B. Z., Evstafiev, O. V., Pollari, P., Comparison of F-region electron density observations by satellite radio tomography and incoherent scatter methods, *Annales Geophysicae*, vol. 14, Issue 12, pp.1422-1428, 1996.

Oliver, W. L., Y. Otsuka, M. Sato, T. Takami, and S. Fukao, A climatology of F region gravity waves propagation over the middle and upper atmosphere radar, *J. Geophys. Res.*, 102, 14449-14512, 1997.

Otsuka, Y., K. Shiokawa, T. Ogawa, and P. Wilkinson, Geomagnetic conjugate observations of medium-scale traveling ionospheric disturbances at midlatitude using all-sky airglow imagers, *Geophys. Res. Lett.*, 31, L15803, doi: 10.1029/2004GL020262, 2004.

Oyama, S., M. Ishii, Y. Murayama, H. Shinagawa, S.C. Buchert, R. Fujii, and W. Kofman, Generation of atmospheric gravity waves associated with auroral activity in the polar F region, *J. Geophys. Res.*, 106, 18543-18554, 2001.

Prikryl, P., Muldrew, D. B., Sofko, G. J., and Ruohoniemi, J. M.: Solar wind Alfvén waves: A source of pulsed ionospheric convection and atmospheric gravity waves, *Ann. Geophysicae*, 23, 401-417, 2005.

Pryse, S. E., C. N. Mitchell, J. A. T. Heaton and L. Kersley, Travelling ionospheric disturbances imaged by tomographic techniques, *Ann. Geophysicae*, **13**, 1325-1330, 1995.

Raymund, T. D., Comparisons of several ionospheric tomography algorithms, *Ann. Geophysicae*, 13, 1254-1262, 1995.

Raymund, T. D., J. R. Austen, S. J. Franke, C. H. Liu, J. A. Klobuchar, and J. Stalker, Application of computerized tomography to the investigation of ionospheric structures, *Radio Sci.*, 25, 771-789, 1990.

Raymund, T. D., S. E. Pryse, L. Kersley, and J. A. T. Heaton, Tomographic reconstruction of ionospheric electron density with European incoherent scatter verification, *Radio Sci.*, 28, 811-817, 1993.

Raymund, T. D., S. J. Franke, and K. C. Yeh, Ionospheric tomography: its limitation and reconstruction methods, *J. Atmos. Terr. Phys.*, 56, 637-657, 1994.

Richmond, A.D., Gravity wave generation, propagation, and dissipation in the thermosphere, *J. Geophys. Res.*, 83, 4131-4145, 1978.

Ruohoniemi, J. M. and Baker, K. B., Large-scale imaging of high-latitude convection with Super Dual Auroral Radar Network HF radar observations, *J. Geophys. Res.*, 103, 20797-20811, 1998.

Saito, A., S. Fukao, and S. Miyazaki, High resolution mapping of TEC perturbations with the GSI GPS network over Japan, *Geophys. Res. Lett.*, 25(16), 3079-3082, 1998.

Samson, J. C., R. A. Greenwald, J. M. Ruohoniemi, and K. B. Baker, High-frequency radar observations of atmospheric gravity waves in the high latitude ionosphere, *Geophys. Res. Lett.*, 16, 875-878, 1989.

Samson, J. C., R. A. Greenwald, J. M. Ruohoniemi, A. Frey, and K. B. Baker, J., Goose Bay radar observations of earth-reflected atmospheric gravity waves in the high-latitude ionosphere, *J. Geophys. Res.*, 95, 7693-7709, 1990.

Sardon, E., A. Rius, and N. Zarraoa, Estimation of the transmitter and receiver differential biases and the ionospheric total electron content from Global Positioning System observations, *Radio Sci.*, 29(3), 577-586, 1994.

