In the past the proceedings of symposia sponsored by the National Institute of Polar Research have been published as “Memoirs of National Institute of Polar Research, Special Issue”.

Since 1987, new series of “Proceedings of the NIPR Symposium” for each discipline have been published as follows:

Proceedings of the NIPR Symposium on Upper Atmosphere Physics
Proceedings of the NIPR Symposium on Polar Meteorology and Glaciology
Proceedings of the NIPR Symposium on Antarctic Geosciences
Proceedings of the NIPR Symposium on Antarctic Meteorites
Proceedings of the NIPR Symposium on Polar Biology

Since 1999, “Proceedings of the NIPR Symposium on Upper Atmosphere Physics” has been renamed to:

“Advances in Polar Upper Atmosphere Research”

Editorial Board
Editor-in-Chief: Okitsugu Watanabe, National Institute of Polar Research (NIPR)
Editors: Takehiko Aso, NIPR
Masaki Ejiri, NIPR
Yoshiyuki Fujii, NIPR
Mitsuo Fukuchi, NIPR
Yoichi Fukuda, Kyoto University
Ken Furuya, University of Tokyo
Yoshikuni Hiroi, Chiba University
Hiroshi Kanda, NIPR
Shun’ichi Kobayashi, Niigata University
Satoru Kojima, Tokyo Woman’s Christian University
Shinji Mae, Asahikawa National College of Technology
Masamichi Miyamoto, University of Tokyo
Yasuhiro Naito, NIPR
Tadahiko Ogawa, Nagoya University
Takayuki Ono, Tohoku University
Kazuo Shibuya, NIPR
Kazuyuki Shiraishi, NIPR
Masashi Takada, Nara Women’s University
Takashi Yamanouchi, NIPR
Foreword

This volume 16 is the fourth volume of the “Advances in Polar Upper Atmosphere Research (APUAR)” which succeeded the “Proceedings of the NIPR Symposium on Upper Atmosphere Physics” in 1999 and includes papers presented at or related to “The Twenty-fifth Symposium on Coordinated Observations of the Ionosphere and the Magnetosphere in the Polar Regions” held at the National Institute of Polar Research (NIPR) in August 2001. The annual Symposium is devoted to discuss significant scientific results of studies on the ionosphere and the magnetosphere in the polar regions. The journal extensively accepts contributed papers which do mark the advance of research works on the polar upper atmosphere physics in recent years.

The present volume contains 14 refereed scientific papers, research notes and reports. We sincerely hope that this will be valuable to those who are interested in the polar ionospheric and magnetospheric physics and to those who are engaged in Arctic and Antarctic research activities.

Editorial Board of APUAR
Editor-in-Chief: Takehiko Aso
Editors: Masaki Ejiri
         Tadahiko Ogawa
         Natsuot Sato
         Takayuki Ono
         Hisao Yamagishi
         Hiroshi Miyaoka
         Yasuhiro Murayama
         Makoto Taguchi
Editorial secretary;
   Nobuko Takahama
Advances in Polar Upper Atmosphere Research

No. 16

CONTENTS

The high-latitude artificial aurora of 21 February 1999: An analysis

\textit{M.J. Kosch, M.T. Rietveld, T.K. Yeoman, K. Cierpka}
\textit{and T. Hagfors}\ldots 1

Evaluation of the region 1 field-aligned current from the low-latitude boundary layer using the 1989 Tsyganenko model

\textit{Takashi Yamamoto, Masao Ozaki and Shoshi Inoue}\ldots 13

Characteristics of negative values of polar cap index as an indicator of reversed convection

\textit{Tsutomu Nagatsuma}\ldots 36

On the energy of the poloidal magnetic field near the ionosphere

\textit{Akimasa Yoshikawa, Masahiro Itonaga and Kiyohumi Yumoto}\ldots 45

Comparison of flow angle variations of \textit{E}-region echo characteristics at VHF and HF

\textit{R.A. Makarevitch, A.V. Koustov, K. Igarashi, K. Ohtaka, T. Ogawa, N. Nishitani, N. Sato, H. Yamagishi and A.S. Yukimatu}\ldots 59

\textit{E} region echoes observed with the Syowa HF radar under disturbed geomagnetic conditions

\textit{Tadahiko Ogawa, Nozomu Nishitani, Natsumo Sato, Hisao Yamagishi and Akira Sessai Yukimatu}\ldots 84

A new meridian imaging spectrograph for the auroral spectroscopy

\textit{Makoto Taguchi, Shoichi Okano, Takeshi Sakanoi, Naoko Koizumi, Takehiko Aso and Masaki Ejiri}\ldots 99

Research Note

Polar mesosphere winter echoes during solar proton events


SC-triggered plasma waves observed by the Akebono satellite in the polar regions and the plasmasphere

\textit{Atsuki Shinbori, Takayuki Ono and Hiroshi Oya}\ldots 126

New model of plasma convection during transpolar arc events

\textit{Hitoshi Nakai, Yosuke Kamide and Mitchell Brittnacher}\ldots 136
Report

New auroral spectrometer using an acousto-optic tunable filter

Kazuo Shiokawa, Yasuo Katoh, Mitsugi Satoh,
Tadahiko Ogawa, Makoto Taguchi and Hisao Yamagishi...146

Polar Patrol Balloon experiment in Antarctica during 2002–2003

Akira Kadokura, Hisao Yamagishi, Natsuo Sato, Masaki Ejiri,
Haruto Hirosawa, Takamasa Yamagami, Shoji Torii,
Fumio Tohyama, Michio Nakagawa, Toshimi Okada
and Edgar A. Bering...157

Calibration of CRL all-sky imagers using an integrating sphere

Masa-yuki Yamamoto, Minoru Kubota, Shu Takeshita,
Mamoru Ishii, Yasuhiro Murayama and Masaki Ejiri...173

A new method for monitoring and removing SuperDARN radar DC offsets

Akira Sessai Yukimatsu, Masaki Tsutsumi, Hisao Yamagishi
and Natsuo Sato...181

Note to contributors...193

Author index...195
The high-latitude artificial aurora of
21 February 1999: An analysis

M.J. Kosch\textsuperscript{1}, M.T. Rietveld\textsuperscript{2}, T.K. Yeoman\textsuperscript{3}, K. Cierpka\textsuperscript{2} and T. Hagfors\textsuperscript{2}

\textsuperscript{1}Communication Systems, Lancaster University, Lancaster LA1 4YR, U.K.
\textsuperscript{2}Max-Planck-Institut für Aeronomie, 37191 Katlenburg-Lindau, Germany
\textsuperscript{3}Physics and Astronomy, Leicester University, Leicester LE1 7RH, U.K.

\textbf{Abstract:} Artificial ionospheric optical emissions can be produced by high-power high-frequency (HF) radio waves. From the observation of HF induced artificial aurora on 21 February 1999, M.J. Kosch \textit{et al.} (Geophys. Res. Lett., 27, 2817, 2000) first noted the equatorial displacement of the optical emission towards the magnetic field line direction. This effect is investigated further and does not appear to be related to either the ion drift or neutral winds. HF coherent radar backscatter observations are presented for the first time in conjunction with artificial aurora. Ray tracing suggests that the optical displacement may be related to the mechanism for producing field-aligned irregularities, which the radars require for receiving backscatter. The altitude of the artificial aurora has been estimated and is shown to be well below the reflection altitude and upper hybrid resonance height of the pump wave.

1. Introduction

High-power high-frequency (HF) radio waves can produce artificial optical emissions in the ionosphere due to electron bombardment of the neutrals. Optical observations are one of the few means of detecting energetic electrons produced from plasma turbulence generated by powerful HF waves in the ionosphere (e.g. Sipler and Biondi, 1972; Haslett and Megill, 1974; Bernhardt \textit{et al.}, 1989a, b). The energisation threshold for the O\textsuperscript{(2D)} 630 nm and the O\textsuperscript{(5S)} 557.7 nm emissions is 1.96 eV and 4.17 eV, respectively (Bernhardt \textit{et al.}, 1989a). However, the effective threshold to excite the O\textsuperscript{(2D)} and O\textsuperscript{(5S)} states is > 3.1 and > 5.4 eV, respectively, due to quenching by N\textsubscript{2} (Djuth \textit{et al.}, 1999). Optical emissions at 630 nm have been observed at mid-latitude HF facilities such as Platteville (Haslett and Megill, 1974) and Arecibo (Sipler and Biondi, 1972; Gordon and Carlson, 1974; Bernhardt \textit{et al.}, 1988), at Moscow (Adeishvili \textit{et al.}, 1978) and the Sura facility near Vasilursk (Bernhardt \textit{et al.}, 1991). From mid-latitude studies, the optical emissions have been attributed to energetic electrons accelerated by Langmuir turbulence (Haslett and Megill, 1974; Sipler \textit{et al.}, 1974; Weinstock, 1975; Gurevich \textit{et al.}, 1985) as well as energetic electrons coming from the tail of a Maxwellian velocity distribution of HF enhanced electron temperature (Mantas, 1994; Mantas and Carlson, 1996; Gurevich and Mrlikh, 1997). Mishin \textit{et al.} (2000) argue that the HF pumped electron energy distribution is not Maxwellian due to the interaction of N\textsubscript{2} regardless of the accelerating mechanism.

At auroral latitudes, an initial attempt to produce an artificial aurora (Stubbe \textit{et al.},
1982) was unsubstantiated by simultaneous independent observations (Henriksen et al., 1984). The first unambiguous high-latitude artificial aurora observations of the O(\(^{1}\)D) emission have been made more recently at Tromsø on 16 February 1999 (Brändström et al., 1999) and 21 February 1999 (Kosch et al., 2000a), using the EISCAT HF facility (Rietveld et al., 1993) in northern Scandinavia, and on 18 March 1999 (Pedersen and Carlson, 2001) using the HAARP HF facility in Alaska. For the first observation, the high latitude artificial aurora is associated with very large electron temperature enhancements amounting to approximately 2500 K above background (Leyser et al., 2000; Gustavsson et al., 2001). Unfortunately, incoherent scatter measurements are unavailable for the other high-latitude artificial auroras published to date. In this paper, we investigate all additional observations associated with the 21 February 1999 artificial aurora. This includes an altitude estimate of the optical emission, HF coherent radar and dynasonde backscatter, F-region ion drifts and neutral winds, and ray tracing of the pump wave.

2. Observations and discussion

On 21 February 1999, the EISCAT HF facility was operated in O-mode at 4.04 MHz with an effective radiated power of 73 MW in the vertical direction. The modulation was a sequence of 4-min on, 4-min off from 1648 to 1828 UT. A co-located advanced ionosonde (dynasonde) made soundings every 4 min from which the real reflection height of the pump wave was obtained using the true-height inversion program POLAN (Titheridge, 1967). The upper hybrid resonance height, which in our case is the height at which the plasma frequency is 240 kHz below the pump wave reflection height, was then computed. The directions of arrival and Doppler shifts of dynasonde backscatter were used to estimate F-region ion drift velocities (Wright and Pitteway, 1994). The CUTLASS pair of radars (Milan et al., 1997), which are part of the SuperDARN network of similar radars (Greenwald et al., 1995), recorded the artificial backscatter generated by HF O-mode pumping (Robinson et al., 1997). In this study, CUTLASS used a 2-min beam scan with 45 km range resolution and operated at 12.4 MHz prior to 18 UT and at 10 MHz thereafter. However, only the Hankasalmi radar, which views predominantly poleward, received backscatter. The Digital All-Sky Imager (DASI) (Kosch et al., 1998) was operated viewing in the zenith above Skibotn (69.35'N, 20.36'E), Norway, which is approximately 50 km south and east of the EISCAT HF facility (69.59'N, 19.23'E). Images were taken using a 116' field-of-view lens, a 630 nm narrowband interference filter and 10 s integration. A Fabry-Perot interferometer located at Skibotn (Kosch et al., 1997) was used to estimate F-region neutral winds (Kosch et al., 2000b). The instrument scans through north, south, east and west at 45° elevation as well as the vertical direction. Geomagnetic conditions were very quiet on 21 February 1999 with \( K_p = 0^* \) during the period of interest, preceded by 9 hours of \( K_p = 0^* \).

A patch of artificial aurora started to appear \(~10\) s after HF switch on, developing to a maximum peak intensity of about 100 Rayleighs above background within \(~60\) s. When the HF pump was switched off the emission decayed within \(~30\) s in agreement with the observations of Brändström et al. (1999). Pedersen and Carlson (2001) also found the rise time to be longer than the decay time. Panel A of Fig. 1 shows the time sequence of optical intensity, spatially averaged over the entire optical emission, in
Rayleighs for 630 nm. The data have been background subtracted, using images during no HF pumping, to remove scattered light from the setting sun. Negative intensities arise from noise in the background subtraction. The square wave represents the HF pump on/off cycle. In the interval 1648–1812 UT, there is a clear correlation between the transmitter being turned on (off) and the optical intensity rapidly increasing (decreasing). The data prior to 1704 UT is badly contaminated by a bright background sky and is not analysed further. The optical emission prior to 1648 UT is caused by phasing up of the HF
transmitters. Panel B of Fig. 1 shows the pump wave reflection altitude (solid line), which ascended from 210 to 290 km throughout the experiment. The estimated error is up to ±10 km caused by the uncertainty in the true-height inversion process and ionogram contamination by the HF pump wave itself.

Figure 2 shows a false-colour sequence of 9 DASI images taken 8 min apart, between 1704 and 1812 UT, each one being the last 10 s integration taken before the end of a 4-min HF on interval. North and east are to the top and right of each image, respectively, and the start time of each integration is overlaid in the images. The white circle shows the modelled HF beam for the −3 dB locus. Assuming no refraction, the effective beam diameter corresponds to ∼15°, which is equivalent to a 65 km diameter at 250 km altitude. The black square outside the circle marks the intersection of the magnetic field line (12.8° zenith angle at 183.3° azimuth) through the HF facility with the pump wave reflection height. The black square inside the circle marks the HF facility’s location. A line joining both squares corresponds to the magnetic meridian direction over EISCAT.

Figure 2 is similar to Fig. 2 of Kosch et al. (2000a) where the HF beam −3 dB locus was mapped to the pump wave reflection altitude. A clear equatorward shift of the optical emission compared to the HF beam projection was seen. In addition, these data showed a small but systematic westward shift of the optical emission compared to the projection of the HF beam, which is noted here for the first time. The westward displacement cannot be explained by refraction of the pump beam because the setting sun would ensure a descending electron density gradient from west to east, which would cause
an eastward refraction of the pump beam. Pedersen and Carlson (2001), whose optical imager was co-located with the HAARP HF facility, also observed the equatorward shift of the optical emission when pumping in the local zenith. However, most importantly, they found the optical emission to be exactly east-west symmetric across the magnetic meridian over the HF facility. This is confirmed by Gustavsson et al. (2001), in their Fig. 6, for 16 February 1999. Since DASI is not co-located with the EISCAT HF facility (50 km separation), and assuming geomagnetic east-west symmetry of the artificial aurora over EISCAT, it is possible to perform “one-legged” triangulation using the DASI images.

Figure 2 shows the HF beam mapped so that the optical emission is symmetrically located on the magnetic meridian over EISCAT. In order to achieve this, the projection altitude had to be reduced from the pump wave reflection altitude by the amount indicated in each image (20–70 km). Panel B of Fig. 1 shows the altitude of the artificial aurora (dashed line), which varied from 180 to 225 km±5 km. With the exception of the first data point at 1708 UT, the height of the artificial aurora is remains remarkably constant (205–225 km) despite the steadily increasing pump wave reflection altitude. Using multi-station observations, Gustavsson et al. (2001) found the maximum optical emission to be 5–25 km below the pump wave reflection altitude. Model calculations by Bernhardt et al. (1989a) found the optical emission altitude to be ~0–25 km below the accelerated electron source region for HF interaction altitudes of 250–300 km. Mid-latitude altitude estimates by triangulation (Haslett and Megill, 1974) show that the optical emission was limited in height to 280±15 km although the pump wave reflection height could be up to 30 km higher. This is consistent with Bernhardt’s et al. (1989a) model. The direction of these results are also consistent with our observations, except that we estimate the vertical displacement to be somewhat greater with an average of ~40 km.

A potential problem with the accuracy of the “one-legged” altitude triangulation is indicated by comparing the radiative lifetime of the O(\(^D\)) emission with theoretical curves by Gustavsson et al. (2001). Their Fig. 8 shows that, for altitudes of 180–225 km, the radiative lifetime of O(\(^D\)) should be ~10–30 s. For 21 February 1999, the mean radiative lifetime is ~35 s. Ignoring the first altitude estimate (180 km), which may be contaminated by scattered sunlight, Gustavsson’s et al. (2001) theoretical estimate becomes ~15–30 s, which is still significantly less than that measured. Uncertainties in the “one-legged” altitude estimate includes errors in inverting the dynasonde data, mentioned earlier, and aligning the DASI images of the artificial aurora, which are noisy and diffuse, onto the EISCAT magnetic meridian. We conclude that our estimates probably constitute the lower bound for the altitude of the artificial aurora.

The large vertical displacement between the artificial aurora and the estimated pump wave reflection altitude is unexpected. If the mechanism of electron acceleration is Langmuir turbulence, then this would maximise at the pump wave reflection altitude. Leyser et al. (2000) suggested that upper hybrid turbulence could be a mechanism for electron acceleration. Panel B of Fig.1 shows the estimated upper hybrid resonance (UHR) height (dotted line), which is 3–8 km below the pump wave reflection altitude. However, the optical emission is also well below the UHR height. It is possible that the energised electrons must travel to lower altitudes where the oxygen density is greater in order to produce an observable optical emission. This situation would be very similar to the natural aurora where the maximum brightness typically occurs near the bottom of the
auroral forms. In the case of the artificial aurora some evidence of vertical extent in the optical emission should be apparent (e.g. ray structures). To date, this is not apparent for both the 16 and 21 February 1999 data sets, where off axis optical recordings were made. Unfortunately, no electron temperature data are available for 21 February 1999. Gustavsson et al. (2001) show electron temperature altitude profiles for 16 February 1999. It seems clear that the electron temperature maximum is close to the pump wave reflection altitude (220–280 km) and temperature enhancements occur at all higher altitudes to beyond 500 km. Unfortunately, the 22 km range resolution makes it difficult to investigate the relationship between electron temperature and optical emission altitude as the artificial aurora appeared only 5–25 km below the pump wave reflection altitude on 16 February 1999. The same difficulty applies when comparing the electron temperature enhancements to the UHR height.

The artificial aurora shown in Fig. 2 is clearly equatorward of the projected pump beam. It should be noted that these images represent steady state conditions as each image is the last 10 s integration of a 4-min HF pump on cycle. Kosch et al. (2000a) showed that the displacement is not due to a geometric effect caused by the equatorward tilt of the magnetic field. Pedersen and Carlson (2001) suggested that the F-region ion drift might explain the displacement because the O(1D) photon is a forbidden transition and is only emitted ∼110 s after energisation of O in free space. Panel C of Fig. 1 shows the F-region ion drift extracted from dynasonde data. The meridional and zonal components are shown as the solid and dashed lines, respectively, with error bars. North and east are defined as positive. For the period of interest, the meridional and zonal velocities are less than 150 and 400 m/s, respectively. During quiet geomagnetic conditions, which prevailed on 21 February 1999 ($K_p=1$), the dynasonde produces accurate estimates of horizontal ion drift (Sedegmore et al., 1998, and references therein). The velocities are consistent with CUTLASS radar Doppler shift data (not shown) (Eglitis et al., 1998), which show the mainly meridional velocity to be less than 50 m/s. Gustavsson et al. (2001) have shown that the effective lifetime of O(1D) is less than 50 s below 250 km due to quenching by N$_2$, O$_2$ and O. For the lowest measured altitude of the artificial aurora (180 km), the horizontal displacement between the zenith and magnetic field line directions is 40 km. For an O(1D) lifetime of 50 s and the maximum equatorward ion drift of 150 m/s, ion transport accounts for a possible maximum 7.5 km equatorward displacement. Using the maximum but less realistic lifetime of 110 s yields a maximum displacement of 16.5 km. Clearly, ion drift does not account for the displacement of the optical emission.

Pedersen and Carlson (2001) also suggested that the F-region neutral wind might explain the displacement of the artificial aurora. Panel D of Fig. 1 shows the thermospheric winds from Fabry-Perot interferometer measurements. The meridional and zonal components are shown as the solid and dashed lines, respectively, with error bars. North and east are defined as positive. The pairs of curves arise because any component of the neutral wind can be estimated twice by scanning the interferometer to opposite look directions. For the period of interest, the meridional and zonal velocities are less than 150 and 120 m/s, respectively. The same reasoning applies here as for the ion drifts. Hence, the neutral winds also cannot account for the displacement of the optical emission. This result is consistent with the observations of 16 February 1999 (Gustavsson et al., 2001).

We conclude that the displacement is probably an important clue to the mechanism of
artificial aurora generation at high latitudes. This conclusion is reinforced by the fact that the
dynasonde sky plots, which show the direction of arrival of backscatter, often show a
distinct equatorward displacement and appear to cluster around the magnetic field line
direction. On 21 February 1999, dynasonde backscatter came from a region up to 100 km
in diameter displaced ~50 km equatorwards of zenith (not shown). This is entirely
consistent with the displacement of the artificial aurora. The reason for this phenomenon
remains unknown (J.W. Wright, private communication).

Bernhardt et al. (2000) found that the motion of the optical emission was a good
indicator of F-region ion drift and neutral wind at low latitudes. Earlier we explained the
small westward displacement of the artificial aurora from the magnetic meridian over the
HF facility in terms of the reduced altitude of the optical emission. However, a westward
ion drift would produce a similar displacement due to the long lifetime of O(\text{D}). Kosch
et al. (2000a) found no zonal drift motion of the optical emission for 21 February 2001.
In addition, Fig. 2 in Kosch et al. (2000a) shows the westward displacement to always be
present whereas Panels C and D of Fig. 1 shows the zonal ion drift and neutral wind,
respectively, to be both eastward and westward at different times. We conclude that the
altitude of the artificial aurora is below the pump wave reflection altitude and the westward
displacement of the optical emission is not due to zonal ion drift or neutral wind. A
future observation of artificial O(\text{S}) aurora will provide definitive evidence as the lifetime
of the 557.7 nm emission is only ~0.7 s.

Panel E of Fig. 1 shows a latitudinal keogram of backscatter power from the
CUTLASS (Hankasalmi) coherent radar. The square wave represents the HF pump on/off
cycle. The dashed line shows EISCAT’s location. The radar’s over-the-horizon ray
path has been corrected for refraction using the ionospheric electron density profile taken
from the dynasonde. This results in a mapping accuracy of less than 1 range gate (45 km)
(Yeoman et al., 2001). There are clear and large enhancements in the backscatter power,
up to 40 dB, during HF pump on periods centered on EISCAT’s latitude. Natural
backscatter exists poleward of EISCAT, which descends in latitude towards the end of the
experiment.

Figure 3 shows the spatial extent of the CUTLASS coherent backscatter power for 15
HF pump on periods. The cross shows the location of the EISCAT HF facility. The
central 9 panels correspond to the optical images shown in Fig. 2. The last 3 panels show
the natural backscatter encroaching from the north. The area covered in each panel is
400 × 400 km, whereas the artificial aurora has a diameter of ~50 km. Hence, it is clear
that the optical emission is very much more localised than the region of artificial backscatter.
Coherent HF backscatter is produced by field-aligned irregularities (FAI), which are
easily stimulated by HF O-mode pumping via resonant mode conversion to electrostatic
waves (Robinson, 1989). Mode conversion occurs efficiently at the UHR height and
artificial FAI are a common feature of HF O-mode pumping (Robinson et al., 1998; Bond
et al., 1997). Electron temperature enhancements, as seen on 16 February 1999 (Leyser et
al., 2000), are also the final result of the conversion of HF waves into plasma waves and
are correlated with FAI and anomalous absorption (Robinson et al., 1997, 1998). Only
about 5% of maximum pump power (~10 MW) is needed to produce FAI (Wright et al.,
2000) whereas the artificial aurora has proven difficult to stimulate even with maximum
power. The spectral width of the CUTLASS backscatter (not shown) is less than 25 m/s,
which is relatively narrow. This is consistent with the pump wave enhanced FAI being an artificial hard target (Eglitis et al., 1998).

Figure 4 shows the ray tracing diagram for a vertically directed 4.04 MHz O-mode pump wave corresponding to the ionospheric conditions prevailing at 1732 UT on 21 February 1999 assuming a horizontally uniform ionosphere. The calculated ray paths are shown for incident angles to the zenith in 3° steps from 15° south to 12° north. The pump wave reflection altitude (dotted line) for 4.04 MHz is 244.9 km and the UHR height (dashed line) is 240.5 km. The orientation of the pump wave E-vector is shown by the arrows for the UHR height. The direction of the magnetic field line passing through the HF facility is shown by the dash-dotted line. The ionospheric electron density profile is shown by the dotted curve and relates to the upper x-axis, which is normalised to the pump
frequency of 4.04 MHz. The critical frequency was 4.8 MHz. All rays within the Spitze cone (±6° at Tromsø) reach the maximum reflection altitude, as expected. By definition, the Spitze angle is the critical zenith angle for which an O-mode wave still reflects at the maximum altitude (Rietveld et al., 1993). For greater zenith angles the O-mode reflection altitude decreases, which is generally expected to reduce the occurrence of artificially stimulated plasma instabilities. Leyser et al. (2000) suggested that upper hybrid turbulence might be the mechanism for producing artificial aurora. Maximum coupling between the pump electromagnetic wave and upper hybrid electrostatic waves occurs when the pump wave electric field vector is perpendicular to the magnetic field line direction. This occurs at the UHR height approximately for the 6° southward ray upward leg and the 9° northward ray downward leg. We note that the 6° southward ray corresponds to the observed position of the artificial aurora. However, no optical emission is observed for the 9° northward ray position. This may be due to the fact that the northward ray has undergone reflection before meeting the condition for optimum coupling and will have lost strength due to absorption and conversion to other wave modes at reflection (e.g. Langmuir waves). In addition, the 9° northward ray is more than 3 dB weaker compared to the zenith (pump beam width ≈15°). Hence, this ray is probably too weak to produce an observable optical emission. This reasoning will be tested in future by scanning the pump beam poleward.

The ray tracing and artificial aurora displacement are both consistent with the observed CUTLASS backscatter. The FAI, which coherent HF radars need for backscatter, are stimulated by electrostatic waves, which are most efficiently generated at the UHR
height where the pump wave electric field vector is perpendicular to the magnetic field. This condition is met for the ray path approximately 6° south of zenith. The large area of FAI produced by O-mode pumping is due to the fact that a significant component of the pump wave electric field vector is perpendicular to the magnetic field over a large area at the UHR height. The very much smaller area of the optical emission implies that only the most efficient coupling will produce an observable artificial aurora. The ray tracing, artificial aurora displacement and CUTLASS backscatter all provide evidence that the mechanism of electron acceleration, leading to the HF induced optical emission, may be linked to upper hybrid turbulence.

3. Conclusions

High-latitude HF induced artificial aurorae appear displaced equatorward towards the magnetic field line direction despite pumping in the local zenith. This effect was first noted on 21 February 1999 by Kosch et al. (2000a), was confirmed on 18 March 1999 (Pedersen and Carlson, 2001), and can be seen in the data from 16 February 1999 (Gustavsson et al., 2001). On 21 February 1999, equatorward meridional F-region ion drifts and neutral winds do not exceed 150 m/s and hence cannot explain the displacement of the optical emission. Dynasonde backscatter also comes predominantly from the region where the optical emission is observed. The reason for this is unknown. Large HF coherent radar backscatter enhancements are associated with the production of artificial aurora. Ray tracing suggests that the optical displacement may be related to the mechanism for producing field-aligned irregularities, which the HF coherent radars require for receiving backscatter. This would be consistent with upper hybrid turbulence as a possible mechanism for the artificial aurora. However, details of the mechanism for particle acceleration, which is necessary for production of the optical emission, remain unclear. In addition, the altitude of the artificial aurora is shown to be well below the HF reflection altitude and upper hybrid resonance height of the pump wave.

Acknowledgments

EISCAT is an international scientific association supported by the research councils of Finland, France, Germany, Japan, Norway, Sweden and the United Kingdom. The authors thank P. Gallop for implementing software by J.W. Wright for extracting ion drift velocities from dynasonde data, C. Davis for running the true-height inversion program and H. Kohl for the ray tracing program.

The editor thanks Drs. C. La Hoz and B. Gustavsson for their help in evaluating this paper.

References


Kosch, M.J., Ishii, M., Nozawa, S., Rees, D., Cierpka, K., Kohsiek, A., Schlegel, K., Fuji, R., Hagfors,


(Received December 18, 2001; Revised manuscript accepted April 30, 2002)
Evaluation of the region 1 field-aligned current 
from the low-latitude boundary layer 
using the 1989 Tsyganenko model

Takashi Yamamoto¹, Masao Ozaki² and Shoshi Inoue³

¹Department of Earth and Planetary Science, University of Tokyo, 3-1, Hongo 7-chome, Bunkyo-ku, Tokyo 113-0033
²Institute of Industrial Science, University of Tokyo, 6-1, Komaba 4-chome, Meguro-ku, Tokyo 153-8505
³Faculty of Education, Iwate University, Morioka, Iwate 020-8550

Abstract: On the basis of the 1989 Tsyganenko magnetic field model, the field-aligned current (FAC) density distribution in the low-latitude boundary layer (LLBL) is numerically evaluated by assuming the entry of magnetosheath particles into the LLBL. The calculated FAC density/intensity profiles are consistent with observations. This FAC generation arises from the divergence of the magnetic drift current density carried by LLBL particles, and the current intensity is then increased with their kinetic energy density. The FAC generation occurs at the inner edge of the LLBL whenever sheath particles penetrate into the magnetosphere, regardless of the entry process. It is also emphasized that the LLBL FAC production is an inevitable consequence of the formation of the magnetopause, because the magnetopause currents act to shield the magnetic field originating from the currents inside the closed region of the magnetosphere, causing the LLBL inner edge to intersect the magnetic drift paths. This simple situation can be illustrated by calculating the distribution of the flux tube volume in the Tsyganenko model.

1. Introduction

Previously, the region 1 current was thought to be connected to a single topological region in the magnetosphere. Using ground-based magnetometer chain measurements, however, Friis-Christensen and Lassen (1991) suggested that the region 1 current actually consists of two separate regions, one (region 1a) connected to the low-latitude boundary layer (LLBL) and one (region 1b) connected to the plasma sheet. Simultaneous particle and magnetic field measurements have shown that the region 1a FAC can be subdivided into the cusp and LLBL parts (Newell et al., 1991; Yamauchi et al., 1993). Moreover, careful observations have established that plasma particles in the LLBL are responsible for the generation of a major portion of the dayside region 1 field-aligned current (FAC) (e.g., Bythrow et al., 1981, 1988; Newell et al., 1991; Woch and Lundin, 1993), while a portion of it may be sustained by particles in the plasma sheet (e.g., Liou et al., 1999).

The viscous interaction between the solar wind and the magnetospheric plasma (Axford and Hines, 1961) had long been believed to be the most likely mechanism for the
FAC generation in the LLBL. The viscous interaction model assumes that the sheared plasma flow in the LLBL produces negative and positive space charges in the postnoon and prenoon sectors, respectively. Coupling to the ionosphere might produce FAC flows away from the polar ionosphere to the boundary layer in the postnoon sector and vice versa in the prenoon sector (e.g., Sonnerup, 1980; Lotko et al., 1987). In a simple model by Sonnerup, the FAC density (intensity) is proportional to the plasma bulk flow speed in the magnetosheath. On the other hand, Yang et al. (1994) proposed a theoretical model for the pressure-gradient-driven FACs in the LLBL. Moreover, Yamamoto et al. (1995) has shown that the required (for FAC generation) inclination of the gradient/curvature drift direction relative to the LLBL inner edge, as was predicted by Yang et al., actually exists in the LLBL region of the Tsyganenko (1989) magnetic field model (hereafter referred to as T89). This fact supports the pressure-gradient-driven mechanism for the generation of LLBL FACs. A number of observations favor the pressure-driven mechanism rather than the viscous interaction model, and this point will be discussed in the subsequent review of LLBL current/particle observations.

Iijima and Potemra (1982) examined the relationship between hourly values of the solar wind density, speed and the interplanetary magnetic field (IMF) and the densities of region 1 FACs in the dayside (0800–1600) MLT sectors. The correlation between the current density and the solar wind speed was shown to be poor, suggesting that the viscous interaction is not likely to play a major role in the LLBL FAC generation. The observation by the Viking satellite (Woch et al., 1993) showed that under northward IMF conditions, the region 1 FAC intensity in the 0400–1030 MLT range is positively correlated with the LLBL ion energy density. Figure 4 of their paper is reproduced as Fig. 1 in the present paper. Assuming that a possible contribution from the plasma sheet to the current intensity is represented by an intercept of the regression line fitted to the data points, the
FAC intensity solely from the LLBL is inferred to be directly proportional to the LLBL ion energy density. This observation then supports the pressure-driven mechanism for the LLBL FAC generation.

Furthermore, Viking observations (Potemra et al., 1987; Bythrow et al., 1987) have revealed a close relation between LLBL particles and FACs, namely that an intense region I current localized in an extremely narrow (<0.1° in invariant latitude) zone is seen near the inner (equatorward) edge of the LLBL particle population. Plate I of Potemra et al. and Fig. 2 of Bythrow et al. are reproduced as Figs. 2 and 3 in the present paper, respectively. Figure 3 is to notify that such a fine current structure cannot be identified by measurements from the low-altitude satellite DMSP, owing to their low spatial resolution.

\[ \text{Fig. 2. Viking observations of the east-west component of the magnetic field (top), the energy-time spectrogram of positive ions (bottom), and the particle pitch angle with 0° denoting precipitating particles (very bottom) (adapted from Plate I of Potemra et al. (1987)). The ion spectrogram also shows ion cones characterized by the “tuning fork” signatures (Gurgiolo and Burch, 1982) beginning at about 2131:30 UT.} \]

\[ \text{Fig. 3. Magnetic field traces from the Viking (top) and DMSP-F7 (bottom) satellites for the same event as Fig. 2 (adapted from Fig. 2 of Bythrow et al. (1987)). The spatial resolution (in invariant latitude) of the Viking data is one order of magnitude higher than the DMSP data.} \]
Fig. 4. (a) Example of (6-s average) data of east-west magnetic field in the LLBL, measured from the DE-1 satellite, showing that the total region 1 current intensity is substantially maintained by an (apparently) single current sheet of narrow width. The magnetic field is calibrated to that at an altitude of 800 km. The spatial resolution of this data is about 0.1' in invariant latitude. (b) Variations of the IMF GSM components as well as the solar wind speed, density and dynamic pressure, which were all 5-min averages and measured from the IMP 8 satellite for a four-hour period just prior to the DE-1 observation of the current structure in (a).

(although these observations from the two spacecraft were not at the same universal and local times.) In the current system in Fig. 2 or 3, downward (region 1) currents with even smaller densities can also be seen inside the region of LLBL particle population, contributing to the total region 1 current intensity. Cases that the total region 1 current intensity
is substantially maintained by an (apparently) single current sheet of narrow width can be found in the (6-s average) data of magnetic field perturbations measured from the Dynamics Explorer-1 (DE-1) satellite. An example of such a current structure in the LLBL is shown in Fig. 4a. The width of the region 1 current is about 0.4' in invariant latitude, even narrower than its average value of 1–2' (Iijima and Potemra, 1978). Simultaneous particle data is not available for this particular case. For one other case of a similar current structure (the search for such current structures is ongoing), the current region is found to be roughly collocated with the LLBL inner edge, although the resolution of available particle data is about 1'. Needless to say, such a spatial relation between the region 1 FAC and the LLBL inner edge is just a prediction from the pressure-gradient-driven model. Note that Lotko and Sonnerup (1995) have opposed this model, partly because FACs generated in the pressure-driven model would flow in $\delta$-function sheets at the inner edge of a crosswise-uniform (LLBL) layer. As shown in Figs. 2 and 4a, however, the FAC does occasionally flow in a $\phi$-function-like sheet. Notably, the region 1 current with an average width of 1–2' has been commonly categorized as the large-scale one, but it sometimes exhibits a signature of localization to a narrow (<0.5') latitudinal zone. (For such examples for the nightside region 1 currents, see Yamamoto et al. (1999)). While physical reasons for the variations in LLBL FAC width will be discussed in Section 5, a precondition for the appearance of a single region 1 current structure of narrow width is assumed to be steadiness of the solar wind, i.e., steady solar wind speed, density and magnetic field. In general, a simplest current structure created in a steady environment surrounding the magnetosphere is expected to provide a clue to the FAC generation mechanism.

The purpose of the present paper is to numerically evaluate the LLBL FAC density/intensity using the T89 magnetic field model, assuming that magnetosheath particles penetrate into the closed magnetosphere to form the LLBL. Importantly, the charge separation responsible for the FAC generation necessarily occurs in the LLBL, which can be intuitively understood from illustration of the direction of average magnetic drift (gradient B drift plus curvature drift) velocity of particles. Conditions that control the thickness (in latitude) of an FAC zone in the LLBL will also be discussed.

2. Model of LLBL

In this section, a simple model of the low-latitude boundary layer is presented on the basis of the T89 model. The $K_p=0$ version of T89 with a nontilted geodipole is used throughout this work unless otherwise stated. While the Tsyganenko models have been used to describe the magnetic field distribution, various plasma regimes such as the LLBL and the plasma sheet, observationally distinguished in the magnetosphere, are not specified in them. In general, those plasma regimes in the Tsyganenko models may then be identified by comparing the ionospheric foot points of their field lines with the distributions of precipitating particles observed from low-altitude satellites. Since the LLBL is adjacent to the magnetopause, the latter should first be determined in T89. According to Kaufmann et al. (1993), the magnetopause (in T89) is defined as the outer boundary of all the closed field lines. (To describe the LLBL, the open-closed boundary crossing the distant tail need not be determined.) To conform with the precipitation patterns of LLBL
particles, statistically determined by Newell and Meng (1994), the inner edge of the LLBL in T89 is defined, on the equatorial plane, as follows. The distance, $\Delta$ (in units of the Earth’s radius, $R_E$), between a point on the inner edge and the magnetopause boundary is given by:

$$\Delta(X) = 0.25 \cdot (2 - X/X_{sub}), \quad \text{for} \quad 0 \leq X \leq X_{sub}$$

$$\Delta(X) = 0.5, \quad \text{for} \quad X > 0$$

(1)

(2)

where the $X$-coordinate is specified on the magnetopause, and $X_{sub}$ is for the subsolar point on it. (In this paper, the solar-magnetospheric coordinates ($X$, $Y$, $Z$) are used.) This profile is consistent with a finding by Mitchell et al. (1987) that the thickness of the LLBL tends to increase with distance from the subsolar point. Figure 5a shows the magnetopause boundary and the above-defined LLBL inner edge on the equatorial plane. (Due to the symmetry of the model field with respect to the $X$-$Z$ plane, various quantities are only shown on the duskside, except in Fig. 7, while the region 1 FAC direction is reversed on the dawnside.) The ionospheric projection of the LLBL inner edge is shown in Fig. 5b. Also indicated are ‘division lines’, each of which divides the LLBL region according to its relative (to $\Delta$) distance from the magnetopause. For example, 1/2 line is located at a distance of $\Delta(X)/2$ from the magnetopause. In T89, the ionospheric projection of the magnetopause (defined as the open/closed boundary) seems to converge to a single point or a short line segment as the computational resolution is increased sufficiently. Namely, in this model the dayside cusp, if defined to be a region of open field lines, has ionospheric projection of negligible dimension. In the present analysis, the ‘near-cusp’ region is formally defined, on the equatorial plane, to be a region between 1/32 line and the magnetopause boundary. The generation of FACs in the open cusp region (e.g., Erlandson et al., 1988; Taguchi et al., 1993; Yamauchi et al., 1993) is beyond the scope of the present paper.

Next, the particle (energy) distribution in the LLBL region is modeled. To this end, the orthogonal curvilinear coordinates ($\xi$, $\eta$) on the equatorial plane are introduced as illustrated in Fig. 5a. The $\xi$-axis is taken to be the magnetopause boundary; the $\xi$-distance is measured from the subsolar point. The $\eta$-axis is locally (at a given $\xi$) defined to be perpendicular to the $\xi$-axis and directed inward; the $\eta$-distance is measured from the $\xi$-axis, i.e., the magnetopause boundary, and it is normalized by the LLBL thickness $\Delta(\xi)$. The particle (energy) distribution is given in terms of the flux tube energy content, $e$, which is defined as the total kinetic energy of plasma particles contained in a flux tube with unit cross-sectional area at the ionospheric height:

$$e = \int_z^\infty \frac{3p}{2} \frac{B_i}{B(s)} ds,$$

(3)

where $p$ is the isotropic plasma pressure, $s$ is the field-aligned distance, and $s_e$ and $s_i$ are the distances to the equator and the ionospheric height, respectively; $B(s)$ and $B_i$ are the magnetic field intensities at distances $s$ and $s_i$, respectively. Assuming that the average energies of protons and electrons, $W^p$ and $W^e$, are constant along field lines, $e$ is written as

$$e = (W^p + W^e) N,$$

(4)
Fig. 5. (a) Magnetopause boundary and LLBL inner edge on the equatorial plane, which are defined using T89, in Section 2. The $\xi$- and $\eta$-axes are indicated. (b) Solid curves labeled $1, 1/2, 1/4, 1/8, 1/16$ and $1/32$ show the ionospheric projections of the division lines at $\eta=1, 1/2, 1/4, 1/8, 1/16$ and $1/32$ on the equatorial plane, respectively. The tailward limits of these lines are determined from the condition that the adiabaticity parameter, $\kappa$, is equal to unity. (For an explanation of this condition, see Section 3.)
where $N$ is the flux tube content defined as follows:

$$N = \int_{s_0}^{s_e} n \frac{B_i}{B(s)} \, ds,$$

where $n$ is the number density of protons or electrons. This equation means that $N$ is the number of particles in a flux tube with unit (ionospheric) cross section. If $n$ is constant along field lines, $N$ is written as $nR_b$, where $R_b$ is the flux tube volume defined as

$$R_b = \int_{s_0}^{s_e} \frac{B_i}{B(s)} \, ds.$$

The model distribution of $\varepsilon$ in the LLBL regions is expressed as a function of $\xi$ and $\eta$:

$$\varepsilon = \varepsilon_0 f(\xi) g(\eta),$$

$$f(\xi) = \left( R_{s0}(\xi)/R_{s0}(\xi_0) \right)^{1/2},$$

$$g(\eta) = 1,$$

$$g(\eta) = \frac{1}{a} \left( b - \tanh \left( \frac{\eta - 1/2}{\delta \eta} \right) \right), \quad \text{for} \quad 0 \leq \eta \leq \eta_0$$

and

$$\left( b - \tanh \left( \frac{\eta - 1/2}{\delta \eta} \right) \right), \quad \text{for} \quad \eta_0 \leq \eta \leq 1.0$$

where $\eta_0$ is fixed at 1/32. In eq. (8), $R_{s0}(\xi)$ is evaluated on $\eta_0$-line, and $\xi_0$ is a value of $\xi = 21.7$ R$_B$, for a point at $X = 0$ on it. Accordingly, $\varepsilon_0$ is the flux tube energy content at $\xi = \xi_0$ in the near-cusp region of $0 \leq \eta \leq \eta_0$. Parameter $\gamma$ is constant (in space), and it is to control the azimuthal ($\xi$-) dependence of $\varepsilon$. A set of eqs. (7)-(9) on $\eta_0$-line is reduced to the standard form of an adiabatic equation of state, i.e., $pR_0 = const$. Physical determination of $\gamma$ is beyond the scope of the present study, whereas this may be involved in the problem of the formation of LLBL or the mechanism by which magnetosheath particles enter it. Fortunately, however, the FAC generation in the LLBL region is essentially attributed to the gradient of $\partial \varepsilon / \partial \eta$, not significantly affected by $\partial \varepsilon / \partial \xi$ (i.e., the value of $\gamma$), so long as $\gamma \geq 0$ as assumed theoretically. This point will be discussed further in the next section. In eq. (10), $\delta \eta$ represents the characteristic length for an energy density falloff, and $a$ and $b$ are determined so that $g(1) = 0$ and $g(\eta_0) = 1$. For $\delta \eta = 1/4$ (which is assumed hereafter), the function $g(\eta)$ is illustrated in Fig. 6, where $b = \tanh 2$ and $a = b - \tanh \left( 4 \cdot (1/32 - 1/2) \right)$.

In the present model, the LLBL is assumed to be closed. However, whether the LLBL is on closed or open field lines have been controversial (e.g., Onsager et al., 1993; Moen et al., 1996; Lockwood, 1997). Lockwood has predicted that the low-altitude precipitation of a mixture of magnetosheath and magnetospheric ions, as assumed to be a signature of the LLBL, can be identified in an open region in the noon sector. He did not argue, however, the entire openness of the LLBL itself at low magnetic latitudes. From direct (in situ) observations of the LLBL, most of the LLBL (not its ionospheric projection nor the region of particle precipitation) is inferred to be on closed field lines even when the IMF $B_Z < 0$. For example, Fig. 1 of Song et al. (1993) shows that the number density in the boundary layer, except at its outer edge, is only a few percent of the density in the sheath transition layer (just outside of the boundary layer). This fact cannot be understood if the boundary layer is open. This is because a considerable amount of magnetosheath particles can flow along the open field lines into the boundary layer. The closed LLBL is basically in conformity with a generally accepted view that the LLBL is
formed regardless of the IMF and the occurrence of the magnetic reconnection at the dayside magnetopause. The reconnection will act to open only an outermost part of the LLBL.

3. FAC generation in LLBL

For the isotropic plasma pressure uniform along the field lines, the (quasi-steady) FAC density $J_\parallel$ at the ionospheric height is given by the following equation (Yamamoto et al., 1996; hereafter referred to as Y96):

$$J_\parallel = e \tilde{v}_{m,i} \cdot \nabla \varepsilon,$$

(11)

where $e$ (\(>0\)) is the electronic charge, a positive value of $J_\parallel$ corresponds to an FAC flowing away from the ionosphere, i.e., an upward FAC, and $\nabla$, denotes the gradient on the ionospheric plane. In this analysis, the ionospheric plane is assumed to be perpendicular to the ambient magnetic field $B_i$. The velocity $\tilde{v}_{m,i}$ is the average of the ionospheric projection of the magnetic drift velocity per unit energy, which is defined as follows:

$$\tilde{v}_{m,i} = \frac{1}{R_b} \int_{\eta_0}^{\eta} \frac{V_{m,i}(s)}{W^p} \frac{B_i}{B(s)} \, ds,$$

(12)

where $V_{m,i}(s)$ is the ionospheric projection of the magnetic drift velocity of a proton fluid with average energy $W^p$. The velocity $\tilde{v}_{m,i}$ is derived from the gradient of $R_b$ (Vasyliunas, 1970; Y96):

$$\tilde{v}_{m,i} = -\frac{2}{3e} \frac{1}{R_b B_i} \hat{b}_i \times \nabla R_b,$$

(13)

where $\hat{b}_i$ is the unit vector parallel to the ionospheric magnetic field $B_i$.

If the FAC density at the equator is defined as $J_{y,e} = (B_e / B_i) J_{y,i}$, i.e., $J_{y,i}$ multiplied by
the mirror ratio of $B_e/B_i$, where $B_e = B(s_e)$, $J_{le}$ is expressed as

$$J_{le} = e(B_e/B_i) \vartheta_{m,e} \cdot \nabla_e \varepsilon,$$  \hspace{1cm} (14)

where $\nabla_e$ denotes the gradient on the equatorial plane. When the field lines as well as the plasma distributions are symmetric with respect to the equator, the FAC density vanishes there. The current density $J_{le}$ is then defined only in a formal sense. The definition of $\vartheta_{m,e}$ (the average of the equatorial projection of the magnetic drift velocity per unit energy) and its relation to $R_b$ are given by

$$\vartheta_{m,e} = \frac{1}{R_b} \int_{s_e}^{s_i} \frac{V_{m,e}(s)}{W_p} \frac{B_i}{B(s)} \, ds,$$ \hspace{1cm} (15)

and

$$\vartheta_{m,e} = -\frac{2}{3e} \frac{1}{R_b B_e} b_e \times \nabla_e R_b,$$ \hspace{1cm} (16)

where $V_{m,e}(s)$ is the equatorial projection of the magnetic drift velocity of a proton fluid with average energy $W_p$, and $b_e$ is the unit vector parallel to the equatorial magnetic field $B_e$.

As is evident from eq. (11) or (14), the FAC direction (i.e., the polarity of charge separation) is determined by the direction of $\vartheta_{m,i(e)}$ relative to that of $\nabla_{i(e)} \varepsilon$. The direction of $\vartheta_{m,i}$ or $\vartheta_{m,e}$ is perpendicular to $\nabla_i R_b$ or $\nabla_e R_b$, respectively (see eq. (13) or (16)). Equivalently, the averages of the magnetic drifts projected to the ionospheric or equatorial plane are parallel to the equicontrous of the flux tube volume $R_b$ plotted on one plane. (Neglecting $\nabla_{i(e)} B_i$, the stream function for $\vartheta_{m,i}$ is given by $(2/3eB_i) \ln R_b$ (Y96), but $(2/3eB_i) \ln R_b$ on the equatorial plane is the stream function for $(B_e/B_i) \vartheta_{m,e}$.) An important factor for the FAC generation is then the equicontrous of $R_b$, which are shown on the equatorial plane, in Fig. 7, where the inner edge of the model LLBL (Fig. 5a) is superposed, and $R_b$ is normalized by $R_{10}$, the value of $R_b$ $(1.32 \times 10^9 \text{ m})$ for the dipole field line with an equatorial distance of $7 R_E$. In contrast with the magnetic drift direction schematically illustrated in Fig. 3 of Yang et al. (1994), all $R_b$-contours in T89 are closed inside the open-closed boundary, i.e., the magnetopause. Figure 8 shows the ‘distortion angles’ of division lines: the distortion angle at a point on one line is defined as the angle between $\vartheta_{m,e}$ and the tangent to it there. The distortion angle (at a given $X$) is greater on more inner division lines, as can be understood as a characteristic of equicontrous closed inside the magnetopause. Similarly to the region 1 FACs on the nightside plasma sheet (Y96), the LLBL FACs are also controlled by the degree of “distortion”. (This point will be again discussed below). For reference, Fig. 9 shows the magnitudes of $\vartheta_{m,i} \times 1 \text{ keV}$ plotted along various division lines mapped to the ionosphere.

A remarkable feature of the $R_b$-contours in Fig. 7 is that some of them cross the LLBL inner edge, namely $\vartheta_{m,e}$ at the LLBL inner edge has a component normal to it. The distortion angles in Fig. 8 are a manifestation of this situation. Such a feature (for all division lines) assures the generation of a region 1 FAC in the whole LLBL region, when the flux tube energy content $\varepsilon$ does not increase with $\xi$ (antisunward), i.e., $\gamma \geq 1$ in eq. (8). This is because in the scalar product of $\vartheta_{m,e} \cdot \nabla_e \varepsilon$ in eq. (14), the term of $\vartheta_{m,e} \nabla_e \varepsilon$ (product of $\xi$-components) is zero or of the same sign as the term $\vartheta_{m,e} \nabla_e \varepsilon$ (product of $\eta$-components) contributing to the generation of the region 1 FAC. Hereafter, $\vartheta_{m,e} \nabla_e \varepsilon$ and
\( \vec{\xi}_{\perp} \) and \( \vec{\xi}_{\parallel} \) are referred to as the 'perpendicular' and 'parallel' contributions to \( J_{\|e} \), respectively. Even in the case of \( \gamma < 1 \), a region I FAC can be generated so long as the increase (with \( \xi \)) of \( \varepsilon \) is sufficiently slow. Since the density and temperature of the 'source' magnetosheath population are expected to decrease with \( \xi \) (or \(-X\)) (Spreiter et al., 1966), a maximum (conceivable) gradient of \( \partial \varepsilon / \partial \xi > 0 \) might be attained when \( \varepsilon \propto R_b \), which corresponds to \( \gamma = 0 \) in eq. (8). For that case (the result not shown), the region I current intensity for \( \xi > 10 \) is relatively small, but the intensity at \( \xi = \xi_0 \) is about 63\% of that in the case of \( \gamma = 5/3 \) and the total current, i.e., the intensity integrated down to the adiabatic limit (for details, see below) is more than that for \( \gamma \geq 1 \). Therefore, almost unconditionally (only if \( \gamma \geq 0 \)), upward FACs are generated from the postnoon LLBL, implying the production of negative space charges. Similarly, downward FACs are from the prenoon LLBL, producing positive ones.

In passing, expressions for the FAC intensity and the total current are obtained for the situation that the perpendicular contribution to \( J_{\parallel e} \) dominates the parallel one. In this case, combination of eqs. (14) and (16) yields
Fig. 8. Distortion angles of division lines at $\eta = 1, 1/2, 1/4, 1/8, 1/16$ and $1/32$, plotted against $X$. (For the definition of distortion angle, see Section 3.) Also $\xi(X)$ is plotted with the ordinate on the right.

Fig. 9. Magnitudes of $\nu_{\text{med}} \times 1 \text{ keV}$ plotted along various division lines mapped to the ionosphere. The abscissa is the distance (in km) from noon. The MLT on 1.0 line is indicated on the abscissa.

$$J_{le} = \pm \frac{2}{3B_i} \frac{1}{R_b} \frac{\partial R_b}{\partial \xi} \frac{\partial \xi}{\partial (\eta_\Delta)} \xi,$$

where the plus and minus signs are for $J_{le}$ on the morning and evening sides, respectively.
Integrating the above equation along the $\eta$-axis gives the current intensity, $I_\eta(\xi)$, in the LLBL. Since $\partial \ln R_b / \partial \xi$ is nearly independent of $\eta$, $I_\eta(\xi)$ is approximately written as

$$I_\eta(\xi) \sim \frac{2}{3B_i} \xi \varepsilon(\xi, \eta_0) \left[ \frac{\partial \ln R_b}{\partial \xi} \right]_{\eta=1/2}.$$  \hfill (18)

Integrating again the above equation along the $\xi$-axis gives the total current, $I_\xi^0$:

$$I_\xi^0 \sim \frac{2}{3B_i} \int_{1/2}^{\xi_0} d\xi \varepsilon(\xi, \eta_0) \left[ \frac{\partial \ln R_b}{\partial \xi} \right]_{\eta=1/2}. \hfill (19)$$

This integration is from the sub-solar point to a point $(\xi_0, \eta_0)$ where the adiabaticity parameter (Büchner and Zelenyi, 1989), $\kappa$, is equal to unity. Here $\kappa$ is the ratio between the radius of the field-line curvature and the Larmor radius of a proton with 1 keV of energy. In a distant tail region of $\xi > \xi_0$, the proton motion is assumed nonadiabatic, i.e., the drift approximation breaks down. Paying attention to the bracketed part of the integrand in eq. (19), which depends directly on the field distribution, $I_\xi^0$ is found to be proportional to

$$\left[ \ln \frac{R_b(\xi_0)}{R_b(0)} \right]_{\eta=1/2}.$$  \hfill (20)

Analogous to the case of the generation of a nightside region 1 FAC (Yamamoto et al., 2001), quantity (20) is assumed to represent the degree of the magnetic field distortion in the LLBL. The LLBL region 1 current thus depends on the field distortion.

Regarding the causality underlying the generation of LLBL FACs, the following two possibilities are suggested. The generation of LLBL FACs may be an inevitable result of the formation of the magnetopause confining the $R_\phi$-contours on the equatorial plane, which would otherwise be distributed in a (dipole-)axially symmetric manner. Note that the magnetopause currents act to shield the magnetospheric fields so that all $R_\phi$-contours can be closed inside the open-closed boundary (see Fig. 7) where $R_b$ increases indefinitely. A second possibility is that the ultimate cause of the LLBL FAC generation is the solar wind distortion of the terrestrial magnetic field lines which would be configured symmetrically (in the same sense as described above) without the influence of the solar wind; more simply, the LLBL FAC generation is attributed ultimately to the loss of field line symmetry.

4. Numerical results

The FAC density and intensity in the LLBL region are numerically calculated using eq. (14) for the distributions of $\varepsilon$ in eqs. (7)-(10) and $\tilde{\nu}_{m,e}$ based on the T89 field. At a point $(\xi_0, \eta_0)$, the plasma density, $n_0$, is assumed to be 1.0 cm$^{-3}$, and the total average energy, $W_0 \equiv W' + W''$, is 0.5 keV. The energy content $e_0$ is then estimated as $n_0 W_0 R_b(\xi_0, \eta_0)$, where $R_b(\xi_0, \eta_0)$ is about 36.9$R_0$. It is straightforward to obtain the FAC density/intensity for other values of parameters $n_0$ and $W_0$, because of its proportionality to these parameters. Parameter $\gamma$ is taken to be 5/3 unless otherwise stated.

Figure 10 shows equicontours of the FAC density $J_\parallel$ on the ionosphere. Figure 11a shows the current intensities in the LLBL, near-cusp and whole regions, defined as
Fig. 10. Equi contours of FAC density $J_{11}$ on the ionosphere, for the flux tube energy content $\varepsilon$ in eqs. (7)-(10) with $\gamma=5/3$, $\delta_B=1/4$ and $\varepsilon_0=-W_0R_0(\xi_0, \eta_0)/1.0 \times 0.5$ (keV cm$^{-3}$) $\times 36.9R_A$. The color bar indicates the FAC density (\mu A/m$^2$).

Figs. 11a, b. (a) Current intensities in the LLBL, near-cusp and whole regions. (For the definitions of these regions, see Section 3.) (b) Current intensity in the LLBL for $\gamma=5/3$, 4/3 and 1. In (a)-(c), the $\xi$-coordinate is given on the abscissas and the position of $\xi_0$ (corresponding to $X=0$) is indicated; $\gamma$ is 5/3 in (a) and (c).
The integral expressions for the FAC densities are given as:

\[ \int_0^{\eta_0} J_{le} d\eta, \int_0^{\eta_0} J_{le}^{(e)} d\eta, \text{ and } \int_0^{\eta_0} J_{le}^{(e)} d\eta, \]

respectively. Figure 11b shows the current intensity in the LLBL for \( \gamma = 5/3, 4/3 \) and 1.0. In Fig. 11c, the FAC densities \( J_{le}(\xi, \eta) \) on various division lines, at the equator, are plotted against \( \xi \); the perpendicular and parallel contributions to \( J_{le} \) (i.e., \( \nu_{le}^{\parallel} \nabla_\perp \xi \) and \( \nu_{le}^{\parallel} \nabla_\parallel \xi \)) are also shown.

\[ \text{Fig. 11c. Current density } J_{le}(\xi, \eta) \text{ on various division lines at the equator: perpendicular and parallel contributions to } J_{le} \text{ (defined in Section 3) are also shown.} \]
From Fig. 10 it is found that the region 1 currents around 14–15 (or 9–10) MLT have large values of $|J_y|$ (e.g., $>8 \mu A/m^2$) in wider latitudinal ranges, being consistent with the statistical distribution of $J_y$ in Fig. 14 of Iijima and Potemra (1978). Figure 11a indicates that in the case of $\gamma=5/3$, the region 1 current intensity in the LLBL is relatively large in a region around $X=0$ (i.e., $\xi = \xi_0 = 21.7$ Re). For smaller values of $\gamma (\geq 1)$, the intensity distribution becomes broader in $\xi$ (see Fig. 11b), because $\xi$ decreases more slowly with $\xi$ or does not decrease. It is found that in the case of $\gamma=5/3$, the perpendicular contribution to $J_y$ dominates the parallel one in a wide $\eta$-range of $\eta > 1/4$ (see Fig. 11c). For smaller values of $\gamma (\geq 1)$, the parallel contribution becomes smaller because of smaller values of $\partial e/\partial \xi$; particularly it vanishes for $\gamma = 1$.

Finally the $K_r$-dependence of the region 1 FAC in the LLBL is briefly discussed. Numerical calculations show that the average current intensity defined as $I_y^a/\xi_a$ as well as the distortion degree, (20), are insensitive to $K_r$. (More exactly, they are not systematically changed with $K_r$.) On the other hand, the observations from the TRIAD satellite (Iijima and Potemra, 1982) have shown that the average densities of the dayside region 1 current increase with the magnitude of southward IMF component, and are better correlated with $P_{9.5}^a (B \sin(\theta/2))^{1/2}$, where $P_{9.5}$ is the solar wind dynamic pressure, $B$ the IMF strength, and $\theta$ the IMF clock angle. This fact may be explained by the fact that magnetosheath particles can be energized, depending on the interplanetary electric and magnetic fields, on the frontside magnetopause just before entering the LLBL region. This issue will be studied in a subsequent paper.

5. Variations of LLBL FAC structure

While the current structure as modelled in the previous sections is assumed to represent, in a sense, a statistically averaged one, satellite observations have shown a variety of the region 1 current structures such as a thin current sheet, multiple current sheets and a broad current zone. In this section, possible physical processes to control the thickness of a current zone (in the LLBL region) are briefly discussed.

5.1. Preconditions for a thin current sheet

As is discussed in Section 1, the DE-1 magnetic field data sometimes show the appearance of an apparently single current sheet which is localized in a narrow ($<0.5^\circ$ in invariant latitude) zone, but can substantially maintain the total region 1 current intensity on a latitudinal line crossing the LLBL region. Here, an attempt is made to identify a specific physical condition required for such a current sheet to emerge. Figure 4b shows variations of the IMF GSM components as well as the solar wind speed, density and dynamic pressure, which were measured by the IMP 8 satellite for a four-hour period just prior to the DE-1 observation of the current structure in Fig. 4a. Remarkably, every quantity in Fig. 4b remains at a nearly constant level. An example of the appearance of a thin current sheet under steady solar wind conditions can be found in the previously published literature: the magnetic field and plasma data from the DE-1 in Fig. 2 of Burch et al. (1983) shows the existence of a downward region 1 current having width of about 0.3° in invariant latitude in the prenoon LLBL region. The ISEE 3 data shows that solar wind parameters had been nearly steady for several hours before this DE-1 observation,
although there are some data gaps. These observations suggest that a steady state of the solar wind may be one precondition for the formation of a thin region 1 current sheet in the LLBL. Note that a thin current sheet does not always appear in steady solar wind conditions, probably because of plasma instabilities in the magnetosphere.

Since the average location of the magnetopause is controlled by the IMF as well as the solar wind dynamic pressure (e.g., Roelof and Sibeck, 1993; Fairfield, 1995; Petrinec and Russell, 1996; Shue et al., 1997), variations of the solar wind parameters will cause the magnetopause to move or oscillate. The following sections describe how the magnetopause oscillation influences the spatial structure of the LLBL region 1 current.

5.2. Thickening of a region 1 current zone

The satellite observations of hot boundary layer plasmas (Sckopke et al., 1981; Lundin and Evans, 1985) have indicated that an injected magnetosheath plasma is embedded in a “halo” of magnetospheric plasma. This structure is schematically shown in Fig. 12a, adapted from Fig. 7 of Sckopke et al. The energy-dependent magnetic drifts of the injected hot particles act to smooth out, in the azimuthal direction, the wavy structure, making the injection boundary of the magnetosheath plasma diffusive. More precisely, even a plasma of the boundary layer proper is thought to be a mixture of plasma sheet and magnetosheath populations: probably, a microscopic filament of sheath particles injected, by a (locally enhanced) $E \times B$ drift, into the magnetosphere will merge with the already mixed plasma, due to the magnetic drift and field-aligned motions. This picture is consistent with the observations that particle energy spectra obtained in respective regions of the magnetosheath, outer boundary layer, inner boundary layer and plasma sheet all

![Diagram of boundary layer plasmas and injection boundary](image)

*Fig. 12.* (a) Schematic illustration of boundary layer plasmas based on the ISEE 1 and 2 satellite observations by Sckopke et al. (1981). (b) Simple model of the injection boundary, roughly parallel to the magnetopause boundary. The open- and closed-headed arrows denote antisunward and sunward plasma flows, respectively.
cross at approximately one point (e.g., Eastman et al., 1976; Song et al., 1990; Traver et al., 1991).

To consider the effect of the magnetopause fluctuation on the region 1 current structure rather than the aforementioned azimuthal 'diffusion' effect, suppose such a laminar structure of the LLBL plasma as shown in Fig. 12b, where the injection boundary is roughly parallel to the magnetopause boundary. When the magnetopause stands still (as tentatively assumed), the distance between the injection boundary and the magnetopause represents (in an average sense) the inherent penetration depth of magnetosheath particles. Then, the injection boundary for the stationary magnetopause is hereafter referred to as the penetration limit of the magnetosheath plasma. (As will be discussed below, the actual injection boundary for the magnetopause in motion is different from the penetration limit.) In the situation in Fig. 12b, as predicted from the theory in Section 3, the region 1 current will arise near the penetration limit, provided that the boundary layer proper is crosswise-uniform. The penetration limit is sharp, but it is assumed to have a finite thickness (of the order of ~0.1° in footpoint latitude), which also gives the minimum thickness of the region 1 current zone. Associated space charges emerge in the region 1 current zone, so that the stagnation line (where the convection flow ceases) appears inside the current zone, because the intensity of the region 2 current is usually smaller than that of the region 1 current. The observations (e.g., Scopke et al., 1981) have shown that the antisunward flow speed in the region of injected magnetosheath often exceeds 100 km/s, but in other regions the flow speeds are on the order of 10 km/s. The above picture of the LLBL may hold true for the case of the southward IMF, but during periods of northward IMF, the convection flow associated with the NBZ field-aligned currents (Iijima et al., 1984; Iijima and Shibaji, 1987) could act to transport the magnetosheath plasma deeper into the magnetosphere, thereby thickening the LLBL (see observations by Mitchell et al., 1987). Hence the following discussion will be focused on the southward IMF case.

A practical point is to what extent a displacement of the magnetopause can change the ionospheric footprints of field lines threading the penetration limit of a magnetosheath plasma. The T89 model is used to assess the relation in displacement between the magnetopause and the footprint of the penetration limit. To obtain a displaced magnetopause (defined as the open-closed boundary), a magnetic field (of magnitude of a few nanoTeslas) in the Z-direction is uniformly added to the T89 field. The penetration limit is assumed to be one of the division lines in the model LLBL region, which are defined in Section 2. Taking 0.6 (division) line or 1.0 line as the penetration limit, Fig. 13 shows the ratio of the displacement of the penetration limit footprint to the magnetopause displacement, plotted against MLT (defined on the ionosphere), where the magnetopause displacement is in units of the Earth's radius and the footprint displacement is in units of the distance of one degree latitude, i.e., about 110 km. The displacement of the magnetopause boundary at a point (on the original boundary) is defined, on the equatorial plane, as the minimum distance from the point to the displaced one, and the footprint displacement is defined similarly on the ionospheric plane.

An examination of ISEE 1 and 2 data during intervals of multiple magnetopause crossings (Song et al., 1988) has shown that the average peak-to-peak amplitude of magnetopause oscillations is about 0.5 R_E under southward IMF conditions, for oscillation periods from ~2 to ~30 min. From Fig. 13, it is then inferred that observed
**Fig. 13.** Ratio of the displacement of the penetration limit footprint to the magnetopause displacement, plotted against MLT (defined on the ionosphere); the penetration limit is taken to be 1.0 line (LLBL inner edge) or 0.6 line. The magnetopause displacement is in units of the Earth’s radius, and the footprint displacements are in units of the distance of one degree latitude, i.e., about 110 km. (For the definitions of these displacements, see Section 5.2.)

**Fig. 14.** Schematic illustration of the injection of magnetosheath particles into the LLBL.

magnetopause oscillations may fluctuate the ionospheric footprint of the penetration limit, with an amplitude of roughly one degree. Here suppose an azimuthally convecting flux tube carrying plasma particles together with field lines threading the LLBL. When the flux tube happens to be within the (instantaneous) penetration limit, in the course of magnetopause oscillations, it will be intruded, at low-latitudes, by newly injected magnetosheath particles (see Fig. 14). When it is earthward of the penetration limit, it will not be subject to such intrusion. The particle distribution in the flux tube close to the average penetration limit can be affected by intermittent loading of magnetosheath particles, if the magnetopause oscillation period is shorter than the time required for the plasma tube to travel over a distance of \( \approx 10 \text{ } R_E \) (a significant fraction of the total length of the LLBL bounded by the adiabaticity limit (see Section 3)). For example, the particle distribution in a flux tube in a stagnant plasma with a speed of 20 km/s will be affected by oscillations with periods shorter than 50 min, while a plasma tube with a high speed of 200 km/s will
be affected only by oscillations with periods less than 5 min. (Either of these period ranges falls within the aforementioned range for the actual observation.) Thus, the magnetopause oscillations as observed by Song et al. could lead to the formation of a transition layer in which the content of a flux tube changes from being highly abundant in magnetosheath particles to lacking in them. From the results shown in Fig. 13 and the findings of Song et al., the ionospheric thickness of this transition layer is inferred to be roughly one degree in latitude. As can be seen from eq. (11), this also gives the width of the (thickened) region 1 current under the influence of magnetopause oscillations.

The above estimation of the width of the region 1 current is somewhat speculative. A more convincing evaluation of this thickening effect would require a numerical simulation for the oscillating magnetopause with a realistically modelled LLBL; this task will be attempted in the future.

5.3. Effect of latitudinal convection

An effect of E x B convection on the latitudinal width of the region 1 current is briefly considered here. This is just the velocity filter effect. For simplicity, suppose such a simple structure of the LLBL as illustrated in Fig. 12b. Since the low-latitude portion of a flux tube in the LLBL region is intruded by sheath particles (see Fig. 14), it takes some time for them to travel along the field lines to low altitudes. If the average field-aligned speeds of injected ions are in the range between 100 and 400 km/s, it takes 160–640 s to travel over a distance of 10 R_e. In such time intervals, the sheath particles are convected poleward, under southward or weakly northward IMF conditions. Therefore, under the assumption that the convection speed has a latitudinal component of 0.1 km/s as viewed on the ionosphere, the inner (equatorward) edge of the population of injected sheath ions just above the ionosphere has a width of roughly 0.5° in latitude, even if the inner edge of the low-latitude (near the equator) portion of that population is strictly sharp. Basically, the FAC width is controlled by the characteristic scale length for a gradient of the flux tube energy content (see eq. (11)). Also noting that the cross section of a flux tube as well as the magnetic drift are greater in magnitude at higher altitudes, particularly in a region near the equator, the LLBL FAC width, under the influence of latitudinal convection, is then estimated as at most ≈0.5° in latitude, although it may depend on the convection speed, the energy distribution of LLBL ions, the field line length, and so on.

At the same time, the velocity filter effect may alter the spatial relation between the inner edge of the precipitating LLBL ions and the associated FAC region. Particularly at low altitudes, the former can be shifted poleward from the latter due to that effect. Such a signature may be discernible in high spatial resolution data of currents and particles. In the Viking observation in Fig. 2, the equatorward edge of LLBL ion precipitation is located slightly poleward of the region of intense downward current. In this case, data of the IMF is not available, but it is inferred to be southward from \( K_p = 5 \).
profile of the LLBL particle population in the framework of the 1989 Tsyganenko model are consistent with observations. This FAC generation is ultimately attributed to the magnetic field distortion in the LLBL region, namely a property of the flux tube volume that it changes significantly along the LLBL inner edge.

Acknowledgments

The work of T. Yamamoto was supported in part by the joint research programs at Radio Atmospheric Science Center, Kyoto University, Uji, Kyoto and the National Institute of Polar Research, Itabashi, Tokyo. The DE-1 6-s average magnetometer data and IMP 8 15-s average data were obtained from the NSSDC SPyCAT data distribution sites (Dieter K. Bilitza and Joseph H. King, respectively). The 5-min averages of solar wind parameters were from the NSSDC IMP 8 Magnetometer Data Server (H. Bridge, A. J. Lazarus, S.M. Petrinec and C.T. Russell).

The editor thanks the referees for their help in evaluating this paper.

References


(Received December 12, 2001; Revised manuscript accepted April 30, 2002)
Characteristics of negative values of polar cap index as an indicator of reversed convection

Tsutomu Nagatsuma

Applied Research and Standards Division, Communications Research Laboratory, Nukui Kitamachi 4-chome, Koganei-shi, Tokyo 184-8795

Abstract: The polar cap (PC) index sometimes shows negative values in the summer hemisphere. The Hall current induced by reversed convection causes the negative values of PC index. This means that the negative PC is an indicator of reversed convection in the polar cap of the summer hemisphere. Using the northern PC index (PCN) for 1995 to 1999, we have statistically examined the occurrence characteristics of negative values of PCN. The results of our data analysis show that a negative value of PCN frequently occurs when the solar zenith angle is less than 75° and when the IMF $B_T$ and $B_Z$ are positive. These results are consistent with those obtained by the previous studies. Further, we found that the occurrence of negative PCN increases with solar wind electric field projected onto the GSM YZ plane ($E_T$), if the clock angle is less than 60°. When the clock angle is from 60 to 80°, the occurrence of negative PCN increases with $E_T$ in the cases when $E_T$ is less than 4 mV/m, and decreases with increasing $E_T$ in the cases of $E_T$ greater than 4 mV/m. Further, the occurrence of negative PCN increases with increasing magnitude of IMF projected onto the GSM YZ plane ($B_T$) and decreases with increasing solar wind velocity. These results suggest that the occurrence of reversed convection in the polar cap is not a simple function of IMF clock angle. Their occurrences are also controlled by the magnitude of $B_T$ and solar wind velocity.

1. Introduction

In the Earth's magnetosphere the plasma convection is driven by the solar wind-magnetosphere interaction. Ionospheric currents are driven by the transmission of magnetic tension from the magnetosphere to the ionosphere. Since the behavior of ionospheric currents is reflected on the ground as magnetic field perturbations, we can estimate the condition of plasma convection in the magnetosphere by analyzing magnetic field variations on the ground. The ionospheric current induced by solar wind-magnetosphere interaction leading to plasma convection is named DP2 current system (Nishida, 1968). The main contributor to horizontal magnetic field variations in the near-pole region is the DP2 current system. Based on this interpretation, Troshichev et al. (1979) proposed a polar cap (PC) index. According to the results of Troshichev et al. (1988), the PC index has a good correlation with the merging electric field $E_m = V_{SW} B_T \sin^2(\theta/2)$ introduced by Kan and Lee (1979), where $V_{SW}$ is the solar wind velocity, $B_T$ is the projection of interplanetary magnetic field (IMF) onto the GSM YZ plane, and $\theta$ is a clock angle of IMF.
The PC index sometimes shows negative values in the summer hemisphere. The occurrence of negative PC corresponds to the case of positive IMF $B_Z$ (e.g., Vennerstrøm et al., 1991), since the flow direction of the Hall current induced by the reversed convection is opposite to the two-cell convection pattern (Maezawa, 1976; Friis-Christensen et al., 1985). Reversed convection corresponds to the lobe cell convection driven by a northward IMF merging at the magnetopause with the southward tail lobe field poleward of the cusps that close within the polar cap (Maezawa, 1976; Crooker, 1992). This suggests that the occurrence of reversed convection in the polar cap can be identified using the polarity of the PC index.

Until now, many ionospheric convection models have been produced as functions of each component of the IMF intensity and clock angle (e.g., Weimer, 1995). These models show that the transition from two-cell to four-cell convection pattern occurs at a fixed clock angle. However, the locations of dayside merging regions expected from anti-parallel merging model depend on the clock angle (Crooker, 1979). And even when the clock angle is the same, the solar wind velocity and magnetic field intensity can be different from one event to the next. The intensity of $B_T$ and solar wind velocity controls the efficiency of the dayside merging process (e.g., Hill, 1975; Mitchell and Kan, 1978). It is expected that these differences in solar wind conditions modulate the occurrence of reversed convection.

In this paper, we have statistically analyzed the characteristics of negative values of PC index and have examined the conditions for the occurrence of reversed convection in the polar cap.

2. Observation

Figure 1 shows the yearly variations in the 15-min northern PC index (PCN) and those in the solar zenith angle both from Qaanaaq (formerly Thule). Occurrence of negative PCN is mostly seen in northern summer for the solar zenith angle less than 90°.

A month-UT plot for the occurrence distribution of negative PCN for 1995 to 1999 is shown in Fig. 2. The size of each dot shows the magnitude of negative PCN. Color contours of solar zenith angle at Qaanaaq are overlaid in this figure. Occurrence of negative PCN is mostly distributed from 12 to 22 UT in summer. Since the magnetic local noon of corrected geomagnetic coordinate at Qaanaaq is about 15 UT, negative PCN is mostly distributed from 9 to 19 MLT. Also, this distribution roughly corresponds to the condition for the solar zenith angle less than 75°.

Based on these results, we use for the rest of our analysis PCN data in summer from 1995 to 1999 for the UT range between 12 and 22 UT. For 1995 to 1997, data from the WIND satellite are used to represent the solar wind parameters. For 1998 to 1999, data from the ACE satellite are used.

Figure 3 shows scatter plots of PCN as a function of IMF clock angle, binned for every 1 mV/m of solar wind electric field projected onto the GSM YZ plane ($E_z$) from 0 mV/m to 8 mV/m. Black lines with filled circles show the average values of PCN, binned for every 20° of clock angle from $-180°$ to $180°$. The calculated $E_m$ is shown as a blue line.

In general, there is a good correspondence between the average value of PCN and $E_m$. 
Fig. 1. (Top) Yearly variation of northern PC index (PCN) at Qaanaaq for 1999. (Bottom) Yearly variation of solar zenith angle at Qaanaaq.

Fig. 2. A month-UT plot for the occurrence distribution of negative PCN. Color contours of solar zenith angle at Qaanaaq are overlaid.
Fig. 3a. Scatter plots of IMF clock angle vs PCN for four ranges of $E_r$, 00–01 mV/m (top left), 01–02 mV/m (top right), 02–03 mV/m (bottom left), and 03–04 mV/m (bottom right). The black lines with filled circles show the average values of PCN. The blue lines show the calculated merging electric field ($E_m$).

Fig. 3b. Same as Fig. 3a but for 04–05 mV/m (top left), 05–06 mV/m (top right), 06–07 mV/m (bottom left), and 07–08 mV/m (bottom right).
However, the average of PCN tends to be saturated when $E_m$ is greater than 5 mV/m (Nagatsuma, 2002). Also, it is clear that negative values of PCN can be seen frequently at positive clock angle (IMF $B_y > 0$). The average of PCN tends to be lower than $E_m$ when the negative value of PCN frequently occurs. On the contrary, the average of PCN tends to be higher than $E_m$ at clock angle from $-90$ to $0^\circ$.

In the case of $E_T$ less than 4 mV/m, negative values of PCN can be seen in a wide range of positive clock angle sector. On the other hand, negative values of PCN only appear at clock angle from 0 to 80° for $E_T$ greater than 4 mV/m.

The occurrence probability of negative PCN are shown as a function of $E_T$ in Fig. 4. In this figure, we examined the data for the cases of positive clock angle ($B_y > 0$). The occurrence probability is calculated from the data binned for every 20° of clock angle from 0 to 180°. Different lines of colors represent each bin of clock angle. The occurrence of negative PCN tends to increase with $E_T$ for the case of clock angle less than 60°. For the case of clock angle from 60 to 80°, the occurrence of negative PCN increases with $E_T$ during $E_T$ less than 4 mV/m. On the other hand, the occurrences of negative PCN decrease with increasing $E_T$ during periods of $E_T$ greater than 4 mV/m. For the cases of clock angle from 80 to 100°, the occurrence of negative PCN decreases with increasing $E_T$.

### 3. Discussion

The sources of horizontal magnetic field variations in the near-pole regions are supposed to be ionospheric Hall currents in the polar cap and the distant field-aligned currents at the poleward rim of the auroral oval (Vennerstrøm et al., 1991). In the winter
hemisphere, Hall currents are negligible since the ionospheric conductivity is low. On the contrary, the effect from the Hall currents is significant in the summer hemisphere with low solar zenith angle, since the ionospheric conductivity is enhanced. The results for the occurrence period of negative PC obtained from our data analysis clearly support this interpretation and are consistent with those from previous studies (Vennerstrøm et al., 1991; Nagatsuma et al., 2002).

The occurrence of negative PCN depends on the polarity of IMF \( B_y \). This is consistent with the results obtained by Vennerstrøm et al. (1991). When \( B_y \) is negative, sunward flow of a lobe cell tends to be formed in the dawn sector. For positive \( B_y \), sunward flow of a lobe cell tends to form in the dusk sector. Since negative PCN is frequently detected in the afternoon sector by enhanced conductivity, a positive \( B_y \) is a favorable condition. The enhancement of PCN during the period of clock angle from \(-90^\circ\) to \(0^\circ\) might be caused by an enhancement of anti-sunward flow of lobe cell for negative \( B_y \).

As shown in Fig. 4, the occurrence probability of negative PCN is a function of \( E_T \) and clock angle. It seems that the clock angle from 60 to 80° is the transition sector for the occurrence of reversed convection in the polar cap. Since \( E_T \) is a product of \( B_T \) and the solar wind velocity (\( V_{SW} \)), we examined the \( B_T \) and \( V_{SW} \) dependence separately for the occurrence of negative PCN.

Figure 5 shows the occurrence probability of negative PCN as a function of \( B_T \). Data are sorted by nine bins of clock angle, which are plotted with different colors. At clock angles less than 60°, the occurrence probability of negative PCN tends to increase with increasing \( B_T \). This result is consistent with that obtained by Weimer (1995).

![Figure 5](image)

Fig. 5. Occurrence probability of negative values of PCN as a function of \( B_T \), binned for every 20° of clock angle from 0 to 180°.
Fig. 6. Occurrence probability of negative values of PCN as a function of $V_{sw}$, binned for every 20° of clock angle from 0 to 180°.

The occurrence probability of negative PCN as a function of $V_{sw}$ is shown in Fig. 6. Data are sorted in nine bins of clock angle, which are represented by lines of different colors. In general, the occurrence probability decreases with increasing $V_{sw}$ and decreases with increasing clock angle.

Figures 5 and 6 show that the occurrence of negative PCN depends on $B_T$ and $V_{sw}$. These results are interpreted as follows. Based on the anti-parallel merging model, reconnection occurs at the region tailward of the cusp and drives reversed convection in the case of northward $B_T$ (Crooker, 1992). Mitchell and Kan (1978) suggests that merging will proceed as long as the flow speed perpendicular to the merging line is less than twice the Alfvén velocity. This means that the flow speed and Alfvén velocity in the magnetosheath controls the efficiency of dayside merging. When $V_{sw}$ is high, the plasma flow perpendicular to the magnetic field at magnetosheath region is also high. In this situation, reconnection would be difficult to proceed in the region far from the stagnation point. On the contrary, reconnection can proceed even in the regions distant from the stagnation point when $V_{sw}$ is low.

The Alfvén velocity at the magnetosheath is proportional to the intensity of IMF. We can consider that the Alfvén velocity is high when $B_T$ is strong since the magnitude of IMF is equal to or greater than $B_T$. So it is expected that the dayside merging progresses when $B_T$ is strong.

It is not clear why the trend of occurrence of negative PCN changes at 4 mV/m of $E_T$ for the case of clock angle from 60 to 80°. According to the anti-parallel merging model, a merging region most distant from the sub-solar point is expected when the IMF points downward or duskward (Crooker, 1979). Under these conditions, the occurrence of
reversed convection strongly depends on $B_T$ and solar wind velocity. It may be possible to interpret $B_T$ and the solar wind velocity dependence of the occurrence of negative PCN in the following way. When $E_T$ is less than 4 mV/m, $B_T$ is large and solar wind velocity is low. In this situation, reversed convection can develop even during periods of large clock angle. When $E_T$ is greater than 4 mV/m, large solar wind speed is expected and reversed convection cannot proceed.

4. Conclusion

We have examined characteristics of negative values of PCN index. The results of our data analysis shows that negative values of PCN frequently occur during the solar zenith angle less than 75° and during the IMF clock angle from 0° to 90° ($B_T > 0$ and $B_Z > 0$). These results are consistent with those in the previous studies. Further, we found that the occurrence probability of negative PCN increases with $E_T$, if the clock angle is less than 60°. When the clock angle is from 60 to 80°, the occurrence of negative PCN increases with increasing $E_T$ until $E_T$ reaches 4 mV/m. On the other hand, the occurrence of negative PCN decreases with increasing $E_T$ when $E_T$ is greater than 4 mV/m. Further, the occurrence of negative PCN increases with $B_T$ and decreases with increasing solar wind velocity. These results suggest that the occurrence of reversed convection in the polar cap is not a simple function of the IMF clock angle. The occurrence conditions are also controlled by the magnitude of IMF projected onto the GSM $YZ$ plane and by solar wind velocity.

Acknowledgments

We are grateful to M. Sugiura for valuable suggestions. We thank the Danish Meteorological Institute for providing the PCN index. We appreciate the efforts by the science teams (Wind/SWE, Wind/MAG, ACE/SWEPAM, ACE/MAG) for making the data available.

The editor thanks the referees for their help in evaluating this paper.

References


(Received January 16, 2002; Revised manuscript accepted June 12, 2002)
On the energy of the poloidal magnetic field near the ionosphere

Akimasa Yoshikawa¹, Masahiro Itonaga² and Kiyohumi Yumoto¹

¹Graduate School of Sciences, Department of Earth and Planetary Science, Kyushu University, Fukuoka 812-8581
²Faculty of Education, Yamaguchi University, Yoshida, Yamaguchi 753-8513

Abstract: The role of the ionospheric Hall effect on the energy balance in the magnetosphere-ionosphere (MI) system coupled through the field-aligned current (FAC) is discussed. FACs lose their stored electromagnetic energy in the ionosphere through Joule dissipation; this process is caused by the closure of the FAC via the ionospheric Pedersen current carried by ions. On the other hand, the ionospheric rotational Hall current carried by electrons cannot be dissipated. However, the ionospheric rotational Hall current can also be excited by an incident FAC, causing it to radiate Poynting fluxes that lead to the growth of a poloidal-type magnetic field in the magnetosphere and atmosphere. From the viewpoint of energy conservation, a few ambiguities in the above statements may be recognized. In this paper, we clarify the energy balance of the electromagnetic disturbances between the magnetosphere, ionosphere and atmosphere. The generation of the Hall current (together with the associated poloidal magnetic field) will be shown to occur during the growth stage, when the electromagnetic energy is pumped through the divergent Hall current, regardless of how slow the growth may have been.

1. Introduction

The field-aligned current (FAC) transmitted to the ionosphere by a shear Alfvén wave is closed by the divergence of the ionospheric Pedersen current. The divergent electric field that is associated with this current generates a rotational ionospheric Hall current, which produces a poloidal-type magnetic perturbation field that is observable on the ground (Nishida, 1964). When the incident field varies with time, the poloidal magnetic field also varies with time, inducing a rotational electric field that is opposite to the Hall current. This rotational electric field in turn drives the divergent Hall current closure via the FAC (Yoshikawa and Itonaga, 1996, 2000). The magnetic perturbation field produced by the divergent Hall current is toroidal.

Thus, when the intrinsically time-varying nature of the incident FAC is taken into account, the energy dissipated in the ionosphere must be dependent on the ionospheric Hall conductivity. When the incident FAC is growing with time, the rotational Hall current also grows; the resulting poloidal magnetic field is either radiated away from the ionosphere or (in the evanescent case) stored in its neighborhood. While the ionospheric reflection coefficient for the static case was given by (Scholer, 1970) as
\[ R_A = \frac{\Sigma_A - \Sigma_P}{\Sigma_A + \Sigma_P}, \]  

(1)

for the time-dependent case, this coefficient must be expressed as

\[ R_A = \frac{\Sigma_A - (\Sigma_P - \Sigma_H MC_{\text{div-rot}})}{\Sigma_A + (\Sigma_P - \Sigma_H MC_{\text{div-rot}})}, \]  

(2)

where \( \Sigma_A \) is the Alfvén wave conductance just above the ionosphere; \( \Sigma_P \) and \( \Sigma_H \) are the height-integrated Pedersen and Hall conductivities, respectively, and \( MC_{\text{div-rot}} \) is the mode conversion ratio from a divergent to a rotational electric field (Yoshikawa and Itonaga, 1996, 2000). In the context of the reflection process described by case (1), perfect impedance matching between the FAC and the ionospheric divergent current occurs when \( \Sigma_P = \Sigma_A \). On the other hand, such perfect impedance matching never happens in case (2) because the effective conductivity for the divergent Hall current \( -\Sigma_H MC_{\text{div-rot}} \) is complex. This implies that the Hall effect controls the energy distribution in the ionosphere (Buchert, 1998; Yoshikawa, 1998; Yoshikawa et al., 1999).

Furthermore, the damping factor for a field-line oscillation (standing FAC system) is estimated for a rectangular magnetosphere model with symmetric ionospheres by the equation

\[ \gamma = \frac{V_A}{2l_1} \ln |R_A|, \]  

(3)

(e.g., Ellis and Southwood, 1983), where \( l_1 \) is the length of the field line and \( V_A \) is the Alfvén velocity. Combining eqs. (2) and (3) shows that the energy dissipation of the field line oscillations is controlled by the ionospheric Hall effect (Yoshikawa et al., 1999). Lysak and Song (2001) and Yoshikawa et al. (2002) independently confirmed that the reflection and mode conversion process described by eq. (2) satisfies the conservation of current and energy.

The involvement of Hall conductivity in energy dissipation and reflection may appear to contradict the conventional idea that the Hall current does not do any work, since it is directed perpendicular to the electric field. The purpose of this paper is to elucidate the role of the Hall effect in the energy conservation process. The divergent Hall current plays a crucial role by extracting electromagnetic energy stored in the FAC and pumping it into the rotational Hall current, which radiates a poloidal field.

In the next section, we will review conduction current and energy dissipation in the ionosphere from the viewpoint of particle motion. In Section 3, we will reconstruct the energy balance of electromagnetic disturbances between the magnetosphere, ionosphere and atmosphere using a sheet ionosphere model, and clarify the role of the Hall effect in the energy balance. In Section 4, we will use the results obtained in Section 2, to discuss the origin of the poloidal magnetic field observed on the ground in the context of the inductive response of the ionosphere. Finally, the results of this paper will be summarized in Section 5.

In this paper, we will focus on the general role of energy flow. To quantitatively evaluate energy flows, the equations describing the evolution of electromagnetic disturbances with boundary conditions must be determined; this topic will be discussed in a
future paper.

2. Motion of charged particles in the ionospheric current and energy dissipation

The motion of charged particles, including collisions between neutral and charged constituents, and their contribution to macroscopic conduction current was established by Hines (1953). In this chapter, the generation of ionospheric currents by the incidence of the FAC will be reconstructed from the viewpoint of particle motions using one of the concepts of Hines’ theory. We will focus on the growth of the incident FAC in the ionosphere and provide a qualitative consideration of how the macroscopic motion of electrons and ions in the ionosphere are driven by the macroscopic electric field and how they generate the ionospheric current. Figure 1 summarizes the relation of motions of the

---

Fig. 1. Schematic illustration of the motions of electrons and ions in the ionospheric currents accompanying the growth phase of the incident FAC’s divergent electric field. In the direct process, the divergent Pedersen current, carried by the ions, is closed via the FAC. On the other hand, the Hall current, carried by the electrons, flows as rotational current.
electron and ion motions and the ionospheric currents accompanying the growth of the incident divergent electric field of the FAC. The same conceptual figures have also been presented in Buchert (1998), Yoshikawa (1998), and Yoshikawa and Itonaga (2000). Detailed explanations of the interactions are presented in Sections 2.1 to 2.3.

In a magnetohydrodynamic (MHD) medium, both electrons and ions are frozen onto the magnetic field line (Alfvén, 1950) and have a bulk velocity of \( \mathbf{v}_\parallel = (\mathbf{E}_\parallel \times \mathbf{B}_0)/B_0^2 \), known as the \( \mathbf{E} \times \mathbf{B} \) drift motion, where \( \mathbf{E}_\parallel \) is the electric field and \( \mathbf{B}_0 \) is the background magnetic field. The subscripts \( \parallel \) and \( \perp \) denote the components of the vector quantity that are parallel and perpendicular to \( \mathbf{B}_0 \), respectively. A twisted magnetic field line generated by some type of magnetospheric generator can be stored as electromagnetic energy. The motion of the twisted field line is associated with the \( \mathbf{E}_\text{div} \times \mathbf{B}_0 \) drift motion of the magnetic field lines, where \( \mathbf{E}_\text{div} \) is the divergent part of the electric field. If \( \mathbf{B}_0 \) is taken to be uniform, the FAC carried by the shear Alfvén wave associated with this \( \mathbf{E}_\text{div} \times \mathbf{B}_0 \) drift motion satisfies the relationship \( (\nabla \times \mathbf{v}_\parallel)_\parallel \approx \nabla \times \mathbf{E}_\text{div}/\mathbf{B}_0 \neq 0 \) (e.g., Vasyliunas, 1984; Southwood and Kivelson, 1991). This divergent electric field leads to a series of phenomena that release the electromagnetic energy stored in the FAC in the magnetosphere-ionosphere-atmosphere and Earth (MIAE) electromagnetically coupled system, as discussed in the following sections.