Shiokawa, K., Y. Otsuka, M. K. Ejiri, Y. Sahai, T. Kadota, C. Ihara, T. Ogawa, K. Igarashi, S. Miyazaki, and A. Saito, Imaging observations of the equatorward limit of midlatitude traveling ionospheric disturbances, *Earth Planets Space*, 54, 57-62, 2002a.

Shiokawa, K., Y. Otsuka, T. Ogawa, N. Balan, K. Igarashi, A. J. Ridley, D. J. Knipp, A. Saito, and K. Yumoto, A large-scale traveling ionospheric disturbance during the magnetic storm of 15 September 1999, *J. Geophys. Res.*, 107(A6), 1088, doi: 10.1029/2001JA000245, 2002b.

Shiokawa, K., et al., Thermospheric wind during a storm-time large-scale traveling ionospheric disturbance, *J. Geophys. Res.*, 108(A12), 1423, doi: 10.1029/2003JA010001, 2003.

Shiokawa, K., Y. Otsuka, T. Tsugawa, T. Ogawa, A. Saito, K. Ohshima, M. Kubota, T. Maruyama, T. Nakamura, M. Yamamoto, and P. Wilkinson, Geomagnetic conjugate observation of nighttime medium-scale and large-scale traveling ionospheric disturbances: FRONT3 campaign, *J. Geophys. Res.*, 110, A05303, doi:10.1029/2004JA010845, 2005.

Sofko, G. J. and C.S. Huang, SuperDARN observations of medium-scale gravity wave pairs generated by Joule heating in the auroral zone, *Geophys. Res. Lett.*, 27, 485-488, 2000.

Stocker, A.J., N.F. Arnold, and T.B. Jones, The synthesis of travelling ionospheric disturbances (TID) signatures in HF radar observations using ray tracing, *Ann. Geophysicae*, 18, 56-64, 2000.

Testud, J., Gravity waves generated during magnetic substorms, *J. Atmos. Terr. Phys.*, 32, 1793-1805, 1970.

Tsugawa, T., A. Saito, Y. Otsuka, and M. Yamamoto, Damping of large-scale traveling ionospheric disturbances detected with GPS networks during the geomagnetic storm, *J. Geophys. Res.*, 108(A3), 1127, doi: 10.1029/2003JA09433, 2003.

Tsugawa, T., A. Saito, and Y. Otsuka, A statistical study of large-scale traveling ionospheric disturbances using the GPS network in Japan, *J. Geophys. Res.*, 109(A6), A06302, doi: 10.1029/ 2003JA010302, 2004.

Tveten, L.H., Ionospheric motions observed with high-frequency backscatter sounders, J. of Research of the National Bureau of Standards, *D. Radio Propagation,* 65D, 115-127, 1961.

Waldock, J. A., and T.B. Jones, Source regions of medium scale traveling ionospheric disturbances observed at mid-latitudes, *J. Atmos. Terr. Phys.*, 49, 105-114, 1987.

Williams, P.J.S., R.V. Lewis, T.S. Virdi, M. Lester, E. Nielsen, Plasma flow bursts in the auroral electrojets, *Ann. Geophysicae*, 10, 835-848, 1992.

Williams, P.J.S., T.S. Virdi, R.V. Lewis, M. Lester, A.S. Rodger, I.W. McCrea, and K.S.C. Freeman, Worldwide atmospheric gravity-wave study in the European sector 1985-1990, *J. Atmos. Terr. Phys.*, 55, 683-696, 1993.

Wilson, B. D., C. H. Yinger, W. A. Feess, and C. Shank, New and improved - the broadcast inter-frequency biases, *GPS World*, 10, 56-66, 1999.

Yau, A. W., James, H. G., Liu, W., The Canadian Enhanced Polar Outflow Probe (e-POP) mission in ILWS, *Advances in Space Research*, Volume 38, Issue 8, p. 1870-1877, 2006.

Yeh, K. C. and Liu, C. H.: Acoustic-gravity waves in the upper atmosphere, *Rev. Geophys. Space Phys.*, 12, 193-216, 1974.