2.1. FAC closure via the divergent pedersen current and the generation of the reflected FAC

Let us begin to reconstruct the generation of ionospheric current from by looking at the particles that are involved. On the MHD time scale, ions that collide with the neutral particles are scattered, causing a reduction in the collective motion of the \( \mathbf{E}_\text{div} \times \mathbf{B}_0 \) drift. The scattering of these ions, which is accelerated by the macroscopic electric field, leads to the absorption of the electromagnetic energy produced by the twisted field lines. Thus, these multiple collisions with neutral particles transform the electromagnetic energy stored in the FAC to thermal energy in the ions and neutral particles (Joule dissipation). A finite mean velocity of the ions in the electric field direction results from an equilibrium between the acceleration caused by the electric field and the deceleration caused by the collisions with neutral particles. This causes a so-called ionospheric divergent Pedersen current, which closes via the FAC. In a divergent current system, the reflected divergent electric field is generated by impedance mismatching of the incident FAC and the divergent Pedersen current generated by the incident electric field. Hence, a Poynting flux associated with this reflected electric field and the associated reflected FAC is generated. In other words, the electromagnetic energy stored in the incident FAC is transformed not only by the Joule dissipation arising from ionic collisions, but also the electromagnetic energy of the \( \mathbf{E}_\text{div} \times \mathbf{B}_0 \) drift motion of the reflected FAC.

2.2. Generation of rotational Hall current by resistance to a counter-electromotive force

On the MHD time scale, electron collisions with particles can be neglected, enabling the electrons to maintain the \( \mathbf{E}_\text{div} \times \mathbf{B}_0 \) drift motion. A reduction in the velocity of the ions from the \( \mathbf{E}_\text{div} \times \mathbf{B}_0 \) drift velocity causes a substantial electron current whose direction is opposite to that of the \( \mathbf{E}_\text{div} \times \mathbf{B}_0 \) drift; this current known as a rotational Hall current. The Hall current does not to undergo any loss because its carriers are unscattered electrons.
To increase a rotational Hall current, however, work must be done by the power supply against the counter electromotive force $E_{\text{rot}}$, which is generated to prevent the growth of the associated magnetic fields: $E_{\text{rot}}$ (direction along $E_{\text{div}} \times B_0$) satisfies $(\nabla \times E_{\text{rot}}) \neq 0$. Furthermore, $E_{\text{rot}}$ accelerates the scattered ions; this leads to Joule dissipation though collisions with the neutral particles and generates the rotational Pedersen current.

The rotational Hall current seems to be naturally excited by the appearance of the $E_{\text{div}} \times B_0$ drift motion of the electrons that results from ion-induced. However, as this current increases, energy is expended as the poloidal magnetic disturbances increase and Joule dissipation from the rotational Pedersen current occurs. The main problem is how and from where such energy is supplied. The original energy source is likely to be the electromagnetic energy stored in the incident FAC. If we can successfully explain how this energy is transferred not only to the Joule dissipation that results from ion collisions and the electromagnetic energy of the reflected FAC but also to the generation of $E_{\text{rot}}$, the role of the ionospheric Hall effect in magnetosphere-ionosphere coupling will become obvious.

2.3. FAC closure via the divergent Hall current and energy absorption from the FAC system

The $E_{\text{rot}} \times B_0$ drift motion of the electrons that accompanies the generation of $E_{\text{rot}}$ will now be discussed. The current generated by the $E_{\text{rot}} \times B_0$ drift motion of the electrons becomes the divergent Hall current, which also closes via the FAC in a manner similar to the closure of the ion-carried divergent Pedersen current.

Whether or not the electromagnetic energy stored by the incident FAC is spent by the ionospheric divergent current depends on the direction of $E_{\text{div}}$, that is, on the sign of $J_{\text{div}} \cdot E_{\text{div}}$. The term $J \cdot E$ in the ionosphere becomes an indicator of the energy exchange between the electromagnetic energy of the MHD fluid in the magnetosphere, the thermal energy in the ionosphere, and the energy in the neutral atmosphere. For an increasing rotational Hall current, the direction of $E_{\text{rot}}$ is opposite to the rotational Hall current; therefore the direction of the divergent Hall current arising from the electrons is the same as that of the divergent Pedersen current arising from the ions. Therefore, the electromagnetic energy stored by the FAC is expended by the generation of $E_{\text{rot}}$ through the generation of the divergent Hall current and the divergent Pedersen current. However, further quantitative consideration using a particle model may not provide any new insights. Accordingly, the Joule dissipation in the context of a thin ionospheric current sheet approximation will be discussed in the next section.

3. Energy balance between magnetosphere and atmosphere through the ionospheric boundary

In this chapter, a thin-sheet approximation will be used to describe the ionosphere. With this approximation, the ambient magnetic field $B_0$ is assumed to penetrate the ionosphere vertically. This model will be used to formulate the energy balance of the electromagnetic disturbances between the magnetosphere and the atmosphere through the ionospheric boundary. In this approximation, the motions of the electrons, ions and neutral particles are incorporated into the ohmic ionospheric current of
\[ J_{\perp}^{\text{ion}} = \sum_{r} E_{\perp}^{\text{ion}} - \sum_{i} \left( E_{\perp}^{\text{ion}} \times b \right), \]  

(4)

where \( b \) is a unit vector along \( B_0 \). Here, the superscripts “mag”, “ion” and “atm” denote the quantities in the magnetosphere, ionosphere and atmosphere, respectively.

### 3.1. Total energy balance in the MIAE system

When the ionosphere is regarded as an infinitely thin conducting layer, the ionospheric current produces a jump in the magnetic field between the magnetosphere and the atmosphere:

\[ \mu_0 J_{\perp}^{\text{ion}} = b \times \left( \lim_{m \to 1} \frac{b^{\text{mag}}}{b^{\text{ion}}} - \lim_{d \to 1} \frac{b^{\text{atm}}}{b^{\text{ion}}} \right), \]

(5)

where \( \mu_0 \) is the magnetic permeability in a free space. Taking the inner product between eq. (5) and \( E_{\perp}^{\text{ion}} \), we get:

\[ J_{\perp}^{\text{ion}} \cdot E_{\perp}^{\text{ion}} = -\lim_{m \to 1} \mu_0 \left( E_{\perp}^{\text{mag}} \times b^{\text{mag}} \right) + \lim_{d \to 1} \mu_0 \left( E_{\perp}^{\text{atm}} \times b^{\text{atm}} \right). \]

(6)

Equation (6) shows that \( J_{\perp}^{\text{ion}} \cdot E_{\perp}^{\text{ion}} \) in the ionosphere balances the Poynting flux in the electromagnetic disturbances from the magnetosphere and the atmosphere into the ionosphere. Here, in the transformation from eqs. (5) to (6), the tangential continuity condition of the electric field is used:

\[ \lim_{m \to 1} E_{\perp}^{\text{mag}} - \lim_{d \to 1} E_{\perp}^{\text{atm}} = E_{\perp}^{\text{ion}}. \]

(7)

We should note that in eq. (6), \((E_{\perp} \times B_{\perp}) > 0\) indicates an upward energy flow along the magnetic field line.

Substituting eqs. (4) for (6) and applying the condition

\[ -\sum_{\nu} \left( E_{\perp}^{\text{ion}} \times b \right) \cdot E_{\perp}^{\text{ion}} = 0, \]

(8)

produces an energy balance equation for the entire system

\[ -\lim_{m \to 1} \mu_0 \left( E_{\perp}^{\text{mag}} \times b^{\text{mag}} \right) + \lim_{d \to 1} \mu_0 \left( E_{\perp}^{\text{atm}} \times b^{\text{atm}} \right) = \sum_{\nu} E_{\perp}^{\text{ion}}. \]

(9)

This shows that the Poynting flux of the electromagnetic field absorbed by the ionosphere is transformed into Joule dissipation by the Pedersen current.

Condition eq. (8) means that the Hall current does not directly contribute to \( J_{\perp}^{\text{ion}} \cdot E_{\perp}^{\text{ion}} \), providing no net energy to the external (dynamical) system. However, for a growth in the ionospheric rotational Hall current accompanying the incident FAC to occur, the increasing counter-electromotive force must be resisted by something, resulting in the rotation of the poloidal-type magnetic disturbance radiation or the storage of magnetic energy as an ionospheric surface wave. Even if the Hall effect cannot do work on an external system, what does it mean on Earth that it contribute to the accumulation of poloidal magnetic energy? Furthermore, as shown by Yoshikawa et al. (1995, 1999), the eigenfrequency and damping factor of a standing field line oscillation, i.e., a standing FAC system, are changed by the Hall conductivity. Why does the Hall effect’s control of the electromag-
magnetic energy stored by the FAC not appear in eq. (9)? To answer these questions, we will discuss the physical meanings of \(-\Sigma_{01}(E_{\perp}^{\text{con}} \times \hat{b}) \cdot E_{\perp}^{\text{con}} = 0\) in the following section.

### 3.2. \(J_{\perp}^{\text{con}} \cdot E_{\perp}^{\text{con}}\) terms

To explicitly consider the ionospheric Hall effect, after Yoshikawa and Itonaga (2000), we will divide the ionospheric current and the electric field into divergent and rotational components: \(J_{\perp}^{\text{con}} = J_{\perp}^{\text{div}} + J_{\perp}^{\text{rot}}\) and \(E_{\perp}^{\text{con}} = E_{\perp}^{\text{div}} + E_{\perp}^{\text{rot}}\), respectively. Physically, \(J_{\perp}^{\text{div}}\) satisfies \(\nabla \cdot J_{\perp}^{\text{div}} = 0\) and closes via the FAC just above the ionosphere; \(J_{\perp}^{\text{rot}}\) satisfies \((\nabla \times J_{\perp}^{\text{rot}}) = 0\) and is associated with the developed poloidal magnetic field. On the other hand, \(E_{\perp}^{\text{rot}}\), which satisfies \(\nabla \cdot E_{\perp}^{\text{rot}} = 0\), shows an accumulated charge density on the ionosphere and \(E_{\perp}^{\text{con}}\), which satisfies \((\nabla \times E_{\perp}^{\text{con}}) = 0\), is an ionospheric inductive electric field. The term \(J \cdot E\) therefore becomes:

\[
J_{\perp}^{\text{con}} \cdot E_{\perp}^{\text{con}} = J_{\perp}^{\text{div}} \cdot E_{\perp}^{\text{div}} + J_{\perp}^{\text{rot}} \cdot E_{\perp}^{\text{rot}}. \tag{10}
\]

The first and second terms of eq. (10) give the energy transformation ratio from electromagnetic and kinetic energy in the ionosphere in a divergent current system and a rotational current system, respectively.

By dividing eq. (5) into the divergent and rotational components, we get the following equations:

\[
J_{\perp}^{\text{div}} = \hat{b} \times \mu_{0}^{-1} \left( \lim_{\mu_{\perp} \to 0} \mathbf{b}_{\text{rot}}^{\text{mag}} - \lim_{\beta_{\perp} \to 0} \mathbf{b}_{\text{rot}}^{\text{am}} \right), \tag{11}
\]

\[
J_{\perp}^{\text{rot}} = \hat{b} \times \mu_{0}^{-1} \left( \lim_{\mu_{\perp} \to 0} \mathbf{b}_{\text{div}}^{\text{mag}} - \lim_{\beta_{\perp} \to 0} \mathbf{b}_{\text{div}}^{\text{am}} \right). \tag{12}
\]

The terms \(\mathbf{b}_{\text{rot}}^{\text{mag}}\) and \(\mathbf{b}_{\text{rot}}^{\text{am}}\) in eq. (11) represent toroidal magnetic field perturbations induced by the FAC and the divergent polarization current in vacuum, respectively. These perturbations produce a torsion in the magnetic flux tube. The terms \(\mathbf{b}_{\text{div}}^{\text{mag}}\) and \(\mathbf{b}_{\text{div}}^{\text{am}}\) in eq. (12) are the poloidal magnetic field perturbations accompanied by the rotational current in the vertical plane of \(\mathbf{B}_{0}\) in the magnetosphere and atmosphere, respectively; these perturbation cause the magnetic flux tube to expand or compress. Equations and show the jump in the toroidal and poloidal magnetic fields by the ionospheric divergent and rotational current, respectively.

Using eqs. (11) and (12), we get the Poynting flux expression in the divergent and rotational current systems can be obtained:

\[
J_{\perp}^{\text{con}} \cdot E_{\perp}^{\text{con}} = S_{\text{FAC}} + S_{\text{TM}}, \tag{13}
\]

\[
J_{\perp}^{\text{con}} \cdot E_{\perp}^{\text{con}} = S_{\text{MS}} + S_{\text{TE}}. \tag{14}
\]

where

\[
S_{\text{FAC}} = -\lim_{\mu_{\perp} \to 0} \frac{(E_{\perp}^{\text{con}} \times \mathbf{b}_{\text{rot}}^{\text{mag}})}{\mu_{0}}, \]

\[
S_{\text{TM}} = -\lim_{\beta_{\perp} \to 0} \frac{(E_{\perp}^{\text{div}} \times \mathbf{b}_{\text{rot}}^{\text{am}})}{\mu_{0}}, \]

\[
S_{\text{MS}} = -\lim_{\mu_{\perp} \to 0} \frac{(E_{\perp}^{\text{con}} \times \mathbf{b}_{\text{div}}^{\text{mag}})}{\mu_{0}}, \]

\[
S_{\text{TE}} = -\lim_{\beta_{\perp} \to 0} \frac{(E_{\perp}^{\text{div}} \times \mathbf{b}_{\text{div}}^{\text{am}})}{\mu_{0}}.
\]

Here, \(S_{\text{FAC}}\) and \(S_{\text{TM}}\) in eq. (13) represent the parallel component of the Poynting vector.
of the FAC (shear Alfvén wave), and the transverse magnetic field (TM) waveguide mode that is absorbed in the ionosphere, respectively. On the other hand, $S_{\text{inc}}^{\text{MS}}$ and $S_{\text{inc}}^{\text{TE}}$ in eq. (14) represent the parallel component of the Poynting vector of a magnetoionic (MS) wave and a transverse electric field (TE) waveguide mode, respectively. $S_{\text{inc}}^{\text{FAC}}$ can be divided into the incident part of $S_{\text{inc}}^{\text{FAC}}$ and the reflected part of $S_{\text{inc}}^{\text{FAC}}$, whereas $S_{\text{inc}}^{\text{FAC}}$ is directly related to magnetic disturbances on the ground.

Furthermore, the ionospheric current eq. (4) can also be separated into divergent eq. (15) and rotational eq. (16) components:

$$J_{\text{div}}^{\text{ion}} = \Sigma \rho E_{\text{div}}^{\text{ion}} - \Sigma \nabla \times \left( E_{\text{rot}}^{\text{ion}} \times \mathbf{b} \right).$$  \hspace{2cm} (15)$$

$$J_{\text{rot}}^{\text{ion}} = \Sigma \rho E_{\text{rot}}^{\text{ion}} - \Sigma \nabla \times \left( E_{\text{div}}^{\text{ion}} \times \mathbf{b} \right).$$  \hspace{2cm} (16)$$

The first and second terms on the right-hand side of eq. (15) are the divergent part of ionospheric Pedersen and Hall currents, respectively, while the first and second terms on the right-hand side of eq. (16) are the rotational part of the ionospheric Pedersen and Hall currents, respectively.

The current expression in eq. (6) thus becomes:

$$J_{\text{div}}^{\text{ion}} \cdot E_{\text{div}}^{\text{ion}} = \Sigma \rho E_{\text{div}}^{\text{ion}} - \Sigma \nabla \times \left( E_{\text{rot}}^{\text{ion}} \times \mathbf{b} \right),$$  \hspace{2cm} (17)$$

$$J_{\text{rot}}^{\text{ion}} \cdot E_{\text{rot}}^{\text{ion}} = \Sigma \nabla \times \left( E_{\text{div}}^{\text{ion}} \times \mathbf{b} \right),$$  \hspace{2cm} (18)$$

Finally, using eq. (13), (14), (17) and (18) the energy-balance equations for divergent-current and rotational-current systems

$$S_{\text{inc}}^{\text{FAC}} + S_{\text{inc}}^{\text{FAC}} + S_{\text{inc}}^{\text{TM}} = \Sigma \rho E_{\text{div}}^{\text{ion}} - \Sigma \nabla \times \left( E_{\text{rot}}^{\text{ion}} \times \mathbf{b} \right),$$  \hspace{2cm} (19)$$

$$S_{\text{inc}}^{\text{MS}} + S_{\text{inc}}^{\text{TE}} = \Sigma \nabla \times \left( E_{\text{div}}^{\text{ion}} \times \mathbf{b} \right),$$  \hspace{2cm} (20)$$

Equation (19) shows the energy balance in the divergent current system. When the FAC is incident from the magnetosphere, the equation shows that its energy flow to the ionosphere ($S_{\text{inc}}^{\text{FAC}}$) is distributed into that of the reflected FAC ($S_{\text{inc}}^{\text{FAC}}$), the atmospheric TM waveguide mode ($S_{\text{inc}}^{\text{TM}}$), the Joule dissipation through the ionospheric-divergent Pedersen current ($\Sigma \rho E_{\text{rot}}^{\text{ion}}$), and the work on the external system through the divergent Hall current ($-\Sigma \nabla \times \left( E_{\text{rot}}^{\text{ion}} \times \mathbf{b} \right)$). On the other hand, eq. (20) shows the energy balance in the rotational current system. This system has work done on it by an external system through the rotational Hall current ($\Sigma \nabla \times \left( E_{\text{div}}^{\text{ion}} \times \mathbf{b} \right)$), and its energy is redistributed into the Poynting flux of a magnetoionic wave ($S_{\text{inc}}^{\text{MS}}$), the atmospheric TE waveguide mode ($S_{\text{inc}}^{\text{TE}}$), and Joule dissipation through the ionospheric rotational Pedersen current ($\Sigma \rho E_{\text{rot}}^{\text{ion}}$). These two equations clearly show that the energy which excites the rotational current system is derived from the divergent current (FAC) system through the ionospheric Hall effect. The relation $-\Sigma \nabla \times \left( E_{\text{rot}}^{\text{ion}} \times \mathbf{b} \right)$ in the total current system can also be written as

$$(-\Sigma \nabla \times \left( E_{\text{rot}}^{\text{ion}} \times \mathbf{b} \right)) \cdot E_{\text{rot}}^{\text{ion}} = 0$$

in divergent system + $\Sigma \nabla \times \left( E_{\text{rot}}^{\text{ion}} \times \mathbf{b} \right)$ in rotational system = 0. (21)
summarizes how the electromagnetic energy in the incident part of the FAC is distributed into the divergent and rotational current systems.

Interestingly, the left hand side of eq. (19) is a summation of the Poynting flux of the incident and reflected shear Alfvén wave and the radiated atmospheric TM waveguide mode; this component shows the net energy that is absorbed into the ionosphere. As shown on the right-hand side of equation, this absorbed energy is used for Joule dissipation and rotational-current system excitation. Thus, the total energy balance can be written as

$$S_{\text{FAC}}^{\text{(in)}} + S_{\text{FAC}}^{\text{(out)}} + S_{\text{TM}}^{\text{(rot)}} = \Sigma_p \left( E_{\text{rot}}^{\text{in}} + E_{\text{rot}}^{\text{out}} \right) - \left( S_{m=1}^{\text{MS}} + S_{z=1}^{\text{TE}} \right).$$  \hspace{1cm} (22)

When the poloidal magnetic field is increasing, the relation $S_{m=1}^{\text{MS}} + S_{z=1}^{\text{TE}} < 0$ is satisfied, which leading to the following relationship:

$$S_{\text{FAC}}^{\text{(in)}} + S_{\text{FAC}}^{\text{(out)}} + S_{\text{TM}}^{\text{(rot)}} > \Sigma_p \left( E_{\text{rot}}^{\text{in}} + E_{\text{rot}}^{\text{out}} \right).$$  \hspace{1cm} (23)

Equation (23) shows that the net electromagnetic energy of the FAC absorbed by the ionosphere is larger than the total Joule dissipation in the ionosphere. In the unsteady phase of MI coupling, Joule dissipations in the ionosphere must be distinguished from the electromagnetic energy of the FAC absorbed in the ionosphere.
4. What causes ground magnetic disturbances?

As shown by eq. (19)+ (20)=(10), the ionospheric Hall currents have not performed any net work on the external system. However, when we consider the MI coupling that divides the ionospheric current into its divergent and rotational components is considered the energy exchange between the divergent and rotational current systems is clearly caused by the Hall effect, with an exchange rate per unit time of $|\sum_{H} E_{\text{div}}^{\text{ion}} E_{\text{rot}}^{\text{ion}}|$. From the viewpoint of energy exchange between the divergent and rotational current systems, the divergent Hall current and rotational Hall currents are one and indivisible. The largest distinction between the divergent and rotational Hall currents is their ability to accumulate magnetic energy. The rotational Hall current can exist as a steady current in the ionosphere and because of its circularity, it accumulates the poloidal magnetic energy as an ionospheric surface wave. The poloidal magnetic energy being stored as the ionospheric surface wave is represented by the following equation:

$$\varepsilon_{\text{pol}}(t) = - \int_{-\infty}^{t} (S_{m-1}^{\text{MS}} + S_{a-1}^{\text{TE}}) dt' = - \int_{-\infty}^{t} J_{\text{rot}}^{\text{ion}}(t) \cdot E_{\text{rot}}^{\text{ion}}(t) dt',$$  \hspace{1cm} (24)

which can exist even in the steady state (when $J_{\text{rot}}^{\text{ion}} \cdot E_{\text{rot}}^{\text{ion}} = 0$) and is a unique function of the developed steady rotational Hall current. On the other hand, the divergent Hall current provides a finite $J_{\text{rot}}^{\text{ion}} \cdot E_{\text{rot}}^{\text{ion}}$ but cannot store the magnetic energy itself. Equation (24) also shows that when the stored poloidal magnetic energy disappears, the sign of the Poynting flux has to satisfy $(S_{m-1}^{\text{MS}} + S_{a-1}^{\text{TE}}) > 0$, which means that $\varepsilon_{\text{pol}}(t)$ must be absorbed into the ionosphere and converted to Joule dissipation and the reflected FAC through $|\sum_{H} E_{\text{div}}^{\text{ion}} E_{\text{rot}}^{\text{ion}}|$

With this in mind, the manner in which magnetic disturbances develop on the ground should be reconsidered; this is a very fundamental and important point. In the past context of MI coupling, ground magnetic disturbances (regardless of their frequency range) observed in the high latitudinal region, such as the aurora electrojet excited by region 1 and 2 current closure (e.g., Ijima and Potemra, 1976), Dp2-type disturbances (e.g., Nishida and Jacobs, 1962), the sudden commencement of storms (Nishida, 1964), and geomagnetic pulsations (e.g., Samson and Rostoker, 1972), are thought to be accompanied by an ionospheric rotational Hall current excited by the Hall effect from the divergent electric field of the incident FAC to the ionosphere. On the other hand, the FAC flowing into the ionosphere was interpreted as being closed via the divergent Pedersen current. This treatment is equivalent to adopting an electrostatic approximation for the driving electric field of the ionospheric current, namely:

$$J_{\perp}^{\text{ion}} \approx \sum_{\text{p}} E_{\text{div}}^{\text{ion}} - \sum_{\text{p}} (E_{\text{div}}^{\text{ion}} \times \mathbf{b})$$  \hspace{1cm} (25)

Under this approximation, however, the existence of a divergent Hall current generated by the ionospheric inductive rotational electric field $E_{\text{rot}}^{\text{ion}}$ is ignored, so from where do the ground magnetic disturbances receive their energy and how do they grow? Equation (25) cannot answer this question. On the other hand, the right-hand side of eqs. (19) clearly indicates that the energy for the excitation of the ground magnetic-field disturbance is taken from the FAC system. Although the essential importance of the divergent Hall current and the physical phenomena that is causes have been extensively investigated (Yoshikawa
and Itonaga, 1996, 2000; Yoshikawa et al., 1999; Buchert and Budnik, 1997; Buchert, 1998; Lysak and Song, 2001), most researchers are of the opinion that such an effect is small and may only become important for high-frequency disturbances, such as Pc 1 pulsations. However, eqs. (19) and (20) show that for low-frequency phenomena to produce ground magnetic field disturbances, a series of physical processes is required whereby the electromagnetic energy of the FAC system is pumped by the finite divergent Hall current and supplied to the rotational current system through the rotational Hall current. It should be emphasized that the rotational Hall current in the steady state does not describe a coupling between the FAC and the rotational current system at that exact time, but shows the results of accumulated couplings.

Finally, a series of physical processes will be described in terms of the equivalent circuit model of the FAC system (e.g., Tamao, 1984). Figure 3 shows an equivalent circuit model of the FAC system coupled to the ionospheric divergent and rotational current. The same conceptual model was shown by Yoshikawa et al. (1999); in their model, however, the existence of the rotational current system was not explicitly represented but was implicitly incorporated into the extra loading coil of the divergent Hall current that
closes via the FAC system. This figure shows that the Pedersen current carried by the ions flows through the resistor of the divergent and rotational parts of the ionospheric circuit, which dissipate the electromagnetic energy of the FAC stored that is as a twist in the magnetic field. On the other hand, the divergent and rotational Hall currents carried by the electrons flow in the mutual inductor of the divergent and rotational part of ionospheric current circuit, respectively. The rotational current system is excited by the electromotive force through the mutual coupling between the divergent and rotational Hall currents. The poloidal magnetic energy, stored as a compressional ionospheric surface wave, is expressed by the current flowing in the loading coil of the magnetosonic wave and the TE waveguide mode in the rotational current system. After passing through an inductive phase, the coupling between the divergent and rotational current systems through the mutual inductor reaches zero. In the divergent system, the generator only works to compensate the Joule loss by the steady divergent Pedersen current. In the rotational system, an electromotive force and Joule dissipation do not occur, and the electrons flow to maintain the lossless rotational Hall current.

5. Summary

In this paper, electromagnetic perturbations produced in the magnetosphere-ionosphere system by the incidence of a shear Alfvén wave are studied, with particular attention given to the role of the ionospheric Hall current. The following results were obtained:

1) When the incident perturbation is developing, the divergent Hall current takes electromagnetic energy from the FAC and supplies it to the rotational Hall current, thereby causing the poloidal magnetic perturbation field to grow. Thus, the dissipation of the electromagnetic energy stored by the FAC in the ionosphere depends on the Hall conductivity.

2) The energy loss from the FAC during the growing phase is larger than the Joule dissipation in the ionosphere, since part of the incident energy is used to build the magnetic field perturbation. In the decaying phase of the FAC, the energy of the poloidal magnetic field propagates backward into the ionosphere and returns to the FAC.

3) In time-dependent situations, the divergent and rotational Hall currents are closely tied with each other. The rotational Hall current is closed in the ionosphere and can build up, but the divergent Hall current is connected to the FAC and flows outwards. This may partly explain why the divergent Hall current has not attracted much attention.

4) In the steady state, coupling between the divergent and rotational current systems does not occur. Although the rotational Hall current is not connected with the energy supply at this stage, it should be noted that this Hall current (together with the associated poloidal magnetic field) has been generated during the growing stage, when the energy was pumped through the divergent Hall current, regardless of how slow the growth may have been. The ionospheric Hall current in the steady state is a product of the inductive Hall effect that was in operation at an earlier time.

The above-mentioned results show that for the growth of poloidal magnetic disturbances on the ground, the finite divergent Hall current pumps up the electromagnetic energy
On the energy of the poloidal magnetic field near the ionosphere

stored by the FAC and forms a rotational current system through the ionospheric Hall effect, no matter how slow the disturbance is.

Since Nishida’s work (1964), the poloidal magnetic field perturbation at the ground level that is associated with the FAC has been interpreted to arise from the ionospheric Hall effect. Although much circumstantial evidence for Nishida’s theory exists, this paper shows that the free energy required for the radiation of the poloidal magnetic field is supplied by the FAC through the work done by the divergent Hall current and its carriers, the electrons. To clarify whether or not Nishida’s theory and our interpretation are correct, a progressive technique or observational equipment that can identify the carriers of the ionospheric current must be developed. By confirming that electron current closure occurs via the FAC when the poloidal magnetic field at ground level is either increasing or decreasing, we would be able to conclude that this event occurs because of the ionospheric Hall effect.

Acknowledgments

Useful discussions with A. Nishida, R. Fuji, R.L. Lysak and D.J. Southwood are gratefully acknowledged. This work was supported by a premium from the President Award of Kyushu University, 2001.

The editor thanks Dr. S. Buchert and another referee for their help in evaluating this paper.

References


(Received January 29, 2002; Revised manuscript accepted June 22, 2002)
Comparison of flow angle variations of $E$-region echo characteristics at VHF and HF

R.A. Makarevitch¹, A.V. Koustov¹, K. Igarashi², K. Ohtaka², T. Ogawa¹, N. Nishitani³, N. Sato⁴, H. Yamagishi⁴ and A.S. Yukimatu⁴

¹Institute of Space and Atmospheric Studies, University of Saskatchewan, 116 Science Place, Saskatoon, S7N 5E2 Canada
²Communications Research Laboratory, Tokyo 184-8795
³Solar-Terrestrial Environment Laboratory, Nagoya University, Honohara, Toyokawa 442-8507
⁴National Institute of Polar Research, Kaga I-chome, Itabashi-ku, Tokyo 173-8515

Abstract: In this study, characteristics of the auroral $E$-region echoes at two significantly different radar frequencies of 12 and 50 MHz are compared. Considered observations were performed at the Syowa Antarctic station in March of 1997 using two HF and one VHF radars at various angles with respect to the magnetic L shells. The diurnal variation of echo occurrence was found to be similar at two frequencies and consistent with previous studies. On the other hand, variation of echo occurrence with L-shell angle $\phi$ was shown to be significantly different at two frequencies. 50-MHz echoes were detected preferentially along the L shell (dominating direction of the electrojet flow) while 12-MHz echoes were detected in a broad range of azimuths with the maximum in echo occurrence at $\phi \approx 40$°–50°. By plotting the Doppler velocity versus L-shell angle, we demonstrate that 12-MHz echoes can be divided into two populations, the high- and low-velocity echoes. The high-velocity echoes were observed mainly along the L shells while the low-velocity echoes were observed at all directions. We also show that the echo populations exhibit different variation of the Doppler velocity with the L-shell angle. We argue that while the 50-MHz echoes are related to the Farley-Buneman and gradient drift plasma instabilities, the 12-MHz echoes can have additional sources, such as the thermo-diffusion instability and/or neutral wind-related plasma instabilities.

1. Introduction

Meter-scale auroral $E$-region irregularities have been traditionally studied with VHF coherent radars (see review papers by Fejer and Kelley, 1980; Haldoupis, 1989; Schlegel, 1996; Sahr and Fejer, 1996). Recent deployment of the Super Dual Auroral Radar Network (SuperDARN or SD) system of coherent HF radars has provided opportunities for studies of ionospheric irregularities in the decameter band (Villain et al., 1987, 1990; Milan and Lester, 1998, 1999, 2001; Milan et al., 2001; Fukumoto et al., 1999, 2000; Jayachandran et al., 2000; Koustov et al., 2001; Makarevitch et al., 2001; Uspensky et al., 2001; Ogawa et al., 2001).

It has been shown that the characteristics of HF echoes are often quite different from
the ones established at VHF. Milan and Lester (1999) were the first who described several new types of HF echoes, in addition to classical type 1 and type 2 echoes that are traditionally related to the Farley-Buneman (F-B) and gradient drift (G-D) plasma instabilities. These authors considered data of the SD Iceland East radar observing predominantly along the direction of the magnetic L shells. Jayachandran et al. (2000), after studying Saskatoon SD radar data at large L-shell angles, proposed to consider separately a new class of HF echoes which they termed "the slow long-lived E-region plasma structures (SLERPs)". SLERPS are very stable, spatially and temporally, echoes with low velocities. Jayachandran et al. (2000) attributed SLERPS' generation to the G-D plasma instability. Ogawa et al. (2001) reported on short-range low-velocity HF echoes observed at the Syowa Antarctic station that were related to the effects of neutral wind and neutral turbulence.

New information on the issue has been given in recent studies by Koustov et al. (2001) and Makarevitch et al. (2001). These authors considered Syowa SD measurements that were supplemented by nearly simultaneous VHF measurements. Analysis of the velocity relationship at HF and VHF showed that the HF echoes could be divided into two categories, the high- and low-velocity echoes. The high-velocity echoes were observed mostly along the L shells while the low-velocity echoes were observed at all directions. The high-velocity echoes were related to the F-B plasma instability while the nature of the low-velocity echoes was not discussed.

Nearly simultaneous VHF and HF coherent observations of E-region echoes are of special interest since they provide information on irregularity characteristics at two quite different scales. Such measurements become even more important if the plasma instabilities that are responsible for the irregularity formation are the same for the two scales. In the previous studies by Koustov et al. (2001) and Makarevitch et al. (2001) data on the echo power and Doppler velocity were discussed. In this paper we expand the above studies by considering much larger data sets and also by considering other echo characteristics such as occurrence rate and HF spectral width. Our main focus is on properties of E-region HF and VHF echoes as a function of the direction of observations.

2. Observations

We use measurements from two SENSU HF SuperDARN radars (Syowa East and Syowa South, ~12 and ~11 MHz, respectively) operated by the National Institute of Polar Research and one VHF coherent radar (~50 MHz) run by the Communications Research Laboratory (CRL). All three radars are located at Syowa (69.0°S, 39.6°E, magnetic latitude ~67°). Figure 1 shows the experimental setup with the near (slant ranges r less than 900 km) field of view (FoV) of the Syowa East (blue sector) and Syowa South (green sector) radars. FoV of the CRL radar for r<1000 km is shown by the yellow sector. The solid red curve represents the line of zero off-perpendicular (aspect angle (α=0) at 110-km altitude, assuming no refraction for radio waves, and the circular dashed white lines are the slant range marks (r=400, 600, and 800 km). The thick black lines in Fig. 1 are the Polar Anglo-American Coordinate System (PACE) magnetic parallels or L shells (Baker and Wing, 1989).

In this study we formulate results in terms of L-shell angle ϕ defined as the angle between the direction of the radar wave vector k and the magnetic L shell. The L-shell
angle within Syowa East and Syowa South FoVs is positive and changes continuously from $\phi \approx 20^\circ$ (eastern edge of HF FoV, Syowa East beam 15, Fig. 1) to $\phi \approx 140^\circ$ (western edge of HF FoV, Syowa South beam 0). The azimuth corresponding to the L-shell angle of $90^\circ$ is located close to the eastern edge of the Syowa South FoV, and we will call the radar directions close to this azimuth “perpendicular directions” or “directions perpendicular to the flow”. The radar directions close to the edges of the common FoV (Syowa East beam 15 and Syowa South beam 0) will be called “parallel directions” or “directions along the flow”, even though the L-shell angle is not exactly zero or $180^\circ$. In this terminology we assume that the flow is predominantly along the L shells.

The Syowa HF radars are analogous to all other SD radars in the network (Greenwald et al., 1995). Each radar scans through 16 successive beam positions in $3.33^\circ$ steps, covering a $\sim 53^\circ$ range of azimuths. One full scan in a normal mode lasts 2 min, so that every 2 min there are data on the spectral power, mean Doppler shift, and spectral width of echoes in each of 75 range cells (starting from a range of 180 km with a 45-km step) in which scattered signal is detected. No interferometer measurements of an angle of echo arrival (or echo height) were available for Syowa. During HF data postprocessing, ground-scattered signals were removed from the records by applying the standard criteria.
that such signals have low (<50 m s\(^{-1}\)) velocity and low (<20 m s\(^{-1}\)) spectral width. For the analysis, power of all radars was corrected on slant range by assuming that there is an \(r^{-3}\) power decrease with range \(r\).

Information on the CRL 50-MHz Doppler radar is given by Igarashi et al. (1995, 1998) and Ogawa (1996). This radar measures the echo power and Doppler velocity (15-km range resolution starting from 120 km) using the double-pulse technique. The radar scans with a 5\(^\circ\) step in a range of geographic azimuths of 48.7\(^\circ\)–218.7\(^\circ\) (geomagnetic azimuths of 5\(^\circ\)–175\(^\circ\)) so that all directions of the SD measurements are covered. One scan through FoV is accomplished in 4 min. Below we refer to the beams of the 50-MHz radar by their geomagnetic azimuth (beam 40 has geographic azimuth of 83.7\(^\circ\)).

3. Database selection

To obtain a data set unbiased by event selection and with maximum possible number of records we decided to consider one month of joint (HF and VHF) measurements. The CRL VHF radar observations were only carried out during winter/equinox seasons of 1995–1997, and this was one of the factors limiting our choice. Analysis of the HF and VHF data sets showed that VHF echoes were more rare than HF echoes. Because of that we selected for the analysis the month of March, 1997, for which 50-MHz echo occurrence was the highest.

Among 31 days in March of 1997, the SD radars worked in the normal (as described in the previous section) mode of operation only during 23 days, and we selected these days. To exclude possible F-region echoes we restricted the slant ranges of all radars by 810 km. And, finally, HF echoes with unusually large spectral width (>500 m s\(^{-1}\)) and HF and VHF echoes with low uncorrected power (<3 dB at HF and <1 dB at VHF) were excluded from the database.

4. Diurnal variation of echo occurrence

We first assess the echo occurrence for all three systems during the selected period. Figure 2 shows the total number of echoes as a function of the universal time at 12 MHz (upper panel) and 50 MHz (bottom panel). Here we marked data obtained by Syowa East and Syowa South as 12 MHz though the radars used slightly different frequencies (we will neglect by this difference below for simplicity of presentation).

The number of echoes in Fig. 2 was computed as follows. Since the FoV of the VHF radar is broader than the combined FoVs of the two HF radars, we considered only those VHF radar directions that were within this combined HF FoV. Then, we computed the total number of echoes for all directions and for all near slant ranges (\(r<810\) km for all radars) for each 30-min period of time. The universal time is approximately equal to the magnetic local time (MLT) for the near FoV of the Syowa radars (\(UT\approx MLT\)), therefore Fig. 2 essentially presents the number of echoes for both frequencies as a function of MLT. For comparison we provide the results from Hanuise et al. (1991) for the radar frequency of 12 MHz (open circles) and data from Saskatoon IGY auroral radar (McNamara, 1972) at 50 MHz (diamonds) in arbitrary scale. For these radars the universal time differs significantly from the magnetic local time so that we shifted their data to account for this.
Fig. 2. Number of echoes in the near field of view detected in all available beams as a function of universal time at 12 MHz (upper panel) and 50 MHz (bottom panel). The open circles in the upper panel represent data of Hanuse et al. (1991) for their HF observations (the scale is arbitrary), and the diamonds in the bottom panel show number of 50-MHz echoes versus time from McNamara (1972).

difference.

One can see that the number of echoes at 12 MHz exhibits a clear maximum around 0100 UT and that there are significantly fewer echoes during the daytime (7–19 UT). This result is in reasonable agreement with the data of Hanuse et al. (1991), except of slight shift in the location of the maxima (~5 MLT vs ~1 MLT).

The similar maximum can be seen at 50 MHz (bottom panel of Fig. 2). The maximum is broader here (1–4 UT), the number of echoes during the evening (18–24 UT) is much less than that at 12 MHz, and there are almost no echoes between 7 and 17 UT. The data of McNamara (1972) show smoother variation with time and the maximum is shifted (~0 MLT vs ours ~3 MLT), but, overall, the shape of the curve is similar.
Fig. 3. Variation of number of echoes in the near field of view with radar beam number and universal time at 12 MHz (upper panel) and 50 MHz (bottom panel).

The absence of echoes during some specific times (1000–1030 UT and 1700–1800 UT in the upper panel) at HF is due to the fact that HF radars simply were not operational during these periods. The reasons for the existence of the local minima at VHF (0300–0400 and 0500–0600 UT in the bottom panel) are not clear. We believe that this is most likely because of the smaller data set at VHF (the number of echoes at HF is several times larger than that at VHF).

Figure 2 gives the overall occurrence in the common FoV. In Fig. 3 we present more detailed information. Here we plotted the number of echoes as a function of the time of the day and radar direction, see surfaces for 12 and 50 MHz (upper and bottom panels, respectively). Notice that the number of echoes for Syowa East is 2–3 times larger than that for Syowa South. However, there is no significant difference between various beams.
for individual radars; the azimuthal distribution of echoes is more or less homogeneous for each HF radar. One can notice that for every UT sector the number of echoes is only slightly larger for central beams of each HF radar.

At 50 MHz the situation is completely different. The number of echoes varies with azimuth of observations/beam number. There are almost no echoes detected for 50-MHz beams 80-100 and the number of echoes is at maximum at the edges of common FoV (50-MHz beams 40 and 140).

5. Echo occurrence within FoV

Our next step is to include into analysis slant range information. We consider here echo occurrence versus range and azimuth for the morning observations between 02 and 06 UT when the echo occurrence is at maximum for both HF and VHF (see previous section). We also expect that for this interval the plasma flow is mostly L-shell aligned.

Figure 4 is a contour plot of number of echoes (azimuth of observations horizontally, slant range vertically) for the near FoVs (180-765 km) of 12-MHz (upper panel) and 50-MHz (bottom panel) radars. The red curve on these plots is the line of perfect aspect angle (\(\alpha = 0\)) at the height of 110 km, calculated for each radar by assuming the ionospheric electron density profile from the International Reference Ionosphere-95 (IRI-95) model (Bilitza, 1997). Dashed yellow lines in Fig. 4 are contours of equal L-shell angle \(\phi\) (contour steps are 20°), computed using the PACE magnetic field model. Data from both SD radars were combined in the upper panel of Fig. 4; the transition between two adjacent SD FoVs occurs at an azimuth of \(\sim 135°\), where the echo occurrence contours change drastically. The FoV of the 50-MHz radar is about 170°-wide; we show 50-MHz data for only those directions that are close to the SD radar directions.

In the 12-MHz part of the diagram one can recognize two distinct areas: close slant ranges (200-500 km), where the distribution of echo occurrence is almost homogeneous with azimuth (except for the noticed earlier difference in number of echoes between two SD radars) and farther ranges (500-765 km) where echoes tend to occur close to the perpendicular directions. There is a clear gap between echoes at closer and farther ranges. For the perpendicular directions, the gap is located at \(\sim 480\) km and it is quite narrow while for the edges of HF FoV (azimuths of 82° and 184°) it is much wider (there are no echoes detected for \(r=500-765\) km). The location of the radar cells with maximum number of echoes follows closely the perfect aspect line, except for the fact that they are shifted slightly to closer ranges. The shift seems to increase towards the edges of the FoV. Interestingly, for the eastern sector (Syowa East) there is a broad global maximum (\(\phi=40-60°\), azimuth=95-115°) with some local peaks (orange spots). For the southern sector the maximum is smaller (500 echoes vs 1000) and its azimuthal location is \(\phi=100-120°\), azimuth=150-170°. Thus the maxima of echo occurrence are somewhere between the parallel and perpendicular directions. In our opinion, echoes from the nearest ranges (\(r<500\) km) are originated from the \(E\)-region heights (aspect angles at the electrojet heights of 110 km are around zero for this area) while echoes from farther ranges are most probably originated from the bottom/central part of the \(F\) region. The far range echoes can also be obtained through the one and half hop propagation mode for which radar waves are refracted in the lower ionosphere and then reflected by the ground towards irregularities.
Fig. 4. Number of echoes for various azimuths of observations (horizontal) and slant ranges (vertical) for 12- (upper panel) and 50-MHz (bottom panel) radars. Red solid curves are the zero aspect angle lines calculated by assuming the International Reference Ionosphere-95 (IRI-95) electron density distribution. Dashed yellow lines are the contours of equal L-shell angle (the step is 20°) calculated by assuming the Polar Anglo-American Conjugate Experiment (PACE) magnetic field model.
At 50 MHz (bottom panel) the echo occurrence distribution is quite different (in agreement with Fig. 3). There are no echoes for slant ranges greater than 650 km for perpendicular directions and the shift between the perfect aspect angle line and the location (in slant range) of maxima in echo occurrence is smaller than that at 12 MHz, especially for the perpendicular directions. But the most striking feature in the 50-MHz diagram is that distribution of echo occurrence with azimuth is not homogeneous as at 12 MHz. The number of echoes increases with departure from the perpendicular directions. It maximizes at L-shell angles of 20° and 150° for the eastern and southern sectors of the VHF radar FoV, respectively.

6. Average echo power within FoV

Occurrence diagram of Fig. 4 reflects only the fact of echo presence, and does not tell us how strong was the echo. We address this issue in this section.

In Fig. 5 we show average power in a format similar to Fig. 4. To produce this diagram, we averaged the range-corrected power and plotted it as a function of azimuth and slant range. Similar to the echo occurrence diagram, here and in the next section we consider only observations between 02:06 UT to limit ourselves to the period of maximum in echo occurrence.

One can see that distribution of 12-MHz power is essentially the same as distribution of number of echoes. One can see the stripe of E-region echoes at short ranges and F-region echoes at farther ranges. Within the stripe of E-region echoes, power decreases gradually by ~2–4 dB from maximum at the edges to the center of FoV. The difference between average power for the Syowa East and the Syowa South radars is ~6 dB.

At 50 MHz, there is quite substantial variation of power with azimuth; the difference in power between parallel and perpendicular directions is ~10 dB and power maximizes at L-shell angles of ~20° and ~150°. We will consider data on spectral power for both HF and VHF in more detail later.

7. Average velocity within FoV

Our next step is to consider variations of the Doppler velocity with range and azimuth, Fig. 6. One can see that along the stripe of E-region echoes, the velocity changes from −400 m s⁻¹ at the eastern edge of the HF radar FoV to +300 m s⁻¹ on the western edge. Similar velocities were measured by the VHF radar. The change of the velocity polarity occurs at the azimuth of ~130–140°, close to the perpendicular directions. Obtained average velocities along and perpendicular to the L shells are consistent with much larger data set considered by Ogawa et al. (2001).

The HF radars measurements show larger velocities at farther ranges, above the line \( \alpha = 0 \). These measurements are very likely affected by the F-region contributions (Makarevitch et al., 2001; Ogawa et al., 2001). The VHF radar echoes also have large average velocities at far ranges (650–765 km) and azimuths of ~180°. We think, however, that these measurements do not necessarily reflect the general tendencies since the number of echoes for these areas is very low (Fig. 4).
8. 12- and 50-MHz velocities: Point-by-point comparison

From the velocity comparison at HF and VHF (the previous section) one can conclude that average HF and VHF velocities are comparable, to a first approximation. Such a conclusion, however, needs to be refined since, as it is well known on the basis of VHF measurements, different types of ionospheric echoes have different mean Doppler
shifts and the variation of Doppler velocity with the L-shell (flow) angle also depends strongly on echo type (Fejer and Kelley, 1980; Haldoupis, 1989). Moreover, Makarevitch et al. (2001) showed that the HF echoes with high- and low-velocities should be considered separately.

In this section we produce a scatter plot of HF velocities versus VHF velocities for 6 azimuths of observations, over a period of 00–24 UT for all 23 days of observations, and consider so called “nearly simultaneous” points, Fig. 7. By definition, some measurement
Fig. 7. Velocity of 12-MHz echoes versus velocity of 50-MHz echoes for 6 azimuths. Average value for the L-shell angle $\phi$ is shown in the top left corner of each panel. For each azimuth, nearly simultaneous measurements at all distances are considered. Red (blue) dots correspond to those echoes whose velocity magnitudes were larger (smaller) than the critical velocity of 320 ms$^{-1}$. Number of red dots/number of blue dots is shown in the right bottom corner of each panel.
at 12 MHz for the azimuth $az_{12}$, slant range $r_{12}$, and time $t_{12}$ was considered to be "nearly simultaneous" with a measurement at 50 MHz at $az_{50}$, $r_{50}$, $t_{50}$ if $|az_{12} - az_{50}| < 5^\circ$, $|r_{12} - r_{50}| < 15$ km, and $|t_{12} - t_{50}| < 60$ s. In total, 552 hours of observations were considered (versus 35 hours in Makarevitch et al., 2001).

We indicated in the top left corner of each panel of Fig. 7 the mean L-shell angle $\phi$ for each specific radar direction. Angle $\phi$ progressively increases as one moves from the easternmost Syowa East beam 15 ($\phi \approx 25^\circ$) to the westernmost Syowa South beam 0 ($\phi \approx 135^\circ$). The other 4 beams considered are Syowa East beams 10, 7, 2, and Syowa South beam 6.

One can clearly see the existence of two echo populations: high- and low-velocity points. We divided all echoes according to the magnitude of 12-MHz Doppler velocity: red (blue) dots are echoes with $V > V_{crit}$ ($V < V_{crit}$), $V_{crit} = 320$ m s$^{-1}$. The number of red/blue points is shown in the right bottom corner of each panel. The difference (vertical distance) between two clouds of points appears to be increasing with $\phi$ for the first four panels. No difference between negative and positive velocities (morning and evening sectors of observations for the first 4 panels, respectively) can be noticed. Two echo types are well seen for both negative and positive HF velocities. There is a drastic difference between variations of number of high- and low-velocity points with angle $\phi$ increase. If low-velocity echoes occur at all directions, the high-velocity echoes are rare at the perpendicular directions (the smallest numbers for high-velocity echoes of 36 and 14 are at $\phi = 70^\circ$ and 110$^\circ$).

The ratio of 12- to 50-MHz velocity magnitudes depends strongly on whether one deals with low- or high-velocity 12-MHz echoes. For the blue dots the velocity magnitude at 50 MHz is almost always larger than that at 12 MHz ($V^{50} \gg V^{12}_{low}$, where index "low" indicates low-velocity 12-MHz echoes). For the red dots the situation seems to be the opposite ($V^{50} < V^{12}_{high}$, red cloud for negative (positive) velocities is shifted to the right (left) from the dotted line of ideal coincidence, index "high" indicates high-velocity 12-MHz echoes).

**9. Doppler velocity versus L-shell angle**

In this section we study high- and low-velocity echoes using a different approach, namely we make scatter plots of HF and VHF velocities versus L-shell angle. Thus for each radar cell (slant range-azimuth) we consider all available data points rather than present one averaged value as in Fig. 6.

Figure 8 shows a scatter plot of velocity versus L-shell angle for 30-min event on March 17, 1997, 0100–0130 UT. This event was studied by Koustov et al. (2001), who considered the variation of the average power and velocity with range. Panel (a) represents data for HF measurements and panel (c) shows data for VHF measurements. The scale for the L-shell angle is given at the bottom of panel (e). For each echo the aspect angle was calculated by assuming one generic electron density profile as explained in Section 5, and we indicate the magnitude of the aspect angle by the color of the point. Color scheme is given in panel (a). Panels (b) and (d) show the histograms of velocity distribution for 12 and 50 MHz, respectively. Total number of points is indicated in the bottom right corner of panels (b) and (d). The scale for panels (b) and (d) is given at the
Fig. 8. Doppler velocity versus L-shell angle at (a) 12 and (c) 50 MHz for March 17, 1997, 0100–0130 UT. The aspect angle for each point is color coded as indicated in panel (a). The horizontal dashed line in (a) corresponds to the critical velocity of $-320$ m s$^{-1}$. The large black dots are maxima of averaged velocity along some radar beams as described in the text. Right panels are histograms of velocity distribution for (b) 12- and (d) 50-MHz radars. Total number of echoes at each frequency is indicated in the bottom right corner of panels (b) and (d). Panel (e) is the histogram of number of echoes (in 5°-wide bins) versus the L-shell angle at 12 MHz. Solid line corresponds to the total number of echoes and the shaded area corresponds to the number of those echoes whose velocity magnitude was larger than the critical value (points below the dashed line in (a)). The scales for panels (a), (c), and (e) are shown at the bottom of panel (e) and the scales for panels (b) and (d) are shown at the bottom of panel (d).
bottom of panel (d).

The most prominent feature in (a) is the existence of two populations of echoes, high- and low-velocity 12-MHz echoes. The low-velocity echoes have relatively small Doppler velocity magnitudes (<200 m s⁻¹) and they present at all L-shell angles Φ. High-velocity points have relatively large velocity magnitudes (>~300 m s⁻¹) and they exist only for Φ<~60°. The presence of two echo types is obvious not only from the panel (a), but also from the histogram (b), where two maxima can be seen, one at ~500 m s⁻¹ and another one at ~0 m s⁻¹.

We believe that these two populations or species correspond to 2 different mechanisms of echo formation, and for this reason we consider data separately for echoes with velocity magnitudes larger and smaller than $V_{\text{crit}} = 320$ m s⁻¹. We indicate the value of $V_{\text{crit}}$ by dashed line in panel (a). In addition to just described separation according to the velocity magnitude, velocities of two echo species seem to have different variations of the Doppler velocity with the L-shell angle. The high-velocity echoes exhibit first increase of the velocity magnitude with Φ for 20°<Φ<35° and then decrease for 35°<Φ<60°. The Doppler velocity of the low-velocity 12-MHz echoes, on the other hand, increases with Φ for all directions exhibiting a “cosine-like” variation.

If one concentrates on the 50-MHz part of the diagram, panels (c) and (d), no distinct echo populations can be identified. The histogram (d) shows rather homogeneous distribution: one can find any specific velocity magnitude with approximately equal probability, except for the very large velocities $V>$600 m s⁻¹. There are some kind of V-structures both for small (Φ<60°) and large (Φ>120°) L-shell angles. For Φ<60°, echoes with small aspect angles (black and dark blue dots) tend to have large velocity magnitudes; they are generally located at the bottom of V-structures.

The V-like structures in Fig. 8 are in no doubt associated with the velocity attenuation with the aspect angle known for VHF echoes (Ogawa et al., 1982; Nielsen, 1986; Kustov et al., 1994). Each V-structure is in fact the data from one of the radar beams. As the range and L-shell angle changes along the beam, so does the aspect angle, reaching at some point perfect condition (α=0). Since phase velocity of plasma waves is expected to have maximum at a perfect aspect angle, V-structures in Fig. 8 are simply a different form of representation of the velocity variation with the aspect angle. V-structures for Φ~90° are not well seen because here the L-shell angle does not change much as one goes along individual radar beams. One can conclude that all “VHF low-velocity points” in (c) are actually echoes with poor aspect angles. Thus Fig. 8 indicates that there is no special “low-velocity” kind of echoes at VHF.

One must bear in mind that values of the aspect angle in Fig. 8 are calculated from the density distribution model. The actual aspect angle conditions might be different from the ones given by the model. For this reason, not all black points are at the bottom of V-structures. Having this in mind, only maximum velocities should be considered for the comparison at HF and VHF since others are significantly affected by the aspect angle attenuation.

Because there is a substantial data spread for each radar cell, to obtain statistically meaningful estimate of the maximum velocity, we performed one more type of analysis. For each radar beam, we considered all points with Doppler velocity magnitudes larger than critical value $V_{\text{crit}}$ (for 12 MHz, points below the dashed line in (a)). We found the
average of these points for each slant range (thus we obtained one point for each vertical line in (a) and (c)) so that the maximum average velocity of high-velocity echoes in each radar beam was calculated. The obtained velocities are plotted in panels (a) and (c) by large black dots. These points should correspond to zero aspect angle measurements since velocity magnitude maximizes at perfect aspect angle.

Now we can compare directly L-shell dependencies for high-velocity echoes at 12 and 50 MHz, large black dots in panels (a) and (c). One can see that, overall, the two trends

Fig. 9. The same as Fig. 8, but for 23 days in March 1997, 0200–0600 UT, nearly simultaneous points (see description in the text).
are similar. The only distinguishable by eye difference between two curves is that 12-MHz curve appears to be shifted with respect to that at 50 MHz. The latter maximizes at $\phi \approx 30^\circ$ and the former at $\phi \approx 35^\circ$. Because of this shift, the velocity magnitudes at HF are slightly larger (smaller) than that at VHF for $\phi > 30^\circ$ ($\phi < 30^\circ$) if one compares velocities for the same value of $\phi$ (drawing vertical line through (a) and (c)). If one compares 50-MHz echoes with low-velocity 12-MHz echoes, it is evident that velocity magnitude at 50 MHz is always larger than that at 12 MHz. Thus symbolically we can write for magnitudes $V^{50} > V^{12}_{\text{low}}$ for all $\phi$, $V^{50} > V^{12}_{\text{high}}$ for $\phi < 30^\circ$, and $V^{50} < V^{12}_{\text{high}}$ for $\phi > 30^\circ$.

Finally, in the panel (e) we present the histogram of number of 12-MHz echoes at various L-shell angles $\phi$. We indicated the total number of echoes for 5'-wide L-shell angle bins by the solid line and the number of high-velocity echoes (with velocity magnitudes larger than $V_{\text{cut}}$) by the shaded area. The difference or the vertical distance between solid curve and shaded area thus represents the number of low-velocity echoes. Total number of points for this particular event is approximately the same for all directions of the FoV. The number of high-velocity points decreases steadily from $\phi = 20^\circ$ to $\phi = 90^\circ$. For $\phi > 90^\circ$ there are no high-velocity echoes. Notice also that the number of low-velocity points (the vertical difference) slightly increases with $\phi$.

In Fig. 9 we present data for the expanded period of 23 days (02-06 UT). Here we used only “nearly simultaneous” measurements (as defined in the previous section) to have diagram readable. One can notice that in Fig. 9 the relative number of points with poor calculated aspect angles is lower than in Fig. 8 and that two peaks in the distribution of velocities, corresponding to two species of 12-MHz echoes, are more distinct. The distribution of the 50-MHz velocities also becomes “smoother”, with maximum at $\approx 0$ ms$^{-1}$. The fact that VHF radar sees some echoes with low velocities is entirely due to the poor aspect angle conditions for these echoes. The peaks in Fig. 8 are less prominent (two at 12 MHz and one at 50 MHz) than in Fig. 9 because of the larger relative number of echoes with poor aspect angles. In Fig. 9e one can also observe that the number of high-velocity echoes decreases with $\phi$, number of low-velocity echoes is almost the same for all $\phi$, except for some difference in eastern and southern sectors.

10. Power and width versus L-shell angle

Since HF echoes can be of two types, we decided to look more carefully at the L-shell variation for the echo power of each type, and, in addition, we study whether there is any difference between spectral widths of these two HF echo types.

Figure 10 shows the L-shell angle variation of the spectral (a) power and (b) width. We used here again small 30-min database for all, not joint points. As in Fig. 7 we labelled each point according to its magnitude of Doppler velocity: red (blue) dots are echoes with velocity magnitude larger (smaller) than $V_{\text{crit}}$. Large red dots and blue crosses in Fig. 10 are averaged values of the spectral power and width calculated using 5'-wide L-shell angle bins for high- and low-velocity HF echoes, respectively. Also in Fig. 10 we show standard deviations and numbers of points used for averaging for each bin and each type of echoes. The average value was not plotted if the number of points in a bin was too small. Direction corresponding to the boundary between two HF radars is at $\phi \approx 80^\circ$ and we connected the average values for each radar separately.
One can notice from Fig. 10 that high-velocity echoes (red dots) tend to have larger power than low-velocity echoes and that there is a tendency for the power to decrease with L-shell angle. The power of low-velocity echoes is approximately the same for all $\phi$, albeit the spread is significant here.

The high-velocity echoes have the same or somewhat smaller widths as the low-velocity echoes for $\phi < \sim 30^\circ$. For $30^\circ < \phi < 60^\circ$, the high-velocity echoes have larger widths. This change occurs because of a general increase in the width of HF high-velocity echoes with $\phi$. Width of the low-velocity echoes is slightly increased at perpendicular directions, $\phi \sim 90^\circ$.

In Fig. 11 we present results for the entire database of joint points (23 days for 02–06 UT). These diagrams generally support our conclusions drawn from the 30-min interval. One can see here more clearly that the power of HF high-velocity echoes is stronger than the power of low-velocity echoes, and that the Syowa East echoes are typically stronger than Syowa South ones, in agreement with the data presented in Fig. 5. We can state that...
a quasi-homogeneous power-azimuth distribution of HF echoes discussed by Koustov et al. (2001) is rather an exception (the difference in power and echo occurrence between eastern an southern sectors of observations will be discussed later in the next section). We also can more clearly see that high-velocity HF echoes have smaller width than low-velocity echoes and the width of low-velocity echoes is about the same at all L-shell angles.

11. Discussion

In this paper we presented statistics for several characteristics of HF and VHF auroral echoes observed from the same location at Syowa, Antarctica. 23 days of two-frequency measurements in March of 1997 were considered. Our main goal was to explore variations of echo characteristics with the azimuth/L-shell angle of observations. More detailed analysis was performed for the morning sector (westward electrojet), between 02 and 06 MLT, where we had about 92 hours of measurements.

We demonstrated that both HF and VHF auroral echoes at short ranges of \( r < 800 \text{ km} \) occur more frequently during the nighttime, in agreement with previous studies, e.g.
McNamara (1972), Hanuise et al. (1991), and Ogawa (1996). We then considered echo occurrence for various azimuths/L-shell angles and found that the distributions were quite different at HF and VHF; VHF echoes were more frequent along the L-shell directions while HF echoes were detected at all azimuths with non pronounced maxima for L-shell angles around 45° and 115°.

How to explain this difference? One would expect predominant detection of coherent echoes along the L shells for the period under study since it is the predominant direction of the electrojet flow, and the F-B instability produces stronger plasma fluctuations along such a direction (e.g., Haldoupis, 1989). This agrees well with the typical velocities observed along L shells for the morning sector, of the order of 350-400 m s⁻¹ (Fig. 6), which is close to the threshold of the F-B instability, and with the fact that the power of 50-MHz echoes is much stronger for these directions (Fig. 5). Our HF observations, however, clearly disagree with this expectation.

One can think that the G-D plasma instability can also be responsible for some of the observed HF (and VHF) echoes. It is well known that the G-D instability gives about the same level of plasma fluctuations at all flow angles and so one would not expect any preferential direction for the echo power/echo occurrence in this case (Haldoupis, 1989; Schlegel, 1996). This prediction is consistent with what our HF observations show except of the existence of 45°/115° preferential directions. One might conclude that at decimeter scales (HF echoes) the F-B instability is not so efficient as at meter scales (VHF echoes) and leads to more homogenous flow angle distribution of density fluctuations.

Still, the issue on the preferential direction of HF echoes cannot be accepted as resolved. From one side, numerical simulations of Janhunen (1994) showed that the power of the F-B waves maximizes not along the flow but at an angle of ~45°. The nonlinear theory by St.-Maurice and Hamza (2001) gives essentially the same result. But on the other hand, why the effect is important at decimeter and not important at meter scales is still unclear.

One can also think about possibility of difference in average heights of HF and VHF echoes as discussed by Koustov et al. (2001) and Makarevitch et al. (2001). In this case, since both the F-B and G-D instabilities are easier to excite along the vector of the relative electron-ion drift, one might have slightly different preferential directions for the instabilities to occur at various heights. One may have up to ~10° azimuthal difference in this case. This is not enough to explain the observed ~25° deviation of the HF echo occurrence maxima from the direction along the flow. In addition, such explanation can be accepted for the Syowa East morning observations but not for the Syowa South observations, for which the direction of current rotation is opposite to the required one.

In our opinion, the preferential detection of HF echoes at ~45°/115° directions indicates that the processes of VHF and HF echo formation might be not exactly the same. The fact that the power of HF echoes did not show any specific maximum at L-shell angles of 45° tells us that the additional sources of HF echoes, if they exist, provide about the same power of echoes as the traditional F-B and G-D instabilities and identification of these additional sources is not possible by looking just at the power of coherent echoes.

In the past, Dimant and Sudan (1997) developed the theory of the thermo-diffusion instability at the electrojet bottom side. This instability is more easily excited at an angle of ~45° with respect to the flow. Thus it could effectively provide a shift of the echo...
occurrence maximum both at VHF and HF. However, since this instability is more efficient at decimeter scales, the effect can only be seen in the occurrence of HF echoes. Other potential sources of irregularities at 90–100 km heights are the neutral wind and neutral turbulence (Kagan and Kelley, 1998, 2000; Gurevitch et al., 1997; Ogawa et al., 2001). Irregularities produced through these processes are controlled by neutral wind that might provide some preferential direction for the echo detection though the question requires further consideration.

The performed in this study velocity comparison supports our major hypothesis that HF coherent echoes might have additional sources. We showed, similar to Makarevitch et al. (2001), that there are two distinct populations of 12-MHz echoes, the high- and low-velocity echoes. New finding of the present study is that these echo populations exhibit completely different dependencies of their Doppler velocity upon the L-shell angle.

We found that low-velocity 12-MHz echoes exist for all radar directions. Their Doppler velocity variation with the L-shell angle, Figs. 8 and 9, can be roughly described by the cosine law if one assumes the maximum velocity of the order of 100–150 m s⁻¹. This would suggest that the G-D instability or secondary F-B waves are responsible for these echoes. However, velocities of these echoes is significantly, up to a factor of 2, smaller than the velocities of simultaneously observed VHF echoes. Not less important is the fact of fairly fixed values of low-velocity population maxima, 100-150 m s⁻¹. These values are suspiciously close to the neutral wind velocities observed at auroral latitudes, e.g. Tsunoda (1988). And finally we failed to reveal an isolated low-velocity echo population in 50-MHz data, Figs. 8 and 9. Here VHF low-velocity echoes seem to be an integral part of the scatter plot.

We have little doubt about the nature of the high-velocity VHF and HF echoes. These are related to the Farley-Buneman plasma instability. Such echoes were observed mostly along the L shells. Their power was generally larger than that of low-velocity echoes, and it was decreasing with the L-shell angle. We also demonstrated using larger data set, Fig. 11, that the high-velocity echoes were more narrow.

One important result concerning these echoes is the velocity change with the L-shell angle, Figs. 8 and 9. The effect is well seen at both HF and VHF. It was thought for a long time that type 1 echoes observed at small flow angles, inside the F-B instability cone, do not exhibit any velocity variation with flow angle (e.g., Nielsen and Schlegel, 1983). Recently Nielsen et al. (2002) showed that the phase velocity for VHF radar frequency of 140 MHz is maximized at zero flow angle and slowly decreases up to flow angle of 40° where it is equal to the ion acoustic speed. Similar feature at HF was discussed recently by Uspensky et al. (2001). In another study by Milan and Lester (2001) it was shown that some HF echoes (type 3 and 4 in their notation) exhibit quite peculiar dependence of the Doppler velocity upon the L-shell angle. Velocity magnitudes of such echoes maximized neither along nor perpendicular to the flow at L-shell angles of 25–35°. We found, Fig. 8, that both VHF and HF velocity magnitudes are largest at the L-shell angles of ~30°, and, in this sense, our results support strongly the result of Milan and Lester (2001). We think, however, that the velocity maximum shift from L-shell angle of 0° is most likely due to some departure of the flow from the L-shell direction, rather than due to some plasma-physical mechanism. Indeed, in the morning sector and at auroral latitudes one might expect additional “curving” of the flow towards the magnetic pole for the two-cell
convection pattern. If the amount of flow curving for the eastern part of FoV was of the order of 30° then our results are in good agreement with the results of Nielsen et al. (2002). One can ask why the low-velocity HF echoes do not show larger velocities for the L-shell angle about 30°. In our opinion, the reason is that these echoes are originated not only from the F-B and G-D instabilities but also from other instabilities that are not so much controlled by the ionospheric electric field.

The rotation of the flow from the direction of magnetic L shells could also be the reason for very small number of the high-velocity HF echoes for the Syowa South radar. Indeed, in a case of exactly L-shell aligned flow, the distributions in Figs. 8 and 9 would be symmetric with respect to the direction ϕ = 90°; both HF radars would detect almost the same number of the high-velocity echoes. The current rotation (electric field rotation) of the order of 30° would shift this direction towards larger L-shell angles (ϕ ≈ 120°) and would shift the F-B instability cone orientation for the Syowa South radar (ϕ ≈ 150°). In this situation, the detection of high-velocity echoes here would be possible only at azimuths larger than 183°, beyond the Syowa South FoV. In the consideration above we assumed that electric field magnitudes are about the same in the Syowa South and Syowa East FoVs. Thus it is not the L-shell angle that controls the generation of the high-velocity HF echoes, but rather the electrojet/electric field direction (and certainly its magnitude).

Statistically, the power measured by the Syowa East radar was larger than the power measured by Syowa South, Fig. 5, by 6–8 dB. This difference between the power of two radars is also obvious in Fig. 11a and for echo occurrence in Fig. 4 where the number of echoes is significantly lower for the Syowa South radar. This fact, unnoticeable on the basis of the 30-min observational period alone (Plate 1c of Koustov et al. (2001) and our Fig. 10a), is most likely due to the fact that the Syowa South radar operated under different technical and observational conditions. It was discovered that the Syowa South radar had higher noise level in the receiver system. Also, to avoid interference with the Syowa East radar (FoVs of both radars slightly overlap) the base sounding frequency for Syowa South was chosen to be ~11 MHz versus ~12 MHz for the Syowa East. The difference between radar frequencies could result in different amount of ray bending and hence to slightly different altitudes of scatter (Koustov et al., 2001; Makarevitch et al., 2001). For these reasons, all conclusions of the present study were drawn on the basis of observations from the Syowa East radar and we used Syowa South data only as a reference.

Finally, we have to note that no special attention has been paid in this study to meteor echoes. It is accepted that their contribution to echo detection at Syowa is not negligible (e.g., Ogawa et al., 1985, 2001). In the previous studies by Fukumoto et al. (1999, 2000), Makarevitch et al. (2001) and Ogawa et al. (2001) attempts have been made to exclude meteor echoes from the data base. Such approach is certainly successful in eliminating isolated in time and space reflections (that are more likely to be truly meteor reflections) but not reflections occurring simultaneously with the auroral ionospheric scatter. Contrary to this approach, in the present study we considered all available echoes at short ranges. We expect, consequently, that some of the low-velocity VHF and especially HF echoes were, in fact, meteor scatter.

Hall et al. (1997) reported on correlation of HF Doppler velocities of short range echoes and neutral winds measured independently by a MF radar. These authors assumed that all detected echoes were meteor scatter. We believe that in our observations there was
a mixture of ionospheric and meteor echoes with predominance of ionospheric echoes. This is supported by several facts. Syowa echoes were more frequent at 1–2 UT and not in the late morning hours and, when occurred, they typically had quite an extensive spatial and temporal coverage. We also analyzed HF/VHF data in the evening sector and found results similar to the ones presented in the present paper for the morning sector. Additional support comes from the revealed features of echo characteristics, for example, velocities were quite often large, >100 m s$^{-1}$, more than typical meteor echo velocities, and the HF spectra were quite broad. We do believe, however, that low-velocity echoes were strongly affected by neutral particle motions, and in this sense our results are in agreement with Hall et al. (1997).

12. Conclusions

In this study a statistical comparison between echo characteristics at two significantly different radar frequencies (12 and 50 MHz) has been performed. We considered one month of nearly simultaneous two-frequency measurements in a broad range of flow and aspect angles. We concentrated on the azimuth/L-shell variation of such echo parameters as echo occurrence, spectral power, mean Doppler shift and spectral width.

We showed that overall echo occurrence within the radars' common FoV varied with the time of the day in a similar manner at both radar frequencies and had maxima at post-midnight hours. However, when specific radar directions were compared, it was found that the echo occurrence depended on the azimuth/L-shell angle of measurements. We explored the nature of this difference using two approaches.

First, we considered data for the morning sector, for which the echo occurrence was the highest, and found that while VHF echoes were more frequently detected along the L shells, HF echoes were observed at all L-shell angles with some broad maximum at angles around 45° and 115°. Azimuthal distribution of the power was found to have a strong maximum along L shells at VHF and to be more or less homogeneous at HF.

We then turned our attention to the velocities of echoes at various L-shell angles. We showed that HF echoes can be divided into two populations, the high- and low-velocity echoes, while VHF echoes are of only one kind, the high-velocity echoes and their aspect-attenuated counterparts. We also explored the behavior of the HF power and spectral width with L-shell angle for these two echo populations and showed that the high-velocity echoes are stronger and more narrow that the low-velocity ones.

We related the high-velocity VHF and HF echoes to the F-B plasma instability. The low-velocity HF echoes were associated with several possible sources. We argued that these echoes are originated not only from scatter on secondary F-B and G-D waves but also on irregularities associated with other plasma processes in the auroral E region. We assumed that the thermo-diffusion instability at the bottom of the electrojet layer or instabilities originated from the neutral wind contribute significantly to the formation of the decameter scale irregularities (HF echoes).

Acknowledgments

A.V.K. acknowledges the Solar-Terrestrial Environment Laboratory of the Nagoya
University for funding during his stay in Japan. The work was also supported by an NSERC grant (Canada) to A.V.K. The authors thank M. Lester, J.-P. St.-Maurice and D. Danskin for discussions of the results and C. Haldoupis for criticism.

The editor thanks Dr. R. Fujii and another referee for their help in evaluating this paper.

References


(Received December 20, 2001; Revised manuscript accepted May 8, 2002)
E region echoes observed with the Syowa HF radar under disturbed geomagnetic conditions

Tadahiko Ogawa\textsuperscript{1}, Nozomu Nishitani\textsuperscript{1}, Natsuo Sato\textsuperscript{2}, Hisao Yamagishi\textsuperscript{2} and Akira Sessai Yukimatu\textsuperscript{2}

\textsuperscript{1}Solar-Terrestrial Environment Laboratory, Nagoya University, Honohara, Toyokawa 442-8507
\textsuperscript{2}National Institute of Polar Research, Kaga 1-chome, Itabashi-ku, Tokyo 173-8515

Abstract: E region radar echo parameters (echo power, Doppler velocity and spectral width) obtained with the Syowa Station HF radar under disturbed geomagnetic conditions are qualitatively analyzed to study how echoing region changes due to HF wave refraction caused by ionospheric disturbance. It is found that with increasing disturbance level, echo ranges become shorter because of wave refraction during propagation due to more enhanced D and E region electron density. When geomagnetic H component variation ($\Delta H$) is less than about $-900$ nT, echoes are returned from the central E region where geomagnetic aspect angle is close to zero. When $\Delta H$ is very high ($\approx -1500$ nT), the echoes are backscattered from the D and lower E regions and their power, Doppler velocity and spectral width are largely suppressed. The results suggest that we must always consider, more or less, wave refraction effect in analyzing near-range E region HF radar echoes.

1. Introduction

The Super Dual Auroral Radar Network (SuperDARN) HF radars can detect coherent echoes backscattered by field-aligned irregularities at slant ranges from 180 to 3000 km (or more), that is, at altitudes including the ionospheric D, E and F regions (Greenwald \textit{et al.}, 1995). A Japanese Antarctic station, Syowa Station (69.0°S, 39.6°E; magnetic latitude 66.2°S; L-value= 6.1), has been operating two SuperDARN radars, called the “Syowa South” and “Syowa East” radars. E region echoes detected with these radars have been analyzed in detail by Kouostov \textit{et al.} (2001) and Makarevitch \textit{et al.} (2001, 2002) who used simultaneous data from a VHF auroral radar at Syowa Station. They compared Doppler velocity and echo power from the HF radars with those from the VHF radar to discuss similarity and dissimilarity between both radar echoes. One of the findings is that HF wave refraction during propagation must be considered in explaining HF echo characteristics (Villain \textit{et al.}, 1984; Uspensky \textit{et al.}, 1994). E region echoes at northern high latitudes have been studied by, for example, Uspensky \textit{et al.} (2001) who also pointed out the important role of ionospheric refraction.

Ogawa \textit{et al.} (2001) have used a large data-set of Doppler velocity obtained for 11 months in 1997 with the Syowa East radar under various geomagnetic conditions to discuss statistical characteristics of the velocity at slant ranges of 180–1200 km and their implications. They found that on average the velocity has a minimum of about 100 m/s at 180–225
km ranges and that it increases monotonically with range to attain a maximum of 300–350 m/s at 400–500 km. Such a range profile is caused by the combined effects of altitude-dependent phase velocities of ionospheric plasma waves and HF wave refraction, and partly by meteor winds. Ogawa et al. (2001) speculated that the low velocity (∼100 m/s) at ranges of 180–225 km may originate in part from neutral winds and/or turbulence of the neutral atmosphere. Ogawa et al. (2002) have found that polar mesosphere summer echoes are another candidate for the low velocity at ranges of 180–315 km. However, other possible reasons to explain the low velocity at short ranges have not been explored.

Refraction of an HF radar wave becomes more severe with increasing electron density in the D and E regions. The density is enhanced during disturbed geomagnetic conditions (e.g., Miyazaki et al., 1981), thus suggesting that HF-scattering area and altitude change depending on geomagnetic disturbance level. Such change has been never considered, in quantitative manner, in the previous E region HF echo studies that pointed out refraction effect. This paper is concerned with E region echoes detected with the Syowa East radar under weakly, strongly and very strongly disturbed geomagnetic conditions, and aims to show qualitatively how the echo region shifts toward the radar due to wave refraction under disturbed conditions.

2. Syowa East radar

Figure 1 displays field of view (FOV) of the Syowa East radar in geographic coordinates. The FOV is covered with 16 narrow beams (beam numbers 0, 1, 2, ..., 15) over an azimuth sector of 52°. Each beam width is believed to be around ∼4.5° for the radar frequencies (10.2–12.2 MHz) used in our observations (Greenwald et al., 1985). The beam direction of beam 0 is almost perpendicular (80°) to the magnetic L-shells and that of beam 15 has an angle of 20° to them. The radar beam in the normal SuperDARN operation is sequentially scanned from beam 0 to beam 15 (16 azimuth bearings) with a step in azimuth of 3.3°, a scan repeat time of ∼120 s, a range resolution of 45 km, and a peak power of around 10 kW. The first range gate is set to 180 km.

Figure 1 also depicts contour lines (solid curves) of the angle (A) between the radar wave vector and the geomagnetic field vector (IGRF95) at an altitude of 110 km around which auroral electrojet intensity is strongest and E region irregularities are most enhanced (e.g., Ogawa et al., 1976). Note that the geomagnetic aspect angle α is defined as 90° − A. It is well known that coherent radar echoes are strongly backscattered from an area where A is between between 89° and 91° and that the strongest backscatter occurs at A = 90°. Figure 1 indicates that such area is located at ground range of around 275 km (slant range of 290 km) on beam 0 and of around 500 km (slant range of 508 km) on beam 15. Previous VHF (50 MHz) radar observations at Syowa Station indicated that echoes on a radar beam nearly perpendicular to the L-shells are strongest at a range of around 280 km (Ogawa and Igarashi, 1982; Ogawa et al., 1989; Koustov et al., 2001). Note that no radar wave refraction during propagation is assumed in the calculations. This assumption is almost valid for VHF waves but not always valid for the SuperDARN HF waves (8–15 MHz) because of possible wave refraction due to enhanced electron density (Villain et al., 1984).

The aspect angle (α) distribution shown in Fig. 1 varies depending on altitude.
Figure 2 plots some ray paths (elevation angles of 4°–30°) from the Syowa radar in altitude-ground range coordinates (Ogawa et al., 2001). The area having $\alpha$ between $-1^\circ$ and $+1^\circ$ is surrounded by the thick solid curves. On beam 0 (Fig. 2a) the radar waves with elevation angles of 20°–23° are backscattered from altitudes of 100–120 km at slant ranges of 250–320 km, while on beam 15 (Fig. 2b) these values are 7°–13° and 400–650 km. The path length crossing the $E$ region altitude becomes longer with increasing beam number.

Radar echo intensity depends on echo range, cross-section of a target, radar antenna pattern and so on. The vertical pattern of the SuperDARN antenna array is quite broad with an average half power beamwidth of $\sim 30^\circ$, and at 8 MHz the sensitivity maximizes at an elevation angle of $35^\circ$ whereas at 20 MHz this value decreases to $15^\circ$ (Greenwald et al., 1985): the sensitivity at angles larger than $60^\circ$ is believed to be very low. This means that the Syowa HF antenna has a maximum gain at around $28^\circ$–$30^\circ$ at 10–12 MHz, meaning that $E$ region echoes are easily detected on low beam numbers.

In this paper we use data of echo power, Doppler velocity and spectral width obtained
Fig. 2. Ray paths with elevation angles of 4°–30° on (a) beam 0 and (b) beam 15 of the Syowa East radar in altitude-ground range coordinates. The area having $\alpha$ between $-1^\circ$ and $+1^\circ$ is surrounded by the thick solid curves. Slant ranges (100 km step) from Syowa Station are also indicated.

by analyzing an auto-correlation function of radar echoes. In particular, the power and width were calculated by assuming an exponential fit to the auto-correlation function, which corresponds to a Lorentzian Doppler spectrum (e.g., Hanuiise et al., 1993).
3. Observations

3.1. Weak disturbance (maximum $\Delta H \approx +100$ nT)

Figure 3 displays range-time variations of the echo power on beams 1, 7 and 14 on January 4, 1998 and the geomagnetic $H$ component (1-min average) observed with a magnetometer at Syowa on that day. The radar frequencies of 10.2–10.6 MHz were used for the observations. The $H$ component is rather quiet ($\Delta H \approx 0$) before 1700 UT (UT $\approx$ MLT and LT = UT + 3 hours at Syowa). At 1850 UT when $\Delta H$ increases to about $+60$ nT, the radar starts to detect echoes: the strongest echoes are located at slant ranges of around 270–360 km on beam 1, 315–405 km on beam 7 and 405–540 km on beam 14. Judging from Figs. 1 and 2, we conclude that these echoes are returned from the $E$ region altitudes of 100–120 km.

With gradually increasing $\Delta H$ due to the growing eastward electrojet, the echo ranges shift toward the radar. At 2015 UT when $\Delta H$ is most enhanced ($\approx +100$ nT), the strongest echoes exist at 270–315 km on beams 1 and 7 and 315–405 km on beam 14, suggesting that the ray paths, in particular, for higher beam number, from the radar to echo targets in the $E$ region are not straight but refracted: we speculate from Fig. 2 that the rays with higher elevation angles are bent downward during propagation in the $D$ and lower $E$ region to satisfy the perpendicularity condition of $\alpha = 0^\circ$ in the central $E$ region. With decreasing $\Delta H$ after 2015 UT, the echo ranges are unchangeable until 2120 UT, after which those on beams 7 and 14 tend to shift a little toward farther ranges. The echoes disappear at around 2135 UT when $\Delta H$ is $+50$ nT. Note that the echoes on beams 1 and 7 at ranges beyond 400 km after 1940 UT are backscattered from the $F$ region.

3.2. Strong disturbance (maximum $\Delta H \approx -900$ nT)

Range-time variations of the echo power on beams 1, 7 and 14 on January 6–7, 1998 are shown in Fig. 4. The radar frequencies varied between 10.2 and 11.5 MHz during the observations. Note that the radar operation was stopped between 1520 and 1600 UT and between 1630 and 1730 UT. The Syowa magnetogram indicates the following $\Delta H$ excursion: quiet before 1530 UT, almost positive (with a maximum of $+200$ nT at 1640 UT) until 0000 UT and negative (with a maximum of $-900$ nT at 0320 UT) until 0730 UT. The movement of the $E$ region echo ranges before 0000 UT is quite similar to the case shown in Section 3.1. The $F$ region echo appears at ranges beyond 450 km. After 0000 UT, with increasing negative excursion of $\Delta H$ (caused by the growing westward electrojet), the echo ranges where the echo intensity has a maximum move toward the radar (the minimum ranges are 225–270 km on all the beams): this movement is larger on higher beam numbers. This fact suggests that the radar wave refraction is more severe for stronger disturbance and higher beam numbers (longer propagation distance).

To investigate in detail echo characteristics during this strong $\Delta H$ disturbance, Figs. 5a, 5b and 5c show range-time variation of the echo power, Doppler velocity (positive toward the radar) and spectral width, respectively, obtained on beam 14 for which the refraction effect is strongest. The $\Delta H$ variation is also displayed in the figure (Fig. 5d). Fig. 5a clearly demonstrates that between 0100 and 0630 UT when $\Delta H$ is disturbed, the strongest echoes appear at ranges of 225–405 km and moreover that the movement of the echo regions is in harmony with the $\Delta H$ variation, that is, the regions move toward (away
Fig. 3. Slant range-time plots of echo power on beams 1, 7 and 14 and geomagnetic H component (1-min average) on January 4, 1998.

from) the radar with increasing (decreasing) $\Delta H$.

The velocities (Fig. 5b) are positive ($\leq +400$ m/s) before 0100 UT and negative ($\leq -600$ m/s) after 0100 UT, corresponding to the westward electron drifts (eastward electrojet) and the eastward electron drifts (westward electrojet), respectively. The radar echo subsidence at around 0000 UT is caused by very weak electric fields (responsible for the irregularity production) near the Harang discontinuity. It is recognized that after 0100 UT
the velocity is faster for stronger echo intensity, a well-known fact for the $E$ region irregularities (e.g., Ogawa and Igarashi, 1982; Milan and Lester, 2001; Makarevitch et al., 2001). The spectral width behavior (Fig. 5c) is interesting. The width after 0200 UT is narrower (wider) for stronger (weaker) echo intensity and faster (slower) velocity. These facts are clearly demonstrated in Fig. 6 in which time variations of the echo power, Doppler velocity and spectral width at the range gate of 270–315 km between 0420 and 0540 UT are

Fig. 4. Same as Fig. 3 but for January 6–7, 1998. Note that the radar operation is stopped between 1520 and 1600 UT and between 1630 and 1730 UT.
Fig. 5. Slant range-time plots of (a) echo power, (b) Doppler velocity (positive toward the radar) and (c) spectral width on beam 14 and (d) geomagnetic H component (1-min average) on January 6–7, 1998.
Fig. 6. Time variations of echo power, Doppler velocity and spectral width at 270-315 km range on beam 14 between 0420 and 0540 UT on January 7, 1998.

displayed. These characteristics are also seen in previous VHF radar observations (e.g., Ogawa and Igarashi, 1982) and HF radar observations (e.g., Milan and Lester, 2001).

3.3. Very strong disturbance (maximum $\Delta H \approx -1500\,\text{nT}$)

A very strong $\Delta H$ disturbance occurred on June 8–9, 1997. Range-time variations of the echo power on beams 0, 7 and 15 and the H component are shown in Fig. 7. The radar frequencies were between 11.2 and 12.2 MHz. $\Delta H$ is quiet before 1630 UT on June 8 and after 1400 UT on June 9. Before 0430 UT and after 0615 UT stronger $E$ region echoes appear at 225–405 km on beam 0, 225–360 km on beam 7 and 225–450 km on beam 15. During these time periods the echo regions occasionally extend to farther ranges beyond 600 km, indicating that these echoes are returned from the $F$ region: there is no demarcation area between the $E$ and $F$ region echoes, contrary to the cases shown in Figs. 3 and 4.

A sharp decrease in $\Delta H$ of about $-1500\,\text{nT}$ is discernible at around 0430 UT. Corresponding to this decrease, the echoes between 0430 and 0515 UT are very weak and are localized within 180–315 km (note that the actual minimum echo range is unknown because of the first range gate of 180 km). Such features are not seen in Figs. 3 and 4. We speculate that largely enhanced $E$ and $D$ region electron density, due to high-energy particle precipitation associated with the strong substorms, caused complete absorption or forward reflection of the radar waves with lower elevation angles that would be backscattered at ranges beyond 315 km without the substorms.

Figure 8 shows range-time variations of the echo power (Fig. 8a), Doppler velocity
Fig. 7. Same as Fig. 3 but for June 9, 1997.

(Fig. 8b) and spectral width (Fig. 8c) on beam 15 together with the $H$ component (Fig. 8d). Between 0430 and 0515 UT the echo powers are less than 18 dB, most velocities are less than $-80$ m/s (note the maximum velocity of $-400$ m/s before 0430 and after 0515 UT) and the spectral widths at 180-225 km are less than 40 m/s. Such characteristics strongly suggest that the echoes are not returned form the central $E$ region but from the lower $E$ and $D$ regions. The echoes may partly include meteor echoes, usually occurring
Fig. 8. Same as Fig. 5 but for June 9, 1997.

at 80–100 km altitudes, that have similar velocity and spectral width (Hall et al., 1997). Note that the meteor echo occurrence at Syowa Station has a maximum at around 0300 UT (0600 LT) and a minimum at around 2000 UT (2300 LT) (Ogawa et al., 1985).
4. Discussion

We have examined $E$ region echo characteristics under three kinds of geomagnetic disturbances, that is, weak disturbance (maximum $\Delta H \simeq +100$ nT), strong disturbance (maximum $\Delta H \simeq -900$ nT) and very strong disturbance (maximum $\Delta H \simeq -1500$ nT). The main results are as follows:

1) Under weakly disturbed conditions ($\Delta H \leq +100$ nT), strong echoes are returned from the central $E$ region. Radar wave refraction during propagation is negligible for lower beam numbers and a little more significant for higher beam numbers, that is, for longer propagation distance to the $E$ region.

2) Refraction effect becomes more pronounced with increasing $\Delta H$. Again, the refraction is most significant for higher beam numbers. Under a very strongly disturbed case (Fig. 7), echo ranges on all beams are limited to within 180 (first range gate)–315 km: in this case, as shown in Fig. 8, echo power, velocity and spectral width are all greatly reduced.

There are two kinds of plasma waves that produce field-aligned $E$ region irregularities. See papers by Haldoupis (1989), Ogawa et al. (2001) and references therein for more details of the $E$ region plasma instabilities. In the central $E$ region phase velocities of the gradient-drift (cross-field) waves are almost equal to $E \times B$ drift and those of the two-stream waves saturate at the local ion acoustic velocity. Below 100 km altitude the gradient-drift phase velocities are more reduced with decreasing altitude and the two-stream waves cannot be excited. Dimant and Sudan (1995) have proposed a new plasma instability that is probably operative below 90 km and has phase velocities less than 100 m/s. This instability, however, requires an electric field exceeding 50 mV/m ($E \times B$ drift of about 1000 m/s).

Relying on these considerations and current observations, Fig. 9 schematically illustrates simplified wave propagation modes (A, B and C) in the $E$ region under different disturbed conditions. In mode A for less disturbed condition, ray paths are almost straight, in particular, for low beam numbers, to achieve perpendicularity ($\alpha = 0^\circ$) in the central $E$ region. Mode B becomes dominant under disturbed condition and for higher beam numbers under weakly disturbed conditions; that is, the rays are refracted downward in the lower $E$ region (because of enhanced electron density) to achieve $\alpha \approx 0^\circ$ in the central $E$ region. Propagation distance (radar range) in this mode should be shorter than that in mode A. In modes A and B, echo power, Doppler velocity and spectral width are not suppressed because the condition $\alpha = 0^\circ$ is satisfied in the central $E$ region.

In mode C for highly disturbed condition, rays are refracted in the $D$ and lower $E$ regions due to greatly enhanced electron density. The rays, however, cannot arrive at the central $E$ region and are backscattered from the lower $E$ region where $\alpha$‘s is probably close to zero. As described above, below 100 km altitude the gradient-drift phase velocity deviates from $E \times B$ drift and decreases rapidly with decreasing altitude. In addition, we can expect that plasma turbulence below 100 km altitude is weaker than that in the central $E$ region, resulting in lower echo power and narrower spectral width below 100 km. The plasma instability proposed by Dimant and Sudan (1995) (see above) might be operative for our strongly disturbed case. Thus, mode C seems to well explain the above item 2: if echoes are returned from the central $E$ region (as in modes A and B), then echo power, Doppler velocity and spectral width are not reduced. Note that Doppler velocity is also
Fig. 9. Schematic illustration of HF wave propagation mode in the E region under weakly (A), strongly (B) and very strongly (C) disturbed conditions.

reduced when backscatter occurs at $\alpha=0^\circ$ in the central E region (e.g., Ogawa et al., 1980, 1982; Fejer et al., 1984).

The importance of wave refraction for the SuperDARN radars has been pointed out by Villain et al. (1984), Uspensky et al. (1994, 2001), Milan and Lester (1998), Koustov et al. (2001) and Makarevitch et al. (2001). This is confirmed by the current results. In our case, with increasing disturbance, the echo ranges become closer to the radar, meaning that ray paths should have higher elevation angles. Such an effect together with the vertical antenna pattern of the radar must be taken into account in considering the echo power reduction at near-ranges.

For a target at a range of 180 km and an altitude of 100 km (Fig. 8), the elevation angle is $34^\circ$ under no wave refraction. When a radio wave (frequency: $f$) enters an ionospheric plasma (plasma frequency: $f_p$) from below with an elevation angle of $\theta_i$, its path is bent downward to have an elevation angle of $\theta_e$. Snell's law tells that for $f=11$ MHz and $\theta_i=40^\circ$, $\theta_e$ is $35^\circ$, $31^\circ$, $24^\circ$ and $7^\circ$ for $f_p=4, 5, 6$ and $7$ MHz, respectively. Thus, it seems easy that the wave vector becomes perpendicular to the geomagnetic field somewhere in the disturbed E region.

Ogawa et al. (2001) speculated that low Doppler velocity ($\leq 100$ m/s) at short ranges (low altitudes) may be partly caused by neutral winds and/or turbulence of the neutral atmosphere. In addition, Ogawa et al. (2002) found that polar mesosphere summer echoes under quiet geomagnetic conditions have low velocity at ranges of 180–315 km. The present results suggest that the low velocity is also caused by field-aligned irregularities in the lower E and D regions under highly disturbed conditions.

5. Conclusions

Case studies of E region HF radar echoes for a range of geomagnetic conditions have been presented. With increasing disturbance level, echo range becomes shorter because of
wave refraction during propagation due to more enhanced $D$ and $E$ region electron density. From analysis of power, Doppler velocity and spectral width of echoes, we have found that the echoes are returned from the central $E$ region, where $\alpha$ is close to zero, when geomagnetic disturbance is less than about $-900$ nT. When highly disturbed ($\approx -1500$ nT), the echoes are returned from the $D$ and lower $E$ regions, resulting in suppression of echo power, Doppler velocity and spectral width. Our results suggest that we must always consider, more or less, the wave refraction effect in analyzing near-range HF radar echoes for the $E$ region study. In the future we need more sophisticated radar operation (shorter first-range gate and higher range resolution) to make more detailed investigations of refraction processes and scattering altitudes. HF ray-tracings using realistic electron density profiles in the $D$ and $E$ regions are necessary to know how the refraction is and where it occurs.

Acknowledgments

We thank all the staff who contributed to the HF radar observations at Syowa Station. This research is supported by the Grant-in Aid for Scientific Research (A:11304029) from Japan Society for the Promotion of Science (JSPS). The Ministry of Education, Culture, Sports, Science and Technology supports the Syowa HF radar system.

The editor thanks Drs. T. Maruyama and M. Pinnock for their help in evaluating this paper.

References


(Received January 24, 2002; Revised manuscript accepted June 6, 2002)
A new meridian imaging spectrograph for the auroral spectroscopy

Makoto Taguchi\textsuperscript{1}, Shoichi Okano\textsuperscript{2}, Takeshi Sakanoi\textsuperscript{3}, Naoko Koizumi\textsuperscript{3}, Takehiko Aso\textsuperscript{1} and Masaki Ejiri\textsuperscript{1}

\textsuperscript{1}National Institute of Polar Research, Kaga 1-chome, Itabashi-ku, Tokyo 173-8515
\textsuperscript{2}Planetary Plasma and Atmospheric Research Center, Tohoku University, Aoba, Aramaki, Aoba-ku, Sendai 980-8578

Abstract: Spectroscopic and monochromatic imaging observations of emissions in the upper atmosphere are mutually complementary. A meridian imaging auroral spectrograph (ASG) that can measure a spectrum in the visible region along a meridian has been developed for research on the auroral physics and the polar upper-atmosphere dynamics. Combination of a fast optical system inherited from a monochromatic all-sky imager, a grism as a dispersive element, and a cooled CCD camera has enabled a wide field-of-view of 180° along a meridian, spectral coverage of 420–730 nm, spectral resolution of 1.5–2.0 nm, and high sensitivity to be obtained. The absolute sensitivity over a full-image field was calibrated using a National Institute of Standards and Technology (NIST) traceable integrating sphere and determined to be 0.06 cts/s/R at a wavelength of 560 nm at the zenith. The ASG was installed at Longyearbyen in March 2000, and routine operation was started in the 2000/2001 winter season. An example of an auroral spectral image is presented in this report to demonstrate the performance of the ASG.

1. Introduction

Spectroscopic investigations have provided significant information on the sources and mechanism of auroral emissions, energy distribution and the flux of auroral particles, and the chemical and dynamical conditions of the upper atmosphere (e.g., Rees et al., 1976; Deehr et al., 1980; Paresce et al., 1983; Ishimoto et al., 1988; Rassoul et al., 1993; Stephan et al., 2001). On the other hand, optical imaging observations, both from the ground and from space platforms, have been a basic approach to researches on the auroral physics, the auroral storm dynamics, and the mapping of auroral particles, etc. (e.g., Akasofu, 1964; Anger et al., 1973, 1987; Ono et al., 1987; Frank and Craven, 1988; Oguti et al., 1990; Newell et al., 1992; Tobiska et al., 1993; Evans et al., 1994; Cummer et al., 2000). Use of highly sensitive CCD devices has drastically improved the sensitivity and time resolution of imaging instruments promoting studies on gravity wave propagations deduced from the patterns seen in successive images of faint airglow (e.g., Mendillo et al., 1997; Taylor et al., 1995). Recently, Rees et al. (2000) succeeded in obtaining an optical image of the visible aurora under near-daytime conditions using a high-resolution imaging optical spectrometer.
An optical instrument combining spectroscopy and imaging techniques makes it possible to investigate the spatial distribution and motion of spectroscopic phenomena, such as aurora and airglow. Grating spectrographs, which are capable of taking an image whose ordinate and abscissa are projections of a slit and wavelength, respectively, have been developed for this purpose. Okamura and Ejiri (1992) applied this technique to auroral observations and studied the differences in the horizontal distribution and the temporal variation of auroral emissions. Swenson et al. (1998) also discussed the vertical profiles of auroral emissions in the visible and near-IR regions using data acquired by an imaging spectrograph. Semeter et al. (1999) analyzed auroral spectroscopic data obtained by a chain of meridional imaging spectrographs to reconstruct an emission field by a tomographic method. Chakrabarti et al. (2001) presented a unique imaging spectrograph with high-spectral resolution and high throughput for ground-based observations of airglow and aurora; they described the performance of the instrument in both laboratory and field experiments. Dymond et al. (2000) developed an ultraviolet spectrograph for a sounding rocket experiment which gives height profiles of O and O₂ densities from spectral images obtained during a rocket flight.

The National Institute of Polar Research, Japan, has developed a monochromatic all-sky imager (ASI) with high sensitivity and high time resolution for studies on aurora and airglow phenomena (Okada et al., 1997). An ASI has been in operation at the South Pole Station since the 1997 austral winter season; two additional ASIs were installed at Syowa Station in 1998. One of the ASIs at Syowa Station was brought back to Japan in 1999 and modified to become an imaging auroral spectrograph (ASG) so that spectral images of aurora and airglow can be obtained. In this paper, we describe instrumentation and performance of the ASG and present example data obtained from observations at Longyearbyen to demonstrate its actual performance.

2. New auroral spectrograph

The performance requirements of the ASG are: 1) a field-of-view that covers the entire meridian, 2) spectral resolution of higher than 2 nm, 3) spectral coverage over the full-range of the visible region, and 4) sensitivity comparable to that of the ASI. Satisfying all of these requirements using a conventional optical layout with a diffraction grating is quite difficult. However, a grism or a transmission grating grooved on a prism surface can be used as a compact dispersive element with both high throughput and moderate spectral resolution, because the light diffracted by a grism does not overlap the path of the incident light. Thus, a grism was used as the dispersive element in the ASG. The main specifications of the ASG are listed in Table 1.

The optical layout illustrated in Fig. 1 is similar to that used in the ASI with the exception that a grism is used as the dispersive element in place of the ASI's interference filter. The optical system is composed of a fast fish-eye lens (f = 6 mm, F1.4), a collimator lens system, a grism, and an imaging lens system. The grism is inserted at the position where the ray becomes parallel. The grooved surface of the grism faces the fore-optics. The optical axis of the fore-optics is normal to the grooved surface. A slit with a width of 42 μm and a length of 20 mm is placed at the focal plane of the fish-eye lens.

A groove frequency of 600 gr/mm and a BK7 glass material were chosen so that the
Table 1. Specifications of the auroral spectrograph.

<table>
<thead>
<tr>
<th>Optical front-end</th>
<th>Fish-eye lens 6 mm F1.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slit width</td>
<td>42 μm</td>
</tr>
<tr>
<td>Grism</td>
<td></td>
</tr>
<tr>
<td>Dimension</td>
<td>68 mm × 68 mm</td>
</tr>
<tr>
<td>Blank material</td>
<td>BK7</td>
</tr>
<tr>
<td>Groove frequency</td>
<td>600 g/mm</td>
</tr>
<tr>
<td>Blaze wavelength</td>
<td>450 nm</td>
</tr>
<tr>
<td>Prism angle</td>
<td>26.75°</td>
</tr>
<tr>
<td>CCD</td>
<td></td>
</tr>
<tr>
<td>Number of pixels</td>
<td>512 × 512</td>
</tr>
<tr>
<td>Pixel size</td>
<td>24 μm</td>
</tr>
<tr>
<td>Quantum efficiency</td>
<td>&gt;80% in visible region</td>
</tr>
<tr>
<td>Temperature</td>
<td>−40°C</td>
</tr>
<tr>
<td>Spectral region</td>
<td>420–730 nm</td>
</tr>
<tr>
<td>Spectral resolution</td>
<td>1.5 nm @ 550 nm</td>
</tr>
<tr>
<td>Field-of-view</td>
<td>180°</td>
</tr>
<tr>
<td>Spatial resolution</td>
<td>0.37–0.98°</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.062 cts/s/R @ 550 nm</td>
</tr>
</tbody>
</table>

Fig. 1. Optics of the ASG. The optical paths are projected on a plane including the optical axes of the fore optics and the focusing optics. The meridional plane is perpendicular to this plane.

A new meridian imaging spectrograph

101

dispersed wavelength range of interest falls just inside the sensitive area of a CCD detector. Both surfaces of the grism are AR-coated. The efficiency of the grism ranges from 40% at 730 nm to more than 75% at 450 nm. These efficiencies are comparable to the peak transmission of the interference filters used in the ASIs for selection of auroral emission lines. Accordingly, the ASG is expected to have high sensitivity that is almost equivalent to that of the ASI. The prism, which contributes in part to the wavelength dispersion, functions as a beam-stirring optical component to minimize the angle between the optical axes of the fore-optics and the focusing optics.

A back-thinned bare CCD (512 by 512 pixels) can take images in a form such that the zenith angle and wavelength are aligned along the rows and lines of the CCD, respectively. The CCD is cooled down to a temperature lower than −40°C using a three-stage Peltier cooler to minimize dark noise. The pixel size of 24 μm matches the monochromatic images produced by the slit's width. Along the meridian, the CCD resolves a 180° field-of-view into approximately 500 pixels, making the spatial resolution along the meridian about 0.36°. The position and tilt of the CCD can be adjusted and fixed by small set-screws. These mechanisms make tuning of the spectral coverage and focusing easier.
Charges accumulated by the CCD are amplified and converted to 14-bit digital signals, which are processed by a computer and stored in a DVD-RAM disk. The preamplifier’s gain can be selected as either a super high gain (6 e−/bit), a high gain (30 e−/bit), or a low gain (60 e−/bit). The gain is normally set at a super high gain, but a high or low gain can be selected when a high signal-to-noise ratio and a wide dynamic range are required for bright auroras.

3. Calibration

3.1. Spatial resolution

The spatial resolution across the meridian was experimentally determined as follows. A rail was placed in such a way that the rail was perpendicular to a plane containing the field-of-view of the ASG. Images of a small spectral lamp, which could be moved along the rail, were taken by changing the position of the lamp in small steps across the field-of-view. The distance between the lamp and the fish-eye lens was set at 3.1 m so that the angle subtended by the emitting area of the lamp at the fish-eye lens was much smaller than the field-of-view expected from the distance between the ASG and the lamp and a CCD pixel size. The results are shown in Table 2. The spatial resolution near the horizon, 0.37°, is close to the originally designed value of 0.36° near the zenith, the spatial resolution became 0.98°.

<table>
<thead>
<tr>
<th>Zenith angle</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>90°</td>
<td>0.37°</td>
</tr>
<tr>
<td>0°</td>
<td>0.98°</td>
</tr>
<tr>
<td>−90°</td>
<td>0.37°</td>
</tr>
</tbody>
</table>

In general, spatial resolution at the outer field-of-view of an optical component tends to be worse than that at the inner field-of-view because of aberrations. However, our experimental results show the opposite findings. The causes of the variation in spatial resolution have not been clarified, but mechanical distortion of the thin slit plate (t = 20 μm) is suspected to be partially responsible. Since the slit plate is held in place by a ring with a diameter of ~20 mm, the outer part of the slit plate can be fixed in the focal plane. If the plate is not sufficiently rigid, however, the central part might be slightly out of the focal plane. A similar effect can occur, if the focal plane is curved.

3.2. Spectral range and resolution

The spectral range and resolution were measured using images of an integrating sphere illuminated by three spectral calibration lamps (Hg, Kr or Ne), and a He-Ne laser. The integrating sphere of the National Institute of Polar Research, Japan serves as an ideal isotropic light source, because it has a large spherical cavity, 2 m in diameter, which is sufficient for the calibration of almost any kind of all-sky optics. Figure 2 shows a raw image of the integrating sphere illuminated by a Hg lamp. In addition to a dominant spectral line at 546.1 nm, two adjacent spectral lines of Hg at 577.0 and 579.1 nm and a weak spectral line at 435.8 nm can be clearly identified. The slight curvature of the
Fig. 2. A raw image of the integrating sphere illuminated by a Hg lamp without any correction. The intensity is scaled using a pseudo-false color code. North and violet are to the top and right of the image, respectively.

spectral lines is caused by the difference in incident angles to the grism between 0° for the beam from the zenith and about 5.6° for the beam from the horizon.

The spectra of the spectral calibration lamps and a He-Ne laser at zenith angles of 0°, +45° and +90° are shown in Fig. 3. The measured spectral range was determined from the positions of the isolated dominant spectral lines whose wavelengths are known. The range always covers 420-730 nm. The full range shifts slightly towards longer wavelengths as the zenith angle increases. Since the glasses used in the ASG optics are not transparent for light with wavelengths shorter than 370 nm, the observed spectrum is free from contamination by higher order spectra.

From these spectral line profiles, the FWHMs of the major spectral lines Hg 546.1 nm, Kr 557.0 and 587.1 nm and He-Ne 632.8 nm were measured to be about 1.7 nm. In addition, the adjacent spectral lines of Hg at 577.0 nm and 579.1 nm were completely resolved. Therefore, the spectral resolution was conservatively estimated to be 2 nm.

To evaluate the limit of the spectral resolution more quantitatively, we performed simulations in which two identical spectral line profiles, which are equivalent to those actually measured by the ASG, are superimposed by changing the separation between their positions by multiples of $\Delta X$ which is the resolution calculated simply from the wave-
length dispersion, the focal length of the imaging optics, and the CCD pixel size, as shown in Fig. 4. In most of the simulation results, the spectral lines were not resolved when they were separated by two \( \Delta X \), e.g., 1.2 nm, however, a distinct minimum between the two peaks was clearly seen when the separation was three \( \Delta X \), e.g., 1.8 nm. Consequently, the spectral resolution is estimated to be 1.5 nm for wavelengths between 540 and 650 nm.
Fig. 4. Simulations in which an actually measured spectral emission line is superimposed on itself after being shifted by one, two or three $\Delta X$s to indicate how the ASG resolves two adjacent spectral lines. The spectral lines of a) Hg 546.1 nm, b) Kr 587.1 nm, c) He-Ne 632.8 nm, and d) Ne 703.2 nm shown in Fig. 3 were selected for these simulations. One $\Delta X$ corresponds to a wavelength difference of about 0.6 nm. For each spectral line, nine combinations of zenith angles (0°, ±45°, and ±90°) and one, two, and three $\Delta X$s are plotted.

The spectral resolution decreases to about 2.0 nm near the shorter and longer wavelength limits as a result of aberration of the overall optics.

A comparison of the spectra at different zenith angles indicates that the spectral resolution has no clear dependence on the zenith angle. Note that the spectral resolution of the ASG can be improved by a factor of two, if the width of the slit is halved and the CCD is replaced by one with 1000 by 1000 12 $\mu$m pixels.
3.3. Sensitivity

The sensitivity of the ASG was calibrated using the same integrating sphere. For this purpose, a halogen tungsten lamp built into the integrating sphere was used to illuminate the inside of the sphere. The intensity of the integrating sphere is calibrated after every one-hundred hours of operation by a secondary calibration standard spectrophotometer which was calibrated by a National Institute of Standards and Technology (NIST) traceable standard lamp. Images were taken at two intensity settings ($0.5 \times 10^{-4}$ and $1.0 \times 10^{-4}$ W/m$^2$/sr/nm at 630 nm) to confirm the linearity of the output. The sensitivity of the ASG at each wavelength was derived from these image data and the absolute spectral intensity of the integrating sphere.

Figure 5 shows the spectral sensitivities for the entire spectral range at zenith angles of 0°, ±45°, and ±90°. The sensitivity at the zenith was at a maximum value of 0.06 cts/s/R at a wavelength of 560 nm, decreasing in value towards both shorter and longer wavelengths. The sensitivities at the zenith angles of ±45° and ±90° were about two thirds and one third of the value at the zenith, respectively. Effects such as the efficiency of the grism, vignetting of the optics, and the spectral response of the CCD are thought to produce the wavelength and zenith angle dependences of the spectral sensitivity.

4. Example of auroral spectrum

The ASG was installed at Longyearbyen, Spitzbergen in March 2000, and observations were started in the 2000/2001 winter season. Images of auroral spectra along the geomagnetic meridian were recorded every three minutes with an exposure time of 15 s. Figure 6 shows an example of an auroral spectral image obtained in the morning on December 6, 2000. The sky was clear at the time of observation. The emission intensities have been calibrated, and the spectral line distortion due to difference in incident angles to the grism has been corrected. Note that the observed wavelength region, 450–765 nm, has been slightly shifted towards a longer wavelength, compared with the range obtained when the ASG was calibrated in Japan.

The field-of-view of the ASG up to about 20° above the northern horizon is blocked
A new meridian imaging spectrograph

Date: 2000-12-06  Location: LYR
09:06:04(UT) / Exp=15(s)

Fig. 6. An auroral spectrum image obtained by the ASG at Longyearbyen on December 6, 2000.

by a nearby building. The southern horizon was illuminated by twilight. Dominant auroral emission lines, such as OI 557.7, 630.0, and 636.4 nm and OII 732/733 nm, are apparent in the northern sky. A strong NaD line emission (589 nm) due to city light contamination is also seen in the low northern sky.

From the data shown in Fig. 6 the intensity distributions of the OI 630.0 nm and OII 732/733 nm emissions along the meridian were plotted in Fig. 7. In the northern sky the profiles of both the OI 630.0 nm and OII 732/733 nm emissions have a similar pattern with three distinct peaks, but the peaks of the OII 732/733 nm emission are located at a slightly higher elevation compared with the peaks of the OI 630.0 nm emission. This indicates that three discrete auroras, in which the altitude regions of ion emissions were higher than those of neutrals, overlapped along the line-of-sight direction of the ASG. To obtain the actual altitude profiles of the emissions, the horizontal distance between the observatory and the auroral arcs must be known, or some assumptions on the altitude of characteristic
auroral emission, such as the OI 557.7 nm emission, must be made. The results of a detailed analysis of these emission profiles will be presented elsewhere.

5. Summary

We have developed an ASG that can measure auroral spectral intensities along the full meridian by producing an image in which the ordinate and abscissa represent the zenith angle of the meridional field-of-view and the spectral wavelength, respectively. The ASG utilizes a grism as its dispersive element and has spectral coverage of 420–730 nm and resolution of 1.5–2.0 nm. The sensitivity at the zenith is comparable to that of the ASI, though the sensitivity at the horizon decreases to about one third of the value at the zenith. The ASG was installed at Longyearbyen in March 2000, and observations were started in the 2000/2001 winter season. An example of a fully corrected and calibrated auroral spectral image is presented.

Acknowledgments

The authors would like to thank Drs. M. Okada and K. Sato for their help in installing the ASG at Longyearbyen. The ASG was calibrated at the calibration facility of the National Institute of Polar Research, Japan.

The editor thanks Drs. R. Fujii and B. Gustavsson for their help in evaluating this paper.

References


(Received May 24, 2002; Revised manuscript accepted July 5, 2002)
Polar mesosphere winter echoes during solar proton events

S. Kirkwood\(^1\), V. Barabash\(^1\), E. Belova\(^1\), H. Nilsson\(^1\), T.N. Rao\(^1\), K. Stebel\(^1\), A. Oseplian\(^2\) and Phillip B. Chilson\(^3\)

\(^1\)Swedish Institute of Space Physics, Box 812, 981 28 Kiruna, Sweden
\(^2\)Polar Geophysical Institute, Halturina 15, Murmansk, Russia
\(^3\)Cooperative Institute for Research in Environmental Science (CIRES) University of Colorado, Boulder, CO, U.S.A.

Abstract: Thin layers of enhanced radar echoes in the winter mesosphere have been observed by the ESRAD 52 MHz MST radar (67°53'N, 21°06'E) during several recent solar proton events. These polar mesosphere winter echoes (PMWE) can occur at any time of day or night above 70 km altitude, whereas below this height they are seen only during daytime. An energy deposition/ion-chemical model is used to calculate electron and ion densities from the observed proton fluxes. It is found that PMWE occurrence correlates well with low values of $\lambda$ (the ratio of negative ion density to electron density). There is a sharp cut-off in PMWE occurrence at $\lambda \sim 10^5$, which is independent of electron density. No direct dependence of PMWE occurrence on electron density can be found within the range represented by the solar proton events, with PMWE being observed at all levels of electron density corresponding to values of $\lambda < 10^5$. Together with results concerning the thickness, echo aspect-sensitivity and echo spectral-width of the PMWE, this observation leads to the conclusion that the layers cannot be explained by turbulence alone. A role for charged aerosols in creating PMWE is proposed.

1. Introduction

VHF radar echoes from the high-latitude mesosphere are well known to be much weaker in winter than in summer. The first studies using the powerful Poker Flat radar in Alaska (location 65.12°N, 147.43°W, with transmitter power 2 MW, and antenna area 40000 m\(^2\)) were reported by Ecklund and Balsley, 1981 and Balsley et al., 1983. Wintertime echoes were seen between 50 and 80 km and reached levels 30 dB lower than summertime echoes, which were concentrated at heights 80–90 km. The strong summertime layers between 80–90 km have since been termed ‘Polar Mesosphere Summer Echoes’, or PMSE, and have been detected and studied by several radars around the world. They are today thought to be caused by the effects of layers of small charged aerosols on the radar refractive index (see e.g. Cho and Röttger, 1997 for a review). Regarding the winter echoes, on the other hand, Balsley et al. (1983) noted that they appeared to be correlated with high-energy particle precipitation enhancing the ionization in the mesosphere, and that vertical profiles of echo power were highly structured on scales of 5–15 km. The echoes were interpreted as due to turbulence caused by breaking gravity waves. Note, however, that the limited height resolution of the Poker Flat radar (2.2 km) precluded
study of the structure on smaller scales, and no systematic study was made of the echo
dependence on ionisation levels in the surrounding atmosphere. Wintertime echoes at
high latitudes have been further studied by Czechowsky et al. (1989), using the SOUSY
mobile radar during a campaign from Andøya, northern Norway in 1983/84 (location
69.17°N, 16.01°E). Results similar to those from Poker Flat were found and it was
confirmed that the echoes were detectable only during periods when electron densities were
enhanced by energetic particle precipitation. It was noted that enhanced electron densities
during the hours of darkness seemed not to lead to observable echoing layers below about
70 km altitude. Also in this case, the echoing structures were said to have vertical extent
of 2–10 km. Although the SOUSY radar operated with high resolution (300 m or better)
and examples of layers as narrow as the resolution can be found in the figures in
Czechowsky et al. (1989), no comment was made on the thinness of the layers and
turbulence due to wave saturation and/or breaking was again proposed as the cause.

In this paper we examine the characteristics of mesospheric winter echoes using the
ESRAD MST radar located at Esrange, near Kiruna in Sweden (67.88°N, 21.10°E). A
detailed description of the radar can be found in Chilson et al., 1999. This radar uses 72
kW transmitter power and a 1600 m² antenna array, together giving about 30 dB (1000
times) less sensitivity than the Poker Flat or SOUSY mobile radars. ESRAD is normally
used to study Polar Mesosphere Summer Echoes (e.g. Kirkwood et al., 1998) and
tropospheric winds and waves (e.g. Récou et al., 1999). However it operates continuous-
ly, monitoring also wintertime radar returns from the stratosphere and mesosphere.
During the winter 2000/2001 several solar proton events occurred and during these events
mesospheric layers were detected by the ESRAD radar. We refer to these layers as Polar
Mesosphere Winter Echoes (PMWE). The good height resolution and detailed analysis
available with the ESRAD spaced-antenna configuration (see e.g. Holdsworth and Reid,
1995) allow new characteristics of the layers to be determined. Because of the long
duration of the solar proton events (several days) and the availability of satellite data on
the precipitating protons, we are able to model realistically the electron density variations
in the mesosphere during both day and night conditions. This allows us to separate
electron density dependence from solar illumination dependence. Together, these analyses
lead to new interpretations of the cause of PMWE.

2. Characteristics of PMWE

Radar returns from the height region 5–100 km are monitored routinely by the
ESRAD radar with 600 m height resolution and time resolution varying from one profile
each 7 min to 1 profile every third minute. For the second half of the winter 2000/2001,
the region 60–80 km was also monitored with 300 m resolution. (1 profile each 3 min).
Examples showing the most intense layers detected in November and April are shown in
the panels second from the top in Figs. 1 and 2a, respectively. The colour scales for the
plots have been chosen so that the highest tops in the background noise are just visible
(blue). The noise can be seen as a generally randomly placed pixels, although sometimes
forming vertical lines due to interference reaching the receivers. The PMWE appear as
layers of enhanced radar echo power above this noise level—generally green, yellow or red
in the plots. The layer thicknesses (FWHM) are, at times, as little as, or less than, the
Fig. 1. Observations and model results for 9–10 November 2000. The x-axis shows the date and time (UT) as a decimal fraction of the day (e.g. 9.5 for 12 UT on 9 November). Top panel: Solar proton fluxes from GOES 10 satellite. 2nd panel: Echo power recorded by the ESRAD radar (colour scale dB). 3rd–5th panels: Modelled densities of electrons, positive cluster ions and negative ions. Colour scale shows log density cm$^{-3}$. 6th panel: Cosmic noise absorption at 30 MHz calculated from the model results (blue) and measured (magenta) by a riometer in Abisko, 80 km WNW from ESRAD.
Fig. 2a. Observations and model results for 2–7 April 2001. The x-axis shows the date and time (UT) as a decimal fraction of the day (e.g. ‘3.5’ for 12 UT on 3 April). Top panel: Solar proton fluxes from GOES satellite. 2nd panel: Echo power recorded by the ESRAD radar (colour scale dB). 3rd–5th panels: Modelled densities of electrons, positive cluster ions and negative ions. Colour scale shows log density cm$^{-3}$. 6th panel: Cosmic noise absorption at 30 MHz calculated from the model results (blue) and measured (magenta) by a riometer in Abisko, 80 km WNW from ESRAD.
resolution of the observations, particularly for the layers seen in April (Fig. 2a). The PMWE are easy to identify in Fig. 1 where they are rather broad in height and generally last several hours at a time. They are less easy to identify in Fig. 2a where they are generally much narrower in height and more sporadic in time. However, careful inspection shows identifiable layers each day, in the midday sector—below 60 km on 3 April, close to 65 km on 4 April, 70 km on 5 April, just above 70 km on 6 April and at about 77 km on 7 April. This last layer, on 7 April, is almost impossible to see in Fig. 2a because of its extreme thinness. It is shown in enlargement in Fig. 2b. Note that the resolution on this day was 300 m and the PMWE for most of the time occupies only one range gate.

Similar layers were seen during all of the solar proton events which occurred between October 2000 and early May 2001 (in October, November, January, March, April and early May). However no layers other than the usual PMSE lying between 75–90 km were seen during the solar proton event which occurred in July 2000, despite proton fluxes 60% higher than the event in Fig. 1. No layers were seen either during solar proton events in August and September 2000. This justifies the term ‘winter’ in the name PMWE. Layers were not seen at times other than during solar proton events. Since such events affect the mesosphere only at high latitudes, this leads to the term ‘polar’ in the name PMWE.

To examine the dependence of PMWE on the density of free electrons and ions in the mesosphere we have used the energy-deposition/ion-chemical model described in Kirkwood and Osepeian, 1995. Tests of the model, including its application to solar proton events, can be found in the same publication. As input to the model we have used a Maxwellian flux-energy spectrum of precipitating protons fitted to the integral fluxes at >10 Mev and >100 MeV measured by the GOES 10 satellite (http://www.sec.noaa.gov/ftpmenu/lists/particle.html). The proton fluxes are shown in the top panel of Figs. 1 and 2a. Model results for the solar proton event of 9–11 November 2000 are shown in the 3rd–6th panels of Fig. 1, and for 2–7 April in Fig. 2a. The lowest panels of Figs. 1 and 2a also show the observed absorption of cosmic radio noise at 30 MHz, from the nearby riometer station of Abisko, Sweden. Comparison between this measured absorption and that calculated on the basis of the model electron density profiles gives a measure of how well our model probably represents the real situation in the atmosphere (at least concerning electron density). In general the model results predict lower absorption than observed. Similar daily variations with constant discrepancies in amount (e.g., up to a factor 2) are likely due to minor inadequacies in the model, i.e., electron densities being underestimated by up to a factor 2 at all heights, or the extension of ionisation to lower heights being underestimated in the model. Large discrepancies and/or rapid time variations in the observed absorption are likely due to precipitation of energetic electrons, in addition to the
protons.

The main features of the model results are:
- peak electron densities reaching ca $4 \times 10^4$ cm$^{-1}$ at about 70 km altitude in the November event, a factor about 10 less in the April event, corresponding to roughly 100 times lower proton fluxes
- above 80 km electron densities vary smoothly with time, following closely the intensity of the proton fluxes
- below 80 km electron densities have a strong daily variation. The explanation for this is to be found in the behaviour of the negative ions which are most persistent at night. This in turn is explained by the fact that electrons are readily removed from negative ions by sunlight and by reactions involving atomic oxygen, which is present in much greater amounts during daytime
- high densities of positive cluster ions below 80 km, reaching $> 10^5$ cm$^{-3}$ (in November), with a strong daily variation. This daily variation is due to the reduction in recombination at night as the number of free electrons is diminished.

A comparison between the 2nd and 3rd panels of Fig. 1 or between the 2nd and 3rd panels of Fig. 2a suggests a correlation of PMWE with high densities of free electrons, with PMWE being absent below 75 km at night, i.e. when and where the electrons attach instead to negative ions. However, a closer comparison between Figs. 1 and 2a shows that the daytime layer on 4 April (centre of Fig. 2a) is present in a background of slightly lower electron density than that which prevails at the same height during the night of 9/10 November. This suggests that the primary parameter controlling the daily variation of PMWE during solar-proton events is not the absolute electron density but rather some effect related to the absence of negative ions (or the presence of atomic oxygen).

To gain more information concerning the dependence of PMWE on electron density or ion composition, all of the radar profiles collected between 1 September 2000 and 30 April 2001 have been searched for statistically significant features as follows: first each height profile (20–100 km with 600 m resolution) is examined and any point lying more than 3 standard-deviations above the mean is identified as a possible PMWE. Next, time continuity is tested by checking whether similarly enhanced signal is present in the following two height profiles, at the same or adjacent altitudes. Each height and time bin satisfying both criteria is recorded as containing PMWE. The detected PMWE and their temporal correspondence to solar proton events is tabulated in Table 1. Only 10% of the detected PMWE are found at times when there is no detectable increase of proton flux. They might be due to enhanced proton fluxes which we cannot detect, or to other sources of ionisation such as high-energy electron precipitation. All detected PMWE are included in Fig. 3 as it is always possible to estimate corresponding values of $\lambda$ according to the solar elevation (see below). PMWE occurring when the proton fluxes are below the detection threshold are not included in Fig. 4 since it is not possible to calculate corresponding electron densities in these cases.

Figure 3 shows the height and solar-elevation for all of the PMWE seen during the whole winter period against a background of contours of $\lambda$, the ratio of negative ion density to electron density. The dependence of on solar elevation shown in Fig. 3 has been calculated using fixed proton fluxes (values at 12 UT on 9 November). Although the
Table 1. Dates and lengths of time when proton fluxes (>10 MeV) were above the detection threshold for the GOES-10 satellite and characteristics of PMWE during these periods. The bottom row summarises PMWE detected when proton fluxes were below the detection threshold.

<table>
<thead>
<tr>
<th>Dates</th>
<th>Total time (hours) with proton flux &gt;0.16 cm(^{-2})s(^{-1}) sr(^{-1})</th>
<th>Maximum proton flux (cm(^{-2})s(^{-1}) sr(^{-1}))</th>
<th>Total time detected PMWE (hours)</th>
<th>Maximum PMWE power (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000 Sep 8</td>
<td>8</td>
<td>0.23</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>2000 Sep 12-21</td>
<td>160</td>
<td>178.75</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>2000 Oct 1</td>
<td>2</td>
<td>0.16</td>
<td>&lt;1</td>
<td></td>
</tr>
<tr>
<td>2000 Oct 7-14</td>
<td>38</td>
<td>0.32</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2000 Oct 16-23</td>
<td>94</td>
<td>4.33</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2000 Oct 25-31</td>
<td>100</td>
<td>6.28</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>2000 Nov 1-6</td>
<td>39</td>
<td>2.15</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>2000 Nov 8-23</td>
<td>355</td>
<td>9741.66</td>
<td>25</td>
<td>22</td>
</tr>
<tr>
<td>2000 Nov 24-30</td>
<td>164</td>
<td>665.66</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>2000 Dec 1-8</td>
<td>96</td>
<td>1.04</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>2000 Dec 21</td>
<td>2</td>
<td>0.16</td>
<td>&lt;1</td>
<td>9</td>
</tr>
<tr>
<td>2000 Dec 28-31</td>
<td>30</td>
<td>0.30</td>
<td>&lt;1</td>
<td>10</td>
</tr>
<tr>
<td>2001 Jan 1-3</td>
<td>5</td>
<td>0.17</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2001 Jan 5-7</td>
<td>27</td>
<td>0.41</td>
<td>&lt;1</td>
<td>9</td>
</tr>
<tr>
<td>2001 Jan 15-16</td>
<td>9</td>
<td>0.21</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>2001 Jan 21-27</td>
<td>104</td>
<td>1.47</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>2001 Jan 28-31</td>
<td>73</td>
<td>29.07</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>2001 Feb 01-02</td>
<td>7</td>
<td>0.21</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2001 Feb 11-12</td>
<td>14</td>
<td>0.28</td>
<td>&lt;1</td>
<td>9</td>
</tr>
<tr>
<td>2001 Feb 18</td>
<td>2</td>
<td>0.19</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2001 Feb 25-28</td>
<td>29</td>
<td>0.67</td>
<td>&lt;1</td>
<td>9</td>
</tr>
<tr>
<td>2001 Mar 10-11</td>
<td>4</td>
<td>0.16</td>
<td>&lt;1</td>
<td>8</td>
</tr>
<tr>
<td>2001 Mar 15-17</td>
<td>5</td>
<td>0.18</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2001 Mar 23</td>
<td>2</td>
<td>0.17</td>
<td>&lt;1</td>
<td>13</td>
</tr>
<tr>
<td>2001 Mar 26-28</td>
<td>32</td>
<td>1.94</td>
<td>&lt;1</td>
<td>8</td>
</tr>
<tr>
<td>2001 Mar 29-31</td>
<td>61</td>
<td>21.70</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>2001 Apr 1-9</td>
<td>210</td>
<td>126.16</td>
<td>7</td>
<td>19</td>
</tr>
<tr>
<td>2001 Apr 10-14</td>
<td>114</td>
<td>286.16</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>2001 Apr 15-17</td>
<td>70</td>
<td>341.25</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>2001 Apr 18-27</td>
<td>177</td>
<td>188.83</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>2001 Apr 28-29</td>
<td>23</td>
<td>15.35</td>
<td>&lt;1</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>2087</td>
<td></td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>2000 1 Sep-2001 31 April flux &lt;0.16 cm(^{-2})s(^{-1}) sr(^{-1})</td>
<td>3721</td>
<td>&lt;0.16</td>
<td>8</td>
<td>14</td>
</tr>
</tbody>
</table>

Absolute values of electron density and negative ion density may vary by orders of magnitude from one event to another, the ratio \(\lambda\) is approximately constant for any particular solar-zenith elevation and height. For example, the 2-3-orders of magnitude difference in absolute electron and ion densities between 9 November in Fig. 1 and 7 April in Fig. 2a corresponds to, at most, a factor 2 change in \(\lambda\). It is very clear that PMWE are seen only where \(\lambda\) is low — i.e. above ca 75 km altitude at any time of day, below that height only during daytime when there is a large proportion of free electrons and a small proportion of negative ions. The only layers which do not seem to fit this pattern very well are those below 55 km altitude. However, the uncertainties in the ion-chemistry
Fig. 3. Detected PMWE as a function of height and solar elevation (deg) plotted against a background (colour) showing the corresponding log ($\lambda$), $\lambda$ being the ratio of negative ion density to electron density. The upper panel is for morning (solar elevation increasing with time), the lower panel for evening.

model is greatest at these low heights so the apparent discrepancy may simply be an artifact of the model. Altogether, the attachment of electrons to form negative ions at night (high $\lambda$) is well correlated with an absence of PMWE.

To further try to separate the dependence on electron density from the dependence on ion composition, a model value of electron density and ion composition is calculated for each detected PMWE, using the corresponding height, solar elevation and solar proton flux. Figure 4 plots detected PMWE as a function of both model electron density and $\lambda$. Individual layers can be seen in several cases as they trace a steady reduction in electron density as $\lambda$ increases, i.e. as the solar elevation decreases. The number of hours of
Fig. 4. Detected PMWE as a function of model values of \( \log(\lambda) \) (\( \lambda \) being the ratio of negative ion density to electron density) and \( \log \) (electron density) calculated using the ion-chemistry model for the appropriate height, solar elevation and solar proton flux (+ and ⬤). The background (colour) shows the number of hours during which radar observations were made for the corresponding electron density/\( \lambda \) conditions. The dark blue areas correspond to ranges of electron density/\( \lambda \) for which we have no observations.

Observations for the different \( \lambda \)/electron density conditions are shown by the coloured background. We are able to compute model values of electron density and \( \lambda \) only when proton fluxes are above the detection threshold for the GOES 10 satellite (ca 0.16 protons cm\(^{-1}\)s\(^{-1}\)sr\(^{-1}\)). This means that only such conditions are represented in Fig. 4, i.e. there is no information on the presence or absence of PMWE for the lowest values of both electron density and \( \lambda \) simultaneously, since such conditions do not occur so long as proton fluxes exceed the GOES 10 detection threshold.

Two symbols are used to indicate PMWE observations: `+` show all detected layers between September 2000 and April 2001, ⬤ show those events where the integral 30 MHz noise absorption from the model is in such close agreement with observed values that we can be sure there is no significant extra ionisation due to precipitating high-energy electrons. In the other cases, energetic electrons from the magnetosphere must be contributing to the electron density profile. It is known from statistical studies that most of the effect from magnetospheric electrons is at heights above 90 km so it is unlikely that they have a significant effect at the heights of the observed layers (50-90 km). However it cannot be categorically ruled out. Despite this uncertainty, it is still rather clear that PMWE have a sharp cut-off as \( \lambda \) increases above 100 (10\(^2\)) even though their environment in terms of electron density can vary by several orders of magnitude at this cut-off. At low values of \( \lambda \), PMWE are seen at electron densities as low as 10\(^2\) cm\(^{-3}\). At high values of \( \lambda \), PMWE are not seen even though electron densities as high as 3×10\(^3\) cm\(^{-3}\) occurred during many hours. If there is a threshold electron density required for PMWE, then it
is below the densities represented by our model during solar proton events.

3. Interpretation of PMWE

Radar signals are scattered by the atmosphere when there are fluctuations in the radar refractive index at appropriate scale sizes (for the ESRAD radar, around 3 m). The radar refractive index, n, depends on neutral density, temperature and humidity, and on electron density (see e.g. Balsley and Gage, 1980).

\[ n = 1 + 77.6 \times 10^{-6}(p/T) + 0.373(e/T^2) - 40.3(Ne/f^2), \]

where \( p \) is atmospheric pressure in mb, \( e \) is the partial vapour pressure of water, in mb, \( T \) is temperature in K, \( Ne \) is electron density m\(^{-3}\), \( f \) is radar frequency.

At 50 km altitude, during our solar proton events, the electron density term is an order of magnitude greater than the neutral density term \((p/T)\) which in turn is about 3 orders of magnitude greater than the water vapour term. At higher altitudes, the importance of the electron density term grows relative to the other terms, reaching about 5 orders of magnitude greater than the neutral density term by 80 km altitude. So to explain our PMWE we must find a mechanism which causes fluctuations in electron density along the direction of the radar beam including scale-sizes of 3 m.

Before discussing possible causes further, it is worth considering what more information about the scattering mechanism we can derive from the properties of the radar echoes. Using the spatial correlation method and the spectral widths of the radar echoes it is possible to estimate the random spread of velocities within the scattering volume (turbulence) and the anisotropy (ratio of vertical thickness to horizontal length) of the scattering structures (Holdsworth and Reid, 1995). Since rather strong signal levels are needed for such analysis, we have been able to estimate these parameters only for the strongest layers seen during our observation period. These are given in Table 2. We will return to these values in the discussion below.

The first explanation to be considered is that the layers are due to turbulence due to gravity-wave breaking or Kelvin-Helmholtz instability (due to wind shear), as proposed for high-latitudes by Balsley et al., 1983 and Czechowsky et al., 1989. Turbulence due to enhanced gravity wave breaking at temperature inversions has also been proposed to cause layers of enhanced radar-echoes. Studies at mid- and low-latitudes (Thomas et al., 1996; Ratnam et al., 2002) have found a close correlation between enhanced radar echoes and strong temperature inversions seen by lidar in the 70–80 km height interval. Lübken (1997), through a succession of sounding rocket experiments, has indeed been able to demonstrate that narrow layers of strong turbulence are a common feature of the winter mesosphere. Further, in the presence of an electron density gradient, neutral turbulence would be expected to cause turbulence also in the electron plasma. However, the layers found by Lübken (1997) were generally a few km thick, much more than the <300 m we observe for the PMWE in Figs. 2a and 2b. The lower latitude radar-echo layers observed by Thomas et al. (1996) to be correlated with temperature inversions have also been rather broad (more than 2 km thick). Typical turbulent velocities implied by the rocket experiments reported by Lübken (1997) were ca 1–2 m s\(^{-1}\). This is close to the values we find for the relatively thiek
Table 2. Radar echo parameters for the three strongest layers observed. The rms turbulent velocity is derived from the doppler spread of the echo spectrum, the aspect sensitivity (and irregularity length/height ratio) are from spatial correlation analysis (see Holdsworth and Reid, 1995).

<table>
<thead>
<tr>
<th>Date</th>
<th>9 November 2000</th>
<th>10 November 2000</th>
<th>7 April 2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>07–13 UT</td>
<td>11–13 UT</td>
<td>07–10 UT</td>
</tr>
<tr>
<td>Layer center height</td>
<td>63–65 km</td>
<td>60–62 km</td>
<td>77–78 km</td>
</tr>
<tr>
<td>Layer thickness</td>
<td>&lt;0.6–2.4 km</td>
<td>&lt;0.6–1.2 km</td>
<td>&lt;0.3–0.6 km</td>
</tr>
<tr>
<td>Height resolution</td>
<td>0.6 km</td>
<td>0.6 km</td>
<td>0.3 km</td>
</tr>
<tr>
<td>Rms turbulent velocity</td>
<td>1.7 m/s</td>
<td>1.7 m/s</td>
<td>0.2 m/s</td>
</tr>
<tr>
<td>Aspect sensitivity</td>
<td>1.3°</td>
<td>1.9°</td>
<td>1.3°</td>
</tr>
<tr>
<td>Irregularity length/height ratio</td>
<td>12</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>Model electron density</td>
<td>3000–24000 cm⁻³</td>
<td>1000–3000 cm⁻³</td>
<td>300–400 cm⁻³</td>
</tr>
<tr>
<td>Model electron density gradient</td>
<td>10⁻⁴–5×10⁻² cm⁻⁴</td>
<td>3×10⁻¹–10⁻² cm⁻⁴</td>
<td>2×10⁻⁴–4×10⁻⁴ cm⁻⁴</td>
</tr>
</tbody>
</table>

November PMWE (Fig. 1 and Table 2) but much more than we find for the very thin PMWE on 7 April 2001 (Fig. 2b and Table 2). The anisotropy we find for all of the PMWE is rather high. This indicates that, if PMWE are due to turbulence, that turbulence must be highly anisotropic. It might also indicate that the echoes are due to Fresnel reflection from multiple sharp layers rather than from turbulence. A further problem with the turbulent-layer hypothesis is the behaviour with respect to electron density. A turbulent layer should simply redistribute the background electron density with the irregularity strength directly related to electron density gradient. The electron density gradients at the relevant heights during solar proton events are, with normal ion-chemistry, directly related to the electron density itself. For constant turbulence intensity, we would expect to see PMWE getting steadily weaker as the electron density (and its gradient) decrease. We would not expect a layer with weak turbulence (7 April, rms turbulent velocity 0.2 ms⁻¹) to be seen at an electron density of only 400 cm⁻³ (gradient 4×10⁻⁴ cm⁻⁴) while a layer with strong turbulence (9 November, rms turbulent velocity 1.7 ms⁻¹) becomes abruptly too weak to detect as soon as the electron density falls to 3000 cm⁻³ (gradient 10⁻² cm⁻⁴).

However, the major problem with the turbulent-layer hypothesis arises when we consider Lübken et al.'s (1993) results concerning their direct observations of the inner scale of turbulence in wintertime turbulent layers. These show that the inner scale length of the turbulence (in the neutral air density) is of the order 4–12 m, so that neutral-density fluctuations start to be attenuated at the shorter scale-sizes (3 m) needed to explain the PMWE. The relationship between the scale-size cut-off in neutral density and that in electron density (to which the radar is sensitive) can be conveniently categorised by the Schmidt number, which is the ratio of kinematic viscosity to electron diffusivity (Driscoll and Kennedy, 1985). When only positive ions are present, electrons are constrained by charge balance to diffuse with the ions (ambipolar diffusion) and ions and electrons have
the same diffusivity. With ion masses close to those of the main neutral molecules in the atmosphere, fluctuations in ion and electron density will then have close to the same scale-size distribution as the neutral air density (the case of Schmidt number 1, Kelley et al., 1987). However, when negative ions are present, the diffusivity of the electrons is effectively increased (Hill, 1978) and the shortest-scale fluctuations in neutral and ion densities should no longer be present in the electron density. We observe PMWE at values of $\lambda$ up to about 100. According to Hill (1978) the electron diffusivity should then be increased by a factor $(1 + \lambda)$, i.e. close to 100. This will reduce the Schmidt number to about 0.01. According to Driscoll and Kennedy (1985) this should put the 'inner scale size' for electron-density fluctuations at about 10 times more than for neutral density fluctuations, i.e. 40-120 m. This is far above the 3 m to which our radar is sensitive. A similar mismatch between the inner scale of turbulent fluctuations and the scale-sizes needed to give radar echoes is found for polar mesosphere summer echoes. The only reasonable way to increase the electron Schmidt number and extend turbulence-produced electron-density fluctuations to short enough scale sizes is thought to be through the presence of heavy, charged aerosols (as reviewed by Cho and Röttger, 1997). Our observations of PMWE at times when no negative ions are present might then similarly be explained by the presence of charged aerosols which increase the Schmidt number to about 100. The disappearance of PMWE at $\lambda \sim 100$ could then be explained by the increased electron diffusivity causing a reduction of the Schmidt number to about 1, at which point the inner-scale size of the electron density fluctuations should increase to above the radar half-wavelength (3 m).

We must also consider other processes than neutral turbulence which might produce fluctuations in electron density with appropriate scale-sizes. For example, we should consider that the height profile of minor constituents of importance for the ion chemistry might be highly structured. There is indeed some evidence that the vertical profile of water vapour in the high-latitude winter mesosphere contains sharp maxima and minima (Khaplanov et al., 1996). To test the effect of a thin layer of enhanced water vapour we have made calculations using the ion-chemistry model described above but with double the normal water vapour concentration in a 1-km thick layer at 62-63 km altitude. The result for the conditions of 3 April 2001 is an electron density depletion (about 20%) within the enhanced water layer. This can be understood as an increase in the ratio of positive cluster ions to positive molecular ions as cluster-ion formation is favoured by increased water vapour. Since cluster ions have shorter lifetimes against recombination with electrons, the net result is a lower electron density. The gradients bounding the electron density depletion layer could, in principle, lead to enhanced radar echoes. However, the model results show that the relative depletion ($\Delta Ne/Ne$) is slightly dependent on the ionisation rate—the electron density depletion increases to 30% (day)–40% (night) if the ionisation rate is increased by a factor 10. The percentage depletion is, if anything, slightly higher at night (high $\lambda$). The net result is a strong dependence of absolute electron-density gradient on electron density. So our observation that PMWE requires low values of $\lambda$ rather than high values of electron density, does not support this explanation either.

A further possibility to consider is that PMWE, in a similar fashion to some types of PMSE, might be due to sharp gradients in electron density caused by layers of charged aerosols. Studies of the aerosol layers by a number of sounding-rocket experiments at the
summer mesopause seem to indicate the presence of both positively and negatively charged aerosols associated with ‘bite-outs’ in the electron density profile, presumably due to scavenging of electrons by the aerosols (e.g. Croskey et al., 2001; Havnes et al., 2001). The sharp gradients in electron density associated with such a ‘bite-out’ can be an effective source of highly aspect-sensitive radar echoes. If such aerosol layers were present in the winter mesosphere they could be expected to cause similar bite-outs at those heights and times of day when electrons are the only negative charge carriers (apart from the aerosols themselves).

Once large numbers of negative ions become available, they can be captured by aerosols in a similar way to electrons. Capture rates are expected to be proportional to the number flux of the charged particles $N \times C$, where $N$ is the number density of charged particles and $C$ is their mean thermal velocity (e.g. Natanson, 1960). The ratio of capture rates for negative ions and electrons, $R$, can then be expressed as:

$$R \propto \frac{\lambda C_e}{C_n} = \lambda \left(\frac{m_e}{m_n}\right)^{1/2},$$

where $m_e$ and $m_n$ are the masses of electron and negative ions, respectively.

Assuming that the main negative ions are $O^-\bullet$, $CO_3^-$ and $NO_3^-$, this gives

$$R \propto \lambda (0.3-0.4) \times 10^{-2},$$

Thus $\lambda \sim 3 \times 10^2$, about the value we find to correspond to the cut-off for PMWE occurrence, corresponds to the situation when electrons and negative ions have equal capture rates. For higher values of $\lambda$, aerosols should preferentially scavange negative ions leading to a negative-ion ‘bite-out’ rather than an electron ‘bite-out’. This would not cause enhanced radar echoes since the echoes require a gradient in electron density, not a gradient in ion density. Any gradient in electron density will diffuse away rapidly due to the very high electron diffusivity in the presence of negative ions (Hill, 1978).

4. Conclusions

Thin layers of enhanced radar-echoes between 50-80 km altitude in the winter mesosphere have been observed by the 52 MHz ESRAD radar during solar proton events. These PMWE (polar mesosphere winter layers) are seen at a wide range of background electron densities but only when the ratio of negative ions to free electrons is expected to be less than about 100. The characteristics of PMWE lead us to conclude that aerosols are most likely involved in creating electron density fluctuations at the $3 \text{ m}$ scale-sizes necessary to produce radar echoes, in the same way as in PMSE (polar mesosphere summer echoes).

Clearly it is not reasonable to propose that water-ice aerosols are involved in PMWE as temperatures in the winter mesosphere are much too high for water saturation. Some other type of aerosol must be proposed, such as the meteoric smoke or dust proposed by Huntsen et al., 1981. Unfortunately, very little is known about the charging properties of such aerosols, if indeed they really are present in the winter mesosphere. However, it is interesting to note in this context that an aerosol layer at the very unusual height of 37 km was reported by several high-latitude lidar stations in mid November 2000, (K.-H. Fricke, personal communication). The layer persisted over several weeks inside the polar vortex.
sinking to about 28 km height before it disappeared in February 2001. The height of the layer seen by the lidars is clearly rather lower than the layers we see with the radar. However, as in the case of PMSE, aerosols responsible for PMWE might well be too small to be detected by lidar but could in some cases grow large enough to be seen by such instruments, at the same time sedimenting to lower altitude.

The possibility that significant numbers of aerosols are present in the winter polar mesosphere, and that they are concentrated into layers, may have considerable significance for chemical processes in the region. Chlorine activation on polar stratospheric clouds (aerosols), which is the precursor to stratospheric ozone destruction in the polar vortex, is one example of such an effect on chemistry which might then also occur in the mesosphere. The possibility of recombination of O and H₂ on aerosol surfaces as a source of mesospheric water vapour, is another (Summers and Siskind, 1999).

Acknowledgments

The ESRAD MST radar is operated jointly by Swedish Institute of Space Physics and Swedish Space Corporation, Esrange. Riometer data from Abisko were provided by Sodankylä Geophysical Institute, Finland. Software for full-correlation analysis of the ESRAD data has been developed at the University of Nebraska. This research has been funded by the Swedish Science Research Council.

The editor thanks Drs. Y. Murayama and A. P. van Eyken for their help in evaluating this paper.

References


Polar mesosphere winter echoes


(Received December 18, 2001; Revised manuscript accepted June 7, 2002)
SC-triggered plasma waves observed by the Akebono satellite in the polar regions and the plasmasphere

Atsuki Shinbori¹, Takayuki Ono¹ and Hiroshi Oya²

¹Department of Geophysics, Graduate School of Science, Tohoku University, Sendai 980-8578
²Fukui University of Technology, Fukui 910-8505

Abstract: Plasma wave phenomena associated with sudden commencements (SCs) were analyzed based on observations conducted with the Akebono satellite, which has been collecting data for more than 13 years (since March 1989). Simultaneous plasma wave observation data for 257 SCs reveal that enhanced plasma waves are observed with an exact one-to-one correspondence with the SCs throughout the entire observation region, including the polar and plasmasphere regions. Electromagnetic whistler mode and ion cyclotron waves are enhanced in the low latitude plasmasphere, while electrostatic whistler mode and electromagnetic ion cyclotron waves are generated in the polar region. The onset times of the SC-triggered plasma waves exhibit a delay or lead time characteristic, compared with the onset times of SCs identified by the Kakioka Magnetic Observatory, with a time resolution of 1 s. By comparing the difference in SCs and enhanced electron plasma waves onset times, the propagation route of the SC disturbances can be identified in the plasmasphere.

1. Introduction

The generation and enhancement of plasma waves in the ULF-VLF range has been reported to be associated with sudden commencements (SCs). Tepley and Wentworth (1962) reported an intensification and an increase in the frequency of Pc1 ULF waves associated with SCs. Kokubun and Oguti (1968) pointed out that SC-triggered ULF events tend to appear in the morning-noon sector. Hirase (1981) demonstrated a clear local time dependence of SC-triggered Pc1 ULF waves. However, the possible effect of the ionosphere depending on the local time has not been evaluated using the ground-based observations. Morozumi (1965) reported that SC-triggered VLF chorus emissions also occur. Hayashi et al. (1968) reported an increase in the frequency of VLF chorus emissions about 30 s or more before the onset of SCs. They attributed the change in the VLF spectra to variations in temperature anisotropy caused by an abrupt increase in the geomagnetic field.

Satellite observations are necessary to clarify the distribution of SC-triggered plasma waves in the magnetosphere, plasmasphere and polar regions without the influence of ionosphere conditions. Mullayaev and Yachmenev (1990) reported the enhancement of whistler mode plasma waves observed based on observations obtained with the GEOS-2 satellite. Gail and Inan (1990) studied 14 VLF emission events observed by the DE1
satellite. However, time resolution limitations (32 s for the DE data) make it difficult to analyze SC-triggered VLF emissions in detail.

In this paper, ELF-VLF plasma wave phenomena associated with SCs are examined using data records collected by the Akebono satellite plasma wave instruments over a 13-year period. The purpose of the present study is to identify the generation of plasma waves associated with SC disturbances in the polar region and the plasmasphere. The study also attempts to use the high-time resolution data obtained from the Akebono satellite to clarify the relationship between the onset of SCs identified on the ground and the modification of energetic particles and the plasma waves in order to understand how the plasma waves and particles in the polar region and plasmasphere respond to SC disturbances.

2. Observation data

The Akebono satellite has been continuously conducting observations for more than 13 years. The satellite was launched on February 21, 1989, and placed in a semi-polar orbit with an inclination of 75° and an initial apogee and perigee of 10500 km and 274 km, respectively. In the present series of studies, the plasma wave data was obtained by the PWS (20 kHz–5.1 MHz) (Oya et al., 1990), VLF (3.16 Hz–17.8 kHz) and ELF (0 Hz–80 Hz) (Kimura et al., 1990) instruments onboard the Akebono satellite. The plasma wave data analysis was combined with an analysis of low energy particle data obtained by the LEP (Mukai et al., 1990) instrument, also installed on the Akebono satellite. The time resolution of the dynamic PWS spectra is 2 s, while the time resolution of the VLF, ELF, and LEP data is 8 s, as specified by the Akebono satellite data-base.

3. Identification of SCs

Between March 1989 and November 2001 (Table 1), 930 SC events were identified in the SYM-H data (Iyemori and Rao, 1996), which has a time resolution of 1 min. SC events were identified by a rapid increase in SYM-H values (more than 5 nT within 10 min). For each SC event, the precise onset time was identified by referring to the $H$-component of the geomagnetic variation measured at the Kakioka Magnetic Observatory; the time resolution of the $H$-component measurement is 1 s. As shown in Fig. 1, the onset time of the SCs can be identified with a time resolution of 1 s. In Fig. 1, the SYM-H data shows an increase of 23 nT per ten minutes from 1222 (UT) on September 22, 1999. The simultaneous $H$-component geomagnetic variation measured at the Kakioka Magnetic Observatory indicates that the increase in the geomagnetic field started at 1222:03 (UT). Thus, the onset times of the SCs can be identified with a time resolution of 1 s. Within the 13-year period of the Akebono satellite operation, from March 1989 to November 2001, two-hundred and fifty-seven PWS observation periods corresponded with the onset time of an SC event. In all the observation cases, a clear plasma wave signature (involving a change in intensity or frequency spectra) appeared in response to each SC. Table 2 shows the event list of these SC-triggered plasma wave phenomena, as observed by the PWS instruments onboard the Akebono satellite. As shown in Fig. 2, the SC-triggered plasma wave events do not have a clear dependence on the geomagnetic local time, altitude or
Table I. SC event list from March 1989 to June 2001, based on the SYM-H index provided by the WDC-C2 for Geomagnetism.

<table>
<thead>
<tr>
<th>year</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989</td>
<td>10</td>
<td>3</td>
<td>7</td>
<td>7</td>
<td>3</td>
<td>9</td>
<td>11</td>
<td>6</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>9</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>1991</td>
<td>3</td>
<td>3</td>
<td>11</td>
<td>5</td>
<td>7</td>
<td>23</td>
<td>12</td>
<td>10</td>
<td>1</td>
<td>9</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>1992</td>
<td>11</td>
<td>12</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>1</td>
<td>9</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>1993</td>
<td>5</td>
<td>5</td>
<td>11</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>7</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>1994</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>1995</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>8</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>1996</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>1997</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>1998</td>
<td>8</td>
<td>0</td>
<td>6</td>
<td>7</td>
<td>9</td>
<td>12</td>
<td>16</td>
<td>11</td>
<td>8</td>
<td>9</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>1999</td>
<td>3</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>8</td>
<td>5</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>2000</td>
<td>2</td>
<td>8</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>17</td>
<td>18</td>
<td>7</td>
<td>11</td>
<td>12</td>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td>2001</td>
<td>13</td>
<td>4</td>
<td>15</td>
<td>18</td>
<td>12</td>
<td>4</td>
<td>3</td>
<td>16</td>
<td>11</td>
<td>21</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1. Identification of the onset of an SC event on September 22, 1999. The SYM-H data shows an increase with a speed of 3.8 nT per ten minutes starting at 1222 (UT). The onset time was identified as 1222:03 (UT), with an accuracy of 1 s, based on the H-component geomagnetic variation observed at the Kakioka Magnetic Observatory.
Table 2. List of SC-triggered plasma wave phenomena observed during the PWS experiment.

<table>
<thead>
<tr>
<th>year</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>1990</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1991</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>1992</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1993</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>1994</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>1995</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1996</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>1997</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1998</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1999</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2000</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2001</td>
<td>6</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

geomagnetic latitude, however, as will be shown in the next section, their propagation modes show a clear dependence on the geomagnetic latitude of the observation point. Therefore, the appearance of plasma wave phenomena corresponded exactly with that of the SCs in any observation region within the plasmasphere, inner magnetosphere and polar ionosphere for each SC event observed during the period of the Akebono satellite’s operation.

4. Plasma waves associated with SCs

4.1. Example of a low latitude (Mlat<45°) event

Figure 3 shows an example of an SC event that occurred at 0720:30 (UT) on March 30, 1990, when the Akebono satellite was located in a low geomagnetic latitude region of the plasmasphere. The dynamic spectra of the VLF plasma waves show a clear intensification associated with the SC near the LHR (lower hybrid resonance) frequency of about 5 kHz. The frequency spectrum of the LHR wave increases above the LHR frequency, with a rising tone signature. The VLF spectra show that the intensity of the plasma waves peaks near the LHR frequency. The enhancement of the VLF plasma waves lasts for about three minutes after the onset of the SC. In this case, a lead time of 40 s in the PWS data was found by comparing the onset time of the enhancement of the whistler mode plasma waves with that of the SC, based on observations performed at Kakioka. In the ELF range, the generation of ion cyclotron waves can be seen above the helium cyclotron frequency, near the oxygen cyclotron frequency. The enhancement of the ELF plasma waves lasts for more than fifteen minutes after the onset of the SC. The onset of the ELF plasma waves occurred about 30 s after the onset of the VLF plasma waves.
4.2. Example of a high latitude (MLat>45°) event

Figure 4 shows an example of an SC event that occurred at 0325:52 (UT) on April 12, 1990 when the Akebono satellite was located in the dayside cusp region. The plasma wave phenomena appear simultaneously with the SC within the 1 Hz to 90 kHz frequency range. Within the frequency range extending from the LHR frequency to the electron cyclotron frequency, the whistler mode plasma waves clearly corresponds with the onset of the SC. The onset of whistler mode plasma wave enhancement occurred at 0325:32 (UT), with a 20-s lead time prior to the onset of the SC. By comparing the intensities of the electric and magnetic fields of the plasma waves, the present plasma wave can be shown to exhibit an electrostatic feature. In general, SC-triggered plasma waves in the whistler mode frequency range resemble the electrostatic whistler mode in the polar region. On the
Fig. 3. Example of an SC event occurring at 0720:30(UT) on March 30, 1990, when the Akebono satellite passed through a low-latitude plasmasphere region. The top panel shows the geomagnetic data provided by the Kakioka observatory; the second panel shows the PWS electric field plasma wave spectrum from 20 kHz to 5.1 MHz, with the white line indicating the electron cyclotron frequency; the third and fourth panels show the VLF electric and magnetic field plasma wave spectra from 3.16 Hz to 17.8 kHz, with the white and red lines indicating the LHR and proton cyclotron frequencies, respectively; the fifth and sixth panels show the ELF electric and magnetic field plasma wave spectra from 0 Hz to 80 Hz, with the red and yellow lines indicating the helium and oxygen cyclotron frequencies, respectively. The enhancement of the electromagnetic whistler mode plasma wave can be seen near the LHR frequency, and the enhancement of the ion cyclotron wave can be seen near the helium frequency in association with the SC.
Fig. 4. Example of an SC event occurring at 0325:52(UT) on April 12, 1990, when the Akebono satellite passed through the cusp region. The formats of the PWS, VLF and ELF spectra are the same as in Fig. 4. The fifth to tenth panels show the LEP data of the electrons and ions for the pitch angles of three sectors and an energy range of 10 eV to 20 keV. In the LEP spectra, 'E-Up', 'E-Perp', and 'E-Down' indicate the electron spectra for the pitch angle ranges of 0°–60°, 60°–120°, and 120°–180°, respectively, followed by the ion spectra for the same pitch angle ranges as for the electrons. Pitch angles of 0° and 180° correspond to the downward and upward directions along the magnetic field line, respectively, since the Akebono satellite is located in the northern hemisphere. In the ELF spectra, the white line shows the proton cyclotron frequency. The enhancement of whistler mode plasma waves and particle fluxes in association with the SC is shown.
other hand, the electromagnetic whistler mode is seen in the low latitude region inside the plasmasphere, as shown in Fig. 2. In the ELF range below 10 Hz, ion cyclotron waves with a magnetic field that is stronger than the electric field intensity are generated near the helium and oxygen cyclotron frequencies. The onset of these enhanced ion plasma waves occurs about 30 s after the SC.

Low-energy electron and ion spectra were available for Akebono satellite observations performed in the high latitude regions. Figure 4 also shows a clear enhancement of low energy electrons and ions associated with the SC-triggered plasma wave phenomena. In the LEP data, the electrons are enhanced in all the pitch angle sectors at 0325:32 (UT) near 80 eV. On the other hand, the ions are enhanced at 0326:52 (UT) near the 1-keV energy range. The enhanced electrons and ions exhibit a peak energy of 800 eV and 10 keV,
respectively. As shown in Fig. 5, the total electron energy flux was clearly enhanced in the downward and upward directions, relative to the magnetic field; on the other hand, the total ion energy flux was more enhanced in the upward direction than in the downward direction, relative to the magnetic field. Interestingly, the upward enhancement of the keV ions in the present case does not coincide with the typical features of cleft observations (Mukai et al., 1990).

5. Discussion and conclusion

Based on Akebono satellite observations collected since 1989, a large number of SC-triggered plasma wave phenomena were identified with a high time resolution. This enabled the relation between SCs identified on the ground and SC-triggered plasma waves observed in the polar ionosphere, plasmasphere and the inner magnetosphere to be clarified with a time resolution of a few seconds. The present analysis shows that the occurrence of plasma wave phenomena corresponds exactly with the occurrence of SCs in the plasmasphere, inner magnetosphere and polar ionosphere. This fact indicates that plasma wave instabilities are very sensitive to SC disturbances, generating electron and ion plasma waves throughout the entire region and producing a significant intensification and modification of the spectra associated with SCs. The present results significantly extend previous knowledge of SC-triggered plasma wave phenomena (Mullayaev and Yachmenev, 1990; Gail and Inan, 1990), thanks to plasma wave observations performed onboard the Akebono satellite and the ground-based, highly sensitive magnetic observations performed at the Kakioka Magnetic Observatory.

The propagation modes of the observed plasma waves vary, depending on the observation position of the Akebono satellite. In the plasmasphere, the electromagnetic propagation modes of whistler and LHR waves are significantly enhanced over long distances. A clear enhancement of the helium and oxygen cyclotron waves is also apparent. In the polar regions, on the other hand, the clear enhancement of electrostatic whistler mode plasma waves is observed over a wide frequency range. The enhancement of electromagnetic ion cyclotron waves in association with SC disturbance is also visible.

By examining the time difference between the onset times of SCs and the enhancement of electron plasma waves, the propagation characteristics of the SC-triggered disturbances can be identified. A detailed analysis of SC-triggered plasma waves could be used to determine the nature of the propagation of SC disturbances in the plasmasphere and polar regions. The results of this delay-time analysis will be described in a future report. Since sufficient satellite observations in the plasmasphere have not yet been obtained, the observation of high-energy particles should be included in future satellite observations performed in the plasmasphere.

Acknowledgments

The Akebono satellite was constructed and launched by the Institute of Space and Astronautical Science (ISAS). The authors thank Prof. Mukai and Dr. Kasahara for providing the LEP, VLF and ELF data sets and their valuable advice and comments. The magnetic field database was obtained from WDC-C2 Kyoto University.
The editor thanks Dr. P.H. Yoon and another referee for their help in evaluating this paper.

References


(Received December 18, 2001; Revised manuscript accepted March 5, 2002)
New model of plasma convection during transpolar arc events

Hitoshi Nakai¹, Yosuke Kamide² and Mitchell Brittnacher³

¹Ibaraki High School, 12-1, Shinjo-machi, Ibaraki 306-0884
²Solar-Terrestrial Environment Laboratory, Nagoya University,
   Honohara, Toyokawa 442-8507
³Geophysics Program, University of Washington,
   Seattle, Washington, U.S.A.

Abstract: A transpolar arc (TPA) event occurred on December 10, 1996, is examined using auroral images from the Polar UVI. Before the appearance of the TPA, the total pressure of the magnetotail observed at $X\sim-24$ $R_E$ had been increasing for about three hours. Since the IMF $B_z$ was strongly positive during this period, the viscous-like interaction probably contributed to the increase in the total pressure of the magnetotail. A clear TPA appeared in association with fluctuations in the IMF $B_z$ component. A subsequent decrease in the IMF $B_z$ component caused a significant equatorward shift of the dawnside auroral oval, probably resulting in the duskward movement of the TPA. The nightside termination of the TPA, however, was fixed at around local midnight of the auroral oval. This observation indicates that a quasi-stable structure existed in the magnetospheric convection pattern. To account for the present observations, a new model for the formation of TPAs is proposed. The model assumes that the interactions between the viscous convection cells and the merging cell occur in the magnetotail, and predicts that the TPA is mapped along magnetic field lines onto a plasma flow channel that branches off from the earthward plasma convection toward the flank of the magnetotail.

I. Introduction

A narrow-band arc with a rather faint luminosity can be occasionally observed extending from dayside to nightside, through the polar cap. The main characteristics of transpolar arcs (TPAs) described by previous studies can be summarized as follows (Frank et al., 1986): (1) TPAs are confined to the region occupied by sunward-convecting, closed field lines; (2) TPAs tend to appear in association with a southward turning and/or change in the polarity of the interplanetary magnetic field (IMF) after a prolonged period of northward IMF; (3) TPAs move duskward (dawnward) when the $B_y$ component of the IMF is positive (negative) and; (4) small auroral breakups are sometimes observed at the same time as TPAs.

TPAs usually remain in the polar cap for more than one hour. This indicates that a quasi-stable structure must exist in the magnetosphere. Frank et al. (1986) hypothesized that TPAs can be mapped along magnetic field lines to a thin plasma structure that bifurcates the tail lobe. If this is the case, a steady, narrow, tailward plasma convection

136
Transpolar arc events

must exist in the neutral sheet region during TPA events (Kan and Burke, 1985). However, as shown in the present study, an intense earthward convection can exist in the magnetotail during TPA events. The relationship between the narrow band of the tailward flow mapped to the TPA and the earthward convection is not clear at present. Furthermore, the stability of the narrow tailward flow surrounded by the earthward convection may be difficult to prove.

In this study, we investigated a TPA event, from its appearance to its disappearance, observed on December 10, 1996. Auroral images obtained by imaging systems onboard the Polar spacecraft and data from the Geotail satellite, the Defense Meteorological Satellite Program (DMSP), the Fast Auroral Snapshot (FAST) satellite, and the Wind satellite were analyzed. A brief description of our observations is presented, and a new model for the formation of TPAs is proposed.

2. Overview of the event

Auroral images obtained by the UVI instrument on Polar (Torr et al., 1995) were inspected to examine the formation and variation of a TPA event occurred between 1430 UT and 1800 UT on December 10, 1996. Data from Wind were used to examine the external conditions during this period (Fig. 1). The propagation time, calculated on the basis of the distance along the X coordinate, was approximately 14 min. However, the variations in the Bz and By components show that Wind was mostly in the away sector during the period of this event. Assuming the garden-hose spiral structure of the IMF, the propagation time is estimated to be approximately 6 min. Data points in Fig. 1 are plotted using the latter time lag. The solar wind speed (540–560 km/s; data not shown) and the dynamic pressure (2.2–3.2 nPa) around the present event were relatively stable. Variations in the total pressure and the magnetic pressure measured by Geotail in the magnetotail are shown in Fig. 2 (Kokubun et al., 1994; Mukai et al., 1994). The total pressure (Pp) scaled at R = 20 Re is plotted in the second panel from the top, assuming that the total pressure is proportional to R−2.40 (Nakai and Kamide, 1994). The inclination of the magnetic field is represented by the ratio of Bx to By in the bottom panel. Geotail was traveling from (−21.6, 5.2, −3.5) to (−26.1, −0.2, −2.6) Re during the time span shown in the figure. The total pressure of the magnetotail continued to increase between 1400 UT and 1555 UT; the elevation angle of the magnetic field decreased during this period, indicating that the magnetotail was recovering from the reduction in the total pressure at ∼1200 UT. The IMF Bz component was continuously positive during this interval (Fig. 1). The values of ε(= VBT2/μ0 sin(θ/2)) were nearly zero during this period except for small enhancements.

The TPA first became visible in an auroral image obtained at 1536 UT. The arc became progressively evident by 1551 UT (upper panel of Fig. 3). While the IMF Bz component was continuously positive for a few hours before 1600 UT, except for several brief southward excursions, the IMF By component was negative with considerable fluctuations between 1510 UT and 1550 UT. Chang et al. (1998) suggested that these changes in the IMF By component may be responsible for the appearance of the TPA (see also Watanabe et al., 2000).

While the Bz component was nearly zero between 1600 UT and 1640 UT, the By
component increased and became steadily positive after 1600 UT, resulting in a gradual increase in the $\epsilon$ parameter from 3.0 $\mu$W/m$^2$ (at 1600 UT) to 9.4 $\mu$W/m$^2$ (at 1700 UT) (Fig. 1). A series of auroral images from Polar show that the nightside portion of the TPA brightened from 1603 UT to 1619 UT (data not shown). Spann et al. (1998) reported a similar auroral activation during a TPA event. They found that an active region appeared to tear away from the oval and move poleward along the TPA. After 1619 UT, our TPA moved duskward, and disappeared at $\sim$1734 UT.

In Fig. 3 auroral UVI images obtained at 1551 UT and 1706 UT are shown with the geomagnetic latitude and magnetic local time coordinates superimposed. The original images were displayed using a false color scale. In this figure, however, the images have been reversed and converted to a gray scale for convenience. The auroral oval and the transpolar arcs at 1551 UT and 1706 UT are delineated by solid and dashed lines, respectively. The solid outline of the 1551 UT image was then overlaid on the 1706 UT image. It is important to note that the nightside termination of the TPA did not move between these two time points, although the central portion of the TPA moved considerably
Fig. 2. Variations in pressure and magnetic field measured in the magnetotail by Geotail. The solid and dashed lines in the top panel represent the total and magnetic field pressures, respectively. In the second panel, $P_{20}$ denotes the total pressure scaled at $R=20\ R_E$. The ratio of the $B_z$ component to the magnetic field magnitude $B_{\text{tot}}$ is plotted in the bottom panel. Geotail traveled from $(−21.6, 5.2, −3.5)$ to $(−26.1, −0.2, −2.6)\ R_E$ during the time span shown in the figure.

Fig. 3. Auroral UV images obtained at 1551 UT and 1706 UT in geomagnetic latitude/geomagnetic local time coordinates. The false color scale has been reversed and converted to a gray scale for convenience. Regions with intense auroral luminosity are outlined by solid and dashed lines for the 1551 UT and 1706 UT images, respectively. The outline of the 1706 UT image is superposed on the 1551 UT image to illustrate the changes in the distribution of the aurora between these two time points.
duskward. The nightside termination, however, eventually moved duskward after 1706 UT (data not shown). Because of the duskward shift of the TPA and the equatorward expansion of the dawnside auroral oval, the polar cap on the dawnside of the TPA was extensively enlarged. Since the $\epsilon$ parameter was moderately enhanced between 1600 UT and 1650 UT (see Fig. 1), this enlargement in the dawnside half of the polar cap probably resulted from the growth of the merging cell, consistent with the positive condition of IMF $B_y$.

3. Convection patterns and open/closed magnetic field configurations

Energy spectrograms of precipitating ions and electrons in the polar cap obtained by the DMSP satellites were examined in detail to evaluate whether the magnetic field lines were open or closed along the orbit of the spacecraft. Data from northern and southern polar-cap crossings by the DMSP F10, F12, and F13 satellites were used in the present study (data not shown). In this paper, the term “polar cap” denotes the ionospheric region located poleward of the auroral UV emissions constituting the auroral oval. The ionospheric regions occupied by closed (or open) field lines are termed as closed (or open) regions hereafter.

On the basis of the energy spectrogram analyses, the magnetic field configurations and the convection pattern in the ionosphere for the periods between $\sim$1600 UT and $\sim$1700 UT are schematically illustrated in Fig. 4. The gray ring indicates the auroral oval. The TPAs are also colored dark gray, representing the horse-collar auroral pattern. The dotted regions correspond with the regions near the dawn or dusk flank of the magnetotail, or the “boundary plasma sheet” in traditional terminology. The resultant region, which is shaped like orchid's petals, indicates the open region. The open region between the TPAs indicates the polar cap portion of the merging cell, which corresponds to the region that expanded between 1551 UT and 1706 UT, as discussed in Fig. 3. On the other hand, the open region between the TPA and the morning (afternoon) auroral oval corresponds to the dawnside (duskside) lobe cell. Since the IMF $B_y$ component was strongly positive during the period of concern, the dawnside patterns (i.e., the dawnside TPA and the dawnside open region) might be severely compressed, making them invisible in the Polar images. The present open/closed patterns are quite similar to those obtained by Henderson et al. (see their Fig. 8). In agreement with them, the dusk- and dawn-side open regions are

Fig. 4. Schematic convection patterns and open/closed magnetic field configurations in the polar region during TPA events. The gray zone indicates the auroral oval and the TPAs. The dotted region is assumed to correspond to the boundary plasma sheet. The white region in the polar cap is the open region. The lines with arrows show the convection patterns. The solid lines indicate the viscous cells. The short-dashed lines are the merging cell. The branches of the merging cell connecting within the TPAs are shown by the long-dashed lines.
assumed not to extend all the way across to the nightside oval.

In Fig. 4, the solid and dotted lines indicate the viscous and merging cells, respectively (Burch et al., 1985). The branches of the merging cell that convect sunward within the TPAs are shown by the dashed lines. Anti-sunward flows of the merging cell are assumed to diverge from the nozzle between the TPAs in the post-midnight sector. Assuming that the IMF \( B_z \) component is moderately positive, the convection lines in Fig. 4 have been drawn so that the merging cell is displaced somewhat downward. Whether or not the nozzle is confined between the TPAs is an important issue in the model of the convection pattern in Fig. 4. Since magnetometer data for the midnight sector are, unfortunately, not available for the time period being examined, the position of the nozzle cannot be confirmed. Pellinen et al. (1990), however, examined active auroras appearing to the west of the nightside termination of a TPA and found that the ion drift velocities were directed eastward near the western edge of the active auroras and westward in the region equatorward of the auroras. Consistent with our model, their results indicate that the region to the west of the TPA’s termination is a portion of the duskside convection cell.

Since the IMF \( B_z \) was negative during the present event, the lobe magnetic field may reconnect with the IMF at the northern magnetopause. Henderson et al. assumed that the lobe field lines which have reconnected with the IMF at high latitudes are convected into the lobe-cell open regions and reconnect at the nightside open/closed boundaries. However, the northern and southern field lines in the lobe cells probably cannot reconnect with each other in the magnetotail, since the lobe cell does not expand to the nightside auroral oval. Therefore, the lobe-cell convection does not significantly influence the nightside convection patterns. The lobe-cell convection patterns are not shown in Fig. 4 for this reason.

Watanabe et al. (2000) suggested that the merged magnetic flux tends to drift along both sides of a TPA, pushing the TPA tailward. In our observations, however, only the polar-cap region between the TPA and the dawnside auroral oval was observed to expand in association with an increase in the dayside reconnection rate (Fig. 3).

4. Interpretations and discussions

Burch et al. (1985) suggested that the plasma convection in the polar region consists of viscous cells, lobe cells, and merging cells. They defined these cells in the following manner. (1) Two “merging” cells, in which closed magnetospheric flux tubes open by merging with solar-wind field lines at the day-side magnetopause, flow anti-sunward across the polar cap, reconnect on the nightside, and return at lower latitudes (the traditional two-cell system). (2) A “lobe” cell is driven by magnetic merging on the polar magnetopause, tailward of the cusp. This cell does not involve a transfer of magnetic flux from the dayside to the night side, but only a stirring and shortening of magnetotail field lines. (3) Two “viscous” cells, one on each side of the polar cap, are driven by a quasi-viscous process on the flanks of the magnetosphere. In this section, we will discuss our observations using this nomenclature.

The IMF \( B_z \) component was continuously positive between 1400 UT and 1555 UT (Fig. 1). The values of \( \varepsilon \) were nearly zero during this period, except for small, brief enhancements, while the total pressure of the magnetotail continued to increase, and the
The elevation angle of the magnetic field decreased (Fig. 2). These observations suggest that the magnetotail was recovering from the reduction in the total pressure at \( \sim 1200 \text{ UT} \), indicating that the magnetic flux was transported from the dayside magnetosphere to the magnetotail to refill the magnetic pressure lost by an unloading process, even when the dayside merging rate was extremely low. The viscous cell is thought to play an important role in this process (Axford and Hines, 1961).

From the observations shown in Fig. 3, the movement of the TPA can be inferred to occur in conjunction with the growth of the merging cell. Since the duskside (duskside) TPA tends to be assimilated with the dayside (duskside) auroral oval in the northern polar cap when the IMF \( B_y \) is positive (negative), only the duskside (duskside) TPA is usually observed when the IMF \( B_y \) is strongly positive (negative). Thus, the TPAs appear to move duskward (dawnward) when the IMF \( B_y \) component is positive (negative) (Frank et al., 1986).

How the sunward convection within TPAs should be mapped to the magnetotail has been a long debated question. Closed field lines threading an area close to the cusp region in the polar cap are generally mapped onto the magnetotail equatorial plane further from Earth. Since plasma flows sunward within TPAs, this implies that a narrow tailward plasma flow exists in the magnetotail (Kan and Burke, 1985). We have, however, found that an intense merging cell can coexist with a TPA. The coexistence of a tailward flow and a predominant earthward convection from the distant-tail neutral line is unlikely.

In our opinion, the difficulty in mapping TPAs on the magnetotail equatorial region cannot be solved as long as one assumes the existence of a tailward plasma channel. One possible alternative is that the earthward plasma flow branches in the magnetotail equatorial plane. In this model, part of the earthward plasma flow in the merging cell is assumed to branch off, leading to the flank of the magnetotail as a result of the interaction between the merging cell and the viscous cell. A conceptual view of this model is shown in Fig. 5. The magnetic field lines (dashed lines) and the plasma convection patterns (solid lines with arrows) in the duskward half of the magnetotail are depicted together with the ionospheric convection pattern shown in Fig. 4. Note that this figure is not scaled. Panels (a) and (b) show the merging cell and the viscous cell, respectively. Panel (c) shows the “branch,” which is assumed to be mapped to the TPA along the magnetic field lines. The “branch” is referred to as the TPA channel hereafter. Although the radial distance at which the earthward flow branches off is not known, it is assumed to occur before the distant neutral line. Given that the spatial scale of the TPA is \( 200 \times 3000 \text{ km}^2 \) and that the \( B_y \) component in the equatorial region is \( 5 \text{ nT} \), the area of the TPA channel is estimated to be \( \sim 180 \ R_e^2 \), based on the conservation of magnetic flux. Therefore, the width of the TPA channel along the \( X \) axis is estimated to be \( \sim 10 \ R_e \).

The mapping of the “lobe” cells into the magnetotail was carefully considered during the formation of our model. Since the IMF \( B_y \) was positive during the period concerned, the geomagnetic field lines reconnected at high latitudes near the cusp were dragged dawnward to form the merging cell. Therefore, such field lines are not likely to be responsible for the duskside open region (Fig. 3). Instead, the field lines from the duskside open region are likely to reconnect with the IMF at the polar magnetopause, tailward of the cusp (Burch et al., 1985). If the closed field lines near the flank of the magnetotail reconnect with the sheath field, some portion of the magnetic field lines in the viscous cell
Fig. 5. Mapping of the ionospheric convection patterns to the plasma convection patterns in the magnetotail. The ionospheric "open/closed" pattern of magnetic field lines is overlaid on the magnetospheric equatorial plane in panels (a)-(c). The flow patterns are shown schematically by the arrows for (a) the merging cell, (b) the viscous cell, and (c) the TPA channel. The dashed lines indicate the magnetic field lines that connect the ionosphere and the magnetotail. The combined patterns in the magnetotail are shown schematically in the panel (d).
would change into those in the lobe cell through the boundary between the lobe and viscous cells. Since the open region of the lobe cell does not reach the nightside auroral oval (Henderson et al., 1996), bifurcation of the lobe's cross section would probably occur only in the near-Earth magnetotail.

Newell and Meng (1992) and Chang et al. (1998) suggested that TPAs often appear in association with a southward turning and/or change in the polarity of the IMF. Our observations were consistent with this suggestion. Additionally, the nighttime half of the TPA was brightened between 1603 UT and 1619 UT in response to an increase in the $\epsilon$ parameter. These observations indicate that the enhancement of the merging cell is a prerequisite for the generation of a TPA. Consistent with this view, our hypothesis predicts that the TPA channel is populated by energetic particles in association with an intensification of the merging cell.

The nozzle of the merging cell is formed between the dawn and dusk viscous cells in our model. Since it is assumed that the branching of the merging cell is caused by its interaction with the viscous cell, the diverging point is fixed upstream of the nozzle point. Therefore, even when the TPA moves duskward in association with the expansion in the open region of the merging cell, the diverging point is thought to be fixed at around the midnight meridian, consistent with the present observation. The diverging point may eventually move duskward, as the nozzle is expanded by the extremely enhanced merging cell.

When the $|B_z|$ component of the IMF is larger than the $|B_z|$ component, the dayside magnetic field reconnection is expected to occur at high latitudes. Merged field lines connected to the southern polar region pass through the equatorial region, draping along the flank of the near-Earth magnetotail. Thus, the closed field lines are likely to be inefficiently driven by the “viscosity” in association with the resulting expansion of the merging cell. This mechanism can explain why the nightside termination of the TPA began to move duskward at $\sim$1706 UT. The dynamics between the viscous cell and the merging cell should be investigated in detail.

Acknowledgments

We deeply thank the staff of the following institutes: Institute of Space and Astronautical Science, Japan (http://www.darts.isas.ac.jp/spdb/index.html), for the magnetic field and plasma data from Geotail; NASA/Goddard Space Flight Center, USA (http://rumba.gsfc.nasa.gov/edweb/istp-public/), for the solar wind data from Wind and the energy spectrometers from FAST; The Johns Hopkins University, Applied Physics Laboratory, USA (http://sd-www.jhuapl.edu/), for the energy spectrometers from DMSP; and World Data Center for Geomagnetism, Kyoto, Japan (http://swdcdb.kugi-kyoto-u.ac.jp/) for the magnetogram data obtained at Kakioka and Phu Thuy. We greatly appreciate the efforts of J.M. Ruohoniemi and F.J. Rich in preparing the data maps for SuperDARN and DMSP ion drift data, respectively. This work was performed by a joint research program at the Solar-Terrestrial Environment Laboratory, Nagoya University.

The editor thanks Drs. N. Sato and M. Pinnock for their help in evaluating this paper.
References


(Received November 1, 2001; Revised manuscript accepted April 16, 2002)
New auroral spectrometer using an acousto-optic tunable filter

Kazuo Shiokawa¹, Yasuo Katoh¹, Mitsugi Satoh¹, Tadahiko Ogawa¹, Makoto Taguchi² and Hisao Yamagishi²

¹Solar-Terrestrial Environment Laboratory, Nagoya University, Honohara, Toyokawa 442-8507
²National Institute of Polar Research, Kaga 1-chome, Itabashi-ku, Tokyo 173-8515

Abstract: This paper reports the performance and capability of a newly developed zenith spectrometer (for measurements of airglow and aurora) that uses an acousto-optic tunable filter (AOTF). The AOTF can scan the pass-band of the spectrometer between 450 and 700 nm with a bandwidth of 2–3 nm by changing the RF driver frequency from 180 to 100 MHz. The absolute sensitivity of the spectrometer is ~0.1–1.5 counts/Rayleigh/s per spectral step. The spectrometer is fully automated. The OI (557.7 nm) airglow line can be clearly identified in test observations of midlatitude airglow performed at Shigaraki Observatory, Japan. Based on an estimate of the signal-to-noise ratio, we conclude that the full auroral spectrum (450–700) nm can be measured by the AOTF spectrometer with a time resolution of ~100 s and a signal-to-noise ratio of ~100 for an auroral emission intensity of 10 kR. An example of the auroral spectra is shown for observations made at Syowa Station in Antarctica.

1. Introduction

The utility of noncollinear acousto-optic tunable filters (AOTFs) for performing spectral measurements has been demonstrated by many researchers over the past two decades (e.g., Chang, 1974, 1977; Sivanayagam and Findlay, 1984; Glenar et al., 1994). AOTFs can be used for spectral measurements from 400 nm to more than 2.5 μm. The devices are solid-state, radio frequency (RF) tunable and are composed mostly of tellurium dioxide (TeO₂) materials.

Figure 1 shows the basic geometry of the AOTF. Incident light passing through the RF-driven AOTF is separated into three beams: (A) a horizontally polarized diffracted beam, (B) a nonpolarized undiffracted beam, and (C) a vertically polarized diffracted beam. Because the acoustic waves propagating in the AOTF work as a grating filter, only a limited bandwidth of 2–3 nm is diffracted into (A) and (C). The wavelength diffracted by the AOTF can be controlled by changing the RF frequency that is transmitted to the AOTF. Thus, the AOTF can be used as a fast-scanning band-pass filter. Moreover, by adjusting appropriate optics and employing a two-dimensional detector, the AOTF can be used as an imaging spectrometer.

Spectral measurements of aurora and airglow have been made mainly using grating
New auroral spectrometer

Fig. 1. Schematic diagram showing the basic geometry of the acousto-optic tunable filter.

spectrometers (e.g., Broadfoot and Kendall, 1968; Okamura and Ejiri, 1992; Johnstone and Broadfoot, 1993; Chamberlain, 1995). Despite the above-mentioned unique capabilities of AOTFs, they have never been used as spectrometers for the measurement of polar aurora, mainly because the small aperture of the AOTFs (maximum \(\sim 10 \text{ mm} \times 10 \text{ mm}\)) results in a small throughput of \(2.4 \times 10^{-3} \text{ cm}^2 \text{ sr}\) for a typical aperture angle of 1.6°. However, the fast-scanning and imaging capabilities of the AOTF are attractive for researchers performing spectral measurement of aurora and other intense atmospheric emissions. In this paper, we report the performance of an AOTF spectrometer that was developed for the measurement of auroral spectra.

2. Instrumentation

Figure 2 shows a schematic diagram of the AOTF spectrometer system. The specifications of the AOTF are listed in Table 1. The AOTF and the RF driver were made by Brimrose Corporation (USA). We developed the appropriate optics and software for the photometer. The optics consists of three identical achromats, each with a focal length of 30 mm and a diameter of 25 mm. The total angle of incident light entering the AOTF is limited to 2.5° by the circular field stop at the first focal point. This almost-parallel incident light provides sufficient separation of the diffracted and undiffracted light spots on the final focal plane. The field-of-view of the front lens is set at 2.5°. The incident light from the front lens passes through the AOTF, which has an aperture of 10 mm \(\times 10\) mm, and is focused on a photomultiplier tube (Hamamatsu R649). The spectrometer is fairly compact, measuring \(\sim 500 \text{ mm} \text{ (height)} \times 150 \text{ mm} \text{ (width)} \times 150 \text{ mm} \text{ (depth)}\).

The photon counts from the photomultiplier are sent to a personal computer through a 30-m signal cable and are recorded onto a hard disk. The shutter just above the field stop is controlled by the personal computer; as a fail-safe system, the shutter is designed to remain closed when daylight is detected by a CdS optical sensor. The personal computer enables the system to be operated automatically by controlling the photon counter, shutter driver, and RF frequency and power of the AOTF driver. The whole system is backed
Fig. 2. Schematic diagram of the AOTF spectrometer system.
up by a power supply (not shown) that can run for more than 10 min in the event of a power failure.

The diffraction of the AOTF produces three light spots on the final focal plane of the photomultiplier tube. To measure the intensity of the monochromatic diffracted light, the undiffracted center spot is masked with a metal plate so that only the two diffracted spots are detected. The diffraction angle is constant independent of the wavelengths of the diffracted beams. The two side spots have a 90° difference in polarization but can not be distinguished from each other on the photomultiplier tube.

3. Calibration

As listed in Table 1, the AOTF has a scanning range of 450–700 nm with a spectral resolution of 0.8 nm (at 450 nm) to 3.0 nm (at 700 nm); the desired scanning range and spectral resolution can be obtained by changing the RF drive frequency from 180 to 100 MHz. We have divided this frequency range into 250 equal steps. The center wavelength, absolute sensitivity, and bandwidth of the spectrometer for these 250 steps are shown in Fig. 3. These calibrations were performed out using a 2-m integrating sphere and a monochromatic light source at the National Institute of Polar Research, Japan. The sensitivities in Fig. 3 were obtained by measuring white light from the integrating sphere, the brightness of which was known. The center wavelength and bandwidth were obtained by measuring monochromatic light with a bandwidth of 0.1 nm for every 0.5-nm step between 450 nm and 700 nm.

A quasi-linear relation between the RF frequency and the center wavelength can be seen in Fig. 3. The fluctuations in sensitivity and bandwidth occur because the spectral resolution of the calibration for determining the transmission function was relatively low (0.5 nm) compared to the band width (2–3 nm). In any case, the sensitivity decreases from 1.5 counts/R/s to 0.1 counts/R/s (per spectral step) and the bandwidth increases from ~2 nm to ~3 nm as the center wavelength increases from 450 nm to 700 nm.

The total angle of the incident light that passes through the AOTF was set at 2.5°. This sets the total throughput $A \Omega$ of the optics at $A \Omega = 2\pi(1 - \cos (1.25\degree)) \times 1 = 1.5 \times 10^{-3}$ sr cm², where $A$ and $\Omega$ are the area (1 cm²) and the solid angle (sr) of the light that passes through the AOTF, respectively. Assuming that the total transmission $T$ of the optics,
including the diffraction efficiency of the AOTF, is 0.5 and the quantum efficiency $q$ of the photomultiplier tube is 0.15 at a wavelength of 400 nm, the ideal sensitivity $S = \left(10^4/4\pi\right) A\Omega Tq$ becomes 8.8 counts/R/s. This value is about 6 times larger than the actual value (1.5 counts/R/s) obtained by the calibration.

We have yet to discover the cause of this discrepancy, but several factors can reduce the ideal sensitivity. First, creating an appropriate optical mask that allows the two diffracted beams to pass and blocks the undiffracted center beam is quite difficult. To avoid contamination from the intense undiffracted beam, the two holes on the optical mask are slightly smaller than the theoretically optimal diameter. A second possible reason for the above discrepancy is that the quantum efficiency of the photomultiplier tube (R649) may be smaller than the expected value because of the non-uniformity of the photoelectric surface. We had to use the area around the edge of the photoelectric surface (8 mm x 5 mm) of R649. Because the two diffracted light spots are separated by about 8 mm, the area around the edge of the photoelectric surface of R649 (8 mm x 5 mm) had to be used. These problems could be avoided by using imaging detectors, such as cooled-CCDs, at the
final focal plane.

Figure 4 shows the average transmission function of the AOTF photometer. This function is obtained by averaging all the transmission functions for the 250 frequency steps and normalized by both the center wavelength and the peak transmission. The vertical bars indicate the standard deviations. A sub-peak transmission about 2.5 nm away from the transmission center (towards the longer wavelength) can be clearly seen in this figure. This sub-peak probably produced by a power overdrive. The RF power used to drive the AOTF might be too strong. If so, this would explain why the measured bandwidth in Fig. 3 is larger than the expected instrumental specification listed in Table 1. Unfortunately, the airglow and auroral measurements performed in Japan and the Antarctica that are described below were performed under this overdrive situation producing broader bandwidths. Further calibrations are necessary to determine the optimal RF power.

4. Observations

We used the AOTF spectrometer described in the present report to observe mid-latitude nocturnal airglow at Shigaraki Observatory (34.8°N, 136.1°E), Kyoto University, in July 1999 and Antarctic aurora at the Syowa Station (69.0°S, 39.6°E, magnetic latitude: −70.3°) between February–October 2000. The airglow and auroral spectra were obtained by scanning the RF frequency from 180 to 100 MHz in 250 equal steps. The exposure time of each step was 0.1 s. Thus, one spectrum was obtained every 25 s. However, the spectrum had to be further averaged because of the noise produced by the photomultiplier tube. The averaging interval varied according to the signal intensity.

Figure 5 shows an example of a zenith airglow spectrum measured at Shigaraki on
Fig. 5. Zenith airglow spectrum measured by the AOTF spectrometer between 1115 and 1215 UT (2015–2115 LT) on July 10, 1999, at the Shigaraki Observatory, Japan. Data obtained from 144 scans (one hour) were averaged. Several lines from city lights (Mercury and Cadmium) and two OI airglow lines (557.7 nm and 630.0 nm) can be identified. The expected locations of other airglow and auroral lines (Hα, Hβ, and Na) are also shown.

Fig. 6. Variation in OI (557.7 nm) airglow intensity measured by the AOTF spectrometer on July 10–11, 1999, at the Shigaraki Observatory. The heavy and light curves are the intensities obtained by averaging 72 scans (7.2-s exposure) and 10 scans (1.0-s exposure), respectively. The data shown by these curves were obtained every 30 min and 250 s, respectively, because one full scan over the 450–700 nm range requires 25 s. Background sky emissions and dark counts were subtracted from the plotted data using data for a neighboring wavelength.
July 10, 1999, averaged over one hour (=25 s × 144 scans). Intense mercury emission lines are identified. These lines probably originated from a light source in an observatory building that was located about 200 m away from the spectrometer. Airglow lines at OI 557.7 nm and 630.0 nm are in the spectrum, but the sodium (Na) line at 589.3 nm is indistinct. The 144-scan average produced a noise level of ~2 counts (rms), as can be seen in the figure.

Using the calibration parameters shown in Fig. 3, we calculated the variations in absolute intensity of the O I (557.7 nm) line for this night. The result is shown in Fig. 6. The intensity increases from ~100 R to ~300 R between 21 LT and 03 LT. Background sky emissions were subtracted from the data by using the counts for a neighboring step. The noise level of the 10-scan average data was ~50 R (rms). The 72-scan average (72 × 25 s = 30 min), which corresponds to an exposure time of 7.2 s per step, produces a reasonable curve for airglow variation in the 100-300 R intensity range.

Figure 7 shows the dynamic spectrum (24 scans = 10-min averages) of auroral emissions measured on April 6, 2000, at Syowa Station, Antarctica. An intense magnetic storm started at 1639 UT on this day. The maximum Dst index (provisional) was −321 nT at 0100 UT on April 7. An intense green line (557.7 nm) was observed throughout the observation. The intensity of the emissions peaked at between 1800–1900 UT (2100–2200 LT). The red line (630 nm) was enhanced at around 2200–2300 UT (0100–0200 LT). Several pictures of a spectacular red aurora were taken around this time. It should be noted in this figure that when the auroral emission lines were enhanced, the background continuum emission was also enhanced (e.g., 1800–1900 UT). The reason for this is...
Fig. 8. Zenith airglow spectrum measured by the AOTF spectrometer between 1820 and 1920 UT (2120–2220 LT) during a magnetic storm on April 6, 2000, at the Syowa Station, Antarctica. Data obtained from 144 scans (one hour) were averaged. Many auroral lines, including O I (557.7 nm and 630.0 nm), can be identified.

Fig. 9. Variations in OI (557.7 nm and 630.0 nm) auroral intensity measured by the AOTF spectrometer on April 6, 2000, at the Syowa Station, Antarctica. The heavy and light curves are the intensities obtained by averaging over 10 scans (1.0-s exposure) and 4 scans (0.4-s exposure), respectively. The data shown by these curves were obtained every 250 and 100 s, respectively, since one full scan over 450–700 nm range requires 25 s. Background sky emissions and dark counts were subtracted from the plotted data using data for a neighboring wavelength.
probably that the measured spectrum was contaminated by zero-order undiffracted emissions arising from multiple scattering in the optics.

The spectrum of the most intense auroral emission at 1820–1920 UT (2120–2220 LT, averaged over one hour = 144 scans) is shown in Fig. 8. Various auroral emission lines and bands can be identified in this spectrum, as indicated. Because of the sub-peak of the filter transmission, shown in Fig. 4, the spectral separation is not very good. For example, the two OI emissions at 630.0 nm and 636.4 nm partially overlap.

Similar to the airglow measurements at Shigaraki, we also calculated the variations in absolute intensity of the OI (557.7 nm and 630.0 nm) on this night. The result is shown in Fig. 9. Background sky emissions were subtracted using the counts for a neighboring step. Thus, the absolute intensity calculation was not affected by the contamination of the zero-order undiffracted beam. The green line (557.7 nm) reached its peak intensity at 1800–1900 UT, while the intensity of the red line (630.0 nm) exceeded that of the green line at around 2200 UT. The figure clearly shows that a 4-scan average (4×25 s = 100 s temporal resolution) already produces a reasonable curve of auroral intensity variation. For example, the curve does not show noise fluctuations at around 20–21 UT, even when the auroral intensity is small.

5. Summary and discussion

We have constructed a compact zenith spectrometer that uses an acousto-optic tunable filter with a wavelength range of 450–700 nm. The bandwidth and sensitivity of the spectrometer are 2–3 nm and 0.1–1.5 counts/R/s/step, respectively. We found that an overdrive of the RF signal causes the bandwidth to broaden with a sub-peak occurring in the transmission function. The spectrometer is automatically operated by a personal computer. Test observations of midlatitude airglow indicate that variations in airglow intensity of a few hundred R can be measured using an exposure time of 7.2 s (72-scan average = 30-min temporal resolution). Automatic measurements of an aurora were performed using the spectrometer at Syowa Station between February and October 2000. Various auroral spectra were obtained from the observations.

One set of full-scan data is obtained every 25 s using an exposure time of 0.1 s/step. This value sets the minimum time resolution of the spectral measurement. Depending on the intensity of the aurora, the averaging time interval can be varied in multiples of 25 s. As shown in Fig. 8, various auroral emission lines are included in the range of measured wavelengths. The auroral emission intensity is much greater than that of airglow. The estimated noise level is 2 counts for 144 scan averages, as shown in Fig. 5. This means that the noise level per scan is 24 counts/0.1 s/step (= 2 × 144). For the 10-kR auroral emission, the output was 1000 counts/0.1 s/step for a sensitivity of 1.0 counts/R/s/step. Thus, a 4-scan (100 s) average produces a signal-to-noise ratio of ~100 ( = 1000/(24/√4)). For the actual measurement shown in Fig. 9, a temporal resolution of 100 s seems to be sufficient to measure variations in intensity of more than 1 kR.

Since the personal computer can arbitrary change the RF frequency, the program can be set to only select the frequencies that correspond to auroral emission lines. In this situation, the actual time resolution becomes much better than that of a full scan.

The automatic operation of the AOTF spectrometer is fairly stable because the
spectrometer does not have any moving parts (for wavelength scans). The AOTF can arbitrary change the center wavelength. These characteristics are suitable for future space-borne measurements of aurora and other atmospheric emissions. Recently, considerable progress has been made in the construction of a highly sensitive imaging detector using cooled-CCD devices. By combining this detector with the AOTF, the spectrometer could be used for imaging spectroscopy of aurora.

Acknowledgments

We would like to gratefully acknowledge T. Katoh of the Solar-Terrestrial Environment Laboratory, Nagoya University, for his skillful support in the construction of the spectrometer hardware. H. Fukunishi and Y. Takahashi of Tohoku University provided the suggestion of using an acousto-optic tunable filter as a spectrometer. The spectrometer was calibrated using the facilities at the National Institute of Polar Research, Japan. Test observations at Shigaraki Observatory were performed out in collaboration with the Radio Science Center for Space and Atmosphere of Kyoto University. The authors wish to thank the members of the 41st Japanese Antarctic Research Expedition for their various support in performing the auroral observations at Syowa Station. This work was supported by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (11440145).

The editor thanks Dr. T. Aso for his help in evaluating this paper.

Reference


(Received February 1, 2002; Revised manuscript accepted April 18, 2002)
Polar Patrol Balloon experiment in Antarctica during 2002–2003

Akira Kadokura1, Hisao Yamagishi1, Natsuo Sato1, Masaki Ejiri1, Haruto Hirosawa2, Takamasa Yamagami2, Shoji Torii2, Fumio Tohyama2, Michio Nakagawa3, Toshimi Okada4 and Edgar A. Bering5

1National Institute of Polar Research, Kaga I-chome, Itabashi-ku, Tokyo 173-8515
2The Institute of Space and Astronautical Science, Sagamihara, Kanagawa 229-8510
3Institute of Physics, Kanagawa University, Kanagawa-ku, Yokohama 221-8686
4Faculty of Engineering, Tokai University, Hiratsuka, Kanagawa 259-1292
5Faculty of Science, Osaka City University, Sumiyoshi-ku, Osaka 558-8585
6Faculty of Engineering, Toyama Prefectural University, Toyama 939-0398
7Physics Department, University of Houston, Houston, Texas, 77204-5506, U.S.A.

Abstract: The first scientific campaign of the Polar Patrol Balloon (PPB) experiment (1st-PPB) was performed at Syowa Station in Antarctica during 1990-1991 and 1992-1993. Based on the fruitful results of the 1st-PPB experiment, the next campaign (2nd-PPB) will be carried out in the austral summer of 2002–2003. This paper summarizes the 2nd-PPB experiment. Four balloons in total will be launched to make astrophysics observations (1 balloon) and upper atmosphere physics observations (3 balloons). The first payload will carry a very sophisticated instrument that will observe primary cosmic-ray electrons in the energy range of 10 GeV–1 TeV. The payloads of the latter 3 flights are identical to each other. They will be launched in as rapid a succession as weather conditions permit to form a cluster of balloons during their flights. Such a “Balloon Cluster” is suitable for observing the temporal evolution and spatial distribution of various phenomena in the various magnetospheric and ionospheric regions and their boundaries that the balloons will traverse during their circumpolar trajectory. The expected flight duration of each balloon is 20 days. Observation data will be obtained mainly by a satellite communication system with a much higher temporal resolution than that used in the 1st-PPB experiment.

1. Introduction

The “Polar Patrol Balloon (PPB)” experiment is a project involving long-term observations in the Antarctica using stratospheric zero-pressure balloons and the stable circumpolar easterly wind that occurs during austral summer. A feasibility study and the development of the balloon technology for the PPB started in 1984 as a 5-year project of the upper atmosphere physics group at the National Institute of Polar Research (NIPR) in collaboration with the Institute of Space and Astronautical Science (ISAS) and nationwide scientists (Nagata et al., 1985; Nishimura et al., 1985). Two and one test flights of the PPB were carried out in 1987 and 1990 by the 28th and 30th Japanese Antarctic Research Expeditions (JARE), respectively, at Syowa Station in Antarctica. These test flights
confirmed the reliability of the predicted long-duration circumpolar trajectory and the effectiveness of the balloon technology for the PPB, including the flight altitude control (auto-ballasting) system, and the ARGOS satellite communication system (Kadokura et al., 1991; Kadokura, 1995; Ejiri et al., 1993; Yamanaka et al., 1988). After these preparatory phases, the first scientific campaign of the PPB (1st-PPB) was performed during 1990–1991 and 1992–1993 by JARE-32 and JARE-34, respectively (Fujii et al., 1989; Ejiri et al., 1995; Nishimura et al., 1994). A total 6 balloons (3 balloons per period) were launched in the 1st-PPB to perform observations of the total force of the geomagnetic field (PPB-1), ionospheric and magnetospheric phenomena (PPB-2, 4, 5), atmospheric ozone and aerosol (PPB-3), and cosmic rays and auroral X-rays (PPB-6), respectively. A summary of the 1st-PPB experiment is given in Table 1, and the trajectories of the 6 balloons are shown in Fig. 1. Observation data were obtained via the ARGOS system when the balloons exited the receiving range of Syowa Station. Twenty and forty ID numbers were assigned to one ARGOS transmitter to transfer the data obtained with a sampling rate of 16 and 32 bytes/30s in the 1990–1991 and 1992–1993 experiments, respectively. The details of the Multi-ID ARGOS transmitter system was described by Fujii et al. (1992). A geomagnetic anomaly arising from the earth’s crust around Antarctica was thoroughly investigated using total force observations performed with a proton magnetometer onboard PPB-1, -2, -4 and -5 (Tohyama et al., 1993). Magnetospheric phenomena were observed using the instruments onboard PPB-2, -4, -5 and -6 (Ebihara et al., 1996; Hiraisima et al., 1999). Flight PPB-6 was particularly successful, allowing a thorough study of auroral X-ray phenomena (Kodama et al., 1995; Suzuki, 1996). Cosmic-ray protons were also observed during the PPB-6 flight (Yamagami et al., 1994). Observations of the ozone and aerosols in the Antarctic ozone hole were successfully performed by PPB-3 for the first time (Kanzawa and Kondo, 1991; Kanzawa et al., 1994; Hayashi et al., 1994). Despite such fruitful scientific

<table>
<thead>
<tr>
<th>PPB no</th>
<th>launching date</th>
<th>flight duration (days)</th>
<th>balloon volume ($x 10^4 m^3$)</th>
<th>payload weight (kg)</th>
<th>ballast weight (kg)</th>
<th>total weight (kg)</th>
<th>control altitude (km)</th>
<th>sampling rate</th>
<th>observation item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dec. 25, 1990</td>
<td>28</td>
<td>25</td>
<td>114.0</td>
<td>152</td>
<td>373.5</td>
<td>28</td>
<td>16byte/30sec</td>
<td>total-B</td>
</tr>
<tr>
<td>2</td>
<td>Jan. 05, 1991</td>
<td>30</td>
<td>32</td>
<td>191.5</td>
<td>152</td>
<td>471.0</td>
<td>28</td>
<td>16byte/30sec</td>
<td>total-B, vector E-field</td>
</tr>
<tr>
<td>3</td>
<td>Sep. 23, 1991</td>
<td>6</td>
<td>5</td>
<td>161.8</td>
<td>145</td>
<td>370.3</td>
<td>18</td>
<td>32byte/2min</td>
<td>ozone, aerosol</td>
</tr>
<tr>
<td>4</td>
<td>Dec. 26, 1992</td>
<td>9</td>
<td>39.7</td>
<td>189.0</td>
<td>150</td>
<td>487.8</td>
<td>28</td>
<td>32byte/30sec</td>
<td>total-B, vector B</td>
</tr>
<tr>
<td>5</td>
<td>Dec. 30, 1992</td>
<td>43</td>
<td>39.7</td>
<td>186.5</td>
<td>150</td>
<td>483.5</td>
<td>28</td>
<td>32byte/30sec</td>
<td>total-B, vector B</td>
</tr>
<tr>
<td>6</td>
<td>Jan. 05, 1993</td>
<td>27</td>
<td>59.5</td>
<td>95.0</td>
<td>150</td>
<td>434.2</td>
<td>30</td>
<td>32byte/30sec</td>
<td>cosmic ray (X-ray, proton)</td>
</tr>
</tbody>
</table>
Fig. 1. Trajectories of the 6 balloons in the 1st-PPB campaign.

and balloon technological results in the 1st-PPB campaign, several objectives were not accomplished in the campaign, and the need for future experiments arose. The payload and balloon configurations of PPB-4 and PPB-5 were identical to each other to enable simultaneous observations at different local times. However, this objective was not fully accomplished because of failures in the altitude control systems of both balloons. Furthermore, the limited data acquisition rate of the Multi-ID ARGOS system was not sufficient for the observation of higher frequency phenomena, such as Pc3 or Pi2 pulsations, or for observations by more sophisticated instruments that required a larger amount of data to be transmitted. Scientific discussions for a future PPB experiment, based on the results of the 1st-PPB, began in 1995. Several workshops for that purpose were held at NIPR in 1995, 1997, and 1999. Eventually, the next PPB project (2nd-PPB) was scheduled as a 3-year project to be performed during 2000-2003. In the following sections, the details of the 2nd-PPB project will be introduced.
2. Overview of the 2nd-PPB project

The 3-year project will be conducted as follows: fabrication of the scientific payload for astrophysics observations in the 1st fiscal year (April 2000–March 2001), fabrication of the scientific payload for geophysics observations in the 2nd fiscal year (April 2001–March 2002), and integration of the total systems for each observation, and launching the payloads at Syowa Station in Antarctica in the 3rd fiscal year (April 2002–March 2003).

The campaign in Antarctica will be performed during late December 2002 to January 2003 by the 44th Japanese Antarctic Research Expedition (JARE-44). The primary launching window at Syowa Station is scheduled for the period from late December 2002 to early January 2003, mainly because of the requirements for a perfect westward circumpolar trajectory. After that period, the wind direction in the stratosphere gradually changes.

---

Fig. 2. Expected PPB trajectory. The contour of the equi-invariant latitude is shown by the gray curve. The location of ground stations, their FOVs above 20° elevation projected at an altitude of 120 km, and the FOVs of the HF-radars in the SuperDARN network are also shown.

A. Kadokura et al.

This document is provided by JAXA.
from easterly to westerly. One 100000 m$^3$ zero-pressure balloon (B100) and three 50000 m$^3$ balloons (B50) will be launched during this period to perform astrophysical and geophysical observations, respectively. The flight altitudes will be controlled to within 35–40 km and 30–35 km for the astrophysical and geophysical observations, respectively, by the auto-ballasting system. The expected life time of each balloon is more than 20 days. The expected trajectory of the balloons is shown in Fig. 2. After their launch from Syowa Station, they will drift westward almost along the equi-geographic latitude line between $-60^\circ$ and $-70^\circ$ around the Antarctica with a circumpolar period of about 2 weeks. Since the geomagnetic dipole axis is tilted from the rotation axis of the earth, they will traverse a large geomagnetic latitudinal range between 50$^\circ$ and 80$^\circ$ invariant latitude (ILAT), shown by the grey-encircled contour in Fig. 2. The location of the ground stations, their field of views (FOVs) above 20$^\circ$ elevation projected at an altitude of 120 km, and the FOVs of the HF-radios in the SuperDARN network (Greenwald et al., 1995) are also shown in Fig. 2. The balloons are expected to float within the FOVs of several ground stations and HF-radios during their circumpolar trajectory. Data will be downlinked via 64 kbps telemetry to several ground stations, including Syowa Station, and via a satellite data link system the Iridium satellite phone system, in which maximum transfer rate is 10 kbps. This transfer rate is about 2 orders of magnitude higher than the Multi-ID ARGOS system used in the 1st-PPB campaign. Accurate time and balloon positions will be obtained using GPS (Global Positioning System). Some basic house-keeping data, including the GPS data, will be obtained via the ARGOS satellite system.

The scientific payload for the astrophysical observations is an instrument for observing the primary cosmic-ray electrons in the energy range of 10 GeV to 1 TeV; this instrument has been named as PPB-BETS (Balloon-borne Electron Telescope with Scintillating Fibers for the PPB experiment). The instrumentation and scientific purpose of the BETS has been briefly described by Torii et al. (1999, 2000). Five instruments will be installed in the balloon payloads for the geophysical observations: an EMW (electromagnetic wave detector) to observe electromagnetic waves in the ULF to LF band range, an EFD (electric field detector) to observe the vector electric field, an MGF (magnetic field detector) to observe the vector geomagnetic field, an AXI (auroral X-ray imager) to observe auroral X-ray emissions, and a TEC (total electron content) to observe the total electron content of the ionosphere. The flight configurations for the astrophysical and geophysical observations are shown in Fig. 3a and 3b, respectively. The estimated payload weight, total ballast weight, total weight, and total lift with 11% of free lift for the astrophysical and geophysical observations are 500, 230, 730, and 810 kg and 340, 160, 500, and 555 kg, respectively. Electric power for the instruments is mainly supplied by a solar battery system. Five solar panel units are installed on each surface of a gondola. The maximum power output from one unit is 45 W, so each surface supplies 225 W. Each balloon has a parachute. If the balloon moves back within the FOV of Syowa Station after the circumpolar trajectory, the payload will be cut-down by a command from the ground and the recovery of the payload will be attempted. For the geophysical observations, two loop antennae for the EMW are set around the balloon surface orthogonal to each other, and the received signals are processed by the EMW main instrument in the sub-gondola. The sub-gondola consists of the EMW instrument and a reel-down. Since the EFD must avoid electric interference from charges on the balloon surface, the main gondola will be
located 100 m below the sub-gondola. The EMW data processed in the sub-gondola is transmitted to the main gondola 100 m below via an RF modem. For the observational requirements of the EFD and MGF, a motor will be used to rotate the main gondola at a steady rate of 2 rpm. In the following sections, further details of the scientific instruments and scientific purposes of each type of observation will be described.

3. Astrophysical observations

Long-term observations by PPB are very useful for observing high-energy primary cosmic-ray electrons because the flux of these electrons is generally very small. The number of electrons expected to be observed during the 20-day PPB observation period is shown in Table 2. The statistical confidence of these observations should be increased by about 100 times, compared with that of usual experiments that last for less than one day. The lifetime of cosmic-ray electrons in the Galaxy is very short because they rapidly lose their energy through synchrotron and inverse Compton processes as they travel. The number of electron sources whose electrons reach Earth decreases as the energy of the electrons increases; SNRs (supernova remnants) and pulsars are known candidates for high-energy electrons sources, and the expected energy spectrum can be predicted by assuming the diffusion parameters. Table 3, taken from Torii et al. (1999), lists such candidates, and Fig. 4 shows the expected energy spectrum (black lines) and observed values in previous experiments (symbols). The expected observation range and values of
Table 2. Expected number of cosmic-ray electrons in the 20-days PPB observation.

<table>
<thead>
<tr>
<th>Energy</th>
<th>&gt; 10 GeV</th>
<th>&gt; 100 GeV</th>
<th>&gt; 1000 GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>electron number</td>
<td>1.1 x 10^5</td>
<td>5.6 x 10^2</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 3. List of SNRs and pulsars as candidates of nearby electrons sources (after Torii et al., 1999).

<table>
<thead>
<tr>
<th>SNR</th>
<th>Pulsar</th>
<th>Distance (kpc)</th>
<th>Age (yr)</th>
<th>Emax (TeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SN 185</td>
<td></td>
<td>0.95</td>
<td>1.8 x 10^5</td>
<td>130</td>
</tr>
<tr>
<td>S 147</td>
<td></td>
<td>0.8</td>
<td>4.6 x 10^5</td>
<td>50</td>
</tr>
<tr>
<td>G 65.3+5.7</td>
<td></td>
<td>0.8</td>
<td>2.0 x 10^4</td>
<td>12</td>
</tr>
<tr>
<td>Cygnus Loop</td>
<td></td>
<td>0.77</td>
<td>2.0 x 10^4</td>
<td>12</td>
</tr>
<tr>
<td>Vela</td>
<td>B 0833-45</td>
<td>0.5</td>
<td>2-3 x 10^4</td>
<td>8-12</td>
</tr>
<tr>
<td>Monogem</td>
<td></td>
<td>0.3</td>
<td>1.0 x 10^5</td>
<td>2.3</td>
</tr>
<tr>
<td>Loop 1</td>
<td></td>
<td>0.17</td>
<td>2.0 x 10^5</td>
<td>1.2</td>
</tr>
<tr>
<td>Geminga</td>
<td>IE 0630-178</td>
<td>0.3</td>
<td>3.4 x 10^5</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Fig. 4. Energy spectrum of the high-energy, cosmic-ray electrons. The expected energy spectrum (black lines) and the expected observation range and values of the PPB-BETS (blue line) are shown with the observed values obtained in previous experiments (symbols).
the PPB-BETS is also shown as a thick, straight blue line in Fig. 4. Observations of the precise energy spectrum using the PPB-BETS should reveal the types of sources responsible for the spectrum and the most reliable combination of diffusion parameters. Precise observations around 10 GeV will also enable us to analyze the details of solar-modulation effects and the reacceleration mechanism in the Galaxy.

The design of the PPB-BETS is basically the same as that for a BETS developed for lower latitude observations (Torii et al., 2000). However, various improvements have been made (Torii et al., 1999). The PPB-BETS instrument consists of a shower detector, which incorporates an imaging calorimeter and a trigger system, a data acquisition system and a telemetry system. The shower detector consists of multiple layers of scintillating fibers and lead plates. Incident high energy electrons induce a shower inside the multilayer detector, which is observed as a three-dimensional characteristic pattern of light emissions. This optical image is intensified by a set of image intensifiers and recorded by a CCD camera system. Two sets of image-intensified CCD camera systems are used to simultaneously image the emission patterns on the X-Z and Y-Z planes, respectively; the three-dimensional pattern, where the Z-axis represents the vertical direction and the X-Y plane represents the horizontal plane, is then reconstructed. Figure 5, taken from Torii et al. (2000), shows an example of the images of the shower on the X-Z plane (panel (a)) and the Y-Z plane (panel (b)). All the data are processed electrically and transmitted via telemetry, hence the recovery of the instrument is not essential for data acquisition. This point is an important advantage for use in the PPB experiment because recovery in the Antarctic region is very difficult, especially for the PPB experiment. For example, an alternative observation technique that utilizes an emulsion chamber (e.g., Nishimura et al., 1997) can be used to observe spectrum up to a few TeV, but the instrument must be recovered in fact to obtain the observation data. The PPB-BETS has several other advantages over the emulsion chamber system, as discussed by Torii et al. (1999, 2000).
4. Geophysical observations

One of the main purposes of the geophysical observations that will be performed in the 2nd-PPB experiment is to observe various phenomena in the magnetospheric and ionospheric boundary layers. As shown in Fig. 2, the PPBs can traverse a large latitudinal range of magnetic coordinates and will float across several interesting regions and boundaries during their circumpolar trajectory. Figure 6 shows these regions and boundaries in the magnetosphere, and Fig. 7 shows the approximate locations of their ionospheric projection along the geomagnetic field lines. From lower latitudes, the PPBs can remain in the region of the plasma sphere, plasmapause, trough, auroral oval (plasma sheet), plasma sheet boundary layer (PSBL), low latitude boundary layer (LLBL), mantle region, cusp region, and polar cap (lobe) region. Small circles on the expected trajectory in Fig. 7 indicate the daily positions of the balloon. The PPBs are expected to remain for about 1–3 days in each region and boundary during their flight. We are planning to launch three balloons in as rapid succession as weather conditions permit (with an interval of less than 1 day) to place all three balloons in adjacent areas of the trajectory (separated by a few hundred kilometers) to form a cluster of balloons. Hence, the three balloons are referred to as a "Balloon Cluster", recalling the successful CLUSTER-2 satellite project in which four satellites were placed in the magnetosphere. The payload configurations of all three balloons are identical. Simultaneous observations with identical sets of instruments within and outside of the target regions and boundaries should enable the spatial distribution and temporal variations of various phenomena occurring around these regions and boundaries to be studied.

A list of the scientific objectives of the geophysical observations in the 2nd-PPB is shown in Table 4. Figure 8, taken from Fujii et al. (1994), shows the electromagnetic parameters across an area of intense auroral activity, westward traveling surge (WTS), observed by a low-altitude polar-orbiting satellite (DE-2). The satellite moved equator-
Fig. 7. Approximate location of the ionospheric projection of the various regions and boundaries in the magnetosphere. The expected trajectory of the PPB is also shown by the thick black line. The circles on the trajectory indicate the daily positions of the balloon.

Table 4. A list of scientific targets of geophysical observation in the 2nd-PPB campaign.

- Phenomena around the PSBL region, associated with substorm
- SAID (sub-auroral ion drift), associated with substorm
- Wave-particle interaction around the plasmapause
- Flux variation of the radiation belt particles, associated with storm
- Phenomena around the cusp, cleft, and mantle region, responding to variation in solar wind parameters
- Response of the polar cap convection to solar wind parameters
- Electromagnetic dynamics in the vicinity of auroral activity (e.g., westward traveling surge, omega band)
- Local time dependence of the relationship among electric field, magnetic variation, and particle precipitation
- Source mechanism for the quasi-periodic VLF emission and particle precipitation in the Pc3 range
- Relationship between the ULF pulsation and auroral X-ray pulsation in the Pc5 range
- Quick response of electric field and magnetic variation at mid-latitude, to variation in the solar wind parameters
- Simultaneous observation of the Global Mode Pc5 with magnetospheric satellites and SuperDARN
- Variation in the atmospheric electric parameters, associated with solar events (e.g., proton event)
- Relationship between the plasma flows observed by the SuperDARN and electric field and magnetic field vectors observed by PPB
- Relationship between the radar echo region observed by the SuperDARN and fluctuation of the total electron content and electric field observed by PPB
ward as time advanced from left to right in Fig. 8. Intense equatorward and poleward electric fields are visible around the poleward and equatorward boundaries of the auroral electron precipitation region, respectively. The former one is frequently observed in the PSBL region and is considered to be closely associated with the process around the X-type neutral line in the tail (e.g., Burke et al., 1994). A characteristic pattern of field-aligned currents (FAC), an intense hydromagnetic wave activity, and a velocity-dispersed ion precipitation structure (VDIS) are also observed around the PSBL region (e.g., Fukunishi et al., 1993; Wygant et al., 2000).

The latter one around the equatorward boundary is called a "sub-auroral ion drift (SAID)" and is usually observed during the late expansion to recovery phase of a substorm predominantly in the pre-midnight, sub-auroral region around the trough region (e.g., Spiro et al., 1979; Anderson et al., 1993; Karlsson et al., 1998). These characteristics of the PSBL region and SAID have been mainly studied using low-altitude satellite observations. To our knowledge, very few observations of their temporal variations have been performed with a higher resolution than that of the satellite revolution period.

Auroral X-ray observations by PPB-6 in the 1st-PPB experiment revealed that the characteristic energy of auroral X-rays reaches a maximum value around the plasmapause region, as shown in Fig. 9 (Suzuki, 1996). Suzuki (1996) suggested that this result is due
Fig. 9. Latitudinal distribution of the characteristic energies of auroral X-rays, observed by PPB-6 in the 1st-PPB experiment (after Suzuki, 1996).

to the characteristics of resonant electrons, which are scattered into the loss-cone by the electron-cyclotron resonance process with whistler mode waves. The location of the plasmapause is known to vary significantly in response to magnetospheric conditions (e.g., Maynard and Chen, 1975). Observations using the “Balloon Cluster” in the 2nd-PPB experiment will enable us to investigate the spatial and temporal variations in wave-particle interactions around the plasmapause region.

The trajectory of the PPB runs across the regions of the outer belt and the slot region of the radiation belt around the southward edge of the South Atlantic Anomaly (SAA), as shown in Fig. 10a (Tsuruuchi, 1998). Energetic particle flux in the outer and inner belts and the slot region increases significantly during strong storm periods, as shown in Fig. 10b. The “Balloon Cluster” will enable the spatial distribution of energetic electron precipitation from the radiation belts and its temporal evolution during the development of a storm to be observed.

In the 2nd-PPB experiment, well-coordinated simultaneous observations with the SuperDARN HF-radar network, Antarctic stations, low-altitude satellites, NOAA, DMSP, AKEBONO, etc., and magnetospheric observations by CLUSTER-2, GEOTAIL, and geosynchronous satellites are also expected.

The EMW has three observational modes: 1) wave form observations in a frequency range of 0.2–4.0 Hz with 10-Hz sampling, 2) intensity observations for 4 specific frequency channels (300 Hz, 600 Hz, 1.2 kHz, and 2.4 kHz) with a 0.5-s resolution, and 3) frequency sweep observations at 4 specific frequencies (5 kHz, 10 kHz, 20 kHz, and 36 kHz) with a 0.5-s resolution. The EFD observes two horizontal and one vertical electric field components using a double spherical probe technique with resolutions of 0.2 mV/m (horizontal)
Fig. 10. Distribution of energetic electron (>300 keV) flux observed by the NOAA satellite during a quiet period (a) and a disturbed period (b) (after Tsuruuchi, 1998). The two horizontal black lines show the latitudinal range of the PPB trajectory.
and 0.8 mV/m (vertical) and a one-second sampling. The atmospheric conductivity and current density are also measured every 10 min. The MGF installs a 3-axis fluxgate magnetometer, 2-axis clinometer, and a set of sunsensors, and measures the geomagnetic field vector with a 0.25 nT resolution and one-second sampling. The AXI contains two kinds of auroral X-ray detectors: an imager that observes the two-dimensional auroral X-ray distribution with a 4×4 resolution and a FOV of 140°, and an omni-directional counter for higher energy X-rays (>200 keV) with a sampling rate of 1 s. The TEC measures the total electron content in the ionosphere along the line-of-sight of a GPS satellite with a GPS dual-frequency receiver and a one-minute sampling period. From the circumpolar trajectory of the PPB, at least three GPS satellites can be expected to be always observed above an elevation of 10°.

5. Summary

An overview of the 2nd-PPB experiment was described in this paper. The 2nd-PPB campaign will be performed during late December 2002 to January 2003 at Syowa Station in Antarctica by JARE-44. In the 2nd-PPB experiment, one B100 balloon will be launched to observe primary cosmic-ray electrons within an energy range of 10 GeV to 1 TeV, and three B50 balloons, known as the “Balloon Cluster”, will be launched to observe the spatial and temporal variations of various phenomena in various magnetospheric and ionospheric regions and boundaries. The scientific purposes and instrument designs that will be used for these observations were also introduced. All of these observations utilize the unique advantages of the PPB experiment, namely its long duration and circumpolar trajectory. The new satellite data acquisition system that will be used in the 2nd-PPB enables more sophisticated observations with much higher temporal resolutions to be performed, compared with the 1st-PPB experiment. A well-coordinated international collaboration with ground-based and satellite-based observations is anticipated.

Acknowledgments

The PPB project has been planned and conducted by the PPB Working Group as a collaborative research program of the NIPR, ISAS, and nationwide scientists, whose members are gratefully acknowledged. The 2nd-PPB project is based on the fruitful results of the previous 1st-PPB project. We would like to express our special thanks to all the members participating in the 1st-PPB project, including the members of JARE-32 and JARE-34, who performed the launching operations at Syowa Station.

The editor thanks the referee for his help in evaluating this paper.

References


Circumpolar observation of bremsstrahlung X-rays by a Polar Patrol Balloon over Antarctica. J Geomagn Geoelectr., 47, 253–266


(Received February 1, 2002; Revised manuscript accepted April 23, 2002)
Calibration of CRL all-sky imagers using an integrating sphere

Masa-yuki Yamamoto\textsuperscript{1}, Minoru Kubota\textsuperscript{1}, Shu Takeshita\textsuperscript{2}, Mamoru Ishii\textsuperscript{1}, Yasuhiro Murayama\textsuperscript{1} and Masaki Ejiri\textsuperscript{2}

\textsuperscript{1}Communications Research Laboratory, 4-2-1 Nukui-kitamachi, Koganei-shi, Tokyo 184-8795
\textsuperscript{2}National Institute of Polar Research, Kaga 1-chome, Itabashi-ku, Tokyo 173-8515

Abstract: As part of an international collaboration with the Geophysical Institute of the University of Alaska, we have developed two all-sky imagers (CRL-ASIs). A sensitivity calibration of the CRL-ASIs was performed using an integrating sphere belonging to the National Institute of Polar Research (NIPR). The two-dimensional sensitivities of the CRL-ASIs produced symmetrical distributions. Using this sensitivity data, we converted arioglow/aurora images into two-dimensional distributions of absolute intensity. The sensitivity of the CRL-ASIs was measured for 13 wavelengths between 427.8 nm and 866.5 nm, and the relationship between the sensitivity and the wavelength was investigated for both imagers. The peak sensitivity occurred at about 550 nm.

1. Introduction

Multiple wavelength observations of aurora/airglow emissions obtained with an all-sky imager, consisting of a bare cooled CCD camera and interference filters with a narrow band passage, have been frequently used as a practical method of producing sensitive images during the past decade (e.g., Taylor et al., 1995; Ejiri et al., 1997; Kubota et al., 2000).

The Joint Program of the Middle and Upper Atmosphere Observation (Alaska Project) was begun in 1995 by the Communications Research Laboratory (CRL) and the Geophysical Institute, University of Alaska Fairbanks, USA. As part of this project, we developed two all-sky imagers (CRL-ASIs). Continuous observation of aurora/airglow emissions by the CRL-ASIs began in October 2000, at the Poker Flat Research Range, Alaska (Latitude: 65.12°N, Longitude: 147.43°E, and Geomagnetic latitude: 65.60°N). Since the two CRL-ASIs are located at the same observation site, we can simultaneously observe auroral images of two independent wavelengths with no time separation. Using this advantage, we plan to estimate the two-dimensional energy deposition rate of precipitating auroral particles. The energy flux and the characteristic energy of the precipitating particles can be estimated using the absolute intensity of 427.8 nm emissions and the ratio of auroral emission intensities of two different wavelengths, respectively (Rees and Luckey, 1974; Rees et al., 1988; Ono, 1993). For this estimation, the absolute intensity distribution of the auroral emissions must be obtained. Therefore, we first
calibrated the sensitivities of the CRL-ASIs using an integrating sphere.

Absolute two-dimensional calibration of an all-sky imaging system can be obtained using an integrating sphere, which has a uniform light source that fills a $2\pi$-steradian field of view. The uniformity of the radiance in the 1.9-m integrating sphere, which belongs to the National Institute of Polar Research (NIPR), is better than $\pm5\%$ for a $2\pi$ field of view (Okano et al., 1998). Shiokawa et al. (2000) reported calibration results for three imagers developed by the Solar-Terrestrial Environment Laboratory of Nagoya University. In their report, they mentioned that filters with highly flattened image-quality surfaces may cause serious Newton's Ring patterns on the final images. In our study, we measured the two-dimensional sensitivities of the CRL-ASIs at 13 wavelengths and investigated the relationship between sensitivity and wavelength.

2. Instrumentation and calibration method

Each CRL-ASI consists of two sections: an optical section, with a wide-field lens (Nikon, 6 mm, F1.4, FOV = 180°), interference filters, and tele-centric lens system; and a detector section, with a bare cooled CCD camera (Pixel Vision, 512×512 pixels, back-illuminated). Both CRL-ASIs have the same configuration. Five filters can be installed in the filter turret of each CRL-ASI. Figure 1 shows one of the CRL-ASIs being calibrated with the integrating sphere (Optronic Laboratories Inc., OL462-80A). The integrating sphere has an internal diameter of 1.9 m. Its specially coated inner surface was carefully manufactured to achieve uniform reflection. The lamp luminosity of the integrating sphere was constantly monitored during the experiment. The transmission profiles

![Fig. 1. CRL-ASI being calibrated with the 1.9-m integrating sphere from NIPR.](image-url)
of the interference filters were measured using the auto spectro-photometer at NIPR (Hitachi, Ltd., U-3300).

During the aurora/airglow emission measurements using the CRL-ASI, the observed value \( C_{\text{AG}}[\text{Count}] \) was given by the following equation (e.g., Shiokawa et al., 2000; Kubota et al., 2001):

\[
\langle C_{\text{AG}} \rangle = a \cdot E_{\text{AG}} \int_0^\infty \langle T_{\text{opt}}(\lambda) \rangle \cdot \langle S_{\text{CCD}}(\lambda) \rangle \cdot \{(\langle I_{\text{AG}}(\lambda) \rangle + \langle I_{\text{BG}}(\lambda) \rangle) \cdot T_{\text{AG}}(\lambda) \} d\lambda,
\]

where \( a[\text{Count}/\text{Rayleigh/s}] \) is a constant, \( E_{\text{AG}}[\text{s}] \) is the exposure time, \( T_{\text{opt}}[\%] \) is the transmission of the optical section without the interference filter, \( S_{\text{CCD}}[\%] \) is the sensitivity of the CCD camera (determined by multiplying the quantum efficiency of the CCD [electrons/photon] and the conversion efficiency [Count/electron]), \( I_{\text{AG}}[\text{Rayleigh}] \) is the absolute intensity of the aurora/airglow emission, \( I_{\text{BG}}[\text{Rayleigh/mm}] \) is the absolute intensity of the background continuum, and \( T_{\text{AG}}[\%] \) is the transmission of the interference filter. The values enclosed by \( \langle \rangle \) indicate the two-dimensional variables at the focal plane. In general, \( T_{\text{AG}}[\lambda] \) becomes a two-dimensional variable because the filter transmission peak is shifted towards shorter wavelength at the edge of an image as a result of oblique incoming light. In eq. (1), however, \( T_{\text{AG}}[\lambda] \) is described as a fixed value at the center of the image. This point will be discussed in greater detail in the Discussion section.

From eq. (1), the following equation can be derived:

\[
\langle C_{\text{AG}} \rangle = a \cdot E_{\text{AG}} \int_0^\infty \langle T_{\text{opt}}(\lambda) \rangle \cdot \langle S_{\text{CCD}}(\lambda) \rangle \cdot \{(\langle I_{\text{AG}}(\lambda) \rangle + \langle I_{\text{BG}}(\lambda) \rangle) \cdot T_{\text{AG}}(\lambda) \} d\lambda = a \cdot \langle T_{\text{opt}}(\lambda_{\text{AG}}) \rangle \cdot \langle S_{\text{CCD}}(\lambda_{\text{AG}}) \rangle \cdot E_{\text{AG}} \langle I_{\text{AG}}(\lambda_{\text{AG}}) \rangle \cdot T_{\text{AG}}(\lambda_{\text{AG}}) \\
+ \langle I_{\text{BG}}(\lambda_{\text{AG}}) \rangle \cdot \int_0^\infty T_{\text{AG}}(\lambda) d\lambda,
\]

where the variation of \( T_{\text{opt}} \) and \( S_{\text{CCD}} \) with different wavelengths within the filter pass band is assumed to be negligible.

To obtain the absolute intensity \( I_{\text{AG}} \), simultaneous measurement of the background intensity \( I_{\text{BG}} \) is needed to enable the background to be subtracted. \( T_{\text{AG}}(\lambda_{\text{AG}}) \), \( \int_0^\infty T_{\text{AG}}(\lambda) d\lambda \), and \( a \cdot \langle T_{\text{opt}}(\lambda_{\text{AG}}) \rangle \cdot \langle S_{\text{CCD}}(\lambda_{\text{AG}}) \rangle \) are required for the calibration:

\( T_{\text{AG}}(\lambda_{\text{AG}}) \) and \( \int_0^\infty T_{\text{AG}}(\lambda) d\lambda \) can be obtained by measuring the transmission profiles of the interference filters, while \( a \cdot \langle T_{\text{opt}}(\lambda_{\text{AG}}) \rangle \cdot \langle S_{\text{CCD}}(\lambda_{\text{AG}}) \rangle \) [Count/Rayleigh/s] can be derived using the integrating sphere in the manner described below.

The integrating-sphere image \( \langle C_{\text{integ}} \rangle[\text{Count}] \) obtained by the CRL-ASIs can be described as

\[
\langle C_{\text{integ}} \rangle = a \cdot \langle T_{\text{opt}}(\lambda_{\text{AG}}) \rangle \cdot \langle S_{\text{CCD}}(\lambda_{\text{AG}}) \rangle \cdot E_{\text{integ}} \int_0^\infty \langle I_{\text{integ}}(\lambda) \rangle T_{\text{AG}}(\lambda) d\lambda,
\]

where \( E_{\text{integ}}[\text{s}] \) is the exposure time and \( \langle I_{\text{integ}}(\lambda) \rangle[\text{Rayleigh/mm}] \) is the luminosity of the integrating sphere. Since \( \langle I_{\text{integ}}(\lambda) \rangle \) has been carefully determined beforehand, \( a \cdot \langle T_{\text{opt}}(\lambda_{\text{AG}}) \rangle \cdot \langle S_{\text{CCD}}(\lambda_{\text{AG}}) \rangle \) can be calculated using eq. (3).
3. Results

Figure 2a shows the two-dimensional distribution of CRL-ASI sensitivity at a wavelength of 557.7 nm. The sensitivity drops symmetrically from the center toward the edge of the image due to vignetting. Figure 2b shows two cross sections of the sensitivity along the horizontal and vertical lines that cross at the center of Fig. 2a. Both cross sections are nearly identical, suggesting that the alignment of the CRL-ASI optical system was well adjusted and that the CRL-ASI can be used for all-sky imaging with excellent results. The sensitivity values along the vertical cross line are slightly higher than the

![Diagram](image)

Fig. 2. (a) Two-dimensional distribution of the CRL-ASI sensitivity (at 557.7 nm). The circle seen in the image represents the joint where the two hemispheres that form the integrating sphere meet. (b) Cross sections of the sensitivity data shown in (a). The solid and broken lines represent the cross sections of the sensitivity data along the horizontal and vertical lines that cross at the center of the image in (a), respectively. The joint where the hemispheres meet appears as a dip at about ±170 pixels. The small ‘bump’ seen at -250 pixels in the vertical cross section represents a localized nonuniformity of the inner surface of the integrating sphere.
values along the horizontal one between \(-240\) and \(-20\) pixels, as seen in Fig. 2b, and a “bump” occurs on the edge of the image at \(-250\) pixels. The higher value seen at the bottom edge of the image seems to be caused by a localized nonuniformity of the inner surface radiance of the integrating sphere. However, the difference between the two cross sections is less than 7% if the “bump” is excluded. Since the integrating sphere consists of two hemispheres, the mating joint of the two hemispheres is seen as small dips at about
\begin{table}
\centering
\begin{tabular}{cccc}
\hline
Center wavelength (nm) & FWHM (nm) & \( T_{\text{AC}} (\lambda_{\text{AC}}) \) & Objective emission \\
\hline
427.9 & 2.8 & 41.0 & \( \text{N}_2^+ \text{ ING}(0.1) \) \\
481.25 & 1.7 & 45.1 & \( \text{Background for H} \beta \) \\
485.75 & 1.7 & 46.8 & \( \text{H} \beta \) \\
557.6 & 2.0 & 52.8 & \( \text{OI} \) \\
572.2 & 2.0 & 61.6 & \( \text{background} \) \\
589.05 & 2.3 & 59.7 & \( \text{Na} \) \\
630.0 & 2.2 & 68.6 & \( \text{OI} \) \\
670.65 & 2.9 & 53.9 & \( \text{N}_2 \) \\
732.05 & 2.1 & 54.8 & \( \text{O}^+ \) \\
770.0 & 3.0 & 59.6 & \( \text{K} \) \\
777.6 & 2.8 & 62.5 & \( \text{OI} \) \\
844.7 & 2.6 & 54.5 & \( \text{OI} \) \\
866.5 & 9.4 & 82.8 & \( \text{O}_2(0.1) \) \\
\hline
\end{tabular}
\caption{Listing of filter specifications. All filters were made by Andover Co. with options of 'type 3', 'block to FIR', and 'non image quality'.}
\end{table}

±170 pixels in the cross sections, which is also seen as a circle on the image. However, this does not seriously affect the results of the observation because the dips can be removed using a smoothing process over a region of 10×10 pixels.

Figure 3a shows the variations in sensitivity between 427.8 nm and 866.5 nm for CRL-ASI system-1, and Fig. 3b shows the same data for CRL-ASI system-2. Each sensitivity value indicates the averaged value for 10×10 pixels at the center of the image (zenith direction). The maximum sensitivity value was observed at about 557.7 nm. A list of the 13 interference filters used in the present investigation is shown in Table 1. The sensitivity of system-2 was somewhat higher than that of system-1, though the two CRL-ASIs were manufactured using the same design. However, the sensitivity distributions of the two CRL-ASIs indicate that their features are similar.

4. Discussions

In eq. (1), \( T_{\text{AC}} (\lambda) \) can be expressed as a fixed value at the image center if \( \lambda_{\text{AC}} \) is situated on the shorter wavelength side of the flat portion of the filter transmission. Most of the Andover Type 3 filters listed in Table 1 satisfy this condition. For the OI (557.7 nm) and H\( \beta \) (485.9 nm) emissions, however, the corresponding filters fail to meet this condition. Nevertheless, the absolute intensity of these emissions can be correctly calibrated at almost any portion of the image except for at the edge.

As is well known, filter transmission shifts with temperature. This may cause the absolute calibration factor to vary when the imagers are used at a different temperature from that in the integrating sphere. In the present case, however, the shift was negligible because the imagers were placed in an air-conditioned room at room temperature in the
observatory building at Poker Flat.

By comparing Figs. 3a and 3b, the sensitivity of CRL-ASI system-2 was found to be about 1.4 times higher than that of CRL-ASI system-1 at each wavelength. This difference seems to be caused by a difference in the gain setting of the conversion efficiency from the number of electrons to the output count rate. In general, the quality of an image can be evaluated by the S/N ratio. The difference in sensitivity between the two CRL-ASIs is not considered to affect the S/N ratio of the image for the following reason. Since the noise of the imager system consists of thermal noise, readout noise, and statistical random noise, the S/N ratio of the image data is determined before the conversion process. Namely, the signal and noise are amplified together at the same rate in the conversion process, so the S/N ratio remains constant and is independent of the conversion process' gain setting.

Using the sensitivity data, the optimum exposure time for the aurora/airglow emissions can be determined for each wavelength. The exposure time is determined by considering the time resolution and S/N ratio. For the present CRL-ASI set-up, the dark noise is about 0.5 [Counts/pixel/s] and the readout noise is about 12 [Counts, rms]. The statistical random noise associated with the quantum effect in the CCD is equal to the square root of the signal count rate. In the case of a typical green aurora with an intensity of 3 [k Rayleigh] when observed at a wavelength of 557.7 nm, a high-quality image with an S/N ratio of 35 can be obtained using a 3-second exposure. For an aurora with an intensity of 500 [Rayleigh] at a wavelength of 844.6 nm, an image with an S/N ratio of 8 can be obtained using a 45-second exposure.

5. Summary

The sensitivity of two CRL-ASIs was calibrated using an integrating sphere. The two-dimensional sensitivities of the CRL-ASIs exhibited symmetrical distributions. Calibration data were obtained for 13 wavelengths between 427.8 nm and 866.5 nm, and the relationship between sensitivity and wavelength was analyzed for each CRL-ASI. The absolute intensity of the aurora/airglow emission can be calculated from the CRL-ASI observations using the results of the accurate sensitivity calibration. The two-dimensional absolute intensity of the aurora emission is useful for estimating the energy flux and characteristic energy of the precipitating auroral particles.

Using the filters listed in Table 1, observations of aurora/airglow emissions at altitudes ranging from the mesopause region to the F-region can be used to quantitatively investigate atmospheric and ionospheric dynamics in this region.

Acknowledgments

The authors would like to express their sincere thanks to an anonymous referee for his/her kind comments and suggestions.

The editor thanks Dr. S. Okano for his help in evaluating this paper.
References


Taylor, M.J., Bishop, M.B. and Taylor, V. (1995): All-sky measurements of short period waves imaged in the OI(557.7 nm), Na(589.2 nm) and near infrared OH and O2(0,1) nightglow emissions during the ALOHA-93 campaign Geophys. Res. Lett., 22, 2833–2836.

(Received December 20, 2001; Revised manuscript accepted April 4, 2002)
A new method for monitoring and removing SuperDARN radar DC offsets

Akira Sessai Yukimatu, Masaki Tsutsumi, Hisao Yamagishi and Natsuo Sato

National Institute of Polar Research, Kaga 1-chome, Itabashi-ku, Tokyo 173-8515

Abstract DC offsets in quadrature outputs (I/Q signals) pertinent to the SuperDARN HF radars can cause problems in determining and analysing auto-correlation functions, cross-correlation functions and raw time series data if they are not negligible. We have developed a new code to monitor and remove DC offsets. To obtain correct DC offset levels and noise levels, the offsets can be best removed using I/Q signals during the receive-only period, i.e., during each clear frequency search stage. We have implemented the new code in the current radar operating system, installed it at the SENSU Syowa radars, and have been obtaining and removing DC offset values continuously since October 2001.

1. Introduction

The Super Dual Auroral Radar Network (SuperDARN) (Greenwald et al., 1995), a network of high-frequency (HF) radars (Greenwald et al., 1985), is a powerful tool used mainly for measuring plasma motions in the ionospheric E- and F-regions in vast fields-of-view that cover most of the polar regions of both hemispheres.

The radars adopt a multi-pulse sounding technique (Farley, 1972) that is specially designed to yield auto-correlation functions (ACFs) of received signals (receiver (Rx) output) using a radar operating system (software) ("ROS") and to obtain physical parameters, such as echo power, line-of-sight (LOS) Doppler velocity and Doppler spectral width, using a real-time "fitacf" algorithm in the ROS (Baker et al., 1988). Cross-correlation functions (XCFs) between the signals received by the main array and those by an interferometer array can also be determined to measure the angles of arrival (elevation angles of received signals) when the interferometer array is available.

DC offsets in the quadrature outputs of the receiver (i.e., I (in-phase) and Q (quadrature-phase) signals) must be properly detected and removed because the existence of non-zero I/Q DC offsets causes systematic errors in the estimation of ACFs, XCFs and the physical parameters in the fitacf algorithm (Yamagishi et al., 1999) as well as in the estimation of echo power ($I^2 + Q^2$) and phase ($\arctan(I/Q)$) using a new raw time series analysis method (Yukimatu and Tsutsumi, 2002). However, the current hardware and software of the SuperDARN radar system does not contain a DC offset removal mechanism. Therefore, we have developed a code to monitor and remove I/Q DC offsets in the ROS (software). We describe the details of this code in the present report.

181
2. I/Q DC offsets and removal methods

2.1. I/Q DC offsets

We will denote the following variables:

- $t$: time,
- $I_i(t)$: I signal in the main receiver,
- $Q_i(t)$: Q signal in the main receiver,
- $I_j(t)$: I signal in the interferometer receiver,
- $Q_j(t)$: Q signal in the interferometer receiver,
- $I_{j0}$: DC offset in $I_j$ signal ($j=1, 2$),
- $Q_{j0}$: DC offset in $Q_j$ signal ($j=1, 2$),
- $Z_j(t)$: a complex signal constructed by $I_j$ and $Q_j$ ($j=1, 2$),
- $\langle A \rangle$: ensemble average of a variable, $A$,
- $A^*$: a complex conjugate of $A$,

and define the following relations:

$$I_{j0} = I_j(t) - I_{j0},$$  \hspace{1cm} (1)
$$Q_{j0} = Q_j(t) - Q_{j0},$$  \hspace{1cm} (2)
$$Z_j(t) = I_j(t) + iQ_j(t),$$  \hspace{1cm} (3)
$$Z_{j0} = I_{j0} + iQ_{j0}.$$  \hspace{1cm} (4)

where $i=\sqrt{-1}$, and $j=1$ or 2.

ACF and XCF can then be defined as

$$\text{ACF}(\tau) = \langle Z_j(t) \cdot Z_j^*(t+\tau) \rangle,$$  \hspace{1cm} (5)
$$\text{XCF}(\tau) = \langle Z_j(t) \cdot Z_j(t+\tau) \rangle,$$  \hspace{1cm} (6)

where $\tau$ is time lag.

In the fitacf algorithm, the “apparent” ACF and XCF (denoted as “ACF$_a$” and “XCF$_a$” in cases where the DC offsets are not taken into account and removed) are obtained as follows:

$$\text{ACF}_a(\tau) = \langle Z_j(t) \cdot Z_j^*(t+\tau) \rangle,$$  \hspace{1cm} (7)
$$\text{XCF}_a(\tau) = \langle Z_j(t) \cdot Z_j(t+\tau) \rangle.$$  \hspace{1cm} (8)

The relationships between the apparent and the true (correct) values for ACF and XCF (denoted as “ACF$_c$” and “XCF$_c$”) can be described as follows:

$$\text{ACF}_c = \langle Z_{1a}(t) \cdot Z_{1a}^*(t+\tau) \rangle$$
$$= \langle (Z_{1a}(t)+Z_{10}) \cdot (Z_{1a}^*(t+\tau)+Z_{10}^*) \rangle$$
$$= \text{ACF}_c + Z_{10} \cdot \langle Z_{1a}(t) \rangle + Z_{10} \cdot \langle Z_{1a}^*(t+\tau) \rangle + |Z_{10}|^2,$$  \hspace{1cm} (9)

$$\text{XCF}_c = \langle Z_{1a}(t) \cdot Z_{2a}^*(t+\tau) \rangle$$
$$= \langle (Z_{1a}(t)+Z_{10}) \cdot (Z_{2a}^*(t+\tau)+Z_{20}) \rangle$$
$$= \text{XCF}_c + Z_{20} \cdot \langle Z_{1a}(t) \rangle + Z_{10} \cdot \langle Z_{2a}^*(t+\tau) \rangle + Z_{10} \cdot Z_{20},$$  \hspace{1cm} (10)

where
SuperDARN radar DC offsets removal method

\[ Z_{\mu}(t) = I_{\mu}(t) + iQ_{\mu}(t), \quad (j = 1, 2) \]  
(11)

\[ Z_{\rho}(t) = I_{\rho}(t) + iQ_{\rho}(t) = Z_{\mu}(t) - Z_{\rho}, \quad (j = 1, 2) \]  
(12)

\[ ACF_c = \langle Z_{\mu}(t) \cdot Z_{\mu}(t + \tau) \rangle, \]  
(13)

\[ XCF_c = \langle Z_{\rho}(t) \cdot Z_{\rho}(t + \tau) \rangle. \]  
(14)

On the other hand, in the raw time series analysis, a simple (and rough) estimation of the apparent echo power and phase ("PWR_a" and "PH_a") can be obtained using the equations

\[ PWR_a = I_j^2 + Q_j^2, \quad (j = 1 \text{ or } 2) \]  
(15)

\[ PH_a = \arctan(I_j/Q_j), \quad (j = 1 \text{ or } 2) \]  
(16)

However, the correct echo power ("PWR_c") and phase ("PH_c") should be calculated using the equations

\[ PWR_c = (I_j - I_0)^2 + (Q_j - Q_0)^2, \quad (j = 1 \text{ or } 2) \]  
(17)

\[ PH_c = \arctan((I_j - I_0)/(Q_j - Q_0)) + 2n\pi. \quad (j = 1 \text{ or } 2, n: \text{ integer}) \]  
(18)

In both methods, if \(|Z_0|\) (I_0 or Q_0) (j = 1 or 2) is not negligible compared with I_{\mu}, Q_{\mu}, |ACF_c| or |XCF_c|, the DC offsets cause systematic errors in the estimation of physical parameters (noise level, echo power, LOS velocity, spectral width and elevation angle, etc.) and can also result in a lower echo detectability, especially for smaller echo powers. We should also note that the elevation angle measurement is especially sensitive to errors caused by the I/Q DC offsets and the frequency dependency of the phase differences between two receivers.

The "fitacf" algorithm can also be described from another point of view. Before the algorithm fits the ACFs to deduce the physical parameters, it determines the "noise ACF (ACF_N)" and subtracts the ACF_N from all the ACFs for all the range gates (Baker et al., 1988). ACF_N is the average value of the ACFs at (a maximum of) 10 range gates where the lag-0 power (i.e., ACF(\tau = 0)) is the smallest among all the observed range gates in a single beam integration time. Therefore, ACF_N is expected to include the DC offset effect, and the process of subtracting the ACF_N seems to remove all the DC offset effects on the ACFs at all the range gates. However, this conclusion is not true because the process can remove the last term in eq. (9), i.e., \(|Z_0|^2\), but the second and third terms in eq. (9), i.e., the effect of the cross terms, differ from range to range; thus, the process cannot remove the effect properly for all ranges. Consequently, the fitacf algorithm suffers from the I/Q DC offset effect unless the offsets are properly removed before constructing the ACFs and XCFs.

2.2. Removal methods

One proposal for removing the DC offsets is to subtract a real part of the ACF averaged over the time lag from each ACF (Greenwald, 2001). However, this method may also subtract real echoes with an LOS velocity close to zero. Such echoes could be produced by, for example, ground backscatters or low ionospheric plasma convective flows. Moreover, if the ensemble average is calculated over large samples, i.e., if the ACFs are averaged over enough long periods, eq. (9) ideally yields
Thus, this method seems to work. In real observations, however, the ACFs are obtained using a limited number of samples. As a result, the cross-term effect (i.e., the second and the third terms) in eq. (9) does not vanish, so eq. (19) is not always correct and the imaginary part of the ACFs (ACF_s in eq. (9)) can also suffer some DC offset effects as a result. Furthermore, the average value of ACF_c is not always guaranteed to be zero. Therefore, this method can not completely and properly remove all of the DC offset effects from the ACFs and is not suitable for studies using ground backscatters, such as research on gravity waves (e.g., Samson et al., 1989).

Determining the I/Q DC offsets using an averaged value of I/Q signals during each pulse sequence or during over the beam integration time period is another possible method for removing the I/Q DC offsets more properly:

\[ I_{0j} = \langle I_j \rangle_{TX}, \quad Q_{0j} = \langle Q_j \rangle_{TX}, \quad (j = 1, 2) \]  

where \( \langle \rangle_{TX} \) is an average during a pulse sequence or a beam integration time.

The previous method did not separately determine the I/Q DC offset values. The present method, however, does attempt to precisely determine and remove the I/Q DC offset values from all the I/Q samples before constructing the ACFs and XCFs. Therefore, this method provides more accurate and reliable results.

However, the DC offsets can be difficult to distinguish from ground backscatters, especially in cases where the ground backscatter consists of an echo with a maximum power among all the range gates, as often happens. Moreover, from the “central limit theorem” in statistics, the standard deviation of the I and Q signals \( (I_j, Q_j \; (j = 1, 2)) \) over the beam integration time \( (\sigma_{I0j}, \sigma_{Q0j} \; (j = 1, 2)) \) is as follows:

\[ \sigma_{I0j} = \frac{\sigma_{Ij}}{\sqrt{n}} \approx \frac{\sigma_{Ij}}{2 \times 10^4} \approx \frac{\sigma_{Ij}}{140}, \quad (j = 1, 2) \]  

where

- \( \sigma_{Ij} \): standard deviation of I_j signals during a beam integration time \( (|I_j| \leq 2^{11}, 0 \leq \sigma_{Ij} \leq 2^{11} \approx 2 \times 10^4 \; \text{(in case of signed 12-bit A/D)}) \)

- \( \sigma_{I0j} \): standard deviation of \( \langle I_j \rangle_{TX} \) (averaged values of I_j during each beam integration time)

- \( n \): number of averaging \( (\text{typically} \; 300 \times 70 \approx 2 \times 10^4) \)

\( n \) is dependent on the sounding modes (e.g., pulse pattern, integration time).

The values of \( \sigma_{I0j} \) and \( \sigma_{Q0j} \) can be large, depending on the amount and the intensity of the real ionospheric and ground backscatters, and will also vary with time (over both short and long time scales). In other words, the precision of the DC offsets may be unstable. Therefore, this method is not the most reliable solution for our purposes.

2.3. New removal method

Determining the I/Q DC offsets using an averaged value of I/Q signals during a non-transmission (receive-only, non-TX or Rx-only) period is a much better solution for
determining the correct DC offsets than the previously described methods because the signals are not affected by any ionospheric or ground backscatters from transmitted pulses, whose values can have large, undesirable standard deviations, although some unexpected and temporal interference noises may still be present. In real operation mode in ROS, there is a non-Tx period called the “clear frequency search (fclr)” stage. During this period, without any pulse transmissions, a limited number of I and Q signals are sampled for a specified 300-kHz frequency band, typically in 5-kHz steps, to determine the quietest frequency and ensure a minimal received noise level. If we could determine the average (expected) values of the I and Q signals at the quietest (chosen) frequency during the fclr stage, these values would represent the expected I and Q DC offsets. This relationship can be expressed as follows:

\[ I_{\rho} = \langle I_{\rho} \rangle_{\rho}, \quad Q_{\rho} = \langle Q_{\rho} \rangle_{\rho}, \quad (j = 1, 2) \]  

(22)

where \( \langle \rangle_{\rho} \) is average during the non-Tx (Rx-only) “fclr” stage.

The relationship between the I/Q signals and the noise level is as follows:

\[ \sigma_{I_j} = \sqrt{\langle (I_j - I_{\rho})^2 \rangle_{\rho}} = \sqrt{\langle I_{\rho}^2 \rangle_{\rho} - I_{\rho}^2}, \quad (j = 1, 2) \]  
\[ \sigma_{Q_j} = \sqrt{\langle (Q_j - Q_{\rho})^2 \rangle_{\rho}} = \sqrt{\langle Q_{\rho}^2 \rangle_{\rho} - Q_{\rho}^2}, \quad (j = 1, 2) \]  
\[ N_c = \sigma_{I_j}^2 + \sigma_{Q_j}^2 = N_{\text{app}} - (I_{\rho}^2 + Q_{\rho}^2) = N_{\text{app}} - |Z_{\rho}|^2, \]  

(23)

(24)

(25)

and

\[ \sigma_{I} = \sigma_{Q}, \quad (j = 1, 2) \]  

(26)

(this is valid statistically, i.e., if the number of averaged items is sufficiently large and the gains of the I and Q channels are exactly the same)

\[ \sigma_{I} = \sigma_{Q} = \sqrt{N_c/2} \]  

(27)

where

\( \sigma_{I_j} \): standard deviation of \( I_j \) signals during the fclr stage \( (j = 1, 2) \),

\( \sigma_{Q_j} \): standard deviation of \( Q_j \) signals during the fclr stage \( (j = 1, 2) \),

\( N_c \): true (correct) Noise value,

\( N_{\text{app}} \): apparent noise level \( (= \langle I^2 + Q^2 \rangle_{\rho} = \langle I_{\rho}^2 \rangle_{\rho} + \langle Q_{\rho}^2 \rangle_{\rho}) \).

However, the number of samples during a typical fclr stage is typically only 200. So, \( I_{\rho} \) and \( Q_{\rho} \) in eq. (22) scatter statistically (even if the DC offset level is temporally constant), i.e., they must be treated as “random variables”.

According to the central limit theorem, they will produce a normal (Gaussian) distribution such that

\[ \sigma_{\langle I \rangle_{\rho}} = \sigma_{I}/\sqrt{\text{num}} = \sigma_{I}/\sqrt{200} \]  

(28)

\[ \sigma_{\langle Q \rangle_{\rho}} = \sigma_{Q}/\sqrt{\text{num}} = \sigma_{Q}/\sqrt{200} \]  

(29)

where

\( \sigma_{\langle I \rangle_{\rho}} \): standard deviation of \( \langle I \rangle_{\rho} \),

\( \sigma_{\langle Q \rangle_{\rho}} \): standard deviation of \( \langle Q \rangle_{\rho} \),

\( \text{num} \): number of averaged items,
\[ \sigma_{d\alpha} = \sigma_{q\alpha} = \sqrt{N_c/2}/\sqrt{200} = \sqrt{N_c/20}. \]  

(30)

If \( N_c \approx 2000 \) (a typical value for many SuperDARN radars using a 12-bit A/D card),

\[ \sigma_{d\alpha} = \sigma_{q\alpha} = \sqrt{N_c/2} \approx 2.2. \]  

(31)

This means that if we determine the I/Q DC offsets from the average I/Q values during the fcfr stages, the values can vary with time with a standard deviation of \( \sim 2 \times \sqrt{N_c/2000} \). This value is not always negligible and is rather dependent on noise levels. The above statistical variation also introduces a new undesired “noise” into the real I/Q signals, even when the real DC offsets are well tuned and exactly equal to zero.

Considering the above, we can conclude that a realistic solution for determining the I/Q DC offsets is to calculate the statistical expectation (average) values of a sufficiently large number of 200-sample I/Q averages during each fcfr stage; that is to say,

\[ I_\beta = \langle I_j \rangle_{fcfr}, \quad Q_\beta = \langle Q_j \rangle_{fcfr}, \quad (j = 1, 2) \]  

(32)

(ensemble averages of \( \langle I_j \rangle_{fcfr}, \langle Q_j \rangle_{fcfr} \) over a sufficiently large number of fcfr stages.)

Fortunately, the I/Q DC offset values are normally expected to vary slowly with time because of, for example, the slow diurnal temperature variation in the radar observation hut where the receivers are located. Therefore, we can replace the ensemble average by a temporal average for a sufficiently long period and can adopt the temporal average of the 200-sample I/Q averages as the true (or expected) I/Q DC offset.

3. Development of new DC offset monitoring and removal code

To realise the new I/Q DC offset monitoring and removal method described in the previous section, the following points should also be considered to create a better algorithm for real-time programming in ROS.

1) the number of 200-average I,Q data points kept in the computer memory should be reduced for faster calculation and lower CPU load (though average values must be calculated using a sufficiently large number (e.g., 100) of data points),

2) the older 200-sample average values should be rather less contributive,

3) apparent large temporal jumps in DC offset values \( \langle I \rangle \) and \( \langle Q \rangle \), due to unexpected, abrupt large interference noise during the fcfr stage, should be avoided, and

4) the CPU load should be reduced as much as possible.

In view of the above, we introduced “oblivion coefficients (obliv)” to improve the algorithm for determining the real DC offset values so that only the latest (most recent) DC offset value and the latest standard deviation of the averaged I,Q signals for each I,Q channel for every frequency band need to be kept in the computer memory (RAM). (This method can be validly applied here because the DC offsets are expected to change very slowly and smoothly with time). The DC offset values are obtained as follows:

\[ I_{j,\text{new}} = \text{obliv} \cdot I_{j,\text{last}} + (1-\text{obliv}) \cdot \langle I_j \rangle_{\text{new}} \quad (j = 1, 2) \]  

(33)

where

\[ 0 < \text{obliv} < 1. \]  

(34)
Note that

\begin{align}
\text{oblv} &= \text{oblv}_1 = \text{e.g.,} 0.90, \quad \text{if } |\langle I_I \rangle_{\text{now}} - I_{\text{fltr}}| \leq 3 \sigma_{\langle I_I \rangle_{\text{last}}}; \\
\text{oblv} &= \text{oblv}_2 = \text{e.g.,} 0.99 \ (> \text{oblv}_1), \quad \text{and} \\
\text{oblv} &< \text{oblv}_2, \\
\sigma_{\langle I_I \rangle_{\text{last}}} &= \sqrt{\langle (\langle I_I \rangle_k - (\langle I_I \rangle_k)^2) \rangle_{\text{now}}} = \sqrt{\langle (\langle I_I \rangle_k)^2 \rangle_{\text{now}} - \langle (\langle I_I \rangle_k \rangle_{\text{now}}^2 \rangle} \\
&= \sqrt{\langle (\langle I_I \rangle_k)^2 \rangle_{\text{now}} - I_{\text{fltr}}^2}, \\
\langle (\langle I_I \rangle_k)^2 \rangle_{\text{now}} &= \text{oblv}_{\text{now}} \times \langle (\langle I_I \rangle_k^2)_{\text{last}} + (1 - \text{oblv}_{\text{now}}) \times \langle (\langle I_I \rangle_k^2)_{\text{now}},
\end{align}

where

$I_{\text{fltr}}$: present real (expected) DC offset value to be determined,
$\langle I_I \rangle_{\text{now}}$: present 200-sample-average of I/Q signals during a fclr stage,
$\langle I_{\text{fltr}} \rangle_{\text{last}}$: DC offset value determined in the last fclr stage,
$\sigma_{\langle I_I \rangle_{\text{last}}}$: standard deviation of the $\langle I_I \rangle_k$ (not $I_I$) at the last time,
$\sigma_{\langle I_I \rangle_{\text{now}}}$: current standard deviation of the $\langle I_I \rangle_k$ (not $I_I$),
$\langle (\langle I_I \rangle_k^2)_{\text{now}}$: current average of $\langle I_I \rangle_k^2$.

As the DC offsets may depend on the operating frequency, we divided the 8–20 MHz SuperDARN operating frequency band into 24 sub-bands in 0.5-MHz steps and used only the offset values in the same frequency sub-band to calculate the expected DC offsets. Using this algorithm and, for simplicity, the case where $\text{oblv}$ is constant (i.e., equal to $\text{oblv}_1$ (e.g., 0.9)),

\[ \text{DC offset}(m) = (1 - \text{oblv}) \sum_{k=0}^{m} (\text{oblv})^k \times \text{DC offset}(m-k), \]

which means that the contribution of the older DC offsets vanishes exponentially (if the DC offset value is assumed to be almost constant or vary only very slowly with time, without any unexpected temporal jumps). In other words, the past DC offset values are averaged with a weighting function, like in eq. (39), so that the older values are less influential or “forgotten” (more “oblivious”) in an exponential manner. The first oblivion coefficient, $\text{oblv}_1$, was selected so that the temporal variation of $I_{\text{fltr}}$, in eq. (33), i.e., the expected DC offset value, would be within 1 A/D unit and could quickly follow the expected real DC offset variations (e.g., temporal variations in room temperature, etc.) over a time scale of typically several minutes. The relationship in eq. (36) (i.e., $\text{oblv}_2 > \text{oblv}_1$) is for reducing the effect of unexpected large interference noises. However, the second oblivion coefficient, $\text{oblv}_2$, should not exactly equal 1.0 because of the possibility of sudden (manual or unexpected) changes in the real DC offset values for safety. The value, $\text{oblv}_2$, determines the time constant for the calculated DC offset values following such sudden changes (although the time constant with this method also depends on the beam integration time and the operating frequency sequences, such as the fixed frequency mode, frequency scan mode, or frequency optimization mode that allows arbitrary jumps in frequency sub-bands).

The current “fclr” function in the ROS only samples 2 A/D channels (the I and Q outputs for the main receiver) to determine the noise level for each test frequency. However, our new “fclr” function uses all 4 A/D channels (the I/Q channels for the main receiver and the ones for the interferometer receiver) to determine all 4 DC offsets.
Furthermore, the real-time DC offset values for the 4 A/D input channels (both the instantaneous 200-sample average and the statistical expected value) can be seen on the "idisplay (real-time ROS quick-look program)" screen (updated for every beam integration time) and the (measured and expected) DC offset values (i.e., \(D_{1c}\) and \(\langle D_{1c}\rangle\)) are recorded in very small "ofs" files for later analysis (if desired). After the obtained DC offset values are subtracted from the I and Q signals, the corrected noise values are determined using our new fclr function. The corrected I and Q sample data are then passed to the functions that construct the ACFs and XCFs, as in the present ROS; the constructed ACFs and XCFs are then passed to the fitacf algorithm. The corrected I and Q samples are also passed to functions that form and process the raw time series.

We developed a code to monitor, remove and record the DC offsets in real-time for the previous radar operating system (called RADOPS/2000) and have migrated this code to the present system (RST/ROS). The new code was installed in August 2001 at the Syowa South and East HF Radars of NIPR for the SuperDARN (SENSU) systems located at Syowa station, Antarctica. The code has been obtaining and removing DC offset values continuously since October 2001.

4. Initial results

Here, we present some initial results obtained with the new DC offset monitoring and removal code installed in the SENSU Syowa radar systems.

Figure 1 shows an example of the one-day DC offset variation in the I channel of the main receiver obtained at SENSU Syowa South radar on Aug. 20, 2001. Each horizontal axis represents the universal time (UT) in hours, and each vertical axis represents the channel output in 12-bit A/D units (12-bit (4096) corresponds to \(\pm 10\) V).

Figure 1A shows the one-day variation of the simple 200-sample-average values during each fclr stage, which corresponds to \(I_{10}\) in eq. (22). Figure 1B shows the standard deviation of \(I_{10}\), which shows that the 200-sample-averages scatter over several A/D units. Figure 1C is the DC offset of the channel determined using the method described in the previous section. The temporal variation in DC offset values is stable with time (within \(~1\) A/D unit) over a short time period and properly determined.

**SENSU Syowa South I1-channel 2001/08/20**

\[<I_{1c}> \quad (200\text{-sample average of } I_{1} \text{ during fclr, } I_{10})\]

\[\sigma_{I_{1}} \quad (\text{Std.dev. of } I_{1} \text{ during each fclr})\]

\[<\langle I_{1c}\rangle> \quad (\text{Determined } I_{1} \text{ DC offset (oblw=0.9, oblw=0.99)})\]

**UT [hour]**

**Fig. 1.** An example of the one-day DC offset data for the I channel of the main receiver, obtained by the SENSU Syowa South radar on August 20, 2001. The horizontal axes represent the universal time in hours. See the text for details.
Fig. 2. Effect of the DC offset removal method on data obtained at the SENSU Syowa East radar between 0800 and 1400 UT on June 25, 2002. See the text for details.
Figure 2 shows the effect of the DC offset removal method on results obtained at the SENSU Syowa East radar between 0800 and 1400 UT on Jun. 25, 2002. The DC offsets are always removed up until 1000 UT. After 1000 UT, two beam scan periods are repeated so that the DC offsets are forced to be added to each I and Q channel during the first normal beam scan period, while the offsets are properly removed during the next beam scan period. The artificially added DC offsets during each first beam scan period were $I_{10}=0$ and $Q_{10}=10 \times \text{bmmnum}$ (in 12-bit A/D units, bmmnum: beam number). Figure 2A (the uppermost left and right panels) shows the temporal variation of the fitted physical parameters of beam 2, i.e., the echo power in dB (top bar), the LOS Doppler velocity in m/s (middle bar), and the Doppler spectral width in m/s (bottom bar) using a color code. The horizontal axis shows the universal time in hours, and the vertical axis shows the range gate number. The left panel shows the effect of the forced DC offset ($Q_{10}=20$) after 1000 UT, while the right panel shows the effect of the DC offset removal code. Figure 2B, 2C, and 2D show the results for beam 3 ($Q_{10}=30$), beam 6 ($Q_{10}=60$), and beam 12 ($Q_{10}=120$), respectively. The noise level ($N_0$) was relatively constant at about 100 (in A/D unit$^2$), throughout the period shown in the figure. Therefore, the relationship among the lag-0 power (denoted here as “lag0pwrdb” in dB, which is equivalent to $10 \log(\text{ACF}(\tau=0))$), the echo power (denoted as “pwrdb” in dB) and the noise level ($N_0$ in A/D unit$^2$), and the relationship between the beam number (bmmnum) and the DC offset power ($10 \times \log(|Z_{10}|^2)$ in dB) for this period are as follows:

\begin{align*}
10 \log(\text{ACF}(\tau=0)) &= \text{lag0pwrdb} = 10 \log(N_0) + \text{pwrdb} = 20 + \text{pwrdb}, \\
10 \log(|Z_{10}|^2) &= 10 \log((10 \times \text{bmmnum})^2) = 20 + 20 \log(\text{bmmnum}).
\end{align*}

We can roughly expect from eq. (9) that, if $N_0 \sim 100$, the ACFs will be difficult to fit if

$$|\text{ACF}(\tau=0)| \lesssim |Z_{10}|^2.$$  

In other words, for $N_0 \sim 100$, 

$$\text{pwrdb} \lesssim 20 \log(\text{bmmnum})$$  

\begin{align*}
\sim -6 \text{ dB} & \quad \text{for bmmnum}=2 \\
\sim -9.5 \text{ dB} & \quad \text{for bmmnum}=3 \\
\sim -15 \text{ dB} & \quad \text{for bmmnum}=6 \\
\sim -22 \text{ dB} & \quad \text{for bmmnum}=12.
\end{align*}

We can see that Fig. 2 roughly agrees with our estimation. Hence, we can use these results to confirm that the DC offset removal code works very effectively, enabling the detectability and the amount of echoes to increase dramatically when large DC offsets are properly removed.

5. Concluding remarks

I/Q DC offsets can cause systematic errors in the estimation of noise levels, ACFs, XCFs, and fitted physical parameters, such as the echo power, LOS Doppler velocity, Doppler spectral width and elevation angle, and also in the estimation of echo power and phase in raw time series analyses. Therefore, we have developed a new code to monitor,
determine remove and record the I/Q DC offset values for the SuperDARN ROS. This code determines the DC offsets during the Rx-only fcbr stage prior to constructing the ACFs and XCFs or before the raw time series analysis and seems to work well. This method will increase the accuracy of SuperDARN fitted physical parameters (especially elevation angles) as well as the ACFs, XCFs and echo power and phase in raw time series. Consequently, the code will increase the detectability and the amount of observed echoes, especially those with a smaller power that is close to the maximum among noise level and DC offset power ($|Z_0|^2$).

However, this method has one limitation in that the calculated DC offsets follow the real DC offset values by a certain time constant if a large, unexpected or manual change in the DC offsets occurs. Furthermore, our method only uses 200-sample data points at the selected operating frequency in each fcbr stage. If a smaller time constant of the DC offset values is desired, we should average a larger number of I/Q samples during each fcbr stage in order to make the average values be statistically more stable. In the typical fcbr stage, the five quietest frequencies out of all 60 5-kHz-step frequencies in the 300-kHz frequency band are selected during the first fcbr stage in order to pass them to the second fcbr stage for choosing the real quietest frequency. Therefore, the 200 samples obtained for each of these five frequencies during the second fcbr stage could then be used to obtain 1000 samples (200 samples/frequency × 5 frequencies = 1000 samples) in order to determine the 1000-sample I/Q average that is statistically more stable than the 200-sample I/Q average. Alternatively, the 50 samples in each of the 60 frequencies in the first fcbr stage could also be added to the 200 samples in each of the 5 quietest frequencies (as determined in the first fcbr stage) measured in the second fcbr stage to obtain 4000 samples (50 samples/frequency × 60 frequencies (in the first fcbr stage) + 200 samples/frequency × 5 frequencies (in the second fcbr stage) = 4000 samples). These methods assume that the DC offsets are almost completely constant within the narrow 300-kHz frequency band. Our method makes this assumption and divides 8–20 MHz band into 24 sub-bands in 500-kHz steps; the above methods also seem to work and might be a better solution. However, the above methods will be affected by possible interference noises more frequently, especially in noisy environments. Our method can be more safely used even in noisy environments, with the above limitation. Nevertheless, using all 1000 samples obtained at the five quietest frequencies might also be a good solution.

Some SuperDARN radar studies have reported that the DC offsets depend on the attenuation levels (as well as the operating frequencies). We feel that this should not occur unless some of the local oscillator signals leak into the IF stage in the receiver, or a similar error occurs. However, since our method can be easily used to determine the DC offset levels for every attenuation level, we will add such a function to our new code in the near future. Recently, some discussion regarding the replacement of the 12-bit A/D card with a 16-bit A/D card has arisen in the SuperDARN community. If this replacement occurs and the input level (gain) of the A/D card does not change, the DC offset values will increase by 16 times. To take advantage of the higher resolution A/D card and to improve the detectability of weaker echoes, the precision (the effective number of averaging) of our DC offset removal code will need to be increased so that the statistical deviation is small enough (<1 A/D bit). Theoretically, this should not present a problem and can
be accomplished with a minor modification of the code.

Acknowledgments

The authors are grateful to K.B. Baker of the National Science Foundation, M. Pinnock of the British Antarctic Survey, and W.A. Bristow and J. Hughes of the University of Alaska, Fairbanks, for their helpful discussions. This research was supported by a Grant-in Aid for Scientific Research (A: 11304029) from the Japan Society for the Promotion of Science (JSPS). The Ministry of Education, Culture, Sports, Science and Technology supports the SENSU Syowa HF radar systems. The 30th-42nd Japanese Antarctic Research Expeditions (JAREs) have performed all of the HF radar operations at Syowa. We especially thank all the JARE members who contributed to the construction and operation of the HF radar systems at Syowa Station.

The editor thanks Dr. T. Ogawa for his help in evaluating this paper.

References


(Received June 5, 2002; Revised manuscript accepted July 2, 2002)
Editorial policy

*Advances in Polar Upper Atmosphere Research* is an annual publication from National Institute of Polar Research (NIPR) including contributed papers on the upper atmosphere physics in the polar regions as well as papers presented at the NIPR Symposium on Coordinated Observations of the Ionosphere and the Magnetosphere in the Polar Regions. *Advances in Polar Upper Atmosphere Research* accepts the following categories of papers; scientific papers, research notes, review papers and reports. Scientific papers are original articles dealing with major problems related to the polar upper atmosphere physics. Research notes are short articles dealing with a significant topic, new findings, and preliminary interpretation, etc. Review papers deal with overview of researches on specific issues. Reports include scientific and technical information on the activities related to polar upper atmosphere research.

Submission of manuscripts

Authors are requested to submit, with their manuscripts, a sheet of submission to indicate the contact address, telephone and fax numbers and e-mail address, etc. In case of more than one author, please indicate to whom the correspondence should be addressed. The format of the sheet is available upon request to:

Editorial Board of Advances in Polar Upper Atmosphere Research
National Institute of Polar Research
Kaga 1-9-10, Itabashi-ku, Tokyo 173-8515, Japan
Tel. +81-3-3962-4756 Fax. +81-3-3962-5742 E-mail: editor@uap.nipr.ac.jp

Three copies of the manuscript prepared according to the Technical details should be submitted. Arrival date of the manuscripts to the Editorial Board is registered as the received date.

Review processes

The received scientific papers and research notes are reviewed by at least two referees. Review papers and reports are reviewed by at least one referee. Referees and the editors may request modification/revision of the papers, and the acceptance of the paper and its final category is determined by the Editorial Committee of NIPR.

Proofs

One set page proofs will be sent to the corresponding author, to be checked for typesetting/editing. The author is not allowed to make changes or corrections that constitute departures from the article in its accepted form. Proofs should be returned within 2 days.

Submission of electronic text

Authors are requested to submit only the final text (after acceptance) on a 3.5" diskette (DOS or ASCII). The name and version of the wordprocessing program should be clearly labeled.

Reprints

It is free of charge for submission and publication. The corresponding author can receive free 50 reprints. Additional reprints can be ordered on a reprint order form.
Acknowledgment to the referee
In your final manuscript, the authors are requested to append the following lines below the acknowledgment paragraph as
"The editor thanks ... and another referee for their help in evaluating this paper."

Copyright agreement
The copyright of any published articles (hereafter referred to as papers) in the Journal belongs to the National Institute of Polar Research. Reproduction, translation and other uses to any extent of papers in this Journal require written permission from the Head, Library of the National Institute of Polar Research and acknowledgment of the source. In exception of the above requirement, use of several figures, tables and brief excerpts from the Journal in scientific and educational publications does not require the permission from the National Institute of Polar Research provided that the source is acknowledged.

Technical details
(1) All manuscripts must be written in English.
(2) All pages of the manuscripts must be typed double spaced on one side of the A4 paper with wide margins, and assembled in the order; title, abstract, text, acknowledgment, references, appendices, tables, figure captions and figures.
(3) An abstract is within 200 words.
(4) Use of the Système International d'Unités (SI) is recommended.
(5) The title page should include the category of contribution, the title, the name(s) of author(s), their affiliations and addresses. In case the title contains more than 50 letters, a running title should also be given.

References
(1) References in the text consist of the surname of the author(s), followed by the year of publication in parentheses. All references cited in the text should be given in the reference list and vice versa.
(2) The reference list should be in alphabetical order.
(3) References should be in the following format:


Tables/Figures
Each table should be typed on a separate sheet of A4 paper and numbered according to their sequence in the text. All figures should be numbered consecutively and referred to in the text. Color figures can be accepted upon request. But the final decision rests with the editorial board.
AUTHOR INDEX

Aso, Takehiko 99
Barabash, V. 111
Belova, E. 111
Bering, Edger A. 157
Brittnacher, Mitchell 136
Chilson, Phillip B. 111
Cierpka, K. 1
Ejiri, Masaki 99, 157, 173
Hagfors, T. 1
Hirosawa, Haruto 157
Igarashi, Kiyoshi 59
Inoue, Shoshi 13
Ishii, Mamoru 173
Itonaga Masahiro 45
Kadokura, Akira 157
Kamide, Yosuke 136
Katoh, Yasuo 146
Kirkwood, S. 111
Koizumi, Naoko 99
Kosch, M. J. 1
Kourov, A. V. 59
Kubota, Minoru 173
Makarevitch, R. A. 59
Murayama, Yasuhiro 173
Nagatsuma, Tsutomo 36
Nagakawa, Michio 157
Nakai, Hitoshi 136
Nilsson, H. 111
Nishitani, Nozomu 59, 84
Ogawa, Tadahiko 59, 84, 146
Ohtaka, Kazuhiro 59
Okada, Toshimi 157
Okano, Shoichi 99
Ono, Takayuki 126
Osepian, A. 111
Oya, Hiroshi 126
Ozaki, Masao 13
Rao, T. N. 111
Rietveld, M. T. 1
Sakano, Takeshi 99
Sato, Natsuo 59, 84, 157, 181
Satoh, Mitsugi 146
Shinbora, Atsuki 126
Shiokawa, Kazuo 146
Stebel, K. 111
Taguchi, Makoto 99, 146
Takeshita, Shu 173
Tohyama, Fumio 157
Torii, Shoji 157
Tsutsui, Masaki 181
Yamagami, Takamasa 157
Yamagishi, Hisao 59, 84, 146, 157, 181
Yamamoto, Takashi 13
Yamamoto, Masayuki 173
Yeman, T. K. 1
Yoshikawa, Akimasa 45
Yukimatu, Akira Sessai 59, 84, 181
Yumoto, Kiyohumi 45

195
Acknowledgments to Referees

The editorial board wishes to thank following referees.

Fujita, S. (Meteorol. Coll.) Gustavsson, B. (NIPR) Hoz, Cesar La (U Tromsø)
Iyemori, T. (Kyoto U) Lester, M. (U Leicester) Maezawa, K. (ISAS)
Ogawa, T. (Nagoya U) Okano, S. (Tohoku U) Omura, Y. (Kyoto U)
Van Eyken, A.P. (SRI) Yoon, P.H. (U Maryland)

The proceedings of past Symposia on Upper Atmosphere Physics were published as follows:
The 1st Symposium: Nankyoku Shiryō (Antarctic Record), 63, 310 p., 64, 176 p., 65, 1-118, 1979
The 2nd Symposium: Nankyoku Shiryō (Antarctic Record), 68, 319 p., 69, 252 p., 1980

Copyright 2002 by the National Institute of Polar Research
9-10, Kaga 1-chome, Itabashi-ku, Tokyo 173-8515
Printed by Sasaki Printing & Publishing Co., Ltd.